Journal of Computational Science 59 (2022) 101547

Contents lists available at ScienceDirect

Journal of Computational Science

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

Humidity forecasting in a potato plantation using time-series neural models

Mercedes Yartu [a], Carlos Cambra [b,*], Milagros Navarro [a], Carlos Rad [a], Ángel Arroyo [b], Álvaro Herrero [b]

[a] Composting Research Group (UBUCOMP), EPS-La Milanera, Universidad de Burgos, C/ Villadiego s/n, 09001, Burgos, Spain

[b] Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Ingeniería Informática, Escuela Politécnica Superior, Universidad de Burgos, Av. Cantabria s/n, 09006, Burgos, Spain

ARTICLE INFO

Keywords:
Precision irrigation
Potato crop
Time series forecast
Supervised learning
Neural networks
Interpolation

ABSTRACT

It is widely acknowledged that, under the frame of sustainable farming, using the minimum water resources is a relevant requirement. In order to do that, precision irrigation aims at identifying the irrigation needs of plantations and irrigate accordingly. Artificial intelligence is a promising solution in this field as intelligent models are able to learn the soil moisture dynamics in the soil-plant-atmosphere system and then generating appropriate irrigation scheduling. This is a complex task as the phenology of plants and its water demand vary with soil properties and weather conditions. The present research contributes to this challenging task by proposing the application of neural networks in order to learn the time-series evolution of irrigation needs associated to a potato plantation. Several of such models are thoroughly compared, together with different interpolation methods, in order to find the best combination for accurately forecasting water needs. In order to predict the soil water content in a potato field crop, in which soil humidity probes were installed at 15, 30, and 45 cm depth during the whole cycle of a potato crop. This innovative study and its promising results provide with significant contributions to address the problem of predicting and managing groundwater for agricultural use in a sustainable way.

1. Introduction

Potatoes (*Solanum tuberosum*), belonging to the solanaceae family of flowering plants, were originated, and first domesticated in the Andes mountains (South America). In terms of agricultural production, potato crop is the third most important food crop in the world after rice and wheat. The EU produced 51.8 million tonnes of potatoes in 2018, with Germany, France, Poland, and Netherlands as main producers [1]. In Spain, potato production reaches 2.24 million tonnes, and 40.3 % of it is located in Castilla y León, mainly in Burgos (4 %) occupying in 2017 around 2400 ha of irrigated land [2].

In the Mediterranean context, irrigation supposes an extraordinary demand for available water, which constitutes an important problem in a context of water scarcity and climatic change. The application of innovative and an appropriate transfer of technologies to an adequate management of irrigation is a key factor to reach a sustainable crop production [3].

Monitoring weather variables and water status in soils are key factors to reach minimum water consumption without compromising crop production. The use of satellite or Unmanned Aerial Vehicles (UAVs) imaginaries, automated weather stations and humidity or water potential probes are important tools to achieve precision irrigation adapted to crop phenology [5]. These issues maximize production avoiding water stresses, lixiviation of nutrients, or the incidence of crop pest and diseases.

In order to contribute to this field, the present paper proposes the application of IoT and Artificial Intelligence (AI) to monitor a potato field crop, located in Cabia (Burgos, Spain), 42°16′57″ N and 3°51′25″W, with sprinkler irrigation, to optimize water use efficiency. More precisely, a meteorological station, together with different sensors were placed in the crop in order to gather data in real time. Additionally, some measurements regarding crop development were taken and are analysed in the present work. As there is no sensor to measure such features, these measurements must be taken manually. Also, imaginary figures are not available on a daily basis and hence they must be interpolated in order to merge such data with those gathered through

* Corresponding author.
E-mail addresses: ccbaseca@ubu.es (C. Cambra), minago@ubu.es (M. Navarro), crad@ubu.es (C. Rad), aarroyop@ubu.es (Á. Arroyo), ahcosio@ubu.es (Á. Herrero).

https://doi.org/10.1016/j.jocs.2021.101547
Received 26 April 2021; Received in revised form 16 October 2021; Accepted 22 December 2021
Available online 1 January 2022
1877-7503/

IoT.

Available soil water content is the key factor that assures crop performance and reduced the risk of crop diseases. Irrigation is for example, a valuable component in the control of some soil-borne pathogens such as *Streptomyces scabies*, the cause of potato common scab [6]. Forecasting soil water content would be an important tool for precision irrigation and also, programming an adequate fertigation of the crop, avoiding detrimental consequences of leaching or pest and diseases incidence. By considering all field collected data, neural networks for time series forecast are applied in order to predict water needs of the crop. With all of these data, models are trained to generate a prediction of the volumetric soil moisture content based on soil moisture, precipitation, and climatic measurements. The models studied in this work could be applied to different intensively irrigated crops such as sugar beet (Beta vulgaris), corn (*Zea mays*) or alfalfa (*Medicago sativa*) in which automated meteorological stations, soil humidity or water potential probes and satellite imaginary are increasingly introduced.

The remaining sections of this study are structured as follows while the methods applied in present study are described in Section 3. Section 4 introduce the real-life problem that is addressed, while the obtained results are presented in Section 5. Finally, both conclusions and future work proposals are discussed in Section 6.

