

Original scientific paper

UDC 004.85(65)
https://doi.org/to_be_assigned

Received: February 9, 2026

Corrected: May 14, 2026

Accepted: June 22, 2026

Khadidja Gherbi^{1*}, Abdelhadi Ammari*, Mohamed Meddi*

** École Nationale Supérieure d'Hydraulique (ENSH). LGEE laboratory. Blida, Algeria*

MACHINE LEARNING ESTIMATION OF DAILY ETo IN NORTHERN ALGERIA

Abstract: Accurate estimation of reference evapotranspiration (ET_o) is essential for efficient water resource management, particularly in agriculture, where precise irrigation planning is critical for sustainability. A correct estimate allows for the precise amount of water needed for crops, preventing waste and overexploitation of resources, and improving irrigation efficiency and water conservation in the long term. This study evaluates four machine learning models such as Support Vector Regression (SVR), Random Forest (RF), Multilayer Perceptron (MLP), and Long Short-Term Memory (LSTM) for predicting ET_o. Model performance was assessed against ET_o values computed using the Penman–Monteith (PM) and Hargreaves Samani (HS) equations, based on NASA POWER reanalysis data and CMIP6 climate datasets. Both data sources were validated against observed meteorological measurements from three weather stations as well as Soumaa, Boukourdane and Mouzaia in northern Algeria, where the results demonstrate a strong correlation between predicted and observed ET_o values, with SVR and RF consistently achieving high accuracy. The Soumaa station, using the NASA POWER-based model, yielded the best performance ($R^2 = 0.92$, MAE = 0.06 mm/day, RMSE = 0.08 mm/day), whereas the Boukourdane station exhibited larger prediction discrepancies. CMIP6-based models also produced robust estimates under varying climatic conditions. The CNRM and IPSL models showed the highest temperature agreement at Boukourdane ($r = 0.8649$) and minimal ET_o bias at Soumaa, while Mouzaia exhibited the largest underestimation using IPSL (−7.3344 mm/day). While NASA POWER data provide superior short-term accuracy, CMIP6-based approaches offer significant value for long-term climate projections. Random Forest and MLP consistently achieved nearly perfect agreement with observed ET_o at Boukourdane and Soumaa, whereas performance decreased at Mouzaia due to stronger local climatic variability and terrain influence. However, the findings fundamentally underscore that integrating machine learning models with climate projections is pivotal for advancing irrigation decision-making and promoting sustainable water resource management under changing climatic conditions.

Keywords: evapotranspiration, CMIP6, NASA POWER, water resource management, irrigation scheduling, bias correction, spatial analysis, prediction models

¹ k.gherbi@ensh.dz (corresponding author)

Khadidja Gherbi (<http://orcid.org/0000-0002-6800-6199>)

Abdelhadi Ammari (<http://orcid.org/0000-0001-7197-5962>)

Mohamed Meddi (<http://orcid.org/0000-0002-9772-7366>)

Introduction

Agriculture in the Mediterranean basin, particularly along the southern coastline, encounters significant challenges attributable to water scarcity and the inequitable distribution of water resources, challenges that are intensified by the ramifications of climate change (Claro et al., 2024). In the context of Algeria, these pressures manifest with particular severity. Along the same lines, the IPCC's sixth assessment report highlights that alterations in the hydrological cycle including changes in rainfall patterns and increased evapotranspiration are expected to intensify water scarcity globally, with Mediterranean regions identified as particularly vulnerable areas (Meddi and Eslamian, 2021; McCarthy et al., 2025). The FAO points out that deteriorating land, high water utilization, and growing water stress are steering agriculture toward unsustainable directions. These developments stress the immediate requirement for the deployment of sustainable water management frameworks and the acceptance of climate-resilient agricultural techniques to guarantee food security (FAO, 2021).

In Algeria, agricultural activities account for 55% of mobilized water resources. However, prevailing irrigation practices often exhibit inconsistencies and fail to adapt to contemporary climatic conditions, resulting in substantial resource wastage (Bouchentouf & Benabdeli, 2021). These challenges stem primarily from inadequate monitoring infrastructure and limited access to reliable meteorological data, which impede precise determination of crop water requirements. Consequently, enhancing the accuracy of evapotranspiration estimates is essential for optimizing irrigation scheduling and improving overall water use efficiency (Sun et al., 2024). The rising scarcity of water resources and the gradual deterioration of groundwater quality highlight the need of precise reference evapotranspiration (ET_0) measurement, which is essential for irrigation planning and agricultural water use optimization. A study by Seraiche et al. (2026) combined hydrochemical assessment with spatial interpolation methods to map groundwater pollution south of Chott Hodna, highlighting the value of geographically based indicators for identifying vulnerable water-resource zones.

Recent studies conducted in Algeria have highlighted the growing importance of spatially based hydroclimatic analyses for addressing water-related challenges across diverse environmental settings. In the Algerian Sahara, Ben Ferhat and Saidou (2025) showed that climatic drought in the Tidikelt region is closely associated with water-balance deficits, increased soil salinity, and the progressive degradation of oasis agriculture. Focusing on evapotranspiration, Bouregaa et al. (2025) analyzed climatic trends and reference evapotranspiration (ET_0) patterns across major cereal-growing areas using AgERA5 data and machine-learning techniques, revealing pronounced spatial differences in ET_0 responses to changing climatic conditions. Similarly, Saci et al. (2025) assessed the performance of several models for monthly ET_0 estimation in the Kebir-West watershed and demonstrated the potential of data-driven methods to improve evapotranspiration predictions where meteorological observations are limited. Collectively, these findings emphasize the need for locally adapted and spatially explicit approaches to support water-resource assessment and management in Algeria, particularly in regions affected by data scarcity and increasing hydroclimatic pressures.

The term reference evapotranspiration (ET_0) refers to the rate of water transfer from a hypothetical reference crop and is vital for effective irrigation planning and hydrological

assessments. The FAO-56 Penman–Monteith (PM) equation is considered the authoritative method for assessing ETo due to its robust physical framework and accuracy (Allen et al., 2001). However, employing the PM equation requires a variety of meteorological data, including solar radiation, air temperature, humidity, and wind velocity, which are challenging to obtain in areas identified for data insufficiency (Aly et al., 2023).

In response to the limitations of data availability, alternative methodologies have been devised. The Hargreaves–Samani (HS) model offers a basic technique that hinges entirely on temperature measurements, thus proving beneficial for locales with restricted meteorological information (Dohler et al., 2020). Moreover, remote sensing and reanalysis datasets, such as NASA POWER, furnish spatially continuous meteorological data that can facilitate ETo estimation in gauged areas (Rodrigues & Braga, 2021). Additionally, climate model outputs derived from the Coupled Model Intercomparison Project Phase 6 (CMIP6) provide invaluable projections for evaluating future water availability across diverse climate scenarios (Al-Hasani et al., 2023).

In contemporary scholarly discourse, machine learning (ML) methodologies have emerged as powerful instruments for the prediction of reference evapotranspiration (ETo), providing both adaptability and high precision without requiring explicit physical interdependencies (Sharafi & Ghaleni, 2024). Numerous algorithms, such as Support Vector Regression (SVR), Random Forest (RF), Multilayer Perceptron (MLP), and Long Short-Term Memory (LSTM) networks, have revealed impressive capabilities in uncovering intricate nonlinear links between atmospheric parameters and ETo (Elbeltagi et al., 2020; Roy et al., 2022; Pereira et al., 2020). These models can be trained on historical datasets and subsequently used to forecast ETo across a spectrum of climatic conditions, making them especially advantageous for locales characterized by climatic heterogeneity. The selection of multiple machine learning models was motivated by their complementary strengths, which enable the modeling of both static nonlinear relationships and temporal dependencies. Notably, Support Vector Regression and Random Forest have demonstrated superior performance in capturing nonlinear and static relationships for ETo estimation (Kumar, 2023), while MLP and LSTM are particularly effective in modeling complex patterns and long-term temporal dynamics in climate time series (Roy et al., 2022). Recent studies also highlight that combining different ML paradigms improves the robustness and accuracy of ETo estimation under diverse climatic conditions (Farag, 2025). Thus, our present approach constitutes a comparative and evaluative framework which aim to demonstrate the strengths and limitations of each methodology.

This strategy holds particular relevance, given the scarcity of prior research that has concurrently highlighted multiple models utilizing both reanalysis datasets and climate projections tailored to northern Algeria.

This geographic region is distinguished by a Mediterranean climate that exhibits considerable spatial and temporal heterogeneity in precipitation and temperature, thereby rendering precise ETo estimation particularly formidable (Achite et al., 2021).

