Fast Detection Method of Antarctic Krill Meat Quality Based on Near Infrared Spectroscopy

Lanlan Zhu[#]
School of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo255000, Shandong, China

Zhongling Hou[#]
Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao266000, Shandong, China
College of Food Science and Technology, Shanghai Ocean University, Shanghai200000, China

Yanbo Wang
School of Food Science and Biotechnology, Zhejiang Gongshang University, Hangzhou 310018, China

Kailiang Leng[*]
Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao266000, Shandong, China
[*]*Corresponding author (Email: lengkl@ysfri.ac.cn)*

Weihong Sun
Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao266000, Shandong, China

Qilong Shi
School of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo255000, Shandong, China

Dongwu Liu
School of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo255000, Shandong, China
[#]*These authors contributed equally to this work as co-first author*

Abstract: In order to protect the health, safety and legal rights of consumers, the quality of shrimp meat must be strictly controlled during the processing, transportation and marketing of shrimp meat. Traditional shrimp meat quality testing is mainly based on chemical methods, which often require multiple chemical instruments and reagents. Samples require pretreatment, which is cumbersome and time-consuming. Based on the above background, the purpose of this paper is to study the rapid detection method of Antarctic krill meat quality based on near-infrared spectroscopy analysis technology. In this paper, a qualitative analysis model of shrimp meat quality using near-infrared spectroscopy is established based on four methods: support vector machine, BP neural network, random forest and width learning. Combined with the principles of modeling methods, the performance of the analysis model is comprehensively evaluated under the same preprocessing method and sample set partition conditions. The best prediction accuracy of the five-fold cross method of the four qualitative analysis models reached 88.48%, 88.57%, 89.05%, and 79.62%, respectively. Then, based on the existing theoretical methods, a method for rapid detection of shrimp freshness using a portable near-infrared spectrometer is studied. By comprehensively assessing the two indicators of sample accuracy and time, a combination mode based on the standard normal variable transformation + discrete Fourier transform method + support vector machine method is determined to construct a qualitative analysis model of the near red spectrum of shrimp meat to achieve real-time online discrimination fresh and refrigerated shrimp. In order to accelerate the process of near-infrared spectroscopy used to achieve fast online detection of shrimp quality.

Key words: Near-infrared Spectroscopy, Antarctic Krill, Rapid Detection, Data Processing, Qualitative Analysis Model

1. Introduction

In order to protect the physical and mental health and legitimate rights and interests of consumers, certain measures must be taken to strictly control the quality of shrimp meat in all aspects of production, processing and sales. At present, the quality inspection methods commonly used in the shrimp market are sensory inspection

and laboratory inspection. However, the existing methods are difficult to meet the fast and accurate demand for shrimp meat quality inspection in the market. Therefore, it is urgent to develop a more rapid, efficient and non-destructive inspection method for meat products. Near-infrared spectroscopy is a substance analysis method based on the difference in the absorbance of near-infrared light at different wavelengths of the substance to be measured. With the continuous progress of near-infrared spectral analysis technology, and the rapid development of data analysis, deep learning and artificial intelligence theory, it has been widely used in the fields of petroleum, chemical industry, agriculture, forestry and food. Compared with the traditional method, this technology has the following advantages: no pretreatment of the sample to be tested, rapid and simultaneous measurement of multiple components, no damage to the sample to be tested, and no threat to the environment and human health. A rapid detection method for shrimp meat quality based on near-infrared spectroscopy analysis technology. By using the collected shrimp meat spectral data, a qualitative analysis model is established to realize the prediction of shrimp meat quality. Due to the problems of strong scattering interference, high noise, and high feature dimensions of shrimp meat spectra, and the corresponding relationship between shrimp meat spectra and their quality is complex, in order to obtain more accurate analysis results, more accurate modeling data and more effective analysis methods are required. Moreover, most of the existing research is limited to laboratory research. The rapid development of the shrimp meat market has put forward higher requirements for theoretical research and technological development. From a technical perspective, both laboratory research and practical application scenarios require more robust preprocessing methods and analysis models as technical support. Therefore, the use of portable near-infrared spectroscopy to quickly identify the quality of shrimp meat, whether it is a pre-processing method or a qualitative analysis model in the later period, requires generalization analysis.

Due to the importance of shrimp meat quality research, many research teams began to study the quality of shrimp meat and achieved good results. Wildan used P. vannamei as a test material and used a combination of enzymolysis method and ultrasonic and enzymatic method to study reduction of tropomyosin (Tm). First, the enzymatic hydrolysis of prawn bromelain Tm was optimized by measuring changes in enzyme activity / substrate mass ratio, culture time, and hydrolysis temperature. Three kinds of hydrolysis methods were used to compare and study the changes of Tm sensitization in shrimp, butterfly shrimp and shrimp. Enzyme-linked immunosorbent assay (ELISA) was used to detect the antiserum of shrimp allergen Tm sensitized animal models. The results showed that the combination of enzymolysis and ultrasound had no significant effect on sensitization of mice ($p > 0.05$). The sensitization of butterfly shrimp treated with ultrasound and bromelain was reduced to 21.05%, and the sensitization of shrimp surimi treated with enzymolysis or combined ultrasound and enzymolysis was reduced to 30.70% and 33.33% respectively [1, 2]. The Zisheng study aimed to investigate the effect of chitosan-carvacrol coating with or without caprylic acid (CAP) on the quality of Pacific white shrimp (Litopenaeus vannamei) for 10 days under refrigeration. The results showed that the chitosan-carotene coating significantly inhibited the increase of total aerobic bacteria plate count (TPC), pH value, and total volatile alkali nitrogen (TVB-N) content of shrimp. The chitosan-carvacrol coating film also delayed the changes of shrimp blackening information and ΔE value, and improved the texture and sensory characteristics of the shrimp. In addition, the addition of CAP enhanced the inhibitory effect of chitosan carvacrol coating on TPC and TVB-N. The addition of CAP also made the texture characteristics of shrimps more preserved [3, 4]. Although the research results are richer now, there are still shortcomings, mainly due to the inability to quickly detect the quality of shrimp meat.