2. Previous work

Thanks to the experimental validation, irrigation needs of the studied plantation could be adjusted. As stated by Shitu et al. [7], different AI approaches and methods have been studied for smart controlling irrigation systems. More precisely, Neural Networks, Genetic Algorithms, and Fuzzy Logic could lead to optimum utilization of irrigation water resources.

Labbé et al. [8] modelled an irrigation decision process for limited water allocation, a very common pattern and challenge caused by climate change [9], and irrigation scheduling for corn plantations. The model consisted of irrigation management rules for different irrigation-related tasks that were derived from farmer surveys and based on the monitoring of their irrigation practices over a 2-year period. This model was incorporated into a simulator engine that, given the context of the decision, was able to predict irrigation schedules and irrigation volumes with an average error ranging from 6 to 13 mm for different farmers, reflecting an error below 6.7 %. Instead of developing a model that captures the farmer's decision individually, using surveys and observations, in this study the Deep Learning and AI were used to capture the agronomist's decision process in irrigation system [8].

Meanwhile, Atsalakis et al. [10] proposed a daily irrigation water demand calculation based on an Adaptive Neuro Fuzzy Inference System (ANFIS). This first-order Sugeno fuzzy model is combined with a back propagation algorithm. It has a better performance (Root Mean Squared Error (MSE) and Mean Absolute Percentage Error) predicting irrigation needs when compared to the Auto Regressive Moving Average models.

The irrigation amount and timing are based on measurements of soil, plant, and climatic variables from which the plant water need is inferred. Working in precision irrigation aims to accurately determine and quantify plant water needs [11].

Khan et al. [12] compared different AI models and their error rates when it comes to irrigation prediction. It was found that among all the models, the 3-fold cross validation multiple decision trees SysFor model gave the best overall results. However, the actual amount of water required by the crop was accurately predicted by neural models. The difference in error percentage between Artificial Neural Networks (ANNs) and SysFor was almost 20 %. The comparison concluded that SysFor, ANNs, and decision tree techniques are the most suitable ones for the task of irrigation prediction.

Adeyemi et al. [13] use a Dynamic Neural Network approach for modelling of the temporal soil moisture fluxes where he obtained results that indicate that the predictive system achieves a water saving ranging between 20 and 46 %, while realizing a yield and water use efficiency similar to that of the rule-based system.

Wunsch et al. [14] applied time-series ANN to obtain groundwater-level forecasts for several wells in three different types of aquifers, namely porous, fractured and karst aquifers in south-west Germany, using precipitation and temperature as input parameters.

Similarly, non-linear time-series neural networks have been previously applied to some different problems ranging from workplace accidents [15] to road transportation [16] and fault detection [17]. Köppen et al. [18] analysed the use of deep learning techniques to analyse data collected from IoT sensors in order to support watering decisions. Differentiating from previous works, present paper proposes time-series neural networks in order to predict the humidity in the under-ground. Table 1 shows a summary of the works previously mentioned.

As a preliminary work in the present research line, different interpolation methods and time-series neural models [4] were combined and benchmarked to predict the soil humidity level in a potato field with only 1 soil moisture probe (15 cm depth). Only meteorological data and the soil moisture humidity itself were used as input data and no information from the crop was managed by the models. The present work extends such research and goes further; neural networks are applied for modelling the temporal (surface) soil moisture fluxes. In order to do that, the records of 3 different soil humidity probes (installed at 15, 30, and 45 cm depth during the whole cycle of a potato crop) are analysed. Neural and interpolation models, as well as the best tuning of them, are comprehensively benchmarked. In general, the case study results are satisfactory and show that neural networks can be a useful tool for

Table 1
Comparison of related works.

Publication	Applied Technology	Research Focus	Results
Shitu [7]	Smart controller and sensors data acquisition were studied as well as some mathematical relation	Review of techniques. Techniques on irrigation water management.	Artificial Neural Network ANN is the most widely used in ET estimation to optimize water utilization
Labbé [8]	Rule Engine (IRMA software)	Analysis of daily simulated water demand	The absolute result of error concerning volumes is smaller than the water applied in a single irrigation position is less than 8.5 %
Atsalakis [10]	Adaptive Neuro-Fuzzy Inferences System (ANFIS)	Prediction results of daily irrigation water demand.	Using ANFIS, they obtained the minimum MSE
Khan [12]	Decision Trees (DT), Artificial Neural Networks (ANNs), Systematically Developed Forest (SysFor) for multiple trees, Support Vector Machine (SVM), Logistic Regression and Evapotranspiration (ETc)	Comparison of water usage predicted by different models to actual water usage	SysFor produce the best prediction
Wunsch [12]	Non Linear Autoregressive Networks with Exogenous Inputs (NARX)	Forecasting groundwater levels	Not provide high resolution in accuracy levels of groundwater
Köeppen [18]	Long Short Term Memory Networks (LSTMs) approach in deep learning techniques	IoT Data Collect from soil and plants and Cloud Computing	Improve the prediction accuracy for the watering system over various soil and plants

predicting humidity. As a result, the best combinations of them are identified in order to accurately forecast the humidity level in the underground so that water demands could be more precisely calculated.

