sustainability

MDPI

Article

Developing a New ANN Model to Estimate Daily Actual Evapotranspiration Using Limited Climatic Data and Remote Sensing Techniques for Sustainable Water Management

Halil Karahan [1,*], Mahmut Cetin [2], Muge Erkan Can [2] and Omar Alsenjar [2]

[1] Department of Civil Engineering, Pamukkale University, Denizli 20160, Turkey
[2] Department of Agricultural Structures and Irrigation, Cukurova University, Adana 01250, Turkey; mcet64@cu.edu.tr (M.C.); merkan@cu.edu.tr (M.E.C.); omarsenjar@yahoo.com (O.A.)
* Correspondence: hkarahan@pau.edu.tr

Abstract: Accurate estimations of actual evapotranspiration (ETa) are essential to various environmental issues. Artificial intelligence-based models are a promising alternative to the most common direct ETa estimation techniques and indirect methods by remote sensing (RS)-based surface energy balance models. Artificial Neural Networks (ANNs) are proven to be suitable for predicting reference evapotranspiration (ETo) and ETa based on RS data. This study aims to develop a methodology based on ANNs for estimating daily ETa values using NDVI and land surface temperature, coupled with limited site-specific climatic variables in a large irrigation catchment. The ANN model has been applied to the two different scenarios. Data from only the 38 days of satellite overpass dates was selected in Scenario I, while in Scenario II all datasets, i.e., the 769-day data were used. An irrigation scheme, located in the Mediterranean region of Turkiye, was selected, and a total of 38 Landsat images and local climatic data collected in 2021 and 2022 were used in the ANN model. The ETa results by the ANN model for Scenarios I and II showed that the R^2 values for training (0.79 and 0.86), testing (0.75 and 0.81), and the entire dataset (0.76 and 0.84) were all remarkably high. Moreover, the results of the new ANN model in two scenarios showed an acceptable agreement with ETa-METRIC values. The proposed ANN model demonstrated the potential for obtaining daily ETa using limited climatic data and RS imagery. As a result, the suggested ANN model for daily ETa computation offers a trustworthy way to determine crop water usage in real time for sustainable water management in agriculture. It may also be used to assess how crop evapotranspiration in drought-prone areas will be affected by climate change in the 21st century.

Keywords: evapotranspiration; artificial neural networks (ANNs); remote sensing (RS); METRIC model; climate change; sustainability

Citation: Karahan, H.; Cetin, M.; Can, M.E.; Alsenjar, O. Developing a New ANN Model to Estimate Daily Actual Evapotranspiration Using Limited Climatic Data and Remote Sensing Techniques for Sustainable Water Management. *Sustainability* **2024**, *16*, 2481. https://doi.org/10.3390/su16062481

Academic Editors: Cheng Li, Fei Zhang, Mou Leong Tan and Kwok Pan Chun

Received: 6 February 2024
Revised: 28 February 2024
Accepted: 28 February 2024
Published: 17 March 2024

1. Introduction

Scientists, as well as practitioners, have been seeking appropriate, integrated, and sustainable approaches to determine reference and actual evapotranspiration amounts in both irrigation and rainfed catchments. Actual evapotranspiration (ETa) measurements in the field or estimation by any respective conventional method [1] are needed to ensure the sustainability of irrigation water management and irrigation scheduling at large-scale irrigation schemes, as well as in farmyards, water-balance works in hydrologic studies, accurate designs of hydraulic structures related to irrigation schemes, etc. By introducing a two-step procedure [2], ETa, in other words, crop evapotranspiration ETc, is estimated under standard conditions [2], i.e., well-watered and optimal agronomic conditions. On the other hand, ETa is an essential component of the water budget, hydrological modeling, and irrigation water management in arid and semi-arid regions of the world. Nevertheless, it is not an easy task to acquire a representative ETa value for it can be determined directly using either a lysimeter or water-balance work/approach which are rather labor-intensive,

time-consuming, and very expensive, thereby threatening the sustainability of freshwater resources. As stated clearly by Allen et al. [2], Rawat et al. [3], Gharbia et al. [4], and Alsenjar et al. [5], among others, ETa describes the physical processes of the amount of water that can occur either through evaporation or transpiration according to climatic conditions, crop types, and soil status. Since direct methods of ETa calculation focus only on one point or parcel, they have limitations in showing the variation in ETa spatially and temporally at a large scale [6], particularly in areas with large irrigation schemes of >100,000 ha as in Turkiye. Therefore, accurately estimating ETa is of the utmost importance, as well as a critical issue in water-balance methods and agricultural water management at the irrigation scheme level. In this regard, the remote sensing (RS)-based surface energy balance models are one of the indirect methods to estimate spatially ETa over large-scale irrigation catchments with a high spatial and temporal resolution [7].

Several researchers have estimated ETa based on the RS-based surface energy balance models, such as the "Surface Energy Balance Algorithm for Land" (SEBAL) [8,9], "Mapping EvapoTranspiration at High Resolution and with Internalized Calibration", i.e., METRIC [10,11], and "Surface Energy Balance System" (SEBS) [12]. These models normally use Landsat satellite imagery or Moderate Resolution Imaging Spectroradiometer (MODIS) which include more required RS data alongside weather parameters [7,10], among others. However, as pointed out by Bachour et al. [13], among others, RS technology has some limitations. Therefore, some of these data cannot be provided due to cloud cover and the unavailability of all relevant climatic data. For example, the implemented METRIC model to estimate ETa requires more of the input parameters of Landsat satellite data (net radiation flux (Rn), soil heat flux (G), sensible heat flux (H), latent heat flux (LE), leaf area index (LAI), land surface temperature (LST), surface albedo (α), and normalized difference vegetation index (NDVI), to name but a few).

In addition, H calculation in the METRIC model is based on the anchoring pixels, i.e., hot and cold, so there are complexities involved in sensitivity selecting anchor pixels for improved estimation of H and LE fluxes [14]. Moreover, Landsat satellite data with a 30 m by 30 m spatial resolution are normally available every 16 days for satellite scenes [7]. As such, the methodology obtained by Allen et al. [10] has been used to estimate ETa based on the METRIC model for each pixel during the satellite image overpass date. In this regard, to determine ETa values based on the METRIC model for a predetermined period, i.e., month or season, Allen et al. [10] applied the procedure by making the interpolation of daily reference evapotranspiration fraction (ETrF) or crop coefficient (Kc) values between two satellite images and multiplying by reference evapotranspiration, i.e., ETo, for each day and then integrated for a given or specific period. However, the accepted method by Allen et al. [10] has some gaps or limitations in showing the effect of precipitation or irrigation practices precisely on the daily growth stages and NDVI of crops between two satellite image overpasses. As known, both ETa and ETo have a nonlinear character in nature [15], and in turn, is a complex phenomenon. Therefore, more nonlinearity exists in the evapotranspiration process due to its stochastic behavior [16]. Thus, over the past few decades, artificial neural networks (ANNs) have been successfully utilized in modeling reference evapotranspiration, i.e., ETo [1,15,16]. As such, the idea of artificial neural networks usage in engineering applications goes back to the 1940s [15]. Due to these reasons, it has been widely used in hydrological practices, particularly in the estimations of reference evapotranspiration, since the early nineties. Contrary to this, the literature review revealed that there have been hardly any studies on estimating METRIC-based actual evapotranspiration values by ANN models.

The most significant merit of the ANN models is to solve complex problems using fewer inputs by adjusting the weights to be able to predict the correct output of the input parameters. Recently, the ANN algorithm has been applied to estimate reference evapotranspiration (ETo) and crop evapotranspiration (ETc) for wheat, maize, and potato in different regions of the world [17–23]. As for ANN-based ET estimation, studies rarely used land surface parameters calculated from RS data, including but not limited to land

surface temperature, vegetation indices, etc. and limited meteorological parameters as inputs in the ANN model [24].

ETa estimation by the ANN model can be beneficial and powerful for using it as an input factor in water-balance calculations at large-scale irrigation catchments since, as claimed by Kumar et al. [1], theoretically, ANN is expected to produce better results than a regression model for the same data length. The novelty of this study is to establish a new methodology for generating daily actual evapotranspiration (ETa) series based on the ANN model using some of the parameters of MODIS data coupled with availably daily moderate spatial and less input of weather variables as compared to the existing methods of ETa-based surface energy balance models estimation. Therefore, this study aims to use an artificial neural network (ANN) approach to estimate daily ETa values in large-scale irrigation catchments using two parameters of MODIS data, i.e., NDVI and LST, coupled with limited site-specific climatic variables at a large-scale irrigation scheme located in the Lower Seyhan Plain irrigation project area with >210,000 ha of land [7]. Furthermore, this paper is the first attempt to generate a new model using the ANN algorithm as an alternative to the existing methods of actual evapotranspiration estimation in a large-scale irrigation district in the Eastern Mediterranean Region of Turkiye. Moreover, this methodology can be generalized for estimating daily ETa using the ANN model to different climate regions and zones of the world.

