electronics

MDPI

Article

# High-Accuracy Estimation of Reference Evapotranspiration Using Classical and AI-Supported System Identification Approaches Under Different Climatic Conditions in Arid Zones

**Wafa Difallah [1,2,*], Boudjema Ouradj [2], Fateh Bounaama [2], Belkacem Draoui [2], Khelifa Benahmed [1] and Khelifa Lammari [2]**

[1] Innovation in Informatics and Engineering Laboratory (INIE LAB), Department of Mathematics and Computer Science, Faculty of Exact Sciences, Tahri Mohammed University of Bechar, Bechar 08000, Algeria
[2] Laboratory of Energetic in Arid Zones (ENERGARID), Faculty of Technology, Tahri Mohammed University of Bechar, Bechar 08000, Algeria
* Correspondence: difallah.wafa@univ-bechar.dz

**Abstract**

Reference evapotranspiration (ET0) is a critical parameter for water resource management and irrigation scheduling. Accurate estimation of ET0 has challenged scientists over the years due to its high sensitivity to climatic variations. Classical methods for estimating ET0 mainly rely on empirical models with a significant number of parameters, which has hampered their use in many cases. Regarding its importance and strong relationship with global food security, this topic has attracted the attention of many researchers. The development of simple models with a low number of parameters or taking advantage of artificial intelligence algorithms has been the aim of different researchers, as it is in this paper, where two approaches are implemented to estimate reference evapotranspiration. The first one is based on the use of classical system identification models, namely linear and nonlinear AutoRegressive models with eXogenous variables (ARX and nonlinear ARX). For the second approach, AI-supported system identification models are used, in which neural networks' performances are used to develop multilayer and deep neural network models for nonlinear system identification. The four models show a high accuracy, with a system fitting exceeded 98%.

**Keywords:** evapotranspiration; system identification; machine learning; deep learning; water

Check for updates

Academic Editor: Domenico Ursino

Received: 20 November 2025
Revised: 14 December 2025
Accepted: 17 December 2025
Published: 13 February 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

The water crisis has a great impact on agricultural activities, specifically irrigation [1–3]. In fact, water shortage is not the only problem threatening agriculture. The accurate estimation of reference evapotranspiration (ET0) is another critical issue that has interested scientists over the years [4–7]. ET0 represents the amount of water lost due to soil water evaporation (E) and plant transpiration to the atmosphere [8,9]. Basically, physical models are used to calculate ET0. These models rely on a high number of climatic parameters, which may not be within reach in many cases. To deal with this challenging issue, researchers tend to adopt alternative methods that rely on simplified physical models or artificial intelligence algorithms.

The researchers in [10] utilized climatic data from 1979 to 2022 to develop multiple linear regression models for ET0 prediction in northwestern Bangladesh. The study proved that ET0 is highly impacted by temperature and solar radiation. The study in [11]

highlighted the importance of relative humidity, temperature, and wind speed in ET0 estimation by applying a traditional system identification technique. The effect of climate was also highlighted in [12] by analyzing the influence of data collected between 1993 and 2022 on calculating ET0 via CROPWAT software (FAO, http://www.fao.org/land-water/databases-and-software/cropwat/en/, accessed on 2 November 2025).

The authors in [13] proved that relative humidity is a key parameter in predicting ET0 in the case of data scarcity. This scarcity was also treated by the authors of [14]. A temperature-based model was proposed to replace the Penman–Monteith model in moist sub-humid and humid climates. In the same context, the implementation of different data requirement models was carried out in [15]. FAO-56 PM, the original and three modified Hargreaves–Samani models (HS and MHS1, MHS2 and MHS3), Trajkovic (TR), and the single-temperature procedure (MaxTET) were compared under different environmental conditions. The results indicated that the precision of each method depends on geographical location. In [16], Priestley–Taylor and 1985 Hargreaves ET0 estimation models were compared with the FAO-56 Penman–Monteith model (FAO-56 PM). The Priestley–Taylor model gives a better fit than the Hargreaves model, which is more suitable as a temperature-based model. The two models were also implemented in [17], and the comparison findings demonstrated their effectiveness over other models. In [18], temperature, wind speed, humidity, and sunshine duration acquired over 30 years were used to develop an adjusted version of the Hargreaves–Samani (HS) equation (HS1). This model demonstrates increased accuracy compared with the HS and FAO-56 PM models. In [19], ET0 was estimated using three evapotranspiration methods: EVAPO, AgSAT, and Evapotranspiration Calculator. In this study, the Thornthwaite model was used as a reference to evaluate the three approaches. In [20], a modified version of the Penman–Monteith equation was evaluated to predict actual evapotranspiration and crop coefficients for sesame crop planting. The study also examined the effectiveness of the FAO Radiation, Priestly–Taylor, Hargreaves–Samani, and Pan Evaporation methods in determining ET0 values compared to the original Penman–Monteith equation. The results showed that the modified version is the most appropriate method for estimating ET0. The estimation of crop evapotranspiration (ETc) was the aim of the study in [21]. The estimation accuracy of 30 ET0 micrometeorological models was evaluated, and the Penman–Monteith model maintained its position as the best model for ET0 estimation. In [22], three Penman–Monteith models were implemented in arid and semi-arid zones. The study highlights the importance of data frequency in ET0 estimation.

Artificial intelligence-based models have also occupied a large space in ET0 estimation. In [23], several machine learning algorithms, including XGBoost, GBM, RF, BT, and Custom Deep Learning, were evaluated, with temperature, wind speed, humidity, and precipitation as inputs and potential evapotranspiration as the output. The XGBoost model showed the best performance based on RMSE and MAE values. In [24], the Penman–Monteith (PM), Hargreaves (HA), and Blaney–Criddle (BC) algorithms were used to evaluate a set of ML algorithms. The authors demonstrated the efficiency of the implemented model compared with standard methods. Artificial Neural Networks (ANN) were demonstrated to be the best method for estimating ET0 compared to multiple linear regression (MLR) and Random Forest (RF) in [25]. A nonlinear relationship between different environmental parameters and ET0 was identified by an Artificial Neural Network (ANN) in [26]. According to the results in [27], ANNs can also demonstrate high reliability in forecasting ET0 values. To improve the system prediction efficiency of ANN algorithms, model hybridization has prevailed as a key solution. In [28], an ANN was coupled with three techniques for data preprocessing to produce five new prediction models.

The impact of input shortages on ET0 system identification performance was addressed in [29]. Different input combinations were tested and compared with measured lysimeter ET. The findings validated the significant influence of temperature on ET0 calculation and the high performance of the ANN in the system identification process with only one climatic input. Using a 35-year dataset of temperature, wind speed, relative humidity, and net radiation, a fuzzy logic-based model was developed to estimate ET0 in the Ogun–Osun River Basin. In [30], a genetic algorithm was used to optimize the neural network's performance in predicting ET0 based on the Penman–Monteith model at 14 weather stations. A genetic algorithm was also used by the authors in [31] with particle swarm optimization and differential evolution algorithms to estimate ET0. The study also analyzes the impact of climatic parameters through the implementation of eleven ET0 prediction models. The same idea was applied in [32], where gradient boosting (GB), Support Vector Machine (SVM), Random Forest (RF), and Ridge Regression (Ridge) were compared with traditional ET0 estimation methods. The data collected between 1960 and 2023 from 31 locations were used to develop the machine learning algorithms. In [33], Random Forest Regressor (RFR), Decision Tree Regressor (DTR), and Support Vector Regressor (SVR) were implemented to estimate local evapotranspiration. $R^2$, RMSE, MAE, and MAPE values showed RFR to be the best ET0 prediction model. The study also underlined the effect of temperature and wind speed on the ML-based estimation process. This result was also confirmed in [34]. The authors of [35] proved that Support Vector Regression (SVR) performs better than Random Forest (RF) in estimating ET0 with rainfall, temperature, wind speed, and relative humidity data.

In [36], the wavelet neural network (WNN) and the Group Method of Data Handling were used to simplify the ET0 estimation process. The findings of Long Short-Term Memory (LSTM) network implementation in [37] demonstrate that good ET0 estimation can be achieved with only a few parameters. The study also evaluates the efficiency of the Gradient Boosting Regressor (GBR), Support Vector Regression (SVR), and Random Forest (RF) models compared with the FAO 56 Penman–Monteith equation. The study in [38] discussed the accuracy of three ET0 forecasting approaches across the FAO Penman–Monteith (FAO-PM) and Hargreaves–Samani (HS) models. The models tested included statistical methods, machine learning, and deep learning techniques. The results indicated that the machine learning and deep learning models were in good agreement with FAO-PM, while the statistical models performed well for HS. In [39], Random Forests (RFs), light gradient boosting (LGB), and extreme gradient boosting (XGB) algorithms were improved by the gray wolf optimizer (GWO) for precise ET0 estimation. The study declared solar radiation as pivotal for ET0 estimation. A Hybrid Convolutional Neural Network with Long Short-Term Memory (H-ConvLSTM) was proposed by researchers in [40]. The system demonstrated high performance compared to other deep learning algorithms, including Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), and Long Short-Term Memory (LSTM), with temperature and relative humidity as system inputs. In [41], the authors utilized ERA5-Land reanalysis data combined with a Random Forest machine learning model to predict daily and 8-day ET0. The data collected from 50 weather stations (2008–2017) were used, and the results indicated that temperature had a significant role in ET0 estimation and water optimization. In [42], five deep learning algorithms were used: CatBoost, LightGBM, LSTM, GRU, and Bi-LSTM. By implementing these models on a Moringa oleifera irrigation dataset in semi-arid regions, CatBoost outcomes prevailed over the other models.

