Support vector machines in wood identification:
the case of three Salix species from Turkey

Kemal TURHAN1,*, Bedri SERDAR2
1Department of Medical Informatics, School of Medicine, Karadeniz Technical University, Trabzon, Turkey
2Faculty of Forestry, Karadeniz Technical University, Trabzon, Turkey

Abstract: The aim of this study was to use a support vector machine (SVM) for the first time as a predictive method for differentiating species of Salix wood through the biometric analysis of their anatomy using wood taken from basal disks of 3 species. The purpose of a SVM is to construct optimal decision boundaries among classes in a decision plane. A decision plane separates a set of objects having different class memberships. In this study, the decision plane has 3 different wood species. Timely and accurate identification of tree species can be crucial in forestry. The similarity of structures in wood anatomy across many species, especially in the case of Salix species, means that they cannot be differentiated anatomically using traditional methods. SVM can be an effective tool for identifying similar taxa with a high percentage of accuracy. A SVM was used to differentiate Salix alba, Salix caprea, and Salix elaeagnos growing in Turkey. These Salix species are sufficiently similar that it is not possible to differentiate between them using traditional anatomical methods. However, a SVM was able to differentiate between the 3 species with a high degree of probability using the biometrics of wood anatomy. For the purposes of classification, a SVM with linear kernel function was designed; it attained an 80.6% success rate in the training group and a 95.2% success rate in the testing group. After feature selection, our SVM was able to classify the 3 species with notable success. If the number of samples were increased, the SVM would return more precise classification results.

Key words: Biometry, classification, Salix alba, Salix caprea, Salix elaeagnos, SVM

1. Introduction
The Salix L. species, which are members of the family Salicaceae, comprise deciduous and dioecious trees or shrubs. There are approximately 450 known species of Salix from all around the world (Argus 1997). Due to the complexity of the genus, however, classification remains difficult, and there is disagreement among authors regarding the exact number of species (Skvortsov 1999; Ohashi 2001; Mabberly 2008). Members of this group are found either naturally or cultivated in wetlands, on riverbanks, and at the sides of streams. Turkey has a wide variety of Salix species. Twenty-eight species grow naturally; 2 of them are endemic while 1 is rare within Turkey (Skvortsov and Edmondson 1970; Güner and Zielinski 1993; Güner 2000; Zielinski and Tomaszewski 2008; Arıhan and Güvenç 2011). Salix is an important group of plants because of their ecological and economic properties. Salix species have many traditional uses because of their role in ethnobotanical practices, such as basketry production (Arıhan and Güvenç 2011). Easy hybridization among willow species creates significant problems for the accurate identification and classification of individual specimens. The taxa of Salix are dioecious, and many species have varied growth periods for flowers and leaves, making it difficult to observe all of the relevant characteristics in a single plant or specimen.

The guiding principle behind support vector machines (SVMs) is constructing optimal decision boundaries between classes to determine the decision plane. Objects can be assigned correctly to the class to which they belong. SVM is a type of machine learning method used widely in many different classification problems, such as microRNA classification (Dubey et al. 2011), genetic cancer classification (Bharathi et al. 2011), estimation of wheat leaf diseases (Tian et al. 2011), and assessment of disease diagnosis (Turhan et al. 2011).

Furthermore, the flowers of Salix are too simple to afford adequate characteristics for taxonomy. When using measured anatomical features of Salix as input variables of SVM, the classification process will be easier thanks to the SVM’s classification ability. Therefore, in addition to morphological characteristics, the use of anatomical data and SVM methods may help to classify Salix taxa more easily.
There are a limited number of published studies describing the use of artificial intelligence or machine-learning methods as alternatives to traditional methods for wood identification. There are several studies in which artificial neural networks (ANNs) have been used for similar classification problems related to forestry. ANNs are widely used in many different classification problems and are inspired by biological nervous systems. Khalid et al. (2008) obtained a high rate of recognition accuracy by designing an automatic wood recognition system based on image processing, feature extraction, and ANNs. Moreover, an ANN was used to differentiate Juniperus cedrus from J. phoenicea var. canariensis on the Canary Islands (Esteban et al. 2009). In another study, an ANN was employed for modeling the moisture absorption and thickness swelling properties of an oriented standard board (Ozsahin 2012).

