

Article

Quantitative Analysis of Coal Quality by a Portable Laser Induced Breakdown Spectroscopy and Three Chemometrics Methods

Youquan Dou ¹, Qingsong Wang ¹, Sen Wang ¹, Xi Shu ¹, Minghui Ni ² and Yan Li ^{2,*}

¹ Nanjing Coal Quality Supervision and Inspection Co. Ltd., China Energy Corporation, Nanjing 210031, China; 12050113@chnenergy.com.cn (Y.D.); wangqs0810@163.com (Q.W.); 15210352762@163.com (S.W.); shuxi007@163.com (X.S.)

² School of Chemistry and Chemical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; mhni22@163.com

* Correspondence: yanli@njust.edu.cn

Abstract: Laser-induced breakdown spectroscopy (LIBS) technology has the characteristics of small sample demand, simple sample preparation, simultaneous measurement of multiple elements and safety, which has great potential application in the rapid detection of coal quality. In this paper, 59 kinds of coal commonly used in Chinese power plants were tested by a lab-designed field-portable laser-induced breakdown spectrometer. The data set division methods and the quantitative analysis algorithm of ash content, volatile matter and calorific value of coal samples were carried out. The accuracy and prediction accuracy of three kinds of dataset partitioning methods, random selection (RS), Kennard–Stone (KS) and sample partitioning based on joint X-Y distances (SPXY), coupled with three quantitative algorithms, partial least squares regression (PLS), support vector machine regression (SVR) and random forest (RF), were compared and analyzed in this paper. The results show that the model featuring SPXY combined with RF has the best prediction performance. The R² of ash content by the RF and SPXY method is 0.9843, the RMSEP of ash content is 1.3303 and the mean relative error (MRE) is 7.47%. The R² of volatile matter is 0.9801, RMSEP is 0.7843 and MRE is 2.19%. The R² of calorific value is 0.9844, RMSEP is 0.7324 and MRE is 2.27%. This study demonstrates that the field-portable LIBS device combining appropriate chemometrics algorithms has a wide application prospect in the rapid analysis of coal quality.

Keywords: laser-induced breakdown spectroscopy; portable device; coal quality analysis; data set partitioning; PLSR; SVR; RF

check for
updates

Citation: Dou, Y.; Wang, Q.; Wang, S.; Shu, X.; Ni, M.; Li, Y. Quantitative Analysis of Coal Quality by a Portable Laser Induced Breakdown Spectroscopy and Three Chemometrics Methods. *Appl. Sci.* **2023**, *13*, 10049. <https://doi.org/10.3390/app131810049>

Academic Editor: Federico Marini

Received: 9 August 2023

Revised: 31 August 2023

Accepted: 2 September 2023

Published: 6 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the most widely used energy sources, coal has been known as “black gold” and “industrial food” for a long time. Although the global environmental degradation made the demand to reduce CO₂ emissions from fossil fuel combustion more and more urgent, coal still takes up a prominent role in global energy systems, and the switching from a traditional energy system to a renewable one requires a long time and every government’s effort in the world. At present, the practical way to reduce CO₂ emission is improving the combustion efficiency of boilers in power stations and other industrial plants. The properties of coal, such as carbon content, ash content, volatile matter, and calorific value, etc., have a close relationship with the safety and economy of boiler operation. Although the traditional off-line analysis method has high analysis accuracy, it has the shortcomings of tedious sample pretreatment, complex analysis procedure, long analysis time and the poor timeliness of the detection results, which is not conducive to the real-time control and optimization of the boiler combustion. Low combustion efficiency of the boiler leads to the increase in cost of power generation and delayed guidance of the boiler unit, resulting in

the energy waste. Therefore, the demand for the development of rapid detection technology for on-site coal analysis is urgent to build an adaptation to the combustion instability and parameter fluctuation caused by complex coal quality.

The routine coal quality analysis methods in a coal-fired power plant are based on national standards of coal quality analysis, such as GB/T 213-2008 [1], GB/T 214-2007 [2] and GB/T 30732-2014 [3], which require a long time and tedious procedure to obtain the value of ash, volatile and calorific value for a single sample. The lag in test results makes it impossible to guarantee real-time coal quality control and boiler efficiency adjustment. Laser-induced breakdown spectroscopy (LIBS) [4] is a rapid spectral detection technology developed rapidly in recent years, which has the advantages of small sample demand, simple sample preparation, simultaneous measurement of multiple elements, and safety [5]. At present, it has been widely used in industrial analysis [6,7], environmental monitoring [8], medical analysis [9], agriculture [10], etc., and its application potential in the rapid detection of coal quality has been widely recognized. The working principle of LIBS follows: a high-energy pulsed laser strikes the substance to be measured, causing high temperature and high-density plasma formed on the sample surface. The laser plasma radiates photons of a specific frequency that come from the characteristic spectral lines of elements that existed in the coal constituent, whose frequency and intensity contain the information of element type and concentration of the analyzed object. The spectral data coupled with a variety of chemometrics algorithms can realize the qualitative and quantitative analysis of the sample.

In industrial production, the ash content, the volatile matter, and the calorific value of coal are directly related to the pricing of coal, the efficiency of power generation and the slagging of boilers. Therefore, the rapid detection of these three indicators is the most concerned issue in on-site detection. Coal ash content is mainly composed of minerals, usually in the form of oxides, such as Al_2O_3 , SiO_2 , Na_2O , K_2O , CaO , MgO and Fe_2O_3 . The volatile matter of coal is mainly from water, carbon, hydrogen oxides, hydrocarbons, and some easily ionized mineral decomposition, like Si, Al, Ca, Fe, Mg, etc. Calorific value is closely related to the concentration of elements releasing heat during coal combustion, especially C, H, O, N and other elements. All in all, these three indicators are all related to the elemental constituent of coal. Given the advantages of LIBS in elemental measurement, many scholars have carried out the research on the application of LIBS technology in coal quality analysis and have achieved certain results [11–14]. To improve the accuracy and precision of the LIBS system in coal analysis, great efforts mainly focus on the experimental apparatus setup [15,16] and calibration model improvement [17,18] to enhance the spectral signal-to-noise ratio and obtain better analysis results [19].

In this study, 59 commonly used coal samples in China power plants were tested by a lab-made field-portable LIBS instrument and the quantitative analysis of ash content, volatile matter and calorific value was carried out by three chemometrics models [20,21]: partial least square regression (PLSR) [22,23], supported vector regression (SVR) [24,25], and random forest (RF) [14]. Necessary data pretreatment and data partitioning methods were also established. The performances of the three calibration models were evaluated and further discussed.

2. Materials and Methods

2.1. Instrumentation

The field-portable laser-induced breakdown spectrometer used in this study was design and assembled in our lab. It mainly includes the following: (1) an 8-channel optical fiber spectrometer: Avantes-Mini2048CL spectrometer (200~1100 nm) with the spectral resolution better than 0.1 nm; (2) compact Nd: YAG pulse laser (1064 nm, pulse width ≤ 8 ns, maximum pulse energy 80 mJ); (3) power supply system: 24 V 20 Ah lithium battery and 220 V AC switchable to meet the needs of field testing; (4) optical coupling device; (5) the sample holder with a 32 mm circular groove is used to carry a 30 mm pulverized coal tablet, and the sample table can be rotated along the designated diameter according to a preset speed to achieve multiple tests of the same sample. The whole system weighs less than 10 kg and is easy to carry for field tests.

