Artificial intelligence-based soil moisture estimation using a combination of in-situ measurements and open-source data

Conference Paper · September 2024

DOI: 10.46354/i3m.2024.foodops.011

3 authors:

Luca Preite
Università di Parma

Arianna Paini
Università di Parma

Giuseppe Vignali
Università di Parma

10th International Food Operations & Processing Simulation Workshop, 011
21st International Multidisciplinary Modeling & Simulation Multiconference

2724-0355 © 2024 The Authors.
doi: 10.46354/i3m.2024.foodops.011

Artificial intelligence-based soil moisture estimation using a combination of in-situ measurements and open-source data.

Luca Preite[1,*], Arianna Paini[1] and Giuseppe Vignali[1]

[1]University of Parma, Department of Engineering for Industrial Systems and Technologies, Parco Area delle Scienze 181/A, Parma, 43124 Italy

[*]Corresponding author. Email address: luca.preite@unipr.it

Abstract

Sustainable solutions must be explored to conserve natural resources such as soil and water, which are being depleted by both natural and anthropogenic causes. The proposed work aims to address water scarcity in agriculture by implementing an artificial intelligence algorithm to estimate soil water content, which can be used to control irrigation. A dedicated living lab focused on a tomato crop is designed to monitor soil and environmental conditions throughout the growing season. In this framework, an IoT network characterizes the acquisition data layer, which collects data on soil (moisture, temperature, and electrical conductivity) and environmental properties (i.e., temperature, carbon dioxide concentration). Additional hourly data on precipitation, solar radiation, wind speed, and reference evapotranspiration are then aggregated by communicating with an open-source API. The resulting dataset is preprocessed to train and test machine learning algorithms, such as linear regression, polynomial regression, random forest regression, and a multilayer perceptron artificial neural network. A heuristic approach is used to optimize hyperparameters and prediction accuracy. In conclusion, the artificial neural network provides better predictive performance, allowing the estimation of the soil moisture with a high degree of accuracy, achieving a coefficient of determination close to 98% and a mean absolute error below 3%.

Keywords: Water Management, Precision Agriculture, Soil Water Content, Machine Learning, Predictive Algorithms

1. Introduction

The agricultural sector needs to face several global issues in the coming years. In this perspective, climate change, soil erosion, water scarcity, population growth, and the instability of the global economic factors are only the most pressing (Bissadu et al., 2024). According to the Food and Agriculture Organization of the United States (FAO), agricultural activities account for 70% of the global water withdrawals, making it one of the sectors with the greatest impact on water scarcity (Food and Agriculture Organization of the United Nations, 2023).

In this framework, future research and investments need to be performed to reduce the pressure on this valuable resource. Therefore, the traditional agricultural practices are less digited and have no control over growing parameters and water used for irrigations (Lou et al., 2024). The application of the 4.0 technologies at the industrial level, provides different advantages that can be also moved to the agricultural in order to properly monitored all the parameters involved through an interconnected physical layer (Codeluppi et al., 2020). Indeed, the scientific literature highlights how several efforts have been carried out to enhance the agricultural sustainability with multiple technologies (Preite et al., 2023b). One of the most important is the internet of things paradigm, as it allows instruments to be smart and interconnected among different levels (Tzounis et al., 2017). Blockchain, digital twin, and artificial intelligence applications are the technologies that can be integrated to the internet of things to gain insights into the collected data and develop efficient decision support systems to optimize all the processes involved and reduce the resource exploitation (Preite et al., 2023a).

The proposed work aims at training and testing different machine learning models to predict the soil moisture in a dedicated living lab, which was focused on a tomato cultivation in open field. Specifically, a proper internet of things network, based on the Low Range Wide Area communication technology provide by the LoRa Alliance, was used to collect data over environmental and soil features and to monitor the irrigations during the growing season (LoRa Alliance, 2024). The data was then preprocessed for learning correlations between the features and preparing a suitable dataset for dealing with the reported regression problems. The most studied algorithms for regression, the multi-layer perceptron artificial neural network, linear and polynomial regression, and the random forest regressor, were implemented to define the best model for the issue investigated.

