Multiple Kinds of Pesticides Detection Based on Back-Propagation Neural Network Analysis of Fluorescence Spectra
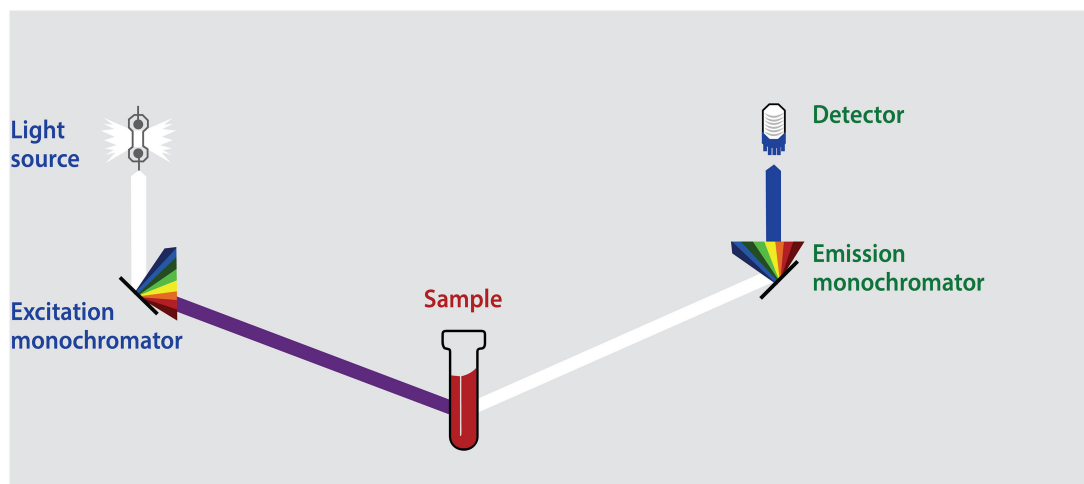
Haiyi Bian
Hua Yao
Guohua Lin
Yinshan Yu
Ruiqiang Chen
Xiaoyan Wang
Rendong Ji
Xiao Yang
Tiezhu Zhu
Yongfeng Ju

Multiple Kinds of Pesticides Detection Based on Back-Propagation Neural Network Analysis of Fluorescence Spectra

Haiyi Bian [1,2] **Hua Yao,** [1] **Guohua Lin,** [1,2] **Yinshan Yu** [1,2]
Ruiqiang Chen, [1,2] **Xiaoyan Wang,** [1,2] **Rendong Ji,** [1,2] **Xiao Yang,** [1,2]
Tiezhu Zhu, [1,2] **and Yongfeng Ju** [1,2]

[1] Faculty of Electronic Information Engineering, Huaiyin Institute of Technology, Huai'an, Jiangsu 223003, China
[2] Jiangsu Engineering Laboratory for Lake Environment Remote Sensing Technologies, Huai'an, Jiangsu 223003, China

Abstract: Fluorescence spectroscopy attracted more and more attention in pesticide residue detection field because of its advantages of non-destructive, non-contact, high speed and no requirement of complex pre-process procedure. However, given that the concentration of the pesticide detected via fluorescence spectroscopy is calculated in accordance with the Beer-Lambert law, this method can only be used to detect samples containing a single kind of pesticide or several kinds of pesticides with completely different fluorescence which is not in accordance with practical cases. In this article, to overcome this disadvantage, back-propagation (BP) neural network algorithm was introduced to detect multiple kinds of pesticides via fluorescence spectroscopy. The results from four kinds of pesticides which are usually used for fruits and vegetables indicated the effectiveness of BP neural network algorithm.

Index Terms: Pesticide residue, fluorescence spectroscopy, BP neural network algorithm.

1. Introduction

In order to increase crop production to solve the global starvation problem, varieties of pesticides have been increasing dramatically in recent decades [1], [2]. The misuse and overuse of pesticides have resulted in heavy environmental pollution, which will become a serious threat to people's health [3]–[5]. In recent years, pesticide residues in vegetables, fruit juice and drinking water attracted considerable attention in the field of food safety [6]–[8]. Many methods have been

developed to detect the pesticides such as solid-phase microextraction method [9], gas chromatography [10], high-performance liquid chromatography [11] and mass spectroscopy [12]. Most of the above methods have been demonstrated to be effective for pesticide detection, however, because of the complex pretreatment process, time-consuming and requirements of chemical reagents which means destructive to samples, these methods are not suitable for on-line detection of the pesticides.

