

Document downloaded from:

<http://hdl.handle.net/10251/37059>

This paper must be cited as:

Martí Pérez, PC.; Gasque Albalate, M.; González Altozano, P. (2013). An artificial neural network approach to the estimation of stem water potential from frequency domain reflectometry soil moisture measurements and meteorological data. *Computers and Electronics in Agriculture*. 91:75-86. doi:10.1016/j.compag.2012.12.001.




The final publication is available at

<http://dx.doi.org/10.1016/j.compag.2012.12.001>

Copyright Elsevier

**AUTHOR QUERY FORM**

 <b>ELSEVIER</b>	<b>Journal: COMPAG</b>  <b>Article Number: 2809</b>	<b>Please e-mail or fax your responses and any corrections to:</b>  <b>E-mail: <a href="mailto:corrections.esch@elsevier.sps.co.in">corrections.esch@elsevier.sps.co.in</a></b>  <b>Fax: +31 2048 52799</b>
--	---	---

Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. Note: if you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file. To ensure fast publication of your paper please return your corrections within 48 hours.

For correction or revision of any artwork, please consult <http://www.elsevier.com/artworkinstructions>.

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Click on the 'Q' link to go to the location in the proof.

<b>Location in article</b>	<b>Query / Remark: <a href="#">click on the Q link to go</a> Please insert your reply or correction at the corresponding line in the proof</b>
<u><a href="#">Q1</a></u>  <u><a href="#">Q2</a></u>	<p>Please confirm that given names and surnames have been identified correctly.</p> <p>Kindly check whether the identification of corresponding author and the e-mail address are okay as typeset, and correct if necessary.</p> <div data-bbox="416 1853 979 1953" style="border: 1px solid black; padding: 5px; margin-top: 20px;"> <p style="color: red;">Please check this box if you have no corrections to make to the PDF file</p> <input data-bbox="868 1868 940 1932" type="checkbox"/> </div>

Thank you for your assistance.

---

**Highlights**

---

► First step approach to estimate stem water potential from soil moisture and standard meteorological variables using ANNs. ► Two principal components are enough to describe systematic variability of data. ► Optimum input combination: temperature, relative humidity, solar radiation and soil moisture at 50 cm. ► Artificial neural networks present higher accuracy than corresponding multi-linear regression models.

---



Contents lists available at SciVerse ScienceDirect

# Computers and Electronics in Agriculture

journal homepage: [www.elsevier.com/locate/compag](http://www.elsevier.com/locate/compag)



## An artificial neural network approach to the estimation of stem water potential from frequency domain reflectometry soil moisture measurements and meteorological data

Pau Martí<sup>a,\*</sup>, María Gasque<sup>b</sup>, Pablo González-Altozano<sup>a</sup>

<sup>a</sup>Departamento de Ingeniería Rural y Agroalimentaria, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain

<sup>b</sup>Departamento de Física Aplicada, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain

### ARTICLE INFO

#### Article history:

Received 26 May 2012

Received in revised form 28 October 2012

Accepted 3 December 2012

Available online xxx

#### Keywords:

Frequency domain reflectometry

Irrigation scheduling

Principal components analysis

Artificial neural networks

### ABSTRACT

Stem water potential seems to be a sensitive measure of plant water status. Nonetheless, it is a labour-intensive measurement and is not suited for automatic irrigation scheduling or control. This study describes the application of artificial neural networks to estimate stem water potential from soil moisture at different depths and standard meteorological variables, considering a limited data set. The experiment was carried out with 'Navelina' citrus trees grafted on 'Cleopatra' mandarin. Principal components analysis and multiple linear regression were used preliminarily to assess the relationships among observations and to propose other models to allow a comparative analysis, respectively. Two principal components account for the systematic data variation. The optimum regression equation of stem water potential considered temperature, relative humidity, solar radiation and soil moisture at 50 cm as input variables, with a determination coefficient of 0.852. When compared with their corresponding regression models, ANNs presented considerably higher performance accuracy (with an optimum determination coefficient of 0.926) due to a higher input-output mapping ability.

© 2012 Published by Elsevier B.V.

### 1. Introduction

Modern irrigated agriculture aims to maximize production while maintaining quality standards, keeping costs as low as possible and preserving the environment. The development of better irrigation procedures is necessary to overcome present and future water restrictions arising from increasing demand and the limited existing resources. So, improving water use efficiency is a crucial issue for sustainable crop production. Precise irrigation scheduling is necessary to achieve this goal. Therefore, the development of more accurate scheduling approaches is essential to guarantee a sustainable water supply for optimal crop growing.

The development of different sensors and techniques to monitor climatic conditions, soil water content or plant water status has led to great advances in irrigation scheduling techniques (Allen et al., 1998; Hanson et al., 2000; Goldhamer and Fereres, 2001; Dane and Topp, 2002; Intrigliolo and Castel, 2004; Jones, 2004). In most of these, soil water status is considered a key factor for planning irrigation doses (Campbell and Campbell, 1982).

Advantages of soil moisture monitoring include determining soil moisture depletion, adequacy of irrigation wetting, patterns of soil moisture extraction due to root uptake of water and trends

in soil moisture content with time during the irrigation season (Hanson et al., 2000). Among the different techniques that allow continuous measurement of the soil water content, the **frequency domain reflectometry** (FDR) probe, with multiple depth capacitance sensors (Paltineanu and Starr, 1997; Fares and Polyakov, 2006), has so far shown excellent performance and is currently a widely used soil moisture sensor for field applications. Therefore, irrigation scheduling has been carried out using FDR measurements in recent studies (Mounzer et al., 2008a, 2008b). However, optimization of irrigation scheduling using soil moisture sensors requires accurate threshold values for individual crops in the considered location. Thresholds for irrigation scheduling using soil water stress indicators are site-specific. Apart from this, there is no clear procedure for the determination of the thresholds. Many plant physiological features respond directly to changes in the plant tissues, rather than to changes in soil water content (Jones, 2004). This makes it difficult to use these probes in irrigation scheduling, especially in regulated deficit irrigation (RDI), where exposing plants to certain levels of stress promotes the accumulation of carbohydrates in fruits (González-Altozano and Castel, 2000).

As an alternative to soil water content monitoring, some studies suggest that plant-based measurements offer higher precision in the application of irrigation water (Jones, 2004). Leaf water potential ( $\psi$ ), measured with a pressure chamber (Peretz et al., 1984), is

\* Corresponding author.

E-mail address: [paumarpe@doctor.upv.es](mailto:paumarpe@doctor.upv.es) (P. Martí).

an optimum indicator of plant water status (Elfvig et al., 1972; Garnier and Berger, 1985; Améglio et al., 1999; González-Altozano and Castel, 1999, 2000). Numerous studies (Choné et al., 2001; Naor, 2006; Ortuño et al., 2006; Gasque et al., 2010) consider the water potential measured in bag-covered (non-transpiring) leaves, named stem water potential ( $\psi_{st}$ ), because it is a significant and more reliable indicator of water status and early water deficit in plants. Further, it offers less variability and seems to be well-related to tree and fruit growth and quality in a wide range of soils and under different irrigation systems. So,  $\psi_{st}$  is a useful and valuable tool for irrigation scheduling in fruit trees, and can be used to set a threshold over which irrigation should be applied, both with conventional irrigation, and with RDI strategies, allowing the saving of water without affecting the production or the quality of the fruit.

Nevertheless, leaf and stem water potential are not easily measurable, are not suited for automatic irrigation scheduling or control, and are destructive as well as time- and labour-consuming. Moreover, it is not possible either to take measurements continuously or to automate the data collection. Thus, the estimation of leaf and stem water potential becomes a task of great importance, as it can play a decisive role in the application of more efficient scheduling strategies. Several authors have proposed models to estimate these parameters. Steppe et al. (2008) proposed a methodology to predict when and how much water is required by the plant based on modelling stem water potential and continuous measurements of sap flow and stem diameter variation. Santos and Kaye (2009) applied near-infrared spectroscopy and multivariate data analysis for the modelling and estimation of leaf water potential in grapevines. Acevedo-Opazo et al. (2010) established and tested a model to extrapolate pre-dawn leaf water potential measured at a reference location to other unsampled locations using a linear combination of spatial ancillary information sources and a reference measurement. Dziki et al. (2010) used remote sensing spectral indices to predict leaf and stem water potential for Satsuma mandarin trees subjected to different drought stress regimes.

Artificial neural networks (ANNs) are connectionist information-processing systems inspired by the biological neural system, with the capability to learn, store and recall information based on a given training data set. They are able to perform non-linear mapping of a multidimensional input space onto another multidimensional output space, requiring no detailed information about the system. As in biological systems, knowledge is acquired during the learning process by adjusting the synaptic connections that exist between the neurons.

They are known to be efficient and less-consuming in the modelling of complex systems compared with other mathematical models such as regression. So far, the application of ANNs in irrigation scheduling issues is very limited. Zakaluk and Ranjan (2007) proposed a model to predict leaf water potential using RGB digital images. Capraro et al. (2008) developed a prediction model to determine the irrigation time necessary to take the soil moisture level up to a user-desired level. To our knowledge, ANNs have still not been used to estimate stem water potential from soil moisture and meteorological data. This paper describes a methodology based on ANNs to estimate stem water potential. In particular, this work describes two procedures to extract more effectively the knowledge contained in a data set, namely, a method to ensure that the training and cross-validating sets are representative, and a 'leave one out' (LOO) procedure to ensure a complete test scan of all data points. This is especially suited to dealing with limited data sets. The experimental data for the study correspond to a commercial orchard of 'Navelina' citrus trees grafted on 'Cleopatra' mandarin.