The prediction performances are evaluated by calculating the coefficient of determination, the mean square error, and the mean absolute error. Summarizing the results, the multi-layer perceptron artificial neural network is the most suitable solution, as it provides a coefficient of determination close to 98% and a mean absolute error of 2.38%.

This paper is organized in different sections. The next section describes the current state of art to highlight gaps, limitations, and future trends in the application of internet of things and artificial intelligence application in agriculture. The methodology is reported in section 3, where the developed living lab and the proposed model are described. Finally, sections 4 and 5 outline the results achieved, provide a discussion over them, and define the conclusion and future development.

2. State of the art

The scientific literature reports some efforts in applying artificial intelligence in agriculture. In this perspective, machine learning algorithms have been investigated for their capabilities to deal with complex problems by gaining insights from multiple features (Araújo et al., 2021). Considering all the agricultural processes, the applications can be divided into three groups: i) pre-harvesting, ii) harvesting, and iii) post-harvesting (Meshram et al., 2021). Therefore, applications in the first group aims at optimizing the water and soil management, the seeding phase and early detecting weed and diseases (Arnal Barbedo, 2019; Arsenovic et al., 2019; Saleem et al., 2019). The main goal of the harvesting operations is to guarantee high food quality and profitability by analyzing physical properties and marketing required parameters, such as color, flavors, size, maturity age and firmness (Chen et al., 2019; Kirk et al., 2020; Onishi et al., 2019). Finally, post-harvesting phase is crucial to preserve the shelf life of the products by controlling some important features like temperature and humidity (Ireri et al., 2019; Li et al., 2018; Piedad et al., 2018).

Focusing on the water management applications, the authors of (Preite & Vignali, 2024) provide a decision support system, based on a classification framework, able to manage the irrigation by assessing environmental and soil data (i.e., ambient temperature, ambient humidity, evapotranspiration, soil moisture, soil temperature, and soil electrical conductivity) aggregated with the weather forecasts. The proposed model achieved high accuracy, close to 99%, in predicting the irrigation activation.

Evapotranspiration is an important phenomenon that influence the water balance between soil, plant, and environment (Pereira et al., 2015). Specifically, it combines the evaporation and transpiration, where the first characterized the evaporation from the soil surface and the capillary fringe. Meanwhile, transpiration considers the amount of water released by the plant to the environment because of their metabolism. The authors of (Feng et al., 2017) tested an artificial neural network and an extreme machine learning algorithm to estimate daily evapotranspiration by assessing the ambient maximum and minimum temperature, mean relative humidity, wind speed at 10-meter height, and solar irradiation. The monthly mean reference evapotranspiration was investigated by developing linear and multivariate adaptive regression (Mehdizadeh et al., 2017). However, different models, such as k-nearest neighbors, support vector machine, multilinear regression, and decision tree were tested and validated, achieving high accuracy (Morellos et al., 2016; Yang et al., 2019). A super learner ensemble and a hybrid neuro-fuzzy model were developed by (Adnan et al., 2021) to estimate evapotranspiration with limited weather data. In order to face the same issue, the combination of machine learning algorithms with optimization meta-heuristic algorithms was deployed in (Yong et al., 2024).

Soil moisture and evapotranspiration were assessed with different algorithms (i.e., support vector machine, artificial neural network, and gradient-boosting decision tree) by taking multispectral and thermal images and collecting data on environmental and soil properties (Alibabaei et al., 2021). The soil moisture has been also predicted by implementing time series forecasting and comparing 4.0 irrigation frameworks with traditional techniques. The authors of (Srivastava et al., 2024) develop a probabilistic framework to deal with the irrigation management, where the soil moisture, leaf area index, and evapotranspiration were estimated by combining a random forest regressor with Long-Short-Term Memory artificial neural network. For the same purpose, in (Singh et al., 2024) the artificial neural network resulted in higher predictive performance by testing 5 different models (i.e., linear regression, support vector machine, decision tree, random forest). Four data mining algorithms, random tree, reduced error-pruning tree, M5P, and random subspace ensemble, were proposed by (Emami et al., 2024). In this framework, the random tree achieved a coefficient of determination close to 97%. An advanced framework based on a deep learning application was reported in (Kara et al., 2024). Specifically, the genetic algorithm was exploited for improving the prediction accuracy provided by a long short-term memory artificial neural network.

