applied sciences

MDPI

Article

Evaluation of a Deep Learning Approach for Predicting the Fraction of Transpirable Soil Water in Vineyards

Khadijeh Alibabaei [1,2,†], Pedro D. Gaspar [1,2], Rebeca M. Campos [3], Gonçalo C. Rodrigues [3,*] and Carlos M. Lopes [3]

1 C-MAST Center for Mechanical and Aerospace Science and Technologies, University of Beira Interior, 6201-001 Covilhã, Portugal
2 Deparment of Electromechanical Engineering, University of Beira Interior, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal
3 Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, 1349-017 Lisboa, Portugal
* Correspondence: gcrodrigues@isa.ulisboa.pt
† Current address: Steinbuch Centre for Computing, Zirkel 2, D-76131 Karlsruhe, Germany.

Abstract: As agriculture has an increasing impact on the environment, new techniques can help meet future food needs while maintaining or reducing the environmental footprint. Those techniques must incorporate a range of sensing, communication, and data analysis technologies to make informed management decisions, such as those related to the use of water, fertilizer, pesticides, seeds, fuel, labor, etc., to help increase crop production and reduce water and nutrient losses, as well as negative environmental impacts. In this study, a Bidirectional Long Short-Term Memory (BiLSTM) model was trained on real data from Internet of Things sensors in a vineyard located in the Douro wine-growing region, from 2018–2021, to evaluate the ability of this model to predict the Fraction of Transpirable Soil Water (FTSW). The model uses historical data, including reference evapotranspiration, relative humidity, vapor pressure deficit, and rainfall, and outputs the FTSW for periods of one, three, five, and seven days. The model achieved an RMSE between 8.3% and 16.6% and an R^2-score between 0.75 and 0.93. The model was validated on an independent dataset collected in 2002–2004 from a different vineyard located in the Lisbon wine-growing region, Portugal, and achieved an R^2-score of 87% and an RMSE of 10.36%. Finally, the performance of the FTSW in the vineyard prediction model was compared with that of the Random Forest model, support vector regression, and linear regression. The results showed that BiLSTM performed better than the RF model on the unseen data, and the BiLSTM model can be considered a suitable model for the accurate prediction of the FTSW.

Keywords: agriculture; FTSW; deep learning; LSTM; BiLSTM; support decision-making algorithms

Citation: Alibabaei, K.; Gaspar, P.D.; Campos, R.M.; Rodrigues, G.C.; Lopes, C.M. Evaluation of a Deep Learning Approach for Predicting the Fraction of Transpirable Soil Water in Vineyards. *Appl. Sci.* **2023**, *13*, 2815. https://doi.org/10.3390/app13052815

Academic Editors: Gniewko Niedbała, Tomasz Wojciechowski and Katarzyna Pentoś

Received: 18 January 2023
Revised: 16 February 2023
Accepted: 17 February 2023
Published: 22 February 2023

1. Introduction

Agriculture plays an important role in environmental degradation due to its extensive use of fertilizers and pesticides, high water and energy consumption, homogenization of the agricultural mosaic, and the loss of biodiversity [1–3]. The new concept of "smart farms" refers to management techniques that use advanced technologies to improve food quantity, quality, and safety while minimizing environmental impacts [4,5]. The purpose of smart farming includes increasing farming sustainability and productivity, improving the knowledge and skills of the agricultural workforce, optimizing costs, reducing waste, creating smart agriculture information and Decision Support Systems (DSSs), reducing greenhouse gas emissions, and developing resilience and adaptation to climate change [4–8].

Internet of Things (IoT) sensors and Machine Learning algorithms (ML) are technology sectors related to smart farming [8]. The IoT is a worldwide network of physical items that are outfitted with sensors and actuators that link to the Internet in real-time in order to be detected, sensed, and controlled remotely [9]. The sensors are used to monitor soil water

content, climate parameters, animal behavior, cultural practices, etc. ML is a subfield of artificial intelligence that analyses data, learns from data, recognizes patterns, and makes predictions with little human intervention. It is used to improve the intelligence of the sensors and analyze the large datasets retrieved from the sensors. Ultimately, it can support the design of DSSs [4,5,10].

A subclass of ML known as Deep Learning (DL) attempts to discover usable representations of the unknown structure in the input distribution, often at many levels, using the learned lower-level features as the input to the higher-level features [11,12]. Compared to traditional ML models, DL algorithms have more hidden layers and can automatically extract features from the input data. Specific applications of DL in agriculture management include disease detection [13–15], fruit detection and yield prediction [16–18], weed detection [19–21], soil management [22,23], and water management [24,25].

In recent years, several studies have attempted to create a DSS for irrigation management using ML and DL algorithms. These algorithms provided estimates for irrigation-related variables such as soil moisture content and evapotranspiration. In a study conducted by Acharya et al. [26], ML algorithms including Multiple Linear Regression (MLR), Support Vector Regression (SVR), Boosted Regression Trees (BRTs), Random Forest Regression (RFR), multiple regression trees (CART), and Artificial Neural Networks (ANNs) were used to predict soil moisture in the Red River Valley of the North in North Dakota and Minnesota, where corn, soybeans, wheat, sugar beets, barley, canola, and potatoes are grown. The best results were obtained with the RFR and BRT algorithms. To predict soil moisture measured by Soil Moisture Active Passive (SMAP) V3 satellite observations using the LSTM method, Fang et al. [**?**] used atmospheric influences, static physiographic attributes such as the sand, silt, and clay fractions, bulk density, and soil capacity and simulated soil moisture from the Noah hydrologic model [**?**]. Using model simulations, the approach largely removes biases and improves predicted soil moisture by achieving low test root-mean-squared errors (0.035) and high correlation coefficients of more than 0.85. Paul and Singh [29] predicted soil moisture for a period of 12 to 13 weeks in advance using machine learning approaches such as linear regression, support vector machine regression, Principal Component Analysis (PCA), and naive Bayes. These methods were applied to four separate datasets from 13 different districts in West Bengal and to four different crops (potato, mustard, rice, and cauliflower). The best model was Support Vector Machine (SVM), followed by linear regression models, which outperformed the PCA and naive Bayes models. In Adeyemi et al. [30], using historical data on soil moisture, rainfall, and meteorological variables, a neural network was trained to predict volumetric soil moisture content one day in advance. Three study sites with land covered in farmland, arable land, and grassland provided the data for building LSTM models to predict soil moisture. An R^2-score greater than 0.94 was achieved in the model evaluation at each test site. Hajjar et al. [31] proposed the use of two nonlinear regression models, Multilayer Perceptron (MLP) and support vector regression, to estimate vineyard soil moisture from digital photographs. Both models were trained using pixels encoded with the RGB color model extracted from digital photographs of the soil, combined with the associated known soil moisture levels, to predict moisture content from newly acquired images. In the analysis of the test data, both approaches were successful in predicting soil moisture and provided good correlation coefficients between observed and predicted soil moisture. An LSTM model was employed by Zhang et al. [32] to predict groundwater level rise in agriculture. They used 14 years of time series data, including information on water table depth, temperature, reference evapotranspiration, rainfall, and water diversion. In terms of forecasting the depth of the water table, the proposed model outperformed the Feed-Forward Neural Network (FFNN) model in terms of the R^2-score.

In other research, the timing and volume of irrigation are directly predicted. A Deep Q-learning (DQN) model was applied by Chen et al. [33] to a rice irrigation decision strategy based on short-term weather forecasts. Daily recorded weather information from three stations during the rice growing season was used as the input of the model. The DQN

irrigation approach showed a significant reduction in irrigation water volume, irrigation time, and drainage water without yield loss compared to the results of the conventional irrigation strategy. The DQN model was used by Alibabaei et al. [34] to schedule the irrigation of a tomato crop in Portugal using climatic Big Data. Soil moisture estimation was used as the state of the environment of the agent. The agent decided when and how much to irrigate based on the predicted next state of the environment (estimate of soil moisture and some climate factors for the next day). Compared to fixed and threshold irrigation, the DQN model showed a reduction in irrigation water volume without affecting production.

Soil water-holding capacity is a key factor in crop growth and yield and plays a critical role in Mediterranean rain-fed crops. Plant available soil water at a given rooting depth can be estimated from the Fraction of Transpirable Soil Water (FTSW) calculated as the ratio between the Available Soil Water (ASW) and Total Transpirable Soil Water (TTSW) [35,36]. The strong and stable relationship between the FTSW and vine leaf water potential allows its use as an indicator of soil water deficit in vineyards [35]. Lebon et al. [37] used a soil–plant–atmosphere system composed of simply defined subsystems and a geometric vine canopy model for estimating the FTSW in the Alsace region, France. Simulations were performed over four seasons and compared with the FTSW data measured with a neutron probe. In general, the model provided a good estimate of the evolution of the FTSW, but it was very sensitive to most canopy parameters and depended on their accurate determination or simulation. In Alentejo, southern Portugal, Lopes et al. [38] used the FTSW to study the effects of vineyard soil management practices combined with deficit irrigation strategies on the vine water status and performance of the Tempranillo grape variety. The study showed that the FTSW is a strong indicator of plant water status, enabling the discrimination of the effects of soil management practices on water use. Phogat et al. [39] used the Pedo-Transfer Function approach (PTF) to estimate Plant Available Water Capacity (PAWC) in the Barossa region, South Australia. They obtained an accurate estimate of the PAWC (which is usually time-consuming and expensive) using the Pedo-Transfer Function (PTF) approach, suggesting that it is possible to obtain estimates of the PAWC at the regional scale from easily measured data. Romas and Martínez-Casasnovas [40] investigated the variability of soil water and its impact on the amount of soil water available to grapevines under various rainfall patterns in the Penedés wine region, Spain. The Soil Water Assessment Tool model (SWAT) was used to simulate soil water during the period of 2000–2012. The FTSW was used as an indicator of water availability. The model enabled determining the effects of soil water content in early and later maturing grape varieties, highlighting the importance of considering water availability in varietal selection. Valdés-Gómez et al. [41] assessed how well the Simulateur mulTIdisciplinaire pour les Cultures Standard (STICS) model captures grapevine phenology, biomass production, yield, and soil water content in two scenarios with different rainfall distributions and water management approaches. A good estimate of the main stages of grapevine phenology and soil water content was obtained. A comparison of the FTSW with grapevine leaf water potential measurements showed that the vine water stress period was correctly estimated. The FTSW can be calculated using soil moisture content data recorded by Internet of Things (IoT) sensors. However, IoT sensors have some drawbacks, including the need for maintenance and the possibility of data pauses. When IoT sensors fail or no sensor is available in the field, a model that can predict the FTSW with minimal inputs is a good tool for estimating the missing data. To the best of our knowledge, there is no study on the estimation of the FTSW using DL models. This literature review highlights the various machine learning and deep learning algorithms that have been applied to irrigation management, with a focus on predicting the soil moisture, irrigation amount and timing, and soil water holding capacity. These studies demonstrated the potential of these algorithms to improve irrigation management and decision-making.

