Prediction of Five Softwood Paper Properties from its Density using Support Vector Machine Regression Techniques

Esperanza García-Gonzalo,[a] António J. A. Santos,[b,c] Javier Martínez-Torres,[d] Helena Pereira,[b] Rogério Simões,[c] Paulino José García-Nieto,[a] and Ofélia Anjos [b,e,*]

Predicting paper properties based on a limited number of measured variables can be an important tool for the industry. Mathematical models were developed to predict mechanical and optical properties from the corresponding paper density for some softwood papers using support vector machine regression with the Radial Basis Function Kernel. A dataset of different properties of paper handsheets produced from pulps of pine (*Pinus pinaster* and *P. sylvestris*) and cypress species (*Cupressus lusitanica*, *C. sempervirens*, and *C. arizonica*) beaten at 1000, 4000, and 7000 revolutions was used. The results show that it is possible to obtain good models (with high coefficient of determination) with two variables: the numerical variable density and the categorical variable species.

Keywords: Vector machine regression; Paper properties; Kraft pulp; Pinus pinaster; Pinus sylvestris; Cupressus lusitanica; Cupressus sempervirens; Cupressus arizonica

Contact information: a: Facultad de Ciencias, Departamento de Matemáticas, Universidad de Oviedo, 33007, Oviedo, Spain; b: Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal; c: Fiber Material ad Environmental Technologies Unit, Universidade da Beira Interior, 6201-001 Covilhã, Portugal; d: Centro Universitario de la Defensa, Academia General Militar, 50090 Zaragoza, Spain; e: Instituto Politécnico de Castelo Branco, Escola Superior Agrária, Apartado 119, 6001-909 Castelo Branco, Portugal;
** Corresponding author: ofelia@ipcb.pt*

INTRODUCTION

Different types of wood fiber raw materials used for paper production may differently impact some of the properties and improve the performance of the final product and consequently allow different end uses (Niskanen *et al.* 1988; Biermann 1996; Santos *et al.* 2012). For a selected raw material, the paper properties are strongly dependent on the pulp refining, which affects the inter-fiber bonding (Wang *et al*. 2003; Gharehkhani *et al*. 2015), and on the individual fiber strength.

Modeling properties is an important step to predict product performance and may help in the management and operating conditions of industrial processes. For instance, paper properties can be predicted using a few, or only one, predictor variables, which is of high practical importance because of the difficulty of determining all paper properties. This possibility has been tested for some hardwood species (*Eucalyptus globulus, Acacia dealbata*, and *A. melanoxylon*) to predict paper properties using paper density as the predictor variable by applying unsupervised classification techniques and multivariable regression techniques (Anjos *et al.* 2015). It would be interesting to know if a similar approach also applies to softwood papers.

Compared with hardwood species, softwood species provide longer and wider pulp fibers. Long fibers can develop more inter-fiber bonds per fiber, which decreases the tensile stress per inter-fiber bond and in consequence enables the production of paper materials with higher mechanical properties, *e.g.*, tensile and tear strength (Niskanen *et al*. 1988). Tear strength is particularly sensitive to fiber length. Long fibers are used in paper materials where mechanical strength is very important, as is the case for sack kraft papers, but they can also be used as reinforcement fibers in other paper grades, such as printing and writing papers, to provide the required wet and dry strength to the paper web (Santos *et al.* 2008; Anjos *et al.* 2011).

Bleached hardwood kraft pulps have a strong market position for printing and writing papers because of their strength, bulk, opacity, and smoothness (Kibblewhite *et al*. 1991). Softwood fibers are principally used in the manufacture of printing and writing grades for their reinforcing properties. Leopold and Thorpe (1968) and Zeng *et al*. (2013) reported an increase in strength for Nordic softwood species through a lower fibril angle and less wall weak points given by kinks and nodes. Therefore, it can be concluded that their superior performance regarding tensile strength is also supported by the higher intrinsic fiber strength of these species.

Support vector regression has been incorporated for solving prediction problems related to wood properties (Mora and Schimleck 2010; Zhang *et al.* 2011; Nascimbem *et al*. 2013) because it can obtain data-driven models that do not need an explicit regression function and because it is able to work with high-dimensional data. As a data-driven methodology, it "learns" from training data and creates a model, and when given new data, it is able to predict the dependent quantity. As in the case of other learning machine methods, the model is not represented by an equation or group of equations, and it works as a black box once the model is created. Since it supports kernels, it can model nonlinear relationships. The regularization parameter provides robustness to the method.

The aim of this study is to build mathematical models using vector machine regression with the Radial Basis Function (RBF) Kernel to predict mechanical and optical properties from the corresponding paper density for some softwood-based papers, namely those made from *Pinus pinaster*, *P. sylvestris*, C*upressus sempervirens*, *C. lusitanica*, and *C. arizonica*. While both pine species are well established as softwood pulp species, the cypress species showed potential to be incorporated into papers with good light scattering/tensile strength and smoothness/tensile strength relationships, although they produce in general lower performing papers in comparison to pine woods, given their lower fiber length and coarseness and higher number of fibers per gram (Esteves *et al.* 2004; Anjos *et al*. 2014; Santos *et al*. 2014).

