IDENTIFICATION OF LOG CHARACTERISTICS IN COMPUTED TOMOGRAPHY IMAGES USING BACK-PROPAGATION NEURAL NETWORKS WITH THE RESILIENT BACK-PROPAGATION TRAINING ALGORITHM AND TEXTURAL ANALYSIS: PRELIMINARY RESULTS

*Qiang Wei**

PhD Candidate

Ying H. Chui†

Professor

Brigitte Leblon

Professor
Department of Forestry and Environmental Management
P.O. Box 44555
28 Dineen Drive
University of New Brunswick
Fredericton, NB E3B 6C2, Canada

Shu Y. Zhang†

Senior Scientist
FPInnovations–Forintek Division
2665 East Mall
Vancouver, BC V6T 1W5, Canada

(Received March 2008)

Abstract. This research addressed the feasibility of identifying internal log characteristics in computed tomography (CT) images of sugar maple and black spruce logs by means of back-propagation (BP) neural networks with a resilient BP training algorithm. Five CT images were randomly sampled from each log. Three of the images were used to develop the corresponding classifier, and the remaining two images were used for validation. The image features that were used in the classifier were gray-level values, textual, and distance features. The important part of the classifier topology, ie the hidden node number, was determined based on the performance indicators: overall accuracy, mean square error, training iteration number, and training time. For the training images, the classifiers produced class accuracies for heartwood, sapwood, bark, and knots of 99.3, 100, 96.7, and 97.9%, respectively, for the sugar maple log; and 99.7, 95.3, 98.4, and 93.2%, respectively, for the black spruce log. Overall accuracies were 98.5% for sugar maple and 96.6% for black spruce, respectively. High overall accuracies were also achieved with the validation images of both species. The results also suggest that using textural information as the inputs can improve the classification accuracy. Moreover, the resilient BP training algorithm made BP artificial neural networks converge faster compared with the steepest gradient descent with momentum algorithm. This study indicates that the developed BP neural networks may be applicable to identify the internal log characteristics in the CT images of sugar maple and black spruce logs.

Keywords: Black spruce, sugar maple, log characteristics, computed tomography (CT) images, artificial neural networks.

* Corresponding author: qiangwf@gmail.com
† SWST Member

Wood and Fiber Science, 40(4), 2008, pp. 620–633

INTRODUCTION

The properties and value of lumber depend on internal log characteristics such as sapwood, heartwood, and knots. Knowledge of internal log characteristics is thus useful in optimizing log breakdown strategies for extracting maximum value. Different sawing strategies can lead to large differences in lumber value recovery (Richards 1980) and lack of information on internal log characteristics impedes optimal sawing (Occeña 1991). Potential gains of approximately 10–15% in lumber value can be achieved based on log internal information (Richards 1977; Lemieux et al 2002). The computed tomography (CT) scan (Hounsfield 1980) is an emerging technique for acquiring information of internal log characteristics nondestructively. CT scanning provides cross-sectional images in planes. A CT image is composed of pixels, or picture elements, which are squares corresponding to the spatial resolution of the image. For each pixel, a brightness measured in terms of gray level (GL) values is acquired from the corresponding volume element (voxel) of the scanned section. The GL values describe the X-ray attenuation of the material. For wood, the attenuation is closely correlated to the wood density. Therefore, the produced CT images of a log reflect the density of the internal log. The scanner is set so that brighter regions (ie regions with higher pixel GL values) correspond to regions of denser material.

Besides optimizing log breakdown strategies, the information derived from CT images has also been used for wood drying and quality control (Pang and Wiberg 1998). To extract reliable internal log information from CT images, correct interpretation of the CT images is crucial. Grundberg (1994) proposed an algorithm that was mainly based on the high contrast between knots and heartwood to obtain the parametric description of the knot structure in pine logs. Zhu et al (1996) analyzed morphological image features such as the shape of knots based on geometric and statistical attributes. Rojas et al (2006) and Wei et al (2008) applied the maximum likelihood classifier to identify clear wood and knots in sugar maple logs. The main drawback of these works is that classification accuracy achieved by these methods is relatively low (70–80% overall accuracy).

Another method that is widely used for processing digital images is the artificial neural networks (ANN) approach (Sjoberg 1995). It was originally developed to model the way the brain performs particular tasks. ANNs can model general nonlinear functions. Among all the ANN types, the feed-forward back-propagation (BP) ANN is commonly used because it is effective for pattern-matching problems and is easy to use (Schmoldt et al 2000). Schmoldt et al (2000) applied this type of ANNs to identify clear wood, bark, and knots in oak, yellow poplar, and black cherry. Nordmark (2002) also used this ANN to identify clear wood and knots in Scots pine, achieving 85–95% overall classification accuracy. The image feature used in these ANN classifiers is GL values. However, to date, important spatial image information like the ones extracted from textural analysis as the input feature for the classifier has not yet been investigated thoroughly. Furthermore, the steepest gradient descent with momentum algorithm (Freeman and Skapura 1991) is widely used as the training algorithm of BP ANNs for identification of log properties. In practice, this algorithm is time-consuming. Riedmiller and Braun (1993) showed that a faster converging BP training algorithm is the resilient BP training algorithm. Such an algorithm has not yet been used in BP ANNs for identifying log characteristics.

The main objective of this study is to investigate the feasibility of feed-forward BP ANNs with the resilient BP training algorithm for identification of selected log characteristics in CT images acquired from sugar maple and black spruce logs. The input features used in the study were not only the raw GL values, but also the textural features computed by a textural analysis. This study is different than those of Schmoldt et al (2000) and Nordmark (2002) because it uses textural features (Table 1) as the inputs and a faster converging training algorithm for the BP ANNs. Moreover, the BP ANN's

Table 1. *Textural features applied in this study.*[a]

Textural feature	Formula[b]
Homogeneity	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{p_{ij}}{[1+(i-j)^2]}$
Contrast	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i-j)^2\times p_{ij}$
Dissimilarity	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(\lvert i-j\rvert\times p_{ij})$
Mean	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} i\times p_{ij}$
Standard deviation	$\sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(p_{ij}\times(i-\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} i\times p_{ij})^2)}$
Entropy	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(-p_{ij}\times\log_e p_{ij})$
Angular second moment	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p_{ij}{}^2$

[a] From Haralick et al (1973).
[b] P_{ij} is unnormalized counts that indicate: how many times are two neighboring pixels separated by a displacement (eg one pixel); one with a gray level (GL) value i and the other with a GL value j. N is the dimension of gray level co-occurrence matrix (GLCM).

performance in hardwoods (sugar maple) and softwoods (black spruce) was compared.