The hypothesis behind this research was that LSTM can detect the relationship between historical, weather data (temperature, relative humidity, rainfall, wind speed, vapor pressure deficit, solar radiation), reference Evapotranspiration (ETo), and the dependent

variable, FTSW, in a vineyard and that these relationships can be utilized to effectively forecast the FTSW for the future or for other vineyards with comparable climate, soil type, and management. The objectives of this study were (1): to assess the ability of the LSTM model to predict the FTSW in conjunction with observed data from IoT sensors deployed in a commercial vineyard in Portugal with minimal inputs, (2) to evaluate the prediction result for another vineyard in a completely different situation with an independent dataset.

The remainder of this paper is organized as follows. An explanation of the models and methods used in this work, the dataset used, and the training configuration is provided in the Materials and Methods Section 2. The results of the proposed model when applied to the validation datasets are presented in Section 3. The summary of the work can be found in Section 4.

2. Materials and Methods

2.1. Recurrent Neural Network

The Recurrent Neural Network (RNN) is a type of neural network designed to process sequential data in a temporal dataset [42]. Long dependence relationships can be captured and stored using RNN models. Unlike other deep learning algorithms, RNN models use a look-back period, which is the number of previous time steps used to predict the next time step. The hidden RNN layers consist of RNN units, which, in turn, consist of RNN cells, with the number of cells in each unit corresponding to the look-back period. The number of hidden layers and units comprises the hyperparameters that should be set before training the model. The output of the cell C_t is determined by Equation (1):

$$h_t = \sigma\left(W_{x_t} x_t + W_{h_t} h_{t-1} + b_t\right) \tag{1}$$

where x_t is the input of the cell, h_{t-1} is the output of the cell C_{t-1}, W_{x_t} is $W_{h_{t-1}}$, and b_t are the weights and biases that should be trained and adjusted during the training of the model to minimize the loss function. The sigmoid function σ is used to add nonlinearity to the model.

The RNN model has the disadvantage that, for large inputs, the gradient of the loss function can approach zero during training, which means that the weights of the first layer cannot be updated. Long Short-Term Memory (LSTM) is an artificial RNN architecture that allows gradients to flow unchanged, avoiding the vanishing gradient problem in classical RNN cells [43].

In the LSTM layer, there are several gates such as the forget gate, the input gate, and the cell memory. These gates help to retain the useful information from the previous output layer and forget the information that is no longer useful to remember [44]. The LSTM cell must first decide which data in the cell to ignore using the forget gate. The forget gate outputs numbers between 0 and 1, where 1 denotes complete preservation of this information by parsing the output of the last cell, h_{t-1}, and the input of the current cell, x_t.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{2}$$

$$\sigma = \frac{1}{1 + \mathrm{e}^{-x}} \tag{3}$$

Equation (4) is used by the input gate i_t to determine which values to update while shielding the unit from unrelated inputs. Then, using Equation (5), the activation function tanh creates a vector of new candidates $\tilde{C}_t$ to add to the state.

$$i_t = \sigma(W_{ix} x_t + W_{ih} h_{t-1} + b_i) \tag{4}$$

$$\tilde{C}_t = \tanh(W_{zx} x_t + W_{zh} h_{t-1} + b_z) \tag{5}$$

$$\tanh = \frac{(\mathrm{e}^{x} - \mathrm{e}^{-x})}{(e^{x} + e^{-x})} \tag{6}$$

The unneeded information of the previous unit is then forgotten by multiplying the forget gate f_t by C_{t-1}, while i_t is multiplied by $\tilde{C}_t$ to preserve the new valuable information (Equation (7)).

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{7}$$

What information is used in the cell to produce the output of the LSTM unit is determined by the output gate o_t through Equation (8).

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{8}$$

Finally, the output of the unit is determined by Equation (9).

$$h_t = o_t \odot \tanh(C_t) \tag{9}$$

The weights (W) and biases (b) of the gates are changed during the LSTM training process to ensure that the loss function reaches the minimum value.

A Bidirectional LSTM (BiLSTM) is a sequence-processing model consisting of two LSTM layers, one of which processes inputs forward and the other backward. Figure 1 shows a bidirectional LSTM layer.

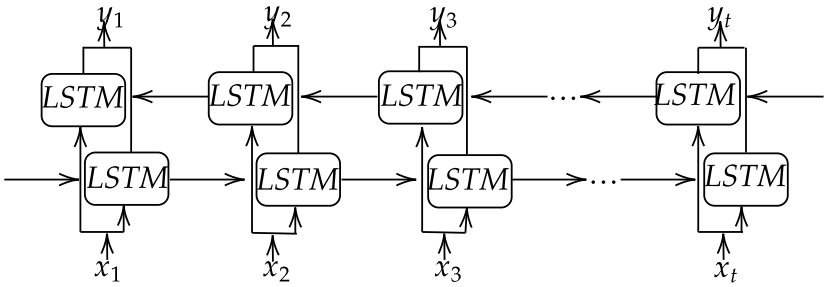


Figure 1. Bidirectional LSTM layer. Reprinted from ref [44].

2.2. Random Forest

A supervised learning technique called Random Forest (RF) uses ensemble learning with decision trees as the framework [45]. The ensemble learning method combines predictions from different machine learning algorithms to provide predictions that are more accurate than those from a single model. Ensemble approaches in statistics and machine learning use numerous learning algorithms to achieve better predictive performance. In the RF algorithm, several decision trees are built throughout the training time in a Random Forest, and the prediction of each tree is the mean of the results. In the case of regression, decision trees start at the root of the tree and follow splits based on the outcomes of the variables until they reach a leaf node. Figure 2 shows the Random Forest algorithm.

There are several hyperparameters that should be established before training the RF model. The hyperparameters include the bootstrap method for sampling data points (with or without replacement); the max depth, i.e., the maximum number of levels in each decision tree; the max features, i.e., the number of features to be considered in the search for the best split; n_estimators, i.e., the number of trees in the forest; the min samples leaf, i.e., the minimum number of data points allowed in a leaf node; and the min samples spilled, i.e., the minimum number of data points placed in a node before the node is split.

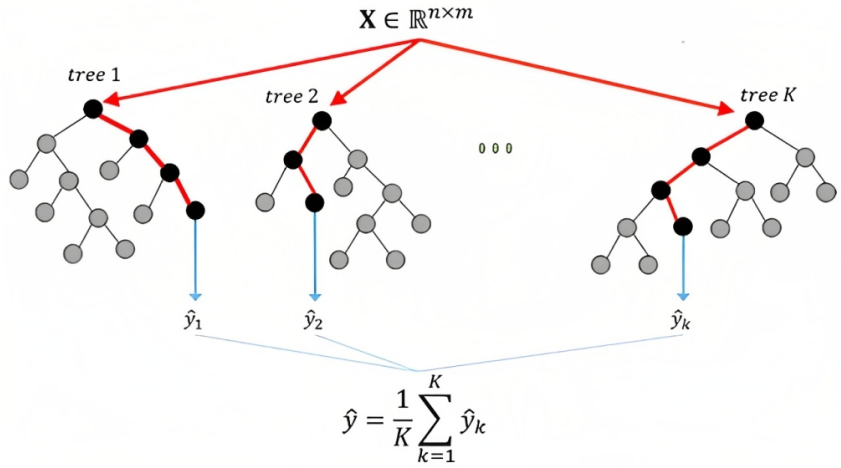


Figure 2. Random Forest's structure. Adapted from ref [46].

2.3. Support Vector Regression

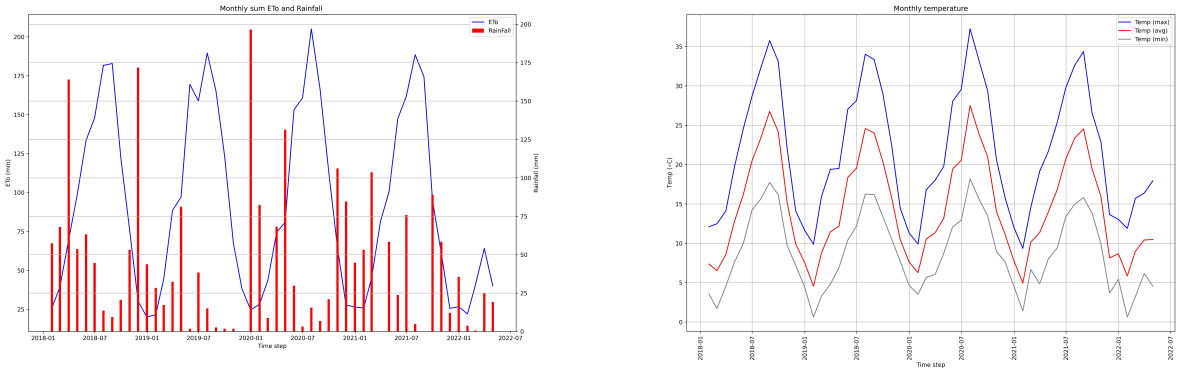
A supervised learning approach called Support Vector Regression (SVR) is used to predict discrete values. Finding the best-fitting line is the basic idea of SVR. In SVR, the hyperplane with the largest number of input points is the line that best fits the data. A set of mathematical operations that accept data as the input and shape them into the desired form is called kernels in the SVR model. They are typically used to locate a hyperplane in a higher-dimensional space. The most-commonly used kernels are the, Radial Basis Function (RBF), nonlinear, polynomial, and linear. Depending on the dataset, any of these kernels can be used.

