NONDESTRUCTIVE SUGAR CONTENT MEASUREMENT IN APPLE BY NIR SPECTRUM ANALYSIS USING NEURAL NETWORK

S. H. Lee	S. H. Noh	W. G. Kim
graduate student	professor	graduate student

Department of Agricultural Engineering, Seoul National University
Suweon 441-744, Korea

ABSTRACT

This study was conducted to develop neural networks of predicting the sugar content of fruits based on the optical densities obtained from a spectrophotometer. Pear, apple and peach were used in investigating the feasibility of the developed neural networks as a nondestructive measurement. A spectrophotometer was used to measure the optical densities of test fruits. The neural networks suggested in this study consisted of multi-layers having one hidden layer and one output layer. The correlation coefficients between the predicted and the measured sugar content for most fruits were high. The neural networks using 2nd derivatives of optical density spectrum produced a better results in predicting the sugar content of fruits. This study contributed to develop a method for nondestructively predicting the sugar content of fruits.

Key words : neural network, NIR, sugar content

INTRODUCTION

At present, the measurement on the internal quality of fruits becomes an emerging issue as the amount of fruit consumption and the demand for better quality fruits have increased. There were few researches on the nondestructive measurement of sugar content even if the sugar content of a fruit is one of the most important factors to determine its internal quality. In other countries, nondestructive measurements on some fruits, such as peach or watermelon, have been studied using visible or near infra red (NIR) spectroscopy which is relatively simple and accurate. The studies of the possibility for nondestructive measurements have been conducted in Korea. But there are many difficulties in implementing a feasible system of measuring the sugar content of fruits, based on

The previous studies have tried to predict the sugar content by using a statistical model along with the difference or ratio of optical density(O.D.) at some specific wavelengths as an input factor for prediction. The reflectance characteristic of sugar content of fruits would change with various factors that are difficult to physically be analyzed, such as the property and the size of epidermal tissues. Therefore, the change in the reflectance characteristic of sugar content will result in decreasing the accuracy of the statistical model for predicting the sugar content.

Many optical sensors have been recently developed through the rapid progress of electronic engineering. In particular, the array sensor acquires the spectrum of reflected light in a wide waveband at a real time. Therefore, it provides a method to relate the sugar content of fruits to the acquired wide range of wavelengths. This study was conducted to predict the sugar content of fruits using a neural network along with various spectrum data.

MATERIALS AND METHODS

Materials

Pear, apple and peach were used as a test fruit in investigating the feasibility of the nondestructive measurement based on the neural network in this study. The test fruits were purchased directly from their related orchards. The number of measurement was 166 for Fuji apple(22 samples), 40 for Miback peach(7 samples) and 80 for Singo pear(9 samples), respectively. In particular, Fuji was tested at two different categories; 72 non-stored(8 samples) and 94 stored apples(14 samples). Such testing conditions for apple were selected to investigate the effect of storage process on the sugar content of fruits.

Measurement system

A spectrophotometer (UV-3101PC, Shimadzu, Japan) in Agricultural Mechanization Center was used to measure the spectral reflectance of fruits. The measurement range of wavelength was 400-2400 nm at 2 nm interval. The range of measurement was from 400nm to 2400nm wavelength with 2nm interval. The measurement of spectral reflectance was conducted at a local area in the middle of test fruits. The sugar contents were measured by using a digital refractometer (Model DBX55, Atago, Japan). Table 1 shows the type and the number of measurement and their sugar contents used in the study.

A neural network for predicting sugar content in fruits

A neural network suggested in this study consisted of multi-layers having one hidden layer and one output layer. The learning process of the neural network was based on general back-propagation algorithm with optical densities obtained from the spectrophotometer as a learning material. The number of hidden unit in the neural network was determined by observing the learning process and selecting its proper number. And the number of output unit was one in the study. The learning process was continued until its mean square error(MSE) was lower than 0.0005. As a target value the sugar content of fruits was linearly transformed into 0 through 1. In the end, each neural network model for apple, peach and pear was constructed by using O.D.s and their second derivatives. The performance of the neural networks for predicting the sugar content was verified by analyzing the correlation coefficients and the mean square errors between the predicted and the measured sugar content.