First, the relationships between stem water potential, soil moisture at different depths and the meteorological variables mean air temperature, wind speed, relative humidity and solar radiation are analyzed using Principal Components Analysis (PCA). In a second stage, the performance accuracy of multiple linear regression (MLR) and ANN models for estimating stem water potential is compared. The development of a general model that can be extrapolated and applied under a wide spectrum of conditions is beyond the scope of the present work.

## 2. Materials and methods

### 2.1. Experimental setup

#### 2.1.1. Plot description and meteorological data

The experiment was carried out from July 2009 to October 2010 on a commercial drip-irrigated plot of 1 ha in Senyera, Valencia (Spain) (39°3'N, 0°30'W, 23 m a.s.l.) planted in 1982 with 'Navelina' orange trees (*Citrus sinensis* L. Osbeck) grafted on 'Cleopatra' mandarin (*Citrus reshni* Hort.) at a spacing of 5 × 5 m.

The soil is deep sandy-loam with pebbles of alluvial origin. According to FAO-UNESCO (1988) and LUCDEME (2001), this soil is classified as Calcaric Fluvisol. Its physical properties are summarized in Table 1. The most common probes usually provide volumetric values of these parameters, which depend on the soil type. The relationship between volumetric moisture and matric potential is required for extrapolation and to carry out a comparison with other soils. Table 2 summarises the hydraulic properties of the soil, at depths of 10 cm, 30 cm and 50 cm, in accordance with Saxton et al. (1986). Table 3 presents the relationships between volumetric moisture and matric potential according to Campbell (1974).

Five meteorological variables, namely mean air temperature  $T$  (°C), wind speed at two meter height  $U$  ( $\text{ms}^{-1}$ ), solar radiation  $R_s$  ( $\text{Wm}^{-2}$ ), air relative humidity RH (%) and effective precipitation  $E_p$  (mm) were recorded every 30 min by an automatic weather station of the Irrigation Technology Service (Valencian Institute for Agricultural Research) in Villanueva de Castellón, less than 1000 m from the experimental plot.

#### 2.1.2. Irrigation scheduling

A Regulated Deficit Irrigation (RDI) treatment was conducted in the experimental plot, seeking to cause wider  $\psi_{st}$  ranges in the considered trees. Accordingly, the RDI consisted of water reduction applied from mid-July until the beginning of September, whereas during the rest of the year crop water demand was fully covered.

Crop evapotranspiration ( $ET_c$ ) was estimated by multiplying reference evapotranspiration ( $ET_o$ ) by a crop coefficient ( $K_c$ ), obtained for this location from Castel (2001) and which depends on the percentage of shaded area in the plot. The crop coefficient was 0.59 in 2009 and 0.54 in 2010. The theoretical irrigation dose ( $I_d$ ) to ensure full irrigation was calculated as  $I_d = ET_c - E_p$ . The irrigation strategy applied during the period of study is shown in Table 4. These doses

**Table 1**  
Physical properties of the soil.

Parameter	10 cm	30 cm	50 cm
Sand	60	54	46
Clay	18	19	22
Organic matter (%)	0.90	0.80	0.70
Electrical conductivity (at 25 °C) (dS/m)	0.99	0.92	0.98
$\rho_{ap}$ ( $\text{g/cm}^3$ )	1.42	1.42	1.40

$\rho_{ap}$ : Bulk density.

**Table 2**  
Hydraulic properties of the soil (Saxton et al., 1986).

Reference	10 cm		30 cm		50 cm	
	$\theta_v$ (%)	$K_H$ (mm/h)	$\theta_v$ (%)	$K_H$ (mm/h)	$\theta_v$ (%)	$K_H$ (mm/h)
Wilting point (–1500 kPa)	11.4	5.79E–08	12.0	1.60E–07	13.8	4.88E–07
Field Capacity (–33 kPa)	22.2	8.88E–04	23.9	2.90E–03	27.3	8.90E–03
Saturation (0 kPa)	46.4	35.11	46.5	29.28	47.1	20.81

$\theta_v$ : Volumetric moisture,  $K_H$ : hydraulic conductivity.

**Table 3**  
Relationship between volumetric soil water content and soil matric potential according to Campbell (1974).

Depth (cm)	Predictive equation	$r^2$	RMSE (kPa)
10	$\Psi_m = -0.457 \cdot (\theta_v / \theta_{vs})^{-5.765}$	0.9997	11.37
30	$\Psi_m = -0.789 \cdot (\theta_v / \theta_{vs})^{-5.573}$	0.9998	9.36
50	$\Psi_m = -1.379 \cdot (\theta_v / \theta_{vs})^{-5.694}$	0.9998	8.26

$\Psi_m$ : Matric potential (kPa),  $\theta_v$ : volumetric moisture (%),  $\theta_{vs}$ : saturated volumetric moisture (%).

**Table 4**  
Deficit scheduling pattern adopted.

Restricted irrigation dose	Data points	Period
40% of $I_d$	1–7	17/07/2009–13/09/2009
100% of $I_d$	8–30	14/09/2009–30/05/2010
80% of $I_d$	31–34	31/05/2010–11/07/2010
32% of $I_d$	35–40	12/07/2010–30/08/2010
80% of $I_d$	41–46	31/08/2010–31/10/2010

$I_d$ : irrigation dose.

were established based on previous studies in the same orchard, where a restriction of 40% during this period did not affect either production or quality (Gasque et al., 2010).

The irrigation system consisted of a double line (1.8 m spaced) of drip-irrigation with eight self-regulating drippers per tree with an average flow of  $7.4 \text{ L h}^{-1}$  per dripper.

### 2.1.3. Soil water status measurements

The volumetric water content through the soil profile was monitored in continuous real time using a multisensor capacitance probe (ENVIROSCAN, Sentek Sensor Technologies) based on FDR. The probe was placed 25 cm from the emitter’s line inside a PVC access tube installed within the wetted area of the RDI trees to record different dynamic variations and a wide range of soil water content within the zone covered by FDR sensors. The probe, installed in the north side of the tree, presented sensors at 10, 30, 50 and 70 cm depth ( $h_1$ ,  $h_2$ ,  $h_3$ , and  $h_4$ , respectively), where the first three sensors covered practically 90% of the active root system (estimated during the installation of the probes). Fig. 1 shows a scheme of the probe installation and the measurement points.

The PVC access tube of the probe was installed carefully to ensure good contact between the tube and the soil. The bottom of the PVC access tube was plugged with a rubber bung to prevent water and water vapour entering the tube. The capacitance probe was properly installed within the active root system zone and the soil water distribution of the emitters was highly uniform.

Data were stored in a data logger and were transmitted via GPRS every 8 h to a central server which was permanently connected. The stored raw data were graphically displayed as volumetric water content. The probe was calibrated prior to installation. The water content in each soil layer, expressed in mm, was calculated by multiplying the measured volumetric moisture content by the thickness (m) of the corresponding soil layer (0.2 m).

### 2.1.4. Tree water status measurements

Midday stem water potential ( $\psi_{st}$ ) was measured weekly around 12 h (GMT) using a pressure chamber following the procedures described by Turner (1981).  $\psi_{st}$  was measured on 4 leaves of the tree, 2 leaves at 2 different heights, (2/4 and 3/4 of the tree height) on the south quadrant of the tree, where the FDR probe was installed. The leaves had been wrapped in bags at least 2 h previously. Afterwards a mean value was calculated. During winter the measurement frequency was occasionally reduced to every fortnight.

## 2.2. Input selection

The complete data set consists of 46 data points corresponding to 46 weekly measurements between 17/07/2009 and 31/10/2010 (Table 4). Moreover, the day of the year ( $J$ ) was also considered. Data were arranged as a matrix of 46 weeks (in rows) by 10 variables, namely  $T$ ,  $U$ ,  $RH$ ,  $R_s$ ,  $J$ ,  $h_1$ ,  $h_2$ ,  $h_3$ ,  $h_4$  and  $\psi_{st}$  (in columns). The individual values of these variables along the period of study are depicted in Fig. 2.

PCA and step-wise MLR were used to study the relationships among observations and variables respectively, as well as to select the inputs to feed the models.

### 2.2.1. Principal component analysis

PCA is a useful technique for explaining the data variability of a matrix, as well as for the interpretation of relationships among observations and variables (Pearson, 1901; Jolliffe, 2002). It allows the identification of underlying correlation structures, often called latent variables. Ample literature can be found about its theoretical fundamentals and applications (Pearson, 1901; Jolliffe, 2002; Jackson, 1991).

Data scaling to unitary variance is commonly applied prior to PCA and is recommended if the variance is very different among variables. This avoids variables with the highest variance having an excessive influence on the components.

One criterion frequently applied to determine how many components should be extracted in order to model the data variability is cross-validation (Diana and Tommasi, 2002; Krzanowski, 1987; Wold, 1978). It considers that one PC provides relevant information if it does not change significantly when several observations are removed. This criterion is implemented in the software SIMCA-P 10.0 (Umetrics AB, Malmö, Sweden) that was used to carry out all the PCA models. This software uses the NIPALS algorithm and considers that a certain component provides systematic information if the goodness-of-fit for that component obtained by cross-validation,  $Q^2$ , is higher than a certain threshold (Eriksson et al., 1999).

Score plots with different combinations of PCs were visually inspected in order to identify outliers. The distance of observations to the PCA model was also checked.