In the study of shrimp quality inspection, near-infrared spectroscopy analysis is a very good method that can solve many quality inspection problems. Therefore, it is widely used in quality inspection research. Representative tobacco samples were determined by near-infrared spectroscopy. A partial least square method was used to establish a calibration model between the spectral data and the total volatile base (TVB) and total volatile acid (TVA) content of tobacco samples, and it was verified with 50 tobacco samples. The predicted root mean square error (RMSEP) of TVB and TVA are 0.020 and 0.009, respectively; the relative standard deviations of TVB and TVA are 1.120% and 0.919%, respectively [5, 6]. António detected and quantified adulteration of camel milk and goat milk using new near-infrared spectroscopy and multivariate analysis. Camel milk samples were collected from the Aldhahira and Sharqia areas of the Sultanate of Oman. Measurements were taken in the wavelength range of 700-2500 nm, with a resolution of 2 cm-1, using 0.2 mm CaF 2 sealed cells, and in near-infrared spectroscopy in absorption mode. Multivariate analysis methods such as PCA, PLS-DA and PLS regression were used to interpret the near-infrared spectral data. The PLS-DA method was used to identify pure milk and adulterated milk samples. For the PLSDA model, the R-squared value is 0.974 and the RMSE is 0.08. In addition, the PLS regression model was used to quantitatively analyze samples with adulteration rates of 0%, 2%, 5%, 10%, 15%, and 20%. The PLS model showed RMSEC = 1.10% and R 2 = 94%. This method is simple, reproducible, and sensitive [7, 8]. Due to the effectiveness of near-infrared spectroscopy analysis, can the near-infrared spectroscopy analysis method be applied to the detection of shrimp meat quality and solve the tedious problem of shrimp meat quality detection.

This article introduces the mechanism of near-infrared spectroscopy generation, the physical and chemical

basis of near-infrared spectroscopy analysis; clarifies the main process of near-infrared spectroscopy analysis of shrimp quality; develops model evaluation methods and indicators for experimental modeling; gives based on random forest and BP neural network, etc. The qualitative analysis model of near-infrared spectroscopy of commonly used methods; the breadth learning algorithm is used to construct the qualitative analysis model of shrimp near-infrared spectroscopy; through a large number of experimental analysis and comparison, the effectiveness of the modeling method is studied. First, the process of collecting shrimp meat spectra and the characteristics of the collected spectra by a portable near-infrared spectrometer is introduced. Then, based on the aforementioned theoretical methods, a portable detection method for shrimp freshness is studied, and a robust analysis is performed to explore the feasibility of using portable near infrared spectrometer to detect shrimp meat quality online.

2. Proposed Method

2.1. Near-infrared Spectrum Generation Mechanism

According to the definition given by the American Society of Testing Materials, near-infrared light refers to electromagnetic waves with a wavelength range of 0.78 to 2.526 microns (wavenumber representation: 12821 to 3959cm^{-1}), and its wavelength (frequency) Between visible light and mid-infrared light, as shown in Figure 1 is a schematic diagram of the near-infrared spectrum in the electromagnetic wave band.

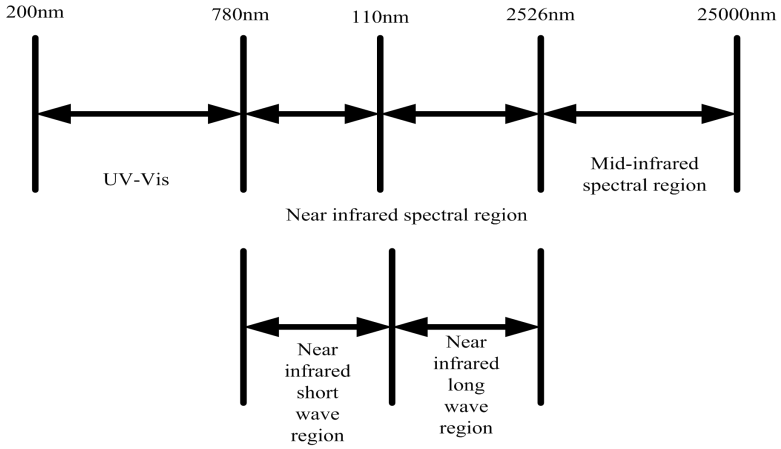


Figure 1. Schematic diagram of the near-infrared spectral wavelength range

Since near infrared light is an electromagnetic wave, it has photon energy E:

$$E = hv \tag{1}$$

Where h is the Planck constant and υ is the frequency of the electromagnetic wave [9, 10].

When continuous-frequency near-infrared light irradiates organic matter and acts, its molecules or atoms absorb photon energy, generate vibration or rotation, and cause energy level transitions. According to the molecular vibration theory of quantum mechanics, the energy of molecular vibration (rotation) is not continuously distributed, but presents discontinuous energy levels. According to the energy level formula of the molecule, there are:

$$E_V = \left(V + \frac{1}{2} \right) hv \tag{2}$$

V is the vibrational quantum number, which can be any natural number; h is the Planck constant; v is the vibration frequency.

In the initial state where energy has not yet been absorbed, the vibration of most molecules is in the ground state with an initial energy of 12 hv [11]. After a certain amount of energy is given to the outside world, the molecular vibration is in the first excited state. At this time, the molecular energy is:

$$E_1 = \frac{3}{2} hv \tag{3}$$

According to the law of conservation of energy, this process must absorb energy Ehv, υ is the frequency of the electromagnetic wave absorbed by the molecule. Since the molecular vibrational energy level satisfies a discrete quantum distribution, the energy absorbed by its transition is also a discrete quantum distribution [12,

13]. When the vibrational levels of different chemical groups change, they absorb electromagnetic waves of a specific frequency. After research, the absorption frequency range of the fundamental frequency vibration (from the ground state to the first excited state) of the hydrogen-containing group (XH) in the organic molecule is in the mid-infrared region (wavelength 2.5 ~ 25 microns), and its frequency doubling (from the frequency range of the absorbed light from the ground state to a higher excited state) and the combined frequency (the two fundamental frequency vibration transitions) is in the near-infrared region (0.078 ~ 0.2526 microns) [14]. When near-infrared light is irradiated to the inside of the sample and subjected to multiple diffuse reflections or transmissions, the detector receives and returns the spectrum carrying sample component information in this frequency range. Measurement and analysis can obtain information related to the characteristics of the sample [15, 16].

