

## Fourier Transform Infrared Spectroscopic Estimation of Residual Lignin Content in Wood Holocellulose (*Pinus brutia* Ten.) Using Artificial Neural Network Modelling

S. YASAR

Department of Forest Products Engineering, Faculty of Forestry,  
Suleyman Demirel University, Isparta-32260, Turkey

Fax: (90)(246)2371810; Tel: (90)(246)2113115; E-mail: syasar@orman.sdu.edu.tr

Wood samples from brutian pine (*Pinus brutia* Ten.) were delignified at different temperatures and different times to generate holocellulose samples of varying residual lignin content. Fourier transform infrared (FTIR) spectra of holocellulose samples were studied in the range of 4000-400  $\text{cm}^{-1}$  at a resolution of 4  $\text{cm}^{-1}$ . The spectral bands at 1508, 1421, 1372, 1265, 1158, 1054 and 1030  $\text{cm}^{-1}$  were used for estimating residual lignin content in the holocellulose samples. Artificial neural network (ANN) modelling was used to predict the amount of residual lignin from FTIR spectral data of holocelluloses. Artificial neural networks trained by Levenberg-Marquardt algorithm were applied for constructing and optimizing calibration models using MATLAB software. For four hidden neurons, three layered ANN model was scored an average relative error of prediction skill of 0.58 %.

**Key Words:** Residual lignin, Holocellulose, FT-IR, Artificial neural network.

### INTRODUCTION

Lignin, a natural complex biopolymer found in wood, is built up of phenylpropane units. Softwood lignins tend to consist mainly of guaiacyl units with minor amounts of syringyl and *p*-hydroxyphenylpropane units, whereas hardwood lignins tend to be composed of guaiacyl and syringyl units with a small portion of *p*-hydroxyphenylpropane units. With a contribution between 20-32 % of wood, lignin is the second most abundant biomass in the plant world after cellulose<sup>1,2</sup>.

Delignification is a process to remove the lignin from wood. Holocellulose is the product after delignification of wood. An ideal delignification should be result in a total removing of lignin without effecting hemicelluloses and cellulose, but there is no delignification technique which can perform this requirement. Therefore, low residual lignin content and minimal loss of hemicelluloses and cellulose in the holocellulose are expectations from delignification of wood<sup>1</sup>.

In the laboratory, chemical methods are applied for direct determination of lignin content, although these methods are laborious and time-consuming. Therefore, methods which can be used easily and cost effective are preferred for measuring

lignin content. For rapid estimation of lignin content, fourier transform infrared (FT-IR) spectroscopy. Calibrated FTIR measurements are able to give quantitative information about lignin content of wood or woodpulp samples<sup>3-7</sup>.

Artificial neural networks (ANNs) are computer based processing models which have a connected structure similar to brain cells. Theoretical background information on ANNs and the applications of ANNs for different systems have been published<sup>8-12</sup>. The applications of ANN for FTIR spectroscopic studies can also be found in the literature<sup>13-17</sup>. An ANN is composed of interconnected processing neurons or nodes organised in layers *i.e.*, input, hidden and output. The interconnections of neurons are quantified with connection weights which can be modified during the training step of ANN. The behaviour of ANN structure is determined by an activation function and a back propagation algorithm used for weights optimization. The connection weights are adapted to minimize the error between the target data values and their predicted data values by the neural network. Once the weights have been determined, the performance of the network is checked on a test set for determining the network's accuracy in predicting unprocessed data sets. For various networks constructed with different numbers of hidden layers and hidden neurons, the training and testing processes are repeated until minimum error value is obtained<sup>18,19</sup>.

In the present work, a typical softwood, brutian pine (*Pinus brutia* Ten.) was chosen. Since application of different delignifying conditions in the acid chlorite process<sup>20</sup> can change residual lignin content<sup>21</sup>, brutian pine wood was delignified using acid chlorite treatment at different temperatures and time periods to maximize the range of residual lignin content in holocelluloses. FTIR spectra from holocelluloses were collected and from these spectra, the rapid estimating residual lignin content was aimed using ANN modelling.

