RIISDS Revista Interdisciplinaria De Ingeniería Sustentable Y Desarrollo Social

Artículo:

Evaluación de modelos de inteligencia artificial para la estimación de la evapotranspiración de referencia en agricultura de precisión bajo clima semiárido

Evaluation of Artificial Intelligence Models for Reference Evapotranspiration Estimation in Precision Agriculture under Semi-Arid Climate Conditions

Gilberto Bojorquez-Delgado[1*], Jesús Bojorquez-Delgado[1], Manuel Alfredo Flores-Rosales[1]

Revista Interdisciplinaria de Ingeniería Sustentable y Desarrollo Social

Recibido: 02 de agosto de 2025
Aceptado: 22 de agosto de 2025
Publicado: 05 de septiembre de 2025

Publicación continua editada por el **Tecnológico Nacional de México / Instituto Tecnológico Superior de Tantoyuca**

Desv. Lindero Tametate, S/N
Col. La Morita
C.P. 92100
Tantoyuca, Veracruz, México.
Teléfono: 789 8931680, Ext.196.

Correo electrónico:
revistadigital@itsta.edu.mx

Sitio WEB
https://itsta.edu.mx/revistadigital

ISSN **2448-8003**
Reserva de derechos al uso exclusivo No. **04-2016-092313253300-203**

Editor responsable:
Dr. Horacio Bautista Santos

[1] Tecnológico Nacional de México – ITS de Guasave, Sinaloa, México.

[*] Autor corresponsal: itsg.gbojorquez@gmail.com

Resumen: La estimación precisa de la evapotranspiración de referencia (ET_0) es esencial para la programación del riego en sistemas agrícolas de alta eficiencia hídrica. Este estudio compara el desempeño de dos enfoques de modelado, específicamente el algoritmo XGBoost y una red neuronal LSTM, para estimar ET_0 diaria en un entorno semiárido del noroeste de México, utilizando datos meteorológicos derivados del sistema NASA POWER entre enero de 2020 y julio de 2024. Ambos modelos fueron entrenados sobre un conjunto homogéneo de 1,827 días y evaluados mediante métricas estándar, incluyendo MAE, RMSE, R^2, NSE, MAPE y el coeficiente d de Willmott, con un análisis diferenciado entre las temporadas seca y húmeda. El modelo XGBoost alcanzó un MAE de 0.145 mm día^{-1} y un R^2 de 0.983, superando en todos los indicadores al modelo LSTM, el cual mostró mayor error absoluto y variabilidad estacional. El análisis SHAP identificó como variables más influyentes a la radiación solar de onda corta, el déficit de presión de vapor y la velocidad del viento, en coherencia con la teoría agrometeorológica. Además, XGBoost mostró mayor estabilidad frente a eventos extremos y menor sesgo sistemático. Estos resultados posicionan a XGBoost como una herramienta precisa, interpretable y computacionalmente eficiente para su integración en sistemas automatizados de toma de decisiones y plataformas IoT agrícolas. No obstante, se reconocen limitaciones geográficas y metodológicas que sugieren la necesidad de validaciones multirregionales y de explorar arquitecturas híbridas en futuros trabajos. Estos hallazgos respaldan la adopción operativa de enfoques basados en XGBoost en la gestión del riego de precisión y resaltan su potencial para mejorar la eficiencia en el uso del agua en sistemas agrícolas de zonas semiáridas.

Palabras clave: Evapotranspiración, Riego, XGBoost, LSTM, SHAP.

Abstract

Accurate estimation of reference evapotranspiration (ET_0) is essential for irrigation scheduling in high-efficiency agricultural systems. This study compares the performance of two modeling approaches—specifically, the XGBoost algorithm and a long short-term memory (LSTM) neural network—for estimating daily ET_0 in a semi-arid region in northwestern Mexico, using meteorological data derived from the NASA POWER system between January 2020 and July 2024. Both models were trained on a homogeneous dataset of 1,827 days and evaluated using standard metrics, including MAE, RMSE, R^2, NSE, MAPE, and Willmott's d coefficient, with separate analyses for dry and wet seasons. The XGBoost model achieved a MAE of 0.145 mm day^{-1} and an R^2 of 0.983, outperforming the LSTM model across all metrics, which exhibited higher absolute error and greater seasonal variability. SHAP analysis identified shortwave solar radiation, vapor pressure deficit, and wind speed as the most influential variables, aligning with agrometeorological theory. Additionally, XGBoost demonstrated greater robustness to extreme events and lower systematic bias. These results position XGBoost as an accurate, interpretable, and computationally efficient tool for integration into automated decision-making systems and agricultural IoT platforms. Nonetheless, geographic and methodological limitations are acknowledged, suggesting the need for multi-regional validations and exploration of hybrid architectures in future research. These findings support the operational adoption of XGBoost-based approaches in precision irrigation management and highlight their potential to enhance water use efficiency in semi-arid agricultural systems.

Keywords: Evapotranspiration, Irrigation, XGBoost, LSTM, SHAP.

Introduction

Precise quantification of reference evapotranspiration (ET_0) forms the fundamental basis of any modern irrigation scheduling algorithm, as it links atmospheric demand with crop water requirements and, ultimately, with agricultural water productivity. Although the Penman–Monteith equation recommended by FAO-56 remains the standard physical reference, its routine application in commercial fields is hindered by the requirement for reliable local measurements of solar radiation, temperature, relative humidity, and wind speed (Liang et al., 2023). Over the past decade, this instrumental gap has spurred a surge in

studies based on machine learning (ML) and deep learning (DL), which claim to match or even surpass the accuracy of energy balance-based methods using reduced datasets (Ayaz et al., 2021).

Initially, multilayer perceptrons showed promise in the arid climates of Iran, reducing mean absolute error (MAE) by approximately 40% compared to the Hargreaves–Samani method (Wu et al., 2020). More recently, gradient boosting methods have gained traction: an XGBoost model optimised using the Nelder–Mead algorithm achieved Nash–Sutcliffe efficiencies (NSE) exceeding 0.95 at 28 stations across the Yangtze River Basin (Li et al., 2024), while a study in a semi-arid climate reported a root mean square error (RMSE) as low as 0.19 mm day^{-1} (Sharafi & Mohammadi Ghaleni, 2024). The integration of satellite products has further enabled XGBoost to upscale instantaneous Landsat-based fluxes to daily ET estimates with errors below 13 W m^{-2} (Yang et al., 2025).

In parallel, advances in deep learning have led to the development of hybrid CNN–LSTM architectures with attention mechanisms, which outperformed the Penman–Monteith method by 18% across 12 subtropical sites (Sarkar et al., 2025). Bidirectional LSTM models have also achieved MAE values below 0.30 mm day^{-1} by incorporating spatial context (Dong et al., 2024). Furthermore, hybrid frameworks have further reduced uncertainty by combining signal decomposition with boosted trees (Liu et al., 2024), or by integrating Penman–Monteith inputs with convolutional encoders (Dong et al., 2024).

Nonetheless, recent reviews reveal persistent knowledge gaps: much of the empirical evidence stems from Asia and the Middle East, leaving semi-arid districts of Latin America virtually unexplored (Liyew et al., 2025). Moreover, direct comparisons between ML/DL models and the Penman–Monteith method using identical datasets remain scarce (Helali et al., 2024). Despite progress in explainable AI, the physical interpretation of "black-box" models remains limited, although recent SHAP analyses have highlighted the dominant role of shortwave radiation and vapour pressure deficit (VPD) (Mushtaq et al., 2024). Seasonal robustness also poses challenges: random forest models have shown performance drops during monsoon peaks in India (Xu et al., 2023), whereas CatBoost and LightGBM appear more resilient during prolonged droughts in Morocco (Yan et al., 2023). Emerging strategies such as quantile regression for ET_0 (Bour et al., 2025), inter-basin transfer learning (Chen et al., 2022), and federated learning on edge devices (Tausif et al., 2025), illustrate the

breadth of innovation, though their applicability under Mexican conditions still requires validation.

In response to these gaps, the present study compiles and analyses a continuous meteorological record spanning 1,827 days (January 2020 to July 2024) obtained from an 8.01-hectare drip-irrigated agricultural field in El Tajito, Guasave—an area representative of the water-stressed agricultural landscape typical of northwestern Mexico's semi-arid zone. The choice of the XGBoost model is grounded in its demonstrated effectiveness in recent studies, attributed to its robust capacity to model complex non-linear relationships, its efficient handling of correlated predictors, and its ability to deliver physically interpretable outputs via SHAP values (Liu et al., 2024; Yang et al., 2025). Conversely, the selection of the LSTM model is justified by its well-established aptitude for capturing complex temporal dependencies in meteorological series and its capacity to retain relevant historical information through sliding time windows—both critical for modelling dynamic hydrological phenomena such as evapotranspiration (Dong et al., 2024; Sarkar et al., 2025).

Accordingly, the specific objectives of this study are to: (i) calibrate and optimise a gradient boosting model (XGBoost) and a recurrent neural network (LSTM) using a seven-day rolling window to leverage short-term historical information; (ii) evaluate and compare the daily prediction accuracy of both models against the standard Penman–Monteith method using widely accepted hydrological metrics (MAE, RMSE, R^2, MAPE, NSE, and Willmott's d) (Bojorquez-Delgado et al., 2023; Graciano-Obeso et al., 2024; Sarkar et al., 2025); and (iii) analyse and interpret the relative importance and interactions of meteorological variables through SHAP values, explicitly linking these empirical findings to underlying agrometeorological theory (Mushtaq et al., 2024). The central hypothesis is that an XGBoost-based approach not only offers superior predictive accuracy compared to LSTM and traditional methods, but also delivers physical interpretability, facilitating operational deployment. In summary, the main contribution of this study lies in providing one of the first systematic evaluations of artificial intelligence models for reference evapotranspiration in a semi-arid region of Mexico, a context largely absent from international literature. By directly comparing XGBoost and LSTM with the Penman–Monteith standard under identical datasets, this work not only fills a geographic and methodological gap, but also demonstrates

the operational feasibility of integrating explainable AI into precision irrigation practices across Latin America's water-stressed agricultural landscapes.

Materials and Methods

This section describes the methodology employed for the acquisition, preparation, and analysis of data, as well as the modelling and evaluation techniques used in the study.

Dataset Construction

The dataset was constructed using publicly available meteorological records retrieved through the 'AG' interface of the NASA POWER v8 service (accessed on 10 February 2025), chosen for its validated global coverage and daily resolution suitable for agricultural studies in regions with limited local instrumentation. In the absence of in-field lysimetric measurements, the reference value was taken as the reference evapotranspiration calculated by the service using the Penman–Monteith equation (ET_o_PM), which has been previously validated for similar latitudes.

Continuous daily observations were extracted between 1 January 2020 and 31 July 2024 for a plot located in the El Tajito ejido, Guasave, Sinaloa (25°39′18″ N, 108°38′14″ W). The requested meteorological variables included maximum, minimum, and mean temperatures; relative humidity; corrected precipitation; wind speed; shortwave and longwave radiation; atmospheric pressure; and dew point temperature. The data, originally in JSON format, were converted into a pandas DataFrame, checked for temporal continuity, and cleaned using physically reasonable ranges based on agrometeorological literature for semi-arid regions (e.g., Tmean between –5 and 45 °C; RH between 10% and 100%). This resulted in a homogeneous and self-consistent dataset of 1,827 days for subsequent analysis. The processed dataset and associated scripts are openly available in Zenodo (Bojorquez Delgado, 2025); DOI: 10.5281/zenodo.16922140).

Computational Environment

Subsequent analyses were conducted on the Google Colaboratory platform using Python 3.11 in a CPU-based environment (Intel Xeon, 2 vCPUs, 13 GB RAM), with GPU acceleration deliberately disabled to ensure result reproducibility under standard computational settings without specialised hardware. The libraries used included: pandas

(1.3), NumPy (1.2), matplotlib (3.4), seaborn (0.11), scikit-learn (0.24), XGBoost (1.5), TensorFlow-Keras (2.6), and SHAP (0.40).

Data Preprocessing

The dataset was examined for missing values, which were imputed using temporal linear interpolation. This method has been shown to be effective in environmental and agroclimatic time series when data gaps are short and the underlying signal is continuous (Lepot et al., 2017; Noor et al., 2015). Predictor variables were standardised using StandardScaler(), trained exclusively on the training set to prevent information leakage into the test set. The target variable (ET_o_PM) was not transformed, preserving its original scale in mm day^{-1} to maintain physical interpretability. To respect the sequential structure of the series, the dataset was split chronologically (70% initial segment for training and 30% final segment for testing), thereby preventing the use of future observations for past predictions.