The objective of this investigation is to address this deficit by scrutinizing the performance of four ML models (SVR, RF, MLP, and LSTM) in the prediction of daily ETo in northern Algeria. The models will be trained and validated using ETo values obtained from the Penman-Monteith (PM) and Hargreaves-Samani (HS) equations, based on NASA POWER

reanalysis data and the Coupled Model Intercomparison Project Phase 6 (CMIP6) climate datasets. The specific aims are to: (1) evaluate the accuracy of NASA POWER and CMIP6 data in replicating observed meteorological variables; (2) compare the predictive performance of various ML models in estimating ETo; and (3) assess the appropriateness of these methodologies for short-term irrigation planning and long-term climate adaptation frameworks. Our regions are facing increasing water scarcity, driven both by a decline in atmospheric inputs since the mid-1970s and by the continuously growing water demand. This demand is particularly pronounced for domestic water supply and, more importantly, for irrigation, which accounts for more than half of the available water resources (Mezouane et al., 2025).

The original contribution of this paper can be considered as a primary investigation concerning the re-examination of local-scale meteorology using NASA POWER and CMIP6 data, a machine learning framework for ETo construction that compares several models including MLP, RF, LSTM, and SVR to reflecting low-dimensional functional relations between input features (e.g., temperature parameter values); combining data-driven and empirical temperature-based approaches within a single framework; and directly relating model performance to irrigation planning and climate change adaptation in semi-arid Mediterranean conditions.

Material and Methods

Study Area

The study zone chosen is the Mitidja Plain, located in northern Algeria (Figure 1). This region is one of the most important agricultural plains in Algeria, characterized by intensive cultivation of cereals, vegetables, and fruit trees especially citrus. It covers an area of 1,450 km² spanning 100 km in length, with a width varying from 8 to 18 km. The climate is Mediterranean, characterized by hot, dry summers and mild, wet winters. Mean annual precipitation ranges from 600 to 800 mm, with most rainfall occurring between October and April. Mean annual temperature varies from 15°C to 18°C, with summer temperatures frequently exceeding 30°C. The Mitidja plain features a Mediterranean climate (Csa), consistent with the Köppen Csa classification (Bellout et al., 2020). We focused in the present study on the west part of the Mitidja plain which remain a large agricultural area, more than the east part more urbanized (Figure. 1).

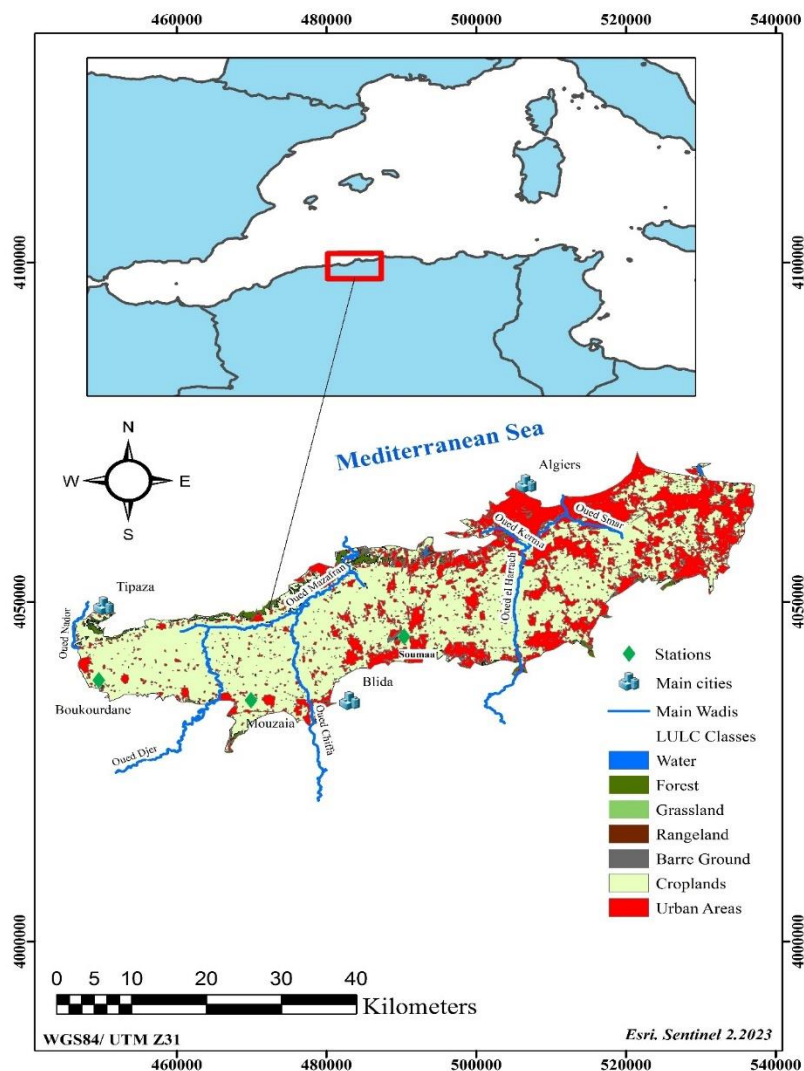


Fig. 1. Geographical location of the Mitidja Plain, northwest Algeria

Three meteorological sites were selected for this study: Boukourdane (west part), Mouzaia (central part), and Soumaa (east part of the considered study zone). These sites are operated by the National Water resources Agency ANRH and provide daily records of temperature, precipitation, humidity, wind speed, and solar radiation. The sites were chosen to represent different microclimatic conditions within the Mitidja Plain, with Mouzaia located at a higher elevation and experiencing greater climatic variability compared to Boukourdane and Soumaa. The Mitidja Plain has a limited number of meteorological stations due to insufficient ground-based observational networks in the region. Three stations had continuous and reliable records for long periods of time. To mitigate this constraint, the present study employs the longest continuous meteorological sites records, reinforced by validated gridded datasets, to ensure spatial consistency.

The three meteorological stations selected for this study were chosen primarily based on the quality of their available data. They have sufficiently continuous, complete, and reliable daily records to allow for the calculation of reference evapotranspiration. Although the number of stations remains limited, the Mitidja Plain, within the study area, exhibits relatively homogeneous climatic conditions, typical of a Mediterranean context. Altitude differences are small, and precipitation does not show significant spatial contrasts. For this reason, the selected stations can be considered representative of the general climatic conditions of the plain.

This methodological choice is consistent with previous work that favors the use of reliable and controlled observed time series, particularly when local meteorological networks are sparse (Allen et al., 1998; Ferreira et al., 2019; Granata, 2019). To mitigate the impact of this constraint and improve the spatial and temporal coverage of the analysis, the data measured at the stations were supplemented with NASA POWER reanalysis data and CMIP6 climate outputs. These datasets provide continuous meteorological information in space and time and are now widely used in hydroclimatic studies and in work on estimating ETo (Eyring et al., 2016; Rodrigues & Braga, 2021).

Data Sources

Observed Meteorological Data

Daily meteorological data for the period in Table 1 – were obtained from the three ANRH stations. The variables included maximum temperature ($T_{\max}$), minimum temperature ($T_{\min}$), mean temperature (T_{mean}), relative humidity (RH), wind speed at 2 m height (u_2), and solar radiation (R_s). Data quality control was performed to identify and remove outliers and missing values. Gaps in the data were filled using linear interpolation for short periods (≤ 3 days) or excluded from the analysis for longer gaps.

Table 1. Characteristics of the meteorological stations and data sources used in this study

Station	Latitude (°N)	Longitude (°E)	Elevation (m)	Data Sources	Periods
Mouzaia	36.467	2.689	94	ANRH, NASA POWER, CMIP6	ANRH: [1988–2008]; NASA POWER: 1981–2020; CMIP6: 2021–2100
Soumaa	36.5184	2.9084	177	ANRH, NASA POWER, CMIP6	ANRH: [1991–2010]; NASA POWER: 1981–2020; CMIP6: 2021–2100
Bou- kourdane	36.5395	2.3080	110	ANRH, NASA POWER, CMIP6	ANRH: [2001–2012]; NASA POWER: 1981–2020; CMIP6: 2021–2100

In the present study, the length of the available time series was constrained by the inherent challenges of data availability and the continuity of reliable meteorological observations in the west part of Mitidja Plain. In North Africa, long-term homogeneous datasets are frequently limited by discontinuities in sites records, changes in instrumentation, and an important historical measurement gap (Ba-ichou et al., 2024b; Mustapha et al., 2025).

