



Functional statistical techniques applied to vine leaf water content determination

C. Ordoñez^a, J. Martínez^{a,*}, J.M. Matías^b, A.N. Reyes^a, J.R. Rodríguez-Pérez^c

^a Department of Environmental Engineering, University of Vigo, 36310 Vigo, Spain

^b Department of Statistics, University of Vigo, 36310 Vigo, Spain

^c Department of Cartography, University of Leon, 22400 Ponferrada, Spain

ARTICLE INFO

Article history:

Received 28 September 2009

Received in revised form 23 February 2010

Accepted 8 March 2010

Keywords:

Reflectance

Functional data

Lineal regression

Radial basis function

ABSTRACT

A statistical analysis of functional data, obtained as reflectance values measured using a hyperspectral sensor, was used to determine water content in vine leaves. Our study was conducted using a sample of 80 vine leaves whose water content was determined by calculating the weight difference between leaves before and after drying in an oven. Two regression models, one linear and the other non-linear, were evaluated and compared: functional linear regression and regression with functional radial basis functions. Compared to traditional methods based on calculating indices that only consider reflectance values in specific wavelengths, the functional approach enables a specific bandwidth or the entire electromagnetic spectrum recorded by the sensor to be taken into account; in other words, the functional approach enables the spectral signature of the leaves to be used. The optimal parameters for each model were determined using a cross-validation procedure and the validity of the approach was tested on a test set drawn from the initial sample. The results obtained demonstrate that water content can be predicted for vine leaves on the basis of their spectral signature, allowing for a certain margin of error. This error was smaller for the non-linear model compared to the linear model.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

For a vine to produce grapes that, in turn, result in a good quality wine, the level of humidity in the leaves needs to be controlled [1]. The most reliable and accurate way to determine humidity is by taking the leaves to a laboratory and calculating weight differences between wet and dry leaves. This method is time consuming, however, and also subject to measurement and sampling errors, especially when data are extrapolated to the whole vine or vineyard. Leaf water content is related to the amount of solar energy reflected by the leaf in specific electromagnetic spectral bands. Consequently, an alternative way of determining humidity – much faster although not as precise as the laboratory method – is to use the spectral signature of the leaves to associate solar energy with water content. Reflectance, especially in the NIR and MIR wavelengths [2–4], depends on the amount of water stored in leaf cells. For wavelengths sensitive to water absorption (760, 970, 1450, 1940 and 2950 nm), leaf reflectance decreases as water content increases [3,5–7]. Recent years have witnessed the development of hyperspectral sensors that are capable of recording reflectance values for very narrow electromagnetic spectral bands. Using such reflectance data, some authors have attempted to measure leaf water content and also other parameters such as water stress and chlorophyll content. Their analyses are generally limited, however, to the construction of indices as algebraic expressions of reflectance values for specific wavelengths [8–10]. In other words, the analyses fail to use all the information contained in the spectral curve.

* Corresponding address: Department of Environmental Engineering, Mines Engineering University, University of Vigo Campus, Lagoas Marcosende, s/n 36310 Vigo, Spain. Tel.: +34 986814052; fax: +34 986811924.

E-mail address: javier.martinez@uvigo.es (J. Martínez).

In hyperspectral remote sensing, measurements of spectral reflectance are often best thought of as functions. Reflectance is a function of wavelengths and is referred to as the spectral signature. Functional data analysis is therefore a suitable approach to analysing this data [11,12]. Furthermore, considering data in terms of curves means that we can take into account not only reflectance values but also other important parameters, such as function derivatives [13].

In this article we evaluate the application of two functional data analysis techniques to the estimation of vine leaf water content, namely functional linear regression [11] and the non-linear functional radial basis functions [14].

Some of the application fields of the functional data analysis include environmental research [15,16], medical research [17,18], sensors [19] and industrial methods [20,21].

The input was data referring to the energy reflected by the leaves in the region of the electromagnetic spectrum located between 700 and 1300 nm, which is the wavelength interval used by other authors for similar problems. An analysis of variance (ANOVA) study of the sample data confirmed this as the optimal region for study.

The remainder of this article is structured as follows: Section 2 provides a basic description of the techniques used; Section 3 describes the results obtained from the application of the techniques to differently sized test samples and discusses the advantages and limitations of each approach. Finally we draw on the findings to provide some conclusions.