Like RF, there are several hyperparameters involved in training an SVR model. The width of the tube around the predicted function is determined by the value of Epsilon (hyperplane). Gamma determines the degree of the desired curvature of a decision boundary, and *C* is a hyperparameter for error control in SVM.

2.4. Datasets Used

The trained dataset was collected from 2018 to 2021 at a commercial vineyard (Colinas do Douro) located in the Douro Superior wine-growing region, Portugal (40.49 N; 6.59 W), at an elevation of 463 to 477 m and a slope of 9%. The vines were spaced 1.0 m within and 2.4 m between east–west-oriented rows and trained on a vertical shoot positioning with a pair of movable wires. Shoots were trimmed at a height of 1.0 m. The climate is of the Mediterranean type with continental influence, presenting very hot and dry summers. The soil is of schist origin with a sandy loam texture, a pH (KCl) of 5.6, and 0.39% of organic matter.

An in situ (Enviroscan) probe composed of four capacitance sensors located at 20, 40, 60, and 80 cm depths and a GPRS communication unit were used to assess soil moisture. An automatic weather station was established near the vineyard to measure the minimum, maximum, and average temperature, as well as the minimum, maximum, and average relative humidity, wind speed, rainfall, and solar radiation. The weather station platform also delivers the ETo (estimated by the Penman–Monteith equation [47]). During the experimental period (2018–2021), the mean air temperature ranged between −2.12 °C and 33.92 °C and the mean total annual rainfall was 763.2, 417.2, 602.8, and 514.8 mm, in 2018, 2019, 2020, and 2021, respectively. The vineyards of Colinas do Douro are drip-irrigated. The irrigation strategy used is sustainable deficit irrigation consisting of a periodical application of 5.5 mm per irrigation event. Figure 3 shows the monthly rainfall, ETo, and temperature (average minimum, maximum, and average) at Colinas do Douro from 2018 to 2021. Further details of the climate data can be found in Table 1.



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

Figure 3. Monthly rainfall and ETo (**a**) and monthly average temperatures (**b**) recorded at the weather station of Colinas do Douro from 2018 to 2021.

Table 1. Climate data details of the weather station of Colinas do Douro from 2018 to 2021. The temporal resolutions of all the variables is daily.

Variables		Year							
		2018		2019		2020		2021	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std
Temp (°C)	Avg	14.97	7.59	14.86	7.05	15.58	7.16	14.87	6.79
	Max	21.79	9.38	22.09	8.84	22.56	9.24	21.96	8.65
	Min	9.43	5.9	8.95	5.64	9.8	5.35	8.96	5.54
RH(%)	Avg	64.72	17.88	63.14	17.43	65.08	19.03	63.97	16.15
	Max	85.05	12.06	84.6	11.86	85.83	13.75	85.57	10.18
	Min	41.2	20.34	38.73	20.77	41.32	21.59	39.82	19.32
WS (m/s)	Avg	1.63	1.16	1.64	1.23	1.52	1.1	1.43	1.06
	Max	4.3	2.21	4.14	2.17	3.98	2.11	3.8	2.04
VPD (kPa)	Avg	0.9	0.85	0.89	0.73	0.93	0.9	0.86	0.72
	Min	0.23	0.27	0.21	0.22	0.23	0.29	0.2	0.2
Rainfall (mm)	Sum	2.1	5.2	1.14	5.22	1.65	5.11	1.41	4.23
ETo (mm)	Daily	3.1	2.03	3.26	2.09	3.13	2.09	3.11	2.03

The abbreviations stand for the following: Max: Maximum, Min: Minimum, Avg: Average, Std: Standard deviation, Temp: Temperature, RH: Relative Humidity, VPD: Vapor Pressure Deficit, WS: Wind Speed, ETo: reference Evaporation.

The ratio between the Available Soil Water at a given time (ASW) and the Total Transpirable Soil Water (TTSW) of a certain crop in a given soil is called the Fraction of Transpirable Soil Water (FTSW) [48]. The ability of the crop to absorb moisture from the soil is usually less than the theoretical volume of soil water calculated by the field capacity minus the permanent wilting point. Therefore, at a given date, the ASW was calculated as the difference between the soil water content on the day of the measurements and the lower limit of water uptake, reached at the end of the season of the driest year (achieved in 2020; Table 1), before the first rainfall. The FTSW was calculated as the ratio of the ASW to the TTSW [35]. The TTSW up to an 80 cm soil depth was estimated as the soil water reserve held between the field capacity and the lower limit mentioned above. Table 2 and Figure 4 show the details of the calculated FTSW from the soil moisture recorded by the sensor and the calculated FTSW at Colinas Do Douro, respectively.

Table 2. Soil moisture data details for each season (from April to October) from Colinas do Douro. VSM indicates Volumetric Soil Moisture content.

Variables		Temporal Resolution	Year							
			2018		2019		2020		2021	
			Max	Min	Max	Min	Max	Min	Max	Min
VSM (%)	20 cm	Daily	31.56	10.85	31.82	11.16	32.52	10.86	32.40	11.18
	40 cm	Daily	28.96	10.89	28.06	10.64	29.75	10.62	29.83	12.01
	60 cm	Daily	31.68	11.05	27.39	10.96	31.93	11.24	29.99	12.27
	80 cm	Daily	33.43	12.56	27.53	12.56	33.29	12.79	28.29	13.43
VSM (mm)	summed	Daily	209.11	81.15	190.04	84.52	217	79.33	202.19	85.66

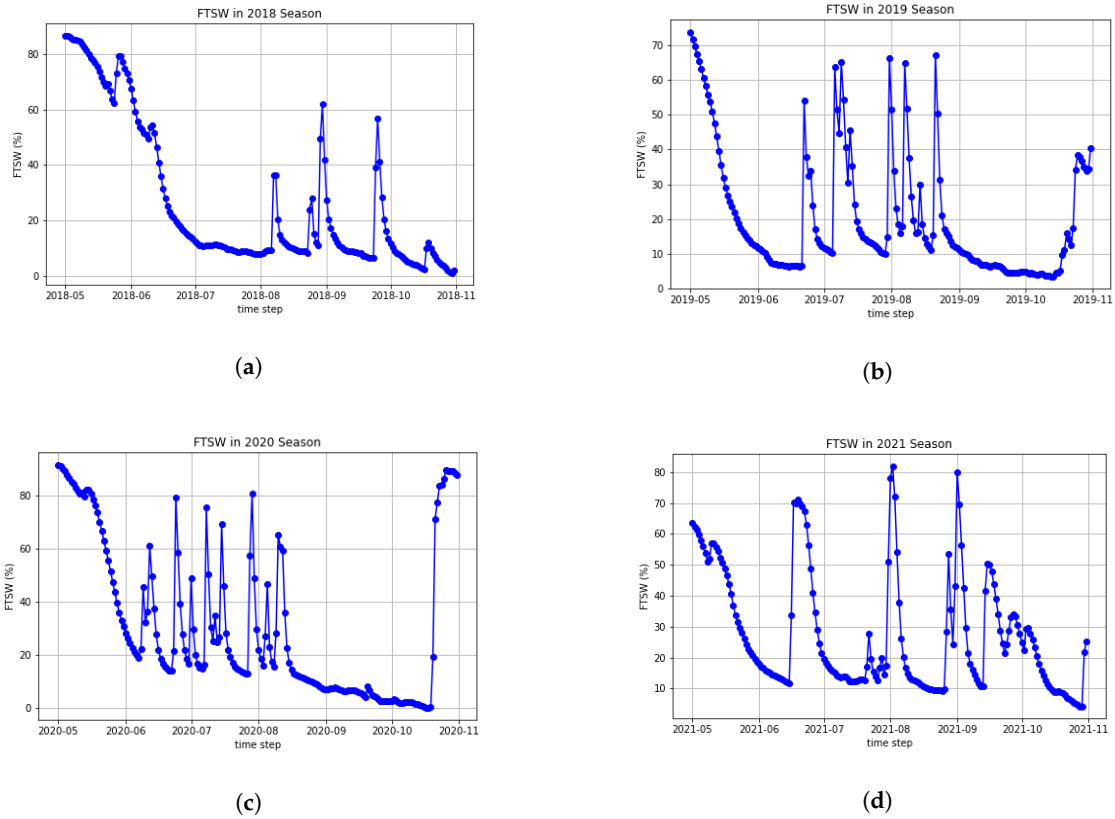


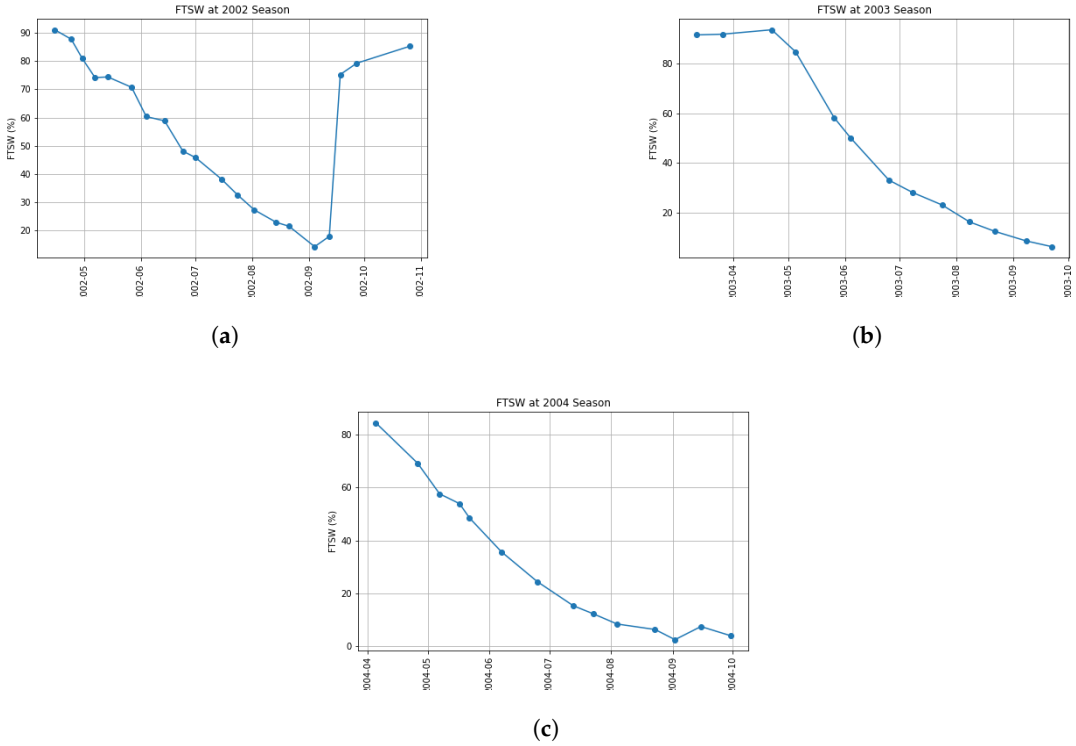
Figure 4. Calculated FTSW of the four seasons for Colinas do Douro from 2018 to 2021. (**a**) FTSW in 2018 season, (**b**) FTSW in 2019 season, (**c**) FTSW in 2020 season, and (**d**) FTSW in 2021 season.