According to the presented studies, researchers have sought to enhance ET0 estimation using different methods, ranging from traditional models to AI-supported models. Control engineering is a crucial aspect in agriculture; with system identification, optimization,

and control, the outcomes of this vital sector can be highly improved. Additionally, swift advancements in artificial intelligence technologies are driving significant transformations in control engineering, contributing to the development of promising system identification methods [43,44]. The incorporation of artificial intelligence (AI) into system identification has profoundly influenced the discipline by improving the capability to model intricate, nonlinear, and dynamic systems [45–48]. Most presented studies used either classical models or artificial intelligence models to estimate ET0 values. In this paper, both approaches are used, in which four ET0 prediction models are presented to address two challenging global crises—water scarcity and food security—especially in arid regions, which experience harsh climatic conditions. Two models are based on the classical system identification process, while the third and fourth models are based on a deep learning approach. These models align well with data shortage and physical consistency requirements for ET0 prediction in arid regions, even with a moderate dataset. They provide an explicit representation of the relationship between the climatic inputs (temperature, relative humidity, wind speed, insolation duration) and the output (ET0). After executing the proposed models, a systematic comparison is conducted to determine the most efficient model for forecasting ET0 values. Regarding the pivotal position of evapotranspiration in the hydrological balance [49–51], these models can greatly optimize irrigation water use by determining the precise crop water requirements, which strongly depends on ET0 [52–54].

## 2. Materials and Methods

### 2.1. Geography and Climate of the Study Area

The testing facility is situated in Bechar city in the southwestern region of Algeria (see Figure 1) at an approximate latitude of 31°37′ N and a longitude of 2°14′ W. This location presents a distinctive opportunity for agricultural research, owing to its unique geographic and climatic features.

**Figure 1.** Location map of the experimental site.

Bechar is characterized by an arid climate that is significantly shaped by the surrounding desert landscape [55–57]. The minimum, mean, and maximum measurements of air temperature, wind speed, relative humidity, and global horizontal irradiance during two different periods are illustrated in Figure 2 below.

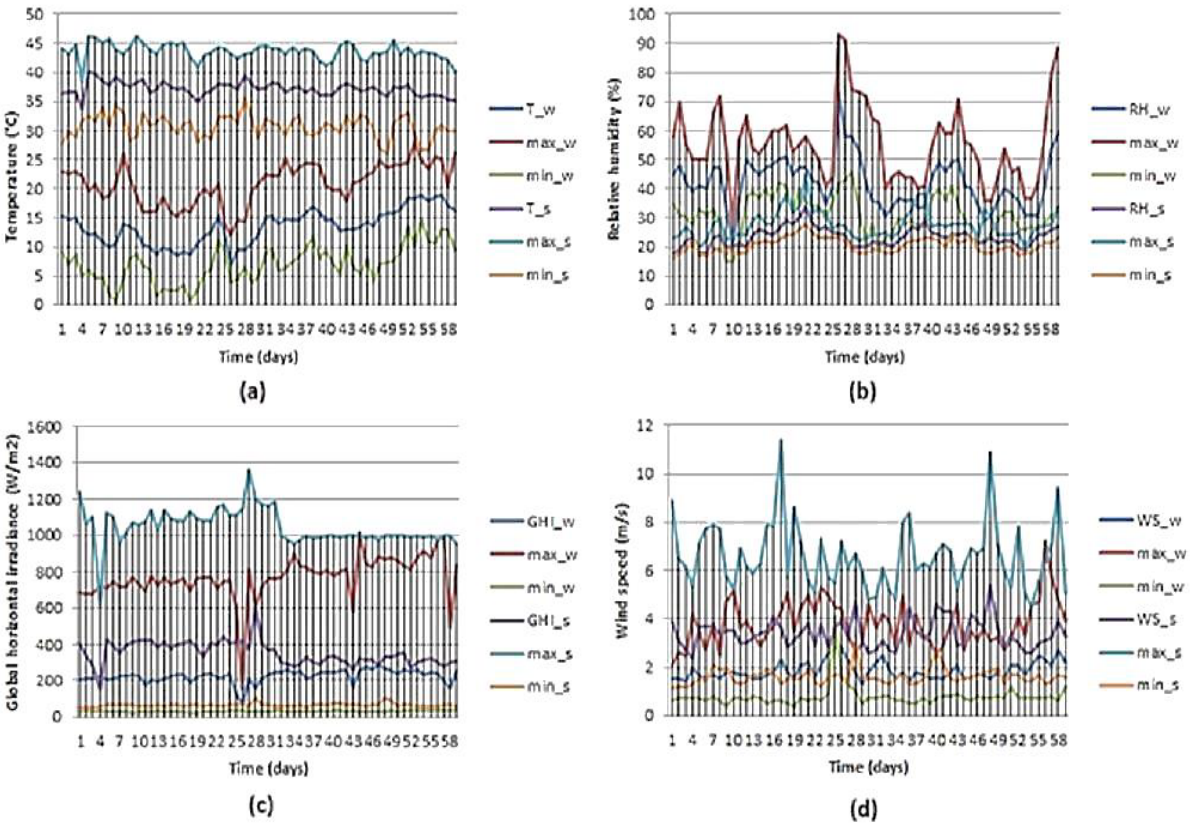


**Figure 2.** Climatic variations in the study area during summer (s) and winter (w): (**a**) temperature, (**b**) relative humidity, (**c**) global horizontal irradiation (**d**) wind speed.

### *2.2. System Design*

The correct determination of evapotranspiration values leads to an optimal irrigation water management process [58–61]. The diagram in Figure 3 illustrates the main steps for developing different reference evapotranspiration models based on classical system identification and AI-supported system identification.

At first, a database is created, composed of temperature (T), relative humidity (RH), wind speed (WS), and insolation duration (I) as inputs and reference evapotranspiration (ET0ref) as the output. ET0ref is calculated with the following multiple linear regression (MLR) formula:

$$\mathrm{ET0} = 0.17\mathrm{T} - 0.06\mathrm{RH} + 1.31\mathrm{WS} + 0.26\mathrm{I} - 0.05 \quad (1)$$

This formula was used in [54], with a dataset consisting of the daily values of the same climatic parameters used in this study. These data were collected over 1272 days by the meteorological station at the experimental site located in an arid region in the south west of Algeria, and which is similar to the study area. This model showed a high correlation with the Penman–Monteith reference model and a coefficient of determination $R^2$ of 0.97.

These data were taken during two different seasons: summer and winter. During these periods, the plants' water needs reach their highest (in summer) and lowest (in winter) levels. This data diversity helps develop a typical ET0 model to be implemented under different climatic conditions. The dataset in this study comprises 2903 hourly samples.

The next step consists of database separation. The following partitioning is adopted:

- For the classical system identification approach: 70% estimation and 30% validation.
- For the AI-supported system identification approach: 70% training, 15% validation, and 15% test.

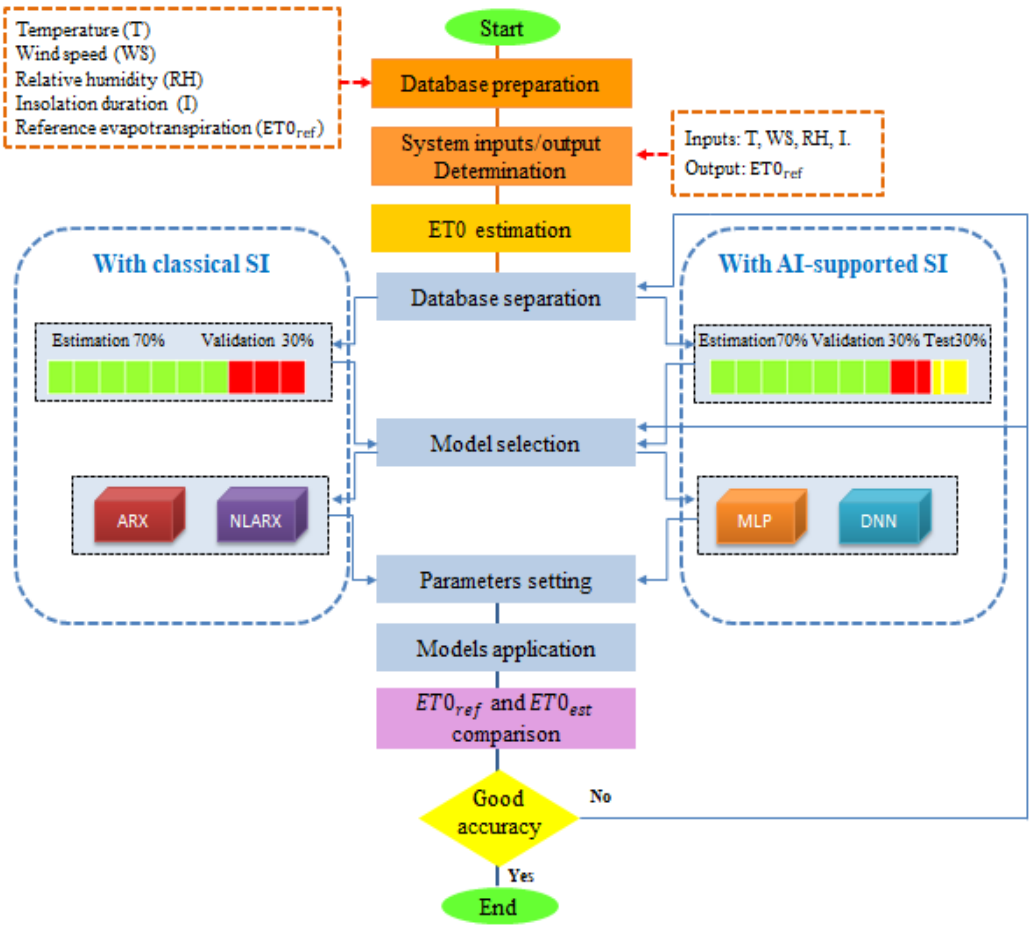


**Figure 3.** Design of the proposed methodology.

The large amount of data is used to provide the models with enough data about the physical system, and the rest of the data is reserved to monitor model performance, avoid overfitting, and evaluate the model's final predictive performance on unseen data.