2. Methodology

2.1. Smoothing

In the functional model [16,21], in most of the applications we do not see the functions $x_i, i = 1, \dots, n$ but only their values $x_i(t_j)$ in a set of n_p points $t_j \in R, j = 1, \dots, n_p$. For the sake of simplicity, we will assume these to be common to all the functions $x_i, i = 1, \dots, n$. These observations may, moreover, be subject to noise, and in this case they take the form: $z_{ij} = x_i(t_j) + \varepsilon_{ij}$, where we assume that ε_{ij} is random noise with zero mean, $i = 1, \dots, n, j = 1, \dots, n_p$.

Therefore, the functional focus first requires the sample functions to be registered, and this requires estimation of each function $x_i \in X \subset F, i = 1, \dots, n$. One approach is to assume that $F = \text{span}\{\phi_1, \dots, \phi_{n_b}\}$ with $\{\phi_k\}$ set of basic functions [11]. For our research we chose a family of B-splines as the set of basic functions, given their good local behaviour. If, for the sake of simplicity, we represent as x any of the functions $x_i, i = 1, \dots, n$ in the sample, we have:

$$x(t) = \sum_{k=1}^{n_b} c_k \phi_k(t). \tag{1}$$

A simple smoother [11] is obtained by minimizing the least squares criterion:

$$\min_{x \in \mathcal{F}} \sum_{j=1}^{n_p} \{z_j - x(t_j)\}^2.$$

However, using this criterion (resubstitution error or empirical risk) may lead sufficiently complex models to overfit to the data (noise or possible measurement errors), negatively affecting their generalization capacity regarding new data not used in the learning process.

It is therefore necessary to control the complexity of the model by using a regularization focus based on the regularized empirical risk criterion [22,23]:

$$\min_{x \in \mathcal{F}} \sum_{j=1}^{n_p} \{z_j - x(t_j)\}^2 + \lambda \Gamma(x) \tag{2}$$

where $z_j = x(t_j) + \varepsilon_j$ is the result of observing x at the point t_j , Γ is an operator that penalizes the complexity of the solution, and λ is a regularization parameter that regulates the intensity of this penalization. In our case, we have used the operator $\Gamma(x) = \int_{\mathcal{T}} \{D^2 x(t)\}^2 dt$ where $\mathcal{T} = [t_{\min}, t_{\max}]$ and D^2 is the second-order differential operator.

Bearing in mind the expansion (1), the above problem (2) may be written as:

$$\min_{\mathbf{c}} \{(\mathbf{z} - \Phi \mathbf{c})^T (\mathbf{z} - \Phi \mathbf{c}) + \lambda \mathbf{c}^T \mathbf{R} \mathbf{c}\}$$

where $\mathbf{z} = (z_1, \dots, z_{n_p})^T, \mathbf{c} = (c_1, \dots, c_{n_b})^T, \Phi$ is the $n_p \times n_b$ matrix with elements $\Phi_{jk} = \phi_k(t_j)$ and \mathbf{R} is the $n_b \times n_b$ matrix with elements:

$$R_{kl} = \langle D^2 \phi_k, D^2 \phi_l \rangle_{L_2(\mathcal{T})} = \int_{\mathcal{T}} D^2 \phi_k(t) D^2 \phi_l(t) dt.$$

The solution to this problem is given by

$$\mathbf{c} = (\Phi^T \Phi + \lambda \mathbf{R})^{-1} \Phi^T \mathbf{z}$$

in such a way that the estimated values of x at the observation points are obtained by means of $\mathbf{x} = \mathbf{S} \mathbf{z}$ where

$$\mathbf{S} = \Phi (\Phi^T \Phi + \lambda \mathbf{R})^{-1} \Phi^T$$

with $\mathbf{x} = (x(t_1), \dots, x(t_{n_p}))^T$.

The selection of the λ forms part of the model selection problem and is usually performed using crossed validation.

2.2. Functional linear regression

The linear regression model can be expressed as [24]:

$$\mathbf{y} = \alpha + \mathbf{X}\boldsymbol{\beta} + \delta \tag{3}$$

where \mathbf{X} is the matrix of inputs, \mathbf{y} is the vector of outputs, α is the translation term, $\boldsymbol{\beta}$ is the vector of regression coefficients and δ is the error in the model.

Once smoothing had been implemented, we had a sample $\{(x_i, y_i)\}_{i=1}^n$, in which x_i are the functional input variables and y_i are the scalar output variables. This enabled a linear regression to be performed taking a functional focus [25]. This regression can be considered, when considering functional data, as an extension of the multivariable model. In this case, (3) can be expressed as follows:

$$\mathbf{y} = \lambda + \int_I \mathbf{x}(t)\gamma(t)dt + \epsilon \tag{4}$$

where λ is a constant that represents the model translation term, $\mathbf{x}(t)$ is the vector of functional inputs, $\gamma(t)$ represents the functional coefficients of the regression, ϵ is the error occurring in the approximation process and I is the defining interval for the functions.

Although there are different approaches to estimating the function $\gamma(t)$ [25], for this research we used functional decomposition for finite-dimension spaces [11], which, with a view to reducing the degrees of freedom of the regression, performs regularization using basis functions, as denoted in (1). Hence, with $\boldsymbol{\phi}$ as a vector of basis functions for length n_β we have:

$$\gamma(t) = \sum_{k=1}^{n_\beta} d_k \phi_k(t) = \boldsymbol{\phi}^T(t)\mathbf{d}.$$

Length n_β is chosen in such a way that the information loss is minimal. Each observed function can be expressed as a function of other basis functions $\boldsymbol{\psi}$:

$$x_i(t) = \sum_{k=1}^{n_p} c_{ik} \psi_k(t) = \mathbf{c}_i^T \boldsymbol{\psi}(t) \implies \mathbf{x}(t) = \mathbf{C}\boldsymbol{\psi}(t)$$

where \mathbf{C} is the matrix of coefficients of the functional input variables $\mathbf{x}(t)$ with regard to the chosen basis functions $\boldsymbol{\psi}$.

Therefore, the prediction $\hat{\mathbf{y}}$ can be expressed as follows:

$$\hat{\mathbf{y}} = \lambda + \mathbf{C}\mathbf{J}_{\phi\psi} \mathbf{d}$$

where $\mathbf{J}_{\phi\psi}$ is the matrix expressed as follows:

$$\mathbf{J}_{\phi\psi} = \int \boldsymbol{\psi}(\mathbf{s})\boldsymbol{\phi}^T(\mathbf{s})d\mathbf{s}.$$

If Fourier functions are used, the matrix $\mathbf{J}_{\phi\psi}$ results in identity, as the Fourier basis functions form an orthonormal base for the functional space. Introducing the notation $\boldsymbol{\zeta} = (\lambda, d_1, \dots, d_{n_\beta})$, $\mathbf{Z} = [1, \mathbf{C}\mathbf{J}_{\phi\psi}]$ the prediction $\hat{\mathbf{y}}$ can be expressed as follows:

$$\hat{\mathbf{y}} = \mathbf{Z}\boldsymbol{\zeta}.$$

The vector of coefficients $\boldsymbol{\zeta}$ is given by applying the minimum squared error (MSE) criterion in the following equation:

$$\mathbf{Z}^T \mathbf{Z} \hat{\boldsymbol{\zeta}} = \mathbf{Z}^T \mathbf{y}.$$

2.3. Functional radial basis functions networks (FRBF)

Functional radial basis function networks for regression are an adaptation of the interpolation approach to problems that takes account of noise in the data [26].

A generalized linear regression model is obtained in which the regression variables are non-linear transformations of the input variables. A number of authors have independently proposed an analogous network using no pre-fixed transfer functions in the hidden layer; instead they are made dependent on a variable-window parameter [27].

In general, the radial basis function architecture takes the form [28]:

$$\hat{\mathbf{y}}(\mathbf{x}) = \sum_{i=1}^n c_i \Psi_i(\mathbf{x}) + c_0 \tag{5}$$

where $\Psi_i(\mathbf{x}) = f(\|\mathbf{x} - \mathbf{w}_i\|)$ is a radial transfer function \mathbf{w}_i are the centres of each node and $\mathbf{c} = \{c_1, \dots, c_n\}$ are the coefficients of the linear transformation of the network output layer.

Using a Gaussian transfer function results in:

$$\Psi_i(\mathbf{x}) \propto \exp \left\{ -\frac{1}{\sigma} \|\mathbf{x} - \mathbf{w}_i\|_{\Sigma_i}^2 \right\} = \exp \left\{ -\frac{1}{\sigma} (\mathbf{x} - \mathbf{w}_i)^T \Sigma_i^{-1} (\mathbf{x} - \mathbf{w}_i) \right\} \tag{6}$$

where \mathbf{w}_i are the chosen centres, generally a subset of the sample, σ is the scaling factor, determined using a cross-validation process and Σ_i is a symmetric positive semi-definite matrix.

Of interest is the particular case in which $\sigma = 1$ and $\Sigma_i = \mathbf{I}$, corresponding to the Euclidean distance and to a spherical transfer function [14].

When considering functional data, the transfer function (6) [14] results:

$$\Psi_i(x(t)) \propto \exp \left\{ -\frac{1}{\sigma} \left(\int \epsilon_i(t)(x(t) - w_i(t))^2 dt \right)^{1/2} \right\} \tag{7}$$

where $\epsilon_i(t)$ is the functional version of the matrix Σ_i when this matrix is diagonal. The function $\epsilon_i(t)$ can be any function, although it seems logical to choose one of the same complexity as the functions $x(t)$, and, more specifically, in the sense as its second derivative.

A typical approximation is to use the expression (1) to express a function as a linear combination of selected basis functions. Considering $\epsilon_i(t) = \mathbf{1}$, therefore, the expression (7) results:

$$\Psi_i(x(t)) \propto \exp \left\{ -\frac{1}{\sigma} ((\mathbf{a} - \mathbf{b}_i)^T \mathbf{H} (\mathbf{a} - \mathbf{b}_i))^{1/2} \right\} \tag{8}$$

where $x(t) = \mathbf{a}^T \boldsymbol{\phi}(t)$, $w_i(t) = \mathbf{b}_i^T \boldsymbol{\phi}(t)$ and $\mathbf{H} = \int \boldsymbol{\phi}(t)^T \boldsymbol{\phi}(t) dt$. The functional radial basis function is therefore reduced to a vectorial radial basis function, bearing in mind the expansion in basis functions and the matrix \mathbf{H} .

The training algorithm is simply a feedforward algorithm, rather than a backpropagation algorithm, as used by multilayer perceptron networks. Hence, the output $\hat{\mathbf{y}}$ of a radial basis function network, in general, is influenced by a non-linear transformation originating in the hidden layer through a radial function and a linear transformation originating in the output layer through a continuous linear function.

3. Application and results

3.1. Data collection

The vineyard block selected for this study was located in the Carneros Region of Napa Valley (California). The coordinate boundaries are (ellipsoid WGS84): SW (38.247104 °N, 122.366210 °W) and NE (38.247982 °N, 122.361995 °W). The vineyard was planted in 1991 and it is composed of a uniform planting of 9321 vines (3.63 ha) of a single cultivar (*Vitis vinifera* L., cv. Pinot Noir cv UC2A).

Leaf reflectance of vine leaves were acquired just after collecting the leaves in the field. The measurements were conducted using a Li-Cor Inc. model 1800-12S External Integrating Sphere (Li-Cor, Inc., Lincoln, NE, USA), coupled with a field spectrometer Analytical Spectral Devices Inc. model FieldSpec Pro ASD (Analytical Spectral Devices, Inc., Boulder, CO, USA), which detects reflectance in the 350–2500 nm spectral region [10].

3.2. Estimating the model parameters

A p -fold cross-validation process was used to estimate the optimal parameters for the regression models. The efficacy of each model was then checked using a test set of a size representative of the sample size. The criterion used was the MSE [29,30]:

$$MSE = \frac{1}{n} \sum_{k=1}^n (\hat{y}_k - y_k)^2. \tag{9}$$

The MSE provides information, expressed in the problem units, on the deviation produced by the regression. Nonetheless, relative error (RE) is used to obtain information that is independent of the problem units. The mean RE value [31,32], was calculated using the following expression:

$$RE = \frac{1}{n} \sum_{k=1}^n \frac{|\hat{y}_k - y_k|}{|y_k|}. \tag{10}$$

Completion of the data capture phase resulted in a sample $\{(\mathbf{x}_i, y_i)\}_{i=1}^{80}$, with $x_i(t_j)$, $j = 1, \dots, 2000$ as the reflectance values obtained using the hyperspectral sensor and y_i as the water values per surface unit for each leaf.

For each point t_j , $j = 1, \dots, 2000$, corresponding to the recording of functions, an ANOVA study [33–35] was implemented so as to obtain the F functions and p -value for the definition domain for each recorded function x_i (ANOVA functional). As can be observed in Fig. 1, (left), the central band (700 to 1300 nm) has a p -value of more than 5%, indicating the consistency of the sample. Fig. 1 (right) also shows the F -test distribution, which is lower than the statistic tabulated in the same region. Given these results, the wavelength interval to be analysed was reduced and the band falling between 700 and 1300 nm was chosen as being of greater scientific interest [5].

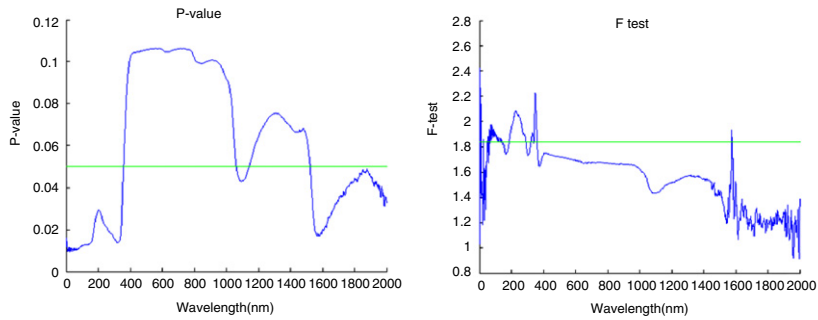


Fig. 1. P-value distribution (left) and F-test distribution (right) obtained in the ANOVA study of the sample data.

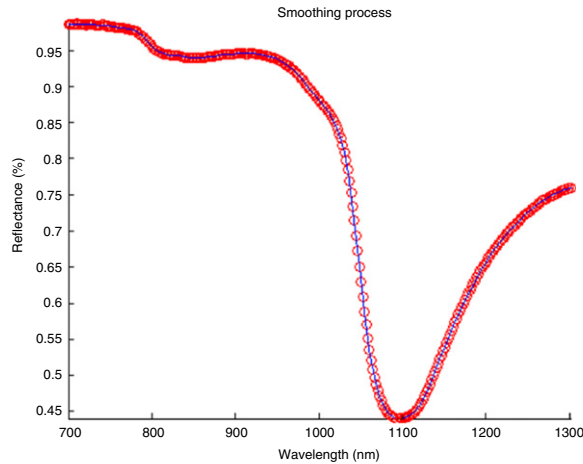


Fig. 2. Smoothing process for an observation. $MSE = 2 \times 10^{-7}$.

Table 1
Results for different test sample sizes using the central band of the sample.

Technique	Test set size (%)	MSE	RE (%)
Linear regression	5	2×10^{-6}	9
	10	6×10^{-6}	17
Functional radial basis functions	5	2.6×10^{-7}	3
	10	4×10^{-7}	4

Smoothing, as described above, was performed on the final sample chosen. Fig. 2 shows discrete values and the adjusted function (represented by circles and a line, respectively) for these points for a specific observation. The resulting MSE was around 10^{-7} .

Obtained in this way, $\{(x_i, y_i)\}_{i=1}^{80}$ where x_i is a functional input and y_i is a scalar output. A 10-fold cross-validation process was implemented for the models in order to obtain the following optimal parameters:

- *Number of basis functions:* Error remained constant from 100 basis functions, so 100 was chosen as the optimal number of Fourier basis functions needed to implement the smoothing process.
- *Scaling factor:* The value of σ was calculated as 1.5.

3.3. Error calculations

Error calculations for each model were performed with test sets representing 5% and 10% of the full sample. This enabled us to determine what influence, if any, sample size had on model results.

Table 1 shows the different values for the error criteria, namely MSE and RE, for the 5% and 10% test samples.

As can be observed, the error corresponding to the radial basis function model is considerably smaller than that for the linear model and is also less sensitive to training sample size.

Fig. 3 shows the differences between estimated and real points (represented by circles and crosses, respectively) for the linear regression model (left) and the radial basis functions model (right).

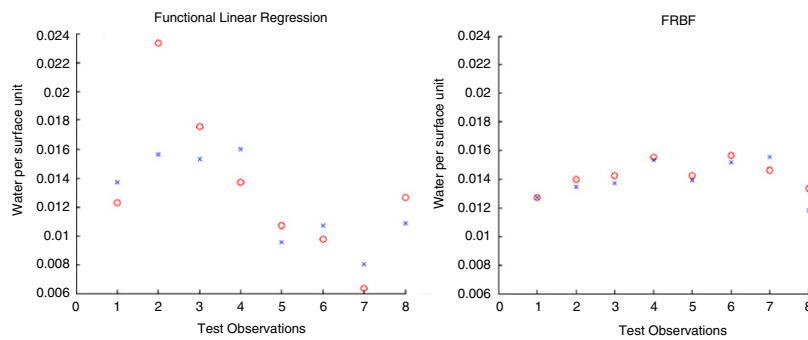


Fig. 3. Comparison of real points (crosses) and estimated points (circles) for the linear regression model (left) and the radial basis functions model (right).

4. Conclusions

Our research examined the feasibility of using functional regression analysis applied to reflectance data captured by a hyperspectral sensor to estimate vine leaf water content per surface unit. The practical interest of our study is that measuring leaf reflectance is faster than measuring water content by analysing weight differences. In addition, the fact that leaves do not have to be collected for analysis in a laboratory means that sampling can be more extensive.

The functional focus of the problem means that – unlike other approaches based on the calculation of indices for specific wavelength values – the entire wavelength spectrum within a given range can be taken into account. Our method enables full advantage to be taken of the capacity of hyperspectral sensors to capture data for bandwidths so narrow that spectral signatures can be represented as curves rather than as vectors of reflectance values.

Our analysis was performed for the central wavelength interval of 700 to 1300 nm – the same interval used by other authors for similar problems. This choice was further reinforced by the fact that an ANOVA study determined this wavelength interval to be the interval to which the data were best adapted.

Of the two models studied, the non-linear model based on functional radial basis functions produces considerably better results than the functional linear regression model. This outcome might indicate the existence of a complex dependency relationship between reflectance and vine leaf water content per surface unit. It might also explain the poor results obtained by some methods based on indices.

The results obtained also indicate that the non-linear model is less influenced by the sample size. Future research will be based on analysing larger data samples and including new functional techniques, such as functional regression models with penalties.

Acknowledgements

J.M. Matías' research is funded by the Spanish Ministry of Education and Science under Grant No. MTM2008-03010.

References

- [1] J.A. Kennedy, M.A. Matthews, A.L. Waterhouse, Effect of maturity and vine water status on grape skin and wine flavonoids, *American Journal of Enology and Viticulture* 53 (4) (2002) 268–274.
- [2] E.R. Hunt, B.N. Rock, Detection of changes in leaf water content using near and middle-infrared reflectances, *Remote Sensing of Environment* 30 (1989) 43–54.
- [3] B.C. Gao, NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sensing of Environment* 58 (1996) 257–266.
- [4] P. Ceccato, S. Flasse, S. Tarantola, S. Jacquemoud, J.M. Grégoire, Detecting vegetation leaf water content using reflectance in the optical domain, *Remote Sensing of Environment* 77 (2001) 22–33.
- [5] W.D. Bowman, The relationship between leaf water status, gas exchange, and spectral reflectance in cotton leaves, *Remote Sensing of Environment* 30 (1988) 249–255.
- [6] J. Peñuelas, J. Piñol, R. Ogaya, I. Filella, Estimation of plant water concentration by the reflectance water index (R900/R970), *International Journal of Remote Sensing* 18 (1997) 2869–2875.
- [7] W.J. Ripple, Spectral reflectance relationships to leaf water stress, *Photogrammetric Engineering and Remote Sensing* 52 (1986) 1669–1675.
- [8] N.H. Broge, E. Leblanc, Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density, *Remote Sensing of Environment* 76 (2001) 156–172.
- [9] D. Haboudane, J.R. Miller, N. Tremblay, P.J. Zarco-Tejada, L. Dextraze, Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture, *Remote Sensing of Environment* 81 (2002) 416–426.
- [10] J.R. Rodríguez-Pérez, D. Riaño, E. Carlisle, S. Ustin, D.R. Smart, Evaluation of hyperspectral reflectance indexes to detect grapevine water status in vineyards, *American Journal of Enology and Viticulture* 58 (3) (2007) 302–317.
- [11] J.O. Ramsay, B.W. Silverman, *Functional Data Analysis*, Springer, New York, 1997.
- [12] L. Bing, Y. Qingzhao, Classification of functional data: a segmentation approach, *Computational Statistics and Data Analysis* 52 (2008) 4790–4800.
- [13] P.J. Zarco-Tejada, C.A. Rueda, S.L. Ustin, Water content estimation in vegetation with MODIS reflectance data and model inversion methods, *Remote Sensing of Environment* 85 (2003) 109–124.
- [14] N. Delannay, F. Rossi, B. Conanguez, M. Verleysen, Functional radial basis function networks, in: *Proceedings of the 2004 European Symposium on Artificial Neural Networks*, 2004, pp. 313–318.

- [15] J.M. Paruelo, F. Tomasel, Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models, *Ecological Modelling* 98 (1997) 173–186.
- [16] J.M. Matías, C. Ordoñez, J. Taboada, T. Rivas, Functional support vector machines and generalized linear models for glacier geomorphology analysis, *International Journal of Computer Mathematics* 86 (2) (2009) 275–285.
- [17] D.A. Dombbeck, M.S. Graziano, D.W. Tank, Functional clustering of neurons in motor cortex determined by cellular resolution imaging in awake behaving mice, *Journal of Neuroscience* 29 (44) (2009) 13751–13760.
- [18] R. Viviani, G. Grön, M. Spitzer, Functional principal component analysis of fMRI data, *Human Brain Mapping* 24 (2) (2005) 109–129.
- [19] D. Wu, S. Huang, J. Xin, Dynamic compensation for an infrared thermometer sensor using least-squares support vector regression (LSSVR) based functional link artificial neural networks (FLANN), *Measurement Science and Technology* 19 (10) (2008) no. 105202.
- [20] J.I. Park, S.H. Baek, M.K. Jeong, S.J. Bae, Dual features functional support vector machines for fault detection of rechargeable batteries, *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 39 (4) (2009) 480–485.
- [21] M. López, J.M. Matías, J.A. Vilán, J. Taboada, Functional Pattern Recognition of 3D Laser Scanned Images of Wood-Pulp Chips, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 4477 (1) (2007) 298–305.
- [22] J.S. Simonoff, *Smoothing Methods in Statistics*, Springer, New York, 1996.
- [23] M.P. Wand, M.C. Jones, *Kernel Smoothing*, Chapman and Hall, London, 1995.
- [24] W. Härdle, *Applied Nonparametric Regression*, Cambridge University Press, Cambridge, 1990.
- [25] T. Tony, P. Hall, Prediction in functional lineal regression, *The Annals of Statistics* 34 (5) (2006) 2159–2179.
- [26] D.S. Broomhead, D. Lowe, Multivariable functional interpolation and adaptative networks, *Complex Systems* 2 (1988) 321–355.
- [27] J.E. Moody, C.J. Darken, Learning with localized receptive fields, in: *Proceedings of the 1988 Connectionist Models Summer School Morgan Kaufmann*, 1988.
- [28] R. González-Grimal, J. Cuevas Tello, Analysis of time series with artificial neural networks, in: *Proceedings of the 2008 Seventh Mexican International Conference on Artificial Intelligence*, 2008.
- [29] J. Martín-Piiego López, L. Ruiz-Maya Pérez, *Estadística I, Probabilidad*, Thomson, Madrid, 2004.
- [30] C.H. Coombs, *A Theory of Data*, Wiley, New York, 1964.
- [31] R. Burden, J. Faires, *Numerical Analysis*, Cengage Learning, 2004.
- [32] G. Stewart, *Afternotes on Numerical Analysis*, S.I.A.M., 1996.
- [33] A. Kagan, G. Letac, Characterization of the normal distribution through the power of a one-way ANOVA, *Journal of Statistical Planning and Inference* 77 (1999) 1–9.
- [34] M. Kendall, A. Stuart, *The Advanced Theory of Statistics-Volume 3*, Charles Griffin & Co Ltd., New York, 1983.
- [35] U. Ulusoy, Application of ANOVA to image analysis results of talc particles produced by different milling, *Powder Technology* 188 (2008) 133–138.