




Review

Machine Learning for Smart Irrigation in Agriculture: How Far along Are We?

Marco Del-Coco [†] , Marco Leo ^{*,†}  and Pierluigi Carcagni [†] 

Institute of Applied Sciences and Intelligent Systems (ISASI), National Research Council (CNR), Via Monteroni snc University Campus, 73100 Lecce, Italy; marco.delcoco@cnr.it (M.D.-C.); pierluigi.carcagni@cnr.it (P.C.)

* Correspondence: marco.leo@cnr.it

[†] These authors contributed equally to this work.

Abstract: The management of water resources is becoming increasingly important in several contexts, including agriculture. Recently, innovative agricultural practices, advanced sensors, and Internet of Things (IoT) devices have made it possible to improve the efficiency of water use. However, it is the application of control strategies based on advanced machine learning techniques that enables the adoption of smart irrigation scheduling and the immediate economic, social, and environmental benefits. This challenging research area has attracted the attention of many researchers worldwide, who have proposed several technological and methodological solutions. Unfortunately, the results of these scientific efforts have not yet been categorized in a thematic survey, making it difficult to understand how far we are from optimal water management based on machine learning. This paper fills this gap by focusing on smart irrigation systems with an emphasis on machine learning. More specifically, the generic structure of a smart agriculture system is presented, and existing machine learning strategies and available datasets are discussed. Furthermore, several open issues are identified, especially in the processing of long-term data, also due to the lack of corresponding annotated datasets. Finally, some interesting future research directions to be pursued in order to build scalable, domain-independent approaches are proposed.

Keywords: smart irrigation; machine learning; water resources management; precision agriculture



Citation: Del-Coco, M.; Leo, M.; Carcagni, P. Machine Learning for Smart Irrigation in Agriculture: How Far along Are We? *Information* **2024**, *15*, 306. <https://doi.org/10.3390/info15060306>

Academic Editors: Vasco N. G. J. Soares and Zahir M. Hussain

Received: 4 April 2024
Revised: 20 May 2024
Accepted: 22 May 2024
Published: 24 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Water Resources Management (WMR) is getting to be an imperative point in different production areas. Seventy percent of freshwater used worldwide is consumed by agriculture, and about 1.2 billion people live in regions with severe water scarcity and shortages that make agriculture difficult, high drought frequency in pastureland and rain-fed areas, or high water stress in irrigated areas [1]. When water is limited, production is reduced, and food and feed prices increase, threatening food security and affordability (https://agriculture.ec.europa.eu/system/files/2023-05/factsheet-agriresearch-water-managment_en_0.pdf (accessed on 23 November 2023).) Water use efficiency can be firstly improved by substituting traditional surface/drip irrigation systems with modern irrigation methods (spray sprinkler, rotor sprinkler, rotary nozzle, and rotators) [2]. Changing the irrigation system is not always possible, and its implementation must consider compatibility with the best services of the farm, the topographic and properties of the soil, crop specs, economic feasibility, and some social constraints. This solution may not completely solve the water scarcity issue and then a further increase in water efficiency can be achieved through a smart use of irrigation systems. Smart irrigation involves the application of water at the right time, in the right amounts, and even at the right spot in the field using existing irrigation systems [3]. This way, it is possible to achieve a balance between the three pillars (economic, social, and environmental) of sustainability. The development of smart irrigation systems is thus attracting the attention of many researchers all around the world.

The fundamental principle of smart irrigation is its conceptual framework, which integrates agricultural methods and cutting-edge sensors and IoT devices to maximize water use efficiency. Sensing soil and/or weather conditions, it is possible to gather real-time data on variables critical for irrigation decisions. Irrigation operations can be performed more efficiently by using automation and control systems that allow for real-time adjustments depending on data collection. Web-based systems for the remote monitoring and management of irrigation [4], and IoT platforms based on edge and cloud computing [5] are some examples of such a kind of smart irrigation approach. A deeper analysis of the historical data, environmental factors, and agricultural technical knowledge can be useful for better forecasting crop water needs. This further step requires the study and exploitation of advanced algorithms that can learn from available data and generalize to unseen future data, falling this way into the scientific sector called machine learning [6].

In recent years, therefore, various technological and methodological solutions have been proposed, with great scientific excitement in proposing machine learning-based solutions capable of optimizing irrigation scheduling. Unfortunately, the results of this scientific fervor have not been categorized yet by a thematic survey, making it difficult to understand how far along we are in terms of optimal water management. Most of the existing surveys have concentrated indeed on the broader application areas of smart and precision agriculture, with particular attention to plant disease detection and classification [7]. An overview of trends in sensors and IoT systems for irrigation can be found in [8] whereas modeling and control, in pipelines for applications in urban and rural agriculture that incorporate artificial intelligence into irrigation systems, was the focus of the systematic review in [9]. Finally, in [10] applications based on machine learning for generic water management (in irrigation, human consumption, management by the municipalities, etc.) were systematically reviewed. Unfortunately, there is no paper specifically focused on technological frameworks for smart irrigation. To fill this gap, this paper introduces a schematic representation of smart irrigation architectures and describes the main technologies already exploited with a deep analysis of machine learning models. It also provides a discussion about the open challenges and viable research pathways, making it possible to understand how far we are from having highly effective smart irrigation frameworks which can be 'affordable' to be put in the fields to achieve optimum water usage.

The remainder of the paper is structured as follows: first, in Section 2, the material and methods used to collect works from the literature are pointed out, and common components of smart irrigation systems are defined and described, also providing an architectural taxonomy. Then, in Sections 3–6, the three layers of a generic technological architecture for smart irrigation are described along with related available datasets for benchmarking existing approaches based on machine learning. Open challenges are then discussed in Section 7, whereas new viable research pathways are mentioned in Section 8, starting from the latest machine learning findings. Finally, Section 9 concludes the paper.

2. Material, Methods and Taxonomy

This paper is a narrative review of the recent literature. However, an initial selection of papers was made according to systematic criteria. A coarse selection of papers was made, indeed, starting from queries into the most common scientific databases, i.e., Elsevier Scopus and Google Scholar. Works published since 2019 were considered. In Scopus, the query TITLE-ABS-KEY ((machine AND learning) AND (smart AND irrigation)) AND PUBYEAR > 2018 returned 367 documents. In Google Scholar, the query "machine learning" and "smart irrigation" returned 818 documents. The documents were therefore selected based on their contents and relevance to this paper's narrative paths. The number of citations (weighted on the publication year) was an important criterion for deciding which works deserve mentioning among similar works. High-quality papers dealing with less debated research lines were also considered independently from the number of citations. After reading the picked-up papers, it was possible to obtain a general definition of a smart irrigation system as a set of sensors, Internet of Things (IoT) technologies, and algorithms.

Sensors sense the environment and the soil conditions, IoT adapts and sends the sensed values to local or remote processing units in which trained machine learning algorithms can forecast agrotechnical indicators useful to decide if irrigation is needed and eventually the optimal amount of water to be provided to the crops. A generic smart irrigation system can be then schematically represented (icons were taken from <https://thenounproject.com>, access on 3 April 2024) as in Figure 1, where it is possible to find the following:

1. A physical layer: it consists of sensors, actuators, processing and storage units interconnected through a communication network.
2. A processing layer: it consists of algorithms used to analyze available knowledge, location, crops, and to provide outcomes depending on the requirement of the decision layer.
3. Datasets and data sources: historical data locally or remotely stored to improve the processing layer's ability to model the problem defined by the decision layer.
4. Decisions layer: the set of services provided to the end user to reach the specific goal. Experts provide specific rules that define actions in response to the processing layer outcomes.

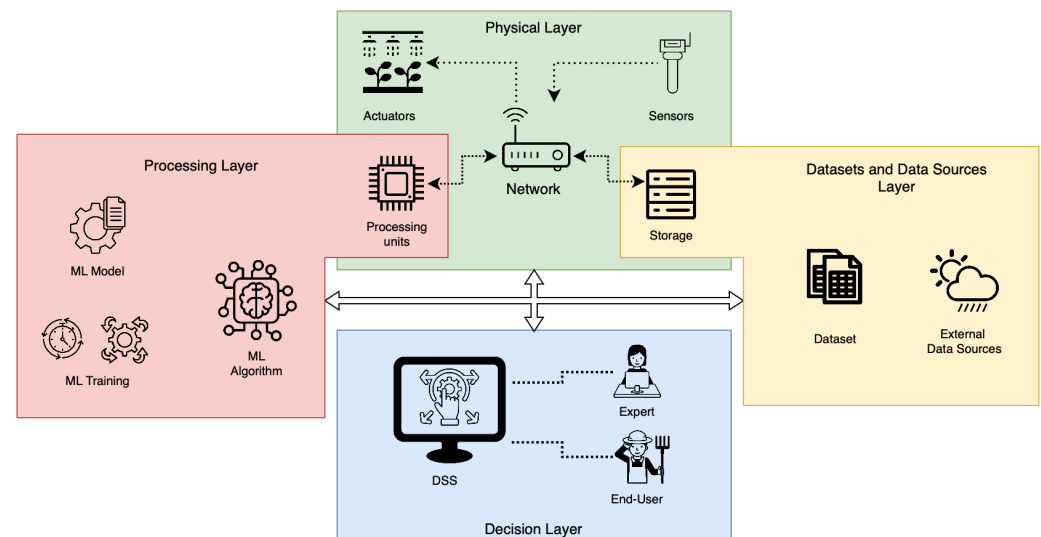


Figure 1. Schema of a generic smart irrigation system: *physical layer* is responsible for retrieving on-field data and operating actions on-field. *Processing units* and *storage*, even if they are part of the physical layer, serve, respectively, the *processing layer* and the *datasets and data sources layer*, resulting in a conceptual superimposition of these layers. The decision layer represents the interface with the expert and end user and communicates with all the other layers.

Each layer will be detailed in the following sections, with a deeper focus on machine learning algorithms and the architectures involved.

3. The Physical Layer

The physical layer comprises locally or remotely situated sensors, actuators, and processing units coupled with a communication network. Physical sensors, such as those that measure temperature, humidity, and soil moisture, are used for monitoring the plants, the soil, and the weather. Irrigation scheduling extensively uses soil moisture sensors, which measure either the soil water potential or the soil water content. Sensors with small footprints that are positioned at various depths can be used to record soil moisture dynamics, improve the accuracy of the measurements, and better understand how crop water usage and irrigation affect variations in soil water content. A wide range of information on the soil's physical, chemical, and mechanical characteristics can be obtained via soil sensors using optical, radiometric, mechanical, acoustic, electrical, electromagnetic, pneumatic, or electrochemical measurements.

