Prediction of Irrigation Water Requirements for Green Beans-Based Machine Learning Algorithm Models in Arid Region

Ali Mokhtar[1] · Nadhir Al-Ansari[2] · Wessam El-Ssawy[1,3] · Renata Graf[4] · Pouya Aghelpour[5] · Hongming He[6] · Salma M. Hafez[1] · Mohamed Abuarab[1]

Abstract

Water scarcity is the most obstacle faced by irrigation water requirements, likewise, limited available meteorological data to calculate reference evapotranspiration. Consequently, the focal aims of the investigation are to assess the potential of machine learning models in forecasting irrigation water requirements (IWR) of snap beans by evolving multi-scenarios of inputs parameters to figure out the impact of meteorological, crop, and soil parameters on IWR. Six models were applied, support vector regressor (SVR), random forest (RF), deep neural networks (DNN), convolutional neural networks (CNN), long short-term memory (LSTM), and Hybrid CNN-LSTM. Ten variables including maximum and minimum temperature, Relative humidity, wind speed, precipitation, root depth, basal crop coefficient, soil evaporation, a fraction of surface wetted and, exposed and soil wetted fraction were used as the input data for models with their combination, 8 input scenarios were designed. Overall models, the best scenario was scenario 4 (relative humidity, wind speed, basal crop coefficient, soil evaporation), however, the best scenario for DNN and RF model was scenario 7 (root depth, basal crop coefficient, soil evaporation, fraction of surface wetted, exposed and soil wetted fraction). While the weakest one was the group of climatic factors in scenario 6 (maximum temperature, minimum temperature, relative humidity, wind speed, and precipitation). Among the models, the hybrid LTSM & CNN was the most accurate and the SVR model had the lowest estimation accuracy. The outcomes of this research work could set up a modeling strategy that would set in motion the improvement of efforts to identify the shortages in IWR forecasting, which sequentially may support alleviation strategies such as policies for sustainable water use and water resources management. The current approach was promising and has research value for other similar regions.

Extended author information available on the last page of the article

1 Introduction

The widespread climatic changes in the twenty-first century and the negative impacts that follow on the available water resources have become one of the most important issues that cast their shadows on the focus of contemporary environmental events and issues (Smith et al. 2012; UNEP 1999). The irrigated agriculture sector represents the largest consumer of water in Egypt, representing 85% of the total water share available to Egypt, which represents 55.5 billion cubic meters. In addition, the biggest problem is that most of the irrigated lands apply surface irrigation systems where the efficiency of application and distribution of water do not exceed 50%, which represents the biggest challenge that must be faced to reduce water losses (Monem 2021). Irrigated agriculture covers about 25% of the cultivated area in the world, yet it contributes 50% of the global food basket. Despite the increase in production with irrigated agriculture, it consumes 67% of the water in agricultural production and 87% of the water is consumed in the irrigation sector (Smith et al. 2012; Shiklomanov 1997). The availability of food is the main element for the survival of the human race, although it faces many challenges, the main cause of which is humans, perhaps the most prominent of which are the rapid population growth and climatic changes resulting from global warming that affect the hydrological cycle (Bellido-Jiménez et al. 2021). Therefore, irrigation scheduling, which is determining the amount of irrigation water added and the time of its addition during the growing season of the crop, is one of the things that must be taken into account under conditions of water scarcity (Kalboussi et al. 2019).

There are many methods used for irrigation scheduling, including visual and thermal monitoring, soil moisture content measuring devices, rainfall estimation devices, and reference evapotranspiration estimation, from which it is possible to estimate the crop evapotranspiration by knowing the crop coefficient, through available climatic data on local basis or through commercial scheduling programs available through the international information network (Zhang et al. 2013). The actual crop evapotranspiration (ETa) is considered one of the main criteria for estimating the amount of water consumed by the crop, and thus greatly affects the hydrological, environmental and agricultural management, and then the ETa is one of the basic elements when designing and managing irrigation networks in the agricultural project. ETa is estimated through field measurements using lysimeters or through empirical equations by estimating the reference evapotranspiration (ETo) and the crop coefficient (Djaman et al. 2018). The water balance method for irrigation scheduling includes estimating of the crop water consumption during the growing season, which includes the amount of water added, whether through irrigation, precipitation and capillary rise, which subtracts the amount of water losses, whether by evaporation from the soil surface, deep percolation the rootzone area, and runoff (Andales et al. 2014; Karam et al. 2019a, b).

The estimation of ETo is one of the most important indicators used in irrigation water management and hydrological studies, and despite the difficult requirements of the empirical equations of ETo, it gives accurate values in addition to its validity on various climatic conditions, which are the main indicators in the empirical equations (Ferreira and da Cunha 2020a, b; Pereira et al. 2015a, b). In light of the limited climatic parameters available for estimating ETo, fast and accurate methods that require least number of meteorological indicators are considered among the most important requirements for irrigation water scheduling (Fan et al. 2021). The Penman–Monteith equation is the standard equation for calculating ETo and is recommended by the United Nations Agriculture and Food Organization.

The Penman Monteith equation requires many daily meteorological parameters to calculate such as maximum and minimum temperature, wind speed, net solar radiation, and relative humidity. However, under the conditions of limited climatic parameters available in the study area, which requires the availability of methods that require less climatic data, while maintaining the accuracy of the calculated values of ETo in the same range to a large extent (Fan et al. 2019). The application of the Penman Monteith equation faces many difficulties, especially in developing countries, which face many challenges represented in the limited monitoring stations and their lack of regular distribution in each country, where the focus is in the areas of spread of irrigated agricultural areas in addition to the lack of ability to measure all the climatic criteria required to apply the Penman–Monteith equation with the required accuracy (Fan et al. 2021).

Recently used computer software's models have shown high accuracy in estimating and forecasting ETo, e.g. Support Vector Machines (SVM) (Fan et al. 2019). In recent years, the use of machine learning programs for ETo estimation has spread with making relationships between the inputs and outputs used in ETo estimation, which are mainly meteorological data, which gives higher accuracy and power to apply machine learning programs in ETo modeling (Ferreira and Cunha 2020a, b; Kumar et al. 2011). The use of both Random Forest (RF) and Generalized Regression Neural Network (GRNN) to predict ETo proved highly accurate under local conditions and across-stations applications, but RF was better than GRNN (Feng et al. 2017). The common method of predicting ETo values using machine learning models is based on the use of former ETo values as input data (Ashrafzadeh et al. 2020; Landeras et al. 2009; Trajkovic et al. 2003). In addition, the use of former daily meteorological data used to estimate ETo through the Penman–Monteith equation in addition to the lagged ETo values greatly enhances the performance of machine learning models and the accuracy of the forecasting ETo.

In addition to what machine learning programs provide an estimate of ETo values for the previous and current agricultural seasons, it gives the possibility to predict the ETo values for the upcoming seasons using the so-called deep learning models, which are restricted to knowing the expected changes in temperature and the amount of precipitation related to global warming, from which it can be make good planning and management of irrigation water for the coming years with an attempt to correct the existing strategies to avoid the negative effects of global warming (Ferreira and da Cunha 2020a, b). By increasing the forecast period from 1 to 30 days, the accuracy of the resulting data decreases (Ferreira and da Cunha 2020a, b). Several deep learning models have been used to forecasting time series, including long short-term memory (LSTM) (Son and Kim 2020; Tian et al. 2018; Zhou et al. 2019), and a one-dimensional convolutional neural network can (1D CNN) (Barzegar et al. 2020). In addition to the above, Barzegar et al. (2020) and Kim et al. (2019) create hybrid software from both LSTM and 1D CNN. Although deep learning programs have proven outstanding performance in many fields, their use in the field of climatic and hydrological studies is still limited.

Therefore, this paper aims to participate in saving water tasks by forecasting irrigation water requirement (IWR) for green bean by an on-field scale experiment, where there wasn't common literature about modeling of irrigation water requirement forecasting in versus of ETo estimation studies. Moreover, the irrigation water requirements forecasting could make irrigation scheduling improved, and so give better solutions for decision makers. This research is based on determining the most important criteria required to forecast IWR of green bean to achieve the best irrigation scheduling in light of the challenges of climatic changes and water scarcity using six models (SVR, RF, DNN, CNN, LSTM and CNN-LSTM) to forecast IWR. To the best of authors' knowledge, the applied approaches

are applying for first time for forecasting of IWR, further, no research has been published in the literature that analyses irrigation water requirement using a combination of climate data, soil data and crop data based novel artificial intelligence approaches applied. Thus, to overcome the gap in literature dealing with the importance of IWR prediction, the main objectives of this study are to 1) evaluate the potential of machine and deep learning models in forecasting IWR of snap beans, 2) develop multi-scenarios of the inputs variables (combinations) to study the impact of meteorological, crop and soil parameters on IWR and, 3) figure out the best scenario with best model performance in IWR prediction.

