agriculture

MDPI

Article

A Hybrid LSTM Approach for Irrigation Scheduling in Maize Crop

Konstantinos Dolaptsis [1], Xanthoula Eirini Pantazi [1,*], Charalampos Paraskevas [1], Selçuk Arslan [2], Yücel Tekin [3], Bere Benjamin Bantchina [4], Yahya Ulusoy [3], Kemal Sulhi Gündoğdu [2], Muhammad Qaswar [5], Danyal Bustan [5,6] and Abdul Mounem Mouazen [5]

1 Laboratory of Agricultural Engineering, Faculty of Agriculture, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; kdolapts@agro.auth.gr (K.D.); paraskevasb@agro.auth.gr (C.P.)
2 Department of Biosystems Engineering, Faculty of Agriculture, Bursa Uludag University, 16059 Bursa, Turkey; sarslan@uludag.edu.tr (S.A.); kemalg@uludag.edu.tr (K.S.G.)
3 Vocational School of Technical Sciences, Bursa Uludag University, 16059 Bursa, Turkey; ytekin@uludag.edu.tr (Y.T.); yahyau@uludag.edu.tr (Y.U.)
4 Department of Biosystems Engineering, Natural and Applied Sciences Institute, Bursa Uludag University, 16059 Bursa, Turkey; ben.bantchina@gmail.com
5 Department of Environment, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium; muhammad.qaswar@ugent.be (M.Q.); d.bustan@qiet.ac.ir (D.B.); abdul.mouazen@ugent.be (A.M.M.)
6 Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Quchan University of Technology, Quchan 94771-67335, Iran
* Correspondence: xpantazi@agro.auth.gr; Tel.: +30-231-099-8868

Abstract: Irrigation plays a crucial role in maize cultivation, as watering is essential for optimizing crop yield and quality, particularly given maize's sensitivity to soil moisture variations. In the current study, a hybrid Long Short-Term Memory (LSTM) approach is presented aiming to predict irrigation scheduling in maize fields in Bursa, Turkey. A critical aspect of the study was the use of the Aquacrop 7.0 model to simulate soil moisture content (MC) data due to data limitations in the investigated fields. This simulation model, developed by the Food and Agriculture Organization (FAO), helped overcome gaps in soil sensor data, enhancing the LSTM model's predictions. The LSTM model was trained and tuned using a combination of soil, weather, and satellite-based plant vegetation data in order to predict soil moisture content (MC) reductions. The study's results indicated that the LSTM model, supported by Aquacrop 7.0 simulations, was effective in predicting MC reduction across various time phases of the maize growing season, attaining R^2 values ranging from 0.8163 to 0.9181 for Field 1 and from 0.7602 to 0.8417 for Field 2, demonstrating the potential of this approach for precise and efficient agricultural irrigation practices.

Keywords: precision agriculture; artificial intelligence; long short-term memory; predictive control; deep learning; moisture content; water management; time series analysis

check for updates

Citation: Dolaptsis, K.; Pantazi, X.E.; Paraskevas, C.; Arslan, S.; Tekin, Y.; Bantchina, B.B.; Ulusoy, Y.; Gündoğdu, K.S.; Qaswar, M.; Bustan, D.; et al. A Hybrid LSTM Approach for Irrigation Scheduling in Maize Crop. *Agriculture* **2024**, *14*, 210. https://doi.org/10.3390/agriculture14020210

Academic Editor: Yufeng Luo

Received: 30 November 2023
Revised: 10 January 2024
Accepted: 23 January 2024
Published: 28 January 2024

1. Introduction

The increased demand for global food production inevitably leads to the increased irrigation and water use of several farming practices, leading to the consumption of 70% of the available freshwater globally [1]. Conventional irrigation scheduling is mostly based on the farmers' estimations and experience resulting in the reduction of available water for irrigation and it usually takes place in an open-loop fashion, where the supplied irrigation volume and the prevailing soil water status are indirectly connected; consequently, they appear ineffective mostly due to their inability to take into account site-specific, spatially variable factors and weather conditions occurring in the irrigated area. For the above reasons, it is crucial to implement sustainable, yet effective water management practices

towards the optimal crop health status, yield, and the minimization of water consumption and cost.

Precision irrigation practices are based on the assessment of the accurate water volume and irrigation period required site-specific to each field, so as to enhance yield productivity while also reducing the farmers' labor costs. The above tasks can be achieved through monitoring the soil moisture levels. Soil moisture is a critical variable in hydrological, climatic, and agricultural processes, influencing everything from plant growth to land-atmosphere interactions. The presence of soil moisture is beneficial for agricultural productivity, as it has the potential to enhance several physiological and biochemical processes associated with crop development and, subsequently, crop yield. Therefore, by acquiring knowledge about the soil moisture content, farmers can obtain valuable insights on optimal timing for sowing and growing crops, adequacy of soil infiltration, and sufficiency of the water supply needed to facilitate crop root growth. The realization and understanding of these procedures increases the need for the investigation of novel approaches that are capable of precisely estimating the level of soil MC in a specific location [2]. For the above reasons, continuous and accurate soil moisture records are very important for research and practical applications. However, obtaining such records can be challenging due to sensor malfunctions, data transmission errors, or environmental interferences [3].

For several decades, the assessment of irrigation needs in agricultural systems has relied on the application of conventional deterministic and empirical models. These models, rooted in established principles and equations, have served as the foundational framework for understanding and quantifying the water requirements essential for optimal crop growth. This approach has been instrumental in shaping irrigation strategies and resource management practices within the agricultural domain, contributing significantly to the sustainable development of crop cultivation methodologies. The applied models are typically based on already known relationships between key variables such as soil moisture, evapotranspiration (ET) rates, environmental conditions, and crop growth stages, however, they often fail to take into consideration some other important variables affecting irrigation needs, including spatial soil variability, which is attributed to natural variation in soil properties and characteristics even within the same investigated field. To overcome these limitations, the effective combination of deterministic and empirical models, coupled with advanced computational technologies such as machine learning (ML), can provide more accurate and adaptive irrigation management approaches.

ML models can effectively predict key variables in irrigation, including ET and MC, with minimal human intervention by capturing their complex relationships, adapting to changing conditions, and incorporating nonlinear and interactive effects, elevating the potential to revolutionize the prediction of soil properties [4]. Wu et al. [5] developed a two-level ensemble model using ML models for the estimation of daily evapotranspiration (ET_o). The first level included Random Forest (RF), Support Vector Regression (SVR), Multilayer Perceptron (MLP), and K-Nearest Neighbors (KNN), while the second layer contained the Linear Regression (LR) model, employed as the meta-learner for results extraction. The applied model demonstrated high accuracy, with coefficient of determination (R^2) values ranging from 0.66 to 0.99. Other common ML approaches for the investigation and prediction of MC dynamics included the application of support vector machines (SVMs) [6] and adaptive neuro-fuzzy inference systems (ANFIS) [7] for simulating times series regarding soil MC, utilizing weather, precipitation, and crop coefficient data as input. The aforementioned traditional ML approaches demonstrated satisfactory performances compared to the deterministic and physical models, however they appeared sensitive to different environmental conditions, unstable in offering reliable predictions for the entire range of soil MC levels, and weak in providing insightful conclusions when employed in different regions than those that they were initially calibrated and developed in. In addition to the weaknesses of traditional ML models, the common ANNs models have been widely applied, but they are also limited in modelling dynamic data and approximating complex processes due to their tendency to lose previously attained and processed

information, too [8]. There are some commonly applied techniques in order to enhance the performance of the traditional ANNs for time series prediction including the addition of extra steps during pre-processing [9] and parameters adjustment with the help of genetic algorithms [10]. However, preprocessing techniques are time-consuming due to their high dependency on time and frequency, and they negatively affect the models' adaptability to unknown environments. The above ambiguities often pose disastrous results to the models' prediction performances especially when dealing with highly causal systems. One further technique to improve the performance of traditional ANNs is the employment of sliding windows. Sliding windows can be employed in an ANN by arranging sequential data into windows of defined length that can either overlap or not overlap with each other. These windows are then used as input for the model. However, these models lack inherent memory mechanisms to retain information from past inputs. Each window is treated independently, and the network does not naturally capture long-term dependencies in sequential data [11]. For the above reasons, robust and scalable data-driven models, in the form of Deep Learning (DL) models, such as convolutional neural networks (CNNs), radial basis function networks (RBFN), long short-term memory networks (LSTM), recurrent neural networks (RNN) and deep belief network (DBN) are employed. Their efficacy lays in their special design's ability to recognize complex patterns through capturing and retaining information among large sequential data, and to manage high-dimensional and spatiotemporal data resulting in accurate predictions. Among the above-mentioned prediction algorithms, the RBFN and LSTM are regarded as highly effective computational tools utilized for learning sequential data and are more suitable for predicting time series. More specifically, the efficiency of the LSTM models can be attributed to their architecture; they are structured of recurrent connected block units that are composed of at least one memory cell in order to store and access the information for a certain period. The existence of at least one memory cell helps to raise their efficiency, while also lowering the gradient error [12]. The self-looped cells of the LSTM model are capable of learning long-term temporal dependencies in sequential data, simultaneously retaining the information from previous time steps [8]. Agyeman et al. [13] developed a predictive control model with discrete actuators for predicting irrigation scheduling. In this approach, the soil–water–atmosphere system is evaluated with the help of an LSTM model, aiming to assess the optimal water uptake in crops. Jetitha and Rajesh [14] proposed a novel approach for irrigation scheduling based on a deep bi-directional LSTM model Deep Data Logger and Irrigation Activator Unit (D^2LIAU). The results indicated that the DBLSTM-based D^2LIAU approach enables efficient irrigation scheduling with a significantly higher degree of reliability in irrigation monitoring since it manages to save water consumption from 21.96% to 63.05% when compared to other controlled irrigation practices.

The main aim of the current study is to demonstrate a hybrid LSTM approach for irrigation scheduling prediction. The irrigation needs have been predicted taking into account soil, weather, and vegetation data. An LSTM model has been employed so as to investigate the predictive capabilities of soil MC reduction per day based on the acquired weather data including average daily air temperature, the total daily solar radiation (SR), the average daily relative humidity (RH), and also the Leaf Area Index (LAI) in two maize fields located in Bakırköy village in Karacabey, Turkey. Maize, a widely cultivated cereal crop, is particularly sensitive to variations in moisture levels [3]. The study employed an LSTM model that integrates data from three distinct sources, including soil sensors, a weather station, and satellite imagery. The data measurements from soil sensors, the weather station and satellite were acquired on a daily basis for the entire investigated period. In order to overcome the lack of MC missing values in Field 1 and the limited number of soil MC measurements obtained from only four soil sensors in Field 2, the water-driven model Aquacrop 7.0 has been used [15,16] in order to compute the daily MC and to acquire data from two extra locations within Field 2 for the entire investigated period. The limited number of soil measurements during the investigated cultivation period has been attributed to some difficulties encountered in establishing a reliable and continuous

data transmission channel between the sensors responsible for MC measurements and the data logger. These transmission issues significantly affected the quantity of the MC data collected. Aquacrop 7.0 is regarded as a reliable tool due to its proven capability of simulating successfully parameters used as input data. These types of data can include soil MC, evapotranspiration (ET), and yield in the occasion that they fail to be measured in a continuous and accurate manner during experimental field measurements [17]. In the current study, the failure of acquiring continuous and accurate records is often attributed to sensor malfunctions and data transmission errors. However, the performance of the LSTM model reveals its significant potential in efficiently predicting the future MC reduction in maize fields, overcoming the model's physical tendency to be influenced by the availability and quality of data which is attributed both to its architecture and the simulation effect of Aquacrop, introducing novel techniques towards more efficient and precise agricultural irrigation practices and patterns. Moreover, the current approach encourages the adoption of site-specific irrigation patterns that enable Variable Rate Irrigation (VRI), paving the way towards enhancing the annual production yields of grains.

2. Materials and Methods

2.1. Study Area and Data Collection

The experimental fields were located in Bakırköy village in Karacabey, Turkey (40°08′ 22.61′ N, 28°22′58.7′ E). An 8.2-ha field named Field 1 and a 2.5-ha field named Field 2 cultivated with maize were investigated (Figure 1). The type of soil in both fields is alluvial, while the soil texture characteristics differ since their percentage content varied between sand, clay, and loam clay. The determination of the mechanical composition of the soil was performed using the Bouyoucos hydrometer method. The results of this method are presented in Section 2.2.4 of this work. The field study was conducted during the 2022 growing season. The seeding took place on 16 May in Field 1 and on 11 May in Field 2. The irrigation system that was utilized was the drip irrigation method.