### 2.2.2. Step-wise multiple linear regression

After the PCA analysis, step-wise MLR was applied to determine the significance level of each independent variable and to fit a

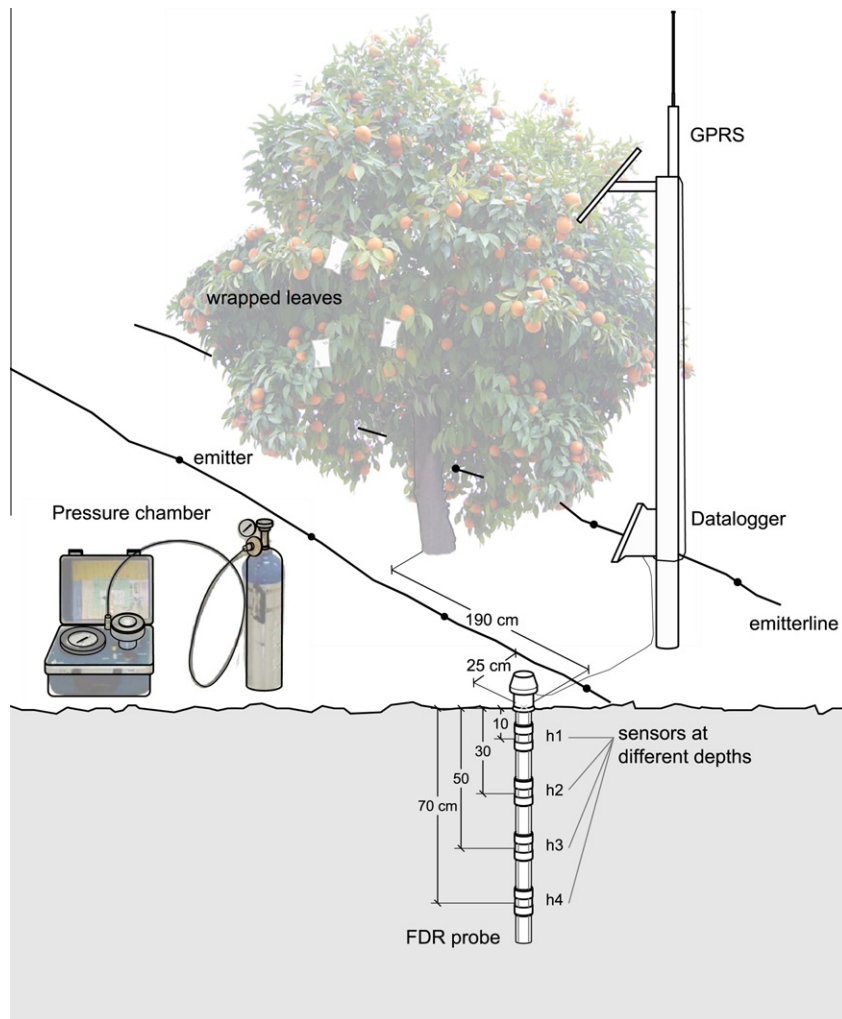


Fig. 1. Scheme of probe installation and measurement points.

282 predictive equation. The software Statgraphics plus 5.1 (StatPoint  
283 Technologies Inc., Warrenton, VA, USA) was used to conduct all  
284 the regression models.

285 2.3. Artificial neural networks

286 2.3.1. Model description

287 In this work, multilayer feed-forward networks with back-prop-  
288 agation are used. The back-propagation algorithm and its variants  
289 are supervised training rules that involve an iterative procedure to  
290 move the weights along the negative gradient of the error function,  
291 so as to minimize the difference between the network's output and  
292 the desired target. In each step, errors are used as inputs to feed-  
293 back connections from which adjustments are made to the synap-  
294 tic weights layer by layer in a backward direction. The weight  
295 associated with each connection is adjusted by an amount propor-  
296 tional to the strength of the signal on the connection and the total  
297 measure of the error. All ANN neurons used were configured on the  
298 model by Haykin (1999). The hidden and output neurons can be  
299 mathematically characterized with the following equations:

302 
$$v_k = \sum_{j=1}^n w_{kj}x_j + b_k \quad (1)$$

303 
$$y_k = \varphi(v_k) \quad (2)$$

306 where  $x_j$  is the input signal;  $w_{kj}$  is the synaptic weight of neuron  $k$ ;  
307  $v_k$  is the linear combiner or summing junction;  $b_k$  is the bias;  $y_k$  is  
308 the output of the neuron and  $\varphi$  is the transfer function. This study  
309 considered two of the most common transfer functions: the hyper-  
310 bolic tangent sigmoid (tansig) and the logarithmic sigmoid (logsig)  
311 functions. If the output layer of the network has sigmoid neurons,  
312 then the output values are limited to a small range. So, linear output  
313 neurons were used.

314 The Levenberg-Marquardt algorithm (Marquardt, 1963) was  
315 used to train the networks as it is suitable to deal with ill-condi-  
316 tioned minimization problems. So, an approximation to the Hes-  
317 sian matrix is used in the Newton-like update:

318 
$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e_k \quad (3)$$

321 where  $\mu$  governs the step size and  $I$  is the unit matrix;  $J$  is the Jaco-  
322 bian matrix that contains first derivatives of the network errors  
323 with respect to the weights and biases, and  $e$  is a vector of network  
324 errors (Hagan and Menhaj, 1994; Hagan et al., 1996).

325 A risk of supervised learning is its tendency to over-fit the infor-  
326 mation contained in the training data set if additional stopping cri-  
327 teria are not considered, which can move on to a lack of  
328 generalizability in the model (Prechelt, 1998). So, in order to avoid  
329 over-fitting of the training data and hence to improve the general-  
330 ization of the networks, the early stopping technique was used in  
331 conjunction with the training algorithm. To this end, training data

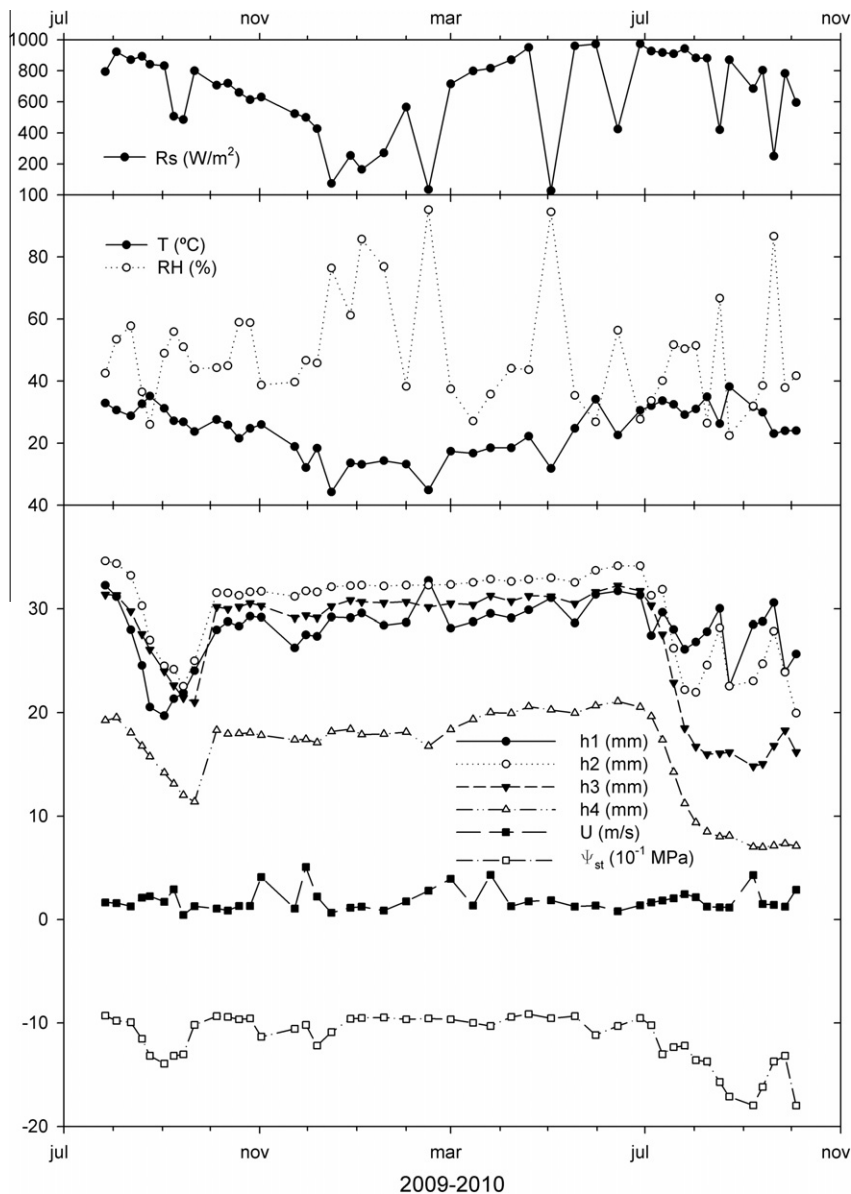


Fig. 2. Individual values of measured variables.

series were divided into 2 groups: one for learning/parameter estimation and one for cross-validation. According to this method, while the chosen error, the mean squared error, of the cross-validation set was lower than its value in the previous iteration, the training of the network proceeded; if not, the training was finished. The training parameters considered for the application of the Levenberg-Marquardt algorithm can be found in Martí et al. (2011). These are default standard values for the current ANN configuration proposed by the software, Matlab version 7.4.0 (Matlab, 2007).

### 2.3.2. ANN model design and implementation

A 'leave one out' procedure (LOO) was applied by leaving each time only a single data point for testing, given that the number of available training patterns is limited. With this method, it is possible to scan the performance of the model throughout the complete data set (Fig. 3). Furthermore, it makes it possible to extract more effectively the knowledge contained in the available data series, as it considers at each stage the smallest test set size to evalu-

ate the model performance. Nonetheless, the smaller the test sets considered are, the more stages (training-test processes) are involved in the LOO procedure, with a subsequent rise in computational costs. In this case, 46 training, cross-validating and test data sets were defined.