2 Materials and Methods

The workflow of this study is showing in Fig. 1. The first phase of the workflow involves the climate, soil and crops data collection and the calculation of the IWR during two seasons. Within the second phase, we applied five machine learning models (SVR, RF, DNN, LSTM and CNN) and hybrid between LSTM and CNN to predict the IWR based on 8 Scenarios (combinations of climate, soil and crop data).

2.1 Climate Conditions

The climatic variables (maximum air temperature, minimum temperature of air, the average air temperature, air relative humidity, and the number of hours of sunshine) were daily recorded during both growing seasons. The maximum air temperature ranged during the growing season (October to December) between 19–30 °C, while minimum air temperatures ranged from 11 to 21 °C during growing season. The total rainfall was negligible
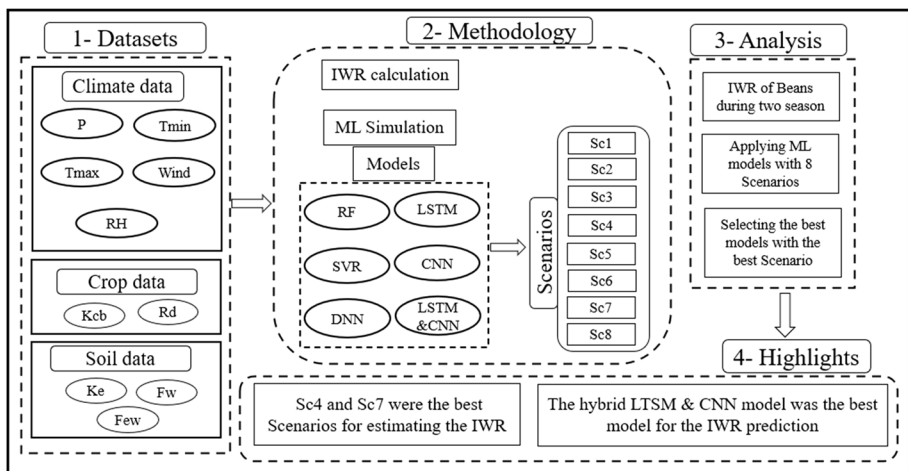

Fig. 1 Flowchart of the research. Note: Tmax: maximum temperature, Tmin: minimum temperature, RH: Relative humidity, WS: wind speed, P: precipitation, Rd: root depth, Kcb: basal crop coefficient, Ke is the soil evaporation, Fw: fraction of surface wetted, Few: Exposed and soil wetted fraction, SVR: support vector regression, RF: random forest, DNN: deep neural network, CNN: convolutional neural networks, LSTM: long short-term memory, Sc: scenarios

(< 20 mm) thus, the irrigation is main variable for the soil water availability. Table 1 summarizes the monthly mean climatic data during the two seasons.

2.2 Machine Learning Implementations

2.2.1 Support Vector Machine (SVM)

SVM is a supervised learning algorithm, and it can also be used as a regression model, maintain all the main features that describe the algorithm (maximal margin). SVR uses a similar theory as SVM for classification method, with a few slightly changes. The main aim is minimizing error, individualizing the hyperplane that increases the tolerance limit, taking in consideration that the part of error is tolerated. The approximated function in the algorithm of SVM is cleared as follows:

$$f(\chi) = \omega\varphi(\chi) + b \tag{1}$$

where $\varphi(x)$ is feature space of higher-dimensional converted from the input vector x, ω is the weights vector and b are a threshold, which were estimated by minimizing the following regularized risk function:

$$R(C) = C\frac{1}{n}\sum_{i=1}^{n} L(L_i, y_i) + \frac{1}{2}\varpi^2 \tag{2}$$

where C is the penalty parameter of the error, d_i is the desired value, n is observations number, and $C\frac{1}{n}\sum_{i=1}^{n} L(d_i, y_i)$ is the empirical error, in which the function L_ε is determined as:

$$L_\varepsilon(d, y) = |d - y| - \varepsilon|d - y| \geq \varepsilon \ or \ 0 \ otherwise \tag{3}$$

where $0.5\, \varpi^2$ is the so-called regularization term and ε is the tube size. The approximated function in Eq. (1) is finally expressed in an explicit form by introducing Lagrange multipliers and exploiting the optimality constraints as:

Year	Climatic variables	Month		
		October	November	December
2017	T$_{min}$. (Rockström et al. 2009)	21	17	12
	T$_{max}$ (Rockström et al. 2009)	30	26.3	20
	T$_{ave}$ (Rockström et al. 2009)	26	21	16
	Relative humidity (%)	59.9	62	53
	Sunshine (h)	11.5	10.5	10.2
2018	T$_{min}$ (Rockström et al. 2009)	18	17	11
	T$_{max}$ (Rockström et al. 2009)	28	26	19
	T$_{ave}$ (Rockström et al. 2009)	23	21	15
	Relative humidity (%)	57	59	56
	Sunshine (h)	11.4	10.9	10.3

Table 1 Average monthly climatic variables for both growing seasons during the experiments

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) k(x, x_i) + b \tag{4}$$

where k (x, x_i) is the kernel function. (Fan et al. 2018) provide detailed information and computation procedures of the SVM algorithm.

2.2.2 Random Forest (RF)

The RF model was developed by Breiman (2001) and used Breiman's "bagging" idea to ensemble a collection of decision trees with controlled variance. The RF model is commonly used for regression and predicting problems. More detailed information can be found at (Ferreira and da Cunha 2020a, b; Breiman 2001).

2.2.3 Deep Neural Network (DNN)

Deep neural networks (DNN) model has become a fundamental but recently still powerful deep learning model (Achieng 2019; Montes-Atenas et al. 2016). The DNN is an artificial neural network (Springmann et al. 2018) with multiple layers between the inputs, hidden layers and output layers to learn more complex nonlinear relationships between inputs and output. The rectified linear unit (ReLu) was applied as activation function which is commonly employed (Ghimire et al. 2019; Achieng 2019).

$$f(a_i^l) = max(0, a_i^l) \tag{5}$$

The loss function in the DNN model is expressed as:

$$loss = \frac{1}{2n} \sum_{i=1}^{n} (T_i - T_i')^2 \tag{6}$$

where n is the number of observation data, and T' is the estimated value by the DNN model, which can be defined below for a three-hidden-layer DNN model with the ReLu activation function:

$$T' = Re\,Lu\,\overline{\varpi}_1(\overline{\varpi}_3[Re\,Lu(\overline{\varpi}_2(Re\,Lu)(\overline{\varpi}_1 + b_1) + b_3)] + b_4 \tag{7}$$

where ω_1, ω_2, ω_3 and ω_4 are the weights in the network, b_1, b_2, b_3 and b_4 are the bias terms.

2.2.4 Convolutional Neural Network (CNN)

CNN consists of a sequence of convolutional layers, the output of which is connected only to local regions in the input. This is achieved by sliding a filter, or weight matrix, over the input and at each point computing the dot product between the two (i.e., a convolution between the input and filter). This structure allows the model to learn filters that are able to recognize specific patterns in the input data. Recent advances in CNNs for time series forecasting include (Mittelman 2015). Zuo et al. (2019) stated that the architecture of a typical CNN consists of a convolutional layer pooling layer, and fully connected layer. Finally, these abstract features are merged through the fully connected layer, and an activation function is used to solve the classification or regression problem (Fig. 2).
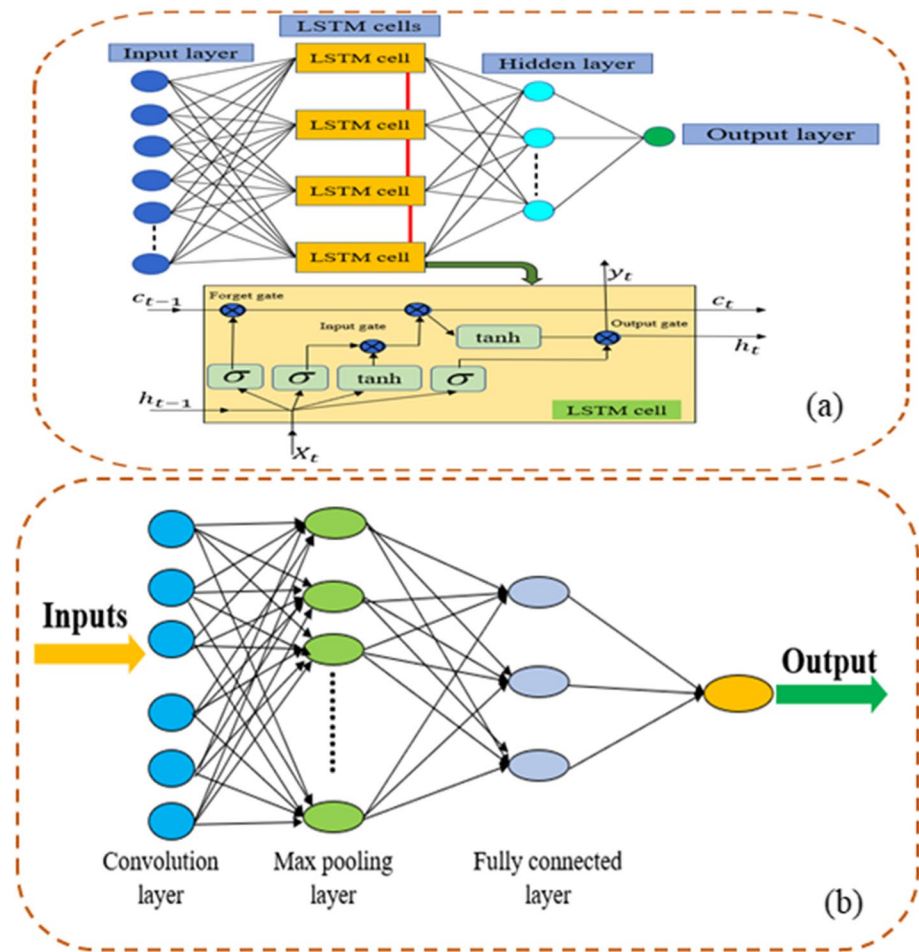
Fig. 2 Architecture of the LSTM model and cell structure (**a**) and the formation of the CNN model (**b**) (Mokhtar et al. 2021a, b)