Machine learning techniques deal with some limitations when applied in agriculture. Therefore, the lack of consistent data for an effective training and the scalability in different operating conditions were the main issues highlighted in the scientific literature (Liakos et al., 2018). This because the growth periods are usually long and there are many parameters that influence the data quality.

3. Materials and Methods

3.1. Living lab and experimental activities

Tomatoes are grown according to the organic rules in a dedicated living lab located in northern Italy. It was designed in collaboration with *Azienda Sperimentale Stuard* within the national framework of Agritech. Specifically, this living lab aims at testing different solution to optimize water consumption in agriculture to address water scarcity issue. Specifically, the experimental activities allow different water stress levels to be tested by deploying three lines of 90 meters, which are continuously monitored during the growing seasons. In this perspective, 100%, 60%, and 30% of the watering recommendation provided by the Irriframe tool have been applied to the rows involved (Irriframe, 2024). As final step, once the harvest has been completed, the productivity performance in term of total, commercial and blessed root production volumes have been investigated to assess the resulting drought stress. Furthermore, the experimental rows are spaced by using a boundary row between them to avoid negative interactions on the collected data. Within this framework, the proposed study aims at gaining valuable insights into the involved parameters by deploying a multilayer framework to real-time collect and process soil and environmental data.

3.2. Water balance at the soil level

Different flow rates influence the water balance at the soil level. Specifically, the soil can be divided into three different layers: i) surface layer, ii) root layer, and iii) above-root-layer. Each layer can be modeled as a tank with a given capacity, where precipitations, irrigations, and water uptake from the groundwater to the root layer are the flowrate entering in the system. Evapotranspiration, percolation, and horizontal runoff water are considered as output flow rates (Battilani & Ventura, 1997).

3.3. Acquisition data layer

A LoRaWAN network has been designed to continuously monitor the growing parameters. Specifically, several end nodes were installed in the field and communicated over radio frequency bands (EU 868 in Europe) with a gateway connected via internet protocol with TheThingsNetwork server (TTN, 2024). Starting from this point, data were stored in databases and processed in multiple processing units to allow data visualization, data pattern discovery and predictive algorithms development. The involved end nodes are described in the follow:

- Three soil sensors (one device for each experimental row) were used to assess soil parameters, such as soil moisture expressed in term of water volume percentage, soil electrical conductivity expressed in milliSiemens per meter, and soil temperature reported in degree Celsius.
- Three water meters were directly installed on the irrigation pipes to track the water volume delivered to each row.
- Three solenoid valves to continuously monitor the irrigation status
- An environmental sensor for gaining information on the carbon dioxide concentration expressed in ppm, ambient temperature, relative humidity, and atmospheric pressure expressed in hPa.

The devices described were employed in Class A and a transmission rate of 10 minutes was set up.

3.4. Open-source API for weather forecasting

As it can be seen in section 4.2. evapotranspiration and weather phenomena play a crucial role in driving the water flow rates across the soil levels. The first is a well-known occurrence which encompasses both the evaporation of the water from the soil and capillary bands of the groundwater and the amount of water released due to the plant metabolic activities. According to the (Pereira et al., 2015), evapotranspiration is highly dependent on the climate conditions and drought status of the crop. In the proposed framework, the evapotranspiration reference, and the weather forecasts in terms of precipitation amount (expressed in millimeters), wind speed at 10-meters height (expressed in km/h), and global tilted irradiance (expressed in Watt per square meter), were hourly gained from an open-source API by leveraging a HTTPS request-response procedure. This platform is based on the national weather services and integrates high-resolution local and global weather models. In addition, the reference evapotranspiration provided by this API, which takes into account the weather influence, was calculated with the Penman-Monteith formula. As second step, the impact of the crop was evaluated by multiplying the reference value with the specific coefficient for tomatoes.

These data were aggregated to the data collected on the field in order to provide a consistent dataset to train and test machine learning algorithms.

3.5. Data preprocessing

Data cleaning and data preprocessing were also performed for maximizing the accuracy in making predictions. In this regard, as the dataset was characterized by many features having different magnitude, data were standardized to normally distribute all the features according to a Gaussian distribution with zero mean and unit variance.

