applied sciences

MDPI

Article

Application of Laser-Induced Breakdown Spectroscopy Combined with Chemometrics for Identification of Penicillin Manufacturers

Kai Wei [1,2], Qianqian Wang [1,2,3,*], Geer Teng [1,2,3], Xiangjun Xu [1,2,3], Zhifang Zhao [1,2] and Guoyan Chen [1,2]

1 School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China; 3120170346@bit.edu.cn (K.W.); 3120185341@bit.edu.cn (G.T.); 3120205322@bit.edu.cn (X.X.); 3120215335@bit.edu.cn (Z.Z.); 3120190515@bit.edu.cn (G.C.)
2 Key Laboratory of Photonic Information Technology, Ministry of Industry and Information Technology, Beijing Institute of Technology, Beijing 100081, China
3 Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing 314033, China
* Correspondence: qqwang@bit.edu.cn

Abstract: Due to the differences in raw materials and production processes, the quality of the same type of drug produced by different manufacturers is different. In drug supervision, determining the manufacturer can help to trace drug quality issues. In this study, a method for the quick identification of drug manufacturers based on laser-induced breakdown spectroscopy (LIBS) was proposed for the first time. We obtained the LIBS spectra from 12 samples of three types of penicillin (phenoxymethylpenicillin potassium tablets, amoxicillin capsules, and amoxicillin and clavulanate potassium tablets) produced by 10 manufacturers. The LIBS characteristic lines of the three types of penicillin were ranked by importance based on the decrease in the Gini index of random forest (RF). Three classifiers—the linear discriminant analysis (LDA), support vector machine (SVM) and artificial neural network (ANN)—were used to identify the different manufacturers of the three types of penicillin. RF-ANN provided the best classification result and an accuracy of 100% in penicillin manufacturer identification. The results show that LIBS combined with chemometrics could be used in the identification of penicillin manufacturers, and this method has application potential in drug quality supervision.

Keywords: laser-induced breakdown spectroscopy; qualitative analysis; chemometric methods

Citation: Wei, K.; Wang, Q.; Teng, G.; Xu, X.; Zhao, Z.; Chen, G. Application of Laser-Induced Breakdown Spectroscopy Combined with Chemometrics for Identification of Penicillin Manufacturers. *Appl. Sci.* **2022**, *12*, 4981. https://doi.org/10.3390/app12104981

Academic Editor: Alberto Milani

Received: 8 April 2022
Accepted: 11 May 2022
Published: 14 May 2022

1. Introduction

Penicillins are an important group of antibiotics widely used in clinical applications as antibacterial agents or bactericides for their high efficiency and low toxicity. Their successful development has greatly improved the ability of humans to fight bacterial infections. Various penicillins have been used in many situations [1]. Many factors can cause quality differences for penicillin. There are many penicillin manufacturers in China, and the same kind of penicillin produced by different manufacturers also shows quality differences, which are caused by differences in raw materials or other factors. In order to ensure the medicinal effects of penicillin, the safety of people and the development of the whole pharmaceutical industry, it is necessary to accurately identify the quality of penicillin, prevent counterfeiting and low-quality products, and ensure the uniformity of varieties and efficacy. It is of great significance to study the identification of drug manufacturers for drug quality control, drug inspection and drug supervision [2].

There are several techniques that can be used to detect pharmaceuticals, such as high-performance liquid chromatography (HPLC), DNA barcoding, near-infrared spectroscopy (NIRS) and Raman spectroscopy. HPLC is mostly used to analyze differences in medicinal components, such as fritillaria alkaloids, and has certain advantages for fundamentally

identifying the similarity of the pharmacological effects [3]. DNA barcoding is suitable for large-scale species identification. Efficient and standardized operation procedures are conducive to the construction of whole-species barcode information libraries [4]. The sample processing for HPLC and DNA barcoding is complex and time consuming, and both techniques are commonly performed in the laboratory [5].

NIRS is a fast technique that is well suited for the detection of pharmaceuticals [6]. Portable NIRS scanners have been used to detect drugs [7]. Scafi et al. tested a variety of drugs to explore the potential of NIRS, which showed that NIRS can be used to identify drugs, even if the composition differences are very small [8]. Raman spectroscopy is often used to identify molecules in chemistry [9]. It can also be used for the rapid analysis of genuine and counterfeit drugs [6]. This technique does not require sample preparation. De et al. used Raman spectroscopy to distinguish genuine and fake artesunate samples [10]. The spectral signals of NIRS and Raman spectroscopy are weak, and they are easily affected by background light [11].

Laser-induced breakdown spectroscopy (LIBS) is a technique based on atomic emission spectroscopy [12,13], and it can be used to analyze all material states (gas, liquid and solid). It allows for the fast, real-time, in situ, and simultaneous detection of multiple elements and is thus widely used in areas including the detection of explosives [14,15], biomedical analysis [16,17], heavy metal detection in soil [18,19], geological analysis [20,21], coal quality detection [22,23] and food safety analysis [24,25].

In drug testing, scholars have performed some studies using the LIBS technique. Aldakheel et al. analyzed the contents of nutritional, harmful and pharmacologically active elements in Rhatany roots, using calibration-free laser-induced breakdown spectroscopy (CF-LIBS), and verified the values obtained using inductively coupled plasma optical emission spectroscopy (ICP-OES). The results from the latter method agree with the LIBS data, showing relatively high accuracy [26]. Dastgeer et al. used LIBS to qualitatively and semi-quantitatively analyze calcium tablets. They identified Ca, Mg, Fe and Zn in the samples, and determined the approximate content of each element using the intensity of the respective spectral peak [27]. Doucet et al. used LIBS combined with chemometrics to predict formulation excipients and active pharmaceutical ingredients in complex pharmaceutical formulations. The results show that the relative deviation of magnesium stearate was less than 4% and that of lactose was less than 15% [28]. Wei et al. introduced learning vector quantization (LVQ) in LIBS for the first time to differentiate *Fritillaria cirrhosa* from non-*Fritillaria cirrhosa*. The classification accuracy was 99.17% [5]. Wang et al. used LIBS and the hyperspectral technique combined with partial least squares discriminant analysis (PLS-DA) to determine the species, origins and ages of ginseng samples. The classification accuracy was above 93%, 94% and 99%, respectively [29]. Zhang et al. applied LIBS with principal component analysis (PCA), linear discriminant analysis (LDA) and support vector machine (SVM) to the spectral analysis of ginkgo leaves at eight different locations in Xi'an. The first 30 principal components extracted using PCA were used as new input variables for identification with LDA and SVM. The classification accuracy was 97.50% and 96.25%, respectively [30]. Zheng et al. used LIBS technology combined with the random forest (RF) algorithm to perform grade classification for Dendrobium. The classification accuracy was 96.46% [31]. However, to the best of our knowledge, there has not yet been any study that has investigated and analyzed chemically synthesized or naturally derived drugs using LIBS combined with machine learning. Considering penicillins as an example, we attempted to identify the different penicillin manufacturers.

In this paper, we propose LIBS technology as a potential tool for identifying penicillin manufacturers for the first time. A total of 12 samples of three types of penicillin (each type produced by four manufacturers) were used in this experiment. LIBS was adopted with feature selection and three commonly used classifiers (LDA, SVM and ANN) to identify penicillin manufacturers. The purpose of this study was to test the feasibility of using LIBS combined with machine learning in penicillin manufacturer identification.