2. Materials and Methods

2.1. Study Area and Its Characteristics

The research area that has gained popularity under the name Akarsu Irrigation District (hereafter, AID, A = 9495 ha $\approx$ 95 km^2, Figure 1) in the studies carried out so far is located in the Lower Seyhan Plain (LSP) in the southeastern part of the Mediterranean region of Turkiye. The LSP shows typical characteristics of a deltaic plain with a rather flat topography (a slope of 1% or less) and a large-scale irrigation and drainage network [5,7]. The Mediterranean climate, characterized by warm and rainy in the winter season, whereas dry and hot in the summer season, prevails utterly in the LSP and, in turn, in the study area. The average annual precipitation of the basin is approximately 650 mm [5]. In the LSP of Turkiye, there is a very remarkable difference in temperature and evaporation values in the irrigation season, i.e., in July and August in particular, compared to those in the winter season, i.e., in December, January, and February [5,7]. As reported by Cetin et al. [7], due to meteorological and geographical factors, the definition of the water year, i.e., hydrological year, varies from region to region. It has been defined as the period, with a length of 365 days, between 1 October of one year and 30 September of the next, as late September to early October is the time for many drainage areas or catchments in Turkiye to have the lowest stream flows and consistent groundwater levels.

2.2. Remote Sensing Data Used

2.2.1. Landsat Satellite Imagery

To run the METRIC model and for the actual evapotranspiration (ETa) estimations, a total of 38 clear-sky Landsat satellite images were downloaded from the USGS website (http://earthexplorer.usgs.gov, accessed on 16 February 2023) (path 175, row 34) and used in this research (Table 1). These images are Landsat 7, Landsat 8, and Landsat 9 with 30 m by 30 m spatial resolution. General characteristics of the Landsat satellite images are given in Table 1. The "Environment for Visualizing Images (ENVIs 5.3)" software program was applied to perform a cloud mask for one satellite image on 1 May 2022, i.e., DOY 121, of grayscale fill, as shown in Table 1. Scan Lines Error for Landsat 7 ETM$^+$ was corrected and/or the gaps were filled before using this data through the implementation of Quantum Geographic Information System (QGIS 3.24.3) and Landsat toolbox by ArcGIS 10.4.1.

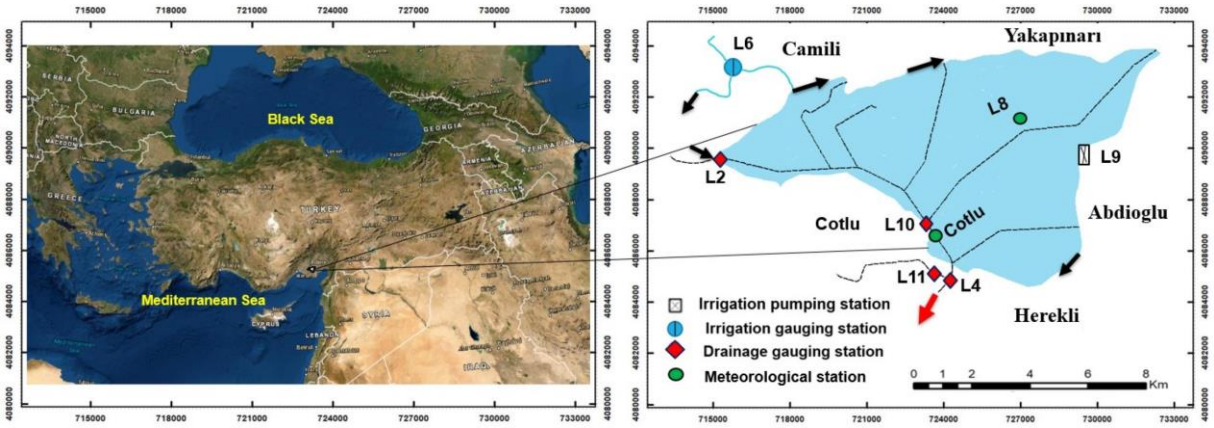

Figure 1. The study area is located in the southeastern Mediterranean region of Turkiye. Dotted lines, red and black arrows represent the drainage canals, and water flow directions, respectively. Meteorological stations are located at L8 and Cotlu. Irrigation water is diverted from L6 and L9 locations into the AID; L2 and L11 stand for drainage water inputs and L4 is the drainage outlet of the catchment.

Table 1. Availability of Landsat 7, Landsat 8, and Landsat 9 scene information in the 2021 and 2022 water years: names of scenes, acquisition dates, and overpass time.

Image	Day of the Year (DOY)	Landsat Scene-ID	Satellite Type	Cloud Cover (%)	Acquisition Dates	Overpass Local Time (AM)
1	260	LC81750342020260LGN00	Landsat 8	1	16 September 2020	11:15:56.5028510
2	300	LE71750342020300SG100	Landsat 7	8	26 October 2020	10:38:56.1154274
3	316	LE71750342020316NPA00	Landsat 7	3	11 November 2020	10:37:51.0228172
4	364	LE71750342020364NPA00	Landsat 7	1	29 December 2020	10:34:21.8153233
5	22	LC81750342021022LGN00	Landsat 8	9	22 January 2021	11:15:49.9861710
6	54	LC81750342021054LGN00	Landsat 8	7	23 February 2021	11:15:43.2139690
7	79	LE71750342021078SG100	Landsat 7	5	19 March 2020	10:28:24.8443048
8	118	LC81750342021118LGN00	Landsat 8	8	28 April 2021	11:15:15.8809360
9	134	LC81750342021134LGN00	Landsat 8	1	14 May 2021	11:15:15.9098560
10	158	LE71750342021158SG100	Landsat 7	1	7 June 2021	10:21:43.6865663
11	182	LC81750342021182LGN00	Landsat 8	4	1 July 2021	11:15:35.1871370
12	190	LE71750342021190SG100	Landsat 7	3	9 July 2021	10:19:04.5579664
13	198	LC81750342021198LGN00	Landsat 8	5	17 July 2021	11:15:36.9021800
14	214	LC81750342021214LGN00	Landsat 8	0	2 August 2021	11:15:45.3409259
15	230	LC81750342021230LGN00	Landsat 8	2	18 August 2021	11:15:51.0643500
16	262	LC81750342021262LGN00	Landsat 8	9	19 September 2021	11:15:59.0216650
17	278	LC81750342021278LGN00	Landsat 8	1	5 October 2021	11:16:04.4435930
18	294	LC81750342021294LGN00	Landsat8	0	21 October 2021	11:16:07.4309270
19	326	LC81750342021326LGN00	Landsat 8	6	22 November 2021	11:16:02.0432969
20	358	LC81750342021358LGN00	Landsat 8	4	24 December 2021	11:15:59.4307040
21	001	LE71750342022001NPA00	Landsat 7	13	1 January 2022	10:03:01.7753300
22	017	LE71750342022017NPA00	Landsat 7	3	17 January 2022	10:01:30.1854341
23	049	LE71750342022049NPA00	Landsat 7	6	18 February 2022	09:58:18.1083401
24	081	LE71750342022081NPA00	Landsat 7	19	22 March 2022	09:55:10.7155745
25	089	LC81750342022089LGN00	Landsat 8	7	30 March 2022	11:15:25.6965150
26	113	LC91750342022113LGN00	Landsat 9	2	23 April 2022	11:15:26.7327310
27	121	LC81750342022121LGN00	Landsat 8	60	1 May 2022	11:15:28.7371250

Table 1. *Cont.*

Image	Day of the Year (DOY)	Landsat Scene-ID	Satellite Type	Cloud Cover (%)	Acquisition Dates	Overpass Local Time (AM)
28	140	LE71750342022140SG100	Landsat 7	12	20 May 2022	09:50:56.6332014
29	157	LE71750342022157SG100	Landsat 7	23	6 June 2022	09:50:06.2292853
30	169	LC81750342022169LGN00	Landsat 8	1	18 June 2022	11:15:53.3072380
31	186	LE71750342022186SG100	Landsat 7	0	5 July 2022	09:42:29.8002719
32	201	LC81750342022201LGN00	Landsat 8	1	20 July 2022	11:15:58.6410700
33	209	LC91750342022209LGN00	Landsat 9	0	28 July 2022	11:15:42.4933480
34	220	LE71750342022220SG100	Landsat 7	29	8 August 2022	09:39:50.3921524
35	237	LE71750342022237SG100	Landsat 7	9	25 August 2022	09:38:18.9640686
36	249	LC81750342022249LGN00	Landsat 8	8	06 September 2022	11:16:15.2109079
37	271	LE71750342022271SG100	Landsat 7	3	28 September 2022	09:34:51.9536731
38	297	LC81750342022297LGN00	Landsat 8	0	24 October 2022	11:16:18.0041120

2.2.2. Moderate Resolution Imaging Spectroradiometer (MODIS) Products

Two parameters of MODIS data, i.e., normalized difference vegetation index (NDVI) and land surface temperature (LST) were downloaded by the Google Earth engine ('MODIS/MOD09GA_006_NDVI'; 'MODIS/061/MOD11A1_LST') to apply the artificial neural network (ANN) for estimating daily ETa values for the entire study area. Typical characteristics of MODIS data are illustrated in Table 2 along with the spatial and temporal resolutions.