Considering the advantages of non-destructive, non-contact and high speed, optical methods such as infrared spectroscopy [13], [14], Raman spectroscopy [15], [16] and fluorescence spectroscopy [17], [18] combined with chemometrics algorithm stand out in the field of qualitative and quantitative detection of chemical components. Considering the absorption of the water, infrared spectroscopy mainly used for the detection of the pesticides in the solid samples. For example, near infrared transmission spectroscopy was used to collect spectral data of 240 lettuce samples and two kinds of pesticides (fenvalerate and triazoline) were identified with the accuracy of training set and test set reaching 98.89% and 95.00%, respectively by Sun *et al.* [19]. For Raman spectroscopy, to realize non-destructive detection, surface sampling is adopted and has proven to be very effective by many researchers [20]– [22]. However, there are still many issues remaining regarding the reported SERS Raman spectroscopy. For example, 1) the fabrication of flexible substrates is time-consuming and their shelf life is limited [20]; 2) the instability and worse repeatability of the SERS substrates from different batches which would affect the Raman intensity. For these reasons, it seems unsuitable for infrared spectroscopy and Raman spectroscopy to realize on-line detection of the pesticides.

Fluorescence spectroscopy, which is a rapid, non-destructive and non-contact technology, has increasingly attracted attention from scholars in the field of pesticide detection [23]–[25]. Moreover, fluorescence is an inexpensive technique with high intrinsic sensitivity, and the instruments involved are easy to perform. For example, Ji *et al.* measured the fluorescence intensity at 356 nm and built the exponential prediction model based on the Beer-Lambert law to predict the concentration of boscalid in grape juice. Their results with a correlation coefficient of 0.9999, an average recovery of 101.7%, and a relative standard deviation of 2.5723% indicated that fluorescence spectroscopy is effective to detect the concentration of single pesticide [24]. However, the limit of the exponential prediction model based on the Beer-Lambert law is that the model can't be used to predict the concentration of multiple pesticides simultaneously. Attempts have been made to overcome this disadvantage. Ibanez *et al.* simultaneously quantified four pesticides (thiabendazole, fuberidazole, carbaryl and naphthyl acetic acid) in vegetables and fruits by measuring excitation-emission fluorescence matrices. Good recoveries between 80% and 115% were acquired which demonstrated that the method enabled to determine the pesticides in complex samples [26]. However, this method required not only the emission fluorescence intensity but also the excitation intensity which increases the data and decreases the speed.

In this work, to overcome the disadvantages of fluorescence spectroscopy applied in the pesticide residues field, BP neural network algorithm was introduced to quantify the concentration of different pesticides simultaneously. Four kinds of pesticides including zhongshengmycin, paclobutrazol, boscalid and pyridaben were used to demonstrate the effectiveness of the method. In total, 151 samples with different concentrations were measured. Among the 151 samples, 51 samples were selected as the training dataset and the remaining 100 samples were used to validate the models. The results indicated that the predicted concentration fitted well with the actual concentration.

2. Principle

BP neural network is one of the most traditional neural network algorithms which has been applied to many fields. The model is built by minimizing the sum of the squared errors between the actual and predicted values which is based on the gradient descent method. The configuration of traditional three layers BP neural network model is shown in Fig. 1.
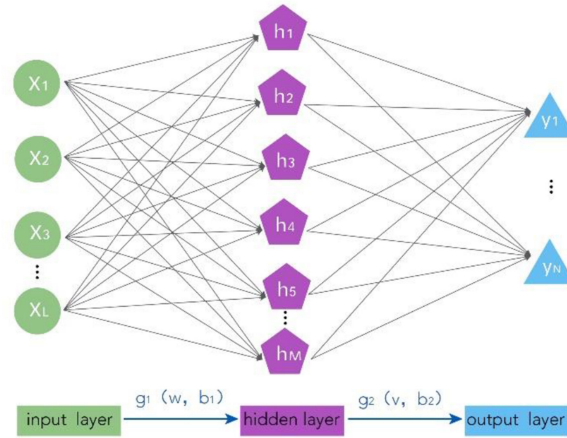
Fig. 1. The configuration of traditional three layers BP neural network model.

The three layers are input layer, hidden layer and output layer. The model between the input layer and hidden layer is built through the following equation [27]–[29]:

$$net_1 = w^T x + b_1, h = g_1(net_1) \tag{1}$$

where x is the input variable, w is the weight factor, b_1 is the offset value, h is the value in the hidden layer, and g_1 is the activation function.

