

# **Research and Application of Quality Detection Method for Corn Oil**

## **Based on Neural Network Algorithm**

Wei-Feng Gao, Ming-fen Wu

Wuyi University, No. 22 Dong Cheng Cun, Wuyi, Jiangmen, P.R. China, 529020

**Keywords:** Neural network algorithm; Corn oil; Infrared spectrum; Quality detection

**Abstract:** The nutritious value of corn oil is gradually recognized by the public, while with the expansion of its production, how to detect its quality has become a tough problem. The research performs data analysis on the infrared spectrum of corn oil to inspect the quality of corn oil by adopting neural network algorithm. Experiments have proved that the method can test the quality of corn oil effectively and compared with other ways, it is qualified with the advantages of high efficiency and accuracy, so it is worthy to be promoted and used.

### **Introduction**

Corn oil, also called maize oil or maize germ oil, is the oil extracted from maize germ. Fat content of maize germ is within the range from 17% to 45% approximately accounting for over 80% of the total content of corn fat. The feature of fatty acid in corn oil is that the content of unsaturated fatty acid reaches the range from 80% to 85% [1]. Corn oil itself doesn't contain cholesterol and it plays a role in dissolving the accumulation of cholesterol in blood, so it can reduce the hardening influence on blood vessel and has positive preventive and therapeutic effect on gerontal disease such as arteriosclerosis and diabetes. Because of the functions of its natural compound vitamin E, it has obvious curative effect and preventive action on all the diseases of heart disease, thrombophlebitis, reproductive function class barriers [2], infantile spinal muscular atrophy and trophism cerebromalacia sickness. Some old people eat a spoonful of corn oil with an empty stomach every day that is taken as an adaptogen [3]. Corn oil is refined after dephosphorization, deacidification, degumming, decolorization, deodorization and dewaxing on the basis of refined corn oil. Corn oil is clear and has faint scent and high lampblack point, so it is very suitable for rapid cooking and frying. For the nutritious value of corn oil is gradually accepted by the public, quality grading method for corn oil is set up step by step. John Dudley believes that most of the effect of genes taking charge of the oil content of corn is additive effect and minority of them is explicit or regulating effect. Therefore, in the breeding of corn variety with high oil, if every grain of corn seeds can be separately done the non-destroyed measurement on the oil content so as to filter those with high oil content, the breeding productivity of high oil can be greatly enhanced. On account of the destructiveness, traditional chemical method cannot be used in nondestructive testing. The occurring of Nuclear Magnetic Resonance Spectrometer has greatly improved the fostering efficiency of high oil corn, however, due to its slow analysis speed, it cannot meet the breeding work demand of seeds in thousands[4], in addition, the purchasing and maintenance expense for it is very high. In recently years, near-infrared technique is gradually becoming mature, which can do non-destructive determination on complex samples without any processing and even do single grain analysis. With rapid speed, consuming no reagent and producing no pollution, it can be taken as a kind of "green analysis" technology [5]. There is no doubt that the development of this kind of instrument will significantly improve the working speed and efficiency of breeding manager, save

manpower and material resources and play a positive promoting role in the breeding of our country's high oil corn.

Prediction model in near-infrared analysis mainly solve the problems in two aspects: the first one is that the relationship between near infrared spectral absorption of the sample and its composition doesn't always conform to "Lambert-Beer" law [6]; the second one is that the non-linear relationship is presented between measured material composition itself and component concentration, when the qualitative change range of a sample is wider, the nonlinear response is more obvious. Spectroscopic data and chemical values of the sample shall be related including MLR, PCR and PLS etc. In the last few years, the application of nonlinear correction technique in spectral analysis is increasing day by day, what's more, BP nerve network and other nonlinear methods hold an important position in multivariate nonlinear correction technology and have achieved good effect in many fields. The research will introduce artificial neural network algorithm to do quantitative analysis and modeling construction on spectroscopic data of corn so as to test the quality of corn oil.