## EXPERIMENTAL

The wood material was milled in a Retsch SK 1 mill. Buchi extraction system B-811 was used for the alcohol extractions of wood sample, Perkin-Elmer Lambda 20 UV/VIS spectrometer for the absorbance value measurements and Perkin Elmer BX FTIR spectrometer for recording of the IR spectra.

Cyclohexane, ethanol, acetic acid, acetone, perchloric acid, sodium hydroxide and KBr for IR spectroscopy were purchased from Merck. Sodium chlorite and acetyl bromide were obtained from Fluka.

**Procedure:** The brutian pine (*Pinus brutia* Ten.) wood sample for analyses was milled to pass 40-100 mesh. The milled wood material was primary extracted with 2:1 cyclohexane:ethanol and secondary with ethanol. The extracted wood sample was further pulverized using agat mortar and pestle to have particle size from 20-45  $\mu\text{m}$ .

For delignifying of the pulverized wood material, the acid chlorite procedure of Browning<sup>20</sup> was adapted. 200 mg of pulverized wood sample (on oven dry basis) was placed in 50 mL flask and 32 mL of distilled water, 0.05 mL glacial acetic acid

and 0.15 g of sodium chlorite were added. Various numbers of the mentioned flask were heated at different temperatures and different times. The time periods were started at 15 min and continued with 15 min intervals to 330 min at a temperature of 70 °C. The same procedure was applied with 30 min intervals at a temperature of 50 °C. When the heating exceeded 30 min, acetic acid and sodium chlorite were added to the reaction mixture and it was repeated after each additional 30 min interval. At the end of each delignification period, the flask was cooled in an ice-bath. Holocellulose (delignified wood) in the flasks were filtered in G-3 crucibles with a light vacuum, washed with cold distilled water and acetone and air-dried.

10 mg of the each holocellulose sample (on oven dry basis) was dispersed in a matrix of 1000 mg KBr, followed by compression of 8 t to form self-supporting pellets. FTIR spectra of the samples were recorded at room temperature. Five scans per sample, totally 165 scans, were collected over the range 4000-400  $\text{cm}^{-1}$  with a spectral resolution of 4  $\text{cm}^{-1}$ .

Residual lignin content of each holocellulose sample was determined by acetyl bromide procedure from Liyama and Wallis<sup>22</sup>. 5 mg of sample (on oven dry basis) was digested with 2.5 mL of a mixture of acetyl bromide and acetic acid (25:75, w/w) and 0.1 mL of 70 % perchloric acid at the temperature of 70 °C for 0.5 h. After digestion, the solution was cooled and transferred to a 50 mL volumetric flask that contained 10 mL of 2 N sodium hydroxide and 12 mL of acetic acid. The solution was made up to 50 mL with acetic acid. A blank was prepared from the same solution without the 5 mg of sample. The absorbance values of solution and blank were measured at 280 nm. The residual lignin content was calculated from the blank-subtracted absorbance value and a specific absorption coefficient of 20  $\text{L g}^{-1} \text{cm}^{-1}$  and expressed in percentage of dry sample weight (w/w) for hemicelluloses. The determination of residual lignin content was carried out in duplicate for each sample.

**Methodology:** The training of ANN was carried out using MATLAB software. The ANN experiments used backpropagation with the Levenberg-Marquardt algorithm for adjusting the weights of networks. For this neural network modelling an input layer, a hidden layer and an output layer were constructed. The inputs consisted of seven neurons in the input layer which were intensities of the bands at 1508, 1421, 1372, 1265, 1158, 1054 and 1030  $\text{cm}^{-1}$  from the FTIR spectra. These bands were chosen based on previous work by Silva *et al.*<sup>7</sup>, where they were found to be informative for rapid determination of softwood lignin content. Baseline for each FTIR spectrum was corrected and defined by connecting the absorbance value at 1876 and 832  $\text{cm}^{-1}$ . The intensity of each selected band was obtained by measuring the peak height defined as absorbance value difference between the band maximum and the baseline. Average band intensities for 5 scans of each holocellulose sample were evaluated. A neuron which was the residual lignin content was used as the output layer (Fig. 1). Neural networks were trained using different number of hidden neurons.