2.2. Coal Samples

Fifty-nine coal samples commonly used in power plants were selected as test samples, with ash content ranging from 5.36% to 55.54%, volatile content ranging from 16.96% to 39.12%, and calorific value ranging from 12.34 to 27.95 MJ/kg. The coal samples were ground into pulverized coal with a particle size of 200 μm . Then, 3 g of pulverized coal was put into an aluminum mold with a diameter of 30 mm and pressed into a pellet under pressure of 30 MPa. For each coal sample, 60 spectra were recorded which were averaged as the original spectral data of the sample shown in Figure 1.

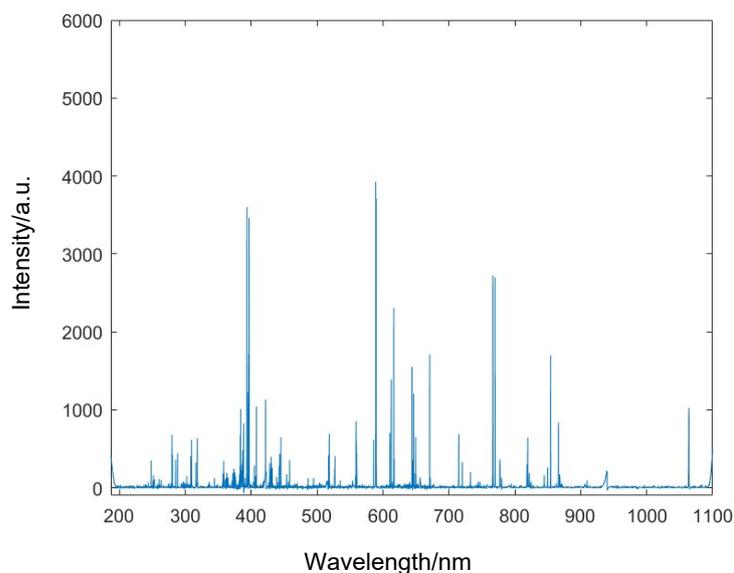


Figure 1. The laser-induced breakdown spectrum of coal sample.

2.3. Spectral Line Selection

Coal ash is the residue left after the high temperature burning of coal, which is mainly composed of minerals, usually in the form of metal oxides and a small amount of sulfocompound, such as Al_2O_3 , SiO_2 , Na_2O , K_2O , CaO , MgO , Fe_2O_3 , etc. The volatile content of coal is mainly composed of carbon, water, hydrocarbons, and Si, Al, Ca, Fe, Mg and other easily ionized metals. The calorific value is closely related to the content of elements that releases heat during coal combustion, especially C, H, O, and inorganic substances and major elements in ash also have a certain impact on the prediction of calorific value. So, the spectral line intensity of C, H, O, Si, Al, Fe, Na, Mg, K, Ca, Ti and Li elements is selected for the quantitative analysis of the ash content, volatile matter and calorific value of coal samples. Table 1 lists the main characteristic spectral lines of the elements in coal selected in this study, using data from the NIST atomic spectrum database [26].

Table 1. Characteristic spectral lines selected for coal industry index analysis.

Element	Wavelength/nm
C	247.8561
H	434.0462; 656.2819; 656.2852
O	407.5862; 441.4905; 777.1944; 777.5388
Si	288.1577
Al	308.2153; 309.2710; 394.4006; 396.1520
Mg	279.5528; 280.2704; 285.2127
Ca	315.8869; 317.9332; 393.3661; 396.8467; 422.6727
Fe	259.9396; 404.5813; 438.35449
K	766.4899; 769.8965
Ti	334.9405; 336.1227; 337.2798; 338.3769
Na	588.9950; 589.5924
Li	460.2898; 670.7775

2.4. Data Partitioning Method

In quantitative analysis, the data set division of training set and testing set influences model performance to some extent. Three partitioning methods, random selection (RS), KS (Kennard–Stone) [27,28] and SPXY (sample partitioning based on joint X-Y distances) [29] were utilized in this paper to compare the performance of the regression model.

The random selection method simply selects a certain number of samples as training sets and the rest as the test set without considering the actual spectral values and physical and chemical values. The key point of this method is to make the selected training samples have overall representativeness. However, if the data set is unevenly distributed and the sample size is small, the selected samples cannot represent the whole sample, which will lead to poor performance of the trained model.

The KS method was proposed by R. W. Kennard and L. A. Tone, which calculates the Euclidean distance between the samples, and samples with large spectral differences are classified as the training set. The KS method ensures the uniformity of spatial distance distribution between selected samples. The samples selected by KS method have good uniformity in spatial distance distribution, which can guarantee the performance of the model. The Euclidean distance $d_{xy}(p, q)$ between $d_x(p, q)$ and $d_y(p, q)$ is formulated as shown in Equation (1),

$$d_{xy}(p, q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2} \quad (1)$$

The SPXY method was first proposed by Galvao et al. and is developed on the basis of the KS method, which considers both the spectral and physicochemical values when calculating the Euclidean distance between samples. $d_x(p, q)$ and $d_y(p, q)$ are divided by the maximum values of the data set $\max_{p, q \in [1, N]}$ in which they are located to ensure that the samples have the same weight in x and y space. Then, the standardized xy distance formula is as shown in Equation (2).

$$d_{xy}(p, q) = \frac{\sqrt{(x_p - x_q)^2}}{\max_{p, q \in [1, N]} d_x(p, q)} + \frac{\sqrt{(y_p - y_q)^2}}{\max_{p, q \in [1, N]} d_y(p, q)} \quad (2)$$

The advantage of the SPXY method lies in the ability to effectively cover the multidimensional vector space, thus improving the predictive ability of the trained model.

In general, the number of training samples and testing samples should be appropriate to both guarantee enough training without either overtraining or undertraining and good prediction performance and the reliability of the model. From an empirical point of view, the overall 59 samples were divided into a training set and testing set according to the ratio of 7:3: namely, 41 samples are selected as the training set and the remaining 18 samples are used as the testing set. In this paper, the influence of three partitioning methods, RS, KS and SPXY, on the performance of the quantitative mode was studied.

2.5. Quantitative Algorithms

In this paper, partial least squares regression (PLSR), support vector machine regression (SVR) and random forest (RF) modeling algorithms are mainly used to establish the corresponding quantitative model. The linear model PLSR is the most commonly used modeling method in quantitative analysis of LIBS, while nonlinear models such as SVR and RF usually have better performance in data mining. It can be seen from the existing work of relevant researchers that the models established by these algorithms have good performance. Therefore, this paper chooses these three methods to model and compares the advantages and disadvantages of each model, so as to choose a better model.