A comparison between BP ANNs with and without textural information was also presented in this study. Meanwhile, the resilient BP training algorithm was compared with the traditional training algorithm, ie the steepest gradient descent with momentum algorithm.

MATERIALS AND METHODS

Materials

In this study, BP ANN classifiers were developed for two species, including one hardwood, sugar maple (*Acer saccharum*), and one softwood, black spruce (*Picea mariana*). The sugar maple sample tree was collected from a natural stand nearby Fredericton, New Brunswick. The tree was 32-yr old with a 19-cm diameter at breast height (DBH) and a height of 12.5 m. The black spruce tree was collected from an initial spacing trial nearby Thunder Bay, Ontario. The sample tree was 48-yr old with a 17.1-cm DBH and a height of 16.8 m. The butt logs from these trees were collected for this study. Both logs were scanned by a Siemens Somatom Plus 4 Volume Zoom CT scanner at the Institut National de la Recherche Scientifique in Quebec City. The scanning conditions were those recommended by Hou et al (2005): 140 kV and 178 mA; slice plane, 10-mm thick; exposure time, 1000 s; and room temperature (approximately 20°C). Refer to Bucur (2003) for more details about CT scanning.

For each log, five CT images were randomly selected from the log, and a corresponding BP ANN classifier was developed using three of the images. They are referenced hereafter as training images. The remaining two images were used for validation, and are referenced hereafter as validation images. The training images and validation images were sampled from different parts of the log. Each image had an eight-bit radiometric resolution and a size of 512 pixels × 512 lines. The internal log characteristics for both species that need to be identified include sapwood, heartwood, and knots. Inner bark of the sugar maple (this log was debarked; therefore, only inner bark remained) and bark of the black spruce log were also considered.

Removal of Image Background

Each raw CT image has a background that represents the air surrounding the log. For the sugar maple images, the GL value of the background was <45 and any such pixels were removed. For the black spruce images, most heartwood pixels also had GL values <45. Euclidean distance between heartwood and pith of the log cross-section is always within a certain range. Therefore, to avoid flagging heartwood pixels as background, in addition to the GL 45 threshold, another threshold, the Euclidean distance of 150, was also used. The pith of the log cross-section was automatically determined using the method proposed by Bhandarkar et al (1999). Any pixels in the black spruce CT images having a GL value <45 and a Euclidean distance between the pixel of interest and the pith of the log cross-section >150 were removed.

Input Feature Selection

The image features that are extracted from CT images and used as the inputs for the classifier play an important role in the classification accuracy. Following Wei et al (2008), nine image features were selected as the input features, including 1) the GL value for each pixel of interest; 2) seven textural features; and 3) the Euclidean distance between the pixel of interest and the pith of the log cross-section (Table 1). The textural features were computed by the method of Haralick et al (1973). It is based on a gray-level co-occurrence matrix within a rectangular moving window. Each pixel of interest was located at the center of a moving window. In this study, the size of the moving window is 5 × 5 pixels. The textural analysis of each image was performed using the TEX program of PCI Geomatica (PCI Geomatica Inc).

Artificial Neural Network Classifier

The development of BP ANN classifiers includes three major steps: defining the training data set, the validation data set, and choosing the training algorithm; selecting the classifier's topology; and training the classifiers. The ANN classifiers were developed using Matlab software.

Defining the training, validation data sets, and selection of the training algorithm. For each classifier, the training data consisted of a set of 1200 vectors. For each of the three training images, an automatically random sample was created with 100 vectors for each of the four log characteristics. A validation set consisted of 800 vectors that were randomly sampled from the two validation images. Each vector consisted of nine components corresponding to the nine input features.

The resilient BP training algorithm was selected as the training method. It can make ANNs converge faster compared with conventional steepest gradient descent with momentum algorithms (Freeman and Skapura 1991; Riedmiller and Braun 1993). Refer to Riedmiller and Braun (1993) for more details about the training procedure using the resilient BP training algorithm.

Classifier's topology selection. As shown in Fig 1, an ANN generally consists of one input layer, one or more hidden layers, and one output layer. Each layer contains a given number of nodes, which are the fundamental processing elements of ANNs. The number of layers and nodes in

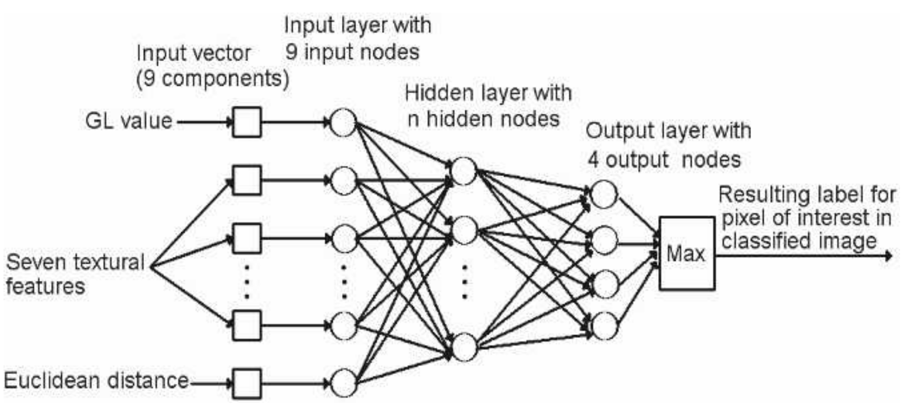

Figure 1. Topology of the feed-forward back-propagation artificial neural networks used in this study (n is the optimal number of hidden nodes).

each layer defines the ANN classifier's topology. The topology has significant influence on the convergence speed and also on the classification accuracy (Nordmark 2002). In this study, the network topology (Fig 1) consists of one input layer with nine input nodes corresponding to the nine input features, and one output layer with four output nodes corresponding to the four log characteristics to be identified. With respect to the number of hidden layers, according to Mas and Flores (2008), one hidden layer with an appropriate number of hidden nodes can produce good classification. Log-sigmoid function was chosen as the transfer function. The learning rate was fixed at 0.01.