Table 2. MODIS data used in the research area.

MODIS Standard Products	Parameter	Spatial Resolution	Temporal Resolution
MOD09GA-Terra	NDVI	500 m by 500 m	Daily
MOD11A1.061-Terra	LST	1000 m by 1000 m	Daily

Daily NDVI and LST datasets were used as casual variables in the modeling practice since NDVI and LST by Landsat or Sentinel-2 are not available daily [5,7] due to the revisit period of the satellites. However, NDVI and LST by MODIS products are available as daily time series even though they have low spatial resolutions. Furthermore, these indices by MODIS data can be used to monitor daily crop development, surface temperature, and spatial distribution over the study area if compared to other remote sensing data which revisit the time of satellite every 3–5 days for Sentinel-2 and 8–16 days for Landsat. These remote sensing data, however, require an interpolation process to fill the gaps of NDVI and LST between two overpass satellite dates to obtain a daily dataset.

2.3. In Situ Meteorological Observations

In this research, hourly and daily climatic variables (minimum and maximum temperatures (Tmin, Tmax), wind speed (U), solar radiation (Rs), minimum and maximum relative humidity values (RHmin and RHmax), and precipitation (P)) acquired from two meteorological stations, i.e., L8 and Cotlu meteorological stations established in the AID (Figure 1), were used in calculations. Before using any climatic data observed in L8 and Cotlu meteorological stations, quality controls, i.e., QC (gaps in the data, outliers, constant values, jumps, etc.) were checked thoroughly. No consistency was found in the meteorological datasets. Climatic datasets cover from 16 September 2020 to 24 October 2022, i.e., 769 daily datasets.

2.4. Reference Evapotranspiration (ETo) Estimation

Reference evapotranspiration (ETo in mm day^{-1} unit) is defined and computed using the standard FAO–Penman–Monteith approach given by Allen et al. [2]. Equation (1) was

developed for short grass; the albedo was 0.23, whilst the aerodynamic resistance was 70 s m^{-1}.

$$ETo = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T+273}u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where ETo is the reference evapotranspiration (mm day^{-1}), R_n is the net radiation at the crop surface (MJ m^{-2} day^{-1}), G is the soil heat flux density (MJ m^{-2} day^{-1}), T is the mean daily air temperature at 2 m height (°C), u_2 is the wind speed at 2 m height (m s^{-1}), e_s is the saturation vapor pressure (kPa), e_a is the actual vapor pressure (kPa), $e_s - e_a$ is the saturation vapor pressure deficit (kPa), Δ slope (kPa $°C^{-1}$) is the vapor pressure curve, and γ the psychrometric constant (kPa $°C^{-1}$).

2.5. METRIC Model

The METRIC model was applied to estimate the following: (a) surface energy balance components (SEBs), i.e., latent heat (LE), net radiation (R_n), sensible heat (H), and soil heat flux (G) in Equation (2), (b) ETa for each pixel, and the whole study area by using Landsat satellite imagery and meteorological stations, i.e., L8 and Cotlu at the time of satellite overpass, primarily based on Allen et al. [10,11] through R-METRIC model using a water package in the *R* program [25] and LandMOD ET mapper-MATLAB [26].

$$LE = R_n - G - H \quad (2)$$

All the energy fluxes are in the unit of watt per meter square (i.e., W m^{-2}). Further information on the METRIC model and its associated equations, i.e., step-by-step ETa calculation, as well as the FAO–Penman–Monteith approach, is given by Allen et al. [10].

2.6. Developing an ANN Model for Actual Evapotranspiration (ETa) Estimation

Artificial Neural Networks (ANNs) are mathematical models that resemble biological neural networks. ANNs can learn from examples and adapt solutions over time by recognizing patterns in data, along with rapidly processing information [27]. In essence, ANNs are tools to mimic the underlying likely relationship between input and output variables in the hand adequately.

Water resources and hydrological processes are often complex, multivariable, and nonlinear. ANNs exhibit a flexible structure to address these complex relationships, making them capable of learning and integrating complex relationships by using various input data. Therefore, in recent times, ANNs have been increasingly utilized in hydrology and water resource management. ANNs might be considered flexible modeling tools and can theoretically model any type of relationship with good accuracy. With ANNs, there is no need to make specific assumptions about the models and the underlying probability distributions or relationships; the underlying relationship is determined solely through data mining procedures.

This data-driven approach is one of the most significant advantages of ANNs in solving various complex real-world prediction problems. ANNs have been used in a whole range of hydrological applications, including reference evapotranspiration estimations and predicting groundwater levels [28,29], flood forecasting simulations [30], rainfall and streamflow modeling [31], and aquifer parameter estimations [32].

Models for calculating ETo and plant water requirements involve a myriad of variables such as meteorological data, soil properties, plant type, and climate conditions. ANNs have a significant advantage in handling these complexities due to their ability to use a large amount of data. In particular, when trained with large datasets, these networks have a better capacity to learn complex relationships and patterns for making logical predictions. Predicting future changes in water resources and plant water requirements due to climate change is becoming increasingly important. In ET modeling, machine learning algorithms are being used more and more as alternatives to traditional methods [33,34]. These algorithms can be used as alternatives to traditional equations for ETc and/or ETo

predictions. They also provide insights into how ET behaves over time and space [35,36]. ANNs can be trained and adapted to be used in different geographical areas, allowing for customized predictions based on different plant species and climatic conditions.

Despite their mentioned advantages, ANNs have some important disadvantages. They require a significant amount of data to learn complex relationships and to determine the optimal network architecture [37]. To create the network structure, the number of hidden layers in the model and the optimal number of neurons in each layer need to be determined. In most of the ANN studies in the literature, a trial-and-error procedure is used to determine the network architecture, which is a time-consuming process. According to Maier and Dandy [38], in most water resources problems, using a single hidden layer is sufficient. Therefore, in this study, a single hidden layer is used in ETa estimation.

To determine the optimal number of neurons in the hidden layer, 80% of the data was used for training the network, and 20% was used for testing. This process was repeated for 100 different randomly selected training and testing datasets, and the Mean Squared Error (MSE) value was calculated. This process was repeated in a loop from 1 to the maximum number of neurons, which is 30 in this study, and the number of neurons that yielded the minimum MSE value was selected as the optimal number of neurons, and the analyses were conducted accordingly [37]. Figure 2 illustrates the typical structure of multi-layer ANNs. The input layer consists of 4 nodes, 2 parameters acquired by the two meteorological stations (Figure 1) installed in the study area, and two variables downloaded from MODIS satellite data. In situ climatic observations are solar radiation (Rs), extraterrestrial radiation (Ra), wind speed (u2) at 2 m height, and the reference evapotranspiration values, i.e., ETo, were calculated by the standard FAO–Penman–Monteith approach through following Allen et al. [2]. The connections between the input layers and 30 hidden layers take different weights and are trained depending on the required output of daily ETa.

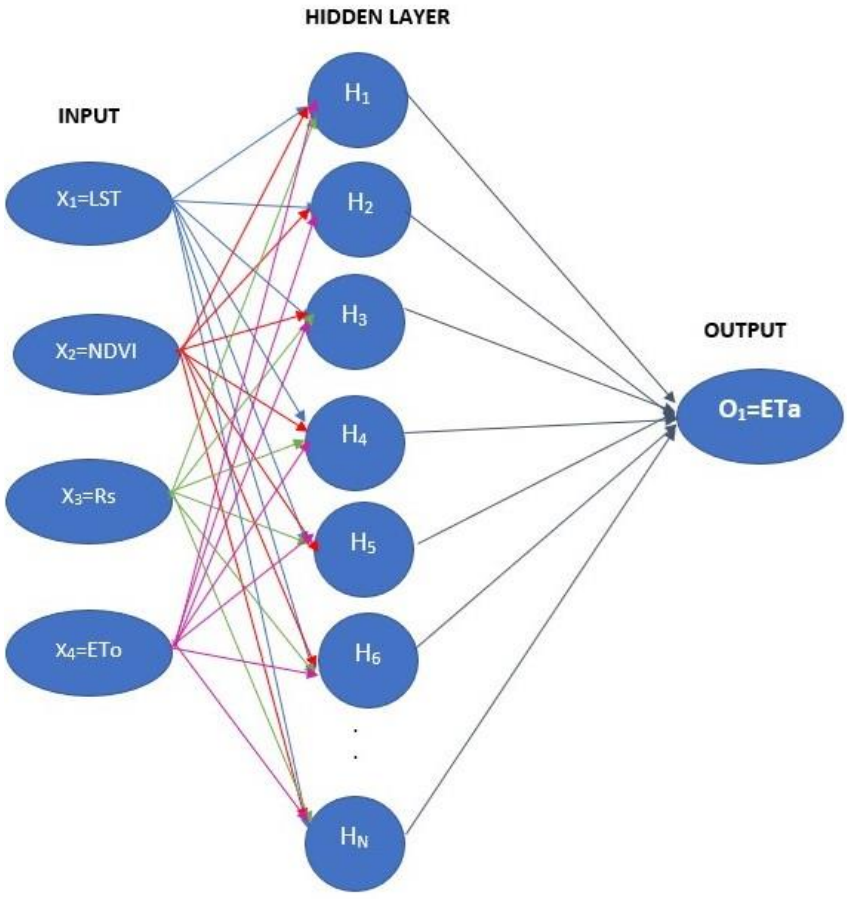

Figure 2. The typical structure of multi-layer ANNs used in this study.