Pearson correlation coefficient was then calculated to discover linear relationship between the independent variables. According to the Equation 1, x_i and y_i are the value of the independent and dependent variables, respectively. In this equation, x_m and y_m represent the corresponding mean values.

$$r = \frac{\sum(x_i - x_m)(y_i - y_m)}{\sqrt{\sum(x_i - x_m)^2 (y_i - y_m)^2}} \quad (1)$$

For the sake of deepening the non-linear relationship between the features involved, the asymmetric, data-agnostic-type predictive power score (PPS), which ranges from 0 (no relationship) to 1 (perfect relationship), was also calculated according to Equation 2. Specifically, this score assesses the relationship between the mean absolute error and a baseline score, defined as a naïve mean absolute error that predicts the median of the target value.

$$PPS = 1 - (MAE/MAE\_Naïve) \quad (2)$$

Table 1 reports the features involved in the dataset with the corresponding units of measurement. For the sake of completeness, there are 16187 samples in the data set.

Table 1. Features description with corresponding unit of measurement.

Independent Features	Units
Soil Electrical Conductivity	[mS/m]
Soil Temperature	[°C]
FAO evapotranspiration	[mm]
Irrigation	[mm]
Precipitation	[mm]
Wind Speed (10 meters)	[km/h]
Ambient humidity	[%]
Ambient temperature	[°C]
Environment Co_2	[ppm]
Global Tilted Irradiance	[W/m^2]

3.6. Processing units: predictive algorithms

As it mentioned before, the main aim of this work is to predict the soil water content, which is a continuous feature. Thus, a regression problem is treated, where the soil moisture represents the target value. In this framework, the performance of four different machine learning algorithms (i.e., linear regression, polynomial regression, random forest, and multi-layer perceptron neural network) were assessed to define the most suitable algorithm for the given application. These were developed with Python using scikit-learn and keras libraries. A heuristic approach was followed to optimize the hyperparameter of each model to maximize the prediction accuracy by exploiting the GridSearchCV class, which assesses different scenarios with the three-fold validation method.

Multiple linear regression was the first model investigated. Matrix correlation and the predictive power scores were used for cutting redundant dependent features to avoid multicollinearity and overfitting. Instead, all the attributes were used for training and the testing the random forest model and the multi-layer perceptron artificial neural network due its intrinsic capability to gain insights into the feature importance.

Each model has been validated by calculating the coefficient of determination (R^2) according to the equation 3, where $\hat{x}$ is the predicted value, $\bar{x}$ is the mean value of the testing data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x-\hat{x})^2}{\sum_{i=1}^{n}(x-\bar{x})^2} \quad (3)$$

In addition, the mean absolute error and the mean square error were evaluated according to equation 4-5. In this case, $\hat{y}_i$ is predicted value of the $i-th$ sample, y_i is the corresponding real value, and n is the total number of samples analyzed.

$$MAE = \frac{1}{n}\sum_{i=0}^{n-1}|y_i - \hat{y}_i| \quad (4)$$

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

4. Results and Discussion

4.1. Data preprocessing

Furthermore, Pearson correlation matrix is reported in Figure 1.

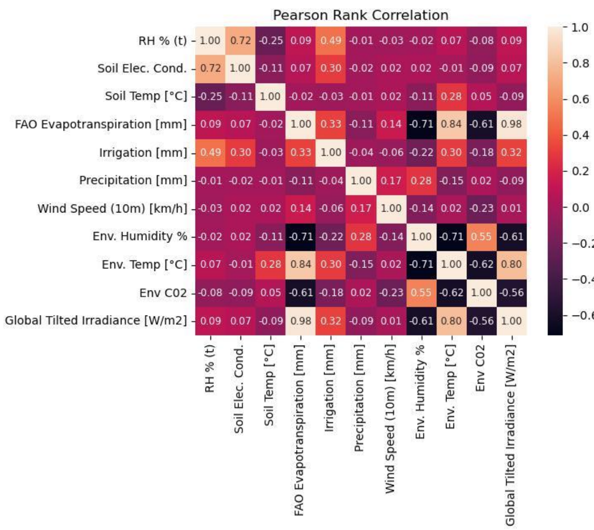

Figure 1. Pearson matrix correlation for the collected data

These sources have been used to reshape the data set to develop the linear regression algorithm, avoiding overfitting and multicollinearity. In this framework, the soil moisture is characterized by a linear correlation with soil electrical conductivity, soil temperature, reference evapotranspiration and the amount of water delivered to the crop. Specifically, soil electrical conductivity and irrigation show the highest Pearson coefficient. Furthermore, the predictive power score allows non-linear relationship to be investigated. As it can be seen in table 2, the same independent features were highlighted as the best predictors.

