

# **Predicting Predawn Leaf Water Potential up to Seven Days Using Machine Learning**

Ahmed A. Fares<sup>( $\boxtimes$ )</sup>, Fabio Vasconcelos, Joao Mendes-Moreira, and Carlos Ferreira

> INESC-TEC, Porto, Portugal ahmed.a.fares@inesctec.pt https://www.inesctec.pt

**Abstract.** Sustainable agricultural production requires a controlled usage of water, nutrients, and minerals from the environment. Different strategies of plant irrigation are being studied to control the quantity and quality balance of the fruits. Regarding efficient irrigation, particularly in deficit irrigation strategies, it is essential to act according to water stress status in the plant. For example, in the vine, to improve the quality of the grapes, the plants are deprived of water until they reach particular water stress before re-watered in specified phenological stages. The water status inside the plant is estimated by measuring either the Leaf Potential during the Predawn or soil water potential, along with the root zones. Measuring soil water potential has the advantage of being independent of diurnal atmospheric variations. However, this method has many logistic problems, making it very hard to apply along all the yard, especially the big ones. In this study, the Predawn Leaf Water Potential (PLWP) is daily predicted by Machine Learning models using data such as grapes variety, soil characteristics, irrigation schedules, and meteorological data. The benefits of these techniques are the reduction of the manual work of measuring PLWP and the capacity to implement those models on a larger scale by predicting PLWP up to 7 days which should enhance the ability to optimize the irrigation plan while the quantity and quality of the crop are under control.

**Keywords:** Precision agriculture *·* Leaf Water Potential *·* Machine Learning

## **1 Introduction**

The best procedure for determining irrigation needs is to measure the crop evapotranspiration  $(Et)$ , i.e., the amount of transpired water in the plant or its estimation. Several methods can be applied to estimate the *Et* but the most popular international method is described in FAO-56 Penman-Monteith (FAO-56) [\[14](#page-11-0)[,17](#page-11-1)[–19\]](#page-11-2). It calculates evapotranspiration reference *ETo*, the Crop Coefficient

G. Marreiros et al. (Eds.): EPIA 2021, LNAI 12981, pp. 39–50, 2021. [https://doi.org/10.1007/978-3-030-86230-5](https://doi.org/10.1007/978-3-030-86230-5_4)\_4

<sup>-</sup>c Springer Nature Switzerland AG 2021

(*Kc*) and Water Stress Coefficient (*Ks*) when the plant culture diverges from its hydric comfort or it is subjected to deficit irrigation, as described in Eq. [1.](#page-1-0)

<span id="page-1-0"></span>
$$
ET = ETo \times Kc \times Ks \tag{1}
$$

*Ks* returns information about the water status inside the plant. This value is hard to calculate because it needs information about root morphology and soil surrounding the roots. To solve this difficulty, different methods to measure plant water status are being used nowadays [\[2](#page-10-0)]. The pressure chamber technique is considered the most accurate procedure available for plant water stress monitorization [\[6\]](#page-10-1). However, this technique requires manual work with a large pressure chamber. So, the implementation of this method on a large scale requires a large number of workers, each one equipped with a pressure chamber, which raises the financial cost of the technique. In vine, changes in water status have a direct effect on grape composition and quality. There is a growing interest in applying deficit irrigation strategies to reach a predetermined water stress level on the crop [\[7](#page-10-2)]. Therefore, this study aims to develop a stand-alone working model using Machine Learning techniques to predict the water stress inside the plant.

Section [2](#page-1-1) presents the state of the art and related work in the same area; Sect. [3](#page-3-0) shows a summary of the data, discussing the problem the client was facing and the experiments to predict Predawn Leaf Water Potential (PLWP); Sect. [4](#page-5-0) explains the experiment; Sect. [5](#page-7-0) shows the results of our models and the discussion about it; Sect. [6](#page-9-0) concludes our work and describes future directions.

