PREDICTING RF HEATING RATE DURING PASTEURIZATION OF GREEN SOFTWOODS USING ARTIFICIAL NEURAL NETWORKS AND MONTE CARLO METHOD

Bogdan Bedelean
Transilvania University of Brasov, Faculty of Wood Engineering
Department of Wood Processing and Wood Products Design
Brasov, Romania

Ciprian Lazarescu, Stavros Avramidis
University of British Columbia, Department of Wood Science
Vancouver, BC, Canada

(Received November 2013)

ABSTRACT

In this paper we tested if the value of the radiofrequencies (RF) heating rate during pasteurization of two green softwoods, namely, red cedar and pine, can be estimated by means of an artificial neural networks model that is solved using the Monte Carlo method (MCM). Based on the proposed approach, the value of the RF heating rate was predicted reasonably with a relative error of 13% for red cedar and 8% for pine. Moreover, the sensitivity analysis revealed that the RF heating rate is more sensitive to the moisture content and less to power density. In addition, the model showed that variability in the RF heating rate is mostly caused by the green moisture content. The proposed model might serve as a tool for estimating the heating time, which is needed for production planning.

KEYWORDS: Pasteurization of green wood, dielectric heating at radio frequencies, RF heating rate, neural networks, Monte Carlo method.

INTRODUCTION

Dielectric heating at radio frequencies (RF) has been studied and recommended as an alternative option to decontaminate green wood traded as logs and timber in order to avoid some disadvantages of the conventional method such as: Long heating time to reach the kill
temperature (hours instead of minutes for RF heating), uneven heating and material surface degrade— if the heating medium is not saturated. The treatment consists in achieving a minimum temperature of 56°C for a minimum duration of 30 continuous minutes throughout the entire profile of the wood (IPPC 2009). Most of the recently published studies deal with factors that affect the RF heating rate (Lazarescu and Avramidis 2011, Watanabe et al. 2011, Lazarescu and Avramidis 2012, Lazarescu et al. 2012, Huang et al. 2013), by analyzing different time-temperature schedules, treatment efficiency and temperature distribution (Lazarescu et al. 2009, 2011, Uzunovic et al. 2013).

Being a potential alternative method for phytosanitary treatment of green wood, it is important for the production planning to estimate the time needed to reach the lethal temperature for different load characteristics such as species, green moisture content, power density (heat generated per unit volume) and position of the board within the load (shell or core). In order to determine the time needed to reach the lethal temperature, the value of the RF heating rate must be known. To this purpose, a one-dimensional deterministic model was developed by Watanabe et al. (2011), to predict the heating rate in the RF pasteurization of green softwoods (pine and red cedar). The model can predict the RF heating rate of boards with various moisture contents under a known power density distribution. However, the model does not take into account the stochastic nature of the RF heating process of wood, where the heating rate is a random variable that can vary from one board to another and even throughout the same board, due to the variability in moisture content and power density. Therefore, the RF heating rate can be treated as a random variable described by a probability density function (PDF) that shows all possible values and their associated probability. In this case, the value of the RF heating rate can be estimated using the mean of PDF, which is the expected value of a random variable (Kozak et al. 2008). In addition to the estimated value of the RF heating rate, the unevenness degree of the heating was assessed using the standard deviation of the distribution.

In this study we hypothesized that the value of the heating rate during RF pasteurization of green wood can be predicted by means of an artificial neural networks (ANN) model that is solved using the Monte Carlo method (MCM). We chose neural networks due to their ability to reveal the nonlinear relationships between the independent and dependent variables of RF wood heating. Also, due to the fact that ANN was successfully applied alone or in combination with ε-regression Support Vector Machine (ε-rSVM) in the field of wood science, namely, to develop models for prediction of thermal conductivity (Avramidis and Iliadis 2005), dielectric loss factor (Avramidis et al. 2006, Iliadis et al. 2013a), drying rates (Wu and Avramidis 2006), mechanical properties of wood (bending strength and stiffness) (Mansfield et al. 2007, 2011), plywood bonding quality (Esteban et al. 2011), density (Iliadias et al. 2013b), final moisture content (Watanabe et al. 2013) and optimization of process parameters (Ozsahin 2013). On the other hand, the Monte Carlo method was used to take into account the stochastic nature of the RF heating of wood, considering that MCM was previously used to incorporate uncertainty and variability in models developed for wood science field (Kayihan 1985, 1993, Cassens et al. 1993, Cronin et al. 2003, Elustondo and Avramidis 2005). An exhaustive review on the application of the Monte Carlo method in wood engineering was elaborated by Taylor et al. (1995).