Test Dataset

Model validation was performed using an independent dataset collected during three growing seasons (2002 to 2004) in a 15-year-old 'Cabernet Sauvignon' (*Vitis vinifera* L.) rain-fed vineyard, grafted on 110 R rootstock, located at Quinta de Pancas, Alenquer, western Portugal ((39.06 N; 9.03 W). The vines were spaced 1.0 m within and 2.5 m between east–west-oriented rows and trained on a vertical shoot positioning with a pair of movable wires. Shoots were trimmed twice, between bloom and veraison, at a height of 1.0 m. The soil presented 24.3% of clay, 20.2% of silt, 55.5% of sand (sandy clay–loam soil), and 0.7% of organic matter. The climate is of the Mediterranean type with Atlantic influence. The weather data were recorded by an automatic weather station located within the experimental vineyard. During the experimental period (2002–2004), the mean air temperature ranged between 10 °C and 23.6 °C and the mean total annual rainfall was 885.3, 941.8, and 564.4 mm, in 2002, 2003, and 2004, respectively. The experimental design was a randomized complete block with three soil management treatments. However, for

validation purposes, only data from one treatment—soil tillage in the inter-row combined with herbicide under the vines—was used, which is comparable to the tillage method used at Colinas do Douro. Soil moisture was assessed with a portable capacitance probe (Diviner 20001: Sentek Environmental Technologies, King Town, Australia). Readings were made periodically between bud break and harvest in 12 access tubes (3 per replication), placed along the row between two contiguous vines, at increments of 0.1 m from the soil surface to a depth of 1.0 m. For each access tube, soil moisture data were computed for an 80 cm depth, and the average was used as the input for the test dataset. The total average soil available water, up to an 80 cm depth, was 124 mm, calculated as the difference between the field capacity and the lowest soil moisture value reached at the end of the season of the driest year (2003), before the first rainfall. For more details, see [49].

Similar to Colinas do Douro, the difference between the soil water content at field capacity and the minimum soil water content was used to determine the TTSW for each access tube, and the amount of ASW was calculated as the difference between the soil water content on the day of measurements and the minimum soil water content. The ratio of the ASW to the TTSW was used to calculate the FTSW [48]. Figure 5 show the FTSW calculated from the soil water content data for Quinta de Pancas.



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

Figure 5. Calculated FTSW for Quinta de Pancas from 2002 to 2004. (**a**) FTSW in 2002 season, (**b**) FTSW in the 2003 season, and (**c**) FTSW in 2004 season.

As the daily ETo value was missing from the Quinta de Pancas dataset, the data were estimated by the Hargreaves–Samani method [50,51]. This method estimates the ETo using only the observed maximum and minimum temperatures and the estimation of the extraterrestrial radiation, expressed by

$$\mathrm{ETo} = 0.0135 \times k_{Rs} \times 0.408 R_a \times (\mathrm{Temp_{avg}} + 17.8) \times (\mathrm{Temp_{max}} - \mathrm{Temp_{min}})^{1/2} \quad (10)$$

where ETo is the reference evapotranspiration (mm day^{-1}); R_a is the extraterrestrial radiation (MJ m^{-2} day^{-1}); 0.0135 is a factor for conversion from the U.S. to the international

system of units; k_{Rs} is the radiation adjustment coefficient ($^{\circ}C^{-0.5}$). The empirical coefficient k_{Rs} was originally taken to be 0.17 $^{\circ}C^{-0.5}$ [50].

2.5. Data Preprocessing

Data normalization aims to convert the values of the dataset into a normal form without distorting the variations in the ranges of values. For many machine learning estimators, this is a typical requirement. In this paper, the missing data were removed from the dataset, and normalization was applied. This involved transferring the data to a new scale using Equation (11).

$$x_{new} = \frac{x_{old} - x_{mean}}{\sigma} \tag{11}$$

where x_{mean} and σ are the mean and standard deviation of the input data.

2.6. Training Configuration

Several BiLSTM models with two BiLSTM layers were trained to estimate the FTSW. These models differed from each other in the number of input variables, the model look-backs (input size), and the length of the output.The length of the output was set to 1, 3, 5, and 7 days for the FTSW forecast, and the number of look-backs was set to 3, 5, and 7 days of the historical data.

Hyperparameters' Selection

The hyperparameters are the variables that determine how the network is trained (e.g., learning rate) or the parameters that determine the structure of the model (e.g., the number of hidden layers). The process of determining the best combination of hyperparameters to maximize model performance is known as hyperparameter tuning (or hyperparameter optimization). It has been demonstrated that changing the number of nodes per layer in LSTM models has a more significant effect on the results than altering the number of layers [52]. To predict the simulated soil moisture, the network designs of LSTM and BiLSTM with two, three, and four LSTM layers and one and two BiLSTM layers were investigated in [44]. The BiLSTM model with two layers was utilized for prediction because it achieved the best accuracy in the above-mentioned study.

The search for the hyperparameters is an optimization problem. The goal is to find the hyperparameters such that they maximize the loss function. There are several methods for solving this problem, such as greedy search [53], random search [54], and Bayesian optimization [55]. In greedy search, a search space must be selected, i.e., a region in which each point represents a different model configuration and each dimension represents a hyperparameter. The algorithm evaluates each position in the grid to find the best hyperparameters. The disadvantage of this method is that the space is discrete and the algorithm's search in each space is time-consuming. To address the time-consuming drawback of greedy search, the random greedy search algorithm randomly selects a value from the search space, but the values are still randomly selected and may result in a hyperparameter that is not the best. Bayesian optimization is used to solve the optimization problem when the objective function is a black box, such as the lossfunction in DL models. Tuning the hyperparameters using Bayesian optimization can reduce the time required to find the best parameter set and improve the generalization performance in the test set. Moreover, the space can be chosen as a continuous space. The other hyperparameters in this study were the number of units in each BiLSTM layer, the learning rate, the learning rate decay, the batch size, and the dropout size. The Bayesian optimization algorithm [55] was used to select these hyperparameters. The set of hyperparameters is shown in Table 3.

Table 3. Hyperparameters selected by the Bayesian optimization algorithm.

Model	Batch Size	Dropout Size	Learning Rate	Decay	No. of Units
FTSW prediction model	40	0.09313	0.003968	0.003456	158

3. Results and Discussion

The permutation feature importance technique [56,57] was used to calculate the importance of the features. The training set for this method is used to train a model, and the validation set is used to measure the increase in the prediction error after a feature is permuted in the validation set, destroying the link between the feature and the true outcome [45]. Since the model depends on the feature for prediction in this scenario, a feature is considered "important" if permuting its values results in an increase in the model error. Figure 6 shows the importance of each feature using the BiLSTM model and permutation feature importance technique. The daily ETo was the most-important feature because the permutation of this feature caused the largest increase in the MSE from 0.3 to 1.75.

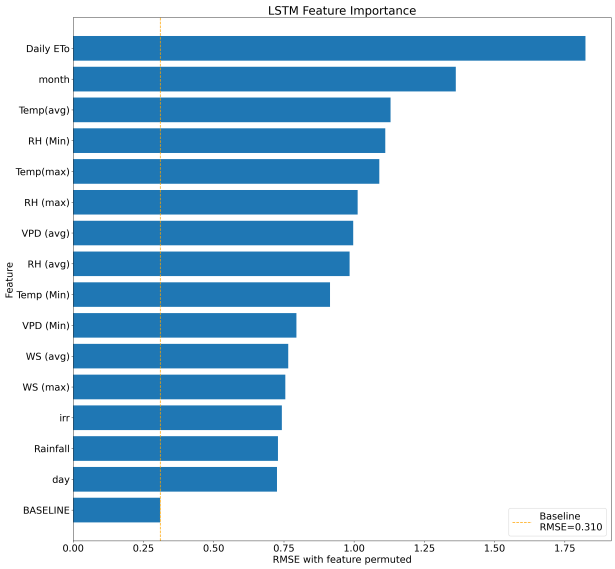


Figure 6. Feature importance using the BiLSTM model and the permutation feature importance technique.

The feature with the highest importance (ETo) was selected as the main independent variable. Then, the other features were added to the group of independent variables one by one in order of importance, and the Variance Inflation Factor (VIF) value for this group of variables was calculated. The VIF [58] with a threshold of five was used to remove the multicollinear variables. The variable is retained as a predictor if the VIF value is below the threshold; otherwise, it is removed from the group. Following these processes, the historical data selected as the input to the FTSW forecast model were the month, average RH, minimum VPD, rainfall, and daily ETo.