## <span id="page-1-1"></span>**2 Background Concept**

Knowing water status response is essential to obtain a balance between the quality of grapes and the yield [\[4\]](#page-10-3). Several indicators can be used to estimate this response. However, Leaf Water Potential (LWP) measured with a pressure chamber is a widely used indicator with an acceptable performance [\[8](#page-10-4)]. These measures can be taken along the day, but implementing it at predawn was favored as it is considered to represent soil water status more accurately since it minimizes the influence of environmental conditions, as shown in Fig. [1](#page-2-0) and demonstrated in [\[5](#page-10-5),[20\]](#page-11-3). However, adverse environmental conditions can affect leaf stomatal opening, which leads to gaps between PLWP and Soil Water Potential(SWP) [\[3](#page-10-6)]. Figure [1](#page-2-0) shows that Water Potential always has a negative value, where values closer to 0 Megapascal (MPa) indicate hydric comfort, while lower values represent water stress. According to  $[21]$ , in Fig. [2,](#page-2-1) after bud burst phenological stage, it is not recommended to put the plant in water stress so it won't affect the bud growth. After Bloom, until Veraison, several restrictions of water can reduce the number of grapes. Between Veraison to Harvest, the water potential has a significant impact on grape size. A controlled reduction of grape size is related to the quality goal of the product. Therefore, it is essential to identify the periods when the crop is less sensitive and define the level of DI to be applied [\[25](#page-11-5)].



<span id="page-2-0"></span>**Fig. 1.** Daily changes in the water potential (Represented as  $\psi$ ) in the soil, leaf and root under normal conditions.



<span id="page-2-1"></span>**Fig. 2.** The optimum PLWP range during different phenological stages. Green - Optimum; Yellow - unfavorable; Red - harmful [\[21](#page-11-4)]. (Color figure online)

Yang et al. [\[26\]](#page-11-6) forecasted daily 7-day-ahead reference crop evapotranspiration (ETo) using the Penman-Monteith (PM) modeled public weather forecasts (including daily maximum and minimum temperatures, weather types, and wind scales, for six stations located in a wide range of climate zones of China were collected). Pelosi et al. [\[27\]](#page-11-7) evaluate the performances of probabilistic daily ETo forecasts with lead times up to 5 days and a spatial resolution of 7 km, computed by using COSMO-LEPS outputs (provided by the European Consortium for small–scale modeling, COSMO). Brillante et al. [\[28\]](#page-11-8) monitored weekly for three years leaf the water potentials Grapevines (Vitis vinifera L. cv Chardonnay) located in eight experimental plots (Burgundy, France). The water stress experienced by grapevine was modeled as a function of meteorological data (minimum and maximum temperature and rainfall, obtained from an on-site weather station) and soil characteristics (soil texture, gravel content, and slope) by a gradient boosting machine. The developed models reached outstanding prediction performance, comparable to the measurement accuracy.

The FAO-56 method [\[14](#page-11-0)] is being used for a long time to compute the crop water requirements and irrigation requirements based on soil, climate, and crop data. Recently, with the increasing availability of high-resolution Normalized Difference Vegetation Index (NDVI) time series, several authors are coupling the FAO-56 method with NDVI images [\[24](#page-11-9)[,29](#page-11-10)]. For instance, the SAMIR (SAtellite Monitoring of IRrigation) tool [\[24](#page-11-9)] is based on the coupling of the FAO-56 dual crop coefficient model with time series of high-resolution NDVI imagery (Normalized Difference Vegetation Index) and can be used to compute spatially distributed estimates of ET and crop water budget at the regional scale. In [\[29\]](#page-11-10) the SAMIR tool was used to estimate regional crop water consumption. In this work, the author explores time series images taken by the SPOT satellite, a commercial high-resolution optical imaging Earth observation satellite system operating from space. The target was to predict the actual basal crop coefficient (Kcb) and the vegetation fraction cover (fc).

## <span id="page-3-0"></span>**3 Materials and Methods**

## **3.1 Experimental Field**

Experiments are carried out using data collected, between 2014 and 2016, from *Herdade do Esporão SA*. regarding vineyard located in 38°23'55.0"N 7◦32 47*.*3W, in the Alentejo region of Portugal with a total area of 452.865 ha. The vineyard is divided into 163 fields called (*Talh˜ao*), according to different soil types, grape varieties (*Casta*), and strategy of irrigation and fertilization. Espor˜ao vineyard is humid mesothermal with dry, hot summer (Csa, Koppen classification), with a mean annual temperature of  $16.5\degree\text{C}$ , mean yearly rainfall of 569 mm.

Usually, PLWP measures are collected using mature leaves located in the middle third of the plant using the pressure chamber method of Scholander [\[1\]](#page-10-7). In order to minimize the bias, each recorded measure is calculated by taking the average of 6 different samples picked from 6 neighbor plants in the same field. This process faces logistic difficulties such as the need for daily manual work done before dawn around 4:00 am to 6:00 am. The number of workers equipped with a Scholander chamber increases linearly with the area of the yard and number of measures.