It has become increasing common to use SVM methods based on statistical learning theory. A system using a wood-structure SVM has been proposed for classifying 4 types of wood knots from images captured from lumber boards (Gu et al. 2009). In a similar vein, an interesting study was developed that used SVM methods to determine errors in the production of plywood (Pham et al. 2007).

There are mainly 2 ways to obtain data in wood anatomy studies. The first is measuring anatomical features with the aid of a microscope, while the other data are obtained automatically by image processing methods. Differentiation of anatomical structures that are similar in both cases requires experts, who are not always available. This system allows persons who are not qualified in wood species detection to perform the diagnosis of wood species with high probability. In these cases the woods are usually in pretty bad condition for obtaining such information because of aging, decay, deterioration, etc.

Despite the use of SVMs to solve various problems in forestry through image processing, no study was found that uses measured anatomical data and SVM. It may not always be possible to obtain anatomical features by image processing methods for unknown wood types. In this case, distinguishing characteristics of woods could only be measured by using microscopic methods. The primary objective of this study was to demonstrate a new method for wood species classification in the absence of precise information about the type of wood.

2. Materials and methods

Identifying wood species according to anatomical structure is a notably effective technique. However, there are no studies that apply artificial intelligence methods, such as SVM or ANN, to this technique. All of the samples (from 83 trees) of the 3 species studied were collected in the natural forests of Turkey. Standard techniques were used in the preparation of the wood sections, which were stained with safranin and alcian blue (Ives 2001); macerations were prepared using Schultze’s method (Normand 1972). In general, quantitative features are based on 25 measurements or counts per character; similarly, the lengths of axial elements are also based on 25 measurements of each cell type. The terminology used is in accordance with the list of microscopic features for hardwood identification (IAWA Committee 1989). Table 1 shows the features measured and the number of measurements obtained using microscopic

<table>
<thead>
<tr>
<th>Features/Salix species</th>
<th>Salix alba</th>
<th>Salix caprea</th>
<th>Salix elaeagnos (mean value)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vessel element</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number in early wood (0.5 mm² sample) (NEW)</td>
<td>35.81 (7.6)</td>
<td>37.96 (10.4)</td>
<td>42.70 (7.9)</td>
</tr>
<tr>
<td>Number of late wood (0.5 mm² sample) (NLW)</td>
<td>57.69 (16.6)</td>
<td>87.14 (20.8)</td>
<td>66.72 (13.2)</td>
</tr>
<tr>
<td>Number of vessels (1 mm² sample) (NV)</td>
<td>93.50 (21.2)</td>
<td>125.10 (24.8)</td>
<td>109.42 (19.1)</td>
</tr>
<tr>
<td>Length (µm) (VEL)</td>
<td>470.26 (62.5)</td>
<td>405.60 (57.1)</td>
<td>392.73 (37.3)</td>
</tr>
<tr>
<td>Diameter of radial (µm) (RDW)</td>
<td>95.77 (14.4)</td>
<td>75.60 (11.9)</td>
<td>73.84 (9.1)</td>
</tr>
<tr>
<td>Diameter of tangential (µm) (TDW)</td>
<td>64.39 (7.5)</td>
<td>51.69 (6.1)</td>
<td>52.24 (4.9)</td>
</tr>
<tr>
<td><strong>Fibers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length (µm) (FB)</td>
<td>949.03 (116.1)</td>
<td>823.81 (99.6)</td>
<td>782.15 (89.7)</td>
</tr>
<tr>
<td>Diameter (µm) (FW)</td>
<td>23.16 (3.1)</td>
<td>18.74 (2.8)</td>
<td>16.30 (2.89)</td>
</tr>
<tr>
<td>Lumen diameter (µm) (LW)</td>
<td>15.96 (2.7)</td>
<td>11.93 (2.1)</td>
<td>10.25 (2.5)</td>
</tr>
<tr>
<td>Thickness of cell wall (µm) (TCW)</td>
<td>3.61 (0.5)</td>
<td>3.42 (0.5)</td>
<td>3.05 (0.3)</td>
</tr>
<tr>
<td><strong>Rays</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of rays/1 mm (NR)</td>
<td>12.83 (1.1)</td>
<td>15.77 (1.1)</td>
<td>15.77 (1.4)</td>
</tr>
<tr>
<td>Height (µm) (RH)</td>
<td>272.02 (35.0)</td>
<td>291.32 (51.9)</td>
<td>256.02 (35.3)</td>
</tr>
<tr>
<td>Width (µm) (RW)</td>
<td>13.42 (1.6)</td>
<td>13.11 (1.1)</td>
<td>13.35 (1.9)</td>
</tr>
</tbody>
</table>
tools. The biometry was conducted on one preparation from each tree, all of which were samples of mature wood taken from a basal disk. The anatomical characteristics of wood samples—vessels (tangential and radial diameters, vessel length, their number in 0.5 mm² early wood, 0.5 mm² late wood, and 1 mm²), fibers (their length, width, thickness of cell walls, lumen diameter), and rays (their number in 1 mm, their heights and widths)—were investigated.