2. Materials and Methods

2.1. LIBS Experimental Setup

The schematic of the experimental setup for LIBS-based penicillin sample detection is shown in Figure 1. We used a He–Ne laser (λ = 632.8 nm) to calibrate the optical path. Laser pulses were generated by using a Q-switched Nd:YAG laser (λ = 1064 nm, pulse frequency f = 1 Hz, beam diameter ϕ = 6 mm, pulse duration τ = 5 ns, and pulse energy E = 40 mJ/pulse). We used three plane mirrors (M1, M2 and M3) to modulate the propagation direction of the laser beam, and the laser beam was focused on the sample surface by using a convex lens with a 100 mm focal length. The penicillin sample was fixed on a three-dimensional motorized stage. The laser interacted with the sample to produce plasma radiation, which was collected in a 600 µm-diameter optical fiber using a lens with a 36 mm focal length. The optical fiber was connected to a dual-channel spectrometer (AvaSpec 2048-2-USB2, Avantes, Apeldoorn, The Netherlands). The resolution of the spectrometer was 0.2–0.3 nm. The wavelength range was 190–1100 nm. The spectrometer was integrated with a charge-coupled device (CCD) camera. A detector was used to detect the laser pulse and trigger the delay generator (DG535). DG535 triggered the CCD after a preset time, and the spectral acquisition delay time was set to 1.28 µs to reduce continuous radiation. The integration time was 2 ms. The Avasoft 7.6 software was used to control the collection of the penicillin LIBS spectra by the spectrometer.

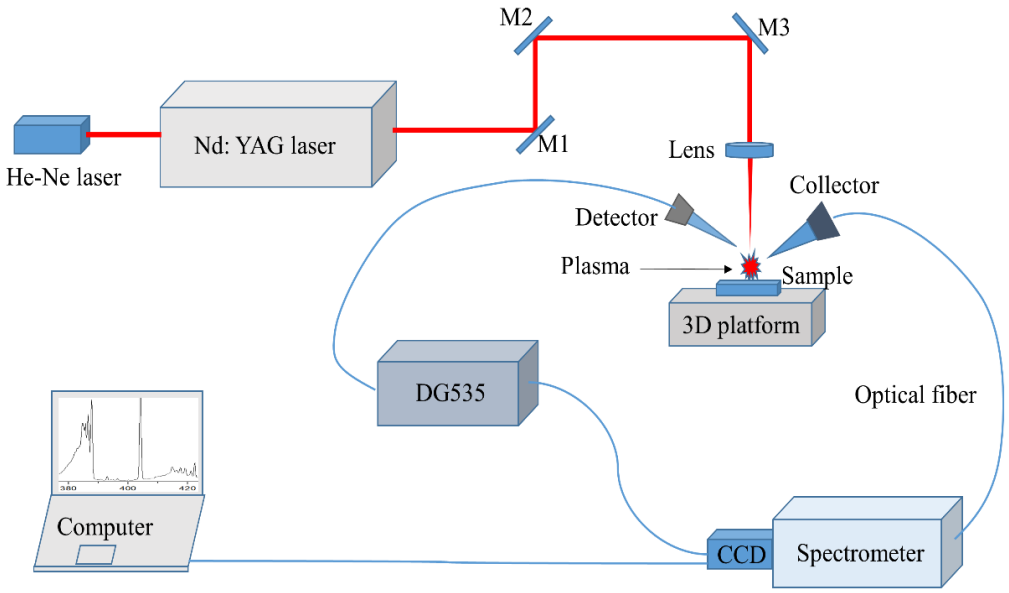

Figure 1. A schematic of the LIBS experimental setup for penicillin sample detection.

2.2. Sample Preparation and Measurement of Penicillin Samples

For each of the three types of penicillin (phenoxymethylpenicillin potassium tablets, PPT; amoxicillin capsules, AC; and amoxicillin and clavulanate potassium tablets, ACPT) used in this experiment, four samples from different manufacturers were purchased in two batches from Beijing Youkangtang Pharmacy Co., Ltd., Beijing, China. For both batches, one box of penicillin was purchased from each manufacturer. The type, molecular formula and manufacturer information of each penicillin sample are listed in Table 1. PPT is available in the form of tablets with a single active ingredient. AC is available in the form of capsules with only one active ingredient. ACPT is available in the form of tablets, and it is a combination of amoxicillin and clavulanate potassium (CP). In the experiment, PPT and ACPT tablets were placed directly on the three-dimensional motorized stage. Each tablet was first ablated with 10 laser pulses at a single location to remove the sugar coating on its surface. Forty LIBS spectra were then collected from each tablet. For each AC capsule, the shell was first removed. Then, 0.5 g of the penicillin powder inside was weighed out on a

piece of weighing paper using an electronic balance (Sartorius, BSA124S-CW, Göttingen, Germany). It was then made into a round pellet with a diameter of ϕ 13 mm and a thickness of 2 mm using a hydraulic machine (HY-12, Tianjin Tianguang Optical Instruments Co., Ltd., Tianjin, China) by applying 20 MPa of pressure for 5 min. The AC pellet was placed on the three-dimensional motorized stage to collect 40 LIBS spectra from one tablet at different positions. In a batch of samples, seven tablets from each manufacturer were selected and their LIBS spectra were measured. These spectra were used as a training set to establish classification models. In the other batch of samples, three tablets were selected from each manufacturer and their spectra were measured. These spectra were used to test the classification models.

Table 1. The type, molecular formula and manufacturer of each penicillin sample.

Type	Molecular Formula	Sample Number	Manufacturer
PPT	$C_{16}H_{17}KN_2O_5S$	PPT1	Chongqing Kerui Pharmaceutical (Group) Co., Ltd., Chongqing, China
		PPT2	North China Pharmaceutical Group Co., Ltd., Shijiazhuang, China
		PPT3	Hainan Sanye Pharmaceutical Factory Co., Ltd., Haikou, China
		PPT4	Taiji Southwest Pharmaceutical Co., Ltd., Chongqing, China
AC	$C_{16}H_{19}N_3O_5S{\cdot}3H_2O$	AC1	Guangzhou Baiyunshan Pharmaceutical Holdings Co., Ltd., Baiyunshan Pharmaceutical General Factory, Guangzhou, China
		AC2	Zhongnuo Pharmaceutical (Shijiazhuang) Co., Ltd., Shijiazhuang, China
		AC3	North China Pharmaceutical Group Co., Ltd., Shijiazhuang, China
		AC4	Chongqing Dikang Changjiang Pharmaceutical Co., Ltd., Chongqing, China
ACPT	$C_{16}H_{19}N_3O_5S{\cdot}3H_2O$ $C_8H_8KNO_5$	ACPT1	Guangzhou Baiyunshan Pharmaceutical Holdings Co., Ltd., Baiyunshan Pharmaceutical General Factory, Guangzhou, China
		ACPT2	Zhuhai United Laboratories (Zhongshan) Co., Ltd., Zhuhai, China
		ACPT3	Hainan Simcere Pharmaceutical Co., Ltd., Haikou, China
		ACPT4	Santa (Zhangjiakou) Pharmaceutical Co., Ltd., Zhangjiakou, China

2.3. Data Analysis of Penicillin Samples

2.3.1. Spectral Selection and Data Preprocessing of Penicillin Samples

Some LIBS spectra could have been different from other spectra due to the variation of the laser pulse energy and the heterogeneity of the sample. These abnormal spectra would not have represented the characteristics of the sample and, if included in the construction of a classification model, would have negatively impacted the classification accuracy of the model. Therefore, before building the classification model, the LIBS spectral data were preprocessed for each sample to eliminate any outliers in the raw dataset.

In this study, the LIBS spectra of each sample were inspected for their cosine similarity to improve the classification performance in later steps. The LIBS spectra of one batch of samples were used as the training set. Seven tablets were selected from each sample, and forty LIBS spectra were collected for each tablet. The LIBS spectra of each sample were treated as vectors in the multi-dimensional space, and we calculated the cosine value of the angle between each spectrum and the mean spectrum. A larger cosine value indicated greater similarity between the specific and mean spectra. The spectra were sorted based on the ascending order of cosine values, and the most similar 50 percent of spectra were selected [16]. For every two spectra, the mean spectrum was computed. In the end, 70 spectra were obtained for each sample, which were used as the training set. The LIBS spectra of the other batch of samples were used as the test set. Three tablets were selected from each sample, and 40 spectra were obtained for each tablet. We preprocessed the spectra in the same way as the training set; 30 spectra were obtained for each sample in the end, and they constituted the testing set.

