Wood Sci Technol (2017) 51:929–942
DOI 10.1007/s00226-017-0915-8

ORIGINAL

Artificial neural network and SIMCA classification in some wood discrimination based on near-infrared spectra

Silvana Nisgoski[1] · André Anastácio de Oliveira[2] · Graciela Inés Bolzon de Muñiz[1]

Received: 11 December 2015 / Published online: 6 May 2017

Abstract Species recognition and identification are very important in the wood industry. The identification of some tree species is complex when only the wood is available, often requiring multiple characterization techniques. In some instances, near-infrared spectroscopy can provide a method for the identification of wood species. However, as the amount of data acquired by near-infrared spectrometers are large, there is a need to use mathematical and computational tools to treat and analyze the data. This paper reports the results of testing an artificial neural network in comparison with SIMCA classification to identify some Brazilian wood species based on near-infrared spectra. The neural network developed did not result in any identification error for a margin of ±2% with the use of a spectral range from 4000 to 10,000 cm^{-1}, while SIMCA produced more than 60% identification error with raw spectral data. The artificial neural network was more efficient than the SIMCA classification and has good potential to be applied for species discrimination.

Introduction

Correct identification of wood species is important since it is related to various properties and cost of materials. Examination of cross sections and surfaces of the material of interest is usually employed to identify wood species (Labati et al. 2009).

✉ Silvana Nisgoski
silvana.ufpr@gmail.com

1 Department of Forest Engineering and Technology, Federal University of Parana, Av. Pref. Lothário Meissner, 900, Jardim Botânico, Curitiba, PR 80210-170, Brazil

2 CESEC, Numerical Methods Graduate Program in Engineering, Polytechnic Center, Federal University of Parana, Jardim das Américas', Curitiba, PR 81531-980, Brazil

Because of excess exploitation, many forest tree species are on the official list of endangered plants in Brazil, including species such as Imbuia/Brazilian walnut (*Ocotea porosa*), Brazilian sassafras (*Ocotea odorifera* (Vellozo) Rohwer) and black cinnamon (*Ocotea catharinensis* Mez), along with other species of the same family (MMA 2008). Species with similar morphological and anatomical features are very difficult to classify (Banerjee et al. 2008), making it necessary to develop and apply fast techniques to obtain more information on the intrinsic characteristics of wood to enable reliable species identification, particularly when little information is available about anatomical or chemical composition. An alternative could be to use an artificial neural network (ANN) combined with near-infrared spectroscopy.

Artificial neural networks were inspired by biological neural networks and consist of many neurons that process information based on nonlinear and multivariable relationships between process parameters (Ozsahin 2012). Their application utilizes their ability to learn from prior examples and to perform generalized identification or recognition of previously unseen patterns. Their limitations depend on the quantity, validity and accuracy of training data (Clark 2003).

In species identification, taxa that are only doubtfully distinct from each other are difficult for any identification system to distinguish. Although a formal botanical key can be better, it often takes a long time to create. Furthermore, a higher level of expertise is required for construction of an effective identification system such as that demonstrated here using the neural network approach (Clark 2003).

For species identification, artificial neural networks have been shown to have good potential to identify species of fish (Simmonds et al. 1996), butterfly (Kang et al. 2012), mosquitoes (Banerjee et al. 2008; Lorenz et al. 2015), jujube tree pathogens (Zhang et al. 2013), wood (Esteban et al. 2009; Ma et al. 2012) and plants from leaf characteristics (Kattmah and Azim 2013). In wood technology, some examples of neural networks application are for classifying images of wood veneer (Packianather and Drake 2000), for identification of wood defects (Pham et al. 2006; Mu et al. 2015), for monitoring MDF milling (Zbiéc 2011), for predicting tensile index and brightness in pulp (Okan et al. 2015) and also for wood classification based on images characteristics (Sundaram et al. 2015).

In this context, some studies have applied infrared spectroscopy for species discrimination based on solid or powdered samples (Adedipe et al. 2008; Russ et al. 2009; Casale et al. 2010; Braga et al. 2011; Pastore et al. 2011; Sandak et al. 2011; Nisgoski et al. 2015). Studies of wood identification applying artificial neural networks combined with near-infrared spectroscopy are scarce. Ma et al. (2012) demonstrated the potential of combining these two techniques. The objective of this paper is to test an artificial neural network (ANN) in comparison with SIMCA classification to identify some Brazilian wood species based on near-infrared spectra.

Materials and methods

Wood samples of *Ocotea porosa, Ocotea odorifera, Nectandra* sp. (Lauraceae) and *Eucalyptus* sp. (Myrtaceae), with dimensions of 2 × 2 × 5 cm^3, obtained from the collection of the Wood Anatomy and Quality Laboratory of Parana Federal

University (UFPR) were used. Material was selected based on its availability in number of samples, species listed in endangered list (*Ocotea porosa*, *O. odorifera*), wood from Lauraceae family much similar with difficult discrimination (*Nectandra* sp.) and wood with no problem of commerce and visually and anatomically distant to other samples (*Eucalyptus* sp.) to test the potential of ANN and SIMCA in wood discrimination in different situations. Sixty (60) physical samples of each species, collected from different boards, were used without identification of age or position within the tree being known.

