Computers and Electronics in Agriculture 142 (2017) 181–189

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Classification models for automatic identification of daily states from leaf turgor related measurements in olive

Rafael Dreux Miranda Fernandes [a,*], Maria Victoria Cuevas [a], Virginia Hernandez-Santana [a], Celia M. Rodriguez-Dominguez [b], Carmen M. Padilla-Díaz [a], José Enrique Fernández [a]

[a] Irrigation and Crop Ecophysiology Group, Instituto de Recursos Naturales y Agrobiología (IRNAS, CSIC), Avenida Reina Mercedes, 10, 41012- Seville, Spain
[b] School of Biological Sciences, University of Tasmania, Private Bag 55, Hobart, Tasmania 7001, Australia

ARTICLE INFO

Article history:
Received 22 June 2017
Received in revised form 1 September 2017
Accepted 5 September 2017
Available online 9 September 2017

Keywords:
Random forest
Leaf patch clamp pressure probe
Supervised classification
Inflection points
Leave-p-out cross validation

ABSTRACT

The leaf patch clamp pressure (LPCP) probe is being used to remotely assess leaf turgor pressure. Recently, different shapes of the LPCP daily curves have been suggested as potential water stress indicators for irrigation scheduling. These curves shapes, called states, have been studied and related to different water stress levels for olives. To our knowledge, the only way to differentiate these curves shapes or states is through the visual observation of the dynamics of the LPCP records during the day, which is highly time-consuming and reduces its potential to automatically schedule irrigation. The aims of this study were: (i) to obtain a random forest model to automatically identify the states from daily LPCP curves recorded in olive trees, by using visually identified states to train the model; (ii) to improve the identification of state II through a second random forest model, relating this state to the midday stem water potential, and; (iii) to obtain a random forest model to identify the states based on ranges of stem water potential. We used LPCP daily curves collected in a commercial olive orchard from 2011 to 2015. The states were visually identified for the days on which concomitant measurements of stem water potential and leaf stomatal conductance were made. We had a data set of 307 LPCP daily curves, being 157 curves in state I, 78 in state II and 71 in state III. The two biggest inflection points of the LPCP curves were used to adjust the models through the use of the R package "randomForest", using the Leave-p-Out Cross-Validation method. With the first model, which was obtained from the whole dataset, its data regarding the inflection points and the visually identified states, we obtained an overall accuracy of 94.37%. With the second model, obtained with the use of the data regarding curves visually identified as state II only, the overall accuracy was of 88.64%. This model was adjusted to be used after the first model, to narrow the stem water potential range of state II curves. Finally, the third model was obtained using the whole dataset and the states established from ranges of stem water potential. This last model did not consider the visual identification, and yielded an overall accuracy of 88.08%. Our results facilitate the use of LPCP probes, since it allows for the automatic identification of the states related to leaf turgor pressure, a key information to schedule irrigation.

1. Introduction

Deficit irrigation (DI) is applied in fruit tree orchards to ensure production in areas with low irrigation water availability. A reliable monitoring of the tree water status has been suggested as an appropriate option to sustainably manage DI strategies. This is definitely necessary in regulated deficit irrigation (RDI), one of the most effective strategies for fruit orchards, in which severe water stress is avoided with enough irrigation during phases of the growing period sensitive to water stress (Fernández, 2014a), while water supplies are reduced on periods when the crop is less sensitive to drought.

A variety of methods based on plant measurements have been successfully used to monitor water status in olive trees (Cuevas et al., 2013; Fernández et al., 2013; Girón et al., 2015; Padilla-Díaz et al., 2016), apple and pear trees (Fernández et al., 2008), and vineyards (Rüger et al., 2011), among other woody crops. However, the use of these methods is still very limited in commercial orchards, mainly because they are expensive and both their installation and data interpretation require training (Fernández, 2014b). One of these methods relies on the use of the leaf patch clamp

* Corresponding author.
E-mail address: rafadreux@gmail.com (R.D.M. Fernandes).

http://dx.doi.org/10.1016/j.compag.2017.09.005
0168-1699/

pressure (LPCP) probe (Zimmermann et al., 2008). This is a user-friendly method suitable to monitor water stress and to schedule irrigation in commercial orchards (Fernández et al., 2011; Padilla-Díaz et al., 2016). The LPCP probe contains a miniaturized pressure sensor integrated into a small metal piece. After being clamped in a leaf, the LPCP probe measures the attenuated pressure (P_p) response of the leaf patch upon the application of a constant clamped pressure (P_{clamp}). Values of P_p are inversely proportional to the leaf turgor pressure, P_c (Ehrenberger et al., 2012). A commercial version of the probe, called ZIM probe (YARA ZIM Plant Technology, Hennigsdorf, Germany) uses a magnet, instead of a clamp, to apply P_{clamp} on the leaf.

From the shape of the daily P_p curves recorded with LPCP probes in olive trees under different water status, Fernández et al. (2011) and Ehrenberger et al. (2012) described three different states, related to the water stress level of the monitored trees: state I accounted for low water stress, state II for moderate water stress, and state III for severe water stress. They both reported that state I corresponded, for olive, to $\Psi_{stem} > -1.2$ MPa (being Ψ_{stem} the midday stem water potential), state II to $-1.2 > \Psi_{stem} > -1.7$ MPa and state III to $\Psi_{stem} < -1.7$ MPa. These Ψ_{stem} ranges for States I to III may be different in orchards with different cultivars and locations (Marino et al., 2016). Findings by Fernández et al. (2011) and Ehrenberger et al. (2012), and a recent work by Padilla-Díaz et al. (2016), confirm that the ranges reported above are valid for 'Arbequina' trees and for our orchard conditions.

