A decision support system for managing irrigation in agriculture

Navarro Hellin, H., Martínez-del-Rincon, J., Domingo Miguel, R., Soto Valles, F., & Torres Sanchez, R. (2016). A decision support system for managing irrigation in agriculture. Computers and Electronics in Agriculture, 124, 121-131. DOI: 10.1016/j.compag.2016.04.003

Published in:
Computers and Electronics in Agriculture

Document Version:
Peer reviewed version

A Decision Support System for managing irrigation in agriculture

H. Navarro-Hellín[a], J. Martínez-del-Rincon[b], R. Domingo-Miguel[c], F. Soto-Valles[d] and R. Torres-Sánchez[e*]

[a] Widhoc Smart Solutions S.L. Parque Tecnológico de Fuente Álamo, CEDIT. Carretera del Estrecho-Lobosillo Km 2 30320, Fuente Álamo. (Murcia), Spain.

[b] The Institute of Electronics, Communications and Information Technology (ECIT), Queens University of Belfast, Belfast BT3 9DT, UK

[c] Producción Vegetal Department, Universidad Politécnica de Cartagena, Paseo Alfonso XIII, 48 30203, Cartagena (Murcia), Spain

[d] Tecnología Electrónica Department, Universidad Politécnica de Cartagena, Campus Muralla del Mar, Doctor Fleming, s/n 30202, Cartagena (Murcia), Spain

[e] Ingeniería de Sistemas y Automática Department, Universidad Politécnica de Cartagena, Campus Muralla del Mar, Doctor Fleming, s/n 30202, Cartagena (Murcia), Spain

**Corresponding author at:* Ingeniería de Sistemas y Automática Department, Universidad Politécnica de Cartagena, Campus Muralla del Mar, Doctor Fleming, s/n 30202, Cartagena (Murcia), Spain.

Tel.: +34 968 325 474

E-mail address: roque.torres@upct.es

Abstract

In this paper, an automatic Smart Irrigation Decision Support System, SIDSS, is proposed to manage irrigation in agriculture. Our system estimates the weekly irrigations needs of a plantation, on the basis of both soil measurements and climatic variables gathered by several autonomous nodes deployed in field. This enables a closed loop control scheme to adapt the decision support system to local perturbations and estimation errors. Two machine learning techniques, PLSR and ANFIS, are proposed as reasoning engine of our SIDSS. Our approach is validated on three commercial plantations of citrus trees located in the South-East of Spain. Performance is tested against decisions taken by a human expert.

Keywords: Irrigation, Decision Support System, water optimization, machine learning.

1 Introduction

The efficient use of water in agriculture is one of the most important agricultural challenges that modern technologies are helping to achieve. In arid and semiarid regions, the differences between precipitation and irrigation water requirements are so big that irrigation management is a priority for sustainable and economically profitable crops (IDAE, 2005).

To accomplish this efficient use, expert agronomists rely on information from several sources (soil, plant and atmosphere) to properly manage the irrigation requirements of the crops (Puerto et al., 2013). This information is defined by a set of variables, which can be measured using sensors, that are able to characterise the water status of the plants and the soil in order to obtain their water requirements. While meteorological variables are representative of a large area and can be easily measured by a single sensor for a vast land extension, soil and plant variables have a large spatial variability. Therefore, in order to use these parameters to effectively schedule the irrigation of the plants, multiple sensors are needed (Naor et al., 2001).

Weather is one of the key factors being used to estimate the water requirements of the crops (Allen et al., 1998). Moreover, it is very frequent that public agronomic management organisms have weather stations spread around the different regions. These weather stations usually provide information of key variables for the agriculture like reference evapotranspiration (ET_0) or the Vapour Pressure Deficit (VPD) that are of great importance to calculate the water requirements of the crops. Using variables related to the climate is the most common approach to create crop water requirement models (Jensen et al., 1970; Smith, 2000; Zwart and Bastiaanssen, 2004). Using these models, based on solely meteorological variables, a decision-making system can determine how a given crop will behave (Guariso et al., 1985).

However, not all the regions have access to an extensive network of weather stations or they may not be nearby a given crop, thus the local micro-climates are not taken into account if only these parameters are used. Besides, irrigation models based only on climate parameters rely on an open loop structure. This means that the model is subject to stochastic events and it may not be able to correct the local perturbations that can occur when a unexpected weather phenomenon occurs (for instance irrigate the crop when it’s already raining) (Dutta et al., 2014; Giusti and Marsili-Libelli, 2015). Finally, monitoring other variables, such as hydrodynamic soil factors or water drainage, might increase the chances that the irrigation predicted by the models is properly used by the plants (Kramer and Boyer, 1995). Therefore, the usage of sensors that measures the soil water status is a key complement to modulate the water requirements of the crops. Soil variables, such as soil moisture content or soil matric potential, are considered by many authors as crucial part of scheduling tools for managing irrigation (Cardenas-Lailhacar and Dukes, 2010; Soulis et al., 2015). The information from soil sensors can be used to create better decision models with closed loop structures that adapt to weather and soil perturbations (Cardenas-Lailhacar and Dukes, 2010; Soulis et al., 2015). This practice, however, has not been widely adopted due to the technological limitations of available soil sensors, which required measured information to be registered and stored, traditionally using wired dataloggers, and limiting the installation flexibility and the real time interaction. This has changed recently with new generation sensors and sensor networks that are more versatile and suited to the agricultural environment (Navarro-Hellín et al., 2015).

Combining climate and soil variables has therefore potential to properly manage irrigation in a more efficient way than other traditional approaches. However, it also entails a series of challenges related with the increased amount of data flow, its analysis and its use to create effective models, in particular when data provided by different sources may seem contradictory and/or redundant. Traditionally, this analysis and modelling is performed by a human expert who interprets the different variables. The need of a human agronomist expert is required due to the complexity introduced by the soil spatial variability, crop species variability and their irrigation requirements over the growth cycle (Maton et al., 2005), which require comparing crops models and local context variables to determine the specific water requirements to achieve a certain goal at a particular location.

The complexity of this problem and the different sources of variability makes than even the best model may deviate from the prediction, which favours the use of close loop control systems combining soil and climate sensors over open loop systems as a way to compensate possible deviations in future predictions.

Human expertise has been proved effective to assist irrigation management but it is not scalable and available to every field, farm and crop and it is slow in the analysis of the data and real time processing. Instead, applying machine learning techniques to replace the manual models and to assist expert agronomists allows the viability of creating automatic Irrigation Decision Support System. Machine learning techniques have been used previously to estimate relevant parameters of the crop (Sreekanth et al., 2015). Giusti and Marsili-Libelli(2015) present a fuzzy decision systems to predict the volumetric water content of the soil based on local climate data. Adeloye et al. (2012), proposed the use of unsupervised artificial neural networks (ANN) to estimate the evapotranspiration also based on weather information solely. King and Shellie (2016) used NN modelling to estimate the lower threshold temperature (Tnws) needed to calculate the crop water stress index for wine grapes. In Campos et al. (2016) the authors presented a new algorithm designed to estimate the total available water in the soil root zone of a vineyard crop, using only SWC sensors, which are very dependent of the location. Taking advantage of the soil information, Valdes-Vela et al.(2015) and Abrisqueta et al.(2015) incorporates the volumetric soil water content, manually collected with a neutron probe, to agro-meteorological data. This information is then fed into a fuzzy logic system to estimate the stem water potential. Other approaches in the literature also make use of machine learning techniques -such as principal component analysis, unsupervised clustering, ANN, etc.- to estimate the irrigation requirements in crops. However they do not specify the quantity of water needed (Dutta et al., 2014), they reduce the prediction to true or false, and/or they are based on open loop structures (Giusti and

Marsili-Libelli, 2015; Jensen et al., 1970; Smith, 2000; Zwart and Bastiaanssen, 2004), only considering the weather information and, therefore, unable to correct deviations from their predictions.

This paper proposes an automated decision support system to manage the irrigation on a certain crop field, based on both climatic and soil variables provided by weather stations and soil sensors. As discussed, we postulate that the usage of machine learning techniques with the weather and soil variables is of great importance and can help to achieve a fully automated close loop system able to precisely predict the irrigation needs of a crop. Our presented system is evaluated by comparing it against the irrigations reports provided by an agronomist specialist during a complete season in different fields.