The model between the hidden layer and output layer is built through the following equation:

$$net_2 = v^T h + b_2, y = g_2(net_2) \tag{2}$$

where v is the weight factor, b_2 is the offset value, y is the output value, and g_2 is the activation function.

Combining Eq. (1) and Eq. (2), the final model between the input layer and output layer can be expressed as

$$y = g_2(net_2) = g_2(v^T g_1(net_1) + b_2)$$
$$= g_2(v^T g_1(w^T x + b_1) + b_2) \tag{3}$$

To optimize the model, loss function is defined as

$$E(\theta) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - y_i)^2 \tag{4}$$

where θ represents the parameter set, Y_i is the actual value, and y_i is the output value predicted by the model.

The value of loss function is calculated and compared with the given value. If the value of loss function is larger than the given value, the weight factor and offset value are adjusted according to the following equation:

$$v^{(k)} = v^{(k-1)} - \eta \frac{\partial E}{\partial v} \tag{5}$$

$$b_2^{(k)} = b_2^{(k-1)} - \eta \frac{\partial E}{\partial b_2} \tag{6}$$

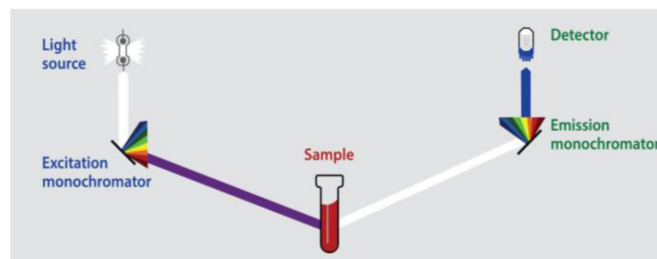$$w^{(k)} = w^{(k-1)} - \eta \frac{\partial E}{\partial w} \tag{7}$$

Fig. 2. The schematic diagram of the fluorescence spectroscopy.

$$b_1{}^{(k)} = b_1{}^{(k-1)} - \eta \frac{\partial E}{\partial b_1} \tag{8}$$

where η is learning rate and k is the number of iterations. The iterations will continue until the value of loss function is smaller than the given value or the validation checks are larger than the set value which means the error of the validation set will not decrease with the training.

In this work, the data in the input layer is the fluorescence spectra obtained from different concentrations. Since the fluorescence spectrum consists of the fluorescence intensities at 801 wavelengths, the number of nodes in the input layer is 801. The data in the output layer is the concentration of the pesticide corresponding to the fluorescence spectra in the input layer.

3. Experiment

3.1 Equipment

All the experiments presented in this work were done on LS55 (Perkin Elmer) with an ozone free pulsed xenon lamp whose wavelength is 200–800 nm. The schematic diagram of the fluorescence spectroscopy was shown in Fig. 2. 260 nm was selected as the excitation light. The emission light ranges from 200 to 600 nm, each 0.5 nm was scanned and recorded by a photomultiplier. The width of the slit was set as 5 nm, while the scanning speed was set as 500 nm/min. The drinking water was contained in a 1.00 cm quartz cell.

3.2 Reagents and Solutions

Zhongshengmycin, paclobutrazol, boscalid and pyridaben were used as the reagents which were dissolved in the distilled water in the present work. The initial concentrations of zhongshengmycin, paclobutrazol, boscalid and pyridaben were 0.144 mg/mL, 0.135 mg/mL, 0.07 mg/mL, 0.016 mg/mL, respectively. The working solutions were prepared by adding different volumes of the different reagents. Totally, 151 samples with different concentrations of different reagents were prepared to collect the fluorescence spectra. The working concentration of zhongshengmycin ranged from 0 mg/mL to 0.0305075 mg/mL, while that corresponding to paclobutrazol, boscalid and pyridaben ranged from 0 mg/mL to 0.0284407 mg/mL, 0.014747 mg/mL and 0.0033391 mg/mL.

3.3 Calibration and Validation Sets

Among the 151 spectra, 51 spectra were used as the training set and the remaining 100 spectra were used as the validation set. The PLS model indicated that the relationship between the fluorescence intensity and concentration was built using the calibration set. To avoid the influence of the abnormal samples to the model, cross-validation was performed first. Since four reagents were used, four PLS models were required to predict the concentration of each reagent independently.