Since the VPD (min) was not present in the Quinta de Pancas dataset, a model without this variable was also trained to evaluate the performance of the model on the Quinta de Pancas dataset. In general, with various input variables and various input and output lengths, fifteen BiLSTM models were trained (see Table 4).

Table 4. Different inputs of the model and different input and output sizes used to train the BiLSTM model. Look-back indicates the number of days used to predict the FTSW, while irr indicates irrigation.

Input Variables	Input Length (Look-back)	Output Length
Month, ETo, RH (avg), rainfall	3, 5,7	1
Month, ETo, rainfall RH (avg), VPD (min)	3, 5, 7	1
Month, ETo, rainfall, RH (avg), VPD (min), irr	3, 5, 7	3, 5, 7

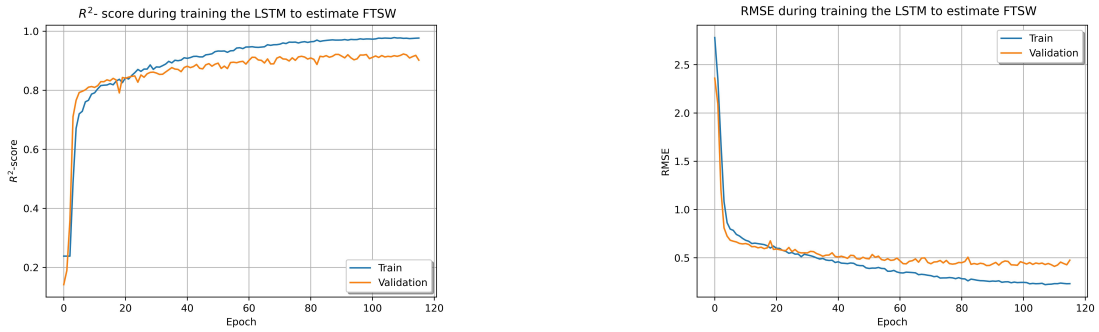
3.1. Training of the FTSW Prediction Model

Figure 7 shows the R^2-score and RMSE during training of the BiLSTM model with input including the VPD (min) and irrigation amount and seven days of look-back.

Early stopping [59] (an optimization strategy intended to lessen overfitting without sacrificing model accuracy) was used. The validation loss of the model was tracked during training, and if it did not improve after a certain number of iterations, the training was stopped. Figure 7 illustrates that the validation loss stopped improving at around Step 110, but the training loss continued to improve.

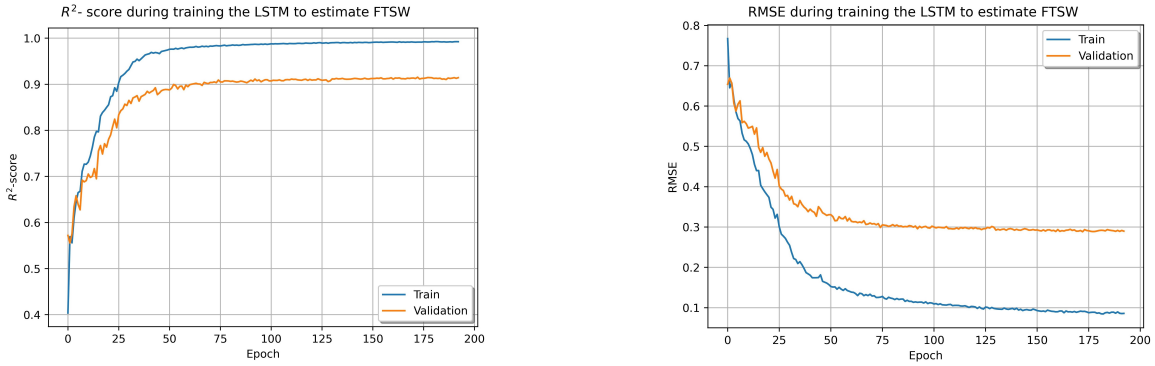
The model with an output length of seven days required more time to complete the training compared to the model with an output length of one day.

Figure 8 shows the R^2-score and RMSE during the training of these BiLSTM models with seven days of look-back and the output size. Figures 7 and 8 show that overfitting during training was more likely when the output size was greater than one and that the model overfit the training set without the early stopping optimization.



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

Figure 7. (**a**) R^2-score and (**b**) RMSE during training of the BiLSTM model with the output of one day, seven days of look-back, and the VPD as the input to the model.



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

Figure 8. (**a**) R^2-score and (**b**) RMSE during training of the BiLSTM model with an output length of seven days, seven days of look-back, and the VPD as the input to the model.

3.2. Performance of the FTSW Prediction Model on Colinas do Douro Dataset

The performance of the models was evaluated using the Root-Mean-Squared Error (RMSE) and the R^2-score. Table 5 shows the performance of the model with different look-backs and a forecast of the FTSW for the last day. The model performance improved when look-backs were increased from one to seven days, with a 14–15% increase in the

R^2-score and a 5.85–7.31% decrease in the RMSE on the Colinas do Douro test set. Therefore, one of the important parameters affecting the BiLSTM model is the look-back of the model.

Table 5. The performance of the BiLSTM model to predict the FTSW one day ahead with different input variables.

Input Variables	Look-Back (Days)	Colinas do Douro			
		Validation Set		Test Set	
		R^2-Score	RMSE (Normalized)	R^2-Score	RMSE
Month, ETo, RH (avg), rainfall,	3	0.75	0.6	0.75	16.6
	5	0.81	0.56	0.81	14.4
	7	0.87	0.54	0.90	10.75
Month, ETo, rainfall RH (avg), VPD (min)	3	0.77	0.50	0.79	15.39
	5	0.83	0.53	0.85	12.78
	7	0.92	0.41	0.93	8.98
Month, ETo, rainfall, RH (avg), VPD (min), irr	3	0.75	0.51	0.79	15.34
	5	0.87	0.46	0.86	12.53
	7	0.96	0.31	0.94	8.03

Moreover, the performance of the BiLSTM model with the VPD (min) in its input features increased by 3–4% in the R^2-score compared to the model trained without these two variables as the input.

To examine the effect of irrigation amount on the FTSW estimation, a model was trained to include irrigation as an input to the model. As shown in Table 5, the R^2-score increased in the range of 1%. Since the number of irrigations in the Colinas do Douro was low (between four and eight times in the period from 2018 to 2021), the influence of this variable on the performance of the model was also low.

Figure 9 compares the observed value of the FTSW with the predicted values for the models with a look-back (input size) of seven days and an output length of one and seven days. The yellow line shows the best fit, and when the points in the scatter plot are closer to the line, better performance was obtained. As can be seen in Figure 4a,b, the points on the scatter plot for the model with the VPD as input were closer to the best-fit line than for the model with the VPD as the input, so better performance was obtained when the VPD was used as the input to the model.

Table 6 shows the performance of the model trained with different look-backs and lengths of output. The model performed better when the look-back was extended, similar to the case where the output length was one, and the performance of the BiLSTM model with the VPD (min) in its input features was better than that of the BiLSTM without this feature. As shown in Tables 5 and 6, the model with an output length of one had a better performance compared to the models with an output size of three, five, and seven days.

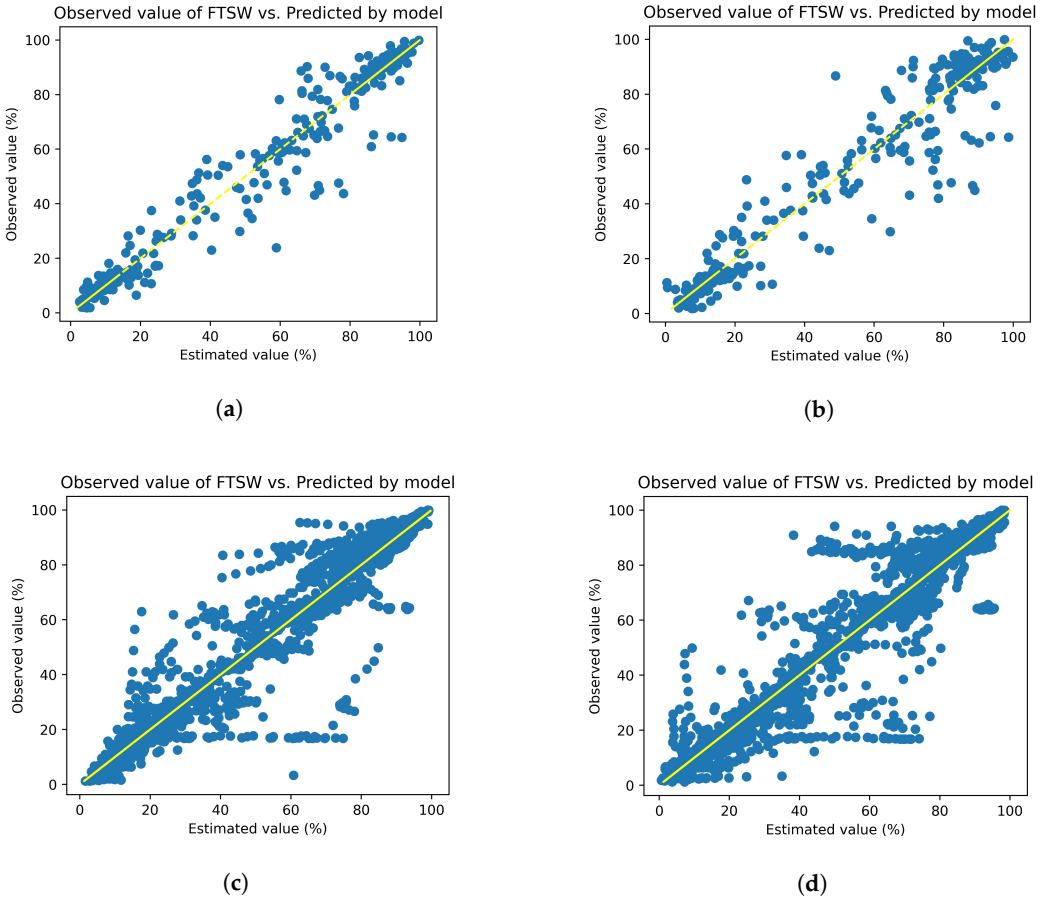


Figure 9. Comparison between the observed and predicted values of the FTSW. The results of the model with the Vapor Pressure Deficit (VPD) as the input are shown in (**a**,**c**), while the results of the model without this variable as the input are shown in (**b**,**d**). A perfect fit is represented as a yellow line.

