IOP Conf. Series: Materials Science and Engineering 466 (2018) 012064 doi:10.1088/1757-899X/466/1/012064

Qualitative Identification of Pesticide Residues in Pakchoi Based on Near Infrared Spectroscopy

Min Li, Xiaoying Zhang and Qiang Jiang

School of Physics and Electronic Engineering, Leshan Normal Uniersity, Leshan 614000, China
Email:cassie_li@163.com

Abstract. This paper provided a new method for rapid and nondestructive classification and identification of vegetable residues. Near Infrared Spectroscopy (NIR) of two kinds of pakchoi were collected. One kind of pakchoi was non-pesticide. And the other was low concentration of pesticide residues . In this paper, wavelet soft threshold was used to pre-treat the spectral data. Then Principal Component Analysis (PCA) was run to reduce the dimensionality of spectral data. Next, Linear Discriminant Analysis (LDA) and K-Nearest Neighbor Classification and Recognition (KNN) was run to classify and discriminate. The experimental results show that the prediction rate is 95%. This method is accurate and effective.

1. Introduction

Pakchoi is a common pakchoi, commonly known as cruciferae, brassica. Brassica evolved from the pakchoi subspecies, is a variety of non-heading pakchoi, annual herbs. Pakchoi is cool and cold, and it is more tolerant to low temperature and high temperature. It can be planted almost all the year round. Pakchoi is rich in protein, dietary fiber, calcium, phosphorus, iron and vitamins, etc. It tastes sweet and fragrant, and is favored by consumers. In the cultivation process of pakchoi, it is easy to suffer insect pests such as pakchoi caterpillar, fly insect, aphid, powdery mildew, soft rot, root swollen disease and so on. In order to ensure the yield and sales performance of pakchoi, vegetable farmers usually spray pesticides from time to time, or even spray pesticides with excessive concentration, resulting in excessive pesticide residue on the surface of pakchoi leaves. It not only harms consumers' health, but also affects the export of pakchoi.

At present, the main methods used in the detection of pesticide residues in vegetables are chromatography, mass spectrometry, and chromatography-mass spectrometry [1-3]. Although the precision of these methods is high, the pretreatment of samples is very tedious, often need to squeeze vegetables into juice, but also need to add some chemical reagents. Because of time-consuming and high detection costs of sample preparation, it is difficult to realize the vegetables residue fast and nondestructively now[4-7]. It is an urgent problem to develop a simple, rapid and non-destructive method for detecting vegetable residues.

Near Infrared Spectroscopy (NIR) is a kind of green technology which can realize rapid nondestructive detection and multi-component simultaneous detection[8,9]. In recent years, with the development of near infrared spectroscopy software and hardware, it has become a hot research topic in food safety detection [10]. Taking pakchoi as the research object and combining with near infrared spectroscopy analysis technology, this paper presented an intelligent algorithm for qualitative identification of two kinds of pakchoi, which were pesticide-free and low concentration pesticides. The accuracy rate reached 95%. A feasible method for rapid non-destructive detection of pesticide

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

residues in vegetables was proposed.

2. Experiment

2.1. Instruments and Reagents

The Agilent Cary 630 FTIR was selected for NIR spectrometer. ATR attenuated total reflectance mode was chosen to collect spectrum. Background scanning were 64 times. Sample scanning were 64 times. Resolution was 8cm-1. Spectrum acquisition also required a computer, and the installation of supporting software: Microlab PC and Resolutions Pro.

The absolute alcohol was selected to clean spectrometer crystal. The pesticides were produced by Shandong Shenda Crop Science and Technology Co., Ltd. The formulation of the pesticides, was micro-emulsion, and the total effective component content was 5%.

2.2. Sample Preparation

In order to ensure the universality of the experiment, pakchoi was planted in the experimental field to ensure that no pesticides had been sprayed before. Two groups of pakchoi were selected, one group was used to produce pesticide-free samples, and the other group was sprayed with 1:500 high-efficiency cyhalothrin solution as samples containing mild pesticides. In order to ensure the uniformity of pesticide ,spraying method was adopted.