3. Applied methods

As previously stated, two kinds of methods have been applied in present paper; on the one hand, interpolation (described in Subsection 3.1) has been applied to predict daily values of some features. On the other hand, neural networks (described in Sub-section 3.2) have been applied to predict the humidity level.

Interpolation has been chosen as the basic and reference technique to compare results and neural models as more advanced models with which to further refine the predictions.

3.1. Interpolation

It is widely known that interpolation consists on generating new data points between a given range of values. In order to do that, several alternatives exist for one-dimensional problems. The following ones have been applied in present study:

Cubic: this is a shape-preserving method for cubic interpolation. Based on the shape of the known data, new values are interpolated by piecewise cubic interpolation, taking into account the values at neighboring grid points.

Spline: each new value calculated by this method is based on a cubic interpolation of the values at neighboring data in each respective dimension. The not-a-knot end conditions are applied.

Makima: this modified version of the Akima cubic Hermite interpolation method [19]. Each new value calculated by this method is based on a piecewise function of polynomials (with degree smaller than or equal to 3). In the Akima formula, the value of the derivative at a given data point is a weighted average of nearby slopes. The weights are defined as:

$$w_1 = |\delta_{i+1} - \delta_i|;\ w_2 = |\delta_{i-1} - \delta_{i-2}| \tag{1}$$

Being δ_i the slope on the interval $[x_i,\ x_{(i+1)}]$. In the modified version, definition of weights is slightly different, as follows:

$$w_1 = |\delta_{i+1} - \delta_i| + \frac{|\delta_{i+1} + \delta_i|}{2};\ w_2 = |\delta_{i-1} - \delta_{i-2}| + \frac{|\delta_{i-1} + \delta_{i-2}|}{2} \tag{2}$$

Thanks to that, when two flat regions with different slopes meet, more importance is given to the side where the slope is closer to zero (horizontal), thus avoiding overshoot.

3.2. Neural models

In order to predict the humidity level, once all data are available (i.e. after interpolation is carried out), 3 neural models for non-linear time-series forecast [20] have been applied, namely: Non-linear Input-Output (NIO), Non-linear Autoregressive (NAR) and Non-linear Autoregressive with Exogenous Input (NARX). These can be seen as feed forward networks [21] in which the input weight has a tap delay line associated with it. Thanks to that, the network has a finite dynamic response to time series input data. The main differences between these 3 models are what data is given to the model in order to predict future values of humidity level. In the case of NIO, it is only the humidity level itself. In the case of NAR, all the other features (described in Section 4) except the humidity level are included. Finally, in the case of NARX, these two data sources are considered in the prediction. As a result, the NARX could be mathematically formulated as:

$$y(t) = f\big(y(t-1), \ldots, y(t-n_y),\ x(t-1), \ldots, x(t-n_x)\big) \tag{3}$$

Being y(t) the variable to be predicted in time instant *t*, *f()* the function to be approximated by the neural model, x(t) an exogenous variable, n_y the maximum number of time delays in the output, and n_x the maximum number of time delays in the input. Consequently, the mathematical formulation for the NAR model is:

$$Y(t) = f(y(t-1), \ldots, y(t-n\ y) \tag{4}$$

As it can be seen, in the case of the NAR model, the exogenous input (x) is not included in the formulation. Differentiating from this model, the predicted variable is replaced by this exogenous one in the NIO formulation:

$$y(t) = f(\ x(t-1), \ldots, x(t-n_x)\ ) \tag{5}$$

To justify better the selection of the two method that are proposed to predict the humidity level diameter range regarding groundwater level.

4. Agronomic setup

Field experiments were conducted from April 16th to October 10th 2019, in a potato field crop of 5 ha, located in the small village of Cabia (Burgos, Spain), (42°16′57″ N, 3°51′25″ W), with a semi-permanent sprinkler irrigation system. Soil was classified as Calcic Luvisol (*LVk*) according to FAO, with loam texture, bulk density 1.26 kg L^{-1}, field capacity 0.31 (w/w), pH (1:5 w/v) 7.6, Electrical Conductivity (1:5 w/v, 25 °C) 0.65 dS m^{-1}, Organic Mater 3.33 %, Total N 0.16 % and lime 16.7 %. Climate in this area is Attenuated Mesomediterranean, according to FAO.

As shown in Fig. 1, an agronomic IoT system was installed in the field, comprising the automatic weather station ATMOS 41 (METTER Group, USA) oriented to North. A soil humidity probe TEROS 10 (METTER) was installed at 15 cm depth, a soil water potential probe TEROS 21 at 30 cm depth and a rain gauge (ECRN 100) were connected to a EM60 G data logger, remotely connected with ZENTRA Cloud System (METER Group, USA) that registered data each 30 min.

Potatoes (*Solanum tuberosum* L. cv. Agria) were planted from April 16th to mid-June, phenological development was assessed according to BBCH-scale and four plants from the center of the plot (20 × 20 m) were removed for laboratory analysis every 15 days. Morphological parameters such as length of aerial plant, number of stems and leaves, length of roots, number and weight of tubers, wet and dry biomass, chlorophyll content with SPAD, and N-content by a combustion autoanalyzer (TruSpec, LECO) were determined. Before harvesting, four sampling locations of 3 m^2 were chosen at random for yield estimation; tubers were classified by considering their diameter in different commercial classes: >80 mm, between 40−80 mm and <40 mm.