Table 2. Predictive Power Score for each feature analyzed

Independent Features	PPS
Soil Electrical Conductivity	0.499
Soil Temperature	0.139
FAO evapotranspiration	0.048
Irrigation	0.339
Precipitation	0.099
Wind Speed (10 meters)	0.092
Ambient humidity	0.076
Ambient temperature	0.083
Environment Co_2	0.035
Global Tilted Irradiance	0.042

4.2. Prediction performance

The dataset has been reshaped for implementing both the linear and polynomial algorithms. Specifically, soil electrical conductivity, soil temperature and the water amount delivered with the irrigations were the variable used to train and test the mentioned algorithms. As a result, the regression function is defined by the Equation 6, where y is response, β_0 is the intercept, X_1, X_2, X_3 are the dependent features described and β_1, β_2, β_3 are the associated weights, and finally ε is the error that encompasses the effect of variables not included in the model.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon \quad (6)$$

As a results of the hyperparameter optimization process, the intercept was computed to enhance the prediction performance of the model. The regression model has been also extended by developing a third order polynomial pipeline for handling more complex patterns between data and achieved higher accuracy. According to the same framework, squared error was the function resulted for the random forest regressor model. The multi-layer perceptron artificial neural network structure, which allows performance to be maximized, is characterized by three hidden layers with 50, 100, and 50 nodes, implementing the hyperbolic tangent as activation functions. The Adaptive Moment Estimation (Adam) algorithm was implemented for the stochastic optimization with a regularization parameter alpha of 0.0001.

Table 3 summarizes the metrics used to assess the model performances (i.e., R^2, mean square error, and mean error). Furthermore, Figure 2 reports the relationship between the real and the predicted values.

Table 3. Evaluation metrics for the implemented algorithms

Models	R^2	MSE	MAE	MAE (%)
Linear Regression	0.62	6.58	2.57	11
Polynomial Regression	0.79	3.84	1.96	8.43
Neural Network	0.98	0.31	0.55	2.38
Random Forest	0.67	5.79	2.41	10.4

In this perspective, the multi-layer artificial neural network results in the best artificial intelligence model for the proposed problem. Therefore, this algorithm achieves a coefficient of determination close to 98% in predicting the soil moisture value with a mean error of 2.38%.

The high predictive performance achieved by the multi-layer perceptron artificial neural network demonstrate that high quality data can be collected for artificial intelligence applications by exploiting an IoT network based on the LoRa communication protocol. Finally, a feasibility study of the proposed LoRaWAN network is available at (Stefanini et al., 2023), where it is stated that the network will have a positive economic effect on the farmers.

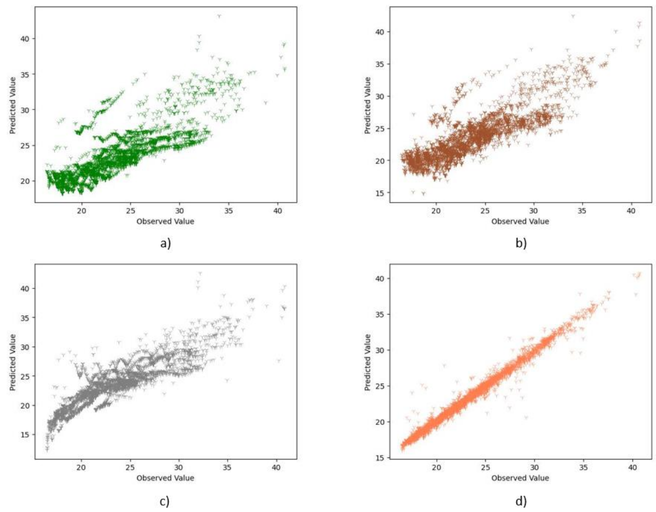

Figure 2. Relationship between the predicted value and the corresponding observed value for the four algorithms tested: a) Linear regression, b) Random Forest, c) Polynomial regression, and d) multi-layer perceptron artificial neural network