2.3. Experimental Description

After spraying pesticide for 24 hours, the pakchoi was picked for spectral data collection. 40 samples of non-pesticide and pesticide-containing Pakchoi leaves were selected to collect near-infrared spectroscopy data. Each spectral data acquisition was strictly carried out in the following steps: the first step was to clean the crystal with absolute alcohol; the second step was to collect the background spectrum; the third step was to put the pakchoi sample on the spectrometer and collect the sample spectrum. Each type of pakchoi collected 40 spectral data, totally collecting 80 spectral data. The dimension of each data was 1942. The range of Near infrared light wave-number was of 7800-400cm-1.

2.4. Spectral Data Analysis

The spectral data processing environment was MATLAB R2016a. The data processing flow is shown in Figure 1. Spectral data preprocessing has two main purposes: one is to reduce noise, reduce the impact of various interferences, improve the robustness of the subsequent analysis model and the accuracy of the prediction results; the other is to compress the data to improve the speed of modeling. Multiplicative Scatter Correction (MSC) is the most commonly used pretreatment method. When the sample absorbance and concentration maintain a good linear relationship, MSC correction effect is good; but in practice, the scattering of sample particles on the near infrared spectrum is not linear, at this time MSC treatment effect is not good. Wavelet analysis is called "mathematical microscope" and has its own unique features in denoising and detail preservation. In this paper, wavelet soft thresholding method was innovatively applied to spectral data preprocessing, which laid a good foundation for subsequent model classification and recognition.

Principal component analysis (PCA), also known as abstract factor analysis, mainly aims to reduce the dimension of data. Data need to be further analyzed, such as discriminant analysis, clustering analysis and so on. PCA uses the idea of data dimensionality reduction to divide the original variables into a few comprehensive variables, and the integrated variables are linear combinations of the original variables, thus eliminating a lot of redundant information. The new variables can maximize the data characteristics of the original variables, and there is no information loss [5]. In this paper, PCA was used to reduce the principal component analysis to 18-dimensional data from 80 samples of two types of Pakchoi and 18 principal components were obtained.

Linear Discriminant Analysis (LDA) , also known as Fisher linear discriminant, is a classical algorithm for pattern recognition. The NIR data of pakchoi were dimensionality reduced by PCA, and then LDA was used for feature extraction effectively. The basic idea of LDA was to project high-dimensional sample vectors into the optimal discriminant vector space to extract classification information and compress the dimension of feature space. After projection, we guaranteed that the sample vectors have the largest inter class distance and the smallest intra class distance in the new subspace.

K-Nearest Neighbor Classification and Recognition (KNN) method calculates the distance between each sample and each training sample one by one. Here we chosen Euclidean distance to find out the most recent K decisions. The accuracy of KNN classification results has a great relationship with K value. There is no rule to follow in the selection of K value at present. Here, the best K value was 3 by experimental verification.

3. Results and Analysis

3.1. Experimental Result

The spectra of 80 pakchoi samples collected were shown in Figure 1. In this paper, three algorithms were used to qualitatively identify two kinds of pakchoi. Algorithm 1: Spectral data without preprocessing, directly been analyzed by PCA, LDA discrimination, and KNN classification. Algorithm 2: Spectral data after MSC preprocessing, and then been analyzed by PCA, LDA discrimination and KNN classification. Algorithm 3: Spectral data after wavelet soft threshold preprocessing, then been analyzed by PCA, LDA discrimination and KNN classification. The spectrum of wavelet soft threshold preprocessing is shown in Figure 2. The comparison among the three algorithms is shown in Table 1.

