Cellulose, lignin and extractive material are mixed in certain proportions by having isolated from lignocellulosic materials, such as Zeyrek stem, hazelnut shell and Scotch pine, respectively. Their higher heating values (HHVs) are determined by using a bomb calorimeter system. Estimated HHVs are calculated by applying these mixture ratios to some Multiple (Non)-Linear Regression (M(N)LR) and Artificial Neural Network (ANN) models from the literature. MLR3 model is developed from the data of this study and this model reveals the highest $R^2$ (0.974), the lowest MAPE (0.012) and RMSE (0.278) values. The closest estimation accuracy to the MLR3 model is obtained from MLR2 ($R^2$:0.972, MAPE:0.066 and RMSE:1.714) in the comparative analysis. MNLR and ANN equations containing quadratic terms are found to show deviations up to 132.6% (ANN3). It is attributed to the lower size and poor homogeneity of the individual group of samples from which model equations are developed.

INTRODUCTION

A rapidly increasing population and industrialization create the need for energy. Biomass energy, a renewable source, is one of the sources that is used to ensure that energy can be provided sustainably without causing environmental pollution. Biomass is quite diverse due to the herbal and animal waste it contains. Herbal biomass contains cellulose, hemicellulose, lignin, lipids, simple sugars, starch, water, hydrocarbons, ash and other components. It is not a standard compound, whose chemical composition, physical properties and thermal properties vary considerably and it is important to determine these properties. One of these properties, the higher heating value (HHV), is an important parameter to consider in order for biomass to be converted into biofuel.

The HHV is important and necessary in order to make an energy analysis of a system. A calorimeter bomb is used to measure HHV. This method is complex and time-consuming due to the need for expensive experimental equipment, the devices that require certain expertise in their use, and the operations. Therefore, the number of studies aimed at developing useful theoretical equations to predict HHVs of lignocellulosic materials based on some analysis results (such as elemental analysis, proximate analysis or structural analysis) that can be obtained with much simpler and more basic instruments

* Corresponding author: incesevilay@gmail.com
available in any laboratory has increased in the last two decades.\textsuperscript{4,5}

These studies are mostly based on the Dulong approach developed for coal and often use elemental analysis results.\textsuperscript{6} While the initial equations developed for the estimation of HHVs of plant biomass focus on the results of elemental analysis,\textsuperscript{7,8} it is revealed that the equations based on the results of proximate analysis and/or structural analysis can also be used for this purpose in later studies. The number of studies based on structural analysis data is very limited.\textsuperscript{9}

HHV estimations, which are made by using structural analysis results, contain several difficulties and conveniences. Cellulose, lignin and extractive substance analysis can be performed relatively easily with simple laboratory equipment.\textsuperscript{10} However, cellulose has a hard and water-insoluble structure in addition to being a fibrous polymer.\textsuperscript{11} Lignin, on the other hand, is a biopolymer with a wide variety of functional groups and ten different bond types.\textsuperscript{12} These features make their structures even more complex. The chemical composition of a lignocellulosic material is closely related to its HHV, and this relationship can be observed by the variability of the HHV estimating equations based on chemical composition.\textsuperscript{13}

The effect of the interaction of variables on the result has been studied by using many modern techniques. The Artificial Neural Network (ANN) approach has been widely used over the last three decades in various applications such as classification, modeling, principal component analysis of data-driven nonlinear data,\textsuperscript{14} estimation of solar radiation and wind speed,\textsuperscript{15,16} geothermal heating systems,\textsuperscript{17} and food drying properties.\textsuperscript{18} In a study using the ANN approach, it has been revealed that time and cost-saving studies can be done with less error and without experimentation to obtain a hydrogen-rich gas mixture from the pyrolysis of lignocellulosic wastes.\textsuperscript{19} Multiple linear regression (MLR) is a technique usually used in statistics. Its purpose is to predict the possible outcome by revealing the relationship (for instance by drawing a graph) between multiple independent variables and a single dependent variable.\textsuperscript{20} Büyüka examined the thermogravimetric behavior of co-combustion of hazelnut shell-coal and walnut shell-lignite mixtures in two separate studies, and tried to estimate the percentage of mass loss with the help of the most appropriate MLR models.\textsuperscript{21,22} He also examined the thermogravimetric behavior of co-combustion of peanut and coal and estimated the mass loss in the combustion process with ANN models in another study.\textsuperscript{23} The method of ANN is used to determine the relationship between the chemical composition and properties of biodiesel.\textsuperscript{24}

According to the literature survey performed in the studies based on structural analysis, natural lignocellulosic samples are mostly used, but there are no studies based on the experimental HHV measurement results of mixtures prepared in certain and desired proportions of components such as pure lignin, cellulose and extractives. It is seen that the average of the lignin contents of the samples (approximately 90 samples) used in the examined articles is between 14.3\%–43.7\% and the average of the extractives is between 3.7–22.6\%.\textsuperscript{13,25–28} However, it is known that there are lignocellulosic materials with lignin and extractives apart from these values.\textsuperscript{29,30} It is thought that the samples with such extraordinary values are not incuded in the study, so this will lead to a relatively narrowing of the range of linearity of the estimation equation obtained. This, in turn, will affect estimation performance of the model.

Lignocellulosic components of plant stems, scapes, branches and leaves are processed in various ways using thermal or biochemical processes to produce biofuels. These components can determine the characteristics such as HHV of the obtained biofuel. The studies in the literature usually consist of the analysis of natural samples and correlating them (components and HHVs) to each other. In this study, the estimating power of the traditional Multiple (Non) Linear Regression (M(N)LR) analysis and ANN-based models previously developed in the literature were compared by using the experimental HHVs of the samples prepared at certain proportions by mixing cellulose, lignin and extractives isolated from natural samples. In addition, whether there was a statistically significant improvement in prediction power with the use of ANN-based equations was investigated. Thus, other known and/or unpredictable factors that might affect the combustion process are minimized, and only the interactions between the energy-intensive components of the lignocellulosic biomass (lignin and extractive content for this study) were taken into account and their effects on the prediction equations were examined more closely.
EXPERIMENTAL

Sampling and preparation for analysis

The powdered samples of Scotch pine (*Pinus sylvestris*), *Linum usitatissimum* L. stem (locally known as Zeyrek) and hazelnut shell (*Corylus avellana*) were kept in a drying-oven at 100°C for 3 hours to dry them. Then, they were ground in an IKEA A11 model laboratory-type mill. The dried and ground biomass samples were separated into various particle sizes with the help of the Retssch AS 200 model vibrating sieve (15–20-minute-sieving time at 60 Hz intensity). The samples with a particle size of 125–250 μm were used in the analyses.

Original wood analyses and preparation of samples

The analyses conducted on the samples and the corresponding standards are presented in Table 1. The procedure that was used in the study about the preparation of the mixture samples is as follows. The amounts of cellulose, lignin and extractive substances weighed out to the nearest 0.1 mg were mixed at predetermined ratios and stored in a desiccator for HHV measurements. The applied mixing ratios were determined based on literature data. Accordingly, the percentage of lignin was changed from 10 to 50% while the percentage of extractive substance was changed from 5 to 45%. Cellulose was used as the carrier medium of lignin and extractives are (such as softwood, hardwood and straw). The extractive substance used was prepared as a solution whose concentration is known exactly due to its viscous feature. Then, that solution was impregnated into the lignin-cellulose mixture in the weighing bowl (in that case, the decomposition vessel of the calorimeter system) and dried in a drying-oven until a constant weight is obtained. The HHVs of the samples were measured by using the IKA C2000 calorimeter device.

Statistical analysis and software

All the experimental works were conducted in two parallels and the results were expressed as an average of these measurements. IBM SPSS statistics for Windows (v. 20.0, IBM Corp., Armonk, NY, USA) and Apple MATLAB R2015b for Mac OSX (v. 8.6, MathWorks, Massachusetts, USA) package programs were used for calculating first and second order multiple regression analysis models.

Selection of regression models

It was found in the literature that there is a linear relationship between carbon content and HHV for different wood species. It is also known that the amount of lignin in different woody biomass species generally varies between 20% and 40%, and it is demonstrated by various studies that it has a positive effect on HHV due to its high carbon content. Extractives, on the other hand, are found in less amounts than other components of biomass. However, they can provide HHV due to their content of complex and variable compound profiles such as waxes, alkaloids, simple sugars, proteins, simple and complex phenolics, pectins, mucilages, resins and terpenes. Therefore, the effect of these two component types is considered important in many studies to predict HHVs from structural component analysis values of lignocellulosic materials. Hence, the models that highlight the contribution of these two structural components are taken into account in the selection of the equations included in the present study. One MLR equation is obtained from the data of this study using the IBM SPSS Statistical program package (MLR3), while two MLR, one MNLR and three ANN equations are selected from the studies in the literature.