5. Conclusion

Several challenges threaten the future of agriculture, such as water scarcity, soil erosion, population growth and others. The first issue has become more important because, according to the FAO, agriculture is one of the sectors with the highest impact on the water withdrawals, accounting for 70% of the total. In this regards, new solutions for reducing the impact on this precious resource have to be investigated. The scientific literature shows how the effectiveness of the 4.0 paradigm can be also applied to the agriculture. As it highlighted in section 2, Internet of Things is an essential technology to create interconnected physical layers that allow important features to be monitored and controlled. In addition, digital twin and artificial intelligence application can also be integrated in order to create an effective virtual layer able to make data-driven decisions. Specifically, different efforts have been carried out in the scientific literature to develop decision support systems to manage irrigations or to estimate suitable parameters, such as evapotranspiration and soil moisture. In this perspective, the latter two features drove the water balance between atmosphere, plant, and soil, as described in the section 3.

The proposed framework aims to strengthen knowledge in applying machine learning algorithms for soil moisture estimation. Specifically, a dedicated living lab, focused on a tomato crop in open field, has been designed for performing different experimental activities. Different 4.0 technologies have been tested through the installation of LoRaWAN network for monitoring environmental and soil properties. The main aim is to aggregate a consistent number of features to virtually modeled the growing conditions and the water balance at the soil level. For this purpose, 3 days-weather forecast data (i.e., precipitations, reference evapotranspiration, wind speed, and the total tilted irradiance) have been gathered from an open-source API. The dataset was then preprocessed for training and testing four machine learning algorithms, i.e., a liner and polynomial regression, a random forest model, and a multi-layer perceptron artificial neural network. Prediction performances have been evaluated by calculating the coefficient of determination (R^2), the mean square error (MSE), and the mean absolute error (MAE). The results achieved reveal how the multi-layer perceptron artificial neural network is able to estimate the soil moisture with high degree of accuracy, with a coefficient of determination close to 98% and a mean absolute error that falls below 3%. Furthermore, the results demonstrate that high quality data for artificial intelligence application can be collected through a LoRaWAN network. Future developments aim to address the scalability issue of machine learning implementations by modeling different operating conditions, different soil configurations, and different crops, with the purpose of providing a consistent solution that can be scaled across multiple applications.

Acknowledgements

This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

Adnan, R. M., Mostafa, R. R., Reza, A., Islam, M. T., Kisi, O., Kuriqi, A., & Heddam, S. (2021). Estimating reference evapotranspiration using hybrid adaptive fuzzy inferencing coupled with heuristic algorithms. *Computers and Electronics in Agriculture*, *191*, 106541. https://doi.org/10.1016/j.compag.2021.106541

Alibabaei, K., Gaspar, P. D., & Lima, T. M. (2021). Modeling Soil Water Content and Reference Evapotranspiration from Climate Data Using Deep Learning Method. *Applied Sciences*, *11*(11), 5029. https://doi.org/10.3390/app11115029

Araújo, S. O., Peres, R. S., Barata, J., Lidon, F., & Ramalho, J. C. (2021). Characterising the Agriculture 4.0 Landscape—Emerging Trends, Challenges and Opportunities. *Agronomy*, *11*(4), 667. https://doi.org/10.3390/agronomy11040667

Arnal Barbedo, J. G. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, *180*, 96–107. https://doi.org/10.1016/J.BIOSYSTEMSENG.2019.02

.002

Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., & Stefanovic, D. (2019). Solving Current Limitations of Deep Learning Based Approaches for Plant Disease Detection. *Symmetry*, *11*(7), 939. https://doi.org/10.3390/sym11070939

Battilani, A. and Ventura, F. (1997). INFLUENCE OF WATER TABLE, IRRIGATION AND ROOTSTOCK ON TRANSPIRATION RATE AND FRUIT GROWTH OF PEACH TREES. Acta Hortic. 449, 521-528.DOI: 10.17660/ActaHortic.1997.449.72. https://doi.org/10.17660/ActaHortic.1997.449.72