Microcontrollers built by open-source electronic prototyping platforms such as ARDUINO (www.arduino.cc/, accessed on 19 May 2024), or even more simplified boards (<https://www.cytron.io/p-maker-uno-simplifying-arduino-for-education>, accessed on 19 May 2024), are commonly used to sample signals coming from analogue sensors.

In a real-time system, one of the crucial issues is the reliability of the physical sensor nodes [11]. A defective sensor node has a significant influence on any real-time system. It is quite challenging, and sometimes impossible, to determine if the physical sensor is communicating accurate values or failing because of outside disturbances like noise. Physical nodes that are not working properly can only be found by manual examination, which is a time-consuming and arduous task. In a smart irrigation system, erroneous values transmitted to the farmer will have a detrimental effect on the system's overall dependability and public perception of the system. Reading wrong data can lead the model to make wrong decisions and, for this reason, fault detection and isolation algorithms are required, but their study is in early infancy [12]

The use of remotely located sensors based on imagery acquired by Unmanned Aerial Vehicles (UAVs) represents a possible alternative to obtaining accurate information about soil and crops, overcoming in this way the aforementioned limitations of physical sensors.

UAV-based methods collect imagery and try to train models by converting the data from the imagery into a variable that has a strong correlation with existing ground measurements [13]. The data collected by UAVs have a high spatio-temporal resolution and can be effectively used to infer soil moisture conditions and crop growth indicators. Several studies in this area have explored the near-infrared and visible bands [14]. Also, the surface soil moisture was proven to be significantly correlated with the brightness of UAV visible images [15]. On the other hand, since transpiration uses a lot of energy and reduces the surface temperature of leaves and vegetation linearly, thermal imaging is especially well suited for the early detection of drought stress [16]. Recent studies have furthermore shown that integrating multisource data can be effectively used for remote soil moisture estimation [17]. Crop health is characterized instead by vegetation indices, which are algebraic combinations of reflectance data acquired from a multispectral sensor [18]. Finally, evapotranspiration, the value that characterizes the loss of water from the soil due to both evaporation from the soil surface and transpiration from the leaves of the plants growing on it, can be also estimated from thermal and multispectral cameras mounted on UAVs [19]. Actuators are additional physical layer components that operate by following directions from the control system. They act on drip irrigation systems, sprinklers, pumps, valves, or sprinkler systems to water the plants as needed based on the findings of the analysis. To ensure that all the components have dependable power sources, power supply facilities are also necessary. Combining mains power, batteries, solar panels, or other renewable energy sources can accomplish this [20].

Some main requirements for the physical components (usually using batteries and requiring having a long working life) are low-energy demand and cheapness (due to the large number of nodes, small additional costs in one node cause large extra expenditures in the overall cost of the system).

Samples from sensors can be processed by a code running on a central control unit (e.g., a workstation), and can be sent to the cloud [21] or directly processed on edge processing units (e.g., a Raspberry Pi (www.raspberrypi.org/), accessed on 19 May 2024), according to the Fog Computing paradigm [22,23].

A communication network must ensure reliability in connecting sensors, actuators, and processing units.

LAN, Wi-Fi, cellular networks, and other wireless communication protocols are only some of the technologies which can be used to guarantee a constant connection between the system's components. The Global System for Mobile Communications (GSM) as the primary controller device, ZigBee (<https://csa-iot.org/all-solutions/zigbee/>, accessed on 19 May 2024) as a data transmission technology, and MQTT (<https://mqtt.org/>, accessed on 19 May 2024) as a messaging protocol are the most suitable, effective, and advantageous wireless technology solutions [24]. Another element of these systems is an application programming interface (API),

which functions as a software bridge between apps and allows users to programmatically interact with sensors or data (e.g., to visualize them). Node-RED and other open-source, low-code alternatives can be used to accomplish this task: for instance InfluxDB (<https://www.influxdata.com/>, accessed on 19 May 2024) to extract insights from time series data, and Node-RED (<https://nodered.org/>, accessed on 19 May 2024) for data queries. Alternative services can be established through Azure IoT (<https://azure.microsoft.com/>, accessed on 19 May 2024) or Amazon Web Services (<https://aws.amazon.com/>, accessed on 19 May 2024), which employ an optional subscription-based model [25].

4. The Processing Layer

In the processing layer, data coming from the physical layer and from available datasets are processed to make predictions which contribute to the decisions about irrigation scheduling. Predictions can lean on predefined rules and formal logic as usually done by traditional ML methods, which are also referred to as symbolic methods. Generally, they need feature engineering, which is the process of manually choosing, extracting, and weighing characteristics from unprocessed data. The underlying models need a small amount of processing power and a finite number of parameters to train.

Alternatively, deep learning enables computers to process data in a manner that mimics the functionality of the human brain. Deep nets do not undertake the process of how the data are made; indeed, they just rely on extensive data to learn how to transform inputs into outputs using numerous rules [26]. Most of the existing machine learning techniques were used to build smart irrigation systems as well. In the following subsections, the techniques used for extracting knowledge at the processing layer in smart irrigation architectures will be presented and discussed.

4.1. Traditional Machine Learning Methods

Based on their characteristics, traditional models—also known as shallow, non-deep models—can be categorized as either parametric or non-parametric [27]. A parametric model is a learning model that, regardless of the quantity of training instances, summarizes data using a fixed set of parameters. In other words, a parametric model assumes a certain data structure, and it will not alter its estimation of the number of parameters it requires no matter how much data it is fed. Conversely, non-parametric machine learning algorithms are those that do not make any strong assumptions about the mapping function's form. They can take any functional form from the training data and learn them without making any assumptions. They work very well when a large amount of data are available with no reliable prior information. Traditional parametric and non-parametric approaches found in the literature with specific applications in smart irrigation are discussed in the following, and they are summarized in Table 1. For each system, the machine learning method, the actual input and estimated output values and performance are reported, providing a quantification of the empirical assessment of the systems.

Linear, polynomial, and logistic regression are some examples of parametric machine learning methods used for smart irrigation. Regression is one of the data mining techniques used to forecast the amount of water required for the next irrigation. Regression can model linear or arbitrary relationships between input and output variables. In the case of non-numerical output (categorical), the methods are reported as logistic regression.

The regression analysis has been widely applied to predict soil moisture, specifically, the amount of water held in the soil in the root zone of the plant, given measured values like ambient temperature and irrigation volume [28]. Methods using the Bayesian theorem for inference are reported as Bayesian models. They estimate posterior probabilities by Bayesian inference using previous information in the form of a prior distribution. This approach has been exploited in [29] to model posterior probability over time of the crop coefficient constrained to radiation. Only three sensors' worth of data—temperature, global radiation, and crop weight—were used as input.

Fuzzy logic helps in modeling situations of the real world which are so complex that they cannot easily be modeled by common binary outputs, e.g., true or false. In those cases, the fuzzy logic proves to be helpful and flexible, and it uses some degrees of truth instead. This approach was used in [30] to manage irrigating through a controlled valve by using soil moisture and temperature as input and to calculate the amount of water to be released.

Among non-parametric machine learning methods, the k-nearest neighbors (kNN) algorithm is a quite simple algorithm widely used for classification and regression. K-nearest neighbors stores all available cases and ranks new cases based on a similarity measure. In [31], it has been demonstrated (K = 3) to be highly effective in a binary classification problem that, starting from multiple sensor inputs (soil humidity sensor, temperature and humidity sensor, and rain sensors), must decide either to deactivate or activate a pump.

Table 1. Most relevant smart irrigation systems relying on traditional machine learning algorithms. Acronyms: ET = evapotranspiration ; (R)MSE = (root) mean squared error; acc = number of correct predictions/number of predictions made.

Works	Method	Inputs (Actual)	Outputs (Estimated)	Performance
[28]	logistic regression	soil moisture weather data type of crop	Irrigation needed Probability	NA
[29]	Bayesian breakModels	temperature global radiation crop weight	ET	RMSE = 0.11
[30]	fuzzy controller	soil moisture air temperature	required amount of water	NA
[31]	K-Nearest Neighbors	soil humidity air temperature air humidity rain	pump act/deact	acc = 0.98
[32]	Decision trees	crop types soil moisture weather conditions weather forecast soil water profile	soil moisture	RMSE = 0.48
[33]	gradient boosting + kmeans	Weather Forecasting Soil Moisture soil temperature Light Radiation Temperature Humidity	soil moisture	acc = 0.97 MSE = 0.20
[34]	gradient boosting regression	weather data	ET	RMSE = 0.16
[35]	K-means + SVM	soil moisture Humidity Temperature Pressure Luminosity	turn on/off sprinkler	acc = 0.98
[36]	Artificial Neural Network	soil moisture air Humidity air Temperature	pump act/deact	acc = 0.97
[37]	Random Forest	Soil moisture	pump act/deact	acc = 0.98

One kind of hierarchical data structure that employs the divide-and-conquer tactic is a decision tree. Due to its high precision and low computation costs, it is widely used. Decisions are made using nodes, and the output or result is represented by leaf nodes. Because of the flexible tree topology, the decision tree model is non-parametric. It expands following data and complexity issues. In [38], information on the soil's moisture content, the humidity level at the moment, and the weather forecast were processed to make decisions about how long irrigation should last and how much water should be used. Categorical data were used by a decision tree to make decisions. A farmer with experience provided sample data needed for the decision tree.

Similarly, in [37], a random forest, i.e., a collection of random decision trees, was used to decide to irrigate or not based on measures of water level of the soil.

Often, decision trees serve as the foundation for gradient boosting approaches, which are machine learning methods that provide a prediction model in the form of an ensemble of weak prediction models—that is, models that, like decision trees, make relatively few assumptions about the data.

An effective example of a boosting machine for smart irrigation can be found in [32]: the final goal was to estimate soil moisture starting from crop types, actual and historical soil moisture values, weather conditions (from in-farm weather stations), and weather forecast data coming from the cloud and soil water profile that is modeled as a feature as well. In [33], it was proved that a combination of gradient boosting techniques for the regression/prediction of soil moisture and k-means for the classification of the irrigation-needed grade can outperform other algorithmic combinations on an edge computing model. Gradient boosting algorithms were also applied in [34] for predicting evapotranspiration from weather data and thereby aiding in irrigation planning.

Other well-known traditional machine learning strategies leverage Support Vector Machines (SVMs). By creating a hyperplane that maximizes the distances separating the data, SVMs are used to divide the data into linear and nonlinear categories. The points that comprise the support vector can be examined to determine the exact location and orientation of the hyperplanes.