3. Results

3.1. Implementation of the ANN Model

The ANN model has been applied to two different scenarios. In the training of Scenario I, data from only the 38 days of satellite observations listed in Table 1 were used, while the data, which were acquired from the interpolation procedure by Allen et al. [10] from the other days were used as test data. In Scenario II, 80% of the total dataset was used for training, and the remaining 20% was used for testing.

For both scenarios, firstly, the network architecture was created, and the change in the MSE value concerning the number of neurons is presented in Figure 3a for Scenario I and Figure 3b for Scenario II. As seen in Figure 3, the optimum number of neurons obtained for Scenario I and II is 5 and 15, respectively.

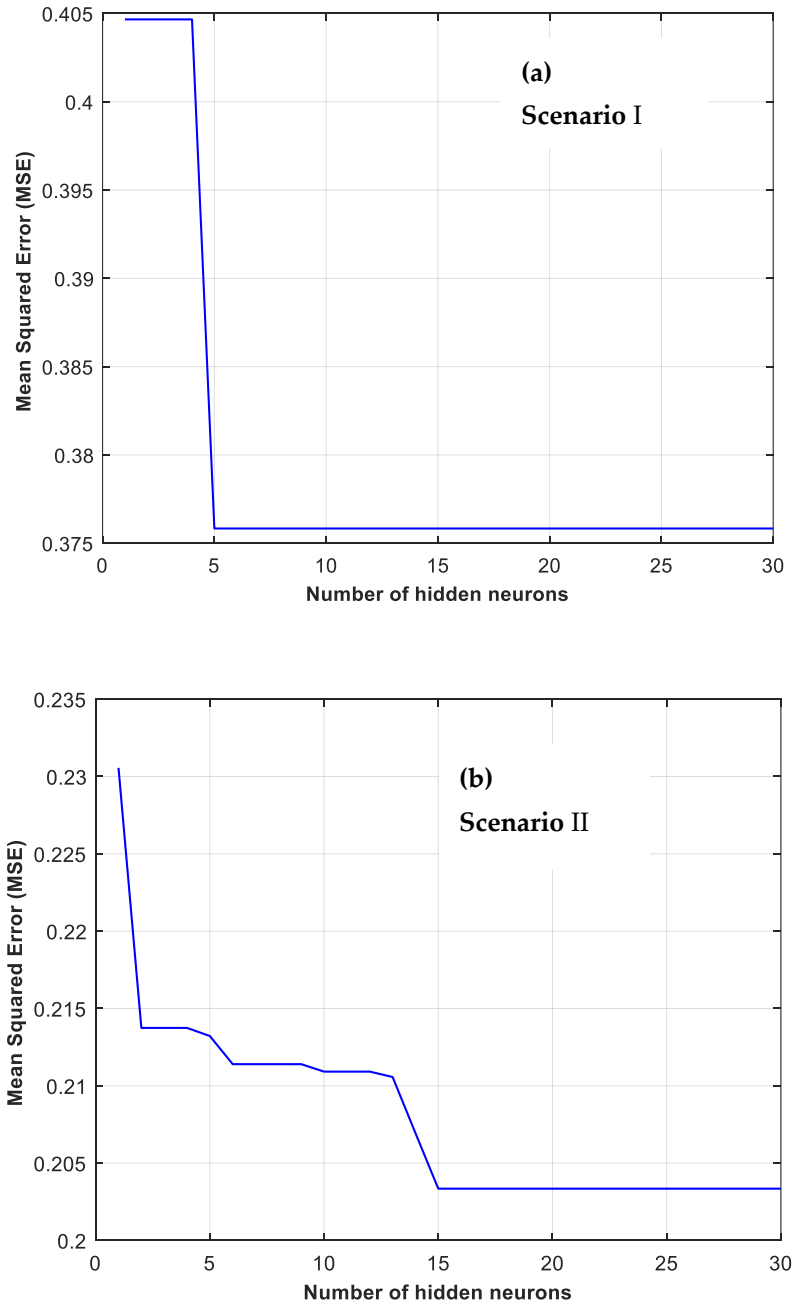

Figure 3. A three-layer feed-forward ANN.

3.1.1. Scenario I

The developed model was applied to Scenario I, and the model results are summarized in Figure 4. As seen in Figure 4, even though the data used in the model training accounts for approximately 5% (N = 38) of the total data, the R-squared value is 0.7547 for the test data (N = 731) and 0.7561 for the total data (N = 769). This indicates that the performance of the developed model is quite good concerning the learning stage and prediction ability; however, it also shows an increased error rate in predicting high ETa values. This result is a natural consequence of the need for a substantial amount of data for ANN models to learn complex relationships in complex problems, as mentioned above. In the following section, this situation will be evaluated in more detail.

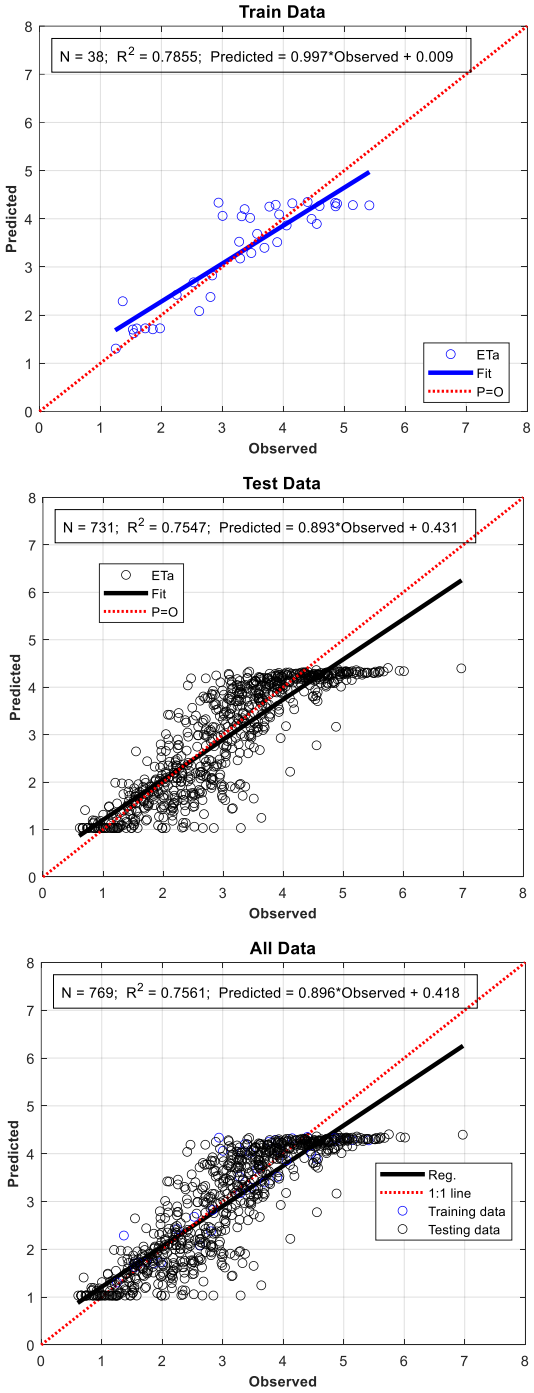

Figure 4. Model results for Scenario I.

3.1.2. Scenario II

Figure 5 shows the model results of the developed model for Scenario II. As seen in Figure 5, the number of data points used in the model training is 615, which constitutes 80% of the total data. During the training phase, the R-squared value is 0.8496, while R-squared values are 0.8055 and 0.8411 for the test data and total data, respectively. These results indicate that the developed model has a very good learning and prediction ability.