EnviroScan capacitance sensor is a complete and stand-alone continuous soil moisture monitoring system. The system consists of a network of probes supporting an array of sensors that monitor changes in soil moisture, which could be installed at various depths [\[9\]](#page-10-8). In the current study, the yard has nine sensors distributed strategically, i.e., each set of homogeneous zones according to the soil type, altitude, and irrigation system has one sensor.

## **3.2 Data Visualization and Summarization**

The first task was to normalize the data due to different timescales used in the recording process. Some of them were recorded every 15 min like humidity sensors; others were recorded daily like PLWP, while the rest have only one reading per year as grape variety (*Casta*) and soil characteristics. The inconsistent information collected from annual variables, i.e., Casta, Regime, Soil, Age, CC (Maximum moisture that the soil supports), CE (Minimum moisture that the soil needs before the plants start dying), TAW (Total Available Water), and vigor or the incoherent readings and missing humidity values are detected (see Table [1\)](#page-4-0). PLWP readings recorded after the harvest date were removed because both showed irregular behavior, and there was no interest in collecting or calculating PLWP that late.

#### **3.3 Problem Definition and Feature Engineering**

The original idea is trying to predict PLWP for the next seven days. While 62.8% of PLWP reading was unknown, the first step was filling the unknown values, and then the predicted values can be passed as input variables to the future prediction models.

<span id="page-4-0"></span>**Table 1.** Data summarization, where DOY- Day of the year; PLWP- Predawn Leaf Water Potential Measures (MPa); Hum.- Humidity Measures at 4 am. In the original dataset we had three variables, for different depths (20, 60 and 100 cm); Age- Age of the plant; W1- Amount of water irrigated on the previous day (mm); ETo- Evapotranspiration on the last day (mm); CC- Maximum moisture that the soil supports; CE- Minimum moisture that the soil needs before the plants start dying; TAW- Total Available Water.

	Min	1stQ	Median	Mean	3rdQ	Max	Missing values%
DOY	126	166	197	196.4	228	261	$0\%$
<b>PLWP</b>	$-0.98$	$-0.43$	$-0.32$	$-0.34$	$-0.22$	$-0.06$	62.8%
Hum.	12.23	12.54	12.65	12.64	12.73	13.34	$0.9\%$
Age	$\overline{2}$	8	11	11.8	13	42	10.3%
W <sub>1</sub>	$\Omega$	$\Omega$	$\Omega$	3.34	$\Omega$	23.27	$0\%$
<b>ETo</b>	1.6	5.6	6.2	6.131	6.9	8.6	$0\%$
CC	0.26	0.28	0.31	0.31	0.33	0.39	$1.6\%$
CE	0.11	0.15	0.16	0.16	0.18	0.23	$1.6\%$
<b>TAW</b>	113.4	136.6	150.5	149.2	162.2	183.7	$1.6\%$

There are several strategies to deal with unknown values. The simplest ways are either to delete the whole records with unknown values or to fill them with given statistics such as the average or the median for quantitative values. On the other hand, there are more complex strategies that normally lead to more accuracy; however, it requires more computing costs [\[16\]](#page-11-11). The five different Machine Learning methods from different regression families were applied to fill all the unknown values of PLWP that are Multivariate Linear Regression (MLR), Multivariate Adaptive Regression Splines (MARS), Support Vector Regression (SVR), Classification and Regression Trees (CART), and Random Forest (RF). The experiments were developed in R computing environment<sup>[1](#page-4-1)</sup> by using  $e1071$ package [\[10\]](#page-10-9) for SVR, *rpart* [\[11](#page-10-10)] for regression trees, *earth* [\[12\]](#page-11-12) for MARS and *RandomForest* [\[13](#page-11-13)] for *RandomForest*. The SVR, MARS, and *RandomForest* are tuned using the function inside the respective package. The MLR and *rpart* are used with the standard hyper-parameters. The 10-fold cross-validation is used as a resampling method to evaluate each method by random partitioning. Each subset is used to evaluate the induced model, which has been trained using the

<span id="page-4-1"></span><sup>1</sup> [https://www.r-project.org/about.html.](https://www.r-project.org/about.html)

remaining nine subsets. This method was applied to data available from 2014. The performance measure is Root Mean Squared Error (RMSE), according to Eq. [2.](#page-5-1)

<span id="page-5-1"></span>
$$
RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} Error_t^2}
$$
 (2)

Where "Error" is the difference between the predicted value of PLWP by model and the measured value of PWLP. The variable *n* is the number of samples. In a general sense, soil stores water; therefore, the water on the soil should be a continuous variable, so the variable "HWater" was created according to Eq. [3.](#page-5-2) Besides, the balance (BAL) variable was derived from the ideal balance of water inside the plant. Equation [4](#page-5-3) shows the difference between irrigation and evapotranspiration.