In order to perform a suitable learning process, a representative collection of input-output examples must be included in the training and cross-validating data sets. To ensure this representativeness inside a limited set of patterns, training and cross-validating data were assigned as follows at each LOO stage. Excluding the test point, the 45 available patterns were sorted according to an increasing range of  $\psi_{st}$ . Afterwards, the total  $\psi_{st}$  range was divided into 5 equally spaced intervals, if at least 2 data points were available per interval. If not, the interval containing only 1 data point was widened until it contained 2 points (reaching the nearest one to the initial point). One pattern was randomly selected per interval and the rest, until the assigned cross-validating percentage was completed, were randomly selected from 2 of the central intervals (those with a higher number of points). The rest of the

350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368



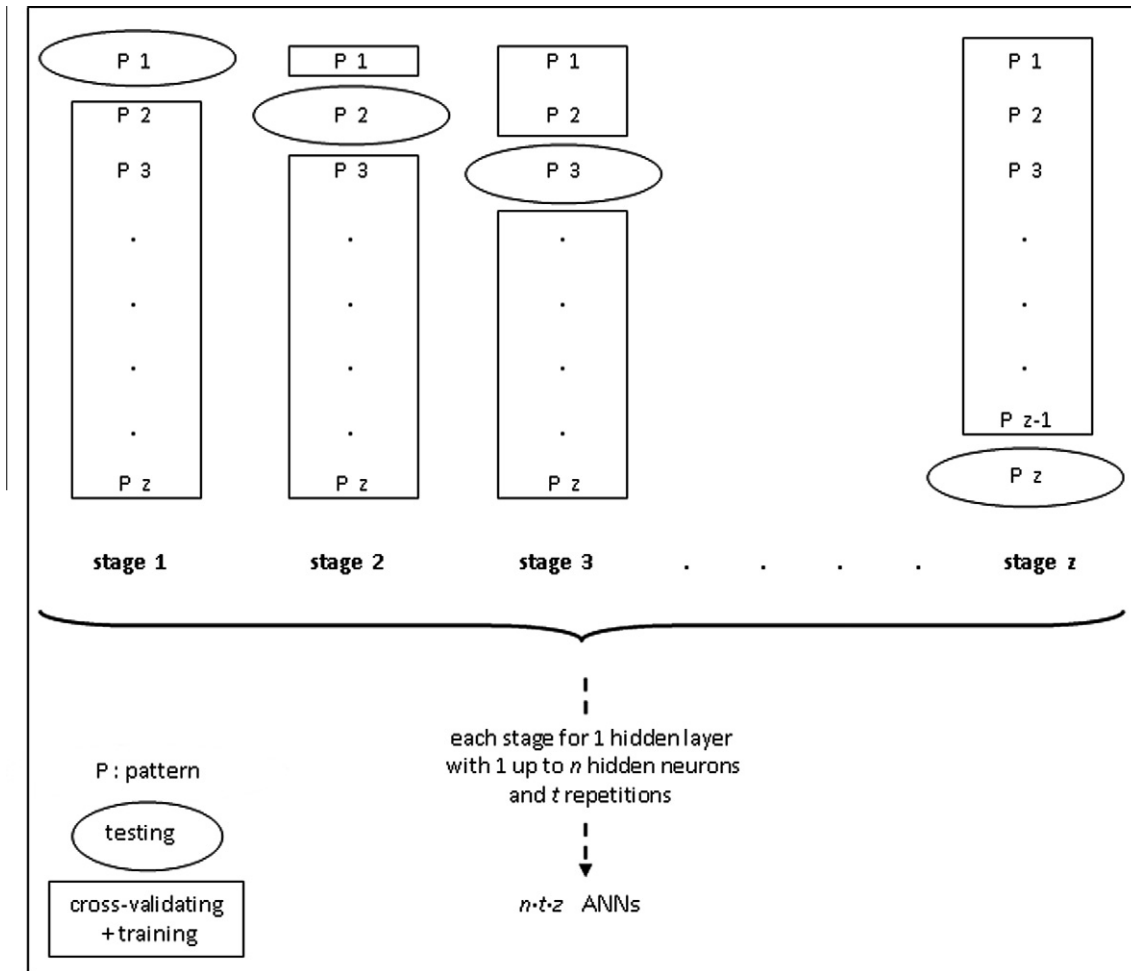


Fig. 3. Scheme of 'leave one out' procedure.

patterns were used for the application of the training algorithm. The random assignment of the cross-validation data is represented in Fig. 4. In this case, 12.5% of the total was used for cross-validating. This size lies in the usual range of many ANN applications. Larger cross-validation data sets would involve an excessive drop in the number of training patterns, which might not be enough for proper application of the training algorithm.

All source data were scaled, which is common practice in ANN modelling, preventing and overcoming the problems associated with extreme values and markedly different orders of magnitude between variables. The assignment of the intervals  $[-1, 1]$  and  $[0, 1]$  for tansig and logsig activation functions, respectively, might limit the extrapolation ability of the ANN, because they involve a saturation of the neuron output range. With these intervals, the neural network cannot produce output values beyond the maximum considered in the data set. Therefore, two different intervals were assigned for each activation function, namely  $[-0.9, 0.9]$  and  $[-0.8, 0.8]$  for the tansig function as well as  $[0.05, 0.95]$  and  $[0.1, 0.9]$  for the logsig function. Outputs were returned to their original values after the simulation. Accordingly,

$$x_s = \frac{(U_x - u_x) \cdot x + (M_x \cdot u_x - m_x \cdot U_x)}{M_x - m_x} \quad (4)$$

where  $x_s$  is the scaled variable;  $x$  is the original variable;  $M_x$  represents the maximum value of the original sample;  $m_x$  represents the minimum value of the original sample;  $U_x$  is the maximum value

assigned in the scaled sample;  $u_x$  is the minimum value assigned in the scaled sample.

For each data set configuration of the LOO stage, different architectures were trained and tested. Each time, the adopted procedure allowed for the selection of the optimum architecture from a set that considered up to  $s$  hidden layers with 1 to  $n$  neurons each, where the different hidden layers always presented the same number of neurons. Further, each architecture was trained  $t$  times to overcome the influence of the initial random assignment of the weights and biases, ensuring the reliability of the estimates. Only one hidden layer was considered, due to the high number of study cases derived from the LOO procedures. This is common practice in many ANN applications. The maximum number of neurons per layer and the number of repetitions were fixed at 20 each. As there is no clear and definitive methodology to deal with the optimum architecture selection in the ANN community, this is usually based on a trial and error process. Within each architecture, the program selects the repetition that provides the best performance of the cross-validation set. Afterwards, the best ANN repetition of each architecture is tested. Finally, the optimum architecture is selected from the analysis of the errors obtained in the 3 data sets, seeking a configuration with high generalizability.

#### 2.4. Model evaluation

Different error parameters were calculated to assess the performance accuracy of the proposed methods. Statistical analysis

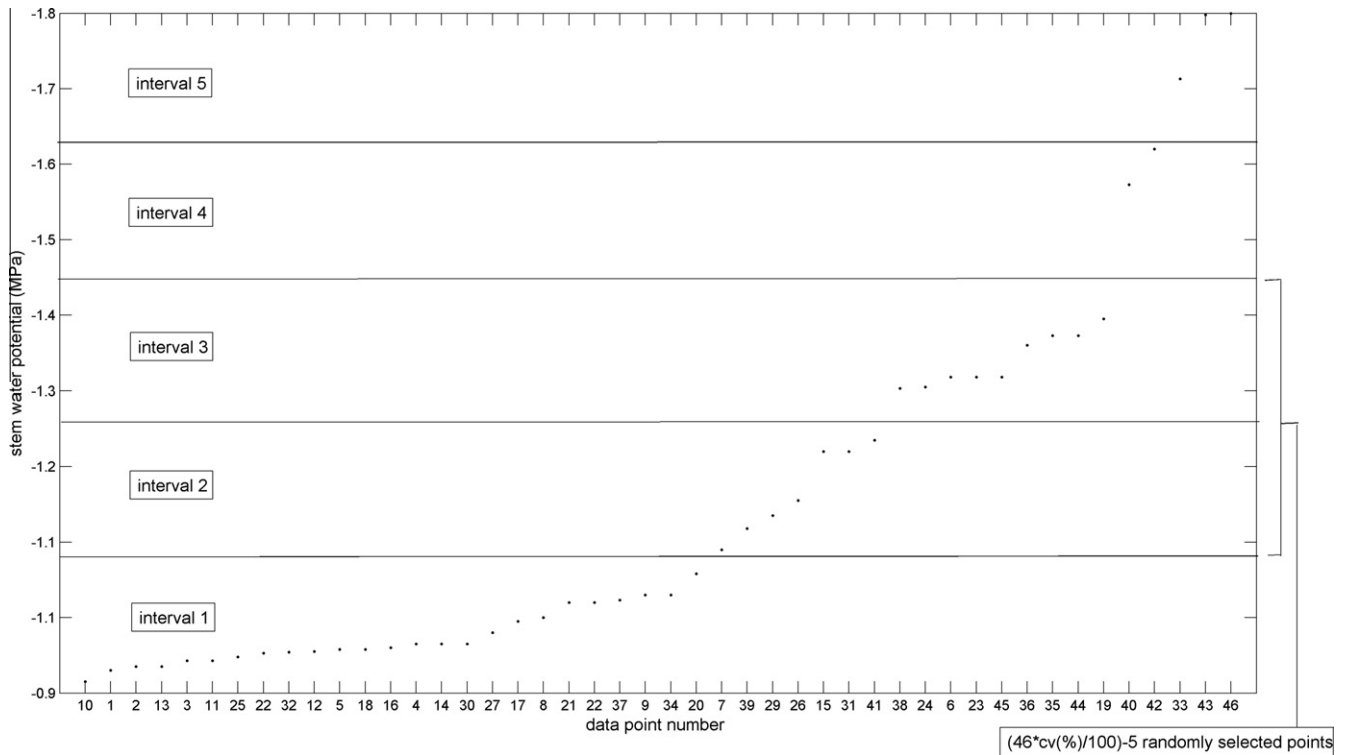


Fig. 4. Implementation of the cross-validation data set.