## Materials and methods

**Materials:** Corn oil: Xiwang corn oil bought in the supermarket. Samples are divided into two sets: the calibration set includes 125 samples and the validation set has 20 samples. The characteristic data of validation set is uniformly distributed within the range of property data for calibration set. Test scope of acid value for calibration set and validation set is from 2.228 to 2.964mg KOH/g. Infrared spectrophotometer: Nicolet Nexus 5700 Fourier transforming infrared spectroscopy

**Detection methods:** The range of spectrum for near infrared spectroscopy in short-wave area is from 780 to 1100nm and it does analysis by adopting the absorbing information of secondary frequency doubling and triple frequency doubling of hydric group in organic compounds or water (such as C-H, O-H, N-H etc.) in the spectral region. The absorption of C-H group for the organic ingredients of corn such as fat, protein and starch and so on in short-wave area is mainly triple frequency doubling and its absorption peak range is from 830 to 1100nm; The absorption of O-H group in short-wave area is mainly secondary frequency doubling and its absorption peak range is from 1000 to 1100nm [7]; The absorption of N-H group in short-wave area is mainly secondary frequency doubling and its absorption peak range is from 980 to 1100nm. Therefore, the absorption information of the corn seed in near infrared shortwave area can be utilized to determine the content of oil, protein, starch etc. in corn seed.

**Computing method:** Feed forward neural network based on error back propagation algorithm is BP neural network. BP neural network is a kind of multilayer mapping neural network with back passing and fixing errors and it is a representative neural network model in a variety of neural network models. BP network has distinctive features: the distributed information storage method; massively parallel processing; self-learning and self-adaptability; strong robustness and fault tolerance. Back propagation algorithm adopted by BP network a learning algorithm that is the most mature and the most widely used one in feed forward neural network at present.

**Determination of the structure of BP neural network:** BP neural network is a typical forward network composed by input level, output level and a number of hidden layers and it is generally believed that network is strictly hierarchical, that is, it is possible for two nodes to be interconnected when and only when they are in adjacent layers. Figure 1 is a structure chart for three-layer BP neural network. Determination of the network structure includes the selection of the number of hidden layers and the determination of the number of nodes in each layer. The input

layer plays a role of buffer memory that adds the data source on the network. The number of nodes depends on the dimensionality of data source, namely, the dimensionality of input feature vector. When neural network is used for pattern classification, the number of nodes in output layer is generally the number of required categories, or the encoding of the output node shall be used to represent the model category, and when neural network is used for functional approximation, the number of output nodes is the output number for required function in general.

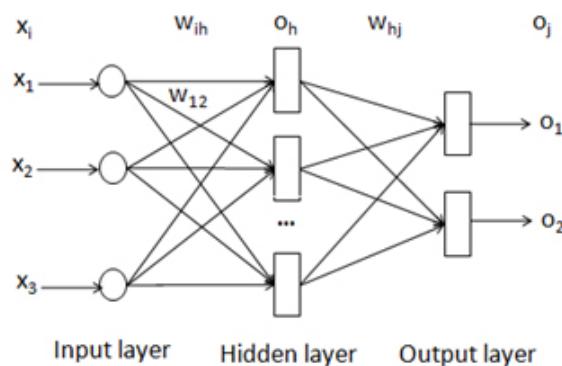


Figure 1 Structure chart for three-layer neural network

In practice, it has been the difficulty and focus to confirm feed forward network structure that how to determine the number of hidden layers and nodes in the layer. It is also the main disadvantage of BP neural network in structure that the number of hidden layers and hidden layer units is difficult to be determined.

**Determination of the number of hidden layers:** For most neural networks, what needs to be done firstly is making sure the number of selection hidden layers. Mallat had proved that when each node has a different threshold, a continuous function in any closed area can be approximated by the network containing a hidden layer. Thus a three-layer BP network can complete any mapping from  $n$ -dimension to  $m$ -dimension. Hidden layer has the abstract function, which means that it can extract features from input. After several experiments, for the neural network model for oil acid value of corn oil, the training frequency and the training time for the network with two hidden layers are much more than just has one hidden layer, but the effect has not been significantly improved. So as for the measurement of acid value of oil, choosing a hidden layer can solve the problem well.