The main challenge for this study is to address these limitations, by developing a robust machine learning framework able to estimate accurately reference evapotranspiration even under conditions of limited ground-based data (Ba-ichou et al., 2024a; Shafizadeh-Moghadam et al., 2023). This is achieved by increasing local station records with gridded NASA POWER meteorological data to ensure spatial and temporal consistency (Mustapha et al., 2025). Furthermore, by integrating CMIP6 climate model projections, the proposed approach extends beyond historical analysis to provide reliable ETo forecasting for future climate scenarios (Mokhtar et al., 2024). This methodology not only mitigates the challenges posed by sparse field observational networks but also provides a more comprehensive and resilient framework for hydrological modeling and agricultural water management in data-scarce regions (Hendy et al., 2023; Mustapha et al., 2025).

NASA POWER Reanalysis Data

NASA POWER (Prediction of Worldwide Energy Resources) provides global meteorological data derived from satellite observations and atmospheric reanalysis models (NASA, 2023). Daily data for the same variables and period were extracted for the coordinates of the three stations using the NASA POWER API. The spatial resolution of NASA POWER data is $0.5^\circ \times 0.625^\circ$, which corresponds to approximately 50 km at the study latitude.

CMIP6 Climate Model Data

Climate projections were obtained from three CMIP6 models: CNRM-CM6-1, IPSL-CM6A-LR, and MPI-ESM1-2-HR. These models were selected based on their performance in simulating Mediterranean climate and their availability in the CMIP6 archive (Daher & Kirtman, 2025). Daily data for temperature, precipitation, humidity, wind speed, and solar radiation were extracted for the historical period and future scenarios (2015–2100) under the SSP2-4.5 and SSP5-8.5 pathways.

While historical climate studies typically require multi-decadal datasets to capture long-term variability, the present research deliberately utilizes a more focused temporal period to address a critical challenge in modern hydrology: the development of high-accuracy models in data-constrained regions (Amani et al., 2024). In many areas, including the Mitidja Plain, while long-term records may exist, they often suffer from significant discontinuities, changes in sensor technology, or missing variables required for the standard FAO-56 Penman-Monteith equation (Ba-ichou et al., 2024b; Shafizadeh-Moghadam et al., 2023).

Therefore, the objective is not to provide a century-long historical reconstruction, but rather to demonstrate a transferable methodology that enables precise ETo estimation and forecasting even when researchers are restricted to limited ground-based observations (Ba-ichou et al., 2024b, 2024a). This approach emphasizes the utility of combining diverse data sources, such as reanalysis products and global climate model outputs, to overcome observational limitations and enhance predictive capabilities for hydrological modeling (Bjarke et al., 2023; Hamouda et al., 2021).

It is important to note that the quantile mapping method was used to correct systematic biases in the datasets. Quantile mapping, which aligns the cumulative distribution functions of the simulated variables with those generated from meteorological data observed during the calibration period, was also used to correct biases in the CMIP6 and NASA POWER variables. This correction was applied independently for each station and each climate variable before the ETo calculation, thus minimizing systematic biases in the model while preserving the relative temporal dynamics of future climate projections.

Model Training and Validation

The categorization of data sets included three parts: training (70%), validation (15%), and testing (15%). The training subset used for model fitting, the validation subset was engaged in hyperparameter tuning, and the testing subset was allocated for a full performance review. The inputs compound daily meteorological variables ($T_{\max}$, $T_{\min}$, RH, U_2 , R_s), while the dependent variable is ET_0 , which was derived from the Penman-Monteith equation.

Data preprocessing necessitated the alteration of the input features to establish a zero mean while ensuring the variance equals one. In the field of Long Short-Term Memory (LSTM) systems, batches of successive days were artfully engineered to uncover temporal connections. Hyperparameter optimization was conducted utilizing a 5-fold cross-validation methodology applied to the training subset.

Reference Evapotranspiration Calculation

Penman–Monteith Equation

The FAO-56 Penman–Monteith equation was used to calculate daily ET_0 (Thiemeßl et al., 2012):

$$ET_0 = \frac{0.408\Delta(R_s - G) + \gamma \frac{900}{T_{mean} + 273} U_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

where ET_0 is reference evapotranspiration (mm/day), R_n is net radiation (MJ/m²/day), G is soil heat flux (MJ/m²/day), T is mean air temperature (°C), u_2 is wind speed at 2 m height (m/s), e_s is saturation vapor pressure (kPa), e_a is actual vapor pressure (kPa), Δ is the slope of the vapor pressure curve (kPa/°C), and γ is the psychrometric constant (kPa/°C).

Standardized evapotranspiration equation by ASCE-EWRI

For standardized PM- ET_0 estimation, the site characteristics such as latitude and elevation must be input, along with daily weather data (Snyder & Eching, 2002). Dew point temperatures are calculated using:

$$T_d = T_{mean} - \frac{100 - R_h}{5} \quad (2)$$

Where R_h is relative humidity (%).

Hargreaves–Samani Equation

The Hargreaves–Samani equation provides a simplified alternative requiring only temperature data (Hargreaves & Samani, 1985):

$$ET_{0-HS} = 0.0023(T_{moyenne} + 17.8) (T_{\max} - T_{\min})^{0.5} R_a \quad (3)$$

where R_a is extraterrestrial radiation (mm/day), calculated based on latitude and day of year.

Machine Learning Models

Four machine learning models were implemented and evaluated for ETo prediction:

Support Vector Regression (SVR)

SVR is a kernel-based method that maps input data into a high-dimensional feature space where linear regression is performed (Smola & Schölkopf, 2004). The radial basis function (RBF) kernel was used, with hyperparameters (C , γ , ϵ) optimized through grid search and cross-validation.

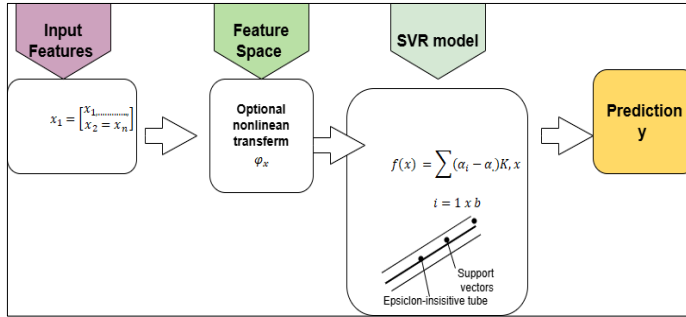


Fig. 2. General Structure of SVR

Random Forest (RF)

RF is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction (Breiman, 2001). The number of trees ($n_estimators$), maximum depth (max_depth), and minimum samples per leaf ($min_samples_leaf$) were optimized using cross-validation.

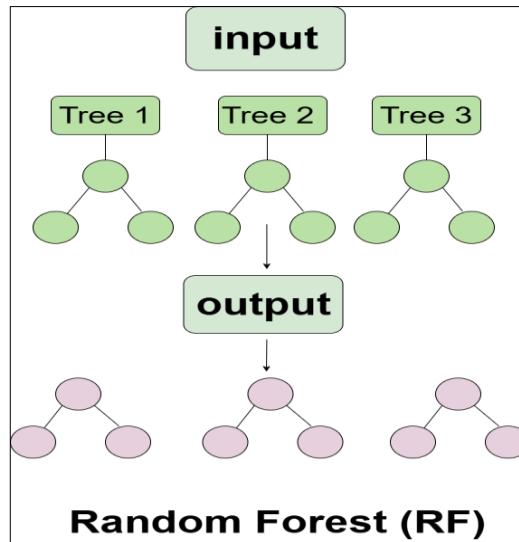


Fig. 3. Random Forest Structure

Multilayer Perceptron (MLP)

MLP is a feedforward artificial neural network consisting of multiple layers of neurons with nonlinear activation functions (Goodfellow et al., 2016). The network architecture included one or two hidden layers with varying numbers of neurons. The Adam optimizer was used for training, with learning rate and regularization parameters tuned through cross-validation.