Public imaginary was captured from the satellite SENTINEL-2B under the scope of the EU Copernicus program. Nine images were obtained corresponding to day 11th to 171st, after plant emergency. From

Fig. 1. Field map of the agronomic IoT system.

them, Normalized Difference Vegetation Index (NDVI) was calculated according to the equation:

$$NDVI = ((NIR\text{-}Red))/((NIR + Red)) \tag{6}$$

Where Red and NIR are the spectral reflectance measurements acquired in the red (visible) and near-infrared regions, respectively. These data correspond to 4 and 8 of SENTINEL-2B bands, respectively. Raster layers were processed using the software QGIS v. 2.18 to obtain an NVDI vector layer. NVDI data were thereafter transformed into basal crop coefficients (K_{cb}) using equation:

$$K_{cb} = 1.44 \times NDVI - 0.1 \tag{7}$$

Crop evapotranspiration was calculated according to FAO Method 56 approach:

〖ET〗 _c = (Kcb + Ks)× 〖ET〗 o

$$ET_c = (K_{cb} + K_s) \times ET_o \tag{8}$$

Where K_s estimates soil evaporation, which is considered zero during the irrigation period as the crop development quickly cover soil surface. The following features are available to apply the neural networks:

1 Air temperature: gathered from the temperature sensor (-40 – 50 °C) in the ATMOS 41 Weather Station (Meter Group, USA), Accuracy +/- 0.5 °C.
2 Precipitation: gathered from the precipitation sensor (0–400 mm/h) in the ATMOS 41 Weather Station (Meter Group, USA), Accuracy +/- 5 %. Daily
3 CCM (Chlorophyl Content Index): CCM-200 plus Chlorophyll Content Meter (OptiSciences, UK) measures optical absorbance in two different wavelengths: 653 nm (Chlorophyll) & 931 nm (Near Infra-Red).
4 Plant height[1] : a Carpenters meter (+/- 1 mm) was used.
5 Plant weight[1]: a weight scale (+/- 1 mg) was used.
6 % Humidity[1]: weight losses after 38 h at 70 °C (+/- 1 °C).
7 Aerial part length[1]: a Ruler lab (+/- 1 mm) was used.
8 Roots length[1]: a Ruler lab (+/- 1 mm) was used.
9 Plant Nitrogen content[1]: aerial part of plants was dried at 70 °C and thereafter, ground in a mill. Samples of 0.2 g were analysed by Dumas method in a TruSpec CN (LECO, USA) with IRD (Infra-Red Detector) and TCD (Thermal Conductivity Detector) for CO_2 and N_2, respectively.
10 Tubers weight per plant[1]: a weight scale (+/- 1 mg) was used.
11 Number of tubers per plant[1]: tubers were visually counted.
12 Tubers humidity[1]: weight losses after 38 h at 70 °C (+/- 1 °C).
13 Percentage of tubers in the 0–40 mm diameter range**Error! Bookmark not defined.**: a squared measurement frame of 40 mm was used.
14 Percentage of tubers in the 40–80 mm diameter range**Error! Bookmark not defined.**: squared measurement frames of 40 and 80 mm were used.
15 Percentage of tubers in the >80 cm diameter range**Error! Bookmark not defined.**: a squared measurement frame of 80 mm was used.
16 Tubers Nitrogen content**Error! Bookmark not defined.**: crushed fresh tubers were dried at 70 °C and thereafter, ground in a mill. Samples of 0.2 g were taken for total N analysis using TruSpec LECO.
17 Soil humidity probe: Teros 10 (Meter Group). It is a capacitance sensor that determines the dielectric permittivity of soil by measuring charge time of a capacitor, which uses that medium as a dielectric. The sensor measures the time to charge a capacitor from a starting voltage, Vi to a voltage Vf with an applied voltage, Vf. Its working frequency (70-MHz) minimizes salinity and textural effects in the soil. This is the data feature to be forecast in the range [0, 1]. Three sensors were installed at the beginning of the experience at 20, 40 and 60 cm connected to a data logger (ZL6, METER).

5. Experiments and results

The results obtained through the different experiments are described in the following subsections. These results have been obtained by applying the interpolation (Cubic, Makima, and Spline) and the neural (NAR, NIO, and NARX) models. During the experimental study of the soil humidity contents at three different soil depths, the analysed data comprises the period April to October 2019.

Each one of the neural models has been tuned with different values of the appropriate parameters:

- Number of input delays: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}
- Number of output delays: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}
- Number of hidden neurons: {1, 5, 10, 15, 20}
- Training algorithm: {1 - Levenberg-Marquardt (LM), 2 - Batch Gradient Descent, 3 - Gradient Descent with Momentum, 4 - Adaptive Learning Rate Backpropagation, 5 - Gradient Descent with Momentum and Adaptive Learning Rate, 6 - Scaled Conjugate Gradient, 7 - Broyden–Fletcher–Goldfarb–Shanno Backpropagation}

As a result, 350 runs have been performed for the NIO and NAR models, and 3.500 for the NARX model. For each one of them, 10 executions have been carried out in order to obtain more statistically significant conclusions. Average Mean Squared Error (MSE) is provided in each case, calculated as the average MSE of all the included runs and executions. This metric is used to compare the models and configurations, in order to select the best-performing one.