It is important to choose a correct number of hidden nodes for the hidden layer. Insufficient hidden nodes will cause the ANN to be unable to learn sufficient information from the training data set for classification. On the other hand, too many hidden nodes are likely to cause the ANN to overfit the training data (Rosin and Fierens 1995). In this study, for each species, 25 different numbers of hidden nodes were tested for the corresponding classifier, from 11 hidden nodes to a maximum of 35 nodes. The hidden node number was selected for each classifier using the following performance indicators: 1) producer's overall accuracy defined as the number of correctly classified pixels divided by the total number of classified pixels (Congalton 1991); it will be called hereafter overall accuracy; 2) mean square error (MSE) defined by Freeman and Skapura (1991):

$$\mathrm{MSE} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(d_{ij}-y_{ij})^2}{m \times n}$$

where d_{ij} is the desired output for class j corresponding to the ith input vector, and y_{ij} is the actual output produced by the classifier for class j corresponding to the same input vector; there are m input vectors and n classes; and 3) the number of training iterations and training time. The selection of the hidden node number involves sixfold crossvalidation (Li et al 1996) as follows: first, the training set was randomly divided into six groups of 200 vectors each and they were fixed. For a given number of tested hidden nodes, at each stage of the process, one of the six groups was reserved as the testing group. The ANN classifier with the corresponding number of hidden nodes was trained on the remaining five groups (the training group). The iteration numbers and the training times were recorded. The trained classifier was applied for classification on the reserved testing group and also on the entire validation set. Overall accuracies and MSEs were computed separately for the testing group and validation set. This process was repeated six times (the testing group and training group were changed accordingly each time). Then an average value for each performance indicator was computed for the tested hidden node number. This procedure was applied to all tested hidden node numbers examined one case at a time. The hidden node number was then selected based on the following rule: choosing the smallest number of hidden nodes that can still provide high overall accuracy and low MSE for both the testing group and the validation set, as well as a small number of training iterations, and a short training time for the training group.

Two major concerned classifiers were then developed in this study, one for each wood species. They are referenced thereafter as the sugar maple classifier and black spruce classifier, respectively. For both the classifiers, textural information was also used as the inputs, and the resilient BP training algorithm was chosen as the training method. The major difference between the two classifiers is that they were developed with different species data.

Classification, Postclassification Procedure, and Accuracy Analysis

Each trained classifier with the selected topology was applied to identify log characteristics in the corresponding log. For each pixel, the nine input feature values of the pixel were input into the classifier to compute the outputs (four outputs in total). The pixel was classified into the class for which the corresponding output node produced a higher output than the other three output nodes did for the pixel. There were occasionally isolated pixels remaining in the image after classification. A 5 × 5 pixels median filter was then used to remove these pixels and defragment the classified image. Areas of each log characteristic in the CT images were also manually delineated with the PCI Geomatica software. They produced the reference images for the corresponding classified images and contain the information of the true class belonging to each pixel. Two types of variables were computed to quantify the classification accuracy for both the training images and the validation images (Congalton 1991): 1) producer's class accuracy. For one class i, it is defined as number of pixels labeled as class i in both the reference image and the classified image divided by total number of pixels of class i in the reference image. It assesses the classification accuracy for each class. It will be called hereafter class accuracy; 2) producer's overall accuracy as defined previously. It is a weighted average classification accuracy for all the classes together, hereafter called overall accuracy. Moreover, for the validation images, the number of false-positive and false-negative pixels was also computed to assess the classification accuracy. For one class i, the amount of false-positive pixels refers to the amount of pixels classified to class i, that do not truly belong to class i; the amount of false-negative pixels refers to the amount of pixels, which actually belong to class i, and are classified to other classes. Good classification performance requires that both the amount of false-positive pixels and false-negative pixels should be low.

The accuracy analysis was undertaken to mainly answer the following questions: 1) For each species, which log characteristics are easily separated from the others? and 2) Did the feed-forward BP ANN produce the same classification accuracy for both species? Sixfold crossvalidation was also applied to the analysis. Each trained classifier produced six estimates of the overall accuracy and six estimates of the class accuracy of each log characteristic for the validation set (Table 2). These estimates were then used as samples in the following statistical analyses, which were performed using Minitab software (Minitab Inc). A one-way analysis of variance (ANOVA) was performed to determine whether there were significant differences among the class accuracies in each species (p values of ANOVA of less than 0.05 [α-level]);

Table 2. *Overall and class accuracies of each BP ANN classifier*[a] *in identifying the four log characteristics in the computed tomography images for the validation set.*

			Class accuracy			
Log	Group (no.)	Overall accuracy (%)	Heartwood (%)	Sapwood (%)	Bark/inner bark[b] (%)	Knots (%)
Sugar maple	1	91.4	100	95.5	88.0	82.0
	2	87.0	100	100	80.5	67.5
	3	91.8	100	100	85.0	82.0
	4	90.5	100	100	84.0	78.0
	5	86.6	100	95.0	83.5	68.0
	6	90.0	100	100	86.5	73.5
	Average	89.5	100	98.4	84.6	75.2
Black spruce	1	91.8	100	77.0	97.5	92.5
	2	92.0	99.5	78.5	97.0	93.0
	3	91.6	99.5	79.5	98.0	89.5
	4	93.3	100	82.0	99.0	92.0
	5	92.0	99.5	80.5	95.5	92.5
	6	94.0	99.5	86.5	98.5	91.5
	Average	92.4	99.7	80.7	97.6	91.8

[a] The inputs for the classifiers include the textural information and the classifiers were trained with the resilient BP training algorithm.
[b] Bark for the black spruce log and inner bark for the sugar maple log.
BP, back-propagation; ANN, artificial neural network.

that is, whether there are some log characteristics separated easily from the others. ANOVA used the 24 estimates of class accuracies (six estimates for each of the four log characteristics) as the statistical samples. If the *p* value of ANOVA is less than 0.05, a Tukey's test (Ghosh and Sharma 1963) was then applied to determine which class accuracies were significantly different from one another; that is, determining which specific log characteristics are separated easily from the others. This test generated several sets of multiple comparison confidence intervals that represent ranges of values derived from sample statistics that are likely to contain the value of an unknown population parameter. There is no significant difference between the class accuracies for each pair of the log characteristics if the confidence interval for the subtraction between the class accuracies of the two characteristics includes 0. A two-sample t-test (Schmoldt et al 2000) was also performed to test whether there was a difference in classification accuracies between the sugar maple classifier and the black spruce classifier. The analysis was performed over six overall accuracy estimates of the sugar maple and of the black spruce (Table 2). There is a distinct difference in the overall accuracy between the two species if the *p* value of the t-test is less than 0.05 (α-level).