2.2.5 Long Short-Term Memory (LSTM)

LSTM network contains different memory blocks which are linked through layers. Each layer includes a set of frequently connected memory pixels and three multiplicative units, namely the input, forget and output gates. LSTM has the ability to automatically store and remove temporal state information.

2.2.6 Hybrid LSTM and CNN

LSTM and CNN were trained with the same input, and they were hybrid to forecast results, the proposed hybrid CNN-LSTM model uses CNN layers for feature extraction from the input data with LSTM layers for sequence learning. The CNN and the LSTM are the most commonly used deep learning models. Our main aim in designing the hybrid model of CNN and LSTM layers is to exploit their characteristics for developing an efficient model for forecasting.

The hyper parameters of convolutional and LSTM layers were the same ones described above for CNN and LSTM models, respectively. Training algorithm, learning rate, batch size and number of training epochs followed the same configurations described for LSTM and CNN.

2.3 Input Combination and Performance Evaluation of the Models

In this study, eight scenarios were consisting of various combinations of the climatic, soil and crop variables to investigate their effects on IWR and evaluate the contribution of each variable (Table 2). This may help improve water resources analysis for data scarcity regions. The two seasons' data were divided into two subsets with 75% data for training and the remaining 25% data for testing.

The mean absolute error (MAE), the root mean square error (RMSE), and mean bias error (Springmann et al. 2018) were used to evaluate the applied models. Moreover, uncertainty with 95% confidence level (U95) and T-Statistic test (Tstat) to evaluate the significant differences between the predicted and calculated yield were used (Behar et al. 2015; Gueymard 2014)(Stone 1994).

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |O_i - P_i| \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n} \sum (P_i - O_i)^2} \tag{9}$$

$$MBE = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i) \tag{10}$$

$$SI = \frac{RMSE}{\overline{O}} \tag{11}$$

Table 2 The summary of the 8 scenarios applied in this study

Scenario	T_{max}	T_{min}	RH	WS	P	R_d	K_{cb}	K_e	F_w	F_{ew}
S1	√	√					√	√		
S2	√	√	√				√	√		
S3	√	√		√			√			
S4			√	√			√	√		
S5			√	√						
S6	√	√	√	√	√					
S7						√	√	√	√	√
S8	√	√	√	√	√	√	√	√	√	√

Tmax maximum temperature, *Tmin* minimum temperature, *RH* Relative humidity, *WS* wind speed, *P* precipitation, *Rd* root depth, *Kcb* basal crop coefficient, *Ke* is the soil evaporation, *Fw* fraction of surface wetted, *Few* Exposed and soil wetted fraction

$$T_{stat} = \sqrt{\frac{(1-n)MBE^2}{RMSE^2 - MBE^2}} \tag{12}$$

$$U_{95} = 1.96\sqrt{(SD^2 + RMSE^2)} \tag{13}$$

$$NSE = 1 - \frac{\sum(P_i - O_i)^2}{\sum(\overline{O} - O_i)^2} \tag{14}$$

$$MSE = \frac{1}{n}\sum(P_i - O_i)^2 \tag{15}$$

$$CC = \frac{\sum_{i=1}^{n}(O_i - \overline{O})(P_i - \overline{P})}{\sqrt{(\sum_{i=1}^{n}(O_i - \overline{O})^2)(\sum_{i=1}^{n}(P_i - \overline{P})^2)}} \tag{16}$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(O_i - \overline{O})(P_i - \overline{P})}{\sqrt{(\sum_{i=1}^{n}(O_i - \overline{O})^2)(\sum_{i=1}^{n}(P_i - \overline{P})^2)}}\right]^2 \tag{17}$$

$$A = 1 - abs(mean\frac{P_i - O_i}{O_i}) \tag{18}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{P_i - O_i}{O_i}| \times 100 \tag{19}$$

$$PSR = \left[\frac{\sum(O_i - P_i)^2}{\sqrt{\sum(O_i - \overline{O})^2}}\right] \tag{20}$$

where $\overline{O}$ represent the average values of the observed IWR, O_i and P_i are the observed and predicted IWR, respectively, and i is the number of observations. SD is the standard deviation of the difference between the measured and estimated value. PSR is ratio of the root mean square error to the standard deviation of measured data. Further, A is the accuracy, R^2 is the coefficient of determination and MAPE is the mean average percentage error.

3 Results

At first, the input variables were arranged according to Table 2. The combination of all 10 inputs, reached 8 input scenarios which must be trained by the models.

The input scenarios were randomized and divided into training and testing phases, then entered the models. The models were implemented, and their estimations were evaluated by the criteria MBE, RMSE and NSE (Table 3).

There are noticeable differences between the models, but also some similarities in relation to individual scenarios. The performance of the models is more favorable when the RMSE and MBE are closer to zero, and NSE are closer to 1. In all models, the NSE values are higher than 0.75, which shows that the models are well performed in IWR estimation. The best estimation of the models are mostly presented by the S4 scenario and after that S7. The weakest accuracy was obtained by SVR model with MBE = 0.007, RMSE = 0.077 mm and NSE = 0.850, under the S4 input scenario. The most accurate estimation was resulted by DNN model under S7 scenario, with MBE = -0.001, RMSE = 0.055 mm and NSE = 0.824. In order to examine the correlation between the actual and predicted IWR values, scatter plots were developed (Fig. 3).

The analysis showed that in the S1 scenario, the highest correlation was obtained for the LSTM and CNN models ($R^2 = 0.9421$, $R^2 = 0.9119$, respectively), the lowest for SVR. S1 scenario takes into account air temperature (Tmax and Tmin) and basal crop coefficient (Kcb) and soil evaporation (Ke). In the second scenario S2, where in addition to S1, relative humidity (RH), the strongest correlation between real and predicted IWR was also shown for the LSTM model ($R^2 = 0.83225$), and the weakest for SVR ($R^2 = 0.6369$). In the S3 scenario (4 parameters: temperature, wind speed and Kcb), the best correlation between actual and predicted IWR values was determined for the LSTM & CNN hybrid model ($R^2 = 0.7618$). This is a relatively low correlation, which indicates that the IWR prediction

Table 3 Evaluating the models' estimation performance

| Model	Index	Input scenario							
		S1	S2	S3	S4	S5	S6	S7	S8
LTSM	MBE (mm)	−0.015	0.003	−0.004	**0.002**	0.01	−0.001	0.002	0.008
	RMSE (mm)	0.075	0.084	0.104	**0.064**	0.087	0.121	0.071	0.072
	NSE	0.929	0.813	0.649	**0.889**	0.101	0.578	0.846	0.859
LTSM + CNN	MBE (mm)	−0.009	0.016	−0.003	**−0.003**	−0.005	0.008	−0.005	0.009
	RMSE (mm)	0.137		0.13	**0.060**	0.117	0.123	0.062	0.069
	NSE	0.768	0.716	0.757	**0.889**	0.595	0.511	0.896	0.866
CNN	MBE (mm)	0.000	0.015	0.007	**0.007**	0.001	0.015	0.006	−0.003
	RMSE (mm)	0.058	0.066	0.13	**0.066**	0.109	0.128	0.109	0.047
	NSE	0.907	0.788	0.497	**0.895**	0.838	0.775	0.842	0.938
DNN	MBE (mm)	0.011	0.010	−0.009	−0.014	0.023	0.003	**−0.001**	0.006
	RMSE (mm)	0.065	0.061	0.085	0.06	0.082	0.08	**0.055**	0.033
	NSE	0.753	0.778	0.57	0.789	0.600	0.621	**0.824**	0.936
SVR	MBE (mm)	0.012	0.029	0.03	**0.007**	−0.032	−0.061	−0.009	−0.024
	RMSE	0.121	0.169	0.194	**0.077**	0.140	0.127	0.108	0.070
	NSE	0.599	0.609	0.539	**0.850**	0.423	0.098	0.719	0.811
RF	MBE (mm)	0.003	−0.034	−0.03	−0.007	−0.018	0.003	**−0.020**	−0.05
	RMSE (mm)	0.072	0.085	0.089	0.117	0.107	0.150	**0.066**	0.108
	NSE	0.854	0.757	0.609	0.539	0.527	0.537	**0.818**	0.293