EXPERIMENTAL

Data

Five softwood species were used: *Pinus pinaster*, *P. sylvestris*, *Cupressus lusitanica, C. sempervirens*, and *C. arizonica*. Pulps were produced that were refined at three refining levels, and paper sheets were produced and analyzed, including a set with zero refining level. The data used for the modeling and the experimental conditions are presented in Anjos *et al*. (2014).

In brief, the wood chips (1000 g oven dry (o.d.) wood) were pulped with a kraft cooking process in a forced circulation digester under 25% effective alkali charge (as

NaOH); 30% sulphidity; 5/1 liquor/wood ratio; 170 °C pulping temperature; 90 min heating time; and 150 min time at temperature. The screened pulps were bleached with the D_0 E_1 D_1 E_2 D_2 (D = ClO_2 and E = extraction with NaOH) elemental chlorine-free sequence (Santos *et al*. 2008). The pulps were PFI beaten at 1000, 4000, and 7000 revolutions under a refining intensity of 3.33 N/mm, and paper handsheets were prepared. The Kappa number ranged from 21.4 to 32.0 for all the pulps and the ISO brightness of the bleached kraft pulps ranged from 65.0 to 85.0%

The measured paper properties and respective standards use for each one were: density (Dens), TAPPI T220 sp-01; Bekk smoothness (Smoo), TAPPI T479 cm-09; tensile index (Tens), ISO 1924-2; stretch (Stre), ISO 1924-2; burst index (Burs), TAPPI T403 om-10; tear index (Tear), TAPPI T220 sp-96; zero-span tensile strength in dry (Zssd) and wet (Zssw) handsheet samples, TAPPI T273 pm-95; opacity (Opac), ISO 2471; brightness (Brig), ISO 2470; and light scattering coefficient (Ligh), ISO 2469. The dataset of the physical properties of handsheets for each species (n=40) is summarized in Table 1 for pine and Table 2 for cypress species.

Table 1. Physical Properties of Handsheets for Pine Species (n=40 for each species)

	P. pinaster		***P. sylvestris***	
	μ±σ	Max-min	μ±σ	Max-min
Dens (g/cm^3)	0.68±0.11	0.80-0.46	0.67±0.13	0.80-0.43
Smoo (Bekk's)	70±61	199-1	73±57	155-6
Tens (N.m/g)	68.3±31.4	118.2-15.0	57.1±24.7	85.3-17.8
Stre (%)	3.3±1.0	5.0-1.6	3.5±0.8	5.0-2.5
Burs (kPa.m^2/g)	3.5±1.9	5.9-0.5	3.5±2.0	5.9-0.4
Tear (mN.m^2/g)	14.3±2.9	19.1-7.8	14.0±2.7	18.2-8.7
Zssd (N.m/g)	144.4±13.1	166.1-119.9	174.5±9.5	191.7-151.0
Zssw (N.m/g)	115.0±6.0	125.3-105.7	88.3±24.7	160.2-65.1
Opac (%)	77.4±3.7	81.9-72.2	80.3±3.1	83.8-75.7
Brig (%)	64.2±8.7	76.4-53.7	68.0±7.7	80.1-53.6
Ligh (m^2/kg)	21.6±7.7	33.9-13.6	25.0±8.2	39.2-16.4

Table 2. Physical Properties of Handsheets for Cypress Species (n=40 for each species)

	C. lusitanica		***C. sempervirens***		***C. arizonica***	
	μ±σ	Max-min	μ±σ	Max-min	μ±σ	Max-min
Dens (g/cm^3)	0.91±0.14	1.04-0.64	0.82±0.13	0.95-0.57	0.88±0.15	1.02-0.53
Smoo (Bekk's)	255±121	431-76	173±95	338-26	200±107	433-43
Tens (N.m/g)	63.9±22.0	93.5-25.7	64.4±25.4	91.4-20.7	58.2±19.5	82.8-23.6
Stre (%)	6.6±1.2	8.0-3.6	5.2±0.8	6.8-3.5	7.1±1.2	9.1-4.3
Burs (kPa.m^2/g)	3.5±1.9	5.9-0.4	3.4±2.0	5.9-0.4	3.3±2.0	5.9-0.4
Tear (mN.m^2/g)	11.1±1.7	14.3-6.1	11.6±1.8	14.9-9.0	10.4±2.2	14.8-5.2
Zssd (N.m/g)	131.7±10.6	151.8-100.7	137.5±16.5	161.9-109.8	115.2±13.3	135.6-91.9
Zssw (N.m/g)	86±12	105-70	91.3±8.6	107.2-75.1	74.1±7.3	95.8-60.2
Opac (%)	77.2±5.6	84.0-68.6	74.2±4.5	80.3-67.9	76.0±4.9	83.3-68.4
Brig (%)	69.1±11.2	86.7-53.7	71.2±10.6	87.5-58.4	65.7±13.3	85.6-50.2
Ligh (m^2/kg)	29.5±16.5	59.1-11.6	27.3±13.3	50.7-13.5	24.9±14.4	49.4-10.8

Data Processing

In this study, an analysis of the data set was performed by using Principal Component Analysis (PCA). The PCA is a technique used to reduce the dimensionality of a data set by selecting the dimensions having the largest variances. This technique is used to find the causes of variability in a data set and to sort them by importance. The first principal component accounts for a percentage of the total variance in the same proportion as the first eigenvalue of the PCA analysis with respect to the sum of all the eigenvalues. Then, the procedure is continued similarly for the other principal components.