The irrigation water is supplied from the buried pressurized irrigation network. Field 1 was located in a 20-ha production field that was divided into five irrigation plots, each with a maximum area of 4.0 ha. Field 1 coincided with three of these plots. Deficit irrigation is common in the region, which was also used in this study. The row spacing of the maize crop was 0.7 m, irrigation lateral line spacing was 1.4 m, and the number of drippers was 18,000 per hectare during the irrigations. Inline linear drippers (Eurodrip -Eolos, Rivulis, Israel) were used in the irrigations with an irrigation intensity of 28.8 t/h-ha at an operating pressure of 3 bar. A manual valve was used to adjust the specified operating pressure (3.0 bar) of the drippers at the end of the secondary irrigation lines to ensure that the water pressure is not below 3.0 bar at the most distant lateral. An analog pressure gauge (Pakkens 0–6 bar) was used to measure the pressure at the end of the secondary (Field 1) and main lines (Field 2). For each irrigation, the farmer changed the irrigation time from 6 h to 10 h per irrigation. The farmer utilizes crop water demands given in the irrigation guide for the Bursa-Karacabey district and the weather station (Metos IMT280) data to determine the application rate for each irrigation. The frequency of the irrigations was affected not only by the technical data but also the availability of the labor since the irrigation system was not automatically controlled. In the early growth stages of the maize crop, the irrigations were done in 2–3 days whereas in the rest of the irrigation season the irrigations were done in 4–10 days depending on the aforementioned conditions. This type of pressurized irrigation system offers several advantages, including a high degree of linear control over the water application, optimal nutrient intake, the optimization of water usage, and the reduction of water resources waste [18,19]).

The quantity and placement of the sensors were made according to the relative MZs constructed for irrigation. MZs were created based on the precise spatial data of predicted soil MC and clay. The aforementioned data were acquired on 28 June 2021 from visible and near-infrared spectral measurements, using the online multi-sensor platform developed by Mouazen [20].

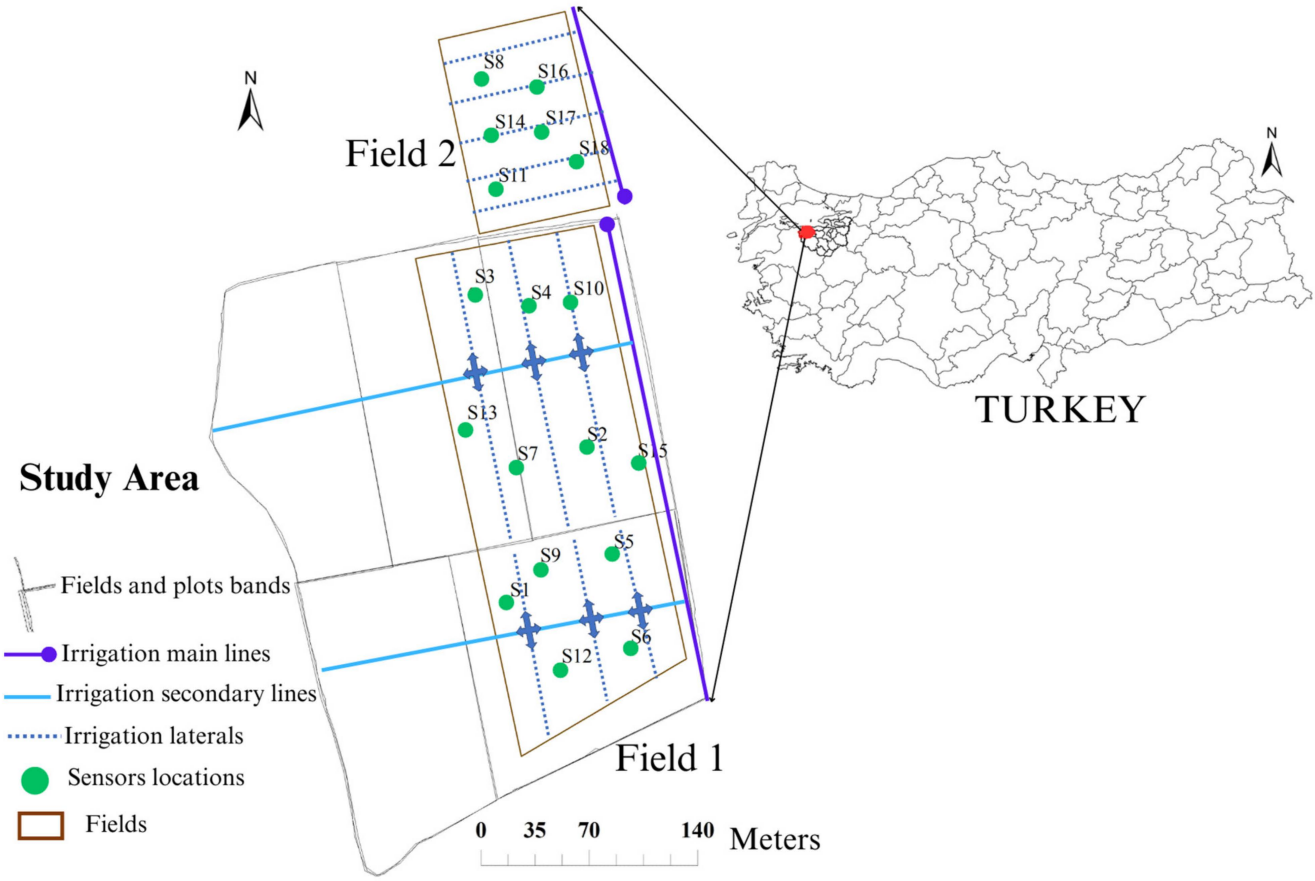


Figure 1. A view of the investigated maize fields (Field1 and Field 2).

The data for the development of the utilized LSTM model for soil moisture prediction were acquired from three distinct sources including soil sensors, a weather station, and satellite-based plant vegetation data. A cloud-based LoRA-WAN wireless network was used so that data from soil sensors could be constantly collected (Figure 2). The study involved the deployment of nine Dragino LSE01 soil sensors in various areas within Field 1, as well as four soil sensors in Field 2. The soil sensors were installed vertically at 30 cm depth to monitor the soil MC ((%) VWC), apparent electrical conductivity (ECa) (uS/cm), and soil temperature (°C). The sensors enable MC measurements using frequency domain reflectometry (FDR), and the accuracy of the MC sensors was reported to be ±3% by the manufacturer [21]. The accuracy of the sensors was checked during the irrigation season once by collecting soil samples at sensor locations, resulting in approximately 3–5% errors in soil moisture measurements. No further investigations were done to check or improve the measurement accuracy of MC sensors. The sensor network gateway was posed about 800 m away from the sensors in a clear view of sight. A commercial weather station (Metos IMT280) located near the farm center and 3.5 km away from the experimental field was used to collect weather data, including air temperature (°C), dew point temperature (°C), solar radiation (W/m^2), vapor pressure (kPa), relative humidity (%), rainfall (mm), wind speed and instant wind speed (km/h), and daily evapotranspiration (mm).

The availability of free-of-cost, high resolution satellite datasets (finer than 30 m) was obtained from sensors such as Sentinel-2 and advanced open-source tools such as Sentinel Application Platform (SNAP). This tool offers accurate and consistent values of LAI.

Sentinel-2 images at 10 m resolution were used to derive LAI using the built-in Biophysical processor, also called the Sentinel-2 Land bio-physical processor (SL2P), within the SNAP Toolbox. The Biophysical processor uses eight reflectance bands (B3, B4, B5, B6, B7, B8A, B11, and B12) [22]. The utilized Sentinel-2 product was Level-2A.

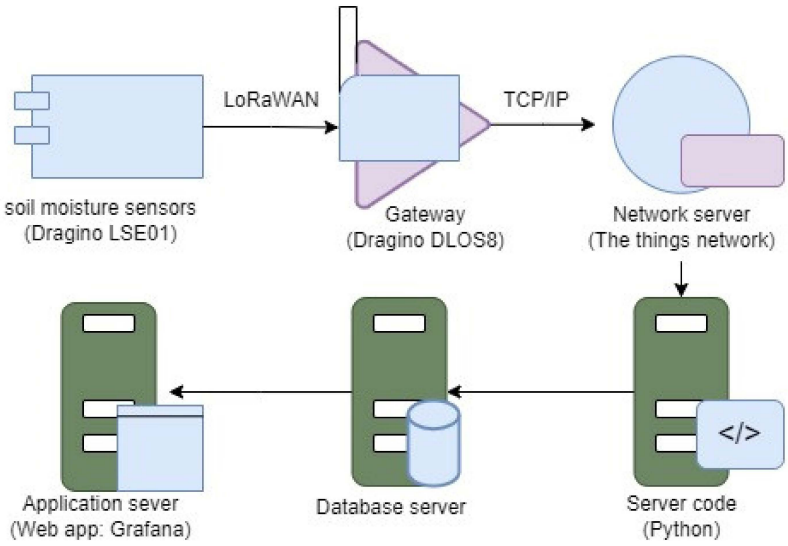


Figure 2. LoRa-WAN used for soil moisture collection.

The LAI values were obtained from satellite images captured on 23, 26 and 28 June, as well as on 1, 3, 8, 13, 16, 18, 21, 23 and 26 July and 2 August 2022. It is worthy to note that deliberate emphasis has been placed on the selection of imagery acquired on days characterized by optimal meteorological conditions. Emphasis was also given to instances featuring clear skies, abundant sunlight, and an absence of cloud cover, in order to ensure the acquisition of high-quality satellite imagery, defining the basis for a precise and reliable extraction of the Leaf Area Index (LAI). Recognizing the necessity for a continuous dataset, a linear interpolation method was implemented for those days that satellite images were not available. The interpolation technique involved leveraging the acquired data points to estimate LAI values for the sensor location pixel for the non-captured days, providing a seamless and comprehensive analysis for the entire monitoring period and bridging any temporal gaps within the dataset. Finally, for the days before the 23 June 2022, the values of LAI were equal to 0 as the crop was at the initial stages of growth. LAI, as an essential biophysical parameter of vegetation, is a dimensionless variable representing the ratio of leaf area to per unit ground surface area. This ratio is associated with gas-vegetation exchange processes including photosynthesis, evaporation, transpiration, and rainfall interception [23].

2.2. Methods

2.2.1. The Long Short-Term Memory Model (LSTM) Process

In the current study, a long short-term memory recurrent neural network (LSTM RNN) has been developed. RNNs belong to a more complex and advanced type of ANNs networks that are well-suited for modeling sequential inputs, such as speech data and natural language data, among others. The structure of RNNs is similar to that of feed-forward neural networks or artificial neural networks (ANNs). However, RNNs differ from ANNs in that they possess cyclic connections, in contrast to the acyclic nature of ANNs. In a typical ANN, information tends to flow only in one direction, moving from the input layer through hidden layers to the output layer. However, the cyclic connections that characterize RNN correspond to the feedback connections between the neurons that allow information to be fed back from the output layer to the hidden layers. This special feedback design enhances the RNN networks' capability to learn from their own predictions based on previously attained information and refine their representations and predictive performance over time. This kind of architecture can be particularly useful when dealing with tasks where temporal dependencies or context are crucial, like in the occasion of time-series data handling. The cyclic connections enable the network to capture and leverage feedback information, making it more effective in tasks like sequence modeling [24]. More

specifically, in the occasion of a RNN, the neurons inside a layer have the capability to establish connections with one another, including the possibility of self-connections, a feature that was previously not allowed in artificial neural networks (ANNs). Due to their inherent cyclic structure, RNNs leverage past inputs to calculate outputs at each iteration, enabling them to retain a memory of prior events. This memory is utilized to facilitate subsequent predictions. The presence of cyclic links enables the integration of prior information into the decision-making process [25]. The diagram in Figure 3 illustrates the architecture of a recurrent neural network.

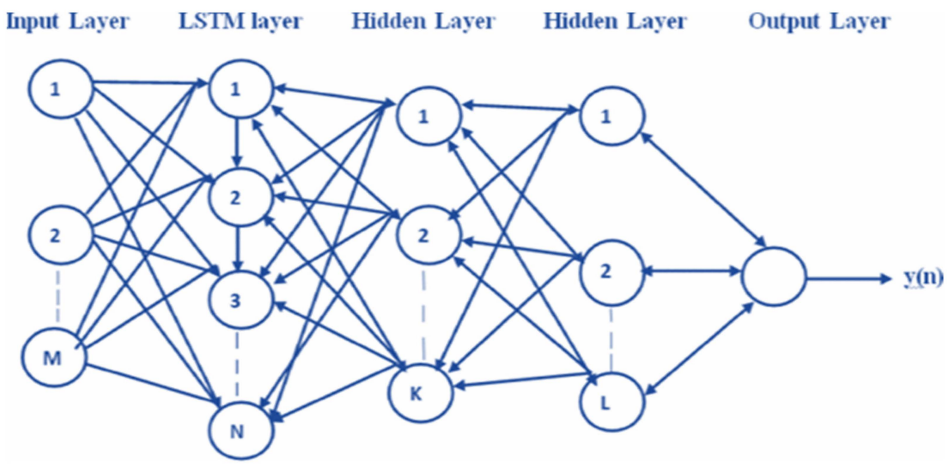


Figure 3. Structure of typical recurrent neural network. Each neuron forms connections with the neurons in its own layer, as well as with the neurons in the preceding and next layers, and with itself.

The LSTM model has been introduced by Hochreiter, and belongs to the broader category of RNN models, but in a more revolutionized version. The main advantage of LSTM over typical RNNs lies in their short-term memory availability. Consequently, as we move on to the next layers, the gradients are preserved so that the weight factors are updated and in the occasion that the short-term memory is in bidirectional means, the information is preserved in both directions [26]. Regarding its architecture, the LSTM model is composed of one block with three gates, namely the Input Gate, the Forget Gate and the Output Gate. Each gate's activation probability ranges from 0 to 1 corresponding to an open gate and a closed gate, respectively. There are some cells capable of recalling information and forwarding it to the next gate, while the hidden layer unit in RNN is replaced by three gates. The main objective of an LSTM model is to get the output variable O_t by utilizing the equations denoted in Equations (1)–(6). In order to control the information flow within the LSTM network, each LSTM unit consists of a hidden state vector h_t, a cell memory state vector C_t, and three gates. The architecture of an LSTM is presented in Figure 4.