After dataset splitting, two models are chosen for each approach: linear and nonlinear ARX for classical system identification and Multilayer Feedforward Neural Networks (MLP) and deep neural networks (DNNs) for artificial intelligence implementation.

ARX (AutoRegressive model with eXogenous variables) is the most widely utilized dynamic model in linear system identification [62]. By using the ARX model, the relationship between the inputs and output of the system can be expressed by a simple B/A transfer function (see Figure 4).

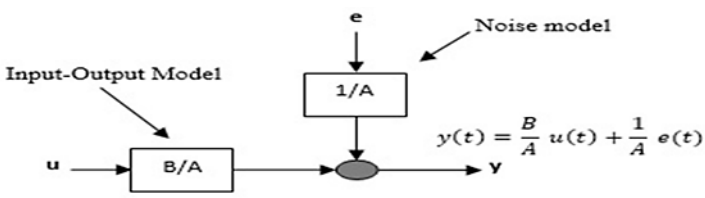


**Figure 4.** ARX model structure.

The nonlinear version of ARX is also used to predict ET0 values as a traditional system identification algorithm. The system is implemented with a multiple inputs/single output architecture (MISO), and it amplifies the regularization of na, nb, and nk parameters, which represent, respectively, the number of previous output terms employed in predicting the current output, the number of previous input terms employed in predicting the current

output, and the time lag from input to output, measured in sample units [58,63]. The general structure of NLARX is presented in Figure 5.

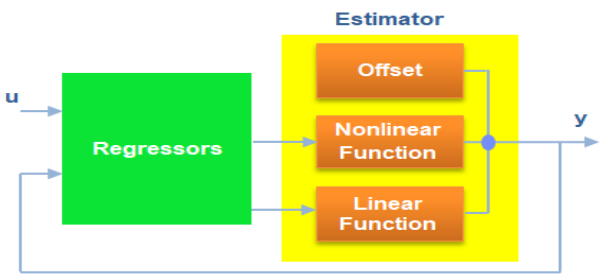


**Figure 5.** NLARX model structure.

MLP (multilayer perceptron) is a special kind of supervised neural network learning algorithm. It is the most used algorithm, featuring a simple implementation process [64–66]. MLP consists of several layers, including an input layer with n variables, hidden layers, and an output layer with m variables. Each layer is made up of a set of neurons. The number of parameters is given by the following relation:

$$\text{Totalparameters} = n.h_1 + \sum_{k=1}^{N_h-1} h_k.h_{k+1} + h_{N_h}.n \quad (2)$$

In which, $N_h$ is the number of hidden nodes in the $i^{th}$ layer [67].

The second AI algorithm is DNN (deep neural network), which is a deeper version of the MLP algorithm with the same number of each type of layer [68]. The architectures of the proposed networks are illustrated in Figure 6 bellow.

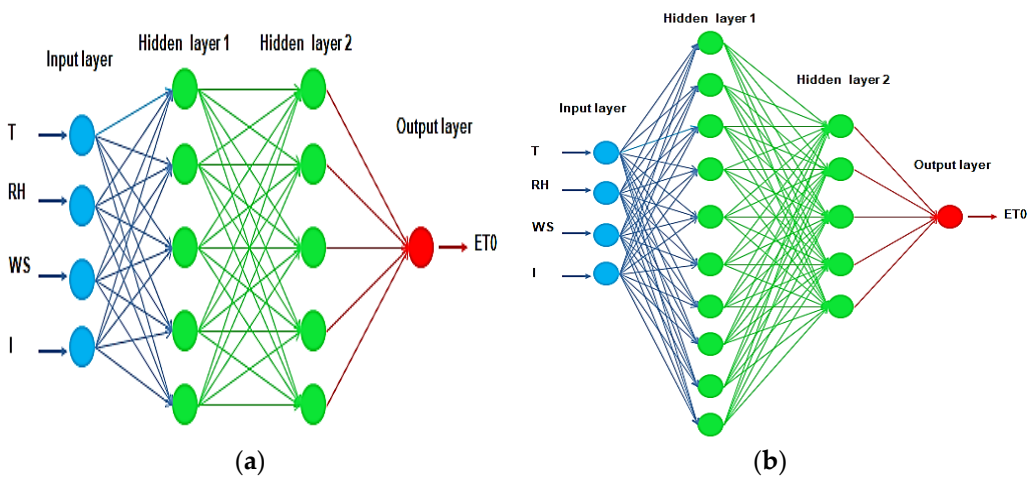


**Figure 6.** Neural network architecture: (**a**) MLP and (**b**) DNN.

To ensure a fair comparison, the same values of na (numerator), nb (denominator), and nk (disturbance) are utilized to estimate the four models. These values are, respectively, 1, [4 1 1 1] and [0 1 0 0].

The rest of the process will be discussed in the next sections.

## 3. Results and Discussion

This section is dedicated to presenting and discussing the findings of the systems' application.

### *3.1. System Identification Models*

#### 3.1.1. Linear ARX Model

The first ET0 model estimation produces the following linear model:

$$m = A(q)y(t) = B(q)u(t) + e(t) \tag{3}$$

$$A(q) = 1 + 0.7182\, q^{-1} \tag{4}$$

$$B1(q) = 1.675 + 0.5721\, q^{-1} - 0.01139\, q^{-2} + 0.007769\, q^{-3} \tag{5}$$

$$B2(q) = 0.2928\, q^{-1} \tag{6}$$

$$B3(q) = -0.1035 \tag{7}$$

$$B4(q) = 0.4391 \tag{8}$$

Upon execution, the identified system demonstrates its precision, achieving a mean squared error (MSE) of 0.0002084 and a final prediction error (FPE) of 0.0002107. The results from the identified model are compared with the calculated values. A strong correlation is observed between the two models for both the estimation and validation datasets (see Figures 7 and 8).

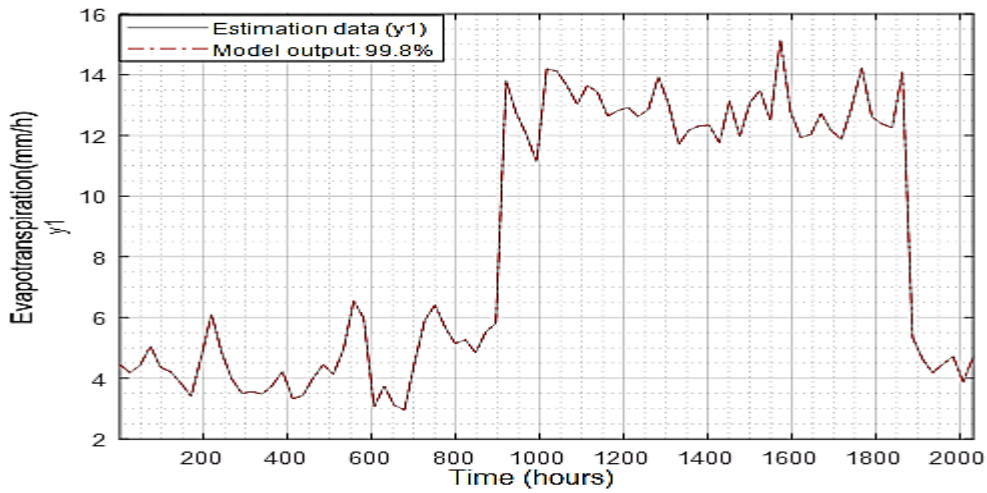


**Figure 7.** Comparison between computed and identified ET0 values using estimation data (1).

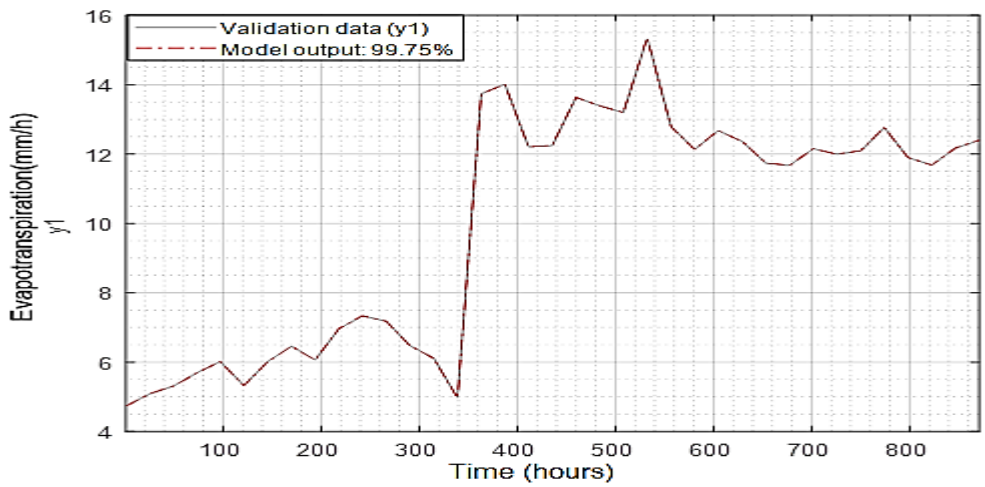


**Figure 8.** Comparison between computed and identified ET0 values using validation data (1).

To support these findings, the system's pole-zero map is extracted (see Figure 9). There are two real poles (approximately at −0.4 and −0.2), both negative and inside the unit circle, which indicates system stability. It is clear that there are no imaginary (complex) parts, indicating the absence of system oscillations. This will guarantee a smooth response to inputs.