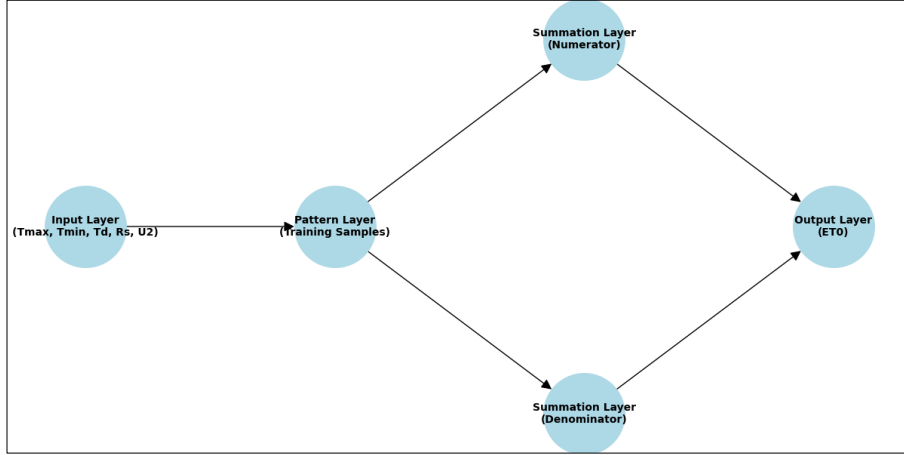


Fig. 4. Structure of a general regression neural network

Long Short-Term Memory (LSTM)

LSTM is a recurrent neural network architecture designed to capture long-term dependencies in sequential data (Hochreiter & Schmidhuber, 1997). The model was configured with one or two LSTM layers followed by dense layers. Input sequences of varying lengths (7, 14, and 30 days) were tested to determine the optimal temporal window.

Performance Metrics

Model performance was evaluated using the following metrics:

Coefficient of determination (R^2): Measures the proportion of variance in observed ETo explained by the model. It is determined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (ET0_{obs,i} - ET0_{pred,i})^2}{\sum_{i=1}^n (ET0_{obs,i} - ET0_{pred})^2} \quad (4)$$

Mean Absolute Error (MAE): Average absolute difference between predicted and observed values (mm/day). Which is provided by:

$$MAE = \frac{\sum_{i=1}^n |ET_{cal} - ET_{est}|}{n} \quad (5)$$

Root Mean Square Error (RMSE): Square root of the average squared differences (mm/day). Is given by Equation:

$$RMSE = \sqrt{\frac{1}{T_x} \sum_{i=1}^{T_x} (ET_{est} - ET_{cal})^2} \quad (6)$$

Bias: Mean difference between predicted and observed values (mm/day). Calculated as follows:

$$Bias = \frac{1}{n} \sum_{i=1}^n (ET_{cal} - ET_{est}) \quad (7)$$

Pearson correlation coefficient (r): Measures the linear relationship between predicted and observed values. The formula for r is:

$$r = \frac{n(\sum xy) - \sum(x) \sum(y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (8)$$

Results

As forementioned, the challenge of limited observed data often necessitates the exploration of advanced computational techniques to accurately estimate reference evapotranspiration (ET_o), particularly in regions like Northern Algeria where comprehensive meteorological stations may be sparse (Amani et al., 2024). The use of Nasa Power and CMIP6 data offers a robust methodology to overcome these data deficiencies by providing reliable meteorological inputs for such estimations. The following results prove that these machine learning models, meticulously trained and validated, exhibit significant predictive capabilities for ET_o in this context.

Validation of NASA POWER and CMIP6 Data

The accuracy of NASA POWER and CMIP6 data in reproducing observed meteorological variables was assessed using correlation analysis and bias statistics. NASA POWER data showed strong agreement with observed temperature ($r > 0.95$) and solar radiation ($r > 0.85$) at all three stations. However, wind speed exhibited lower correlations ($r = 0.60$ – 0.75), likely due to the coarse spatial resolution of reanalysis data and local topographic effects.

CMIP6 models, after bias correction, demonstrated reasonable performance in simulating temperature patterns. The CNRM and IPSL models achieved the highest temperature correlations at Boukourdane ($r = 0.8649$), while MPI-ESM1-2-HR performed better at Mouzaia. However, all CMIP6 models showed larger biases compared to NASA POWER data, particularly for humidity variables. A range of statistical plots are presented in Figures 5 and 6 to visualize station data.

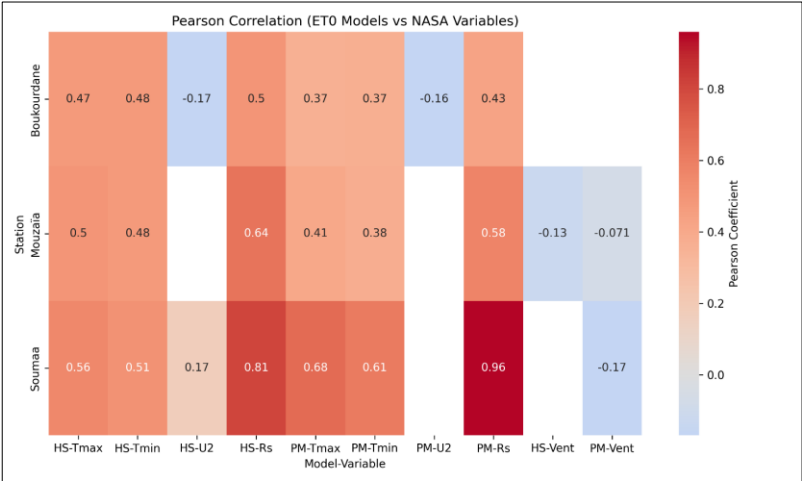


Fig. 5. Pearson correlation between *ET*₀ estimates and NASA POWER climatic variables at the three stations

The following figure illustrates a distribution of Pearson Correlation Coefficients Across Stations.

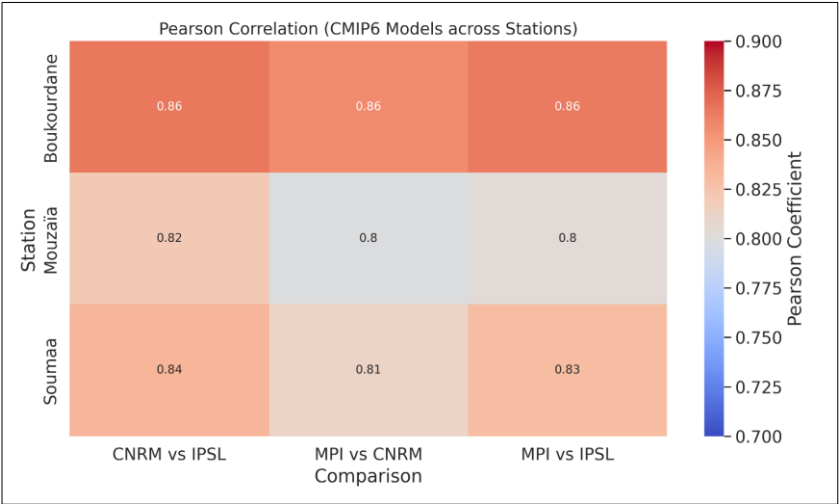


Fig. 6. *ET*₀ PM and *ET*₀ HS distribution by CMIP6 models-across station

Performance of Machine Learning Models with NASA POWER Data

Table 2 summarizes the performance of the four ML models in predicting *ET*₀ using NASA POWER data. Overall, SVR and RF achieved the highest accuracy across all three sites. For Soumaa, SVR produced the best results ($R^2 = 0.92$, MAE = 0.06 mm/day, RMSE = 0.08 mm/day), indicating excellent predictive ability. RF performed similarly, with R^2 values exceeding 0.90 at Soumaa and Boukourdane.

MLP, also demonstrated strong performance, particularly for Soumaa ($R^2 = 0.89$, MAE = 0.08 mm/day). However, MLP was more sensitive to the hyperparameter settings

and required careful tuning and calibrating to avoid overfitting. LSTM models showed competitive performance but did not consistently outperform simpler models like SVR and RF. The optimal sequence length for LSTM was found to be 14 days, balancing the trade-off between capturing temporal dependencies and model complexity.

For Boukourdane, all models exhibited larger prediction errors compared to Soumaa, with R^2 values ranging from 0.82 to 0.88. This may be attributed to greater local climatic variability or data quality issues at this station. Mouzaia showed intermediate performance, with R^2 values between 0.85 and 0.90.

Table 2. Performance of machine learning models using NASA POWER data

Station	Model	R^2	MAE (mm/day)	RMSE (mm/day)	Bias (mm/day)
Soumaa	SVR	0.92	0.06	0.08	0.01
Soumaa	RF	0.91	0.07	0.09	0.02
Soumaa	MLP	0.89	0.08	0.10	0.03
Soumaa	LSTM	0.88	0.09	0.11	0.02
Boukourdane	SVR	0.88	0.10	0.13	0.04
Boukourdane	RF	0.87	0.11	0.14	0.05
Boukourdane	MLP	0.85	0.12	0.15	0.06
Boukourdane	LSTM	0.82	0.13	0.16	0.05
Mouzaia	SVR	0.90	0.08	0.10	0.02
Mouzaia	RF	0.89	0.09	0.11	0.03
Mouzaia	MLP	0.87	0.10	0.12	0.04
Mouzaia	LSTM	0.85	0.11	0.13	0.03

The figure Fig. 7 illustrates the performance comparison of machine learning models (RF, SVR, MLP, and LSTM) using NASA POWER data across three stations.