Near-infrared spectra were acquired using a Bruker Tensor 37 spectrometer (Bruker Optics, Ettlingen, Germany, www.bruker.com) equipped with an integrating sphere operating in reflectance mode. A total of 64 scans were acquired for each spectrum with a resolution of 4 cm^{-1} and a spectral range of 10,000–4000 cm^{-1}, in other words, a set of 1500 wavenumbers. The samples were placed on top of integrating sphere, and four spectra were obtained from different points from each face; transversal, radial and tangential, resulting in a total of 12 separate spectra for each physical sample. All spectra were identified by sample name and used for further analysis without averaging.

The ANN used was developed with MATLAB (version 7.10 with Neural Network Toolbox 6, from MathWorks, Natick, MA, www.mathworks.com). The architecture employed consisted of an artificial neural network with one hidden layer, using back propagation learning with the Levenberg–Marquardt algorithm for training and the gradient descent method with moment to update the weights and biases. The performance index used was the mean square error (MSE) with the target criteria being less than or equal 10^{-10}, and for the network output an error tolerance of ±2%.

The input layer was composed of an array of size $P_{ij}(1500 \times 1)$ belonging to the set of samples, where its elements represent the absorbance for each of the 1500 wavenumbers acquired in each spectra. The input layer, the absorbance by wavenumber, belongs to interval [0, 1], due to the treatment given by the Fourier transform, a FTIR feature. All data were randomly presented to neural network (Fig. 1), and the goal is to distribute any possible noises and optimize the results. The network output was adjusted by a linear activation function to discrete intervals [1, 2, 3, 4], chosen for convenience. Each of four numbers refers to a type of wood, as shown in Table 1. Reference values indicate the desired output of the network for a given species, for example by analyzing the spectra of the samples relative to a wood, if the network provides a value close to 1, considering a tolerance of ±2%, then the species analyzed is probably *Nectandra* sp.

As shown in Table 1, the initial model developed using ANN used approximately 60% of the available samples, and so for further analysis a subset of 70% of these samples was selected to create an ANN training set, leaving 20% for validation via cross validation and 10% for the test set. The 40% remaining samples were used exclusively for testing the fully trained ANN.

In search of the best model, the number of neurons in the hidden layer was alternated empirically from one to fifteen. For each choice of neuron number in the hidden layer, the network was trained repeatedly ten times in order to find the best result, based on randomly obtained weights and biases.

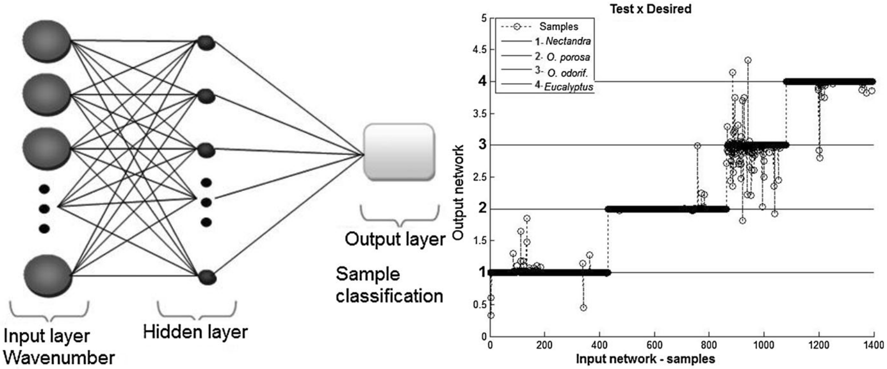

Fig. 1 Representation of architecture applied to network

Table 1 Samples used for neural network analysis and SIMCA classification

Reference	Species	Test samples	Training samples	Total samples
1	*Nectandra* sp.	288	431	719
2	*Ocotea porosa*	288	432	720
3	*Ocotea odorifera*	218	218	436
4	*Eucalyptus* sp.	211	316	527
Total		1005	1397	2402

The Unscrambler X (version 10.1, from CAMO Software AS, Oslo, Norway, www.camo.com) multivariate analysis program was used to develop the SIMCA model. Exploratory modeling was done by analyzing the score and loading graphs obtained by principal component analysis (PCA), and SIMCA classification was performed with the sample numbers listed in Table 1.

For ANN and SIMCA, data were analyzed in raw form and also two pretreatment were tested: second derivative of Savitzky–Golay (polynomial order = 2, smoothing point = 3) and multiplicative scatter correction (MSC).

Results and discussion

The mean spectra of the four woods studied (Fig. 2) and the principal component analysis performed on the individual spectra (Fig. 3) show their similarity. Informative wavenumbers are correlated with the presence of polysaccharides, lipids and protein, which are related to cell structure and are resumed in Table 2.

