CHEMICAL ENGINEERING TRANSACTIONS
VOL. 65, 2018
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Guest Editors: Eliseo Ranzi, Mario Costa

ISBN 978-88-95608-62-4; **ISSN** 2283-9216
DOI: 10.3303/CET1865127

Hyperspectral Analysis of Biomass Element Content in Straw

Junlong Fang, Heng Li, Jinglong Zhen, Yu Nie, Chentao Zhang

College of Electrical and Information, Northeast Agricultural University, Harbin 150030, China
jlfang34658@163.com

In this paper, a quantitative analysis model is established with the help of hyperspectral imaging technology and the least squares method to study the biomass content in straw. The results show that the optimal selection of spectral dimensional data can be achieved through a competitive adaptive weighted sampling algorithm. In the experiment, the correlation coefficient of nitrogen in the verification set is 0.923 and the correlation coefficient of oxygen in the verification set is 0.876. Given that the prediction results of these two elements are relatively good and thus can be applied in practice. The practicality of other elements is poor. However, in the quantitative analysis model, the prediction results of these five elements are not satisfactory and the practicality is poor.

1. Introduction

With the development of science and technology, the productivity in rural areas has been further improved and the number of straw crops in rural areas has also presented a growing trend. The rational use of biomass content in straw is not only beneficial to the secondary utilization of energy, but also can promote the sustainable development of rural agriculture. Based on this, this paper uses the hyperspectral imaging technology to analyze the biomass content in straw.

2. State of the art

Hyperspectral image technology is a detection technique widely used in non-destructive testing of agricultural products in recent years (Edelman et al., 2015). Compared with conventional machine vision and near-infrared spectroscopy, hyperspectral imagery can simultaneously obtain external image information such as seed shape and color, as well as spectral information reflecting the internal chemical composition of the seed, thereby realizing a multi-faceted and multi-angle analysis of the object to be measured. Hyperspectral is widely used in non-destructive testing of agricultural products due to its many advantages. However, there are many bands of hyperspectral data. When using full-band modeling analysis, it not only increases the storage space of data, but also increases the actual amount of calculations. The real-time performance of the test results is affected. In fact, there is a high degree of correlation and redundancy between adjacent bands of a hyperspectral image, so that not every band has the same value for image processing (ElMasry and Nakauchi, 2016). Through the band selection algorithm, the least and most representative band subsets are combined to form a new hyperspectral image space to approximately replace the original hyperspectral image space, thereby achieving the purpose of reducing the dimension of hyperspectral image data. At the same time, the impact on the overall recognition accuracy is also relatively small, which is the ultimate goal of the selection of hyperspectral image bands. Since the industrial revolution, the process of urbanization in many countries of the world has entered a stage of rapid development. The number of urban population is increasing and the scale of cities is expanding. This is not a subjective process, but a result of the development of productive forces, the progress of science and technology, and the optimization and upgrading of industrial structures. Urbanization makes social division clearer (Guo et al., 2017). Scale effect and agglomeration effect are given full play, and the huge energy of urbanization is released. Siciliano linked urbanization strategies to changes in land use and associated impacts on rural communities and agro-ecosystems in a rural area of China, his study showed that urbanization strategies brought pressure on environment. Sargeson concluded that the urbanization leaded to violent land expropriation. Since the reform

Please cite this article as: Junlong Fang , Heng Li , Jinglong Zhen , Yu Nie , Chentao Zhang , 2018, Hyperspectral analysis of biomass element content in straw, Chemical Engineering Transactions, 65, 757-762 DOI: 10.3303/CET1865127