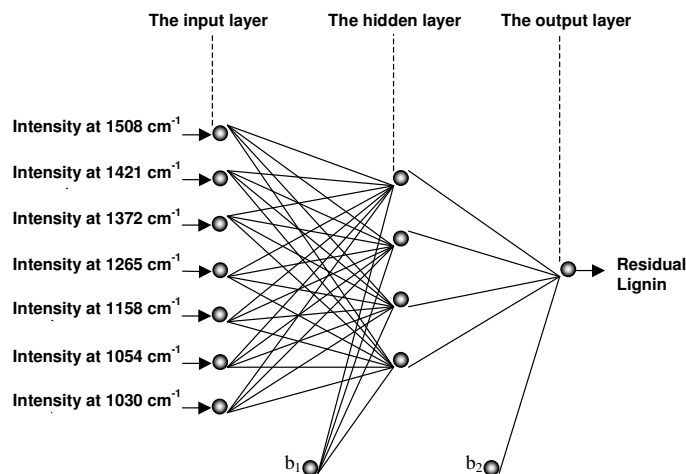


Fig. 1. Network architecture used in the fourier transform infrared (FTIR) spectroscopic method for estimation of residual lignin content

The 33 data pairs were divided randomly into three sets. Each data pair consists of the intensities of the bands at 1508, 1421, 1372, 1265, 1158, 1054 and 1030  $\text{cm}^{-1}$  from FTIR spectra and the residual lignin content, as input and output variables for ANN, respectively. Twenty two data pairs were used for training and 7 data pairs for testing. Four data pairs were used to show validation of the selected ANN model (Table-1).

TABLE-1  
TRAINING, TESTING AND VALIDATION DATA SETS FOR ANN

Sample	DT (min)	DT* (°C)	Data set	Sample	DT (min)	DT* (°C)	Data set
1	15	70	Training	18	180	50	Validation
2	30	70	Testing	19	195	70	Training
3	30	50	Training	20	210	70	Training
4	45	70	Training	21	210	50	Training
5	60	70	Training	22	225	70	Testing
6	60	50	Validation	23	240	70	Training
7	75	70	Testing	24	240	50	Testing
8	90	70	Training	25	255	70	Training
9	90	50	Training	26	270	70	Training
10	105	70	Training	27	270	50	Training
11	120	70	Training	28	285	70	Training
12	120	50	Testing	29	300	70	Validation
13	135	70	Testing	30	300	50	Training
14	150	70	Training	31	315	70	Testing
15	150	50	Training	32	330	70	Training
16	165	70	Training	33	330	50	Training
17	180	70	Validation	–	–	–	–

DT = Delignification time; DT\* = Delignification temperature.

When the logistic function was used as the activation function in a network, data should be normalized, so all data were normalized into a range 0.1-0.9 using the following equation:

$$X_N = 0.1 + \frac{0.8(X - X_{\min})}{(X_{\max} - X_{\min})} \quad (1)$$

where  $X_N$  is the normalized value of the input or the output data,  $X$  is original value of the data and  $X_{\max}$  and  $X_{\min}$  are the maximum and the minimum original values of the data. After training, the neural network was tested by introducing the test data set. Actual data (targets) were then compared with produced data by the network. If the network predicted data were in close agreement with actual data, then network topology was accepted. If it was not in agreement, the training process was repeated. The overall performance of the ANNs was evaluated in terms of root mean square (RMS) error according to the following equation:

$$\text{RMS} = \sqrt{N^{-1} \sum_{i=1}^N (X'_i - X_i)^2} \quad (2)$$

where  $N$  is the number of data,  $X'_i$  is the target value and  $X_i$  is the output value produced by the network.