**MATERIAL AND METHODS**

**Experimental data**

In this study the dataset needed to develop and validate the model was obtained from a previous work done by Lazarescu et al. (2012). The work consisted in heating up twenty western red cedar (*Thuja plicata* Donn.) and twenty pine (*Pinus contorta* Douglas ex Loudon subsp. *
Contorta) boards, 40 x 90 mm in cross section and 2 m long from initial temperature to a target temperature of 60°C in a radio frequency vacuum dryer (RFV) that operates at 6.8 MHz. The temperature was monitored in six areas of interest that were located in the same vertical plane (Fig. 1). Two out of six sensors (FISO Technologies Inc., Quebec, Canada) were inserted in exterior areas (shell), namely #2 and #5. The other sensors were inserted in interior areas (core), namely, #1, #6, #7 and #8. In addition, the moisture content (MC) and power density (PD) for each area of interest were measured and calculated, respectively, according to the methodology described by Lazarescu et al. (2012). Each board was heated up at five different moisture contents and lengths.

The heating rate (HR) of each area of interest was calculated based on the temperature gradient (ΔT), which is equal to the difference between initial and target temperature, divided by the time (Δt) needed to achieve the target temperature of 60°C (Eq. 1)

\[
HR = \frac{\Delta T}{\Delta t} \quad (^\circ C \text{ min}^{-1})
\]  

The gathered dataset was divided randomly in three subsets. The first subset (500 values) was used to develop the ANN model; the second subset (100 values) to test the artificial neural network model and the third subset (260 values) to solve the ANN model using the MCM.

**ANN model development and validation**

A multilayer feedforward network (MLFF) structure was developed in this work by means of the NeuralWorks Predict Software (NeuralWare Inc., PA, USA) and the recommendations mentioned in the user guide of Predict (NeuralWare 2009). In this architecture, the neurons are arranged in layers, namely, input, hidden and output layer, in such a way that signals flow from one layer to another (Fig. 2) (Sablani 2006, Iliadis et al. 2013b). Each layer contains a number of artificial neurons or processing elements. According to Iliadis et al. (2013b), Ozsahin (2013), the task of one artificial neuron j is to receive input signals (xi) weighted by connection weights (wij) from all neurons located in the previous layer, to sum these weighted signals in order to produce the net input of neuron (netj), to add the bias (w0) to the netj, and to transmit the output value to the neurons from the next layer or presented as the output of network (Yj). The output value is computed by applying a mathematical (usually nonlinear) function, known as activation or transfer function (Palmer et al. 2006). The final result for a neuron is described by Eq. (2) (Iliadis et al. 2013b).
The number of neurons in the input and output layers was equal to the number of independent variables and dependent variables, respectively. The number of neurons in the hidden layer was found by trial and error approach. The standard back propagation algorithm was used for the network training. Once the network was trained, it was tested with the unseen dataset.