and opening up, China's economic development has made great achievements, and people's living standards have also been improved. However, the economic and social development between urban and rural areas is not coordinated. The traditional urbanization is taking economic development as the main target, the industrialization as the main line, and the local government as the leading (Rahman et al., 2016). It belongs to the high cost of low-income urbanization. Traditional urbanization brought about a series of problems of structural imbalance, spatial imbalance, industrial structure imbalance and urban disease. New urbanization is not only the key to the realization of modernization, but also to achieve the potential of building a moderately prosperous society. Industry is a natural process from agriculture to industry, and there is a general law. However, under different systems, the different stages of industrialization have different development paths and patterns. With the improvement of the level of industrialization, urban land area has been expanding and the population has been increasing (Luciani et al., 2015). Accordingly, with the acceleration of the urbanization process, it brings a series of problems such as traffic congestion, environmental pollution and chaos. Chemical industry is an important part of industrial development, which plays an important role in the national economy. It is also the basic industry and pillar industry in many countries (Liu et al., 2014). The pace and scale of the development of the chemical industry have a direct impact on the various sectors of the society and economy. At the same time, the chemical industry is also a big polluter. In the processing, storage, use and waste disposal and other links, chemical products are likely to produce a large number of toxic substances and affect the ecological environment, thus endangering human health. The sustainable development of chemical industry is of great practical significance to human economic and social development (Ma et al., 2014). Urbanization and industrialization are two wheels of economic advance. If there is no industrialization, urbanization will lose the power of development. If there is no urbanization, industrialization will lose the support of development. New urbanization and new industrialization are put forward in the new situation. It is not only the socialist path with Chinese characteristics of promoting the building of a well-off society and realizing the dream of China, but also a comprehensive development strategy of promoting new industrialization and new urbanization. In particular, the current construction of new urbanization requires the coordinated development of the chemical industry with sustainable development (Naganathan et al., 2016). The sustainable development of chemical industry requires the development of green industry in chemical industry. First, we should adopt the principle of cleaner production, use the clean energy and raw materials, and apply the advanced technology and equipment to improve the comprehensive utilization of management and other measures to reduce pollution from the source to improve resource efficiency, in order to reduce or eliminate the hazards to human health and the environment.

3. Research principles and methods

A total of 188 straw samples were collected, of which 89 were rice; 39 were wheat; 21 were corn; and 39 were rape. The samples came from the producing areas in Central and Southwest China (Hubei, Hunan, Sichuan, Chongqing, Guizhou, and Yunnan Province) covering various factors such as region, variety, climate, and species. The collected straw samples were first spread in the open air to a dry state (moisture content was about 10%) and then the samples were crushed with a pulverizer and placed in an oven (45±5) °C for about 8 hours. The moisture was controlled at 5.5% ± 1%. Then, a hammer cyclone mill was used to grind, passing through a 40-mesh screen. The ground samples were placed in a labeled ziplock bag and stored at room temperature (around 25°C) for the collection of hyperspectral images and the measurement of basic chemical elements (N, C, H, S, O). The hyperspectral imaging system was used in this experiment for the collection of sample images, as is shown in Figure 1.

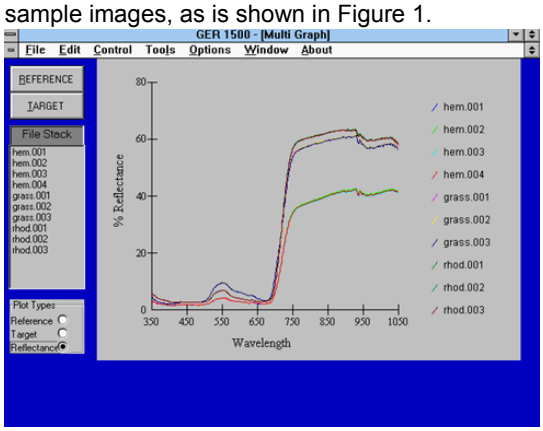

Figure 1: Reflectance hyperspectral imaging system

This system is mainly composed of hyperspectral imager (SPECIM, V10E, Finland), collection camera obscura, height-adjustable working stage, ventilation fan, high-precision electronic control mobile platform and specialized computer. In the reflectance hyperspectral imaging system, the ring illuminator is placed on the illuminator holder in the camera obscura, which can provide the light source of 400 to 2500 nm spectral band. In order to avoid the interference of external light sources and other noise, the entire operation is completed in a camera obscura. The exposure time of the camera will affect the sharpness of the image and the movement speed of the high-precision electronic control mobile platform will affect the shape of the captured image. To obtain clear and undistorted images, these parameters need to be set. Based on the principle that the image signal is lower than the saturation value, the camera exposure time is determined as 0.1 s and the image resolution ratio is 250 pixels x 250 pixels; the movement speed is determined by collecting the images of the standard black and white squared paper at different collection speed. The speed that satisfies the undistorted condition is 2 mm/s. After 30 minutes of preheating, the hyperspectral imager can start the collection of images. The grinded and preserved straw samples are placed in a specialized spectral collection sample pool for image collection.