The following parameters are used to evaluate the performance of each model: (1) fitting degree (R^2), (2) root mean square error of calibration (RMSEC), (3) root mean square error of prediction (RMSEP), (4) mean absolute error (MAE), and (5) mean relative error (MRE). R^2 represents the linear relationship between the reference value of the index and

the predicted value of the model. The closer R^2 is to 1, the higher the correlation between the spectral signal and the index content, and the more significant the regression effect. $R^2 > 0.7$ indicates that the data are reliably represented; $R^2 > 0.9$, indicating a good fitting effect. $RMSEC$ and $RMSEP$ are used to measure the deviation between the index reference value and the predicted value of the model, which is to evaluate the performance of the model from the perspective of collation. MAE and MRE represent the mean value of absolute error and the mean value of relative error between the predicted value of the model and the reference value, respectively, which are used to evaluate the accuracy of model prediction from a single measurement. The smaller the four values above, the better the prediction performance of the model.

The above evaluation indexes are defined as

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \tag{3}$$

$$RMSEC = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \tag{4}$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_p} (y_i - \hat{y}_i)^2}{n_p}} \tag{5}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{6}$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \tag{7}$$

where y_i is the true value, \hat{y}_i is the predicted value, \bar{y}_i is the mean of the sample true value, n is the number of training samples, and n_p is the number of testing samples.

3. Results and Discussion

3.1. Predicted Results of PLSR

The principle of PLSR is to build a linear regression model between the input and output data by injecting the input data into a new space, which takes advantage of the multiple linear regression analysis, principal component analysis, and typical correlation analysis. Generally, full spectrum data or partial spectral band data can be used as the input variables of the PLSR model. Full spectrum data contain a great mass of information, but there are also a lot of noise data, leading to the increased iteration time. In order to reduce the interference of irrelevant information and improve the iteration efficiency, the band spectral data containing the main characteristic spectral lines of the index-related elements are chosen as the input variables of the PLSR model shown in Table 2.

Table 2. Input variables of PLSR.

Industrial Index	Elements	Input Variables
Ash content	Si, Al, Mg, K, Ti, Fe, Na, Ca, Li	287.917~288.505; 308.055~308.615; 308.845~309.827; 394.060~394.714; 395.448~396.508; 279.268~279.683; 765.247~766.938; 768.723~770.204; 335.712~336.326; 438.056~438.580; 588.252~590.309; 392.915~393.815; 396.508~397.483; 422.186~423.116; 669.629~671.148
Volatile matter	C, CN, H, O, Al, Mg, Ca, Fe	247.332~248.302; 380.918~390.123; 648.881~661.868; 773.839~779.961; 394.060~394.714; 395.448~396.508; 279.268~279.683; 392.915~393.815; 396.508~397.483; 438.056~438.580
Calorific value	C, C ₂ , CN, H, O, Al, Mg, Ca, Fe, Si	247.332~248.302; 470.238~473.889; 380.918~390.123; 432.243~432.850; 648.881~661.868; 406.911~408.268; 441.046~441.939; 773.839~779.961; 394.060~394.714; 395.448~396.508; 279.268~279.683; 392.915~393.815; 396.508~397.483; 422.186~423.116; 438.056~438.580; 287.917~288.505

The number of principal components of the PLSR model was determined by the remaining one cross-validation. The fitting prediction results of the ash content model are shown in Figure 2 and Table 3.

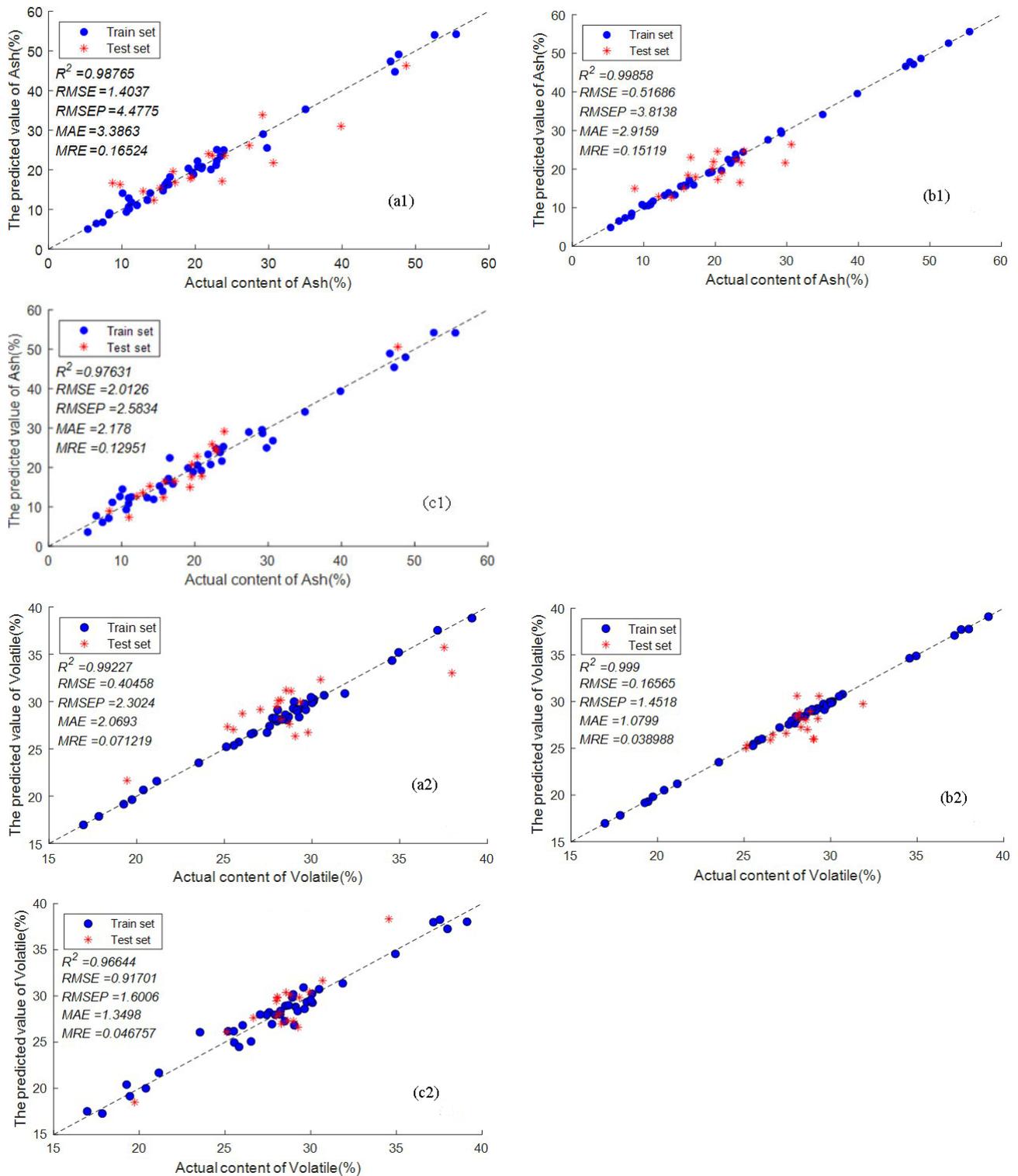


Figure 2. Cont.