In [35], sensors (soil moisture, humidity, temperature, pressure, and luminosity) are used to collect soil and surrounding data. At the beginning, features are selected by a correlation-based criterion and clustered by the K-means algorithm, keeping, in this way, similar data together. The classification model is built using Support Vector Machines, and the binary output allows to decide to turn on/off the sprinkler according to the supervised set of training samples on rise crops.

Neural networks (NNs) are one of the most widely used machine learning models for data processing. Neural networks are collections of algorithms that attempt to identify the relationships between two or more pieces of data in a way that mimics the way the human brain operates. The structure of a neural network is similar to that of the biological brain, with neurons receiving inputs with weights that can be adjusted along the edges of the incoming data. The most predictive model often uses the activation function, sigmoid, or Relu, to represent the non-linearity of the data. They are used also for smart irrigation purposes.

As an example, a three-stage network trained using the Levenberg–Marquardt algorithm was employed in [36]. The input sensors provided the first stage's three primary parameters, which were air temperature, soil moisture, and humidity. Ten nodes in the hidden layer, which made up the middle stage, generated the values of learned neural weights that were needed to assist in making the ultimate decision about how to design an irrigation system. The output layer, the last stage, oversees the decision of whether to turn the water pump motor on or off. Finally, various shallow classifiers (KNN, logistic regression, neural networks, SVM, and Naïve Bayes) were compared in [31] concerning categorical soil moisture, temperature, and humidity data. A value of "0" indicated that pumping had to be stopped, while a value of "1" indicated that pumping had to be activated.

4.2. Deep Learning Methods

Deep learning offers several advantages over traditional machine learning, such as the ability to learn from raw data without much preprocessing, capture complex and nonlinear relationships, scale well with large and diverse datasets, and perform well in domains where human expertise is limited. Additionally, its performance can be improved with more data and computational power. In this subsection, the most relevant deep architectures exploited to aid in irrigation processes are reported, and Table 2 summarizes them. For each system, the machine learning strategy, the actual input and estimated output values, and performance are reported, providing a quantification of the empirical assessment of the systems.

Table 2. Most relevant works relying on deep learning architectures. In table: LSTM = Long Short-Term Memory, Bid-LSTM = bidirectional LSTM, CNN = convolutional neural networks, Mask-RCNN = mask region-based convolutional neural network, GNN = Graph Neural Networks, GWL = groundwater level. Acronyms: MAE = mean absolute error; RMSE = root mean square error; acc = number of correct predictions/number of predictions done.

Works	Method	Input (Actual)	Output (Estimated)	Performance
[39]	LSTM	temperature, humidity, and soil moisture	temperature, humidity, soil moisture	RMSE = 2.35%
[40]	CNN	field's images	soil moisture	MAE = 1.44% RMSE = 2.74%
[41]	CNN	field's images	soil moisture	RMSE = 2.01%
[42]	Mask-RCNN	field's images	soil moisture	Validation loss = 0.8
[43]	CNN	field's images	soil moisture	acc = 0.97
[44]	TFT	Multivariate environmental Sources	soil moisture	MAE = 2.75 RMSE = 3.34
[45]	Bid-LSTM	air temperature, air humidity, wind speed, precipitation data, Soil moisture, electrical conductivity, temperature	soil moisture, electrical conductivity	MAE = [0.79, 4.32] RMSE = [1.41, 5.03], MAE = [0.68, 4.38] RMSE = [1.12, 6.52]
[46]	GNN	Ground Water	Ground Water	MAE = 0.67 RMSE = 1.14
[47]	ResNet GoogleNet	Images	Texture-water class	acc = 0.99
[48]	LSTM	volumetric soil moisture, soil temperature, climate data, and rainfall)	volumetric soil moisture	RMSE = 1.2%

Deep neural networks that can work with variable length sequences and do not require fixed-size time windows are known as recurrent neural networks (RNNs) [49]. This makes them an effective method for handling sequence-dependent data. The hidden neurons are gradually coupled by feedback loops in these densely connected networks. A state

vector found in the hidden units retains the memory of every element in a sequence that came before it. RNNs may learn long-term dependencies across numerous time steps and generalize across input sequences rather than just particular patterns thanks to their internal memory. For this reason, RNNs are a good fit for a variety of natural language-processing applications, including forecasting and time series analysis. It makes precise projections of the future possible by utilizing time series data patterns.

Long Short-Term Memory (LSTMs) networks [50] refer to a particular type of RNN architecture that can sustain long-lived dependencies. The basic concept behind LSTM networks is a memory cell that stores the time state of the network across several time steps (more than 1000). Non-linear gateways control the flow of information by selectively storing and removing information in this memory cell. The use of LSTM-based neural networks was exploited in [39] to obtain predictive values of temperature, humidity and soil moisture. The Bid-LSTM network is a deep bidirectional long short-term memory network that processes sequential input in both the forward and backward directions. By fusing bidirectional processing with the power of LSTM, the model can represent the input sequence's past and future context. It was utilized in [45] to enhance forecasts of soil electrical conductivity and moisture (SM), offering a useful guide for fertilization and irrigation. In [48], an LSTM was also employed to forecast the volumetric soil moisture content for a given day, the duration of irrigation, and the spatial distribution of water needed to irrigate the field.

One type of deep learning technique that is particularly helpful for computer vision applications like image recognition and classification is convolutional neural networks. Through the capture of critical features in early layers and intricate patterns in subsequent levels, they are intended to understand the spatial hierarchy of features. They have been exploited for smart irrigation to estimate soil moisture from the in situ field or remotely (from satellite [40], Unmanned Aerial Vehicle [42] or airborne [41] acquired images [43]). Some works make use of CNNs to identify probable plant illnesses associated with irrigation systems [51]. Conversely, traditional CNN architectures (specifically, GoogleNet and ResNet) have demonstrated significant potential in identifying irrigation needs for agricultural fields with different soil texture classes and lighting in soil frames segmented in close-up photos [47].

To address the issue of neural machine translation, or sequence transduction, transformer designs were created. This refers to any activity that converts an input sequence into an output sequence [52]. They are based on the encoder–decoder strategy and have been largely exploited for speech recognition and text-to-speech transformation.

Modern time series forecasting models like the temporal fusion transformer [53] are based on the architecture of classical transformers. Their strength lies in their efficient architecture and their capacity to incorporate static (i.e., time-invariant) variables. Interpretability can be maintained by using temporal fusion transformers to simulate intricate relationships between various temporal (historic or future) and static inputs. Specifically, in [44], they have been effectively compared with LSTM and other state-of-the-art techniques for soil moisture forecasting in smart agriculture with diverse multivariate, local, and non-local sources.

Graph Neural Networks (GNNs) represent a new and rapidly expanding class of neural network models that can represent intricate interactions within sensor topologies and have shown state-of-the-art performance in multiple IoT learning tasks [54]. A class of deep learning techniques called Graph Neural Networks (GNNs) is made to make inferences on graph-described data. In [46], for example, they have been used recently to accurately estimate groundwater levels, where each well is represented as a node in the graph.

5. The Datasets

The above-discussed methodologies highlighted as reliable soil moisture forecasts mainly rely on the analysis of multivariate time series and the correlation between the

stations of a sensor network. Including additional parameters like soil temperature, soil composition, underground water availability, current and forecast weather conditions, and exploiting the mutual influence of different areas in a field under monitoring enables a better understanding of the complex dynamics driving soil moisture variations. In particular, air humidity, air temperature, and of course, rain events directly influence the amount of ingoing and outgoing soil water.

Unfortunately, the collection of such data requires multiple expensive sensors that frequently occur in failures, making the collection of desired data challenging.

In the next paragraph, the main datasets and data sources concerning the aforementioned data will be presented, discussing their main characteristics and their availability.

5.1. Soil Moisture

Soil moisture represents the water content of the soil. It is usually expressed in terms of volume or weight, and its measurement can be made by means of in situ probes, remote sensing, or model-based approaches. Water that enters a field during rain events or irrigation is removed by evaporation, transpiration, runoff, and drainage [55]. The amount of the water that evaporates into the atmosphere directly from the field's surface is known as *evaporative water loss*, whereas *transpiration* refers to the amount of water that passes into the atmosphere from the plant itself. Runoff and drainage are regarding, respectively, the water that flows on the surface to the edge of the field and the water that flows through the soil downward.

Soil moisture products include all the solutions devoted to systematically measuring the soil water content. These products can be broadly classified into three categories: remote sensing-based products, model simulations-based products, and in situ measurements-based products [56].

Remotely sensed products that retrieve data from active/passive satellite microwave observations are characterized by a strong sensitivity to soil moisture, the capacity to observe, at the same time, the weather context, and the advantages of short-term access. For these reasons, they have become a valuable way to estimate soil moisture at large scales. Among remote sensing-based solutions, it is worth mentioning the merged approaches that estimate soil moisture by blending multiple separately released microwave remotely sensed products [57]. Unfortunately, space-borne microwave instruments for soil moisture retrievals usually work on the coarse spatial resolution of $\sim 10^1$ – $\sim 10^3$ km², and the retrieval quality is affected by multiple spatially and temporally variable factors, such as weather conditions and land cover conditions [58]. Also, the retrieval algorithm and the instrument characteristics affect the reliability of the measurement [59]. These reasons make remote sensing suitable for monitoring soil moisture at the meteorological scale of spatial variability but almost useless for precision agriculture applications.

Model-derived soil moisture represents a way to estimate the soil moisture that mainly relies on a mathematical model marginally driven by measurements and includes, among others, NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA), European Center for Medium-Range Weather Forecasts (ECMWF), and NASA's Global Land Data Assimilation System (GLDAS). Unfortunately, such an approach does not provide the accuracy required by a reliable water management system.

On the other side, ground-based sensor networks show more reliable characteristics and can be modeled both on large and small scales depending on the purposes. The calibration and validation of models devoted to soil moisture forecast, as well as the validation of remote sensing-based products, have primarily relied on large-scale networks. Small-scale networks best fit the needs of smart agriculture. The first attempt devoted to offering a unique access point for multiple, globally available ground-based soil moisture datasets was the Global Soil Moisture Data Bank (GSMDB) [60,61]. Another attempt, the first international initiative, regards FLUXNET, which includes the monitoring of land-atmosphere exchanges of carbon (C), energy, and water but unfortunately does not include soil moisture measurements in all sites [62]. In recent years, the most reliable

solutions devoted to the collection and sharing of this kind of data have provided research and industry with a huge amount of data. (See Table 3). In the next paragraphs, we will discuss the most relevant datasets selected depending on their size, data reliability, and completeness. Finally, a discussion on the control of data quality will be provided.