The PCA analysis was conducted using the PCA MATLAB function princomp with previously normalized data (for each variable, the mean was subtracted and the result divided by the standard deviation).

The prediction of paper properties requires several steps: 1) choosing a model; 2) optimizing the model parameters; and 3) obtaining property predictions. The first step is to select a combination of input properties that affect the output property.

If $\mathbf{x} = (x_1, x_2, \ldots, x_k)$ is a vector that comprises the input variables, k being the number of such variables, and if we have n observations,

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n), \tag{1}$$

where y is the output property or the variable to predict, the problem arises, according to the theory of support vector machines (SVM) (Vapnik 1999), as

(a) finding a function,

$$f(\mathbf{x}) = \boldsymbol{\omega}^T \phi(\mathbf{x}) + b \tag{2}$$

and (b) using this function to fit the training data, where $\boldsymbol{\omega}$ is the vector that contains the so-called weights that affect each predictor, b is a real number, and ϕ is a non-linear mapping function.

A non-linear problem can be mapped into a higher dimension space, using an inner product (kernel), where a linear regression can be performed. The SVM theory affirms that the solution to problem (2) is the same as the solution of the equation,

$$\begin{aligned} &\min_{\omega,b,\xi} \left(-\boldsymbol{\omega}^T \boldsymbol{\omega} + C \sum_{i=1}^{n} \xi_i \right) \\ &\textit{subject to} \quad y_i \left(\boldsymbol{\omega}^T \phi(\mathbf{x}_i) + b \right) \geq 1 - \xi_i \\ &\qquad \xi_i \geq 0 \\ &\qquad (i = 1, 2, \ldots, n) \end{aligned} \tag{3}$$

where ξ_i is the error between observed and predicted values, that is,

$$\max \left\{ 0, | y_i - f(\mathbf{x}_i) | - \varepsilon \right\} \tag{4}$$

with $\varepsilon > 0$ determines an insensitivity zone around the fitted model where error is not taken into account; and C is the regularization or penalty parameter that weights the error in the function that is minimized, that is to say, the parameter C controls the trade-off between the margin and the size of the slack variables.

Thus, $C\sum_{i=1}^{n}\xi_i$ values are the losses on the training set.

Once C, ε, and the parameters of the kernel have been selected, this problem has a unique solution. Different parameters will give different solutions or models; therefore, the parameters must be tuned to optimize the model. Different optimization methods can be used. In this paper, we used particle swarm optimization (PSO) (Kennedy and Eberhart 1995) for tuning SVM with RBF kernel parameter, σ.

The input variables $\mathbf{x} = (x_1, x_2, \ldots, x_k)$ were density and species. This last variable is represented as a dummy explanatory variable. There were 40 (n) observations for each case. An independent study was conducted for each of the independent variables y: Bekk smoothness, tensile index, stretch, burst index, tear index, zero-span tensile strength in dry; and wet, handsheet samples, opacity, brightness, and light scattering coefficient. The first part of the study was the selection of the model, *i.e.*, a search of the optimal parameters for the SVM model was performed using the PSO algorithm. For this, we began with 20 different random parameter sets of parameter (C,ε,σ). The cross-validation coefficient of determination of the SVM with RBF kernel model was found for each set. Thus, a nonlinear model was built. Based on these cross-validation values, the PSO proposes 20 new sets of models. After a number of iterations when the ending criteria are met, the best set, *i.e.* the one that achieves the model with the highest cross-validation coefficient of determination, is chosen as the optimal model. Then we proceeded to construct the model and predict the values for the output variables (smoothness, tensile index, and so on).

RESULTS AND DISCUSSION

Tables 1 and 2 show that the dataset contained a high variation and range for the values of the paper properties, which is usually needed for construction of good models.

Principal Component Analysis

All the variables except tear index (Tear) and zero-span tensile strength in dry (Zssd) and wet (Zssw) presented a medium to high correlation coefficient with paper density (Dens). Additionally, the variables tear index (Tear) and zero-span tensile strength in dry (Zssd) and wet (Zssw) samples were weakly correlated with the other variables except with burst index (Burs) variable.

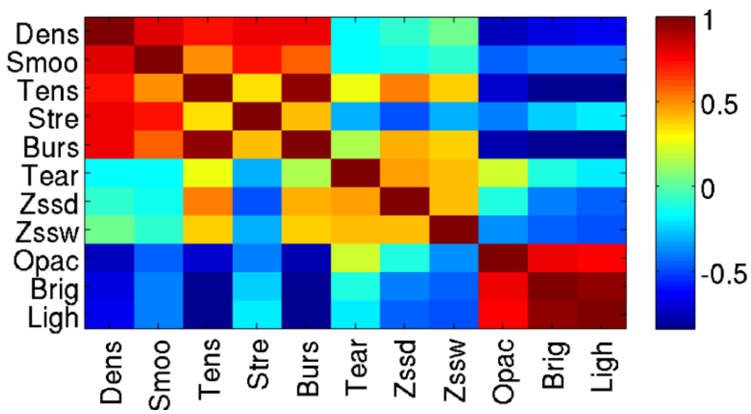


Fig. 1. Partial correlation coefficients between variables (scale on the right)

The correlation matrix for all the samples (Fig. 1) shows a close relationship between the variables opacity (Opac), brightness (Brig), and light scattering coefficient (Ligh), with a high correlation index with the variables tensile index (Tens) and burst index (Burs) and a good, but not so high, correlation with density. The variables Dens and Smoo were highly correlated.