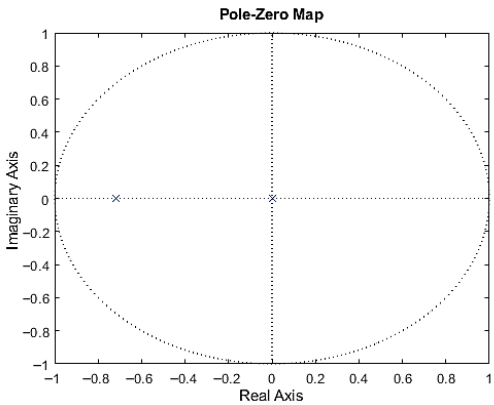


**Figure 9.** Pole-zero map of the obtained system (×: pole).

### 3.1.2. Nonlinear ARX Model

According to the results in Figures 10 and 11, it can be concluded that the nonlinear model could be a promising alternative to the reference ET0 model, given that FPE = 0.0001136 and MSE = 0.0001099.

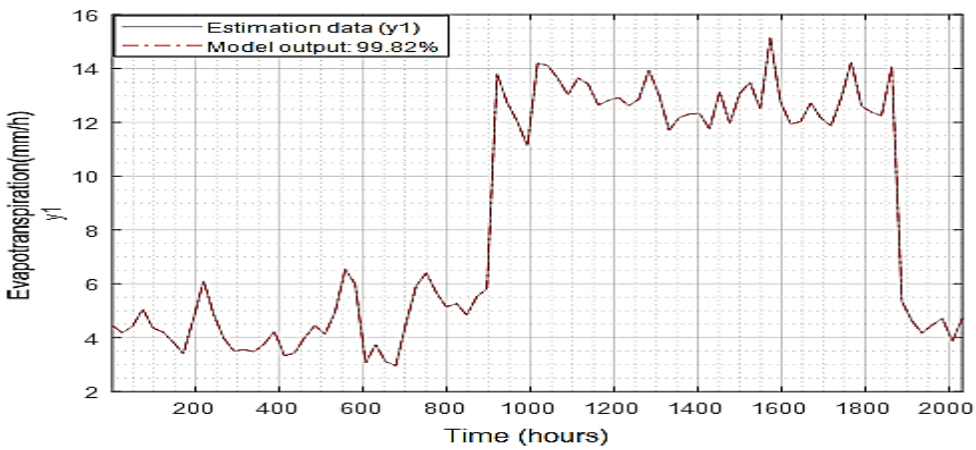


**Figure 10.** Comparison between computed and identified ET0 values using estimation data (2).

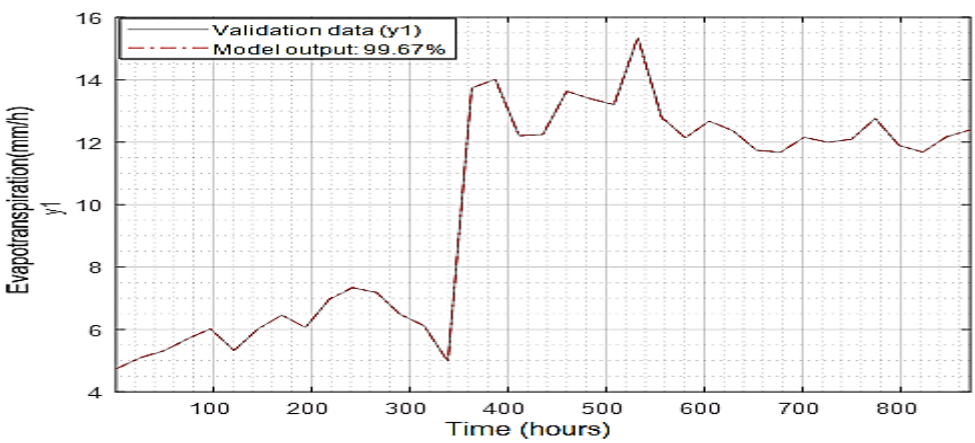


**Figure 11.** Comparison between computed and identified ET0 values using validation data (2).

### 3.2. AI-Supported System Identification Models

In this section, neural network performances are utilized to develop multilayer and deep neural network models for nonlinear system identification. As illustrated in Figures 12 and 13, the developed AI models demonstrate high accuracy, reaching 99.73% and 99.83% fits on training data for the multilayer perceptron (MLP) and deep neural networks (DNNs), respectively.

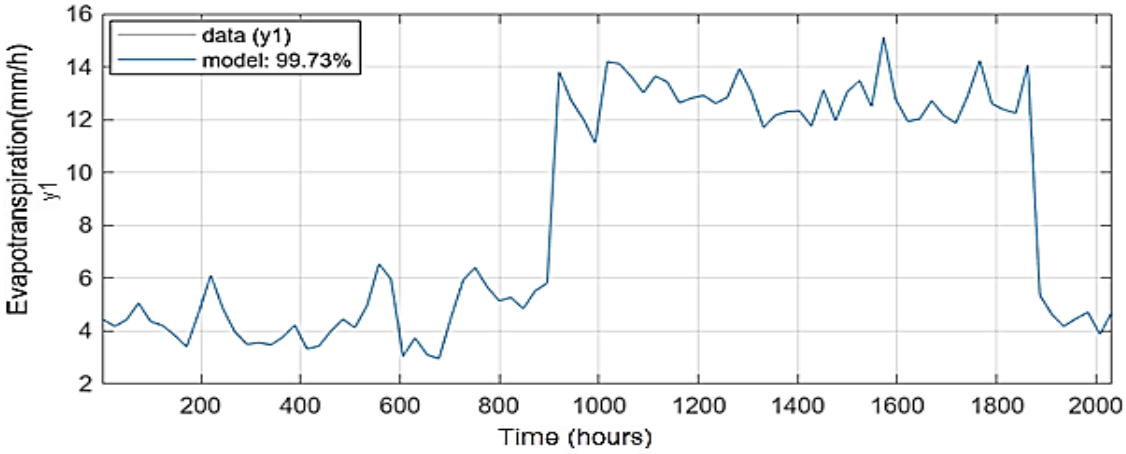


**Figure 12.** Comparison between computed and MLP-identified ET0 values.

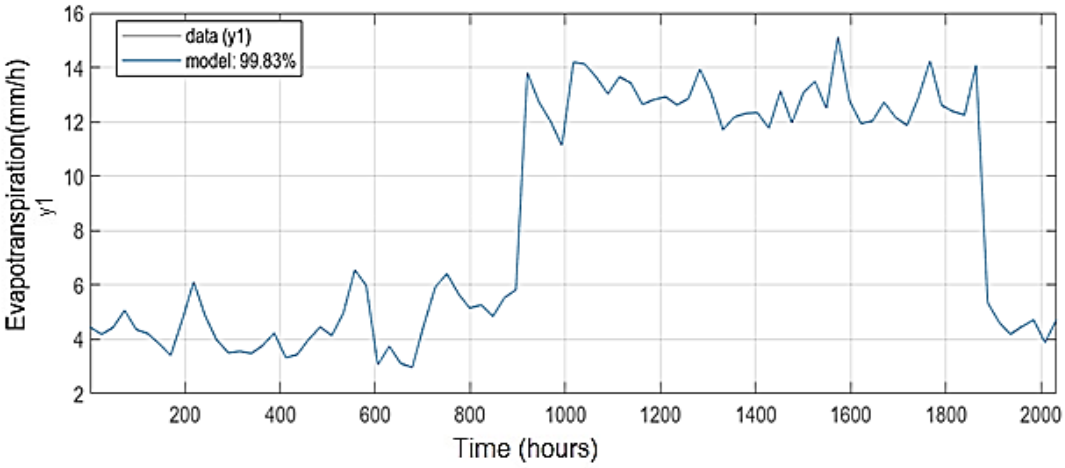


**Figure 13.** Comparison between computed and DNN-identified ET0 values.

For network development, the database is separated into three subsets. The training set, which is the largest portion, provides the system with the necessary data to learn. A comparison between the training and validation datasets is conducted to observe the system performance.

According to the obtained results in Figures 14 and 15, it is obvious that the developed models are in good agreement with the measured data for both the MLP and DNN algorithms.