* The bold rows show the best performance of each model
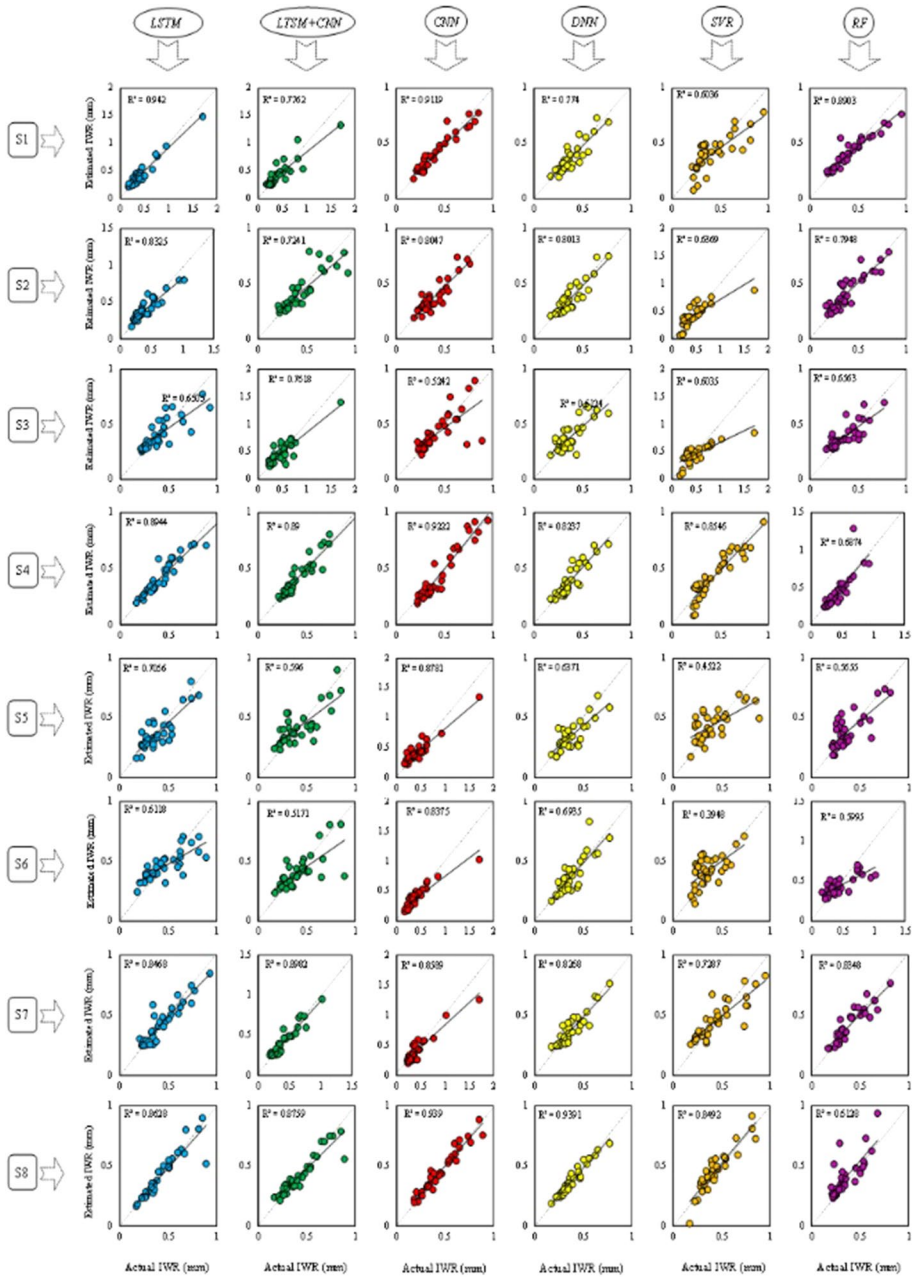
Fig. 3 Scatter plots to investigate the correlations between actual and predicted IWR values

is also influenced by RH and Ke, and there is Kcb and wind speed, in addition to air temperature, is not sufficient this parameter. In the S3 scenario, the results are not promising, a relatively low R^2 was obtained, which means that the predictors used support the prediction of IWR within 50–76%. The weakest correlation was shown by the CNN model ($R^2 = 0.5$)

but the LSTM & CNN hybrid already gives a good correlation result between actual and predicted IWR values.

In the S4 scenario, where wind speed, RH, Kcb and Ke are considered, the R^2 value ranges from 0.6874 for the RF model to 0.922 for the CNN model. Apart from RF, the remaining models showed a high correlation of actual and predicted IWR, which emphasizes that the selection of predictors in the discussed scenario is optimal for the prediction. In the S5 scenario, considering only the two predictors RH and WS, the best correlation for the current and predicted IWR values was obtained for the CNN model ($R^2=0.8781$), while weak correlation applies to the SVR ($R^2=0.4522$) and RF ($R^2=0.5655$) models. In the next scenario S6, where the group of climatic factors (Table 3) was considered as Input, the highest R^2 was obtained for the CNN model (0.837), while the weakest result in the SVR model ($R^2=0.3948$). In this variant, the remaining models showed a weak ($R^2=0.5$–0.6) correlation between actual and predicted IWR values, which indicates that climatic factors alone are not sufficient for predicting of irrigation requirement (IR), predicting irrigation demand (IR). In the S7 scenario, which takes into account only non-climatic parameters root depth, k_{cb}, k_e, f_w and f_{ew} (Table 3), in almost all models a similar dimension of the correlation strength between actual and predicted IWR values ($R^2>0.8$) was obtained, except for the SVR model ($R^2=0.7287$). In the last scenario S8, in which all predictors were taken into account (Table 3), the highest $R^2>0.9$ were obtained for the CNN and DNN models, and the weakest correlations concerned the RF model ($R^2=0.6128$). In the next step, for the test period, combo charts (combined line-bar chart) were used to compare the scenarios based on the CC and U_{95} criteria and in the second variant A and PSR (Fig. 4).

Comparison charts (combo-graphs) show that the values of CC and U_{95} show wide differentiation. The analyzes show that the highest CC=0.97 was determined for the S1 scenario in the LTSM model. However, in the case of the CNN model, for most of the scenarios (except S3, which takes into account temperature, wind speed and crop factor), the CC coefficient is above CC=90. The uncertainty of the scenarios adopted in modeling was additionally determined for the limits of the 95% confidence interval (U_{95}), which means the probability of obtaining the result lying close to the expected value within the range defined by this uncertainty. S5 and S6 scenarios do not work for LTSM and LTSM & CNN, SVR and RF models. However, they are slightly better for IWR prediction using the CNN model itself. The S1 scenario seems to be the most optimal, as it achieves satisfactory performance for all models, except for the SVR model. The S4 scenario is also favorable, with the exception of the RF model. However, in the case of the remaining scenarios, a differentiated assessment of the CC and U_{95} criteria was obtained.

In the second case, the accuracy (A) and PSR. Metrics were used to assess the correctness of the scenarios used. These criteria are used to obtain more relevant information from the time series. Model accuracy (A) is defined as the percentage of correct predictions for the test data, it denotes the degree of compliance of the actual value with the arithmetic mean of the results (predicted value) obtained for the marked scenario and prediction model. For accuracy, results above 80% are considered good. In some scenarios and models, the obtained values were higher than 1, which means that the prediction results in these variants may be burdened with a random error, which can be considered a random error, i.e., a type of measurement error that does not result from systematic and repeatable factors.

In the case of the CNN and DNN models in particular, overestimated A values were obtained, which may be related to a random error. For the S1-S4, the best accuracy was demonstrated by the SVR model. For the S5 forecast, the LTSM model was the most accurate, while for the other S6-S7 forecasts, the first 4 models presented the most accurate.

Fig. 4 Combo-graphs for comparison between the scenarios, based on the criteria CC, U_{95}, A and PSR

Another analyzed RSR indicator is strongly related to the RMSE criterion, as it is understood as the ratio of the RMSE and the standard deviation of the measured data. Better performance is shown by models with a lower RSR value, which at the same time means a

lower RMSE. A very good model is classified when $0.00 \leq RSR \leq 0.50$, while unsatisfactory when $RSR > 0.70$. The analyzes show that the lowest RSR values (≤ 0.50), covering very good ranges of classification, were achieved by all models for the S4 scenario (Input: RH, WS, Kcb, Ke in Table 2), and then by four models: LTSM, CNN, DNN and RF for the S1 scenario (Fig. 4).