Better correlations were found when the species were analyzed separately. This is in accordance with the PCA results, where the two first components explain 82% of the total variation (Fig. 2). Similar findings have been reported by Anjos *et al*. (2015).

Variables Opac, Brig, and Ligh were clustered together in the positive region of the first component axis, with almost zero second principal components, which means a high correlation between the variables of this cluster. Burs and Tens presented a similar situation, being highly correlated, and with zero second principal components, thus they are expressed in terms of the first principal component. Variables from both clusters were highly correlated but with negative sign because they were in a symmetric position with respect to the y-axis. Dens clustered with Smoo and Stre. The variables Tear, Zssd, and Zssw clustered together, thus Smoo and Stre were strongly related and similarly the members of the other cluster.

The inverse relationship between density and light scattering coefficient has been found previously (Batchelor and He 2005; Hubbe *et al*. 2008; Anjos *et al*. 2015). Some authors have also reported a trend of increasing tensile strength with increasing paper density and of decreasing tensile strength with decreasing fiber strength (Seth and Kingsland 1990; Santos *et al*. 2006; Vainio and Paulapuro 2007; Anjos *et al*. 2014).

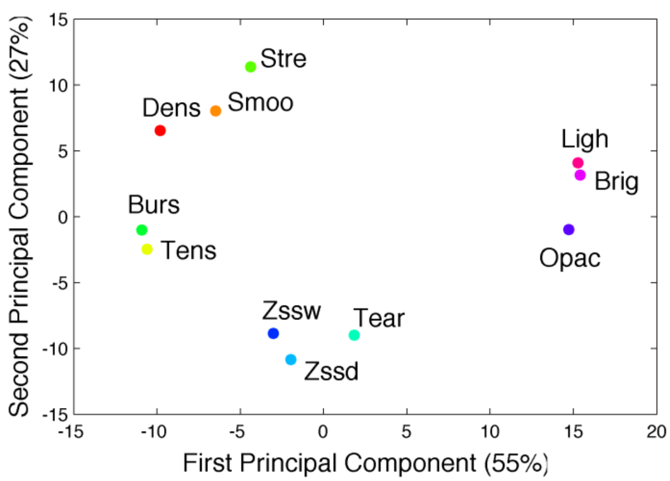


Fig. 2. Principal component analysis for the paper properties of the five species with four pulp refining levels

In Fig. 2 the data of both the different refining levels and the different species are merged together, and a role of the refining levels is not revealed. Figure 3 represents the same data regarding the paper properties but revealing the effect of the refining for the different species. The PCA indicates that the first principal component separates the samples by their refining level: the zero level corresponds to the clusters of the left, the first level to the clusters in the middle and the second and third refining level clusters are quite near and appear as very close clusters at the right side in Fig. 3. The second principal component is closely related with the different species. Refining improves internal fibrillation and swelling (Garcia *et al*. 2002), enhancing fiber flexibility, collapsibility, and fiber-fiber bonds (Fardim and Duran 2003). The extent and intensity of these behaviors are species-dependent and also depend on the pulp fiber characteristics (Paavilainen 1993; Santos *et al*. 2008; Anjos *et al*. 2011, 2014).

The results show that the variable species should be one of the predictive variables. Moreover, higher differences were observed when comparing *Cypress* and *Pinus* species, in agreement with previous results (Anjos *et al*. 2014; Santos *et al*. 2014).

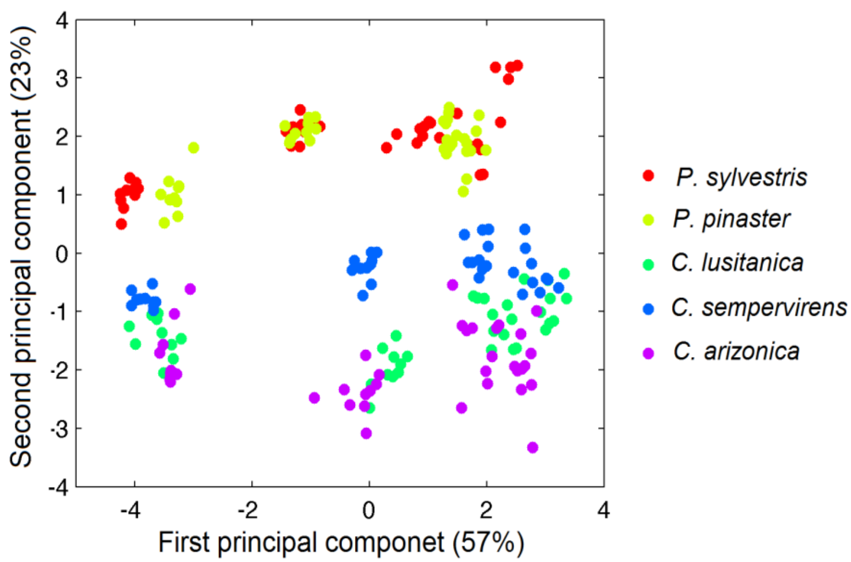


Fig. 3. Principal component analysis for the normalized data of the paper properties of *P. pinaster*, *P. sylvestris*, *C. lusitanica*, *C. sempervirens*, and *C. arizonica* for the four levels of refining