2 System Structure

An irrigation advice system is based on the concept of predicting the waters needs of the crops in order to irrigate them properly. Traditionally this decision has been taken by an experienced farmer or an expert agricultural technician. Figure 1 shows the flow diagram of which the proposed system is based.

In this schema, an expert agronomist is in charge of analysing the information from different sources: Weather stations located near the crops that collect meteorological data, Crop and Soil characteristics (type, age, size, cycle, etc.) and Soil sensors installed in the crop fields. The expert analyses the information to provide an irrigation report, which indicates the amount of water needed to irrigate properly the crops in the upcoming week. To make this decision making process manageable, the information needed to create the irrigation report on the next week is only the information of the current week.

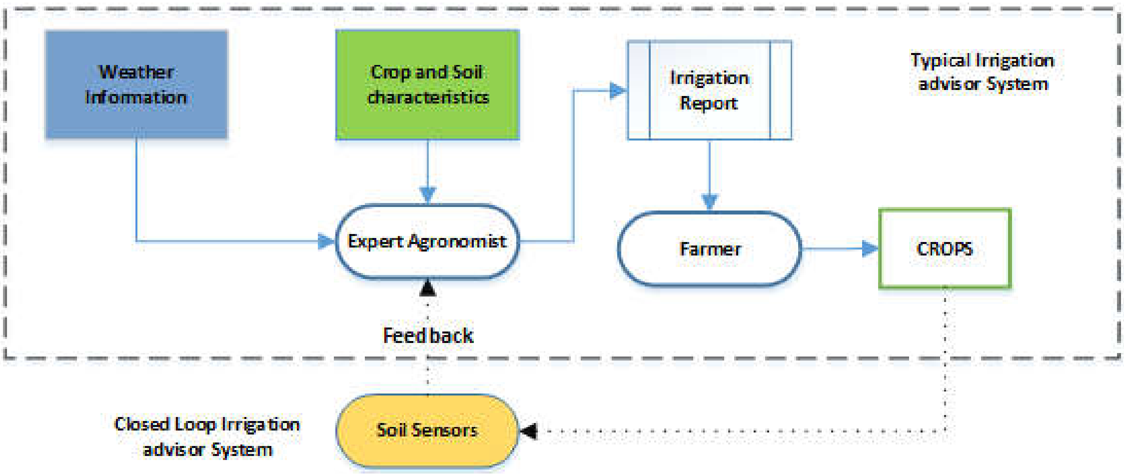

Figure 1: Flow Diagram of the proposed system

Based on this concept, our Smart Irrigation Decision Support System (SIDSS) is proposed. In order to evaluate the performance and validity of our approach, the decision system will use the same information used by the expert agronomist and will output the water requirements for the upcoming week. This will ensure a fair comparison between the decisions taken by a human expert and the SIDSS. To accomplish this, the machine learning system must be trained with historical data and irrigations reports of the agronomist, using the irrigation decisions taken in these reports as the groundtruth of the system. The aim of the system is to be as accurate as possible to this groundtruth. Several machine learning techniques were applied and evaluated to achieve the best performance. Figure 2 shows a diagram of the SIDSS.

The Irrigation Decision System is composed of three main components: a collection device that gathers information from the soil sensors, weather stations that provide agrometeorological information and the SIDSS that, when trained correctly, is able to predict the irrigation requirements of the crops for the incoming week. Table 1 shows the set of possible input variables of the system.

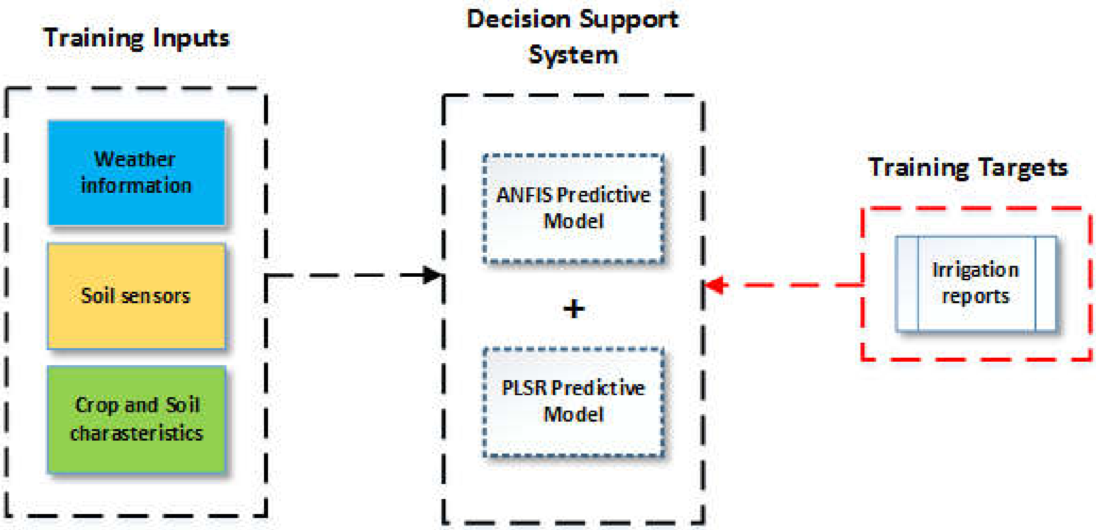

Figure 2: Training inputs and targets of SIDSS

	Name	Symbol	Category
1	Volumetric Water Content depth 1	VWC1	Soil Sensors
2	Volumetric Water Content depth 2	VWC2	
3	Volumetric Water Content depth 3	VWC3	
4	Soil Water Potential	SWP	
5	Soil Temperature	ST	
6	Rainfall	RF	Weather Stations
7	Wind Speed	WS	
8	Temperature	T	
9	Relative Humidity	RH	
10	Global Radiation	GR	
11	Dew Point	DP	
12	Vapour pressure Deficit	VPD	
13	Crop Evapotranspiration	ET_c	Crop and Soil Characteristics + Weather Stations

Table 1: Set of possible input variables of the system

2.1 Collection device and soil sensors

The information from the soil sensors is gathered using our own developed device that has been proved to be completely functional for irrigation management in different crops and conditions (Navarro-Hellin et al., 2015). This device is wireless, equipped with a GSM/GPRS modem, and is completely autonomous, so that the installation procedures are accessible to any farmer.

Figure 3 shows the collection device installed in a lemon crop field located in the South-East of Spain.

Figure 3: Device installed in a lemon crop field.

The device allows to fully configure the recording rates of all the embedded sensors. In our experiments, a sampling rate of 15 minutes was set, since this gives a good balance between providing enough information to support a correct agronomic decision and maintaining the autonomy of the device with the equipped solar panel and battery (López Riquelme et al., 2009; Navarro-Hellin et al., 2015). The information is received, processed and stored in a relational database.

2.1.1. Soil Sensors

The soil control variables used to provide SIDSS with relevant information are matric potential (Ψ_m) and volumetric soil water content (θ_v), which are common in irrigation management (Jones, 2004). By using these variables, the irrigation can be scheduled for maintaining soil moisture conditions equivalent or close to field capacity in order to satisfy the required crop water requirements. Likewise, they can be used to maintain soil water content or soil matric potential under certain reference values proper of regulated deficit irrigation strategies. Both Ψ_m and θ_v are used to decide the irrigation frequency and to adjust the gross irrigation doses.

Soil matric potential was measured with MPS-2 sensors (Decagon devices, Inc., Pullman, WA 99163 - USA), while volumetric soil water content was measured with both 10-HS (Decagon devices, Inc., Pullman, WA 99163 - USA) and Enviroscan (Sentek Pty. Ltd., Adelaide, Australia) sensors

Besides both previous soil sensors, another sensor is used. A pluviometer (Rain-o173-matic small, Pronamic Ltd., Ringkøbing, Denmark) was used under the dripper to provide accurate estimation of the amount of water applied and the irrigation run time. The information provided by this sensor was used to ensure that the farmer is following the instruction of the agronomic reports provided by the expert. Table 2 summarizes the variables measured by the soils sensors.