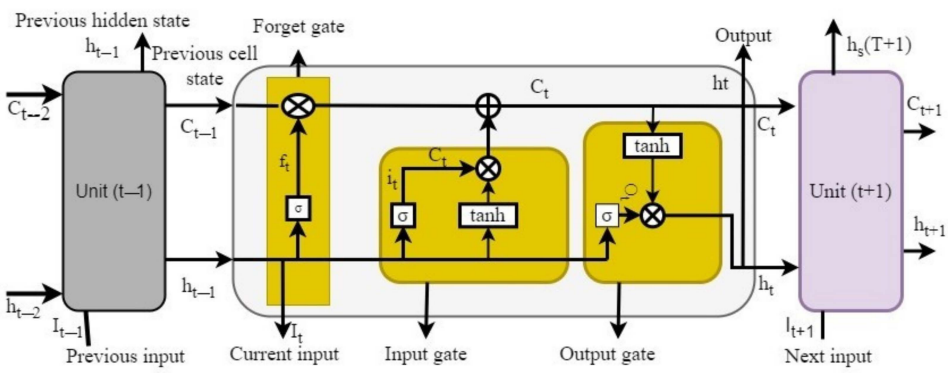


Figure 4. Architecture of an LSTM cell, comprising three special gates: Input Gate (i), Forget Gate (f), and Output Gate (O). These gates control which information is going to be retained, forgotten, and maintained, and the volume of information.

The first gate (f_t) is called the Forget Gate, which controls to what extent should specific information from the prior cell memory state (C_{t-1}) be disregarded. The Forget Gate's output vector is represented in Equation (1) as follows:

$$f_t = \sigma\left(w_f I_t + z_f h_{t-1} + b_f\right), \tag{1}$$

where f_t denotes the sigmoid activation function that takes values in the set {0,1}, I_t is the input sequence, w_f and z_f are the adjustable weight parameters, and b_f is the bias vector. At the initial time instant (t = 0), the variables h_t and C_t are initialized with a vector equal to zero as specified by the user-defined parameters of the network. Then, the cell memory state vector undergoes an update through the use of a hyperbolic tangent layer, namely the tanh function, which is mathematically demonstrated in Equation (2) as follows:

$$C_t = tanh(w_c I_t + z_c h_{t-1} + b_c), \tag{2}$$

where $C_t \in \{-1,1\}$, and w_c, z_c, and b_c are additional coefficients that can be learned. Furthermore, the output of the second gate, known as the Input Gate, regulates the information that is utilized so as to update the cell memory state C_t at the present time. This is expressed in Equation (3) as follows:

$$i_t = \sigma(w_i I_t + z_i h_{t-1} + b_i), \tag{3}$$

where $i_t \in \{0,1\}$ is sigmoid function, and w_i, z_i, b_i are another set of learnable coefficients for the input gate. Using the outcomes of Equations (1) and (3), the cell memory state vector is modified as follows:

$$C_t = f_t \odot C_{t-1} + i_t \odot C_t, \tag{4}$$

where $\odot$ represents the multiplication between the gate and cell memory state vector elements. The first term of Equation (4) can be interpreted as the information that must be forgotten from the previous cell memory state vector C_{t-1} and the information that must be stored f_t close to one). Similarly, the second term can be interpreted as the degree to which new input information must be stored (i_t close to 1) and the information that can be neglected (i_t close to 0).

The third and final gate is the Output Gate (O_t), which regulates the volume of information acquired from the current cell memory state vector, transferred to the next hidden state h_t, and is represented in Equation (5) as follows:

$$O_t = \sigma(w_o I_t + z_o h_{t-1} + b_o), \tag{5}$$

where $O_t \in \{0,1\}$ is sigmoid function, and w_o, z_o, and b_o are the set of learnable coefficients for the output gate. In addition, the new hidden state h_t is calculated using Equations (4) and (5), and it can be expressed as follows:

$$h_t = tanh(C_t) \odot O_t. \tag{6}$$

Then, the final output of the Long Short-Term Memory (LSTM) layer is forwarded to a completely connected dense layer that consists of a single neuron. The ultimate predicted output, denoted as y_{pred}, is calculated by Equation (7):

$$y_{pred} = w_d h_k + b_d, \tag{7}$$

where h_k is the output of the final LSTM layer, w_d and b_d are the weight coefficients and bias term of the fully connected dense layer [27].

2.2.2. LSTM Model Training and Hyperparameter Tuning

The soil measurements concerned the following variables: electrical conductivity (ECa), soil temperature, and moisture content (MC). In this study, MC has been solely utilized as the target variable that the model seeks to predict. More specifically, the daily MC

reduction was determined through utilizing the value of MC for each sensor on a daily basis. The soil sensor data were acquired from mid-June to 3 August. However, sporadic missing values occurred as a result of transmission failures in the data logger. In order to address this issue and acquire a thorough dataset regarding the sowing dates, specifically on 11 May for Field 2 and on 16 May for Field 1, the Aquacrop program was employed to simulate the moisture content on a daily basis. Furthermore, in Field 2, due to the limited number of measurements obtained from the four soil sensors, the Aquacrop program was employed to acquire data from two extra locations inside the field throughout the entire period. The weather measurements were obtained by the utilization of a weather station. The meteorological data that was used consisted of the average daily air temperature, the total daily solar radiation, and the average daily relative humidity. Furthermore, the utilization of Sentinel-2 satellite images was employed to generate the leaf area vegetation index (LAI). The data (measured and interpolated) that were used for training the LSTM network are visually presented in Appendix A. The objective of the model is to forecast the reduction of MC for the next day using the MC reduction, the weather data, and Leaf Area Index (LAI) of the previous three days. The model underwent testing for a duration spanning from 18 June 2022 to 3 August 2022, corresponding to four distinct phases. During the initial phase, the prediction of a decrease in MC between the 18 and 28 June 2022 was made using data from the months of May (corresponding to the sowing date) up through to the 17 June, which were utilized for training purposes. Similarly, during the period from 29 June 2022 to 12 July 2022, the training process was conducted using data from May until 28 June. During the third phase, the model underwent training using data from the sowing date until the 12 July 2022. The purpose of this training was to enable the model to generate predictions for the period between the 13 and 24 July 2022. In the final phase, the entire available dataset was utilized to evaluate the performance of the model in predicting the MC reduction between the 25 July 2022 and the 3 August 2022 (Table 1).

Table 1. Date periods throughout the investigated cultivating period divided into four district phases for training and testing of the LSTM model.

Model	Phase			
	Phase 1	**Phase 2**	**Phase 3**	**Phase 4**
Training	11 May–17 June 2022	11 May–28 June 2022	11 May–12 July 2022	11 May–24 July 2022
Testing	18 June–28 June 2022	29 June–12 July 2022	13 July–24 July 2022	25 July–3 August 2022

During each phase, the available training dataset was divided into a window of three. The term "window of three" denotes that the model takes a sliding window of size three across the sequential data in order to predict the next target value. This window starts at the first data point and moves one step at a time through the sequence [26]. Furthermore, in each phase, the available dataset was partitioned into two subsets. Specifically, 80% of the dataset was assigned for training the model, while the remaining 20% was reserved for conducting an independent test. Through testing on unseen data, the model's ability to generalize to the desired timeframe can be evaluated. In this study, the time series approach is essential because it accounts for temporal dependencies and seasonal variations that may affect irrigation requirements. By training and testing on discrete time periods, the model does not rely on future data during the training phase, preventing data leakage and obtaining objective performance metrics. Once the separation is completed, the data underwent pre-processing and fusion procedures prior to being fed to the model. Standardization was used to transform the features of the dataset so as to take a mean of 0 and a standard deviation of 1. Feature level fusion was selected due to the combination of features extracted from multiple sources. The selection of the appropriate hyperparameters used to train the model was also a significant issue that needed to be addressed, as they tend to have a substantial impact on the efficacy of the models. Tuning the hyperparameters is crucial, and can be accomplished through a variety of techniques, such as grid search, random search, and Bayesian optimization [28]. The grid search technique has been

chosen for the tuning of hyperparameters. In this method, a variety of hyperparameters are evaluated, and finally the one with the highest performance on the validation set is chosen. The different hyperparameters that were tested (first layer size, second layer size, optimizer, dropout rate, number of epochs, batch size) for training each model are demonstrated in Table 2.

Table 2. Tested values of the different deep learning hyperparameters for their tuning by grid search.

Model	Hyperparameters	Values
LSTM	Number of fully connected layers	1, 2, 3
	First layer size	16, 32, 64, 100
	Second layer size	16, 32, 64, 100
	Optimizer	Adam, SGD, RMSprop
	Dropout rate	0, 0.1, 0.2
	Number of epochs	40, 80, 120
	Batch size	8, 16, 32

Regarding the selection of the optimization algorithm, the Adaptive Moment Estimation (Adam) algorithm appeared to be more performant compared to other tested algorithms, such as SGD (Stochastic Gradient Descent) and RMSprop (Root Mean Square Propagation), since the Adam's learning rates can easily adjust to the varying gradients during the training of sequences when employing an LSTM model [29,30]. The selected Adam algorithm for training was set at a batch size of 16 and a total of 80 epochs. In addition, the optimal dropout rate was determined to be equal to 0, and the initial LSTM layer was selected with a size of 64, while the second LSTM layer was chosen to have a size of 32. An attention mechanism was placed between the two LSTM layers in order to enhance the model's understanding of sequences. The method accomplishes this by computing relevance scores between a current time step and all other time steps. These scores are transformed into attention weights which were used to combine information from all time steps, creating an enriched dataset representation. This enriched information is integrated into the initial LSTM's output and passed to the next LSTM layer. This process allows the model to emphasize important parts of the sequence, capture long-range dependencies, handle various sequence lengths, and produce more accurate predictions. The created attention mechanism tends to improve the model's performance when dealing with sequence-related tasks through enabling the model to focus on relevant information and to interpret complex patterns [31].

Additionally, certain callbacks were involved to further optimize the Grid Search method and the overall performance of the training process. Under this framework, three primary callback functions were employed including Early Stopping, Model Checkpoint, and ReduceLROnPlateau. The Early Stopping callback has been chosen for monitoring the validation loss and also for finalizing the training process in case no progress is indicated throughout a specified period of epochs, also known as "patience". In the current study, a patience level of 6 epochs has been selected. The Model Checkpoint callback was employed to store the model's weights after each epoch, keeping the epoch with the highest validation accuracy, to ensure that the most optimal model will be used for future deployment. ReduceLROnPlateau was employed to dynamically adjust the learning rate of the optimizer based on the observed behavior of the loss function during training. In the current study, the learning rate was set equal to 0.0021. The data processing and models development have been implemented using the TensorFlow library (version 2.10) within a Python-based programming environment (Python version 3.9).

2.2.3. Evaluation Metrics for the LSTM Model

The model was evaluated, using standard evaluation metrics that are commonly used in regression problems, such as the coefficient of determination (R^2) and the root mean squared error ($RMSE$). R^2 ranges from 0 to 1 and it is interpreted as the percentage of the variation in the dependent variable that can be attributed to the independent variables in

the model. The formula for the R^2 metric, also known as the coefficient of determination, is given in Equation (8) as follows:

$$R^2 = 1 - \left(\frac{SSR}{SST}\right), \tag{8}$$

where SSR (Sum of Squared Residuals) represents the sum of the squared differences between the predicted values and the mean of the dependent variable. It measures the unexplained variation in the dependent variable by the regression model. On the other hand, SST (Total Sum of Squares) represents the total sum of squared differences between the actual values and the mean of the dependent variable. It measures the total variation in the dependent variable. $RMSE$ represents the standard deviation of the residuals, i.e., the differences between the predicted and actual values. A lower $RMSE$ value indicates better accuracy and a closer fit of the model to the data. The formula for the Root Mean Square Error ($RMSE$) metric is given in Equation (9) as follows:

$$RMSE = \sqrt{\left(\frac{1}{n} \cdot \sum (y\_pred - y\_actual)^2\right)}, \tag{9}$$

where n is the total number of data points, y_pred represents the predicted values from the mode and y_actual represents the actual values of the dependent variable [32].

2.2.4. AQUACROP Model Simulations

Precision agriculture has gained momentum with the integration of simulation models and ML approaches. A precise estimation of maize's water requirements can substantially augment water usage efficiency. In this work, Aquacrop v7.0 has been used [15] to simulate the soil MC so as to overcome the soils sensor's failure to acquire constant or adequate measurements in the two investigated maize fields. The MC data derived from Aquacrop have been integrated to the existent MC datasets acquired from soil sensor measurements, then were employed for the training of the LSTM neural network to determine the maize's water requirements.

