USE OF FOURIER TRANSFORM INFRARED SPECTROSCOPY AND ARTIFICIAL NEURAL NETWORKS TO PREDICT THE WOOD DENSITY OF CEDRUS LIBANI A. RICH

Article *in* Fresenius Environmental Bulletin · May 2021

3 authors:

Ali ihsan Kaya
Mehmet Akif Ersoy University

Ahmet Çifci

Muhammer Ilkuçar
Mehmet Akif Ersoy University

USE OF FOURIER TRANSFORM INFRARED SPECTROSCOPY AND ARTIFICIAL NEURAL NETWORKS TO PREDICT THE WOOD DENSITY OF *CEDRUS LIBANI* A. RICH.

Ali Ihsan Kaya[1,*], Ahmet Cifci[2], Muhammer Ilkucar[3]

[1]Department of Design, Vocational School of Technical Sciences, Burdur Mehmet Akif Ersoy University, 15100 Burdur, Turkey
[2]Department of Electrical-Electronics Engineering, Faculty of Engineering and Architecture, Burdur Mehmet Akif Ersoy University, 15030 Burdur, Turkey
[3]Department of Management Information Systems, Faculty of Management, Mugla Sitki Kocman University, 48300 Mugla, Turkey

ABSTRACT

The aim of this study was to measure the wood density of *Cedrus libani* A. Rich. samples from its Fourier Transform Infrared Spectroscopy spectrum. 40 density values were obtained by using 3600 properties belonging to *C. libani* tree in laboratory environment. Since 1045 properties between 832-1876 from 3600 properties were found to be sufficient to determine the density, 1045 data between 832 and 1876 were used for training and testing of the network. Data used as attribute were normalized between 0.1 and 0.9. 20% of the data were used as the test set and the remaining 80% of the data are used as the training set. This analysis indicated that Fourier Transform Infrared Spectroscopy combined with Artificial Neural Network can be used to measure the density of wood in less effort and in less time than other laboratory methods.

KEYWORDS:
Artificial neural network, *Cedrus libani*, Fourier transform infrared spectroscopy, wood density

INTRODUCTION

Density is one of the characteristics that gives a quick information about the physical and chemical properties of wood materials. Conventional methods used for determination of chemical contents (cellulose, hemicellulose, lignin and extractives) as well as density determination require destructive methods and are time consuming [1-3]. Conventionally, density and chemical content determination methods are used to perform statistical analyses on samples taken from different regions in the forest [4]. This is particularly true for determining chemical properties using conventional chemical analysis methods involving long and tedious procedures [1-3].

New methods should be developed to analyze a large number of samples with minimum delay and to quickly evaluate wood properties. Therefore, several studies have been carried out to develop effective methods that can quickly yield results in such determinations. Fourier Transform Infrared Spectroscopy (FTIR) spectra is a recently developed method used to determine the physical and chemical properties of the tree and has been widely studied in the literature [5-9]. Due to its rapid and non-destructive nature, FTIR spectroscopy has become a popular method for density determination and chemical analysis [10]. Therefore, FTIR spectroscopy has found the widest application in the agricultural and food industry [11].

The structure of wood has been studied in many studies based on vibrational spectroscopic techniques [12]. In all of these techniques, the molecular structure of wood can be directly analyzed without strenuous preparation. Thus, it is possible to obtain information about the molecular level interactions between wood polymer components in their natural state. Moreover, these techniques are both time-saving and non-destructive [5-9]. Appropriate techniques for the rapid assessment of physical properties (density and grain angles [13]) are rapidly evolving. With researches, the use of absorbance bands corresponding to functional groups in biomass materials and the presence of the relationship between the bands and their chemical components is increasing [14]. With FTIR spectrum, the characteristic absorption bands of C-O, O-H, C-H and N-H groups, especially in the chemical structure of the wood material, can be determined quickly and analyzed clearly [12]. The success of the analysis is to establish a linear relationship between chemical components and density in the structure of the tree [15].

This paper describes the development of empirical models using projection to estimate the density of *Cedrus libani* A. Rich wood from FTIR spectrum of wood samples using principal component analysis.