RESULTS AND DISCUSSION

Selection of the Hidden Node Number and Image Classification

The box plots representing the four performance indicators are presented as a function of the number of hidden nodes in Fig 2, for the sugar maple classifier, and in Fig 3 for the black spruce classifier. For the sugar maple classifier, 26 hidden nodes produced good performance in-

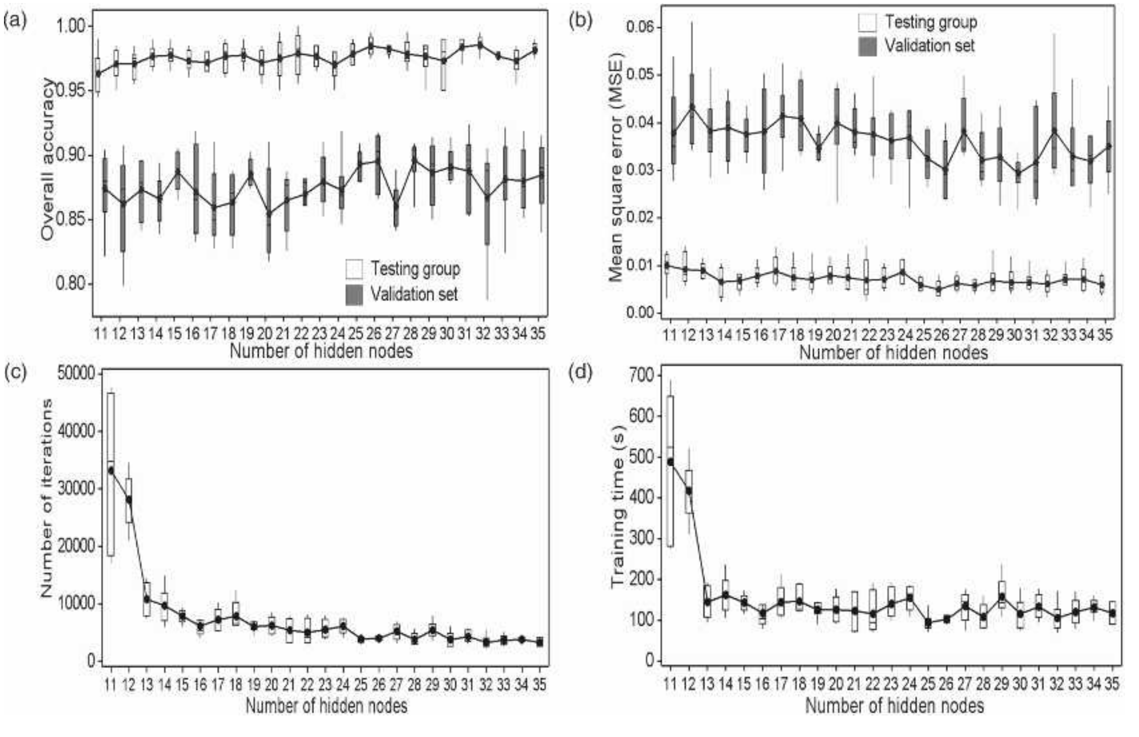

Figure 2. Performance indicators for selecting the hidden node number of the sugar maple classifier (each black dot within a box indicates an average value for a performance indicator produced by the classifier with the corresponding tested number of hidden nodes): (a) overall accuracy for the testing group and validation set; (b) mean square error for the testing group and validation set; (c) training iteration number; and (d) training time.