Aquacrop, developed by the Food and Agriculture Organization (FAO), Rome, Italy, of the United Nations, is a widely accepted deterministic model that primarily focuses on addressing water needs under different environmental conditions. The model is based on the balance between water availability and demand and is specialized in handling environments where water availability is limited. For maize, specific parameters such as root depth, canopy cover, and soil characteristics are used as input data into Aquacrop. The software has built-in templates for maize, but for better accuracy, it is essential to adjust settings based on local conditions [33,34]. One of the significant features of Aquacrop is its ability to utilize basic soil parameters and transform it into hydraulic properties with the help of pedotransfer functions, which are essential for performing accurate simulation regarding soil MC.

The hydraulic properties required for AQUACROP v7.0 were estimated using pedotransfer functions, derived from the soil texture data. In Tables 3 and 4 the soil texture data and the hydraulic properties (residual soil water content θ_r, saturated soil water content θ_s, α and n empirical fitting parameters and saturated hydraulic conductivity K_s) were estimated using pedotransfer functions [35] for the investigated fields Field 1 and 2. These data are used to assess the water content of field capacity and permanent wilting point using the van Genuchten equation [36].

Table 3. Estimated soil texture and hydraulic parameters using pedotransfer functions in different soil sensor locations in Field 1.

Loc #	Silt	Clay	Sand	θ_r	θ_s	α	n	K_s	USDA Texture
		(%)		(cm^3/cm^3)		(1/cm)	(−)	(cm/Day)	
Loc 2	29.99	37.36	32.68	0.0822	0.4411	0.0116	1.4204	8.77	Clay Loam
Loc 3	28.79	39.36	32.01	0.0820	0.4429	0.0108	1.4388	10.41	Clay Loam
Loc 5	29.48	36.75	34.25	0.0837	0.4453	0.0122	1.4052	8.46	Clay Loam
Loc 7	30.05	38.21	31.51	0.0811	0.4388	0.0110	1.4346	9.42	Clay Loam
Loc 9	30.62	38.90	30.49	0.0798	0.4370	0.0107	1.4455	9.82	Clay Loam
Loc 10	29.74	35.87	34.65	0.0839	0.4445	0.0126	1.3969	7.97	Clay Loam
Loc 12	30.97	37.27	31.30	0.0805	0.4361	0.0113	1.4296	8.60	Clay Loam
Loc 13	30.29	40.06	29.50	0.0789	0.4356	0.0100	1.4608	10.79	Clay
Loc 15	30.07	38.71	31.02	0.0806	0.4381	0.0108	1.4413	9.80	Clay Loam

Table 4. Estimated soil texture and hydraulic parameters using pedotransfer functions in different soil sensor locations in Field 2.

Loc #	Silt	Clay	Sand	θ_r	θ_s	α	n	K_s	USDA Texture
		(%)		(cm^3/cm^3)		(1/cm)	(−)	(cm/Day)	
Loc 8	35.70	29.13	35.17	0.0824	0.4319	0.0162	1.3479	5.23	Clay Loam
Loc 11	35.24	29.20	35.56	0.0828	0.4332	0.0161	1.3463	5.30	Clay Loam
Loc 14	34.54	31.84	33.63	0.0814	0.4330	0.0144	1.3751	5.80	Clay Loam
Loc 16	35.01	30.14	34.85	0.0823	0.4331	0.0155	1.3565	5.44	Clay Loam
Loc 17	32.86	30.74	36.40	0.0840	0.4387	0.0152	1.3521	6.03	Clay Loam
Loc 18	32.51	32.54	34.95	0.0831	0.4383	0.0141	1.3726	6.38	Clay Loam

Clay loam soil has a balanced texture that includes a mixture of sand, silt, and clay particles. This balance allows it to retain moisture while still providing good drainage. Also, this type of soil generally has moderate to good infiltration properties. It can absorb water without excessive runoff, which is beneficial for crop growth. One of the strengths of clay loam soil is its ability to retain moisture. The fine clay particles help hold water in the soil, which can be advantageous during dry periods.

Soil, crop, and meteorological parameters from both investigated fields were used as input data for Aquacrop (Tables 3 and 4). The simulated MC data derived from Aquacrop have been integrated to the initial MC dataset acquired from the soil sensors, which was utilized for training the employed LSTM model.

3. Results and Discussion

In order to assess and predict the irrigation needs for each of the investigated fields, the reduction of MC within a defined period, from 18 June 2022 to 3 August 2022 needed to be assessed and predicted as well. The dataset was split into four distinct phases in response to fluctuations of crop irrigation needs and also to facilitate the repeated training and testing loops of the LSTM model to ensure that the developed model will be refined with the most pertinent data for every temporal interval, in order to yield the accurate predictions. The model predicts the decrease in MC in millimeters (mm) converting them using the volumetric water content (% VWC). Table 5 presents the results of the LSTM predictions for each of the four investigated phases regarding Field 1. These findings reveal

the model's effectiveness in estimating the reduction in MC within the four designated time periods.

Table 5. Coefficient of determination (R^2) and Root Mean Square Error (RMSE) results of Field 1 on test data.

	Phase 1	**Phase 2**	**Phase 3**	**Phase 4**
R^2	0.8264	0.8163	0.8422	0.9181
$RMSE$	0.9280	1.033	0.7220	0.4908

The findings derived from the LSTM predictive model demonstrate promising performance in estimating the decrease in MC in maize in Field 1 across the four discrete phases. The obtained results demonstrate an upward trend in the efficacy of the model as we move through the phases, a phenomenon that can be attributed to the increasing size of the training dataset. During the initial phase, the coefficient determination (R-squared, R^2) was determined to be 0.8264, which suggests a significant degree of reliability in forecasting the decrease in moisture content. This phase corresponds to the time period from 16 May to 28 June. During this phase the daily reduction of MC was low and exhibited a steady trend. As a result, the model exhibited very good performance during this phase, as the patterns observed in the testing dataset closely mirrored those present in the training data. Moving to the second phase, the R^2 remains similarly high to that at phase 1, at 0.8163. While this value is slightly lower than the first phase, it remains quite impressive, since during this period the reduction in MC was more variable, which can explain the minor reduction in performance compared to the first phase. In the third phase, the model's performance improves more, with an R^2 equal to 0.8422. This increase in performance may be attributed to the continued expansion of the training dataset, giving the model the opportunity to capture a wider range of MC reduction patterns. The final phase demonstrates the highest R^2 of 0.9181. This substantial increase regarding the model's predictive accuracy indicates that the model has become exceptionally adept at forecasting MC reduction, leading to the general observation that the larger training dataset, that conveyed data from all the three previous phases, contributed significantly to enhancing the model's predictive ability for the fourth phase.

Figure 5 illustrates the comparison between the observed values and the predicted values generated by the LSTM model for each of the nine soil sensor locations throughout the first phase in Field 1.

As observed in Figure 5, the reduction in MC at each location is generally minimal, with the exception of the final (ninth) day. This phenomenon occurs due to the relatively low ET rates during these eight days, leading to a subtle reduction of MC. Consequently, the model exhibits a high performance across all locations throughout the initial eight-day period. On the contrary, on the ninth day of the first phase, a slight deviation between the observed and predicted MC value is indicated.

In Figure 6 the comparison between the observed values and the predicted values derived from the LSTM model for each of the nine soil sensor locations during the second phase conducted in Field 1 are depicted. Each panel of the illustration portrays a comparison between the three sensors.

The three panels in Figure 6 exhibit a noticeable variation between the observed and forecasted values for all sensors for the whole duration. In comparison to the first figure, it is evident from the graph depicting the MC reduction over the timeframe including 29 June 2022 to 12 July 2022, that there has been a notable decrease in MC. The considerable rate of MC reduction can be attributed to the significantly elevated levels of evapotranspiration that have been prevalent during this period. One particularly noteworthy observation is that on day 7, the most substantial decrease in MC from all the nine locations in the maize field was recorded. This day appears to have experienced conditions that contribute directly to MC reduction, including high temperatures, increased solar radiation, and lower

relative humidity levels. Figure 7 depicts the comparison between the observed values and the predicted values of MC for the third phase.

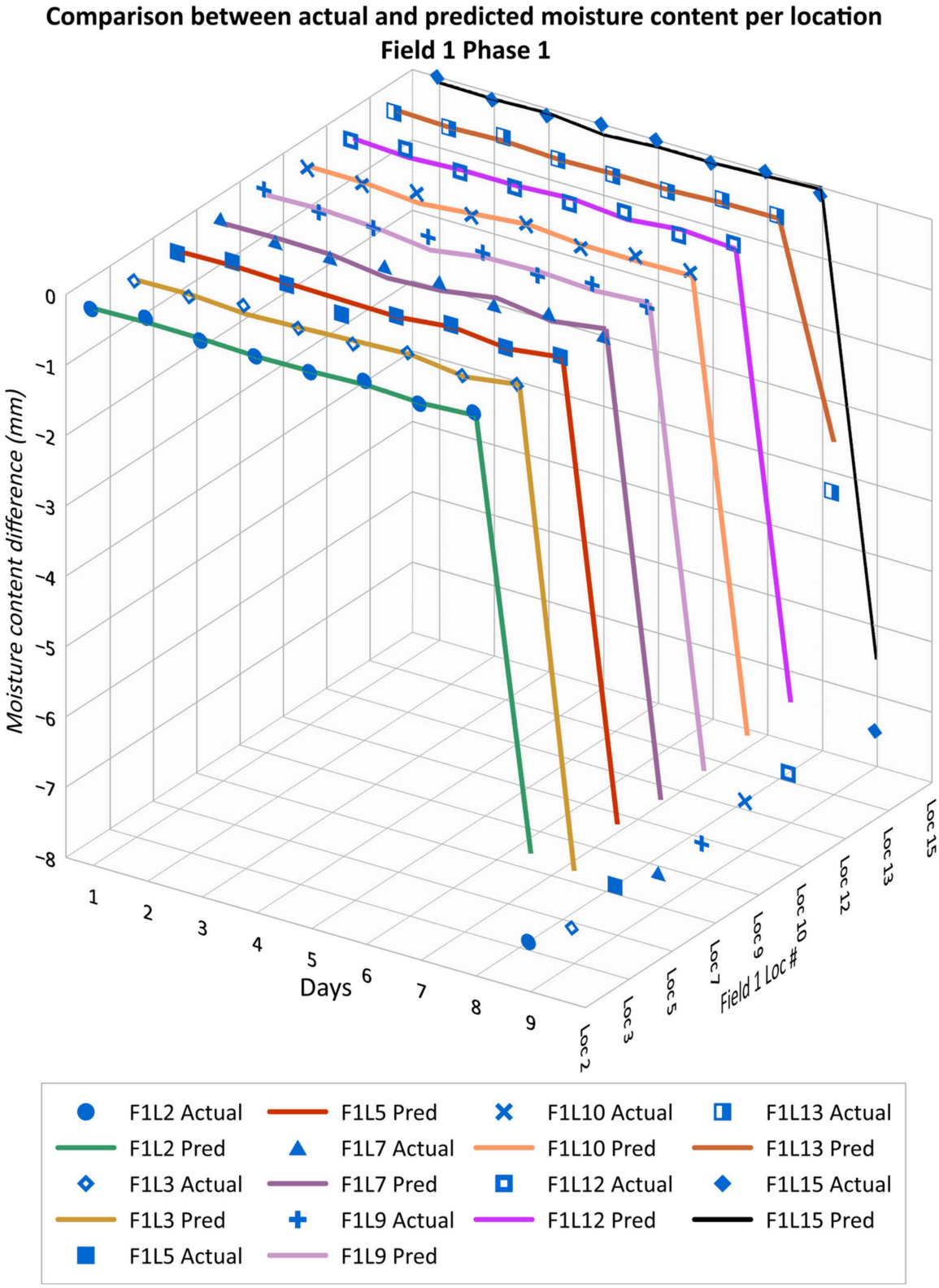
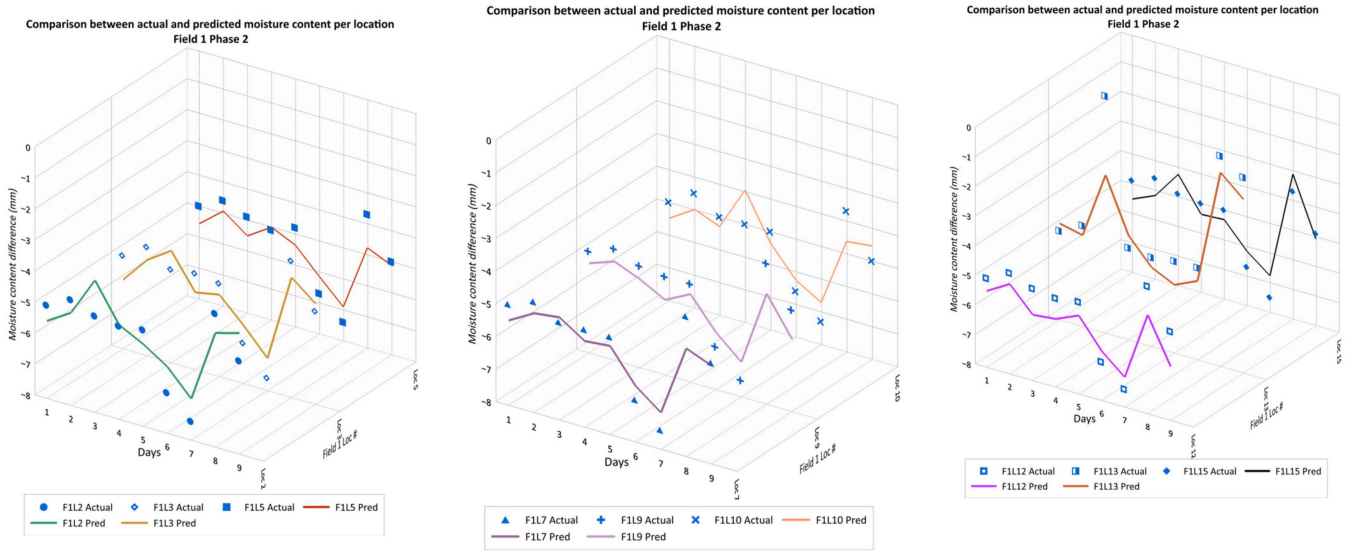
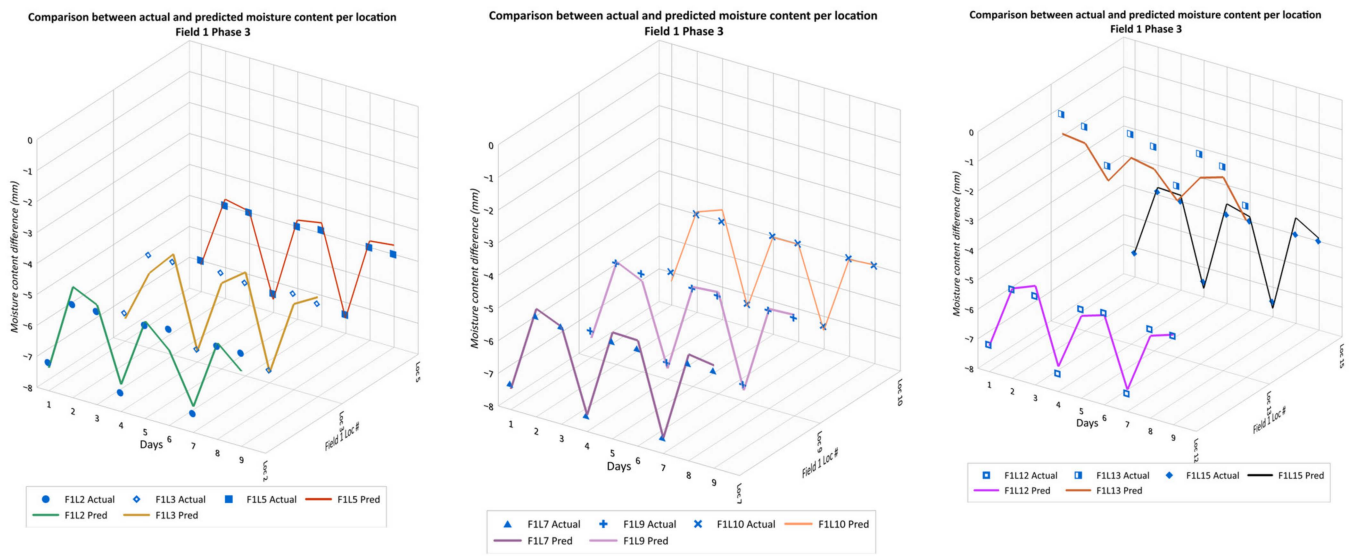