TABLE 1

The Initial Value of the Parameters for BP Neural Network Model

Parameters	Value
The number of nodes in the hidden layer	8
Learning rate (Mu)	0.001
Mean square error (Goal)	10^{-10}
Number of iteration	5000

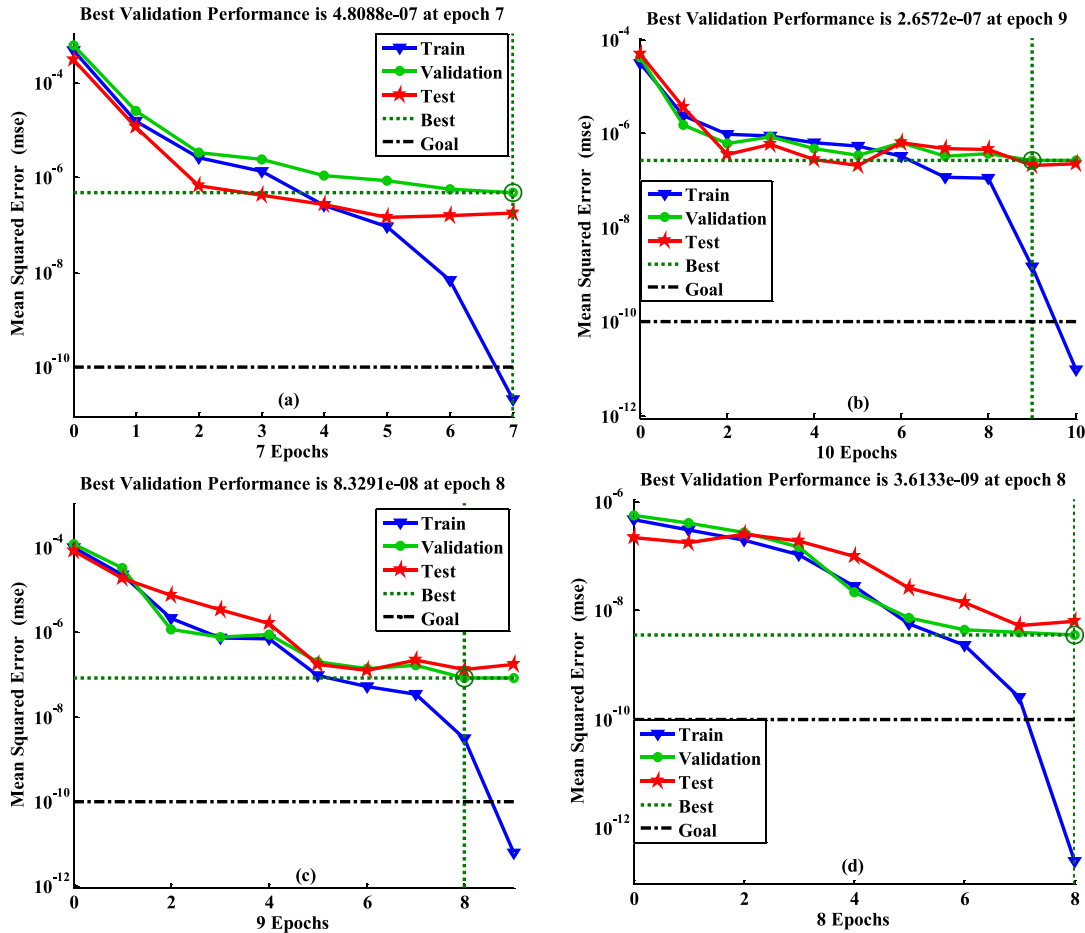
Fig. 3. The performance of the model for the training dataset. (a) for zhongshengmycin; (b) for paclobutrazol; (c) for boscalid; (d) for pyridaben.

3.4 The Parameters of the Model

Since four kinds of pesticides were selected and no chemical reaction would occur which means the fluorescence intensity of each pesticide will not be affected by other pesticides. As a result, the final fluorescence intensity can be regarded as the sum of the fluorescence intensity of each pesticide. Considering the Beer-Lambert law, when the concentration of the pesticide is low, the fluorescence intensity can be expressed as [30]