Fisher (1936) published the first pattern recognition related algorithm in 1936. Subsequently, in other studies SVMs were further developed out of statistical learning theory (Vapnik et al. 1974; Vapnik 1979). Using SVM is a highly successful approach to linear or nonlinear classification problems.

SVMs characterize an extension to nonlinear models of the generalized depiction algorithm developed by Vapnik and Lerner (1963). The SVM algorithm is built on the statistical learning theory and the Vapnik–Chervonenkis (VC) dimension (Vapnik and Chervonenkis 1974). Statistical learning theory, describing the properties of machines learning to make reliable estimates, was reviewed (Vapnik 1979, 1995, 1998). In the current formulation, the SVM algorithm was developed at AT&T Bell Laboratories by Vapnik et al. (Lipkowitz and Cundari 2007).

Conceptual and mathematical foundations of the SVM will be given below with the great support of the review that gives information about all aspects of the method (Lipkowitz and Cundari 2007).

2.1. SVM classification for linearly separable data

The optimum separation hyperplane (OSH) is the hyperplane with the maximum margin for a given limited set of learning patterns. The optimization problem based on the notations from Figure 1 will establish the conditions necessary to determine the maximum separation hyperplane. A hyperplane \( w.x + b = 0 \) can be represented as a pair \((w, b)\). The linear classifier is characterized by the set of pairs that satisfy the following inequalities for any pattern in the training set. Suppose the training set consisting of elements

\[
X = \{x_i, y_i\}, \quad i = 1, 2, \ldots, m \quad \text{and} \quad y_i \in \{-1,1\}:
\]

\[
\begin{align*}
\quad w.x_i + b > +1 & \quad \text{if} \quad y_i = +1 \\
\quad w.x_i + b < -1 & \quad \text{if} \quad y_i = -1 
\end{align*}
\]  

(1)

Compacted equation:

\[
y(wx_i + b) \geq +1
\]

(2)

Such a hyperplane \((w, b)\) is a classifier that correctly separates all patterns from the training set:

\[
\text{class}(x_i) = \begin{cases} 
+1 & \text{if} \quad w.x_i + b > 0 \\
-1 & \text{if} \quad w.x_i + b > 0 
\end{cases}
\]  

(3)