Sensor	Measured data	Variable name	Range	Resolution	Supply voltage range	Output	URL
10HS	soil moisture	VWC1,VWC2,VWC3	0 to 57 % VWC	0.08%VWC	3-15 VDC	0.3-1.25 V	http://www.decagon.com/
MPS-2	soil matric potential and temperature	SWP ST	-10 to -500 kPa -40º to +50 ºC	0.1 kPa 0.1°C	6-15 VDC	SDI-12	http://www.decagon.com/
Enviroscan	soil moisture	VWC1,VWC2,VWC3	0 to 65% VWC	0.003 %VWC	8-32 VDC	4-20 mA	http://www.sentek.com.au/

Table 2: Soil sensors technical information

2.2 Weather Stations

Experiments took place in the Region of Murcia, Spain. In this region, there is a network of 45 agro-meteorological stations located in irrigated areas, the Agricultural Information Network System of Murcia (SIAM), funded by the EU and installed to help estimate the reference evapotranspiration (ET_0) and the irrigation needs of crops after a severe drought between 1979 and 1985.

The variables measured by the stations are the following:

Temperature (T),Relative humidity (RH),Global radiation (GR),Wind speed (WS), Rainfall (RF), Dew point (DP),Vapour pressure deficit (VPD).

These variables, measured by the different stations, are publicly available and can be downloaded from the SIAM website (SIAM, 2015). The weather stations are tested and calibrated periodically according to the manufacturer's specifications.

The amount of water required to compensate the evapotranspiration loss from the cropped field is defined as crop water requirement. Therefore, knowing the reference crop evapotranspiration is of key importance to estimate the crop´s water requirements. Using the FAO Penman-Monteith formulation (Allen et al., 1998), the daily reference crop evapotranspiration (ET_0) can be calculated by means of the weather information. The crop evapotranspiration under standard condition (ET_c) can be calculated using the single crop coefficient approach shown below:

$$ETc = K \cdot ETo \quad [1]$$

where K_c is the crop coefficient and depends on multiple factors, namely, the crop type, climate, crop evaporation and soil growth stages.

2.3 Smart Irrigation Decision Support System

The decision support system is the component in charge of taking the final decision on the amount of water to be irrigated, or equivalently, the number of minutes to irrigate considering constant water flow. This decision is taken automatically on the basis of the information provided by the sensors and the usage of machine learning and pattern recognition techniques. The aim of this component, therefore, is to mimic a human expert in the decision making process of weekly optimising the irrigation, which could assist the farmer.

Applying machine learning techniques such as Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA) allow us to visualize the information to perform an initial exploratory analysis. Figure 4 shows the LDA of the input, array containing the sensorial variables, and output, the estimated irrigation time need, used in the system. The output was divided in classes (18), each one representing the weekly irrigation time by increments of 150 minutes, from 0 to 2,700 minutes. From this figure, it can be noticed that discrete classification in classes will be hard to accomplish due to the high number of classes necessary to precisely quantise the irrigation estimation. This is due to the fact that the variable to estimate - either the amount of water or the watering time- has an intrinsic continuous nature, since the expected output can take any real value between 0 and infinity. Therefore, conventional classifiers aiming categorical outputs - such as LDA (Fisher, 1938), SVM (Belousov et al., 2002), ANN etc. are not optimal for this application. Instead, methodologies based on regression (Wold et al., 1984),and/or fuzzy logic (Zadeh et al., 1996) allow us to estimate a more suited continuous variable.

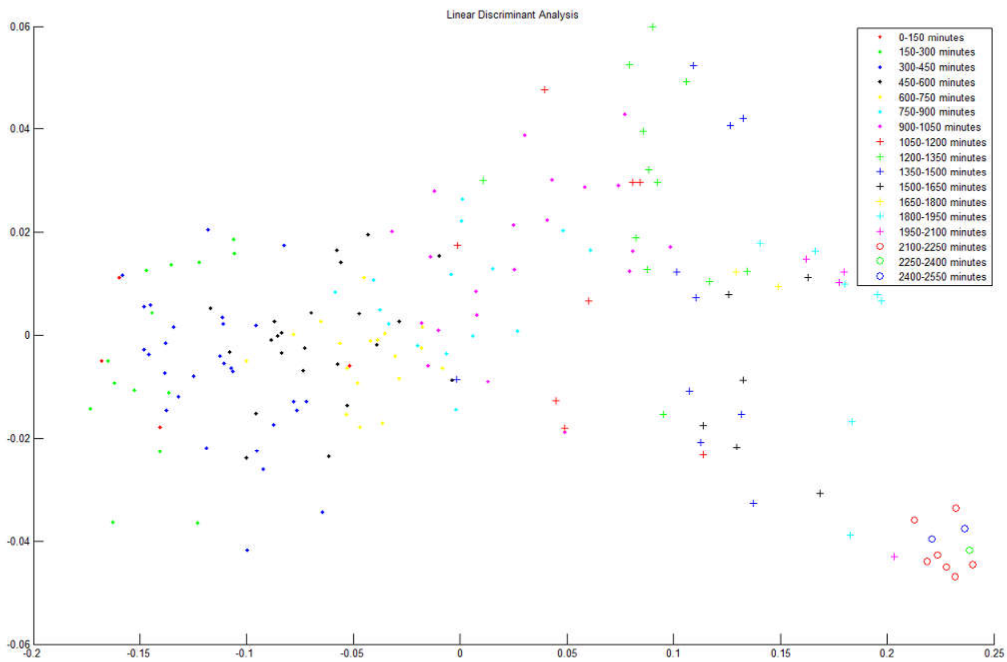

Figure 4: Linear Discriminant Analysis for 18 irrigation time intervals

In this section, we propose two different techniques, each belonging to one of the previous families, to estimate the weekly required amount of water. As described in the introduction and experimental sections, both modelling techniques require a supervised training set in order to learn the irrigation model.

2.3.1. Partial Least Square Regression

Partial Least Square Regression (PLSR) (Wold et al., 2001) is a statistical method that seeks the fundamental relations between predictor and response variables. Predictor variables, X, are defined as the observable variables that can be measured and input into the decision system. Response variable Y are the outputs or estimates that must be deducted from the input.

The relationship between both variable sets, and linear multivariate regression model, is found by projecting both predicted and observable variables into a new space, where latent variables are estimated to model the covariance structure between the predictor space and the observation space.

This PLSR model is developed from a training set $D=\{X, Y\}$ of S samples, which is composed of the predictor matrix $X=[x_1, ..x_i,..x_S]^T$ and the response matrix $Y=[y_1, ..y_i,..y_S]^T$. x_i is a column vector of K elements, that can contain all the sensor and weather variables measured at a given week i:

$$xi=[VWC1,VWC2,VWC3,MP,ST,ET_c,RF,WS,T,RH,GR,DP,VPD]^T \qquad [2]$$

and y_i is another column vector of M elements, containing the corresponding variables to be estimated at that week i. Since in our application this is only the irrigation time recommended at that week, y_i is reduced to a scalar and $M=1$:

$$y_i=minutes\ of\ irrigation$$

PLSR constructs new predictor latent variables, known as components, which are linear combinations of the original predictor observable variables. These components are created to explain the observed variability in the original predictor variables, while simultaneously considering the response variable. That is, the estimated latent variables are linear combinations of predictor variables that have higher covariance with *Y*. Using the latent variables leads to a regression models able to fit the response variable with fewer components.

The PLSR learning model can be expressed as:

$$X = T \cdot P^T + E \qquad [3]$$

$$Y = U \cdot Q^T + G \qquad [4]$$

where *T* and *U* are the projections -aka scores- of *X* and *Y* into a smaller *L*-dimensional latent space respectively, *P* and *Q* are the orthogonal projection matrices -aka loading matrices- and *E* and *G* the error residuals. *P* and *Q* can be obtained by eigendecomposition of the original matrices.