The performance of fifteen ANN models was analyzed using both Pearson’s correlation coefficient (R) and the coefficient of determination (R²) in order to figure out the optimal structure of the ANN model, respectively, which variables can be used in the input layer and the number of neurons in the hidden layers. The combination of input variables that generated the highest R and R² values was chosen. The chosen ANN structure was analyzed using the mean relative error (MRE) and the previously mentioned performance measures (R and R²) – Eqs. (3), (4) and (6) (Avramidis and Iliadis 2005, Wu and Avramidis 2006, Benli 2013, Watanabe et al. 2013). Moreover, the predicted values were plotted against the experimental ones.

\[
Y_j = w_0 + \sum_{j=1}^{n} x_j w_j
\]  
(2)

Pearson’s R correlation =
\[
\frac{\sum_{i=1}^{n} (p_i - \bar{p})(a_i - \bar{a})}{\sqrt{\sum_{i=1}^{n} (p_i - \bar{p})^2 \sum_{i=1}^{n} (a_i - \bar{a})^2}}
\]  
(3)

\[
R^2 = 1 - \frac{\text{SSE}}{\sum_{i=1}^{n} p_i^2}
\]  
(4)

\[
\text{SSE} = \sum_{i=1}^{n} (a_i - p_i)^2
\]  
(5)

\[
\text{MRE} = \frac{1}{n} \sum_{i=1}^{n} \frac{100 \cdot |a_i - p_i|}{a_i}
\]  
(6)

where:  
n - the number of data points,  
\(a_i\) - actual values of heating rate (°C.min⁻¹),  
\(p_i\) - predicted value of heating rate (°C.min⁻¹),  
\(\bar{a}\) - mean of experimental values (°C.min⁻¹),  
\(\bar{p}\) - mean of predicted values, (°C.min⁻¹),  
SSE – sum squared errors.

**Solving the ANN model using the Monte Carlo method**

The previously developed ANN model was solved using the direct Monte Carlo method (MCM). Generally speaking, the method comprises the following steps (Cronin and Gleeson 2006):

- selection of a deterministic model, which in our case was the ANN model;
- statistical analysis of the input random variables in order to figure out the mean, standard deviation and the distribution that best describe the data;
- generation of random values for the stochastic variables based on the statistical parameters of the fitted distributions;
- running the model for a large number of times till no significant difference is observed in the value of the output variable;
- storing the solution of each iteration;
- statistical analysis of the output variables using the mean and standard deviation.
For our model, we considered that the green moisture content (MC) and power density (PD) are two stochastic input variables (Fig. 2) because they both fluctuate randomly during RF heating from one board to another or even throughout the same board. The moisture content variability is due to the biologic origin of wood. On the other hand, the variability of the power density is due to both moisture content and temperature random dispersion inside the load (Zhao 2006). The use of direct MCM is conditioned by the requirement that the predictive variables are independent (Taylor et al. 1995, Cronin and Gleeson 2006). We assumed these variables as independent based on the variance inflation factor (VIF) that was lower than 10, namely, 1.51 for red cedar and 1.85 for pine. This let us conclude that multicollinearity will not cause problems in the model estimation (Chatterjee and Hadi 2006). The VIF was computed using Eq. (7).

$$VIF = \frac{1}{1 - R^2}$$ (7)

In order to generate the random values for MC and PD, the probability density function (PDF) for each stochastic input variable was figured out using the distribution fitting tool from the MATLAB Statistics ToolBox (MathWorks, Massachusetts, USA). The normal, log-normal, and Weibull distributions were the candidates both for green moisture content and power density. The chi-square goodness of the fit test was performed in MATLAB (MathWorks, Massachusetts, USA) in order to check if the selected PDF best describes the data. The random values were generated based on the parameters of the selected distribution using the random number generator tool from MATLAB Statistics ToolBox (MathWorks, Massachusetts, USA). The generated values that were not in the same range with the experimental data were removed. Thus, the range of moisture content was 8.29-136.2 (red cedar) and 4.84-96.82 % (pine), and the power density was between 41.46 – 274.49 (red cedar) and 47.17-231.33 kW.m⁻³ (pine). The model was executed till the change of the output stabilized in acceptable tolerance. The mean (M) and standard deviation (SD) were used to summarize the simulation results. In addition, a graphical comparison using frequency histograms was employed to compare simulated with experimental data.