Multiple (non)linear regression analyses (M(N)LRAs)

Regression analysis is a statistical technique for estimating the relationship between variables that have cause and effect connection. The main focus of univariate regression is to analyze the relationship between a dependent variable and an independent variable and formulate the linear relationship equation between them. Regression models that include one dependent variable and more than one independent variable are called multiple linear regression. In the study, multiple linear regression (MLR) was used to explain the effect of variables such as lignin and extractive contents on HHV. The relationships between these dependent and independent variables can be linear or non-linear (MLR or MNLR). Nonlinear regression (MNLR) is a form of regression analysis that is modeled by a function which is a nonlinear combination of data and include one or more independent variables. Multiple linear regression (MLR) analysis has been used in many studies in the literature to explain the effect of cellulose, lignin and extractive contents on the HHV. In some studies, Multiple Nonlinear Regression Analysis (MNLR) in which equations containing quadratic terms that take into account the interaction contributions of these components to HHV are obtained is used.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moisture content (wt%)</td>
<td>ASTM D2016-74</td>
</tr>
<tr>
<td>Ash</td>
<td>ASTM D1102-84</td>
</tr>
<tr>
<td>Volatile compounds</td>
<td>ASTM E997-82</td>
</tr>
<tr>
<td>Fixed carbon</td>
<td>ASTM E870-82</td>
</tr>
<tr>
<td>Cellulose</td>
<td>Kürschner-Hoffer's nitric acid method</td>
</tr>
<tr>
<td>Lignin</td>
<td>TAPPI T222-om02</td>
</tr>
<tr>
<td>Extractives</td>
<td>TAPPI T204-om88</td>
</tr>
<tr>
<td>HHV</td>
<td>DIN 51900-2</td>
</tr>
</tbody>
</table>

Table 1

The analyses conducted on the samples and the corresponding standards
Artificial neural networks (ANNs), inspired by the information processing technique of the human brain, have recently found a place in many different engineering applications. In particular, ANN has received significant attention due to its high performance in classification, modeling and estimating applications. With ANN, the way the simple biological nervous system works is imitated. In other words, it is the digital modeling of biological neuron cells and the synaptic bond that these cells establish with each other. An artificial neuron is the core processing unit of an ANN and can be implemented in a variety of ways. Figure 1 depicts the general architecture of an artificial neuron. In this manner, the input from the previous layer neuron's output is multiplied by the weight value. Usually, initial weights are assigned at random. As can be seen in Figure 1, the inputs, weights (W), activation function (f) and output are the basic components that make up a simple artificial neuron.

In Figure 1, the output of a neuron can be described as follows:

\[ \text{output} = f \left( \sum W_i \cdot \text{input} \right) \]  

(1)

Neurons form networks by connecting to each other in various ways. These networks have the capacity to learn, memorize and reveal the relationship between data. In other words, ANNs produce solutions to problems that normally require a human’s natural abilities to think and observe. ANN takes data samples instead of whole datasets to arrive at solutions, which saves both time and money. ANNs are considered fairly simple mathematical models to improve existing data analysis technologies. Figure 2 shows a simple ANN structure. The most basic task of ANN is to learn the structure in the sample data set and make generalizations to fulfill the desired task. For this purpose, a wide variety of network structures and models have been developed in artificial neural networks. Multilayer Perceptron (MLP), Radial Based Function Networks (RBF), Learning Vector Quantization (LVQ), Hopfield Networks, SOM Networks and Adaptive Resonance Theory Networks (ART) are some of the ANN models used in different fields.

RESULTS AND DISCUSSION

Results of original wood analysis

The original wood analyses conducted on the raw materials and the results obtained are given in Table 2 below.

<table>
<thead>
<tr>
<th>Material</th>
<th>Moisture (wt. %)</th>
<th>Ash (wt. %)</th>
<th>Volatile compounds (wt. %)</th>
<th>Fixed carbon (wt. %)</th>
<th>Cellulose (wt. %)</th>
<th>Lignin (wt. %)</th>
<th>Extractives (wt. %)</th>
<th>HHV (MJ/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zeyrek stem</td>
<td>4.24±0.08</td>
<td>11.82±0.05</td>
<td>62.70±0.04</td>
<td>21.20±0.06</td>
<td>36.07±0.04</td>
<td>22.42±0.08</td>
<td>2.46±0.34</td>
<td>18.33±0.06</td>
</tr>
<tr>
<td>Hazelnut shell</td>
<td>5.29±0.06</td>
<td>1.96±0.07</td>
<td>65.12±0.09</td>
<td>27.62±0.05</td>
<td>32.90±0.03</td>
<td>52.78±0.06</td>
<td>1.64±0.42</td>
<td>21.03±0.05</td>
</tr>
<tr>
<td>Scotch pine powder</td>
<td>4.49±0.11</td>
<td>1.43±0.08</td>
<td>73.31±0.08</td>
<td>20.82±0.07</td>
<td>35.35±0.06</td>
<td>30.29±0.05</td>
<td>5.07±0.28</td>
<td>20.18±0.03</td>
</tr>
</tbody>
</table>