5.1. Results for soil probe 1

This subsection presents the experimental results for the Probe 1 (at 15 cm depth). Results are represented according to the coefficient of determination and MSE as quantitative statistical measures.

Fig. 2 shows the results (average MSE) obtained by the Cubic, Makima, and Spline interpolations when applying the NAR, NIO and NARX models. The Cubic-NARX MSE value (5.1297) is out of range when compared to the other results. So, in order to keep the plot at an appropriate scale, this value has been truncated to 0.

It is worth highlighting from Fig. 2 the great variability in the results when applying the different methods to each one of the components.

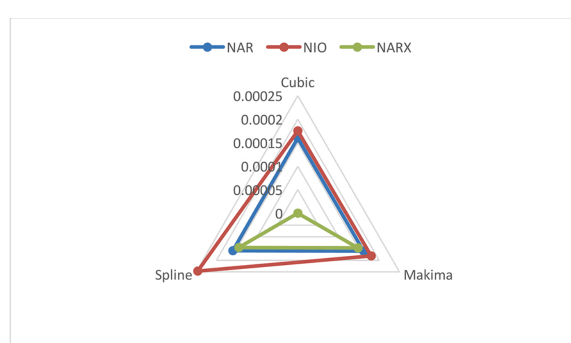

Fig. 2. Radar plot of the MSE average values on the soil moisture probe 1 per interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO, and NARX) model.

[1] Interpolated by means of the methods described in Subsection 2.1. All features are interpolated by means of same method each time.

This is especially important in the case of the NARX model. An extremely high error rate (truncated in Fig. 2) is obtained by this model when applying the cubic interpolation. However, when combined with the Spline method, the lowest MSE rate (0.0001458) is obtained. On the contrary, the lowest average MSE rates for this probe are obtained with Makima interpolation. For the rest of neural models, error rates are pretty similar, slightly varying between the different methods.

In addition to the previous results (average MSE), the lowest values or the error are shown in Fig. 3. They are presented in a similar way, by the interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO and NARX) models. As it happened for the average MSE, the Cubic-NARX lowest MSE value (0.00429) is out of range when compared to the other results, and it has been truncated to 0 in the figure.

When analysing these results, it can be said that the trends are very similar to those in the average results (Fig. 2); there are big differences regarding the NARX model that obtained both the highest and lowest (when combined with the Spline interpolation) MSE values. Additionally, the Makima interpolation method obtained the lowest average (for the 3 neural models) MSE rates for this probe.

As previously mentioned, the lowest MSE value for Probe 1 (0.0001458) has been obtained when combining the Spline interpolation together with the NARX neural model. In order to study the effect of the parameter values for such combination, Fig. 4 presents the average values obtained with these two methods when using different values for the NARX parameters (number of input delays, number of hidden neurons, and training algorithms).

From Fig. 4, it can be highlighted that, regarding the number of input delays, the more input delays, the lower MSE. As it can be seen, the two lowest MSE values are obtained with 9 and 10 input delays, being 9 the lowest one. That is, a wide temporal window lets the NARX model to accurately predict the target value. An opposite trend can be observed regarding the number of hidden neurons; the lowest MSE value is obtained with only 1 hidden layer. Finally, the best result has been obtained by NARX when being trained with the LM (number 1 in Fig. 4) algorithm. Worst results are associated to the Batch Gradient Descent and Gradient Descent with Momentum training algorithms (2 and 3 respectively in Fig. 4).

5.2. Results for soil probe 2

In a similar way to Section 5.1, this subsection presents the experimental results (obtained MSE rates) for the Probe 2 (at 30 cm depth). Firstly, Fig. 5 shows the results (average MSE) obtained by the Cubic, Makima, and Spline interpolations when applying the NAR, NIO and NARX models to this probe 2.

In general terms, it can be said about Fig. 5 that quite homogeneous average results are obtained by all interpolation methods. Regarding the neural model, it can be said that the lowest average MSE rates are obtained by NARX for all the three interpolation methods. The second best is NIO and the worst (highest MSE rates) is NAR. It can be concluded that including the series to be predicted themselves (endogenous input) is a successful strategy for the probe 2.

In addition to the previous results (average MSE), the lowest values or the error for the probe 2 are shown in Fig. 6. They are presented in a similar way, by the interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO and NARX) models.

Differentiating from the average results, the lowest error rates significantly vary from one interpolation method to the other ones. The Cubic interpolation clearly outperforms the remaining methods for all the neural models, while the Makima errors are slightly smaller than the Spline ones. On the other hand, the NARX neural model is the most

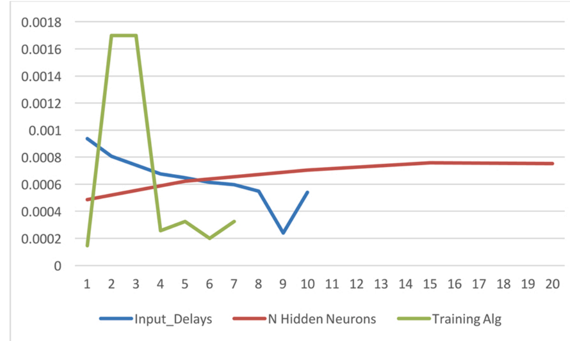

Fig. 4. Averaged MSE results obtained by Spline and NARX methods on the soil moisture probe 1. a) per the number of input delays, b) per hidden neurons, and c) per training algorithm.