<span id="page-5-2"></span>
$$
HWater = \frac{\sum_{i=1}^{3} \frac{1}{i} \times Humidity_{t-i} \times W_{t-i}}{\sum_{i=1}^{3} \frac{1}{i}}
$$
(3)

<span id="page-5-3"></span>
$$
BAL = W_{t-1} - ETo_{t-1} \tag{4}
$$

### <span id="page-5-0"></span>**4 Experiments**

#### <span id="page-5-4"></span>**4.1 Fill the Gaps**

**Variables Selection.** The best variables were chosen using the *rfcv* function from *randomForest* package to perform a 10-fold cross-validation over the data from year 2014 with all the variables. Afterward, a cut point was chosen to select the most important variables to avoid overfitting and complexity of the system without compromising the accuracy.

The *varImpPlot* function from *randomForest* R package [\[13](#page-11-13)] has been used to know the variable importance, as it is shown in Fig. [3.](#page-6-0) DOY was the most important feature. According to the plant life cycle, the behavior of PLWP could differ during phenological stages. It can explain the importance of DOY since the dates of phenological stages were not available. Humidity in different depths also seems to be important, representing the absolute quantity of water in the soil. Since PLWP is a measure that represents the water inside the plant, the importance of this variable to the model makes sense. Other variables that seem to be important are *Casta*, Age, and W1. These are the variables that distinguish between plant characteristics and irrigation strategies. The ETo considers weather information to calculate the amount of water lost by the plant and seems to have some importance. The CE, CC, TAW, and Soil variables are correlated, and all of them are describing the soil characteristics. We can conclude that the soil type is important to predict PLWP, which is further supported since the available water is different for different types of soil [\[22\]](#page-11-14). According to this way of calculating variable importance, vigor and regime seem to be the least important variables.



<span id="page-6-0"></span>**Fig. 3.** Variable importance measures by *varImpPlot* function

**Methodologies.** The objective was to build a stand-alone working model to fill the gaps of PWLP for the previous agricultural cycles. This model will be used to train the prediction models on the one hand and also will be used to fill the gaps of the following agricultural cycles whenever needed.

Figure [4](#page-7-1) shows the proposed cascading (CAS) technique to train a model for the prediction of PLWP and then predicted PLWP would be used as a variable for the prediction of the next day PLWP. To be able to do that, the algorithm was split into two different tasks: (1) fill the missing data, and (2) build model 2. So, instead of ignoring all the daily data with no values of PLWP, these values are predicted using model 1 afterward. These predictions are used as a variable to model 2. The big modification on the model 2 are the values of PLWP for the previous 3 days ( $T_i$  represents  $PLWP_{t-i}$ ; *t* is the current day and *i* the number of previous days) that are from model 1 and the respective modification of PLWP between two days  $(C_i = PL\hat{W}P_{t-i} - PL\hat{W}P_{(t-1)-i}).$ 

#### **4.2 Seven Days Prediction**

Like filling the gaps models, the random forest was chosen to train seven models to predict PLWP for the next seven days, one model per day.

**Variables Selection.** The available variables were *ETo*, Field data, Irrigation, and PLWP of the previous seven days.

The variable *Mean*7*DaysCurrent* was the most critical variable in all the seven models, which means the predicted value of a day is highly dependent on the weighted average of the seven days before that day, giving higher weights for closer days. Also, irrigation variables like Water*xxx* (the average of the irrigation



<span id="page-7-1"></span>**Fig. 4.** Proposed cascading algorithm

of the three days before that day), EToW*x* (the difference between ETo and irrigation for the day *x*) and Bal*xxx* (the average of the difference between ETo and irrigation of the three days before that day), were good candidates with moderate importance in predicting the following days.

**Methodologies.** The value of PLWP always depends on DOY, as mentioned in Sect. [4.1.](#page-5-4) We had split the experiment into two parts and compared the performance. (1) Creating one model for all phenological stages. (2) Creating three models (one model per phenological stage).