* The percentages of moisture, ash, volatile compounds, extractives and fixed carbon were calculated based on air dried samples, while the percentages of lignin and cellulose were calculated based on extractive free, air dried samples.

Table 2

The results of original wood analysis of the samples (wt, %) and HHVs (MJ/kg).
Extractive content based models

HHV values of cellulose (from Zeyrek), lignin (from hazelnut shell) and extractive material (from Scotch pine) were 15.530±0.339, 22.582±0.427 and 35.310±0.190 MJ/kg, respectively. It seems that higher lignin percentage caused a higher HHV value of the hazelnut shell. This was followed by the HHV of Scotch pine powder, which had the second higher lignin value. In addition, Scotch pine powder had the highest percentage of extractives among the three samples. It is a known fact that natural lignocellulosic samples with a high percentage of lignin and extractive material show higher HHV values.5,9,10

Results of HHV values of sample mixtures

Mixture ratios and corresponding HHV values for each sample that was prepared are given in Table 3.

According to the literature studies, while the lignin value increases, the HHV value also increases. However, according to the data of this study, it is seen that the HHV value increases while the lignin value decreases (see Table 3). It is seen that the reason for this might be the increasing rate of extractive material, which is another energy-intensive component. It is known that the components such as waxes, alkaloids, resins and terpenes in the composition of the extractives provide high HHVs, although it varies according to the kind of biomass.9 Due to the high variability of its content, it can cause large deviations in the HHV prediction models in which it is included.25,28

Comparative evaluation of M(N)LR and ANN regression models

The regression models and related statistical data obtained from the literature and the results of this study to be used to determine the effect of lignin and extractive substance ratios on HHVs are presented in Table 4, and Figure 3 below, respectively.

Table 4 and Fig. 3 show that R² values increased and MAPE values decreased as one goes from MLR1 to MLR3. The R² value of the MNLR equation (0.563) is smaller than those of these equations, while the MAPE value (0.128) is higher. It is also seen that the R² values of the ANN equations were lower than those of MLR equations, and the MAPE values were much higher. However, a more fundamental problem here was that the R² values of the ANN1 and ANN2 equations, which were relatively comparable to the others, corresponded to a negative correlation. The graphs of the experimental HHVs against the predicted values drawn for each model are given below to see the situation in this respect.

As can be seen in Figure 4, the relationship calculated by using the MNLR model was a negative relationship. Therefore, since the expected positive linear relationship could not be obtained, the calculated statistical parameters did not fully define the fitting of the model.

Table 3

Lignin, cellulose and extractives ratios and measured HHVs of the mixture samples

<table>
<thead>
<tr>
<th>Sample Code</th>
<th>Cellulose (% w/w±SD)</th>
<th>Lignin (% w/w±SD)</th>
<th>Extractives (% w/w±SD)</th>
<th>HHV (MJ/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>44.99±0.01</td>
<td>49.99±0.01</td>
<td>5.02±0.01</td>
<td>20.71±0.049</td>
</tr>
<tr>
<td>S2</td>
<td>44.99±0.02</td>
<td>45.01±0.03</td>
<td>10.00±0.04</td>
<td>21.96±0.074</td>
</tr>
<tr>
<td>S3</td>
<td>44.97±0.00</td>
<td>40.05±0.00</td>
<td>14.98±0.00</td>
<td>22.66±0.056</td>
</tr>
<tr>
<td>S4</td>
<td>44.99±0.03</td>
<td>35.06±0.02</td>
<td>19.95±0.04</td>
<td>22.81±0.200</td>
</tr>
<tr>
<td>S5</td>
<td>45.00±0.03</td>
<td>30.11±0.02</td>
<td>24.89±0.04</td>
<td>23.16±0.152</td>
</tr>
<tr>
<td>S6</td>
<td>45.07±0.05</td>
<td>25.00±0.03</td>
<td>29.93±0.06</td>
<td>24.52±0.316</td>
</tr>
<tr>
<td>S7</td>
<td>44.57±0.03</td>
<td>20.86±0.01</td>
<td>34.57±0.03</td>
<td>25.32±0.131</td>
</tr>
<tr>
<td>S8</td>
<td>44.85±0.03</td>
<td>15.39±0.01</td>
<td>39.76±0.03</td>
<td>26.12±0.260</td>
</tr>
<tr>
<td>S9</td>
<td>44.96±0.05</td>
<td>10.22±0.01</td>
<td>44.82±0.05</td>
<td>26.37±0.066</td>
</tr>
</tbody>
</table>
Table 4  
Regression equations used and corresponding statistical data