Since the *X*-scores *T* are meant to be good predictors of *Y*, it can be approximated that:

$$Y = T \cdot Q^T + F \qquad [5]$$

Being *F* a new residual. This reduces the problem to find a set of weights *W* such that *T*=*X***W* predicts *X* and *Y* reasonably well. As mentioned, these orthogonal coefficients should maximise the correlation between *X* and *Y* while explaining the variance of *X*:

$$\max_w Corr^2(Y, X) \cdot Var(X\ ) \qquad [6]$$

P and *Q* can be solved by applying a Least Square Estimator(LSE) so:

$$Q^T = (T^T \cdot T)^{-1} \cdot T^T \cdot Y \qquad [7]$$

$$P^T = (T^T \cdot T)^{-1} \cdot T^T \cdot X \qquad [8]$$

Finally, by rewriting the previous equation, it can be derived that:

$$Y = T \cdot Q^T + F = X \cdot W \cdot Q^T + F = X \cdot B + F \quad [9]$$

Being B the PLSR regression coefficients. Once these coefficients have been learned, responses y* for new observation x* can be estimated by applying the learning model:

$$y^* = x^* \cdot B + f \quad [10]$$

assuming an estimation error *f*.

We favour the use of PLSR among other regression techniques due to its suitability when the number of predictors is bigger than the number of response variables, the responses are noisy and there is a high probability of having multicollinearity among the predictor variables. The multicollinear phenomenon happens when those variable are highly correlated, due to redundancy between sensors and or between meteorological factors. As it can be noticed, all these factors appear in our irrigation problem.

2.3.2. Adaptive Neuro Fuzzy Inference Systems

Adaptive Neuro Fuzzy Inference Systems (ANFIS) (Jang, 1993) is a fuzzy inference system for systematically generating fuzzy rules from a given input/output *D* dataset. This machine learning technique combines advantages from fuzzy logic and artificial neural networks. On the one hand, it allows us to represent an element not only into categories but also into a certain degree of membership functions, which allows mimicking the characteristics of human reasoning and decision making. On the one hand, it can be trained and so can self-improve in order to adjust the membership functions parameters directly from data (Wang et al., 2006).

The ANFIS architecture consists in a five-layer feedforward neural network (Figure 5) whose parameters are updated using a combination of gradient descent and LSE in a two-pass learning algorithm.

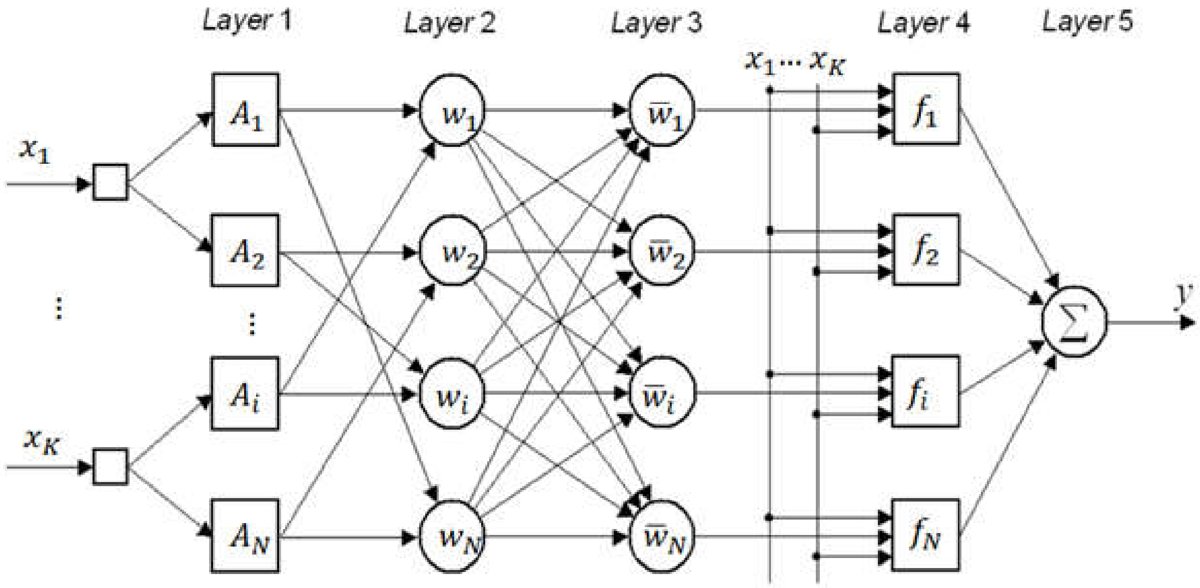

Figure 5: Example of ANFIS architecture for a input x with K variables and a 1-variable output y

In a first forward pass step, neuron outputs are calculated layer by layer and some internal consequent parameters are identified by the least squares estimator (LSE) to obtain the final single output. The forward pass operation at layer 1 defines the fuzzy membership for each input variable *X*. Assuming a Gaussian distribution function $N(c_n, \sigma_n)$, the output of this layer is given by:

$$O_{1,n} = \mu_{An}(x) = e^{-\frac{(x-c_n)^2}{2\sigma_n^2}} \quad [11]$$

Layer 2 is a multiplicative layer, which calculates the firing strength of the rules as a product of the previous membership grades.

$$O_{2,n} = \mathrm{w}_n = \prod_{k_n} \mu_{Ak_n}(x) \quad [12]$$

Layer 3 is a normalising layer, where:

$$O_{3,n} = \overline{\mathrm{w}}_n = \frac{\mathrm{w}_n}{\sum_j \mathrm{w}_j} \quad [13]$$

Layer 4 applies a node function:

$$O_{4,n} = \overline{\mathrm{w}}_n \cdot f_n = \overline{\mathrm{w}}_n \cdot \left(\sum_k p_k^n x_k + r^n\right) \quad [14]$$

where p^n and r^n are consequent parameters estimated using LSE

Finally, layer 5 is the output layer that provides the overall estimation *y* as a summation of all incoming signals. For the case *M=1*, where only one output variable is estimated:

$$O_{5,1} = \sum_n \overline{\mathrm{w}}_n \cdot f_n \quad [15]$$

After the forward pass has been completed, an initial estimation is provided by the ANFIS network. Since initial premise parameters $c_n, \sigma_n$are initialised randomly, the initial estimation will differ greatly from the desired values *Y*. This error or difference between the desired output *y* and the estimated output $O_{5,1}$ for a given training sample $\{x_i, y_i\}$ can be expressed as:

$$E_i = (y_i - O_{5,1})^2 \quad [16]$$

To correct this deviation, a second learning step, or backward pass, attempts to minimise the estimated error by modifying the value of the premise parameters until the desired and estimated outputs are similar. This process is performed using backpropagation, where the error is propagated back over the layers and decomposed into the different nodes using the chain rule. Gradient descend is used as optimisation technique to update the premise parameters while the consequent parameters are kept fixed until the next iteration.

This double step learning process is repeated iteratively for every single sample in the training set until the estimated error is smaller than a given threshold, i.e. convergence is achieved, or a maximum number of iterations –epocs- are reached. The ANFIS implementation used in this work is taken from the Fuzzy logic toolbox (Inc, 2016), by Mathworks where the parameter Radii used to train was a scalar of value 0.75 and the average number of epochs used to train was 1500.

3 Experimental setup

The system was evaluated in three commercial plantations of lemon trees in the Region of Murcia, located in the semiarid zone of the South-East of Spain where the water is very scarce and drip irrigation is commonly used. The irrigation criteria followed was to maximize the yield.

Plantation 1. Fino lemon trees (*Citrus limon* L. Burm. fil cv. 49) on *C. macrophylla* Wester, growing in a soil with a low water retention capacity. The soil is characterized by a deep and homogeneous sandy - clay - loam texture. The irrigation water had an electrical conductivity (EC) of 2200 μS cm^{-1}. The orchard consist of 11 year old lemon trees with an average height of 3.5 m. Tree spacing was 7.0 m x 5.5 m, with an average ground coverage of about 47%. Two drip irrigation lines (0.8 m apart) were used for each tree row. There were 4 emitters (4 L h^{-1}) on both sides of each tree. One sensor node was installed in the 5.5 ha orchard, with a soil matric potential sensor (MPS-2, Decagon devices, Inc., Pullman, WA 99163 - USA) at a depth of 30 cm and three soil moisture sensors at a depth of 20, 40 and 80 cm (Enviroscan, Sentek Pty. Ltd., Adelaide, Australia) located 20 cm from a representative dripper and 2.25 m from the trunk.

According to the nearest weather station of SIAM, located about 5 km from the orchard, the climate was typically Mediterranean. Thus, over this period (2014), the annual rainfall for the area was 210 mm and ET_0 was 1395 mm. The average wind speed was 1.66 m/s, generally light wind and sometimes moderate.