MATERIALS AND METHODS

Sample Preparation. The samples were collected from randomly selected areas in Elmalı and Finike district of Antalya for this study. The height of the test trees varies from 14.8 to 16.2 meters, the diameter ranges from 35 to 65 cm and the age of the test trees ranges from 30 to 55 years. Samples were taken by removing 1 m trunk parts of the test trees from heights of 2-4 m. Test samples of 2 x 2 x 3 cm for density determination were cut from these samples. 30 test samples of juvenile wood and 30 test samples of mature wood were prepared. Test samples were prepared to be smooth fibrous, knotless, without cracks and defects. After pre-drying, the samples were stored in the air-conditioning chamber until they reached 65 ± 5% relative humidity and 12 ± 2% equilibrium humidity at 20 ± 2 ºC and their density was measured. The samples determined density were processed with a Retsch SK1 mill for FTIR analysis and passed through 40-100 mesh sieves. FTIR spectra were recorded on a Perkin Elmer (Spectrum BX) spectrometer in the wavelength range from 4000 and 400 cm^{-1} (Figure 1).

Artificial Neural Network. Although Artificial Neural Network (ANN) algorithm dates back to the 1950s, it has become one of the popular machine learning algorithms especially in recent years. ANN is frequently used in problems such as text/sentiment analysis, image processing, sound analysis and autonomous driving. ANN can perform clustering, classification, pattern recognition and prediction in many disciplines.

ANN is a computational model inspired by the working principle of brain cells [16]. As shown in Figure 2, it consists of perceptrons (nodes) that collect the signals received and generate an output.