Table 6. The performance of the BiLSTM-trained model to predict the FTSW with different look-backs and lengths of the output.

Input Variables	Look-Back	Output	Colinas do Douro			
			Validation Set		Test Set	
			R^2-Score	RMSE	R^2-Score	RMSE
Month, ETo, RH (avg), rainfall, irr	3	3	0.61	0.60	0.58	21.37
	5	5	0.80	0.46	0.74	16.77
	7	7	0.32	0.89	0.90	10.35
Month, ETo, RH (avg), rainfall, irr VPD (min)	3	3	0.69	0.57	0.64	19.68
	5	5	0.44	0.80	0.81	14.54
	7	7	0.31	0.88	0.92	9.19

3.3. Performance of the FTSW Model on Quinta de Pancas Dataset

As was mentioned, the ETo for Quinta de Pancas was calculated using Equation (10). First, Equation (10) was evaluated using the Colinas do Douro dataset, and the validation yielded an RMSE of 1.10 mm and an R^2-score of 0.92%. The evaluation was performed to ensure that the formula can estimate the ETo with reasonable error.

Figure 10 shows the monthly accumulated rainfall and estimated ETo using Equation (10) for Quinta de Pancas from 2002 to 2004. Note that the recording of the climate data for this dataset was not continuous. The value of zero for the ETo in the plot shows these missing data. A comparison of Figures 3 and 10 shows that the maximum monthly ETo totals at Colinas do Douro were in the range of 175–200 for each season and between 140 and 155 for Quinta de Pancas.

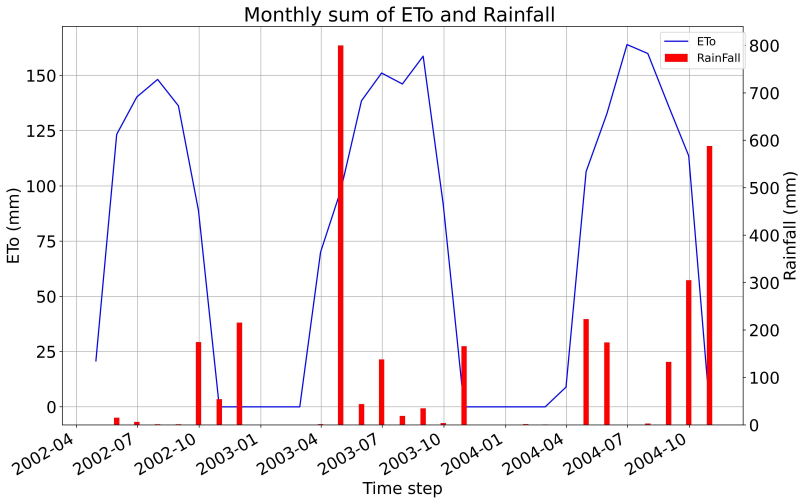


Figure 10. Monthly sum of rainfall and estimated ETo for Quinta de Pancas from 2002 to 2004.

After ETo estimation, the BiLSTM model was used to predict the FTSW for Quinta de Pancas. Table 7 shows the performance of the FTSW prediction model for Quinta de Pancas. Compared to Colinas do Douro, the R^2-score decreased by 6% and the RMSE increased by 1.39%. Part of this decrease may be due to the error in the estimation of the ETo (see Table 7), but in general, the model was able to predict the FTSW at Quinta de Pancas with a reasonable prediction performance.

Table 7. Performance of the trained BiLSTM model to predict the FTSW one day ahead on the Quinta de Pancas test dataset.

Location	RMSE	R^2-Score
Quinta de pancas	10.22	0.87
Colinas do Douro	8.43	0.93

Figure 11 compares the observed values of the FTSW with the value predicted by the BiLSTM model. The points in the scatter plot must move to the yellow line to obtain a good approximation. The fact that the points lie along the line shows that the model represented the relationships very well, as can be observed.

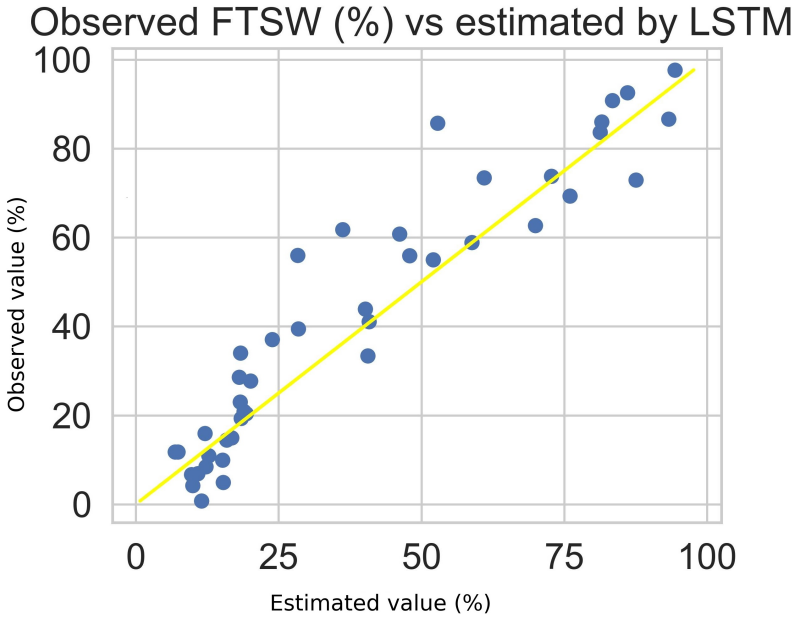


Figure 11. Comparison between observed and predicted values of the FTSW at Quinta de Pancas.

3.4. Comparison of FTSW Prediction Using the LSTM Model Performance with Other Models

The LSTM model with a seven-day look-back and a one-day output length was compared with the RF, SVR, and Linear Regression (LR) models. To achieve the optimal performance of the RF and SVR models for the dataset, the hyperparameters of the RF model were selected using a random grid search. The hyperparameters selected for RF and SVR are listed in Table 8. Note that Sqrt indicates that the maximum number of features is equal to the Square root of the number of features.

Table 8. Hyperparameters selected for RF and SVR using random grid search.

SVR		RF	
Hparam	**Value**	**Hparam**	**Value**
C	1.83	n_estimators	118
kernel	RBF	bootstrap	False
bootstrap	False		
gamma	scale	max depth	110
epsilon	0.26	max_features	Sqrt
-	-	Min samples leaf	2
-	-	Min samples split	2

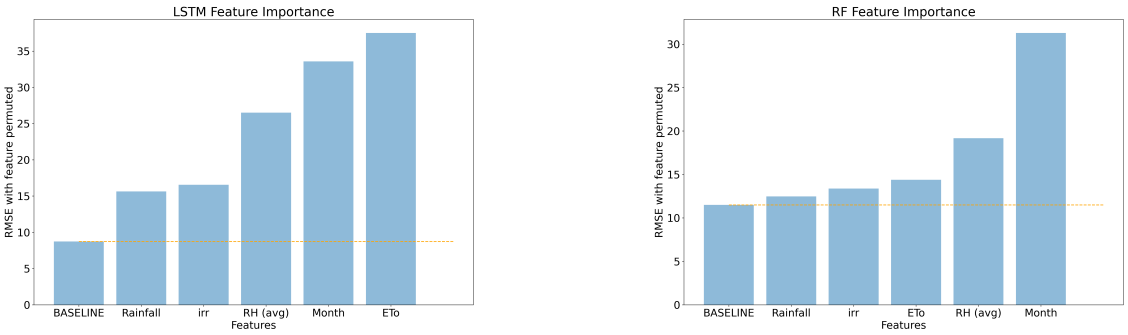
Table 9 shows the performance of BiLSTM, RF, SVM, and LR on the Colinas do Douro and Quinta de Pancas test datasets. In general, the BiLSTM model performed better on both sets of test data. This was expected given that evapotranspiration directly affects soil moisture [60,61] and that the BiLSTM model prioritizes this variable. RF achieved the second-best performance, and the most-valuable performance was achieved with LR. Since the relationship between the FTSW and the other independent variables is not linear, LR was expected to achieve the best performance.

Moreover, the R^2 score of the BiLSTM model on Quinta de Pancas decreased by 8% compared to the R^2 score of the BiLSTM model on Colinas do Douro, while this deterioration for RF and SVR was more than 16%. Therefore, the BiLSTM model performed better on the unobserved data and generalized better to the unseen data.

Table 9. Comparison of the performance of the trained BiLSTM model with RF in predicting the FTSW one day ahead on the Quinta de Pancas and Colinas do Douro validation sets.

Model	Quinta de Pancas		Colinas do Douro	
	RMSE	**R^2-Score**	**RMSE**	**R^2-Score**
BiLSTM	10.36	0.87	8.3	0.93
RF	16.8	0.67	12.88	0.85
SVR	19.43	0.58	14.15	0.82
LR	22.5	0.43	18.68	0.69

To better assess the results, the two best models, i.e., BiLSTM and RF, were compared in terms of feature selection and model variance. First, these two models were compared in terms of feature importance. The permutation feature importance technique was used for both models to investigate which feature was more important in predicting the FTSW by each model. Figure 12a,b show the importance of each feature with the LSTM model and RF, respectively. As the figure shows, in the BiLSTM model, the most-important feature was the ETo and then the month, but in RF, the month was selected as the most-important feature and the ETo as the third-most-important feature for predicting the FTSW. The ETo had the strongest correlation (−0.56) with the FTSW, and LSTM correctly captured this correlation, noting that it was the most-important feature between the independent variables. This fact showed that LSTM was more reliable at predicting the FTSW than RF.