Most of array sensors can measure the spectral reflectance in a wide range of wavelengths, up to 1100nm wavelength. The optical densities in the range of 400-1100 nm were used for developing a neural network in the study. In general, the optical densities contains a lot of variations which depend on measurement precision of spectrophotometer. Thus, the O.D.s should be preprocessed to remove the variations. In this study, the O.D.s obtained from the sensor were preprocessed using moving average as a smoothing technique. In applying moving average, the segment size was 10 nm in this study.

As another input factor into the neural network, the second derivatives of the O.D.s were obtained by the following equation (1).

$$D_x = O.D_{x-4nm} - O.D_x + O.D_{x+4nm} \text{ ------------------------ } (1)$$

All wavelengths in the range of 400-1100 nm would not be used as an input into the neural network because too many inputs reduce the stability of neural network. By the multiple regression analysis, some useful combinations of 3 ~ 10 wavelengths explaining more than 80 % of the measured sugar content were obtained and used as an input to the neural networks.

RESULTS AND DISCUSSION

Correlation analysis between sugar content and O.D. for Singo pear

Fig. 1 shows the correlation coefficients between the sugar content and the O.D. at each wavelength for pears. The maximum correlation coefficient was very low with 0.3. The correlation was also very low even at the wavelengths of 838, 888, 913, 978, and 1005 nm which were known as sugar absorption waveband.

Fig 2 shows the correlation coefficients between the sugar content and the second derivative of the O.D. at the segment size of 10 nm and the gap size of 4 nm. In this case, the maximum correlation coefficient was -0.7667. The results showed that the second derivatives of O.D.s produced more precise prediction of sugar content than the O.D..

The prediction of sugar content by the neural network using O.D. spectrum

The learning process of the neural network using the O.D.s was performed on 3 units of the hidden layer until the MSE of the learning process got down to 0.0005. Table 2 shows the input wavelengths and the correlation coefficients for the prediction of sugar contents by the neural network using O.D. spectrum. The prediction was performed with 36 apples, 20 peaches and 40 pears. The correlation coefficient between the predicted and measured sugar content was 0.8806 for apple and the MSE was 0.65. The results for apple showed that the prediction of the sugar content by the neural network using O.D.s would be feasible. However, The correlation coefficient was 0.7751 for peach and 0.8202 for pear. Thus the precision of the prediction for pear and peach show a little lower than that for apple.

The prediction of sugar content by the neural network using 2nd derivatives of O.D. spectrum

The 2nd derivatives of O.D. spectrum were obtained by using equation (1). The same learning process of a neural network was performed on the basis of the 2nd derivatives. Table 3 shows the input wavelengths and the correlation coefficients for the prediction of sugar content by the neural network using 2nd derivatives of O.D. spectrum. The proper number of wavelengths necessary for the neural network varied depending on the type of fruit. On the basis of 3 wavelengths, the correlation coefficient for non-stored apples was 0.9054 and the MSE was 0.56, which showed that the precision of prediction would be very high. In case of stored apples in store, the number of wavelengths was 10 and the correlation coefficient was 0.9016. In addition to the separated analysis for apples, the neural network was tested for the case without separating the type of apple storage. The correlation coefficient for the prediction of sugar content was 0.8733 for apples. The results showed that the neural network using 2nd derivatives of O.D. spectrum produce a better prediction than the one using O.D. spectrum. As shown in Table 3, the correlation coefficient for prediction was 0.9011 at 3 wavelengths for peaches and 0.8114 at 3 wavelengths for pears.

Table 4 shows the results for the prediction of sugar content based on multiple regression analysis. The wavelengths in the table were obtained from applying a stepwise technique for selecting independent variables in multiple regression analysis. In predicting the sugar content by multiple regression analysis, the number of wavelengths was 3 for all fruits used in the study on consideration of the stability of analysis. The correlation coefficients for most fruits were high except peaches, so that the multiple regression models were feasible for predicting the sugar content of fruits. However, the coefficients on multiple regression models were a little lower than those on the neural networks using 2nd derivatives of O.D. spectrum. Based on the results of the study, the neural networks suggested in the study are very feasible for predicting the sugar content of fruits.