Table 3. Soil moisture datasets: a summary of the most relevant soil moisture dataset. ISMN data have variable coverage, time interval, sensor depth, and weather availability since they come from several networks that are spread worldwide.

Dataset	Coverage	No. of Site	Time Interval	Sensors Depth	Weather	Additional Data
ISMN [63,64] https://ismn.earth/en/ Last access: 19 May 2024	Variable	~450	Variable	Variable	Variable	Variable
CAF [65] https://goo.gl/JYAIT3 Last access: 19 May 2024	<1 Km ²	42	2007–2016	up to 150 cm	Yes	Field Info
MSMMN [66] https://www.oznet.org.au Last access: 19 May 2024	82,000 Km ²	38	2001–current	up to 90 cm	Yes	No

5.1.1. ISMN

The International Soil Moisture Network (ISMN) (<https://ismn.earth/en/>—Last Access: 19 May 2024) [63,64] was born from the cooperation of the *Global Energy and Water Cycle Experiment*, the *Group of Earth Observations*, the *Committee on Earth Observation Satellites* and the financial effort of the European Space Agency (ESA). The ISMN represents a data lake for in situ soil moisture networks data devoted to overcoming the issues related to data accessibility, timeliness in data delivery, and data heterogeneity. ISMN does this by collecting and harmonizing soil moisture datasets from 80 worldwide distributed networks (consisting of more than 3000 sites) and making them available through a centralized data portal. All individual stations measure soil moisture, but many of them also provide measurements of soil and air temperature, soil suction, precipitation, snow water equivalent depth, and surface temperature, enabling complex analyses of soil moisture dynamics. Other than that, it is worth noting that most other networks, before being shared with the ISMN, first undergo extensive data quality inspection [67].

It is finally worth mentioning that ISMN also provides additional external data regarding, soil characteristics, consistent climate and land cover. More precisely, soil information is obtained from the Harmonized World Soil Database (HWSD) with a 30" resolution, land cover is retrieved from ESA's Climate Change Initiative with a 300 m resolution, and finally the Köppen-Geiger database provides climate classification.

Among the available networks, we can mention the one located within the Shandian River basin, referred to as SMD-SDR. Such a network is designed to match multiple scales and, consequently, multiple needs spreading from the application of smart irrigation algorithms to the assessment of remote SM products. More precisely, network stations are distributed in a three-sampling-scale nested structure (100 km, 50 km and 10 km). Each station is equipped with five sensors at different measuring depths (3, 5, 10, 20, and 50 cm) and with a sampling interval varying between 10 and 15 min. Out of the 34 stations, 20 are equipped with the HOBO rain gauge, mostly at small scale and medium scale. The observation period of this study ranges from September 2018 to December 2020.

5.1.2. CAF Dataset

The Cook Agronomy Farm (CAF) dataset [65] provides a 9-year record of hourly soil moisture of 42 stations located in the 37 ha of the experimental R.J. The CAF dataset represents an example of a soil moisture dataset explicitly designed for smart agriculture

applications. The CAF dataset aims to represent a no-tillage annual cropping system in regional dryland. More precisely, the dataset was obtained by dividing the farm into three experimental sub-fields. Additionally, each sub-field was divided into wide strips devoted to account for the crop rotation.

The 42 stations were installed in 2 successive campaigns (12 in 2007 and 230 in 2009), and each of them monitors soil water content at five depths—0.3, 0.6, 0.9, 1.2, and 1.5 m—for a total of 210 sensors. Locations have been carefully chosen to maximize variability in several relevant parameters (e.g., elevation, slope, and insolation). The 42 sensor sites span lag distances between 60 m and 905 m. The dataset provides 14,127,840 hourly readings; out of these, 5,121,639 are missing a water content reading, and 5,070,945 are missing a temperature reading. Each specific sensor shows a missing record rate between 9% and 79% with an average value of 35% across all sensors. Volumetric water content values less than 0 or greater than 1 (sensor failure) have been replaced with *NA* values.

The dataset provides data regarding the cropping history of each sub-field and strip for each year, a digital elevation model, particle size, and the bulk density to 1.5 m depth at each of the 42 instrumented locations. Finally, meteorological data are collected by an on-site station recording values like air temperature, relative humidity, dew point temperature, soil temperature, rainfall, wind speed and direction, solar radiation, and leaf wetness at 2 m of height. The dataset is publicly available at the dataset website [<https://goo.gl/JYAIT3>].

5.1.3. MSMMN Dataset

The Murrumbidgee Soil Moisture Monitoring Network (MSMMN) [66] dataset is a soil moisture dataset from the 82,000 km² Murrumbidgee River Catchment in Australia. The sensor network includes 38 soil moisture-monitoring sites across the Murrumbidgee Catchment that are mainly concentrated in three subareas.

The network inception started in 2001 and counts two different generations of sensors. Soil moisture sensors were installed in the upper 90 cm of the profile (0–30, 30–60, and 60–90 cm), and measures are sampled every 5 or 60 s and then averaged to 30 or 20-minute measurements depending on the specific site. Furthermore, other sensors complete the data, providing information about soil temperature and precipitation. Additional weather data like air temperature, air pressure, relative humidity, wind speed, and downward short-wave and long-wave radiation are available for all sites.

In order to minimize the amount of corrupted data, measurements from all soil moisture monitoring sites have been carefully inspected to identify and remove errors or frag missing data.

5.1.4. Data Quality Control and Interpretation

In the discussion above, we defined a network as a set of stations managed by an entity (e.g., organizations and partnership), where the number of stations ranges between one and several hundred, and each network has been designed for different applications, uses different sensors, and consequently shows different characteristics and features. Other than that, depending on the specific area of the globe where the network is located, the soil moisture regime varies dramatically. Even specific meteorological events can cause specific soil moisture dynamics that have to be specifically considered. In other words, although it is unrealistic to consider a single typical behavior of logged soil moisture, most of the readings have several features in common, and their understanding is fundamental for detecting errors in soil moisture measurements. Indeed, most of the logged raw data contain errors such as outliers, breaks, signal saturation, or missing data due to unresponsive sensors. Finally, it is mandatory to check the changes in soil moisture as a function of changes in the soil temperature and precipitation.

Among the specific behaviors, the fluctuation related to temperature is one of the most significant. Most of the soil moisture probes show a pronounced sensitivity to temperature because of the positive relationship between electric conductivity and temperature. This effect leads to artificial diurnal soil moisture fluctuations that are more evident in the upper

soil layers where the temperature fluctuation is greater [67]. Another phenomenon related to the temperature regards frozen soils and the cyclic behavior of freeze–thaw events. In the first cases, a lower value of soil water content is recorded because of the significantly lower dielectric conductivity (reference of most of the SM probes) of ice than liquid water. For the same reason, even in the absence of rain or irrigation events, there will be alternating low and high levels of soil moisture records during thawing (during the day) and freezing (at night).

Common sources of error, such as random noise, which is an intrinsic component of any measurement, as well as the previously listed outliers, breaks, signal saturation, and unresponsive sensors, all affect the quality of data. Spikes usually come out from temporary sensor failure or reduced current supply. On the other hand, brakes usually cause a semi-permanent offset of measurement value, where the main difficulty is represented by the understanding of the truth represented by the period before or after. The source of brakes (jumps or drops) can also be caused by a current supply reduction but can also be associated with real sudden changes in environmental conditions. Constant values can also affect the data reliability and are usually caused by values exceeding the upper limit of sensitivity of the sensor (high signal) or longer sensor dropouts (low signal). Finally, erroneous calibration could cause systematic biases, whereas instrument drift (a gradual systematic change over time caused by oxidation of the sensor rods, salinization, increasing soil compaction, etc.) shows non-existing changes in climatologist conditions.

An estimate of the station data quality, as well as the understanding of specific soil moisture dynamics and their meaning, is of particular interest to those researcher who wish to train algorithms and models devoted to forecasting soil moisture trends to drive water management policy toward keeping a constant and high level of health for soil and crops.

5.2. Weather

Most of the above-mentioned datasets do not come with meteorological data that, anyway, represent pivotal information in approaches exploiting multivariate time series analysis for soil moisture forecasting and smart irrigation purposes. Indeed, weather, as stated above, is the main factor driving the soil water content in addition to irrigation. More precisely, historical data can be exploited for training and evaluation purposes and allow the enrichment of soil moisture datasets, whereas forecasts are highly useful at the inference step to obtain reliable predictions of soil moisture trends. There exists a wide set of online services providing both historical and current weather conditions as well as weather forecasts. Many of them are completely free for non-commercial use, others allow a limit on the number of requests, and others provide some specific services (historical data or long-term forecasts) only for payment plans.

5.2.1. Open-Meteo

Open-Meteo (<https://open-meteo.com>—Last Access: 19 May 2024) is an open-source weather platform offering free access for non-commercial use. It offers 80 years of historical weather data: hourly weather forecasts for up to 16 days, global weather models and regional models with respectively up to 11 km and 1.5 km resolution, and weather model updates every hour for both Europe and North America.

Among the high-resolution local and global weather models integrated in Open-Meteo, we can cite NOAA GFS with HRRR, DWD ICON, MeteoFrance Arome&Arpege, ECMWF IFS, JMA, GEM HRDPS, and MET Norway. Such models are free to download but challenging to process (they require expertise in grid systems, projections, and the fundamentals of weather predictions). Open-Meteo takes care of the daily download and processing of these models, making them easily available using simple API requests.

The API is available at no cost for non-commercial use and ensures top-notch forecast accuracy. The API utilizes a vast set of local weather models with frequent updates, ensuring that the most precise forecast is generated for any location globally. The API is

provided as a REST service but wrapping in Python, Typescript and Swift is also available, making their integration in most scientific environments easily affordable.

5.2.2. OpenWeather Map

OpenWeather (<https://openweathermap.org/>—Last Access: 23 May 2024) provides local minute forecasts, historical data, current state, and short-term and even annual forecast weather data. Their weather products are available using reliable APIs (provided as REST service) that follow industry standards and are compatible with different kinds of enterprise systems. Historical weather goes back for one month to more than 40 years, depending on the subscription plan. Additionally, they offer customized weather maps with fifteen weather layers that show forecast, historical, and present weather information as well as global precipitation maps created using satellite and radar data.

OpenWeather provides different pricing plans; at the time of writing, the free one allows 1,000,000 calls for months, on current weather and short-term forecasts, whereas the fee plans include additional services, including historical data.