Figure 5. Comparison between the actual and the predicted MC per location in Field 1 during 1st phase.



(**a**) (**b**) (**c**)

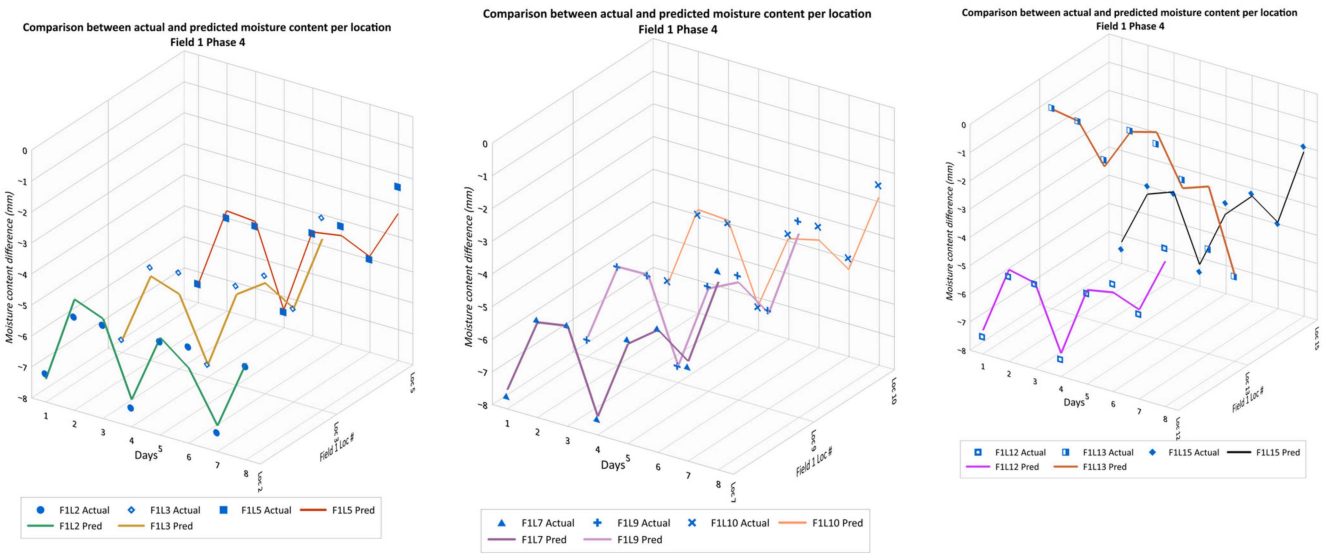
Figure 6. (**a**) Comparison between actual and predicted MC for locations 2, 3 and 5 during 2nd phase; (**b**) Comparison between actual and predicted MC for locations 7, 9 and 10 during 2nd phase; (**c**) Comparison between actual and predicted MC for locations 12, 13 and 15 during 2nd phase.



(**a**) (**b**) (**c**)

Figure 7. (**a**) Comparison between actual and predicted MC for locations 2, 3 and 5 during 3rd phase; (**b**) Comparison between actual and predicted MC for locations 7, 9 and 10 during 3rd phase; (**c**) Comparison between actual and predicted MC for locations 12, 13 and 15 during 3rd phase.

The performance of the LSTM model in the third phase appears to be significantly high for all sensors, except sensor 13. This notably abnormal behavior can be explained by the high content of clay in the soil, which contributed to the significant reduction of MC on a daily basis, and is comparatively lower in this particular location. The presented figure illustrates a consistent and significant daily decrease in MC across all locations. Figure 8 presents a visual representation of the comparison conducted between actual and predicted MC for different locations during the fourth phase.



(**a**) (**b**) (**c**)

Figure 8. (**a**) Comparison between actual and predicted MC for locations 2, 3 and 5 during 4th phase; (**b**) Comparison between actual and predicted MC for locations 7, 9 and 10 during 4th phase; (**c**) Comparison between actual and predicted MC for locations 12, 13 and 15 during 4th phase.

As previously indicated, the estimated values closely align with the observed values in this phase. Moreover, it is noteworthy that the reduction in MC during this phase is of a higher magnitude compared to the previous phases (1st, 2nd and 3rd). The performance of the LSTM model on the testing dataset for Field 2 is shown in Table 6.

Table 6. Coefficient of determination (R^2) and Root Mean Square Error (RMSE) results of Field 2 on test data.

	Phase 1	**Phase 2**	**Phase 3**	**Phase 4**
R^2	0.8392	0.7602	0.7992	0.8417
$RMSE$	1.0242	1.5349	1.2582	0.6682

In contrast to the results observed for Field 1, in the occasion of Field 2 the LSTM model exhibits lower precision in estimating MC reduction. This difference can be attributed to multiple factors, with the limitation of data availability being the main factor. Field 2 demonstrates a comparatively constrained dataset compared to Field 1, so it would be reasonable to assume that this would have an immediate negative impact on the model's ability to accurately capture and extrapolate the patterns of MC reduction. However, it is noteworthy to mention that the performance trends observed in Field 2 seem to follow a similar pattern to those of Field 1. In the first phase, Field 2 exhibited an R^2 value of 0.8392, which indicates a good level of reliability. This phase corresponds to a period from 11 May to 28 June, which is characterized by a small and steady reduction in MC, making it easier for the model to make accurate predictions. However, during the second phase, the value of R^2 decreases to 0.7602. Such a decrease of R^2 in Field 2 can possibly be attributed to the increased variability or to the complexity of MC reduction patterns during this period, similar to Field 1. In the third phase, the model's performance improves slightly with an R^2 equal to 0.7992. Finally, in the fourth phase, Field 2 demonstrates an R^2 equal to 0.8417. The performance of this phase concerning Field 2 closely resembles that of Field 1, indicating the model's improvement, attributed to the inclusion of a bigger training dataset that incorporates data from the three previous phases.

Figure 9 illustrates the comparison between the observed values and the predicted values generated by the model for each of the six soil locations throughout the first phase in Field 2.

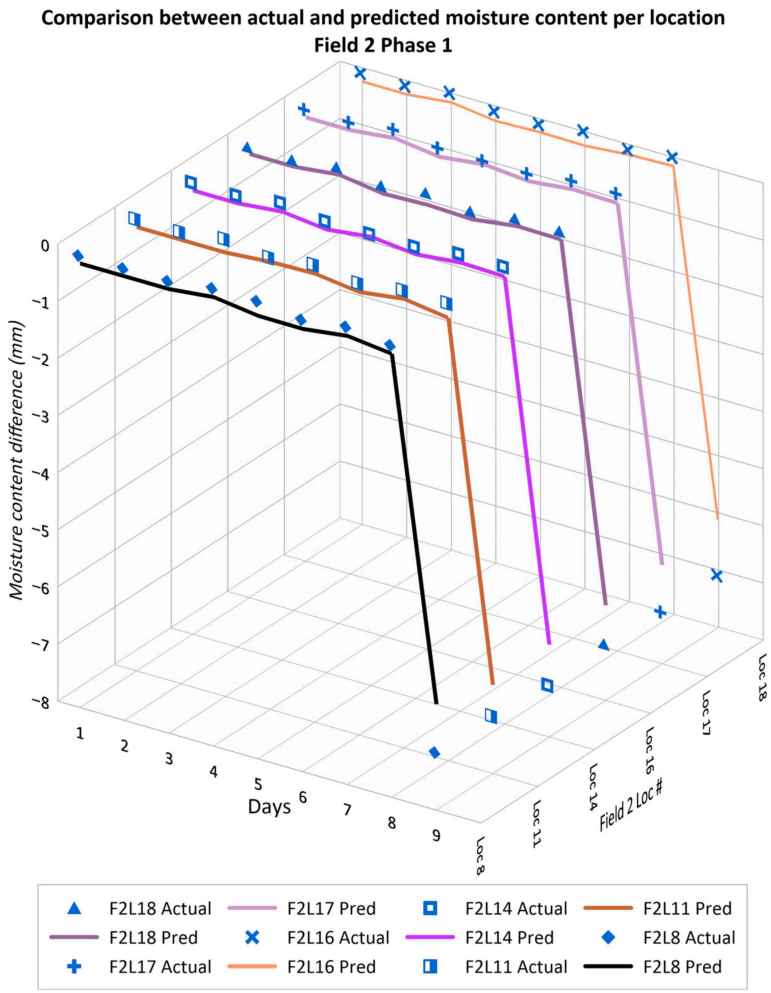
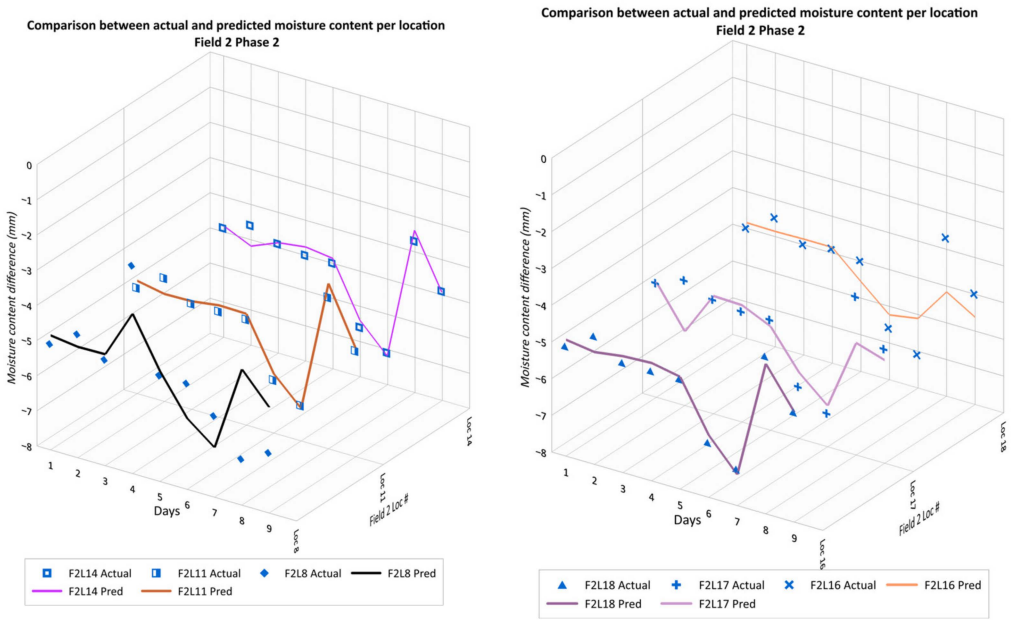


Figure 9. MC comparison per location in Field 2 during phase 1.