**Determination of node number in hidden layer:** In BP neural networks, the selection of the number of nodes on each floor has great effect on network performance. Therefore, the number of nodes in the layer needs to be selected appropriately. If the number of nodes in hidden layer is too small, the local minimum is more, and the network training effect cannot be reached. However, if the number of hidden nodes is too much, network fitting function is too complex and there is easily the over fitting, which leads to the poor generalization ability for the network, that means the training effect is very well, while predictive effect may be badly poor. Hence many scholars have done a lot of researches and explorations on how to solve the optimal number of hidden nodes and proposed various feasibility plans. Two estimation methods for nodes number of single hidden layer can refer to the reference.

(1) Formula for the calculation of node-number

$$h=0.43mn+0.12n^2+2.54m+0.77n+0.35+0.51$$

In the formula:  $h$  for the number of nodes in hidden layer;  $m$  for the number of input nodes;  $n$  for the number of output nodes.

## (2) Dynamically adjusting the number of hidden layer nodes

Typically there are two methods: one is putting enough hidden layer nodes at the beginning, and then the inoperative ones shall be gradually removed after learning until they cannot shrink. Another is putting less hidden layer nodes, after certain times of learning, if it is still unsuccessful, the number of nodes shall be increased until a reasonable number is reached.

For neural network model of acid value of corn oil, method (1) should be used at first to calculate out the number of hidden layer, then method (2) shall be adopted to do adjustment, and finally the getting number of hidden layer nodes is 5.

## (3) Learning rules

Delta learning rules shall be adopted and the adjustment of weight value will be done according to the direction of the error function gradient descent. Please refer to formula 1:

$$\begin{aligned}\Delta W_{jk} &= -\eta \frac{\alpha E}{\alpha W_{jk}}, \quad j = 0, 1, \dots, m \\ \Delta V_{ij} &= -\eta \frac{\alpha E}{\alpha V_{jk}}, \quad i = 0, 1, \dots, n; k = 1, 2, \dots, m\end{aligned}\quad (1)$$

In the formula: minus shows gradient descent; constant represents the scale factor reflecting the learning rate in training and two are used to derive weight value, and the specific formulas can refer to formulas( 2 ):

$$\begin{aligned}\Delta W_{jk} &= \eta(d_k - O_k)O_k(1 - O_k)y_j \\ \Delta V_{ij} &= \eta[\sum_{k=1}^1(d_k - O_k)O_k(1 - O_k)W_{jk}]y_j(1 - y_j)x_i\end{aligned}\quad (2)$$

Threshold adjustment value is the correction error of corresponding layer multiplied by the scale factor, please refer to formula( 3 ):

$$\begin{aligned}\theta_k &= \eta(d_k - O_k)O_k(1 - O_k) \\ \gamma_j &= \eta[\sum_{k=1}^1(d_k - O_k)O_k(1 - O_k)W_{jk}]y_j(1 - y_j)\end{aligned}\quad (3)$$

## Results and discussion

**Analysis of near-infrared spectra of samples:** Infrared spectroscopy will be used to observe the spectrum chart with the range of wavelength from 950 to 1650nm and the near infrared spectroscopy of 125 corn oil samples. From the near infrared spectroscopy of corn oil, the spectral data experiences the first-order differential so as to eliminate the effect of light scattering and differential width is 15 points. It can be seen that this batch of corn oil has greater absorption in about 1200nm and 1500nm and has smaller absorption in about 1300nm. Spectrum analysis system built in the software can give the quantitative expression of the main component of this batch of corn oil.

**Determined results and discussion of BP neural network and comparison model parameter:** BP neural network, multiple linear regression, principal component regression and partial least square method shall be adopted to set up “oil-acid value” model to compare their modeling effect. As for “oil-acid value” model, the main factor scores gotten from partial least squares and principal components analysis should be taken by comparison and regarded as the correction effect of input variable for BP neural network; And due to that the active ingredients of this batch of corn oil gained from partial least squares and principal component analysis are similar, it is believed that the main factor scores obtained from these two methods have no significant difference, so the “oil-acid

value” model can just use principal component analysis method for data dimension reduction. Based on the prediction effect of BP neural networks on the acid value of oil as well as the predictions of models with certain cases of interference, this section tries to get the application value of BP neural network in the near infrared analysis of oil. Determined by the above, BP network structure includes the three-tier, which contains only one hidden layer; the number of hidden layer nodes is 6; Delta learning rule is employed and activation functions of the hidden layer and the output layer are respectively the S Function and the linear function, which has established a BP neural network. Number of training samples is 70 and monitoring number of samples is 55, and gradient method is used.