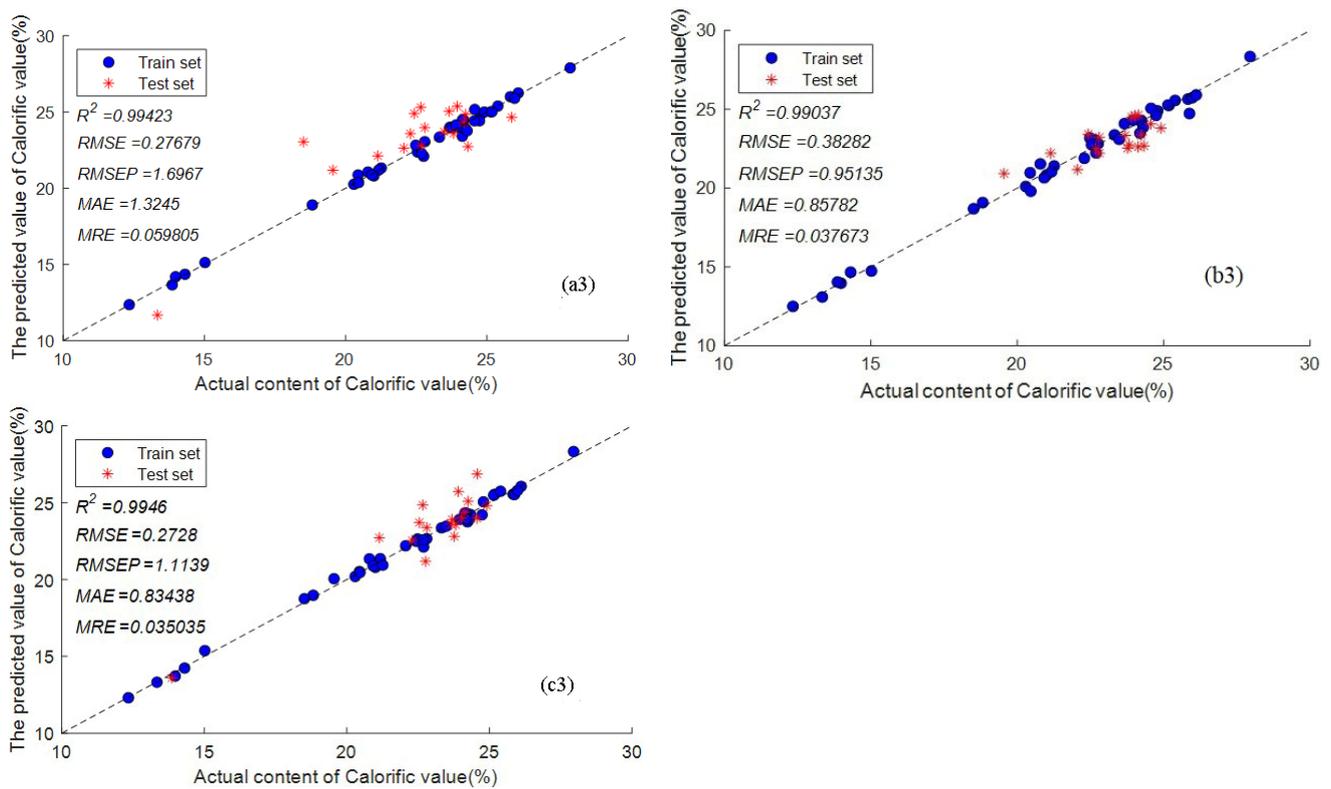


Figure 2. The prediction results of ash content, volatile matter and calorific value of coal by PLSR based on (a1–a3) RS; (b1–b3) KS; (c1–c3) SPXY.

Table 3. The prediction results of PLSR model with different partitioning methods.

Index	Partitioning Method	R ²	RMSEC	RMSEP	MAE	MRE
Ash content	RS	0.9877	1.4037	4.4775	3.3863	0.1652
	KS	0.9986	0.5169	3.8138	2.9159	0.1512
	SPXY	0.9763	2.0126	2.5834	2.1780	0.1295
Volatile matter	RS	0.9923	0.4046	2.3024	2.0693	0.0712
	KS	0.9990	0.1657	1.4518	1.0799	0.0390
	SPXY	0.9664	0.9170	1.6006	1.3498	0.0468
Calorific value	RS	0.9942	0.2768	1.6967	1.3245	0.0598
	KS	0.9904	0.3828	0.9514	0.8578	0.0377
	SPXY	0.9946	0.2728	1.1139	0.8344	0.0350

It can be seen from Figure 3 and Table 3 that the prediction results of the PLSR model with KS and SPXY partitioning methods are better than that with the RS method. It is easy to understand that the data set samples selected by the random method make it difficult to have a good overall representation, while the KS method and the SPXY method select the samples with large spectral differences as the training set based on Euclidean distance and leave the samples with small spectral differences as the test set, which is in essence to perform like cluster analysis. In addition, the KS method is better than the SPXY method in terms of volatile matter and calorific value prediction. The R² of PLSR based on the KS data set partitioning method of volatile matter is 0.9990, the RMSEP of the volatile matter is 1.4518, the MRE is 3.90%, the R² for calorific value regression is 0.9904, the RMSEP of the heating value is 0.9514, and the MRE is 3.77%. However, for the ash content, the model performs relatively poorly: the R² is 0.9763, RMSEP is 2.5834 and MRE is 12.95%.