$$I = \phi I_0 \left(1 - \exp\left(-\varepsilon c d\right)\right) \tag{9}$$

where I_0 is the intensity incident on the sample, ϕ is the fluorescence efficiency and ε is the absorption coefficient, c is the concentration of the sample, d is the optical path length. Since
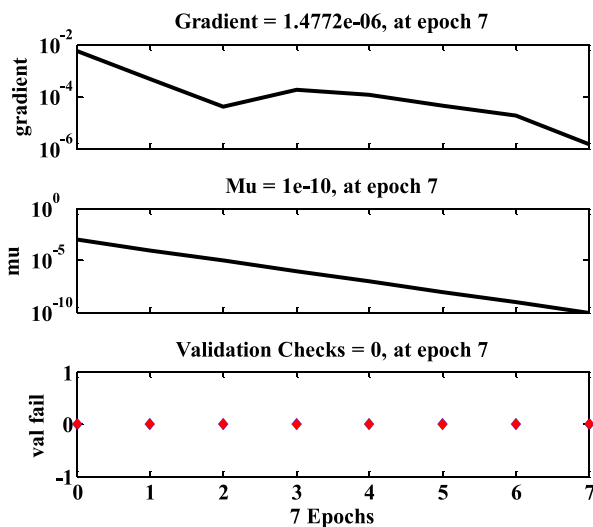
Fig. 4. The training stage of the model.
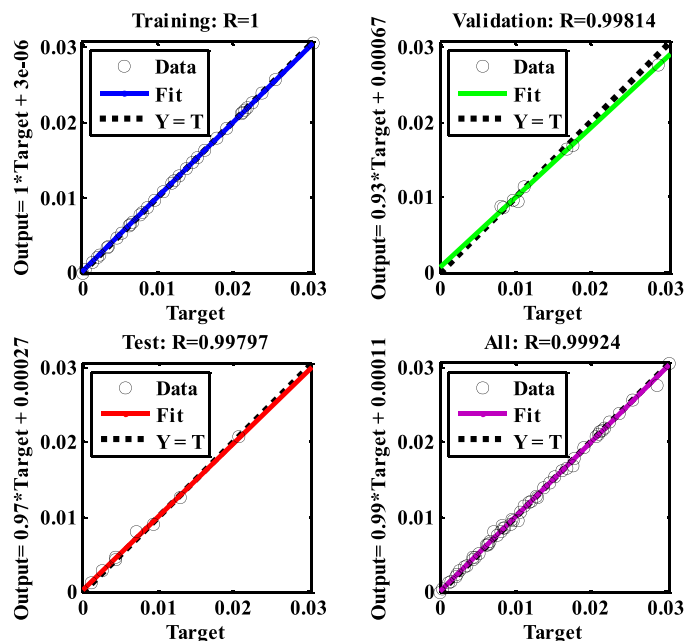
Fig. 5. The relationship between the actual concentration and the output of the model for the inner-train, inner-validation and inner-test.

the light source for each measurement is the same, the value of I_0 can be treated as the same for each pesticide. Given that a 1.00cm quartz cell was used to contain the samples and the refractive index of each pesticide sample can be treated as 1.33 (the refractive index of water) because of the low concentration, the optical path lengths for each pesticide sample is the same. Equation (9) can be revised as

$$I = \phi' \left(1 - \exp\left(-\varepsilon' c\right)\right) \tag{10}$$

Thus, there are only two parameters related to the pesticide: ϕ' and ε'. Considering that there are four kinds of pesticides, the fluorescence efficiency and absorption coefficient for each pesticide

Fig. 6. The validation of the model.