## <span id="page-7-0"></span>**5 Results and Discussion**

## **5.1 Algorithms Comparison and Variable Importance**

Five different algorithms were used to choose the model that fits better with the year 2014. It was concluded that Support Vector Regression (SVR) with RMSE of 0.0812 and Random Forest with RMSE of 0.0791 obtained the best results, followed by MARS, RPart, and MLR with RMSE of 0.101, 0.105, and 0.110, respectively. Random Forest usually receives a good performance, and the induced models are of straightforward interpretation. Besides, we don't need to worry about tuning a large set of parameters that increase the computational costs [\[15](#page-11-15)]. Therefore, the random forest has been chosen as the best of the five algorithms tested for the current study while passing only two parameters *importance* is true and *nTrees* is 2000 trees for filling gaps model and 600 trees for each of the seven days models.

While *HW ater* and *BAL* are calculated based on the amount of irrigation, it seems that there could be a chance to use them in understanding and control the performance of PLWP. The right panel depicts the previous values PLWP has of high significance, especially the previous day (T1). Moreover, the changes (C1 and C2) seem to have some importance.

#### **5.2 Models Validation**

Specialists from *Herdade do Espor˜ao* agreed that an absolute error up to 0.2 Megapascal (MPa) is acceptable. An interesting observation is that all the algorithms seem to have very similar results, and all of them seem to have more accurate results when the measured values are higher than *<sup>−</sup>*0*.*7 MPa. At the same time, values smaller than  $-0.7$  MPa are being predicted with values higher than their actual values. This situation can be explained by the limit between healthy and non-healthy plant behavior, supporting the idea of changing the plant behavior when the plant starts dying. Contrarily, PLWP above *<sup>−</sup>*0*.*2 MPa, the plant doesn't feel any limitation in water uptake [\[23](#page-11-16)], and it can be observed that our prediction is weak in this range.

Accepting a 0*.*2 MPa error, we can conclude that all the methods have around 98% of the predictions inside this range, except persistence. i.e. 98% of the predicted values have absolute error less than 0*.*2 MPa. Moreover, when we decrease the maximum acceptable error to 0*.*1 MPa, we got around 80% of accuracy.

#### **5.3 Error Analysis**

**Fill the Gaps.** This result shows that it exists a slight asymmetry in our predictions. In each day, it was collected more than one measure of PLWP. Thus it is possible to calculate RMSE/day. Also, it shows that cascading Random Forest has better results for all the scenarios, and all trials have beat the dummy persistence algorithm.

**Predict Seven Days.** To check the performance of the models, we have tested two different approaches. a) create one model per phenological stage, b) create only one model which covers the whole year. Then, Root Mean Square Error has been calculated as a validation metric. Also, the holdout methodology has been used as a validation technique. The idea was to keep the phenological stages consistent within a single year and across different years.

Tables [2](#page-9-1) shows Root Mean Square Error when a single model was created for the whole year. It is understood that the accuracy of the models was decreasing while we tried to predict farther days. But at the same time, it was not reducing dramatically, and in the worst case to predict the 7th day, the Root Mean Square Error is higher than the error of predicting the current day by only 15%.

On the other side, Table [3](#page-9-2) shows Root Mean Square Error when three models per year were created representing the three phenological stages. The results show a more extensive range of errors resulting from having different phenological stages in different years, especially in the transition from a stage to the following.

<span id="page-9-1"></span>

	Train 2014+2015	Train $2014 + 2016$	Train $2015+2016$
	Test 2016	Test 2015	Test 2014
$T+1$	0.084	0.086	0.084
	$T+2$ 0.083	0.087	0.084
	$T+3$ 0.083	0.088	0.084
	$T+4$   0.085	0.089	0.085
$T+5$	$\mid 0.086$	0.091	0.087
$T+6$	0.087	0.092	0.088
$T+7$	0.089	0.092	0.089