The LTSM, LTSM & CNN and DNN models additionally achieved high performance in the S7 scenario, and the CNN and DNN models also in the S2 scenario. The SVR model was the only one that achieved very good quality in only one S4 scenario. The SVR model was the only one that achieved very good quality in only one S4 scenario. In most models, the lowest RSR was achieved in those scenarios where the lowest values also reached the RMSE criterion (Table 3), which proves that the scenarios seem to be the most optimal for IWR prediction. Unsatisfactory results were achieved in those scenarios where the RSR criterion > 0.70. The most common scenarios are: S3 (Input: Tmax, Tmin, WS, Kcb), S5 (Input: RH, WS) and S6 (Input: Tmax, Tmin, RH, WS, P), thus mostly scenarios that include climatic parameters as input to the model (Fig. 4). In the next stage (Fig. 5), radar charts were developed to compare the models in each scenario based on the SI criterion.

The SI values obtained for the models in various S1-S8 scenarios. According to the value of the SI criteria, the performance of models can be divided into several levels. In the discussed case, the $SI < 0.1$ was not obtained, which would classify the models as "excellent". Good (if $0.1 < SI < 0.2$) and satisfactory (if $0.2 < SI < 0.3$) performance of most models was obtained for the S8 scenario, which takes into account all the inputs: climate, vegetation characteristics and soil data, and for the scenarios: S4 (takes into account 4 components RH, WS, Kcb and Ke) and the S7 scenario (Root depth, Kcb, Ke, Fw and Few). Satisfactory performance was achieved by the models in the S1 scenario, which takes into account temperature data, Kcb and Ke, and in the S5 scenario, which includes only two outputs, namely RH and WS. Poor performance was achieved by the SVR model for the S2 and S3 scenarios, and the RF and CNN models for the S6 scenario, which takes into account only climatic data. The models in the S6 scenario show the least satisfactory performance (Fig. 5).

The LTSM model, according to the Scatter index, achieved the best performance for S4 and the weakest for S6, the LTSM & CNN hybrid model for S4 and S1 / S6 respectively, the CNN model for S1 / S4 and S6 respectively, the DNN model for S4 / S7 and S5, the SVR model for S4 and S3, while the RF model for the S7 and S6 scenarios. Figure 5 shows that, apart from the distinguished single cases related to specific scenarios, the prediction accuracy in most of the developed models regarding the SI value was good ($0.1 < SI < 0.2$) and satisfactory ($0.2 < SI < 0.3$). Figure 6 shows the MAE values for 6 different models for 2 options: under and over estimation.

The MAE is a measure of performance (model quality) used to evaluate the performance of a model after finalization. Investigating the models under and over estimations, i.e., the study of underestimating and overestimating model estimates was carried out using mean absolute error (MAE). In the context of machine learning, absolute error refers to the magnitude of the difference between the prediction of an observation and the true value of that observation. The values on two different data sets indicate that MAE values in under estimation in almost all models were higher than after Over estimation, except for the RF model, where higher MAE was achieved in over estimation. The RF approach is another widely used decision tree method (Cavallo et al. 2017; García Nieto et al. 2017, 2018) which, using the classification and regression tree procedure, combined with random node optimization and packing (Breiman 2001) builds a forest of uncorrelated trees.
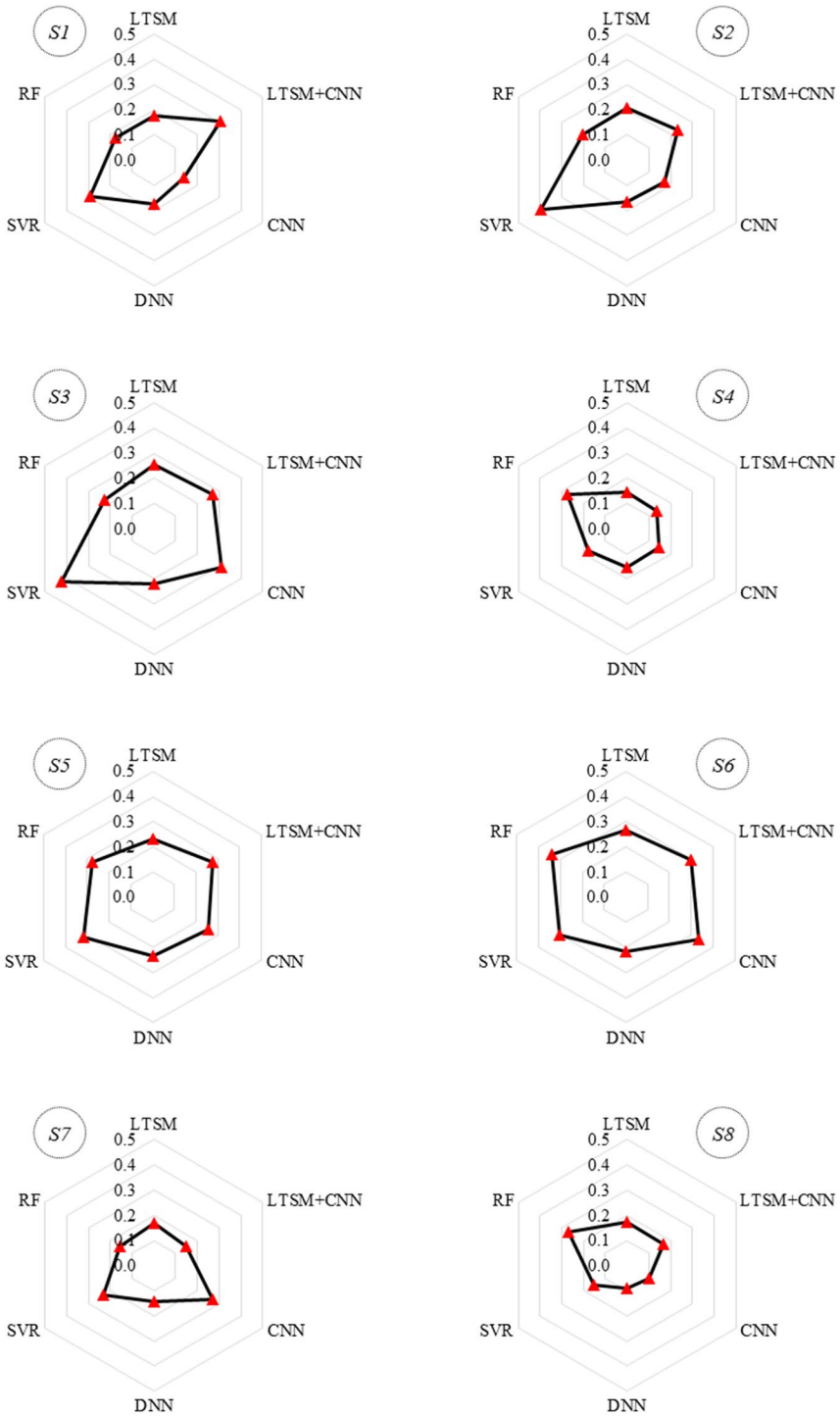
Fig. 5 Radar charts to compare the models in each scenario separately, based on the SI criterion
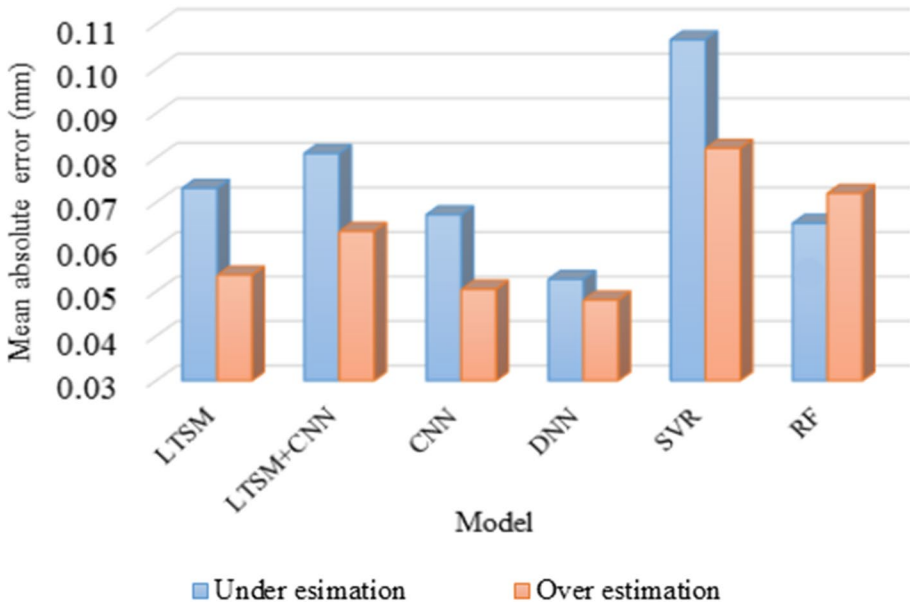
Fig. 6 Investigating the models' under and over estimations

However, in the case of noise-rich datasets, RF models tend to over-fit the data. Moreover, over-selecting too many characteristic values also has a greater impact on RF decision making, which affects goodness of fit (Naghibi et al. 2017). The MAE values decrease significantly after over estimation in the case of the following models: LTSM, LTSM & CNN, CNN and SVR) (they are within 0.05). This proves that over estimation can reduce the MAE values for most tested models. The smallest MAE differences between the two variants (under and over estimation) were obtained in the DNN model, while the largest in the SVR model. The use of the LTSM & CNN hybrid model reduced the MAE between the two options Under and over estimation.