Finally, different forecasting models and the above parameters will be employed to establish “corn oil- acid value” model, in addition, nature of the validation set needs to be predicted. Prediction results for different models are shown in table 1.

Table 1 Predictions of “corn oil -acid value” model in different models

standard deviation	Project			
The standard deviation of prediction for calibration set ( % )	BP Neural Network	PLS	MLR	PCR
	0.1294	0.1643	0.1364	0.1920

From the table 1, it can be seen that the accuracy of the prediction results of BP neural network and multivariate linear regression method is prior to that of principal component regression and partial least square. The nature change range of corn oil is wide and the change relationship between its spectrum and nature is non-linear, thus nonlinear modeling method can more truly reflect the relationship of sample spectrum and nature.

## Conclusion

For that there is a non-linear relationship between the spectral response and the chemical measurement value of soybean oil level four, better accuracy will be gotten by adopting BP neural network model to do forecasting. Moreover, according to “mapping network existence” theorem, three-layer BP neural network can fit any continuous function with arbitrary precision, which has strong nonlinear modeling capabilities; in addition, establishing model by using BP neural network doesn’t need to know in advance the specific form of the model, so it is suitable to be applied to solve complex mapping problems.

As for systems with wider sample property change ranges, better prediction accuracy can be obtained from BP neural network and multivariate linear regression method rather than from principal component regression and partial least square. Unified modeling can be set up among systems with wider sample property change ranges by employing BP neural network method, thus the model application scope can be extended. For some sample systems, multiple linear regression method is more accurate than BP neural network method, but that is oversensitive to minor changes of unknown samples in spectral measurement and the modeling anti-interference performance is poor. When sample property change range is wide, BP neural network method shall be more accurate and have better performance in anti-interference, which is a good prediction model.

## References

- [1] Xihong Zhao, Tao Song, Jian Li, Nanjing Zhong and Weicheng Hu, 2015. Validation of Method in Microbial Limit Tests for Two Types of Health Foods. Advance Journal of Food Science and Technology, 7(11): 872-874.
- [2] Moghaddam, T. M., Razavi, S. M., & Taghizadeh, M. (2013). Applications of hyperspectral imaging in grains and nuts quality and safety assessment: a review. Journal of Food Measurement and Characterization, 7(3), 129-140.
- [3] Ellis, D. I., Brewster, V. L., Dunn, W. B., Allwood, J. W., Golovanov, A. P., & Goodacre, R. (2012). Fingerprinting food: current technologies for the detection of food adulteration and contamination. Chemical Society Reviews, 41(17), 5706-5727.
- [4] Peng, D., Bi, Y., Ren, X., Yang, G., Sun, S., & Wang, X. (2015). Detection and quantification of adulteration of sesame oils with vegetable oils using gas chromatography and multivariate data analysis. Food Chemistry, 188, 415-421.
- [5] Ramzi, M., Kashaninejad, M., Salehi, F., Mahoonak, A. R. S., & Razavi, S. M. A. (2015). Modeling of rheological behavior of honey using genetic algorithm-artificial neural network and adaptive neuro-fuzzy inference system. Food Bioscience, 9, 60-67.
- [6] Caceres, J. O., Moncayo, S., Rosales, J. D., de Villena, F. J. M., Alvira, F. C., & Bilmes, G. M. (2013). Application of laser-induced breakdown spectroscopy (LIBS) and neural networks to olive oils analysis. Applied spectroscopy, 67(9), 1064-1072.
- [7] Taghadomi-Saberi, S., Omid, M., Emam-Djomeh, Z., & Ahmadi, H. (2014). Evaluating the potential of artificial neural network and neuro - fuzzy techniques for estimating antioxidant activity and anthocyanin content of sweet cherry during ripening by using image processing. Journal of the Science of Food and Agriculture, 94(1), 95-101.