Support Vector Machine (SVM) Regression

A support vector machine regression using the LIBSVM library (Chang and Lin 2011) with the RBF kernel was performed with its MATLAB code. With the linear model it was also possible to obtain good models with only two variables: the numerical variable density and the categorical variable species. MATLAB function Linear Model.fit was used for this task. However, for the higher refining levels, a lower predictability was found. Given this result, two independent variables were used to create regression models for the dependent variables Smoo, Tens, Stre, Burs, Tear, Zssd, Zssw, Opac, Brig, and Ligh.

The parameters for the SVM with a RBF Kernel are C, ε, and σ. Maximum values of 1 and 10 for σ and C, respectively, were used to avoid oscillations in the predicted regression function. The parameters were tuned with the algorithm particle swarm optimization (PSO), with the criterion that the mean R^2 for a cross-validation with 10 folds was the maximum. The chosen version was the 2011 Standard PSO (Clerc 2012) with MATLAB code. The search space parameter limits are represented in Table 3.

Table 3. Search Space Parameter Limits for PSO while Tuning SVM Parameters with a RBF Kernel Model

Parameter	Lower boundary	Upper boundary
C	10^{-4}	10^{1}
σ	10^{-4}	10^{0}
ε	10^{-6}	10^{0}

The stopping criteria for the PSO algorithm was 20 iterations without improvement in a R^2 value with six decimals, with a maximum number of 1000 iterations. The parameters obtained for each RBF-SVM model, with their corresponding R^2 cross-validation values, are presented in Table 4. The cross-validation R^2 was also obtained for 5, 15, and 20 folds for the same parameters to check the robustness of the models. It can be observed that there is little variation related with the number of cross-validation folds. For all the variables but one, Tear, cross-validation R^2 higher than 0.8 were obtained, and higher than 0.9 for most of the independent variables.

Table 4. RBF-SVM Model Parameters and their corresponding R^2 Cross-Validation Value

				R^2 cross-validation			
Variable	***C***	**σ**	**ε**	**5 fold**	**10 fold**	**15 fold**	**20 fold**
Smoo	9.797	0.997	0.0477	0.847	0.854	0.855	0.854
Tens	4.876	0.599	0.0396	0.950	0.951	0.952	0.952
Stre	3.005	0.801	0.0945	0.903	0.913	0.914	0.914
Burs	8.629	0.696	0.0237	0.957	0.960	0.959	0.959
Tear	8.723	1.000	0.0542	0.704	0.694	0.703	0.701
Zssd	1.546	0.996	0.0000	0.921	0.924	0.922	0.923
Zssw	6.431	0.743	0.0550	0.937	0.939	0.936	0.936
Opac	9.999	1.000	0.0368	0.933	0.937	0.937	0.940
Brig	7.961	1.000	0.0217	0.957	0.955	0.955	0.957
Ligh	4.484	0.985	0.0171	0.980	0.980	0.980	0.979

For comparative purposes, a linear model was also created for each variable. We began predicting a linear model from each variable using as predictors the qualitative variable Species, together with the numerical variables Smoo and Dens. It emerged that good models, with coefficients of determination over 0.8, could be obtained using only the variables Species and Dens.

The comparison of both models shows that SVM models noticeably improved the obtained R^2 in relation to the linear models for all the variables (Table 5) from a mere 10% decrease in the RMSE for variable Stre to almost 100% for the variable Tear. Nevertheless, a good model was not achieved for the variable Tear because this parameter is strongly related to the fiber length and inter-fiber bonding (Anjos *et al*. 2011), which was not measured in this work.