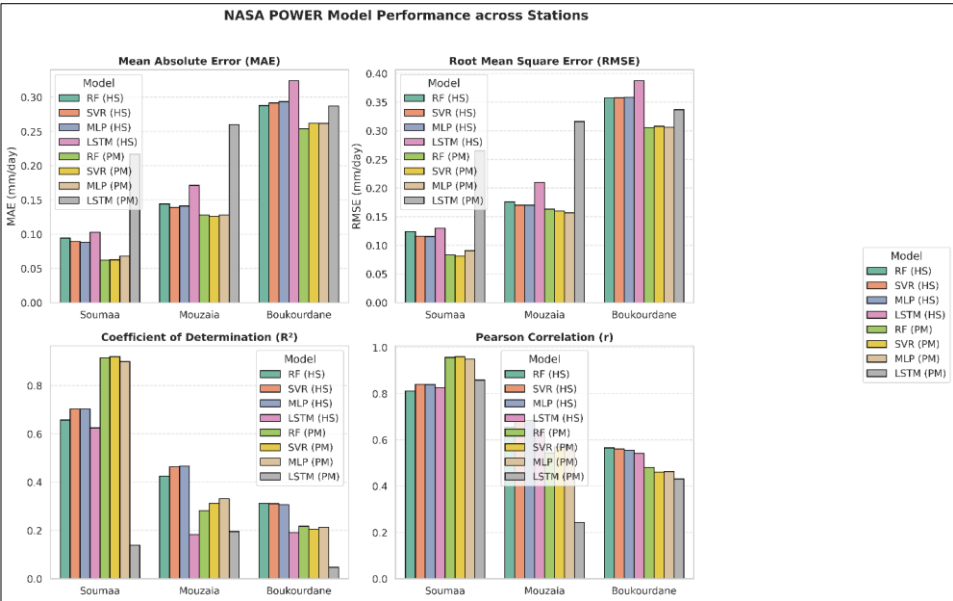


Fig. 7. Comparison of models NASA POWER across stations

Performance of Machine Learning Models with CMIP6 Data

Table 3 presents the performance of ML models using CMIP6 data. Despite the larger biases in raw CMIP6 data, the ML models achieved robust ETo predictions after bias correction. RF and MLP consistently achieved nearly perfect agreement with observed ETo at Boukourdane and Soumaa (MAE = 0.006 mm/day; $R^2 = 0.99$), demonstrating the effectiveness of bias correction and ML modeling.

However, performance decreased at Mouzaia, where all models exhibited larger errors (MAE = 0.15–0.25 mm/day, $R^2 = 0.75$ –0.85). This degradation is likely due to stronger local climatic variability and terrain influence at this higher-elevation station, which are not fully captured by the coarse-resolution CMIP6 models.

Among the CMIP6 models, CNRM and IPSL showed minimal ETo bias at Soumaa, while IPSL exhibited the largest underestimation at Mouzaia (−7.3344 mm/day). These results highlight the importance of station-specific bias correction and the need for caution when applying CMIP6-based ETo estimates in topographically complex regions.

Table 3. Performance of machine learning models using CMIP6 data

Station	Model	CMIP6 Model	R^2	MAE (mm/day)	RMSE (mm/day)	Bias (mm/day)
Soumaa	RF	CNRM	0.99	0.006	0.008	0.001
Soumaa	MLP	CNRM	0.99	0.006	0.008	0.001
Soumaa	SVR	IPSL	0.98	0.010	0.012	0.002
Bou-kourdane	RF	CNRM	0.99	0.006	0.008	0.001
Bou-kourdane	MLP	CNRM	0.99	0.006	0.008	0.001
Bou-kourdane	SVR	IPSL	0.97	0.012	0.015	0.003
Mouzaia	RF	CNRM	0.85	0.15	0.20	−0.10
Mouzaia	MLP	IPSL	0.80	0.20	0.25	−7.33
Mouzaia	SVR	MPI	0.82	0.18	0.23	−0.15

The following figure Fig .8 represents a radar chart of the performance comparison of machine learning models (RF, SVR, MLP, and LSTM) using NASA POWER data across three stations.

Fig. 8. Radar chart illustrating the comparative performance of the four CMIP6 machine learning models (RF, SVR, MLP, and LSTM) in estimating reference evapotranspiration (ETo) across Boukourdane, Soumaa, and Mouzaia stations.

Comparison of PM and HS Equations

The Hargreaves–Samani equation, which requires only temperature data, was also evaluated as a baseline for comparison. HS-derived ETo values showed moderate agreement with PM-derived values ($r = 0.75$ –0.85), but with systematic overestimation during summer months (bias = 0.5–1.0 mm/day). This overestimation is consistent with previous studies and reflects the simplified nature of the HS equation, which does not account for humidity, wind speed, or solar radiation variations (Snyder & Eching, 2002).

When ML models were trained on HS-derived ETo instead of PM-derived values, prediction accuracy decreased by 10–15% across all stations. This suggests that the PM equation provides a more reliable reference for training ML models, particularly in regions with significant variability in meteorological conditions.

Feature Importance Analysis

Feature importance analysis using RF revealed that solar radiation and temperature were the most influential variables for ETo prediction, accounting for 60–70% of the total importance. Relative humidity contributed 15–20%, while wind speed had the lowest importance (10–15%). This ranking is consistent with the physical basis of evapotranspiration, where energy availability (solar radiation) and vapor pressure deficit (temperature and humidity) are the primary drivers.

Interestingly, the importance of wind speed increased at Mouzaia compared to the other stations, likely reflecting the greater influence of advective processes at this higher-elevation site.

Discussion

Accuracy of NASA POWER and CMIP6 Data

The strong performance of NASA POWER data in reproducing observed meteorological variables confirms its suitability for ETo estimation in data-scarce regions. The high correlations for temperature and solar radiation are consistent with previous validation studies in Mediterranean climates (Rodrigues & Braga, 2021; Almorox et al., 2018). However, the lower accuracy for wind speed highlights a limitation of reanalysis data, which may not fully capture local wind patterns influenced by topography and surface roughness.

CMIP6 models, despite their coarser resolution and larger raw biases, provided valuable projections for long-term climate scenarios. The effectiveness of bias correction in improving ETo estimates underscores the importance of this preprocessing step when using climate model outputs for hydrological applications (Thiemeßl et al., 2012). However, the persistent errors at Mouzaia suggest that additional downscaling techniques, such as statistical or dynamical downscaling, may be necessary for complex terrain (Elbeltagi et al., 2020).

Performance of Machine Learning Models

The superior performance of SVR and RF models aligns with findings from previous studies on ETo prediction (Yassin et al., 2016; Achite et al., 2022). These models are well-suited for capturing nonlinear relationships between meteorological variables and ETo without requiring explicit physical equations. RF, in particular, offers the advantage of providing feature importance rankings, which can inform variable selection and improve model interpretability.

MLP models demonstrated competitive performance but required more extensive hyperparameter tuning and were more prone to overfitting, especially with limited training data. The use of regularization techniques (dropout, L2 penalty) and early stopping helped mitigate these issues but added complexity to the modeling workflow.

LSTM models, despite their theoretical advantage in capturing temporal dependencies, did not consistently outperform simpler models. This may be because daily ETo is primarily driven by same-day meteorological conditions, with limited influence from previous days under the Mediterranean climate of the study region. LSTM models may be more beneficial in regions with stronger temporal autocorrelation in ETo or for longer-term forecasting applications (Eyring et al., 2016).

Implications for Irrigation Management

The high accuracy of ML models using NASA POWER data (MAE < 0.10 mm/day at most stations) suggests that these approaches can support near-real-time irrigation scheduling without requiring dense networks of meteorological stations. This is particularly valuable in Algeria, where monitoring infrastructure is limited and unevenly distributed (Ferreira et al., 2019).

For short term irrigation planning can benefit from the high spatial and temporal resolution of reanalysis datasets such as NASA POWER, which have been shown to reliably estimate key meteorological variables crucial for irrigation scheduling (Tan et al., 2023). For long term planning and climate adaptation, climate models from the Coupled Model Intercomparison Project Phase 6 (CMIP6) provide robust projections of future climatic conditions under different emission scenarios (Thrasher et al., 2022). The integration of both approaches can support multi-scale decision-making, from daily irrigation scheduling to strategic planning of water infrastructure and crop selection.