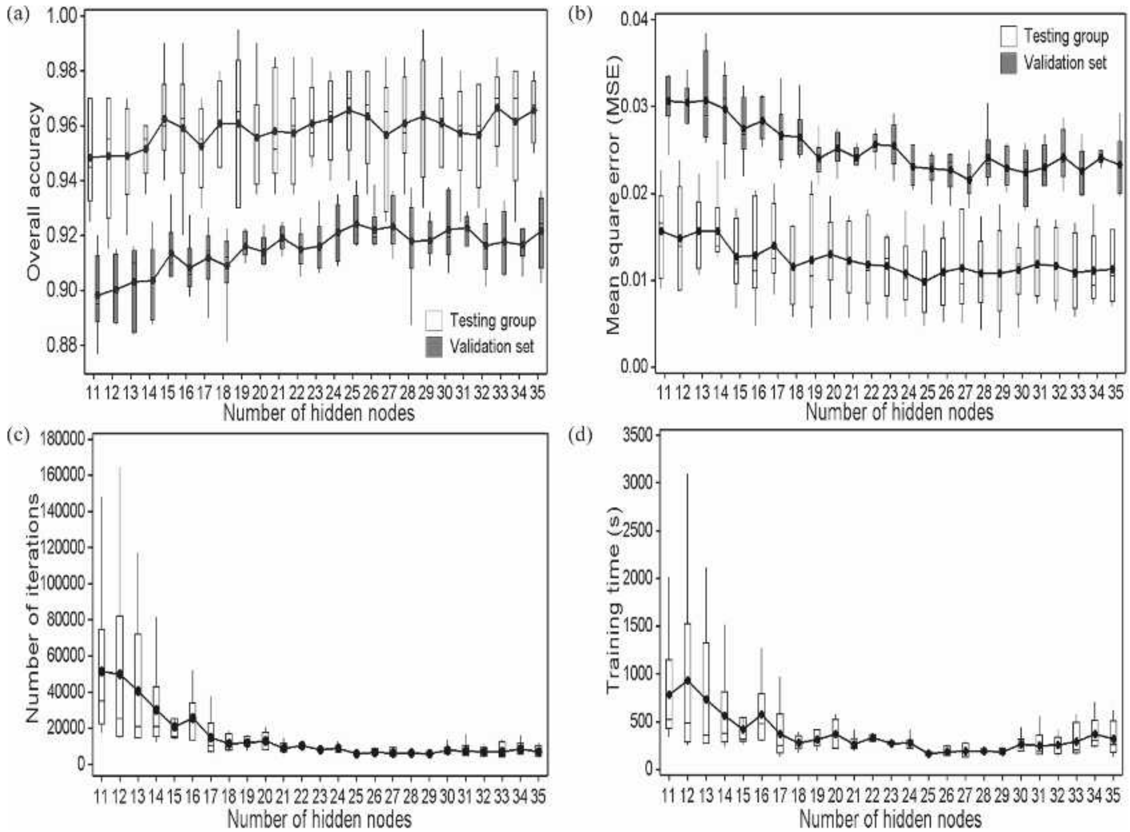

Figure 3. Performance indicators for selecting the hidden node number of the black spruce classifier (each black dot within a box indicates an average value for a performance indicator produced by the classifier with the corresponding tested number of hidden nodes): (a) overall accuracy for the testing group and validation set; (b) mean square error for the testing group and validation set; (c) training iteration number; and (d) training time.

dicators, although it did not produce the best one for all indicators (Fig 2). With 26 hidden nodes, the mean overall accuracies for the testing group and for the validation set were 98.5 and 89.5%, respectively. The mean values for MSE were 0.005 and 0.030, respectively. The training time was 101.5 s, and the training iteration number was 4011. For the black spruce classifier, 25 hidden nodes usually produced the best performance indicators (Fig 3). In this case, the mean overall accuracies for the testing group and for the validation set were 96.6 and 92.4%, respectively. The mean values for MSE were 0.010 and 0.023, respectively. The training time was 163.4 s and the training iteration number was 5831. Therefore, 26 and 25 hidden nodes were used for the sugar maple classifier and the black spruce classifier, respectively, in the subsequent analyses. An example of classified CT images is given in Fig 4b for the sugar maple classifier and Fig 5b for the black spruce classifier. Both classified images were then filtered by a 5 × 5 pixels median filter (Figs 4c and 5c).

Accuracy Analysis

For the training images of the sugar maple log, the overall accuracy was 98.5%; and class accuracies for sapwood, heartwood, inner bark, and knots were 100, 99.3, 96.7, and 97.9%, respectively. For the validation images, the overall accuracy was 89.5%, and class accuracies were 98.4, 100, 84.6, and 75.2%, respectively. For the

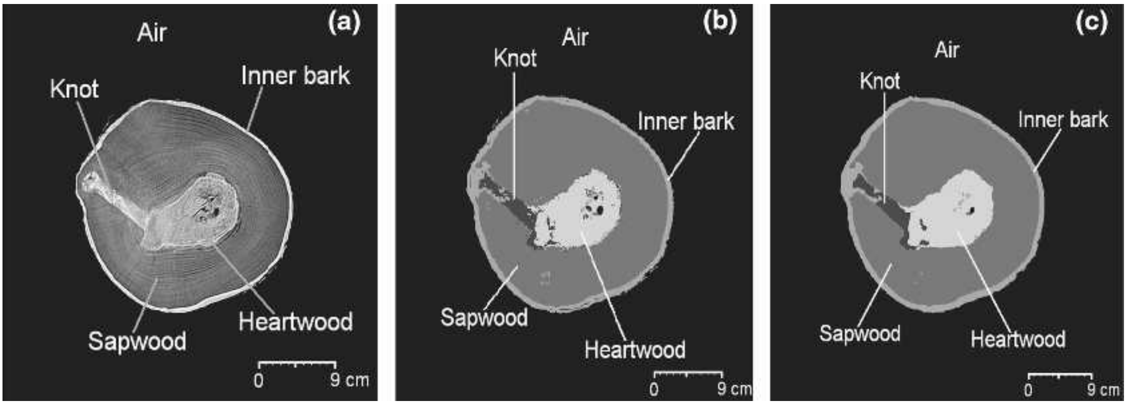

Figure 4. Example of cross-section computed tomography (CT) images for the sugar maple log: (a) raw CT image; (b) classified image by the sugar maple classifier; and (c) classified image filtered using a 5 × 5 pixels median filter.

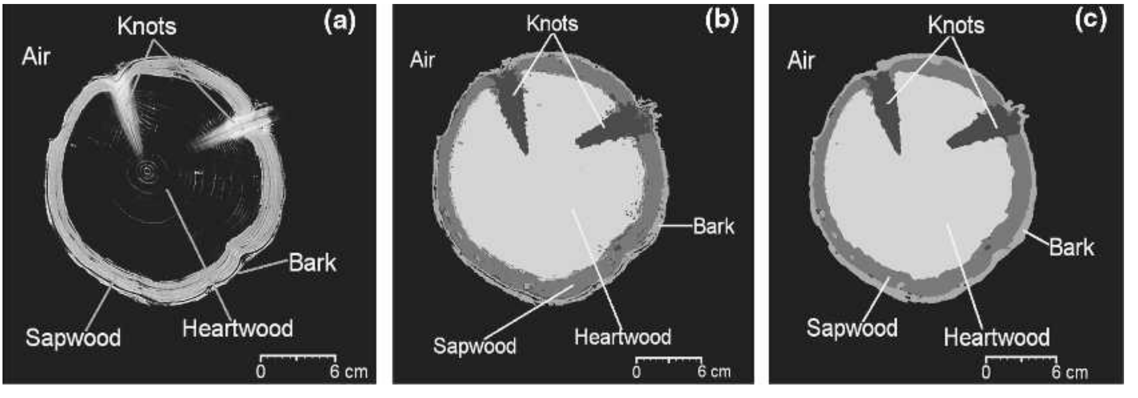

Figure 5. Example of cross-section computed tomography (CT) images for the black spruce log: (a) raw CT image; (b) classified image by the black spruce classifier; and (c) classified image filtered using a 5 × 5 pixels median filter.

