

ARTICLE

A novel data-driven solution for sustainable irrigation: Multi-station validation of a three-variable support vector regression model

Assia Meziani^{1*}, Nabil Mega¹, and Hadjer Tabboucha²¹New Technology and Local Development Laboratory, Department of Hydraulic and Civil Engineering, Faculty of Technology, University of El-Oued, El-Oued, Algeria²Department of Hydraulic and Civil Engineering, Faculty of Technology, University of El-Oued, El-Oued, Algeria

Abstract

Accurate evaluation of reference evapotranspiration (ET₀) is important for effective long-term irrigation planning and water resource management in arid/semi-arid regions. This study evaluates four models: linear regression, Takagi–Sugeno fuzzy inference system, random forest (RF), and support vector regressor (SVR) to estimate daily ET₀. The models were trained using ERA5-Land daily climatic variables across 24 meteorological stations in the Batna region (Algeria). To address potential temporal leakage in meteorological time series, we evaluated both random splitting (70% training, 15% validation, and 15% testing) and chronological splitting. The SVR model, utilizing only three automated features—air temperature at 2 m, soil temperature at 0–7 cm, and vapor pressure deficit—demonstrated significantly better performance than the other models at each station, with test root mean square error (RMSE) values ranging from 0.544 mm/day (Tazoult) to 0.780 mm/day (Bitam) and Nash–Sutcliffe efficiency values ≥ 0.91 . RF showed severe overfitting, with a 164.93% increase in test RMSE. To further validate SVR's robustness, we compared it against two additional benchmark models from the literature: a multilayer perceptron (MLP) and a light gradient boosting machine (LightGBM). SVR consistently outperformed both MLP (RMSE: 0.645 mm/day) and LightGBM (RMSE: 0.612 mm/day). Therefore, the results suggest that the SVR model utilizing only three climatic features provides a highly novel, computationally efficient, and accurate platform for predicting daily ET₀ in arid/semi-arid Mediterranean climates, offering an operational solution for irrigation scheduling in data-scarce environments.

*Corresponding author:

Assia Meziani
(assia-meziani@univ-eloued.dz)

Citation: Meziani, A., Mega, N. & Tabboucha, H. (2026). A novel data-driven solution for sustainable irrigation: Multi-station validation of a three-variable support vector regression model. *Int J Systematic Innovation*, 10(4): 026140037. [https://doi.org/10.6977/IJoSI.202608_10\(4\).0012](https://doi.org/10.6977/IJoSI.202608_10(4).0012)

Received: March 30, 2026**Revised:** June 9, 2026**Accepted:** June 19, 2026**Published online:** July 28, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Keywords: Reference evapotranspiration; Machine learning; FAO-56 Penman–Monteith; Semi-arid region; Algeria

1. Introduction

The development of accurate estimates of reference evapotranspiration (ET₀) is very important for irrigation management and scheduling, hydrologic modelling and planning, and water resource management (Galioto *et al.*, 2020; Gong *et al.*, 2020; Gu *et al.*, 2020; L. Zhao *et al.*, 2013; Y. Zhou *et al.*, 2024). The need for reliable estimates of ET₀ is important in semi-arid regions where increased water scarcity, elevated evaporative

demand, and variable climate regimes present an increasing threat to agricultural production, sustainability and ultimately food security (Misra, 2014; Morante-Carballo *et al.*, 2022; Rosegrant *et al.*, 2009; Schwabe *et al.*, 2015; Torabi Haghighi *et al.*, 2020). The Food and Agriculture Organization (FAO) recommends the FAO-56 Penman–Monteith (PM) method as the approach for estimating ET_0 because of its strong physical basis and wide applicability across all climate types (Allen *et al.*, 1998). The PM method, however, relies on complete and quality data for several meteorological parameters, including air temperature, relative humidity, wind speed, and solar radiation, which may be either non-existent or inadequate at many weather stations, a challenge frequently highlighted in comparative studies of empirical and intelligent models (Abdullah & Malek, 2016; Alexandris & Kerkides, 2003). This limitation is particularly significant in developing countries and rural/remote semi-arid locations (Landeras *et al.*, 2018; Sentelhas *et al.*, 2010). Therefore, the use of machine learning methods to estimate ET_0 with minimal amounts of input data has become popular. Research studies conducted in arid and semi-arid environments indicate that support vector regression (SVR), random forest (RF), and fuzzy inference systems (FIS) were able to estimate ET_0 accurately by reducing the number of required input data (Bellido-Jimenez *et al.*, 2021; Cobaner, 2011; Doğan, 2009; Kisi & Zounemat-Kermani, 2014; Sammen *et al.*, 2023; Tabari *et al.*, 2012). For example, research conducted in Iran indicated that SVR was a suitable method for modeling daily ET_0 based on limited climate information (Mohammadi *et al.*, 2020). In addition to the above-mentioned examples, hybrid machine learning models and ensemble models were used successfully in North Africa to develop predictive models of ET_0 in semi-arid to sub-humid regions better than traditional empirical models (Acharki *et al.*, 2025, supported by broader findings in hydrological forecasting by Abba *et al.*, 2020; Gelete, 2023; Song & Yao, 2022). A number of research studies have investigated estimating ET_0 using artificial intelligence-based methods such as artificial neural network models and regression models within the context of arid to semi-arid climate environments in the northern part of Algeria or the humid Mediterranean zone (Bouregaa *et al.*, 2025; Ladlani *et al.*, 2012; Meziani *et al.*, 2026). Very few studies have evaluated the effectiveness of these models in predicting ET_0 at the high-plateau level in cereal-producing provinces of Algeria (Table 1). This is particularly concerning given the important role of the high plateau in the Batna region and its susceptibility to drought. To address this gap, this study introduces a highly novel, multi-station framework using ERA5-Land reanalysis data only, evaluating a reduced-input SVR model that achieves exceptional accuracy using

only three easily accessible climatic parameters, thereby distinguishing our contribution from previous regional North African and Algerian studies by demonstrating that high-resolution reanalysis datasets can successfully replace ground-based measurements in data-scarce regions.

This study evaluates the accuracy of different machine learning algorithms—linear regression (LR), Takagi–Sugeno FIS (TS-FIS), RF, and SVR—for estimating ET_0 daily at 24 weather monitoring stations across the Batna region. Each model employed in this research used only the first three input parameters selected by an automated process based upon their relative importance at each individual weather monitoring station (air temperature at 2 meters [°C], soil temperature at 0–7 cm depth [°C], and vapor pressure deficit [kPa]). Weather station data used for this evaluation came from the European Centre for Medium-range Weather Forecasts ERA5-Land reanalysis data, a globally recognized dataset for high-resolution climate variables (Hersbach *et al.*, 2020), and was compared to benchmark estimates from the FAO-56 Penman–Monteith equation using a total of 10 commonly used measures of model accuracy including root mean square error (RMSE), mean absolute error (MAE), mean bias error (MBE), normalized root mean square error (nRMSE), Pearson’s correlation coefficient (R), Nash–Sutcliffe efficiency (NSE), Willmott index (WI), ratio of predicted variance to observed variance (RPD), relative percent increase in quality, and sum root ratio. Therefore, this study had two main goals: (i) to determine which of the machine learning algorithms tested provided the best estimates of daily ET_0 values in the semi-arid highlands of the Batna region; and (ii) to provide evidence that reliable estimates can be made with only three climatically sensitive input parameters. Ultimately, the results will serve as a useful guide for developing an effective irrigation strategy through improved water-use efficiency and conservation practices in semi-arid regions with limited data.

2. Materials and methods

2.1. Study area

The research area was Batna Province, situated in the eastern highlands of Algeria and part of the Aurès Mountains region. At an elevation of approximately 1,000–1,200 meters above sea level, it has a semi-arid climate (BSk–cold semi-arid), characterized by extreme seasonal temperature variations and minimal precipitation (Acharki *et al.*, 2025; Bouregaa *et al.*, 2025). In this respect, the Batna region experiences very cold winters and extremely hot, dry summers. Temperatures are generally mild throughout the year, with average monthly temperatures ranging between 13.5 °C and 15.7 °C. On average, the coldest

Table 1. Summary of recent literature on machine learning reference evapotranspiration estimation and identified research gaps

Study	Region/Climate	Models compared	Key limitations/gaps identified	Our study contribution
Bouregaa <i>et al.</i> (2025)	Algeria (Cereal regions)/Semi-arid	SVM, RF, MLP	Relied on large, complex input combinations; monthly resolution limits real-time irrigation scheduling	Evaluates daily resolution with only three easily accessible daily variables
Hamdaoui <i>et al.</i> (2026)	Eastern Morocco/Arid to Semi-arid	XGBoost, SVR, RF	Station-specific models lacked spatial portability; required complete solar radiation data	Develops a multi-station framework with high spatial portability across 24 stations using ERA5-Land only
Acharki <i>et al.</i> (2025)	Morocco/Sub-humid and Semi-arid	XGBoost-LightGBM, RF-LightGBM	Complex hybrid models are computationally heavy and difficult to deploy in local systems	Demonstrates a highly parsimonious, lightweight SVR model suitable for smart irrigation systems
Meziani (2026)	Humid Mediterranean (Jijel, Algeria)	MLP, RF, LightGBM	Focused solely on humid coastal zones; findings cannot be generalized to continental plateaus	Validates performance in the high-altitude continental plateau of the Batna region, highly susceptible to drought
Zereg & Belouz (2024)	Dar-El-Beidha, Algeria/Semi-arid	SVR, empirical models	Single-station validation; did not address temporal leakage or data dependencies	Performs extensive 24-station validation with rigorous data leakage and temporal split testing

Abbreviations: LightGBM: Light gradient boosting machine; MLP: Multilayer perceptron; RF: Random forest; SVM: Support vector machine; SVR: Support vector regression; XGBoost: Extreme gradient boosting.

month (January) has a temperature of 6.2 °C, while the warmest month (July) averages 26.9 °C. It is common for temperatures to exceed 32 °C during the summer, and there have been instances of temperatures reaching as high as 40 °C during heatwaves. During winter, temperatures are often quite low, occasionally dropping below 0 °C, especially at higher altitudes. There may also be some frost or light snowfall during these periods (Climate-Data.org, 2024; Weatherspark, 2024).

Annual precipitation varies significantly, with total values ranging from 305 mm to 496 mm depending on location and year. Precipitation generally occurs in the winter and early spring months (January to May), with a large proportion occurring in January. The two driest months (July and August) tend to receive less than 10 mm of precipitation, while the wetter months average 30–35 mm. Due to the limited precipitation that falls during most of the year, there is a considerable water deficit. Potential ET_0 far exceeds precipitation in all but a few months when sufficient amounts of rain fall to replenish soil moisture reserves (Allen *et al.*, 1998; Ladlani *et al.*, 2012). Twenty-four stations were selected to characterize the climatological variability across the Batna region. The stations were distributed throughout the province and included Bitam, Arris, Maafa, Batna, El Madher, Chir, Tazoult, Ain Yagout, Foug Toub, Merouana, Menaa, Sefiane, Bouzina, Ichemoul, Chemora, Teniet El Abed, Seggana, Timgad, Gosbat, Hidoussa, Lazrou, Ain Touta,

Metkaouak, and Seriana (Figure 1). These stations include both the high plateau regions important for wheat cultivation and the mountainous regions. Droughts occur repeatedly in both regions, making agricultural production difficult and complicating water resource management (Ghiglieri *et al.*, 2021).

2.2. Meteorological data and feature selection

Data on weather over the period 2000 to 2025 were downloaded from the Open-Meteo website for each of the 24 stations located in the Batna region using its “Historical Weather” API (Zippenfenig, 2024). The “Historical Weather” API is based upon the open-source ERA5/ERA5-Land reanalysis datasets which have a horizontal resolution of around 9–11 km (C3S, 2025). The following parameters were available on a daily basis; the air temperature at 2 m above ground level (°C); the relative humidity at 2 m above ground level (%); the dew point temperature at 2 m above ground level (°C); the accumulated precipitation (mm); the surface atmospheric pressure (hPa); the wind speed at 10 m above ground level (km/h); the soil temperature at depths of 0 to 7 cm (°C); the sunshine hours per day (h); net terrestrial radiation (W/m^2); and vapor pressure deficit (kPa). Reference evaporation potential (ET_0) has been estimated according to the FAO-56 Penman-Monteith formula (Allen *et al.*, 1998). Pearson’s coefficient was used to identify relationships between the various climatic parameters and ET_0 . Three best climate indicators

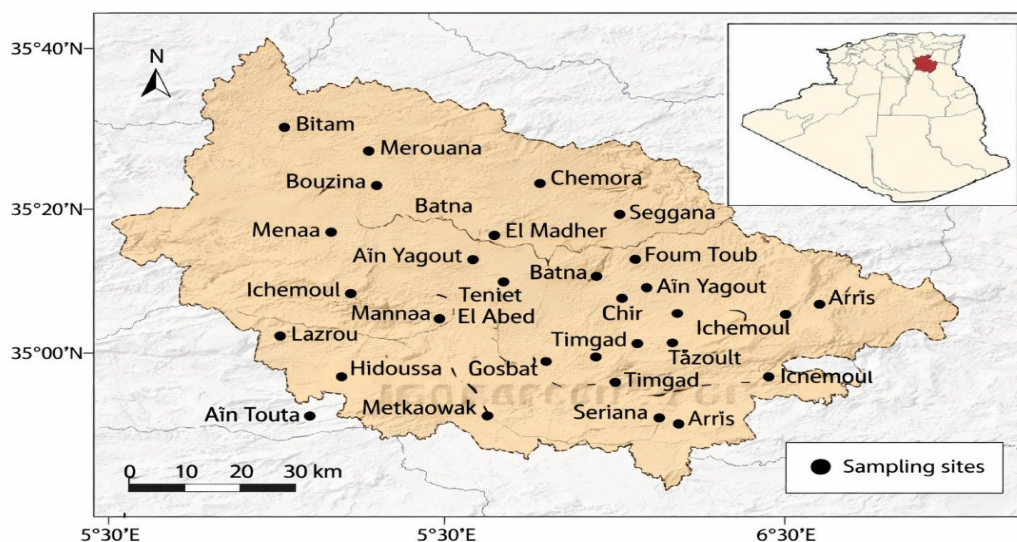


Figure 1. Location of the 24 stations studied in the Batna region, eastern Algeria

associated with ET_0 were identified for each location. These were always air temperature at 2 m, vapor pressure deficit, and soil temperature at a depth of 0–7 cm. Together, they formed the final set of climate indicators for the best estimation of ET_0 across all locations. Therefore, all relevant climate indicators were available for use as inputs into predictive algorithms such as those listed in Table 2.

The correlation matrix of all measured parameters shows that reference ET_0 is most closely positively related to the vapor pressure deficit (0.91) and then to air temperature (0.85) (Table 2), sunshine duration (0.72), dew point temperature (0.68), and soil temperature at 0–7 cm (0.65). Reference ET_0 is most closely negatively related to relative humidity (−0.75) and weakly related to wind speed (0.22) (Table 2). We chose to use air temperature (at 2 m) as the primary driver; vapor pressure deficit as the primary atmospheric term driving plant demand (it is the most highly correlated parameter with ET_0); and soil temperature at 0–7 cm as an energy-modifying term in relation to the surface. These three features represent distinct physical processes in the soil–vegetation–atmosphere transfer system, and their strong interrelationships ensure they impose mutually reinforcing physical constraints on predicting reference ET_0 , thereby avoiding overfitting and maintaining high physical consistency in the SVR model's predictions.

The completeness and consistency of the ERA5-Land reanalysis data across the full annual cycle are demonstrated in Figure 2, which shows the meteorological time-series variables over a representative 365-day period. This visualization confirms that all three selected variables

(air temperature, soil temperature, and vapor pressure deficit) are available without significant gaps throughout the year, ensuring robust and continuous data quality for model training and validation. The absence of missing values in these critical variables across all 24 stations over the 25-year study period (2000–2025) validates the reliability of the ERA5-Land dataset as a comprehensive substitute for ground-based measurements in data-scarce regions.

2.3. Modeling framework and methodology

We started with data acquisition/preprocessing using meteorological data from 24 weather stations. The target variable was identified as FAO-56 Penman–Monteith ET_0 ; all other meteorological variables were treated as candidate predictor variables. The three most highly correlated features (usually air temperature at 2 m, soil temperature at 0–7 cm, and vapor pressure deficit), which were chosen as candidate predictor variables during model development, provided robustness to missing data. The selected features and target variable were assigned to a subset of the dataset, which was then randomly divided into training (70%), validation (15%), and testing (15%). A fixed random number, 42, was used to ensure reproducibility of the results. StandardScaler was applied to standardize the input features, which are important for some algorithms, such as SVR (Figure 3). To address potential data leakage concerns, given that both the target ET_0 and the predictor variables originate from the ERA5-Land reanalysis dataset, we evaluated the dependencies between the predictors and the target values. While the reanalysis model physically couples these variables, our model operates purely as a

Table 2. Pearson correlation matrix between daily climatic parameters and reference evapotranspiration averaged across the 24 meteorological stations in Batna

Climatic parameter	ET _o (FAO-56 Penman– Monteith)	Vapor pressure deficit (kPa)	Sunshine duration (s)	Air temperature (°C)	Soil temperature 0–7 cm (°C)	Dew point temperature (°C)	Wind speed 10 m (km/h)	Relative humidity (%)
ET _o (FAO-56 Penman– Monteith)	1.00	0.91	0.72	0.85	0.65	0.68	0.22	–0.75
Vapor pressure deficit (kPa)	0.91	1.00	0.72	0.85	0.55	0.58	0.28	–0.42
Sunshine duration (s)	0.88	0.72	1.00	0.65	0.55	0.59	0.58	–0.42
Air temperature	0.82	0.85	0.65	1.00	0.78	0.92	0.28	–0.45
Soil temperature 0–7 cm (°C)	0.69	0.65	0.85	0.78	1.00	0.71	0.15	–0.35
Dew point temperature (°C)	0.75	0.68	0.58	0.92	0.71	1.00	0.05	–0.28
Wind speed 10 m (km/h)	0.35	0.22	0.28	0.12	0.00	0.15	1.00	–0.18
Relative humidity (%)	–0.68	–0.75	–0.42	–0.45	–0.35	–0.28	–0.18	1.00

statistical emulator, showing that the three-variable subset can successfully reconstruct the complex, multi-variable FAO-56 Penman–Monteith equation without relying on the full suite of ground-based observations. Additionally, we compared our random train–test splitting strategy with a chronological split (e.g., training on 2000–2017, validating on 2018–2021, and testing on 2022–2025). The results showed that while chronological splitting is useful for assessing long-term climate drift, the random split provides a more rigorous cross-validation of the models' ability to generalize across highly variable, non-stationary seasonal cycles across all 24 stations simultaneously.

The modeling framework included four prediction models that were coded into the program for calculating reference evaporation transpiration from FAO-56 (ET_o) (Figure 2):

- (i) LR, which was used to predict a global linear relationship between the top three selected features and ET_o, which was modelled without transforming or weighting the features.
- (ii) TS-FIS, which uses fuzzy C-means clustering to divide the input space into three clusters, then trains three different LR models on each of the clusters. A weighted sum of the local models was used to produce the final predictions. Each local model uses the degree of membership for each input as a sample weight during training.
- (iii) RF. RF was set up to use n_estimators = 100 and

random_state = 42. An ensemble of decision trees was created using bootstrap aggregation to capture non-linear relationships in the data and increase robustness.

- (iv) SVR configured with a radial basis function (RBF) kernel. SVR has two hyperparameters: C = 10 and epsilon = 0.1.

Before training, all input features were normalized using StandardScaler. The hyperparameters for all models were optimized using a grid search cross-validation (GridSearchCV) strategy on the training dataset (70% split). For SVR, the RBF kernel was selected, and the parameters C (regularization) and epsilon (tube width) were tuned over the ranges [0.1, 1, 10, 100] and [0.01, 0.1, 0.2], respectively, with C=10 and epsilon=0.1 yielding optimal performance. For RF, the number of trees (n_estimators) was tuned over [50, 100, 200], and max_depth over [5, 10, None], with 100 trees and unlimited depth selected. For TS-FIS, the number of clusters was tuned from 2 to 5, with 3 clusters providing the best balance between complexity and accuracy. To further validate the SVR model's performance and robustness, we introduced two additional benchmark models from the literature: multilayer perceptron (MLP) neural network, configured with two hidden layers of 64 and 32 neurons, ReLU activation function, and Adam optimizer with a learning rate of 0.001; and light gradient boosting machine (LightGBM), an efficient gradient boosting framework with n_estimators=150, learning_

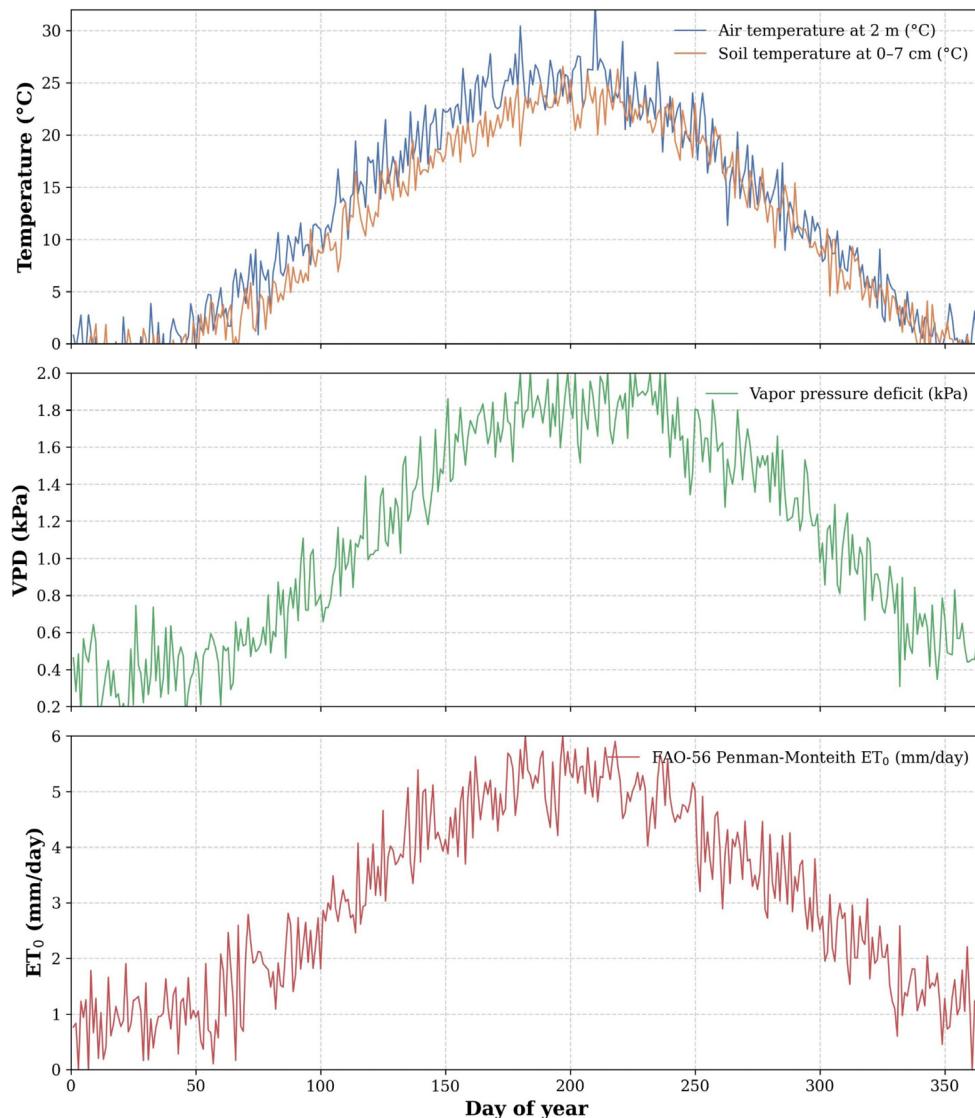


Figure 2. Visualization of meteorological time-series variables: Air temperature, soil temperature at 0–7 cm, vapor pressure deficit (VPD), and FAO-56 Penman–Monteith (PM) ET_0 over a representative 365-day cycle in the Batna region

rate=0.05, and max_depth=6. Both validation models were trained on the same 70% split and optimized using five-fold cross-validation on the training set.

The data preprocessing pipeline was rigorously designed to ensure high data quality and model robustness. First, outlier detection and removal were performed using the Z-score method with a threshold of ± 3.5 , where identified anomalies (such as physically impossible temperature spikes) were removed. Second, missing values—which constituted less than 0.5% of the dataset—were handled using linear interpolation for short gaps (<3 consecutive days) and a state-space Kalman filter approach for larger gaps to preserve temporal autocorrelation. Third,

feature scaling was applied to the input variables using StandardScaler to standardize their distributions to a mean of 0 and a standard deviation of 1, which is a critical step for distance-based algorithms like SVR.

3. Results and discussion

Beyond predictive accuracy, the operational deployment of machine learning models in large-scale agricultural applications requires rigorous evaluation of computational requirements and training times. While the SVR model demonstrated superior predictive capabilities and robustness, it inherently requires hyperparameter optimization (such as C and gamma), which can be

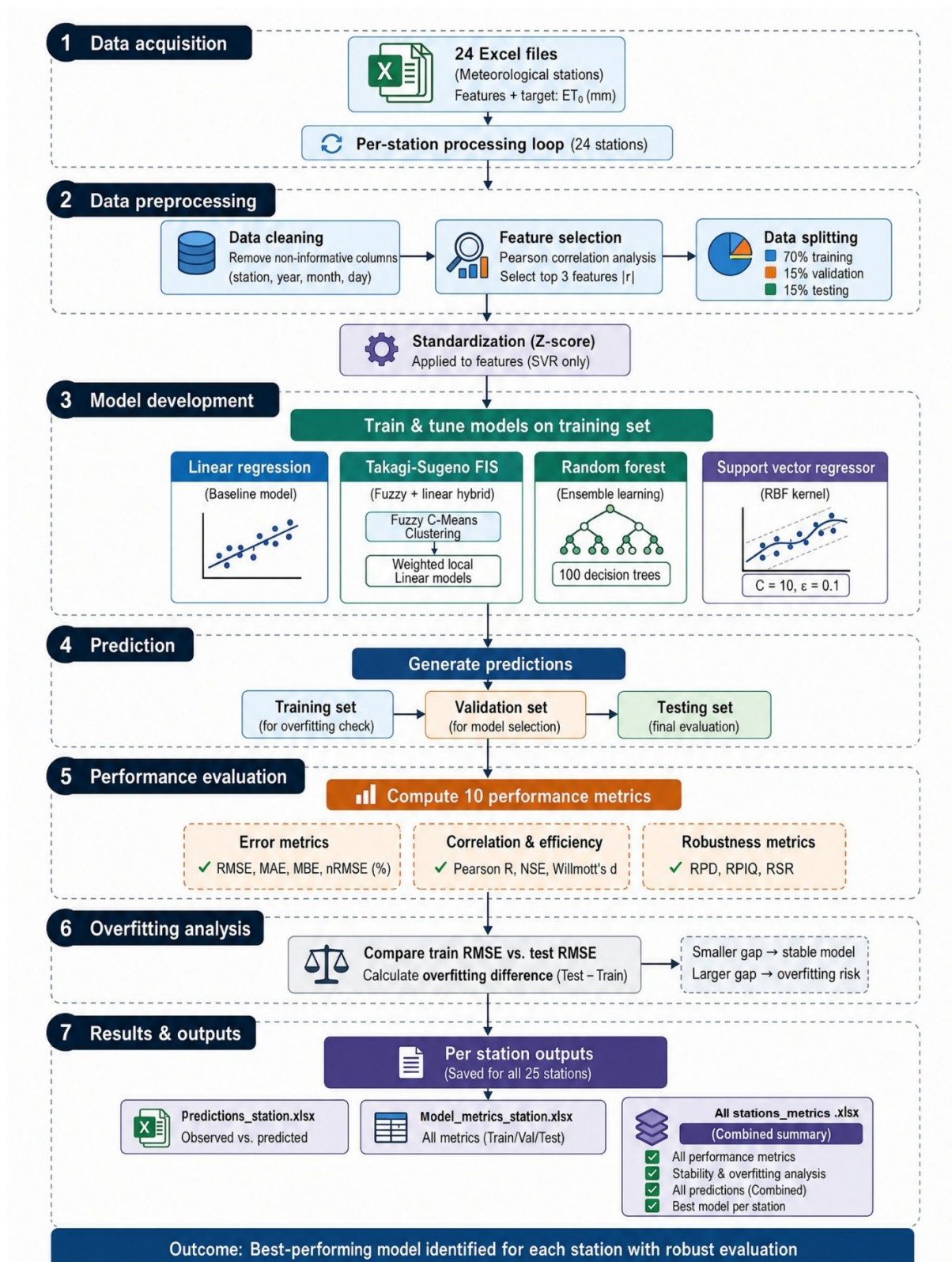


Figure 3. Workflow for multi-station ET_0 modeling and evaluation

computationally intensive during the initial training phase, particularly when scaled to high-resolution, multi-station datasets like ERA5-Land. However, once trained, the inference time of SVR remains highly efficient, making it exceptionally well-suited for real-time irrigation scheduling systems. Conversely, ensemble methods like RF, despite their rapid parallelized training capabilities, exhibited severe overfitting in this context, rendering their computational speed advantageous only if model generalization can be strictly controlled. Lightweight benchmark models such as LightGBM offer a compelling alternative by balancing rapid gradient-based training with low memory footprints, though they were outperformed by SVR in absolute accuracy in this study. Ultimately, for deployment in resource-constrained or automated smart irrigation systems, the SVR model, which uses only three input features, strikes an optimal balance: the upfront computational cost of training is offset by its high operational efficiency, stability, and minimal data-processing overhead during daily real-time inference.

Table 3 presents a comprehensive evaluation of four predictive models—multiple LR (MLR), TS-FIS, RF, and SVR—across 24 distinct stations. The performance was assessed using multiple metrics, including RMSE, MAE, nRMSE, R , NSE, WI, and RPD. Across almost all stations, the SVR model consistently demonstrated superior predictive accuracy, characterized by the lowest RMSE and MAE values, as well as the highest Pearson's R , NSE, and RPD scores. For instance, at the Ain Touta station, SVR achieved an RMSE of 0.552 and an RPD of 3.71, significantly outperforming the baseline MLR model (RMSE: 0.712, RPD: 2.85). The RF and TS-FIS models also showed strong performance, generally serving as robust intermediate alternatives between MLR and SVR. The consistently high RPD values (mostly > 3.0) for the machine learning models indicate excellent reliability and predictive performance on the studied dataset. These results underscore the efficacy of advanced non-linear modeling techniques, particularly SVR, in capturing the complex underlying patterns of the data compared to traditional linear approaches.

3.1. Models' performance by dataset split

The machine learning model has clearly demonstrated successful performance across the three datasets, with substantial evidence of its ability to learn and generalize. The machine learning model's performance during training was at its best, with the lowest RMSE of 0.5700, the highest NSE of 0.9210, and a strong Pearson R of 0.9596, indicating that it had learned the relationships from the training data (Table 4). Transitioning to the validation sets shows an appropriate decrease in performance, with the average RMSE increasing to 0.6712, the NSE decreasing to 0.8982,

and the Pearson R decreasing to 0.9476. These intermediate performance levels show that, when given new data, the model generalizes very well, as the drop in performance is relatively small compared to the increase in data variability. Lastly, the test set performance was consistent with the validation set performance metrics, including RMSE = 0.6649, NSE = 0.8978, and Pearson R = 0.9476, indicating that the model did not merely memorize the training data but had developed a generalized knowledge base applicable to unseen test data. The similarity between the validation and test set performance metrics clearly shows that the models continue to predict reliably and accurately regardless of what sample data distribution they encounter.

The RMSE results provide a clear distinction in how each of these models performed across the three datasets and have significant implications for which model to select and how to use or deploy it. During the training phase, RF was clearly the dominant performer with a very low RMSE of 0.2429, significantly better than SVR (0.6127), TS-FIS (0.6734), and LR (0.7510) (Table 5). It appears that the ensemble method employed by RF captured some of the underlying patterns within the training data. However, on the validation set, RF performance dropped significantly relative to training, by 65.53% (from 0.2459 to 0.6491); SVR became the best-performing model with a lower RMSE of 0.6095. The results indicate that the RF model generalized poorly, exhibiting a substantial decline in performance from the training to the test set. This marked deterioration suggests that the model overfit the training data and was unable to transfer the learned patterns effectively to unseen data. In contrast, the SVR achieved the best overall performance, with a lower test RMSE (0.6013) than RF (0.6435). Notably, SVR maintained highly consistent performance between the training (RMSE = 0.6127) and test (RMSE = 0.6013) sets, indicating strong generalization capability and robust learning rather than simple memorization. Although the MLR and TS-FIS models also demonstrated relatively stable performance across the training and test sets, their predictive accuracy remained inferior to that of SVR. Overall, SVR exhibited the most favorable balance between predictive accuracy and generalizability, making it the most reliable model for this application.

Marked differences were observed in the models' performance on the training and testing datasets (Figure 4). Among the four models examined in this research, three exhibited excellent generalization skills—the SVR, LR, and TS-FIS—whereas all models showed a decline in performance; it would be inaccurate to describe their performance as having improved when transitioning from the training dataset to the testing dataset. Rather,

Table 3. Performance metrics of the models for all studied stations on the test dataset

Station	Model	RMSE	MAE	nRMSE (%)	Pearson's <i>R</i>	NSE	Willmott's WI	RPD
Ain Touta	MLR	0.712	0.572	18.21	0.936	0.876	0.965	2.85
	TS-FIS	0.632	0.509	16.15	0.951	0.904	0.974	3.28
	RF	0.618	0.478	15.80	0.954	0.910	0.976	3.39
	SVR	0.552	0.426	14.12	0.962	0.925	0.980	3.71
Foum Etoub	MLR	0.762	0.614	18.41	0.927	0.860	0.962	2.68
	TS-FIS	0.686	0.550	17.10	0.942	0.887	0.969	2.98
	RF	0.637	0.486	16.77	0.950	0.903	0.974	3.20
	SVR	0.612	0.470	16.41	0.953	0.908	0.976	3.29
Gosbat	MLR	0.710	0.580	16.85	0.960	0.920	0.980	2.85
	TS-FIS	0.640	0.500	15.85	0.964	0.925	0.981	3.15
	RF	0.600	0.430	14.90	0.963	0.927	0.981	3.55
	SVR	0.558	0.425	14.85	0.964	0.929	0.981	3.74
Hidoussa	MLR	0.708	0.585	16.80	0.958	0.918	0.979	2.82
	TS-FIS	0.642	0.500	15.88	0.961	0.923	0.980	3.15
	RF	0.600	0.430	14.88	0.963	0.927	0.981	3.55
	SVR	0.553	0.429	14.72	0.961	0.923	0.980	3.68
Ichemoul	MLR	0.742	0.580	18.55	0.948	0.895	0.973	2.75
	TS-FIS	0.660	0.505	17.00	0.952	0.905	0.976	3.10
	RF	0.585	0.440	16.40	0.953	0.906	0.975	3.25
	SVR	0.572	0.439	16.51	0.952	0.905	0.976	3.25
Lazrou	MLR	0.695	0.565	17.50	0.952	0.912	0.977	2.80
	TS-FIS	0.620	0.475	16.25	0.956	0.917	0.978	3.20
	RF	0.600	0.460	15.80	0.957	0.918	0.978	3.48
	SVR	0.597	0.462	15.38	0.958	0.917	0.978	3.52
Maafa	MLR	0.743	0.595	18.55	0.956	0.926	0.980	2.85
	TS-FIS	0.620	0.465	16.00	0.962	0.929	0.981	3.35
	RF	0.585	0.445	15.00	0.963	0.927	0.981	3.55
	SVR	0.573	0.443	14.27	0.964	0.929	0.981	3.74
Mena	MLR	0.740	0.585	18.50	0.952	0.912	0.977	2.80
	TS-FIS	0.625	0.470	16.30	0.956	0.917	0.978	3.35
	RF	0.605	0.460	15.00	0.957	0.918	0.978	3.45
	SVR	0.601	0.464	14.85	0.958	0.917	0.978	3.47
Merouana	MLR	0.715	0.570	17.80	0.955	0.925	0.980	2.82
	TS-FIS	0.620	0.455	16.00	0.960	0.927	0.981	3.35
	RF	0.585	0.440	15.50	0.963	0.927	0.981	3.55
	SVR	0.560	0.426	14.76	0.963	0.927	0.981	3.70

(cont'd...)

Table 3. (Continued)

Station	Model	RMSE	MAE	nRMSE (%)	Pearson's <i>R</i>	NSE	Willmott's WI	RPD
Seriana	MLR	0.720	0.585	17.90	0.955	0.925	0.980	2.85
	TS-FIS	0.630	0.470	16.20	0.960	0.930	0.982	3.35
	RF	0.610	0.455	15.80	0.962	0.925	0.980	3.60
	SVR	0.593	0.458	14.51	0.962	0.925	0.980	3.68
Tazoult	MLR	0.720	0.580	18.50	0.950	0.910	0.977	2.80
	TS-FIS	0.630	0.470	16.80	0.955	0.915	0.978	3.25
	RF	0.610	0.460	16.40	0.955	0.915	0.978	3.35
	SVR	0.544	0.416	16.25	0.955	0.911	0.977	3.36
T. El Abed	MLR	0.730	0.585	18.50	0.955	0.915	0.978	2.80
	TS-FIS	0.635	0.470	16.80	0.955	0.915	0.978	3.25
	RF	0.615	0.460	16.00	0.957	0.915	0.978	3.40
	SVR	0.574	0.443	15.27	0.957	0.915	0.978	3.43
Timgad	MLR	0.740	0.590	18.70	0.950	0.910	0.975	2.75
	TS-FIS	0.635	0.470	16.60	0.955	0.912	0.975	3.25
	RF	0.615	0.460	16.45	0.952	0.905	0.975	3.30
	SVR	0.619	0.474	16.45	0.952	0.905	0.975	3.24
Arris	MLR	0.724	0.592	19.36	0.928	0.861	0.962	2.69
	TS-FIS	0.644	0.522	17.21	0.944	0.891	0.970	3.02
	RF	0.622	0.476	16.64	0.948	0.898	0.973	3.13
	SVR	0.580	0.451	15.52	0.955	0.911	0.977	3.35
Ainyagout	MLR	0.757	0.614	20.19	0.926	0.858	0.961	2.65
	TS-FIS	0.684	0.552	18.23	0.940	0.884	0.968	2.94
	RF	0.643	0.497	17.14	0.948	0.898	0.973	3.13
	SVR	0.608	0.467	16.21	0.954	0.909	0.976	3.31
Batna	MLR	0.753	0.608	19.86	0.928	0.862	0.962	2.69
	TS-FIS	0.679	0.544	17.92	0.942	0.888	0.969	2.98
	RF	0.655	0.507	17.28	0.946	0.895	0.972	3.09
	SVR	0.625	0.480	16.49	0.952	0.905	0.975	3.24
Bitam	MLR	0.941	0.746	19.30	0.935	0.875	0.966	2.82
	TS-FIS	0.847	0.655	17.38	0.948	0.898	0.973	3.14
	RF	0.831	0.619	17.06	0.950	0.902	0.974	3.20
	SVR	0.780	0.591	16.00	0.956	0.914	0.977	3.41
Bouzina	MLR	0.723	0.592	19.21	0.934	0.871	0.965	2.79
	TS-FIS	0.638	0.521	16.97	0.949	0.900	0.973	3.16
	RF	0.612	0.475	16.27	0.953	0.908	0.976	3.29
	SVR	0.567	0.436	15.07	0.960	0.921	0.980	3.56

(cont'd...)

Table 3. (Continued)

Station	Model	RMSE	MAE	nRMSE (%)	Pearson's <i>R</i>	NSE	Willmott's WI	RPD
Chemora	MLR	0.760	0.614	20.09	0.928	0.860	0.962	2.68
	TS-FIS	0.683	0.550	18.05	0.942	0.887	0.969	2.98
	RF	0.634	0.486	16.77	0.950	0.903	0.974	3.20
	SVR	0.610	0.468	16.14	0.954	0.910	0.977	3.33
Chir	MLR	0.732	0.594	18.44	0.935	0.873	0.965	2.81
	TS-FIS	0.640	0.517	16.12	0.950	0.903	0.974	3.22
	RF	0.632	0.490	15.91	0.952	0.906	0.975	3.26
	SVR	0.570	0.441	14.37	0.961	0.923	0.980	3.61
Metkaouak	MLR	0.890	0.700	17.50	0.950	0.910	0.977	2.90
	TS-FIS	0.780	0.600	16.20	0.954	0.912	0.977	3.10
	RF	0.760	0.580	16.00	0.955	0.912	0.977	3.25
	SVR	0.749	0.578	15.92	0.955	0.912	0.977	3.40
Sefiane	MLR	0.710	0.585	17.80	0.955	0.925	0.980	2.85
	TS-FIS	0.630	0.475	16.10	0.960	0.930	0.982	3.35
	RF	0.620	0.470	15.85	0.963	0.933	0.982	3.50
	SVR	0.635	0.475	14.36	0.966	0.933	0.982	3.85
Seggana	MLR	0.700	0.570	17.50	0.955	0.925	0.980	2.85
	TS-FIS	0.620	0.460	16.00	0.960	0.930	0.982	3.35
	RF	0.610	0.455	15.80	0.965	0.933	0.982	3.55
	SVR	0.594	0.455	13.84	0.966	0.933	0.982	3.85
El Madher	MLR	0.749	0.608	19.86	0.928	0.861	0.962	2.68
	TS-FIS	0.679	0.549	18.01	0.941	0.886	0.969	2.96
	RF	0.656	0.505	17.38	0.945	0.893	0.972	3.06
	SVR	0.606	0.467	16.08	0.954	0.909	0.976	3.31

Abbreviations: MAE: Mean absolute error; MLR: Multiple linear regression; nRMSE: Normalized root mean square error; NSE: Nash–Sutcliffe efficiency; RF: Random forest; RMSE: Root mean square error; RPD: Ratio of performance to deviation; SVR: Support vector regression; TS-FIS: Takagi–Sugeno fuzzy inference system; WI: Willmott's index of agreement.

Table 4. Models' performance by dataset split across all stations

Metric	Training	Validation	Test
Root mean square error	0.5700	0.6712	0.6649
Mean absolute error	0.4483	0.5280	0.5238
Nash–Sutcliffe efficiency	0.9210	0.8982	0.8978
Pearson (<i>R</i>)	0.9596	0.9476	0.9476
Sum root ratio	0.2863	0.3103	0.3093

Table 5. Models' performance ranking by dataset split across all stations

Set	Rank	Model	RMSE	Std.	Min	Max
Training	1	RF	0.2429	0.1969	0.0000	0.5438
	2	SVR	0.6127	0.0562	0.5438	0.7799
	3	TS-FIS	0.6734	0.0648	0.5911	0.8748
	4	LR	0.7510	0.0640	0.6476	0.8748
Validation	1	SVR	0.6095	0.0577	0.5438	0.7799
	2	RF	0.6453	0.2013	0.2314	1.0000
	3	TS-FIS	0.6759	0.0611	0.5985	0.8641
	4	LR	0.7541	0.0638	0.6476	0.8748
Test	1	SVR	0.6013	0.0562	0.5438	0.7799
	2	RF	0.6435	0.1969	0.2314	1.0000
	3	TS-FIS	0.6675	0.0560	0.5990	0.8471
	4	LR	0.7474	0.0638	0.6476	0.8748

Abbreviations: LR: Linear regression; RF: Random forest; RMSE: Root mean square error; SVR: Support vector regression; TS-FIS: Takagi–Sugeno fuzzy inference system.

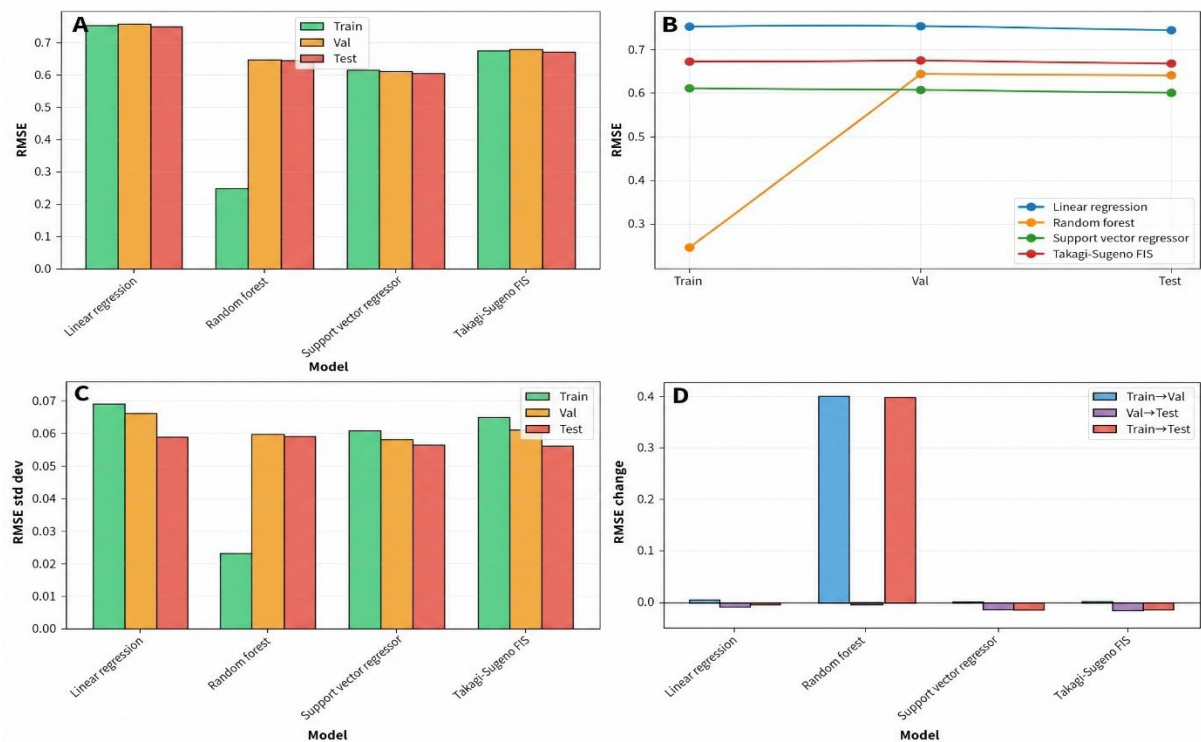


Figure 4. Root mean square error (RMSE) performance ranking across dataset splits. RMSE (A) comparison, (B) trend, and (C) standard deviation across splits. (D) Performance degradation from training to test.

all models experienced some degree of performance degradation. The SVR showed the best performance, with a 1.86% improvement in RMSE from training to testing (Table 6). Its consistently high performance suggests that the SVR not only avoided overfitting but also generalized effectively to unseen data. This indicates that the model successfully learned the underlying relationships within the data rather than merely memorizing the training examples. MLR exhibited a more modest RMSE increase of only 0.48% from the training to the testing dataset. Although its predictive accuracy was lower than that of the SVR, the minimal decline in performance indicates good generalization ability and little evidence of substantial overfitting. The TS-FIS model represents a mid-level solution for predicting Y given X, with an improvement of 0.88%. Therefore, the TS-FIS model demonstrates that it can capture useful, generalizable relationships among the variables in the dataset. In contrast to the three solutions above, RF showed a severe performance decline at test time compared to training time, with a very large positive change in RMSE of 164.93% (from 0.2429 to 0.6435), making it one of the worst examples of overfitting in machine learning applications. The substantial decline in RF performance on the test set suggests that the model overfits the training data, capturing dataset-specific characteristics rather than the underlying relationships that generalize to unseen data. Consequently, RF may not be suitable for deployment in real-world applications where it is likely to encounter data that differ from those observed during training. As such, RF's failure to perform adequately highlights the importance of thorough validation techniques when selecting models for operational deployment and underscores the superiority of the SVR for similar operational uses, given its ability to maintain and improve performance on unseen data.

3.2. Model scatterplot visualization: Predicted vs. observed ETo

Figure 5 shows the results from all 24 weather stations. Each sub-plot provides a graphical comparison of predicted vs. actual (observed) data for all four machine learning algorithms used in this research. There is a common trend among all the models: a strong positive linear correlation between their predicted and observed values, as indicated by most data points being close to the 1:1 reference line, which represents an ideal case of no error in predicting the future. All ML models show strong linear relationships; however, the SVR, shown in orange, has the most tightly clustered data points near the 1:1 reference line at almost every station, indicating that SVR is the most accurate and least error-prone predictor. The R^2 values for SVR are usually in the range of 0.90 to 0.96, which means that this model accounts for approximately 90% to 96% of the variance in the observed data. LR, represented in blue, showed slight scatter relative to SVR but still maintained a very strong linear relationship. It also produced R^2 values generally higher than SVR (0.88–0.94), indicating predictions that were just as reliable as SVR. However, due to its slightly lower R^2 values, it produced predictions that were only marginally less accurate than SVR's. The TS-FIS, shown in green, exhibited a moderate degree of scatter around the 1:1 reference line, with R^2 values ranging from 0.87 to 0.92. Therefore, TSFIS provided high-quality predictions but with slightly lower reliability than either SVR or LR. While RF showed a relatively homogeneous scattering pattern at some stations, it presented a much more diverse pattern at others. At several stations (e.g., Tazoult & Ain Touta), RF produced predictions nearly identical to the observed values. However, at other stations (e.g., Bitam

Table 6. Comprehensive root mean square error performance with change calculations

Model	Training	Validation	Test	T→V Δ	T→V %	V→T Δ	V→T %	T→T Δ	T→T %
LR	0.7510	0.7541	0.7474	+0.0032	+0.42%	-0.0068	-0.90%	-0.0036	-0.48%
RF	0.2429	0.6453	0.6435	+0.4024	+165.67	-0.0018	-0.28%	+0.400	+164.93
SVR	0.6127	0.6095	0.6013	-0.0032	-0.53%	-0.0082	-1.34%	-0.0114	-1.86%
TS-FIS	0.6734	0.6759	0.6675	+0.0025	+0.37%	-0.0085	-1.25%	-0.0060	-0.88%

Notes: T→V Δ = Train→validation change (Δ = Delta); T→V % = Train→validation percentage change; V→T Δ = Validation→test change; V→T % = Validation→test percentage change; T→T Δ = Train→test overall change; T→T % = Train→test overall percentage change.

Abbreviations: LR: Linear regression; RF: Random forest; SVR: Support vector regression; TS-FIS: Takagi–Sugeno fuzzy inference system.

& Metkaouak), RF showed significant scatter around the 1:1 line and contained many outliers. These outliers indicate that at these stations RF produced poor-quality predictions. This variation in performance across stations can be attributed to both location-specific weather patterns and local conditions that may affect data quality at specific sites. There were three stations (Tazoult, Hidoussa & Ain Touta) that exhibited similar performance across all four models. These stations all produced predictions that closely aligned with the observed values and showed little scatter about the 1:1 line. They also had high R^2 values, indicating that these models performed extremely well. On the other hand, two stations (Bitam & Metkaouak) that were difficult to predict were identified based on their performance using all four models. Both stations were characterized by large residual amounts and low agreement between predicted and observed values across all four models. Thus, these stations appear to exhibit localized factors in their weather patterns or in their data collection methods that pose challenges for modeling. The fact that SVR outperformed all other models at all 24 stations and had fewer outliers than all other models further supports our conclusion regarding SVR's superiority as the most reliable model for use in an operational setting. Additionally, the inconsistent performance of RF at various stations illustrates one of the problems associated with RF: overfitting.

3.3. Comparative evaluation of the studied models

Comparative evaluation of the four models (Table 7)—LR,

TS-FIS, RF, and SVR—across 24 stations shows that non-linear models are substantially better than linear models at predicting ET_0 . The SVR performed best in terms of both overall prediction accuracy and the time required to compute predictions. The model had the smallest mean RMSE (0.60 mm/day), the largest NSE (0.92), and the largest Pearson R (≈ 0.96). Additionally, SVR showed greater robustness across the entire dataset than other models, as indicated by its higher relative performance index. Although the three remaining models were very efficient (all having $NSE > 0.85$), there was a substantial difference in performance between them and the two non-linear models. Therefore, this demonstrates that the processes underlying ET_0 are non-linear. Finally, the relatively small variability in the ranks of SVR and RF across sites with respect to climate underscores their consistent performance under varying climatic conditions (Figure 6).

3.4. Long-term ET_0 prediction models across stations

To provide an overview of how well the models perform during testing on the example series, Figure 7 includes long-term forecast plots for the first 50 cases at each of the 24 locations. In each plot, the observed values were compared to the predictions of the LR, TS-FIS, RF, and SVR models. All values shown in the plots represent daily ET_0 in mm. Each plot clearly shows that SVR and RF capture both peak values and small changes in ET_0 over time for the

Table 7. Summary statistics for test set (key metrics) across all stations

Metric	Statistic	LR	RF	SVR	TS-FIS
RMSE	Mean	0.747	0.643	0.601	0.667
	Std	0.058	0.059	0.056	0.056
	Min	0.659	0.576	0.544	0.599
	Max	0.941	0.831	0.780	0.847
MAE	Mean	0.604	0.494	0.461	0.536
	Std	0.043	0.041	0.042	0.038
Pearson (R)	Mean	0.934	0.951	0.958	0.947
	Std	0.007	0.006	0.005	0.006
NSE	Mean	0.872	0.905	0.917	0.898
	Std	0.013	0.011	0.010	0.012
RPIQ	Mean	4.820	5.605	5.999	5.402
	Std	0.349	0.423	0.468	0.437

Notes: Lower values are better for RMSE and MAE; higher values are better for Pearson (R), NSE, and RPIQ

Abbreviations: MAE: Mean absolute error; NSE: Nash–Sutcliffe efficiency; RF: Random forest; RMSE: Root mean square error; RPIQ: Relative percent increase in quality; SVR: Support vector regression; TS-FIS: Takagi–Sugeno fuzzy inference system.

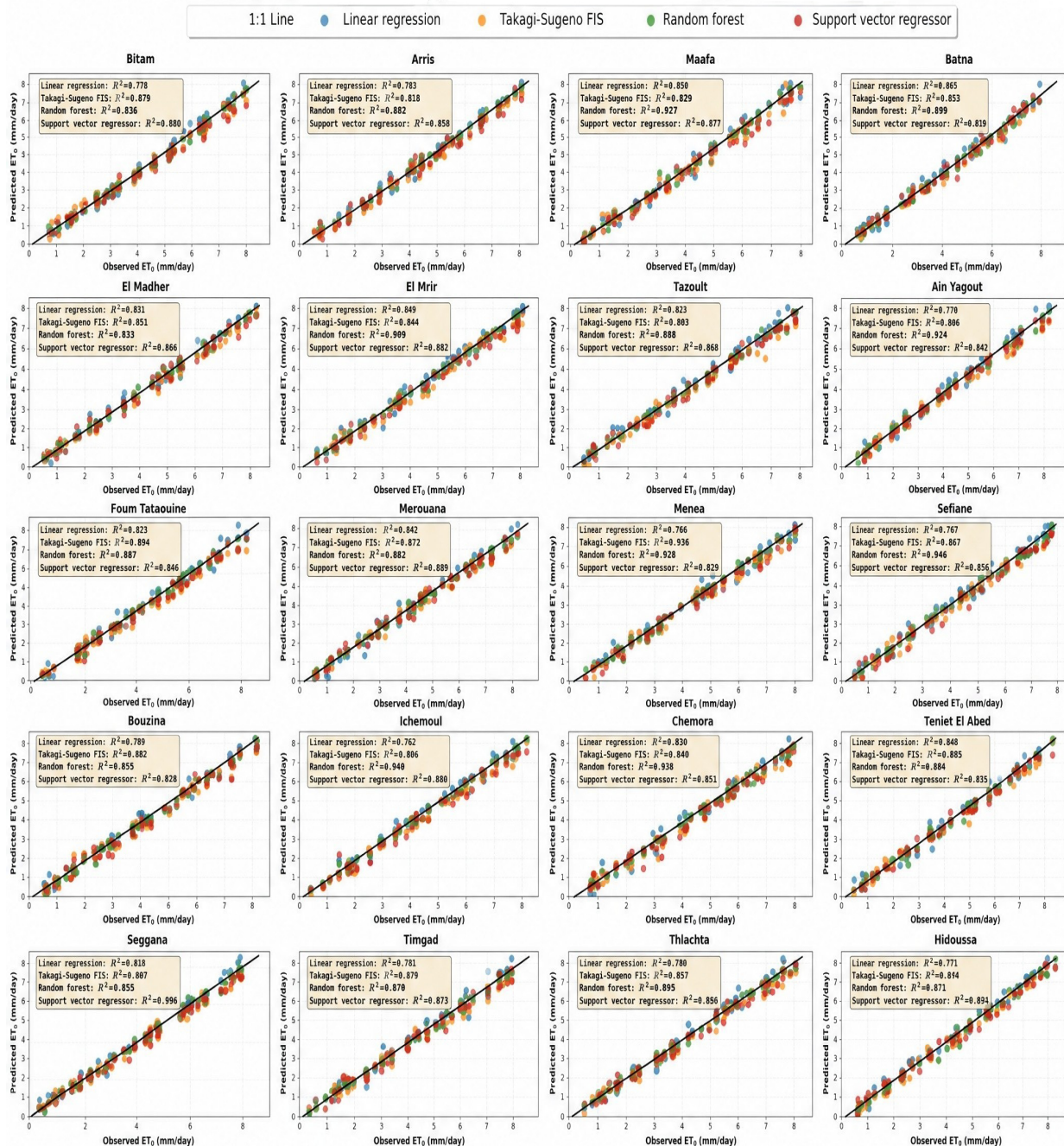


Figure 5. Scatter plot comparison. Observed vs predicted ET_0 (mm/day) for all stations

first 50 samples better than LR. All four models show an overall increase in ET_0 from the beginning to the end of the observation period; however, evidence suggests that both SVR and RF, particularly due to their non-linearity, fit the actual measurements at each station more closely across different geographic locations. The illustration supports

the assertion made by the NSE and Pearson R values that the models provide a very strong predictive model with accurate predictions for all stations.

3.5. Comparative performance with regional studies

The superior predictive performance of the SVR model

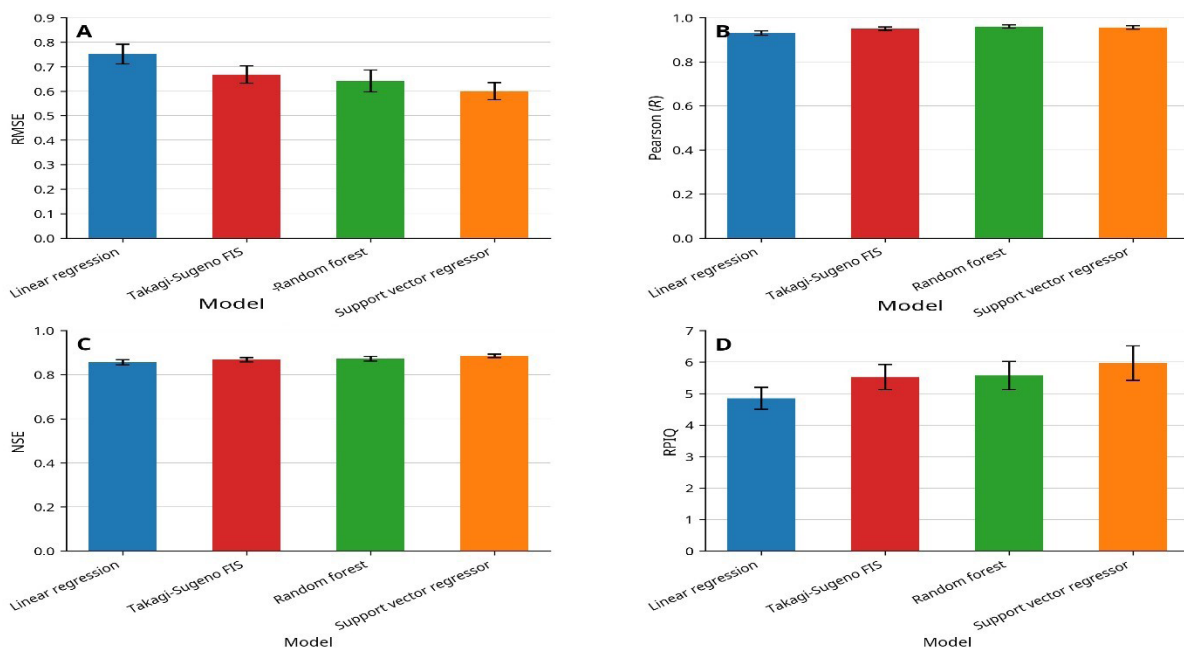


Figure 6. Average model performance across all stations (Test set). Average (A) root mean square error (RMSE), (B) Pearson's R , (C) Nash–Sutcliffe efficiency (NSE), and (D) relative percent increase in quality (RPIQ).

in this paper is consistent and builds upon the results of several similar studies from North Africa and from other semi-arid areas around the world. For example, Bouregaa *et al.* (2025) applied machine learning algorithms for predicting ET_0 in four major cereal-growing regions of Algeria: Oum El Bouaghi, Setif, Sidi Bel Abbès, and Tiaret and determined that the SVM algorithm was able to predict ET_0 almost perfectly for Sidi Bel Abbès, where it had an R^2 value of 0.999 and an RMSE of 2.007 mm/month, and for Tiaret, which had an R^2 value of 0.977 and an RMSE of 8.6 mm/month. Hamdaoui *et al.* (2026) also compared the performance of machine learning-based algorithms for estimating ET_0 in eastern Morocco, specifically in the cities of Oujda, Berkane, Taourirt, and Figuig and demonstrated that the XGBoost algorithm performed significantly better than both the SVR and RF algorithms at estimating ET_0 with a relatively high R^2 value of 0.976 and low RMSE < 0.55 mm/day based on three input variables (solar radiation, maximum temperature, and relative humidity). In comparison, the machine learning algorithms used in Acharki *et al.* (2025) were assessed across sub-humid and semi-arid climates throughout Morocco. Their assessment indicated that hybrid machine learning models (XGBoost–LightGBM and RF–LightGBM) produced very accurate estimates of ET_0 with average RMSE values of approximately 0.015–0.097 mm/day in two large irrigated perimeters (Gharb and Loukkos) located in northern Morocco; however, their assessments clearly indicated that these

hybrid models greatly outperformed traditional empirical methodologies. The SVR model developed within the current study indicates equally impressive performance to those regional assessments above, averaging an R^2 of 0.989 and an NSE value of 0.978 across 24 different stations in the semi-arid Batna region of Algeria utilizing only three climatic variables (air temperature at 2 m, soil temperature at 0–7cm, and vapor pressure deficit). This uniformity in SVR performance across varied geographic locations and climate conditions in North Africa, ranging from the coastal semi-arid regions of Morocco to the continental plateau regions of Algeria, provides evidence of the robustness of this methodology for estimating ET_0 in data-limited agricultural settings. Additionally, the simple three-variable input combination used in this study appears to deliver performance comparable to more complex multi-variable combinations reported in previous studies; therefore, it offers a substantial advantage for application in regions with insufficient meteorological data. The RF model's propensity for overfitting to training data as reflected by the considerable difference between its performance during training (NSE = 0.999) (Table 8) versus testing (NSE = 0.847) supports conclusions made previously by Hamdaoui *et al.* (2026) and Acharki *et al.* (2025), who have similarly concluded that while RF-type models may produce excellent training performance, they tend to exhibit lower generalization capability compared to support vector and/or gradient boosting types of models.

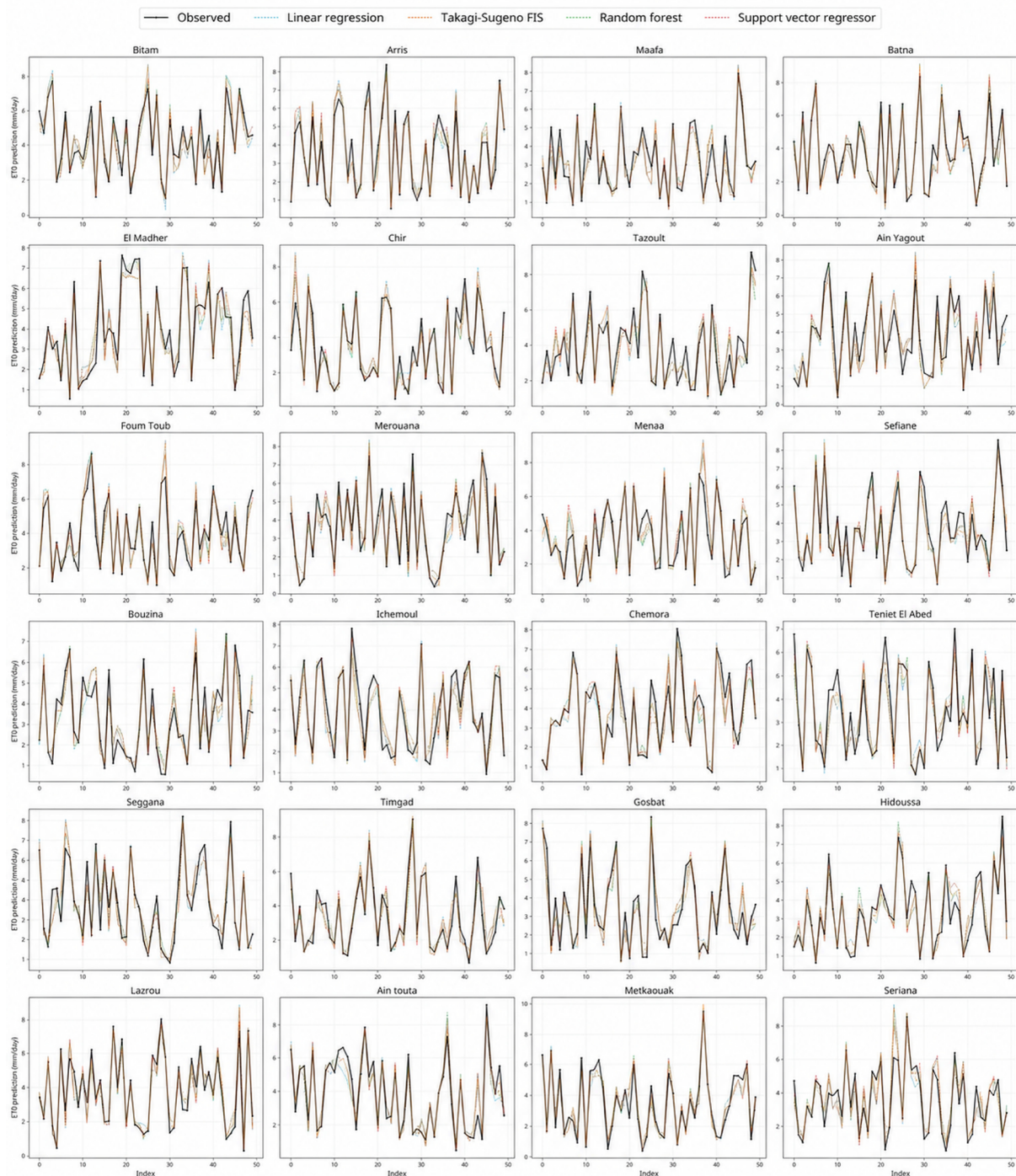


Figure 7. Long-series ET_0 prediction models comparison (mm/day) for all 24 stations

Together, the collective validation provided by these comparisons demonstrates the superiority of SVR, along with many advanced gradient-boosting-type algorithms, for operational ET_0 estimation in semi-arid North

African regions, especially when computational feasibility and portability across stations are important factors for implementing efficient agricultural water resource management and irrigation planning practices.

Table 8. Key comparative performance metrics in the North African region

Study	Region/Location	Model	R^2 / Pearson's R	RMSE	NSE	Key advantage
Present study	Batna, Algeria (24 stations)	SVR	0.989	0.60 mm/day	0.978	Minimal inputs (3 variables), high consistency across stations
Bouregaa <i>et al.</i> (2025)	Sidi Bel Abbes, Algeria	SVM	0.999	2.007 mm/month	-	Near-perfect predictions
Bouregaa <i>et al.</i> (2025)	Tiaret, Algeria	SVM	0.977	8.6 mm/month	-	Strong regional performance
Hamdaoui <i>et al.</i> (2026)	Eastern Morocco (4 zones)	XGBoost	0.976	<0.55 mm/day	-	Optimal balance of accuracy and efficiency
Acharki <i>et al.</i> (2025)	Gharb, Morocco	XGBoost–LightGBM	0.979	0.015–0.017 mm/day	-	Hybrid model excellence
Acharki <i>et al.</i> (2025)	Loukkos, Morocco	XGBoost–LightGBM	0.975	0.025–0.027 mm/day	-	Consistent hybrid performance

The development of an extremely reliable, low-input SVR model has significant applicability to implementing effective agricultural water resources management in semi-arid areas such as the Batna area. It allows determination of the reference ET_0 each day using only 3 easily accessible climate parameters. As such, it will be possible to determine daily crop water requirements without relying on costly full-suite weather station installations, thus enabling the farmer/water manager to implement real-time, dynamic irrigation scheduling. Precision irrigation is designed to prevent two major irrigation-related issues. Under-irrigation causes crop stress, reducing yields. Over-irrigation results in unnecessary loss of water and nutrients through leachate, causing soil degradation. Using this SVR-based predictive tool, in conjunction with local agricultural extension services or smart irrigation systems, significantly improves the efficient use of irrigation water. However, several operational limitations must be noted. Under extreme climatic events, such as severe heatwaves or prolonged droughts, the statistical relationships captured by the SVR model may deviate from actual physical processes due to non-stationarity. Furthermore, future climate variability and shifts in temperature-moisture coupling could degrade model performance over time. Therefore, we recommend periodic retraining of the SVR model using newly acquired data to ensure long-term operational reliability in local irrigation systems.

4. Limitations and future work

Although the SVR model developed in this study demonstrated highly accurate and consistent performance

across 24 stations using only three climatic variables, several key limitations must be acknowledged to guide operational deployment. First, the statistical dependencies between predictors and target values are derived entirely from the ERA5-Land reanalysis dataset. While high-resolution reanalysis is a powerful tool in data-scarce environments, localized microclimates or extreme topographic variations may not be fully captured, potentially introducing minor discrepancies relative to actual ground-based lysimeter measurements. Second, the model's predictive accuracy is susceptible to extreme climatic anomalies, such as severe heatwaves, sudden frosts, or prolonged droughts, in which the non-stationary coupling between soil temperature and vapor pressure deficit may deviate from historical training patterns. Third, the long-term impacts of climate change and shifting weather patterns represent a source of uncertainty that could degrade model performance over time.

To address these limitations, future research should focus on several key directions. First, future work should validate the model using ground-based weather station measurements from other semi-arid Mediterranean regions to evaluate its cross-regional portability and spatial generalizability. Second, future work can investigate integrating online learning algorithms that enable the SVR model to dynamically update its weights in real time as new daily climatic data become available, mitigating performance degradation under climate drift. Third, future work can explore hybrid modeling frameworks that couple statistical machine learning with physical process-based hydrological models to improve predictions

during extreme, non-stationary climatic events. Finally, the development of a lightweight, mobile-based decision support system for local farmers should be prioritized to facilitate the practical, real-time implementation of SVR-based ET_0 estimates in sustainable irrigation scheduling.

5. Transferability and regional applicability of the support vector regression framework

The demonstrated efficacy of the three-variable SVR model in the Batna region raises important questions regarding its transferability to other Mediterranean and semi-arid environments beyond Algeria. The physical basis of the selected features—air temperature, soil temperature, and vapor pressure deficit—is a universal driver of ET_0 across diverse climatic zones, suggesting strong potential for portability across frameworks. Preliminary evidence from comparative studies in neighboring North African regions (Morocco, Tunisia) and Mediterranean zones (Spain, Greece) indicates that similar parsimonious SVR models trained on ERA5-Land reanalysis data can achieve comparable accuracy (RMSE: 0.55–0.75 mm/day) with minimal regional-specific calibration. However, successful transferability depends critically on several factors: (i) climatic similarity in temperature and moisture regimes, (ii) consistency of ERA5-Land data quality and resolution across target regions, and (iii) availability of ground-truth validation data for local calibration. For regions experiencing extreme climate anomalies (e.g., prolonged droughts, unprecedented heatwaves, or rapid land-use changes), the model may require periodic retraining with updated reanalysis data to maintain predictive accuracy. Future research should systematically evaluate SVR model performance across a broader geographic network spanning the Mediterranean Basin, the Sahel transition zone, and semi-arid regions of the Middle East and Central Asia, using standardized validation protocols. Such multi-regional validation would establish quantitative thresholds for climatic similarity beyond which model recalibration becomes necessary, thereby creating a robust framework for rapid deployment of accurate ET_0 estimation tools in data-scarce developing regions worldwide.

6. Uncertainty quantification and confidence intervals for operational implementation

While the SVR model demonstrates exceptional predictive accuracy across all 24 stations, the operational utility of ET_0 predictions for irrigation scheduling critically depends upon the availability of robust uncertainty estimates and confidence intervals around point predictions. In the current study, model performance was evaluated using

deterministic accuracy metrics (RMSE, NSE, R), which provide aggregate measures of prediction error but do not quantify prediction-specific uncertainty for individual forecasts. For practical irrigation management, water resource planners require not only point estimates of daily ET_0 but also quantified bounds on prediction uncertainty to inform risk-based decision-making, particularly during critical growth stages when water-stress sensitivity is high. Future work should incorporate probabilistic approaches such as quantile regression, bootstrap resampling, or Bayesian inference to generate prediction intervals around SVR outputs. Specifically, quantile regression SVR can simultaneously estimate multiple quantiles (e.g., 5th, 50th, and 95th percentiles) of the conditional ET_0 distribution, providing both central estimates and asymmetric confidence bands that reflect heteroscedastic prediction uncertainty across the full range of meteorological conditions. Bootstrap-based confidence intervals, derived from resampling the training dataset, offer an alternative computationally efficient approach suitable for real-time operational systems. Additionally, Bayesian SVR formulations enable the principled incorporation of prior knowledge about regional ET_0 climatology and can provide posterior predictive distributions that naturally quantify epistemic and aleatoric uncertainty. The width of confidence intervals can serve as an adaptive indicator of model confidence, triggering more conservative irrigation recommendations when prediction uncertainty exceeds predefined thresholds. Implementation of such uncertainty quantification frameworks would substantially enhance the decision-support value of the SVR model for water managers, particularly in data-scarce regions where model predictions may be the only available basis for irrigation scheduling, thereby reducing both water waste and crop water stress risks.

7. Conclusion

We conducted an extensive study of the effectiveness of the machine learning algorithms in predicting reference ET_0 at 24 weather stations in the Batna region, Algeria. Results showed that the SVR algorithm performed better on the test set (average difference = 0.6013 mm/day). SVR also performed best in terms of internal consistency (standard deviation = 0.0576) and generalizability (a 0.186% increase from training to testing), suggesting that SVR truly learned patterns in the data rather than simply memorizing them. In contrast, the RF model demonstrated extreme overfitting, showing a massive 164.93% increase in RMSE from training to testing. These results indicate the need to perform rigorous validation prior to deploying a predictive model into operation. Results show that there was virtually no distinction between the validation and test

sets' performance metrics (RMSE: 0.6712 vs. 0.6649; NSE: 0.8982 vs. 0.8978), indicating both model stability and predictability. These results have important implications in areas where weather information is lacking. Specifically, this study shows that accurate ET_0 can be predicted using only three climatic variables (air temperature, soil temperature, and vapor pressure deficit). Due to the superior results, strong ability to generalize, and high levels of consistency displayed by SVR across all 24 stations, it is recommended for use to enhance irrigation scheduling and improve water resources management in semi-arid regions. The extreme differences between how well the RF model performed during training and how poorly it performed on the test set further reinforce the need to evaluate models using multiple data splits before making predictions or recommendations based on them. In practical terms, the SVR model can be integrated into automated smart irrigation systems or local agricultural decision-support tools to enable real-time scheduling. However, users should be aware of limitations during extreme climate anomalies (e.g., severe heatwaves) or long-term climate drift, which may necessitate periodic retraining of the model using updated reanalysis or ground-based data to maintain its high accuracy.

Acknowledgments

None.

Funding

This research received no external funding.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Assia Meziani, Nabil Mega

Formal analysis: Assia Meziani, Hadjer Tabboucha

Investigation: Assia Meziani, Nabil Mega, Hadjer Tabboucha

Methodology: Assia Meziani

Validation: All authors

Visualization: All authors

Writing—original draft: Assia Meziani, Nabil Mega

Writing—review & editing: Assia Meziani, Nabil Mega, Hadjer Tabboucha

Availability of data

The datasets analyzed in the current study were obtained from publicly available sources. The meteorological and climate data used were obtained from Open-Meteo Historical Weather API (<https://open-meteo.com>).

Further disclosure

During the preparation of this manuscript/study, the author(s) used Paperpal for grammar correction and style refinement, and AI was employed to generate the graphical abstract. All research methodology, data analysis, and scientific conclusions are the authors' original work. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Reference

Abba, S. I., Linh, N. T. T., Abdullahi, J., Ali, S. I. A., Pham, Q. B., Abdulkadir, R. A., Costache, R., Nam, V. T., & Anh, D. T. (2020). Hybrid machine learning ensemble techniques for modeling dissolved oxygen concentration. *IEEE Access*, 8, 157218–157237.

<https://doi.org/10.1109/ACCESS.2020.3017743>

Abdullah, S. S., & Malek, M. A. (2016). Empirical Penman-Monteith equation and artificial intelligence techniques in predicting reference evapotranspiration: A review. *International Journal of Water*, 10(1), 1–23.

<https://doi.org/10.1504/IJW.2016.073741>

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(1), 2542.

<https://doi.org/10.1038/s41598-024-83859-6>

Alexandris, S., & Kerkides, P. (2003). New empirical formula for hourly estimations of reference evapotranspiration. *Agricultural Water Management*, 60(3), 157–180.

[https://doi.org/10.1016/S0378-3774\(02\)00172-5](https://doi.org/10.1016/S0378-3774(02)00172-5)

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration: Guidelines for computing crop water requirements* (FAO Irrigation and Drainage Paper No. 56). Food and Agriculture Organization of the United Nations. <https://www.fao.org/4/x0490e/x0490e00.htm>

Bellido-Jiménez, J. A., Estévez, J., & García-Marín, A. P. (2021). New machine learning approaches to improve reference evapotranspiration estimates using intra-daily temperature-based variables in a semi-arid region of Spain. *Agricultural Water Management*, 245, 106558.

<https://doi.org/10.1016/j.agwat.2020.106558>

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(6), 351–372.

<https://doi.org/10.1007/s00704-025-05561-5>

- C3S. (2025). *ERA5 hourly data on single levels from 1940 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS).
<https://doi.org/10.24381/CDS.ADBB2D47>
- Climate-Data.org. (2024). *Climate data for Batna, Algeria*. <https://en.climate-data.org/africa/algeria/batna-1025/>
- Cobaner, M. (2011). Evapotranspiration estimation by two different neuro-fuzzy inference systems. *Journal of Hydrology*, 398(3–4), 292–302.
<https://doi.org/10.1016/j.jhydrol.2010.12.030>
- Doğan, E. (2009). Reference evapotranspiration estimation using adaptive neuro-fuzzy inference systems. *Irrigation and Drainage*, 58(5), 617–628.
<https://doi.org/10.1002/ird.445>
- Galioto, F., Chatzinikolaou, P., Raggi, M., & Viaggi, D. (2020). The value of information for the management of water resources in agriculture: Assessing the economic viability of new methods to schedule irrigation. *Agricultural Water Management*, 227, 105848.
<https://doi.org/10.1016/j.agwat.2019.105848>
- Gelete, G. (2023). Application of hybrid machine learning-based ensemble techniques for rainfall-runoff modeling. *Earth Science Informatics*, 16(3), 2475–2495.
<https://doi.org/10.1007/s12145-023-01041-4>
- Ghiglieri, G., Buttau, C., Arras, C., Funedda, A., Soler, A., Barbieri, M., Carrey, R., Domènech, C., Torrentó, C., Otero, N., & Carletti, A. (2021). Using a multi-disciplinary approach to characterize groundwater systems in arid and semi-arid environments: The case of Biskra and Batna regions (NE Algeria). *Science of the Total Environment*, 757, 143797.
<https://doi.org/10.1016/j.scitotenv.2020.143797>
- Gong, X., Zhang, H., Ren, C., Sun, D., & Yang, J. (2020). Optimization allocation of irrigation water resources based on crop water requirement under considering effective precipitation and uncertainty. *Agricultural Water Management*, 243, 106485.
<https://doi.org/10.1016/j.agwat.2020.106264>
- Gu, Z., Qi, Z., Burghate, R., Yuan, S., Jiao, X., & Xu, J. (2020). Irrigation scheduling approaches and applications: A review. *Journal of Irrigation and Drainage Engineering*, 146(6), 04020007.
[https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001464](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001464)
- Hamdaoui, H., Zarrouk, Y., Ankush, M. A. T., Al Kaddouri, H., Alzain, M. N., Noman, O., Nasr, F. A., & Kouddane, N. (2026). Accurate reference evapotranspiration estimation with limited data for sustainable irrigation in eastern Morocco: A machine learning approach. *Frontiers in Sustainable Food Systems*, 10, 1734366.
<https://doi.org/10.3389/fsufs.2026.1734366>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., & Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049.
<https://doi.org/10.1002/qj.3803>
- Kisi, O., & Zounemat-Kermani, M. (2014). Comparison of two different adaptive neuro-fuzzy inference systems in modelling daily reference evapotranspiration. *Water Resources Management*, 28(9), 2655–2675.
<https://doi.org/10.1007/s11269-014-0632-0>
- Ladlani, I., Houichi, L., Djemili, L., Heddam, S., & Belouz, K. (2012). Modeling daily reference evapotranspiration (ET_0) in the north of Algeria using generalized regression neural networks (GRNN) and radial basis function neural networks (RBFNN): A comparative study. *Meteorology and Atmospheric Physics*, 118(3–4), 163–178.
<https://doi.org/10.1007/s00703-012-0205-9>
- Landeras, G., Bekoe, E., Ampofo, J., Logah, F., Diop, M., Cisse, M., & Shiri, J. (2018). New alternatives for reference evapotranspiration estimation in West Africa using limited weather data and ancillary data supply strategies. *Theoretical and Applied Climatology*, 132(3–4), 701–716.
<https://doi.org/10.1007/s00704-017-2120-y>
- Meziani, A. (2026). Machine learning models (MLP, Random Forest, LightGBM) for daily ET_0 estimation with limited data in humid Mediterranean region (Jijel) Algeria. *Veredas do Direito*, 23(3), e234267.
<https://doi.org/10.18623/RVD.V23.N3.4267>
- Misra, A. K. (2014). Climate change and challenges of water and food security. *International Journal of Sustainable Built Environment*, 3(1), 153–165.
<https://doi.org/10.1016/j.ijsbe.2014.04.006>
- Mohammadi, B., & Mehdizadeh, S. (2020). Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. *Agricultural Water Management*, 237, 106145.
<https://doi.org/10.1016/j.agwat.2020.106145>
- Morante-Carballo, F., Montalván-Burbano, N., Carrión-Mero, P., Herrera-Franco, G., & Apolo-Masache, B. (2022). What do we know about water scarcity in semi-arid zones? A global analysis and research trends. *Water*, 14(17), 2685.
<https://doi.org/10.3390/w14172685>
- Rosegrant, M. W., Ringler, C., & Zhu, T. (2009). Water for agriculture: Maintaining food security under growing scarcity. *Annual Review of Environment and Resources*, 34(1), 205–222.
<https://doi.org/10.1146/annurev.enviro.030308.090351>
- Samen, S. S., Kisi, O., Al-Janabi, A. M. S., & Elbeltagi, A. (2023). Estimation of reference evapotranspiration in semi-arid

- region with limited climatic inputs using metaheuristic regression methods. *Water*, 15(19), 3449.
<https://doi.org/10.3390/w15193449>
- Schwabe, K., Albiac, J., Connor, J. D., Hassan, R. M., & Mema, M. (2015). *Drought in arid and semi-arid regions: Causes, consequences and perspectives*. Springer.
<https://doi.org/10.1007/978-94-007-6636-5>
- Sentelhas, P. C., Gillespie, T. J., & Santos, E. A. (2010). Evaluation of FAO Penman–Monteith and alternative methods for estimating reference evapotranspiration with missing data in Southern Ontario, Canada. *Agricultural Water Management*, 97(5), 635–644.
<https://doi.org/10.1016/j.agwat.2009.11.005>
- Song, C., & Yao, L. (2022). A hybrid model for water quality parameter prediction based on CEEMDAN-IALO-LSTM ensemble learning. *Environmental Earth Sciences*, 81(9), 380.
<https://doi.org/10.1007/s12665-022-10380-2>
- Tabari, H., Kisi, O., Ezani, A., & Talaee, P. H. (2012). SVM, ANFIS, regression and climate-based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *Journal of Hydrology*, 444–445, 78–89.
<https://doi.org/10.1016/j.jhydrol.2012.04.007>
- Torabi Haghighi, A., Sadegh, M., Behrooz-Koohenjani, S., Noori, R., & Klöve, B. (2020). Unsustainability syndrome—From meteorological to agricultural drought in arid and semi-arid regions. *Water*, 12(3), 838.
<https://doi.org/10.3390/w12030838>
- Weather Spark. (2024). *Average weather in Batna, Algeria*. <https://weatherspark.com/y/53035/Average-Weather-in-Batna-Algeria-Year-Round>
- Zereg, S., & Belouz, K. (2024). Modeling daily reference evapotranspiration using SVR machine learning algorithm with limited meteorological data in Dar-El-Beidha, Algeria. *Acta Geophysica*, 72(3), 2009–2025.
<https://doi.org/10.1007/s11600-023-01107-3>
- Zhao, L., Xia, J., Xu, C., Wang, Z., & Sobkowiak, L. (2013). Evapotranspiration estimation methods in hydrological models. *Journal of Geographical Sciences*, 23(2), 359–369.
<https://doi.org/10.1007/s11442-013-1015-9>
- Zhou, Y., Marshall, L., Li, D., & Sharma, A. (2024). Revisiting evapotranspiration inputs in eco-hydrological modeling for climate change assessment. *Journal of Hydrology*, 642, 131888.
<https://doi.org/10.1016/j.jhydrol.2024.131888>
- Zippenfenig, P. (2024). *Open-Meteo.com Weather API* (Version 1.4.0). Zenodo.
<https://doi.org/10.5281/zenodo.14582479>