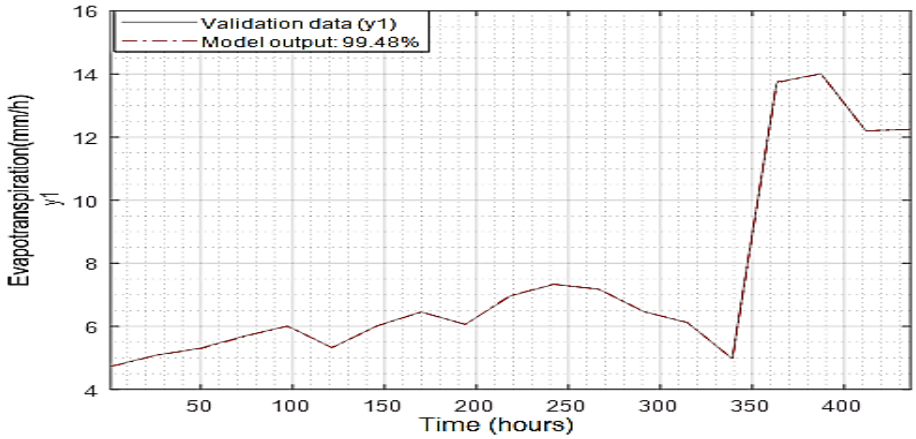


**Figure 14.** Comparison between MLP-identified ET0 values and validation data.

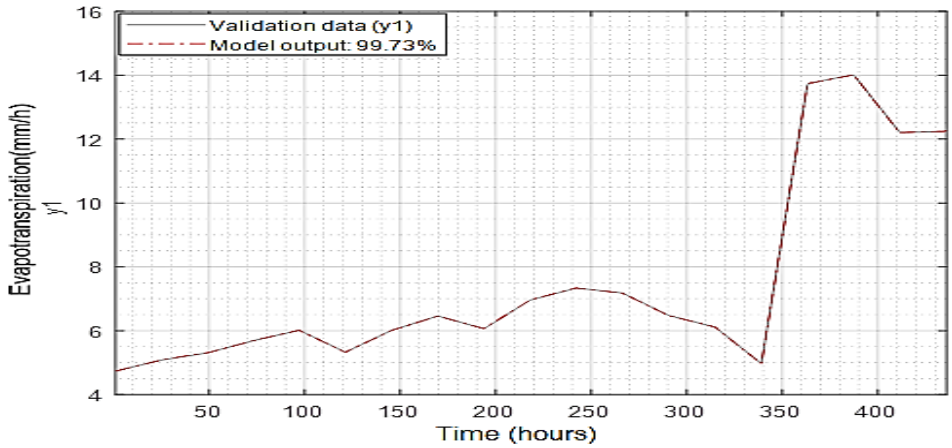


**Figure 15.** Comparison between DNN-identified ET0 values and validation data.

The final portion of data is used to evaluate the final predictive performance of the implemented models on previously unencountered data. The results in Figures 16 and 17 confirm stable performance with an average fit exceeding 98%, which enhances the validation of the applied models.

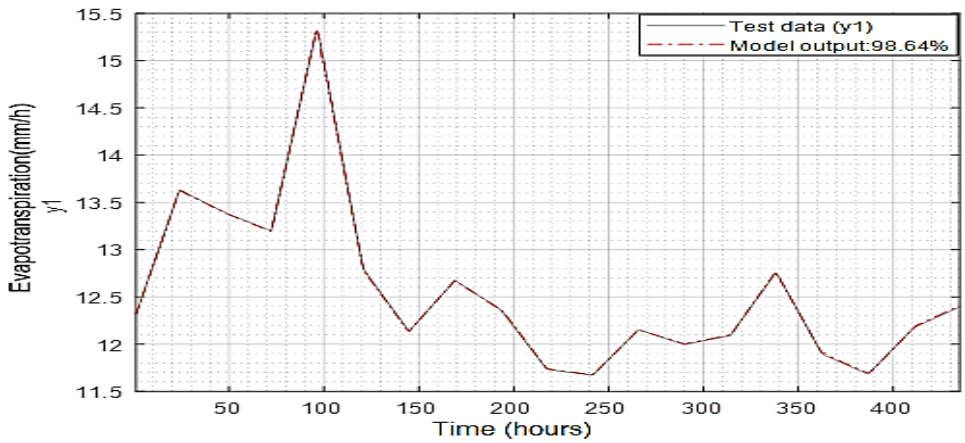


**Figure 16.** Comparison between MLP-identified ET0 values and test data.

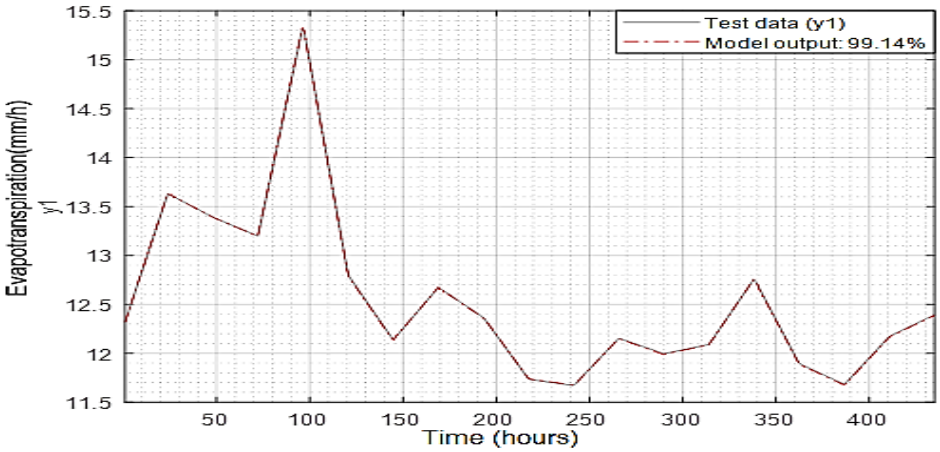


**Figure 17.** Comparison between DNN-identified ET0 values and test data.

This database split ensures precise evaluation of the obtained models with overfitting prevention. In addition, parameter tunning and MaxIterations regularization are implemented to control the system fitting. Given the dataset size and the network's architecture, sigmoid and tanh activation functions are used to model the smooth, bounded nonlinear properties of the studied system. These functions demonstrated high stability for handling physical systems with continuous relationships that do not exhibit sharp sparsity behaviors. Model parameters are optimized using the Sequential Quadratic Programming (SQP) algorithm of the fmincon function for nonlinear constrained minimization. The parameters settings and the main obtained results for both the multilayer and deep neural networks are summarized in Table 1 below.

**Table 1.** Parameters and results of MLP and DNN implementation.

| | | MLP | DNN |
|---|---|---|---|
| Network architecture | | 2 layers, with 10 and 5 neurons, respectively | 2 layers, each with 5 neurons |
| Activation function | | sigmoid, tanh | sigmoid, tanh |
| Optimization function | | SQP(fmincon) | SQP (fmincon) |
| Number of iterations | | 10 | 10 |
| Number of function evaluations | | 37 | 29 |
| System fitting (%) | Training | 99.73 | 99.83 |
| | Validation | 99.48 | 99.73 |
| | Test | 98.64 | 99.14 |
| FPE | | $7.43217 \times 10^{-5}$ | 0.000157467 |

### *3.3. Comparative Analysis of the ET0 Estimation Models*

Considering the results obtained after applying the proposed models, it can be said that it is possible to rely heavily on these models, as the compatibility rate exceeded 98% for the four models, whether classical or AI-supported models.

There are noticeable rises and falls in the ET0 values, which are sometimes sharp, as is the case in Figures 7, 10, 12 and 13, where the ET0 value decreased from more than 14 mm to less than 5 mm by the 1850th hour and increased from 5 mm to more than 13 mm by the 340th hour, as shown in Figures 14 and 15. This is due to the diversity of the database between values taken in summer and in winter; this confirms that the selected systems are capable of adapting to different climatic conditions and sudden changes. As for AI-

supported models, despite the reliance on a moderate network architecture for parameter regularization, the max number of equations limitation, and the database segmentation, which had a significant impact on the assessment of system feasibility across different database categories, accurate and consistent results were obtained. To further analyze and compare the system's performance, MSE (mean squared error), RMSE (root mean squared error), and $R^2$ (coefficient of determination) were calculated using the following equations:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(X_i - Y_i)^2 \tag{9}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(X_i - Y_i)^2} \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(X_i - Y_i)^2}{\sum_{i=1}^{m}(\overline{Y} - Y_i)^2} \tag{11}$$

in which $X_i$ is the predicted ith value, $Y_i$ is the calculated $i$th value, and $\overline{Y}$ is given by the relation bellow [69]:

$$\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i \tag{12}$$

The values in Table 2 represent the performance metrics of the four implemented models: ARX, NLARX, MLP, and DNN. All the models show excellent system fits, exceeding 99.6% with both training (estimation) and validation datasets. For FPE, the nonlinear models (NLARX, MLP, DNN) have lower FPE values than linear ARX, with a slight advantage for MLP, which indicates better equilibrium between fit and model complexity.

**Table 2.** Results of SI and AI-supported SI implementation.

| | | **ARX** | **NLARX** | **MLP** | **DNN** |
|---|---|---|---|---|---|
| System fitting (%) | Estimation/Training | 99.80 | 99.82 | 99.73 | 99.83 |
| | Validation | 99.75 | 99.67 | 99.48 | 99.73 |
| FPE | | 0.0002107 | 0.0001136 | $7.43217 \times 10^{-5}$ | 0.000157467 |
| MSE | | 0.0002 | 0.0002 | 0.0004 | 0.0002 |
| RMSE | | 0.0141 | 0.0149 | 0.0194 | 0.0129 |
| $R^2$ | | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

It can be seen that the MLP model has the highest error and RMSE values among the models, but these are still very low. The DNN model has the best RMSE value, followed by the ARX and NLARX models. Generally, all the obtained values are very low, which confirm the prediction accuracy of the four models.

In terms of $R^2$, all the models have the same value ($R^2$ = 1.0000), confirming the perfect correspondence between predicted and measured data.

To further analyze the obtained results, a scatter plot comparing predicted ET0 with measured ET0 for the four system identification models is generated in Figure 18. Each model's predictions are denoted by distinct markers, with the ideal fit depicted using a red dashed line.