A sensitivity analysis was conducted in order to see how the results might change due to the uncertainty in the input variables (MC and PD) and, also, to see if the behavior of the model makes sense when changing the input parameters (Datta and Rakesh 2009). The sensitivity coefficient (S) was calculated as the ratio of the change in the output variable to the corresponding change in the input variable (Cronin and Gleeson 2006). The sensitivity analysis was also performed to see the change in the output variable (HR) if the input probability distribution is assumed to be normal both for MC and PD instead of being the fitted distribution (Law and Kelton 1991). Also, we wanted to reveal which input variable generates the higher variation in the RF heating rate (Carro-Corrales et al. 2002). The last mentioned task was performed by running the ANN-MCM model in two trials. In the first trial the power density was kept constant at 125 (red cedar) and 108 kW.m⁻³ (pine) and therefore we assumed that variability exists only in the moisture content. The value of variance obtained in this case was reported to the value of variance when both independent variables were not constant. In the second trial, the value of the moisture content was kept constant at 42 (red cedar) and 45 % (pine).
RESULTS AND DISCUSSION

ANN model

The optimal structure of the multilayer feed forward network was the one that contained 3 neurons in the input layer (species, moisture content and power density), 4 neurons in the hidden layer, having a hyperbolic-tangent transfer function and 1 neuron in the output layer, with an exponential transfer function (Tab. 1 and Fig. 2).

An important finding was that the position of the area within the heating assembly does not affect the accuracy of the RF heating rate prediction. The model works better without this input variable (Tab. 1).

Tab. 1: Performance indicators for different architectures of the ANN model.

<table>
<thead>
<tr>
<th>Input variables</th>
<th>Number of neurons in the hidden layer</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>0.35</td>
</tr>
<tr>
<td>P</td>
<td>0</td>
<td>-0.06</td>
</tr>
<tr>
<td>MC</td>
<td>4</td>
<td>0.54</td>
</tr>
<tr>
<td>PD</td>
<td>0</td>
<td>0.26</td>
</tr>
<tr>
<td>S x P</td>
<td>2</td>
<td>0.33</td>
</tr>
<tr>
<td>S x MC</td>
<td>0</td>
<td>0.65</td>
</tr>
<tr>
<td>S x PD</td>
<td>7</td>
<td>0.47</td>
</tr>
<tr>
<td>P x MC</td>
<td>8</td>
<td>0.54</td>
</tr>
<tr>
<td>P x PD</td>
<td>0</td>
<td>0.20</td>
</tr>
<tr>
<td>MC x PD</td>
<td>6</td>
<td>0.54</td>
</tr>
<tr>
<td>S x P x MC</td>
<td>9</td>
<td>0.64</td>
</tr>
<tr>
<td>S x MC x PD</td>
<td>4</td>
<td>0.66</td>
</tr>
<tr>
<td>S x P x PD</td>
<td>8</td>
<td>0.47</td>
</tr>
<tr>
<td>P x MC x PD</td>
<td>2</td>
<td>0.54</td>
</tr>
<tr>
<td>S x P x MC x PD</td>
<td>2</td>
<td>0.65</td>
</tr>
</tbody>
</table>

S- species; P-position (shell or core); MC – moisture content; PD – power density.

The descriptive statistics of the chosen input variables is presented in Tab. 2 for red cedar and in Tab. 3 for pine. The performance of the ANN model was better in case of red cedar \((R=0.61)\) than in case of pine \((R=0.34)\) and much better than in the case of using the multiple linear regression (Figs. 3 and 4, Tab. 4). The mean, standard deviation, minimum and maximum values of the dataset used to test the ANN model were roughly within the same range with the ones for the training data set (Tabs. 2 and 3).

Tab. 2: Descriptive statistics of input and output variables used to train and test the ANN model – red cedar.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Training data set ((n=258))</th>
<th>Testing data set ((n=48))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>MC</td>
<td>44.88</td>
<td>26.88</td>
</tr>
<tr>
<td>PD</td>
<td>121.30</td>
<td>46.57</td>
</tr>
<tr>
<td>HR</td>
<td>3.61</td>
<td>1.89</td>
</tr>
</tbody>
</table>
Tab. 3: Descriptive statistics of input and output variables used to train and test the ANN model – pine.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Training data set (n=242)</th>
<th>Testing data set (n=52)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>MC</td>
<td>43.02</td>
<td>19.74</td>
</tr>
<tr>
<td>PD</td>
<td>106.44</td>
<td>38.57</td>
</tr>
<tr>
<td>HR</td>
<td>2.40</td>
<td>1.24</td>
</tr>
</tbody>
</table>

Fig. 3: Experimental data versus values predicted by the ANN model in case of red cedar.