In LIBS measurement, the fluctuation of the laser pulse energy can affect the stability of the laser beam. There is a certain degree of fluctuation between the spectra, which affects

the repeatability of the collected penicillin LIBS spectra, and it in turn affects subsequent analysis and identification. To reduce the influence of the above fluctuation on the penicillin LIBS spectra, it is necessary to normalize the obtained penicillin LIBS spectra. Maximum normalization was used in this study. For each spectrum, first, we calculated the integrated intensity of each spectral line, and then, we divided the integrated intensity of all the spectral lines by the maximum integrated intensity among the spectral lines. Maximum normalization can effectively reduce the fluctuation between the spectra and reduce the difference between the same type of spectra. Maximum normalization is also a commonly used normalization method in LIBS spectral preprocessing [32–34], but there might be a better choice. Our purpose was not to compare the preprocessing methods, but to compare the performance of classification algorithms.

2.3.2. Methods

Some methods have been applied to the classification and recognition of LIBS spectra, such as LDA [35–37], SVM [38–40] and ANN [41,42]. In this study, they were used to build models for the identification of different penicillin manufacturers. Some publications have introduced these three algorithms, and we only provide a brief description in this section.

LDA is a classic linear discriminant analysis method [17]. It can quickly and simply handle two-label and multi-label classification problems without the requirement of parameter optimization [5].

SVM can handle both linear and nonlinear classification by transforming the kernel functions. The basic concept of the SVM classification method is to find the best hyperplane to maximize the classification interval from the point in the sample dataset to this hyperplane, that is, to maximize the boundary between classes [16]. SVM has a variety of kernel functions. In this study, the commonly used radial basis function (RBF) was used as the kernel function. The penalty parameter c and the RBF kernel function parameter g were optimized using the particle swarm optimization (PSO) algorithm.

ANN is a nonlinear classifier. The basic structure of the commonly used multi-layer feedforward neural network, the multi-layer perceptron, includes three layers: the input layer, where the number of neurons is equal to the number of input variables; the hidden layer, where the number of neurons is optimized based on Equation (1); and the output layer, where the output of the neuron represents the predicted categories. The number of neurons in the hidden layer, l, is determined by the following equation [17]:

$$l = \sqrt{n + m} + \alpha \quad (1)$$

where n is the number of neurons in the input layer, a is a constant between one and ten, and m is the number of neurons in the output layer.

3. Results and Discussion

3.1. LIBS Spectra

The mean spectrum of 70 LIBS spectra in the training set sample from each penicillin manufacturer is shown in Figure 2. We analyzed the spectral lines of penicillin with a wavelength of 200–900 nm and intensities greater than 500 counts. The corresponding wavelengths of the elements or molecular bands in the penicillin LIBS spectra are listed in Table 2. We used the National Institute of Standards and Technology (NIST) atomic emission database to identify the elements corresponding to these spectral lines. It can be observed from Figure 2 that the LIBS spectral intensities of AC are lower than those of PPT and ACPT. This might be because the AC samples were prepared through powder compression, and their density and hardness were lower than those of the PPT and ACPT tablet samples.

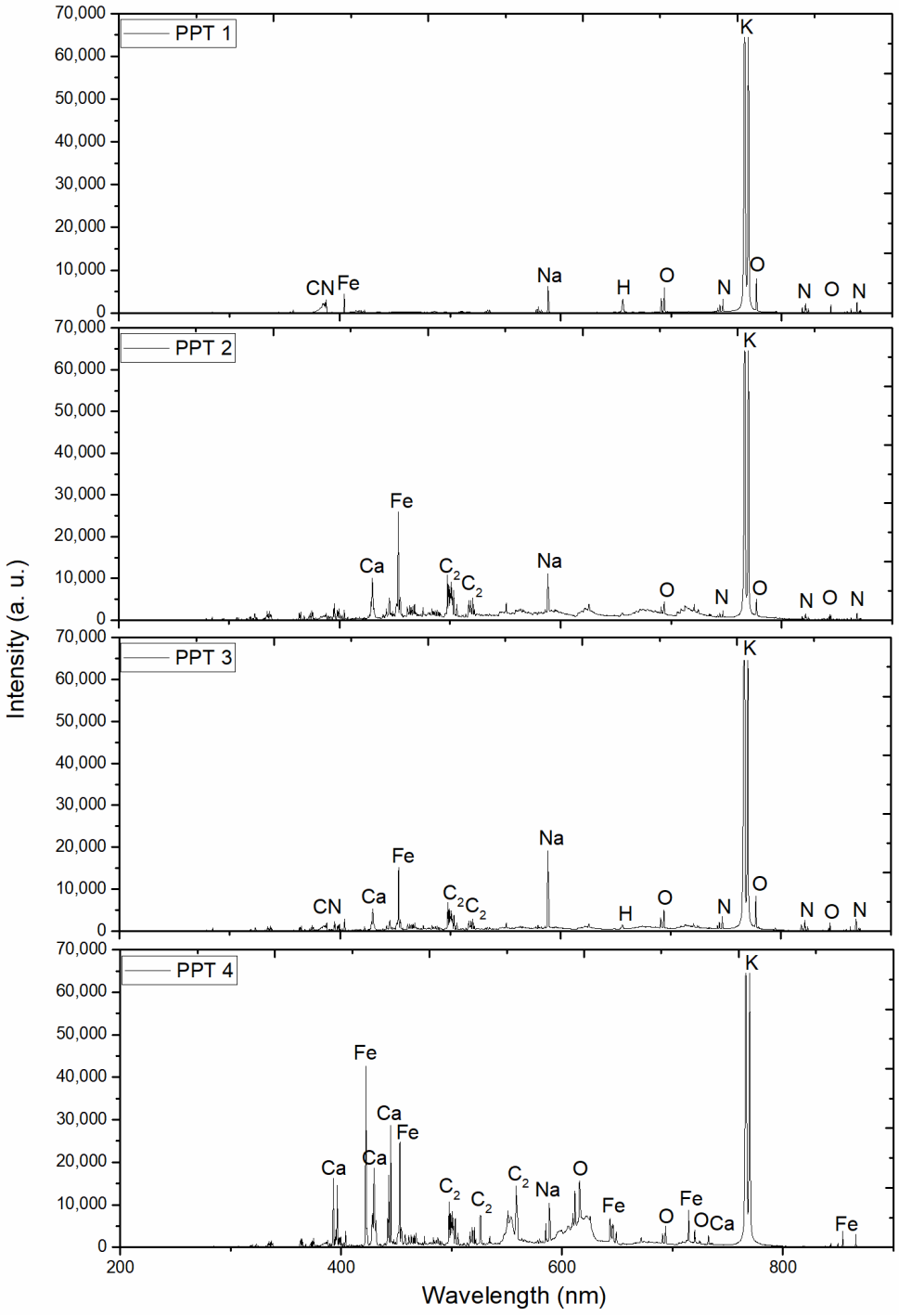

Figure 2. *Cont*.