The N, C, H, S, O elements of the sample (mass fraction) are measured according to the standards of the American Society for Testing and Materials (ASTM) and the instrument used is an EA 3000 elemental analyzer (Euro Vector Company, Italy). Three parallels are made on each sample and the average value is taken as the content of the sample.

The element quantitative analysis model is established using spectral dimension and image dimension data respectively. First, the average spectrum of the region of interest in the hyperspectral image of the straw sample is extracted as the modeling spectrum of the straw element analysis model and the pre-possessing of the original spectrum is conducted using the remove the trend of transformation (Detrend), first-order derivative process (FD), multiple scattering correction (MSC) and its combination processing. The quantitative analysis model of straw element content is established combined with partial least squares (PLS).

On this basis, the competitive adaptive reweighted sampling (CARS) is used to select the characteristic spectrum with higher contribution rate and an optimized element analysis model based on spectral dimensional data is established. Spectral element-based data element analysis model is established. Secondly, the independent component analysis (ICA) is conducted on the images of straw hyperspectral samples to obtain a series of images containing from image information to noise. Through observation and analysis, relevant images containing element information are selected and the characteristic spectrum is obtained using weight coefficient method. Based on the sample IC image and the extracted characteristic spectral information, the straw element analysis model is established based on ICA-PLS.

In order to evaluate the predictability and practicality of the quantitative analysis model, the following parameters are introduced: correlation coefficient of calibration set cross-validation and root mean square error; correlation coefficient in the verification set and root mean square error. Meanwhile, the validation set is used to conduct relative analysis on error RPD. RPD = SD / RMSEP. SD is the standard deviation of sample element content in the validation set and RMSEP is the root mean square error of the validation set, which are then used for further evaluation of the model.

4. Research results and analysis

Chapter 2 The Monte Carlo algorithm is used to remove abnormal samples. The number of samples eliminated of N, C, H, S, O elements is 10, 5, 13, 8 and 12 respectively; the sampling is conducted using concentration gradient method and principal component analysis. The statistical results of the elemental content of the sample set, calibration set and verification set after removing the abnormal sample are shown in Table 1.

Table 1: Statistics of elemental analysis index and gross calorific value in straw

	Correction set		Validation set		
	n	SD/%	n	Average/%	SD/%
N	120	0.63	58	1.22	0.61
C	122	2.47	61	44.32	2.53
H	116	0.70	59	6.24	0.68
S	116	0.14	64	0.24	0.11
O	117	2.23	59	38.50	2.35

According to Table 1, the mass fraction of the C and O elements in the sample set is relatively high, ranging from 34.33% to 52.81% and 33.29% to 43.75% respectively; the average mass fraction of N, H and S

elements is 1.12%, 6.48% and 0.25% respectively, which are relatively low content. Due to the wide sources and varieties of samples, the variable coefficient of each element is relatively large. The variable coefficient of N, H and S elements is 52.29%, 10.72% and 49.30% respectively, which indicates that the dispersion degree of sample element content data is relatively higher and the source of samples is wide. Also, these samples are representative and this experiment has practical significance.