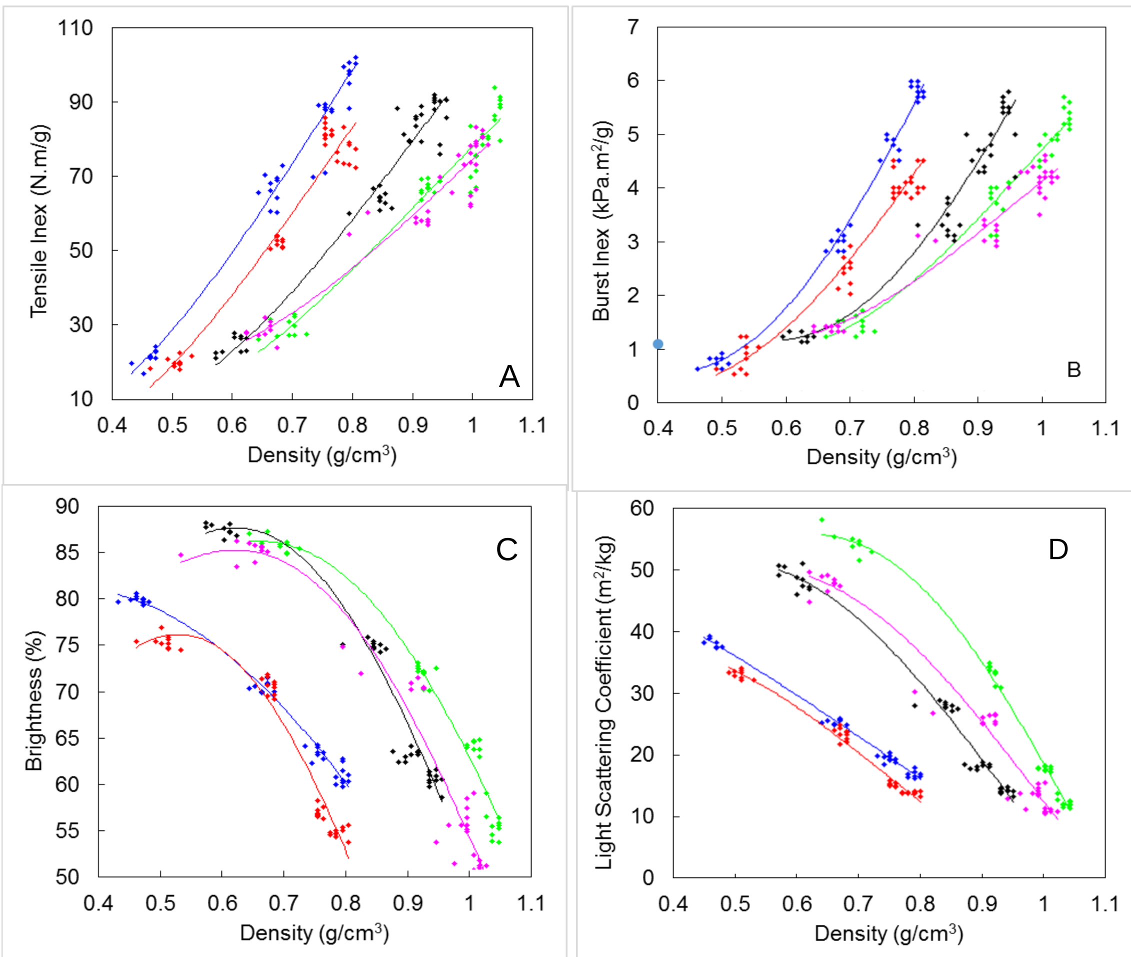
Table 5. Comparison of the R^2 obtained with SVM Models and Linear Models

Variable	**R^2 SVM**	**R^2 lin. mod.**	**RMSE SVM**	**RMSE lin.mod.**	**% decrement RMSE**
Smoo	0.854	0.807	$4.0437\ 10^{1}$	$5.1410\ 10^{1}$	21%
Tens	0.951	0.916	$5.1561\ 10^{0}$	$7.1214\ 10^{0}$	26%
Stre	0.913	0.905	$5.1885\ 10^{-1}$	$5.7747\ 10^{-1}$	10%
Burs	0.960	0.910	$3.0341\ 10^{-2}$	$4.7785\ 10^{-2}$	37%
Tear	0.694	0.633	$3.0341\ 10^{-2}$	$1.9042\ 10^{0}$	98%
Zssd	0.924	0.903	$6.1613\ 10^{0}$	$7.3251\ 10^{0}$	16%
Zssw	0.939	0.904	$3.9496\ 10^{0}$	$5.2626\ 10^{0}$	25%
Opac	0.937	0.807	$1.1132\ 10^{0}$	$2.1426\ 10^{0}$	48%
Brig	0.955	0.898	$2.0691\ 10^{0}$	$3.4690\ 10^{0}$	40%
Ligh	0.980	0.932	$1.6792\ 10^{0}$	$3.2781\ 10^{0}$	49%

With the resulting SVM models, the predicted values of the dependent variable for the density of each species were calculated.

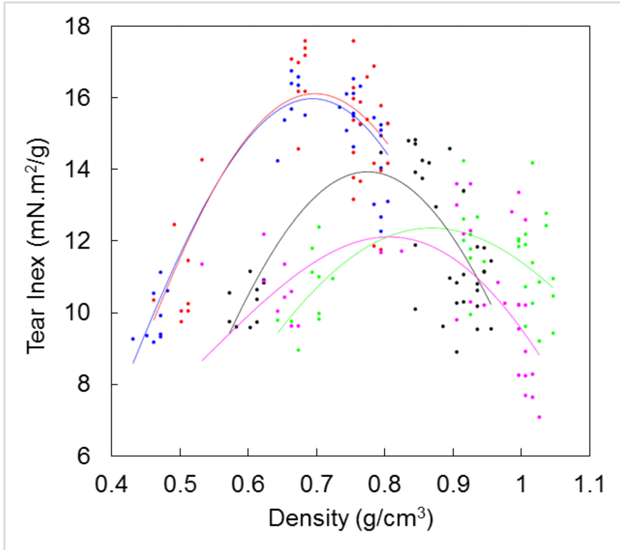
The regression with SVM creates a data-driven model and thus does not need any assumption for the regression function. The model relies on data to find patterns that can be extended to a wider range of data, *i.e.*, it maps the input-output relationship in the observed data, expecting that this learned mapping allows the prediction of outputs from new input data.