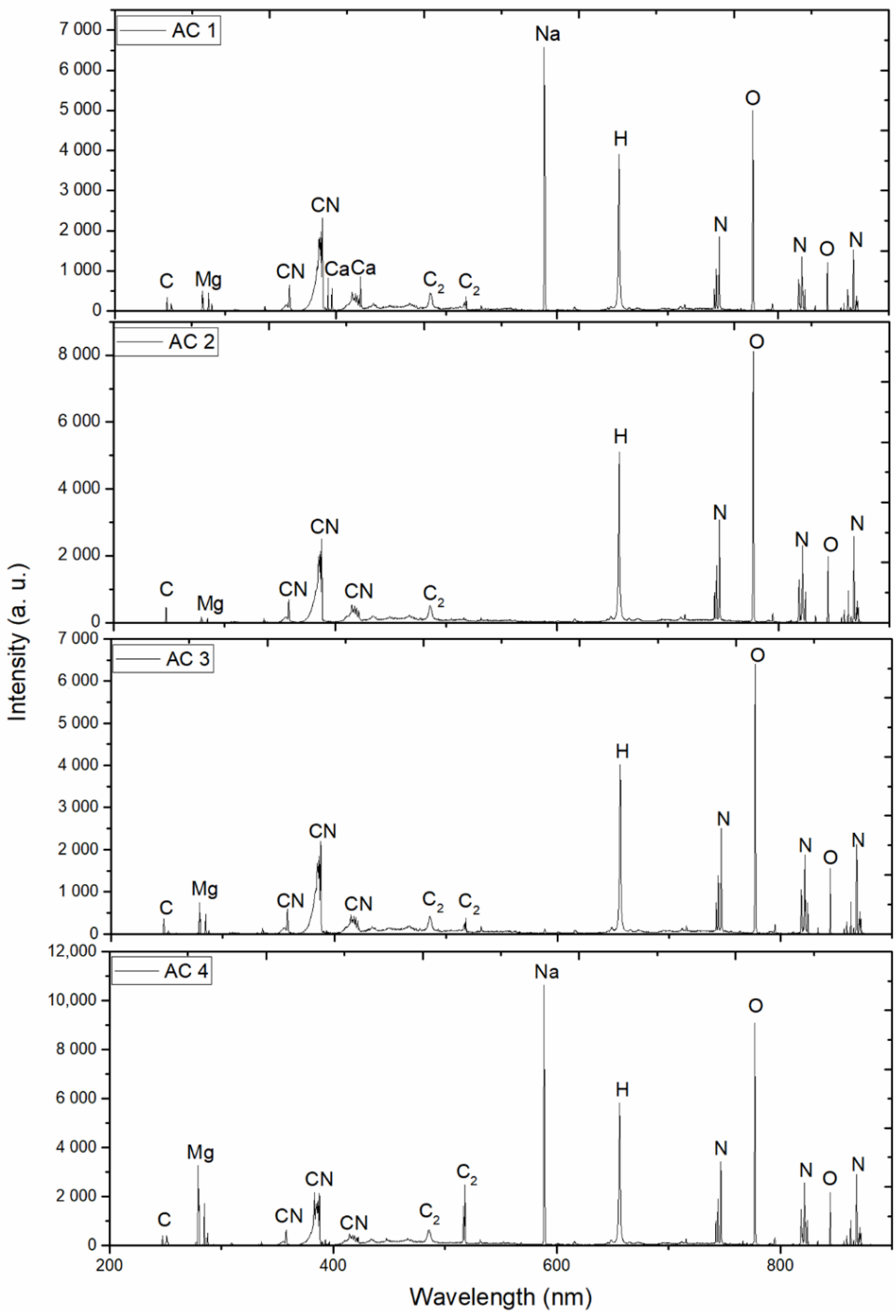

Figure 2. *Cont*.

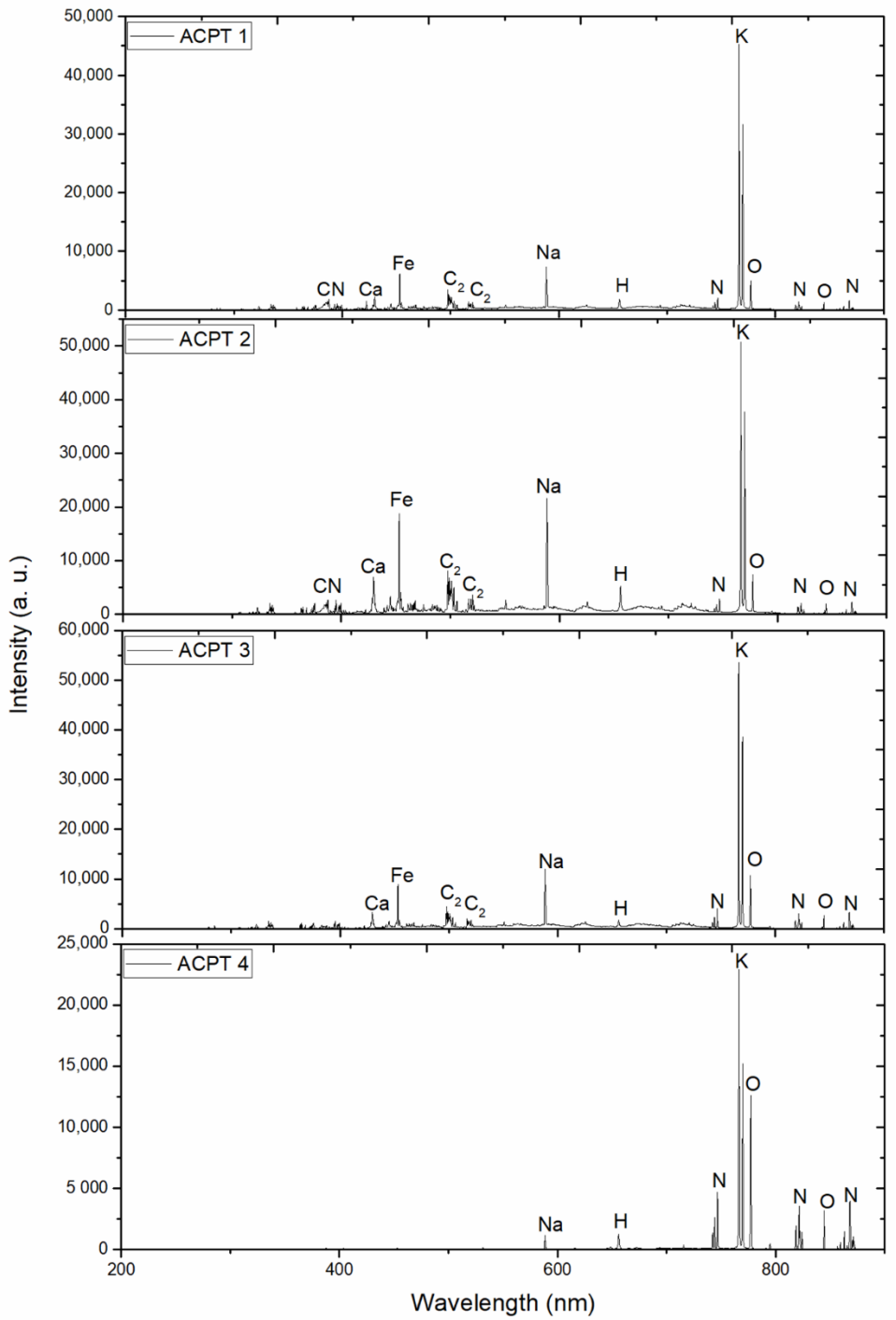

Figure 2. The mean spectrum of 70 LIBS spectra in the training set sample from each type of penicillin manufacturer.

It can be observed from Table 1 that the main elements shared by the three types of penicillin are C, H, O, N and S. The element of S was not detected in the LIBS spectra of the three types of penicillin. This is mainly because the content of S is relatively low, and the element of S is difficult to excite. There were two sources of C, H, O and N. On one hand, they came from the penicillin samples themselves, and on the other hand, as the penicillin LIBS experiments were performed in the air, the intensities of the spectral lines of these elements might be partially triggered by the excitation of the air.

Table 2. Corresponding wavelengths of the elements or molecular bands in the penicillin LIBS spectra.