training images of the black spruce log, the overall accuracy was 96.6%; and class accuracies for sapwood, heartwood, inner bark, and knots were 95.3, 99.7, 98.4, and 93.2%, respectively. For the validation images of the black spruce log, the overall accuracy was 92.4%; and class accuracies were 80.7, 99.7, 97.6, and 91.8%, respectively. As expected, better classification was achieved on the training images because the classifiers were trained using these images. For the validation images, the classifiers produced low class accuracy for knots in sugar maple and sapwood in black spruce. This indicates that the training vectors may be not sufficient to contain enough possible variations in log CT images. In this study, only three CT images of each species were used to train the classifier. More log CT images are needed for both training and validation in future work to develop robust classifiers. For the validation sets, the amount of false-positive and false-negative pixels are given in Table 3. Overall, this study shows that the BP ANN classifiers are able to identify the considered internal log characteristics fairly accurately. The classification accuracy achieved in this study is as good as that of Schmoldt et al (2000) (the overall accuracy of 85–95% for oak, yellow poplar, and black cherry) and Nordmark (2002) (the overall accuracy of 93–95% for Scots pine).

For the sugar maple log, the ANOVA analysis produced the p value of less than 0.05 (α-level) indicating there were significant differences among the class accuracies. Tukey's test was

Table 3. *The amount of false-positive pixels and false-negative pixels for the validation set.*

(a) Sugar maple

		Amount of false-positive pixels and false-negative pixels							
		Heartwood		Sapwood		Inner bark		Knots	
Classifier	Group (no.)	A[a]	B[a]	A	B	A	B	A	B
Classifier with textural information (trained with the resilient BP training algorithm)	1	0	0	9	0	24	36	36	33
	2	0	3	0	0	39	62	65	39
	3	0	1	0	0	30	35	36	30
	4	0	4	0	2	32	38	44	32
	5	0	14	10	0	33	60	64	33
	6	0	5	0	0	27	48	53	27
	Average	0	5	3	0	31	47	50	32
Classifier without textural information (trained with the resilient BP training algorithm)	1	0	8	0	0	53	42	50	53
	2	0	12	0	11	76	44	65	74
	3	0	0	0	10	74	54	64	74
	4	0	5	1	14	71	34	52	71
	5	0	13	0	12	96	35	57	93
	6	0	5	0	0	74	36	41	74
	Average	0	7	0	9	74	41	55	73
Classifier with textural information (trained with the steepest gradient descent with momentum algorithms)	1	0	2	0	0	88	46	48	88
	2	0	2	0	0	77	56	58	77
	3	0	3	0	0	95	37	40	95
	4	0	2	0	0	92	45	47	92
	5	0	5	0	0	90	49	54	90
	6	0	2	0	0	80	45	47	80
	Average	0	3	0	0	87	46	49	87

(b) Black spruce

		Amount of false-positive pixels and false-negative pixels							
		Heartwood		Sapwood		Bark		Knots	
Classifier	Group (no.)	A	B	A	B	A	B	A	B
Classifier with textural information (trained with the resilient BP training algorithm)	1	0	1	46	9	5	10	15	46
	2	1	2	43	11	6	8	14	43
	3	1	0	41	17	4	8	21	42
	4	0	2	36	6	2	10	16	36
	5	1	2	39	14	9	9	15	39
	6	1	2	27	8	3	10	17	28
	Average	1	2	39	11	5	9	16	39
Classifier without textural information (trained with the resilient BP training algorithm)	1	3	6	36	46	25	47	53	18
	2	1	2	43	11	6	8	14	43
	3	3	6	64	38	24	46	46	47
	4	2	5	47	49	26	44	52	29
	5	3	4	29	44	21	45	52	12
	6	4	7	48	39	16	46	53	29
	Average	3	5	45	38	20	39	45	30
Classifier with textural information (trained with the steepest gradient descent with momentum algorithms)	1	0	5	16	23	4	8	33	17
	2	0	4	15	16	4	18	30	11
	3	0	9	19	19	3	18	38	14
	4	0	8	13	21	4	18	40	9
	5	1	3	15	18	4	22	34	11
	6	0	6	20	19	4	20	36	15
	Average	0	6	16	19	4	17	35	13

[a] A represents the amount of false-negative pixels; B represents the amount of false-positive pixels.
BP, back-propagation.

then performed (Table 4a). There was a statistical difference between class accuracies of each pair of the log characteristics, except for the sapwood–heartwood pair. Class accuracies for both sapwood and heartwood are higher than that for knots and inner bark in sugar maple (Table 2). That means sapwood and heartwood of the studied sugar maple log are easier to identify compared with inner bark and knots. This is mainly attributable to GL values of the log characteristics that are directly related to their physical properties, including density and MC. Previous studies (Lamb and Marden 1970; Einspahr and Harder 1975; Rojas et al 2005) have shown that the sapwood in sugar maple generally has a lower density than knots and heartwood. The GL value of a characteristic is proportional to its density. Therefore, sapwood can be separated easily from knots and the other characteristics because of its lower GL values. Knots always have the highest densities resulting in the highest GL values. As a result of a higher MC, some parts of the inner bark had higher GL values (Fig 4a) that are almost comparable with those of knots. When knots are close to these parts of the inner bark, it becomes difficult to distinguish the knots from the inner bark because they have similar GL values and Euclidean distances. As a result, incorrect classification may happen. Compared with the inner bark, the heartwood had lower GL values and Euclidean distances. This also makes heartwood easy to identify from the other characteristics.