In all species, along PC1 (Fig. 3), there is a distinction of two groups, due to spectral data being acquired from the transverse and radial–tangential faces. In *Ocotea porosa* it is much more evident and may be the result of wood anatomical characteristics, with more distinction of growth rings, frequency of vessels and cell oil. Spectra collected on the transverse section were more often correctly classified

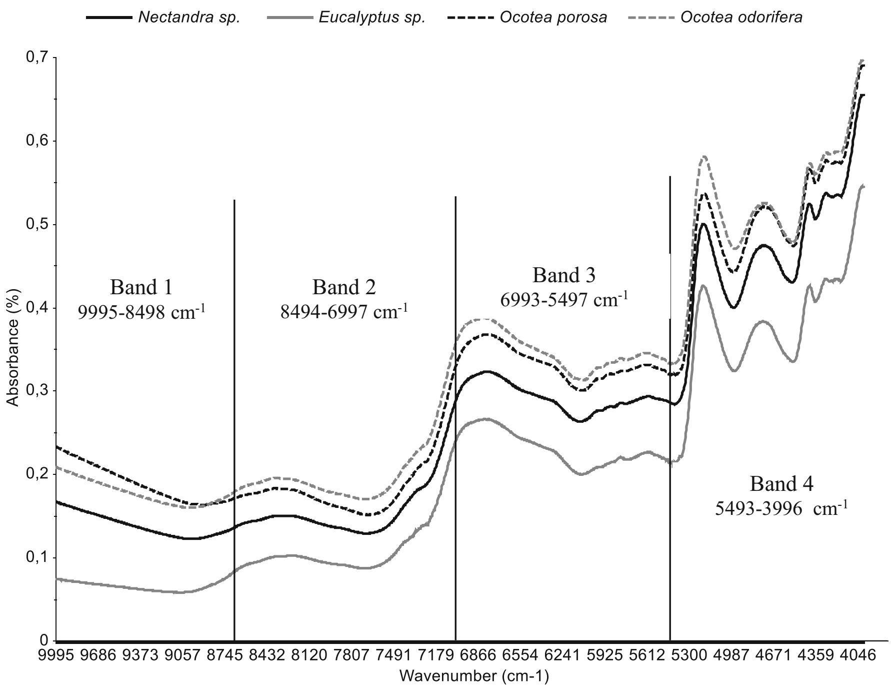

Fig. 2 Mean spectra of raw data from studied material

in SIMCA analysis than when spectra were collected from the radial–tangential face, and this result is similar to anatomical identification, where the transverse section provides the most important information in species identification. In the Lauraceae family, the distinction between *Ocotea* and *Nectandra* based on wood anatomy is difficult, and the characteristic odor from *Ocotea porosa* and *Ocotea odorifera* is important for correct information, but in some *Nectandra* species it is only the crystal type (raphides, prismatic or acicular crystals) and occurrence that is able to provide correct identification.

The original spectra and pretreatments using second derivative or multiplicative scatter correction (MSC) were analyzed by both ANN and SIMCA techniques. In the instance of the SIMCA models, the first two PCs and the number of training samples for each species in Table 1 were used. Individual models were based on the NIPALS algorithm and validated with leverage correction. In the ANN analysis, the number of neurons in the hidden layer and the training error are given in Table 3. One important detail is the number of neuron in hidden layer; small numbers indicate gain in speed of learning of ANN.

A table was constructed after classification using either ANN or SIMCA with either the complete spectral range (Table 4) or divided into regions (Tables 5, 6, 7, 8, 9). In SIMCA classification, there was an overlapping classification, resulting in some samples not being uniquely classified, and in fact some individual samples

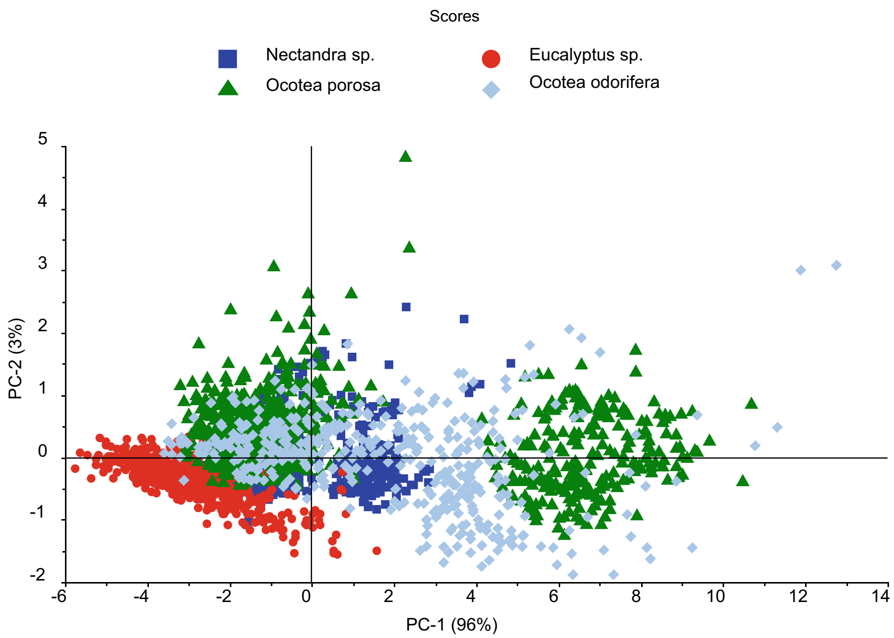

Fig. 3 Principal component analysis of raw data

Table 2 Informative wavenumbers from wood spectra (Tsuchikawa and Siesler 2003; Yonenobu and Tsuchikawa 2003; Schwanninger et al. 2011)