results are shown in terms of the mean squared error (MSE), to penalize large errors more than small ones, the mean absolute error (MAE), the relative root mean squared error RMSE, to introduce a relative measure of the error, and the determination coefficient ( $r^2$ ), obtained according to equations 5, 6, 7 and 8, respectively,  $x_i$  being the actual value of  $\psi_{st}$  and  $\hat{x}_i$  the estimation, while  $\sigma_{x_i}$  and  $\sigma_{\hat{x}_i}$  are the standard deviations of the observed and predicted  $\psi_{st}$  values, respectively. The units of MSE and MAE are  $\text{MPa}^2$  and  $\text{MPa}$ , respectively.

$$\text{MSE} = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (5)$$

$$\text{RMSE} = \frac{1}{x} \cdot \left( \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \hat{x}_i)^2 \right)^{0.5} \quad (6)$$

$$\text{MAE} = \frac{1}{n} \cdot \sum_{i=1}^n |x_i - \hat{x}_i| \quad (7)$$

$$r^2 = \left( \frac{n \sum x_i \hat{x}_i - (\sum x_i)(\sum \hat{x}_i)}{\sqrt{n(\sum x_i^2) - (\sum x_i)^2} \sqrt{n(\sum \hat{x}_i^2) - (\sum \hat{x}_i)^2}} \right)^2 \quad (8)$$

The MSE was also used as the performance function for the application of the training algorithm. This function was chosen because it is a non-negative, differentiable function, because of its statistical properties, and because it is better understood than other measures.

### 3. Results and discussion

The characteristics of the first 4 principal components are shown in Table 5.  $R_x^2$  is the proportion of the data variance explained by a given component. The threshold  $Q^2$  value to consider

Table 5  
PCA performance indicators.

PC	$R_x^2$	Cumulative $R_x^2$	$\lambda$	$Q^2$	$Q_{\text{limit}}^2$
1	0.458	0.458	4.58	0.337	0.111
2	0.215	0.673	2.15	0.213	0.12
3	0.107	0.780	1.07	-0.232	0.131
4	0.0889	0.869	0.889	-0.271	0.145

Eigenvalue ( $\lambda$ ), goodness-of-fit ( $R_x^2$ ), cumulative values, goodness-of-fit by cross-validation ( $Q^2$ ) and threshold value.

the component as relevant is also shown. In order to check if the results of Table 5 are conditioned by influential observations, different score plots were inspected, but no outliers appear with anomalous high scores. Although the first 4 components explain 86.9% of the total data variance, only the first two principal components (PC1 and PC2) satisfy the cross-validation criterion because their  $Q^2$  values are higher than the threshold considered by the software. PC1 is the linear combination of the original variables and explains the highest amount of the total data variability. In this case, PC1 explains 45.8% of the total data variance, while PC2 explains 21.5%. Another criterion often applied in PCA is to focus attention on those PCs with an eigenvalue higher than 1. Based on this criterion, PC3 might be regarded as a relevant component too. However, PC3 does not satisfy the cross-validation criterion, and therefore it is more convenient to focus only on PC1 and PC2.

The projections of individuals over PC1 and PC2 are usually called  $t[1]$  and  $t[2]$  scores, respectively. Fig. 5 shows the projections over the directions determined by PC1 and PC2. This figure shows that one observation is out of Hotelling's  $T^2$  ellipse considering a significance level of 0.05. This result indicates that this point presented a slightly different performance from the rest. Taking into account that the multivariate model that best describes the data variability is based on PC1 and PC2, it is of interest to check if any point might be regarded as an outlier according to this model. For this purpose, the distances of observations to the model with

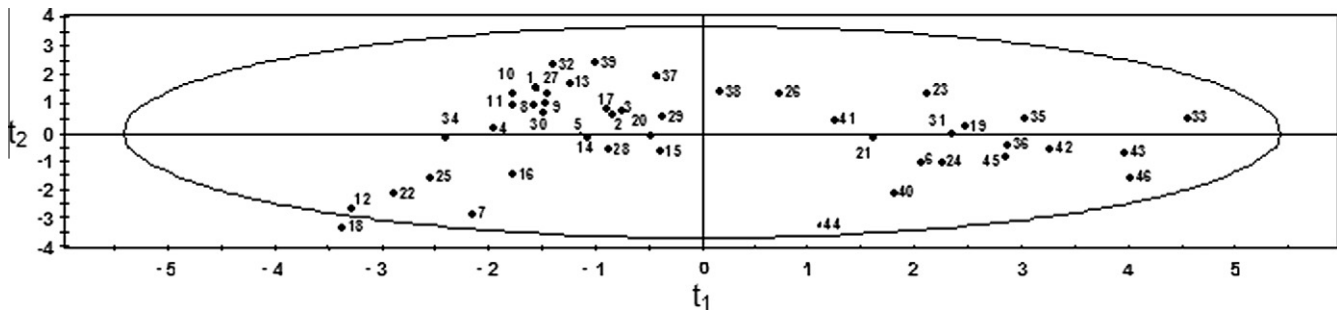


Fig. 5. Score plot of PC1 - PC2 ( $t_2$  vs.  $t_1$ ) with Hotelling's  $T^2$  ellipse at  $\alpha = 0.05$ .

two principal components were obtained, as shown in Fig. 6. Considering a significance level  $\alpha = 0.05$ , there is a pattern (28) with a distance slightly greater than the critical threshold calculated by the software. At  $\alpha = 0.01$ , this point still appears as a slight outlier. Nevertheless, this pattern was not excluded from the data set, because this abnormal performance is only very slight (i.e. a value that results in a frequency  $\leq 1\%$ ).

In order to evaluate possible relationships among variables, Fig. 7 shows the loading plot of the first two components. Loadings are the weights or contributions of variables in the formation of a given component. As observed,  $h_2$ ,  $h_3$  and  $h_4$  appear very close to each other and could therefore be correlated. On the other hand,  $h_1$  lies clearly apart from the former ones considering PC1. A possible reason for this might be a higher variation in the  $h_1$  values because the moisture sensor is very close to the surface, where soil water content changes are more marked and fluctuating.

Understanding the correlation structure of the input variables is important for correct interpretation of the MLR results. Table 6 presents the correlation coefficients (above the main diagonal) and  $p$ -values (below the diagonal) within the considered variables. These results confirm that the correlation within  $h_2$ ,  $h_3$  and  $h_4$  is high and statistically significant, with correlation values ranging from 0.891

( $h_2-h_4$ ) to 0.983 ( $h_3-h_4$ ), always with  $p$ -values  $< 0.0001$ .  $h_1$  and  $h_2$  are also correlated, but considerably less than the former ones ( $r = 0.677$ ,  $p < 0.0001$ ). This confirms the conclusions drawn from the loading plot. Attending to  $\psi_{st}$ , the most important correlations were detected between  $\psi_{st} - T$  ( $r = -0.461$ ,  $p = 0.001$ ), and between  $\psi_{st} - h_i$  (for  $h_1$ ,  $r = 0.428$ ,  $p = 0.003$ , for  $h_2$ ,  $r = 0.811$ ,  $p < 0.0001$ , for  $h_3$ ,  $r = 0.879$ ,  $p < 0.0001$ , for  $h_4$ ,  $r = 0.860$ ,  $p < 0.0001$ ), where  $h_3$  appears as the most correlated one. Finally, additional correlations were detected between temperature-relative humidity ( $r = -0.598$ ,  $p < 0.0001$ ), temperature-solar radiation ( $r = 0.756$ ,  $p < 0.0001$ ) and relative humidity-solar radiation ( $r = -0.805$ ,  $p < 0.0001$ ), which makes sense according to their physical meaning.

Next, step-wise MLR (forward and backward) was applied to determine if  $\psi_{st}$  is correlated with the aforementioned input variables. This analysis revealed that the optimum predictive equation of  $\psi_{st}$  considered temperature, relative humidity, solar radiation and soil moisture  $h_3$  as input variables. In parallel, three additional equations (Table 7) were obtained replacing, respectively,  $h_3$  by another depth  $h_i$  of the moisture sensors in the optimum input combination to find whether the differences in their performance accuracies are important or not. All regression coefficients of the best predictive equations are statistically significant, except  $p^d$  in

497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518



Fig. 6. Distance of observations to the model built with the first two principal components.