## RESULTS AND DISCUSSION

In the FTIR measurements, although the band at  $1508 \text{ cm}^{-1}$  is characteristic to only lignin<sup>4,23-26</sup>, because of different ratios of carbohydrates, the intensities of the bands at  $1421$ ,  $1372$ ,  $1265$ ,  $1158$ ,  $1054$  and  $1030 \text{ cm}^{-1}$  are affected<sup>7</sup>. All these pure lignin and carbohydrate affected bands (Fig. 2) were used as inputs which are independent variables for the performed neural network models.

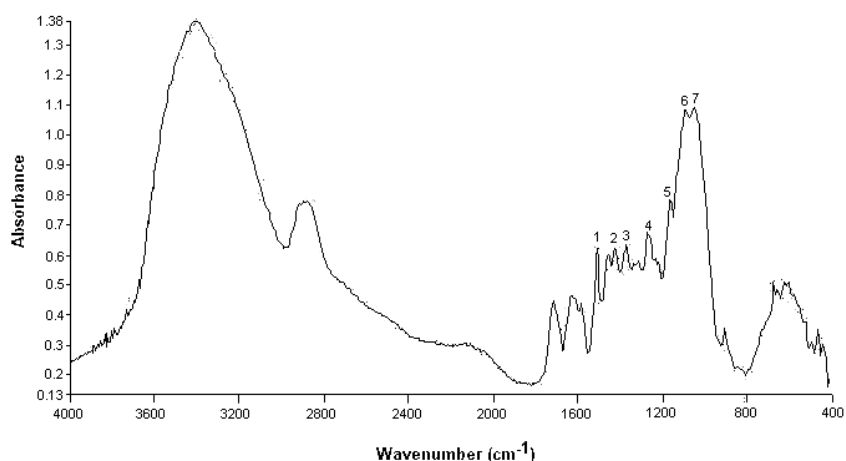


Fig. 2. FTIR spectrum of the brutian pine (*Pinus brutia* Ten.) 1:  $1508 \text{ cm}^{-1}$ , 2:  $1421 \text{ cm}^{-1}$ , 3:  $1372 \text{ cm}^{-1}$ , 4:  $1265 \text{ cm}^{-1}$ , 5:  $1158 \text{ cm}^{-1}$ , 6:  $1054 \text{ cm}^{-1}$ , 7:  $1030 \text{ cm}^{-1}$

The residual lignin contents of the holocellulose samples, between 15-330 min, are shown in Fig. 3. In the figure, the residual lignin content of each sample was expressed in percentage of dry sample weight (w/w). These values of the residual lignin contents were the output variables of neural networks. The residual lignin content values were varied between 6.51-26.52 %, 8.27-25.75 % and 9.32-25.13 % for the training, testing and validation of ANN, respectively.

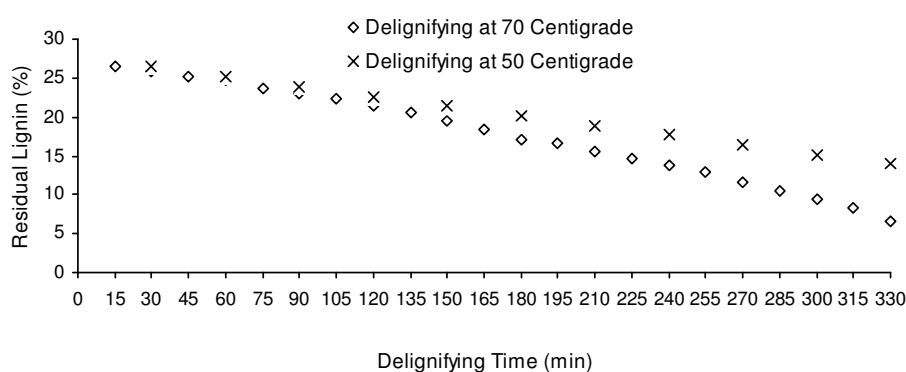


Fig. 3. Residual lignin contents (% dry matter) of the holocelluloses from brutian pine (*Pinus brutia* Ten.)