For the black spruce log, the p value of the ANOVA was also less than 0.05 (α-level) indicating that there were significant differences among the class accuracies. Tukey's test results show that there was a statistical difference between class accuracies of each pair of the log characteristics, except for the heartwood–bark pair (Table 3b). The class accuracies for heartwood and bark were higher than that for sapwood (Table 2). All these suggest the heartwood and the bark of the studied black spruce log are easier to identify as compared with sapwood. According to Panshin and Zeeuw (1971), the heartwood of a black spruce log generally has lower density than knots, sapwood, and bark. Moreover, average MC of heartwood and sapwood in black spruce (green wood) is estimated to be approximately 52 and 113%, respectively (Forest Products Laboratory 1999). These cause the darkness of the heartwood in CT images (Fig 5a), therefore making the heartwood easily separable from the other characteristics. Bark is located far from the pith as compared with sapwood and heartwood. That means that the bark pixels have higher Euclidean distances. Therefore, the bark is also separated easily from the other characteristics. Although the sapwood was rather easily separated from other characteristics of the sugar maple log, this log characteristic was difficult to distinguish from the other characteristics in the black spruce log. Carefully checking of the sapwood pixels in the black spruce CT images shows a distinct variation in

Table 4. *Confidence interval for the difference between the class accuracies of each pair of the log characteristics as estimated by a Tukey's test.*

(a) Sugar maple

	Heartwood	Sapwood	Inner bark	Knots
Heartwood	—			
Sapwood	−0.044, 0.076	—		
Inner bark	−0.214, −0.094	−0.199, −0.078	—	
Knots	−0.309, −0.188	−0.293, −0.172	−0.154, −0.034	—

(b) Black spruce

	Heartwood	Sapwood	Bark	Knots
Heartwood	—			
Sapwood	0.160, 0.221	—		
Bark	−0.051, 0.010	0.139, 0.200	—	
Knots	−0.109, −0.048	0.081, 0.142	−0.088, −0.027	—

GL values throughout the sapwood (some sapwood pixels were very dark compared with others). This variation, which may be related to the MC influence, causes a high class spectral variability and thus a low class accuracy for the sapwood.

For the t-test comparing classification accuracies of both classifiers, the *p* value of the t-test was 0.02. It was less than 0.05 (α-level), indicating that there was a difference in the classification accuracy between the two classifiers. This suggests that each BP ANN developed in this study provides a different performance for the related species. Higher overall accuracy (92.4% for the validation images) was achieved with the black spruce classifier compared with the sugar maple classifier (89.5%). This was mainly caused by the low class accuracy of the knots' region (75.2%) in the validation images of the sugar maple log.

Other Comparisons

Comparison between back-propagation artificial neural networks with and without textural information. In this study, for each species, a corresponding BP ANN classifier without textural information was also developed using the same three training images and validated with the same two validation images. They are referenced thereafter as sugar maple classifier without textures and black spruce classifier without textures, respectively. Comparing with BP ANN classifiers with textural information (ie sugar maple classifier and black spruce classifier), the inputs for the BP ANN classifiers without textural information did not include the seven textural features (Table 1). For the training images of the sugar maple log, sugar maple classifier without textures produced an overall accuracy of 93.4%, which is less than the overall accuracy of 98.5% produced by sugar maple classifier. Class accuracies for sapwood, heartwood, inner bark, and knots were 98.7, 99.6, 86.3, and 88.6%, respectively. For the validation images, the overall accuracy was 83.9%, which is also less than the overall accuracy of 89.5% produced by sugar maple classifier; class accuracies were 99.8, 100, 63.0, and 72.6%, respectively. For the training images of the black spruce log, black spruce classifier without textures produced the overall accuracy of 93.4%, which is less than the overall accuracy of 96.6% produced by black spruce classifier. Class accuracies for sapwood, heartwood, bark, and knots were 94.2, 98.7, 93.7, and 88.0%, respectively. For the validation images, the overall accuracy was 86.0%, which is also less than the overall accuracy of 92.4% produced by black spruce classifier; class accuracies were 77.8, 98.7, 90.2, and 77.5%, respectively. The amounts of false-positive pixels and negative pixels are given in Table 4. Taking classification on both sugar maple images and black spruce images into consideration, the results suggest that the classification performance of BP ANNs with textural information is better than that of BP ANNs without textural information.

Comparison between the resilient back-propagation training algorithm and the steepest gradient descent with momentum training algorithm. For each species, a corresponding BP ANN classifier trained with the traditional training algorithm, ie the steepest gradient descent with momentum algorithm, was also developed and validated using the same five images. They are referenced thereafter as sugar maple classifier with traditional algorithm and black spruce classifier with traditional algorithm, respectively. As shown in Table 5, the resilient BP training algorithm made BP ANN classifiers converge faster compared with the steepest gradient descent with momentum algorithm indicating the superiority of the resilient BP training algorithm. For the training images of the sugar maple log, sugar maple classifier with traditional algorithm produced the overall accuracy of 94.8%, which is less than the overall accuracy of 98.5% produced by sugar maple classifier. Class accuracies for sapwood, heartwood, inner bark, and knots were 99.2, 99.6, 84.9, and 95.5%, respectively. For the validation images, the overall accuracy was 83.0%, which is also less than the overall accuracy of 89.5% produced by sugar

Table 5. *Training time of each developed BP ANN classifier (with textural information) using different training algorithms.*

Training algorithm	Group (no.)	Training time (s)	
		Sugar maple	black spruce
The resilient BP training algorithm	1	89.1	212.7
	2	112.6	139.3
	3	106.9	152.1
	4	95.4	170.3
	5	110.0	176.0
	6	95.2	130.0
	Average	101.5	163.4
The steepest gradient descent with momentum algorithm	1	724.2	539.4
	2	671.3	605.7
	3	907.2	576.0
	4	674.2	631.0
	5	551.1	589.1
	6	649.1	532.0
	Average	696.2	578.9

BP, back-propagation; ANN, artificial neural network.

maple classifier; class accuracies were 100, 100, 56.5, and 75.5%, respectively. For the training images of the black spruce log, black spruce classifier with traditional algorithm produced the overall accuracy of 95.2%, which is less than the overall accuracy of 96.6% produced by black spruce classifier. Class accuracies for sapwood, heartwood, bark, and knots were 95.4, 100, 97.4, and 88.2%, respectively. For the validation images, the overall accuracy was 93.1%, which is greater than the overall accuracy of 92.4% produced by black spruce classifier; class accuracies were 91.8, 99.9, 98.4, and 82.4%, respectively. The amount of false-positive pixels and negative pixels are listed in Table 4.