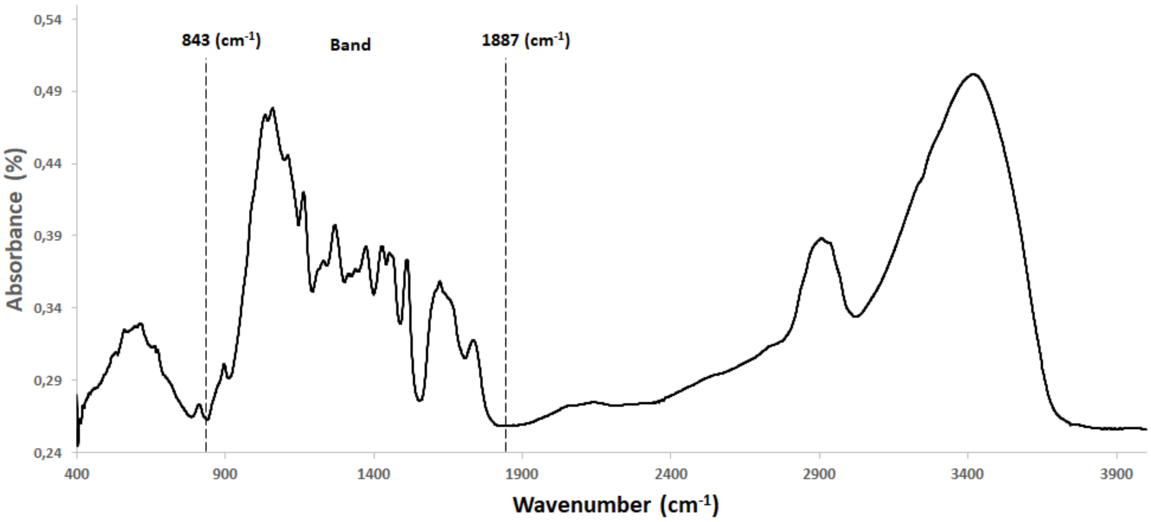


FIGURE 1
Typical FTIR spectra of *Cedrus libani* A. Rich. wood samples

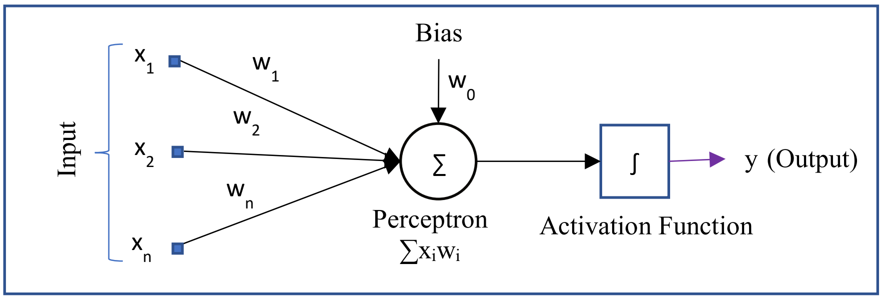


FIGURE 2
Perceptron (node) structure

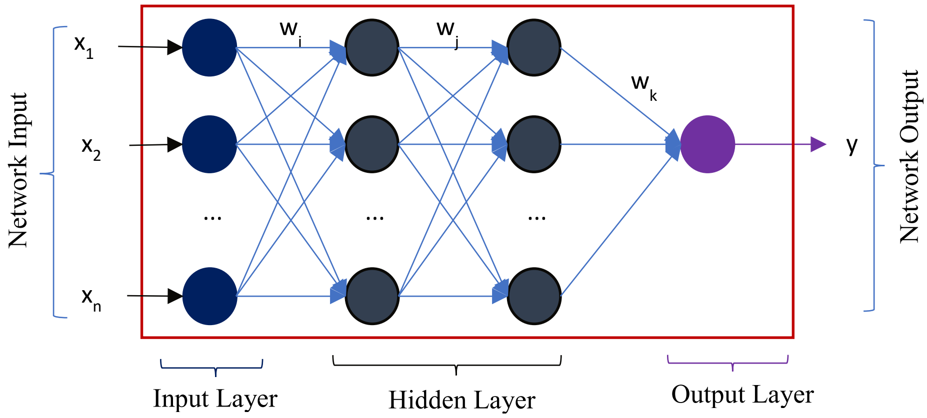


FIGURE 3
ANN architecture

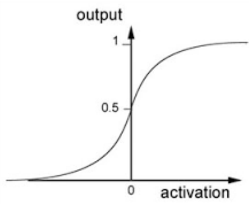
The perceptrons are interconnected and each connection has a weight value (w_i), as indicated in Figure 3. Perceptron output value is linear. On the other hand, since this value is in the form of $\sum x_i w_i$, it has a structure that can be big values. Therefore, the node is passed through the activation function to convert the output values to a nonlinear shape and to draw these values to a certain range. Although there are many different ANN structures in the literature, the multi-layer ANN architecture is used in our problem.

A typical ANN structure consists of three layers, called the input layer, hidden layer, and output layer. The number of input and output layer nodes depends on the nature of the problem. For instance, our data set has got 1045 attributes so the number of input layer occur 1045 nodes. The number of hidden layers can be one or more. There are no rules about how many hidden layers to use, it differs according to the problem. The problem is solved with different number of hidden layers to determine the most appropriate value. Another case is the number of nodes on hidden layers. There is no specific rule for determining this number. Grid search or random search is tried to be determined by.

(1) Activation Function. There are many different activation functions used in the literature. Figure 4 shows the most commonly used activation functions:

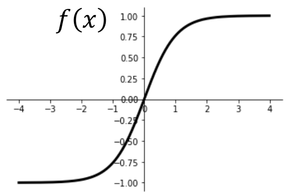
a. Sigmoid Function:

$$f(x) = \frac{1}{1+e^{-Net}}$$



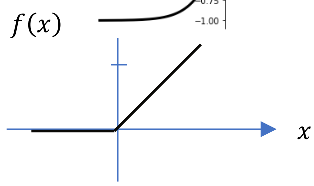
b. Hiperbolik Tanjant Function (Tanh):

$$f(x) = \frac{e^{Net} - e^{-Net}}{e^{Net} + e^{-Net}}$$



c. Rectified Linear Unit (ReLU) :

$$f(x) = \begin{cases} f(x) & if\ f(Net) > 0 \\ 0 & if\ f(Net) \leq 0 \end{cases}$$



d. Leaky Rectified Linear Unit (Leaky ReLU) :

$$f(x) = \begin{cases} f(x) & if\ f(Net) > 0 \\ a\, f(x) & if\ f(Net) \leq 0 \end{cases}$$