Fig. 4: Experimental data versus values predicted by the ANN model in case of pine.

Tab. 4: Values of performance indicators used to compare the ANN model with the multiple linear regression.

<table>
<thead>
<tr>
<th>Species</th>
<th>ANN model</th>
<th>Multiple regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MRE %</td>
<td>R</td>
</tr>
<tr>
<td>Red cedar</td>
<td>37.15</td>
<td>0.61</td>
</tr>
<tr>
<td>Pine</td>
<td>66.94</td>
<td>0.34</td>
</tr>
</tbody>
</table>

ANN – MCM model simulation results

The fitted distributions of the green moisture content were in good accordance with the ones mentioned in other studies, namely, Weibull in case of red cedar and log-normal for pine (Kayihan 1993, Cronin et al. 2002, Elustondo and Avramidis 2002). The power density followed a log-normal distribution both for red cedar and pine. The mean, standard deviation, fitted distribution and parameters of distribution for the input variables are presented in Tab. 5.

Tab. 5: Statistical parameters that were used to generate random values for the independent variables.

<table>
<thead>
<tr>
<th></th>
<th>Red cedar</th>
<th>Pine</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MC</td>
<td>PD</td>
</tr>
<tr>
<td>Mean</td>
<td>41.72 %</td>
<td>124.92 kW.m⁻³</td>
</tr>
<tr>
<td>SD</td>
<td>22.66 %</td>
<td>44.69 kW.m⁻³</td>
</tr>
<tr>
<td>Fitted distribution</td>
<td>Weibull</td>
<td>Log-normal</td>
</tr>
<tr>
<td>Parameters of distribution</td>
<td>a = 47.26</td>
<td>μ = 4.76</td>
</tr>
<tr>
<td></td>
<td>b = 1.97</td>
<td>ζ = 0.37</td>
</tr>
</tbody>
</table>

a – scale factor, b – shape factor, μ - log location, ζ - log scale
The expected value of the heating rate predicted by the ANN-MCM model (using 10000 iterations) was 3.36 (for red cedar) and 2.23°C.min⁻¹ (for pine) (Tab. 6). This means that the simulated results will be by 13 (red cedar) and 8 (pine) lower than the experimental values, namely: 3.84 (for red cedar) and 2.42°C.min⁻¹ (for pine). Based on the predicted values, the time to reach the lethal temperature (56°C) for a temperature difference of 50°C between initial and target temperature is 14.88 min for red cedar instead of 13.02 and 22.42 instead of 20.66 min for pine. The error between the predicted and measured heating time can be considered reasonable even for RF heating, which is a fast heating method, considering that two or three minutes differences between experimental and predicted value don’t play an important role in the production planning. In addition, the about 2 minutes longer time estimated by the model can be overcome by a safety factor that is recommended to be used in order to alleviate the concern of under-treatment (Uzunovic et al. 2013).

The prediction of the standard deviation was less accurate than expected. The ANN-MCM model predicted a lower variation of the RF heating rate in-between the boards, and therefore, a much more uniform heating of the boards than in real life (Tab. 6, Figs. 5 and 6). The discrepancy between the actual and the predicted variation might be due to the low performance of the ANN deterministic model. However, the accurate prediction of the standard deviation is recognized as a difficult task, as also pointed out by Carro-Corrales et al. (2002).

Tab. 6: Mean and standard deviation values predicted by the model vs. measured values.