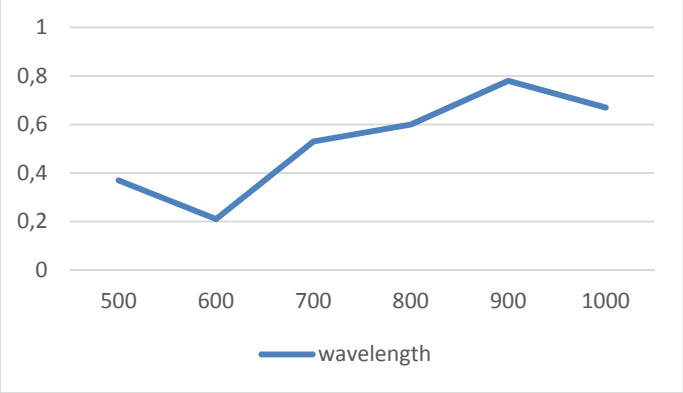

Figure 2: Wavelength variation curve

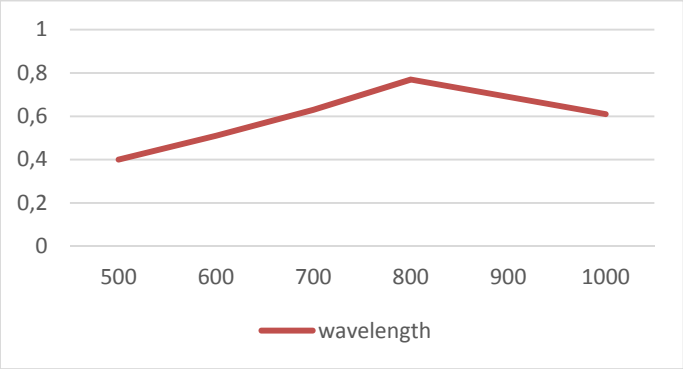

Figure 3: Variable selection results based on competitive adaptive reweighted sampling (CARS)

The average spectrum of the region of interest in the hyperspectral image of the straw sample is extracted as the modeling spectrum of the straw element analysis model and the PLS is applied in the full spectrum range (380 to 1100 nm) and the quantitative analysis model of straw element content is established combined with the pre-processing algorithm of various spectrums. The modeling results are shown in Table 2.

Table 2: Results of elemental analysis model based on full spectrum and partial least squares (PLS)

	Calibration set		Validation set		
	R_{cv}	RMSECV/%	R_p	RMSEP/%	RPD
N	0.926	0.234	0.901	0.217	2.81
C	0.673	1.891	0.666	1.384	1.83
H	0.657	0.537	0.649	0.529	1.29
S	0.717	0.092	0.688	0.063	1.74
O	0.863	1.100	0.856	1.105	2.13

It can be seen from Table 2 that the correlation coefficient (Rp) of N element of the testing model in the verification set is 0.901; the root mean square error (RMSEP) is 0.217%; the relative analysis error (RPD) is 2.81, showing good prediction ability. The correlation coefficient (Rp) of O element of the testing model in the verification set is 0.856; the root mean square error (RMSEP) is 1.105%; the relative analysis error (RPD) is 2.13, showing good prediction effect. The correlation coefficient of C, H and S in the verification set is less than 0.70, which indicates that the established model cannot achieve quantitative analysis.

The competitive adaptive weighted sampling algorithm (CARS) is used to select a certain number of variables from the full-band variables to establish the model. The effect of the model is compared to eliminate the variable with low contribution and finally the optimal spectral variable is selected to build the model. The distribution of the variables selected by each element of the CARS algorithm is shown in Figures 2 and 3.
It can be seen from the analysis of Figures 2 and 3 that the CARS variable optimization algorithm is used to select the sensitive variables of different elements of the straw. Although the variables are distributed over the entire spectrum, they are mainly concentrated in the near-infrared spectral region. The variables in the optical band range are relatively less, so the near-infrared band is very suitable for the quantitative analysis of straw elements.
Based on the spectral variables selected by CARS, the PLS is used to establish the analysis model of various elements of the straw. The modeling results are shown in Table 3.

Table 3: The modeling results

	Calibration set		Validation set		
	R_{cv}	RMSECV/%	R_p	RMSEP/%	RPD
N	0.954	0.185	0.923	0.196	3.11
C	0.716	1.725	0.721	1.364	1.85
H	0.859	0.511	0.792	0.596	1.14
S	0.764	0.083	0.669	0.062	1.77
O	0.871	1.060	0.876	1.015	2.32