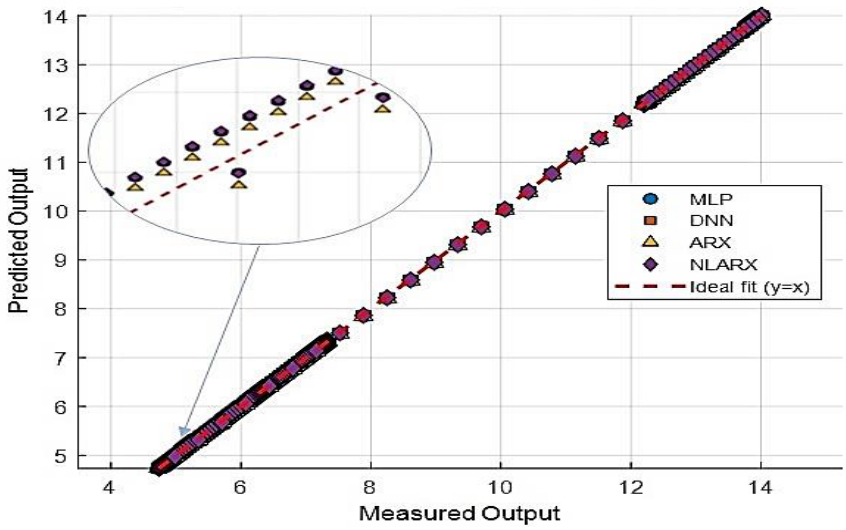


**Figure 18.** Comparison of the four system identification models.

All models demonstrate a high agreement with the ideal fit, as indicated by the clustering of points in close proximity to the y = x reference line, and the negligible deviation from the ideal line for all models reflects both high accuracy and robustness, with no apparent outliers.

The findings show that modern nonlinear and deep learning approaches (MLP, DNN, NLARX) do not provide substantial benefits over the classical linear ARX model within this dataset, as all models achieve nearly perfect fits. This conclusion is further supported by the performance metric values in Table 2, which are consistent with the graphical results and confirm the reliability of all models.

A challenging issue in system modeling and deep learning implementation is data availability. Given the dataset size in this study, another validation strategy is needed to demonstrate that the near perfect values of $R^2$ and low errors are not the result of overfitting or a specified dataset partition. The K-fold cross-validation method is adopted. By training and validating the models on different data subsets, overfitting is improbable.

The results in Table 3 display the values of $R^2$, RMSE, and MSE of the four models after executing K-fold CV with 5 folds (5-fold CV).

**Table 3.** K-fold CV summary of the four models.

| | DNN | MLP | NLARX | ARX |
|---|---|---|---|---|
| $R^2$ (mean ± std) | 1.000 ± 0.000 | 1.000 ± 0.000 | 1.000 ± 0.000 | 1.000 ± 0.000 |
| RMSE (mean ± std) | 0.016 ± 0.004 | 0.012 ± 0.004 | 0.019 ± 0.008 | 0.020 ± 0.006 |
| MSE (mean) | 0.0003 | 0.0002 | 0.0004 | 0.0004 |

For every fold, the models achieve nearly perfect goodness of fit with identical mean coefficients of determination, $R^2 = 1.000 \pm 0.000$.

These results are further supported by the bar plots for $R^2$ in Figure 19 and confidence interval line plots in Figure 20, which show negligible variation across folds.

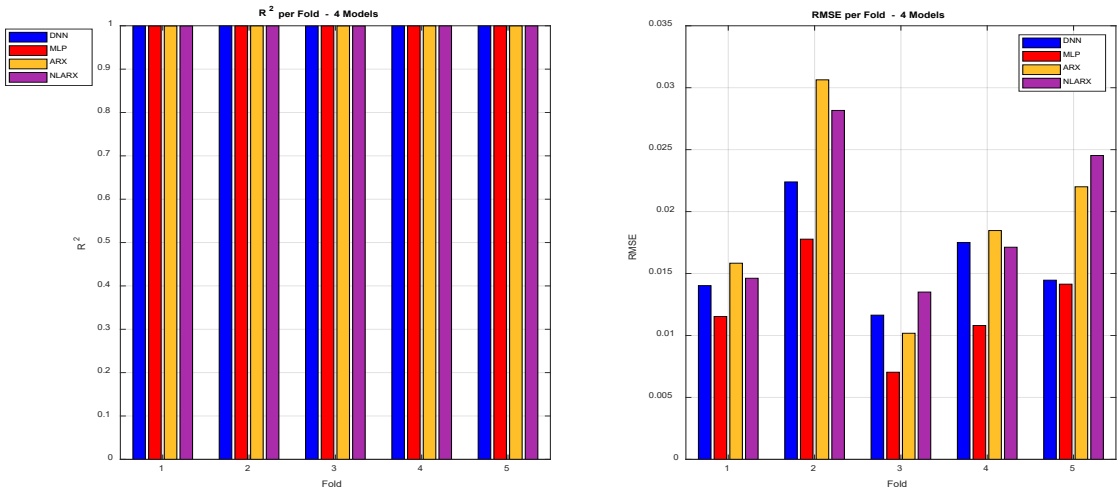


**Figure 19.** $R^2$ and RMSE cross-validation values of the four models over the 5 folds.

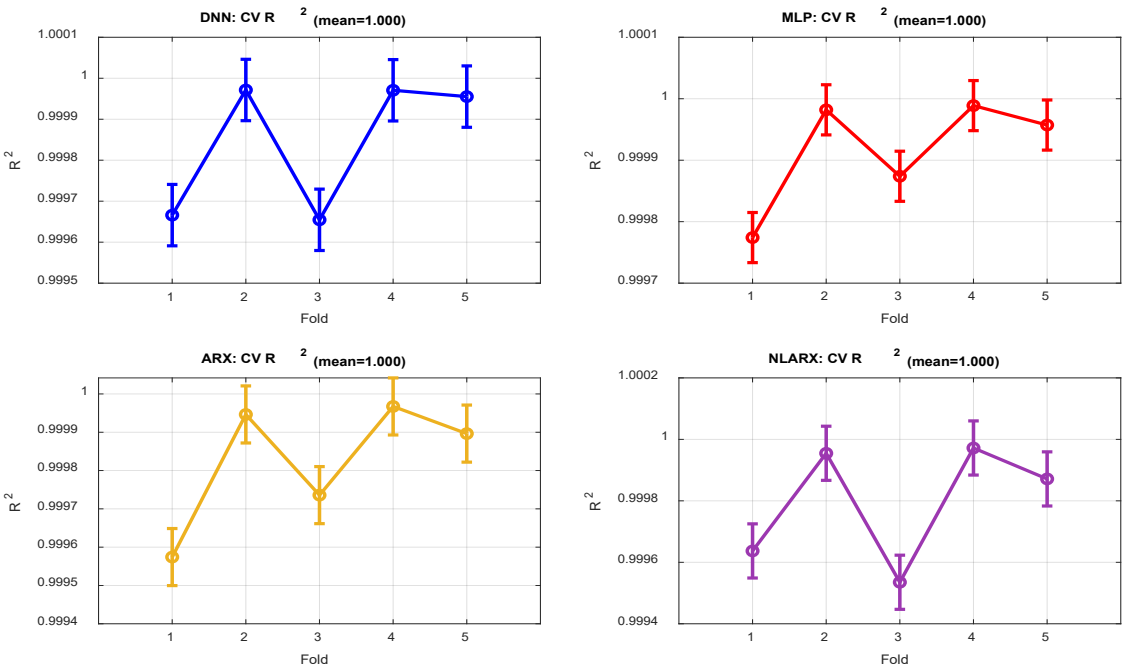


**Figure 20.** $R^2$ Cross-validation values with confidence intervals for the four models.

The lowest prediction error is offered by MLP with an RMSE of 0.012 ± 0.004 and an MSE of 0.0002, followed by DNN (RMSE 0.016 ± 0.0040.016 ± 0.004, MSE 0.0003). This error rose very slightly with the ARX and NLARX identification models, as illustrated in the bar plots for RMSE in Figure 19. Overall, these findings confirm the stability and reproducibility of the performance metrics and justify the high accuracy of the four models on the dataset.

## 4. Conclusions

In this paper, different models are implemented to estimate reference evapotranspiration values, which have a significant role in determining and controlling plants' water requirements. These models provide a key solution to deal with water crisis and food security, especially in arid regions.

The impact of artificial intelligence is studied through the comparison of two classical system identification models and AI-supported system identification models.

In order to develop a resistant model, a heterogeneous database is used; it covers the different climatic conditions that plants face during their growth.

The implemented models show high and very close levels of reliability, varying between 99.67% and 99.82% for classical system identification models and 99.83% and 98.64% for AI-supported system identification models, regardless of the climatic circumstances.

This study's findings demonstrate that integrating artificial intelligence with traditional systems does not necessarily lead to better results for system fitting. Despite the neural network's adaptation to match the data size and the database split, which significantly contributed to preventing overfitting, the classical model showed superiority over the AI-supported models, albeit by a small or marginal difference.

The four implemented models show high performance, with slight advantages in FPE for nonlinear models and RMSE for the DNN model. These results indicate that while the linear model provides the better system fitting ratio, the nonlinear models can improve prediction accuracy and generalization. Given that the four models reach similarly high accuracy, the choice between the two approaches is constrained by data availability and the nature of the studied system. For example, classical systems perform better with small and moderate datasets when modeling physical systems.

To deal with dataset constraints and further confirm the effectiveness of the proposed models, future studies will focus on the use of different climatic parameters and a larger database, as well as the reference Penman–Monteith ET0 model for output calculation, to further improve the precision of ET0 estimation models. This will lead to optimal irrigation water use and improve the agricultural sector.