Plantation 2 and 3. 40 and 35 year old lemon trees (*Citrus limon* L. Burm. Fil) cv. Fino and cv. Verna respectively, grafted on sour orange (*Citrus aurantium* L.*)*, growing in a soil with a medium water retention capacity. The soil is clay sandy loam texture and the irrigation water had an electrical conductivity (EC) of 1600 $\mu S\ cm^{-1}$ during all season except in summer which was of 2285 $\mu S\ cm^{-1}$. The tree spacing was 7.0 m x 6.75 m and 6.75 m x 6.75 m and the average ground coverage about 57% and 50%, respectively. One drip irrigation line was used for each tree row. There were 8 and 6 emitters of 4 $L\ h^{-1}$ per tree, respectively. One sensor node was installed in the Fino orchard (≈15 ha) and another in Verna orchard (≈ 23 ha), each with two soil matric potential sensor (MPS-2, Decagon devices, Inc., Pullman, WA 99163 - USA) at a depth of 25 and 45 cm and three soil moisture sensors at a depth of 25, 45 and 70 cm (10HS, Decagon devices, Inc., Pullman, WA 99163) located 20 cm from a representative dripper and the vertical canopy projection.

According to the nearest SIAM's weather station, located about 7 km from the orchards, the climate was also typically Mediterranean. Over this period (2014), the annual rainfall for the area was 150 mm and ET_0 was 1250 mm. The average wind speed was 1.4 m s^{-1}, i.e. light wind generally.

The decision of selecting these three plantations is based on the fact that all of them are mature lemon trees and therefore their water irrigation requirement differences depend mainly of environmental conditions (soil and atmosphere) rather than the plant. Besides, all the plantations use drip emitters of 4 L h^{-1} so estimating the irrigation runtime of the week instead of the water volume will be a correct approach.

Drip irrigation provides a fixed volume of water per hour; the pressure is maintained using pressure compensating emitters. The Irrigation frequency is calculated taking into account that only a certain amount of water depletion is allowed before the next replenishment is scheduled. Thus, the run time (gross irrigation dose) is determined to be equivalent to the previous amount of water depletion. The experts only need to calculate the irrigation run time (minutes) and the number of watering times per week or day depending on the time of year or crop development stage. The main goal of the system, also reflected by the expert agronomist in his reports, is to maximize the yield (maximum production per crop surface) with an optimum water management.

Since information from the weather stations, soil sensors and crops characteristics has different sampling periods, the first step is pre-process this information. After analysing several methods and time intervals it was decided that the best option was to calculate the week average value for each of the sensors or weather stations variable except for the rainfall where the total amount of rainfall during the week is used instead. The week average fits better than others method like the daily average due to the fact, that the irrigation reports from the expert agronomist are already fixed, limited and done weekly. Besides, adding more input will make the data sparser, making more difficult to find patterns in the feature space, requiring a higher amount of data to train the system accordingly.

The input obtained will be a one dimensional vector x_i for each week in which the columns are the different variables or inputs of our system.

The target vector will be the water requirements of the crops in the following week y_i. This information has been extracted from the agronomist expert weekly reports in order to be used as groundtruth for comparison as for supervising the learning process.

Three datasets are available, each dataset represent a different plantation. Data was collected from January 2014 until June 2015. Each plantation dataset has 74 weeks of data, which makes a total of 224 weeks of data. To accomplish a proper analysis of the system, we have divided the experiment in two different scenarios. Both scenarios differ from the other on the training and testing split.

Two machine learning methods are applied on each scenario, a method based on PLSR and a method based on ANFIS. The performances of both methods in the different scenarios are analysed.

4 Experimental results and discussion

4.1 Scenario 1

In this scenario, we aim to successfully predict the irrigation needs of one or several plantation, based on the information provided by the collection device and learned knowledge from a historical archive of the previous year irrigation reports. This is of obvious usefulness in real life.

We will demonstrate this capability by predicting the irrigation needs of year 2015 for the three plantations based on the information of the year 2014. The training set is therefore composed by all 2014 weeks of data belonging the three plantations, while the test set is composed by all 2015 weeks belonging to the three plantations.

The information given to the system, or input vector, is a critical part of the design. On the one hand using unnecessary features may make the system perform poorly due to redundant information and noise. On the other hand, using too few features may not provide all the required information. Therefore, among all the available features explained in Table 1, they will not all be necessary. Table 3 shows the features subsets selected for each test. Among all possible sets of features, only combinations with logical sense, according to an expert agronomist were chosen a priori for the different experiments. Performance of the different sets is shown in Figure 6.

Feature set	Variables
F1	VWC1,VWC2,VWC3,SWP,ST,ET_c,RF
F2	VWC1, ET_c,RF
F3	SWP,ST,ET_c,RF
F4	SWP,ET_c,RF
F5	SWP,ST,ET_c
F6	VWC1,SWP,ET_c
F7	VWC1,SWP,ET_c,RF
F8	VWC1,SWP,ST,ET_c
F9	VWC1,VWC2,VWC3,SWP,ET_c
F10	VWC1,SWP
F11	VWC1,VWC2,VWC3,SWP
F12	SWP, ET_c

Table 3: Features subset and variables associated

Figure 6: Performance of the different sets of variables for Linear Regression and ANFIS

The set that accomplish the best performance for both methods is F6, with and error of 155.1 and 121.1 min week^{-1} for PLSR and ANFIS respectively. In order to put this error into context, it can be noticed that 2.5 extra hours of irrigation represent around 10% of the total time in summer months -and up to 20% in spring and autumn months-, being 10% error considered as an acceptable error in agriculture (Bos et al., 2004). Therefore, this feature set F6 will be the input vector of the system. It can be noticed that including the rain as input of the system (F7), increases the error. In the Region of Murcia, the rainfall are extremely low (around 210 mm per year) and usually being concentrated in a few days of the year, being the weekly total rain in most cases 0. With this information only available for the year 2014, the system didn't have enough information to be trained properly and developed in unpredictable results. However we understand that in other regions the rainfall could be really useful to increase the performance of the system. Besides, considering the water retention capabilities of the soil, part of the rainfalls would be considered in the next irrigation report.

Figure 7 shows the water irrigation pattern over time predicted by the PLSR and ANFIS respectively when using feature set 6.

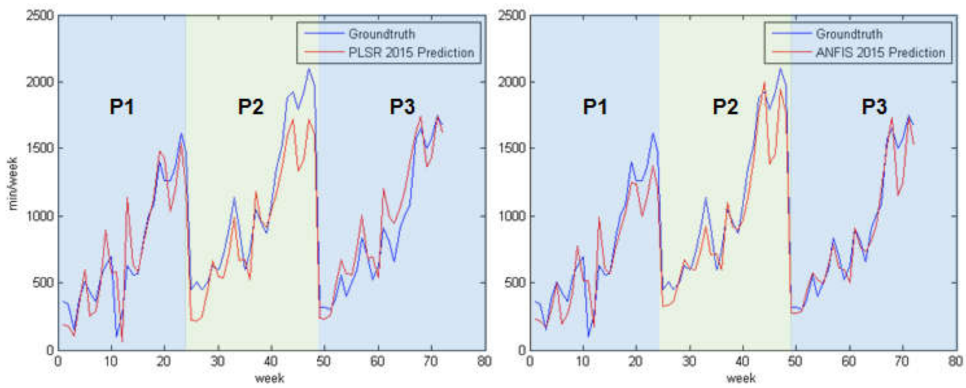

Figure 7: Prediction of the water irrigation pattern using soil and weather information for the different plantations (Plantation 1:Week 1-24, Plantation 2:Week25-48, Plantation 3: Week 49-72)

The weekly errors for predicting the irrigation needs during the year 2015 in the three plantations are 155.1 and 121.1 min week^{-1} for PLSR and ANFIS respectively. The standard deviation for PLSR is 120.7. In the case of ANFIS, the standard deviation is 105.2. The total amount of time needed to irrigate the crops in the three plantations in 2015 is 65,641 minutes. ANFIS method estimates this value in 60,506 minutes and PLSR estimates 63,240 minutes. As conclusion, ANFIS performance is better than PLSR for each individual week water requirement estimation. However, PLSR estimation also follows the irrigation pattern accurately and estimates the total amount of water required more accurately over time than ANFIS, which seems to be more conservative in the water usage. Looking at the higher peaks of water requirement in the graphs, PLSR may overestimate the water needs while ANFIS is more accurate in general. It is important to note that in agronomy the most important point is not only the amount of water plants need but when they need it (Allen et al., 1998). Following this criterion, the performance of ANFIS is much better than PLSR for this scenario.