4 Discussion

The assessment of the amount of irrigation for individual types of crops requires the analysis of a large number of parameters, which are often difficult to obtain through direct measurement and observation are often considered burdensome and costly for farmers. Hardly any country has a good system for measuring and recording the total water consumption of crops, and more than 40% of crops are grown under irrigation conditions. Hence, in order to assess the future situation in terms of obtaining high yields and maintaining rationalization in water abstraction for crops, it is necessary to model irrigation water requirements (Pulido-Calvo and Gutierrez-Estrada 2009; Döll and Siebert 2002). Modeling of the irrigation water needs as a function of the irrigated area, climatic conditions and type of crops has a broader context, providing the basis for estimating the future impact of climate change, as well as demographic, socio-economic and technological changes.

In this study, various machine learning models (SVM, RF, DNN, CNN, LSTM and hybrid LSTM & CNN) are investigated to predict the IWR indispensable for the green bean crops in Egypt. Importantly, it presents the adaptation of typical deep neural network models, which include the DNN model and the convolutional neural network, long short-term memory network, and finally a combination of some of their variants. An important element in the study was the selection of the scenario and model that obtains the best performance for the conditions declared in the scenario. Indication of the best predictive model for IWR, with the given input data for modeling, is of great importance for the selected research area that experiences extreme droughts and the related needs for intensive irrigation of crops. Mokhtar et al. (2021a, b) emphasizes that the droughts in the research area in 2015–2019 covered even 80–90% of the area, and 30% in 2010–2019 were affected by extreme and very extreme droughts.

The reason for water shortages in the root zone is primarily water consumption by crops and evapotranspiration, while rainfall and irrigation ensure water supply (Andales et al. 2011; 2014; Karam et al. 2019a, b). However, it should be emphasized that the actual evapotranspiration of green bean crops, one of the dominant components in the water balance, significantly reduces the size of its harvest, which on the other hand, with increasing evaporation, will require increased irrigation of the area. Due to the complexity of the evapotranspiration process, it is most often estimated on the basis of meteorological data (Pereira et al. 2015a, b; Ferreira and Cunha 2020a, b). However, in this case, monitoring of all parameters is often insufficient (El Bilali and Taleb 2020).

In the performed IWR prediction for green bean crops in Egypt, the accurate prediction of IWR was largely dependent on the number of input variables and their impact on the modeling results under certain scenario variants. According to Krupakar et al. (2016), the sources of parameters influencing the irrigation process tend to differ depending on the area where agriculture is grown. In addition, there are new characteristics that we obtain thanks to new technologies for terrestrial and satellite measurements. The sources that make irrigation water demand prediction possible can be broadly categorized into meteorological factors, crop input and agricultural parameters. Meteorological factors are important in determining the irrigation water needs of crops as rain is the main source of water in some areas (e.g., India). Wind speed also affects the amount of irrigation needed, although not as much as the temperature (O'Toole and Hatfield 1983; Schlenker and Roberts 2009; Krupakar et al. 2016).

In the presented work, the climate data was adapted to the S6 scenario (Table 2), which took into account the following parameters: maximum and minimum temperature, relative humidity, wind speed and precipitation. In the case of having information only from the group of climatic factors (treated as input data), taking into account the S6 scenario, the best prediction was shown by the CNN model ($R^2 = 0.8275$), and the weakest SVR ($R^2 = 0.3$).

One of the most important criteria for determining irrigation needs is soil type, and differences are seen when the amount of water required for different crops on different soil types is compared. The models take into account factors including the type of crop itself (Tolk et al. 1999) or other crop area factors that change over time and affect the water requirements of plants (Allen 2000). According to Krupakar et al. (2016) the RNN model is robust enough to map the differences in observations that occur due to various factors such as global warming, faulty hardware, etc. In the presented study, if only data on crops and soils were considered as predictors: root depth, basal crop coefficient, soil evaporation, fraction of surface wetted and exposed and soil wetted fraction, in the absence of climate
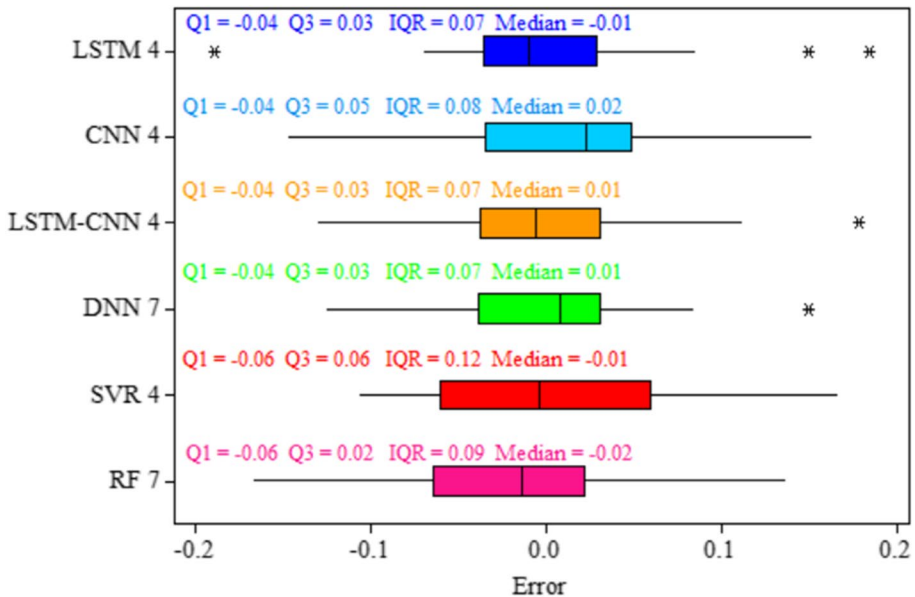
Fig. 7 Boxplots showing the distribution of the estimation errors of the IWR in the test section for the best model-scenarios. Q1 and Q3 are lower and upper quartile of errors, and IQR is interquartile range for each model

data, as presented in the S7 scenario, the best results were obtained for a hybrid model (LSTM & CNN) and separately for LSTM and CNN.

The box plot was developed to show the error distribution in the test section (Fig. 7). For IWR prediction, the best models turned out to be LSTM, CNN, LSTM & CNN and SVR for the S4, which takes into account the RH, WS, Kcb and Ke parameters, and the DNN and RF models for the S7 scenario, which includes root depth, Kcb, Ke, Fw and Few. In the group with the S4, the LSTM and LSTM & CNN models performed best because they have the lowest values of Q1, Q3 and IQR (Fig. 7), and the SVR model the worst. However, for the S7 scenario, the DNN model turned out to be the best for IWR prediction. In the RF 7 model, the lowest Q3 was obtained among the others, but the IQR was not among the lowest and amounted to 0.09. The LSTM 4 and LSTM & CNN 4 have the same IQR as the DNN 7 model and slightly lower compared to the RF 7 model (Fig. 7).

In all the models used, the error distribution is around zero, while the median line shows the normality of the error distribution, primarily for the LSTM & CNN hybrid model S4. Additionally, it should be noted that the models developed for the S4 and S7 scenarios obtained the best results in the assessment at the previous stage (Table 2). These scenarios are intermediate between the use of all input variables (S8) and the smallest number of input data represented by the S5 scenario (RH and WS data).

However, none of the models in the S6 scenario, which takes into account only climatic data, was included in the group of the best models distinguished on the basis of distribution of the estimation errors of the IWR in the test section for the best model-scenarios (Fig. 7). In turn, research by Krupakar et al (2016) showed that ANNs models perform well in prediction based on training but have a lower ability to infer semantic meaning from the sequential flow. SVMs models are designed to be tested on normal data, hence they cannot capture the essential sequential information. According to Xie and Tang (2010), the factors influencing

the amount of irrigation are complex and non-linear, and the support vector machine (SVM) has many advantages for non-linear small samples. A forecasting model of irrigation water requirement based on least squares support vector machine (LS-SVM) was used in Tarim River Basin. These authors demonstrated in their research that the prediction model based on the LS-SVM has excellent generalizability and low error, which provides an efficient method of forecasting irrigation water requirements (Xie and Tang 2010).