For the hyperplane that describes the linear classifier (i.e., where \(w.x + b = 0\)), the distance between the origin and the hyperplane \(H\) is \(|b|/||w||\). We consider the patterns from the class \(-1\) that satisfy the equality \(w.x + b = -1\) and that determine the hyperplane \(H_1\); the distance between the origin and the hyperplane \(H\) is equal to \(|-1-b|/||w||\). Similarly, we consider the patterns from the class \(+1\) that satisfy the equality \(w.x + b = +1\) and that determine the hyperplane \(H_2\); the distance between the origin and the hyperplane \(H\) is equal to \(|+1-b|/||w||\). Hyperplanes \(H_1\), \(H_2\), and \(H\) are parallel and no training patterns are located between hyperplanes \(H_1\) and \(H_2\). Based on the above considerations, the margin of the linear classifier \(H\) (the distance between hyperplanes \(H_1\) and \(H_2\)) is \(2/||w||\).

Consequently, the wider the margin, the smaller is \(d_v\), which is a nonnegative integer, called the Vapnik–Chervonenkis (VC) dimension of a classifier, determining the capacity of a classifier. From these considerations, it follows that the optimum separation hyperplane is obtained by maximizing \(2/||w||\), which is equivalent to minimizing \(||w||^2/2\). The problem of finding the optimum separation hyperplane is represented by the identification of the linear classifier \((w, b)\), which satisfies

\[
\begin{align*}
\quad w.x_i + b & \geq +1 \quad \text{if} \quad y_i = +1 \\
\quad w.x_i + b & \leq -1 \quad \text{if} \quad y_i = -1 
\end{align*}
\]  

(4)

for which \(||w||\) has the minimum value.

Based on the considerations presented above, optimum separation hyperplane conditions from Eq. (4) can be formulated into the following expression, which represents a linear SVM:

\[
\text{minimize } f(x) = \frac{||w||^2}{2}
\]

(5)

with the constraints \(g_i(x) = y_i(w.x_i + b) - 1 \geq 0, \ i = 1, \ldots, m\). The optimization problem from Eq. (5) represents the minimization of a quadratic function under linear constraints (quadratic programming), a problem studied comprehensively in optimization theory. Details on quadratic programming can be found in almost any textbook.
on numerical optimization, and efficient implementations exist in many software libraries. However, Eq. (5) does not represent the actual optimization problem that is solved to determine the OSH. Based on the use of a Lagrange function, Eq. (5) is transformed into its dual formulation. All SVM models (linear and nonlinear, classification and regression) are solved for the dual formulation, which has important advantages over the primal formulation [Eq. (5)]. The dual problem can be easily generalized to linearly nonseparable learning data and to nonlinear support vector machines.

An appropriate way to solve constrained minimization problems is by using a Lagrangian function of the problem defined in Eq. (5):

\[ \minimize L_p(w, b, \lambda) = \frac{1}{2} \| w \|^2 \]

\[ -\sum_{i=1}^{m} \lambda_i y_i w \cdot x_i - \sum_{i=1}^{m} \lambda_i y_i b + \sum_{i=1}^{m} \lambda_i \]

(6)

Here \( \lambda = (\lambda_1, \lambda_2, ..., \lambda_m) \) is the set of Lagrange multipliers of the training (calibration) patterns with \( \lambda_i \geq 0 \), and \( P \) in \( L_p \) indicates the primal formulation of the problem. The Lagrangian function \( L_p \) must be minimized with respect to \( w \) and \( b \), and maximized with respect to \( \lambda_i \), subject to the constraints \( \lambda_i \geq 0 \). This is equivalent to solving the Wolfe dual problem (Fletcher, 1987), namely to maximize \( L_d \) subject to the constraints that the gradient of \( L_p \) with respect to \( w \) and \( b \) is zero, and subject to the constraints \( \lambda_i \geq 0 \).