Both Fernández et al. (2011) and Ehrenberger et al. (2012), and also Padilla-Díaz et al. (2016), analyzed the correlation between the state determined from a visual analysis of the daily P_p curves and the stem water potential (Ψ_{stem}) measured with a Scholander-type pressure chamber. Padilla-Díaz et al. (2016) used the biggest data set and found that 68.3% of the trees in state I showed $\Psi_{stem} > -1.2$ MPa and that 81.8% of those days in state III presented $\Psi_{stem} < -1.7$ MPa. However, only 32.8% of the trees in state II days showed values of ψ_{stem} between -1.2 and -1.7 MPa. Their results agree with findings already stated by Fernández et al. (2011), in the sense that that states I and III are easily defined from a visual analysis of the curves, and that both states are reasonably well correlated to Ψ_{stem}. The correlation of state II with Ψ_{stem}, however, is low, which could be partly explained by this state being more difficult to be visually identified by the user. A proper identification of state II, however, is important when using the LPCP probes for scheduling irrigation, as it represents a moderate stress level between state I and III. Thus, the appearance of state II may advice for modifying the irrigation dose or irrigation frequency (Fernández et al., 2011; Padilla-Díaz et al., 2016).

An option to improve the identification of states is to use classification methods usually applied for automatic sensed data. Among them, decision tree classification techniques have substantial advantages for automatically sensed data classification problems because of their flexibility, intuitive simplicity and computational efficiency (Fayyad and Irani, 1992; Hampson and Volper, 1986). Moreover, decision trees are strictly nonparametric and do not require assumptions regarding the distribution of the input data. In addition, decision trees handle nonlinear relations between features and classes, allow missing data and are capable of handling both numeric and categorical inputs in a natural fashion (Fayyad and Irani, 1992; Hampson and Volper, 1986). To improve the prediction accuracy, random forests have been used as a classification method because they are based on a specific algorithm that uses hundreds of decision trees to achieve a better prediction. Thus, we hypothesize that the identification of states from daily P_p curves can be substantially improved using the random forest classification method. Other classification methods imply some assumptions, such as data normality, and some even require that independent variables are identically distributed. The linear discriminant analysis (LDA), support vector machine (SVM) and artificial neural networks (ANNs) could also be used but they imply relevant assumptions and this curtails their applicability.

This work had three main objectives: (i) to adjust a random forest model to automatically identify the daily states from P_p daily curves recorded in a commercial olive orchard; (ii) to improve the identification of state II curves through a second random forest model; and (iii) to adjust a random forest model to identify the states based on Ψ_{stem} ranges only, without any visual identification. We used a large data set of P_p daily curves collected along five years from the olive orchard in which Fernández et al. (2011) and Padilla-Díaz et al. (2016) worked, and identified characteristic points and tendencies of the P_p variations during the day as predictors to train the model, together with the states visually identified.

2. Material and methods

2.1. Experimental orchard and irrigation treatments

Measurements were made by Fernández et al. (2011) and Padilla-Díaz et al. (2016) in a super-high-density olive orchard (*Olea europaea* L., cv. Arbequina), located at 25 km to the south-east of Seville (37°15′N, −5°48′W). It had trees planted at 4 m × 1.5 m (1667 trees ha^{-1}), at the top of 0.4 m high ridges oriented N–NE to S–SW. The annual average precipitation (P) and potential evapotranspiration (ET_o) in the area are 501.2 mm and 1498.1 mm, respectively (period 2002–2016). The olive trees were five years old in 2011, when the experiments started.

In 2011 and 2012 the olive trees were submitted to three irrigation treatments: (i) full irrigation (FI), in which trees were irrigated for the whole irrigation season to replace 100% of the irrigation needs (IN); (ii) 60% regulated deficit irrigation (60RDI), in which the trees were irrigated with 60% of the total IN, varying depending on the phenological phase water requirements, and; (iii) 30% regulated deficit irrigation (30RDI), in which the trees were irrigated with 30% of the total IN, varying depending on the phenological phase water requirements.

From the results of 2011 and 2012, Fernández (2014b) concluded that the best irrigation strategy should be between 30 and 60% of the IN. Therefore, from 2013 there were two treatments: (i) full irrigation (FI), in which trees were daily irrigated for the whole irrigation season to replace 100% of the irrigation needs (IN), and; (ii) a regulated deficit irrigation treatment ($45RDI_{CC}$), in which the total water supplied along the season was aimed to replace 45% of IN. Both treatments were scheduled with the crop coefficient approach (Allen et al., 1998). See Fernández (2014b) for details. In 2014 and 2015, the olives were also submitted to FI and to $45RDI_{CC}$, but in addition we had a $45RDI_{TP}$ treatment, for which we also used the 45RDI strategy but scheduled from leaf turgor pressure related measurements, i.e. from the outputs of LPCP probes. See Padilla-Díaz et al. (2016) for details.

The treatments were distributed in a randomized block design with four 12 m × 16 m plots per treatment. Each plot contained eight central trees surrounded by 24 border trees. All measurements were taken in the central trees of each plot. The irrigation seasons lasted from May until the end of October or mid-November (see next section for details).

2.2. Plant water status assessment

During the five experimental irrigation seasons, one leaf per tree from two representative trees in three plots per irrigation