The neural network models with logistic function were trained and tested. The number of hidden neurons of the network was determined by evaluating the performance defined by RMS error of the network models (Table-2). The NN2 model with 4 hidden neurons was chosen to predict the residual lignin content because the RMS error for testing data set was the lowest value.

TABLE-2  
COMPARISON OF THE PERFORMANCES OF THE NEURAL NETWORK MODELS

Model	RMS error	
	Residual lignin	
	Training	Testing
NN1 7-3-1	0.022512	0.049672
NN2 7-4-1	0.010182	0.014216
NN3 7-5-1	0.010413	0.031405
NN4 7-6-1	0.010865	0.047836
NN5 7-7-1	0.004599	0.075956

The predicted values of residual lignin content from the NN2 model and the actual (target) values are shown in Fig. 4 for the training and testing data set by linear regression. The results showed that agreement between the predicted and the actual residual lignin contents was excellent with correlation  $R^2$  values of 0.9982 and 0.9963, respectively.

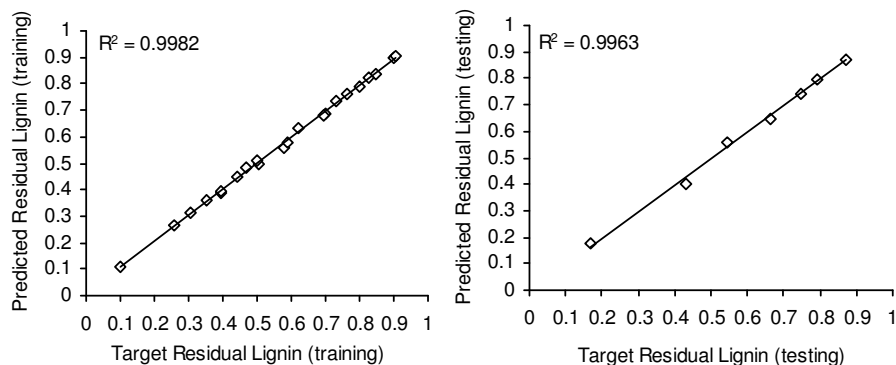


Fig. 4. The prediction performances of NN2 model for residual lignin training and testing data set

The experimental data pairs of brutian pine were used for validation step of model NN2. This run for validation is listed in Table-3. All of the predicted values were very close to actual values with an average per cent relative error of 0.58. It was therefore found that ANN can be used to successfully predict residual lignin content.

TABLE-3  
PREDICTION OF VALIDATION DATA SET BY NN2 MODEL

Delignification temperature and period	Residual lignin (%)		
	Actual	Predicted	RE %
70 °C, 180 min	17.15	17.25	0.58
70 °C, 300 min	9.32	9.38	0.64
50 °C, 60 min	25.13	25.24	0.44
50 °C, 180 min	20.14	20.27	0.65

## Conclusion

A three-layered ANN model was developed and tested for its ability to estimate the residual lignin content of wood samples taken from brutian pine (*Pinus brutia* Ten.). Samples were derived by delignifying at different temperatures and different times. Residual lignin contents of the holocellulose (delignified wood) samples were used as output variables for ANN. Input variables were the band intensities at 1508, 1421, 1372, 1265, 1158, 1054 and 1030  $\text{cm}^{-1}$ , taken from the FTIR spectra of the holocellulose samples. It was shown that the ANN could reliably predict residual lignin content with high correlation  $R^2$  values (0.9982 and 0.9963) between actual and predicted values of the training and testing sets, as well as with a small average prediction error (0.58 %) between actual and predicted values of a validation set.