Type	Element	Line/Band Wavelength (nm)
PPT	CN	358.0, 384.8, 385.7, 386.5, 388.0
	Ca	393.1, 396.8, 420.7, 422.6, 429.9, 443.1, 732.6
	Fe	404.1, 445.1, 453.2, 577.7, 579.6, 580.7, 582.8, 648.9, 714.8, 849.8, 853.8
	C_2	497.7, 498.5, 499.6, 500.1, 501.0, 503.2, 505.9, 516.8, 518.7, 520.6, 522.0, 550.9, 553.9, 558.5
	Na	588.9
	O	611.3, 615.8, 645.8, 671.8, 691.0, 693.6, 720.2, 777.4, 844.7
	H	656.2
	N	742.3, 744.2, 746.8, 818.8, 821.6, 822.3, 824.2, 862.9, 865.6, 868.0
	K	766.4, 769.9
AC	C	247.8
	Mg	279.1, 279.7, 284.7
	CN	358.0, 384.8, 385.7, 386.5, 388.0, 414.5, 416.2, 417.6, 419.1
	Ca	393.1, 396.8, 420.7, 422.6
	C_2	497.7, 498.5, 499.6, 500.1, 501.0, 503.2, 505.9, 516.8, 518.7, 520.6, 522.0
	Na	588.9
	O	777.4, 844.7
	H	656.2
	N	742.3, 744.2, 746.8, 818.8, 821.6, 822.3, 824.2, 862.9, 868.0
ACPT	CN	358.0, 384.8, 385.7, 386.5, 388.0
	Ca	393.1, 395.7, 422.6, 429.9
	Fe	445.1, 453.2
	C_2	497.7, 498.5, 499.6, 501.0, 503.2, 505.9, 516.8, 518.7, 520.6, 522.0, 550.9, 553.9, 558.5
	Na	588.9
	O	691.0, 693.6, 777.4, 844.7
	H	656.2
	N	742.3, 744.2, 746.8, 818.8, 821.6, 822.3, 824.2, 862.9, 868.0
	K	766.4, 769.9

We can see that the LIBS spectra of PPT and ACPT contain the element of K from Figure 2, while the LIBS spectra of AC do not. K is a metallic element, which is easily excited by lasers. The spectral lines of K in the LIBS spectra of PPT were saturated, and in the subsequent identification of different PPT manufacturers, the spectral lines of K were not used as the input data for the classification algorithms.

Some trace elements (Mg, Ca, Fe and Na) were also detected in the LIBS spectra of penicillin. The element of Mg was not detected in the samples of PPT or ACPT, but it was detected in the samples of AC. The Ca element was detected in the samples of PPT2, PPT3, PPT4, AC1, ACPT1, ACPT2 and ACPT3, but not in other samples. The element of Fe was detected in the samples of PPT, ACPT1, ACPT2 and ACPT3; however, it was not detected in the samples of AC or ACPT4. The Na element was detected in the samples of PPT, AC1, AC4 and ACPT, but not in other samples. The types and quantities of trace elements in penicillin produced by different manufacturers are different, which is also a standard for identifying different penicillin manufacturers.

3.2. Feature Selection

The LIBS spectra of penicillin contain a large amount of information, much of which may not be relevant for the identification of different penicillin manufacturers. In previous literature, it was confirmed that it is necessary to perform feature selection on LIBS spectral data to filter interference or redundant information, to facilitate the analysis of LIBS spectral data and improve the accuracy of spectral classification [17,43]. In this study, we evaluated the importance of each PPT, AC and ACPT spectral line for the manufacturer identification based on the decrease in the Gini index of RF. A significant decrease in the Gini index indicates that the spectral line is more important for classification. The RF algorithm was

introduced in a previous study [16]. As shown in Figure 2, there were 44 spectral peaks in the LIBS spectra of PPT. We calculated the integrated intensities of the 44 spectral peaks and normalized the 44 data using the maximum normalization method. The 44 normalized data were used as the inputs of RF to obtain 44 Gini indices.

For the same type of penicillin, the spectral lines of different elements had different effects on the identification of different penicillin manufacturers. The importance of each spectral line for the three types of penicillin is shown in Figure 3. In the LIBS spectra of PPT, the C_2 (497.7–505.9 nm) molecular band had the highest importance weight, and the spectral line of Na (588.9 nm) ranked second. The spectral lines of Na (588.9 nm) and Mg (279.7 nm and 279.1 nm) have relatively high importance in the LIBS spectra of AC. In the LIBS spectra of ACPT, the spectral lines of Na (588.9 nm) and K (766.4 nm) have great influence for the identification of different ACPT manufacturers. In the LIBS spectra of the three types of penicillin, the spectral line of Na (588.9 nm) has a relatively significant impact on the manufacturer identification.

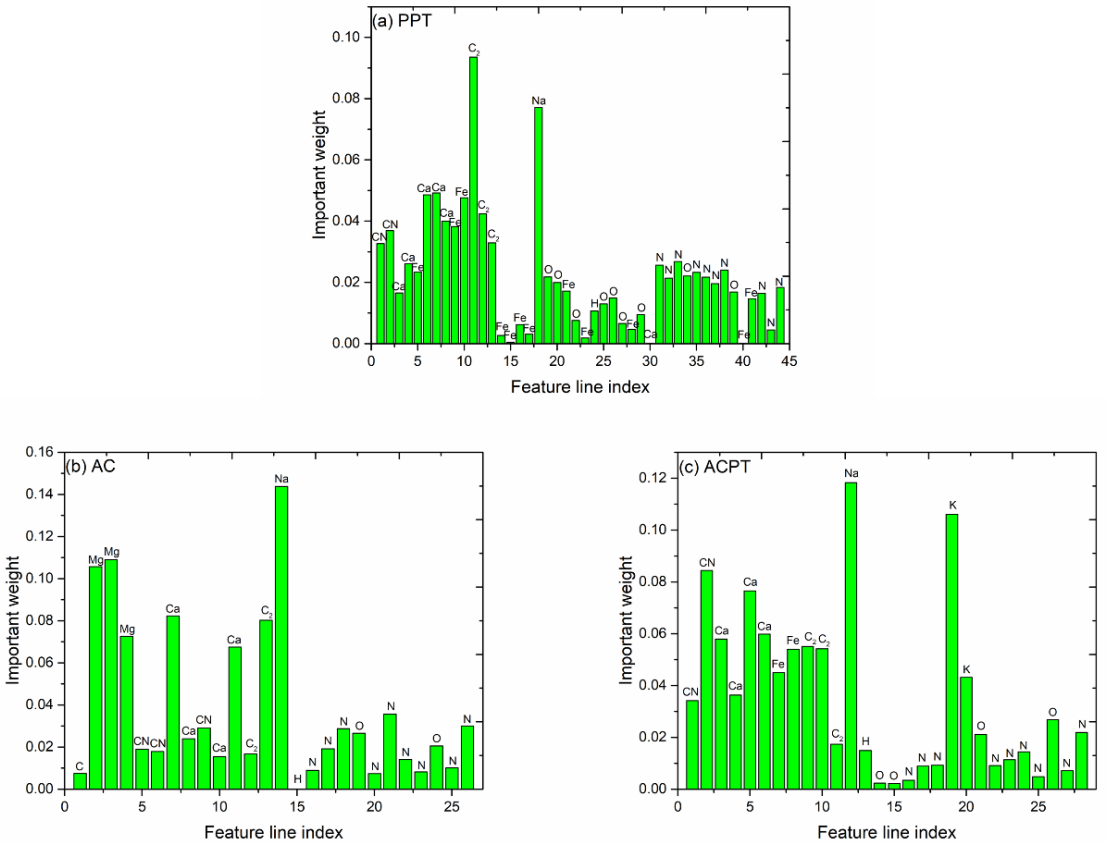

Figure 3. Importance of each (**a**) PPT, (**b**) AC and (**c**) ACPT spectral line for manufacturer identification based on the decrease in the Gini index of RF.

3.3. Identification of Penicillin Manufacturers

In the manufacturer identification for each type of penicillin, the most important spectral lines were selected as the inputs of three classifiers, LDA, SVM and ANN. The number of selected characteristic spectral lines affects the classification accuracy of the model. Therefore, after ranking the spectral lines based on their importance from high to low, we selected different numbers of penicillin feature lines, from the top one to the first twenty-five, as input data. Figure 4 shows the relationship between the test set accuracies of the three models—RF-LDA, RF-SVM and RF-ANN—and the number of characteristic spectral lines when these models were established using the optimal parameters. We used a computer (model: Lenovo, R9000X 2021; CPU: AMD Ryzen 7 4800 H with Radeon

Graphics) to deal with this particular problem. When the top n feature lines were used as the inputs of the classifiers for modeling, the classification accuracies obtained and the test time of one spectrum were as listed in Table 3.