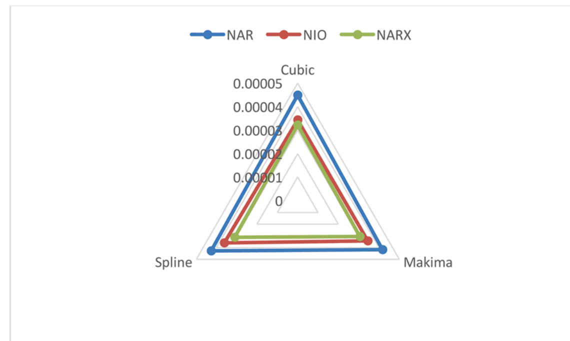

Fig. 5. Radar plot of the MSE average values on the soil moisture probe 2 per interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO, and NARX) model.

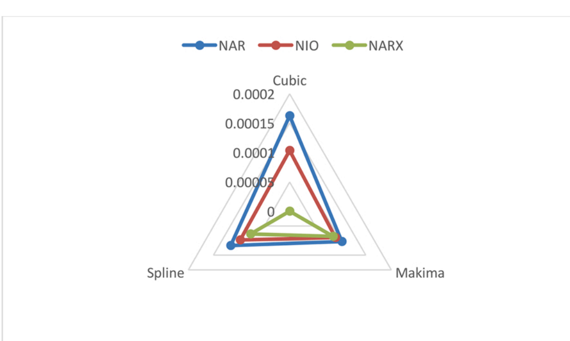

Fig. 3. Radar plot of the MSE average values on the soil moisture probe 1 per interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO, and NARX) model.

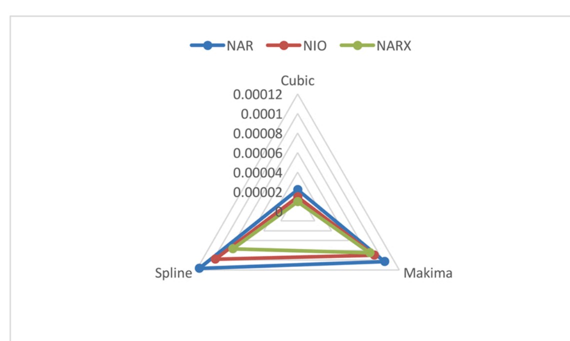

Fig. 6. Radar plot of the MSE average values on the soil moisture probe 2 per interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO, and NARX) model.

accurate one for all the interpolation methods. As it happened for the average results, NIO is the second best and NAR the worst one.

The lowest MSE value for this probe (9.8512 E-06) is obtained by the combination of the Cubic interpolation and the NARX neural model. In order to study the effect of the parameter values for such combination, Fig. 7 presents the average values obtained with these two methods when using different values for the NARX parameters (number of input delays, number of hidden neurons, and training algorithms).

After analysing the results shown in the Fig. 7, it can be said that the trends are quite similar to those pointed out for the probe 1 (Fig. 4): the more input delays, the less hidden neurons, the lower MSE. Additionally, the LM algorithm is the best training option for the NARX model.

5.3. Results for soil probe 3

Finally, the experimental results (obtained MSE rates) for the Probe 3 (at 45 cm depth) are presented in this subsection. Firstly, Fig. 8 shows the results (average MSE) obtained by the Cubic, Makima, and Spline interpolations when applying the NAR, NIO and NARX models to this probe 3.

In the case of this probe (3), results are not as clear as in the previous cases. There is neither an interpolation method nor a neural model that outperforms the other ones in all situations. For the Makima and Cubic interpolation, NARX has obtained the lowest MSE values while it has been NIO in the case of the Spline Interpolation. In addition to the previous results (average MSE), the lowest values or the error for the probe 3 are shown in Fig. 9. They are presented in a similar way, by the interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO and NARX) models.

It can be seen in the Fig. 9 that the trends are similar to those revealed in the Fig. 8: none of the interpolation and neural methods clearly outperforms the other ones in all situations. It is worth mentioning that for the probe 3, differentiating from the two previous probes (more superficial), NIO and Makima interpolation is the combination of methods obtaining the lowest error rate (4.4058 E-08). In order to study the effect of the parameter values for such combination, Fig. 10 presents the average values obtained with these two methods when using different values for the NIO parameters (number of input delays, number of hidden neurons, and training algorithms).

When analysing the evolution according to the input delays, in the case of probe 3, increasingly higher values are associated to increasingly lower error rates. Consequently, 10 is the value for such parameter that has obtained the lowest error rate. The trend is not so smooth in the case of the number of hidden neurons; although the error increases while adding hidden neurons, it abruptly decreases for 20 hidden neurons, actually being the value with the lowest average MSE value. Once again, LM is the training algorithm outperforming all the other ones.