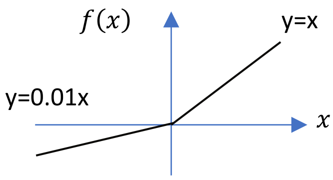


FIGURE 4
Activation functions

TABLE 1
Laboratory measurement information of *Cedrus libani* A. Rich. wood samples

Wood Density	Attribute (gr/cm^3)							
	4000	3999	3998	3997	…	402	401	400
0,444113198	55,17	55,17	55,17	55,16	…	55,37	53,76	52,53
0,449684211	69,65	69,65	69,64	69,64	…	69,51	68,29	67,32
0,444332405	72,62	72,62	72,62	72,62	…	70,57	69,66	69,34
0,467568338	66,69	66,7	66,7	66,7	…	64,8	64,1	64,08
0,472610097	65,46	65,45	65,45	65,45	…	64,56	63,67	63,46
0,452662268	72,73	72,73	72,72	72,72	…	70,86	69,45	69,07
0,447699653	74,64	74,63	74,63	74,63	…	73,65	72,64	72,15
0,458361648	58,37	58,37	58,36	58,36	…	57,51	56,71	56,35
0,442526316	68,31	68,3	68,3	68,3	…	67,86	67,04	66,57
0,454947368	71,69	71,69	71,69	71,69	…	70,8	70,23	69,95
0,483594798	72,4	72,4	72,4	72,4	…	73,03	72,42	71,94
0,471294855	67,69	67,68	67,68	67,67	…	65,58	64,17	63,79
0,468445165	63,55	63,55	63,55	63,54	…	62,25	61,58	61,15
0,449194415	56,22	56,22	56,22	56,22	…	54,52	54,05	54,32
0,468743287	73,83	73,82	73,82	73,82	…	72,95	72,14	72,03
0,478967097	70,08	70,07	70,07	70,06	…	68,07	67,18	66,93
0,43211162	56,81	56,81	56,8	56,8	…	56,37	55,62	55,3
0,455511958	73,28	73,28	73,27	73,27	…	73,2	72,46	72,24
0,454947368	71,29	71,29	71,27	71,28	…	70,51	70,15	70,09
0,478183301	65,36	65,37	65,35	65,36	…	64,74	64,59	64,83
0,478183301	65,64	65,65	65,63	65,63	…	63,81	62,49	61,96
0,497434052	72,49	72,48	72,49	72,49	…	70,66	69,98	70,12
0,43211162	60,19	60,2	60,19	60,19	…	59,31	58,81	58,68
0,451362371	69,12	69,1	69,1	69,09	…	65,5	63,89	63,88
0,468743287	66,5	66,49	66,49	66,5	…	65,22	64,61	65,2
0,481164339	69,11	69,12	69,1	69,12	…	69,62	69,55	69,93
0,444113198	71,64	71,64	71,64	71,63	…	72,2	70,6	69,47
0,456534251	72,62	72,62	72,62	72,62	…	70,57	69,66	69,34
0,447699653	67,79	67,8	67,77	67,77	…	66,8	65,15	63,99
0,466950404	75,43	75,39	75,4	75,4	…	71,41	69,69	70,4
0,466950404	72,82	72,76	72,79	72,78	…	67,01	64,75	65,67
0,456726593	68,77	68,78	68,75	68,76	…	65,17	64,21	63,86
0,43211162	67,83	67,79	67,82	67,82	…	63,92	63,35	65,09
0,42188781	71,88	71,77	71,86	71,84	…	65,76	63,89	66,9
0,434308863	71,04	71,04	70,97	71,02	…	71,73	69,18	64,01
0,446729915	70,15	70,05	70,11	70,03	…	70,76	70,87	70,41
0,436506105	68,81	68,77	68,8	68,81	…	62,29	62,41	64,96
0,413270173	71,84	71,78	71,81	71,79	…	68,64	65,69	65,8
0,425691225	62,74	62,73	62,72	62,72	…	62,14	61,18	60,56
0,413270173	74,77	74,81	74,75	74,76	…	74,2	73	70,8

The sigmoid function output value is bounded in the range [0,1]. Therefore, it ignores negative values. Another point is that after a few derivatives of the sigmoid function, the derivative result will converge towards zero (the vanishing gradient problem). The Tanh function output value ranges from -1 to 1, so that it takes into account negative data. However, as in sigmoid, convergence towards zero is also present here. The ReLU activation function is used to avoid convergence to zero. This function can be calculated very quickly when giving linear output for positive values, but ignores negative values. The Leaky ReLU activation function can be used to allow ReLU to leak negative values to a certain extent.