<table>
<thead>
<tr>
<th>Prediction Models</th>
<th>Regression equation</th>
<th>Ref.</th>
<th>MSE</th>
<th>R²</th>
<th>RMSE</th>
<th>MAPE</th>
<th>MAE</th>
<th>D-W*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNL R</td>
<td>HHV=20.137+0.008L+0.131E-0.004E²</td>
<td>37</td>
<td>17.326</td>
<td>0.563</td>
<td>4.162</td>
<td>0.128</td>
<td>3.214</td>
<td>-</td>
</tr>
<tr>
<td>MLR1</td>
<td>HHV=14.337+0.122L+0.135E²</td>
<td>25</td>
<td>8.102</td>
<td>0.962</td>
<td>2.846</td>
<td>0.097</td>
<td>2.411</td>
<td>-</td>
</tr>
<tr>
<td>MLR2</td>
<td>HHV=15.605+0.074L+0.172E²</td>
<td>26</td>
<td>2.938</td>
<td>0.972</td>
<td>1.714</td>
<td>0.066</td>
<td>1.602</td>
<td>1.6193</td>
</tr>
<tr>
<td>MLR3</td>
<td>HHV=28.179+0.880L+1.019E²</td>
<td>This study</td>
<td>0.077</td>
<td>0.974</td>
<td>0.278</td>
<td>0.012</td>
<td>0.247</td>
<td>2.181</td>
</tr>
<tr>
<td>ANNI</td>
<td>HHV=17.0000+0.0337689[L]-0.370197[E]+0.0217811[EL]-0.0001237[EL²]-0.00026906[LE²]</td>
<td>26</td>
<td>124.799</td>
<td>0.864</td>
<td>11.171</td>
<td>0.334</td>
<td>8.431</td>
<td>1.8845</td>
</tr>
<tr>
<td>ANN2</td>
<td>HHV=18.9886-0.0821496[L]-0.788837[E]+0.047379[LE]-0.009044862[El²]-0.000301034[LE²]+0.00146961[LE²]</td>
<td>26</td>
<td>195.114</td>
<td>0.779</td>
<td>13.968</td>
<td>0.382</td>
<td>9.721</td>
<td>1.8455</td>
</tr>
<tr>
<td>ANN3</td>
<td>HHV=20.7165-0.141898[L]-1.78999[E]+0.0763139[LE]-0.0006305[EL²]-0.00249472[EL²]+0.00194996[LE²]+0.1129688[E²]</td>
<td>26</td>
<td>0.003</td>
<td>0.813</td>
<td>53.576</td>
<td>1.379</td>
<td>35.296</td>
<td>1.8179</td>
</tr>
</tbody>
</table>

*: D-W values are taken from the relevant literature, L: Lignin percentage, E: Extractives percentage.

When the MLR equations were evaluated among themselves, MLR3, which revealed a positive linear relationship, gave the highest R² (0.974), the lowest MAPE (0.012) and RMSE (0.278) values, which were expected in terms of representing the sample. The standard error of the estimation in the model was 0.339. In addition, 46.26% of the variation in HHV could be explained by the lignin content, while 53.74% could be explained by the extractive content. While these values were 62.4% and 37.6%, respectively, in the MLR2 model, which revealed the closest prediction performance to this model, these values are 56.4% and 43.6% in the MLR1 model. A greater part of the change in HHV, in MLR1 and MLR2 models was caused by lignin, while it was caused by the extractives in the model obtained from this study (MLR3). Among the possible reasons for this situation, the extractive material used in the current study was obtained from a single type (Scotch pine) sample, and the extractive ratios were up to 45% in contrast to the samples used in MLR1 and MLR2 studies (generally below 15%).
It is important that there is no autocorrelation in the creation of MLR models. Durbin-Watson statistic is used to determine this situation. It is also applied to optimize model stability and distinguish important independent variables from unimportant ones. If there is a random distribution in the model, the value of the D-W statistic will be close to 2. In general, this value should be between 1.5–2.5. The value of the D-W coefficient obtained in this study is 2.181, which reveals that there is no autocorrelation in the study. The obtained D-W coefficient stands out as the best value among the examined models.