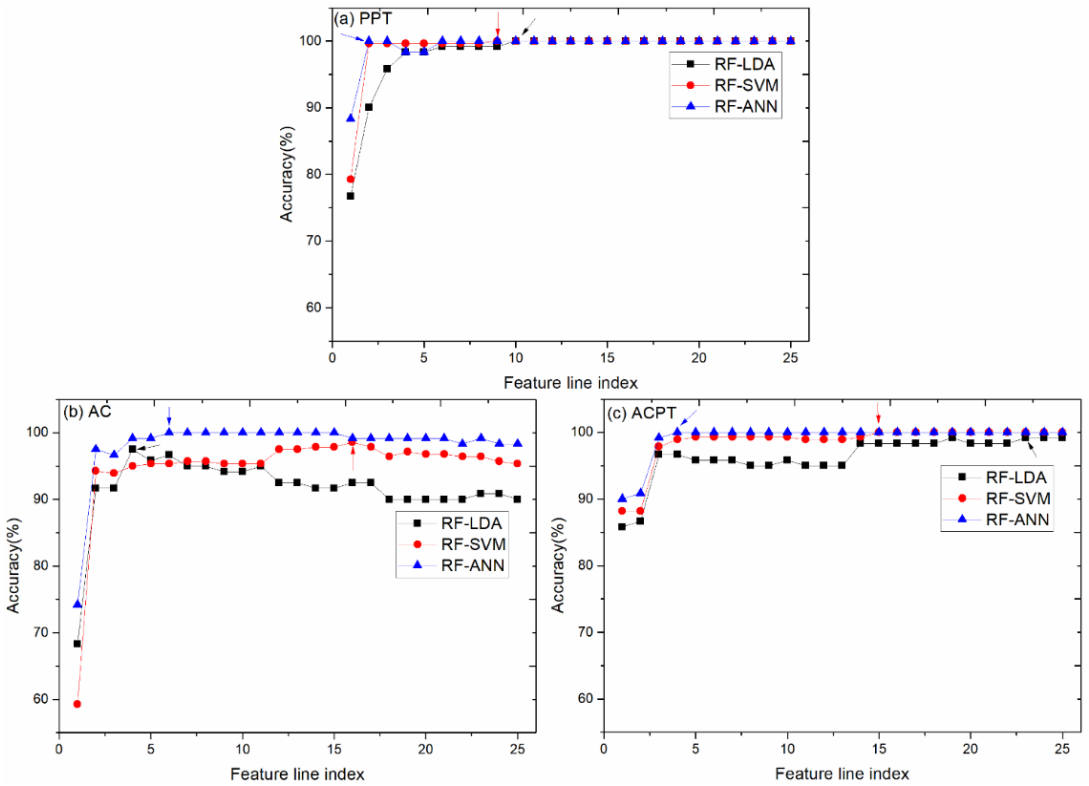

Figure 4. Test set classification accuracies of RF-LDA, RF-SVM and RF-ANN models for the three types of penicillin samples ((**a**) PPT, (**b**) AC and (**c**) ACPT) with different numbers of characteristic spectral lines.

Table 3. The best classification accuracy and corresponding analysis time of each model.

Type	Model	Top n Feature Lines	Test Time (s)	CCR
PPT	LDA	10	0.0017	100%
	SVM	9	0.0036	100%
	ANN	2	0.0049	100%
AC	LDA	4	0.0015	97.50%
	SVM	16	0.0047	98.33%
	ANN	6	0.0054	100%
ACPT	LDA	23	0.0021	99.17%
	SVM	15	0.0045	100%
	ANN	4	0.0051	100%

Figure 4a shows the test set classification accuracies of the PPT manufacturer identification using the three models—RF-LDA, RF-SVM and RF-ANN—with different numbers of characteristic spectral lines. The classification accuracy of the RF-LDA model reached 100% with the top ten important characteristic spectral lines. When inputting the top nine important characteristic spectral lines and using the RF-SVM model (c = 0.1, g = 19.4909), the classification accuracy was 100%. The classification accuracy of the RF-ANN model (l = 4) was 100% when using the top two important characteristic spectral lines. The LIBS spectra of four different PPT manufacturers were correctly classified. The RF-ANN model provided an excellent classification result when using the top two important characteristic spectral lines. It can be seen from Figure 3a that the order of the important weights of

the first two important characteristic variables is C_2 (497.7–505.9 nm) > Na (588.9 nm). C_2 emission was principally fragmented directly from the samples. The molecular structure of PPT contained some carbon–carbon bonds (C-C) and carbon–carbon double bonds (C=C). After the high-energy laser interacted with the samples, some molecular fragments were excited, and C_2 mainly came from these molecular fragments [44]. The element of Na came from the sodium citrate ($C_6H_5Na_3O_7$), which could be used as an antioxidant, in the prescription of PPT [45].

For AC manufacturer identification, Figure 4b shows that, when inputting the top four important characteristic spectral lines and using RF-LDA to build the model, the classification accuracy reached 97.50%. Three LIBS spectra of AC3 were misclassified into AC2. It can also be seen from Figure 2 that the LIBS spectrum of AC3 was more similar to that of AC2 compared to the spectra of AC1 and AC4. The classification accuracy of the RF-SVM model (c = 60.1465, g = 39.7502) was 98.33% with the top 16 important characteristic spectral lines. Two LIBS spectra of AC3 were misclassified into AC2. Compared with LDA and SVM, the ANN classifier has higher accuracy when inputting the same number of characteristic spectral lines. When inputting the top six important characteristic spectral lines and using the RF-ANN (l = 13) model, the classification accuracy reached 100%, successfully identifying different AC manufacturers. It can be seen from Figure 3b that the order of the important weights of the first six important characteristic variables is Na (588.9 nm) > Mg (279.7 nm) > Mg (279.1 nm) > Ca (393.1 nm) > C_2 (516.8–518.7 nm) > Mg (284.7 nm). The elements of Na and Mg came from the lubricants sodium dodecyl sulfate ($C_{12}H_{25}SO_4Na$) and magnesium stearate ($C_{36}H_{70}MgO_4$), respectively, in the prescription of AC [46]. The element of Ca might come from impurities or additives introduced by different manufacturers during the production process for AC.

Figure 4c shows that, when inputting the top 23 important characteristic spectral lines and using RF-LDA to build the model, the classification accuracy reached 99.17%. One LIBS spectrum of ACPT3 was misclassified into ACPT1. It can also be seen from Figure 2 that the LIBS spectrum of ACPT3 is more similar to that of ACPT1 than the spectra of ACPT2 and ACPT4. The classification accuracy of the RF-SVM model (c = 7.2070, g = 6.2123) reached 100% with the top fifteen important characteristic spectral lines. When inputting the top four important characteristic spectral lines and using the RF-ANN (l = 10) model, the classification accuracy reached 100%. The LIBS spectra of four different ACPT manufacturers could be correctly classified by using the SVM and ANN models. It can be seen from Figure 3c that the order of the important weights of the first four important characteristic variables is Na (588.9 nm) > K (766.4 nm) > CN (384.8 nm–388.0 nm) > Ca (422.6 nm). The elements of Na and Ca might come from impurities or additives introduced by different manufacturers during the production process for ACPT. The element of K came from clavulanate potassium in the prescription of ACPT [47]. The molecular band of CN had two sources. The molecular structure of the ACPT sample was destroyed by the high-energy laser, producing some CN fragments. In addition, CN could also be formed by a chemical reaction between the C in the sample and the N in the atmosphere [44].