(2) Loss Function. ANN learning is the process of calculating the weight values that get the output values (feed forward) as close as possible to the expected actual value. The amount of difference (error) between the output value obtained and the actual value is reflected back to the weight values of the network and the weight values are changed (back propagation). The output error amount can be measured with the various types of loss functions (Mean Squared Error-MSE, Mean Absolute Error-MAE, Root Mean Squared Error-RMSE, etc.). In Equations (1), (2) and (3), MSE, MAE, RMSE are given, respectively. The loss function value is a parameter that affects learning because it will be reflected back to the weight values of the network.

$$Mean\ Squared\ Error\ (MSE) = \frac{1}{n}\sum_{i=1}^{n}(p^i - y^i)^2 \qquad (1)$$

$$Mean\ Absolute\ Error\ \ (MAE) = \frac{1}{n}\sum_{i=1}^{n}\left(\left|p^i - y^i\right|\right) \qquad (2)$$

$$Root\ Mean\ Squared\ Error\ (RMSE) = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n}(p^i - y^i)^2} \qquad (3)$$

where p is the predicted value, y is the actual value.

RESULTS AND DISCUSSION

40 density values were obtained by using 3600 properties of *C. libani* tree in the laboratory (Table 1). The absorbance values in the range of 1887 and 843 cm^{-1} were used for training and testing of the network (Figure 1). The density values of the data were rounded to three digits after the comma. The dataset was normalized between [0.1, 0.9]. 80% of the dataset was assigned as the training set while 20% of the dataset was assigned as the testing set.

The initial weight values of the network were determined randomly. No dropout was needed during learning. The hyper parameters used in the network are given in Table 2.

ANN was tested with different network models to obtain the best performance with the most economical network structure (Table 3). Thus, the complexity of the network and the computations will be reduced, increasing the possibility of operating on simple devices (such as microcontroller) of a smaller scale (RAM, CPU, etc.). Since the output data set is continuous, it is not appropriate to carry out analyzes such as validity and consistency with the help of confusion matrix.

TABLE 2
Hyper parameters of ANN model

Optimizer	Learning Rate	Decay	Momentum	Epoch	Hidden Layer Activation Function	Output Layer Activation Function
SDG	0.001	0.0001	0.6	1000	ReLU	Sigmoid

TABLE 3
Various ANN model for training and testing process

Model		Training Loss Function	Test MSE
NN1	1045-520-1	Figure 5(a)	0.023
NN2	1045-260-1	Figure 5(b)	0.020
NN3	1045-180-1	Figure 5(c)	0.021
NN4	1045-90-1	Figure 5(d)	0.022
NN5	1045-45-1	Figure 5(e)	0.023
NN6	**1045-22-1**	**Figure 5(f)**	**0.020**
NN7	1045-11-1	Figure 5(g)	0.022
NN8	1045-6-1	Figure 5(h)	0.035