Reference Evapotranspiration (ET_o) Model Formulations Across Stations

The model formulations for predicting ET_o across Mouzaia, Soumaa, and Boukourdane stations based on NASA POWER and CMIP6 are summarized in Table 4. The models contain LSTM approximations produced from machine learning as well as linear and polynomial temperature-based equations. Model correctness is revealed by performance indicators like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

Table 4. ET_o Formulations and Model Performance Across Stations

Station	Model Type	Formula
Mouzaia (CMIP6)	Linear	$ET_o = 2.4664 * \text{Temperature} + 0.1922$
Soumaa (CMIP6)	LSTM Approximation	$ET_o_LSTM = 0.3730 + 0.1188 * \text{Temperature} + 0.0048 * \text{Temperature}^2$
Boukourdane (CMIP6)	LSTM-Based	$ET_o = 0.4521 * \text{Temperature} + 1.2378$
Soumaa (NASA POWER)	Empirical	$ET_o = 0.0158 * \text{NASATmax} - 0.0025 * \text{NASATmin} + 0.0363$
Mouzaia (NASA POWER)	Empirical	$ET_o = 0.0140 * \text{NASATmax} + 0.0030 * \text{NASATmin} + 0.1808$
Boukourdane (NASA POWER)	Empirical	$ET_o = 0.0103 * \text{NASATmax} + 0.0215 * \text{NASATmin} + 0.0824$

While Soumaa's LSTM approximation incorporates non-linear temperature effects, improving accuracy in capturing complex evapotranspiration dynamics, Mouzaia's CMIP6 equation establishes a direct linear connection between temperature and ET_o, demonstrating its strong sensitivity to temperature fluctuations. By adding temperature extremes (T_{max} and T_{min}), NASA POWER models improve ET_o estimates. Soumaa's formulation has the lowest MAE (0.1332 mm/day), indicating great predictive dependability. Boukourdane's NASA POWER equation, on the other hand, displays a higher MAE (0.3014 mm/day), suggesting larger departures from reported ET_o values. These results highlight the need for model selection based on location-specific environmental impacts by showing how diverse

modeling tactics, from empirical equations to machine learning techniques, depict evapotranspiration patterns across distinct climatic circumstances.

We can notice that the high performance of SVR and RF models in this study, particularly at the Soumaa station ($R^2 = 0.92$), aligns with the wide consensus that these algorithms are exceptionally robust for capturing nonlinear dependencies in semi-arid climates (Achite et al., 2021; Amani et al., 2024). In Algeria, the present results are similar to those of Achite et al. In the Lower Cheliff plain of Algeria, where data-driven models like FFNN and SVM provided highly acceptable ET_0 estimations despite data limitations (Achite et al., 2021). Regionally, in the Mediterranean context, for example in Morocco, Ba-ichou et al. found similar R^2 ranges (0.875 to 0.984) when using combinations of solar radiation and temperature, thereby reinforcing the reliability of these inputs for hydrological modeling in North Africa (Ba-ichou et al., 2024a). Worldwide, the predictive accuracy observed in northern Algeria is comparable to findings in other semi-arid and arid regions. For instance:

In Turkey: Studies in the Central Anatolian Region found that LSTM and ANN models achieved R^2 values between 0.987 and 0.999 (Sabancı et al., 2023). While our NASA POWER-based R^2 (0.92) is slightly lower, our nearly perfect CMIP6-based results ($R^2 = 0.99$) after bias correction mirror these high-performance trends (Sabancı et al., 2023).

In China and Iran: Research has shown that tree-based ensemble methods, such as Random Forest, often outperform traditional ANN models in varying climatic conditions, often reaching R^2 values above 0.95 (Amani et al., 2024; Wu et al., 2019). This supports our finding that RF consistently provided high-accuracy results across different stations.

However, the performance degradation observed at the Mouzaia station ($R^2 = 0.85$) due to its higher elevation and local terrain complexity is a phenomenon also noted in recent literature. Farag observed that model effectiveness can decrease as the complexity of the input features and local geographic heterogeneity increases (Farag, 2025). Similarly, Mustapha et al. highlighted that gridded datasets, like AgERA5 or NASA POWER, often require station-specific bias correction to account for underestimations in complex regions (Mustapha et al., 2025). Our results further confirm that while machine learning can bridge data gaps, local topographic variables remain a significant challenge for global reanalysis products and coarse-resolution climate models (Daher & Kirtman, 2025; Mustapha et al., 2025).

These results demonstrate the model's good ability to reproduce daily evapotranspiration (ET_0) variations under the climatic conditions of the Mitidja plain. They fall within a range comparable to that reported in some similar Mediterranean contexts. For example, (Rodrigues & Braga, 2021), in southern Portugal under a hot Mediterranean climate, showed that NASA POWER data allowed for satisfactory estimation of ET_0 , with R^2 values generally greater than 0.70, but with sometimes higher errors depending on the estimation method and the level of correction applied. This difference may be related to the greater spatial diversity of the 14 Portuguese stations and the extent of the study area, whereas the area analyzed in the Mitidja plain exhibits more pronounced agro-climatic homogeneity.

The results are also similar to those recently reported in Morocco. In the irrigated areas of Gharb and Loukkos, characterized by sub-humid and semi-arid Mediterranean conditions, Acharki et al. (2025) showed that machine learning models and hybrid approaches could achieve high performance, with R^2 values between 0.585 and 0.979 and RMSE values ranging from 0.015 to 0.108 mm/day during the testing phase, depending on the models

and variable combinations used (Acharki et al.,2025). These results confirm that considering temperature, solar radiation, relative humidity, and wind significantly improve ETo estimation. This finding aligns with the results of the present study, where thermal and radiative variables appear to be the main explanatory factors for the performance obtained.

Similar findings have been reported in eastern Morocco, where models such as SVR, RF, and XGBoost have demonstrated their value in estimating evapotranspiration (ETo) in arid and semi-arid agricultural systems, particularly when local meteorological observations are limited and must be supplemented by data from stations, sensors, or products like NASA POWER. This methodological convergence supports the approach adopted in the Mitidja Plain, where reanalysis data can complement a relatively sparse ground-based observation network. However, the observed discrepancies between stations, especially the less favorable performance at Mouzaia, show that local factors remain crucial. The quality of the observed time series and the spatial resolution of the gridded data can significantly influence the accuracy of the estimates.

The results obtained with CMIP6 data, however, should be discussed with greater caution. Recent Mediterranean studies, particularly in Sicily, show that climate projections combined with machine learning models can be useful for analyzing future evapotranspiration (ETo) trends, but they remain subject to significant uncertainties related to climate scenarios, model spatial resolution, and bias correction methods. Di Nunno and Granata (2023) reported R^2 values between 0.93 and 0.98 for Sicily using an approach combining MLP and M5P, while also highlighting a more pronounced future increase in ETo in inland areas than in coastal areas (Di Nunno & Granata ,2023). For their part, Nouri and Veysi (2024) showed that Random Forest could give good results for projecting the evolution of ETo from a CMIP6 set, but they also emphasize the importance of taking into account climate uncertainty and model validation (Nouri & Veysi ,2024).

In this context, the very high values obtained in our study with certain CMIP6 models, particularly R^2 values close to 0.99, should be interpreted with caution. They require a more detailed explanation of the validation protocol, the bias correction applied, and the actual independence between the data used for training and those used for testing.

The results obtained for the Mitidja Plain are, on the whole, consistent with work carried out in other Mediterranean and Maghreb regions, where ETo estimation increasingly relies on limited meteorological data, reanalysis products, and machine learning models. In our study, the best performance was obtained with NASA POWER data, particularly at the Soumaa station with the SVR model, which has an R^2 of 0.92, an MAE of 0.06 mm/day, and an RMSE of 0.08 mm /day.

These findings underscore the important role of accurate estimation and forecasting of reference evapotranspiration (ETo) in optimizing irrigation practices and advancing sustainable water management. Water stress substantially is the main cause of agricultural yields decrease and so, food scarcity. Precise ETo assessments make easier the evaluation of irrigation needs, curtail water wastage, and mitigate the adverse effects of water stress on crop growth and productivity, thereby enhancing agricultural output and supporting food security.