CONCLUSIONS

This study mainly focused on developing two feed-forward BP ANN classifiers with a resilient BP training algorithm to identify four internal log characteristics in sugar maple and black spruce, respectively. GL values, textural, and distance features were selected as input features for each classifier. The optimum number of hidden nodes was 26 for the sugar maple classifier and 25 for the black spruce classifier. Both classifiers produced fairly accurate classification for the corresponding species. Statistical analyses show that higher classification accuracy was achieved with the black spruce classifier. The proposed feed-forward BP ANN classifiers (using the selected input features) may be feasible to identify log characteristics in sugar maple CT images and black spruce CT images, respectively. The results also suggest that the classification performance of BP ANNs with textural information is better than that of BP ANNs without textural information; the resilient BP training algorithm made BP ANN classifiers converge faster compared with the steepest gradient descent with momentum algorithm. One drawback of BP ANNs is that the presented procedure of choosing the hidden node number is time-consuming. The classifiers were developed on a single log for each species. There is the need to test the classifiers with a high number of sugar maple and black spruce logs. Moreover, the classified images were only filtered by a median filter. A more advanced post-classification procedure (eg merging classified pixels to 3-D logs) should be developed in future work to improve classification accuracy.

ACKNOWLEDGMENTS

We thank the Canada Foundation for Innovation for supporting the acquisition of the CT scanner used in this study, and FPInnovations–Forintek Division, NSERC Discovery Grant programs and NBIF for providing the funding toward research stipends and costs.

REFERENCES

Bhandarkar SB, Faust TD, Tang MJ (1999) Catalog: A system for detection and rendering of internal log defects using computer tomography. Mach Vis Appl 11(4):171–190.

Bucur V (2003) Nondestructive characterization and imaging of wood. Springer Series in Wood Science. Springer-Verlag, Berlin, Germany. 337 pp.

Congalton R (1991) A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens Environ 37(1):35–46.

Einspahr DW, Harder M (1975) Hardwood bark properties important to the manufacture of fiber products. IPC Techn Pap Ser No. 11:11–28.

Forest Products Laboratory (1999) Wood handbook. University Press of the Pacific, Madison, WI. 464 pp.

Freeman JA, Skapura DM (1991) Neural networks algorithms, applications and programming techniques. Addison-Wesley Publishing Company. 401 pp.

Ghosh MN, Sharma D (1963) Power of Tukey's test for non-additivity. JR Stat Soc [Ser A] 25(1):213–219.

Grundberg S (1994) Scanning for internal defects in logs. Licentiate Thesis, Luleå University of Technology. 121 pp.

Haralick RM, Shanmugam K, Dinstein I (1973) Texture characteristics for image classification. IEEE T Syst Man Cyb 3(6):610–621.

Hou ZQ, Zhang SY, Hu LJ (2005) Determining wood density and its 3-D distribution in a log using computed tomography scanning technique. Pp 427–423 *in* Proc 14th International Symposium on Non-destructive Testing of Wood, May 2–4, 2005, Shaker Verlag, Germany.

Hounsfield GN (1980) Computed medical imaging. Forest Sci 210(4465):22–28.

Lamb FM, Marden RM (1970) Variation in density of sugar maple sapwood and heartwood. USDA For Prod Lab Res Note NC-90.

Lemieux H, Beaudoin M, Zhang SY, Grondin F (2002) Improving structural lumber quality in a sample of *Picea mariana* logs sawn according to the knots. Wood Fiber Sci 34(2):266–275.

Li P, He J, Abbott L, Schmoldt DL (1996) Labeling defects in CT images of hardwood logs with species-dependent and species-independent classifiers. Pp 113–126 *in* Proc IAPR TC-8 Workshop on Machine Perception Applications, September 2–4, 1996, Technical University Graz, Graz, Austria.

Mas JF, Flores JJ (2008) The application of artificial neural networks to the analysis of remotely sensed data. Int J Remote Sens 29(3):617–663.

Nordmark U (2002) Knot identification from CT images of young *Pinus sylvestris* sawlogs using artificial neural networks. Scand J For Res 17(1):72–78.

Occeña LG (1991) Computer integrated manufacturing issues related to the hardwood log sawmill. J Forest Eng 3(1):39–45.

Pang S, Wiberg P (1998) Model predicted and CT scanned moisture distributed in a *Pinus radiata* board during drying. Holz Roh Werkst 56(1):9–14.

Panshin AJ, Zeeuw CD (1971) Textbook of wood technology: Volume 1. 3rd ed. McGraw-Hill Book Company, New York, NY. 705 pp.

Richards DB (1977) Value yield from simulated hardwood log sawing. Forest Prod J 27(12):47–50.

——— (1980) Lumber value for computerized simulation of hardwood log sawing. USDA For Prod Lab Res Pap FPL-356, Madison, WI.

Riedmiller M, Braun H (1993) A direct adaptive method for faster back-propagation learning: the RPROP algorithm. Pp 828–845 *in* Proc IEEE International Conference on Neural Network, March 28–April 1, 1993, San Francisco, CA.

Rojas G, Condal A, Beauregard R, Verret D, Hernandez RE (2006) Identification of internal defect of sugar maple logs from CT images using supervised classification methods. Holz Roh Werkst 64(4):295–303.

———, Hernández RE, Condal A, Verret D, Beauregard R (2005) Exploration of physical properties of internal characteristics of sugar maple logs and relationships with CT images. Wood Fiber Sci 37(4):591–604.

Rosin PL, Fierens F (1995) Improving neural network generalization. Pp 1255–1257 *in* Proc Internation Geoscience and Remote Sensing Symposium–IGARSS, July 10–14, 1995, Firenze, Italy.

Schmoldt DL, He J, Abbott AL (2000) Automated labeling of log features in CT imagery of multiple hardwood species. Wood Fiber Sci 32(3):287–300.

Sjoberg J (1995) Nonlinear black box modeling in system identification: A unified overview. Automatica 31(12): 1691–1724.

Wei Q, Leblon B, Chui YH, Zhang SY (2008) Identification of selected log characteristics from computed tomography images of sugar maple logs using maximum likelihood classifier and textural analysis. Holzforschung 62(4):441–447.

Zhu DP, Conners RW, Schmoldt DL, Araman PA (1996) A prototype vision system for analyzing CT imagery of hardwood logs. IEEE T Syst Man Cyb 26(4):522–532.