CONCLUSIONS

The sugar content of fruits is an important factor to determine their internal quality. This study was conducted to develop neural networks of predicting the sugar content of fruits based on the optical densities obtained from a spectrophotometer. Pear, apple and peach were used as test fruits to investigate the feasibility of the neural networks developed in the study. The neural network using the second derivatives of optical density spectrum produced more precise prediction of sugar content than the optical density and the conventional multiple regression model. Based on the results of the study, the developed neural networks were very feasible for predicting the sugar content of fruits. This study can contribute to develop a method for nondestructively predicting the sugar content of fruits.

REFERENCES

1. Kim, Y. H., C. S. Kim, S. B. Kim, M. S. Kim, and K. C. Shin. 1992 Correlation between Non-destructive quality evaluation parameters and spectral reflectance of apple. Journal of the Korean Society for Agricultural Machinery. No 17(4) :pp 370-381.
2. Kim. H. U. 1994. Non-destructive determination of sugar content in fruits by near infrared spectroscopy. M.S. thesis
3. Noh. S. H. 1995. Development of technology for nondestructive detection of internal qualities of apple.
4. Lee. K. H. 1993. Measurement of freshness of apple by NIR spectroscopy. M. S. thesis.
5. ORIEL catalog. 1994. ORIEL corporation

Table 1. Fruit samples for calibration and prediction of sugar content

Fruit	Apple				Peach		Pear	
	Non-stored		Stored					
Data set	Cal.	Pre.	Cal.	Pre.	Cal.	Pre.	Cal.	Pre.
No. of measurement	36	36	47	47	20	20	40	40
Sugar content range	11.7 ~ 17.5	11.8 ~ 17.6	8.1 ~ 16.6	8.3 ~ 17.1	11.2 ~ 16.1	9.8 ~ 15.8	9.6 ~ 15.6	9.7 ~ 16.7

Table 2. Sugar content prediction by neural network for apple, peach and pear with O.D spectrum

Item	wavelength(nm)	correlation coefficient r by neural network (mean squared error)
Nonstored apple	444, 604, 784, 794, 824	0.8806(0.65)
Peach	496, 656, 936, 1056, 1096	0.7751(1.31)
Pear	414, 464, 484, 684, 874	0.8202(0.80)

Table 3. Result of Sugar content prediction by neural network for fruit with 2nd derivatives spectrum

Item	Wavelength(nm)	Correlation coefficient r (mean squared error)
Nonstored apple	580, 644, 930	0.9054(0.56)
Stored apple	472, 524, 536, 538, 648 670, 694, 800, 828, 854	0.9016(0.72)
Total apple	406, 582, 650, 698, 720 780, 856, 930, 1000, 1024	0.8773(0.73)
Peach	596, 836, 866, 986, 1046	0.9011(0.71)
Pear	480, 594, 934	0.8114(0.66)

Table 4. Result of Sugar content prediction by multiple regression model for fruit with O.D and 2nd derivatives spectrum

Item	Selected wavelength(nm) with O.D spectrum	Correlation coefficient r (mean squared error)	Selected wavelength(nm) with 2nd derivatives spectrum	Correlation coefficient r (mean squared error)
Non-stored apple	444, 604, 784, 794, 824	0.8430(0.80)	580, 644, 930	0.8371(0.75)
Stored apple	434, 664, 684, 954, 1024	0.8295(0.86)	648, 690, 800	0.8017(0.90)
Total apple	684, 834, 1044, 1074, 1094	0.8304(0.83)	650, 656, 954	0.8544(0.84)
Peach	496, 656, 936, 1056, 1096	0.7364(1.25)	480, 538, 582	0.6225(2.29)
Pear	414, 464, 484, 684, 874	0.8526(0.97)	480, 594, 934	0.8062(0.67)