Some examples of the graphical representation of these models are presented in Fig. 4. Figure 5 shows the graphical representation for the worst model.



P. sylvestris: ● observed, — predict; *P. pinaster*: ● observed, — predict;
C. lusitanica: ● observed, — predict; *C. sempervirens*: ● observed, — predict;
C. arizonica: ● observed, — predict

Fig. 4. Fitted curves for density *vs*. (A) tensile index, (B) burst index, (C) brightness, and (D) light scattering coefficient as the dependent variables



P. sylvestris: ● observed, — predict; *P. pinaster*: ● observed, — predict; *C. lusitanica*: ● observed, — predict; *C. sempervirens*: ● observed, — predict; *C. arizonica*: ● observed, — predict

Fig. 5. Fitted curves for density *vs*. tear index as the dependent variable

CONCLUSIONS

1. It is possible to use support vector machine regression with Radial Basis Function kernel techniques to establish prediction models for some paper properties of softwood pulps based on their paper density.
2. Principal component analysis of paper handsheet properties showed that the pulp refining level strongly affects them, as does the species.
3. Tear index could not be predicted with sufficient accuracy, probably because tear is a function of fiber length, which was not measured.
4. PCA showed that Opac, Brig, and Ligh, Opac properties are related between them, as well as Burs and Tens properties.
5. SVM models noticeably improved the obtained R^2 in relation to the linear models for all the variables from a mere 10% decrease in the RMSE for variable Stre to almost 100% for the variable Tear.

ACKNOWLEDGMENTS

Centro de Estudos Florestais is a research unit funded by Fundação para a Ciência e a Tecnologia (Portugal) within UID/AGR/00239/2013. The research unit Fiber Materials and Environmental Technologies is funded by Fundação para a Ciência e a Tecnologia (Portugal) within UID/Multi/00195/2013.

REFERENCES CITED

Anjos, O., García-Gonzalo, E., Santos, A. J. A., Simões, R., Martínez-Torres, J., Pereira, H., and García-Nieto, P. J. (2015). "Using apparent density of paper from hardwood kraft pulps to predict sheet properties, based on unsupervised classification and multivariable regression techniques," *BioResources* 10(3), 5920-5931. DOI: 10.15376/biores.10.3.5920-5931.

Anjos, O., Santos, A. J. A., and Simões, R. (2011). "Effect of *Acacia melanoxylon* fibre morphology on papermaking potential," *Appita J.* 64(2), 185-191.

Anjos, O., Santos, A. J. A., Simões, R., and Pereira, H. (2014). "Morphological, mechanical and optical properties of cypress," *Holzforschung* 68(8), 861-995. DOI: 10.1515/hf-2013-0125.

Batchelor, W., and He, J. (2005). "A new method for determination the relative bonded area," *TAPPI J.* 4(6), 23-38.

Biermann, C. J. (1996). *Handbook of Pulping and Papermaking*, 2nd ed., Academic Press, San Diego, CA.

Chang, C. C., and Lin, C. J. (2011). "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology* 2(3), 27. DOI: 10.1145/1961189.1961199.

Clerc, M. (2012). "Standard particle swarm optimisation. From 2006 to 2011," Technical Report, assessed on June 17, 2015. (http://clerc.maurice.free.fr/pso/SPSO_descriptions.pdf).

Esteves, S., Santos, A., Morais, C., Anjos, O., Simões, R., Pereira, H. (2004). "Chemical and anatomical characterization of Cypress wood for pulp production," ICECFOP1 – 1st International Conference on Environmentally – Compatible Forest Products. Porto, Portugal, pp. 427.

Fardim, P., and Duran, N. (2003). "Modification of fiber surface during pulping and refining as analyzed by SEM, XPS and ToF-SIMS," *Colloid. Surface. A* 223(2), 263-276. DOI: 10.1016/S0927-7757(03)00149-3.

Garcia, O., Torres, A., Colom, J., Pastor, F., Diaz, P., and Vidal, T. (2002). "Effect of cellulase-assisted refining on the properties of dried and never-dried eucalyptus pulp," *Cellulose* 9(2), 115-125. DOI: 10.1023/A:1020191622764.

Gharehkhani, S., Sadeghinezhad, E., Kazi, S. N., Yarmand, H., Badarudin, A., Safaei, M. R., and Zubir, M. N. M. (2105). "Basic effects of pulp refining on fiber properties - A review," *Carbohydr. Poly.* 115, 785-803. DOI: 10.1016/j.carbpol.2014.08.047.

Hubbe, M. A., Pawlak, J. J., and Koukoulas, A. A. (2008). "Paper's appearance: A review," *BioResources* 3(2), 627-665. DOI: 10.15376/biores.3.2.627-665.

ISO 1924-2 (1994). "Paper and board, Determination of tensile properties, Part 2: Constant rate of elongation method," International Organization for Standardization, Geneva, Switzerland.