Band (cm^{-1})	Component
~8370	CH second overtone and CH_3–groups
~7000	Amorphous region of cellulose
6900–6850	C–H combination of aromatics, phenolic OH group of lignin and extractives
6622	O–H first overtone and interchain H-bond of cellulose
6287	Crystalline regions of cellulose
5974	Aromatic ring of lignin
5800	Furanose/pyranose functional group present in hemicellulose
5587	Crystalline and semi-crystalline cellulose
4739	Cellulose
4700	CH deformations attributed to water
4686	Acetyl groups, lignin and extractives
4198	Holocellulose
4014	C–H and C–C stretching and cellulose

may have been classified to two or more species. Each species therefore has both its right and wrong classified number of samples presented in the tables.

When the analysis was performed with the complete spectral range (Table 4), ANN was able to correctly classify all the samples across all wavenumbers within an error tolerance of ±2% when using raw spectral data and MSC pretreatment. In

Table 3 Training error and number of neurons in the hidden layer in ANN analysis

Spectra	Raw		2nd derivative		MSC	
	Error	Neuron	Error	Neuron	Error	Neuron
Full	3.38×10^{-15}	8	3.68×10^{-02}	6	6.23×10^{-08}	4
Band 1	2.91×10^{-02}	1	1.19×10^{-01}	12	1.52×10^{-02}	10
Band 2	7.29×10^{-03}	3	9.33×10^{-02}	1	1.10×10^{-03}	6
Band 3	3.79×10^{-3}	3	1.07×10^{-02}	4	1.40×10^{-03}	7
Band 4	2.56×10^{-07}	2	4.00×10^{-03}	8	6.26×10^{-06}	13
Band 2–4	1.11×10^{-02}	12	2.79×10^{-02}	9	8.91×10^{-10}	5

the case of the second derivative, a total of 24% of samples was not classified. SIMCA classification, using raw spectral data, performed best for *Ocotea porosa* and *Ocotea odorifera,* but many samples were incorrectly identified. In samples of *Nectandra* sp., the best result was observed after applying a MSC pretreatment, with a correct identification of 95.5% of samples. For *Eucalyptus* sp., the best result was with second derivative pretreatment resulting in a correct identification of 46% of samples. It is interesting that ANN has obtained very satisfactory results using the raw spectrum, while SIMCA requires prior treatment for best results.

When the analysis was performed using only wavelengths from 9995 to 8498 cm^{-1} (Table 5), ANN produced the best classification with MSC pretreatment, but this region in the near-infrared spectra does not provide relevant information about the species' chemical composition, producing mostly noise, which can explain the high rate of unidentified samples in *Ocotea porosa* and *O. odorifera*. In the SIMCA classification, the best result is for *Nectandra* sp. with MSC pretreatment, with 95.1% of samples correctly classified. In this spectral region, in ANN 21% (MSC) to 65% (raw) of samples were not classified, and in SIMCA many samples were non-unique classified. The best results were obtained by applying the filter MSC, to both ANN and SIMCA. Although this spectrum band presents no great relevance, the MSC filter accentuated the most significant absorbance variations. The identification methods are sensitive to these variations, and it was noticed that the ANN showed to be more sensitive to the identification of these variations.

When the analysis was performed with wavelengths from 8494 to 6997 cm^{-1} (Table 6), some wavelengths associated with cellulose, lignin and extractives were selected and the ANN performed better. In this case, the best network performance occurred with data with a MSC pretreatment, and only 1.5% of samples were unclassified. For SIMCA, with raw data, there was a great deal of confusion between species, including *Eucalyptus,* a species from the Myrtaceae family, which is anatomically different to species of the Lauraceae family. In SIMCA classification, samples of *Ocotea odorifera* classified as *Nectandra* sp. was an interesting result, as spectra acquired from the transverse plane were not classified in this case. The best result was achieved using MSC for *Nectandra* sp., second derivate for *Eucalyptus* sp. and raw data for *Ocotea porosa* and *Ocotea odorifera.* In this spectrum band, the ANN was more sensitive to the absorbance variations compared

Table 4 Classification table for ANN and SIMCA for the complete spectral region (4000–10,000 cm^{-1})