Another factor that is important to analyse in this research is the use of soil sensors in addition to weather stations to close the loop. We consider that using this kind of sensors to estimate the water requirements of the crops improves the accuracy of the estimation and helps to deal with local disturbances. Since this is one of our main contributions and differences with other proposed automatic irrigation systems, a detailed analysis of the contribution of these variables is needed to validate our hypotheses and facilitate comparison with previous research systems. Therefore, the input vector was changed, using only weather information to train the system and predict the irrigation time. Table 4 shows the weekly average error for different sets of input vectors.

The weather-only input vector that performs best is produced using ET_0 exclusively, so this is used in the following analysis as representative of the weather-only prediction systems. Figure 8 shows the results of PLSR and ANFIS methods using the ET_c in comparison to the F6 system.

System	Input Vector	Weekly Error (minutes)	
		PLSR	ANFIS
Soil + weather variables (F6)	VWC1,SWP,ET_c	155.1	121.1
Only weather variables	ET_c	175.3	159.6
	ET_c,RF	178.4	163.6
	ET_c,RF,WS	378.4	379.5

Table 4: Summary of the performance of the different subsets

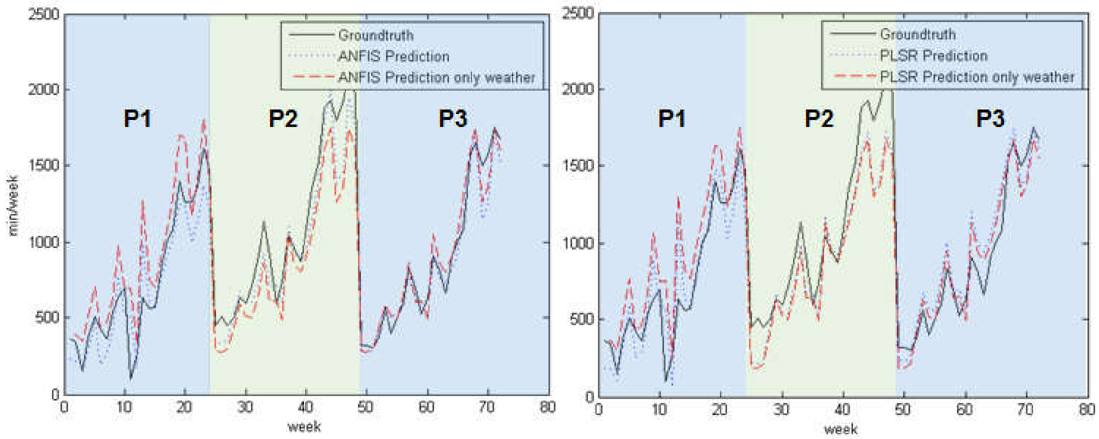

Figure 8: Prediction of the water irrigation pattern using weather information for the different plantations (Plantation 1:Week 1-24, Plantation 2:Week25-48, Plantation 3: Week 49-72)

The error in PLSR using only weather information is 175.3 minutes week^{-1} with a standard deviation of 147.6. In the case of ANFIS, the error is 159.6 minutes week^{-1} with a standard deviation of 146.6.

Although in general the shape of the graph is quite similar to the one using both soil and weather. The use of soil sensors gives a fine adjustment increasing the accuracy of the estimation for both PLSR and ANFIS reasoning engines.

It can be concluded that a much better performance in the weekly irrigation estimation (around a 22% smaller weekly average error) is achieved when adding soils sensor information to the weather information.

Next, a cross-validation strategy is applied to the scenario to validate how the results will generalise to an independent dataset. In cross validation, the complete dataset of the three plantations is divided in training and testing sets. The method used to cross-validate the information is Leave one out (LoO CV), a particular case of the Leave-p-out cross-validation (LpO CV). (Kohavi, 1995; Picard and Cook, 1984) that involves using 1 observation as the testing set and the remaining observations as the training set. This process is repeated the number of samples times (n) changing the test sample each time to validate the system with all the samples. Cross validation method was used for both PLSR and ANFIS.

Figure 9 shows the results of this LoO Cross-Validation method for PLSR and ANFIS respectively using the set F6 as input vector.

The error in PLSR is 277.8 minutes $week^{-1}$ with a standard deviation of 153.2. In the case of ANFIS, the error is 87.6 minutes $week^{-1}$ with a standard deviation of 102.9. The total amount of time needed to irrigate the crops for the 189 weeks in the three plantations is 214,020 minutes. The ANFIS method estimates this value on 213,180 minutes and PLSR estimates 213,960 minutes. Table 5 summarizes the result of the experiments.

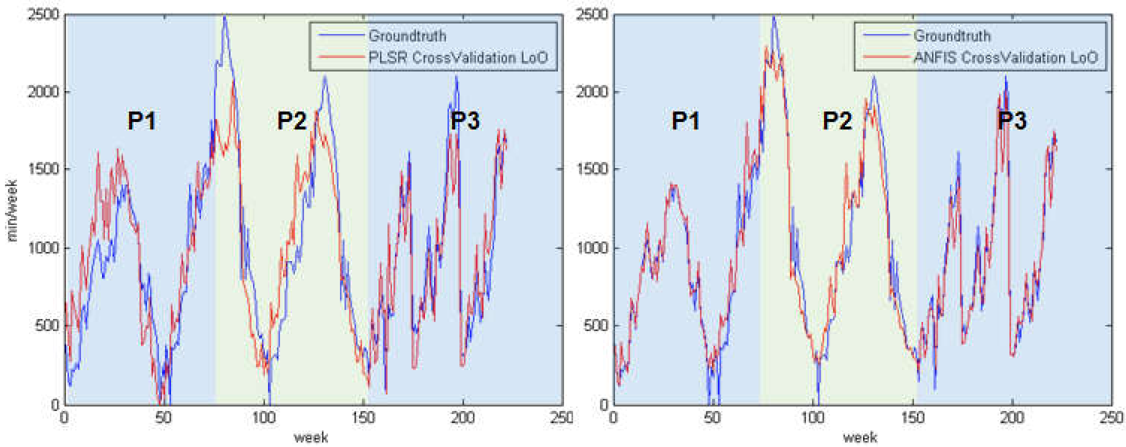

Figure 9: Cross-Validation LoO prediction for Linear Regression and ANFIS. (Plantation 1:Weeks 1-52 and 157-180, Plantation 2: Weeks 53-104 and 181-204, Plantation 3: Weeks 105-156 and 205-229)

RESULTS	Average weekly error (min.)	
	With soil sensors	No soil sensors

Scenario 1	Predict 2015	PLSR	155.1	175.3
		ANFIS	121.1	159.6
	Cross-Validation	PLSR	277.8	295.7
		ANFIS	87.6	211.9

Table 5: Scenario 1 results summary

Similar conclusions are extracted using Cross-Validation. Both PLSR and ANFIS systems are really close to the groundtruth in the total amount of water estimated but it is clear that ANFIS performs much better than PLSR if we consider the weekly error. It is also confirmed that using soil sensors in addition to weather information results in a better performance for both ANFIS and PLSR methods.

The improvement on ANFIS performance during cross validation is explained by the larger amount of training data regarding the “predict 2015” experiment. This behaviour is expected due to the nature of neural networks, which require large amount of data to be trained in comparison with other machine learning techniques and we predict than having a historical archive longer for training could results in a further improvement.

Although we are validating our systems with the three plantations described before as case of study, in principle, our methodology has been designed to be independent of the crop, terrain and location of the plantation, aiming to propose a general close-loop automatic irrigation estimator. In practical terms, this means that to apply our system to new plantations, training data in the form of sensor and weather weekly data as well as irrigation reports provided by and expert agronomist for the new plantation will be needed. Since these reports can be expensive and compiling a substantial amount of weekly reports is time consuming and must be planned in advance, it is important to know how big the dataset must be and how the performance may improve with the number of training weeks.