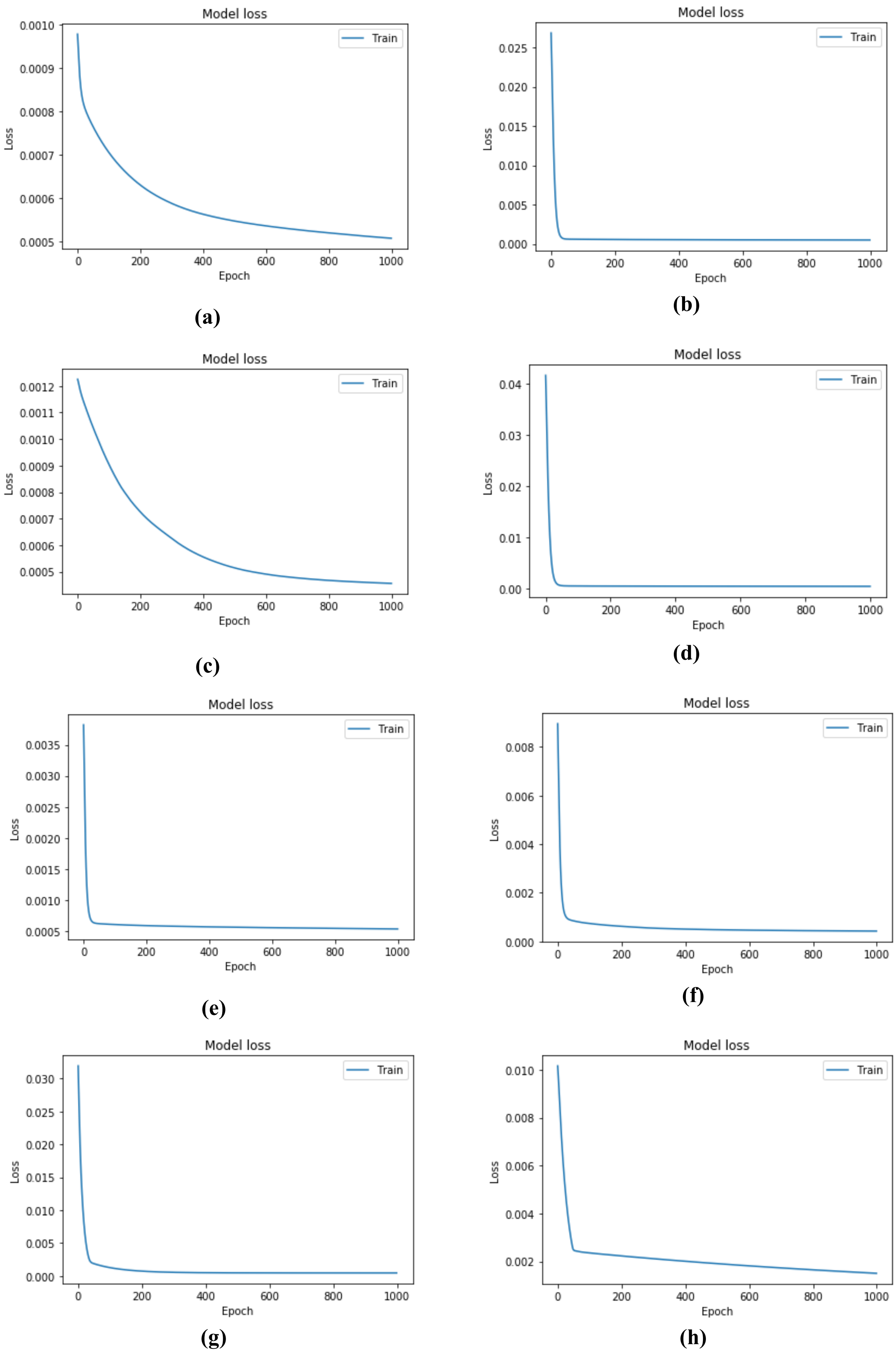


FIGURE 5
Training loss function (a) model NN1, (b) model NN2, (c) model NN3, (d) model NN4, (e) model NN5, (f) model NN6, (g) model NN7, (h) model NN8

CONCLUSION

In this study, the determination of the wood density of *Cedrus libani* A. Rich. from FTIR test data was estimated by multilayered ANN. For this purpose, different ANN parameters were tried with different hyper parameters and the most suitable structure was tried to be obtained. When using SDG as the learning algorithm and 1045-22-1 as the network architecture, the test data was estimated with an error of 2%. Since the actual output data is continuous, the analysis such as validity and consistency which can be obtained with the help of confusion matrix has not been performed. The output error amount is considered as the performance criterion of the network. The proposed approach can be considered as a time-saving, non-destructive and alternative way.

ACKNOWLEDGEMENT

This study was published as an abstract paper in International Conference on Science and Technology (ICONST 2019) hold from August 26 to 30, 2019, in Prizren, Kosovo.

REFERENCES

[1] Alves, A., Schwanninger, M., Pereira, H., Rodrigues, R. (2006) Analytical pyrolysis as a direct method to determine the lignin content in wood: Part 1: Comparison of pyrolysis lignin with Klason lignin. Journal of Analytical and Applied Pyrolysis. 76(1-2), 209-213.