Solving the SVM problem is equivalent to finding a solution to the Karush–Kuhn–Tucker (KKT) conditions. The dual problem \( L_d \):

\[ \maximize L_d(w, b, \lambda) = \]

\[ \sum_{i=1}^{m} \lambda_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \lambda_i \lambda_j y_i y_j x_i \cdot x_j \]

subject to \( \lambda_i \geq 0, i = 1, ..., m \), \( j = 1, ..., m \)

and \( \sum_{i=1}^{m} \lambda_i y_i = 0 \)

(7)

2.2. Prediction for new patterns

After training, the classifier is ready to find the class membership for new patterns not included in the training set. In this case, the support vectors from the training set and the corresponding values of the Lagrange multipliers \( \lambda_i \) will be used for prediction:

\[ \text{class}(x_k) = \text{sign} \left( \sum_{i=1}^{m} \lambda_i y_i x_i \cdot x_k + b \right) \]

(8)

Patterns that are not support vectors (\( \lambda_i = 0 \)) do not affect the classification of new patterns. It is only necessary to compute the dot product between \( x_k \) (\( k^{th} \) learning pattern to be classified and not included in the training set) and every support vector. This results in a significant saving of computational time whenever the number of support vectors is small compared with the total number of patterns in the training set. Moreover, Eq. (8) can be easily adapted for nonlinear classifiers that use kernels.

Both the primal \( L_p \) and the dual \( L_d \) Lagrangian functions are derived from the same objective functions but with different constraints, and the solution is found by minimizing \( L_p \) or by maximizing \( L_d \). The training patterns from the SVM solution that have \( \lambda_i \geq 0 \) represent the support vectors. The training patterns that have \( \lambda_i = 0 \) are not important in obtaining the SVM model, and they can be removed from training without any influence on the SVM result. As will be seen below, any SVM model is completely defined by the set of support vectors and the corresponding Lagrange multipliers.

It can be seen below how the SVM classification for linearly separable patterns can be adapted to accommodate classification errors in the training set. The resulting SVM classification will still be linear and it will be able to compute an optimum separation hyperplane even for cases like that in Figure 2, which cannot be fully separated with a linear classifier.

To obtain an optimum linear classifier for nonseparable data (Figure 3), a penalty is introduced for misclassified data, denoted with \( \xi_i \) and called a slack variable. This penalty related to any pattern in the training is zero for patterns classified correctly, and has a positive value that increases with the distance from the corresponding hyperplane for patterns that are not situated on the correct side of the classifier.

\[ \minimize \frac{\| w \|^2}{2} + C \sum_{i=1}^{m} \xi_i \]

with the constraints

\[ y_i (w \cdot x_k + b) \geq +1 - \xi_i \quad i = 1, ..., m \]

\[ \xi_i \geq 0, \xi_i \quad i = 1, ..., m \]
where C is a constant that can be adjusted by the user, and can either increase or decrease the penalty for classification errors. A large C gives a higher penalty to classification errors, consequently minimizing the number of misclassified patterns. A small C maximizes the margin so that the OSH is less sensitive to errors from the learning set.

To solve the above constrained quadratic optimization problem, the approach based on Lagrange multipliers can be followed [Eq. (6)]. Lagrange multipliers can be defined $\lambda = (\lambda_1, \lambda_2, ..., \lambda_m)$ for each constraint $y_i(w.x_i + b) \geq +1 - \xi_i$ and the Lagrange multipliers $M = (\mu_1, \mu_2, ..., \mu_m)$ for each constraint $\xi_i \geq 0, \forall = 1, ..., m$. With these notations, the primal Lagrangian function of this problem is minimize $L_p(w, b, \lambda, M) = \frac{1}{2} \|w\|^2 + C\sum_{i=1}^{m} \xi_i - \sum_{i=1}^{m} \lambda_i [y_i(w.x_i + b) - 1 + \xi_i]$ 

(10)

$-\sum_{i=1}^{m} \mu_i \xi_i$

where $\lambda = (\lambda_1, \lambda_2, ..., \lambda_m)$ is the set of Lagrange multipliers of the training (calibration) patterns.