In turn, El Bilali and Taleb (2020), using correlation coefficients, showed the inaccuracy of the SVR model in predicting the quality of irrigation water for training and validation processes sequentially. Artificial neural network (Springmann et al. 2018), multiple linear regression (MLR), decision tree, random forest (RF), stochastic gradient descent (SGD) and adaptive boosting (AdaBoost) demonstrated good prediction accuracy. Krupakar et al. (2016) emphasize that the models described for an irrigation prediction system are generally developed and trained using a large amount of historical data (training data) on objects or characteristics that influence the amount of water required for irrigation. The new RNN LSTM model proposed by them can be used to obtain better results, richer in semantic correlations found in the parameters used in real-time data. When testing machine learning models in predicting irrigation water, it is emphasized that they are efficient tools for accurately predicting irrigation only using parameters that can be directly measured in a short time. There have been a lot of hybrid models that have been used for time series prediction (Ruan 1997; Kavousi-Fard 2013). Pulido-Calvo et al. (2003) developed a hybrid model of combining CNNs with fuzzy membership function, optimized by a genetic optimization algorithm to test the efficacy of irrigation demand prediction, a Pulido-Calvo and Gutierrez-Estrada (2009) presented improved irrigation water demand forecasting using a soft computing hybrid model.

For the prediction of irrigation for green bean crops in Egypt, the LSTM & CNN hybrid model was proposed, which enhances the information potential for IWR prediction. Deep learning concerns multilayer neural networks, which simultaneously act as a generator of diagnostic features for the analyzed process and the final function of a classifier or regression system. It also makes it possible to improve the accuracy of the system operation. The recurrent network called LSTM (Long Short-Term Memory) is a recursive network characterized by a long memory of short patterns (Schmidhuber 2015; Greff et al. 2017). The levels of the LSTM and CNN models are often combined when forecasting time series. This allows the LSTM layer to account for sequential dependencies across time series, while the CNN layer additionally informs about this process through the use of extended convolutions. Unlike LSTM, the CNN model does not keep memory of previous time series patterns, and instead can only train have based on data entered over a specific period of time. Krupakar et al. (2016) propose a novel methodology of using a sequence learning based recurrent neural network (RNN) model that uses the LSTM activation function to model for irrigation requirement so that it doesn't have memory problems on long input streams.

The IWR prediction results for Egyptian green bean crops, performed for different scenario variants, show the models' ability to predict irrigation potential. The proposed predictive irrigation planning approach, compared to field studies, makes it possible to reduce or prevent water stress in crops. Predicting future IWR values from the results of predictive models can improve the real-time irrigation planning system depending on the meteorological situation and terrain factors (plants, soil). Sustainable irrigation aims to adjust water availability and demand in terms of quantity and quality, space and time, at reasonable cost and with acceptable environmental impacts, whereby forecasting water demand is proving to be a valuable management tool.

5 Conclusions

Water and food need crisis is facing all over world, furthermore agriculture consume the largest amount of water to cover plant requirement of water. Therefore, the current research applies machine models in order to predict the irrigation water requirement for one of the most important crops in Egypt (green bean) by actual field experiment, which there weren't common literature about modeling of irrigation requirement forecasting in versus ETo evaluation studies.

The main conclusions were reported as follow; the scenario S4 (relative humidity, wind speed, basal crop coefficient, the soil evaporation), is the best scenario for SVR, LSTM, CNN and LSTM & CNN models, while the best scenario in DNN and RF model was in scenario (S7) (root depth, basal crop coefficient, the soil evaporation, fraction of surface wetted, exposed and soil wetted fraction). The climatic factors in scenario (S6) presented the highest R^2 for the CNN model and the weakest result in the SVR model. Further, by taking into account only the two predictors relative humidity and wind speed, the performance of all models ranged from very good to satisfactory fit for all models, while it was an unsatisfactory fit for SVR.

On contrast, the hybrid LTSM & CNN model was the best model for the IWR prediction. Finally, the models applied can improve the irrigation planning with few available data. The current approach was promising and has research value for other similar regions. Therefore, the future work will focus much more in predicting the irrigation water requirements in other crops in other regions to draw a general picture of the predicting of the irrigation water requirements in order to ensure the water resources management under the impact of the climate changes.

Data Availability The researchers provided the experiment requirements of raw materials, machine learning algorithms models, and climate data through their own financial resources, while the data for previous studies and research was done through the Cairo University platform, which provides research on a regular basis.

Declarations

Ethics Approval and Consent to Participate All authors are approval in participation in this research.

Consent for Publication All authors are consent of publication of the research paper.

Conflicts of Interest The authors declare no conflicts of interest.

Financial Interests The authors declare they have no financial interests.

References

Achieng KO (2019) Modelling of soil moisture retention curve using machine learning techniques: Artificial and deep neural networks vs support vector regression models. Comput Geosci 133:104320. https://doi.org/10.1016/j.cageo.2019.104320

Allen RG (2000) Using the FAO-56 dual crop coefficient method over an irrigated region as part of an evapotranspiration intercomparison study. J Hydrol 229:27–41

Andales AA, Bauder TA, Arabi M (2014) A Mobile Irrigation Water Management System Using a Collaborative GIS and Weather Station Networks. In: Ahuja LR, Ma L, Lascano R (eds) Practical Applications of Agricultural System Models to Optimize the Use of Limited Water. Advances in Agricultural Systems Modeling, Volume 5. ASA-CSSA-SSSA, Madison, Wisconsin, pp 53–84

Andales AA, Chávez JL, Bauder TL (2011) Irrigation Scheduling: The Water Balance Approach (Doctoral dissertation, Colorado State University. Libraries)

Ashrafzadeh A, Kişi O, Aghelpour P, Biazar SM, Masouleh MA (2020) Comparative study of time series models, support vector machines, and GMDH in forecasting long-term evapotranspiration rates in Northern Iran. J Irrig Drain Eng 146:04020010. https://doi.org/10.1061/(asce)ir.1943-4774.0001471

Barzegar R, Aalami MT, Adamowski J (2020) Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. Stoch Environ Res Risk Assess 34:415–433. https://doi.org/10.1007/s00477-020-01776-2

Behar O, Khellaf A, Mohammedi K (2015) Comparison of solar radiation models and their validation under Algerian climate - The case of direct irradiance. Energy Convers Manag 98:236–251. https://doi.org/10.1016/j.enconman.2015.03.067

Bellido-Jiménez JA, Estévez J, García-Marín AP (2021) New machine learning approaches to improve reference evapotranspiration estimates using intra-daily temperature-based variables in a semi-arid region of Spain. Agric Water Manag 245. https://doi.org/10.1016/j.agwat.2020.106558

Breiman L (2001) Random forests. Random For 45:1–122. https://doi.org/10.1201/9780429469275-8

Cavallo DP, Cefola M, Pace B, Logrieco AF, Attolico G (2017) Contactless and non-destructive chlorophyll content prediction by random forest regression: A case study on fresh-cut rocket leaves. Comput Electron Agr 140:303–310

Djaman K, O'Neill M, Owen CK, Smeal D, Koudahe K, West M, Allen S, Lombard K, Irmak S (2018) Crop evapotranspiration, irrigation water requirement and water productivity of maize from meteorological data under semiarid climate. Water (Switzerland) 10. https://doi.org/10.3390/w10040405

Döll P, Siebert S (2002) Global modeling of irrigation water requirements. Water Resour Res 38(4):8-1–8-10

El Bilali A, Taleb A (2020) Prediction of irrigation water quality parameters using machine learning models in a semi-arid environment. J Saudi Soc Agric Sci 19:439–451

Fan J, Ma X, Wu L, Zhang F, Yu X, Zeng W (2019) Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. Agric Water Manag 225. https://doi.org/10.1016/j.agwat.2019.105758

Fan J, Wang X, Wu L, Zhou H, Zhang F, Yu X, Lu X, Xiang Y (2018) Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. Energy Convers Manag 164:102–111. https://doi.org/10.1016/j.enconman.2018.02.087

Fan J, Zheng J, Wu L, Zhang F (2021) Estimation of daily maize transpiration using support vector machines, extreme gradient boosting, artificial and deep neural networks models. Agric Water Manag 245:106547. https://doi.org/10.1016/j.agwat.2020.106547

Feng Y, Cui N, Gong D, Zhang Q, Zhao L (2017) Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. Agric Water Manag 193:163–173. https://doi.org/10.1016/j.agwat.2017.08.003

Ferreira LB, da Cunha FF (2020a) Multi-step ahead forecasting of daily reference evapotranspiration using deep learning. Comput Electron Agric 178:105728. https://doi.org/10.1016/j.compag.2020.105728

Ferreira LB, da Cunha FF (2020b) Multi-step ahead forecasting of daily reference evapotranspiration using deep learning. Comput Electron Agric 178:105728. https://doi.org/10.1016/j.compag.2020.105728

García Nieto PJ, García-Gonzalo E, Bové J, Arbat G, Duran-Ros M, Puig-Bargués J (2017) Modeling pressure drop produced by different filtering media in microirrigation sand filters using the hybrid ABC-MARS-based approach, MLP neural network and M5 model tree. Comput Electron Agr 139:65–74

García Nieto PJ, Sánchez Lasheras F, García-Gonzalo E, de Cos Juez FJ (2018) PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study. Sci Total Environ 621:753–761