**Table 2.** Root Mean Square Error considering one model for all phenological stages

**Table 3.** Root Mean Square Error considering 3 models (one model per stage)

<span id="page-9-2"></span>

	Train $2014 + 2015$	Train $2014 + 2016$	Train $2015+2016$
	Test 2016	Test 2015	Test 2014
$T+1$	0.068 0.108 0.138		$0.082$ $0.098$ $0.103$   $0.080$ $0.093$ $0.098$
$T+2$	$0.067$ 0.106 0.146	$0.082$ $0.097$ $0.105$ $0.079$ $0.090$ $0.098$	
$T+3$	$0.068$ 0.107 0.149		$0.083$ $0.098$ $0.108$   $0.078$ $0.091$ $0.104$
$T+4$	$\mid 0.069 \; 0.112 \; 0.156 \mid$	$0.084$ 0.099 0.111 0.080 0.093 0.106	
		$T+5$ 0.070 0.113 0.154 0.085 0.102 0.117 0.081 0.095 0.116	
$T + 6$		$0.071$ 0.114 0.165 0.085 0.104 0.121 0.080 0.096 0.121	
$T+7$		$0.073$ $0.120$ $0.172$ $0.086$ $0.103$ $0.122$ $0.081$ $0.095$ $0.128$	

## <span id="page-9-0"></span>**6 Conclusion and Future Work**

In this project, several models were developed to predict PLWP at a specific time in the vineyard, from the flowering phenological stage until the maturation stage. An easy to collect information like grape varieties and soil type, moisture, and meteorological information was considered. The results showed the possibility to predict PLWP instead of physical examination that consumes time and money. Specialists from *Herdade do Esporão S.A.* defined the maximum acceptable error rate to be 0*.*2 MPa, so at this point, we conclude that the objective was accomplished by having around 98% of the predictions with error rates less than 0*.*2 MPa. Regarding different strategies, it seems that the cascading approach brings a slight improvement, so considering computational cost versus benefit does not seem to worth it.

Also, we have been able to forecast PLWP for the following seven days with an accuracy lower than predicting the current point of time by only 15%, which is considered an original work that should be followed by future enhancement.

#### **6.1 Future Work**

While the results look promising, we believe it could be even enhanced using more information regarding plagues, stomatal opening, root morphology, phenological stages, and NDVI (Normalized Difference Vegetation Index) information. Also, evolved models could improve themselves over time and include data from other vineyards to generalize these models.

As future work, we could focus on optimizing irrigation plans using our forecasting models once time and quantity of irrigation water are considered important decision variables.