The right side of the Lagrangian function (to find a solution to KKK conditions), obtaining the dual problem is maximize $L_d(w, b, \lambda, M) = \sum_{i=1}^{m} \lambda_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \lambda_i \lambda_j y_i x_i x_j$ 

(11)

subject to $0 \geq \lambda_i \geq C, i = 1, ..., m$

and $\sum_{i=1}^{m} \lambda_i y_i = 0$

The classification of new patterns uses the optimum values for w:

$$\text{class} (x_i) = \text{sign} \left( \sum_{i=1}^{m} \lambda_i y_i x_i . x_k + b \right)$$

(12)

2.3. Mapping patterns to a feature space

The separation surface may be nonlinear in many classification problems, but SVMs can be extended to deal with nonlinear separation surfaces by using feature functions $\phi(x)$. The SVM extension to nonlinear datasets is based on mapping the input variables into a feature space of a higher dimension (Figure 4) and then performing a linear classification in that higher dimensional space.

We consider a soft margin SVM in which the variables $x$ are substituted with the feature vector $\phi(x)$, which represents an optimization problem similar to that from Eq. (9). Using this nonlinear SVM, the class of a pattern $x_k$ is determined with Eq. (13).

$$\text{class} (x_i) = \text{sign} \left( \sum_{i=1}^{m} \lambda_i y_i \phi (x_i) . \phi (x_i) + b \right)$$

(13)

In this study SVM was implemented using Statistica, a very user friendly statistical software package, according to the model described above. Statistica supports a number of kernels ($\phi$) for use in SVM models. These are listed below.

- Linear $x_i ^* x_j$
- Polynomial $(\gamma x_i . x_j + \text{coefficient})^{\text{degree}}$
- Radial based function (RBF) $\exp(-\gamma |x_i * x_j|^2)$
- Sigmoid $\tanh (\gamma x_i . x_j + \text{coefficient})$

In this study, classification of wood types was implemented by using linear kernel function and performed in 2 phases.

3. Results

Thirteen biometric patterns as predictor values were used to construct support vectors in the first phase. Inputs and outputs of SVM training by using the statistical tools are given below.

3.1. First training phase and SVM results for 13 biometric patterns

Dependent variable to be classified: Salix species
Input variables as predictors: VEL, FB, FW, LW, TCW, NR, RH, RW, NEW, Nlw, NV, RDW, TDW

Sample size = 62 (Train), 21 (Test), 83 (Overall)

Kernel type: Linear, (for mapping variable to feature space),
Capacity constants: (for slack variable for SVM) = 10
Number of support vectors = 35 (11 bounded)
Support vectors per class: 12 (Salix alba), 14 (Salix caprea), 9 (Salix elaeagnos)
Classification accuracy (%) = 88.710 (Train), 80.952 (Test), 86.747 (Overall)

The success of the test class was found to be 80.952% using all of the features. In this context, feature selection was performed to predict the power of each individual biometric pattern. The power of a feature is indicated by a positive correlation between the biometric pattern and wood species. For feature selection, a chi-square test was performed. The results are given in Table 2.

FW, LW, NR, NLW, TDW, and RDW were selected for the new SVM analysis due to their high predictive powers.

3.2. Second training phase and SVM results for 6 biometric patterns
Dependent variable to be classified: Salix species
Input variables as predictors: FW, LW, NR, NLW, RDW, TDW
Sample size = 62 (Train), 21 (Test), 83 (Overall)
SVM results
Kernel type: Linear (for mapping variable to feature space),
Capacity constants: (for slack variable for SVM) = 10
Number of support vectors = 35 (16 bounded)

Support vectors per class: 10 (Salix alba), 15 (Salix caprea), 10 (Salix elaeagnos)
Classification accuracy (%) = 80.645 (Train), 95.238 (Test), 84.337 (Overall)
The calculated probabilities of belonging to a particular species are displayed in the confidence column (Table 3). When these possibilities are examined, Salix alba and Salix