2.2. Diffuse Reflection Spectrum Analysis Technology

Near Infrared Spectroscopy (NIRS) is based on the relationship between the absorbance of the near-infrared light of a sample to be measured and the number of particles of light-absorbing substances, that is, a specific frequency (or wavelength) in the sample. The absorbance of electromagnetic waves is directly proportional to the number of particles of a specific substance [17, 18]. Considering the proportionality factor and the optical path in the proportional relationship, the strict mathematical relationship is called Lambert-Beerlaw [19]. Its expression is:

$$A = -\lg \frac{I}{I_0} = \varepsilon bC \tag{4}$$

Among them, I0 is the parallel and uniform incident light intensity with a wavelength of λ, ε is the molar absorption coefficient, I is the light intensity after transmission spectrum, A is the absorbance, b is the optical path, and C is the concentration of the sample to be measured.

The Lambert-Beer law states that for monochromatic light of a certain wavelength λ, the absorbance A of the substance to be measured is proportional to the optical path b and the concentration C, and its proportionality constant ε is called the molar absorption coefficient. Therefore, the absorbance A of the test substance can be regarded as a multivariate function [20, 21]. For light with a specific wavelength and the substance to be measured (specific absorption coefficient ε), it can be concluded that A has a linear relationship with the concentration C of the substance to be measured, which can be used for quantitative analysis; when the optical path b and the concentration C are constant, at this time, the relationship between the A and the absorption wavelength of a specific substance is the absorption spectrum, which can be used for qualitative analysis of the spectrum [22]. From the above analysis, we know that Lambert-Beer law provides a theoretical basis for quantitative or qualitative analysis of the spectrum.

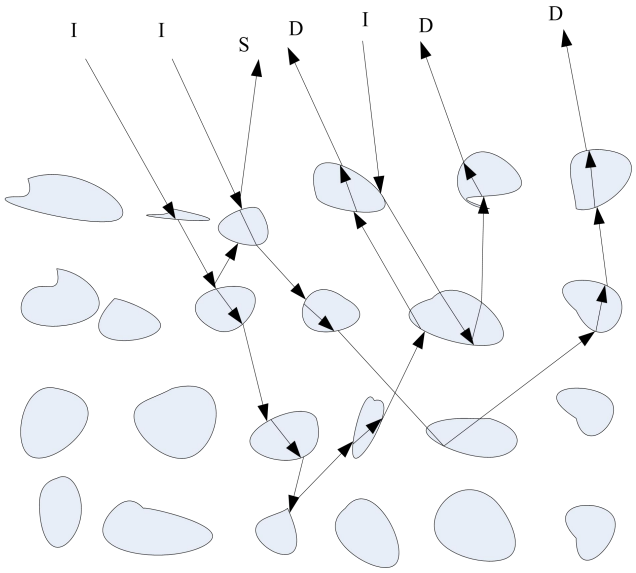


Figure 2. Schematic reflection of shrimp meat sample particles (I-incident light, D-diffuse reflection light, S-specular reflection light)

When a beam of near-infrared light is directed to a sample composed of solid particles, various effects (reflection, absorption, transmission, etc.) occur between the particles of the material and pass through the

sample. The effect of the sample cross-section in this process is shown in Figure 2. When the light is directed to the solid particle sample, specular and diffuse reflections will occur on the surface and inside of the sample. As shown in Figure 2, only the specularly reflected light on the surface of the sample particles did not enter the inside of the sample, and these near-infrared light did not carry the sample information [23]. Diffuse reflected light is the light that enters the inside of the sample through the light source, and interacts with the molecular groups in the sample particles (fully absorbs near-infrared light), and then returns to the surface of the sample, which carries a large amount of components with the sample to be measured structure-related information. The near-infrared light received and returned by the detector includes not only diffusely reflected light, but also specularly reflected light. The resulting spectrum contains interference information that is not related to the nature of the sample.

Since the applicable range of Lambert-Beer law is limited to the transmission absorbance spectrum, the relationship between the diffuse reflectance absorbance spectrum and the concentration of the absorbing component does not satisfy the above law. Through a lot of research on the diffuse reflection mechanism of light, an approximate formula for the diffuse reflectance of the sample to be measured can be derived:

$$R = 1 + \frac{K}{S} - \sqrt{\left(\frac{K}{S}\right)^2 + 2 \cdot \frac{K}{S}} \tag{5}$$

The variable relationship reflected by the above formula is also called "Kubelka-Munk" function. Define diffuse reflection absorbance based on equation (5):

$$A = -\log\left[\frac{1}{R_\infty}\right] = -\log\left[1 + \frac{K}{S} - \sqrt{\left(\frac{K}{S}\right) + 2 \cdot \frac{K}{S}}\right] \tag{6}$$

In the formula, A is the absorbance of the test substance; K is the absorption coefficient, which mainly depends on the chemical composition and molecular structure of the shrimp meat sample; S is the scattering coefficient, which mainly depends on the particle size of the shrimp meat sample. It can be known from equation (6) that the functional relationship curve between A and KS passes through the origin.

2.3. Near-infrared Spectrometer

Near-infrared spectrometer is an instrument used to measure the absorption of near-infrared light by a substance. Its main components are radiation source, spectroscope, sample cell, detector, acquisition and analysis system. The Fourier near-infrared spectrometer is designed based on the principle of optical interference and Fourier transform. 6 is a detector, 7 is a signal amplifier, and A / D and D / A are analog-to-digital and digital-to-analog converters, respectively. After the light source is emitted from 1, the interference light is formed by the structure of the Michelson interferometer, and is converted into an electrical signal by the detector. The abscissa of the signal is the moving distance of the moving mirror. Then, the discrete Fourier transform is performed in a computer, and then the converted digital signal is converted into an analog signal to obtain the final (with the wavelength as the abscissa) spectrum chart.