Therefore, as final experiment to obtain an estimation of the required amount of training data for a new crop/plantation, the complete dataset was divided in different percentages of training and testing. Figure 10 shows the weekly error of both PLSR and ANFIS methods with respect to the training dataset percentage.

According to the figure, it is noticeable that ANFIS performance is much better than PLSR if there are enough samples to train the system. In cases where the percentage of samples for training is low (less than 25% of the data, i.e less than 4 months of data for a given field), PLSR overperforms ANFIS. This case is relevant for new plantations without historical data of previous reports. In such situations, the PLSR predictive model may be used in early stages, before switching to ANFIS once enough samples to train the system properly are collected.

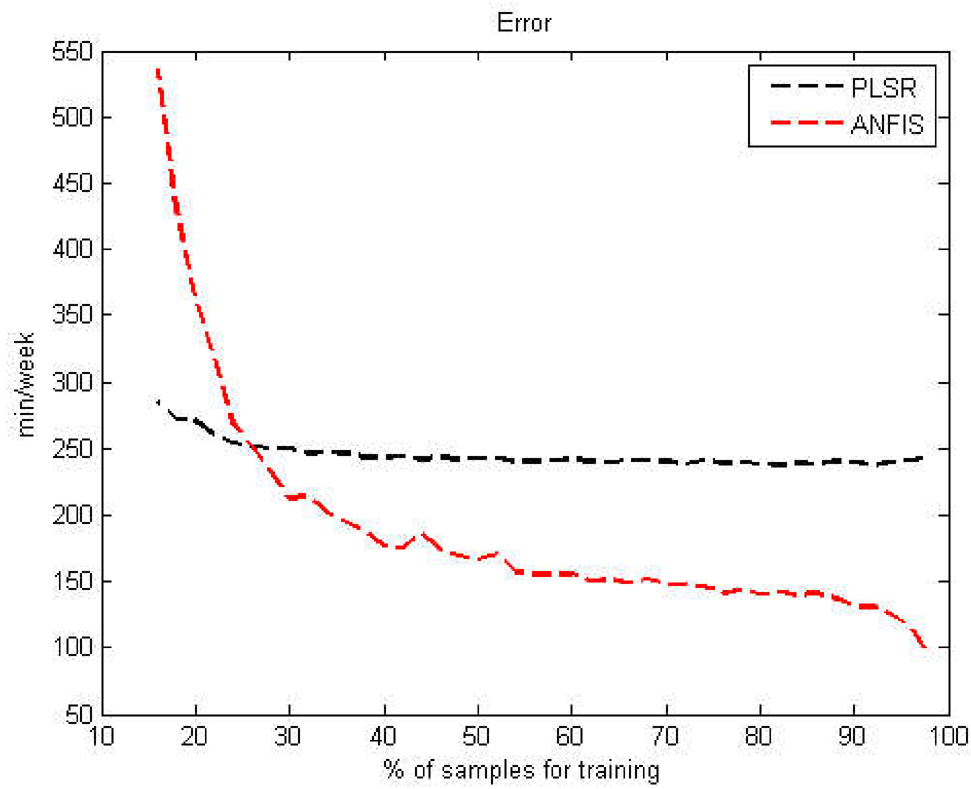

Figure 10: Performance comparison for Linear Regression and ANFIS with respect to the % of samples used to train

4.2 Scenario 2

The goal is to predict the irrigation of a plantation based on its weather and soil measured variables but using a SIDSS system trained exclusively with other fields. This will be the hardest scenario as it will be necessary to predict the irrigation needs of a field with no previous irrigation reports of that specific plantation. This scenario attempts to show the potential of our methodology to create a universal irrigation estimator of a given crop -in our case, lemon trees- for any given plantation, independently of the location and/or terrain. A lower performance can be expected in comparison to what could be achieved by retraining the system with information of the plantation (scenario 1), which is sacrificed for the benefit of not having to generate manual irrigation report for new plantations. Cross validation, specifically leave-one_plantation-out is applied in validation. Thus, 2014 and 2015 data from two of the plantations are used for training, while the remaining plantation data (2014+2015) is used for testing. This is repeated 3 times, leaving a different plantation out of the training set each time, and the results averaged.

Table 6 shows the error and standard deviation of this scenario for PLSR and ANFIS using different features vector used to compare the performance.

Method	Features vector	Test Plantation 1		Test Plantation 2		Test Plantation 3		Total	
		Weekly Error (min)	Std	Weekly Error (min)	Std	Weekly Error (min)	Std	Average Weekly Error (min)	Average Std
PLSR	VWC1+ SWP+ETc	364.1	205.6	179.4	141.2	227.5	185.8	257.0	177.5
ANFIS	VWC1+ SWP+ETc	373.2	300.7	175.4	129.8	421.4	495.5	323.3	308.6
PLSR	SWP+ETc	182.2	133.3	176.2	120.9	224.9	172.38	194.4	142.2
ANFIS	SWP+ETc	200.8	140.1	156.5	126.9	234.8	192.6	197.4	153.2

Table 6: Scenario 2 results summary

The best feature vector F6 used in scenario 1 is used as input. In this case PLSR outperforms ANFIS with an average error of 257.0 minutes in comparison with 323.3 minutes for ANFIS. However, we noticed that, in this scenario, removing the VWC1 sensor results in a better performance for both methods as a universal estimator. This is explained because the VWC sensor is very dependent on the soil where it is installed and, as both algorithms were trained with a sensor installed in a different plantation than the one that is predicting, the provided information introduces noise and does not help the system to estimate properly the water need. This does not happen, however, with the SWP sensor, which quantifies the tendency of water to move from one area to another in the soil and it is less dependent on the soil installed. Removing the VWC sensor results in a better performance of the system obtaining an average weekly error of 194.4 minutes with PLSR and 197.4 minutes with ANFIS. This result proves that there is certain potential to develop a universal estimator using our system for a given crop, although this means an increase of the average error. This error could be reduced if more than 2 plantations of the same crop were available for training. Both PLSR and ANFIS performs similarly, being PLSR slightly better.

5 Conclusions

This paper describes the design and development of an automatic decision support system to manage irrigation in agriculture. The main characteristic of system is the use of continuous soil measurements to complement climatic parameters to precisely predict the irrigation needs the crops, in contrast with previous works that are based only on weather variables or doesn't specify the quantity of water required by the crops. The use of real-time information from the soil parameters in a closed loop control scheme allows adapting the decision support system to local perturbations, avoiding the accumulative effect due to errors in consecutive weekly estimation, and/or detecting if the irrigation calculated for the SIDSS has been performed by the farmer. The analysis of the performance of the system is accomplished comparing the decisions taken by a human expert and the decision support component. Two machine learning techniques, PLSR and ANFIS, have been proposed as the basis of our reasoning engine and analysed in order to obtain the best performance.

The experiments have taken place in three commercial plantations of citrus trees located in the South-East of Spain. A first experimental scenario shows a comparison of the system´s performance using soil sensors in addition to the weather information for predicting year 2015 using 2014 information to train the system. The usage of soil sensor in the three plantations accomplished a 22% less of weekly error in comparison to the performance of using only weather information.

A second scenario shows the potential of our system as universal estimator for a given crop, i.e the use case of installing the system in a new plantation, not having previous information of it. For this application, VWC sensors should be removed due to their high dependence with the soil type. Although, as expected, the estimation error increases in this scenario, it does not require historical data from agronomical reports to be retrained, which implies a significant advantage, in particular for new plantations in early stages. If more training data from a bigger variety of field were available, a better performance in this scenario could be expected. Another possible improvement for this scenario will be the addition of a VWC to get a better performance than using only the matric potential sensors. However, in order to use the VWC sensor in this scenario, a precise study of the soil textures of the plantation will be required to extrapolate the VWC sensor information to similar soil textures where the DSS was trained.

For future research, we aim to extend and evaluate the system in plantations different than citrus and analyse the performance under several conditions and regions. Thus, adding the weather forecast as input of the SIDSS could help to improve the next week irrigation schedule and consider the predicted rainfall in our estimation. Similarly, past rainfall information, that did not prove beneficial in our system due to the region of Murcia characteristics, may become a good factor to improve the accuracy of the system in regions with a more regular and predictable raining pattern. We also aim to capture a bigger dataset that will allow us to generate more general models towards a universal irrigation estimator of a given crop. This dataset will also explore the use of multiple sensors per plantation in order to address inhomogeneous ground conditions in the different plantation as well as provide more input information to the system for a better reasoning.