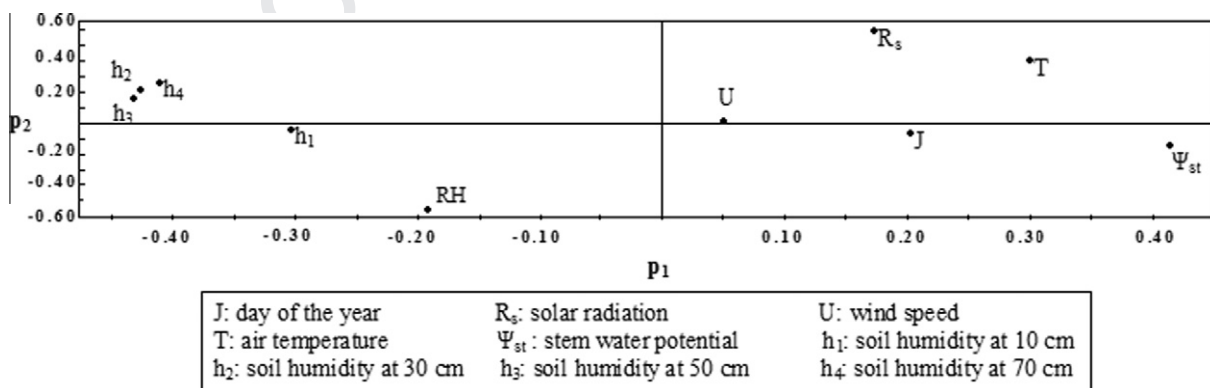


Fig. 7. Loading plot of PC1 and PC2.

**Table 6**

Correlation matrix of the input variables considered. Correlation coefficients are indicated above the main diagonal, and *p*-values below the diagonal.

	<i>T</i>	<i>R<sub>s</sub></i>	<i>U</i>	RH	<i>h</i> <sub>1</sub>	<i>h</i> <sub>2</sub>	<i>h</i> <sub>3</sub>	<i>h</i> <sub>4</sub>	<i>J</i>	$\psi_{st}$
<i>T</i>		0.756	-0.078	-0.598	-0.327	-0.400	-0.424	-0.334	0.230	-0.461
<i>R<sub>s</sub></i>	<0.0001		0.043	-0.805	-0.219	-0.153	-0.124	-0.021	-0.005	-0.102
<i>U</i>	0.607	0.775		-0.158	-0.019	-0.093	-0.054	0.060	-0.020	-0.169
RH	<0.0001	<0.0001	0.294		0.316	0.174	0.122	0.054	-0.127	0.225
<i>h</i> <sub>1</sub>	0.27	0.144	0.898	0.032		0.677	0.434	0.426	-0.319	0.428
<i>h</i> <sub>2</sub>	0.006	0.310	0.539	0.248	<0.0001		0.912	0.891	-0.316	0.811
<i>h</i> <sub>3</sub>	0.003	0.413	0.719	0.420	0.003	<0.0001		0.983	-0.299	0.879
<i>h</i> <sub>4</sub>	0.023	0.889	0.690	0.720	0.003	<0.0001	<0.0001		-0.328	0.860
<i>J</i>	0.123	0.975	0.897	0.400	0.031	0.032	0.044	0.026		-0.328
$\psi_{st}$	0.001	0.498	0.263	0.133	0.003	<0.0001	<0.0001	<0.0001	0.026	

Variables: stem water potential ( $\psi_{st}$ ), temperature (*T*), solar radiation (*R<sub>s</sub>*), relative humidity (RH), wind speed (*U*), soil moisture at 10 cm (*h*<sub>1</sub>), soil moisture at 30 cm (*h*<sub>2</sub>), soil moisture at 50 cm (*h*<sub>3</sub>), soil moisture at 70 cm (*h*<sub>4</sub>).

**Table 7**

Predictive equations of stem water potential according to different soil moisture depths and associated statistical parameters.

<i>h<sub>i</sub></i> Considered	Predictive equation <sup>a</sup>								
-	$\psi_{st} = -0.97459 - 0.00059T^2 + 0.00048R_s - 0.06119U$								
<i>h</i> <sub>1</sub>	$\psi_{st} = -1.28074 - 0.02373T + 0.00051R_s + 0.00047h_1^2$								
<i>h</i> <sub>2</sub>	$\psi_{st} = -2.67028 - 0.00029T^2 + 0.00430RH + 0.00051R_s + 0.03928h_2$								
<i>h</i> <sub>3</sub>	$\psi_{st} = -2.39467 - 0.00021T^2 + 0.00493RH + 0.00045R_s + 0.03170h_3$								
<i>h</i> <sub>4</sub>	$\psi_{st} = -2.09284 - 0.00024T^2 + 0.00454RH + 0.00039R_s + 0.03920h_4$								
	<i>p</i> <sup>b</sup>	<i>p</i> <sup>c</sup>	<i>p</i> <sup>d</sup>	<i>p</i> <sup>e</sup>	<i>r</i> <sup>2</sup>	MSE (MPa <sup>2</sup> )	MAE (MPa)	RMSE	
-	<0.0001	0.003	0.036	-	0.412	0.03493	0.1446	0.162	
<i>h</i> <sub>1</sub>	<0.0001	0.003	0.016	-	0.436	0.03343	0.1425	0.159	
<i>h</i> <sub>2</sub>	0.001	0.018	0.001	<0.0001	0.766	0.01386	0.0910	0.102	
<i>h</i> <sub>3</sub>	0.003	0.001	<0.0001	<0.0001	0.852	0.00880	0.0681	0.081	
<i>h</i> <sub>4</sub>	0.001	0.004	0.002	<0.0001	0.832	0.00996	0.0767	0.087	

<sup>a</sup> Variables: stem water potential,  $\psi_{st}$  (MPa), temperature, *T* (°C), solar radiation, *R<sub>s</sub>* (W/m<sup>2</sup>), relative humidity, RH (%), wind speed, *U* (m/s), soil moisture at 10 cm, *h*<sub>1</sub> (mm), soil moisture at 30 cm, *h*<sub>2</sub> (mm), soil moisture at 50 cm, *h*<sub>3</sub> (mm), soil moisture at 70 cm, *h*<sub>4</sub> (mm).

<sup>b</sup> *p*-Value associated with the first variable appearing in the equation.

<sup>c</sup> *p*-Value associated with the second variable in the equation.

<sup>d</sup> *p*-Value associated with the third variable.

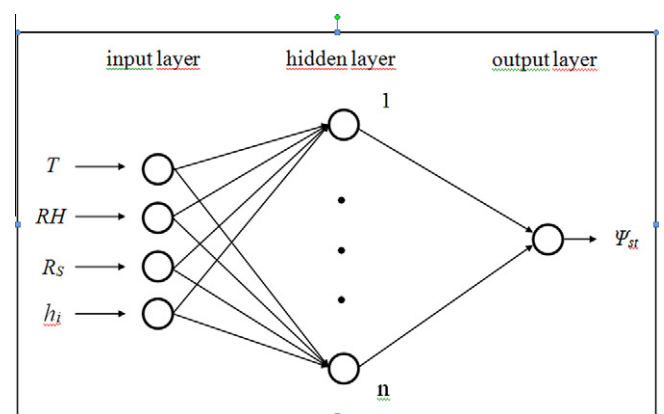
<sup>e</sup> *p*-Value associated with the fourth variable.

*h*<sub>1</sub> (by 0.016) and *p*<sup>c</sup> in *h*<sub>2</sub> (by 0.018). However, these variables were included because they are not significant by a very small amount and because these two models are only presented to carry out a comparison with the optimum ones, those relying on *h*<sub>3</sub> and *h*<sub>4</sub>. This table also presents the statistical indicators corresponding to the associated estimations. For each equation it was checked that residuals followed an approximately normal distribution and no outliers were detected, in agreement with the PCA results. The models which consider *h*<sub>3</sub> and *h*<sub>4</sub> present similar performance accuracy (0.852 vs 0.832 in the *r*<sup>2</sup> and 0.081 vs 0.087 in the RMSE, respectively). The consideration of *h*<sub>2</sub> instead of the former ones involves a decrease in accuracy of 0.07–0.09 in the *r*<sup>2</sup>, 0.00390–0.00506 in the MSE, 0.0143–0.0229 in the MAE and 0.015–0.021 in the RMSE in comparison to *h*<sub>4</sub> and *h*<sub>3</sub>. The performance of the model that considers *h*<sub>1</sub> is unacceptable, with an *r*<sup>2</sup> of 0.436 and an RMSE of 0.159. This is mainly due to a low *h*<sub>1</sub> –  $\psi_{st}$  correlation. As stated before, *h*<sub>1</sub> values might present higher variability than the records of the other sensors, due to its proximity to the soil surface, which could explain this clearly lower correlation. Accordingly, the consideration of *h*<sub>3</sub> or *h*<sub>4</sub> might be preferable to *h*<sub>1</sub> and *h*<sub>2</sub> as more reliable indicators of the real soil water content available for the tree. Although the *h*<sub>3</sub> –  $\psi_{st}$  correlation is slightly higher than between *h*<sub>4</sub> –  $\psi_{st}$ , *h*<sub>4</sub> might be as valid as *h*<sub>3</sub> for developing the MLR model, because both coefficients are very similar. Finally, a model relying exclusively on all meteorological parameters was also obtained. The regression coefficient of the wind speed is not statistically significant (by 0.026). The coefficient of the relative humidity was not considered, because it is not significant by a very large amount. Nevertheless, this equation was included to make

evident the performance improvement achieved when considering *h*<sub>3</sub> or *h*<sub>4</sub> (or even *h*<sub>2</sub>), i.e., information about the soil moisture at root level.

Finally, seeking more meaningful and accurate MLR models, five new input variables were studied, namely vapour pressure deficit (VPD), which is a function of RH and *T*, the accumulated moisture values *h*<sub>2</sub>–*h*<sub>3</sub> and *h*<sub>2</sub>–*h*<sub>3</sub>–*h*<sub>4</sub>, and the mean moisture values of *h*<sub>2</sub>–*h*<sub>3</sub> and *h*<sub>2</sub>–*h*<sub>3</sub>–*h*<sub>4</sub>. Nevertheless, none of these models improved the accuracy of the optimum input combination presented in Table 7.