5.3. The Missing Dataset

The described panorama shows that, despite the huge number of available datasets, just a few of them fit the scale of the “smart irrigation” problem and just the CAF dataset [65] is made with a well-structured experiment setup and provides information related to agriculture practices and type of crops and fields. Unfortunately, even this one lacks data related to irrigation, quality/quantity of crops, and additional details. Other than that, it refers to a specific location and small set of crops, limiting in this way the generalization capabilities of the developed models. Finally, weather data availability is an additional limitation. Weather services, as discussed above, are usually based on the combination of several models, and cannot ensure 100% reliable data that, especially in case of precipitation events, play a pivotal role in the interpretation of soil moisture variations. The considerations above make it clear that to enable the deployment of general-purpose smart irrigation systems, the deployment of ad hoc datasets is mandatory. From the perspective of the agricultural context, a desirable dataset should fit the scale of a typical farm field and provide measurements from multiple sites onto the field to make the use of approaches exploiting spatial correlation possible. Weather stations providing, at minimum, information like rain gauge, air temperature, soil temperature, and solar radiation are required to validate and understand soil moisture dynamics in each specific location. The experimental setup should follow the modality of the CAF dataset [65], considering multiple crops and agriculture methodologies in a single location. This way, such aspects should be included as informative parameters driving the decisions of the system. Another key element is the tracking of the irrigation process. A dataset providing data regarding the irrigation instances, irrigation modalities, and the employed amount of water should be pivotal in understanding the response of the field to each specific process. Other than that, considering the application of different irrigation strategies in different sectors of the field would be also useful to optimize the irrigation process as a function of available water reserve, soil moisture forecast and crop health. Finally, this kind of process should be reproduced in several locations world-wide in order to consider different environments, season dynamics, and agriculture trends. A bonus should be specific attention paid to sensor failures in order to minimize missing data, a non-negligible limitation in the application of multivariate time series approaches, especially in the situation relying on spatial correlation. All these characteristics should allow the creation of complete and rich datasets playing the role of a game-changer in the context of smart irrigation.

6. Decisions Layer

A smart irrigation system tries to imitate a human expert in helping farmers with making decisions on irrigation optimization.

If the learning process has been carried out on reliable data for the considered field and crops, i.e., data can be considered fully representative of the complex system made up of the soil, the environment and the plants, the decision can be directly made on the outcomes of the processing layer. In these cases, a thresholding operation, or the putting in place of the provided categorical classification, is enough to make the final decision about irrigation. In the case of thresholding, stored values describing the minimum requirement for the considered observed parameters (e.g., soil moisture) are used as a reference. This allows the system to better fit the conditions of the specific field under observation. In the case of categorical outputs from the processing layer (e.g., irrigation needed yes/no), the actual conditions of the field are ignored. However, it is most common that the training is performed on available data not perfectly fitting the actual conditions in the monitored field, and then a further step becomes necessary. This step is included in an additional decision layer. The decision layer takes as input the numerical outcomes estimated by the processing layer (e.g., forecasted soil moisture) and mixes them with models describing the agricultural processes (e.g., evapotranspiration, plant growth) and additional information/knowledge useful to make the best decision at that specific moment for that specific field to reach the predefined application goal. The integration of the decision layer for an optimal irrigation system could be schematically represented as in Figure 2.

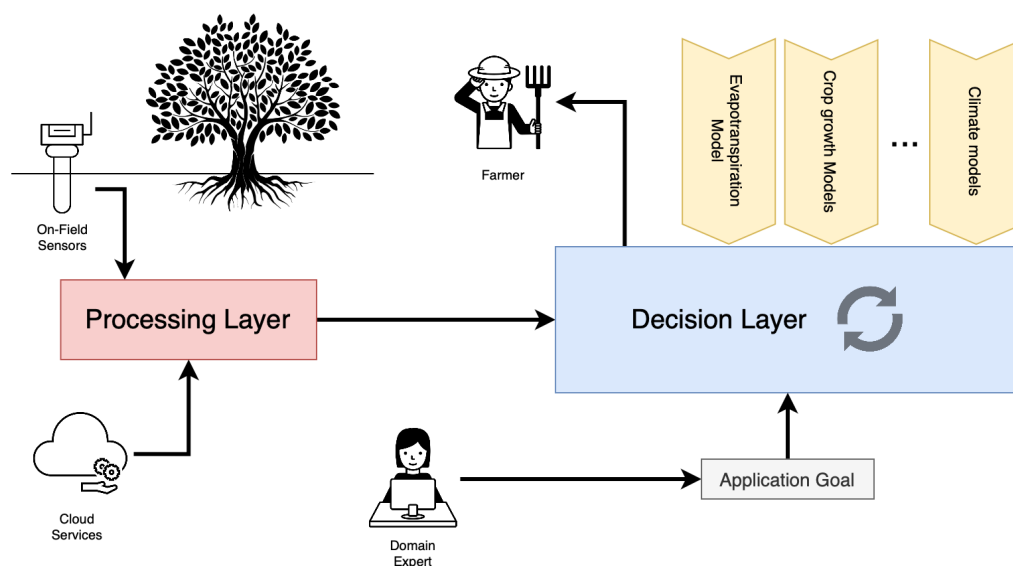


Figure 2. Physical layer exploits outcomes of the Processing layer, the information provided by the models describing the specific environment and, depending on the goals provided by the domain expert, provides the farmer with a proposal on an irrigation plan.

Models describe physical environmental phenomena that occur in the field and its surroundings. In general, rainfall runoff models [68] are the standard tools for investigating hydrological processes. Transient modeling, which enables the evaluation of water evolution (both quality and quantity), as well as the creation of hypothetical scenarios, can be used to integrate the temporal variable into the system and improve decision-making [69]. On the other hand, crop models can also play important roles in setting up decisions for assessing and estimating crop water needs [70].

Additionally, evapotranspiration (ET) models can be used to calculate agricultural water requirements. These models account for the transfer of vaporized water from the land to the atmosphere through plant transpiration and soil evaporation [71]. Crop growth models and climate simulation models can also provide additional important elements that can help in automatically making irrigation decisions [72].

It is increasingly important to use crop and hydrological models in combination to compensate for each other's shortcomings in order to quantitatively characterize many aspects of agricultural production and, ultimately, increase the efficiency of water resource

utilization [73]. The Water Balance Models [74] combine the available observations with the fundamental knowledge to characterize the system's behavior through the application of scientific methodologies. According to Water Balance Models, evapotranspiration, capillary rise, deep percolation, and the surface runoff effect are examples of water outflows that contribute to changes in soil moisture content over time, as water inflows like irrigation and rainfall are [75].

The **application goal** is the motivation for why an intelligent irrigation system is being designed and installed. Common goals include reducing water waste, reducing crop disease, or increasing crop yield [76]. In other words, the application objective defines the cost function to be optimized. Some other important aspects that could help in setting up the decision layer are the groundwater quality indexes, the quantity of water in reservoirs [77], and the amount of human labor required to irrigate [78]. Summing up, the application goals define the cost function to be minimized/maximized by the smart irrigation system, and it can depend on one of several aspects among the aforementioned ones.

After setting up the application goal, the available hydrogeological and functional models, and the eventual additional application aspects, the decision rules can be consequently defined, in general, through logic diagrams.

In this regard, an effective proof of concept can be found in [79] in which information on soil moisture distribution and real-time crop water uptake fed a water-saving scheme for making irrigation decisions.

The use of (deep) reinforcement learning can be an intermediate level of automation of the decision-making task: the goal of reinforcement learning is to find the optimal strategy to maximize the sum of long-term rewards to the agent devoted to making decisions. This interaction between the agent and training environment is repeated until the agent converges to an optimal strategy, for instance, for choosing the next day's irrigation amount [80].

7. Discussion and Open Challenges

From this study, it emerged that there are several issues to be addressed to step forward with reliable and scalable smart irrigation systems. Machine learning models exploited in this research area focus on the prediction of data series values. The main issue related to machine learning is the lack of data to accomplish knowledge extraction in the specific and complex application domains. Available datasets on soil moisture are mostly designed for environmental monitoring and lack most of the information that could drive the design of a complete and reliable smart irrigation system. First of all, the sensor networks scale is usually oriented to monitor large areas and does not fit the standard dimension of agriculture fields. Training machine learning solutions exploiting geographical information on these kinds of datasets does not guarantee their reliability in monitoring small-scale sensor networks in an agricultural context.

Another issue regards data inconsistency. In most of the datasets, the sensor type and the number of sensors change site by site, making the selection of input to the machine learning algorithms challenging. Meteorological data are available on a limited number of sites, making the necessity of online weather services mandatory. Unfortunately, these resources are usually based on satellite data and models that usually do not fit the desired fine grain required by the context. Other than that, no information concerning field management in terms of agriculture practices is provided. Only the CAF dataset [65], which covers a farm field use case, contains information regarding crops in different sub-fields but does not track irrigation sessions, crop health, and water resource availability, information that is pivotal in the design of a complete and reliable decision support system for water management in smart irrigation. The irrigation process is characterized by geographical, seasonal, and agricultural conditions (type of crops, irrigation method, etc.). Collecting domain-specific datasets can lead to overfitting with performance dropping when new unseen data, acquired with different surrounding conditions, are provided.

Unfortunately, transfer learning, domain adaptation and other common tricks in machine learning have not been investigated adequately yet [81].

Even though the benefit of using smart irrigation systems emerged from the state of the art, the performance evaluation (i.e., the quantification of the advantages of using them concerning traditional irrigation approaches) is another critical point: it is often based on numerical comparisons between predicted and actual data values following common error quantification metrics, such as mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE) [82]. It is worth noting that these error metrics may not capture the nuances of forecasting accuracy, especially when dealing with intermittent demand, outliers, or imbalanced datasets. Unfortunately, domain-specific metrics able to assess the effectiveness of decisions about in-field actions (e.g., how much water is necessary concerning the amount of water provided) have not been carried out since it is particularly challenging to obtain data.

The empirical assessment of smart irrigation systems is usually carried out by simulation [83] or, more hopefully, by twin fields [84].

To achieve this objective, digital twin technologies have also been utilized. These technologies enable the automatic bidirectional flow of data between a physical object and its virtual counterpart. In the case of a smart irrigation system, its sensors and actuators are linked to their virtual representations within the digital twin framework [85,86].

Understanding the specific goals and requirements of the smart irrigation domain (water saving, crop health, water quality, groundwater quantity, etc.) and developing or selecting evaluation metrics that are meaningful and relevant to this domain is not trivial and requires strong knowledge sharing. Moreover, incorporating domain knowledge into the evaluation process relies upon interaction among several domain experts who can provide valuable insights into the data, features, and potential biases that may affect the performance of machine learning models in the specific domain. This process leads to some problems of interaction, which can hinder effective collaboration and knowledge sharing. They arise from communication barriers (jargon, terminology differences, and disciplinary silos), and differing perspectives. The key aspects to overcome these drawbacks could be to start domain-aware feature engineering involving domain experts in identifying relevant variables, domain-specific patterns, and meaningful representations of the data. Testing crop growth patterns, for instance, can aid in mitigating soil depletion, while varying weather and soil conditions may influence decisions regarding irrigation.