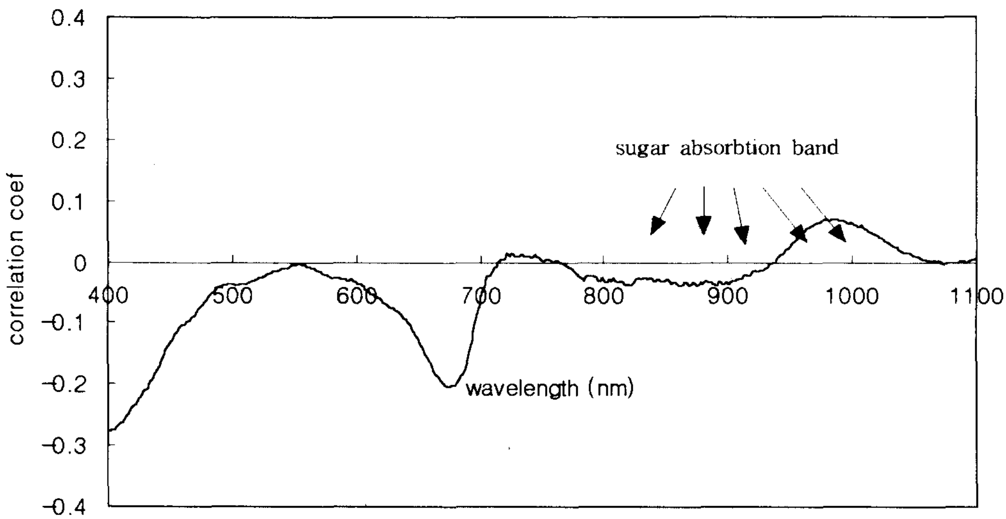


Fig 1. Correlation efficient between sugar content and O.D. for 'Singo' pear

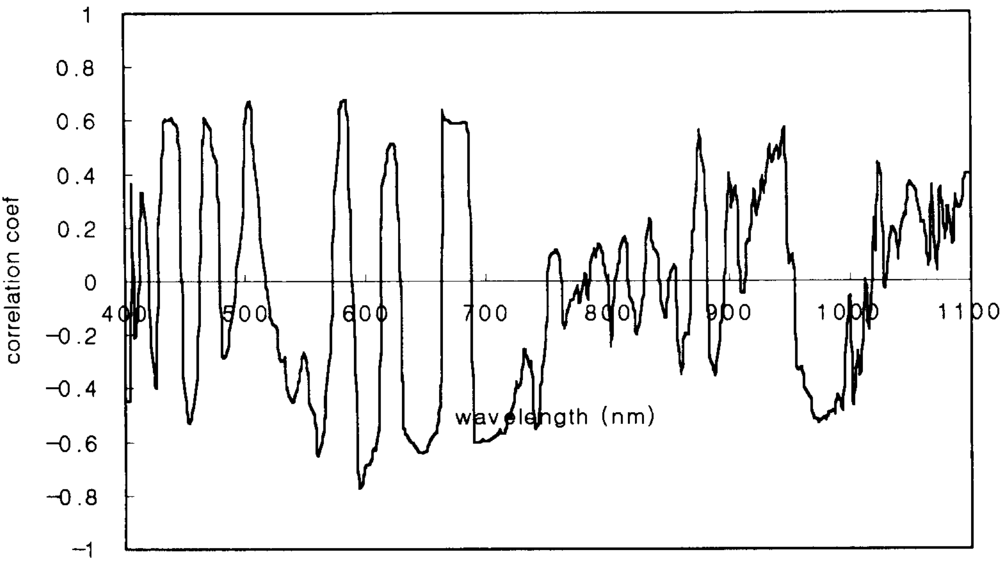
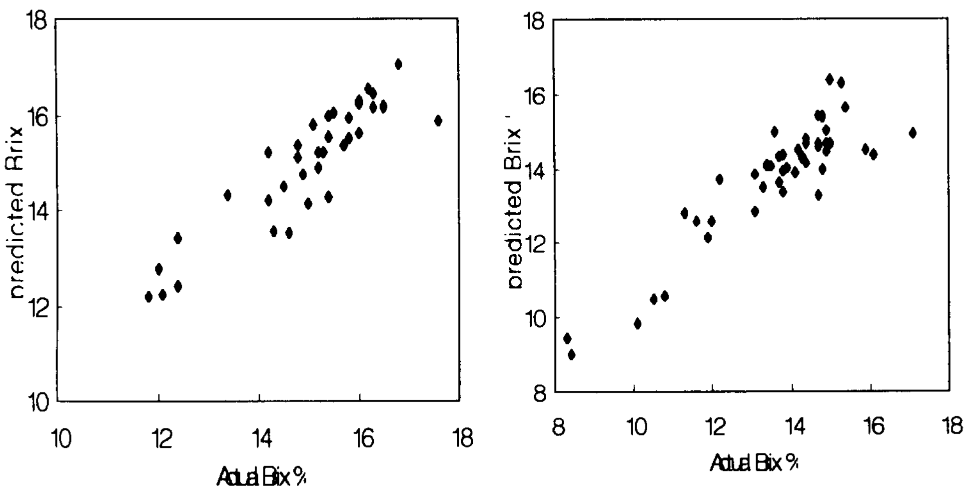