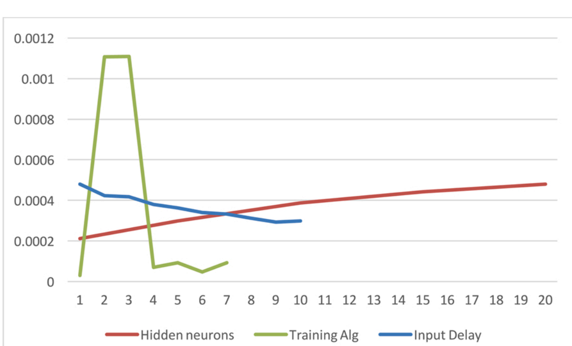

Fig. 7. Averaged MSE results obtained by Spline and NARX methods on the soil moisture probe 2. a) per the number of input delays, b) per hidden neurons, and c) per training algorithm.

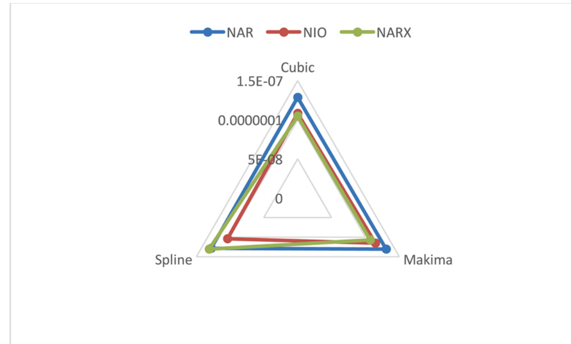

Fig. 8. Radar plot of the MSE average values on the soil moisture probe 3 per interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO, and NARX) model.

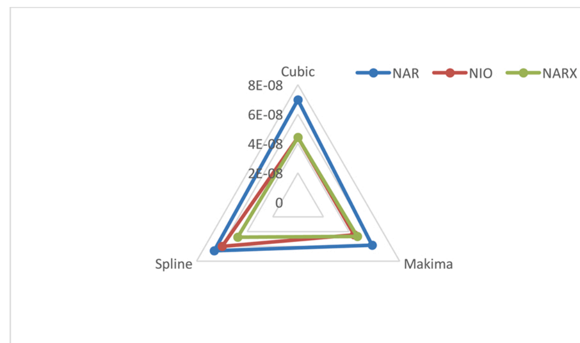

Fig. 9. Radar plot of the MSE average values on the soil moisture probe 3 per interpolation (Cubic, Makima, and Spline) and neural (NAR, NIO, and NARX) model.

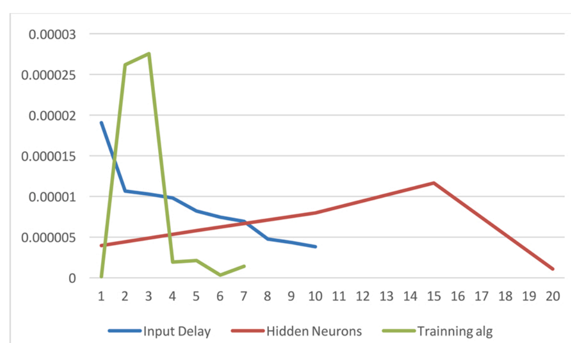

Fig. 10. Averaged MSE results obtained by Spline and NARX methods on the soil moisture probe 3. a) per the number of input delays, b) per hidden neurons, and c) per training algorithm.

6. Conclusions and future work

In the present study, different interpolation methods and time-series neural models have been combined and benchmarked in order to accurately predict the soil humidity level in a potato field. According to the experimental results previously presented, irrigation needs of the studied plantation could be successfully adjusted.

After analyzing these results, interesting conclusions can be derived from both the AI and agricultural perspective. Taking into account the interpolation and neural methods, it can be said that there is not any of such methods that clearly outperforms the alternate ones for all probes. NARX is the best neural model for the 2 most superficial probes (1 and 2 in Section 5). However, the Spline interpolation outperforms the other ones for the probe 1 while it was the Cubic one for the probe 2. The deepest probe entails different results from the previous ones: NIO is the neural model that has obtained the lowest error rate when combined with the Maxima interpolation.

As it is known, the parameter tuning of each model must be adjusted to each case as there is not a given combination of parameters that always leads to best results. However, there are some general guidelines that are coherent with the results obtained for the three probes: the most accurate predictions are obtained with a high number of input delays (9–10) and the LM training algorithm.

On the other hand, it can be concluded that the prediction error varies according to the depth of probes as the lowest error has been obtained for the probe 3. Authors believe that this can be cause by the smaller impact of weather conditions and irrigation system faults on the deepest probe. However, this phenomenon requires further investigation.

Actually, the activity of country-based institutional services involved in helping farmers to manage irrigation practices, are based only in weather predictions, being difficult to also take into account information about the available water content of the soil. The installation of non-expensive soil stations, with soil humidity probes located in reference soil profiles and covering wide irrigation areas, and the use of time series neural networks for data analysis, would greatly improve soil-water content monitoring and irrigation predictions. The main irrigated crops in the European context, such as potato, sugar beet, corn, alfalfa, and many other horticultural activities will obtain a clear benefit reducing irrigation and fertilization cost, avoiding crop diseases, and avoiding nutrient losses that are actually responsible for eutrophication.