Species	Pretreatment	Classification					
		ANN			SIMCA		
		Right	Wrong	Not classified	Right	Wrong	Not classified
Nectandra sp.	Raw	288	0	0	5	273	10
	2nd derivative	280	0	8	169	97	22
	MSC	288	0	8	275	0	13
Eucalyptus sp.	Raw	211	0	0	1	210	0
	2nd derivative	206	0	5	97	114	0
	MSC	211	0	0	0	0	211
Ocotea porosa	Raw	288	0	0	70	203	15
	2nd derivative	198	0	90	13	271	4
	MSC	288	0	0	0	113	175
Ocotea odorifera	Raw	218	0	0	131	76	11
	2nd derivative	79	0	137	1	216	1
	MSC	218	0	0	0	11	207

Wrong classification represents one sample misclassified and samples with multiple classifications

Table 5 Classification table for ANN and SIMCA with band 1 (9995–8498 cm^{-1})

Species	Pretreatment	Classification					
		ANN			SIMCA		
		Right	Wrong	Not classified	Right	Wrong	Not classified
Nectandra sp.	Raw	53	0	235	0	279	9
	2nd derivative	58	3	227	0	288	0
	MSC	288	0	0	274	0	14
Eucalyptus sp.	Raw	180	0	31	2	209	0
	2nd derivative	166	0	45	14	197	0
	MSC	211	0	0	0	0	211
Ocotea porosa	Raw	63	2	223	53	215	20
	2nd derivative	89	3	196	7	281	0
	MSC	182	0	106	0	0	10
Ocotea odorifera	Raw	24	28	166	107	104	7
	2nd derivative	42	12	164	0	218	0
	MSC	116	0	102	0	170	48

Wrong classification represents one sample misclassified and samples with multiple classifications

to the band 1. It can be concluded that in this band there is more meaningful information for the distinction of the woods. It is difficult to identify which wavenumber is more significant for the distinction, because this information is processed within the black box of the ANN (Sussillo and Barak 2013).

Table 6 Classification table for ANN and SIMCA with band 2 (8494–6997 cm^{-1})

Species	Pretreatment	Classification					
		ANN			SIMCA		
		Right	Wrong	Not classified	Right	Wrong	Not classified
Nectandra sp.	Raw	65	0	223	0	286	2
	2nd derivative	170	0	118	111	94	83
	MSC	288	0	0	260	0	28
Eucalyptus sp.	Raw	129	0	82	0	209	2
	2nd derivative	197	0	14	171	40	0
	MSC	211	0	0	0	164	47
Ocotea porosa	Raw	134	0	154	93	191	4
	2nd derivative	64	0	224	0	287	1
	MSC	287	0	1	0	79	209
Ocotea odorifera	Raw	5	60	153	131	76	11
	2nd derivative	35	16	167	0	217	1
	MSC	204	0	14	0	35	183

Wrong classification represents one sample misclassified and samples with multiple classifications

Table 7 Classification table for ANN and SIMCA with band 3 (6993–5497 cm^{-1})

Species	Pretreatment	Classification					
		ANN			SIMCA		
		Right	Wrong	Not classified	Right	Wrong	Not classified
Nectandra sp.	Raw	233	0	55	93	191	4
	2nd derivative	280	1	7	111	108	69
	MSC	288	0	0	209	0	79
Eucalyptus sp.	Raw	210	0	1	2	188	21
	2nd derivative	211	0	0	171	40	0
	MSC	211	0	0	0	0	211
Ocotea porosa	Raw	177	0	111	51	234	3
	2nd derivative	283	0	5	98	187	3
	MSC	288	0	0	0	0	288
Ocotea odorifera	Raw	218	0	0	105	103	10
	2nd derivative	202	2	14	72	146	0
	MSC	202	0	16	0	0	218

Wrong classification represents one sample misclassified and samples with multiple classifications

When the analysis was performed with wavelengths from 6993 to 5497 cm^{-1} (Table 7), some wavelengths associated with water, cellulose, lignin and extractives were selected by the model and the ANN produced a better classification with MSC pretreatment for all species, with only 1.5% of samples not identified. The best

Table 8 Classification table for ANN and SIMCA with band 4 (5493–3996 cm^{-1})

Species	Pretreatment	Classification					
		ANN			SIMCA		
		Right	Wrong	Not classified	Right	Wrong	Not classified
Nectandra sp.	Raw	288	0	0	23	259	6
	2nd derivative	288	0	0	175	28	85
	MSC	288	0	0	261	0	27
Eucalyptus sp.	Raw	211	0	0	0	210	1
	2nd derivative	211	0	0	197	0	14
	MSC	211	0	0	0	0	211
Ocotea porosa	Raw	288	0	0	92	193	3
	2nd derivative	288	0	0	233	38	17
	MSC	288	0	0	0	0	288
Ocotea odorifera	Raw	213	1	4	184	22	12
	2nd derivative	217	0	1	194	17	7
	MSC	218	0	0	0	0	218

Wrong classification represents one sample misclassified and samples with multiple classifications

Table 9 Classification table for ANN and SIMCA with band 2–4 (8494–3996 cm^{-1})