Ghimire S, Deo RC, Raj N, Mi J (2019) Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction. Energies 12. https://doi.org/10.3390/en121 22407

Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J (2017) LSTM: A search space odyssey. IEEE Trans Neural Netw Learn Syst 28(10):2222–2232

Gueymard CA (2014) A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. Renew Sustain Energy Rev 39:1024–1034. https://doi.org/10.1016/j.rser.2014.07.117

Kalboussi N, Roux S, Boumaza K, Sinfort C, Rapaport A (2019) About modeling and control strategies for scheduling crop irrigation. IFAC-PapersOnLine 52:43–48. https://doi.org/10.1016/j.ifacol.2019.11.007

Karam F, Mouneimne A, Aichouche F, Rapaport A, Harmand J (2019a) Processing soil water content measurements to estimate crop water requirements and optimize irrigation supply. IFAC-PapersOnLine 52:49–53. https://doi.org/10.1016/j.ifacol.2019.11.008

Karam F, Mouneimne A, Aichouche F, Rapaport A, Harmand J (2019) Processing Soil Water Con- tent Measurements to Estimate Crop Water Requirements and Optimize Irrigation Supply. IFAC Workshop on Control Methods for Water Resource Systems – CMWRS, International Federation of Automatic Control (IFAC), Delft, pp 49–53. https://doi.org/10.1016/j.ifacol.2019.11.008 (hal-02408187)

Kavousi-Fard A (2013) A new fuzzy-based feature selection and hybrid TLA–ANN modelling for short-term load forecasting. J Exp Theor Artif Intell 25:543–557

Kim N, Ha KJ, Park NW, Cho J, Hong S, Lee YW (2019) A comparison between major artificial intelligence models for crop yield prediction: Case study of the midwestern United States, 2006–2015. ISPRS Int J Geo-Information 8:2006–2015. https://doi.org/10.3390/ijgi8050240

Krupakar H, Jayakumar A (2016) A review of intelligent practices for irrigation prediction. arXiv preprint arXiv:1612.02893

Kumar M, Raghuwanshi NS, Singh R (2011) Artificial neural networks approach in evapotranspiration modeling: A review. Irrig Sci 29:11–25. https://doi.org/10.1007/s00271-010-0230-8

Landeras G, Ortiz-barredo A, López JJ (2009) Forecasting weekly evapotranspiration with ARIMA and artificial neural network models. J Irrig Drain Eng 135(3):323–334

Mittelman R (2015) Time-series modeling with undecimated fully convolutional neural networks. arXiv preprint arXiv:1508.00317

Mokhtar A, Jalali M, Elbeltagi A, Al-Ansari N, Alsafadi K, Abdo H G, Sammen SS, Gyasi-Agyei Y, Rodrigo-Comino J, He H (2021b) Estimation of SPEI Meteorological Drought using Machine Learning Algorithms. IEEE Access

Mokhtar A, Jalali M, He H, Al-Ansari N, Elbeltagi A, Alsafadi K, Abdo HG, Sammen SS, Gyasi-Agyei Y, Rodrigo-Comino J (2021a) Estimation of SPEI meteorological drought using machine learning algorithms. IEEE Access 9:65503–65523. https://doi.org/10.1109/ACCESS.2021.3074305

Monem, K.M.A.S.A., 2021. Political, technical and economical contexts of the government of Egypt's approaches towards the Nile basin countries

Montes-Atenas G, Seguel F, Valencia A, Bhatti SM, Khan MS, Soto I, Becerra Yoma N (2016) Predicting bubble size and bubble rate data in water and in froth flotation-like slurry from computational fluid dynamics (CFD) by applying deep neural networks (DNN). Int Commun Heat Mass Transf 76:197–201. https://doi.org/10.1016/j.icheatmasstransfer.2016.05.031

Naghibi SA, Ahmadi K, Daneshi A (2017) Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. Water Resour Manag 31(9):2761–2775

O'Toole JC, Hatfield JL (1983) Effect of wind on the crop water stress index derived by infrared thermometry. Agron J 75:811–817

Pereira LS, Allen RG, Smith M, Raes D (2015b) Crop evapotranspiration estimation with FAO56: Past and future. Agric Water Manag 147:4–20. https://doi.org/10.1016/j.agwat.2014.07.031

Pereira LS, Paredes P, Rodrigues GC, Neves M (2015a) Modeling malt barley water use and evapotranspiration partitioning in two contrasting rainfall years. Assessing AquaCrop and SIMDualKc models. Agric Water Manage 159:239–254

Pulido-Calvo I, Roldán J, López-Luque R et al (2003) Demand forecasting for irrigation water distribution systems. J Irrig Drain Eng 129:422–431

Pulido-Calvo I, Gutierrez-Estrada JC (2009) Improved irrigation water demand forecasting using a softcomputing hybrid model. Biosyst Eng 102:202–218

Rockström J, Kaumbutho P, Mwalley J, Nzabi A, Temesgen M, Mawenya L, Barron J, Mutua J, Damgaard-Larsen S (2009) Conservation farming strategies in East and Southern Africa: Yields and rain water productivity from on-farm action research. Soil Tillage Res 103:23–32

Ruan D (ed) (1997) Intelligent hybrid systems: fuzzy logic, neural networks, and genetic algorithms (pp. 161-187). Boston: Kluwer Academic Publishers

Schlenker W, Roberts MJ (2009) Nonlinear temperature effects indicate severe damages to US crop yields under climate change. Proc Natl Acad Sci 106:15594–15598

Schmidhuber J (2015) Deep learning in neural networks: An overview. Neural Netw 2015(61):85–117

Shiklomanov IA (1997) Assessment of water resources and water availability in the world. Comprehensive Assessment of the Freshwater Resources of the World

Smith PC, Calanca P, Fuhrer J (2012) A simple scheme for modeling irrigation water requirements at the regional scale applied to an Alpine river catchment. Water (Switzerland) 4:869–886. https://doi.org/10.3390/w4040869

Son H, Kim C (2020) A deep learning approach to forecasting monthly demand for residential-sector electricity. Sustain 12:3103. https://doi.org/10.3390/SU12083103

Springmann M, Clark M, Mason-D'Croz D, Wiebe K, Bodirsky BL, Lassaletta L, De Vries W, Vermeulen SJ, Herrero M, Carlson KM (2018) Options for keeping the food system within environmental limits. Nature 562:519–525

Stone RJ (1994) A nonparametric statistical procedure for ranking the overall performance of solar radiation models at multiple locations. Energy 19:765–769. https://doi.org/10.1016/0360-5442(94)90014-0

Tian C, Ma J, Zhang C, Zhan P (2018) A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. Energies 11. https://doi.org/10.3390/en11123493

Tolk JA, Howell TA, Evett SR (1999) Effect of mulch, irrigation, and soil type on water use and yield of maize. Soil Tillage Res 50:137–147

Trajkovic S, Todorovic B, Stankovic M (2003) Forecasting of reference evapotranspiration by artificial neural networks. J Irrig Drain Eng 129:454–457. https://doi.org/10.1061/(asce)0733-9437(2003)129:6(454)

United Nations Environment Programme (UNEP) (1999) Global Environmental Outlook 2000. Earthscan, London

Xie F, Tang D-S (2010) Forecasting Model of Irrigation Water Requirement Based on Least Squares Support Vector Machine," 2010 International Conference on Intelligent Computation Technology and Automation, pp 335–338. https://doi.org/10.1109/ICICTA.2010.91

Zhang Q-tao, Xia Q, Liu CCK, Geng S (2013) Technologies for efficient use of irrigation water and energy in China. J Integr Agric 12:1363–1370. https://doi.org/10.1016/S2095-3119(13)60544-4

Zhou Y, Chang FJ, Chang LC, Kao IF, Wang YS (2019) Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. J Clean Prod 209:134–145. https://doi.org/10.1016/j.jclepro.2018.10.243

Zuo W, Zhou F, Li Z, Wang L (2019) Multi-resolution CNN and knowledge transfer for candidate classification in lung nodule detection. Ieee Access 7:32510–32521

Authors and Affiliations

Ali Mokhtar[1] · Nadhir Al-Ansari[2] · Wessam El-Ssawy[1,3] · Renata Graf[4] · Pouya Aghelpour[5] · Hongming He[6] · Salma M. Hafez[1] · Mohamed Abuarab[1] ⃝

✉ Nadhir Al-Ansari
nadhir.alansari@ltu.se

✉ Mohamed Abuarab
mohamed.aboarab@agr.cu.edu.eg

[1] Department of Agricultural Engineering, Faculty of Agriculture, Cairo University, Giza 12613, Egypt

[2] Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, 97187 Lulea, Sweden

[3] Irrigation and Drainage Department, Agricultural Engineering Research Institute, Giza 12613,

Egypt

[4] Department of Hydrology and Water Management, Institute of Physical Geography and Environmental Planning, Adam Mickiewicz University, Poznań, Poland

[5] Water Engineering Department, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

[6] School of Geographic Sciences, East China Normal University, Shanghai 210062, China