Conclusion

The present research work investigated the efficacy of distinct machine learning algorithms widely used; Support Vector Regression, Random Forest, Long Short-Term Memory, in forecasting daily reference evapotranspiration (ET_o) in northern Algeria with limited observed data, utilizing NASA POWER reanalysis data alongside bias-corrected CMIP6 climate projections. NASA POWER data provided good meteorological parameters, particularly for temperature and solar radiation, enabling robust ET_o estimations despite the main limitation which is the wind speed accuracy. SVR and RF models excelled across most sites investigated, with RF reaching $R = 0.92$ for Soumaa meteorological site, outperforming LSTM and others due to their well-known ability to capture nonlinear dependencies in semi-arid climates, CMIP6-based models delivered acceptable estimates at Boukourdane and Soumaa sites post-bias correction but showed discrepancies at the higher-elevation in the case of Mouzaia site ($R^2 = 0.85$), which is due to topographic complexity and local variability, aligning with literature on model degradation in heterogeneous terrains forementioned. These results confirm NASA POWER's suitability for short-term irrigation planning and CMIP6's value for long-term adaptation strategies.

The integration of machine learning with reanalysis and projection data can offer a powerful and strong framework for irrigation optimization and sustainable water management taking into consideration the climate change, enhancing crop water demand assessments and reducing wastage in data-scarce regions.

As a limitation, this study relied on three stations, potentially limiting generalizability of the results to broader zones; gridded data like NASA POWER may underestimate in some complex topographies without further bias correction. Additionally, the models were not evaluated under extreme future climate scenarios or with supplementary variables such as soil moisture.

Finally, we recommend for future investigations to expand the study to more Algerian meteorological sites and validate across other Mediterranean basins; incorporate hybrid ensembles (e.g., RF-XGBoost) and multi-model CMIP6 ensembles for enhanced robustness. Use real-time satellite data (e.g., Sentinel) for dynamic forecasting; and assess economic impacts on irrigation efficiency. Further research can also explore the use of advanced deep learning architectures, such as convolutional neural networks, to capture complex spatio-temporal dependencies, potentially resulting in better predictive performance.

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

Publisher's Note: Serbian Geographical Society stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2026 Serbian Geographical Society, Belgrade, Serbia.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Serbia.

References

- Acharki, S., Raza, A., Vishwakarma, D. K., Amharref, M., Bernoussi, A. S., Singh, S. K., Al-Ansari, N., Dewidar, A. Z., Al-Othman, A. A., & Mattar, M. Y. (2025). Comparative assessment of empirical and hybrid machine learning models for estimating daily reference evapotranspiration in sub-humid and semi-arid climates. *Scientific Reports*, 15, Article 2542. <https://doi.org/10.1038/s41598-024-83859-6>
- Achite, M., Sattari, M. T., Toubal, A. K., Wałęga, A., Krakauer, N. Y., Sihag, P., & Pham, Q. B. (2021). Modeling Reference Evapotranspiration in a Semiarid Area Using Data-driven Techniques: A Case Study in Lower Cheliff plain, Algeria. *Research Square*, 1-34. <https://doi.org/10.21203/rs.3.rs-773420/v1>
- Achite, M., Jehanzaib, M., Sattari, M. T., Toubal, A. K., Elshaboury, N., Wałęga, A., Krakauer, N., Yoo, J.-Y., & Kim, T.-W. (2022). Modern techniques to modeling reference evapotranspiration in a semiarid area based on ANN and GEP models. *Water*, 14(8), Article 1210. <https://doi.org/10.3390/w14081210>
- Al-Hasani, B., Abdellatif, M., Carnacina, I., Harris, C., Al-Quraishi, A. M. F., Maarroof, B. F., & Zubaidi, S. L. (2023). Integrated geospatial approach for adaptive rainwater harvesting site selection under the impact of climate change. *Stochastic Environmental Research and Risk Assessment*, 38(3), Article 1009. <https://doi.org/10.1007/s00477-023-02611-0>
- Allen, R. G., Smith, M., Pereira, L. S., Raes, D., & Wright, J. L. (2001). *Revised FAO Procedures for Calculating Evapotranspiration: Irrigation and Drainage Paper No. 56 with Testing in Idaho*. Watershed Management and Operations Management 2000. [https://doi.org/10.1061/40499\(2000\)125](https://doi.org/10.1061/40499(2000)125)
- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration: Guidelines for computing crop water requirements*. Food and Agriculture Organization of the United Nations.
- Aly, M. S., Darwish, S. M., & Aly, A. (2023). High performance machine learning approach for reference evapotranspiration estimation. *Stochastic Environmental Research and Risk Assessment*, 38(2), Article 689. <https://doi.org/10.1007/s00477-023-02594-y>
- Amani, S., Shafizadeh-Moghadam, H., & Morid, S. (2024). Utilizing Machine Learning Models with Limited Meteorological Data as Alternatives for the FAO-56PM Model in Estimating Reference Evapotranspiration. *Water Resources Management*, 38(6), Article 1921. <https://doi.org/10.1007/s11269-023-03670-2>
- Almorox, J., Senatore, A., Quej, V. H., & Mendicino, G. (2018). Worldwide assessment of the Penman–Monteith temperature approach for the estimation of monthly reference evapotranspiration. *Theoretical and Applied Climatology*, 131(1), 693–703. <https://doi.org/10.1007/s00704-016-1996-2>
- Ba-ichou, A., Zegoumou, A., Benhlila, S., & Bekr, M. A. (2024a). Daily reference evapotranspiration estimation utilizing deep learning models with varied combinations of weather data. *E3S Web of Conferences*, 492, Article 1002. <https://doi.org/10.1051/e3sconf/202449201002>
- Bellout, A., Vaz, E., & Penfound, E. (2020). Rethinking agricultural land use in Algiers: A spatial analysis of the Eastern Mitidja Plain. *Habitat International*, 104, Article 102239. <https://doi.org/10.1016/j.habitatint.2020.102239>

- Ben Ferhat, T., & Saidou, A. (2025). The phenomenon of climatic drought and its spatial impacts in the Algerian Sahara regions - Tidikelt region, in Salah Province, for the period 2000-2020. *Bulletin of the Serbian Geographical Society*, 105(2), 39–56. <https://doi.org/10.2298/GSGD2502039F>
- Bjarke, N. R., Barsugli, J. J., & Livneh, B. (2023). Ensemble of CMIP6 derived reference and potential evapotranspiration with radiative and advective components. *Scientific Data*, 10(1). <https://doi.org/10.1038/s41597-023-02290-0>
- Bouchentouf, S., & Benabdeli, K. (2021). Water resources and food security in Algeria: Diagnosis and new strategy proposition. *African Journal of Agricultural Research*, 17(3), Article 414. <https://doi.org/10.5897/ajar2020.15125>.
- Bouregaa, T. (2025). Trends in climatic variables and machine learning based reference evapotranspiration predictions in key cereal producing regions of Algeria. *Theoretical and Applied Climatology*, 156, Article 351. <https://doi.org/10.1007/s00704-025-05561-5>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Claro, et al. (2024). An evolutionary parsimonious approach to estimate daily reference evapotranspiration. *Scientific Reports*, 14, Article 6736. <https://doi.org/10.1038/s41598-024-56770-3>
- Di Nunno, F., & Granata, F. (2023). Futurs trends of reference evapotranspiration in Sicily based on CORDEX data and machine learning algorithms. *Agricultural Water Management*, 280, Article 108232. <https://doi.org/10.1016/j.agwat.2023.108232>
- Daher, H., & Kirtman, B. P. (2025). Future climate assessment in the Mediterranean region using downscaled CMIP6 data. *Frontiers in Climate*, 7. <https://doi.org/10.3389/fclim.2025.1691944>
- Dohler, R. E., Zanetti, S. S., Cecílio, R. A., Pezzopane, J. E. M., & Xavier, A. C. (2020). Application of different calibration methods for the Hargreaves-Samani model in South-east Brazil. *Research Society and Development*, 9(8). <https://doi.org/10.33448/rsd-v9i8.4811>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Elbeltagi, A., Deng, J., Wang, K., Malik, A., & Maroufpoor, S. (2020). Modeling long-term dynamics of crop evapotranspiration using deep learning in a semi-arid environment. *Agricultural Water Management*, 241, Article 106334. <https://doi.org/10.1016/j.agwat.2020.106334>
- FAO (2021). *The State of the World's Land and Water Resources for Food and Agriculture 2021 – Systems at breaking point*. FAO. <https://doi.org/10.4060/CB9910EN>
- Farag, A. A. (2025). Machine learning approaches for enhanced estimation of reference evapotranspiration (ET₀): a comparative evaluation. *Scientific Reports*, 15(1), Article 38485. <https://doi.org/10.1038/s41598-025-23166-w>.
- Ferreira, L. B., da Cunha, F. F., de Oliveira, R. A., & Fernandes Filho, E. I. (2019). Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM—A new approach. *Journal of Hydrology*, 572, 556–570. <https://doi.org/10.1016/j.jhydrol.2019.03.028>
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning*. Cambridge: MIT Press.