For PPT, AC and ACPT manufacturer identification, the RF-LDA model failed to reach 100% accuracy in AC and ACPT manufacturer identification; RF-SVM failed to achieve 100% accuracy in AC manufacturer identification; RF-ANN achieved 100% accuracy in the manufacturer identification of PPT, AC and ACPT. LDA is more suitable for the classification of linear data. However, the penicillin LIBS data in this study are nonlinear data. Part of the data of the penicillin LIBS spectra fall on the support vectors of SVM. Compared with that of ANN, the learning ability of SVM is insufficient. Through the repeated learning and training of known information, ANN can better adjust the connection weights of neurons, has strong robustness and fault tolerance to noise, and can fully approximate complex nonlinear relationships. Therefore, the established model is more suitable for nonlinear systems and unclear functional relationships. Overall, compared with RF-LDA and RF-SVM, RF-ANN was the most suitable model for penicillin manufacturer identification in this study.

4. Conclusions

The study mainly focused on the feasibility of using LIBS technology combined with chemometrics to identify different penicillin manufacturers. We analyzed the LIBS spectra of different PPT, AC and ACP manufacturers and used the decrease in the Gini index of RF to evaluate the important weight of each spectral line for the manufacturer identification of three types of penicillin. In the identification of four different PPT manufacturers, the importance weight of Na ranked second. In the identification of four different AC and ACPT manufacturers, the importance weight of Na ranked first. This indicated that Na played an important role in the identification of penicillin manufacturers. We used three classification algorithms—LDA, SVM and ANN—to identify different manufacturers of PPT, AC and ACPT, respectively. The RF-ANN model achieved 100% classification accuracy for all three test sets, successfully accomplishing penicillin manufacturer identification. The results show that LIBS technology combined with chemometrics could be used to identify different penicillin manufacturers. In future work, we will test the robustness and generalization of these classification models in the application of drug detection.

Author Contributions: Supervision, Q.W.; methodology, G.T.; software, X.X.; formal analysis, Z.Z.; validation, G.C.; writing—review and editing, K.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (No. 62075011).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data in this study can be requested from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Girmatsion, M.; Mahmud, A.; Abraha, B.; Xie, Y.; Cheng, Y.; Yu, H.; Yao, W.; Guo, Y.; Qian, H. Rapid detection of antibiotic residues in animal products using surface-enhanced Raman Spectroscopy: A review. *Food Control* **2021**, *126*, 108019. [CrossRef]
2. Li, L.Q.; Pan, X.P.; Feng, Y.C.; Yin, L.H.; Hu, C.Q.; Yang, H.H. Deep convolution network application in identification of multi-variety and multi-manufacturer pharmaceutical. *Spectrosc. Spect. Anal.* **2019**, *39*, 3606–3613.
3. Liu, Y.G.; Tan, P.; Li, F.; Qiao, Y.J. Study on the aconitine-type alkaloids of Radix Aconiti Lateralis and its processed products using HPLC-ESI-MSn. *Drug Test. Anal.* **2013**, *5*, 480–484. [CrossRef]
4. Nithaniyal, S.; Vassou, S.L.; Poovitha, S.; Raju, B.; Parani, M. Identification of species adulteration in traded medicinal plant raw drugs using DNA barcoding. *Genome* **2017**, *60*, 139–146. [CrossRef]
5. Wei, K.; Cui, X.T.; Teng, G.E.; Khan, M.N.; Wang, Q.Q. Distinguish *Fritillaria cirrhosa* and non-*Fritillaria cirrhosa* using la-ser-induced breakdown spectroscopy. *Plasma Sci. Technol.* **2021**, *23*, 085507. [CrossRef]
6. Zhang, H.; Hua, D.; Huang, C.; Samal, S.K.; Xiong, R.; Sauvage, F.; Braeckmans, K.; Remaut, K.; De Smedt, S.C. Materials and Technologies to Combat Counterfeiting of Pharmaceuticals: Current and Future Problem Tackling. *Adv. Mater.* **2020**, *32*, e1905486. [CrossRef]
7. Wilson, B.K.; Kaur, H.; Allan, E.L.; Lozama, A.; Bell, D. A New Handheld Device for the Detection of Falsified Medicines: Demonstration on Falsified Artemisinin-Based Therapies from the Field. *Am. J. Trop. Med. Hyg.* **2017**, *96*, 1117–1123. [CrossRef]
8. Scafi, S.H.F.; Pasquini, C. Identification of counterfeit drugs using near-infrared spectroscopy. *Analyts* **2001**, *126*, 2218–2224. [CrossRef]
9. Paudel, A.; Raijada, D.; Rantanen, J. Raman spectroscopy in pharmaceutical product design. *Adv. Drug Deliv. Rev.* **2015**, *89*, 3–20. [CrossRef]
10. De Veij, M.; Vandenabeele, P.; Hall, K.A.; Fernandez, F.M.; Green, M.D.; White, N.J.; Dondorp, A.M.; Newton, P.N.; Moens, L. Fast detection and identification of counterfeit antimalarial tablets by Raman spectroscopy. *J. Raman Spectrosc.* **2007**, *38*, 181–187. [CrossRef]
11. Dong, P.K.; Zhao, S.Y.; Zheng, K.X.; Wang, J.; Gao, X.; Hao, Z.Q.; Lin, J.Q. Rapid identification of ginseng origin by laser induced breakdown spectroscopy combined with neural network and support vector machine algorithm. *Acta Phys. Sin.* **2021**, *70*, 040201.
12. Gaft, M.; Nagli, L.; Gorychev, A.; Raichlin, Y. Atomic and molecular emission of beryllium by LIBS. *Spectrochim. Acta B* **2021**, *182*, 106233. [CrossRef]
13. Dietz, T.; Gottlieb, C.; Kohns, P.; Ankerhold, G. Comparison of atomic and molecular emission in laser-induced breakdown spectroscopy for the quantification of harmful species in cement-based materials. *Spectrochim. Acta B* **2019**, *161*, 105707. [CrossRef]