According to these results, only the best two combinations of inputs, namely *T*, RH, *R<sub>s</sub>* and *h*<sub>3</sub>, and *T*, RH, *R<sub>s</sub>* and *h*<sub>4</sub>, were consid-



**Fig. 8.** Scheme of ANN configurations tested (where *h<sub>i</sub>* can be *h*<sub>3</sub> or *h*<sub>4</sub>).

ered to feed the neural network models (Fig. 8). Another alternative might have been to feed the ANNs with the scores of the significant principal components. Nevertheless, this study aims to find the most significant input variables to estimate  $\psi_{st}$  and to compare the corresponding MLR and ANN models. These models might be applied by future users more easily. The error parameters corresponding to the four cases defined in the methods section (2 activation functions and 2 scaling intervals) are summarised in Table 8. Within configurations considering  $h_3$ , the logsig activation function enables a better performance ( $r^2$  between 0.902 and 0.926) than the tansig function ( $r^2$  around 0.883). Moreover, data scaling in the interval [0.05,0.95] seems preferable to the interval [0.1, 0.9], although accuracy differences are more slight than before (only 0.024 in the  $r^2$ , 0.00100 in the MSE, 0.0032 in the MAE and 0.005 in the RMSE). On the other hand, the performance differences within the configurations considering  $h_4$  are lower. Logsig and tansig functions provide similar performances, ranging between 0.871 and 0.890 of  $r^2$  in the former, and between 0.883 and 0.897 of  $r^2$  in the latter. The same conclusions can be drawn on the basis of the error parameters.

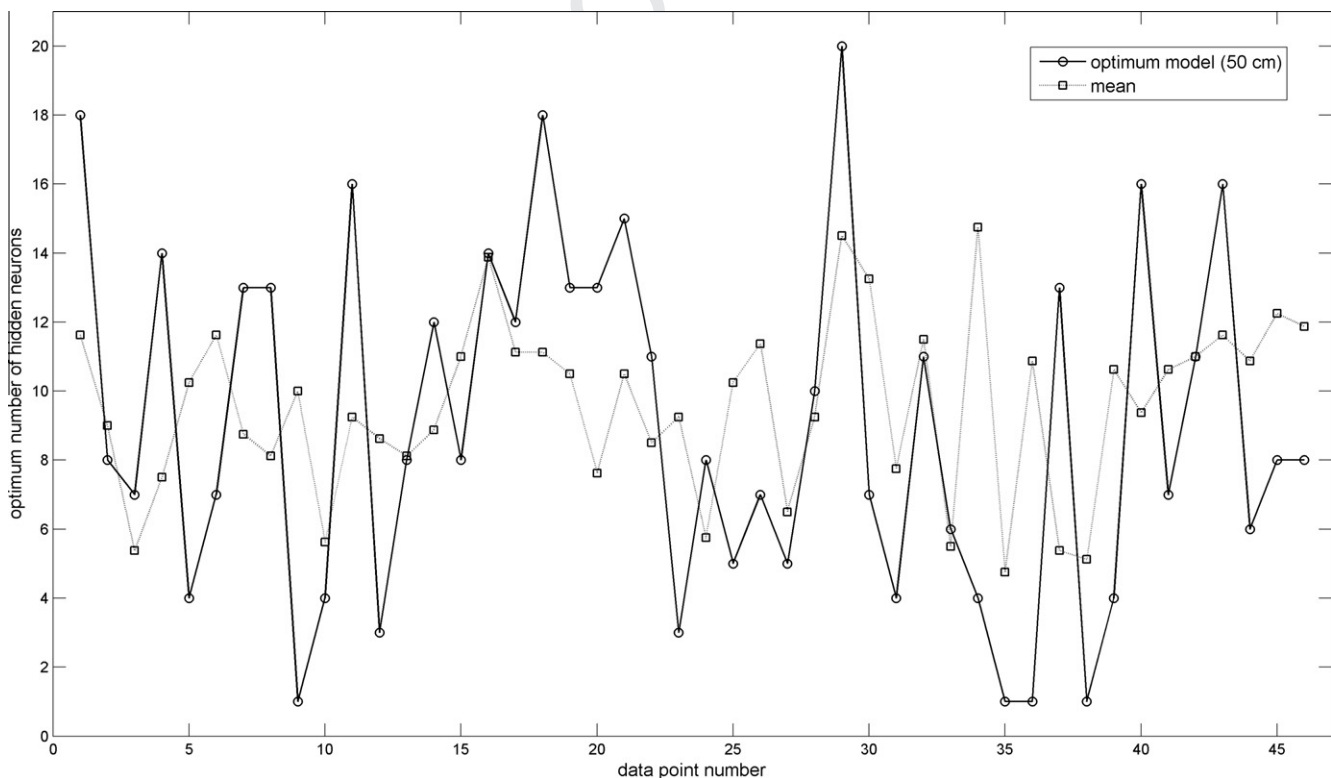
The optimum  $h_4$  configuration corresponds to the logsig transfer function and the [0.1, 0.9] scaling interval, while the optimum

$h_3$  configuration corresponds to the logsig function and the [0.05, 0.95] interval. Comparing these configurations, the ANN fed with the  $h_3$  input provides greater accuracy than the ANN fed with the  $h_4$  input (with improvements of 0.029 in the  $r^2$ , 0.00122 in the MSE, 0.0133 in the MAE and 0.007 in the RMSE), in agreement with the regression results. Finally, comparing regression models with their corresponding neural network optimum configurations, ANN estimations present considerably better indicators than regression estimations. Within models fed with  $h_3$ , ANNs involve an improvement of 0.074 in the  $r^2$ , 0.00377 in the MSE, 0.0178 in the MAE and 0.019 in the RMSE. Within models fed with  $h_4$ , ANNs involve an improvement of 0.065 in the  $r^2$ , 0.00371 in the MSE, 0.0131 in the MAE and 0.018 in the RMSE. This is due to the higher mapping ability of the neural networks.

The optimum number of hidden neurons per stage of the LOO process (per test data point) is represented in Fig. 9. Here, two cases are considered: the optimum model (ANN with logsig transfer functions in the hidden neurons, scaled in the interval [0.05, 0.95] and fed with  $h_3$  data) and the average results of the 8 cases considered in Table 8. The optimum number of hidden neurons presents a fluctuating trend in the former case, in general with a rather high (more than 13) or low (less than 8) number of neu-

**Table 8**  
Statistical parameters of ANN models performed.

Depth considered	Activation function	Range scaling	$r^2$	MSE (MPa <sup>2</sup> )	MAE (MPa)	RMSE
$h_3$	Tansig	[-0.9, 0.9]	0.883	0.00709	0.0605	0.073
		[-0.8, 0.8]	0.884	0.00700	0.0616	0.072
	Logsig	[0.05, 0.95]	0.926	0.00503	0.0503	0.062
		[0.1, 0.9]	0.902	0.00603	0.0535	0.067
$h_4$	Tansig	[-0.9, 0.9]	0.890	0.00664	0.0620	0.071
		[-0.8, 0.8]	0.871	0.00766	0.0717	0.076
	Logsig	[0.05, 0.95]	0.883	0.00711	0.0648	0.073
		[0.1, 0.9]	0.897	0.00625	0.0636	0.069



**Fig. 9.** Comparison between selected architectures of the optimum model ( $h_3$ ) and mean number of hidden neurons in the 8 ANN configurations.

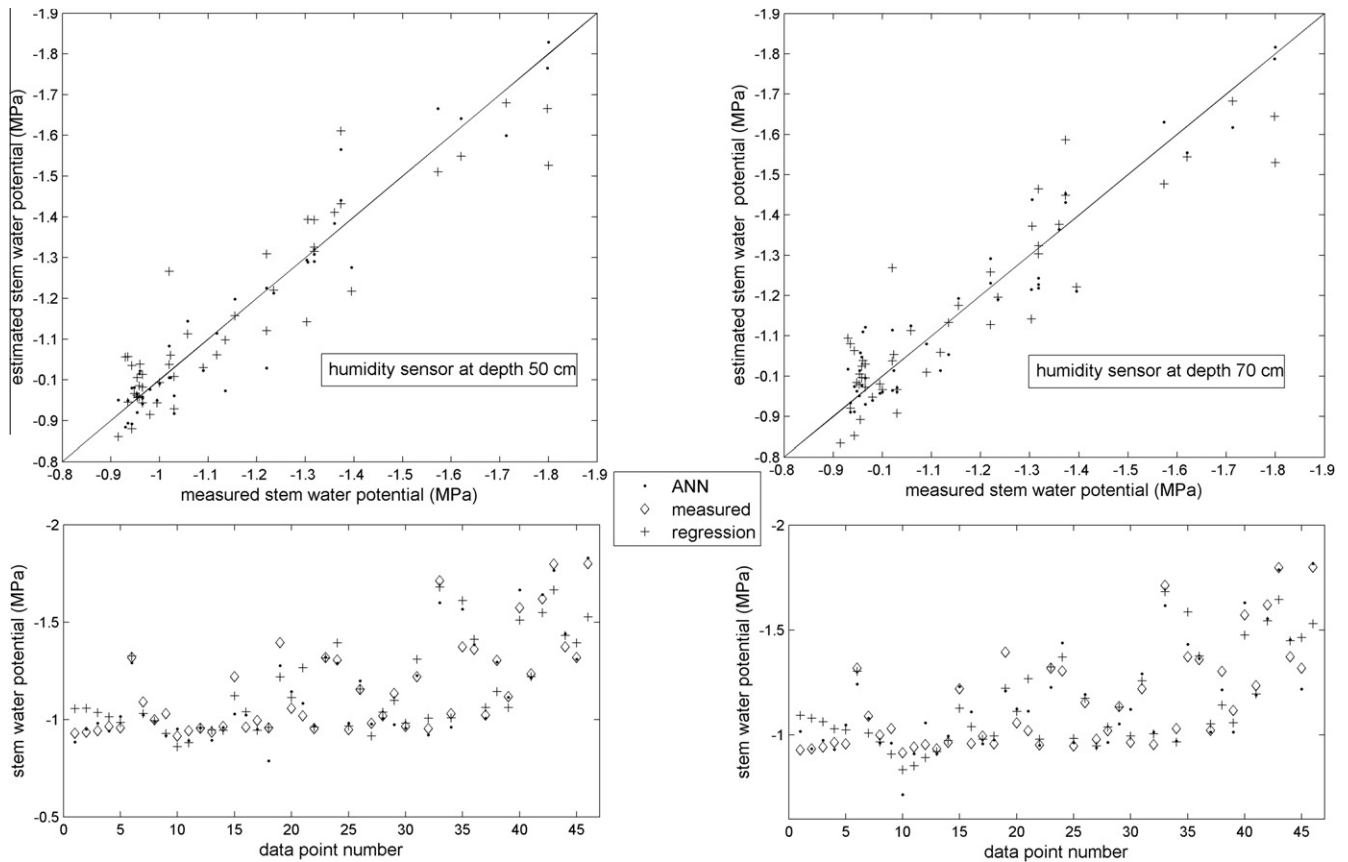


Fig. 10. Comparison of regression and ANN models' performance.