Strictly related to the definition of affordable metrics is the problem of assessing the reliability of smart irrigation solutions: an acceptable level of error in irrigation scheduling depends on several specific factors, and it is very difficult to determine if an error score is acceptable or not for a specific application. Solutions to the above issues pass through explainable and transparent machine learning models, which can allow domain experts to validate and understand the underlying reasoning. User feedback would hopefully refine the technological solutions, but this has been only simulated so far [87]. Finally, smart irrigation systems should be scalable and adaptable to different farm sizes, crop types, and environmental conditions. They should also integrate with other agricultural technologies and management systems to support holistic farm management.

8. New Research Horizons

By carefully considering these factors and tailoring the evaluation process to the specific application domain, researchers and practitioners can ensure that machine learning techniques are effectively evaluated and deployed to address real-world challenges and opportunities. On the one hand, a feasible way forward is to take advantage of the continuous improvements in machine learning methods. Researchers and practitioners continue to develop new methods and tools to address these challenges and improve the effectiveness of time series forecasting techniques. Below are some examples of current techniques that have not been considered for smart irrigation so far.

Since GNNs operate under the premise that a node's state is influenced by the states of its neighbors, they enable each node in a graph to be aware of the context of its neighborhood by distributing information through structures. Theoretically, these kinds of structures in smart irrigation systems could better identify underlying correlations between variables and also help comprehend the underlying dynamics of values measured on fields in nearby locations. Integrating graph networks with various temporal modeling frameworks (e.g., LSTM) allows the capture of both intra-series (temporal) dependencies and inter-series (spatial) dynamics [88]. Unfortunately, forecasting is severely constrained by distinct spatial and temporal modeling, which inherently contradicts the unified spatiotemporal inter-dependencies in the real world [89]. Introducing hypervariate graphs, such as the Fourier Graph Neural Network (FourierGNN) [90] which views each series value as a graph node regardless of variates or timestamps and represents sliding windows as space-time fully connected graphs, is an intriguing idea to get around the limitations mentioned above. By using graph learning networks, it would be possible to learn the hidden dependencies between variables, enhancing in this way the multivariate time series forecasting [91]. They can handle complicated real-world patterns like those seen in a smart irrigation scenario by describing relationships between variables as static long-term and dynamic short-term patterns, where the short-term patterns reflect the dynamic nature of the multivariate time series.

Transformer-based models (e.g., Crossformers for cross-dimension dependency modeling) have demonstrated significant promise in representing temporal dependencies (cross-time dependencies) as well [92]. However, they fail to account for cross-dimensional dependencies, which are crucial for forecasting in intelligent irrigation applications. The nature of the permutation-invariant self-attention mechanism inevitably results in temporal information loss [93].

Using foundation models may be a practical way to enhance time series forecasting [94]. They are learning models based on the Transformer architecture and its core attention module, with billions of parameters, trained on large amounts of data to enable their use in several use cases. Their latest achievements in computer vision and natural language are quite impressive. Foundation models could be effectively exploited for smart irrigation since they can identify universal patterns within extensive time series data from varied sources, even in the presence of temporal distribution shifts. On the other hand, they require huge amounts of data which must be collected as a preliminary step. This data availability is, at this moment, the main limitation of their exploitation in smart irrigation tasks.

Among foundation models, the Mamba architecture should be explored for the specific smart irrigation domain since it is particularly suited for modeling long sequences, and efficient implementations are also publicly available [95].

The drawbacks of data can be addressed by using powerful generative approaches. Generative models aim to create new, artificial samples based on a dataset of real data by learning the underlying probability distribution. Diffusion models [96] are a class of deep learning-based generative models that have gained popularity recently in advanced machine learning research. They are employed extensively in text, video, and image synthesis due to their exceptional ability to produce samples that closely mirror the observed data. The idea of diffusion has been expanded to time series applications in recent years, and numerous potent models have been created [97]. Their ability to work without constraining the target distributions made them preferable with respect to other generative approaches [96]. In particular, in a smart irrigation system, they could be effectively exploited in the case of sensor failures or missing historical records since they have been well suited for the time series imputation task. Finally, unsupervised systems could also be exploited for anomaly detection in incoming data. They can effectively detect sensor failures, which helps make the right decisions about irrigation. For example, the use of Autoencoders [98] and GANs [99] should be further investigated considering the recent promising results in [100,101].

9. Conclusions

This paper reported a narrative survey of recent works dealing with machine learning architectures for smart irrigation purposes. Throughout the paper, existing strategies have been discussed and open issues have been pointed out, with also prospective research lines to be pursued to build scalable, domain-independent approaches. From the literature, it emerges that significant efforts have been put into building smart architectures by combining ICT and IoT technologies, i.e., by concentrating efforts on the physical layer. On the other hand, the processing level and related datasets require much more effort to reach maturity. The availability of reliable and curated datasets is the key point put in place for the exploitation of recent machine learning approaches, and it could be the breaking point towards a fully automated smart irrigation system that, in the future, should also include the application models into the training process. This will lead to end-to-end machine learning strategies that are completely, fully data-driven. The study paves the way for our next work, which will deal with the implementation of a new machine learning architecture, leveraging up-to-date strategies and datasets collected in different living labs in the Mediterranean area.

Funding: This research was undertaken as part of the project PURECIRCLES: Maximizing resource use efficiency within the water-nutrient-energy nexus for sustainable agriculture in marginal environments” MEL n. 1825, funded under the Notice PRIMA 2022 (Partnership for Research and Innovation in the Mediterranean Area) art. 185 of TFUE, approved by the European Parliament and Council 2017/1324, 4 July 2017. (Project code PRIMA22-00082, MIUR dd n.3153 del 01 03 2023, CUP B83C22009530005).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. FAO. *The State of the World's Land and Water Resources for Food and Agriculture—Systems at Breaking Point*; FAO: Rome, Italy, 2021.
2. Gamal, Y.; Soltan, A.; Said, L.A.; Madian, A.H.; Radwan, A.G. Smart Irrigation Systems: Overview. *IEEE Access* **2023**. [[CrossRef](#)]
3. Bwambale, E.; Abagale, F.K.; Anornu, G.K. Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: A review. *Agric. Water Manag.* **2022**, *260*, 107324. [[CrossRef](#)]
4. Capraro, F.; Tosetti, S.; Rossomando, F.; Mut, V.; Vita Serman, F. Web-based system for the remote monitoring and management of precision irrigation: A case study in an arid region of Argentina. *Sensors* **2018**, *18*, 3847. [[CrossRef](#)]
5. Zamora-Izquierdo, M.A.; Santa, J.; Martínez, J.A.; Martínez, V.; Skarmeta, A.F. Smart farming IoT platform based on edge and cloud computing. *Biosyst. Eng.* **2019**, *177*, 4–17. [[CrossRef](#)]
6. GS Campos, N.; Rocha, A.R.; Gondim, R.; Coelho da Silva, T.L.; Gomes, D.G. Smart & green: An internet-of-things framework for smart irrigation. *Sensors* **2019**, *20*, 190. [[CrossRef](#)]
7. Altalak, M.; Ammad uddin, M.; Alajmi, A.; Rizg, A. Smart agriculture applications using deep learning technologies: A survey. *Appl. Sci.* **2022**, *12*, 5919. [[CrossRef](#)]
8. García, L.; Parra, L.; Jimenez, J.M.; Lloret, J.; Lorenz, P. IoT-based smart irrigation systems: An overview on the recent trends on sensors and IoT systems for irrigation in precision agriculture. *Sensors* **2020**, *20*, 1042. [[CrossRef](#)]
9. Vallejo-Gómez, D.; Osorio, M.; Hincapié, C.A. Smart Irrigation Systems in Agriculture: A Systematic Review. *Agronomy* **2023**, *13*, 342. [[CrossRef](#)]
10. Ahmed, A.A.; Sayed, S.; Abdoulhalik, A.; Moutari, S.; Oyedele, L. Applications of machine learning to water resources management: A review of present status and future opportunities. *J. Clean. Prod.* **2024**, *441*, 140715. [[CrossRef](#)]
11. Hamami, L.; Nasserredine, B. Application of wireless sensor networks in the field of irrigation: A review. *Comput. Electron. Agric.* **2020**, *179*, 105782. [[CrossRef](#)]
12. Jihani, N.; Kabbaj, M.N.; Benbrahim, M. Sensor fault detection and isolation for smart irrigation wireless sensor network based on parity space. *Int. J. Electr. Comput. Eng.* **2023**, *13*, 1463. [[CrossRef](#)]
13. Ahansal, Y.; Bouziani, M.; Yaagoubi, R.; Sebari, I.; Sebari, K.; Kenny, L. Towards smart irrigation: A literature review on the use of geospatial technologies and machine learning in the management of water resources in arboriculture. *Agronomy* **2022**, *12*, 297. [[CrossRef](#)]
14. Zhang, Y.; Han, W.; Zhang, H.; Niu, X.; Shao, G. Evaluating soil moisture content under maize coverage using UAV multimodal data by machine learning algorithms. *J. Hydrol.* **2023**, *617*, 129086. [[CrossRef](#)]
15. Lu, F.; Sun, Y.; Hou, F. Using UAV visible images to estimate the soil moisture of steppe. *Water* **2020**, *12*, 2334. [[CrossRef](#)]
16. Moradi, S.; Bokani, A.; Hassan, J. UAV-based smart agriculture: A review of UAV sensing and applications. In Proceedings of the 2022 32nd International Telecommunication Networks and Applications Conference (ITNAC), Wellington, New Zealand, 30 November–2December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 181–184.