14. Wang, Q.; Teng, G.; Li, C.; Zhao, Y.; Peng, Z. Identification and classification of explosives using semi-supervised learning and laser-induced breakdown spectroscopy. *J. Hazard. Mater.* **2019**, *369*, 423–429. [CrossRef]
15. Moros, J.; Serrano, J.; Sánchez, C.; Macías, J.; Laserna, J.J. New chemometrics in laser-induced breakdown spectroscopy for recognizing explosive residues. *J. Anal. At. Spectrom.* **2012**, *27*, 2111–2122. [CrossRef]
16. Teng, G.E.; Wang, Q.Q.; Zhong, H.W.; Xiangli, W.T.; Yang, H.F.; Qi, X.L.; Cui, X.T.; Idrees, B.S.; Kai, W.; Khan, M.N. Discrimination of infiltrative glioma boundary based on laser-induced breakdown spectroscopy. *Spectrochim. Acta B* **2020**, *165*, 105787. [CrossRef]
17. Wang, Q.; Xiangli, W.; Chen, X.; Zhang, J.; Teng, G.; Cui, X.; Idrees, B.S.; Wei, K. Primary study of identification of parathyroid gland based on laser-induced breakdown spectroscopy. *Biomed. Opt. Express* **2021**, *12*, 1999–2014. [CrossRef]
18. Kim, K.-R.; Kim, G.; Kim, J.-Y.; Park, K.; Kim, K.-W. Kriging interpolation method for laser induced breakdown spectroscopy (LIBS) analysis of Zn in various soils. *J. Anal. At. Spectrom.* **2013**, *29*, 76–84. [CrossRef]
19. Wang, T.; He, M.; Shen, T.; Liu, F.; He, Y.; Liu, X.; Qiu, Z. Multi-element analysis of heavy metal content in soils using laser-induced breakdown spectroscopy: A case study in eastern China. *Spectrochim. Acta B* **2018**, *149*, 300–312. [CrossRef]
20. Huang, W.; He, C.; Wang, Y.; Zhao, W.; Qiu, L. Confocal controlled LIBS microscopy with high spatial resolution and stability. *J. Anal. At. Spectrom.* **2020**, *35*, 2530–2535. [CrossRef]
21. Yu, J.; Hou, Z.; Sheta, S.; Dong, J.; Han, W.; Lu, T.; Wang, Z. Provenance classification of nephrite jades using multivariate LIBS: A comparative study. *Anal. Methods* **2017**, *10*, 281–289. [CrossRef]
22. He, T.; Liang, J.; Tang, H.; Zhang, T.; Yan, C.; Li, H. Quantitative analysis of coal quality by mutual information-particle swarm optimization (MI-PSO) hybrid variable selection method coupled with spectral fusion strategy of laser-induced breakdown spectroscopy (LIBS) and fourier transform infrared spectroscopy (FTIR). *Spectrochim. Acta B* **2021**, *178*, 106112. [CrossRef]
23. Palásti, D.; Metzinger, A.; Ajtai, T.; Bozóki, Z.; Hopp, B.; Kovács-Széles, E.; Galbács, G. Qualitative discrimination of coal aerosols by using the statistical evaluation of laser-induced breakdown spectroscopy data. *Spectrochim. Acta B* **2019**, *153*, 34–41. [CrossRef]
24. Stefas, D.; Gyftokostas, N.; Kourelias, P.; Nanou, E.; Kokkinos, V.; Bouras, C.; Couris, S. Discrimination of olive oils based on the olive cultivar origin by machine learning employing the fusion of emission and absorption spectroscopic data. *Food Control* **2021**, *130*, 108318. [CrossRef]
25. Nespeca, M.G.; Vieira, A.L.; Júnior, D.S.; Neto, J.A.G.; Ferreira, E.C. Detection and quantification of adulterants in honey by LIBS. *Food Chem.* **2019**, *311*, 125886. [CrossRef]
26. Aldakheel, R.; Gondal, M.; Nasr, M.; Dastageer, M.; Almessiere, M. Quantitative elemental analysis of nutritional, hazardous and pharmacologically active elements in medicinal Rhatany root using laser induced breakdown spectroscopy. *Arab. J. Chem.* **2020**, *14*, 102919. [CrossRef]
27. Nisar, S.; Dastgeer, G.; Shafiq, M.; Usman, M. Qualitative and semi-quantitative analysis of health-care pharmaceutical products using laser-induced breakdown spectroscopy. *J. Pharm. Anal.* **2018**, *9*, 20–24. [CrossRef]
28. Doucet, F.R.; Faustino, P.J.; Sabsabi, M.; Lyon, R.C. Quantitative molecular analysis with molecular bands emission using laser-induced breakdown spectroscopy and chemometrics. *J. Anal. At. Spectrom.* **2008**, *23*, 694–701. [CrossRef]
29. Zhao, S.; Song, W.; Hou, Z.; Wang, Z. Classification of ginseng according to plant species, geographical origin, and age using laser-induced breakdown spectroscopy and hyperspectral imaging. *J. Anal. At. Spectrom.* **2021**, *36*, 1704–1711. [CrossRef]
30. Zhang, D.; Ding, J.; Feng, Z.; Yang, R.; Yang, Y.; Yu, S.; Xie, B.; Zhu, J. Origin identification of Ginkgo biloba leaves based on laser-induced breakdown spectroscopy (LIBS). *Spectrochim. Acta B* **2021**, *180*, 106192. [CrossRef]
31. Zheng, P.C.; Zheng, S.; Wang, J.M.; Liao, X.Y.; Li, X.J.; Peng, R. Study on grade identification of dendrobium by LIBS. *Spectrosc. Spect. Anal.* **2020**, *40*, 941–944.
32. Pořízka, P.; Klus, J.; Hrdlička, A.; Vrábel, J.; Škarková, P.; Prochazka, D.; Novotny, J.; Novotny, K.; Kaiser, J. Impact of La-ser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy. *J. Anal. Atom. Spectrom.* **2017**, *32*, 277–288. [CrossRef]
33. Zorov, N.B.; Gorbatenko, A.A.; Labutin, T.A.; Popov, A.M. A review of normalization techniques in analytical atomic spec-trometry with laser sampling: From single to multivariate correction. *Spectrochim. Acta B* **2010**, *65*, 642–657. [CrossRef]
34. Castro, J.P.; Pereira-Filho, E.R. Twelve different types of data normalization for the proposition of classification, univariate and multivariate regression models for the direct analyses of alloys by laser-induced breakdown spectroscopy (LIBS). *J. Anal. At. Spectrom.* **2016**, *31*, 2005–2014. [CrossRef]
35. Li, W.T.; Zhu, Y.N.; Li, X.; Hao, Z.Q.; Guo, L.B.; Zeng, X.Y.; Lu, Y.F. In situ classification of rocks using stand-off laser-induced breakdown spectroscopy with a compact spectrometer. J. *Anal. At. Spectrom.* **2018**, *33*, 461–467. [CrossRef]
36. Martino, L.J.; Angelo, C.A.D.; Marnelli, C.; Cepda, R. Identification and detection of pesticide in chard samples by laser-induced breakdown spectroscopy using chemometric methods. *Spectrochim. Acta B* **2020**, *177*, 106031. [CrossRef]
37. Gibbons, E.; Léveillé, R.; Berlo, K. Data fusion of laser-induced breakdown and Raman spectroscopies: Enhancing clay mineral identification. *Spectrochim. Acta B* **2020**, *170*, 105905. [CrossRef]
38. Yao, M.; Fu, G.; Chen, T.; Liu, M.; Xu, J.; Zhou, H.; He, X.; Huang, L. A modified genetic algorithm optimized SVM for rapid classification of tea leaves using laser-induced breakdown spectroscopy. *J. Anal. At. Spectrom.* **2020**, *36*, 361–367. [CrossRef]
39. Lin, J.; Lin, X.; Guo, L.; Guo, Y.; Tang, Y.; Chu, Y.; Tang, S.; Che, C. Identification accuracy improvement for steel species using a least squares support vector machine and laser-induced breakdown spectroscopy. *J. Anal. At. Spectrom.* **2018**, *33*, 1545–1551. [CrossRef]

40. Lin, X.; Sun, H.; Gao, X.; Xu, Y.; Wang, Z.; Wang, Y. Discrimination of lung tumor and boundary tissues based on laser-induced breakdown spectroscopy and machine learning. *Spectrochim. Acta B* **2021**, *180*, 106200. [CrossRef]
41. Yang, Y.; Li, C.; Liu, S.; Min, H.; Yan, C.; Yang, M.; Yu, J. Classification and identification of brands of iron ores using laser-induced breakdown spectroscopy combined with principal component analysis and artificial neural networks. *Anal. Methods* **2020**, *12*, 1316–1323. [CrossRef]
42. Li, L.-N.; Liu, X.-F.; Yang, F.; Xu, W.-M.; Wang, J.-Y.; Shu, R. A review of artificial neural network based chemometrics applied in laser-induced breakdown spectroscopy analysis. *Spectrochim. Acta B* **2021**, *180*, 106183. [CrossRef]
43. Vrábel, J.; Kepes, E.; Duponchel, L.; Motto-Ros, V.; Fabre, C.; Connemann, S.; Schreckenberg, F.; Prasse, P.; Riebe, D.; Junjuri, R.; et al. Classification of challenging Laser-Induced Breakdown Spectroscopy soil sample data—EMSLIBS contest. *Spectrochim. Acta B* **2020**, *169*, 105872. [CrossRef]
44. Xu, F.; Ma, S.; Zhao, C.; Dong, D. Application of Molecular Emissions in Laser-Induced Breakdown Spectroscopy: A Review. *Front. Phys.* **2022**, *10*, 7. [CrossRef]
45. Leng, C.J.; Zhang, T.L.; Xu, J. Prescription Technology Study on Penicillin V Potassium Tablets. *China Pharm.* **2013**, *22*, 56–57.
46. Li, T.Y.; Li, G.L.; Zhong, X. Preparation and study of amoxicillin dispersible tablet. *Jilin Med. J.* **2008**, *29*, 2158–2160.
47. Ouyang, Y.H.; Guo, X.J. Preparation of Dispersible Tablets of Amoxicillin and Clavulanate Potassium. *Pharm. Today* **2013**, *23*, 664–666.