[2] Syverud, K., Leirset, I., Vaaler, D. (2003) Characterization of carbohydrates in chemical pulps by pyrolysis gas chromatography/mass spectrometry. Journal of Analytical and Applied Pyrolysis. 67(2), 381-391.

[3] Wenming, H., Chongyun, X., Yi, N. (2010) The study of rapid predicted of moisture content and air-dried density of wood by using near infrared spectroscopy technology. China Pulp & Paper Industry. 31(6), 12-16.

[4] Mora, C.R., Schimleck, L.R., Isik, F. (2008) Near infrared calibration models for the estimation of wood density in *Pinus Taeda* using repeated sample measurements. Journal of Near Infrared Spectroscopy. 16, 517-528.

[5] Silva, J.C., Nielsen, B.H., Rodrigues, J.C., Pereira, H., Wellendorf, U. (1999) Rapid determination of the lignin content in Sitka spruce (*Picea sitchensis* (Bong.) Carr.) wood by Fourier transform infrared spectroscopy. Holzforschung. 53, 597-602.

[6] Meder, R., Gallagher, S., Mackie, K.L., Bohler, H., Meglen, R.R. (1999) Rapid determination of the chemical composition and density of *Pinus radiata* by PLS modelling of transmission and diffuse reflectance FTIR spectra. Holzforschung. 53, 261-266.

[7] Haukson, J.B., Bergqvist, G., Bergsten, U., Sjostrom, M., Edlund, U. (2001) Prediction of basic wood properties for Norway spruce. interpretation of near infrared spectroscopy data using partial least squares regression. Wood Science and Technology. 35, 475-485.

[8] Kelley, S.S., Rials, T.G., Snell, R., Groom, L., Sluiter, A. (2004) Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood. Wood Science and Technology. 38, 257-276.

[9] Yasar, S. (2010) Fourier transform infrared spectroscopic estimation of residual lignin content in wood holocellulose (*Pinus brutia* Ten.) using artificial neural network modelling. Asian Journal of Chemistry. 22(4), 2853-2860.

[10] Wesley, I.J., Larroque, O., Osborne, B.G., Azudin, N., Allen, H., Skerritt, J.H. (2001) Measurement of gliadin and glutenin content of flour by NIR spectroscopy. Journal of Cereal Science. 34, 125-133.

[11] Delwiche, S.R. (1998) Protein content of single kernels of wheat by near-infrared reflectance spectroscopy. Journal of Cereal Science. 27(3), 241-254.

[12] Uner, B., Karaman, İ., Tanriverdi, H., Özdemir, D. (2009) Prediction of lignin and extractive content of *pinus nigra arnold. var. pallasiana* tree using near infrared spectroscopy and multivariate calibration. Journal of Wood Chemistry and Technology. 29(1), 24-42.

[13] Evans, R., Downes, G., Menz, D., Stringer, S. (1995) Rapid measurement of variation in tracheid transverse dimensions in a radiata pine tree. Appita Journal. 48(2), 134-138.

[14] Schwanninger, M., Rodrigues, J.C., Gierlinger, N., Hinterstoisser, B. (2011) Determination of lignin content in Norway spruce wood by Fourier transformed near infrared spectroscopy and partial least squares regression. Part 1: Wavenumber selection and evaluation of the selected range. Journal of Near Infrared Spectroscopy. 15, 319-329.

[15] Li, X., Sun, C., Zhou, B., He, Y. (2015) Determination of hemicellulose, cellulose and lignin in moso bamboo by near infrared spectroscopy. Scientific Reports. 5, 1-11.

[16] Suzuki, K. (2011) Artificial neural networks - methodological advances and biomedical applications. InTech series of numerical analysis and scientific computing, Rijeka, Croatia

Received: **06.12.2019**
Accepted: **28.12.2020**

CORRESPONDING AUTHOR

Ali Ihsan Kaya
Department of Design,
Vocational School of Technical Sciences,
Burdur Mehmet Akif Ersoy University,
15100 Burdur – Turkey

e-mail: aikaya@mehmetakif.edu.tr