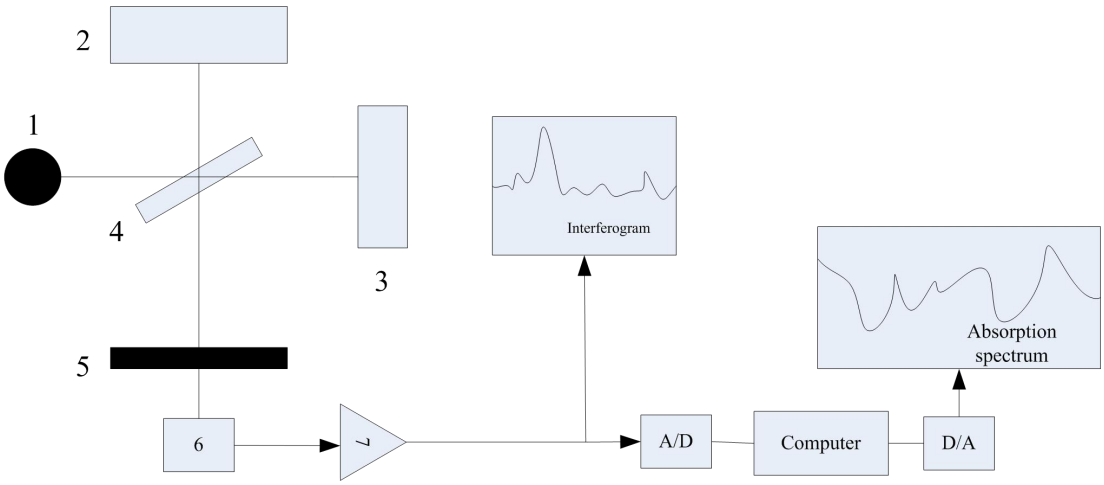


Figure 3. Schematic diagram of the Fourier near-infrared spectrometer

At present, NIR technology has become one of the indispensable and important analysis methods for product quality monitoring and analysis in the production process of industry, agriculture and so on. This is inseparable from the related features of NIR technology. Its unique advantages include the following aspects:

1) Convenient detection. Because the absorption intensity of NIR spectrum is weak, for most samples, detection and analysis can be performed without any pretreatment. Does not damage the sample, does not require reagents, and has no pollution to the environment. For example, for solid samples, a diffuse reflection measurement method can be used to directly detect and analyze the sample without damaging the sample and requiring no chemical reagents. It is an environmentally-friendly analytical technique. However, in order to obtain more accurate measurement results, samples are sometimes required, such as crushing or grinding.

2) Fast analysis speed and high efficiency. It only takes a few seconds for the sample to be detected by NIR technology (using a Fourier transform near-infrared spectrometer), and even a traditional grating-type near-infrared spectrometer can be used for detection in just a few minutes. At the same time, the reproducibility and reproducibility of NIR are better than traditional analysis methods.

3) Multiple components can be detected simultaneously. Under the same detection method, NIR can quantitatively measure each component contained in the sample at the same time. For example, in the milk powder test mode, the protein, moisture, and fat content can be measured at the same time, which greatly simplifies the determination process. Although different components have different degrees of influence on the test results, the chemical quantitative calculation software in NIR technology can use mathematical methods eliminating this effect is the main reason why NIR technology can be widely applied.

4) NIR instrument has low cost, which is very suitable for online analysis in the production process, which can effectively guarantee the stability of product quality. The optical material used in the near-infrared spectrometer is glass or quartz, and the cost of measuring accessories is low. NIR technology can also be transmitted through quartz light solder, suitable for remote online analysis of toxic materials or harsh environments.

However, NIR technology has its limitations:

1) The qualitative and quantitative analysis of NIR technology depends almost entirely on the established mathematical model. But mathematical models often need to be established separately for different samples, which requires a lot of manpower and material resources. Moreover, the model is not built once and for all. In practical applications, if you encounter non-model samples, you need to adjust the model according to the composition and properties of the sample to be measured.

2) The model has high requirements on the matching of the instrument, that is, the model cannot be used on another NIR instrument unless the performance parameters of the two instruments are the same. Even if the instrument is replaced with an accessory, Kenneng will have a huge impact on the model.

3) The absorption coefficient of the sample in the near-infrared region is small, and the detection limit of the sample is high, usually 0.1%, which is not suitable for the analysis of trace amounts of samples.

3. Experiments

3.1. Main Process of Antarctic Krill Meat Analysis by Near Infrared Spectroscopy

The process of near-infrared spectroscopy analysis of Antarctic krill meat mainly includes: collection and preparation of experimental samples, cleaning of modeling data, construction of qualitative analysis models, and collection and prediction of unknown samples, as shown in Figure 4.

(1) Collection and preparation of experimental samples

Several shrimp meat samples of different quality were collected in the market and breeding grounds, minced with a meat grinder to form a meat emulsion, each sample was weighed by a precision electronic scale, and the uniformity of the samples was strictly controlled. All samples were packed in ziplock bags and placed in a refrigerator at 4 ℃ for refrigeration. Finally, the Fourier near-infrared spectrometer and portable near-infrared spectrometer were used to collect and record the sample data.

(2) Modeling data cleaning mainly includes denoising and dimensionality reduction processing of shrimp meat sample data. Mean centering, smoothing of movement, standard normal transform, scattering correction and derivative method are used for denoising; principal component analysis and Fourier transform are used for feature extraction. The whole process is based on a unified verification platform for support vector machines.

(3) Construction of qualitative analysis model using the cleaned data, a shrimp meat quality discriminant model based on support vector machine, random forest, BP neural network and width learning system is established. The spectral data of the test samples are used as the input of the qualitative analysis model, and the output results are used as the evaluation basis to analyze the performance of the model.

(4) Collection and prediction of unknown samples. The trained model is introduced into a portable spectrometer analysis system. A portable near-infrared spectrometer was used to collect the unknown shrimp meat sample spectrum and input it into the model to obtain prediction results.