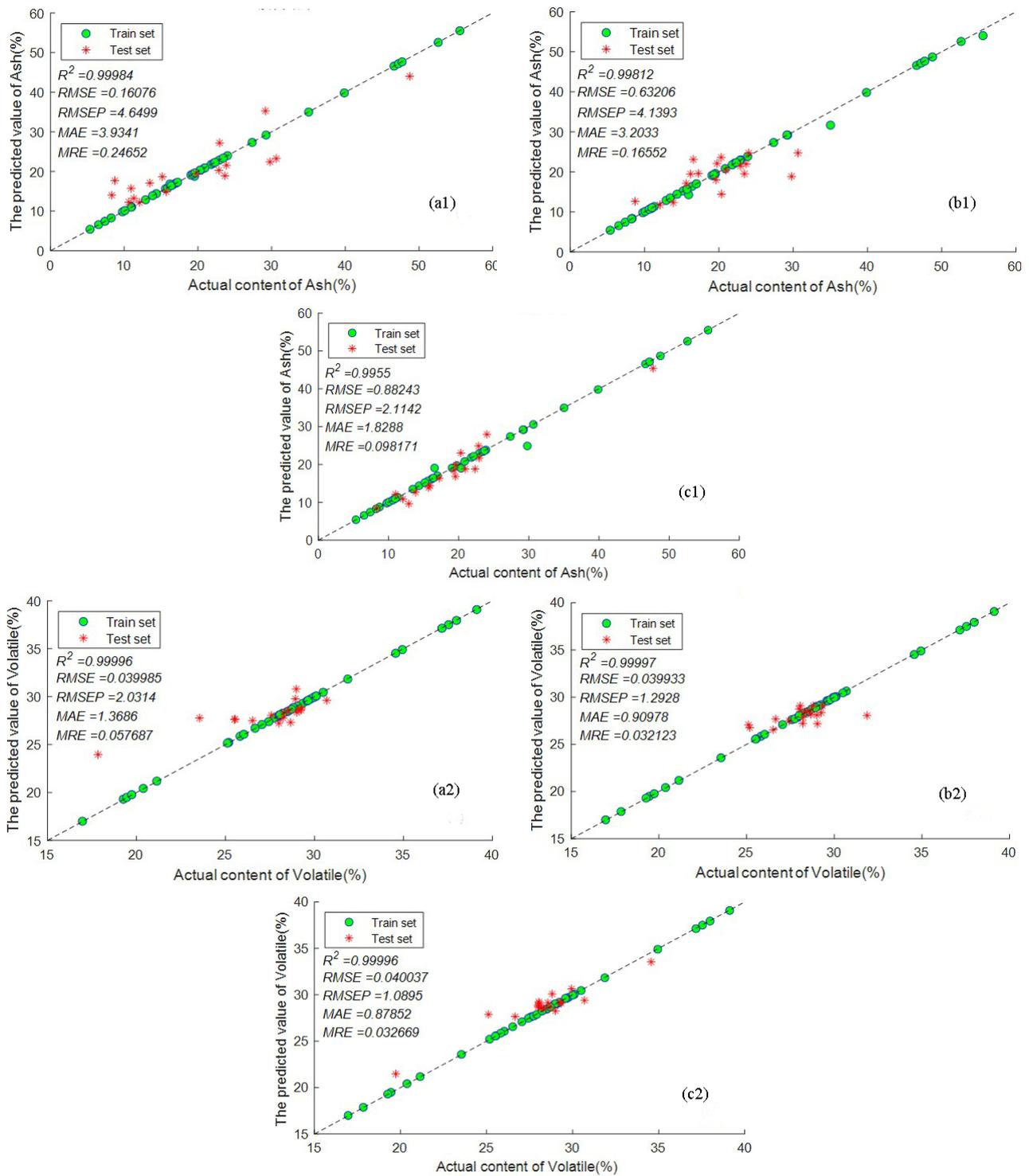


Figure 3. Cont.

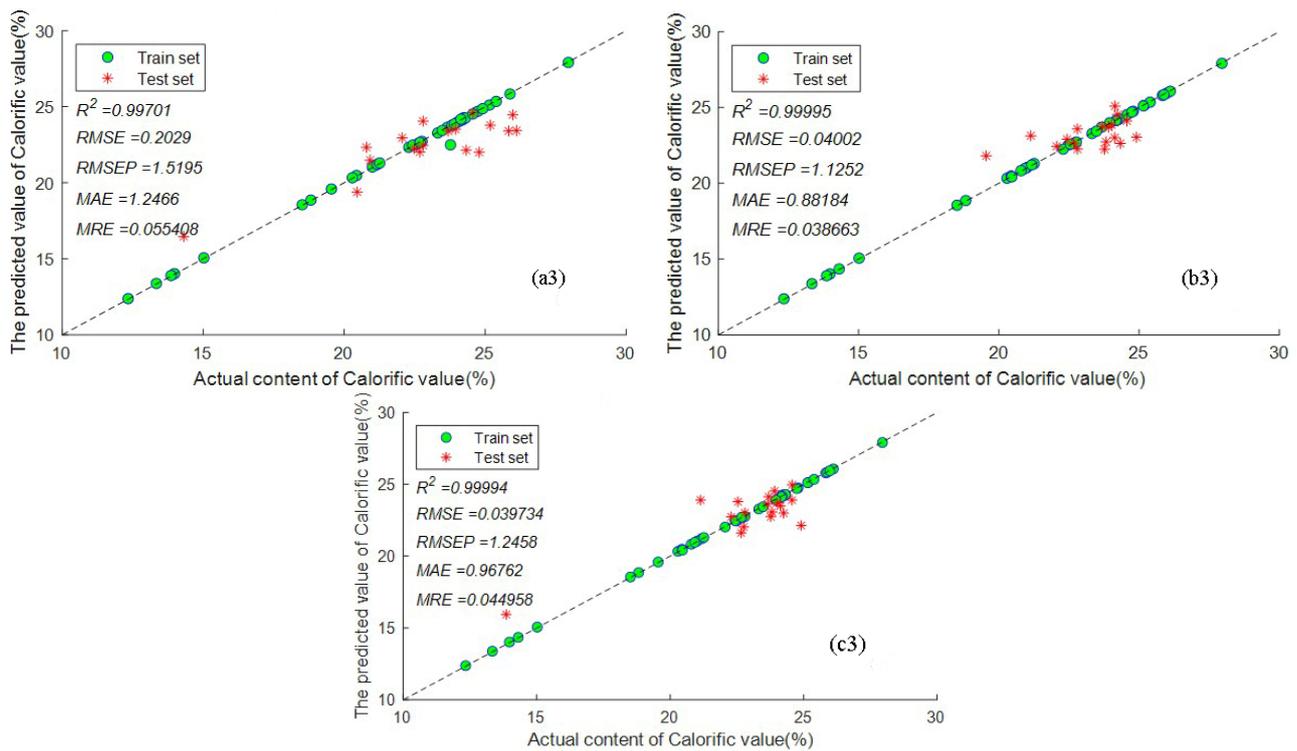


Figure 3. The prediction results of ash content, volatile matter and calorific value of coal by SVR based on (a1–a3) RS; (b1–b3) KS; (c1–c3) SPXY.

3.2. Prediction Results of SVR

The ζ —SVR model using a Gaussian radial basis function (RBF) kernel was established in this paper with parameters of penalty factor C , kernel function parameter g , and parameter ζ controlling the number of support vectors. C and g greatly affect the accuracy of the SVR model for coal quantitative analysis. The size of C represents the degree of punishment to the sample when exceeding the empirical error. The larger the value C is, the higher the degree of punishment will be. However, if the value of C is too large, the overfitting state will occur. Whereas the too small value of C will lead to the underfitting state. The size of g determines the feature space of kernel function and can be used to characterize the complexity of the subspace distribution of sample data. If the value of g is too small, the model will be relatively complex, and the generalization ability cannot be guaranteed. If g is too large, the accuracy of the model is low.

Among the optimization algorithms of C and g , the particle swarm optimization algorithm (PSO) does not need any derivative operation on the solution function and can overcome the non-convergence problem of the traditional optimization algorithms due to the ill condition or singular gradient matrix. Compared with other algorithms, a particle swarm optimization algorithm retains the global search strategy for population, and its unique memory enables it to dynamically track the current search situation and adjust its search strategy. Since the velocity-displacement model adopted by the particle swarm optimization algorithm is simple to operate, it is used in this paper to optimize the parameters of the SVR model [30,31]. The initial parameters of the particle swarm optimization algorithm were set as $c_1 = 1.5$, $c_2 = 1.7$, particle population as 20 and iteration time as 100. Based on the input variables and the optimal model parameters sought by the particle swarm optimization algorithm, the quantitative SVR model of ash content, volatile matter and calorific value under different data set partitioning methods was established. The results are shown in Figure 3 and Table 4.

Table 4. The prediction results of the SVR model under different partitioning methods.