**Acknowledgments.** This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020.

## **References**

- <span id="page-10-7"></span>1. Scholander, P.F., Bradstreet, E.D., Hemmingsen, E.A., Hammel, H.T.: Sap pressure in vascular plants. Science **148**, 339–346 (1965)
- <span id="page-10-0"></span>2. Jones, H.G.: Monitoring plant and soil water status: established and novel methods revisited and their relevance to studies of drought tolerance. J. Exp. Bot. **58**, 119– 130 (2007)
- <span id="page-10-6"></span>3. Tonietto, J., Carbonneau, A.: A multicriteria climatic classification system for grape-growing regions worldwide. Agric. For. Meteorol. **124**, 81–97 (2004)
- <span id="page-10-3"></span>4. Acevedo-Opazo, C., Ortega-Farias, S., Fuentes, S.: Effects of grapevine (Vitis vinifera L.) water status on water consumption, vegetative growth and grape quality: an irrigation scheduling application to achieve regulated deficit irrigation. Agric. Water Manag. **97**, 956–964 (2010)
- <span id="page-10-5"></span>5. Yamane, T., Shibayama, K., Hamana, Y., Yakushiji, H.: Response of containergrown girdled grapevines to short-term water-deficit stress. Am. J. Enol. Vitic. **60**, 50–56 (2009)
- <span id="page-10-1"></span>6. Acevedo-Opazo, C., Tisseyre, B., Guillaume, S., et al.: The potential of high spatial resolution information to define within-vineyard zones related to vine water status. Precision Agric. **9**, 285–302 (2008)
- <span id="page-10-2"></span>7. Acevedo-Opazo, C., Tisseyre, B., Ojeda, H., Ortega-Farias, S., Guillaume, S.: Is it possible to assess the spatial variability of vine water status? OENO One **42**, 203–219 (2008)
- <span id="page-10-4"></span>8. Am´eglio, T., et al.: Significance and limits in the use of predawn leaf water potential for tree irrigation. Plant Soil **207**, 155–167 (1999)
- <span id="page-10-8"></span>9. Wels, C., O'Kane, M., Fortin, S.: Assessment of water storage cover for Questa tailings facility, New Mexico. In: Proceedings of the 9th Annual Conference of the American Society for Surface Mining Reclamation, Albuquerque, New Mexico (2001)
- <span id="page-10-9"></span>10. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C.-C., Lin, C.-C.: "The e1071 package" in Misc Functions of Department of Statistics (e1071), TU Wien (2006)
- <span id="page-10-10"></span>11. Therneau, T., Atkinson, B.: rpart: recursive partitioning and regression trees. R package version 4.1-15 (2019)
- <span id="page-11-12"></span>12. Milborrow. S.: Derived from mda:mars by T. Hastie and R. Tibshirani., "earth: Multivariate Adaptive Regression Splines" (2011)
- <span id="page-11-13"></span>13. Liaw, A., Wiener, M.: Classification and regression by randomForest. R News **2**, 18–22 (2002)
- <span id="page-11-0"></span>14. Allan, R.G., Pereira, L.S., Raes, D., Smith, M.: Crop evapotranspiration - Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, vol. 300, p. D05109. FAO, Rome (1998)
- <span id="page-11-15"></span>15. Chen, E.: Choosing a Machine Learning Classifier (2011)
- <span id="page-11-11"></span>16. Torgo, L.: Data Mining with R: Learning with Case Studies, 1st edn. Chapman and Hall/CRC (2016)
- <span id="page-11-1"></span>17. Suleiman, A.A., Hoogenboom, G.: Comparison of Priestley-Taylor and FAO-56 Penman-Monteith for daily reference evapotranspiration estimation in Georgia. J. Irrig. Drain. Eng. **133**, 175–182 (2007)
- 18. Mutziger, A.J., Burt, C.M., Howes, D.J., Allen, R.G.: Comparison of measured and FAO-56 modeled evaporation from bare soil. J. Irrig. Drain. Eng. **131**, 59–72 (2005)
- <span id="page-11-2"></span>19. de Jabloun, M., Sahli, A.: Evaluation of FAO-56 methodology for estimating reference evapotranspiration using limited climatic data: application to Tunisia. Agric. Water Manag. **95**, 707–715 (2008)
- <span id="page-11-3"></span>20. Ribeiro, A.C., Sá, A., Andrade, J.L.: Avaliação do stresse hídrico em videiras submetidas a diferentes regimes hídricos. In: VI Congresso Ibérico de Agro-Engenharia (2011)
- <span id="page-11-4"></span>21. Ojeda, H.: Riego cualitativo de precisi´on en la vid. Revista Enologia **1**, 14–17 (2007)
- <span id="page-11-14"></span>22. Cassel, D.K., Nielsen, D.R.: Field capacity and available water capacity. In: Methods of Soil Analysis: Part 1-Physical and Mineralogical Methods, pp. 901–926 (1986)
- <span id="page-11-16"></span>23. Van Leeuwen, C., et al.: Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes? OENO One **43**, 121–134 (2009)
- <span id="page-11-9"></span>24. Lepage, M., et al.: SAMIR a tool for irrigation monitoring using remote sensing for evapotranspiration estimate. Marrakech. Melia (2009)
- <span id="page-11-5"></span>25. Fernandes-Silva, A., Oliveira, M., Paço, T.A., Ferreira, I.: Deficit irrigation in Mediterranean fruit trees and grapevines: water stress indicators and crop responses. In Irrigation in Agroecosystems. IntechOpen (2018)
- <span id="page-11-6"></span>26. Yang, Y., et al.: Short-term forecasting of daily reference evapotranspiration using the Penman-Monteith model and public weather forecasts. Agric. Water Manag. **177**, 329–339 (2016)
- <span id="page-11-7"></span>27. Pelosi, A., Medina, H., Villani, P., D'Urso, G., Chirico, G.B.: Probabilistic forecasting of reference evapotranspiration with a limited area ensemble prediction system. Agric. Water Manag. **178**, 106–118 (2016)
- <span id="page-11-8"></span>28. Brillante, L., Bois, B., Mathieu, O., Lévêque, J.: Electrical imaging of soil water availability to grapevine: a benchmark experiment of several machine-learning techniques. Precision Agric. **17**(6), 637–658 (2016). [https://doi.org/10.1007/s11119-](https://doi.org/10.1007/s11119-016-9441-1) [016-9441-1](https://doi.org/10.1007/s11119-016-9441-1)
- <span id="page-11-10"></span>29. Saadi, S., et al.: Monitoring irrigation consumption using high resolution NDVI image time series: calibration and validation in the Kairouan Plain (Tunisia). Remote Sens. **7**(10), 13005–13028 (2015)