**Author Contributions:** Conceptualization, W.D., F.B., B.D., K.B. and B.O.; data curation, B.O. and B.D.; methodology, W.D., F.B. and K.B.; software, W.D. and F.B.; validation, W.D., F.B. and B.O.; formal analysis, W.D., F.B. and B.D.; writing—original draft, W.D.; writing—review and editing, W.D., K.L., F.B. and B.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

## References

1. Malamos, N.; Koutsoyiannis, D. Global reference evapotranspiration clustering and its relation to the Köppen-Geiger climate classification. *J. Hydrol.* **2025**, *660*, 133342. [CrossRef]
2. Zhang, Y.; He, T.; Liang, S.; Ma, Y.; Yao, Y. A novel approach for estimating evapotranspiration by considering topographic effects in radiation over mountainous terrain. *Agric. For. Meteorol.* **2025**, *366*, 110468. [CrossRef]
3. Ali, A.; Haq, Z.U.; Sheraz, K. Review on Estimation and Forecast of Reference Evapotranspiration. *Int. J. Agric. Sustain. Dev.* **2025**, *7*, 65–92. [CrossRef]
4. Mousavi, R.; Sabziparvar, A.A.; Marofi, S.; Pak, N.A.E.; Heydari, M. Calibration of the Angström-Prescott solar radiation model for accurate estimation of reference evapotranspiration in the absence of observed solar radiation. *Theor. Appl. Climatol.* **2015**, *119*, 43–54. [CrossRef]
5. Chandel, A.K.; Khot, L.R.; Stöckle, C.O.; Kalcsits, L.; Mantle, S.; Rathnayake, A.P.; Peters, T.R. Canopy Transpiration Mapping in an Apple Orchard Using High-Resolution Airborne Spectral and Thermal Imagery with Weather Data. *AgriEngineering* **2025**, *7*, 154. [CrossRef]
6. Elnashara, A.; Shojaeezadeha, S.A.; Weber, T.K.D. A Multi-model approach for remote sensing-based actual evapotranspiration mapping using Google Earth Engine (ETMapper-GEE). *J. Hydrol.* **2025**, *657*, 133062. [CrossRef]
7. Xu, C.; Hu, X.; Tian, J.; Guo, X.; Lei, J. Estimating the Grape Basal Crop Coefficient in the Subhumid Region of Northwest China Based on Multispectral Remote Sensing by Unmanned Aerial Vehicle. *Horticulturae* **2025**, *11*, 217. [CrossRef]
8. Tayyaba, B.; Khan, M.U.G.; Waheed, T.; Al-Otaibi, S.; Saba, T. TMRE: Novel Algorithm for Computing Daily Reference Evapotranspiration Using Transformer-Based Models. *Comput. Mater. Contin.* **2025**, *83*, 2851–2864. [CrossRef]

9. Arshad, A.; Zia, U.l.H.; Khurram, S. Evaluation of different models to estimate reference evapotranspiration at different climatic zones of Pakistan. *Spectr. Eng. Sci.* **2025**, *3*, 634–647.
10. Prodhana, M.M.H.; Guptaa, L.; Sunnya, A.M.; Islam, M.B. Multiple linear regression-based modeling of reference evapotranspiration for northwestern Bangladesh. *Environ. Ecosyst. Sci.* **2025**, *9*, 41–48. [CrossRef]
11. Mohan, S.; Arumugam, N. Relative importance of meteorological variables in evapotranspiration: Factor analysis approach. *Water Resour. Manag.* **1996**, *10*, 1–20. [CrossRef]
12. Dinpashoh, Y.; Allahverdipour, P. Monitoring and predicting changes in reference evapotranspiration in the Moghan Plain according to CMIP6 of IPCC. *Environ. Water Eng.* **2025**, *11*, 47–56.
13. Asolo, O.O.; Alabi, H.A.; Ewemoje, T.A. Estimation of reference evapotranspiration using fuzzy logic: A case study of the Ogun-Osun River Basin. *J. Res. Agric. Food Sci.* **2025**, *2*, 65–72. [CrossRef]
14. Garbanzo, G.; Céspedes, J.; Temudo, M.; Ramos, T.B.; Cameira, M.d.R.; Pereira, L.S.; Paredes, P. Addressing Weather Data Gaps in Reference Crop Evapotranspiration Estimation: A Case Study in Guinea-Bissau, West Africa. *Hydrology* **2025**, *12*, 161. [CrossRef]
15. Rodrigues, G.C.; Braga, R.P. A Simple Application for Computing Reference Evapotranspiration with Various Levels of Data Availability—ETo Tool. *Agronomy* **2021**, *11*, 2203. [CrossRef]
16. Adjei, F.O.; Adjei, K.A.; Obuobie, E. Calibration of the Hargreaves and Priestley-Taylor equations for estimating reference evapotranspiration in the Densu River Basin in Ghana. *Afr. J. Environ. Sci. Technol.* **2025**, *19*, 131–142.
17. Zhang, J.; Wang, L.; Cheng, G.; Jia, L. Applicability of Evapotranspiration Models and Water Consumption Characteristics Across Different Croplands. *Agronomy* **2025**, *15*, 1441. [CrossRef]
18. Figueiró, L.S.d.P.; Bonfá, C.S.; da Costa Santos, L. Calibration and evaluation of the Hargreaves-Samani equation for estimating reference evapotranspiration: A case study in Northern Minas Gerais, Brazil. *Rev. Gest. Secr.* **2025**, *16*, e4638. [CrossRef]
19. Lykhovyd, P.V.; Vozhehova, R.A.; Piliarska, O.O. Comparison of smart applications for evapotranspiration assessment automation with manual estimation using the Thornthwaite method. *Agrology* **2025**, *8*, 89–95. [CrossRef]
20. Bhosale, M.S.; Kadam, S.A.; Danwale, N.J.; Mane, M.S.; Malunjkar, V.S.; Joshi, J.K. Estimation of Actual Evapotranspiration and Crop Coefficient to Improve Water Use Efficiency of Sesamum (*Sesamum indicum* L.) in Kharif Season Using Digital Weighing Lysimeter. *Asian J. Env. Ecol.* **2025**, *24*, 67–78. [CrossRef]
21. Ratshiedana, P.E.; Abd Elbasit, M.A.M.; Adam, E.; Chirima, J.G. Evaluation of Micrometeorological Models for Estimating Crop Evapotranspiration Using a Smart Field Weighing. *Water* **2025**, *17*, 187. [CrossRef]
22. Alazba, A.A.; Mattar, M.A.; El-Shafei, A.; Radwan, F.; Ezzeldin, M.; Alrdyan, N. Daily Reference Evapotranspiration Derived from Hourly Timestep Using Different Forms of Penman–Monteith Model in Arid Climates. *Water* **2025**, *17*, 2272. [CrossRef]
23. Akar, F.; Katipoğlu, O.M.; Yeşilyurt, S.N.; Taş, M.B.H. Evaluation of Tree-Based Machine Learning and Deep Learning Techniques in Temperature-Based Potential Evapotranspiration Prediction. *Pol. J. Environ. Stud.* **2023**, *32*, 1009–1023. [CrossRef] [PubMed]
24. Youssef, M.A.; Peters, R.T.; El-Shirbeny, M.; Abd-ElGawad, A.M.; Rashad, Y.M.; Hafez, M.; Arafa, Y. Enhancing irrigation water management based on ETo prediction using machine learning to mitigate climate change. *Cogent Food Agric.* **2024**, *10*, 2348697. [CrossRef]
25. Paul, J.C.; Paramaguru, P.K.; Swain, S. Application of innovative data-driven techniques for modeling reference evapotranspiration. *J. Soil Water Conserv.* **2025**, *24*, 47–53. [CrossRef]
26. Kumar, M.; Raghuwanshi, N.S.; Singh, R.; Wallender, W.W.; Pruitt, W.O. Estimating evapotranspiration using artificial neural network. *J. Irrig. Drain. Eng.* **2002**, *128*, 224–233. [CrossRef]
27. Trajkovic, S.; Todorovic, B.; Stankovic, M. Forecasting of reference evapotranspiration by artificial neural networks. *J. Irrig. Drain. Eng.* **2003**, *129*, 454–457. [CrossRef]
28. Wu, C.L.; Chau, K.W.; Li, Y.S. Methods to improve neural network performance in daily flows prediction. *J. Hydrol.* **2009**, *372*, 80–93. [CrossRef]
29. Sudheer, K.P.; Gosain, A.K.; Ramasastri, K.S. Estimating actual evapotranspiration from limited climatic data using neural computing technique. *J. Irrig. Drain. Eng.* **2003**, *129*, 214–218. [CrossRef]
30. Coutinho, E.R.; Madeira, J.G.F.; da Silva, R.M.; Delgado, A.R.S.; Coutinho, A.L.G.A. Genetic Algorithms Applied to Optimize Neural Network Training in Reference Evapotranspiration Estimation. *Rev. Bras. Meteorol.* **2025**, *40*, e40240009. [CrossRef]
31. Li, Z.; Zhao, Z.; Xing, L.; Zhao, L.; Cui, N.; Cai, H. Estimating daily reference crop evapotranspiration in Northeast China using optimized empirical models based on heuristic intelligence algorithms. *Agronomy* **2025**, *15*, 599. [CrossRef]
32. Zhao, J.; Zhong, H.; Wang, C. Estimation of reference crop evapotranspiration in the Yellow River Basin based on machine learning and its regional and drought adaptability analysis. *Agronomy* **2025**, *15*, 1237. [CrossRef]
33. Tausif, M.; Iqbal, M.W.; Bashir, R.N.; AlGhofaily, B.; Elyassih, A.; Khan, A.R. Federated learning based reference evapotranspiration estimation for distributed crop fields. *PLoS ONE* **2025**, *20*, e0314921. [CrossRef] [PubMed]
34. Jayashree, T.R.; Reddy, N.V.S.; Acharya, U.D. Modeling daily reference evapotranspiration from climate variables: Assessment of bagging and boosting regression approaches. *Water Resour. Manag.* **2023**, *37*, 1013–1032. [CrossRef]