Aknowledgments

The development of this work was supported by the Spanish Ministry of Science and Innovation through the projects MICINN, AGL2010-19201-C04-04 and MINECO, AGL2013-49047-C2-1-R. We would like to thank Widhoc Smart Solutions S.L. and Queen's Belfast University for letting us use their facilities and equipment to carry out the tests.

References

Abrisqueta, I., Conejero, W., Valdés-Vela, M., Vera, J., Ortuño, M.F., Ruiz-Sánchez, M.C., 2015. Stem water potential estimation of drip-irrigated early-maturing peach trees under Mediterranean conditions. Computers and Electronics in Agriculture 114, 7–13. doi:10.1016/j.compag.2015.03.004

Adeloye, A.J., Rustum, R., Kariyama, I.D., 2012. Neural computing modeling of the reference crop evapotranspiration. Environ. Model. Softw. 29, 61–73. doi:10.1016/j.envsoft.2011.10.012

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration: Guidelines for

Computing Crop Water Requirements. FAO Irrigation and Drainage Paper No. 56.

Belousov, A.I., Verzakov, S.A., von Frese, J., 2002. A flexible classification approach with optimal generalisation performance: support vector machines. Chemom. Intell. Lab. Syst. 64, 15–25. doi:10.1016/S0169-7439(02)00046-1

Bos, M.G., Burton, M.A., Molden, D.J. (Eds.), 2004. Irrigation and drainage performance assessment: practical guidelines. CABI, Wallingford.

Campos, I., Balbontín, C., González-Piqueras, J., González-Dugo, M.P., Neale, C.M.U., Calera, A., 2016. Combining a water balance model with evapotranspiration measurements to estimate total available soil water in irrigated and rainfed vineyards. Agricultural Water Management 165, 141 – 152. doi:http://dx.doi.org/10.1016/j.agwat.2015.11.018

Cardenas-Lailhacar, B., Dukes, M.D., 2010. Precision of soil moisture sensor irrigation controllers under field conditions. Agric. Water Manag. 97, 666–672. doi:10.1016/j.agwat.2009.12.009

Dutta, R., Morshed, A., Aryal, J., D'Este, C., Das, A., 2014. Development of an intelligent environmental knowledge system for sustainable agricultural decision support. Environ. Model. Softw. 52, 264–272. doi:10.1016/j.envsoft.2013.10.004

Fisher, R.A., 1938. The Statistical Utilization of Multiple Measurements. Ann. Eugen. 8, 376–386. doi:10.1111/j.1469-1809.1938.tb02189.x

Giusti, E., Marsili-Libelli, S., 2015. A Fuzzy Decision Support System for irrigation and water conservation in agriculture. Environ. Model. Softw. 63, 73–86. doi:10.1016/j.envsoft.2014.09.020

Guariso, G., Rinaldi, S., Soncini-Sessa, R., 1985. Decision support systems for water management: The Lake Como case study. Eur. J. Oper. Res. 21, 295–306. doi:10.1016/0377-2217(85)90150-X

IDAE, 2005. Ahorro y Eficiencia Energética en Agricultura de Regadío. Madrid.

Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. 23, 665–685. doi:10.1109/21.256541

Jensen, M.E., Robb, D.C.N., Franzoy, C.E., 1970. Scheduling Irrigations Using Climate-Crop-Soil Data. J. Irrig. Drain. Div. 96, 25–38.

Jones, H.G., 2004. Irrigation scheduling: advantages and pitfalls of plant-based methods. J. Exp. Bot. 55, 2427–2436. doi:10.1093/jxb/erh213

King, B.A., Shellie, K.C., 2016. Evaluation of neural network modeling to predict non-water-stressed leaf temperature in wine grape for calculation of crop water stress index. Agricultural Water Management 167, 38 – 52. doi:http://dx.doi.org/10.1016/j.agwat.2015.12.009

Kohavi, R., 1995. A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection, in: Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI’95. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 1137–1143.

Kramer, P.J., Boyer, J.S., 1995. Water relations of plants and soils. Academic Press, Inc.

López Riquelme, J.A., Soto, F., Suardíaz, J., Sánchez, P., Iborra, A., Vera, J.A., 2009. Wireless Sensor Networks for precision horticulture in Southern Spain. Computers and Electronics in Agriculture 68, 25–35. doi:10.1016/j.compag.2009.04.006

Maton, L., Leenhardt, D., Goulard, M., Bergez, J.-E., 2005. Assessing the irrigation strategies over a wide geographical area from structural data about farming systems. Agric. Syst. 86, 293–311. doi:10.1016/j.agsy.2004.09.010

Naor, A., Hupert, H., Greenblat, Y., Peres, M., Kaufman, A., Klein, I., 2001. The Response of Nectarine Fruit Size and Midday Stem Water Potential to Irrigation Level in Stage III and Crop Load. J. Am. Soc. Hortic. Sci. 126, 140–143.

Navarro-Hellín, H., Torres-Sánchez, R., Soto-Valles, F., Albaladejo-Pérez, C., López-Riquelme, J.A., Domingo-Miguel, R., 2015. A wireless sensors architecture for efficient irrigation water management. Agric. Water Manag., New proposals in the automation and remote control of water management in agriculture: agromotic systems 151, 64–74. doi:10.1016/j.agwat.2014.10.022

Picard, R.R., Cook, R.D., 1984. Cross-Validation of Regression Models. Journal of the American Statistical Association 79, 575–583. doi:10.2307/2288403

Puerto, P., Domingo, R., Torres, R., Pérez-Pastor, A., García-Riquelme, M., 2013. Remote management of deficit irrigation in almond trees based on maximum daily trunk shrinkage. Water relations and yield. Agric. Water Manag. 126, 33–45. doi:10.1016/j.agwat.2013.04.013

SIAM, 2015. Red del Sistema de Información Agrario de Murcia. URL siam.imida.es

Smith, M., 2000. The application of climatic data for planning and management of sustainable rainfed and irrigated crop production. Agric. For. Meteorol. 103, 99–108. doi:10.1016/S0168-

1923(00)00121-0

Soulis, K.X., Elmaloglou, S., Dercas, N., 2015. Investigating the effects of soil moisture sensors positioning and accuracy on soil moisture based drip irrigation scheduling systems. Agric. Water Manag. 148, 258–268. doi:10.1016/j.agwat.2014.10.015

Sreekanth, M.S., Rajesh, R., Satheeshku, J., 2015. Extreme Learning Machine for the Classification of Rainfall and Thunderstorm. Journal of Applied Sciences 15, 153–156. doi:10.3923/jas.2015.153.156

Valdés-Vela, M., Abrisqueta, I., Conejero, W., Vera, J., Ruiz-Sánchez, M.C., 2015. Soft computing applied to stem water potential estimation: A fuzzy rule based approach. Comput. Electron. Agric. 115, 150–160. doi:10.1016/j.compag.2015.05.019

Wang, Z., Palade, V., Xu, Y., 2006. Neuro-Fuzzy Ensemble Approach for Microarray Cancer Gene Expression Data Analysis, in: 2006 International Symposium on Evolving Fuzzy Systems. Presented at the 2006 International Symposium on Evolving Fuzzy Systems, pp. 241–246. doi:10.1109/ISEFS.2006.251144

Wold, S., Ruhe, A., Wold, H., Dunn, I., W., 1984. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. SIAM J. Sci. Stat. Comput. 5, 735–743. doi:10.1137/0905052

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst., PLS Methods 58, 109–130. doi:10.1016/S0169-7439(01)00155-1

Zadeh, L.A., Klir, G.J., Yuan, B., 1996. Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A Zadeh, Advances in Fuzzy Systems — Applications and Theory. WORLD

SCIENTIFIC.

Zwart, S.J., Bastiaanssen, W.G.M., 2004. Review of measured crop water productivity values for irrigated wheat, rice, cotton and maize. Agric. Water Manag. 69, 115–133. doi:10.1016/j.agwat.2004.04.007