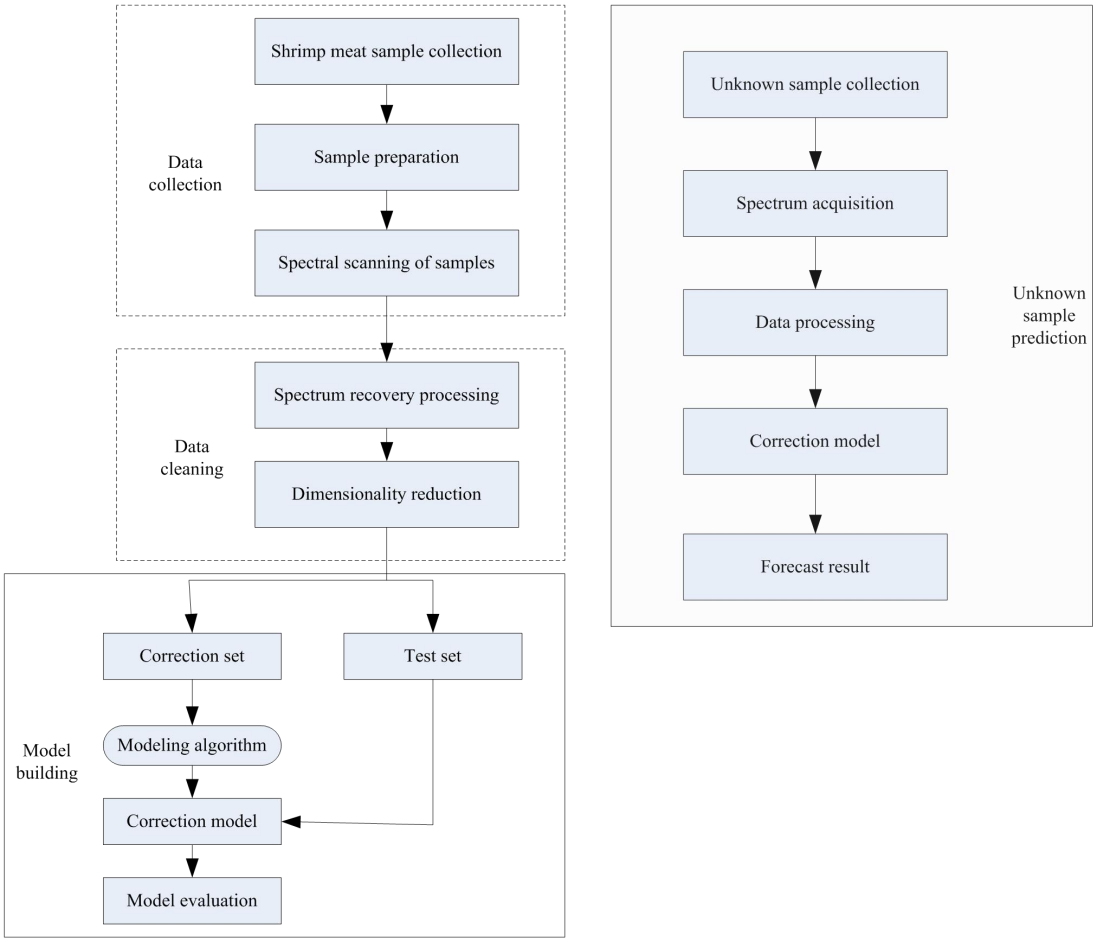


Figure 4. Basic process of near-infrared spectroscopy

3.2. Collection of Antarctic Krill Meat Samples by Near Infrared Spectroscopy

Samples of shrimp from healthy and sick shrimp were collected from markets and farms. The sample was prepared according to the national standard "Meat and Meat Products: Sampling Method" GB / T9695. The shrimp meat spectrum was collected using an Antaris Ⅱ Fourier transform near-infrared spectrometer. It is 4cm-1 and the number of wavelength points is 1557. As shown in Table 1, a total of 105 groups of spectra were collected, of which 52 were healthy shrimp and 53 were dead shrimp. The software TQAnalyst and MATLABR2014a were used to analyze and process the spectral data.

Table 1. Number and number of shrimp of different quality

Quality	Health	Die
Number of samples	52	53
Sample number	1-52	53-105

3.3. Evaluation Methods and Indicators of Experimental Models

This experiment uses cross-validation to judge the merits of the model. Cross-validation methods include simple cross-validation and K-fold cross-validation. The so-called simple cross-validation method is to divide the data sample into two subsets according to a certain ratio, one of which is used as the training set of the model and the other is used as the test set to evaluate the pros and cons of the model; and K-fold cross-validation is to divide all samples into K Disjoint subsets, using any subset as the test set of the model, and the remaining samples as the training set of the model, train K models, and calculate the average value of the simple cross-validation accuracy obtained by K predictions as the model Forecast accuracy. For K-fold cross validation, its model accuracy can be expressed by the formula:

$$Acc = \frac{1}{K}\sum_{n=1}^{k} Acc_n \tag{7}$$

In particular, when the size of K is the number of samples in the model, the cross-validation method at this

time is also called the leave-one cross-validation method. In the leave-one cross-validation method, each sample is used as a test set in turn, and all the other samples are used as a training set. Then at this time, the ratio of the number of correctly predicted samples to the number of all sample sets is the accuracy of the prediction, that is,

$$Acc = \frac{N_{correct}}{N_{sample}} \tag{8}$$

4. Discussion

4.1. Materials and Spectral Analysis

In order to reflect the generality and representativeness of the materials, this experiment purchased about 2kg of fresh quality shrimp meat from the vegetable market and crushed the shrimp meat to a minced meat with a mixer to bring it back to the laboratory for a total of two days. In order to maintain the stability of the instrument during the experiment, the spectrometer and computer were connected to warm up for 30 minutes before the acquisition, and background correction was performed every 30 minutes during the measurement. The relevant parameters of the instrument are set as follows: the integration time is 10.7ms; the wavelength range is 908.1nm ~ 1676.2nm (wavenumber: 5965.9cm^{-1} ~ 11012cm^{-1}); the number of scans is 5; the wavelength number is 125.

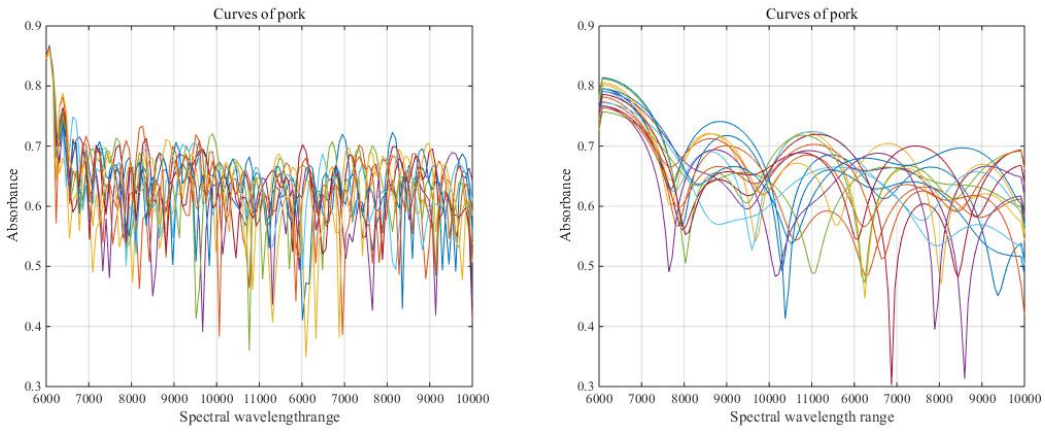
The shrimp meat purchased on the first day was divided into 42 groups of samples (including 22 groups of head and chest meat and 20 groups of tail foot meat), which were filled into glass Petri dishes and compacted, and the spectra at 4 different points were measured and the average value was taken as the spectrum of this sample. After all the samples are processed, they are stored in the refrigerator uniformly and the temperature is controlled at about 4 ℃. After storing for about 24 hours, the sample spectral data was measured in the same way. The second purchase of fresh shrimp was divided into 31 groups (including 13 head and chest meat groups and 18 foot and foot meat groups). It should be noted that the second purchase of shrimp meat samples did not collect spectral data after 24 hours of storage.