Species	Pretreatment	Classification					
		ANN			SIMCA		
		Right	Wrong	Not classified	Right	Wrong	Not classified
Nectandra sp.	Raw	253	0	35	0	286	2
	2nd derivative	231	0	57	115	91	82
	MSC	288	0	0	259	0	29
Eucalyptus sp.	Raw	203	0	8	0	210	1
	2nd derivative	170	0	41	205	4	2
	MSC	211	0	0	0	1	210
Ocotea porosa	Raw	204	0	84	127	160	1
	2nd derivative	196	0	92	171	102	15
	MSC	288	0	0	0	0	288
Ocotea odorifera	Raw	159	0	59	44	163	11
	2nd derivative	74	0	144	99	112	7
	MSC	218	0	0	0	0	218

Wrong classification represents one sample misclassified and samples with multiple classifications

result was due to the existence, in this spectrum range, of a greater amount of significant information for the distinction. However, it was still necessary to use a filter to highlight the variations of absorbance. For SIMCA, a significant confusion between species occurred, including a classification of some species from the Lauraceae family as *Eucalyptus*. Samples of *Ocotea porosa* classified as either

Nectandra sp. and/or *Eucalyptus* sp. were from radial and tangential sections, indicating that some characteristics of the surface have influence on the spectra. The best result used MSC pretreatment for *Nectandra* sp., second derivative for *Eucalyptus* sp. and *Ocotea porosa* and raw data for *Ocotea odorifera*. Many samples were non-unique and not classified by SIMCA in this band.

When the analysis was performed with wavelengths from 5493 to 3996 cm^{-1} (Table 8), some wavelengths associated with water, cellulose, lignin and extractives were selected and the classification by ANN was similar for all spectra. Note that there is no significant difference between the results obtained with the data processed by applying a filter and those that were used in raw form. The results with raw data and pretreatment were similar apart from *O. odorifera* that performed better with MSC. There was little confusion in four samples of *Ocotea odorifera* using raw spectra, which is natural and can be explained by some irregularity on the surface. SIMCA classification produced substantial confusion for raw data. The spectra of samples of *Ocotea porosa* classified as *Nectandra* sp. were not from the transverse face. In this band, the best result for SIMCA was achieved with preprocessing data, with MSC for *Nectandra* sp. and second derivative for the other species.

When the analysis was performed with wavelengths from 8494 to 3996 cm^{-1} (Table 9), most of the information about wood composition was present and the classification by ANN was similar for all spectra and MSC pretreatment resulted in all samples being correctly classified. In the case of the SIMCA classification, substantial confusion was still observed and the best classification with raw data was for *Ocotea porosa*. The spectra of samples of *Ocotea odorifera* classified as *Nectandra* sp. were not from the transverse face. Pretreatment increased the classification performance, with the best pretreatment for *Nectandra* sp. being MSC and second derivative for other species.

In near-infrared analysis of solid material, a higher number of samples are indicated because surface, shape and particle size can influence the results in discrimination of species (Brunner et al. 1996; Hein et al. 2010; Nisgoski et al. 2015). Literature reports on the efficient use of pretreatment with second derivative in wood discrimination (Sandak et al. 2011; Zhang et al. 2014; Horikawa et al. 2015; Hwang et al. 2016; Muñiz et al. 2016) and also the division of spectral range region from 4249 to 6100 cm^{-1} showed the distinction of wood species similar to mahogany (Pastore et al. 2011), from 4000 to 6200 cm^{-1} presented adequate results in discriminating six provenances of Sugi in south Brazil (Nisgoski et al. 2016) and 4000–5000 cm^{-1} plus 5500–6200 cm resulted in separation of wood and charcoal from "Angelim" Brazilian group (Muñiz et al. 2016).

Ma et al. (2012) present correct identification rates of 97–99% of wood samples with ANN and near-infrared spectra. Other studies on biological material discrimination with infrared showed the potential of ANN, but in mid-region (400–4000 cm^{-1}) and powder samples. In *Fusarium* species identification, Nie et al. (2007) trained ANN with first ten principal component scores from second derivative of FTIR spectra in each input layer and obtained a R^2 of 0.99. Further, in this study only PCA analysis failed because the main variations in the spectra were not related to variation between species. For *Ephedra* species discrimination, based on near-infrared spectra of powder samples, ANN presented 95% of prediction accuracy. When the analysis evaluated different habitats of *E. sinica* and samples collected at different times of day, ANN reached 100 and 93.3% of prediction accuracy, respectively (Fan et al. 2010).

Division of mid-spectra in different regions with more influence in species discrimination is also present. For *Campylobacter* species identification, Mouwen et al. (2006) obtained adequate results with a four-layer ANN for discrimination of genotypes. Also spectral windows between different wavenumbers in mid-infrared were applied to input layer for *Listeria* species discrimination (Rebuffo et al. 2006), and results showed 96% of accuracy in differentiation.

ANN also presented potential for application to species distinction based on anatomical features of wood (Esteban et al. 2009) and leaf images (Pandolfi et al. 2009).

Cell dimensions in input layer of a network resulted in 92% of probability of differentiation of two *Juniperus* species (Esteban et al. 2009). Based on leaf images of 17 accessions of *Camellia sinensis,* with different origin and varieties, Pandolfi et al. (2009) showed the potential of ANN and commented that the limitations are the same as of a human expert, which involves number and accuracy of training data: better train and learn when data present rich variation.