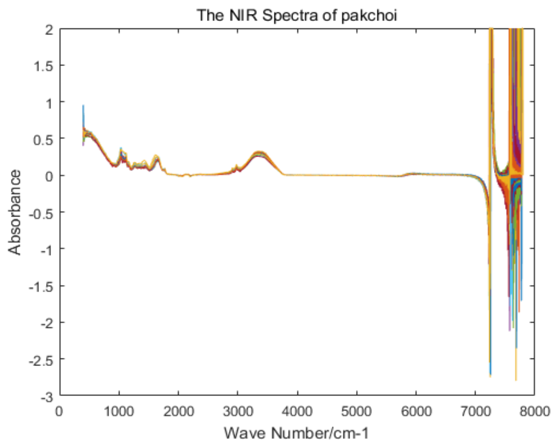

Figure 1. Primary NIR spectra of Pakchoi

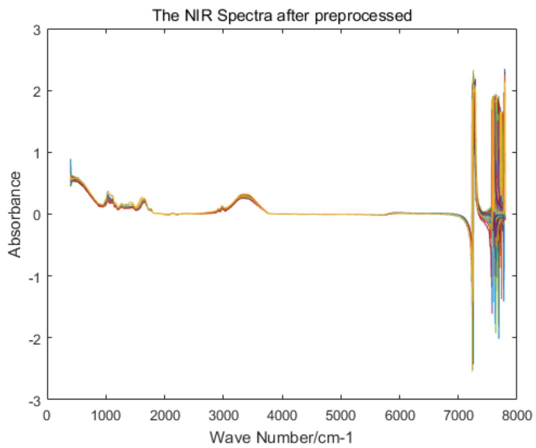

Figure 2. Spectrum of Pakchoi after preprocessing with wavelet soft threshold

Table 1. Qualitative identification results of pakchoi

Algorithms	Number of Categories	Total sample number	Training set sample size	Test set sample size	Prediction rate(%)
Algorithm 1	2	80	60	20	90
Algorithm 2	2	80	60	20	90
Algorithm 3	2	80	60	20	95

3.2. Experimental Analysis

From Table 1, we can see that the accuracy of qualitative discrimination of algorithm 1 and algorithm 2 is 90%, lower than that of algorithm 3, and the accuracy of algorithm 3 can reach 95%. It is obvious that wavelet soft threshold preprocessing is superior to MSC preprocessing and non preprocessing. Wavelet preprocessing can preserve the details of the spectrum as much as possible while denoising, which improves the stability and correct recognition rate of subsequent modeling.

4. Conclusion

Pakchoi as the research object, this paper combined with near infrared spectroscopy analysis technology to Identify non-pesticide and low concentration of pesticide residues in two types of pakchoi. This paper presented an algorithm for spectral preprocessing based on wavelet soft threshold, combined with PCA analysis, LDA discrimination and KNN classification. The prediction rate reached 95%, which realized fast and accurate qualitative identification of pakchoi residues. This paper provided a new idea for qualitative classification and identification of vegetable residues and had a certain reference value.

5. Acknowledgments

This work was supported by Sichuan Provincial Education Department (NO: 18ZA0231).

6. Reference

[1] Wenxiu Li, Kexin Xu, 2004, *Determination of Pesticide Residues in Vegetables by Infrared Spectroscopy*. Spectroscopy and Spectral Analysis. 24(10) P1202-1204

[2] Minfa Liu, Lingbiao Zhang, Songlei Wang,et al. 2014, *Identification of Pesticide Species on the Surface of Chinese Date by Near Infrared Hyperspectral Spectroscopy*, Food Research and Development, 35 (15)P81-86.

[3] chaoJin Wang, Qi Cai. 2006, *Detection Trend of Pesticide Residues in Agricultural Products*.Modern Scientific Instruments, 01P106-108.

[4] Fei Shen, Zhanke Yan, Zunzhong Ye,et al. 2009,*Application of Near Infrared Spectroscopy in the Determination of Phoxim Pesticide Residues*, Spectroscopy and Spectroscopy, 29 (9)P2421-2424.

[5] Xiangyang Zhou, Chunzhong Lin, Xiangna Hu. 2004, *Rapid diagnosis of organ phosphorus pesticide residues in vegetables by near infrared spectroscopy* .Food Science, 25 (05) P151-154.

[6] Jun Sun, Xin Zhou, Hanping Mao. 2016, *Detection of pesticide residues in lettuce based on fluorescence spectroscopy*, Journal of Agricultural Engineering, 32 (19): 302-307.

[7] Sergio Armenta. 2007, *Partial least squares-near infrared determination of pesticides in commercial formulations*. Vibrational Spectroscopy, 44(2) P273-278.

[8] Josep F. Ventura-Gayete, Sergio Armenta, Salvador Garrigues, et al. 2006, *Multicommutation-NIR determination of Hexythiazox in pesticide formulations. Tlanta,* 68 P1700-1706.

[9] Jingzhu Wu, Hui Li, Cuiling Liu, et al. 2010, *Rapid qualitative determination of vegetable residues based on near infrared spectroscopy*. Food Industry Science and Technology, 31 (10)P 377-379.

[10] Cuiling Liu, Guang Zheng, Xiaorong Sun, et al. 2010,*Studies on the application of near infrared spectroscopy in the detection of pesticide residues*, Journal of Beijing Business University (Natural Science Edition), 28 (4) P52-55.