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

Figure 12. (**a**) shows the importance of the input features for predicting the FTSW with the BiLSTM model, and (**b**) shows the importance of the input features for predicting the FTSW with the RF model. The shift of a feature causing the most changes in the RMSE shows the importance of the feature.

Moreover, in both models, the irrigation amount was more important than rainfall. The reason was that, during the irrigation period (summer period), the variations in the FTSW were mainly caused by the water added by irrigation, as rainfall was very low or even absent. Figure 13 shows an example (2019) of the irrigation events (black triangles), rainfall, and the FTSW (0–80 cm). As can be seen in the figure, the irrigation events had a greater impact on the soil moisture than rainfall.

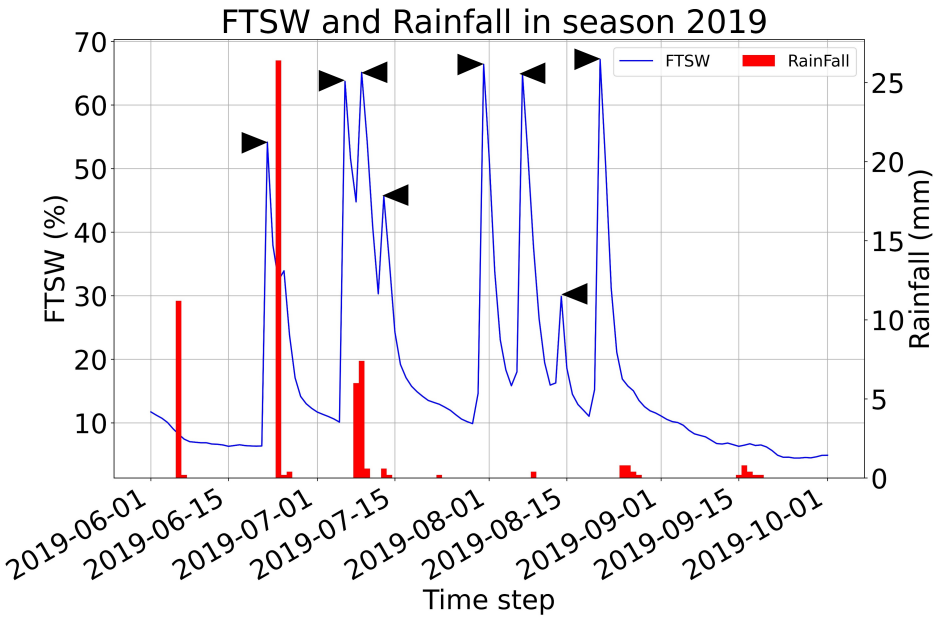


Figure 13. Irrigation event (black triangles), rainfall, and the FTSW (0–80 cm) in 2019 at Colinas do Douro.

Figure 14 shows the boxplots for the fifth, twenty-fifth, seventy-fifth, and ninety-fifth percentiles for the observed and predicted FTSW with BiLSTM and RF using the Colinas do Douro dataset. While the whiskers cover the values from the 5th to the 95th percentile, the boxes indicate the 25th and 75th percentiles. The mean of the measured data in the test phase is represented by the solid line within each box. As you can see from the figure, the FTSW predicted by the RF model was within the box between 20 and 80%, but the boundaries of the box for the observed FTSW and those predicted by the BiLSTM are almost identical (between 15 to 85). Moreover, in both models, the mean of the predicted FTSW was close to the mean of the observed values.

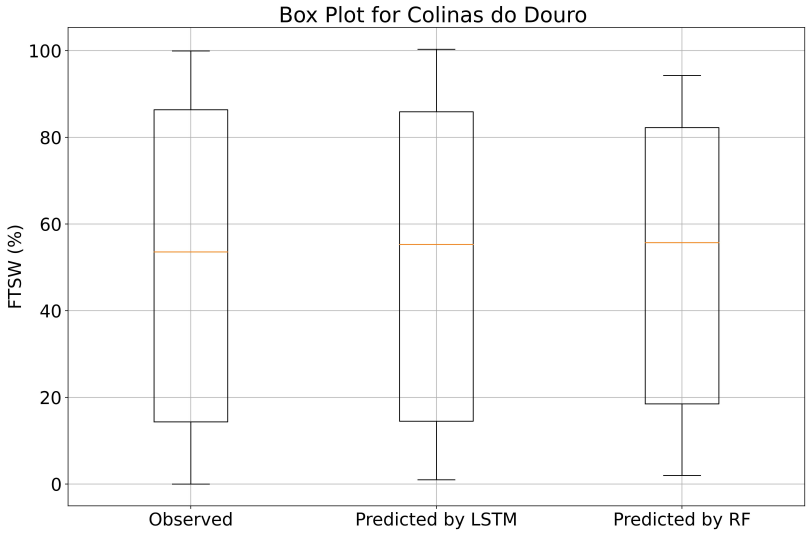


Figure 14. Boxplots showing the dispersion of the observed and predicted FTSW during the testing phase for the RF and BiLSTM models. The box shows the interquartile range (25–75th from the 5th to 95th percentile). The solid line within the box indicates the mean.

4. Conclusions

The FTSW is a variable that can help farmers schedule irrigation and reduce water waste. In this study, the propensity of the BiLSTM model to forecast the FTSW for vineyards

was investigated. The data from two different vineyards, Colinas do Douro and Quinta de Pancas, were used as the training and test datasets, respectively. Using the feature importance method and eliminating the collinearity between variables, the ETo, the month of the year, the VPD (min), the RH (avg), rainfall, and irrigation were selected as the inputs to the model. The model achieved an R^2-score ranging from 87% to 94% on the test set.

The performance of the BiLSTM model for FTSW prediction was compared with the RF, SVR, and LR models. The degradation of the R^2-score of the Colinas do Douro and Quinta de Pancas test series in RF and SVR was twice that of the BiLSTM model, showing that the BiLSTM could better generalize on the unseen data.

The results showed that BiLSTM was able to find nonlinear relationships between historical climate and soil moisture data and was able to generalize the results for grape-growing terroirs. To use the model more broadly for other environments with different conditions, it should be re-trained with the new data.

Author Contributions: Data curation: K.A., P.D.G., and R.M.C.; Formal analysis: K.A., P.D.G., R.M.C., G.C.R., and C.M.L.; Funding acquisition: P.D.G.; Investigation: K.A.; Methodology: K.A., P.D.G., R.M.C., G.C.R., and C.M.L.; Project administration: P.D.G.; Resources: K.A., P.D.G., and C.M.L.; Software: K.A.; Supervision: P.D.G.; Validation: K.A., P.D.G., R.M.C., G.C.R., and C.M.L.; Visualization: K.A.; Writing—original draft: K.A.; Writing—review and editing: K.A., P.D.G., R.M.C., G.C.R., and C.M.L. All authors have read and agreed to the published version of the manuscript.

Funding: The work is supported by by the R&D Project BioDAgro – Sistema operacional inteligente de informação e suporte á decisão em AgroBiodiversidade, project PD20-00011, promoted by Fundação La Caixa and Fundação para a Ciência e a Tecnologia, taking place at the C-MAST - Centre for Mechanical and Aerospace Sciences and Technology, Department of Electromechanical Engineering of the University of Beira Interior, Covilhã, Portugal.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: P.D.G. acknowledges Fundação para a Ciência e a Tecnologia (FCT—MCTES) for its financial support via the project UIDB/00151/2020 (C-MAST).

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

References

1. Clark, M.; Tilman, D. Comparative analysis of environmental impacts of agricultural production systems, agricultural input efficiency, and food choice. *Environ. Res. Lett.* **2017**, *12*, 064016. [CrossRef]
2. Baiano, A. An Overview on Sustainability in the Wine Production Chain. *Beverages* **2021**, *7*, 15. [CrossRef]
3. Rodrigo-Comino, J. Five decades of soil erosion research in "terroir". The State-of-the-Art. *Earth-Sci. Rev.* **2018**, *179*, 436–447. [CrossRef]
4. Sundmaeker, H.; Verdouw, C.N.; Wolfert, J.; Freire, L.P. Internet of Food and Farm 2020. In *Digitising the Industry*; Vermesan, O., Friess, P., Eds.; River Publishers: Gistrup, Denmark, 2016; pp. 129–150.
5. Lohchab, V.; Kumar, M.; Suryan, G.; Gautam, V.; Das, R.K. A Review of IoT based Smart Farm Monitoring. In Proceedings of the 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 20–21 April 2018; pp. 1620–1625. [CrossRef]
6. Nayyar, A.; Puri, V. Smart Farming: IoT Based Smart Sensors Agriculture Stick for Live Temperature and Moisture Monitoring Using Arduino, Cloud Computing & Solar Technology. In Proceedings of the International Conference on Communication and Computing Systems (ICCCS-2016), Gurgaon, India, 9–11 September 2016; pp. 673–680. [CrossRef]
7. De-Pablos-Heredero, C.; Montes-Botella, J.L.; García-Martínez, A. Sustainability in Smart Farms: Its Impact on Performance. *Sustainability* **2018**, *10*, 1713. [CrossRef]
8. Said Mohamed, E.; Belal, A.; Kotb Abd-Elmabod, S.; El-Shirbeny, M.A.; Gad, A.; Zahran, M.B. Smart farming for improving agricultural management. *Egypt. J. Remote Sens. Space Sci.* **2021**, *24*, 971–981. [CrossRef]
9. Verdouw, C.; Wolfert, S.; Tekinerdogan, B. Internet of Things in agriculture. *CAB Rev.* **2016**, *11*, 1–12. [CrossRef]
10. Doshi, J.; Patel, T.; kumar Bharti, S. Smart Farming using IoT, a solution for optimally monitoring farming conditions. *Procedia Comput. Sci.* **2019**, *160*, 746–751. The 10th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2019) / The 9th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2019) / Affiliated Workshops. [CrossRef]

11. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
12. Bengio, Y. Deep Learning of Representations for Unsupervised and Transfer Learning. In *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*; Guyon, I., Dror, G., Lemaire, V., Taylor, G., Silver, D., Eds.; PMLR: Bellevue, WA, USA, 2012; Volume 27, *Proceedings of Machine Learning Research*, pp. 17–36.
13. Kerkech, M.; Hafiane, A.; Canals, R. Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Comput. Electron. Agric.* **2020**, *174*, 105446. [CrossRef]
14. Jiang, P.; Chen, Y.; Liu, B.; He, D.; Liang, C. Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks. *IEEE Access* **2019**, *7*, 59069–59080. [CrossRef]
15. Karthik, R.; Hariharan, M.; Anand, S.; Mathikshara, P.; Johnson, A.; Menaka, R. Attention embedded residual CNN for disease detection in tomato leaves. *Appl. Soft Comput.* **2020**, *86*, 105933.
16. Silver, D.L.; Monga, T. In Vino Veritas: Estimating Vineyard Grape Yield from Images Using Deep Learning. In *Advances in Artificial Intelligence*; Meurs, M.J., Rudzicz, F., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 212–224.
17. Aguiar, A.S.; Magalhães, S.A.; dos Santos, F.N.; Castro, L.; Pinho, T.; Valente, J.; Martins, R.; Boaventura-Cunha, J. Grape Bunch Detection at Different Growth Stages Using Deep Learning Quantized Models. *Agronomy* **2021**, *11*, 1890. [CrossRef]
18. Ghiani, L.; Sassu, A.; Palumbo, F.; Mercenaro, L.; Gambella, F. In-Field Automatic Detection of Grape Bunches under a Totally Uncontrolled Environment. *Sensors* **2021**, *21*, 3908. [CrossRef]
19. Assunção, E.; Gaspar, P.D.; Mesquita, R.; Simões, M.P.; Alibabaei, K.; Veiros, A.; Proença, H. Real-Time Weed Control Application Using a Jetson Nano Edge Device and a Spray Mechanism. *Remote Sens.* **2022**, *14*, 4217. [CrossRef]
20. Wang, A.; Xu, Y.; Wei, X.; Cui, B. Semantic Segmentation of Crop and Weed using an Encoder-Decoder Network and Image Enhancement Method under Uncontrolled Outdoor Illumination. *IEEE Access* **2020**, *8*, 81724–81734. [CrossRef]
21. Lottes, P.; Behley, J.; Milioto, A.; Stachniss, C. Fully Convolutional Networks With Sequential Information for Robust Crop and Weed Detection in Precision Farming. *IEEE Robot. Autom. Lett.* **2018**, *3*, 2870–2877. [CrossRef]
22. Li, C.; Zhang, Y.; Ren, X. Modeling Hourly Soil Temperature Using Deep BiLSTM Neural Network. *Algorithms* **2020**, *13*, 173. [CrossRef]
23. Yu, F.; Hao, H.; Li, Q. An Ensemble 3D Convolutional Neural Network for Spatiotemporal Soil Temperature Forecasting. *Sustainability* **2021**, *13*, 9174. [CrossRef]
24. Saggi, M.K.; Jain, S. Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. *Comput. Electron. Agric.* **2019**, *156*, 387–398. [CrossRef]
25. Loggenberg, K.; Strever, A.; Greyling, B.; Poona, N. Modelling Water Stress in a Shiraz Vineyard Using Hyperspectral Imaging and Machine Learning. *Remote Sens.* **2018**, *10*, 202. [CrossRef]
26. Acharya, U.; Daigh, A.L.M.; Oduor, P.G. Machine Learning for Predicting Field Soil Moisture Using Soil, Crop, and Nearby Weather Station Data in the Red River Valley of the North. *Soil Syst.* **2021**, *5*, 57. [CrossRef]
28. Fang, K.; Shen, C.; Kifer, D.; Yang, X. Prolongation of SMAP to Spatiotemporally Seamless Coverage of Continental U.S. Using a Deep Learning Neural Network. *Geophys. Res. Lett.* **2017**, *44*, 11030–11039. [CrossRef]
28. Ek, M.B.; Mitchell, K.E.; Lin, Y.; Rogers, E.; Grunmann, P.; Koren, V.; Gayno, G.; Tarpley, J.D. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res. Atmos.* **2003**, *108*, 1–12. [CrossRef]
29. Paul, S.; Singh, S. Soil Moisture Prediction Using Machine Learning Techniques. In Proceedings of the 3rd International Conference on Computational Intelligence and Intelligent Systems, CIIS 2020, Tokyo Japan, 13–15 November 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–7. [CrossRef]
30. 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]
31. Hajjar, C.S.; Hajjar, C.; Esta, M.; Chamoun, Y.G. Machine learning methods for soil moisture prediction in vineyards using digital images. *E3S Web Conf.* **2020**, *167*, 02004. [CrossRef]
32. 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]
33. Chen, M.; Cui, Y.; Wang, X.; Xie, H.; Liu, F.; Luo, T.; Zheng, S.; Luo, Y. A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. *Agric. Water Manag.* **2021**, *250*, 106838. [CrossRef]
34. 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]
35. Pellegrino, A.; Lebon, E.; Voltz, M.; Wery, J. Relationships between plant and soil water status in vine (*Vitis vinifera* L.). *Plant Soil* **2005**, *266*, 129–142. [CrossRef]
36. Rallo, G.; Provenzano, G. Modelling eco-physiological response of table olive trees (*Olea europaea* L.) to soil water deficit conditions. *Agric. Water Manag.* **2013**, *120*, 79–88. Soil and Irrigation Sustainability Practices. [CrossRef]
37. Lebon, E.; Dumas, V.; Pieri, P.; Schultz, H.R. Modelling the seasonal dynamics of the soil water balance of vineyards. *Funct. Plant Biol.* **2003**, *30*, 699. [CrossRef]
38. Lopes, C.M.; Santos, T.P.; Monteiro, A.; Rodrigues, M.L.; Costa, J.M.; Chaves, M.M. Combining cover cropping with deficit irrigation in a Mediterranean low vigor vineyard. *Sci. Hortic.* **2011**, *129*, 603–612. [CrossRef]

39. Phogat, V.; Petrie, P.R.; Collins, C.; Bonada, M. Plant available water capacity of soils at regional scale: Analysis of fixed and dynamic field capacity. *Pedosphere* **2022**, *in press*. [CrossRef]
40. Ramos, M.; Casasnovas, J.M. Soil water variability and its influence on transpirable soil water fraction with two grape varieties under different rainfall regimes. *Agric. Ecosyst. Environ.* **2014**, *185*, 253–262. [CrossRef]
41. Valdés-Gómez, H.; Celette, F.; García de Cortázar-Atauri, I.n.; Jara-Rojas, F.; Ortega-Farías, S.; Gary, C. Modelling soil water content and grapevine growth and development with the stics crop-soil model under two different water management strategies. *Oeno One* **2009**, *43*, 13–28. [CrossRef]
42. Marhon, S.A.; Cameron, C.J.F.; Kremer, S.C., Recurrent Neural Networks. In *Handbook on Neural Information Processing*; Bianchini, M., Maggini, M., Jain, L.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 29–65. [CrossRef]
43. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
44. 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]
45. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
46. Dang, V.H.; Hoang, N.D.; Nguyen, L.M.D.; Bui, D.T.; Samui, P. A Novel GIS-Based Random Forest Machine Algorithm for the Spatial Prediction of Shallow Landslide Susceptibility. *Forests* **2020**, *11*, 118. [CrossRef]
47. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration—Guidelines for Computing Crop Water Requirements FAO Irrigation and Drainage Paper 56*; FAO—Food and Agriculture Organization of the United Nations: Rome, Italy, 1998.
48. Sinclair, T.; Ludlow, M. Influence of Soil Water Supply on the Plant Water Balance of Four Tropical Grain Legumes. *Funct. Plant Biol.* **1986**, *13*, 329. [CrossRef]
49. Monteiro, A.; Lopes, C.M. Influence of cover crop on water use and performance of vineyard in Mediterranean Portugal. *Agric. Ecosyst. Environ.* **2007**, *121*, 336–342. [CrossRef]
50. Hargreaves, G.H.; Samani, Z.A. Reference Crop Evapotranspiration from Temperature. *Appl. Eng. Agric.* **1985**, *1*, 96–99. [CrossRef]
51. Rodrigues, G.C.; Braga, R.P. Estimation of Reference Evapotranspiration during the Irrigation Season Using Nine Temperature-Based Methods in a Hot-Summer Mediterranean Climate. *Agriculture* **2021**, *11*, 124. [CrossRef]
52. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
53. Hsu, C.W.; Chang, C.C.; Lin, C.J. *A Practical Guide to Support Vector Classification*; Technical Report; Department of Computer Science, National Taiwan University: Taibei, Taiwan, 2003.
54. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
55. Bergstra, J.; Bardenet, R.; Bengio, Y.; Kégl, B. Algorithms for Hyper-Parameter Optimization. In Proceedings of the 24th International Conference on Neural Information Processing Systems, NIPS'11, Granada, Spain, 12–15 December 2011; Curran Associates Inc.: Red Hook, NY, USA, 2011; pp. 2546–2554.
56. Molnar, C. A Guide for Making Black Box Models Explainable. In *Interpretable Machine Learning*, 2nd ed.; Lulu Press, Inc.: Morrisville, NC, USA, 2022.
57. Wei, P.; Lu, Z.; Song, J. Variable importance analysis: A comprehensive review. *Reliab. Eng. Syst. Saf.* **2015**, *142*, 399–432. [CrossRef]
58. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer Publishing Company, Incorporated: Berlin/Heidelberg, Germany, 2014.
59. Prechelt, L., Early Stopping—But When? In *Neural Networks: Tricks of the Trade: Second Edition*; Montavon, G., Orr, G.B., Müller, K.R., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 53–67.
60. Brocca, L.; Morbidelli, R.; Melone, F.; Moramarco, T. Soil moisture spatial variability in experimental areas of central Italy. *J. Hydrol.* **2007**, *333*, 356–373. [CrossRef]
61. Cosh, M.H.; Stedinger, J.R.; Brutsaert, W. Variability of surface soil moisture at the watershed scale. *Water Resour. Res.* **2004**, *40*, 1–9. [CrossRef]