ISO 2469 (2007). "Measurement of diffuse radiance (reflectance) factor," International Organization for Standardization, Geneva, Switzerland.

ISO 2470 (2008). "Measurement of diffuse blue reflectance factor (ISO brightness)," International Organization for Standardization, Geneva, Switzerland.

ISO 2471 (2008). "Determination of opacity (paper packing) – diffuse reflectance method," International Organization for Standardization, Geneva, Switzerland. Kennedy, J., and Eberhart, R. (1995). "Particle swarm optimization," in: *Proc. IEEE Inter. Conf. Neural Networ.* 4, 1942-1948.

Kibblewhite, R. P., Bawden, A. D., and Hughes, M. C. (1991). “Hardwood market kraft fibre and pulp qualities,” *Appita J*. 44(5), 325.
Leopold, B., and Thorpe, J. (1968). “Effect of pulping on strength properties of dry and wet pulp fibres from Norway spruce,” *TAPPI J*. 51(7), 304.
Mora, C. R., and Schimleck, L. R. (2010). “Kernel regression methods for the prediction of wood properties of *Pinus taeda* using near infrared spectroscopy,” *Wood Sci. Technol.* 44(4), 561-578. DOI: 10.1007/s00226-009-0299-5.
Nascimbem, L., Ribeiro Rubini, B., and Jesus Poppi, R. (2013). “Determination of quality parameters in moist wood chips by near infrared spectroscopy combining PLS-DA and support vector machines,” *J. Wood Chem. Technol.* 33(4), 247-257. DOI: 10.1080/02773813.2013.783075.
Niskanen, K., Kajanto, I., and Pakarinen, P. (1988). “Paper structure,” in: *Paper Physics*, K. Niskanen (ed.), Papermaking Science and Technology series, Fapet Oy, Helsinki, Finland.
Paavilainen, L. (1993). “Conformability, flexibility and collapsibility of sulphate pulp fibers,” *Paperi ja Puu* 75(9-10), 689-702.
Santos, A. J. A., Amaral, M., Vaz, A., Anjos, O., and Simões, R., (2008). “Effect of *Eucalyptus globulus* wood density on papermaking potential,” *TAPPI J.* 7(5), 25-32.
Santos, A. J. A., Anjos, O., and Simões, R. (2006). “Papermaking potential of *Acacia dealbata* and *Acacia melanoxylon*,” *Appita J*. 59(1), 58-64.
Santos, A. J. A., Anjos, O., Morais, M. C., Diogo, G., Simões, R., and Pereira, H. (2014). “Characterization of cypress wood for kraft pulp production,” *BioResources* 9(3), 4764-4774. DOI: 10.15376/biores.9.3.4764-4774.
Santos, A., Amaral, M., Gil, N., Anjos, O., Pereira, H., Simões, R. (2012). “Influence on pulping yield and pulp properties of wood density of *Acacia melanoxylon*,” *J. Wood Sci.* 58(6), 479-86. DOI: 10.1007/s10086-012-1286-2.
Seth, R., and Kingsland, M. (1990). “The reinforcing properties of softwood kraft pulps,” *Pulp Pap.-Canada* 91(7), 68-72, 74-75.
TAPPI T205 om-88 (1988). “Forming handsheets for physical tests of pulp,” TAPPI Press, Atlanta, GA.
TAPPI T220 sp-96 (1996). “Physical testing of pulp handsheets,” TAPPI Press, Atlanta.
TAPPI T220 sp-01 (2001). “Physical testing of pulp handsheets,” TAPPI Press, Atlanta.
TAPPI T273 pm-95 (1995). “Wet zero-span tensile strength of pulp, test method,” TAPPI Press, Atlanta, GA.
TAPPI T403 om-10 (2010). “Bursting strength of paper,” TAPPI Press, Atlanta, GA.
TAPPI T479 cm-09 (2009). “Smoothness of paper (Bekk Method),” TAPPI Press, Atlanta, GA.
Vainio, A., and Paulapuro, H. (2007). “Interfiber bonding and fiber segment activation in paper,” *BioResources* 2(3), 442-458. DOI: 10.15376/biores.2.3.442-458
Vapnik, V. N. (1999). “An overview of statistical learning theory,” *IEEE T. Neural Networ.* 5(10), 988-999. DOI: 10.1109/72.788640
Wang, X., Maloney, T. C., and Paulapuro, H. (2003). “Internal fibrillation in never dried and once-dried chemical pulps,” *Appita J*. 56(6), 455-459.
Zeng, X., Vishtal, A., Retulainen, E., Sivonen, E., and Fu, S. (2013). “The elongation potential of paper. How should fibres be deformed to make paper extensible?,” *BioResources* 8(1), 472-486. DOI: 10.15376/biores.8.1.472-486

Zhang, Y., Li, Y., Zhang, H., Zhang, H., and Li, P. (2011). "Modeling wood density of larch by near-infrared spectrometry and support vector machine," *Appl. Mech. Mater*. 55-57, 433-438. DOI: 10.4028/www.scientific.net/AMM.55-57.433

Article submitted: August 4, 2015; Peer review completed: November 22; Revised version received and accepted: December 5, 2015; Published: January 11, 2016.
DOI: 10.15376/biores.11.1.1892-1904