As a proposal for future work, authors suggest applying some other AI models to improve the humidity forecast. Additionally, more input features from the agronomic field may be considered.

Data availability

Raw data and other supplementary material are available at the following repository: osf.io/7zpcj.

No data was used for the research described in the article.

Data will be made available on request.

Authorship conformation form

All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.

Declaration of Competing Interest

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

Acknowledgements

This work was financed by a grant agreement between Lab-Ferrer (METER Group) and the UBUCOMP research group at the University of Burgos. Authors are grateful to the farmer Mr. José María Izquierdo for providing the experimental field and the monitoring of irrigation.

References

[1] Agricultural Production Crops, https://ec.europa.eu/eurostat/statistics-explained/index.php/Agricultural_production_-_crops#Potatoes_and_sugar_beet, last accessed 09/02/2020.
[2] Yearly Statistics, https://www.mapa.gob.es/es/estadistica/temas/publicaciones/anuario-de-estadistica/2018/default.aspx?parte=3&capitulo=07&grupo=3&seccion=2, last accessed 09/02/2020.
[3] L.S. Pereira, T. Oweis, A. Zairi, Irrigation management under water scarcity, Agric. Water Manag. 57 (2002) 175–206.
[4] M. Yartu, C. Cambra, M. Navarro, C. Rad, A. Arroyo, A. Herrero, Neural models to predict irrigation needs of a potato plantation. 15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020), pp 600–613.
[5] D. Althoff, F.C.G. Alvino, R. Filgueiras, C.C. Aleman, F.F. da Cunha, Evapotranspiration for irrigated agriculture using orbital satellites, Biosci. J. 35 (2019) 670–678.
[6] G.R. Dixon, Water, irrigation, and plant diseases, CAB Rev. 10 (2015) 1–18, 009.
[7] A. Shitu, M. Tadda, A. Danhassan, Irrigation water management using smart control systems: a review, Bayero J. Eng. Technol. 13 (2018).
[8] F. Labbé, P. Ruelle, P. Garin, P. Leroy, Modelling irrigation scheduling to analyse wa-ter management at farm level, during water shortages, Eur. J. Agron. 12 (2000) 55–67.
[9] A. Fry, Water: Facts and trends. World Business Council for Sustainable Development, 2006.
[10] S. Andriyas, M. McKee, Recursive partitioning techniques for modeling irrigation be-havior, Environ. Model. Softw. 47 (2013) 207–217.
[11] G. Atsalakis, C. Minoudaki, N. Markatos, A. Stamou, J. Beltrao, T. Panagopoulos, Daily irrigation water demand prediction using adaptive neuro-fuzzy inferences systems (anfis), Proc. 3rd IASME/WSEAS International Conference on Energy, Environment, Ecosystems and Sustainable Development (2007) 369–374. WSEAS.
[12] R.J. Smith, J.N. Baillie, I. Futures, Defining precision irrigation a new approach to irrigation management, in: Proceedings of the Irrigation and Drainage Conference 2009, Victoria, Australia, 18–21 October, 2009, pp. 18–21.
[13] M.A. Khan, M.Z. Islam, M. Hafeez, Evaluating the performance of several data mining methods for predicting irrigation water requirement, AusDM (2012) 199–208.
[14] O. Adeyemi, I. Grove, S. Peets, Y. Domun, T. Norton, Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling, Sensors 18 (2018) 3408.
[15] A. Wunsch, T. Liesch, S. Broda, Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX), Hydrology 567 (2018) 743–758, https://doi.org/10.1016/j.jhydrol.2018.01.045. In press.
[16] S. Contreras, M.Á. Manzanedo, Á. Herrero, A hybrid neural system to study the interplay between economic crisis and workplace accidents in Spain, J. Univers. Comput. Sci. 25 (2019) 667–682.
[17] C. Alonso de Armiño, M.Á. Manzanedo, Á. Herrero, Analysing the intermeshed patterns of road transportation and macroeconomic indicators through neural and clustering techniques, Pattern Anal. Appl. 23 (2020) 1059–1070.
[18] N. Sirimorok, M. As, K. Yoshida, M. Köppen, Smart watering system based on framework of low-bandwidth distributed applications (LBDA) in cloud computing. INCoS 2020, in: Advances in Intelligent Systems and Computing, vol. 1263, Springer, 2020.
[19] S.A. Taqvi, L.D. Tufa, H. Zabiri, A.S. Maulud, F. Uddin, Fault detection in distillation column using NARX neural network, Neural Comput. Appl. 32 (2020) 3503–3519.
[20] H. Akima, A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points, ACM Trans. Math. Softw. 4 (1978) 148–159.
[21] I.J. Leontaritis, S.A. Billings, Input-output parametric models for non-linear systems Part Ideterministic non-linear systems, Int. J. Control 41 (1985) 303–328.

Carlos Cambra Baseca is a PhD Lecturer at the Polytechnic School of the University from Burgos (Spain). Its main areas of research interests are related to applied artificial intelligence, Precision Agriculture, networks of IoT sensors and artificial vision, as well as some aspects of robotics. Carlos Cambra has published articles in international and prestigious journals and congresses.