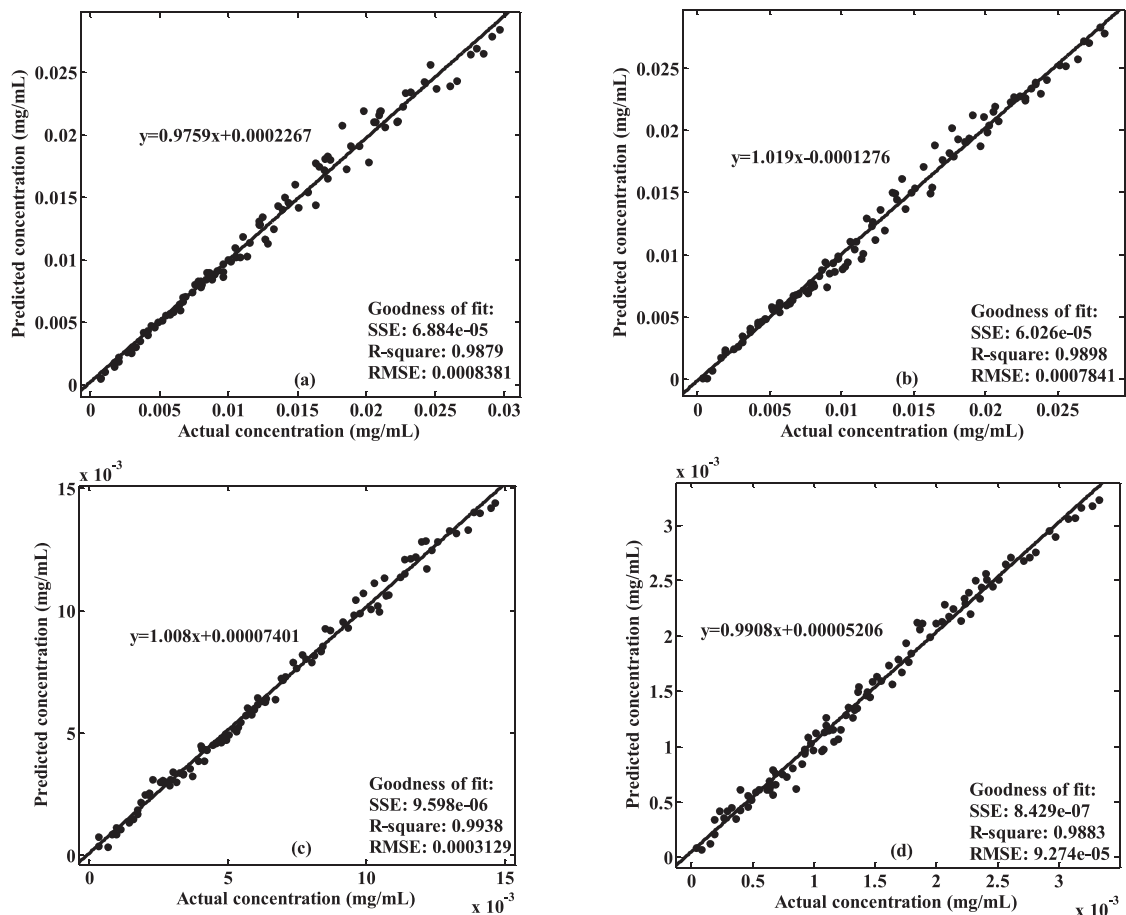may be different. Thus, the number of the node in the hidden layer is selected the value around 8. Matlab R 2013b was used to build the model. The value of the relative parameters was labeled in Table 1. The learning rate (Mu) was set as 0.001. The minimal error of the training was set as e^{-10}. The number of the iteration was set as 5000.

4. Results

The samples in the training dataset were used to build the model. To evaluate the model, neural network algorithm will divide the training dataset into three parts: train, validation and test. To be different from the training and validation dataset for the model, the train, validation and test divided from the training dataset was called as inner-train, inner-validation and inner-test. Fig. 3(a)–(d) shown the performance of the model for zhongshengmycin, paclobutrazol, boscalid and pyridaben. The best inner-validation performances (mean squared error) were 4.8088×10^{-7}, 2.6572×10^{-7}, 8.3291×10^{-8} and 3.6133×10^{-9} for zhongshengmycin, paclobutrazol, boscalid and pyridaben. In Fig. 3(a), even though the mean squared error of the validation is larger than the value 10^{-10} which is the goal, the value 4.8088×10^{-7} mg/mL is smaller than the 0.0003564 mg/mL which is lowest concentration of zhongshengmycin (except 0 which means the pure water). The mean squared error for the train and test were both lower than 4.8088×10^{-7} mg/mL which means that the model is suitable.

Fig. 4 shown the training stage of the model. The value of Mu decreased with the increase of epoch which means that the performance of the model got better with the increase of epoch. The value of validation checks means the times that the error of the validation is larger than that of the training. This parameter is used to prevent the model overfitting. If the value of the validation checks is larger than 6, it means that the error of the test would not be improved by increasing the epochs.

Fig. 5 shown the results obtained by the model for the inner-train, inner-validation and inner-test. The circle sign represented the output of the model. The dotted line represented the theoretical relationship between the actual concentration (Target) and the output of the model. The solid line, which was achieved by fitting, represented the actual relationship between the actual concentration (Target) and the output of the model. R-square coefficients were used to evaluate the performance of the fitting. The R-square coefficients for them were 1, 0.99814 and 0.99797, respectively. The average R-square coefficient for all the samples in the training dataset was 0.99924. All these results indicated a good linear relationship between the output predicted by the model and the actual concentration according to Wang's research [31], if the R-square of the model is larger than 0.9756, then the model indicates a good linear relationship between the predicted and actual values. The dotted line and solid line were overlapped for training, validation and test which means that the performance of the model was in accordance with the theoretical value.