Index	Partitioning Method	R ²	RMSE	RMSEP	MAE	MRE
Ash content	RS	0.9998	0.1608	4.6499	3.9341	0.2465
	KS	0.9981	0.6321	4.1393	3.2033	0.1655
	SPXY	0.9955	0.8824	2.1142	1.8288	0.0982
Volatile matter	RS	0.9999	0.0400	2.0314	1.3686	0.0577
	KS	0.9999	0.0399	1.2928	0.9098	0.0321
	SPXY	0.9999	0.0400	1.0895	0.8785	0.0327
Calorific value	RS	0.9970	0.2029	1.5195	1.2466	0.0554
	KS	0.9999	0.0400	1.1252	0.8818	0.0387
	SPXY	0.9999	0.0397	1.2458	0.9676	0.0450

It can be seen from Figure 3 and Table 4 that the prediction results of the SVR model with KS and SPXY partitioning methods are better than those with the RS method. The R^2 of the SVR model based on the SPXY partitioning method of ash content is 0.9955, $RMSEP$ is 2.1142, and MRE is 9.82%. The R^2 of the volatile matter is 0.9999, $RMSEP$ is 1.0895 and MRE is 2.19%, The R^2 of the calorific value is 0.9999, $RMSEP$ is 1.1252 and MRE is 3.87%.

3.3. Predicted Results of RF

Using the input variables listed in Table 3 and default parameters, the RF quantitative models of ash content, volatile matter and calorific value under different data set partitioning methods were established. The predicted results of the RF model are shown in Figure 4 and Table 5.

Table 5. The prediction results of the RF model under different partitioning methods.

Index	Partitioning Method	R ²	RMSE	RMSEP	MAE	MRE
Ash content	RS	0.9861	1.7693	3.2348	2.4389	0.1348
	KS	0.9886	1.6545	4.1274	3.1410	0.1585
	SPXY	0.9843	1.9235	1.3303	1.1760	0.0747
Volatile matter	RS	0.9818	0.8324	1.9561	1.4596	0.0515
	KS	0.9813	0.8541	1.3220	0.9841	0.0354
	SPXY	0.9801	0.9540	0.7843	0.6001	0.0219
Calorific value	RS	0.9815	0.5646	1.8893	1.5568	0.0790
	KS	0.9849	0.6379	1.0612	0.7843	0.0354
	SPXY	0.9844	0.6381	0.7324	0.4941	0.0227

It can be seen from Figure 4 and Table 5 that the prediction result of the RF model with the SPXY partitioning method is the best. The R^2 of ash content is 0.9843, the $RMSEP$ is 1.3303 and the mean relative error (MRE) is 7.47%. The R^2 of volatile matter is 0.9801, $RMSEP$ is 0.7843 and MRE is 2.19%; the R^2 of calorific value is 0.9844, $RMSEP$ is 0.7324 and MRE is 2.27%.

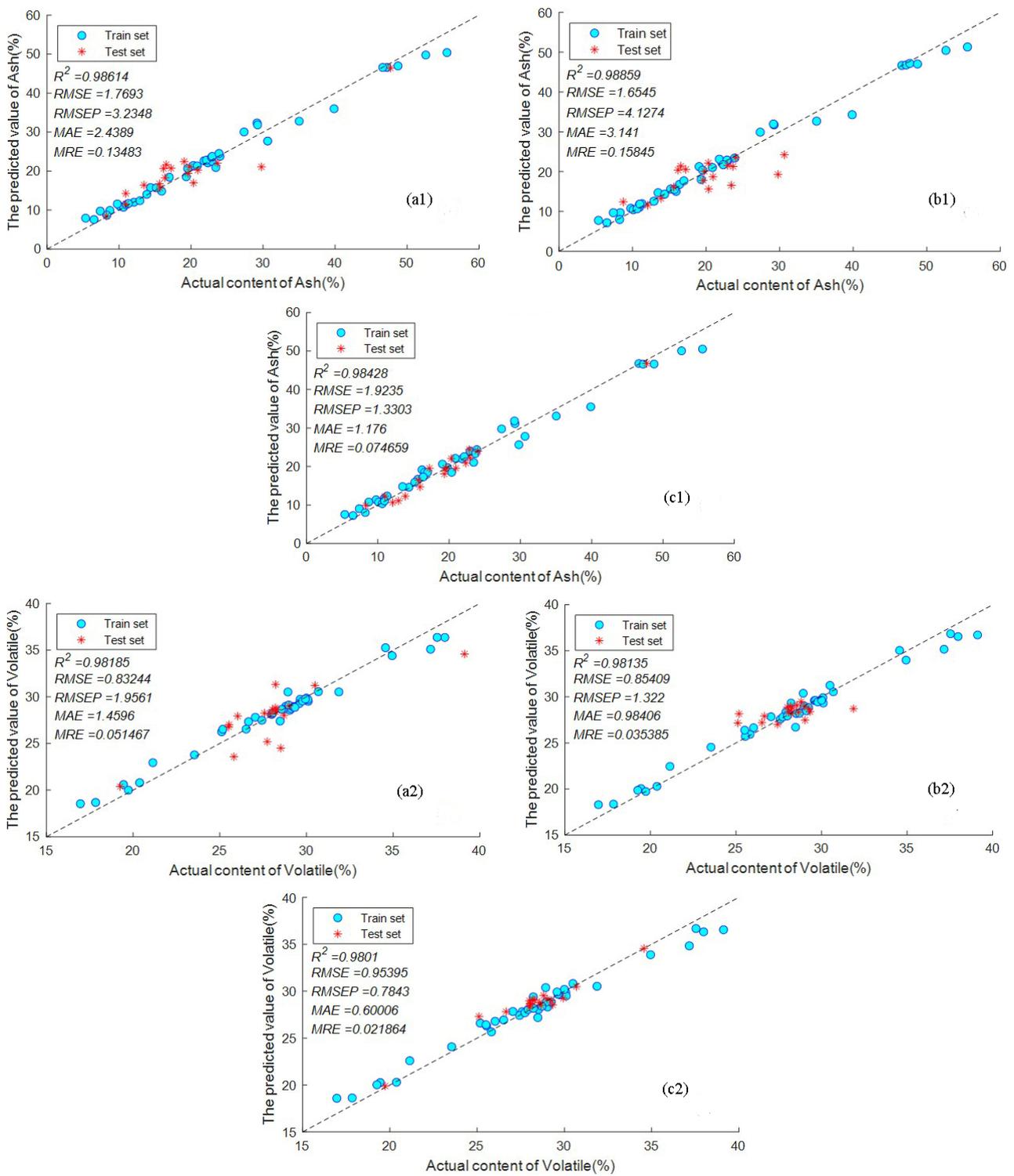


Figure 4. Cont.