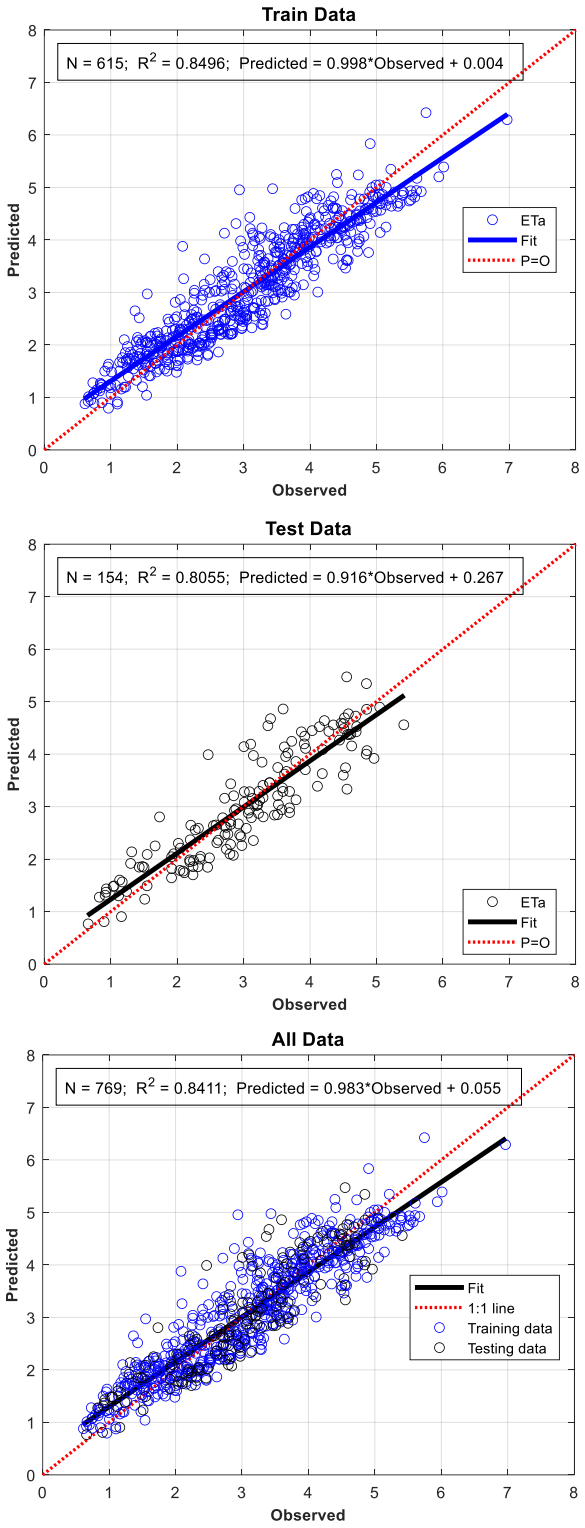

Figure 5. Model results for Scenario II.

4. Discussion

The primary objective of this study was to develop a methodology for estimating daily ETa values for large irrigation areas based on artificial neural networks using some parameters of MODIS data and in situ climatic data with ETo values. It is important to highlight that ETo estimations will be an easy task, provided that in situ climatic data required by the standard FAO–Penman–Monteith approach [2,39–44] are available at the study site. In our study, there are two meteorological stations for collecting data. However, the problem is that Landsat satellite images are available with a 16-day repeat cycle. Therefore, ETa values for the days of no satellite overpass need to be estimated by using an appropriate methodology such as ANNs. To this end, the developed methodology consists of two main parts. In the first stage of the methodology, daily ETo values calculated using the standard FAO–Penman–Monteith approach detailed in Allen et al. [2] are combined with MODIS data to create the input data for the proposed ANN model. These values are based on observations from a meteorological station established in the study area. Additionally, observed ETa values are calculated based on Allen et al. [10,11] using L8 and Cotlu meteorological station data for each pixel at satellite overpass times and other times. This completes the dataset required for the training and testing stages of the developed ANN model.

After completing the dataset, calculations are performed in two different scenarios. In Scenario I, only data acquired for the days with satellite imagery are used for model training, while a randomly selected 80% portion of the total dataset is used in Scenario II. It should be noted that the input and output layers of both scenarios are the same, as shown in Figure 2. As can be seen in Figure 3, the optimal number of neurons in the hidden layer is determined to be 5 for Scenario I and 15 for Scenario II. After determining the appropriate network architecture, detailed analyses were conducted for both scenarios.

In the model training, the goal was to minimize the MSE values as the objective function, and the results obtained for training, testing, and the entire datasets were provided graphically in the previous section. In the relevant graphs, the number of data points used, the R^2 value, and the regression relationships between predicted values and observed values are presented.

As seen in Figure 4, in Scenario I, even though approximately 5% of the total dataset was used for model training, high R^2 values of 0.7855 (N = 38), 0.7547 (N = 731), and 0.7561 (N = 769) were achieved for training, testing, and the entire datasets, respectively. However, it can be observed from the relevant graphs that despite the overall high model performance, the model's prediction ability in Scenario I decreases unexpectedly if the ETa values are approximately 4 mm day^{-1} or greater, indicating conspicuous underestimates of high actual evapotranspiration rates. The underestimation behavior of the model in Scenario I indicates evidently that the model fails to mimic the actual evapotranspiration of 4 mm day^{-1} or greater. This behavior can be seen more clearly in Figure 6, where the model results are presented graphically for both scenarios in comparison with ETa values calculated based on Allen et al. [10,11,42,43]. As clearly visible in the respective Figure, although the general trend is captured in both scenarios, it is noteworthy that the error rate in predicting values, especially those greater than four, increases in Scenario I. It is believed that this is due to the low number of satellite observations (only 38 daily data with very few extreme values) available for training the ANN model. Scatter plots provided in Figure 5 for Scenario II, as well as the temporal ETa values presented in Figure 6, support our postulate. Most studies in the literature also support this argument by allocating 80% of the total dataset for training and the remaining 20% for testing when training ANN models.

However, there are some challenges encountered in both the provision of Landsat satellite data and the implementation of the procedure suggested by Allen et al. [10]. These issues of Landsat satellite data belong to the availability of satellite images every 8–16 days, and cloud coverage, especially for the rainy seasons. Furthermore, using the METRIC model, as suggested by Allen and other scholars, requires well-trained experts with a strong background in handling RS data and choosing the anchor pixels, i.e., hot and cold

pixels. Therefore, in order to determine the optimal dataset to be used in training the ANN model, ratios of different training data used to the total data (TRNRatio) were employed, and the change in model performance is presented in Figure 7. As seen clearly from Figure 7, the optimal TRNRatio value was 0.20. Although the model performance, i.e., R-squared values for training, testing, and all datasets, increases, to some extent, up to this value, there is a significant decrease in test performance for larger TRNRatio values, indicating overfitting. Consequently, the model becomes more and more dependent on the data, which is considered a weakness of the ANN models. This situation can be observed from the scatter plots provided in Figure 8. After evaluating the analysis results given in Figures 7 and 8, it can be concluded that a TRNRatio value of 0.10 can yield ETa values that are quite satisfactory from a practical perspective without increasing the number of satellite data.