Fig 2. Correlation efficient between sugar content and 2nd derivatives spectrum for 'Singo' pear



(a) 36 non-stored apples (b) 47 stored samples

Fig 3. Result of sugar content prediction for stored and non-stored 'Fuji' apple by neural network

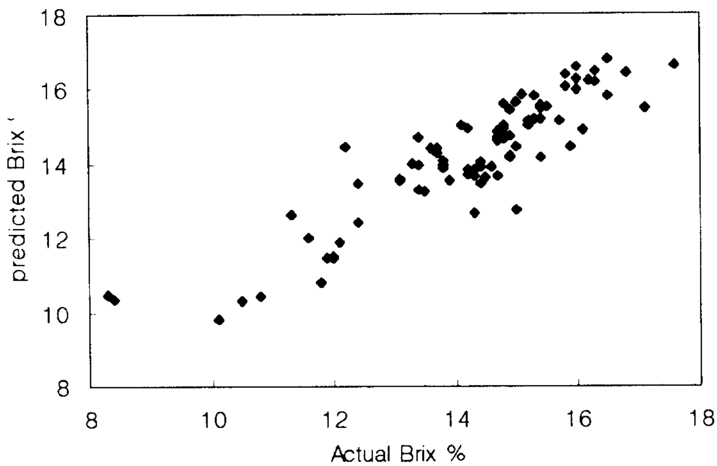
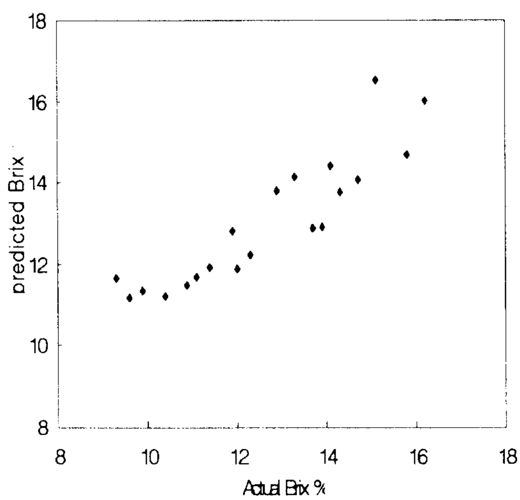
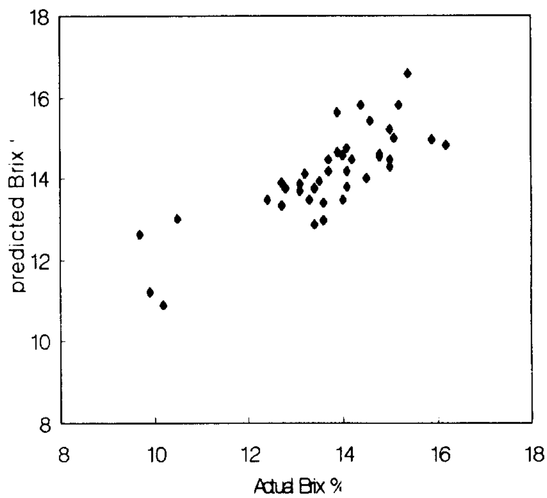


Fig 4. Result of sugar content prediction for total 'Fuji' apple (83 samples) by neural network



(a) peach



(b) pear

Fig 5. Result of sugar content prediction for peach and pear by neural network