rons. By contrast, the number of optimum hidden neurons ranges in general between 7 and 12 in the mean case. Accordingly, a maximum of around 12 neurons might be enough to reach a suitable input-output mapping.

Finally, Fig. 10 shows a comparison between ANN and regression estimations using the models fed with  $h_3$  and  $h_4$  inputs, respectively. The two figures on the left side correspond to the models fed with  $h_3$  (sensor at 50 cm depth) and the other two figures on the right side correspond to the models fed with  $h_4$  (sensor at 70 cm depth). Attending to the figures on the left side, the ANN estimations present lower dispersion than regression estimations. Further, regression and ANN estimations seem to present separately a similar accuracy within all  $\psi_{st}$  ranges. Finally, estimations do not show a clear trend in terms of over-/underestimation. The figures on the lower side present the same estimations throughout the year. Nevertheless, it is difficult to find a relationship between the accuracy of the models and the day of the year or  $\psi_{st}$  range. Attending to the figures on the right side, the estimations obtained using the models fed with  $h_4$  inputs present the same trends as the model relying on  $h_3$ .

A robust methodology for dealing with  $\psi_{st}$  estimation was presented, particularly adapted to deal more effectively with limited data sets. The present experimental data set allows no generalization of the proposed models for other conditions. It also does not allow conclusions to be drawn on the physical relationships between the studied variables. So, the study should be regarded as a first step approach presenting a promising modelling tool in this field. Further research should analyze these relationships considering a data set that allows for generalization of the physical relationships encountered and/or for estimation under other conditions. An accurate ANN-based model, combined with a system providing remote downloaded FDR measurements could be used

to obtain  $\psi_{st}$  values more directly without the need for time- and labour-consuming measurements, allowing automation of (estimated)  $\psi_{st}$  data collection. Hence, generally, when properly trained with a representative dataset, the ANN-based approach might be used to provide  $\psi_{st}$  values from automated FDR and meteorological measurements as part of a decision-support aid to achieve more efficient irrigation scheduling.

#### 4. Conclusions

An ANN-based approach was applied to estimate stem water potential from soil moisture data at different depths and standard meteorological variables. This approach was adapted to deal effectively with a limited data set. Prior to ANN application, PCA and MLR were used to study the relationships between observations and variables, as well as to select the inputs to feed the models. Two principal components account for the systematic data variation. The optimum predictive equation of  $\psi_{st}$  considered temperature, relative humidity, solar radiation and soil moisture at 50 cm as input variables. The performance of the ANN models confirmed the conclusions drawn from the regression analysis. Compared with their corresponding regression models, ANNs presented considerably higher performance accuracy, due to a higher input-output mapping ability. The accurate mapping between the considered input and target variables, combined with an accurate soil moisture monitoring setup can be used as a decision-support aid to achieve more efficient irrigation scheduling.

#### Acknowledgments

The authors are grateful to TECVASA, which obtained a subsidy from the Conselleria de Agricultura, Pesca y Alimentación de la

663 Generalitat Valenciana (DOCV 5493, 19 April 2007, [no. exp.:](#)  
664 2007TAHAVAL00018), and to the Valencian Institute for Agricul-  
665 tural Research (IVIA) for providing the meteorological data for this  
666 study.

667 **References**

668 Acevedo-Opazo, C., Tisseyre, B., Taylor, J.A., Ojeda, H., Guillaume, S., 2010. A model  
669 for the spatial prediction of water status in vines (*Vitis vinifera* L.) using high  
670 resolution ancillary information. *Precision Agriculture* 11, 358–378.  
671 Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration –  
672 Guidelines for computing crop water requirements. FAO Irrigation and Drainage  
673 Paper No. 56. FAO, Rome, Italy, p. 465.  
674 Améglio, T., Archer, P., Cohen, M., Valancogne, C., Dudet, F.A., Daya, S., Cruziat, P.,  
675 1999. Significance and limits in the use of predawn leaf water potential for tree  
676 irrigation. *Plant and Soil* 207, 155–167.  
677 Campbell, G., 1974. A simple method for determining unsaturated conductivity  
678 from moisture retention data. *Soil Sci.* 117, 311–314.  
679 Campbell, G.S., Campbell, M.D., 1982. Irrigation scheduling using soil moisture  
680 measurements: theory and practice. In: Hillel, D.J. (Ed.), *Advances in Irrigation*,  
681 vol. 1. Academic Press, New York, pp. 25–42.  
682 Capraro, F., Patino, D., Tosetti, S., Schugurensky, C., 2008. Neural network-based  
683 irrigation control for precision agriculture. ICNSC 2008. In: IEEE International  
684 Conference on Networking, Sensing and Control, pp. 357–362.  
685 Castel, J.R., 2001. Consumo de agua por plantaciones de cítricos en Valencia.  
686 *Fruticultura Profesional* 1, 27–32 (in Spanish).  
687 Choné, X., Van Leeuwen, C., Dubourdieu, D., Gaudillere, J.P., 2001. Stem water  
688 potential is a sensitive indicator of grapevine water status. *Annals of Botany*–  
689 London 87, 477–483.  
690 Dane, J.H., Topp, G.C. (Eds.) 2002. *Methods of Soil Analysis. Part 4: Physical*  
691 *Methods.* Soil Science Society of America Book Series, vol. 5. Madison, WI, p.  
692 1692.  
693 Diana, G., Tommasi, C., 2002. Cross-validation methods in principal components: a  
694 comparison. *Statistical Methods and Applications* 11 (1), 71–82.  
695 Dzikiti, S., Verreyne, J.S., Stuckens, J., Strever, A., Verstraeten, W.W., Swennen, R.,  
696 Coppin, P., 2010. Determining the water status of Satsuma mandarin trees  
697 [*Citrus Unshiu Marcovitch*] using spectral indices and by combining  
698 hyperspectral and physiological data. *Agricultural and Forest Meteorology*  
699 150, 369–379.  
700 Elfving, D.C., Kaufmann, M.R., Hall, A.E., 1972. Interpreting leaf water potential  
701 measurements with a model of soil–plant–atmosphere continuum. *Physiologia*  
702 *Plantarum* 27, 161–168.  
703 Eriksson, L., Johansson, E., Kettaneh-Wold, N., Wold, S., 1999. Introduction to Multi-  
704 and Megavariate Data Analysis Using Projection Methods (PCA & PLS). *Umetrics*  
705 AB, Umea, Sweden.  
706 FAO-UNESCO, 1988. *Soil Map of the World. Revised legend.* 1:5.000.000. World Soil  
707 Resources Report 60. FAO, Roma.  
708 Fares, A., Polyakov, V., 2006. Advances in crop management using capacitive  
709 sensors. *Advances in Agronomy* 90, 43–77.  
710 Garnier, E., Berger, A., 1985. Testing water potential in peach trees as an  
711 indicator of water stress. *The Journal of Horticultural Science &*  
712 *Biotechnology* 60, 47–56.  
713 Gasque, M., Granero, B., Turégano, J.V., González-Altozano, P., 2010. Regulated  
714 deficit irrigation effects on yield, fruit quality and vegetative growth of  
715 ‘Navelina’ citrus trees. *Spanish Journal of Agricultural Research* 8 (S2), S40–S51.  
716 Goldhamer, D.A., Fereres, E., 2001. Irrigation scheduling protocols using  
717 continuously recorded trunk diameter measurements. *Irrig. Sci.* 20, 115–125.  
718 González-Altozano, P., Castel, J.R., 1999. Regulated Deficit Irrigation in ‘Clementina  
719 de Nules’ citrus trees. I. Yield and fruit quality effects. *The Journal of*  
720 *Horticultural Science and Biotechnology* 74 (6), 706–713.  
721 González-Altozano, P., Castel, J.R., 2000. Regulated deficit irrigation in ‘Clementina  
722 de Nules’ citrus trees. II. Vegetative effects. *The Journal of Horticultural Science*  
723 *and Biotechnology* 75 (4), 388–392.  
724 Hagan, M.T., Menhaj, M.B., 1994. Training feedforward Networks with the  
725 Marquardt Algorithm. *IEEE Trans Neural Networks* 5 (6), 989–993.

Hagan, M.T., Delmuth, H., Beale, M., 1996. *Neural Network Design.* PWS Publishing  
Company, Boston, MA.  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000