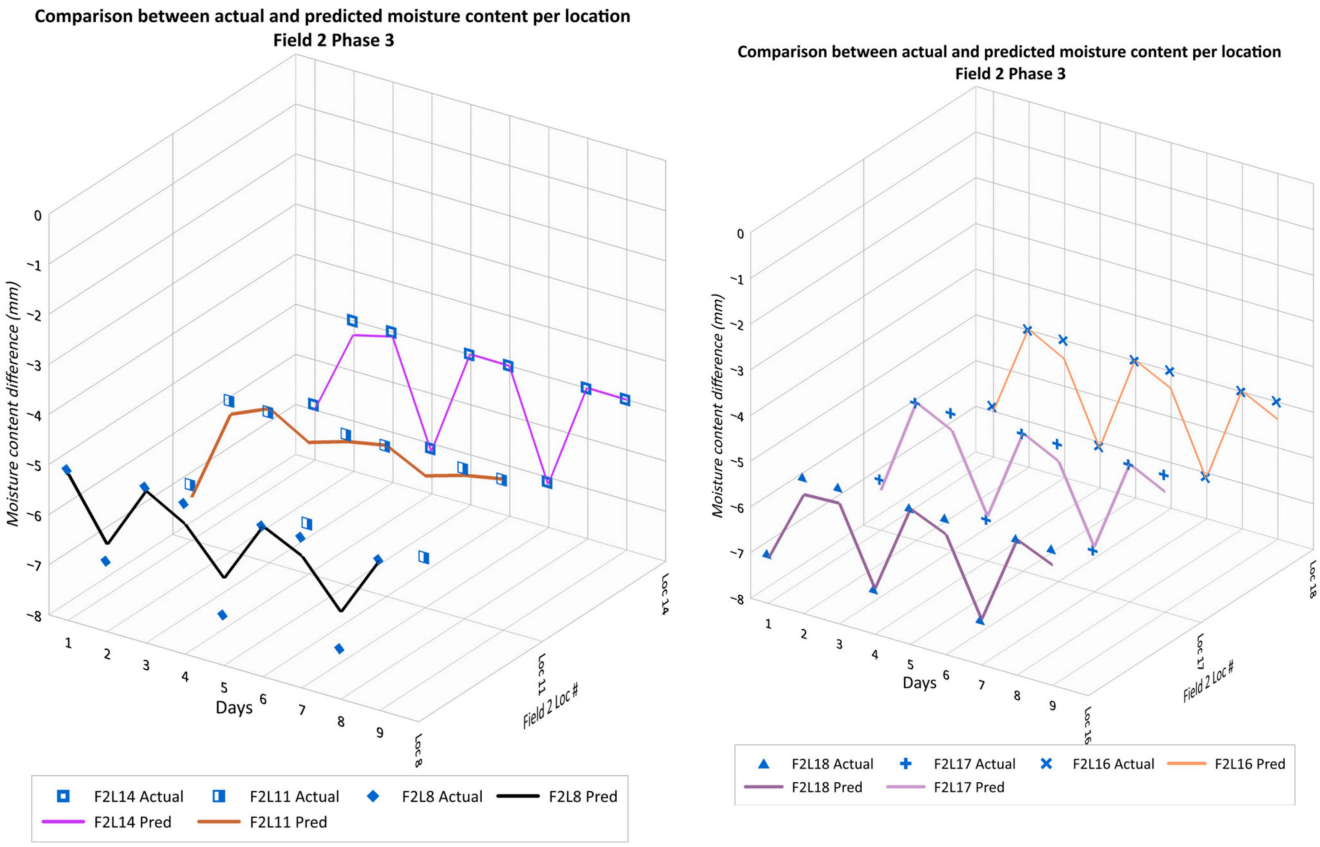
Figure 10 depicts the comparative outcomes between the actual and predicted MC after the second phase in Field 2.

At this phase, the model indicates satisfactory results in predicting the reduction of MC in locations 11, 14, and 18. For the other locations, the model appeared less accurate in predicting the MC reduction values for day eight. The uniformity in soil MC reduction during the initial five days, followed by variations in the subsequent four days across all six sensor locations, can be explained due to the occurrence of several interacting factors during the four different phases that contributed to the reduction of MC. More specifically, weather conditions, including but not limited to air temperature, humidity and solar radiation, hold a pivotal role over the rate of evapotranspiration, which captures both plant transpiration and soil evaporation. The constancy in moisture reduction during the first five days may be attributed at least partially to the consistent meteorological conditions prevailing during the investigated time period. Finally, the progressive growth of maize plants over time, followed by the extension and deepening of their root systems, contributes to the inevitable decrease of MC since the plants have a greater ability to capture water and nutrients from deeper soil layers. Consequently, the maize crop development tends to significantly alter the soil moisture kinetics patterns. Figures 11 and 12 illustrate the comparison of the observed values and the predicted values produced by the model for each soil location during the third and fourth phases in Field 2.



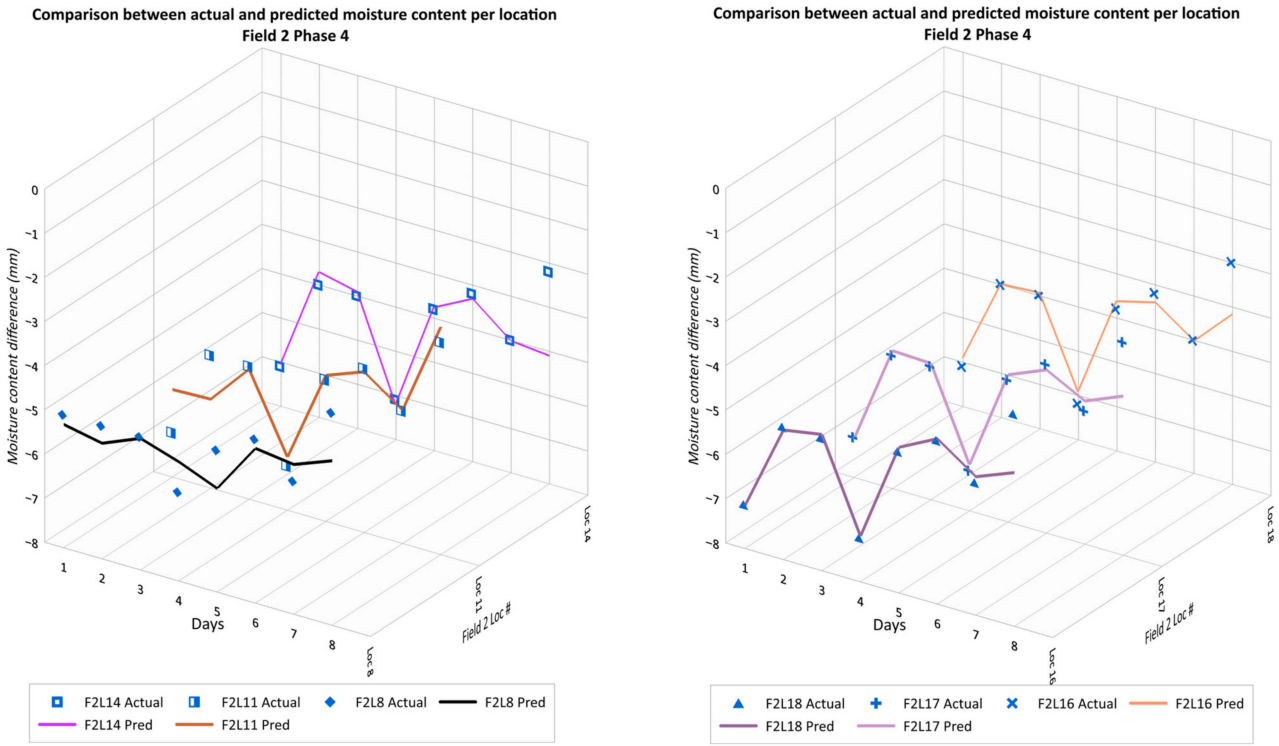
(**a**) (**b**)

Figure 10. Comparison between actual and predicted MC for locations (**a**) 8, 11 and 14 during 2nd phase and (**b**) 16, 17 and 18 during 2nd phase.



(**a**) (**b**)

Figure 11. Comparison between actual and predicted MC for locations (**a**) 8, 11 and 14 during 3rd phase and (**b**) 16, 17 and 18 during 3rd phase.



(**a**) (**b**)

Figure 12. Comparison between actual and predicted MC for locations (**a**) 8, 11 and 14 during 4th phase and (**b**) 16, 17 and 18 during 4th phase.

Similar to Field 1, the daily changes are substantial due to the heightened evapotranspiration during these stages and the greater frequency of irrigation by the farmer. Additionally, the model's performance exhibits improvement during phases 3 and 4, similar to what was observed in Field 1.

In summary, it can be concluded that the LSTM model exhibited encouraging potential forecasting the reduction of MC in maize fields. The results validated that the performance of the LSTM model was greatly influenced by the availability and quality of data confirming previously conducted research in the field [37], which utilized a Bi-LSTM model, and demonstrated a notable R^2 improvement, benefiting from a comprehensive dataset spanning 14 years. A more extended period of investigation would have ideally offered a vast array of data for model training through the extensive historical datasets [38] to ensure robust predictive capabilities. However, in the current study, irrigation needs have been effectively predicted despite the challenge that has been encountered due to the soil data measurements' discontinuity, revealing that predictions for irrigation scheduling can be made without the existence of extensive historical datasets that often require a span of several years to decades to yield high predictive performances. The high performance of the LSTM model is attributed to its special asset among traditional RNNs, characterized by its special gates that are responsible for controlling the flow, the volume, and the type of the previously attained and processed information that is going to be maintained and forwarded from the memory cell to the next hidden state. The utilization of the Aquacrop 7.0 model to simulate soil MC data due to data limitations in the investigated fields, equipped the LSTM model with a consistent and adequate volume of data in order to perform precise predictions for irrigation scheduling in maize crops, confirming its effectiveness towards the accurate and reliable simulation of moisture content (MC) [17] that is also mirrored in our findings.

For both investigated fields, the LSTM model becomes more efficient in predicting MC reduction as we move through the four investigated phases, since the dataset is enriched

with data from previous phases. The high performance of the model has been assisted by the sufficient number of soil sensors which provided a satisfactory volume of measurements. Comparing the performance of the LSTM model between the two investigated fields, the results appear slightly lower for Field 2 compared to those for Field 1, indicating an R^2 ranging from 0.7602 to 0.8417 can be explained by the limited data availability. However, Field 2 continues to show similar behavior in its predictive ability, close to that of the LSTM model demonstrated for Field 1. It is worth noting that despite the lack of a sufficient number of sensors, there were only six for Field 2, the observed variations in MC reduction occur mostly due to the same factors that are mostly related to the weather conditions and the expected crop growth over time.

Incorporating these methods into a comprehensive system, such as the predictive irrigation scheduling system we employed, underscores the value of integrating real-time data with advanced modelling techniques. This integration allows for more effective irrigation strategies, aligning with the broader goals of sustainable agriculture and water resource management. The employed LSTM model has been proven capable of maintaining its predictive accuracy, highlighting the potential for real-time, season-specific irrigation management. This advancement is particularly significant in precision agriculture, where timely and accurate predictions can lead to more efficient water usage and improved crop management strategies.

The further enhancement of the Aquacrop model with more comprehensive field data, including precise measurements of root distribution and Leaf Area Index (LAI), could substantially elevate its accuracy in simulating crop water needs. This improvement will enable more precise irrigation scheduling, optimizing water usage while ensuring optimal crop growth and yield. In the future, additional studies will include the integration of detailed data regarding the root systems and canopy characteristics, that are expected to assist the model in reflecting the actual water uptake and evapotranspiration patterns of crops in a better and more detailed manner, leading to more precise predictions regarding the irrigation needs.