35. Nasir, M.A.M.; Harun, S.; Zainuddin, Z.M.; Kamal, M.R.; Rose, F.Z.C. Machine learning techniques for reference evapotranspiration and rice irrigation requirements prediction: A case study of Kerian irrigation scheme Malaysia. *Paddy Water Environ.* **2025**, *23*, 671–683. [CrossRef]
36. Nasrolahia, A.; Sabzevari, Y.; Eslamian, S. Estimation of reference evapotranspiration using experimental models and approaches based on neural network (case study: Khorramabad City, Iran). *Water Pract. Technol.* **2025**, *20*, 1437–1449. [CrossRef]
37. Ayaz, A.; Rajesh, M.; Singh, S.K.; Rehana, S. Estimation of reference evapotranspiration using machine learning models with limited data. *AIMS Geosci.* **2021**, *7*, 268–290. [CrossRef]
38. Lahkim-Bennani, M.A.; Bonaccorso, B.; Tribak, A.A.; Rossi, A. Reference evapotranspiration forecasting in the North Etna aquifer: A comparative analysis of statistical, deep learning, and machine learning models. *Theor. Appl. Climatol.* **2025**, *156*, 492. [CrossRef]
39. Heramb, P.; Rao, K.V.R.; Subeesh, A.; Srivastava, A. Predictive modelling of reference evapotranspiration using machine learning models coupled with grey wolf optimizer. *Water* **2023**, *15*, 856. [CrossRef]
40. Haris, A.; Marimin, M.; Wahjuni, S.; Setiawan, B.I. H-ConvLSTM to estimate reference evapotranspiration from air temperature and relative humidity. *IEEE Access* **2025**, *13*, 128789–128801. [CrossRef]
41. Zolfaghari, A.A.; Raeesi, M.; Longo-Minnolo, G.; Consoli, S.; Dyck, M. Daily reference evapotranspiration prediction in Iran: A machine learning approach with ERA5-land data. *J. Hydrol. Reg. Stud.* **2025**, *59*, 102343. [CrossRef]
42. Boukhali, Y.; Kabbaj, M.N.; Benbrahim, M. Evapotranspiration estimation using deep learning models for robust sensor performance. *Frankl. Open* **2025**, *12*, 100335. [CrossRef]
43. Yu, P.; Wan, H.; Zhang, B.; Wu, Q.; Zhao, B.; Xu, C.; Yang, S. Review on system identification, control, and optimization based on artificial intelligence. *Mathematics* **2025**, *13*, 952. [CrossRef]
44. Ahmed, A.; del Rio-Chanona, E.A.; Mercangöz, M. Comparative study of machine learning and system identification for process systems engineering dynamics. *Ind. Eng. Chem. Res.* **2025**, *64*, 4450–4478. [CrossRef]
45. Kühl, N.; Schemmer, M.; Goutier, M.; Satzger, G. Artificial intelligence and machine learning. *Electron. Mark.* **2022**, *32*, 2235–2244. [CrossRef]
46. Sadeghi, T.S.; Ourang, S.A.H.; Sohrabniya, F.; Sadr, S.; Shobeiri, P.; Motamedian, S.R. Performance of artificial intelligence on cervical vertebral maturation assessment: A systematic review and meta-analysis. *BMC Oral Health* **2025**, *25*, 187. [CrossRef]
47. Mishra, A. A comprehensive review of artificial intelligence and machine learning: Concepts, trends, and applications. *Int. J. Sci. Res. Sci. Technol.* **2025**, *11*, 126–142.
48. Sarker, I.H. Machine learning: Algorithms, real world applications and research directions. *SN Comput. Sci.* **2021**, *2*, 160. [CrossRef]
49. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. Crop evapotranspiration: Guidelines for computing crop water requirements. In *FAO Irrigation and Drainage Paper 56*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 1998; pp. 11–13.
50. Miralles, D.G.; Holmes, T.R.H.; De Jeu, R.A.M.; Gash, J.H.; Meesters, A.G.C.A.; Dolman, A.J. Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 453–469.
51. Anderson, M.C.; Allen, R.G.; Morse, A.; Kustas, W.P. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sens. Environ.* **2012**, *122*, 50–65. [CrossRef]
52. Pereira, L.S.; Allen, R.G.; Smith, M.; Raes, D. Crop evapotranspiration estimation with FAO56: Past and future. *Agric. Water Manag.* **2015**, *147*, 4–20. [CrossRef]
53. Doorenbos, J.; Pruitt, W.O. Guidelines for predicting crop water requirements. In *FAO Irrigation and Drainage Paper 24*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 1977; pp. 1–4.
54. Laaboudi, A.; Mouhouche, B.; Draoui, B. Neural network approach to reference evapotranspiration modelling from limited climatic data in arid regions. *Int. J. Biometeorol.* **2012**, *56*, 831–841. [CrossRef] [PubMed]
55. Kifouche, A.; Kaddour, A.; Lalmi, D.; Chenini, N.; Alkhafaji, M.A.; Chambashi, G.; Kaid, N.; Menni, Y. Enhancing real-time earth–air heat exchanger outlet temperature forecasting in arid climates using artificial neural network: A case study from Bechar, Algeria. *Int. J. Low-Carbon Technol.* **2024**, *19*, 2493–2501. [CrossRef]
56. Belkendil, A.; Habi, M.; Yakhlefoune, M.; Sarra, A.; Mekkaoui, A.; Tariq, A.; Jemai, S.; Zeghmar, A.; Bougrine, H.; Zouidi, M.; et al. Integrated GIS-based approach for groundwater potential mapping in arid regions: A case study of Kenadsa-Bechar using AHP and frequency ratio methods. *Groundw. Sustain. Dev.* **2025**, *30*, 101461. [CrossRef]
57. Mebarki, S.; Kendouci, M.A.; Bendida, A. Monitoring the spatial evolution of groundwater quality during its diversion in the drinking water supply network in arid areas, case of Bechar city (Algeria Sahara). *Appl. Water Sci.* **2024**, *14*, 118. [CrossRef]
58. Difallah, W.; Bounaama, F.; Draoui, B.; Benahmed, K.; Laaboudi, A. Model identification and control of evapotranspiration for irrigation water optimization. *Comput. Mater. Contin.* **2022**, *70*, 1749–1767. [CrossRef]
59. Aly, M.S.; Darwish, S.M.; Aly, A.A. High performance machine learning approach for reference evapotranspiration estimation. *Stoch. Environ. Res. Risk Assess.* **2024**, *38*, 689–713.

60. Marquez-Alvarez, T.; Bedolla, J.H.; Pardo-Loaiza, J.; Lara-Ledesma, B.; Domínguez-Sánchez, C. Multi standardized precipitation evapotranspiration index (Multi-SPEI-ETo): Evaluation of 40 empirical methods and their influence in SPEI. *Agriculture* **2025**, *15*, 703. [CrossRef]
61. Karuppanan, S.; Ramasamy, S.; Lakshminarayanan, B.; Anuthaman, S.N. An effective machine learning model for the estimation of reference evapotranspiration under data-limited conditions. *Res. Agr. Eng.* **2025**, *71*, 22–37.
62. Xie, J.; Li, C.; Li, N.; Li, P.; Wang, X.; Gao, D.; Yao, D.; Xu, P.; Yin, G.; Li, F. Robust Autoregression with Exogenous Input Model for System Identification and Predicting. *Electronics* **2021**, *10*, 755. [CrossRef]
63. Zhang, Q.; Ljung, L. Multiple steps prediction with nonlinear ARX models. *IFAC Proc. Vol. 2004*(37), 309–314. [CrossRef]
64. Irnanda, M.D.; Pratama, A.S.; Putra, F.A.; Sugiyanto, S. Application of Artificial Neural Network (MLP) for Multivariate Analysis of Stunting Causes in Indonesia. *J. Appl. Inform. Comput.* **2025**, *9*, 1785–1792. [CrossRef]
65. Popescu, M.C.; Balas, V.E.; Perescu-Popescu, L.; Mastorakis, N. Multilayer Perceptron and Neural Networks. *WSEAS Trans. Circuits Syst.* **2009**, *7*, 579–588.
66. Isabona, J.; Imoize, A.L.; Ojo, S.; Karunwi, O.; Kim, Y.; Lee, C.-C.; Li, C.-T. Development of a Multilayer Perceptron Neural Network for Optimal Predictive Modeling in Urban Microcellular Radio Environments. *Appl. Sci.* **2022**, *12*, 5713. [CrossRef]
67. Chan, K.Y.; Abu-Salih, B.; Qaddoura, R.; Al-Zoubi, A.M.; Palade, V.; Pham, D.S.; Ser, J.D.; Muhammad, K. Deep neural networks in the cloud: Review, applications, challenges and research directions. *Neurocomputing* **2023**, *545*, 126327. [CrossRef]
68. Çeki¸s, I.K.; Ayrancı, B.; Salman, F.N.; Özçelik, I. IP Spoofing Detection Using Deep Learning. *Appl. Sci.* **2025**, *15*, 9508. [CrossRef]
69. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef] [PubMed]


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