17. Guo, J.; Bai, Q.; Guo, W.; Bu, Z.; Zhang, W. Soil moisture content estimation in winter wheat planting area for multi-source sensing data using CNNR. *Comput. Electron. Agric.* **2022**, *193*, 106670. [[CrossRef](#)]
18. Guan, Y.; Grote, K. Assessing the Potential of UAV-Based Multispectral and Thermal Data to Estimate Soil Water Content Using Geophysical Methods. *Remote Sens.* **2023**, *16*, 61. [[CrossRef](#)]
19. Niu, H.; Hollenbeck, D.; Zhao, T.; Wang, D.; Chen, Y. Evapotranspiration estimation with small UAVs in precision agriculture. *Sensors* **2020**, *20*, 6427. [[CrossRef](#)]
20. Reddy, V.S.; Harivardhagini, S.; Sreelakshmi, G. IoT and Cloud Based Sustainable Smart Irrigation System. *E3S Web Conf.* **2024**, *472*, 01026. [[CrossRef](#)]
21. Abdikadir, N.M.; Abi Hassan, A.; Abdullahi, H.O.; Rashid, R.A. Smart Irrigation System. *Int. J. Electr. Electron. Eng.* **2023**, *10*, 224–234. [[CrossRef](#)]
22. Puliafito, C.; Mingozzi, E.; Longo, F.; Puliafito, A.; Rana, O. Fog computing for the internet of things: A survey. *ACM Trans. Internet Technol. (TOIT)* **2019**, *19*, 1–41. [[CrossRef](#)]
23. Cordeiro, M.; Markert, C.; Araújo, S.S.; Campos, N.G.; Gondim, R.S.; da Silva, T.L.C.; da Rocha, A.R. Towards Smart Farming: Fog-enabled intelligent irrigation system using deep neural networks. *Future Gener. Comput. Syst.* **2022**, *129*, 115–124. [[CrossRef](#)]
24. Motamedi, B.; Villányi, B. Design of a Smart Irrigation using wireless communication protocols in Greenhouse. In Proceedings of the 2022 IEEE 22nd International Symposium on Computational Intelligence and Informatics and 8th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Science and Robotics (CINTI-MACRO), Budapest, Hungary, 21–22 November 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 000179–000184.
25. Yasin, A.; Delaney, J.; Cheng, C.T.; Pang, T.Y. The Design and Implementation of an IoT Sensor-Based Indoor Air Quality Monitoring System Using Off-the-Shelf Devices. *Appl. Sci.* **2022**, *12*, 9450. [[CrossRef](#)]
26. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
27. Gautam, A.; Singh, V. Parametric Versus Non-Parametric Time Series Forecasting Methods: A Review. *J. Eng. Sci. Technol. Rev.* **2020**, *13*, 165–171. [[CrossRef](#)]
28. Mishra, P.; Somkunwar, R.K. Smart Irrigation with Water Level Indicators Using Logistic Regression. In Proceedings of the 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 26–28 May 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5.
29. Kocian, A.; Carmassi, G.; Cela, F.; Chessa, S.; Milazzo, P.; Incrocci, L. IoT based dynamic Bayesian prediction of crop evapotranspiration in soilless cultivations. *Comput. Electron. Agric.* **2023**, *205*, 107608. [[CrossRef](#)]
30. Alomar, B.; Alazzam, A. A smart irrigation system using IoT and fuzzy logic controller. In Proceedings of the 2018 Fifth HCT Information Technology Trends (ITT), Dubai, United Arab Emirates, 28–29 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 175–179.
31. Tace, Y.; Tabaa, M.; Elfilali, S.; Leghris, C.; Bensag, H.; Renault, E. Smart irrigation system based on IoT and machine learning. *Energy Rep.* **2022**, *8*, 1025–1036. [[CrossRef](#)]
32. Togneri, R.; dos Santos, D.F.; Camponogara, G.; Nagano, H.; Custodio, G.; Prati, R.; Fernandes, S.; Kamienski, C. Soil moisture forecast for smart irrigation: The primetime for machine learning. *Expert Syst. Appl.* **2022**, *207*, 117653. [[CrossRef](#)]
33. Premkumar, S.; Sigappi, A. IoT-enabled edge computing model for smart irrigation system. *J. Intell. Syst.* **2022**, *31*, 632–650. [[CrossRef](#)]
34. Ponraj, A.S.; Vigneswaran, T. Daily evapotranspiration prediction using gradient boost regression model for irrigation planning. *J. Supercomput.* **2020**, *76*, 5732–5744. [[CrossRef](#)]
35. Kumar, G.K.; Bangare, M.L.; Bangare, P.M.; Kumar, C.R.; Raj, R.; Arias-González, J.L.; Omarov, B.; Mia, M.S. Internet of things sensors and support vector machine integrated intelligent irrigation system for agriculture industry. *Discov. Sustain.* **2024**, *5*, 6. [[CrossRef](#)]
36. Al-Faydi, S.N.; Al-Talb, H.N. IoT and Artificial Neural Network-Based Water Control for Farming Irrigation System. In Proceedings of the 2022 2nd International Conference on Computing and Machine Intelligence (ICMI), Istanbul, Turkey, 15–16 July 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–5.
37. Chaudhry, R.; Rishiwal, V.; Yadav, P.; Singh, K.R.; Yadav, M. Automatic Smart Irrigation Method for Agriculture Data. In *Towards the Integration of IoT, Cloud and Big Data: Services, Applications and Standards*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 57–73.
38. Patil, C.; Aghav, S.; Sangale, S.; Patil, S.; Aher, J. Smart irrigation using decision tree. In *Proceedings of the International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications: ICMISC 2020*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 737–744.
39. Sami, M.; Khan, S.Q.; Khurram, M.; Farooq, M.U.; Anjum, R.; Aziz, S.; Qureshi, R.; Sadak, F. A deep learning-based sensor modeling for smart irrigation system. *Agronomy* **2022**, *12*, 212. [[CrossRef](#)]
40. Hegazi, E.H.; Yang, L.; Huang, J. A convolutional neural network algorithm for soil moisture prediction from Sentinel-1 SAR images. *Remote Sens.* **2021**, *13*, 4964. [[CrossRef](#)]
41. Seo, M.G.; Shin, H.S.; Tsourdos, A. Soil moisture retrieval from airborne multispectral and infrared images using Convolutional Neural Network. *IFAC-PapersOnLine* **2020**, *53*, 15852–15857. [[CrossRef](#)]

42. Albuquerque, C.K.; Polimante, S.; Torre-Neto, A.; Prati, R.C. Water spray detection for smart irrigation systems with Mask R-CNN and UAV footage. In Proceedings of the 2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), Trento, Italy, 4–6 November 2020; IEEE: Piscataway, NJ, USA, 2020, pp. 236–240.
43. Singh, N.; Ajaykumar, K.; Dhruw, L.; Choudhury, B. Optimization of irrigation timing for sprinkler irrigation system using convolutional neural network-based mobile application for sustainable agriculture. *Smart Agric. Technol.* **2023**, *5*, 100305. [[CrossRef](#)]
44. Deforce, B.; Baesens, B.; Diels, J.; Asensio, E.S. Harnessing the power of transformers and data fusion in smart irrigation. *Appl. Soft Comput.* **2024**, *152*, 111246. [[CrossRef](#)]
45. Gao, P.; Xie, J.; Yang, M.; Zhou, P.; Chen, W.; Liang, G.; Chen, Y.; Han, X.; Wang, W. Improved soil moisture and electrical conductivity prediction of citrus orchards based on IOT using Deep Bidirectional LSTM. *Agriculture* **2021**, *11*, 635. [[CrossRef](#)]
46. Bai, T.; Tahmasebi, P. Graph neural network for groundwater level forecasting. *J. Hydrol.* **2023**, *616*, 128792. [[CrossRef](#)]
47. Kurtulmuş, E.; Arslan, B.; Kurtulmuş, F. Deep learning for proximal soil sensor development towards smart irrigation. *Expert Syst. Appl.* **2022**, *198*, 116812. [[CrossRef](#)]
48. Kashyap, P.K.; Kumar, S.; Jaiswal, A.; Prasad, M.; Gandomi, A.H. Towards precision agriculture: IoT-enabled intelligent irrigation systems using deep learning neural network. *IEEE Sens. J.* **2021**, *21*, 17479–17491. [[CrossRef](#)]
49. Medsker, L.; Jain, L.C. *Recurrent Neural Networks: Design and Applications*; CRC Press: Boca Raton, FL, USA, 1999.
50. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [[CrossRef](#)]
51. Kanmani, R.; Muthulakshmi, S.; Subitcha, K.S.; Sriranjani, M.; Radhapoorani, R.; Suagnya, N. Modern irrigation system using convolutional neural network. In Proceedings of the 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 19–20 March 2021; IEEE: Piscataway, NJ, USA, 2021; Volume 1, pp. 592–597.
52. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–11.
53. Lim, B.; Arik, S.Ö.; Loeff, N.; Pfister, T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *Int. J. Forecast.* **2021**, *37*, 1748–1764. [[CrossRef](#)]
54. Dong, G.; Tang, M.; Wang, Z.; Gao, J.; Guo, S.; Cai, L.; Gutierrez, R.; Campbel, B.; Barnes, L.E.; Boukhechba, M. Graph neural networks in IoT: A survey. *ACM Trans. Sens. Netw.* **2023**, *19*, 1–50. [[CrossRef](#)]
55. Wallace, J.; Batchelor, C. Managing water resources for crop production. *Philos. Trans. R. Soc. B: Biol. Sci.* **1997**, *352*, 937–947. [[CrossRef](#)]
56. Zheng, J.; Zhao, T.; Lü, H.; Shi, J.; Cosh, M.H.; Ji, D.; Jiang, L.; Cui, Q.; Lu, H.; Yang, K.; et al. Assessment of 24 soil moisture datasets using a new in situ network in the Shandian River Basin of China. *Remote Sens. Environ.* **2022**, *271*, 112891. [[CrossRef](#)]
57. Dorigo, W.; Gruber, A.; De Jeu, R.; Wagner, W.; Stacke, T.; Loew, A.; Albergel, C.; Brocca, L.; Chung, D.; Parinussa, R.; et al. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sens. Environ.* **2015**, *162*, 380–395. [[CrossRef](#)]
58. Vinnikov, K.Y.; Robock, A.; Qiu, S.; Entin, J.K. Optimal design of surface networks for observation of soil moisture. *J. Geophys. Res. Atmos.* **1999**, *104*, 19743–19749. [[CrossRef](#)]
59. Ochsner, T.E.; Cosh, M.H.; Cuenca, R.H.; Dorigo, W.A.; Draper, C.S.; Hagimoto, Y.; Kerr, Y.H.; Larson, K.M.; Njoku, E.G.; Small, E.E.; et al. State of the art in large-scale soil moisture monitoring. *Soil Sci. Soc. Am. J.* **2013**, *77*, 1888–1919. [[CrossRef](#)]
60. Robock, A.; Vinnikov, K.Y.; Srinivasan, G.; Entin, J.K.; Hollinger, S.E.; Speranskaya, N.A.; Liu, S.; Namkhai, A. The global soil moisture data bank. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 1281–1300. [[CrossRef](#)]
61. Robock, A.; Mu, M.; Vinnikov, K.; Trofimova, I.V.; Adamenko, T.I. Forty-five years of observed soil moisture in the Ukraine: No summer desiccation (yet). *Geophys. Res. Lett.* **2005**, *32*. [[CrossRef](#)]
62. Baldocchi, D.; Falge, E.; Gu, L.; Olson, R.; Hollinger, D.; Running, S.; Anthoni, P.; Bernhofer, C.; Davis, K.; Evans, R.; et al. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Am. Meteorol. Soc.* **2001**, *82*, 2415–2434. [[CrossRef](#)]
63. Dorigo, W.; Van Oevelen, P.; Wagner, W.; Drusch, M.; Mecklenburg, S.; Robock, A.; Jackson, T. A new international network for in situ soil moisture data. *Eos, Trans. Am. Geophys. Union* **2011**, *92*, 141–142. [[CrossRef](#)]
64. Dorigo, W.; Wagner, W.; Hohensinn, R.; Hahn, S.; Paulik, C.; Xaver, A.; Gruber, A.; Drusch, M.; Mecklenburg, S.; van Oevelen, P.; et al. The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1675–1698. [[CrossRef](#)]
65. Gasch, C.; Brown, D.; Campbell, C.; Cobos, D.; Brooks, E.; Chahal, M.; Poggio, M. A field-scale sensor network data set for monitoring and modeling the spatial and temporal variation of soil water content in a dryland agricultural field. *Water Resour. Res.* **2017**, *53*, 10878–10887. [[CrossRef](#)]
66. Smith, A.B.; Walker, J.P.; Western, A.W.; Young, R.; Ellett, K.; Pipunic, R.; Grayson, R.; Siriwardena, L.; Chiew, F.H.; Richter, H. The Murrumbidgee soil moisture monitoring network data set. *Water Resour. Res.* **2012**, *48*. [[CrossRef](#)]
67. Dorigo, W.; Xaver, A.; Vreugdenhil, M.; Gruber, A.; Hegyiova, A.; Sanchis-Dufau, A.D.; Zamojski, D.; Cordes, C.; Wagner, W.; Drusch, M. Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone J.* **2013**, *12*, 1–21. [[CrossRef](#)]