It can be seen from the comparison of Table 2 and Table 3 that the number of variables participating in the modeling decreases significantly and the stability and prediction performance of the model both improve when CARS algorithm is used to optimize the quantitative analysis model of straw elements. The models of N and O elements are the optimal and 24 variables are selected to establish the model for N element. The correlation coefficient of the validation set (Rp) is 0.923; the root mean square error (RMSEP) is 0.196%; and the relative analysis error (RPD) is 3.11. Only 10 variables are selected to establish the model for O element. The correlation coefficient of the validation set (Rp) is 0.876; the root mean square error (RMSEP) is 1.105%; and the relative analysis error (RPD) is 2.32. It can be seen that the model predictive ability of N and O elements has been improved significantly and can be used in practical application.
Comparing the model prediction results of the CARS-PLS for the C, H and S elements with the model prediction results established for the entire waveband, although the correlation coefficient (Rp) and relative analysis error (RPD) of the validation set have increased, the correlation coefficients (Rp) both are less than 0.80, indicating poor prediction effect. From the above analysis of the sensitivity variables selected for the detection of each element, it is known that the near-infrared spectral region is more conducive to the quantitative analysis of straw elements. Therefore, the use of near-infrared hyperspectral imaging may improve the accuracy of the quantitative detection of C, H, and S elements.
Based on the sample IC image and the extracted spectral information, quantitative analysis model of straw elements based on ICA-PLS is established combined with partial least squares (PLS) algorithm. The prediction effect of O element is the best, whose correlation coefficient (Rp) in the verification set is 0.808; root mean square error (RMSEP) is 0.932%; relative analysis error (RPD) is 2.04, indicating poor prediction effect. The model prediction effect of N, C, H, S and O in biomass straw based on ICA-PLS is worse than that of full spectrum-PLS and CARS-PLS. In addition, except O, the correlation coefficient (Rp) of the validation set of the remaining four elements is less than 0.80. The results show that the quantitative analysis model of straw element based on ICA-PLS algorithm cannot be used in practical application.

5. Conclusion

With the help of the hyperspectral system, this paper analyzes the biomass content in straw based on spectral dimension and least square method. It can be from the research results that the optimal selection of spectral dimension data can be achieved with the help of competitive adaptive weight sampling algorithm. In this experiment, the correlation coefficient of nitrogen in the verification set is 0.923 and the correlation coefficient of oxygen in the verification set is 0.876. Given that the prediction results of these two elements are relatively good and thus can be applied in practice. The practicality of other elements is poor. However, in the quantitative analysis model, the prediction results of these five elements are not satisfactory and the practicality is poor.

Acknowledgments

Supported by the national science and technology support program (2014BAD06B04-1-09); China Post Doctoral Fund (2016M601406); Heilongjiang Post Doctoral Fund (LBHZ15024).

References

Edelman G.J., Leeuwen T.G., Aalders M.C. 2015, Visualization of latent blood stains using visible reflectance hyperspectral imaging and chemometrics, Journal of forensic sciences, 60(1), 234-245, DOI: 10.1111/1556-4029.12591

ElMasry G.M., Nakauchi S. 2016, Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality-A comprehensive review, Biosystems engineering, 142, 53-82, DOI: 10.1016/j.biosystemseng.2015.11.009

Guo D., Zhu Q., Huang M., Guo Y., Qin J. 2017, Model updating for the classification of different varieties of maize seeds from different years by hyperspectral imaging coupled with a pre-labeling method, Computers and Electronics in Agriculture, 142, 1-8, DOI: 10.1016/j.compag.2017.08.015.

Liu D., Sun D.W., Zeng X.A. 2014, Recent advances in wavelength selection techniques for hyperspectral image processing in the food industry, Food and Bioprocess Technology, 7(2), 307-323, DOI: 10.1007/s11947-013-1193-6.

Luciani V., Bonifazi G., Rem P., Serranti S. 2015, Upgrading of PVC rich wastes by magnetic density separation and hyperspectral imaging quality control, Waste Management, 45, 118-125, DOI: 10.1016/j.wasman.2014.10.015.

Ma W.K., Bioucas-Dias J.M., Chan T.H., Gillis N., Gader P., Plaza A.J., Chi C.Y. 2014, A signal processing perspective on hyperspectral unmixing: Insights from remote sensing, IEEE Signal Processing Magazine, 31(1), 67-81, DOI: 10.1109/MSP.2013.2279731.

Naganathan G.K., Cluff K., Samal A., Calkins C.R., Jones D.D., Meyer G.E., Subbiah J. 2016, Three dimensional chemometric analyses of hyperspectral images for beef tenderness forecasting, Journal of Food Engineering, 169, 309-320, DOI: 10.1016/j.jfoodeng.2015.09.001.

Rahman N.S.F.A., Najib A.F.A. 2017, Selection of the most practical Malaysian port for enhancing the Malaysia-China Kuantan Industrial Park business trade, International Journal of Shipping and Transport Logistics, 9(4), 500-525, DOI: 10.1504/IJSTL.2017.084829.