Bissadu, K. D., Sonko, S., & Hossain, G. (2024). Society 5.0 enabled agriculture: Drivers, enabling technologies, architectures, opportunities, and challenges. *Information Processing in Agriculture*. https://doi.org/10.1016/J.INPA.2024.04.003

Chen, M., Li, P., Chen, J., Bresilla, K., Grappadelli, L. C., Manfrini, L., Demetrio Perulli, G., Boini, A., & Morandi, B. (2019). *Single-Shot Convolution Neural Networks for Real-Time Fruit Detection Within the Tree*. https://doi.org/10.3389/fpls.2019.00611

Codeluppi, G., Cilfone, A., Davoli, L., & Ferrari, G. (2020). LoRaFarM: A LoRaWAN-Based Smart Farming Modular IoT Architecture. *Sensors*, *20*(7), 2028. https://doi.org/10.3390/s20072028

Emami, S., Rezaverdinejad, V., Dehghanisanij, · Hossein, Emami, H., & Elbeltagi, A. (2024). Data mining predictive algorithms for estimating soil water content. *Soft Computing, Volume 28, pages 4915–4931, (2024).* https://doi.org/10.1007/s00500-023-09208-3

Feng, Y., Peng, Y., Cui, N., Gong, D., & Zhang, K. (2017). Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Computers and Electronics in Agriculture*, *136*, 71–78. https://doi.org/10.1016/J.COMPAG.2017.01.027

Food and Agriculture Organization of the United Nations. (2023). 2050: A third more mouths to feed. Retrieved from: https://www.fao.org

Ireri, D., Belal, E., Okinda, C., Makange, N., & Ji, C. (2019). A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. *Artificial Intelligence in Agriculture*, *2*, 28–37. https://doi.org/10.1016/J.AIIA.2019.06.001

Irriframe, 2024. Retrieved from: https://www.irriframe.it. Accessed in June 2024

Kara, A., Pekel, E., Ozcetin, · Erdener, Gazi, ·, & Yıldız, B. (2024). Genetic algorithm optimized a deep learning method with attention mechanism for soil moisture prediction. *Neural Computing & Applications, Volume 36, pages 1761–1772, (2024).* https://doi.org/10.1007/s00521-023-09168-7

Kirk, R., Cielniak, G., & Mangan, M. (2020). L*a*b*Fruits: A Rapid and Robust Outdoor Fruit Detection System Combining Bio-Inspired Features with One-Stage Deep Learning Networks. *Sensors*, *20*(1), 275. https://doi.org/10.3390/s20010275

Li, J., Chen, L., & Huang, W. (2018). Detection of early bruises on peaches (Amygdalus persica L.) using hyperspectral imaging coupled with improved watershed segmentation algorithm. *Postharvest Biology and Technology*, *135*, 104–113. https://doi.org/10.1016/J.POSTHARVBIO.2017.09.007

Liakos, K., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine Learning in Agriculture: A Review. *Sensors*, *18*(8), 2674. https://doi.org/10.3390/s18082674

LoRa Alliance, 2024. LoRaWAN Specification v1.1. Retrieved from: https//resources.lora-alliance.org/technical-specifications/lorawan-specifications-v1-1. Accessed in June 2024.

Lou, Y., Xing, W., Hao, W., Mei, X., Feng, L., Sun, Z., Hu, N., Noellemeyer, E., Le Cadre, E., Minamikawa, K., Muchaonyerwa, P., AbdelRahman, M., Machado Pinheiro, É. F., de Vries, W., Liu, J., Chang, S. X., & Zhou, J. (2024). Climate-smart agriculture: Insights and challenges. *Climate Smart Agriculture*, *1*(1), 100003. https://doi.org/10.1016/J.CSAG.2024.100003

Mehdizadeh, S., Behmanesh, J., & Khalili, K. (2017). Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. *Computers and Electronics in Agriculture*, *139*, 103–114. https://doi.org/10.1016/J.COMPAG.2017.05.002