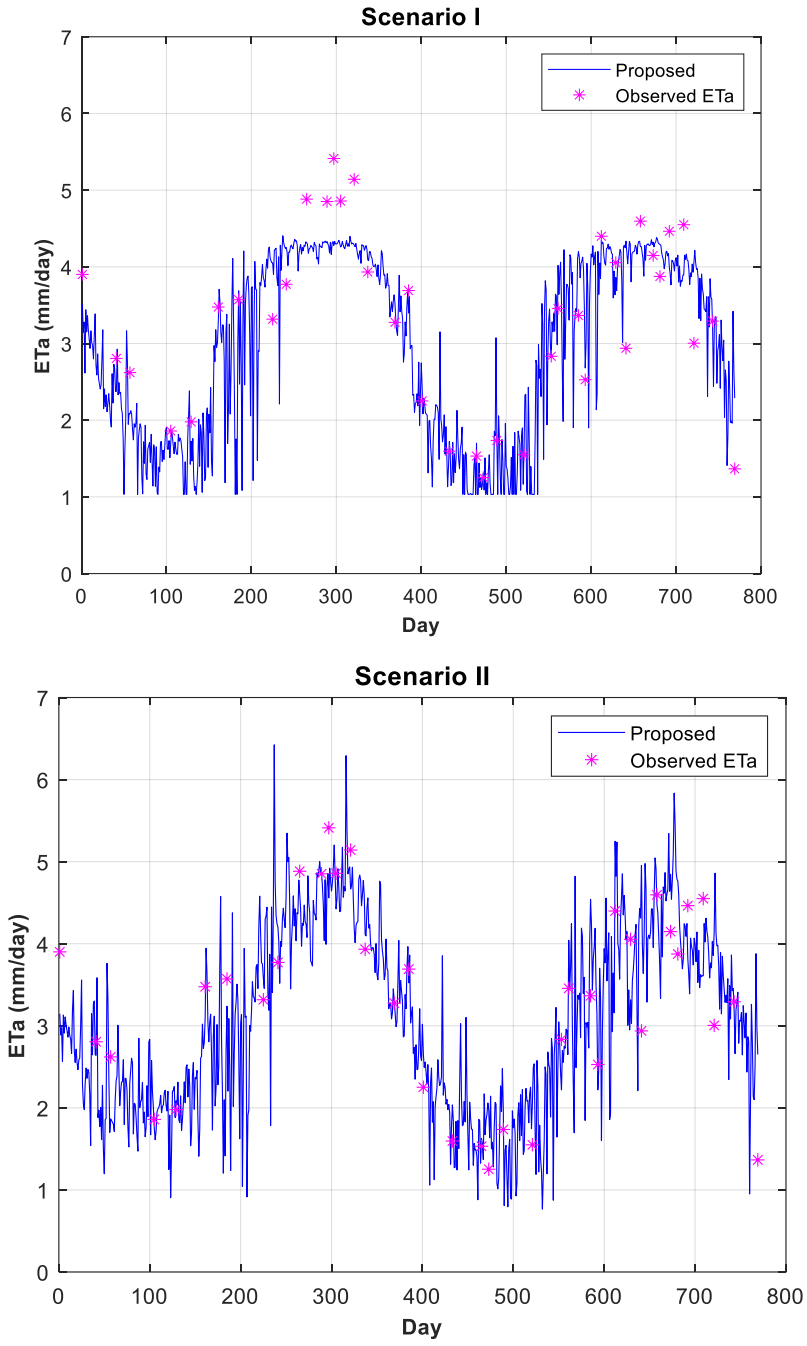

Figure 6. Model results for Scenario I and II.

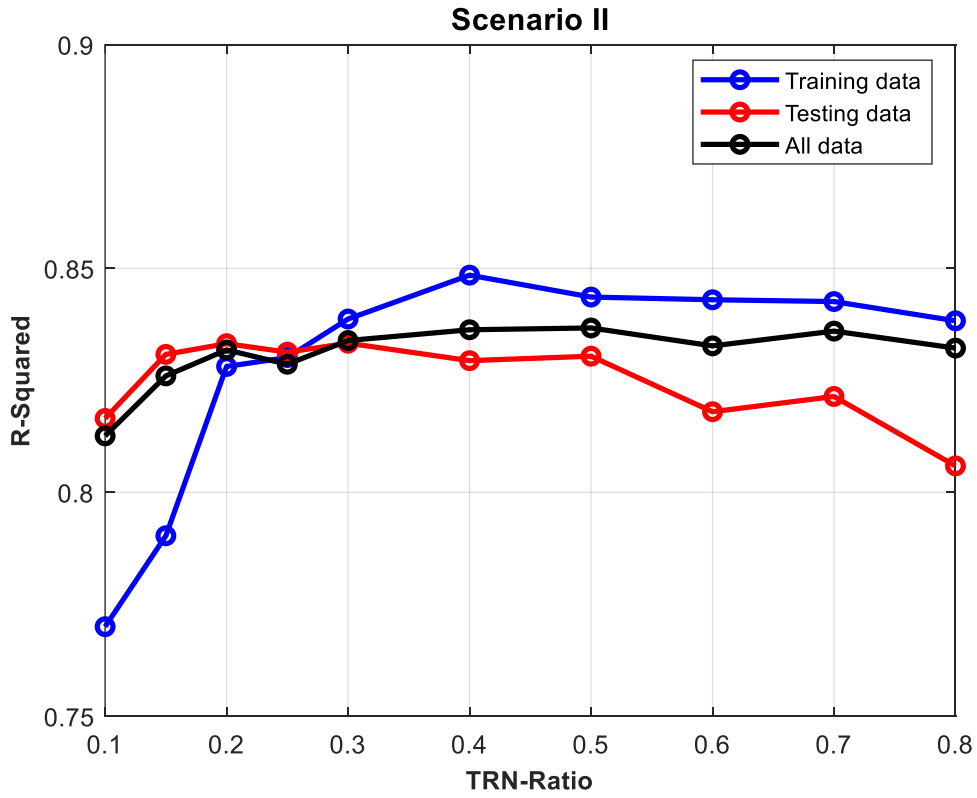

Figure 7. The impact of the data ratio used in the training of the model on the model's performance.