68. Beven, K.J. *Rainfall-Runoff Modelling: The Primer*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
69. Cid-Escobar, D.; Folch, A.; Ferrer, N.; Katuva, J.; Sanchez-Vila, X. An assessment tool to improve rural groundwater access: Integrating hydrogeological modelling with socio-technical factors. *Sci. Total Environ.* **2024**, *912*, 168864. [[CrossRef](#)]
70. Tolomio, M.; Casa, R. Dynamic crop models and remote sensing irrigation decision support systems: A review of water stress concepts for improved estimation of water requirements. *Remote Sens.* **2020**, *12*, 3945. [[CrossRef](#)]
71. Ghat, I.; Mackey, H.R.; Al-Ansari, T. A review of evapotranspiration measurement models, techniques and methods for open and closed agricultural field applications. *Water* **2021**, *13*, 2523. [[CrossRef](#)]
72. Sun, Z.; Di, L. A Review of Smart Irrigation Decision Support Systems. In Proceedings of the 2021 9th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Shenzhen, China, 26–29 July 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–4.
73. You, Y.; Wang, Y.; Fan, X.; Dai, Q.; Yang, G.; Wang, W.; Chen, D.; Hu, X. Progress in joint application of crop models and hydrological models. *Agric. Water Manag.* **2024**, *295*, 108746. [[CrossRef](#)]
74. Gasser, L.; Le Gall, F.; Abily, M. Water efficiency in smart cities: Optimising irrigation for public green spaces. *LHB* **2024**, *110*, 2294076. [[CrossRef](#)]
75. Abioye, A.E.; Abidin, M.S.Z.; Mahmud, M.S.A.; Buyamin, S.; Mohammed, O.O.; Otuoze, A.O.; Oleolo, I.O.; Mayowa, A. Model based predictive control strategy for water saving drip irrigation. *Smart Agric. Technol.* **2023**, *4*, 100179. [[CrossRef](#)]
76. Touil, S.; Richa, A.; Fizir, M.; Argente Garcia, J.E.; Skarmeta Gomez, A.F. A review on smart irrigation management strategies and their effect on water savings and crop yield. *Irrig. Drain.* **2022**, *71*, 1396–1416. [[CrossRef](#)]
77. El Bilali, A.; Taleb, A.; Brouziyne, Y. Groundwater quality forecasting using machine learning algorithms for irrigation purposes. *Agric. Water Manag.* **2021**, *245*, 106625. [[CrossRef](#)]
78. Champness, M.; Vial, L.; Ballester, C.; Hornbuckle, J. Evaluating the Performance and Opportunity Cost of a Smart-Sensed Automated Irrigation System for Water-Saving Rice Cultivation in Temperate Australia. *Agriculture* **2023**, *13*, 903. [[CrossRef](#)]
79. Liao, R.; Zhang, S.; Zhang, X.; Wang, M.; Wu, H.; Zhangzhong, L. Development of smart irrigation systems based on real-time soil moisture data in a greenhouse: Proof of concept. *Agric. Water Manag.* **2021**, *245*, 106632. [[CrossRef](#)]
80. Alibabaei, K.; Gaspar, P.D.; Assunção, E.; Alirezazadeh, S.; Lima, T.M. Irrigation optimization with a deep reinforcement learning model: Case study on a site in Portugal. *Agric. Water Manag.* **2022**, *263*, 107480. [[CrossRef](#)]
81. Risheh, A.; Jalili, A.; Nazerfard, E. Smart Irrigation IoT solution using transfer learning for neural networks. In Proceedings of the 2020 10th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 29–30 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 342–349.
82. Routis, G.; Roussaki, I. Low Power IoT Electronics in Precision Irrigation. *Smart Agric. Technol.* **2023**, *5*, 100310. [[CrossRef](#)]
83. Froiz-Míguez, I.; Lopez-Iturri, P.; Fraga-Lamas, P.; Celaya-Echarri, M.; Blanco-Novoa, Ó.; Azpiliceta, L.; Falcone, F.; Fernández-Caramés, T.M. Design, implementation, and empirical validation of an IoT smart irrigation system for fog computing applications based on Lora and Lorawan sensor nodes. *Sensors* **2020**, *20*, 6865. [[CrossRef](#)] [[PubMed](#)]
84. Zia, H.; Rehman, A.; Harris, N.R.; Fatima, S.; Khurram, M. An experimental comparison of IOT-based and traditional irrigation scheduling on a flood-irrigated subtropical lemon farm. *Sensors* **2021**, *21*, 4175. [[CrossRef](#)]
85. Alves, R.G.; Maia, R.F.; Lima, F. Development of a Digital Twin for smart farming: Irrigation management system for water saving. *J. Clean. Prod.* **2023**, *388*, 135920. [[CrossRef](#)]
86. Manocha, A.; Sood, S.K.; Bhatia, M. IoT-digital twin-inspired smart irrigation approach for optimal water utilization. *Sustain. Comput. Inform. Syst.* **2024**, *41*, 100947. [[CrossRef](#)]
87. Saikai, Y.; Peake, A.; Chenu, K. Deep reinforcement learning for irrigation scheduling using high-dimensional sensor feedback. *arXiv* **2023**, arXiv:2301.00899.
88. Jin, M.; Koh, H.Y.; Wen, Q.; Zambon, D.; Alippi, C.; Webb, G.I.; King, I.; Pan, S. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. *arXiv* **2023**, arXiv:2307.03759.
89. Sharma, A.; Singh, S.; Ratna, S. Graph Neural Network Operators: A Review. *Multimed. Tools Appl.* **2024**, *83*, 23413–23436. [[CrossRef](#)]
90. Yi, K.; Zhang, Q.; Fan, W.; He, H.; Hu, L.; Wang, P.; An, N.; Cao, L.; Niu, Z. FourierGNN: Rethinking multivariate time series forecasting from a pure graph perspective. *Adv. Neural Inf. Process. Syst.* **2024**, *36*.
91. Li, Z.L.; Zhang, G.W.; Yu, J.; Xu, L.Y. Dynamic graph structure learning for multivariate time series forecasting. *Pattern Recognit.* **2023**, *138*, 109423. [[CrossRef](#)]
92. Zhang, Y.; Yan, J. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting. In Proceedings of the Eleventh International Conference on Learning Representations, Kigali, Rwanda, 1–5 May 2023.
93. Zeng, A.; Chen, M.; Zhang, L.; Xu, Q. Are transformers effective for time series forecasting? In Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 11121–11128. [[CrossRef](#)]
94. Miller, J.A.; Aldosari, M.; Saeed, F.; Barna, N.H.; Rana, S.; Arpinar, I.B.; Liu, N. A survey of deep learning and foundation models for time series forecasting. *arXiv* **2024**, arXiv:2401.13912.
95. Gu, A.; Goel, K.; Ré, C. Efficiently modeling long sequences with structured state spaces. *arXiv* **2021**, arXiv:2111.00396.
96. Lin, L.; Li, Z.; Li, R.; Li, X.; Gao, J. Diffusion models for time-series applications: A survey. *Front. Inf. Technol. Electron. Eng.* **2023**, *25*, 1–23. [[CrossRef](#)]
97. Meijer, C.; Chen, L.Y. The Rise of Diffusion Models in Time-Series Forecasting. *arXiv* **2024**, arXiv:2401.03006.

98. Park, M.H.; Chakraborty, S.; Vuong, Q.D.; Noh, D.H.; Lee, J.W.; Lee, J.U.; Choi, J.H.; Lee, W.J. Anomaly Detection Based on Time Series Data of Hydraulic Accumulator. *Sensors* **2022**, *22*, 9428. [[CrossRef](#)]
99. Ehrhart, M.; Resch, B.; Havas, C.; Niederseer, D. A Conditional GAN for Generating Time Series Data for Stress Detection in Wearable Physiological Sensor Data. *Sensors* **2022**, *22*, 5969. [[CrossRef](#)]
100. Benameur, R.; Dahane, A.; Kechar, B.; Benyamina, A.E.H. An Innovative Smart and Sustainable Low-Cost Irrigation System for Anomaly Detection Using Deep Learning. *Sensors* **2024**, *24*, 1162. [[CrossRef](#)]
101. Cheng, W.; Ma, T.; Wang, X.; Wang, G. Anomaly detection for internet of things time series data using generative adversarial networks with attention mechanism in smart agriculture. *Front. Plant Sci.* **2022**, *13*, 890563. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.