On the other hand, in SIMCA, misclassifications occurred with a greater frequency. It can be the result of the point (cell type) where spectra were acquired or irregularities on wood surface, since the samples were only sawn. Some examples of adequate results from SIMCA classification are in studies with red and white oak (Adedipe et al. 2008) and thermally modified wood of spruce, beech and ash (Bächle et al. 2012).

Even though a few numbers of species were studied, the objective of the study was achieved, ANN presented potential for species discrimination based on NIR spectra of solid wood specimens and can be a rapid and effective tool in forest commerce.

Conclusion

SIMCA classification is sensitive to surface orientation of the samples, resulting in considerable non-unique and non-classification, and is not recommended for discrimination of *Ocotea porosa, Ocotea odorifera, Nectandra* sp. and *Eucalyptus* sp. species. MSC and second derivative pretreatment had a good influence on SIMCA classification. The disadvantage of the use of filters is due to the fact that their choice is generally empirical, making the classification process of species slower. In this aspect, ANN becomes more advantageous due to its large sensibility perception of changes in absorbance between species without the need for prior treatment data. Considering the large amount of data and low linearity between them, the ANN demonstrated good potential for species discrimination based on near-infrared spectral analysis. In this study, the best performance came from an artificial neural network using the tan–sigmoid transfer function and eight neurons in the hidden layer. For ANN analysis, the use of the spectral range from 4000 to

10,000 cm^{-1} is recommended. In this case, the neural network developed did not result in any identification error for a margin of ±2%.

References

Adedipe OE, Dawsin-Andoh AB, Slahor J, Osborn AL (2008) Classification of red oak (*Quercus rubra*) and white oak (*Quercus alba*) wood using a near infrared spectrometer and soft independent modelling of class analogies. J Near Infrared Spectrosc 16(1):49–57

Bächle H, Zimmer B, Wegener G (2012) Classification of thermally modified wood by FT-NIR spectroscopy and SIMCA. Wood Sci Technol 46(6):1181–1192

Banerjee AK, Kiran K, Murty USN, Venkateswarlu C (2008) Classification and identification of mosquito species using artificial neural networks. Comput Biol Chem 32(6):442–447

Braga JWB, Pastore TCM, Coradin VTR, Camargos JAA, Silva ARD (2011) The use of near infrared spectroscopy to identify solid wood specimens of *Swietenia macrophylla* (CITES appendix II). IAWA J 32(2):285–296

Brunner M, Eugster R, Trenka E, Bergamin-Strotz L (1996) FT-NIR spectroscopy and wood identification. Holzforschung 50(2):130–134

Casale M, Schimleck LR, Espeyd C (2010) Classification of pernambuco (*Caesalpinia echinata* Lam.) wood quality by near infrared spectroscopy and linear discriminant analysis. J Near Infrared Spectrosc 18(6):435–442

Clark JY (2003) Artificial neural networks for species identification by taxonomists. Biosystems 72(1–2):131–147

Esteban LG, Fernández FG, Palacios PP, Romero RM, Cano NN (2009) Artificial neural networks in wood identification: the case of two *Juniperus* species from the Canary Islands. IAWA J 30(1):87–94

Fan Q, Wanga Y, Sun P, Liu S, Li Y (2010) Discrimination of *Ephedra* plants with diffuse reflectance FT-NIRS and multivariate analysis. Talanta 80:1245–1250

Hein PRG, Lima JT, Chaix G (2010) Effects of sample preparation on NIR spectroscopic estimation of chemical properties of *Eucalyptus urophylla* S.T. Blake wood. Holzforschung 64:45–54

Horikawa Y, Tazuru SM, Sugiyama J (2015) Near-infrared spectroscopy as a potential method for identification of anatomically similar Japanese diploxylons. J Wood Sci 61:251–261

Hwang SW, Horikawa Y, Lee WH, Sugiyama J (2016) Identification of Pinus species related to historic architecture in Korea using NIR chemometric approaches. J Wood Sci 62:156–167

Kang SH, Song SH, Lee SH (2012) Identification of butterfly species with a single neural network system. J Asia Pac Entomol 15(3):431–435

Kattmah G, Azim GA (2013) Fig (*Ficus carica* L.) identification based on mutual information and neural networks. Int J Imag Graph Signal Proc 5(9):50–57

Labati RD, Gamassi M, Piuri V, Scotti F (2009) A low-cost neural-based approach for wood types classification. In: CIMSA 2009—International conference on computational intelligence for measurement systems and applications, Hong Kong, China. doi:10.1109/CIMSA.2009.5069947

Lorenz C, Ferraudo AS, Suesdek L (2015) Artificial neural network applied as a methodology of mosquito species. Acta Trop 152:165–169

Ma MY, Wang GY, Huang AM, Zhang ZY, Xiang YH, Gu X (2012) Study on artificial neural network combined with near infrared spectroscopy for wood species identification. Spectrosc Spectr Anal 32(9):2377–2381