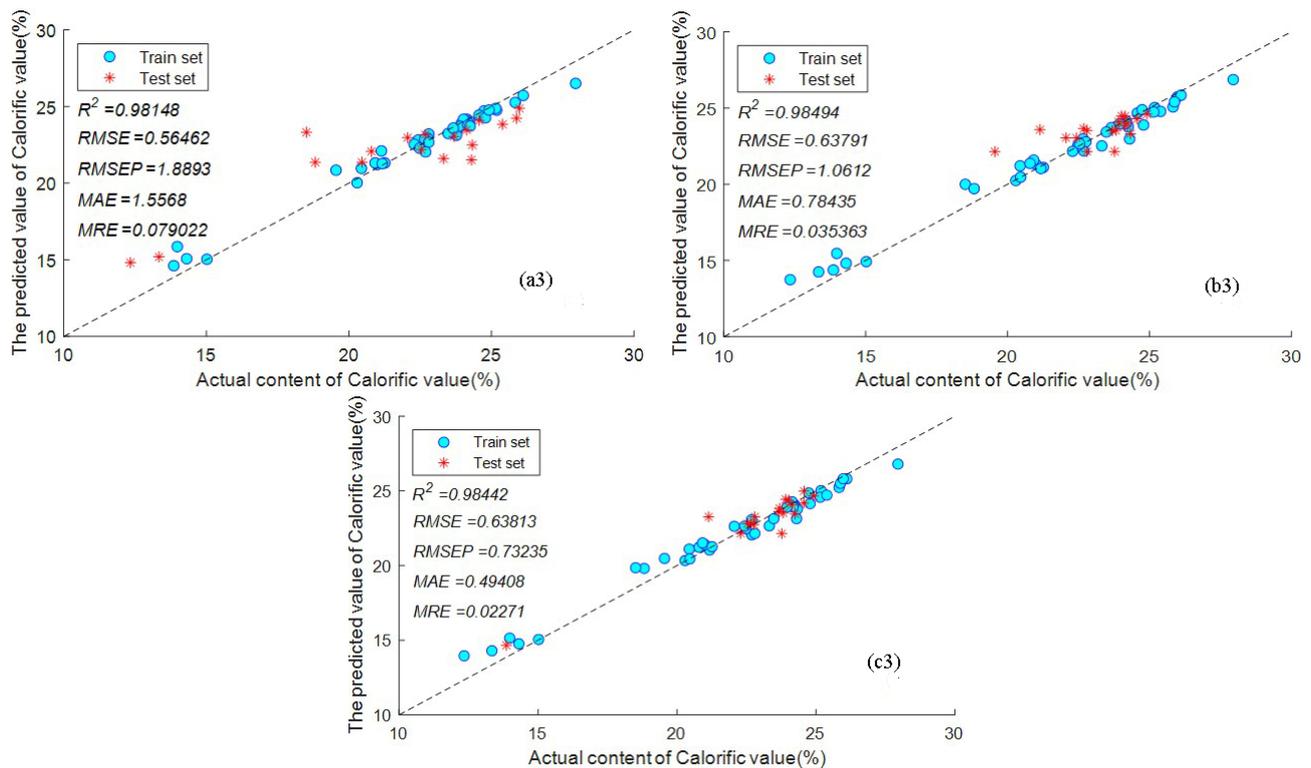


Figure 4. The prediction results of ash content, volatile matter and calorific value of coal by RF based on (a1–a3) RS; (b1–b3) KS; (c1–c3) SPXY.

4. Conclusions

The rapid detection and quantitative analysis of coal quality can guide the adjustment of boiler combustion conditions in real time, which is of great significance to promote the clean and efficient utilization of coal and the energy conservation and CO₂ emission reduction in the thermal power industry. Laser-induced breakdown spectroscopy (LIBS) technology has the characteristics of small sample demand, simple sample preparation, and simultaneous measurement of multiple elements, which has great application potential in the rapid analysis of coal quality. In this paper, the field-portable laser-induced breakdown spectrometer coupled with three chemometrics methods are used to quantitatively analyze the ash content, volatile content and calorific value of coal commonly used in China power plants. The LIBS data set partitioning methods, random selection (RS), Kennard–Stone (KS), and sample partitioning based on joint X–Y distances (SPXY), and the quantitative models, partial least square regression (PLSR), support vector regression (SVR) and random forest (RF), were combined in this paper to establish the appropriate quantitative model for coal industrial indicator prediction. The model performance and the prediction accuracy of the methods are systematically analyzed and compared. The results show that SPXY data set partitioning combined with the RF model has better prediction performance than that of the PLSR and SVR models coupled with RS and KS methods. The R^2 of ash content by the RF and SPXY method is 0.9843, the RMSEP of ash content is 1.3303 and the mean relative error is 7.47%. The R^2 of volatile matter is 0.9801, RMSEP is 0.7843 and MRE is 2.19%. The R^2 of calorific value is 0.9844, RMSEP is 0.7324 and MRE is 2.27%. The study demonstrates that laser breakdown spectroscopy (LIBS) coupled with the sample partitioning method based on joint X–Y distances (SPXY) and random forest (RF) model could be a prosperous way to attain the on-site quick analysis of the ash content, volatile and calorific value. The field-portable laser-induced breakdown spectrometer used in this study coupled with the chemometrics method can rapidly and simultaneously analyze the ash, volatile and calorific value using a single instrument. The key point of the application of this method is the effective data preprocessing and quantitative algorithm optimization which requires a

large amount of sample testing and training. And the consistency of the sample surface also influences the predicting result, which makes it necessary to keep the surface smooth and consistent for all samples during testing. On this basis, this technology can be used as a real-time and rapid detection method for coal for a variety of occasions, such as coal mines, coal transportation, coal enterprises, etc.