## REFERENCES

1. D. Fengel and G. Wegener, *Wood-Chemistry, Ultrastructure, Reaction*, Walter de Gruyter, Berlin (1984).
2. E. Sjöström, *Wood Chemistry, Fundamentals and Applications*, Second Edition, Academic Press, Inc., San Diego (1993).
3. J.L. Grandmaison, J. Thibault, S. Kaliaguine and P.D. Chantal, *Anal. Chem.*, **59**, 2153 (1987).
4. J. Rodrigues, O. Faix and H. Pereira, *Holzforschung*, **52**, 46 (1998).
5. T.P. Schultz and D.A. Burns, *TAPPI J.*, **73**, 209 (1990).
6. T.P. Schultz, M.C. Templeton and G.D. McGinnis, *Anal. Chem.*, **57**, 2867 (1985).
7. J.C. Silva, B.H. Nielsen, J.C. Rodrigues, H. Pereira and U. Wellendorf, *Holzforschung*, **53**, 597 (1999).
8. P.D. Wasserman, *Neural Computing: Theory and Practice*, Van Nostrand Reinhold, New York (1989).
9. J. Zupan and J. Gasteiger, *Anal. Chim. Acta*, **248**, 1 (1991).
10. J. Zupan and J. Gasteiger, *Neural Networks for Chemists-An Introduction*, VCH, Weinheim (1993).
11. J. Zupan, *Acta Chim. Slov.*, **41**, 327 (1994).
12. C.M. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford (1995).
13. R. Goodacre, E.M. Timmins, R. Burton, N. Kaderbhai, A.M. Woodward, D.B. Kell and P.J. Rooney, *Microbiology*, **144**, 1157 (1998).
14. H.E. Johnson, R.J. Gilbert, M.K. Winson, R. Goodacre, A.R. Smith, J.J. Rowland, M.A. Hall and D.B. Kell, *Genetic Progr. Evol. Mach.*, **1**, 243 (2000).
15. F. Ehrentreich, *Anal. Chim. Acta*, **427**, 233 (2001).
16. P. Lasch, W. Haensch, E.N. Lewis, L.H. Kidder and D. Naumann, *Appl. Spectr.*, **56**, 1 (2002).
17. W. Yong, B. Link, R. O'Malley, J. Tewari, C.T. Hunter, C.A. Lu, X. Li, A.B. Bleecker, K.E. Koch, M.C. McCann, D.R. McCarty, S.E. Patterson, W.D. Reiter, C. Staiger, S.R. Thomas, W. Vermerris and N.C. Carpita, *Planta Med.*, **221**, 747 (2005).
18. G. Zhang, B.E. Patuwo and M.Y. Hu, *Int. J. Forecast.*, **14**, 35 (1998).
19. R.B. Boozarjomehry and W.Y. Svrcek, *Comp. Chem. Eng.*, **25**, 1075 (2001).
20. B.L. Browning, *Methods of Wood Chemistry, Volume 2*, Wiley-Interscience, New York (1967).
21. D. Fengel, H. Ucar and G. Wegener, *Das Papier*, **6**, 233 (1979).
22. K. Liyama and A.F.A. Wallis, *J. Sci. Food Agric.*, **51**, 145 (1990).
23. O. Faix, in eds.: S.Y. Lin and C.W. Dence, *Fourier Transform Infrared Spectroscopy, Methods in Lignin Chemistry*, Springer-Verlag, Berlin, pp. 83-109 (1992).
24. H.L. Hergert, in eds.: K.V. Sarkanen and C.H. Ludwig, *Infrared Spectra, Lignins: Occurrence, Formation, Structure and Reactions*, Wiley-Interscience, New York, pp. 267-297 (1971).
25. A.J. Michell, A.J. Watson and H.G. Higgins, *TAPPI J.*, **48**, 520 (1965).
26. N.L. Owen and D.W. Thomas, *Appl. Spec.*, **43**, 451 (1989).

(Received: 24 April 2009;

Accepted: 22 December 2009)

AJC-8210