Considering the model has demonstrated effective with the training dataset, to demonstrate the stability of the model, the remaining 100 samples in the validation dataset were used to test the model. Fig. 6(a)–(d) were the results for zhongshengmycin, paclobutrazol, boscalid and pyridaben samples in the validation dataset. The R-square coefficients for zhongshengmycin, paclobutrazol, boscalid and pyridaben were 0.9879, 0.9898, 0.9938 and 0.9883 which were all larger than 0.9756. The slopes of the fitting lines for each pesticide were 0.9759, 1.019, 1.008 and 0.9908 which were close to the theoretical value 1. The R-square coefficients and the slopes of the fitting lines demonstrated that BP neural network model performed well in predicting the concentration of the pesticides.

5. Conclusion

In this work, BP neural network algorithm was introduced to realize the simultaneous detection of multiple pesticides with fluorescence spectroscopy to overcome the disadvantages of exponential prediction model and excitation-emission fluorescence matrices methods. Four traditional three layers BP neural network models were tested with four kinds of pesticides and all the results indicated the promising application of the fluorescence spectroscopy in the field of pesticide residue detection. In the future, we hope to improve the BP neural network model with four nodes in the output layer which means that only one model is required to predict the concentration of four pesticides.

References

[1] F. P. Carvalho, "Agriculture, pesticides, food security and food safety," *Environ. Sci. Policy*, vol. 9, no. 7–8, pp. 685–692, 2006.
[2] A. G. Hornsby, T. M. Buttler, and R. B. Brown, "Managing pesticides for crop production and water quality protection: Practical grower guides," *Agriculture, Ecosyst. Environ.*, vol. 46, no. 1-4, pp. 187–196, 1993.
[3] M. R. Bonner and M. C. R. Alavanja, "Pesticides, human health, and food security," *Food and Energy Secur.*, vol. 6, no. 3, pp. 89–93, 2017.
[4] O. A. Ibigbami, A. F. Aiyesanmi, A. J. Adesina, and O. K. Popoola, "Occurrence and levels of chlorinated pesticides residues in cow milk: A human health risk assessment," *J. Agricultural Chemistry Environ.*, vol. 8, no. 1, 2019, Art. no. 58.
[5] A. Sabarwal, K. Kumar, and R. P. Singh, "Hazardous effects of chemical pesticides on human health-cancer and other associated disorders," *Environ. Toxicology Pharmacology*, vol. 63, pp. 103–114, 2018.
[6] J. Sun *et al.*, "Dual functional PDMS sponge SERS substrate for the on-site detection of pesticides both on fruit surfaces and in juice," *Analyst*, vol. 143, no. 11, pp. 2689–2695, 2018.

[7] T. Mu *et al.*, "Detection of pesticide residues using Nano-SERS chip and a smartphone-based Raman sensor," *IEEE J. Sel. Topics Quantum Electron.*, vol. 25, no. 2, Mar./Apr. 2019, Art. no. 5200206.

[8] B. Jamshidi, E. Mohajerani, and J. Jamshidi, "Developing a Vis/NIR spectroscopic system for fast and non-destructive pesticide residue monitoring in agricultural product," *Measurement*, vol. 89, pp. 1–6, 2016.

[9] J. W. Li, Y. L. Wang, S. Yan, X. J. Li, and S. Y. Pan, "Molecularly imprinted calixarene fiber for solid-phase microextraction of four organophosphorous pesticides in fruits," *Food Chemistry*, vol. 192, pp. 260–267, 2016.

[10] B. Chen *et al.*, "Determination of 27 pesticides in wine by dispersive liquid–liquid microextraction and gas chromatography–mass spectrometry," *Microchem. J.*, vol. 126, pp. 415–422, 2016.

[11] A. Goon *et al.*, "A simultaneous screening and quantitative method for the multiresidue analysis of pesticides in spices using ultra-high performance liquid chromatography-high resolution (Orbitrap) mass spectrometry," *J. Chromatography A*, vol. 1532, pp. 105–111, 2018.

[12] G. H. Torosyan, E. K. Armudjyan, and V. A. Davtyan, "Determination of malathion in water using liquid chromatography and mass spectroscopy," *Analytical Chemistry An Indian J.*, vol. 18, no. 1, 2018, Art. no. e101.

[13] M. T. Sánchez, K. Flores-Rojas, J. E. Guerrero, A. G.-Varo, and D. P.-Marin, "Measurement of pesticide residues in peppers by near-infrared reflectance spectroscopy," *Pest Manage. Sci. Formerly Pesticide Sci.*, vol. 66, no. 6, pp. 580–586, 2010.