Author Contributions: Conceptualization, Y.D. and Y.L.; methodology, Q.W. and X.S.; software, S.W., M.N. and Y.L.; writing—original draft preparation, Y.D.; writing—review and editing, Y.D., Q.W., S.W., X.S., M.N. and Y.L.; supervision, Y.L.; project administration, Y.D.; funding acquisition, Y.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the China National Energy Group Science and Technology Project (Project No. GJNY-22-90) and National Natural Science Foundation of China (Project No. 516761000).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Acknowledgments: We want to thank all participants and assistants for their support.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. GB/T 213-2008; Determination of Calorific Value of Coal. Standardization Administration of People's Republic of China: Beijing, China, 2008.
2. GB-T 214-2007; Determination of Total Sulfur in Coal. Standardization Administration of People's Republic of China: Beijing, China, 2007.
3. GB/T 30732-2014; Proximate Analysis of Coal by Instrumental Method. Standardization Administration of People's Republic of China: Beijing, China, 2014.
4. Miziolek, A.W.; Palleschi, V.; Schechter, I. *Laser Induced Breakdown Spectroscopy; (LIBS): Fundamentals and Applications*; Cambridge University Press: Cambridge, UK, 2006.
5. Zhang, Y.J.; Xiong, Z.; Ma, Y.W.; Zhu, C.W.; Zhou, R.; Li, X.Y.; Li, Q.; Zeng, Q.D. Quantitative Analysis of Coal Quality by Laser-induced Breakdown Spectroscopy Assisted with Different Chemometric Methods. *Anal. Methods* **2020**, *12*, 3530–3536. [[CrossRef](#)] [[PubMed](#)]
6. Gottfried, J.L.; Harmon, R.S.; Lucia, F.C.; Miziolek, A.W. Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geomaterial classification. *Spectrochim. Acta Part B* **2009**, *64*, 1009–1019. [[CrossRef](#)]
7. Aragon, C.; Aguilera, J.A. Direct analysis of aluminum alloys by CSigma laser induced breakdown spectroscopy. *J. Anal. Chim. Acta* **2018**, *1009*, 12–19. [[CrossRef](#)] [[PubMed](#)]
8. Bhatt, B.; Angeyo Kalambuga, H.A.; Dehayem-Kamadjeu, A. LIBS development methodology for forensic nuclear materials analysis. *Anal. Methods* **2018**, *10*, 791–798. [[CrossRef](#)]
9. Brunnbauer, L.; Gajarska, Z.; Lohninger, H.A. Limbeck A critical review of recent trends in sample classification using Laser-Induced Breakdown Spectroscopy (LIBS). *Trend Anal. Chem.* **2023**, *159*, 116859. [[CrossRef](#)]
10. Rajavelu, H.; Vasa, N.J.; Seshadri, S. Effect of ambiance on the coal characterization using laser-induced breakdown spectroscopy (LIBS). *Appl. Phys. A* **2020**, *126*, 395. [[CrossRef](#)]
11. Liu, K.; He, C.; Zhu, C.W.; Chen, J.; Zhan, K.P.; Li, X.Y. A review of laser-induced breakdown spectroscopy for coal analysis. *Trend Anal. Chem.* **2021**, *143*, 116357. [[CrossRef](#)]
12. Iqbal, S.M.Z.; Uddin, Z.; Umar, Z.A.; Ahmed, N.; Ahmed, R.; Baig, M.A. Analysis of Lakhra Coal by Calibration Free Laser-Induced Breakdown Spectroscopy (CF-LIBS) and Comparison of Self-Absorption Correction Procedures. *Anal. Lett.* **2022**, *55*, 11–23. [[CrossRef](#)]
13. Sheta, S.; Afgan, M.S.; Hou, Z.; Yao, S.C.; Zhang, L.; Li, Z.; Wang, Z. Coal analysis by laser-induced breakdown spectroscopy: A tutorial review. *J. Anal. Atom. Spectrom.* **2019**, *34*, 1047–1082. [[CrossRef](#)]
14. Legnaioli, S.; Campanella, B.; Pagnotta, S.; Poggialini, F.; Palleschi, V. Determination of Ash Content of coal by Laser-Induced Breakdown Spectroscopy. *Spectrochim. Acta Part B* **2019**, *155*, 123–126. [[CrossRef](#)]
15. Ma, Y.W.; Zhang, W.; Xiong, Z.; Cui, H.; Li, Q.; Zhou, R.; Zhang, Y.; Li, X.; Zeng, X.; Li, Q. Accurate sulfur determination of coal using double-pulse laser-induced breakdown spectroscopy. *J. Anal. Atom. Spectrom.* **2020**, *35*, 1458–1463. [[CrossRef](#)]

16. Petrovic, J.; Savovic, J.; Rankovic, D.; Kuzmanovic, M. Quantitative Analysis of Coal by Laser-Induced Breakdown Spectroscopy Using TEA CO₂ Laser as the Excitation Source. *Plasma Chem. Plasma Process.* **2022**, *42*, 519–533. [[CrossRef](#)]
17. Zhang, W.H.; Zhuo, Z.; Lu, P.; Tang, J.; Tang, H.; Lu, J.; Xing, T.; Wang, Y. LIBS analysis of ash content, volatile matter, and calorific value in coal by partial least squares regression based on ash classification. *J. Anal. Atom. Spectrom.* **2020**, *35*, 1621–1631. [[CrossRef](#)]
18. Wang, W.; Kong, W.; Shen, T.; Man, Z.; Zhu, W.; He, Y.; Liu, F. Quantitative analysis of cadmium in rice roots based on LIBS and chemometrics methods. *Environ. Sci. Eur.* **2021**, *33*, 37. [[CrossRef](#)]
19. Idris, N.; Pardede, M.; Jobiliong, E.; Lie, Z.S.; Hedwig, R.; Suliyanti, M.M.; Kurniawan, D.P.; Kurniawan, K.H.; Kagawa, K.; Tjia, M.O. Enhancement of carbon detection sensitivity in laser induced breakdown spectroscopy with low pressure ambient helium gas. *Spectrochim. Acta Part B* **2019**, *151*, 26–32. [[CrossRef](#)]
20. Palleschi, V. *Chemometrics and Numerical Methods in LIBS*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2022.
21. Ewusi-Annan, E.; Delapp, D.M.; Wiens, R.C.; Melikechi, N. Automatic preprocessing of laser-induced breakdown spectra using partial least squares regression and feed-forward artificial neural network: Applications to Earth and Mars data. *Spectrochim. Acta Part B* **2020**, *171*, 105930. [[CrossRef](#)]
22. Li, X.; Yang, Y.; Li, G.; Chen, B.; Hu, W. Accuracy improvement of quantitative analysis of calorific value of coal by combining support vector machine and partial least square methods in laser-induced breakdown spectroscopy. *Plasma Sci. Technol.* **2020**, *22*, 074014. [[CrossRef](#)]
23. Sánchez-Esteva, S.; Knadel, M.; Kucheryavskiy, S.; de Jonge Lis, W.G.; Rubæk, H.; Hermansen, C.; Heckrath, G. Combining Laser-Induced Breakdown Spectroscopy (LIBS) and Visible Near-Infrared Spectroscopy (Vis-NIRS) for Soil Phosphorus Determination. *Sensors* **2020**, *20*, 5419. [[CrossRef](#)]
24. Andres, J. Retraction: Novel techniques for enhancing the performance of support vector regression chemo-metric in quantitative analysis of LIBS spectra. *J. Anal. Atom. Spectrom.* **2019**, *34*, 2477. [[CrossRef](#)]
25. Zhang, Y.; Dong, M.; Cheng, L.; Wei, L.; Cai, J.; Lu, J. Improved measurement in quantitative analysis of coal properties using laser induced breakdown spectroscopy. *J. Anal. Atom. Spectrom.* **2020**, *35*, 810–818. [[CrossRef](#)]
26. Kramida, A.; Ralchenko, Y.; Reader, J. NIST Atomic Spectra Database (ver. 5.10). 2022. Available online: <https://physics.nist.gov/asd> (accessed on 3 January 2023).
27. Kennard, R.W.; Stone, L.A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148. [[CrossRef](#)]
28. Morais, C.L.M.; Santos, M.C.D.; Lima, K.M.G.; Martin, F.L. Improving data splitting for classification applications in spectrochemical analyses employing a random-mutation Kennard-Stone algorithm approach. *Bioinformatics* **2019**, *35*, 5257–5263. [[CrossRef](#)] [[PubMed](#)]
29. Wu, W.; Walczak, B.; Massart, D.L. Artificial neural networks in classification of NIR spectral data: Design of the training set. *Chemom. Intell. Lab. Syst.* **1996**, *33*, 35–46. [[CrossRef](#)]
30. Gad, A.D. Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. *Arch. Comput. Methods Eng.* **2022**, *29*, 2531–2561. [[CrossRef](#)]
31. Huang, S.H.; Tian, L.; Zhang, J.S.; Chai, X.L.; Wang, H.L.; Zhan, H.L. Support Vector Regression Based on the Particle Swarm Optimization Algorithm for Tight Oil Recovery Prediction. *ACS Omega* **2021**, *6*, 32142–32150. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.