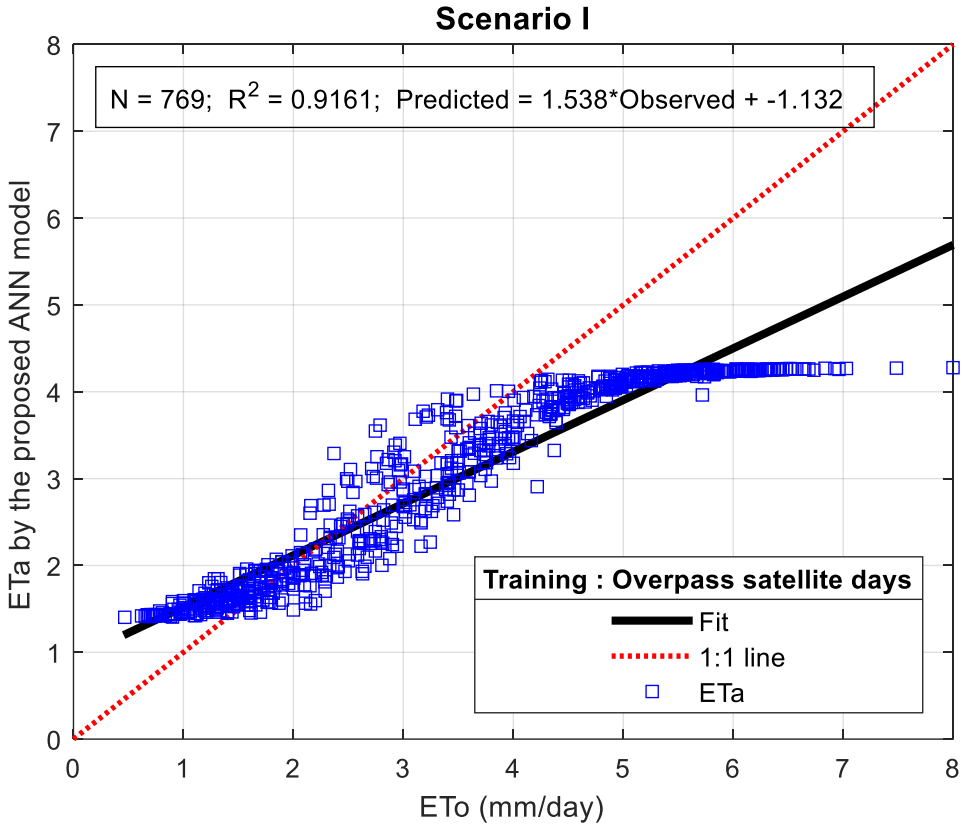

Figure 8. *Cont.*

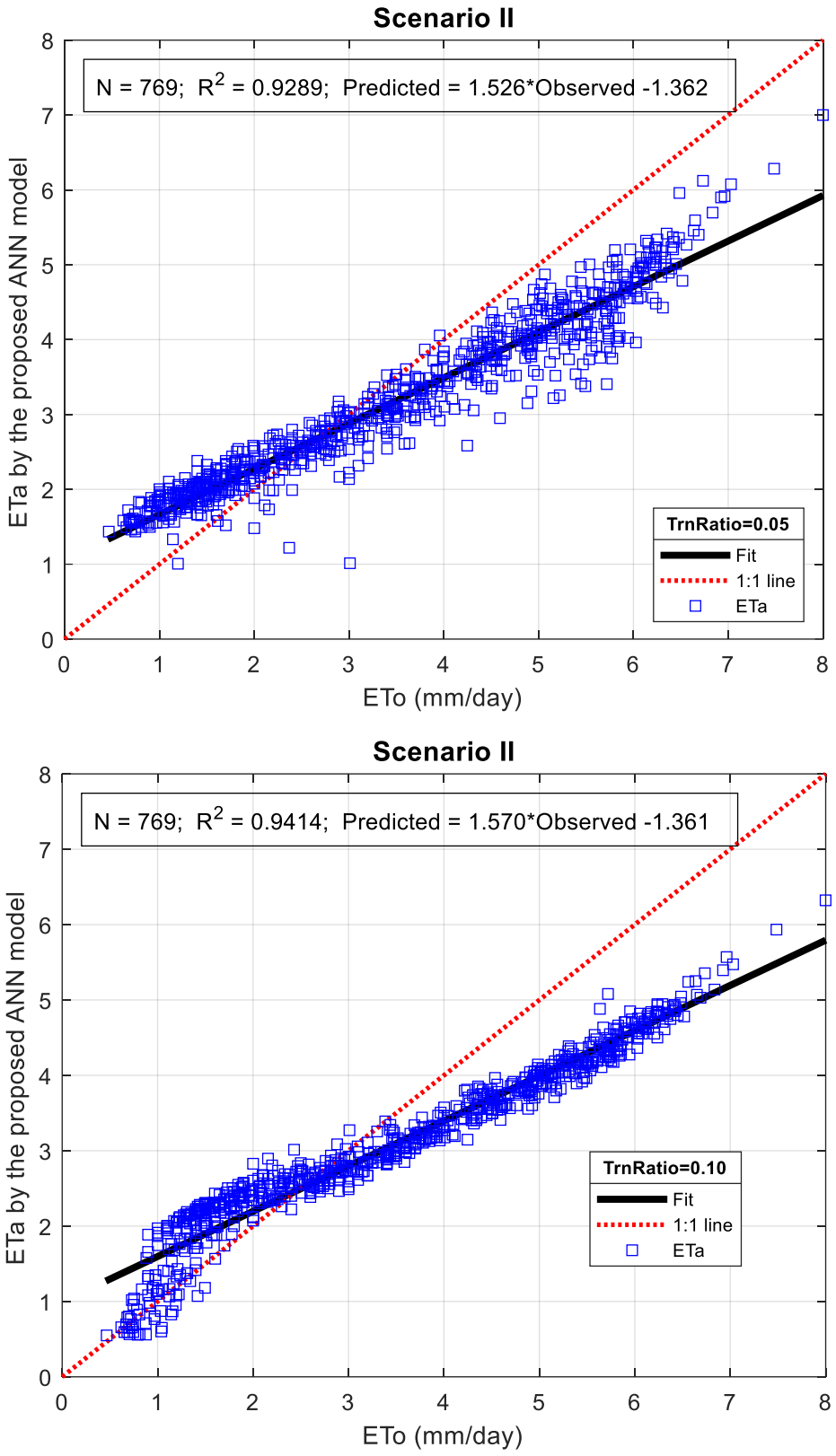

Figure 8. Relationship between ETo by the *FAO–Penman–Monteith* approach and *ETa* by the *ANN model* for Scenario I and II in the Akarsu Irrigation District, i.e., the AID, located in the Eastern Mediterranean Region of Turkiye.

In this study, ETo by FAO–Penman–Monteith method was calculated using the meteorological parameters from two ground-based observations (L8 and Cotlu meteorological stations shown in Figure 1) established in the AID. In addition, the estimated ETa by the ANN model for two scenarios was validated to the reference ET by using a simple linear regression approach (ETa = a + b*ETo) over the AID. As can be seen in Figure 8, the results demonstrate that the R^2 value is 0.9161 for Scenario I if we use the whole dataset (N = 769) for training. On the other hand, in Scenario II with TRNRatios of 0.05 and 0.10, R^2 values reach 0.9289 and 0.9414, respectively. In the Cukurova region, where our study is based, several studies have been conducted to estimate ETa by direct methods [7]. For example, Cetin et al. [7] compared the estimated ETa-METRIC of the second crop of soybean to the ETa-lysimeter. They concluded that the ETa-METRIC values were more closely aligned with the lysimeter observations. Additionally, as indicated by the findings reported by Cetin et al. [7], there is a good agreement between ETc values for the crops of citrus, wheat, potato, lettuce, first-crop corn, and peanut, as well as second-crop soybean, and ETa by the MERIC model during 2021 water year. All these could lead us to conclude that the estimated ETa results by the METRIC model, as well as the proposed ANN model, were consistent with the ETa values by direct and indirect techniques in the same study region in both this research and prior studies.

5. Conclusions

This study presents a novel approach that allows daily ETa estimation using a new ANN model in a large-scale irrigation scheme using limited climatic data and MODIS satellite data. The study demonstrated that the daily ETa predicted values are comparable with the ETa which is estimated by the RS-based surface energy balance models. Furthermore, this methodology is the first attempt to generate a new model using the ANN algorithm as an alternative to the existing methods of ETa estimation in a large-scale irrigation district in the Eastern Mediterranean Region of Turkiye. The daily ETa estimation results of the ANN model in the two implemented scenarios showed a reasonable degree of agreement with ETa values computed by the METRIC model. Thus, this work can be considered a significant contribution to obtaining reliable ETa results without the need for lengthy and labor-intensive processes for complex equations, as in RS-based surface energy balance models. This study revealed that the proposed ANN model is a powerful tool for estimating daily ETa values in large-scale irrigation schemes using limited meteorological observations and some of the parameters of remote sensing in arid and semi-arid regions, as well as in different climate regions and zones of the world. Moreover, the proposed methodology in this research may help water authorities, practitioners, and end users estimate actual evapotranspiration for sustainable water management, especially in Mediterranean countries, where freshwater resources are scarce in an increasingly globalized world.

Author Contributions: Conceptualization, H.K.; Methodology, H.K., M.C. and O.A.; Software, H.K.; Validation, H.K.; Formal analysis, M.E.C. and O.A.; Investigation, H.K. and M.C.; Resources, O.A.; Data curation, M.C., M.E.C. and O.A.; Writing—original draft, H.K., M.C. and O.A.; Writing—review & editing, H.K. and O.A.; Supervision, H.K.; Project administration, H.K. and M.E.C. All authors have read and agreed to the published version of the manuscript.

Funding: The authors thank the Scientific and Technological Research Council of Turkiye [TUBITAK, Project number: 122Y007] and the Scientific Research Projects (BAP) Coordination Unit of Cukurova University [Project Number: FDK-2022-14907] for obtaining financial support for this work.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available on request from the authors.