The quality and quantity information of the experimental samples is shown in Table 2. A total of 115 sets of data were collected and expressed as matrix X115x125. In these two different types of samples, there are two obvious peaks: 1212nm (8250.825cm^{-1}) and 1453nm (6882.312cm^{-1}). From the perspective of the spectral absorption mechanism, the two peaks are generated and the absorption group CH. It is closely related to the stretching vibration of CH2, and CH and CH2 represent the content of specific substances in shrimp meat, so it can be concluded that by analyzing the content of the substance and dividing the sample feature space according to a certain threshold, the test result can be obtained. The nature of shrimp meat samples.

Table 2. Collection of experimental samples

Quality	Fresh shrimp		24h frozen shrimp	
Category	Head breast	Tail foot meat	Head breast	Tail foot meat
Number	35	38	22	20

The spectrograms of the head and tail meat are shown in Figure 5 (the left is fresh head and tail meat, and the right is the head and tail meat left for 24 hours).



(a) (b)

Figure 5. Pork spectra for the first and second days

The spectrum curve of the collected sample is shown in Figure 5. From the comparison of the figure, it can be found that the two samples with different properties (the spectrum of the healthy shrimp meat sample and the

spectrum of the dead shrimp sample on the right) have the same curve trend, which reflects that the fresh shrimp and the dead shrimp sample they are share common characteristics. However, different characteristics appear in some bands, which reflects the different physical and chemical characteristics of samples with different properties. For example, in the 5900-5600cm-1 (1694.9-1785nm) spectral region, the spectrum of healthy shrimp and dead shrimp shows different characteristic peaks. The absorption band in this spectral region depends on the stretching vibrations of CH and CH2 related to fatty acid content. Among them, the prominent peak at 1710nm originates from the first overtone of CH stretching; in the 4400-4200cm-1 (2272.73-2380.95nm) spectral region, compared with the healthy shrimp meat spectrum, the absorbance of sick and dead shrimp meat is smaller, and this spectrum absorption band in the region is related to the CH vibration of the fatty acid. The absorbance values at 2280nm, 2325nm, and 2352nm correspond to the stretching of the CH group and the deformation of the CH group. These spectral characteristics not only contain the characteristic information that distinguishes dead shrimp from healthy shrimp, but also explains the rationality of the spectral differences between the two quality shrimps from the perspective of the spectral absorption mechanism.

Fresh shrimp meat samples and frozen shrimp meat samples stored for 24 hours are shown in Figure 6.

(a) (b)

Figure 6. Pictures of shrimp meat samples on the first and second days (the first picture on the left and the second day on the right)

It can be seen from Figure 6 that it is difficult to discern the difference between the shrimp samples on the first and second days. From Figure 6, it can be seen that in the data of the first day and the second day, the spectra of head breast meat and tail foot meat show a large difference, so it can effectively distinguish the two types of shrimp meat. (The two figures respectively show that the upper part is the head and breast meat, and the lower part is the tail and foot meat). However, there is no obvious difference between the left and right pictures, so whether the different freshness of shrimp can be effectively distinguished needs further analysis and discussion.

4.2. Analysis of Modeling Results

In order to verify the validity of the classification and analysis of samples with different properties of the spectrum, principal component analysis was performed on all sample data. The contribution rate of the principal components is shown in Table 3. The contribution rate of the first 8 principal components reaches 99.998%, of which the first 3 principal components can represent 99.938% of the original spectral information, which can analyze most of the information in the original spectral data, and visualize the 3 principal component information with a three-dimensional coordinate chart. The ground is shown in Figure 7.

Table 3. Table of contribution ratios of the first 8 principal components

Main ingredient	Contribution rate (%)	Cumulative contribution rate (%)
PC1	97.437	97.437
PC2	1.422	98.870
PC3	1.046	99.927
PC4	0.023	99.986
PC5	0.014	99.983
PC6	0.0056	99.996
PC7	0.0015	99.961
PC8	0.0012	99.987

It can be found through observation that the fresh head chest and the frozen head chest have strong separability, the distance between the two clusters is small, and the distance between the clusters is large; and

the fresh tail foot and the frozen tail foot have a large overlap and are difficult to distinguish. Therefore, further analysis and discussion are needed through modeling.

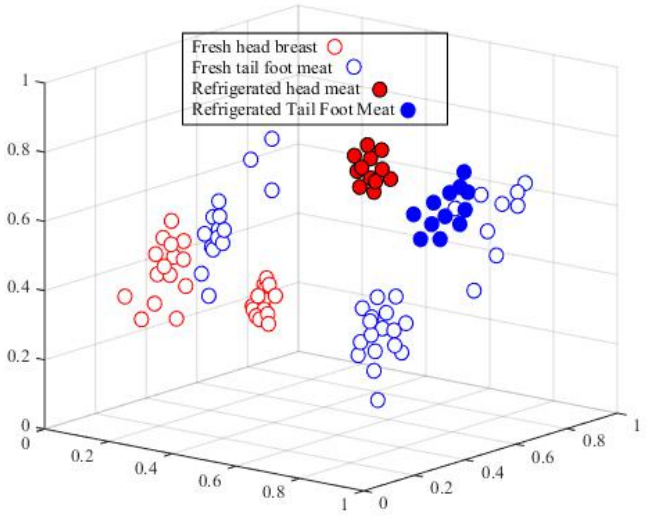


Figure 7. Scatter plot of 3 principal components after principal component analysis

It can be seen from the figure that the spectral data of the sample after the principal component transformation has a certain separability, which indicates that the data after dimensionality reduction can reflect the category characteristics of the sample to a certain extent. However, very few sample points of different categories still overlap each other, and further classification models need to be established. The distribution of shrimp meat data shows the same characteristics, that is, the overlap between frozen head breast meat and fresh head breast meat is high, and subsequent analysis models need to be established for further classification.