<table>
<thead>
<tr>
<th>RF heating rate, °C.min⁻¹</th>
<th>Probabilistic model</th>
<th>Experimental runs</th>
<th>Relative error %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Red cedar</td>
<td>3.36</td>
<td>1.00</td>
<td>3.84</td>
</tr>
<tr>
<td>Pine</td>
<td>2.23</td>
<td>0.35</td>
<td>2.42</td>
</tr>
</tbody>
</table>

Sensitivity analysis results

The results revealed that the value of the RF heating rate in case of red cedar was more sensitive to the moisture content than to the power density (Tab. 7). With pine, the heating rate was sensitive only to the moisture content: A change by 10 % of the power density value did not affect at all the value of the RF heating rate; increasing the change in PD from ±10 to ±50 %
caused non-significant change (by 2 %) in the output variable. Consequently, we concluded that the model is less sensitive to power density in the case of pine. This could be due to the fact that red cedar has better dielectric properties than pine (Lazarescu et al. 2012).

Tab. 7: Results of the sensitivity analysis.

<table>
<thead>
<tr>
<th>Input parameter</th>
<th>Species</th>
<th>Change in input variable (%)</th>
<th>Input PDF</th>
<th>Change in output variable (%)</th>
<th>Sensitivity coefficient (S), (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC, %</td>
<td>Red cedar</td>
<td>±10</td>
<td>Weibull</td>
<td>±4.5</td>
<td>± 0.45</td>
</tr>
<tr>
<td></td>
<td>Pine</td>
<td></td>
<td></td>
<td>±3</td>
<td>± 0.3</td>
</tr>
<tr>
<td>PD, kW.m⁻³</td>
<td>Red cedar</td>
<td>±10</td>
<td>Lognormal</td>
<td>±2</td>
<td>±0.2</td>
</tr>
<tr>
<td></td>
<td>Pine</td>
<td>±10</td>
<td>Lognormal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±50</td>
<td>Lognormal</td>
<td>±2</td>
<td>±0.04</td>
</tr>
<tr>
<td>PD, kW.m⁻³</td>
<td>Red cedar</td>
<td>-</td>
<td>Normal*</td>
<td>-2.40</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Pine</td>
<td>-</td>
<td>Normal**</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>MC, %</td>
<td>Pine</td>
<td>-</td>
<td>Normal ***</td>
<td>-2.76</td>
<td>-</td>
</tr>
<tr>
<td>PD, kW.m⁻³</td>
<td></td>
<td></td>
<td>Normal ****</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

* (M=41.72, SD=22.66), ** (M=124.92, SD=44.69), *** (M=45.16, SD=20.27), **** (M=107.75, SD=36.12)

The sensitivity analysis also revealed that when the PDF of the input variables was changed from the fitted distribution to an estimated normal distribution, the error of the model increased by about 2.5 %. In addition, the parametric study revealed that the moisture content had the greatest effect on the final variation of the RF heating rate. The moisture content accounted for about 90 % of the total variance in the RF heating rate, in case of both species. Variability in power density had a lower effect on the final dispersion of the heating rate (about 10 % of the total variance). Therefore, sorting green boards into different moisture content groups prior to pasteurization – similar to the methodology recommended by Elustondo et al. (2010) for wood drying – might be an option to solve the problem of unequal heating during RF pasteurization, in addition to a slower heating up to the lethal temperature (as mentioned by Uzunovic et al. 2013).

CONCLUSIONS

This work showed that the value of the RF heating rate during pasteurization of green softwoods can be reasonably predicted, namely, with a relative error of 13 for red cedar and 8 % for pine, by means of an ANN model that is solved using the Monte Carlo Method. The proposed modeling tool revealed that the variation in the RF heating rate is mostly caused by the moisture content. Therefore, sorting boards into different moisture content groups prior to pasteurization might represent a solution to minimize the degree of uneven heating. Although the use of the model is limited by the experimental conditions under which the dataset was gathered, it is obvious that a probabilistic model can give us more information needed both for production planning and to optimize dielectric (RF) heating, as a swift and environment-friendly method to decontaminate green wood.
ACKNOWLEDGMENT

This work was financially supported by a Strategic Grant from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES