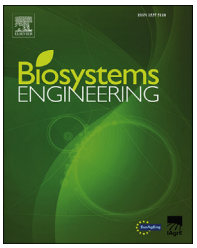
BIOSYSTEMS ENGINEERING XXX (2017) 1–11

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110



Special Issue: Crop Water Status

Research Paper

# Leaf water content estimation by functional linear regression of field spectroscopy data

*José R. Rodríguez-Pérez* [a], *Celestino Ordóñez* [b], *Ana B. González-Fernández* [a], *Enoc Sanz-Ablanedo* [a], *José B. Valenciano* [c], *Victoriano Marcelo* [d,*]

[a] *Grupo de Investigación en Geomática e Ingeniería Cartográfica (GEOINCA), Escuela Superior y Técnica de Ingeniería Agraria, Universidad de León, Avenida de Astorga, s/n, 24401, Ponferrada, León, Spain*

[b] *Departamento de Explotación y Prospección Minera, Universidad de Oviedo, C/Gonzalo Gutiérrez Quirós, s/n, 33600, Mieres, Asturias, Spain*

[c] *Departamento de Ingeniería y Ciencias Agrarias, Escuela Superior y Técnica de Ingeniería Agraria, Universidad de León, Avenida de Portugal 41, 24071 León, Spain*

[d] *Departamento de Ingeniería y Ciencias Agrarias, Escuela Superior y Técnica de Ingeniería Agraria, Universidad de León, Avenida de Astorga, s/n, 24401, Ponferrada, León, Spain*

ARTICLE INFO

*Article history:*
Published online xxx

*Keywords:*
Functional linear regression
Field spectral reflectance
Plant water stress
Equivalent water thickness
*Vitis vinifera* L.

Grapevine water status is critical as it affects fruit quality and yield. We assessed the potential of field hyperspectral data in estimating leaf water content ($C_w$) (expressed as equivalent water thickness) in four commercial vineyards of *Vitis vinifera* L. reflecting four grape varieties (Mencía, Cabernet Sauvignon, Merlot and Tempranillo). Two regression models were evaluated and compared: ordinary least squares regression (OLSR) and functional linear regression (FLR). OLSR was used to fit $C_w$ and vegetation indices, whereas FLR considered reflectance in four spectral ranges centred at the 960, 1190, 1465 and 2035 nm wavelengths. The best parameters for the FLR model were determined using cross-validation. Both models were compared using the coefficient of determination ($R^2$) and percentage root mean squared error (%*RMSE*). FLR using continuous stretches of the spectrum as input produced more suitable $C_w$ models than the vegetation indices, considering both the fit and degree of adjustment and the interpretation of the model. The best model was obtained using FLR in the range centred at 1465 nm ($R^2 = 0.70$ and % *RMSE* = 8.485). The results depended on grape variety but also suggested that leaf $C_w$ can be predicted on the basis of spectral signature.



* *Corresponding author.*
E-mail addresses: jr.rodriguez@unileon.es (J.R. Rodríguez-Pérez), ordonezcelestino@uniovi.es (C. Ordóñez), agonf@unileon.es (A.B. González-Fernández), enocsanz@unileon.es (E. Sanz-Ablanedo), joseb.valenciano@unileon.es (J.B. Valenciano), v.marcelo@unileon.es (V. Marcelo).

http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017
1537-5110/

## 1. Introduction

Water plays an important role in plant physiology, as it conditions yield and quality of crops such as grapes (*Vitis vinifera* L.). Water stress induces stomata closure to reduce transpiration, which, in turn, also reduces photosynthesis and carbon assimilation. The management of water deficits by controlling grapevine vigour and improving grape maturity could be an efficient strategy for producing a high-quality wine (Chaves et al., 2010). Water content estimation is, therefore, an important issue in managing vineyards.

Several techniques are available for water content estimation in crops. The main ground-based method used in viticulture is leaf water potential, which requires measurement of sap pressure in the xylem (Scholander, Hammel, Bradstreet, & Hemmingsen, 1965). However, this is a destructive and laborious method for estimating water content, especially as variations in water potential are often related to soil type (Chone, Van Leeuwen, Dubordieu, & Gaudillère, 2001). Thus, although it provides the most accurate assessments of plant water status, it is not feasible for estimates involving large areas (Oumar & Mutanga, 2010).

Since water has some absorption maxima in the infrared region of the spectrum centred at the 970, 1200, 1440 and 1950 nm wavelengths (Palmer & Williams, 1974), it is possible to assess plant water status using non-destructive remote sensing technologies. These are faster than the water potential method, and so offer a cost/time ratio advantage; moreover, spatial patterns of water plant content can also be detected by imagery (Moshou, Pantazi, Kateris, & Gravalos, 2014; Xue & Su, 2017).

The use of remote sensing to monitor crop growth and development is attracting interest from researchers and commercial organisations alike. This interest is primarily driven by opportunities for cost-effective generation of spatial data capable of supporting precision agriculture (Hall, Lamb, Holzapfel, & Louis, 2002). To date, limited use has been made of this technology in the grape and wine sector, whether for research or commercial monitoring purposes. This article describes the key principles of remote sensing, reviews the current status of remote sensing in viticulture and discusses remote sensing's potential as an integrated management tool for vineyards. Sims and Gamon (2003) classified remote sensing methods as follows: (1) vegetation index calculation using mathematical formulae for reflectance at several wavelengths; (2) continuum removal (CR) of the spectral signature and analysis of depth and area in the dip below the continuum; and (3) water content fitting to spectral reflectance over a range of two wavelengths mainly centred on the water absorption maxima.

Spectroscopic determination of leaf water content has been explored by Cheng, Rivard, and Sánchez-Azofeifa (2011) and Ustin, Riaño, and Hunt (2012), while a number of studies have analysed vine water status estimation using remote sensing. Strever (2005) assessed water stress in vines by field spectroscopy, finding important differences depending on vine vigour and concluding that the spectral reflectance of higher vigour and lower vigour vines was related to leaf water content and pigment, respectively. Serrano, González-Flor, and Gorchs (2010) studied the feasibility of using field spectral measurement to estimate vine water status at both leaf and canopy levels, reporting strong correlations for the water index (WI) and stomatal conductance ($g_s$), with coefficient of determination ($R^2$) values over 0.80. Note, however, that this result was obtained for potted plants subjected to varying degrees of water availability. In the field they demonstrated a correlation between predawn water potential and the normalized difference vegetation index (NDVI), achieving $R^2 = 0.57$. Serrano, González-Flor, and Gorchs (2012) related berry yield and quality with hyperspectral reflectance indices at canopy level, estimating berry yield by NDVI and WI ($R = 0.57$ and $R = 0.61$, respectively), and suggesting that total soluble acidity and the total soluble solids/total soluble acidity ratio might be estimated by WI when vineyards were experiencing moderate to severe water deficits.

Field spectroscopy is an effective technique for assessing the canopy density of vines. Dobrowski, Ustin, and Wolpert (2002) observed strong correlations between leaf area per metre of canopy and narrow vegetation indices, achieving $R^2$ values of 0.87, 0.92 and 0.79 for the ratio vegetation index (RVI), NDVI and perpendicular vegetation index (PVI), respectively. These authors recommended RVI for vineyard remote sensing applications, since it is more linearly related to canopy density and contains the same information as the NDVI. Similar results were found for imagery at vineyard level: the RVI was linearly correlated ($R^2$ values of between 0.68 and 0.88) with pruning weight for growing seasons (Dobrowski, Ustin, & Wolpert, 2003). Dobrowski, Pushnik, Zarco-Tejada, and Ustin (2005) linked vine physiological status and photosynthetic functioning with reflectance fluorescence indices (RFIs) calculated in the red-edge spectral region at canopy level. They indicated that RFIs were more suitable than the photochemical reflectance index (PRI) and NDVI indices for tracking photosynthetic status and plant stress, especially for rapid changes in vine status.

The use of sensors in applications to grapevines is complicated by the fact that a vineyard has a temporally and spatially changing environment that affects light interactions with leaves and grapes (Strever, Bezuidenhout, Zorer, Moffat, & Hunter, 2012). Nevertheless, the literature cited above would support the usefulness of passive reflectance measurements in monitoring vine physiological status, so remote sensing methods for estimating water content merit further study.

A different approach to tackling the problem of water content estimation from reflectance is based on considering spectral signatures as continuous curves instead of discrete values. A review of potential applications of this kind of functional data analysis to chemometrics is provided by Aguilera, Escabias, Mariano, Valderrama, and Aguilera-Morillo (2013) and by Saeys, Keteleare, and Darius (2008). In exemplifying the use of functional models, Saeys et al. (2008) concluded that functional data analysis is comparable to partial least squares regression (PLSR) in terms of predictive ability. Reiss and Ogden (2007) introduced functional versions of principal component regression and PLSR to NIR spectral analyses of both real and simulated data, concluding that functional models offer advantages over non-functional approaches. Dias, García, Ludwig, and Saraiva (2015) also used functional data techniques to calibrate and predict NIR spectral data. Ordóñez, Martínez, Matías, Reyes, and

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

Rodríguez-Pérez (2010) estimated vine leaf water content using functional linear regression (FLR) and functional radial basis functions, concluding that the complex dependency relationship between reflectance and vine leaf water content might explain the poor results obtained with methods based on indices. Ordóñez, Rodríguez-Pérez, Moreira, and Sanz (2013) used FLR and non-parametric functional methods to predict certain vine leaf chemical characteristics (moisture, dry mass and nitrogen, phosphorus, potassium, calcium, iron and magnesium concentrations) from electromagnetic reflectance between 350 and 2500 nm, reporting that non-parametric methods yielded better results and that moisture and phosphorous were the best predicted components. However, non-parametric methods have the drawback that since they do not allow for a physical interpretation of the model, there is a risk of overfitting.

This work reports the results of functional analysis of the vine leaf spectrum, using VIS/NIR spectroscopy at the leaf level, as the basis for rapid and non-destructive assessment of water content. The results are compared with those for ordinary least squares regression (OLSR) and vegetation indices.

## 2. Material and methods

### 2.1. Study site and experimental setup

The study vineyard was located in Bierzo DO in northern Spain (550 m mean height above sea level and 42°36′N, 6°42′W; Datum: WGS84). Measurements of leaf water content were made for four varieties (Mencía, Cabernet Sauvignon, Merlot and Tempranillo) of 18-year-old vines (*V. vinifera* L.; rootstock: 1103 Paulsen). The vines were vertical shoot-positioned with two pairs of wires and mean row spacing was 1.1 m × 2.8 m.

A regular grid of 20 m × 30 m was defined to select the 162 data vines (47 Cabernet Sauvignon, 45 Mencía, 27 Merlot and 43 Tempranillo), corresponding to 14 vines $ha^{-1}$. Three leaves per vine were marked, giving a total of 486 leaf samples. All the mature leaves had the same relative location on the shoot (the opposite side of the first cluster from the bottom). In accordance with Santos and Kaye (2009), field data were collected between berry set and veraison (on 17 July 2012).

### 2.2. Workflow

The methodology involved three main steps: (1) spectral data collection, (2) leaf data collection, and (3) statistical analysis. The field and laboratory measurement sequence for each leaf was as follows: reflectance measurements were made, the leaf was picked, placed in a plastic bag and stored in a cooler, fresh mass ($M_f$) was measured, the leaf was dried and finally, dry mass ($M_d$) was calculated. From the spectral data, two different transformations were made to obtain the CR intervals and the vegetation indices. FLR and OLSR were used to estimate $C_w$.

### 2.3. Spectral data

#### 2.3.1. Field spectroscopy measurements

Reflectance measurements of the 486 leaves were made using an ASD FieldSpec 4 spectroradiometer (Analytical Spectral Devices Inc., Boulder, CO, USA) that detects reflectance in the 350–2500 nm spectral region. The spectroradiometer was coupled with a leaf clip and a plant probe to ensure correct data acquisition at leaf level. Three spectral measurements per leaf were made from the adaxial surface and the mean reflectance value for each point was saved. Measurements were made avoiding leaf veins and spots and applying the same criteria: thus, the first measurement was made on the right part of the leaf, the second in the centre and the third on the left part of the leaf. Each capture saved was the average of three spectral measurements, meaning that there were three spectral signatures per sample. The spectroradiometer was calibrated against the white panel face following ASD Inc. (2012) recommendations and was recalibrated before measuring the first leaf of each vine.

#### 2.3.2. Spectral data pre-processing

The field reflectance data were pre-processed using ViewSpect Pro 6.0 (Analytical Spectral Devices, Inc., Boulder, CO, USA) and SAMS 3.2 (Center for Spatial Technologies and Remote Sensing-CSTARS, University of California, Davis CA, USA; http://cstars.metro.ucdavis.edu/resources/software/), obtaining an average spectral signature per leaf sample. The spectral data processing resulted in the calculation of the narrow-band vegetation indices derived from the spectral signatures, the CR transformation and the functional analysis.

#### 2.3.3. Vegetation indices

Eleven vegetation indices, which take into account the wavelengths most related to spectrum water absorption, were calculated to estimate water content using OLSR. Table 1 shows the vegetation indices calculated for this research.

#### 2.3.4. Continuum removal

CR was used to enhance the absorption characteristics of the spectrum (Kokaly & Clark, 1999). CR transformation normalises reflectance values to a common baseline, thereby allowing individual absorption features to be compared and highlighting and identifying absorption features of interest. The CR calculation requires the target regions to be identified, and in this study they were determined by the main water absorption features located at 970, 1200, 1440 and 1950 nm (Kokaly, Asner, Ollinger, Martin, & Wessman, 2009; Sims & Gamon, 2003). Table 2 shows the wavelengths for the four zones ($Z_i$) where CR was calculated.

### 2.4. Leaf data collection

Immediately after reflectance measurement, three 6.16-$cm^2$ disks were cut from each leaf and weighed using a SNUG-150 precision scale (Xiamen Jadever Scale Co., Xiamen, Fujian, China) in order to determine $M_f$ for each sample. The leaf disks were then dried in an oven at 65 °C for 72 h, after which $M_d$ was calculated. Water content was determined by the equivalent water thickness ($C_w$), obtained by calculating the difference between fresh and dry mass ($M_f - M_d$) per unit of leaf area ($A_L$) according to the following equation (Datt, 1999):

$$C_w = (M_f - M_d)/(\rho_w \times A_L) \quad (1)$$

where $\rho_w$ is the density of pure water (1 g $cm^{-3}$).

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

**Table 1 – Spectral indices.**

| Vegetation index | Acronym | Equation | Reference |
|---|---|---|---|
| Red/green index | RGI | $RGI = \frac{R_{695}}{R_{554}}$ | (Fuentes, Gamon, Qiu, Sims, & Roberts, 2001) |
| Structure intensive pigment index | SIPI | $SIPI = \frac{R_{800}-R_{445}}{R_{800}+R_{680}}$ | (Peñuelas, Baret, Filella, 1995) |
| Water index | WI | $WI = \frac{R_{900}}{R_{970}}$ | (Peñuelas, Pinol, Ogaya, Filella, 1997) |
| Simple ratio water index | SRWI | $SRWI = \frac{R_{1350}}{R_{870}}$ | (Zarco-Tejada & Ustin, 2001) |
| Normalized difference vegetation index | NDVI | $NDVI = \frac{R_{858}-R_{645}}{R_{858}+R_{645}}$ | (Rouse, Haas, Schell, & Deering, 1974) |
| Normalized difference water index | NDWI | $NDWI = \frac{R_{870}-R_{1260}}{R_{870}+R_{1260}}$ | (Gao, 1996) |
| Floating position water band index | fWBI | $fWBI = \frac{R_{900}}{\min(R_{930-980})}$ | (Strachan, Pattey, & Boisvert, 2002) |
| Shortwave infrared water stress index | SIWSI | $SIWSI = \frac{R_{858.5}-R_{1640}}{R_{858.5}+R_{1640}}$ | (Fensholt & Sandholt, 2003) |
| Normalized difference infrared index | NDII | $NDII = \frac{R_{835}-R_{1650}}{R_{835}+R_{1650}}$ | (Hardisky, Klemas, & Smart, 1983) |
| Zarco–Tejada–Miller Index | ZTM | $ZTM = R_{750}/R_{710}$ | (Zarco-Tejada & Ustin, 2001) |
| Photochemical reflectance index | PRI | $PRI = \frac{R_{570}-R_{531}}{R_{570}+R_{531}}$ | (Gamon, Peñuelas,Field 1992) |

$R_\lambda$: Reflectance at λ wavelength.

**Table 2 – Definitions of spectral zones for continuum removal calculations.**

| Zone | Interval (nm) | Range | Central wavelength (nm) |
|---|---|---|---|
| $Z_1$ | 860–1065 | 205 | 960 |
| $Z_2$ | 1114–1265 | 151 | 1190 |
| $Z_3$ | 1265–1668 | 403 | 1465 |
| $Z_4$ | 1830–2240 | 410 | 2035 |

Total specific leaf fresh mass ($C_{fm}$, Equation (2)) and specific leaf mass ($C_{dm}$, Equation (3)) were calculated for $M_f$ and $M_d$, respectively, per unit of leaf area:

$$C_{fm} = M_f/A_L \tag{2}$$

$$C_{dm} = M_d/A_L \tag{3}$$

### 2.5. Statistical analysis

#### 2.5.1. Functional analysis: mathematical model

We constructed a mathematical model based on FLR to estimate water content from the reflectance. We used this technique because data collected by the spectroradiometer can be considered as samples of continuous curves, so it could be assumed that the underlying curve-generation process was smooth and that the measured data were dependent. It was more appropriate, in order to preserve the dependence relationships in the analysis, to operate at a global level using functional regression rather than on individual wavelengths (e.g., using vegetation indices).

In functional regression analysis, which is an extension of ordinary linear regression analysis, the covariates are functions instead of scalar values (Ramsay & Silverman, 2002). The model takes the form:

$$y = \alpha + \int_S x(s)\beta(s)ds + \varepsilon = \alpha + \langle x, \beta\rangle + \varepsilon \tag{4}$$

where y represents the response variable; x(s) represent the covariates, which are real functions; $\alpha$ and $\beta(s)$ are the regression coefficients, represented by a real value and real functions, respectively; and $\varepsilon$ is the error term.

Since discrete data instead of functions are normally available, x(s) and $\beta(s)$ are approximated by means of decomposition into basis functions:

$$\begin{aligned} x(s) &= \sum_{k=1}^{n} a_k\phi_k = \mathbf{a}^T\phi(s) \\ \beta(s) &= \sum_{p=1}^{m} b_p\psi_p = \mathbf{b}^T\psi(s) \end{aligned} \tag{5}$$

where **a** and **b** are vectors of coefficients, and $\phi(s)$ and $\psi(s)$ are the basis functions. These functions can be polynomial, exponential, B-splines or Fourier functions, among others.

Substituting the expressions in Equation (5) in Equation (4), an estimate of y is obtained as follows:

$$\hat{y} = \hat{\alpha} + \mathbf{a}^T\mathbf{J}\mathbf{b} \tag{6}$$

where the *nxm* matrix **J** is given by:

$$\mathbf{J}_{\phi\psi} = \int_T \psi(s)\phi(s)ds = \langle\psi, \phi\rangle \tag{7}$$

The prediction $\hat{\mathbf{y}}$ in Equation (6) can be expressed as:

$$\hat{y} = \mathbf{Z}\zeta \tag{8}$$

where $\mathbf{Z} = [1, \mathbf{a}^T\mathbf{J}_{\phi\psi}]$ and $\zeta = (\alpha, b_{1,\ldots,}b_m)$

The estimate of the vector of regression coefficients $\hat{\zeta}$ is obtained by minimising the residual sum of squares $[y - \alpha - \int_S X(s)\beta(s)ds]^2$. The result can be expressed as:

$$\hat{\zeta} = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T y \tag{9}$$

To avoid overfitting, a regularisation of the solution is obtained by means of a penalty function that prevents excessive local fluctuation in the estimated function. Given any twice differentiable function $\omega$, it is possible to define the penalised residual sum of squares as:

$$P_\lambda(\alpha, \beta) = \left[y - \alpha - \int_S X(s)\beta(s)ds\right]^2 + \lambda\int_S [D^2\omega(s)]ds \tag{10}$$

where the operator $D^2$ represents the second derivative, and where the smoothing parameter $\lambda > 0$ controls the trade-off between roughness and infidelity in the observed data (Ramsay & Silverman, 2002). When the smoothing value $\lambda$ is close to zero, we are mainly concerned with the fit to the data; a complex and difficult to interpret regression coefficient $\beta(s)$ is typically the result and overfitting of the model is common. With increased $\lambda$ we obtain a smoother solution, which means

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

that the regression coefficient is easier to interpret. However, since larger smoothing parameter values produce excessively smooth solutions that approach standard linear regression, it is very important to locate a compromise value for this parameter.

The solution to the regularisation problem with a roughness penalty is analogous to that of Equation (9):

$$\hat{\zeta} = \left(\mathbf{Z}^{T}\mathbf{Z} + \lambda\mathbf{R}\right)^{-1}\mathbf{Z}^{T}\mathbf{y} \tag{11}$$

where R is a $m \times m$ matrix with elements $R_{jk} = \int_S D^2\psi_j(s)D^2\psi_k(s)ds = \langle D^2\psi_j, D^2\psi_k\rangle$.

The smoothing parameter $\lambda$ can be chosen subjectively if we have some *a priori* knowledge of the relationship between the response and explanatory variables; alternatively, it can be chosen by means of cross-validation, that is, by minimising a cross-validation score defined as:

$$CV(\lambda) = \sum_{j=1}^{N}\left(y_j - \alpha_{\lambda}^{(-j)} - \int_S X_j(s)\beta_{\lambda}^{(-j)}ds\right)^2 \tag{12}$$

where N is the sample size, and where $\alpha_{\lambda}^{(-j)}$ and $\beta_{\lambda}^{(-j)}$ are the estimates of $\alpha$ and $\beta$ obtained by minimising the penalised residual sum of squares based on all the data except $(x_j, y_j)$.

#### 2.5.2. Ordinary least squares regressions

Using vegetation indices, $C_w$ was estimated by OLSR, a linear regression model that can be used to model a single response variable that has been recorded, at least, at an interval scale. The relationship between $C_w$ and vegetation indices can be represented using an equation that indicates the best fitted line.

#### 2.5.3. Validation

The regression models were compared using observed and predicted values. They were validated by means of the leave-one-out cross-validation method, using two comparison criteria, namely, the highest cross-validated coefficient of determination ($R^2$) and the least error (root mean square error (*RMSE*) and *%RMSE* (*RMSE* expressed as a percentage of the mean value of the variable) as per Equation (13):

$$\%RMSE = (RMSE/\bar{x}) \times 100 \tag{13}$$

where *RMSE* is the root-mean-square error of the cross-validation calibration and $\bar{x}$ is the mean of the predicted values for $C_w$, $C_{fm}$ and $C_{dm}$. Accurate models should reduce the *RMSE* by at least 2% (Clevers, Kooistra, & Schaepman, 2008).

## 3. Results

### 3.1. Spectral and leaf data

Leaf disk mass for the different varieties was not very different in terms of fresh and dry matter and water content (Table 3). For this reason, it was difficult to correlate spectral information and leaf water content.

Spectral signatures for the leaves with highest, medium and lowest $C_w$ for the studied varieties showed similar trends for all four varieties (Fig. 1). The differences in reflectance values in the visible region (350–650 nm) were small, but were greater for higher wavelengths. Since differences were greatest in the range 650–1400 nm, these were the most suitable intervals for detecting variations in water content. Tempranillo was the variety which showed the greatest differences, followed by Mencía, Merlot and Cabernet, in that order.

### 3.2. Vegetation indices

$R^2$ values and errors (*RMSE* and *%RMSE*) obtained in estimating $C_w$ using the vegetation indices as spectral data indicated that $R^2$ values were higher than 0.5 only for Tempranillo, for the shortwave infrared water stress index (SIWSI) ($R^2 = 0.53$) and for the normalized difference infrared index (NDII) ($R^2 = 0.52$) (Table 4). No vegetation indices were suitable for predicting $C_w$ for Mencía, Merlot or Cabernet, which would suggest that these indices are not suitable for estimating water content.

### 3.3. Functional linear regression

Regression values obtained for prediction of $C_w$ using FLR as the fitted method showed — as with the vegetation indices — that Tempranillo was the variety that gave the highest regression values — notably the models for $Z_3$ and $Z_4$, both with $R^2 = 0.7$ (Table 5). Both these models had the same *%RMSE* (8.485), but the model based on $Z_4$ had a smaller smoothing parameter (Sp) than that based on $Z_3$ (2.5 and 4, respectively); hence, the model for $Z_4$ was easier to interpret than the model for $Z_3$. The suitability of $Z_3$ and $Z_4$ for determining $C_w$ was confirmed for the other grape varieties. $Z_4$ was the most appropriate interval for Merlot, with $R^2 = 0.61$, the smallest *%RMSE* and the same Sp as $Z_3$. Mencía gave the highest regression value in $Z_4$ ($R^2 = 0.54$), but since the difference in *%RMSE* between $Z_3$ and $Z_4$ was greater than 2%, $Z_3$ was the most appropriate interval. On the other hand, $Z_4$ gave a lower Sp than $Z_3$, so $Z_3$ and $Z_4$ were both suitable for predicting water

**Table 3 – Leaf variables statistics.**

| | Cabernet | | | Mencía | | | Merlot | | | Tempranillo | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $C_w$ | $C_{dm}$ | $C_{fm}$ | $C_w$ | $C_{dm}$ | $C_{fm}$ | $C_w$ | $C_{dm}$ | $C_{fm}$ | $C_w$ | $C_{dm}$ | $C_{fm}$ |
| SD | 0.017 | 0.010 | 0.021 | 0.017 | 0.008 | 0.021 | 0.018 | 0.008 | 0.021 | 0.025 | 0.012 | 0.033 |
| Mean | 0.156 | 0.072 | 0.228 | 0.162 | 0.062 | 0.224 | 0.142 | 0.062 | 0.203 | 0.165 | 0.066 | 0.231 |
| Range | 0.100 | 0.051 | 0.111 | 0.087 | 0.046 | 0.100 | 0.087 | 0.043 | 0.125 | 0.133 | 0.062 | 0.192 |

**Variables**: $C_w$: equivalent water thickness (kg m$^{-2}$); $C_{fm}$: total specific leaf fresh mass (kg m$^{-2}$); $C_{dm}$: specific leaf mass (kg m$^{-2}$). **Statistics**: SD: standard deviation.

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

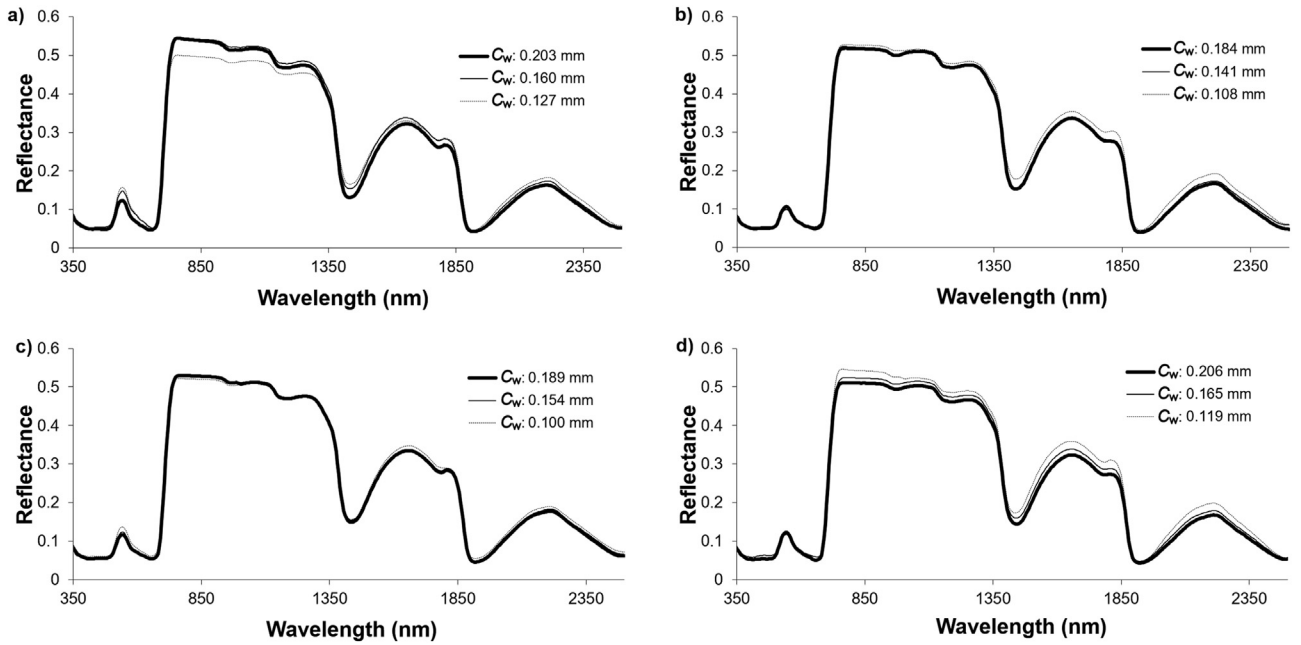



**Fig. 1 – Spectral signatures for three leaves with high (—), medium (—) and low (- - -) equivalent water thickness values ($C_w$) for (a) Mencía, (b) Merlot, (c) Cabernet and (d) Tempranillo.**

content for Mencía. Cabernet gave $R^2$ values lower than 0.5 for all the spectral intervals, indicating that this variety was not suitable for estimating water content.

Plots of $C_w$ versus wavelength were created to investigate the relationship between measured water content and leaf reflectance. The results varied depending on grape variety and the smoothing parameter applied to the functional algorithm. Since the best results were obtained for Tempranillo, these are the only results reported here. In Figs. 2 and 3, the continuous line represents the regression coefficient value $\beta(\lambda)$ obtained for the $C_w$ estimate with respect to the reflectance value for each wavelength, and the broken line shows the confidence interval for the regression coefficient.

$\beta(\lambda)$ values obtained for $C_w$ estimates for the Tempranillo variety, using CR-transformed reflectance values as the predictor variable (from 860 to 2240 nm) and using FLR as the fitting method, indicated the largest absolute regression coefficient values for wavelength intervals near 1500–1900 nm and 2100–2250 nm, although with opposite signs (Fig. 2). Higher water content was associated with lower reflectances in the first interval and with higher reflectances in the second interval. The lowest wavelengths, especially those in the

**Table 4 – Determination coefficients ($R^2$) obtained by ordinary least square regression between water content ($C_w$) and spectral information (vegetation indices).**

| | Cabernet | | | Mencía | | | Merlot | | | Tempranillo | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $R^2$ | *RMSE* | *%RMSE* | $R^2$ | *RMSE* | *%RMSE* | $R^2$ | *RMSE* | *%RMSE* | $R^2$ | *RMSE* | *%RMSE* |
| fWBI | 0.23 | 0.017 | 10.897 | 0.42 | 0.013 | 8.025 | 0.31 | 0.015 | 10.563 | 0.40 | 0.018 | 10.909 |
| NDII | 0.25 | 0.016 | 10.256 | 0.43 | 0.013 | 8.025 | 0.33 | 0.014 | 9.859 | 0.52 | 0.016 | 9.697 |
| NDVI | 0.00 | 0.019 | 12.179 | 0.00 | 0.017 | 10.494 | 0.03 | 0.017 | 11.972 | 0.01 | 0.023 | 13.939 |
| NDWI | 0.18 | 0.017 | 10.897 | 0.29 | 0.014 | 8.642 | 0.13 | 0.016 | 11.268 | 0.14 | 0.022 | 13.333 |
| PRI | 0.00 | 0.019 | 11.999 | 0.00 | 0.017 | 10.534 | 0.04 | 0.017 | 12.016 | 0.00 | 0.025 | 14.968 |
| RGI | 0.01 | 0.019 | 12.179 | 0.15 | 0.016 | 9.877 | 0.04 | 0.017 | 11.972 | 0.06 | 0.023 | 13.939 |
| SIPI | 0.01 | 0.019 | 12.179 | 0.03 | 0.017 | 10.494 | 0.02 | 0.017 | 11.972 | 0.01 | 0.023 | 13.939 |
| SIWSI | 0.25 | 0.016 | 10.256 | 0.43 | 0.013 | 8.025 | 0.33 | 0.014 | 9.859 | 0.53 | 0.016 | 9.697 |
| SRWI | 0.28 | 0.016 | 10.256 | 0.46 | 0.013 | 8.025 | 0.37 | 0.014 | 9.859 | 0.43 | 0.018 | 10.909 |
| WI | 0.25 | 0.016 | 10.256 | 0.44 | 0.013 | 8.025 | 0.35 | 0.014 | 9.859 | 0.42 | 0.018 | 10.909 |
| ZTM | 0.00 | 0.019 | 12.179 | 0.00 | 0.017 | 10.494 | 0.01 | 0.017 | 11.972 | 0.01 | 0.023 | 13.939 |

$R^2$: coefficient of determination (cross-validation); *RMSE*: root mean square error (cross-validation); *%RMSE*: percentage root mean square error (cross-validation) in relation to the average value of the variable. **Variable**: $C_w$: equivalent water thickness (kg $m^{-2}$). **Vegetation indices**: fWBI: floating position water band; NDII: normalized difference infrared index; NDVI: normalized difference vegetation index; NDWI: normalized difference water index; RPI: photochemical reflectance index; RGI: red/green index; SIPI: structure intensive pigment index; SIWSI: shortwave infrared water stress; SRWI: simple ratio water index; WI: water index; ZTM: Zarco–Tejada–Miller index.

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

**Table 5 – Determination coefficients ($R^2$) and *RMSE* values obtained by functional linear regression to estimate water content ($C_w$) from spectral information for specific spectral regions (see Table 2).**

| | Cabernet | | | | Mencía | | | | Merlot | | | | Tempranillo | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $R^2$ | Sp | *RMSE* | %*RMSE* | $R^2$ | Sp | *RMSE* | %*RMSE* | $R^2$ | Sp | *RMSE* | %*RMSE* | $R^2$ | Sp | *RMSE* | %*RMSE* |
| $Z_1$ | 0.27 | 1 | 0.016 | 10.256 | 0.49 | 4 | 0.012 | 7.407 | 0.47 | 1 | 0.013 | 9.155 | 0.65 | 1 | 0.015 | 9.091 |
| $Z_2$ | 0.33 | −1 | 0.015 | 9.615 | 0.50 | 4 | 0.012 | 7.407 | 0.47 | 5 | 0.013 | 9.155 | 0.66 | 3.5 | 0.015 | 9.091 |
| $Z_3$ | 0.40 | 1 | 0.015 | 9.615 | 0.52 | 8 | 0.012 | 7.407 | 0.62 | 2.5 | 0.017 | 11.972 | 0.70 | 4 | 0.014 | 8.485 |
| $Z_4$ | 0.34 | 5.5 | 0.015 | 9.615 | 0.54 | 3 | 0.016 | 9.877 | 0.61 | 2.5 | 0.011 | 7.746 | 0.70 | 2.5 | 0.014 | 8.485 |

$R^2$: coefficient of determination (cross-validation); Sp: smoothing parameter (cross-validation); *RMSE*: root mean square error (cross-validation); %*RMSE*: percentage root mean square error (cross-validation) in relation to the average value of the variable. **Variables:** $C_w$: equivalent water thickness (kg m$^{-2}$); $C_{fm}$: total specific leaf fresh mass (kg m$^{-2}$); $C_{dm}$: specific leaf mass (kg m$^{-2}$). **Spectral intervals**: $Z_1$: continuum removal for the 860–1065 nm spectral interval; $Z_2$ continuum removal for the 1114–1265 nm spectral interval; $Z_3$: continuum removal for the 1265–1668 nm spectral interval; $Z_4$: continuum removal for the 1860–2240 nm spectral interval.

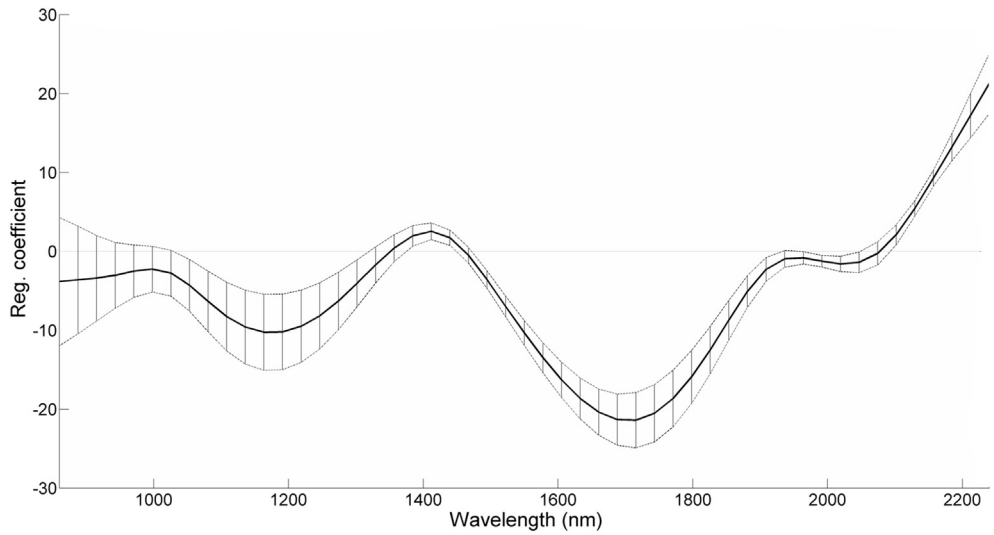



**Fig. 2 – Regression coefficient values for equivalent water thickness ($C_w$) estimates for Tempranillo using functional linear regression in the 860–2240 nm wavelengths ($Z_1 + Z_2 + Z_3 + Z_4$). Dependent variable values multiplied by 1000.**

interval 860–1000 nm, corresponded to confidence intervals around zero, so this region does not provide useful information for estimating $C_w$.

Results for $Z_1$, with the sign of the regression coefficient changing from positive to negative depending on wavelength, indicated that, for the first part of the $Z_1$ interval (860–900 nm), the $\beta(\lambda)$ values were not different from zero and, moreover, remained low until 1100 nm, from which point they increased rapidly (Fig. 3(*a*)). This would indicate a strong variation in the effect of reflectance on $C_w$ according to wavelength.

Regression values obtained for $C_w$ estimates for Tempranillo, using FLR as the fitting method for $Z_2$, showed that the regression coefficient decreased as the wavelength value increased, and became negative from 1140 nm (Fig. 3(*b*)). The lowest confidence interval value was obtained at 1190 nm, indicating less uncertainty in coefficient estimation. In the interval 1114 nm–1265 nm, the response for $C_w$ was negative when the wavelength value increased, mainly from 1190 nm. This would indicate that $C_w$ and reflectance were inversely related.

Regression values obtained for $C_w$ estimates for Tempranillo, using FLR as the fitting method for $Z_3$, showed a strong increase in $\beta(\lambda)$ with $\lambda$ until 1375 nm (Fig. 3(c)). Thereafter, the slope of coefficient of regression tended to decrease and its absolute value was smaller, indicating a lower effect of reflectance on $C_w$. $Z_3$ appeared to be most suitable in relating water content and reflectance by FLR, given the strong response of $C_w$, especially in the interval 1265–1375 nm.

In the $Z_4$ interval, the relationship between $C_w$ and wavelength was more sinuous and unstable (Fig. 3(*d*)), even though $R^2$ reached 0.7, as happened with $Z_3$ (Table 5). This could indicate a complex relationship between reflectance and $C_w$, but may also be partially explained by overfitting. In this case, using this functional regression coefficient with a new dataset could produce a poor estimate for $C_w$.

## 4. Discussion

### 4.1. *Spectral and leaf data*

Spectral data collected with a field spectroradiometer were used to estimate water content in four vineyard plots, given the well-known relationship between leaf reflectance and leaf

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

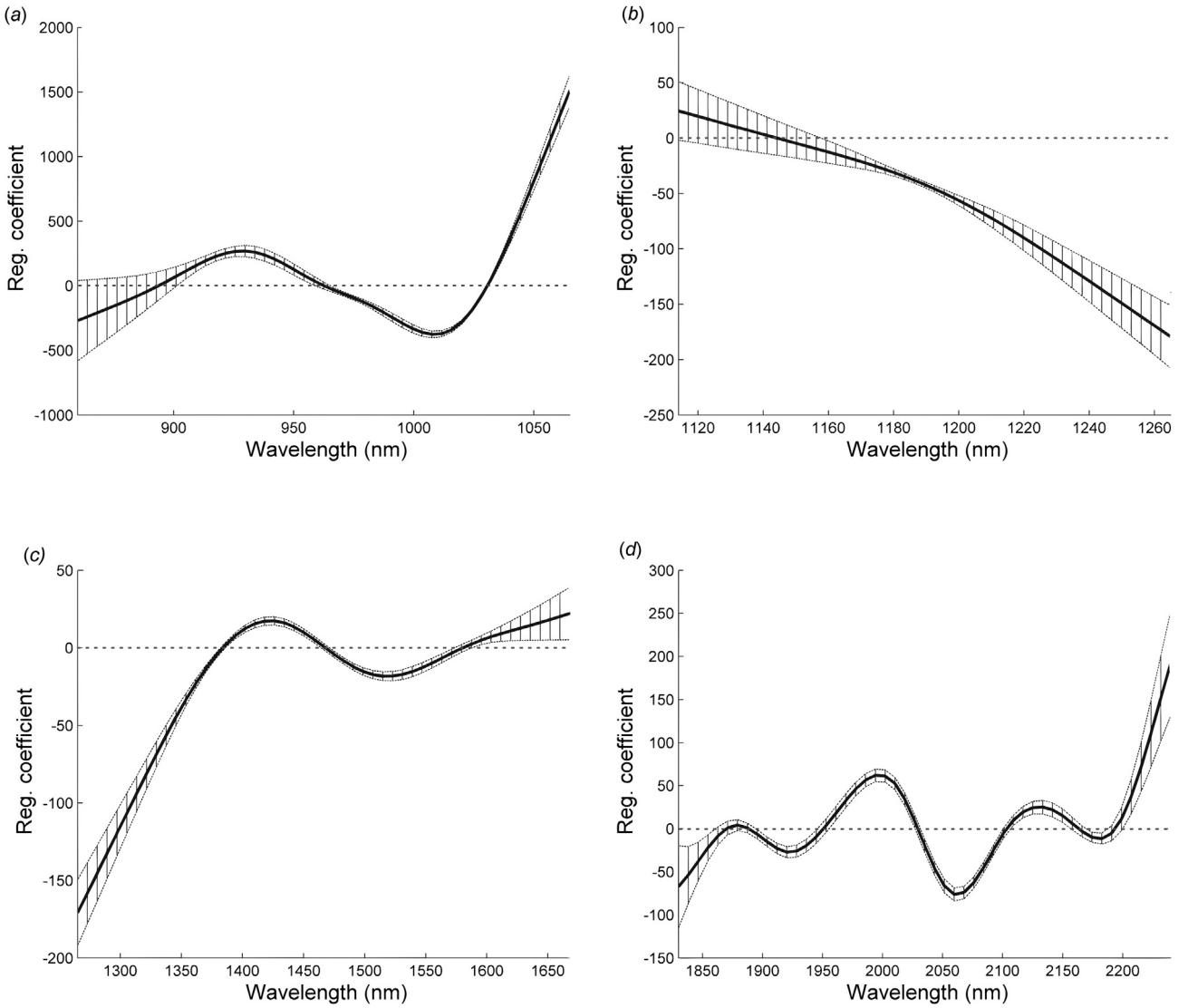



**Fig. 3 – Regression coefficients for equivalent water thickness ($C_w$) estimates for Tempranillo using functional linear regression for (*a*) $Z_1$ (860–1065 nm), (*b*) $Z_2$ (1114–1265 nm), (*c*) $Z_3$ (1265–1668 nm), and (*d*) $Z_4$ (1830–2240 nm). Dependent variable values multiplied by 1000.**

composition (Ordóñez et al., 2013; Xue & Su, 2017). The results for $C_w$ estimation differed depending on grape variety. Tempranillo was the variety with the largest leaves and also had the highest $C_w$ and $C_{fm}$ values, but not the highest $C_{dm}$ value; it was thus the variety with the greatest water content and the greatest difference between fresh and dry matter content in the sample. These results were confirmed by representation of the spectral signatures for the leaves with the highest, medium and lowest $C_w$ values (Fig. 1). Spectral signatures for Tempranillo showed the highest differences for $C_w$ (Fig. 1). Those results corroborate the conclusion that Tempranillo was the most vigorous variety, as reported in studies of the same vine samples (González-Fernández, Marcelo, Valenciano, & Rodríguez-Pérez, 2012). Moreover, Tempranillo gave the highest correlation values for spectral data and $C_w$. This corroborates the findings of Strever (2005), who indicated that water content determination was influenced by the variety and its vegetative state. In this regard, Diago, Fernandes, Millan, Tardáguila, and Melo-Pinto (2013) developed a method for grapevine identification based on spectroscopy imagery acquired with a hyperspectral camera, classifying the Tempranillo, Grenache and Cabernet Sauvignon varieties as a function of the reflectance properties of their leaves and obtaining identification higher than 92% for all the varieties. De Bei et al. (2011) also found different values for estimated water content depending on the variety studied, while Strever (2005) showed that low-vigour vines compared to high-vigour vines produced spectral responses to water content estimates in shorter wavelengths.

On the other hand, $C_w$ estimations based on reflectance for Cabernet Sauvignon were not feasible, possibly because this variety is isohydric (Schultz, 2003), unlike Tempranillo, Merlot (Sánchez de Miguel, De la Fuente, Linares, Lissarrague, 2007; Gutiérrez, 2014) and Mencía (Baeza et al., 2011), which are anisohydric. In a water restriction period, Tempranillo, Merlot and Mencía would continue using available water, whereas

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

Cabernet would vary growth and physiology to conserve water, causing its leaves to be drier than in normal conditions. This is corroborated by Table 3, which shows that Cabernet was the variety with the highest $C_{dm}$, indicating that its leaves contained more dry matter than the other varieties. According to studies by Strever (2012), this situation may be due to a mixing of the water content and pigment signals, which would change the wavelengths most suitable to estimating water content.

### 4.2. *Vegetation indices*

In the regression performed with the vegetation indices, a relationship between $C_w$ and the vegetation indices was found for Tempranillo. Zhao et al. (2009) and Wang, Hunt, Qu, Hao, and Daughtry (2013) showed the usefulness of the NDII and SIWSI for the determination of water content in leaves. The WI, as one of the most studied vegetation indices, has been demonstrated to correlate highly with leaf water content. However, it has always been studied in crops cultivated in controlled conditions (Serrano et al., 2012). Serrano et al. (2010) also indicated that the WI is not an effective index to estimate water content for non-irrigated vines.

The relationship between vegetation indices and $C_w$ could be improved by optimising the wavelengths when computing spectral indices, as done by Verrelst et al. (2016) and Rivera, Verrelst, Delegido, Veroustraete, and Moreno (2014). Verrelst et al. (2016) obtained a suitable relationship to estimate the leaf area index and leaf chlorophyll from an empirical spectral index obtained by optimising wavelengths. Rivera et al. (2014) demonstrated that Gaussian process regression band analysis to optimise wavelengths could correctly predict leaf chlorophyll and the green leaf area index, because this tool identifies the most informative bands for a variable and determines the least number of bands that preserve a high predictive accuracy.

### 4.3. *Functional linear regression*

The models obtained using spectral ranges were more suitable for estimating $C_w$ than those obtained with the vegetation indices, for three main reasons: (1) functional data obtained a better prediction in terms of $R^2$ and *RMSE*; (2) functional regression coefficients allowed a better interpretation of the relationship between water content and reflectance; and (3) considering the spectral signature as a function enabled the dependence between reflectance values to be taken into account. Our findings are corroborated by previous research, at the same study site (González-Fernández, Rodríguez-Pérez, Marabel, & Álvarez-Taboada, 2015), that demonstrated that using spectral ranges pointed to stronger relationships than methods that use narrow bands.

The FLR results indicate that suitable spectral ranges for water content estimation in grapevines were $Z_3$ and $Z_4$ (centred on 1465 nm and 2035 nm, respectively). This is consistent with the findings of Santos and Kaye (2009) and Zhang, Li, and Zhang (2012), who indicated that 1400 nm ($Z_3$) and 1900 nm ($Z_4$) were the intervals most correlated with leaf water content. These bands correspond to the first overtone of O–H excitation for $H_2O$ and the combination of O–H and H–O–H deformation. Other bands associated with water content are also associated with cellulose and other organic leaf components (Shenk, Workman, & Westerhaus, 2001).

The FLR results were somewhat poorer for $Z_2$, given the values of $R^2$ and *RMSE*. However, the regression coefficient was smoother — even more so than those obtained for the rest of the zones, especially for $Z_4$. It could be said, nonetheless, that the result for $Z_2$ was satisfactory in terms of simplicity and interpretability. What remains clear from Fig. 3(*b*) is that longer wavelengths correspond to negative increases in water concentration.

Regarding the use of spectral ranges for all the varieties, $Z_4$ obtained a value for $R^2$ that was similar to that for $Z_3$, but the regression coefficient distribution for estimates of $C_w$ fitted using FLR and $Z_3$ was more stable and easier to interpret than that obtained with $Z_4$. As mentioned earlier, it is very difficult to determine whether the regression coefficient reflects the real relationship between reflectance and $C_w$ or whether there is a problem of overfitting. This is consistent with De Bei et al. (2011) who, using an ASD FS3 spectroradiometer, identified water absorption bands in the 1400–1450 nm region for Chardonnay, Cabernet Sauvignon and Shiraz leaves.

We report $R^2$ values of 0.70 for Tempranillo fitted by FLR to estimate water content. Research conducted in the same study area by González-Fernández et al. (2015), using PLSR as the statistical method, demonstrated that the most suitable models for predicting $C_w$ obtained $R^2$ values of 0.68. This agrees with De Bei et al. (2011), who obtained lower values using PLSR to estimate water in vineyards than we obtained in our study.

Finally, we suggest that models to estimate water content in vineyards based on using FLR as the fitting method and normalising reflectance using CR may be more suitable than models obtained using narrow bands or other statistical methods that use spectral ranges.

## 5. Conclusions

Using functional analysis regression, we analysed four different reflectance bands between 860 nm and 2240 nm to estimate the water content of leaves representing four grape varieties (Mencía, Cabernet Sauvignon, Merlot and Tempranillo) and to determine whether the results depended on wavelength. Our research demonstrates that field spectroscopy data processed by functional analysis regression was better able to predict leaf water content than ordinary least squares regression with vegetation indices. Furthermore, this method enabled a better understanding of the relationship between reflectance and leaf water content as a function of wavelength.

The best models were achieved for bands $Z_3$ (1265–1668 nm) and $Z_4$ (1830–2240 nm), previously continuum-removal transformed. However, the more complex relationship revealed for $Z_4$ made it more difficult to establish a simple interpretation of this dependence. The suitability of the models varied according to the grape variety, with the highest $R^2$ value obtained for Tempranillo.

We suggest that the detailed information obtained from the functional regression coefficients regarding the

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017

relationship between reflectance and leaf water content could be the basis for future implementation of a precise, rapid and non-destructive approach to detecting water stress in vineyards.

## Acknowledgements

This work was supported by the University of León [grant number 2016/00145/001-T102]. The authors acknowledge the assistance of the Ribas del Cúa SA winery and thank Bonsai Advanced Technologies SL for providing equipment and support.

## REFERENCES

Aguilera, A. M., Escabias, M., Mariano, J., Valderrama, M., & Aguilera-Morillo, C. (2013). Functional analysis of chemometric data. *Open Journal of Statistics*, 3(5), 334–343. http://dx.doi.org/10.4236/ojs.2013.35039.

ASD Inc. (2012). *FieldSpect4 user manual*. Boulder, Colorado USA http://support.asdi.com/Document/Viewer.aspx?id=163. Online; (Accessed 8 February 2016).

Baeza, P., Sánchez-Lirio, M., Verdugo, J., Tejerina, M. C., López-Pavón, C., Martín-Cuadrado, L., et al. (2011). Respuesta agronómica de diferentes cultivares tintos de vid en zona templada. *Agricultura*, 940, 328–332 (in Spanish).

Chaves, M. M., Zarrouk, O., Francisco, R., Costa, J. M., Santos, T., Regalado, A. P., et al. (2010). Grapevine under deficit irrigation – hints from physiological and molecular data. *Annals of Botany*, 105, 661–676. http://dx.doi.org/10.1093/aob/mcq030.

Cheng, T., Rivard, B., & Sánchez-Azofeifa, A. (2011). Spectroscopic determination of leaf water content using continuous wavelet analysis. *Remote Sensing of Environment*, 115, 659–670. http://dx.doi.org/10.1016/j.rse.2010.11.001.

Chone, X., Van Leeuwen, C., Dubordieu, D., & Gaudillère, J. P. (2001). Stem water potential is a sensitive indicator of grapevine water status. *Annals of Botany*, 87, 477–483. http://dx.doi.org/10.1006/anbo.2000.1361.

Clevers, J. G. P. W., Kooistra, L., & Schaepman, M. E. (2008). Using spectral information from the NIR water absorption features for the retrieval of canopy water content. *International Journal of Applied Earth Observation and Geoinformation*, 10, 388–397. http://dx.doi.org/10.1016/j.jag.2008.03.003.

Datt, B. (1999). Remote sensing of water content in eucalyptus leaves. *Australian Journal of Botany*, 47, 909–923. http://dx.doi.org/10.1071/BT98042.

De Bei, R., Cozzolino, D., Sullivan, W., Cynkar, W., Fuentes, S., Dambergs, R., et al. (2011). Non-destructive measurement of grapevine water potential using near infrared spectroscopy. *Australian Journal of Grape and Wine Research*, 17, 62–71. http://dx.doi.org/10.1111/j.1755-0238.2010.00117.x.

Diago, M. P., Fernandes, A. M., Millan, B., Tardáguila, J., & Melo-Pinto, P. (2013). Identification of grapevine varieties using leaf spectroscopy and partial least squares. *Computers and Electronics in Agriculture*, 99, 7–13. http://dx.doi.org/10.1016/j.compag.2013.08.021.

Dias, R., García, N. L., Ludwig, G., & Saraiva, M. (2015). Aggregated functional data model for near-infrared spectroscopy calibration and prediction. *Journal of Applied Statistics*, 42, 127–143. http://dx.doi.org/10.1080/02664763.2014.938224.

Dobrowski, S. Z., Pushnik, J. C., Zarco-Tejada, P. J., & Ustin, S. L. (2005). Simple reflectance indices track heat and water stress-induced changes in steady-state chlorophyll fluorescence at the canopy scale. *Remote Sensing of Environment*, 97, 403–414. http://dx.doi.org/10.1016/j.rse.2005.05.006.

Dobrowski, S. Z., Ustin, S. L., & Wolpert, J. A. (2002). Remote estimation of vine canopy density in vertically shoot-positioned vineyards: Determining optimal vegetation indices. *Australian Journal of Grape and Wine Research*, 8, 117–125. http://dx.doi.org/10.1111/j.1755-0238.2002.tb00220.x.

Dobrowski, S. Z., Ustin, S. L., & Wolpert, J. A. (2003). Grapevine dormant pruning weight prediction using remotely sensed data. *Australian Journal of Grape and Wine Research*, 9, 177–182. http://dx.doi.org/10.1111/j.1755-0238.2003.tb00267.x.

Fensholt, R., & Sandholt, I. (2003). Derivation of a shortwave infrared water stress index from MODIS near-and shortwave infrared data in a semiarid environment. *Remote Sensing of Environment*, 87, 111–121. http://dx.doi.org/10.1016/j.rse.2003.07.002.

Fuentes, D., Gamon, J., Qiu, H. L., Sims, D., & Roberts, D. (2001). Mapping Canadian boreal forest vegetation using pigment and water absorption features derived from the AVIRIS sensor. *Journal of Geophysical Research*, 106, 33565–33577. http://dx.doi.org/10.1029/2001JD900110.

Gamon, J. A., Peñuelas, J., & Field, C. B. (1992). A narrow-wave band spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, 41, 35–44.

Gao, B. (1996). NDWI – a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257–266. http://dx.doi.org/10.1016/S0034-4257(96)00067-3.

González-Fernández, A. B., Marcelo, V., Valenciano, J. B., & Rodríguez-Pérez, J. R. (2012). Relationship between physical and chemical parameters for four commercial grape varieties from the Bierzo region (Spain). *Scientia Horticulturae*, 147, 111–117. http://dx.doi.org/10.1016/j.scienta.2012.09.009.

González-Fernández, A. B., Rodríguez-Pérez, J. R., Marabel, M., & Álvarez-Taboada, F. (2015). Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression. *Scientia Horticulturae*, 188, 15–22. http://dx.doi.org/10.1016/j.scienta.2015.03.012.

Gutiérrez, G. I. (2014). *Caracterización del intercambio gaseoso en cuatro cultivares vitícolas (Isohídricos y Anisohídricos) en el valle del Maule* (Master dissertation). Talca, Chile: Universidad de Talca http://dspace.utalca.cl/handle/1950/9967 (in Spanish). Online; (Accessed 12 May 2016).

Hall, A., Lamb, D. W., Holzapfel, B., & Louis, J. (2002). Optical remote sensing applications in viticulture – a review. *Australian Journal of Grape and Wine Research*, 8, 36–47. http://dx.doi.org/10.1111/j.1755-0238.2002.tb00209.x.

Hardisky, M. A., Klemas, V., & Smart, R. M. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of Spartina alterniflora canopies. *ISPRS Journal of Photogrammetry and Remote Sensing*, 49, 77–83.

Kokaly, R. F., Asner, G. P., Ollinger, S. V., Martin, M. E., & Wessman, C. A. (2009). Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment*, 113, 78–91. http://dx.doi.org/10.1016/j.rse.2008.10.018.

Kokaly, R. F., & Clark, R. N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise linear regression. *Remote Sensing of Environment*, 67, 267–287. http://dx.doi.org/10.1016/S0034-4257(98)00084-4.

Moshou, D., Pantazi, X., Kateris, D., & Gravalos, I. (2014). Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier. *Biosystems Engineering*, 117, 15–22. http://dx.doi.org/10.1016/j.biosystemseng.2013.07.008.

Ordóñez, C., Martínez, J., Matías, J. M., Reyes, A. N., & Rodríguez-Pérez, J. R. (2010). Functional statistical techniques applied to vine leaf water content. *Mathematical and Computer Modelling, 52*(7–8), 116–1122. http://dx.doi.org/10.1016/j.mcm.2010.03.008.

Ordóñez, C., Rodríguez-Pérez, J. R., Moreira, J. J., & Sanz, E. (2013). Using hyperspectral spectrometry and functional models to characterize vine-leaf composition. *IEEE Transactions on Geosciences and Remote Sensing, 51*(5), 2610–2618. http://dx.doi.org/10.1109/TGRS.2012.2217344.

Oumar, Z., & Mutanga, O. (2010). Predicting plant water content in *Eucalyptus grandis* forest stands in KwaZulu-Natal, South Africa using field spectra resampled to the Sumbandila Satellite Sensor. *International Journal of Applied Earth Observation and Geoinformation, 12*, 158–164. http://dx.doi.org/10.1016/j.jag.2010.02.002.

Palmer, K., & Williams, D. (1974). Optical properties of water in the near infrared. *Journal of the Optical Society of America, 64*, 1107–1110. http://dx.doi.org/10.1364/JOSA.64.001107.

Peñuelas, J., Baret, F., & Filella, I. (1995). Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica, 31*, 221–230.

Peñuelas, J., Pinol, J., Ogaya, R., & Filella, I. (1997). Estimation of plant water concentration by the reflectance water index WI (R900/R970). *International Journal of Remote Sensing, 18*, 2869–2875. http://dx.doi.org/10.1080/014311697217396.

Ramsay, J. O., & Silverman, B. W. (2002). *Applied functional data analysis. Methods and case studies*. New York: Springer-Verlag.

Reiss, P., & Ogden, T. (2007). Functional principal component regression and functional partial least squares. *Journal of the American Statistical Association, 102*, 984–996. http://dx.doi.org/10.1198/016214507000000527.

Rivera, J. P., Verrelst, J., Delegido, J., Veroustraete, F., & Moreno, J. (2014). On the semi-automatic retrieval of biophysical parameters based on spectral index optimization. *Remote Sensing, 6*, 4927–4951.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In *Proceedings of third earth resources technology Satellite-1 symposium* (Vol. 1, pp. 309–317).

Saeys, W., Keteleare, B. De, & Darius, P. (2008). Potential applications of functional data analysis in chemometrics. *Journal of Chemometrics, 22*(5), 335–344. http://dx.doi.org/10.1002/cem.1129.

Sánchez de Miguel, P., de la Fuente, M., Linares, R., & Lissarrague, J. R. (2007). Effects of water potential and relative humidity on leaf photosynthesis response to PAR light in Cabernet Sauvignon and Tempranillo cultivars during berry maturation. In *Procedings of XVth Internacional Gesco Symposium* (Vol. 1, pp. 602–608).

Santos, A. O., & Kaye, O. (2009). Grapevine leaf water potential based upon near infrared spectroscopy. *Scientia Agricola, 66*, 287–292. http://dx.doi.org/10.1590/S0103-90162009000300001.

Scholander, P. F., Hammel, H. J., Bradstreet, A., & Hemmingsen, E. A. (1965). Sap pressure in vascular plants. *Science, 148*, 339–346.

Schultz, H. R. (2003). Differences in hydraulic architecture account for near-isohydric and anisohydric behaviour of two field-grown *Vitis vinifera* L. cultivars during drought. *Plant, Cell & Environment, 26*, 1393–1405. http://dx.doi.org/10.1046/j.1365-3040.2003.01064.x.

Serrano, L., González-Flor, C., & Gorchs, G. (2010). Assessing vineyard water status using the reflectance based water index. *Agriculture, Ecosystems & Environment, 139*, 490–499. http://dx.doi.org/10.1016/j.agee.2010.09.007.

Serrano, L., González-Flor, C., & Gorchs, G. (2012). Assessment of grape yield and composition using the reflectance based water index in Mediterranean rainfed vineyards. *Remote Sensing of Environment, 118*, 249–258. http://dx.doi.org/10.1016/j.rse.2011.11.021.

Shenk, J. S., Workman, J., & Westerhaus, M. O. (2001). Application of NIR spectroscopy to agricultural products. In D. A. Burns, & E. D. Ciurczak (Eds.), *Handbook of near-infrared analysis (practical spectroscopy series)* (2nd ed., pp. 419–474). New York: Marcel Dekker.

Sims, D. A., & Gamon, J. A. (2003). Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: A comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sensing of Environment, 84*, 526–537. http://dx.doi.org/10.1016/S0034-4257(02)00151-7.

Strachan, I. B., Pattey, E., & Boisvert, J. B. (2002). Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. *Remote Sensing of Environment, 80*, 213–224. http://dx.doi.org/10.1016/S0034-4257(01)00299-1.

Strever, A. E. (2005). Estimating water stress in *Vitis vinifera* l. using field spectrometry: A preliminary study incorporating multispectral vigour classification. In *Proceedings of the conference FRUTIC 05, information and technology for sustainable fruit and vegetable production. Montpellier, France*.

Strever, A. E. (2012). *Non-destructive assessment of leaf composition as related to growth of the grapevine* (Vitis vinifera *L. cv. Shiraz)* (PhD diss.). Stellenbosch, South Africa: Stellenbosch University.

Strever, A. E., Bezuidenhout, D., Zorer, R., Moffat, T., & Hunter, J. J. (2012). Optical and thermal applications in grapevine (*Vitis vinifera* L.) research – an overview and some novel approaches. *SAIEE African Research Journal, 103*, 55–60.

Ustin, S. L., Riaño, D., & Hunt, E. R. (2012). Estimating canopy water content from spectroscopy. *Israel Journal of Plant Sciences, 60*, 9–23. http://dx.doi.org/10.1560/IJPS.60.1-2.9.

Verrelst, J., Rivera, J. P., Gitelson, A., Delegido, J., Moreno, J., & Camps-Valls, G. (2016). Spectral band selection for vegetation properties retrieval using Gaussian processes regression. *International Journal of Applied Earth Observation and Geoinformation, 52*, 554–567.

Wang, L., Hunt, E. R., Qu, J. J., Hao, X., & Daughtry, C. S. T. (2013). Remote sensing of fuel moisture content from ratios of narrow-band vegetation water and dry-matter indices. *Remote Sensing of Environment, 129*, 103–110. http://dx.doi.org/10.1016/j.rse.2012.10.027.

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors, 2017*, 1–17. http://dx.doi.org/10.1155/2017/1353691.

Zarco-Tejada, P. J., & Ustin, S. L. (2001). Modeling canopy water content for carbon estimates from MODIS data at land EOS validation sites. In *Proceedings of geoscience and remote sensing symposium* (Vol. 1, pp. 342–344). http://dx.doi.org/10.1109/IGARSS.2001.976152.

Zhang, Q., Li, Q., & Zhang, G. (2012). Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection. *Journal of Spectroscopy, 27*, 93–105. http://dx.doi.org/10.1155/2012/276795.

Zhao, S., Qin, Q., You, L., Yao, Y., Yang, N., & Li, J. (2009). Application of two shortwave infrared water stress indices to drought monitoring over Northwestern China. In *Proceedings of IEEE International Geoscience & remote sensing Symposium (IGARSS 2009). Cape Town, South Africa*. http://dx.doi.org/10.1109/IGARSS.2009.5417809.

Please cite this article in press as: Rodríguez-Pérez, J. R., et al., Leaf water content estimation by functional linear regression of field spectroscopy data, *Biosystems Engineering* (2017), http://dx.doi.org/10.1016/j.biosystemseng.2017.08.017