Theoretical HHV values calculated by applying the mixture ratios prepared in this study to the equations selected from the literature are given in Table S1 in comparison with the experimental HHV values. The percentage relative error between the experimental and calculated HHVs of the MNLR and MLR equations did not exceed 15.71% for the samples in which the extractive substance value was below 30% (Table S1). However, it is noteworthy that larger estimation deviations occurred in samples where the extractive substance value exceeds 30% (especially in the MNLR model). When the relevant studies in the literature were examined, was noticed that the extractive substance percentages were generally below 30%. The lignin and extractive substance contents of the samples used in all three studies varied between 14.66–57.36%, and 1.23–28.28%, respectively, and these changes did not show homogeneity in each group of the samples. In addition, considering the limited number of samples in each group (N= 12, 17 and 11) from which the relevant
estimation equations were obtained, it could be concluded that the derived equations reflected the character of these samples in a way. Sheng and Azvedo (2005) also obtained the results supporting this idea in their study.

When the model fit of the ANN equations in Fig. 4 and Table 4 were considered, ANN1 and ANN2 models showed a negative relationship, while ANN3 showed a positive linear relationship but had very large deviations in the estimation results (RMSE=53.576, MAPE=1.379). Considering the experimental and calculated HHVs of the ANN-based equations, it was seen that the percentage relative error increased in samples with 20% or more extractive substance percentage (Table S1). For instance, the relative error rates of the ANN equations reached 40% for samples with extractive substance ratios of 30% and above, and even reached 132.6% for the ANN3 model.

While it is possible that the inclusion of quadratic terms in the equations can make a positive contribution to the accuracy of the estimation results calculated from the model obtained in the relevant study using a limited number of samples, it is obvious that these models will create large deviations when applied to samples where the distribution of lignin and extractive substance ratios are more homogeneous as in the current study. In fact, the main purpose of deriving such equations is to obtain the highest estimation model fit by using the least number of independent variables. It can even be seen from the low $R^2$ and high MAPE values obtained from the MNLR equation that the contribution of such terms can significantly alter the estimation accuracy. It is worthy of note that a quadratic term ($E^2$) that is not in the other MLR equations is included in this equation.

Since the samples, from which the estimation models are derived, are limited, their estimation performances are also limited. Moreover, the addition of quadratic terms to the model causes large deviations in the estimation results obtained when the model is applied to different samples, unless it is studied with sufficiently large and homogeneous samples as much as possible.

**CONCLUSION**

Experimental HHVs of the study samples obtained by mixing cellulose, lignin and extractive substances, respectively, isolated from lignocellulosic wastes such as zeyrek stem, hazelnut shell and Scotch pine in certain proportions were determined. The estimation ability of traditional M(N)LR and ANN based models previously developed in the literature were examined comparatively by using these HHVs and lignin-extractive substance ratios. When these models were applied to samples with different ratios of lignin and extractives than their own samples, they revealed large deviations in their estimation results (exceeding 130%). It has been shown that having quadratic terms in models with limited sample size (<20 for the models included in the study) revealed large deviations in estimation results when the model was applied to different samples. This demonstrated the importance of the sample size being as large as possible, and the component (in this case lignin and extractives) ratio variability being homogeneous in deriving equations for the estimation of HHVs of lignocellulosic wastes. The MLR3 model obtained from this study data showed the highest model fit.

In future studies, lignin, cellulose and extractive material values of various lignocellulosic materials and their experimental HHVs are planned to be collected from the literature. Then, the agreement of the experimental HHVs of these data with the theoretical HHVs calculated using the model obtained from the current study (MLR3) will be investigated.

**Acknowledgements.** The author thanks Dr. Fikret Akdeniz for his contributions and support at every stage of the study, Rojda Naz Ali Beyoglu for her support in the design of the graphical abstract, and Dr. Hüseyin Ertap for his support in the graphic drawing, and Dr. M. Ali Karabulut for his theoretical scientific support.

**REFERENCES**