MMA (2008) Ministry of Environment. Normative Instruction 06, September 23, 2008. Brasília, Ministry of Environment. (**In Portuguese**)

Mouwen DJM, Capita R, Alonso-Calleja C, Prieto-Gómez J, Prieto M (2006) Artificial neural network based identification of *Campylobacter* species by Fourier transform infrared spectroscopy. J Microbiol Methods 67:131–140

Mu H, Zhang M, Qi D, Ni H (2015) The application of RBF neural network in the wood defect detection. Int J Hybrid Inf Technol 8(2):41–50

Muñiz GIB, Carneiro ME, Batista FRR, Schardosin FZ, Nisgoski S (2016) Wood and charcoal identification of five species from the miscellaneous group known in Brazil as “angelim” by near-ir and wood anatomy. Mad Ciencia y Tecnol 18(3):505–522
Nie M, Zhang WQ, Xiao M, Luo JL, Bao K, Chen JK, Li B (2007) FT-IR spectroscopy and artificial neural network identification of *Fusarium* species. J Phytopathology 155:364–367
Nisgoski S, Carneiro ME, Muñiz GIB (2015) Influencia de la granulometría de la muestra en la discriminación de especies de *Salix* por infrarrojo cercano (Influence of sample granulometry on discrimination of Salix species by near infrared). Mad Ciencia y Tecnol 17(1):195–204
Nisgoski S, Schardosin FZ, Batista FRR, Muñiz GIB, Carneiro ME (2016) Potential use of NIR spectroscopy to identify *Cryptomeria japonica* varieties from southern Brazil. Wood Sci Technol 50(1):71–80
Okan OT, Deniz I, Tiryaki S (2015) Application of artificial neural networks for predicting tensile index and brightness in bleaching pulp. Mad Ciencia y Tecnol 17(3):571–584
Ozsahin S (2012) The use of an artificial neural network for modeling the moisture absorption and thickness swelling of oriented strand board. BioResources 7(1):1053–1067
Packianather MS, Drake PR (2000) Neural networks for classifying images of wood veneer. Part 2. Int J Adv Manuf Technol 16(6):424–433
Pandolfi C, Mugnai S, Azzarello E, Bergamasco S, Mais E, Mancuso S (2009) Artificial neural networks as a tool for plant identification: a case study on Vietnamese tea accessions. Euphytica 166:411–421
Pastore TCM, Braga JWB, Coradin VTR, Magalhães WLE, Okino EYA, Camargos JAA, De Muñiz GIB, Bressan OA, Davrieux F (2011) Near infrared spectroscopy (NIRS) as a potential tool for monitoring trade of similar woods: discrimination of true mahogany, cedar, andiroba and curupixá. Holzforschung 65(1):73–80
Pham DT, Soroka AJ, Ghanbarzadeh A, Koc E, Otri S, Packianather M (2006) Optimising neural networks for identification of wood defects using the bees algorithm. In: 2006 IEEE International conference on industrial informatics, doi:1-4244-9701-0/06/$20.00 c_ 2006 IEEE
Rebuffo CA, Schmitt J, Wenning M, von Stetten F, Scherer S (2006) Reliable and rapid identification of *Listeria monocytogenes* and *Listeria* species by artificial neural network-based fourier transform infrared spectroscopy. Appl Environ Microbiol 72(2):994–1000
Russ A, Firesova M, Gigac J (2009) Preliminary study of wood species identification by NIR spectroscopy. Wood Res 54(4):23–32
Sandak A, Sandak J, Negri M (2011) Relationship between near-infrared (NIR) spectra and the geographical provenance of timber. Wood Sci Technol 45(1):35–48
Schwanninger M, Rodrigues JC, Fackler K (2011) A review of band assignments in near infrared spectra of wood and wood components. J Near Infrared Spectrosc 19(5):287–308
Simmonds EJ, Armstrong F, Copland PJ (1996) Species identification using wideband backscatter with neural network and discriminant analysis. ICES J Mar Sci 53:189–195
Sundaram M, Abitha J, Raj MM, Ramar K (2015) Wood species classification based on local edge distributions. Optik 126(21):2884–2890
Sussillo D, Barak O (2013) Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. Neural Comput 3(25):626–649
Tsuchikawa S, Siesler HW (2003) Near-infrared spectroscopy monitoring of the diffusion process of deuterium-labeled molecules in wood. Part I. softwood. Appl Spectrosc 57(6):667–674
Yonenobu H, Tsuchikawa S (2003) Near-infrared spectroscopic comparison of antique and modern wood. Appl Spectrosc 57(11):1451–1453
Zbiéc M (2011) Application of neural network in simple tool wear monitoring and identification system in MDF milling. Drvna Ind 62(1):43–54
Zhang W, Teng G, Wang C (2013) Identification of jujube trees diseases using neural network. Optik 124(11):1034–1037
Zhang X, Yu H, Li B, Li WJ, Li X, Bao C (2014) Discrimination of *Pinus yunnanensis*, *P. kesiya* and *P. densata* by FT-NIR. J Chem Pharm Res 6(4):142–149