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

References

1. Kumar, M.; Raghuwanshi, N.S.; Singh, R. Artificial Neural Networks Approach in Evapotranspiration Modeling: A Review. *Irrig. Sci.* **2011**, *29*, 11–25. [CrossRef]
2. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration Guidelines for Computing Crop Water Requirements*; FAO Irrigation and Drainage Paper 56; FAO: Rome, Italy, 1998.
3. Rawat, K.S.; Bala, A.; Singh, S.K.; Pal, R.K. Quantification of Wheat Crop Evapotranspiration and Mapping: A Case Study from Bhiwani District of Haryana, India. *Agric. Water Manag.* **2017**, *187*, 200–209. [CrossRef]
4. Gharbia, S.S.; Smullen, T.; Gill, L.; Johnston, P.; Pilla, F. Spatially Distributed Potential Evapotranspiration Modeling and Climate Projections. *Sci. Total Environ.* **2018**, *633*, 571–592. [CrossRef] [PubMed]
5. Alsenjar, O.; Cetin, M.; Aksu, H.; Akgul, M.A.; Golpinar, M.S. Cropping Pattern Classification Using Artificial Neural Networks and Evapotranspiration Estimation in the Eastern Mediterranean Region of Turkey. *J. Agric. Sci.* **2023**, *29*, 677–689. [CrossRef]
6. Zhang, X.C.; Wu, J.W.; Wu, H.Y.; Li, Y. Simplified SEBAL Method for Estimating Vast Areal Evapotranspiration with MODIS Data. *Water Sci. Eng.* **2011**, *4*, 24–35. [CrossRef]
7. Cetin, M.; Alsenjar, O.; Aksu, H.; Golpinar, M.S.; Akgul, M.A. Comparing Actual Evapotranspiration Estimations by METRIC to In-situ Water Balance Measurements Over an Irrigated Field in Turkey. *Hydrol. Sci. J.* **2023**, *68*, 1162–1183. [CrossRef]
8. Bastiaanssen, W.G.M.; Menenti, M.; Feddes, R.A.; Holtslag, A.A.M. A Remote Sensing Surface Energy Balance Algorithm for Land (SEBAL): 1. Formulation. *J. Hydrol.* **1998**, *212–213*, 198–212. [CrossRef]
9. Bastiaanssen, W.G.M.; Pelgrum, H.; Wang, J.; Ma, Y.; Moreno, J.F.; Roerink, G.J.; van der Wal, T. A Surface Energy Balance Algorithm For Land (SEBAL): 2. Validation. *J. Hydrol.* **1998**, *212–213*, 213–229. [CrossRef]
10. Allen, R.; Tasumi, M.; Morse, A.; Trezza, R.; Wright, J.L.; Bastiaanssen, W.; Kramber, W.; Lorite, L.; Robison, C.W. Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration, METRIC (applications). *J. Irrig. Drain. Eng.* **2007**, *133*, 395–406. [CrossRef]
11. Allen, R.G.; Tasumi, M.; Trezza, R. Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration, METRIC (model). *J. Irrig. Drain. Eng.* **2007**, *133*, 380–394. [CrossRef]
12. Su, Z. The Surface Energy Balance System (SEBS) for Estimation of Turbulent Heat Fluxes. *Hydrol. Earth Syst. Sci.* **2002**, *6*, 85–100. [CrossRef]
13. Bachour, R.; Walker, W.; Ticlavilca, A.; McKee, M.; Maslova, I. Estimation of Spatially Distributed Evapotranspiration Using Remote Sensing and A Relevance Vector Machine. *J. Irrig. Drain. Eng.* **2014**, *140*, 04014029. [CrossRef]
14. Singh, R.K.; Irmak, A. Treatment of Anchor Pixels in the METRIC Model for Improved Estimation of Sensible and Latent Heat Fluxes. *Hydrol. Sci. J.* **2011**, *56*, 895–906. [CrossRef]
15. ASCE Task Committee on Application of Neural Networks in Hydrology. Artificial Neural Network in Hydrology. I: Preliminary Concepts. *J. Hydrol. Eng.* **2000**, *5*, 115–123. [CrossRef]
16. ASCE Task Committee on Application of Neural Networks in Hydrology. Artificial Neural Network in Hydrology. II: Hydrologic Application. *J. Hydrol. Eng.* **2000**, *5*, 124–137. [CrossRef]
17. Bruton, J.M.; McClendon, R.W.; Hoogenboom, G. Estimating Daily Pan Evaporation with Artificial Neural Networks. *Trans.* **2000**, *43*, 491–496. [CrossRef]
18. Odhiambo, L.O.; Yoder, R.E.; Yoder, D.C.; Hines, J.W. Optimization of Fuzzy Evapotranspiration Model Through Neural Training with Input-Output Examples. *Trans. ASAE* **2001**, *44*, 1625–1633. [CrossRef]
19. Kumar, M.; Raghuwanshi, N.; Singh, R.; Wallender, W.W.; Pruitt, W.O. Estimating Evapotranspiration Using Artificial Neural Network. *J. Irrig. Drain. Eng.* **2002**, *128*, 224–233. [CrossRef]
20. Dehbozorgi, F.; Sepaskhah, A.R. Comparison of Artificial Neural Networks and Prediction Models for Reference Evapotranspiration Estimation in a Semi-Arid Region. *Arch. Agron. Soil. Sci.* **2011**, *58*, 477–497. [CrossRef]
21. Khoshhal, J.; Mokarram, M. Model for Prediction of Evapotranspiration Using MLP Neural Network. *Inter. J. Environ. Sci.* **2012**, *3*, 1000–1009.
22. Abrishami, N.; Sepaskhah, A.R.; Shahrokhnia, M.H. Estimating Wheat and Maize Daily Evapotranspiration Using Artificial Neural Network. *Theoret Appl. Climatol.* **2019**, *135*, 945–958. [CrossRef]
23. Yamac, S.S.; Todorovic, M. Estimation of Daily Potato Crop Evapotranspiration Using Three Different Machine Learning Algorithms and Four Scenarios of Available Meteorological Data. *Agric. Water Manag.* **2020**, *228*, 105875. [CrossRef]
24. Virnodkar, S.S.; Pachghare, V.K.; Patil, V.C.; Jha, S.K. Application of machine learning on remote sensing data for sugarcane crop classification: A review. In *ICT Analysis and Applications. Lecture Notes in Networks and Systems*; Fong, S., Dey, N., Joshi, A., Eds.; Springer: Singapore, 2020; Volume 93. [CrossRef]
25. Olmedo, G.F.; Ortega-farias, S.; Fonseca-luengo, D. Tools and Functions to Estimate Actual Evapotranspiration Using Land Water: Toolsand Functions to Estimate Actual Evapotranspiration Using Land Surface Energy Balance Models. *R J.* **2016**, *8*, 352. [CrossRef]
26. Bhattarai, N.; Quackenbush, L.J.; Im, J.; Shaw, S.B. A New Optimized Algorithm for Automating Endmember Pixel Selection in the SEBAL and METRIC Models. *Remote Sens. Environ.* **2017**, *196*, 178–192. [CrossRef]
27. Jain, S.K.; Nayak, P.C.; Sudheer, K.P. Models for Estimating Evapotranspiration Using Artificial Neural Networks, and Their Physical Interpretation. *Hydrol. Process.* **2008**, *22*, 2225–2234. [CrossRef]

28. Coppola, E., Jr.; Szidarovszky, F.; Poulton, M.; Charles, E. Artificial Neural Network Approach for Predicting Transient Water Levels in a Multilayered Groundwater System Under Variable State, Pumping, and Climate Conditions. *J. Hydrol. Eng.* **2003**, *8*, 348–360. [CrossRef]
29. Daliakopoulos, I.N.; Coulibaly, P.; Tsanis, I.K. Groundwater Level Forecasting Using Artificial Neural Networks. *J. Hydrol.* **2005**, *309*, 229–240. [CrossRef]
30. Karahan, H.; Iplikci, S.; Yasar, M.; Gurarslan, G. River Flow Estimation from Upstream Flow Records Using Support Vector Machines. *J. Appl. Math.* **2014**, *2014*, 714213. [CrossRef]
31. Luk, K.C.; Ball, J.E.; Sharma, A. A Study of Optimal Model Lag and Spatial Inputs to Artificial Neural Network for Rainfall Forecasting. *J. Hydrol.* **2000**, *227*, 56–65. [CrossRef]
32. Garcia, L.A.; Shigidi, A. Using Neural Networks for Parameter Estimation in Groundwater. *J. Hydrol.* **2006**, *318*, 215–231. [CrossRef]
33. Granata, F. Evapotranspiration Evaluation Models Based on Machine Learning Algorithms: A Comparative Study. *Agric. Water Manag.* **2019**, *217*, 303–315. [CrossRef]
34. Tikhamarine, Y.; Malik, A.; Kumar, A.; Souag-Gamane, D.; Kisi, O. Estimation of Monthly Reference Evapotranspiration Using Novel Hybrid Machine Learning Approaches. *Hydrol. Sci. J.* **2019**, *64*, 1824–1842. [CrossRef]
35. Madugundu, R.; Al-Gaadi, K.A.; Tola, E.; Hassaballa, A.A.; Patil, V.C. Performance of the METRIC Model in Estimating Evapotranspiration Fluxes Over an Irrigated Field in Saudi Arabia Using Landsat-8 Images. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6135–6151. [CrossRef]
36. Tang, D.; Feng, Y.; Gong, D.; Hao, W.; Cui, N. Evaluation of Artificial Intelligence Models for Actual Crop Evapotranspiration Modeling in Mulched and Non-mulched Maize Croplands. *Comput. Electron. Agric.* **2018**, *152*, 375–384. [CrossRef]
37. Karahan, H.; Ayvaz, M.T. Simultaneous Parameter Identification of a Heterogeneous Aquifer System Using Artificial Neural Networks. *Hydrogeol. J.* **2008**, *16*, 817–827. [CrossRef]
38. Maier, H.R.; Dandy, G.C. Neural Networks for the Prediction and Forecasting of Water Resources Variables: A Review of Modelling Issues and Applications. *Environ. Model. Softw.* **2000**, *15*, 101–124. [CrossRef]
39. Antonopoulos, V.Z.; Antonopoulos, A.V. Daily Reference Evapotranspiration Estimates by Artificial Neural Networks Technique and Empirical Equations Using Limited Input Climate. *Comput. Electron. Agric.* **2017**, *132*, 86–96. [CrossRef]
40. Ferreira, L.B.; Cunha, F.F. New Approach to Estimate Daily Reference Evapotranspiration Based on Hourly Temperature and Relative Humidity Using Machine Learning and Deep Learning. *Agric. Water Manag.* **2020**, *234*, 106–113. [CrossRef]
41. Mattar, M.A. Using Gene Expression Programming in Monthly Reference Evapotranspiration Modeling: A Case Study in Egypt. *Agric. Water Manag.* **2018**, *198*, 28–38. [CrossRef]
42. Santos, P.A.B.; Schwerz, F.; Carvalho, L.G.; Baptista, V.B.S.; Marin, D.B.; Ferraz, G.A.S.; Rossi, G.; Conti, L.; Bambi, G. Machine Learning and Conventional Methods for Reference Evapotranspiration Estimation Using Limited-Climatic-Data Scenarios. *Agronomy* **2023**, *13*, 2366. [CrossRef]
43. Chen, X.; Yu, S.; Zhang, H.; Li, F.; Liang, C.; Wang, Z. Estimating the Actual Evapotranspiration Using Remote Sensing and SEBAL Model in an Arid Environment of Northwest China. *Water* **2023**, *15*, 1555. [CrossRef]
44. Abdel-Fattah, M.K.; Kotb Abd-Elmabod, S.; Zhang, Z.; Merwad, A.R.M.A. Exploring the Applicability of Regression Models and Artificial Neural Networks for Calculating Reference Evapotranspiration in Arid Regions. *Sustainability* **2023**, *15*, 15494. [CrossRef]