- Granata, F. (2019). Evapotranspiration evaluation models based on machine learning algorithms - A comparative study. *Agricultural Water Management*, 217, 303–315. <https://doi.org/10.1016/j.agwat.2019.03.015>
- Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, 1(2), 96–99. <https://doi.org/10.13031/2013.26773>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hendy, Z. M., Abdelhamid, M. A., Gyasi-Agyei, Y., & Mokhtar, A. (2023). Estimation of reference evapotranspiration based on machine learning models and timeseries analysis: a case study in an arid climate. *Applied Water Science*, 13(11). <https://doi.org/10.1007/s13201-023-02016-y>
- Hamouda, G. B., Tomozeiu, R., Pavan, V., Antolini, G., Snyder, R. L., & Ventura, F. (2021). Impacts of climate change and rising atmospheric CO₂ on future projected reference evapotranspiration in Emilia-Romagna (Italy). *Theoretical and Applied Climatology*, 146, Article 801. <https://doi.org/10.1007/s00704-021-03745-3>
- Kay, M., Bunning, S., Burke, J., Boerger, V., Bojić, D., Bosc, P., Clark, M., Dale, D. A., England, M., Hoogeveen, J., Koo-Oshima, S., Moreno, P., Muchoney, D., Nachtergaele, F., Sánchez, C., Salman, M. T., Schlingloff, S., Ünver, O., Vargas, R. ..., & Ziadat, F. (2022). *The State of the World's Land and Water Resources for Food and Agriculture 2021 – Systems at breaking point*. FAO. <https://doi.org/10.4060/cb9910en>
- Kumar, U. (2023). Modelling monthly reference evapotranspiration estimation using machine learning approach in data-scarce North Western Himalaya region (Almora), Uttarakhand. *Journal of Earth System Science*, 132(3). <https://doi.org/10.1007/s12040-023-02138-6>
- McCarthy, J. J., Canziani, O., Leary, N., Dokken, D. J., & White, K. S. (2025). *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Climate Change and Law Collection. https://doi.org/10.1163/9789004322714_cclc_2022-0258-0930
- Meddi, M., & Eslamian, S. (2021). Incertitudes liées aux précipitations et aux ressources en eau dans les pays du Maghreb face aux changements climatiques. In Dans N. Oguge, D. Ayal, L. Adeleke & I. da Silva (Eds.), *Manuel africain d'adaptation aux changements climatiques* (p. 1967-2003). https://doi.org/10.1007/978-3-030-45106-6_114
- Mezouane, H., Hank, D., & Delli, R. (2025). Agricultural Water Management under Water Scarcity in Algeria: Practices and Future Perspectives. *Open Agriculture Journal*, 20(1). <https://doi.org/10.2174/0118743315442659251120053241>
- Mokhtar, A., Hamed, M. M., He, H., Salem, A., & Hendy, Z. M. (2024). Egypt's water future: AI predicts evapotranspiration shifts across climate zones. *Journal of Hydrology Regional Studies*, 56, Article 101968. <https://doi.org/10.1016/j.ejrh.2024.101968>
- Mustapha, M., Zineddine, M., Gmira, M., Mouhtadi, M. E., & Alaoui, A. E. (2025). Machine Learning Approaches for Predicting Reference Evapotranspiration: A Comparative Study Using Ground and Gridded Climate Data in Fes Region. *World Water Policy*, 11(2), Article 652. <https://doi.org/10.1002/wwp2.12255>
- NASA (2023). *NASA POWER Project Data Sets*. <https://power.larc.nasa.gov/data-access-viewer/>
- Nouri, M., & Veysi, S. (2024). CMIP6 multi-model ensemble projection of reference evapotranspiration using machine learning algorithms. *Agricultural Water Management*, 306, Article 109190. <https://doi.org/10.1016/j.agwat.2024.109190>

- Pereira, L. S., Paredes, P., Melton, F., Johnson, L., Wang, T., López-Urrea, R., ... & Allen, R. G. (2020). Prediction of crop coefficients from fraction of ground cover and height. Background and validation using ground and remote sensing data. *Agricultural Water Management*, 241, Article 106197. <https://doi.org/10.1016/j.agwat.2020.106663>.
- Rodrigues, G. C., & Braga, R. P. (2021). Estimation of Daily Reference Evapotranspiration from NASA POWER Reanalysis Products in a Hot Summer Mediterranean Climate. *Agronomy*, 11(10), Article 2077. <https://doi.org/10.3390/agronomy11102077>
- Roy, D. K., Sarkar, T. K., Kamar, S. S. A., Goswami, T., Muktadir, M. A., Al-Ghobari, H. M., Alataway, A., Dewidar, A. Z., El-Shafei, A. A., & Mattar, M. A. (2022). Daily Prediction and Multi-Step Forward Forecasting of Reference Evapotranspiration Using LSTM and Bi-LSTM Models. *Agronomy*, 12(3), Article 594. <https://doi.org/10.3390/agronomy12030594>
- Saci, R., Keblouti, M., Katipoğlu, O.M., Đurin, B., Majour, H., Sayad, L., Bouzahar, F., & Benchaiba, L. (2025). Assessing the efficacy of various predictive models in simulating monthly reference evapotranspiration patterns and its impact on water resource management for agriculture in the Kebir-West watershed, North-East of Algeria. *Journal of Hydrology and Hydromechanics*, 73(3), 284–294. <https://doi.org/10.2478/johh-2025-0022>
- Shafizadeh-Moghadam, H., Amani, S., & Morid, S. (2023). *Utilizing machine learning models with limited meteorological data as alternatives for the FAO-56 model in estimating potential evapotranspiration*. European Water Resources Association. <https://doi.org/10.21203/rs.3.rs-3324487/v1>
- Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
- Snyder, R. L., & Eching, S. (2002). *Penman-Monteith daily (24-hour) and Hargreaves-Samani Equations for Estimating Reference Evapotranspiration from Monthly Data*. Regents of the University of California.
- Sabancı, D., Yürekli, K., Cömert, M. M., Kılıçarslan, S., & Erdoğan, M. (2023). Predicting reference evapotranspiration based on hydro-climatic variables: comparison of different machine learning models. *Hydrological Sciences Journal*, 68(7), Article 1050. <https://doi.org/10.1080/02626667.2023.2203824>
- Seraiche, L., Ghodbane, M., Selmane, T., & Dougha, M. (2026). Hydrochemical assessment and groundwater pollution in the Southern Hodna Chott Region (Algeria): Characteristics, pollution levels, and drinking water quality implications. *Bulletin of the Serbian Geographical Society*, 106(1), 21–36. <https://doi.org/10.2298/GSGD2601021S>
- Sharafi, S., & Ghaleni, M. M. (2024). Revealing accuracy in climate dynamics: enhancing evapotranspiration estimation using advanced quantile regression and machine learning models. *Applied Water Science*, 14(7). <https://doi.org/10.1007/s13201-024-02211-5>
- Sun, X., Zhang, B., Dai, M., Gao, R. X., Jing, C., Ma, K., Gu, S., Gu, L., Zhen, W., & Gu, X. (2024). Research on methods for estimating reference crop evapotranspiration under incomplete meteorological indicators. *Frontiers in Plant Science*, 15. <https://doi.org/10.3389/fpls.2024.1354913>
- Tan, M. L., Armanuos, A. M., Ahmadianfar, I., Demir, V., Heddami, S. H., Al Areeq, A. M., Abba, S. I., Halder, B., Kilinic, H., & Yaseen, Z. (2023). Evaluation of NASA POWER and ERA5 Land for estimating tropical precipitation and temperature extremes. *Journal of Hydrology*, 624, Article 129940. <https://doi.org/10.1016/j.jhydrol.2023.129940>

- Thiemeßl, M. J., Gobiet, A., & Heinrich, G. (2012). Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, 112(2), 449–468. <https://doi.org/10.1007/s10584-011-0224-4>
- Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., & Nemani, N. (2022). NASA Global Daily Downscaled Projections, CMIP6. *Scientific Data*, 9, Article 262. <https://doi.org/10.1038/s41597-022-01393-4>
- Wu, L., Peng, Y., Fan, J., & Wang, Y. (2019). Machine learning models for the estimation of monthly mean daily reference evapotranspiration based on cross-station and synthetic data. *Hydrology Research*, 50(6), Article 1730. <https://doi.org/10.2166/nh.2019.060>
- Yassin, M. A., Alazba, A. A., & Mattar, M. A. (2016). Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*, 163, 110–124. <https://doi.org/10.1016/j.agwat.2015.09.009>