4. Conclusions

The basis of the present study has been the soil moisture data which exhibited non-stationary time series, characterized by a regular periodic pattern with significant fluctuations. The above challenge has been faced through the application of the LSTM model. The LSTM model has been developed to forecast the reduction of soil MC, focusing on the needed amount of water so as to assess and cover the required irrigation needs in maize crops. The efficiency of the LSTM model has been evaluated based on its forecasting abilities on the impact of weather data and satellite extracted data on the reduction of MC in four distinct time phases. Based on the results demonstrated above, the performance of the proposed model has been estimated as high for both investigated fields despite the limited availability of soil sensors in Field 2 compared to Field 1. This demonstrates the fact that the LSTM model is a powerful tool for forecasting soil moisture dynamics in grain fields. Moreover, the LSTM's adaptability and robustness in predicting the MC reduction, encourages its potential applicability to a variety of other precision agriculture practices that require effective handling of sequential data such as crop growth monitoring and crop yield prediction. The effectiveness of the LSTM model in accurately predicting the soil MC decrease holds the potential of revolutionizing agriculture, offering several benefits to farmers, agronomists, and policymakers. More specifically, through the adoption of such site-specific irrigation patterns enabled by RNN models, Variable Rate Irrigation (VRI) application schemes can be adopted, securing and enhancing yield annual productivity in grains. Through its integration to decision support systems (DSSs), the LSTM model can provide a variety of real-time recommendations in the precision agriculture domain, due to the model's efficiency of taking into account a wide range of factors, such as weather conditions and historical data, enabling dynamic and efficient irrigation. The above-mentioned ability can be considered as a special asset of the LSTM model in securing and improving

decision-making in the field of precision agriculture. Consequently, farmers can optimize their irrigation scheduling practices based on accurate forecasts and policymakers are given the opportunity to use these forecasts for evidence-based decisions towards efficient irrigation water management and sustainable agriculture in general. Based on the results of this study, future work could focus on the design of a predictive irrigation approach, where crucial components of crop growth are integrated so as to have a better understanding of the future irrigation needs of crops based on the growth stage.

Author Contributions: Conceptualization, K.D., X.E.P. and C.P.; methodology, K.D. and X.E.P.; software, K.D. and C.P.; validation, X.E.P.; formal analysis, K.D., X.E.P. and C.P.; data curation, K.D. and C.P.; writing—original draft preparation, K.D., X.E.P. and C.P.; writing—review and editing, X.E.P., K.D., C.P., M.Q., D.B. and A.M.M.; supervision, X.E.P.; resources: S.A., Y.T., K.S.G., Y.U. and B.B.B.; investigation S.A. and Y.T.; project administration, A.M.M.; funding acquisition, A.M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been co-financed by the European Regional Development Fund (ERDF), the Greek national fund General Secretariat for Research and Innovation through the Partnership Agreement for the Development Framework 2014–2020 (Action: ERANETS—2021A), the Scientific and Technological Research Council of Turkey—TUBITAK (contract Nr: 120N787) and Fonds Wetenschappelijk Onderzoek (FWO) (Contract no. S007621N).

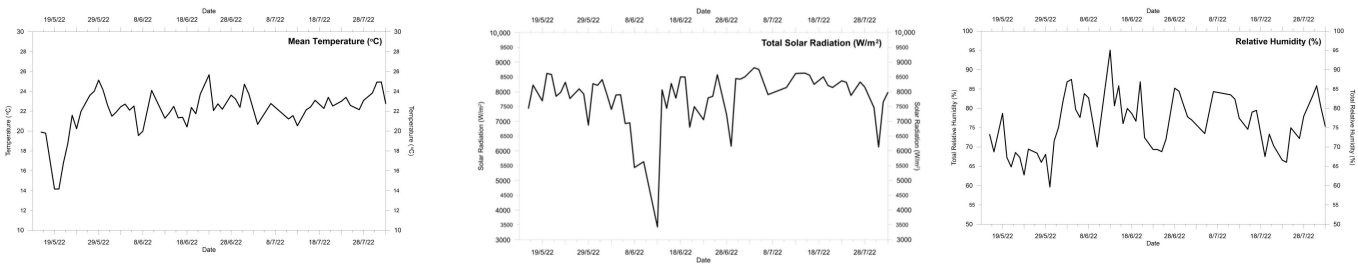
Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

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

Appendix A

In the following figures the data (measured and interpolated) that were used for training the LSTM network are visually presented.



(**a**) (**b**) (**c**)

Figure A1. Daily values of mean temperature (**a**), total solar radiation (**b**) and relative humidity (**c**), observed over the study period.

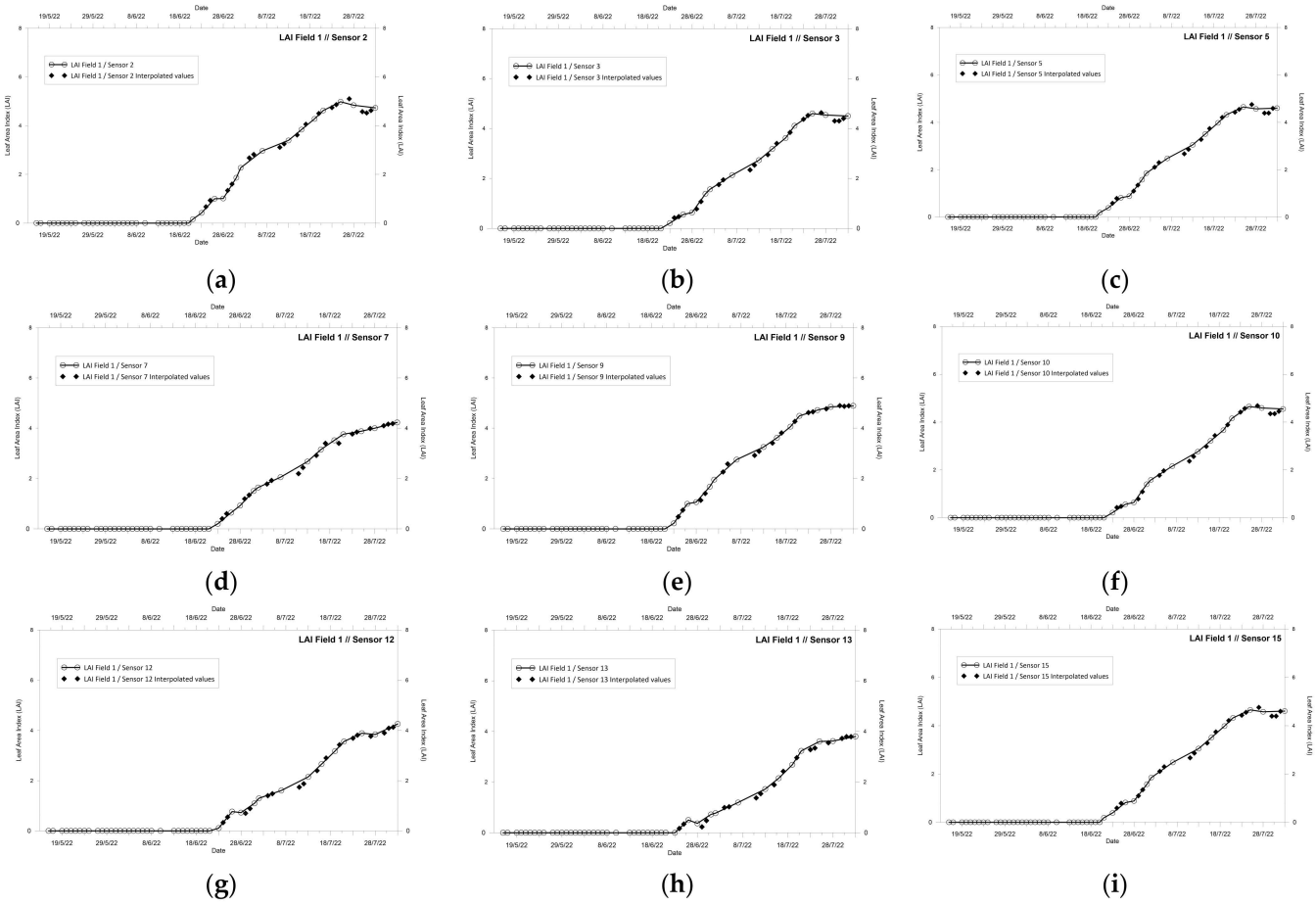


Figure A2. Leaf Area Index measured and interpolated values for sensors 2, 3, 5, 7, 9, 10, 12, 13 and 15 (respectively **a**,**b**,**c**,**d**,**e**,**f**,**g**,**h**,**i**) of Field 1.

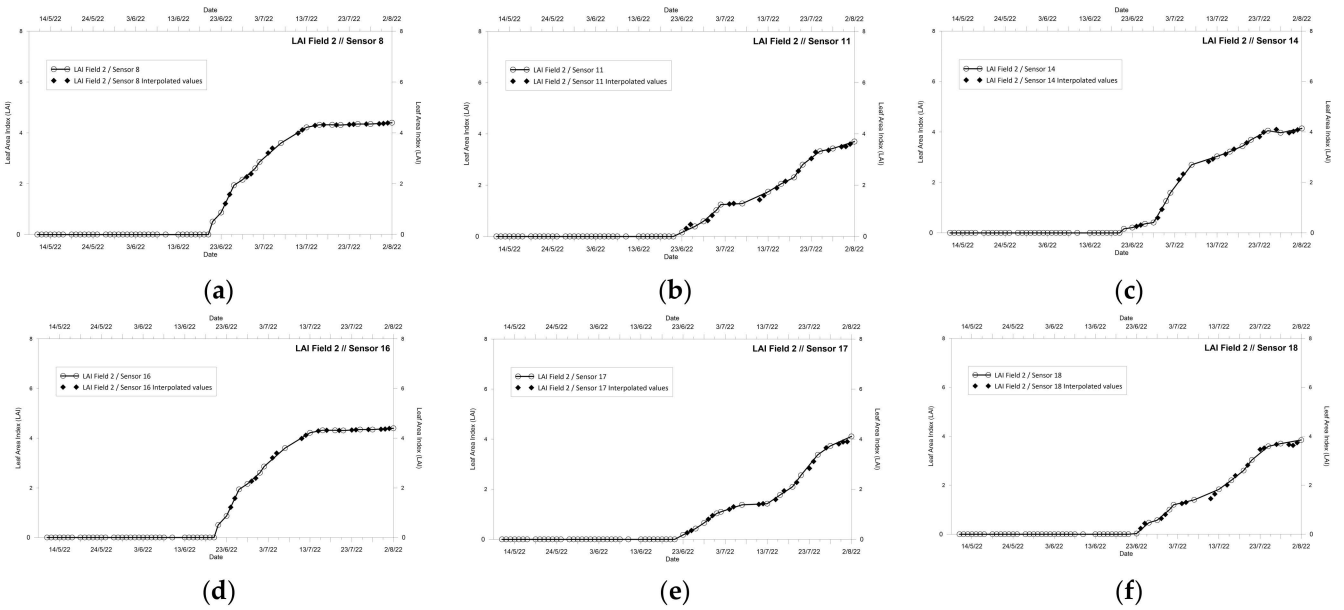


Figure A3. Leaf Area Index measured and interpolated values for sensors 8, 11, 14, 16, 17 and 18 (respectively **a**,**b**,**c**,**d**,**e**,**f**) of Field 2.

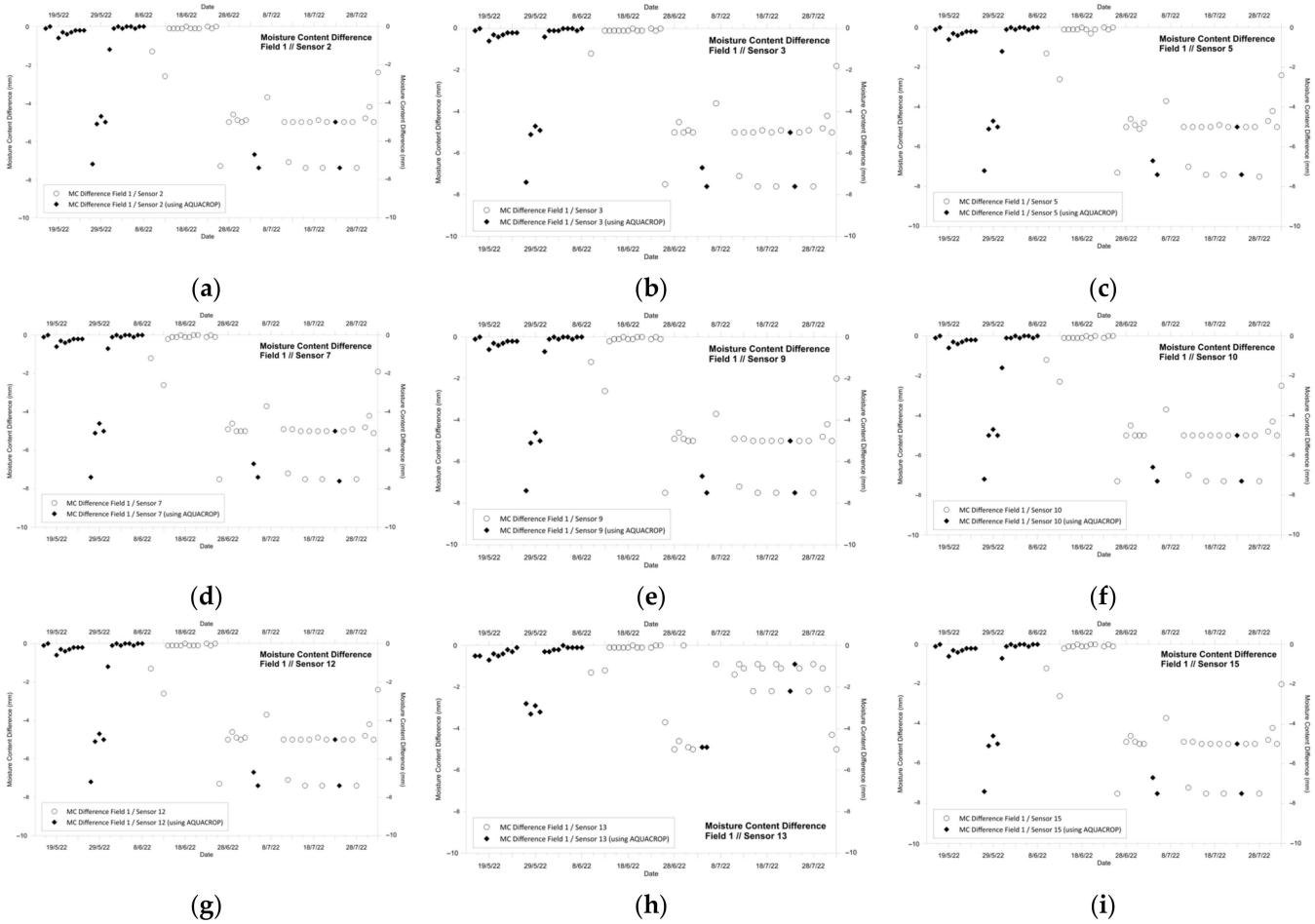


Figure A4. Moisture Content difference using measured and simulated by AQUACROP values for sensors 2, 3, 5, 7, 9, 10, 12, 13 and 15 (respectively **a**,**b**,**c**,**d**,**e**,**f**,**g**,**h**,**i**) of Field 1.