Meshram, V., Patil, K., Meshram, V., Hanchate, D., & Ramkteke, S. D. (2021). Machine learning in agriculture domain: A state-of-art survey. *Artificial Intelligence in the Life Sciences*, *1*, 100010. https://doi.org/10.1016/J.AILSCI.2021.100010

Morellos, A., Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., & Mouazen, A. M. (2016). Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering*, *152*, 104–116. https://doi.org/10.1016/J.BIOSYSTEMSENG.2016.04.018

Onishi, Y., Yoshida, T., Kurita, H., Fukao, T., Arihara, H., & Iwai, A. (2019). An automated fruit harvesting robot by using deep learning. *Robomech Journal, 6. Article number:13 (2019).* https://doi.org/10.1186/s40648-019-0141-2

Pereira, L. S., Allen, R. G., Smith, M., & Raes, D. (2015). Crop evapotranspiration estimation with FAO56: Past and future. *Agricultural Water Management*, *147*, 4–20. https://doi.org/10.1016/J.AGWAT.2014.07.031

Piedad, E., Larada, J. I., Pojas, G. J., & Ferrer, L. V. V. (2018). Postharvest classification of banana (Musa acuminata) using tier-based machine learning. *Postharvest Biology and Technology*, *145*, 93–100. https://doi.org/10.1016/J.POSTHARVBIO.2018.06.004

Preite, L., Solari, F., & Vignali, G. (2023a). A digital model application to optimize water consumption in agriculture. *Proceedings of the International Food Operations and Processing Simulation Workshop, FOODOPS*, *2023-September*. https://doi.org/10.46354/i3m.2023.foodops.006

Preite, L., Solari, F., & Vignali, G. (2023b). Technologies to Optimize the Water Consumption in Agriculture: A Systematic Review. *Sustainability*, *15*(7), 5975. https://doi.org/10.3390/su15075975

Preite, L., & Vignali, G. (2024). Artificial intelligence to optimize water consumption in agriculture: A predictive algorithm-based irrigation management system. *Computers and Electronics in Agriculture*, *223*, 109126. https://doi.org/10.1016/J.COMPAG.2024.109126

Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant Disease Detection and Classification by Deep Learning. *Plants*, *8*(11), 468. https://doi.org/10.3390/plants8110468

Singh, T., Kundroo, M., & Kim, T. (2024). WSN-Driven Advances in Soil Moisture Estimation: A Machine Learning Approach. *Electronics*, *13*(8), 1590. https://doi.org/10.3390/electronics13081590

Srivastava, S., Kumar, N., Malakar, A., Sruti, ·, Choudhury, D., Chittaranjan Ray, ·, & Roy, T. (2024). A Machine Learning-Based Probabilistic Approach for Irrigation Scheduling. *Water Resources Management. Volume 38, pages 1639–1653, (2024).* https://doi.org/10.1007/s11269-024-03746-7

Stefanini, R., Preite, L., Bottani, E., Belli, L., Davoli, L., Ferrari, G., & Vignali, G. (2023). Selection of 4.0 sensors for small holders: the compromise between the advantages and the costs of the implementation. *Proceedings of the International Food Operations and Processing Simulation Workshop, FOODOPS*, *2023-September*. https://doi.org/10.46354/i3m.2023.foodops.007

TTN, 2024. The Things Network platform. https://www.thethingsnetwork.org. Accessed in June 2024.

Tzounis, A., Katsoulas, N., Bartzanas, T., & Kittas, C. (2017). Internet of Things in agriculture, recent advances and future challenges. *Biosystems Engineering*, *164*, 31–48. https://doi.org/10.1016/J.BIOSYSTEMSENG.2017.09.007

Yang, M., Xu, D., Chen, S., Li, H., & Shi, Z. (2019). Evaluation of Machine Learning Approaches to Predict Soil Organic Matter and pH Using vis-NIR Spectra. *Sensors*, *19*(2), 263. https://doi.org/10.3390/s19020263

Yong, S. L. S., Ng, J. L., Huang, Y. F., Ang, C. K., Ahmad Kamal, N., Mirzaei, M., & Najah Ahmed, A. (2024). Enhanced Daily Reference Evapotranspiration Estimation Using Optimized Hybrid Support Vector Regression Models. *Water Resources Management*. https://doi.org/10.1007/s11269-024-03860-6