<table>
<thead>
<tr>
<th>Variables</th>
<th>Chi-square</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FW</td>
<td>73.50212</td>
<td>0.000000</td>
</tr>
<tr>
<td>LW</td>
<td>70.97164</td>
<td>0.000000</td>
</tr>
<tr>
<td>NR</td>
<td>70.87531</td>
<td>0.000000</td>
</tr>
<tr>
<td>NLW</td>
<td>53.31215</td>
<td>0.000007</td>
</tr>
<tr>
<td>TDW</td>
<td>54.12006</td>
<td>0.000018</td>
</tr>
<tr>
<td>RDW</td>
<td>51.96943</td>
<td>0.000038</td>
</tr>
<tr>
<td>FB</td>
<td>41.22333</td>
<td>0.001418</td>
</tr>
<tr>
<td>NV</td>
<td>32.71389</td>
<td>0.008063</td>
</tr>
<tr>
<td>VEL</td>
<td>31.53180</td>
<td>0.024964</td>
</tr>
<tr>
<td>TCW</td>
<td>27.74544</td>
<td>0.066036</td>
</tr>
</tbody>
</table>

Table 2. Chi-square test results.

<table>
<thead>
<tr>
<th>Type of Salix</th>
<th>SVM prediction (95.2%)</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Salix alba</td>
</tr>
<tr>
<td>Salix alba</td>
<td>Salix alba</td>
<td>0.7</td>
</tr>
<tr>
<td>Salix alba</td>
<td>Salix alba</td>
<td>0.7</td>
</tr>
<tr>
<td>Salix alba</td>
<td>Salix alba</td>
<td>0.7</td>
</tr>
<tr>
<td>Salix alba</td>
<td>Salix alba</td>
<td>0.7</td>
</tr>
<tr>
<td>Salix alba</td>
<td>Salix alba</td>
<td>0.7</td>
</tr>
<tr>
<td>Salix alba</td>
<td>Salix alba</td>
<td>0.7</td>
</tr>
<tr>
<td>Salix caprea</td>
<td>Salix caprea</td>
<td>0.0</td>
</tr>
<tr>
<td>Salix caprea</td>
<td>Salix caprea</td>
<td>0.0</td>
</tr>
<tr>
<td>Salix caprea</td>
<td>Salix caprea</td>
<td>0.3</td>
</tr>
<tr>
<td>Salix caprea</td>
<td>Salix caprea</td>
<td>0.0</td>
</tr>
<tr>
<td>Salix caprea</td>
<td>Salix caprea</td>
<td>0.0</td>
</tr>
<tr>
<td>Salix caprea</td>
<td>Salix caprea</td>
<td>0.0</td>
</tr>
<tr>
<td>Salix caprea</td>
<td>Salix caprea</td>
<td>0.0</td>
</tr>
<tr>
<td>Salix elaeagnos</td>
<td>Salix elaeagnos</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 3. The probability of belonging to a class according to the SVM analysis.

1 Incorrect classification
caprea exhibit more similar features. The same similarity can be observed between Salix elaeagnos and Salix caprea. Table 3 shows that the model is able to distinguish between Salix alba and Salix elaeagnos more easily.

4. Discussion
An ANN was used to differentiate Juniperus cedrus and J. phoenicea var. canariensis, both from the Canary Islands, and their anatomical properties were very similar to each other. The proposed neural network was satisfactory for the desired purpose and enabled J. cedrus and J. phoenicea var. canariensis to be differentiated with a 92% probability (Esteban et al. 2009). We developed a Multi-Layer Perceptron AN using the Statistica program with 6 input variable (FW, LW, NR, NLW, TDW, RDW), 10 hidden layers, and 3 outputs (for the species). Its test performance was 80%.

The main limitation of this study was the small number of species sampled. When the number of species sampled is increased, a more successful and accurate classification can be developed. SVM and other artificial intelligence methods can be used in combination to solve more complex classification problems related to wood identification. Obtaining the biometric patterns requires tedious and time-consuming procedures. In future studies, using software and technology to automate the collection of biometric patterns could accelerate the process.

References