In order to select the best number of principal components, the first l highest-scoring principal component (l = 1, 2, 3, 4, 5, 6, 7, 8) is used as the input of the SVM model and the classification is discriminated. The accuracy rate of 50% cross-validation is used as the basis for selecting the best principal component. As shown in Figure 8, when l> 3, the accuracy rate is maintained at about 87.5%. The contribution rate of the first four principal component scores has reached more than 99.86%, which can analyze almost all the information of the spectrum, so the first four principal components are selected as the input of the model in subsequent experiments.

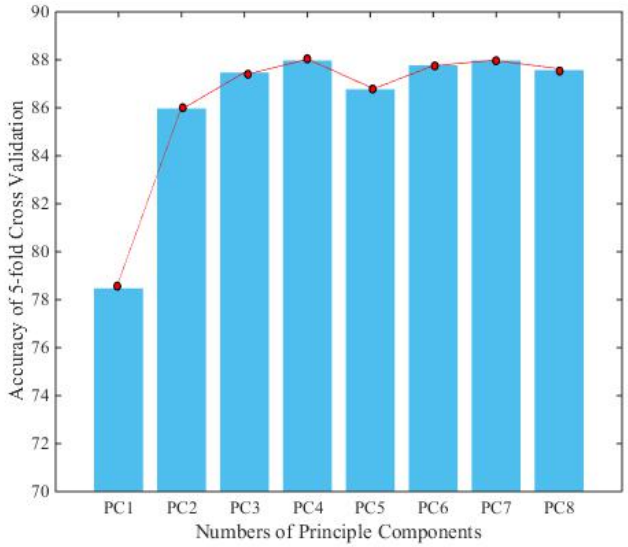


Figure 8. The effect of the number of principal components on the accuracy of SVM 5-fold cross-validation

From the perspective of dimensionality reduction, using the original spectrum modeling, whether it is a single accuracy index or the overall accuracy rate, is more stable than the dimensionality-reduced data modeling

effect, but the time consumed can be more than double the other two methods; from considering different qualitative model perspectives, although the highest positive class (healthy shrimp) samples of the random forest qualitative model have a 5-fold average correct rate higher than the support vector machine model, but in combination with the specific application background, a negative class (refrigerated shrimp) sample is required only with high recognition accuracy can consumers' rights be protected. Therefore, from a practical perspective, it can be concluded that the random forest model is inferior to the support vector machine model. At the same time, it can be found from the table that the overall recognition accuracy of SNV + FT + SVM reaches 96.78%, the accuracy of negative samples reaches 96.90%, and the model training + prediction time is 4.07s, indicating that the learning method can better fit the non-linear relationship between the spectral data and the label of the sample property indicates that the portable near-infrared spectrometer combined with machine learning has a good prospect in the detection of shrimp freshness.

5. Conclusions

This article takes Antarctic krill meat as the research object, and uses near-infrared spectroscopy to carry out a rapid analysis of the main quality indicators of calcium, protein, moisture and oil content in Antarctic krill meat. The correlation coefficients and relative analysis error values of the prediction results of the calibration index and validation set of the model on the various research indicators show that the model can accurately predict the calcium, protein, oil content and moisture content in the Antarctic krill meat, and an infrared spectrometer can be used to detect the quality of Antarctic krill. Combined with the chemical measurement of the quality index, the conditions affecting the results of the model: spectral scattering processing, mathematical processing parameters, and regression techniques were investigated, and the optimal modeling condition parameters were selected to establish a quantitative detection model of Antarctic krill meat quality.

The near-infrared spectrum of shrimp meat quality analysis based on machine learning (including BP neural network, random forest, support vector machine, and width learning) has a higher accuracy rate (the accuracy of width learning is weak, compared to the other three types) The method is not suitable for establishing an analysis model for the time being), and MSC + SNV + random forest achieved the highest average accuracy rate of 50% off 89.05%, which has certain applicability. However, because the support vector machine model is more accurate for detecting negative samples than the random forest model, and the accuracy of detecting negative samples is more important for ensuring food safety. After comprehensive consideration of various factors, this experiment considers the standard to be positive. The combination mode of state variable transform + discrete Fourier transform method + support vector machine is more suitable for constructing a qualitative analysis model of shrimp near-infrared spectrum than random forest and BP neural network.

In the practical analysis of near-infrared spectroscopy, a narrow range model can be established for 50 to 100 calibration samples, and at least 150 calibration samples are required for a wide range or open sample group. The sample composition content range directly affects the accuracy of the analysis and the scope of the model. In this paper, a near-infrared analysis model of Antarctic krill meat quality was established using near-infrared spectroscopy analysis technology. The selected samples came from all over the country and were concentrated in the Central Plains. The range of moisture and protein content was relatively broad. Protein and crude fat the narrow range of calcium content and calcium content makes the calibration model not ideal. This subject is limited by the number of samples and other factors, and the prediction effect can be improved by adding the number of calibration samples.

Acknowledgements

This work was supported by National Key Research and Development Program of China (Grant No. 2018YFC1406800), National Natural Science Foundation of China (Grant No.31571915).

References

[1] Wildan, D. M., Affandi, R., Pratiwi, N. T. M., Krisanti, M., Ayu, I. P., & Iswantari, A. (2017) "Evaluation of Karst Water Quality as an Early Reference of Land Suitability Mapping for Vaname Shrimp (Litopenaeusvannamei) Culture Media", I*OP Conference Series Earth and Environmental Science*, 54(1), pp.e012070.

[2] Winkler-Moser, J. K., Singh, M., Rennick, K. A., Bakota, E. L., Jham, G., Liu, S. X., & Vaughn, S. F. (2015) "Detection of Corn Adulteration in Brazilian Coffee (Coffea arabica) by Tocopherol Profiling and NIR Spectroscopy", *Journal of Agricultural & Food Chemistry*, 63(49), pp.10662-10668.