[14] G. Xiao *et al.*, "Detection of pesticide (chlorpyrifos) residues on fruit peels through spectra of volatiles by FTIR," *Food Analytical Methods*, vol. 8, no. 5, pp. 1341–1346, 2015.

[15] H. Y. Bian, Y. L. Zhang, W. R. Gao, and J. Gao, "Fourier based partial least squares algorithm: New insight into influence of spectral shift in "frequency domain," *Opt. Express*, vol. 27, pp. 2926–2936, 2019.

[16] H. Bian and J. Gao, "Error analysis of the spectral shift for partial least squares models in Raman spectroscopy," *Opt. Express*, vol. 26, pp. 8016–8027, 2018.

[17] B. Lin *et al.*, "Modification-free carbon dots as turn-on fluorescence probe for detection of organophosphorus pesticides," *Food Chemistry*, vol. 245, pp. 1176–1182, 2018.

[18] M. V. Navarro, M. A. Cabezón, and P. C. Damiani, "Simultaneous determination of pesticides in fruits by using second-order fluorescence data resolved by unfolded partial least-squares coupled to residual bilinearization," *J. Chemistry*, vol. 2018, pp. 321746–3217481, 2018.

[19] M. Wu *et al.*, "Application of deep brief network in transmission spectroscopy detection of pesticide residues in lettuce leaves," *J. Food Process Eng.*, vol. 42, no. 3, 2019, Art. no. e13005.

[20] T. B. Pham *et al.*, "Detection of Permethrin pesticide using silver nano-dendrites SERS on optical fiber fabricated by laser-assisted photochemical method," *Sci. Reports*, vol. 9, no. 1, pp. 1–10, 2019.

[21] Y. Wang *et al.*, "Grating-like SERS substrate with tunable gaps based on nanorough Ag nanoislands/moth wing scale arrays for quantitative detection of cypermethrin," *Opt. Express*, vol. 26, no. 17, pp. 22168–22181, 2018.

[22] J. Chen *et al.*, "Flexible and adhesive surface enhance Raman scattering active tape for rapid detection of pesticide residues in fruits and vegetables," *Analytical Chemistry*, vol. 88, no. 4, pp. 2149–2155, 2016.

[23] X. Gong *et al.*, "Screening pesticide residues on fruit peels using portable Raman spectrometer combined with adhesive tape sampling," *Food Chemistry*, vol. 295, pp. 254–258, 2019.

[24] Y. Y. Yuan, S. T. Wang, Q. Cheng, D. M. Kong, and X. G. Che, "Simultaneous determination of carbendazim and chlorothalonil pesticide residues in peanut oil using excitation-emission matrix fluorescence coupled with three-way calibration method," *Spectrochimica Acta Part A: Mol. Biomolecular Spectrosc.*, vol. 220, 2019, Art. no. 117088.

[25] S. Ma *et al.*, "Fluorescence detection of boscalid pesticide residues in grape juice," *Optik*, vol. 180, pp. 236–239, 2019.

[26] A. P. Pagani and G. A. Ibañez, "Pesticide residues in fruits and vegetables: High-order calibration based on spectrofluorimetric/pH data," *Microchem. J.*, vol. 149, 2019, Art. no. 104042.

[27] L. Liu, "Recognition and analysis of motor imagery EEG signal based on improved BP neural network," *IEEE Access*, vol. 7, pp. 47794–47803, 2019.

[28] H. Tang, M. Lei, Q. Gong, and J. Wang, "A BP neural network recommendation algorithm based on cloud model," *IEEE Access*, vol. 7, pp. 35898–35907, 2019.

[29] J. Wang, J. Fang, and Y. Zhao, "Visual prediction of gas diffusion concentration based on regression analysis and BP neural network," *J. Eng.*, vol. 2019, no. 13, pp. 19–23, 2019.

[30] K. Kumar, M. Tarai, and A. K. Mishra, "Unconventional steady-state fluorescence spectroscopy as an analytical technique for analyses of complex-multifluorophoric mixtures," *TrAC Trends Analytical Chemistry*, vol. 97, pp. 216–243, 2017.

[31] L. Wu *et al.*, "Application of magnetic solvent bar liquid-phase microextraction for determination of organophosphorus pesticides in fruit juice samples by gas chromatography mass spectrometry," *Food Chemistry*, vol. 176, pp. 197–204, 2015.