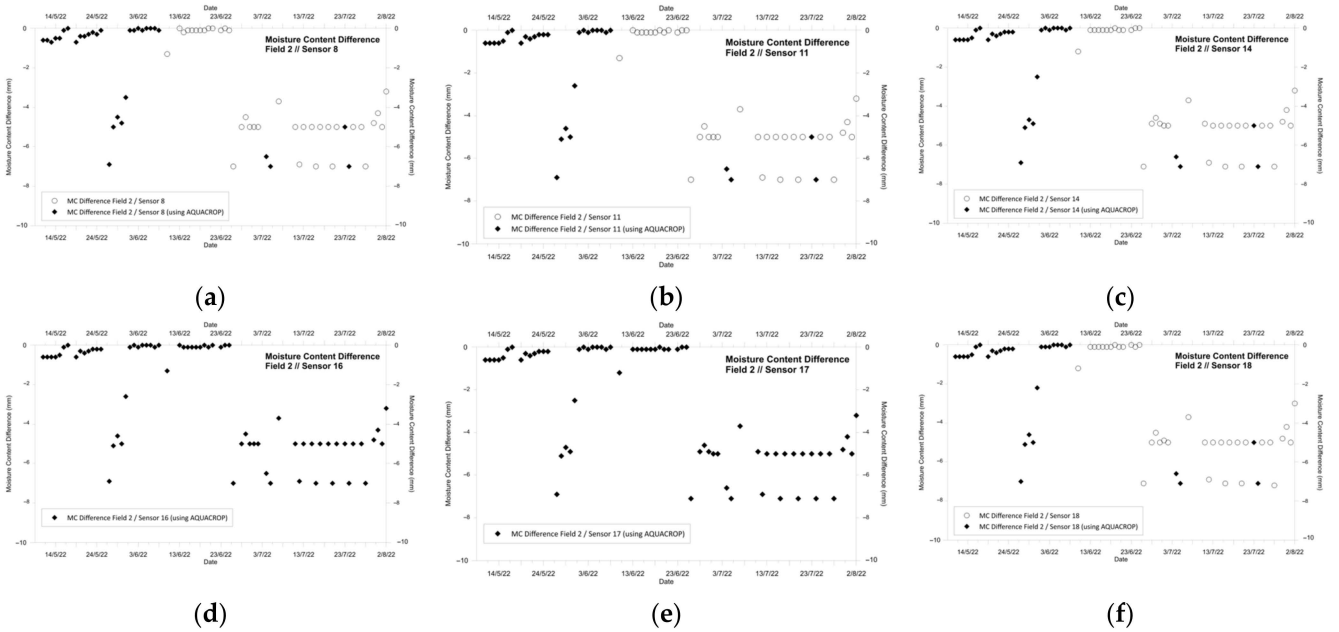


Figure A5. Moisture Content difference using measured and simulated by AQUACROP values for sensors 8, 11, 14, 16, 17 and 18 (respectively **a**,**b**,**c**,**d**,**e**,**f**) of Field 2.

References

1. Ramachandran, V.; Ramalakshmi, R.; Kavin, B.P.; Hussain, I.; Almaliki, A.H.; Almaliki, A.A.; Elnaggar, A.Y.; Hussein, E.E. Exploiting IoT and its enabled technologies for irrigation needs in agriculture. *Water* **2022**, *14*, 719. [CrossRef]
2. Guan, Y.; Grote, K.; Schott, J.; Leverett, K. Prediction of soil water content and electrical conductivity using random forest methods with UAV multispectral and ground-coupled geophysical data. *Remote Sens.* **2022**, *14*, 1023. [CrossRef]
3. Dumedah, G.; Coulibaly, P. Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data. *J. Hydrol.* **2011**, *400*, 95–102. [CrossRef]
4. Lin, S.L. Application of machine learning to a medium Gaussian support vector machine in the diagnosis of motor bearing faults. *Electronics* **2021**, *10*, 2266. [CrossRef]
5. Wu, T.; Zhang, W.; Jiao, X.; Guo, W.; Hamoud, Y.A. Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration. *Comput. Electron. Agric.* **2022**, *184*, 106039. [CrossRef]
6. Karandish, F.; Šimůnek, J. A comparison of numerical and machine-learning modeling of soil water content with limited input data. *J. Hydrol.* **2016**, *543*, 892–909. [CrossRef]
7. Hong, Z.; Kalbarczyk, Z.; Iyer, R.K. Using a wireless sensor network and machine learning techniques. In Proceedings of the 2016 IEEE International Conference on Smart Computing (SMARTCOMP), St. Louis, MO, USA, 18–20 May 2016.
8. Adeyemi, O.; Grove, I.; Peets, S.; Domun, Y.; Norton, T. Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling. *Sensors* **2018**, *18*, 3408. [CrossRef]
9. Tsang, S.W.; Jim, C.Y. Applying artificial intelligence modeling to optimize green roof irrigation. *Energy Build.* **2016**, *127*, 360–369. [CrossRef]
10. Gu, J.; Yin, G.; Huang, P.; Guo, J.; Chen, L. An improved back propagation neural network prediction model for subsurface drip irrigation system. *Comput. Electr. Eng.* **2017**, *60*, 58–65. [CrossRef]
11. Ramos, M.M.P.; Del Alamo, C.L.; Zapana, R.A. Forecasting of meteorolog1ical weather time series through a feature vector based on correlation. In Proceedings of the Computer Analysis of Images and Patterns: 18th International Conference, Salerno, Italy, 3–5 September 2019; Part I 18; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 542–553.
12. Farzad, M.; Shafieifar, M.; Azizinamini, A. Experimental and numerical study on bond strength between conventional concrete and Ultra High-Performance Concrete (UHPC). *Eng. Struct.* **2019**, *186*, 297–305. [CrossRef]
13. Agyeman, B.T.; Naouri, M.; Appels, W.; Liu, J. Irrigation management zone delineation and optimal irrigation scheduling for center pivot irrigation systems. *IFAC Pap.* **2023**, *56*, 9906–9911. [CrossRef]
14. Jenitha, R.; Rajesh, K. Intelligent irrigation scheduling scheme based on deep bi-directional LSTM technique. *Int. J. Environ. Sci. Technol.* **2023**, *21*, 1905–1922. [CrossRef]
15. FAO. *AquaCrop Stand-Alone Program, Version 7.0. 2022*; Food and Agriculture Organization of the United Nations, Land and Water Division: Rome, Italy, 2022; Available online: https://www.fao.org/aquacrop/software/aquacropplug-inprogramme/en/ (accessed on 8 October 2022).
16. Sandhu, R.; Irmak, S. Performance of AquaCrop model in simulating maize growth, yield, and evapotranspiration under rainfed, limited and full irrigation. *Agric. Water Manag.* **2019**, *223*, 105687. [CrossRef]
17. Zhao, H.; Di, L.; Guo, L.; Zhang, C.; Lin, L. An Automated Data-Driven Irrigation Scheduling Approach Using Model Simulated Soil Moisture and Evapotranspiration. *Sustainability* **2023**, *15*, 12908. [CrossRef]
18. Türkler, L.; Akkan, T.; Akkan, L.Ö. Detection of Water Leakage in Drip Irrigation Systems Using Infrared Technique in Smart Agricultural Robots. *Sensors* **2023**, *23*, 9244. [CrossRef] [PubMed]
19. Sharu, E.H.; Ab Razak, M.S. Hydraulic Performance and Modelling of Pressurized Drip Irrigation System. *Water* **2020**, *12*, 2295. [CrossRef]
20. Mouazen, A.M. Soil survey device. In *International Publication Published under the Patent Cooperation Treaty (PCT)*; International Publication Number: WO2006/015463; World Intellectual Property Organization, International Bureau: Geneva, Switzerland, 2006.
21. Dragino. LSE01-LoRaWAN Soil Moisture & EC Sensor User Manual. Available online: https://www.dragino.com/products/agriculture-weather-station/item/277-se01-lb.html (accessed on 17 August 2023).
22. Kganyago, M.; Mhangara, P.; Alexandridis, T.; Laneve, G.; Ovakoglou, G.; Mashiyi, N. Validation of sentinel-2 leaf area index (LAI) product derived from SNAP toolbox and its comparison with global LAI products in an African semi-arid agricultural landscape. *Remote Sens. Lett.* **2020**, *11*, 883–892. [CrossRef]
23. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Daloye, A.M.; Erkbol, H.; Fritschi, F.B. Crop monitoring using satellite/UAV data fusion and machine learning. *Remote Sens.* **2020**, *12*, 1357. [CrossRef]
24. Bianchi, F.M.; Maiorino, E.; Kampffmeyer, M.C.; Rizzi, A.; Jenssen, R. *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis*; Springer: Cham, Switzerland, 2017.
25. Nguyen, A.D.; Le Nguyen, P.; Vu, V.H.; Pham, Q.V.; Nguyen, V.H.; Nguyen, M.H.; Nguyen, T.H.; Nguyen, K. Accurate discharge and water level forecasting using ensemble learning with genetic algorithm and singular spectrum analysis-based denoising. *Sci. Rep.* **2022**, *12*, 19870. [CrossRef]
26. Kratzert, F.; Klotz, D.; Brenner, C.; Schulz, K.; Herrnegger, M. Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 6005–6022. [CrossRef]
27. Ishfaque, M.; Dai, Q.; Haq, N.U.; Jadoon, K.; Shahzad, S.M.; Janjuhah, H.T. Use of recurrent neural network with long short-term memory for seepage prediction at Tarbela Dam, KP, Pakistan. *Energies* **2022**, *15*, 3123. [CrossRef]

28. Belete, D.M.; Huchaiah, M.D. Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *Int. J. Comput. Appl.* **2022**, *44*, 875–886. [CrossRef]
29. Yu, C.; Qi, X.; Ma, H.; He, X.; Wang, C.; Zhao, Y. LLR: Learning learning rates by LSTM for training neural networks. *Neurocomputing* **2020**, *394*, 41–50. [CrossRef]
30. Zhang, J.; Hu, F.; Li, L.; Xu, X.; Yang, Z.; Chen, Y. An adaptive mechanism to achieve learning rate dynamically. *Neural Comput. Appl.* **2019**, *31*, 6685–6698. [CrossRef]
31. Yin, H.; Jin, D.; Gu, Y.H.; Park, C.J.; Han, S.K.; Yoo, S.J. STL-ATTLSTM: Vegetable price forecasting using STL and attention mechanism-based LSTM. *Agriculture* **2020**, *10*, 612. [CrossRef]
32. Achieng, K.O. Modelling of soil moisture retention curve using machine learning techniques: Artificial and deep neural networks vs. support vector regression models. *Comput. Geosci.* **2022**, *133*, 104320. [CrossRef]
33. Kour, K.; Gupta, D.; Rashid, J.; Gupta, K.; Kim, J.; Han, K.; Mohiuddin, K. Smart Framework for Quality Check and Determination of Adulterants in Saffron Using Sensors and AquaCrop. *Agriculture* **2023**, *13*, 776. [CrossRef]
34. Wang, F.; Xue, J.; Xie, R.; Ming, B.; Wang, K.; Hou, P.; Zhang, L.; Li, S. Assessing growth and water productivity for drip-irrigated maize under high plant density in arid to semi-humid climates. *Agriculture* **2022**, *12*, 97. [CrossRef]
35. Schaap, M.G.; Leij, F.J.; van Genuchten, M.T. ROSETTA: A Computer Program for Estimating Soil Hydraulic Parameters with Hierarchical Pedotransfer Functions. *J. Hydrol.* **2001**, *251*, 163–176. [CrossRef]
36. Van Genuchten, M.T. A Closed-Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Sci. Soc. Am. J.* **1980**, *44*, 892–898. [CrossRef]
37. Zhang, J.; Zhu, Y.; Zhang, X.; Ye, M.; Yang, J. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *J. Hydrol.* **2018**, *561*, 918–929. [CrossRef]
38. Alibabaei, K.; Gaspar, P.D.; Lima, T.M. Modeling soil water content and reference evapotranspiration from climate data using deep learning method. *Appl. Sci.* **2021**, *11*, 5029. [CrossRef]