[3] Luo, Z., Qin, Y., & Ye, Q. (2015) "Effect of Nano-Tio2-LDPE Packaging on Microbiological and Physicochemical Quality of Pacific White Shrimp during Chilled Storage", I*nternational Journal of Food Science & Technology*, 50(7), pp.1567-1573.

[4] Kühnel, A., & Bogner, C. (2017) "In−Situ Prediction of Soil Organic Carbon by Vis–NIR Spectroscopy:

An Efficient Use of Limited Field Data", *European Journal of Soil Science*, 68(5), pp.689-702.
[5] Jiang, J. F., Zhao, M. Y. (2005) "Rapid Determination of Total Volatile Acid and Base in Tobacco with NIR Spectroscopy", *Tobacco Science & Technology*, 49(9), pp.50-56.
[6] Han, S., Zhang, W., Li, X., Li, P., & Liu, J. (2016) "Determination of Three Alcohols in Chinese Dukang Base Liquor by FT-NIR Spectroscopy", *Food Analytical Methods*, 9(8), pp.2194-2199.
[7] Santos, A. J., Anjos, O., & Pereira, H. (2015) "Estimation of Acacia melanoxylon unbleached Kraft pulp brightness by NIR spectroscopy", *Forest Systems*, 24(2), pp.10-20.
[8] Ying, Y., Wei, W., Xuan, C., & Mingjie, X. (2015) "Detection of Moldy Corns with FT-NIR Spectroscopy Based on SVM", *Journal of the Chinese Cereals & Oils Association*, 30(5), pp.143-146.
[9] Chen, Z. G., Li, X., & Fan, X. J. (2016) "Method for the Discrimination of the Variety of Potatoes with Vis/NIR Spectroscopy", *Guang Pu Xue Yu Guang Pu Fen Xi = Guang Pu*, 36(8), pp.2474-2478.
[10] Qin, H., Ma, J. Y., Chen, S. J., Yan, Y. L., Li, W., Wang, P., & Liu, J. (2016) "Identification of Haploid Maize Kernel Using NIR Spectroscopy in Reflectance and Transmittance Modes: A Comparative Study", *Spectroscopy and Spectral Analysis*, 36(1), pp.292-297.
[11] Li, W. L., & Qu, H. B. (2016) "Research Progress on Standardization Study of NIR Spectroscopy Based Method for Quality Control of Traditional Chinese Medicine", *China Journal of Chinese Materia Medica*, 41(19), pp.3511-3514.
[12] Huang, X., Guan, C., Ding, R., & Lü, R. (2015) "Freshness Evaluation of Sea Bass Using Multi-Sensor Information Fusion Based on Olfactory Visualization and NIR Spectroscopy Technique", *Transactions of the Chinese Society of Agricultural Engineering*, 31(8), pp.277-282.
[13] Biancolillo, A., De Luca, S., Bassi, S., Roudier, L., Bucci, R., Magrì, A. D., & Marini, F. (2018) "Authentication of an Italian PDO Hazelnut ("Nocciola Romana") by NIR Spectroscopy", *Environmental Science & Pollution Research*, 25(29), pp. 28780-28786.
[14] Zhang, D. F., Liu, Z. C., Zhang, T. G. (2017) "Intrinsic Quality Evaluation of Tipping Paper by FT-NIR Spectroscopy Pattern Recognition", *Tobacco Science & Technology*, 50(11), pp.39-47.
[15] Xu, J., Huang, F., Li, Y., Chen, Z., & Wang, Y. (2016) "Rapid Detection of Total Nitrogen Content in Soy Sauce Using NIR Spectroscopy", *Czech Journal of Food Sciences*, 33(6), pp.518-522.
[16] Steidle Neto, A. J., Moura, L. D. O., Lopes, D. D. C., Carlos, L. D. A., Martins, L. M., & Ferraz, L. D. C. L. (2017) "Non−Destructive Prediction of Pigment Content in Lettuce Based On Visible–NIR Spectroscopy", *Journal of the Science of Food & Agriculture*, 97(7), pp.2015-2022.
[17] Calero, A. M., Muñoz, E., Pérez-Marin, D., Riccioli, C., Pérez, L., & Garrido-Varo, A. (2018) "Evolution of Frying Oil Quality Using Fourier Transform Near-Infrared (FT-NIR) Spectroscopy", *Applied Spectroscopy*, 72(10), pp. 1001-1013.
[18] Dematte, J. A., Araujo, S. R., Fiorio, P. R., Fongaro, C. T., & Nanni, M. R. (2015) "VIS-NIR-SWIR Spectroscopy in Soil Evaluation Along a Toposequence in Piracicaba", *Revista Ciencia Agronomica*, 46(4), pp.679-688.
[19] Dai, L., Liu, C., Han, X., Wang, L., Tan, C., Yan, Z., & Xu, Y. (2017) "Dopant Occupancy and UV-VIS-NIR Spectroscopy of Mg (0, 4, 5 And 6 Mol. %), Pp. Dy: Linbo 3 Crystal", *Modern Physics Letters B*, 31(25), pp.1750232.
[20] Tomassetti, M., Marini, F., Bucci, R., Coppa, A., & Campanella, L. (2017) "Comparison of NIR Spectroscopy Coupled to Chemometrics and Derivative Thermogravimetry for Relative Dating of Human Fossil Bones", *Journal of Thermal Analysis & Calorimetry*, 130(1), pp. 559-565.
[21] Jiong, G. E., Wang, J., Wang W. M., & Zhang, J. P. (2006) "Application of NIR Spectroscopy in Discriminating Counterfeit Cigarettes from Genuine", *Tobacco Science & Technology*, 50(12), pp.77-84.
[22] Zhu, H. Y., Fu, X. P., You, G. R., & He, J. C. (2015) "Application of NIR Spectroscopy for Nondestructive Qualitative and Quantitative Analysis of Lotus Seeds", *Spectroscopy & Spectral Analysis*, 35(10), pp. 2752-2756.
[23] Shi, R. J., Pan, X. Z., Wang, C. K., Liu, Y., Li, Y. L., & Li, Z. T. (2015) "Prediction of Cadmium Content in the Leaves of Navel Orange in Heavy Metal Contaminated Soil Using VIS-NIR Reflectance Spectroscopy", *Guang Pu Xue Yu Guang Pu Fen Xi = Guang Pu*, 35(11), pp.3140-3145.