In addition, the selection of a seven-day sliding window for the LSTM model was supported by the partial autocorrelation function of the ET_o_PM series and is consistent with previous studies that applied lagged input sequences to predict short-term evapotranspiration dynamics(Yin et al., 2020).

The predictive performance of the models was evaluated using a set of widely recognised statistical indices in hydrology and agroclimatology, including the Nash–Sutcliffe Efficiency (Nash & Sutcliffe, 1970) and Willmott's d coefficient (Willmott, 1981), which provide complementary assessments of model accuracy and agreement with observed values.

Exploratory Analysis

Initial exploratory analyses were performed using time series plots, histograms, and correlation matrices to examine distributions, trends, and inter-variable correlations.

Predictive Modelling

Two advanced predictive techniques were implemented to model reference evapotranspiration (ET_o_PM). The first was an XGBoost model, developed using temporal cross-validation through TimeSeriesSplit with five folds. Hyperparameters were systematically optimised with GridSearchCV, testing different numbers of trees (50 and 100 estimators), maximum depths (3 and 5), and learning rates (0.01 and 0.1). This configuration ensured that the model was both parsimonious and well adapted to the seasonal variability of the dataset.

In parallel, a long short-term memory (LSTM) recurrent neural network was constructed to capture temporal dependencies inherent in agroclimatic series. The model employed a sliding window of seven consecutive days as input, a setting supported by prior studies that demonstrated the utility of 5–7-day windows for improving evapotranspiration estimation and forecasting performance (Jia et al., 2023; Matos et al., 2025; Soylu et al., 2012). The architecture consisted of a single LSTM layer with 32 units and a tanh activation function, followed by a Dropout(0.2) layer to reduce overfitting, and a linear Dense(1) output layer to generate daily predictions. Training was conducted with the Adam optimiser for a maximum of 50 epochs and a batch size of 16, reserving 20% of the training data for internal validation. An early stopping criterion with a patience of five epochs was applied to prevent unnecessary iterations, while fixed random seeds (tf.random.set_seed(42), np.random.seed(42)) guaranteed reproducibility across runs.

This dual-model strategy allowed the evaluation of a tree-based boosting approach against a deep learning architecture specifically designed for sequential data, thereby providing a rigorous comparative framework for assessing the potential of explainable machine learning versus recurrent neural networks in the estimation of reference evapotranspiration under semi-arid conditions.

Performance Evaluation

The predictive performance of both models was assessed using standard metrics commonly applied in agroclimatic studies: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), Mean Absolute Percentage Error (MAPE), mean bias, Nash–Sutcliffe Efficiency (NSE), and Willmott's d coefficient. The use of NSE (Nash & Sutcliffe, 1970) and Willmott's d (Willmott, 1981) follows their established application in hydrological and agrometeorological model evaluation. Additionally, separate analyses were conducted for the dry season (November–April) and the wet season (May–October).

Study Limitations

A key limitation of this study is that model training and validation were conducted at a single site. Although the results indicate internal robustness, the geographic reproducibility of the model has yet to be assessed in other semi-arid and irrigated environments. This

methodological constraint should be addressed in future research through external validation studies that examine the model's generalisation across diverse agroclimatic contexts**.**

Results and Discussion

This section presents a critical analysis of the results obtained by comparing the predictive performance of a gradient boosting model (XGBoost) and a recurrent neural network (LSTM) in estimating reference evapotranspiration (ET_0) under semi-arid climatic conditions. Using a multi-year dataset that integrates key thermodynamic and radiative variables, the performance of each approach is quantified through hydrological metrics of absolute and squared error, efficiency, and bias, and their consistency across dry and wet seasons is examined. In addition, the relative importance of predictors and the physical coherence of the modelled responses are evaluated in order to determine which architecture provides more reliable and operationally useful estimates for high-precision irrigation scheduling.

As shown in Figure 1, the daily time series of reference evapotranspiration (ET_0_PM) and mean temperature (Tmean) from January 2020 to July 2024 exhibits a clear seasonal pattern: between June and August, ET_0_PM reaches its maximum values (6–10 mm day^{-1}) in conjunction with elevated daily mean temperatures (28–34 °C), whereas during the December–February period, both indicators drop to their minimum levels (2–4 mm day^{-1} and 15–20 °C, respectively). Furthermore, small interannual anomalies are reflected in ET_0_PM variations of up to 0.5 mm day^{-1}, suggesting the influence of fluctuations in radiation and atmospheric humidity.

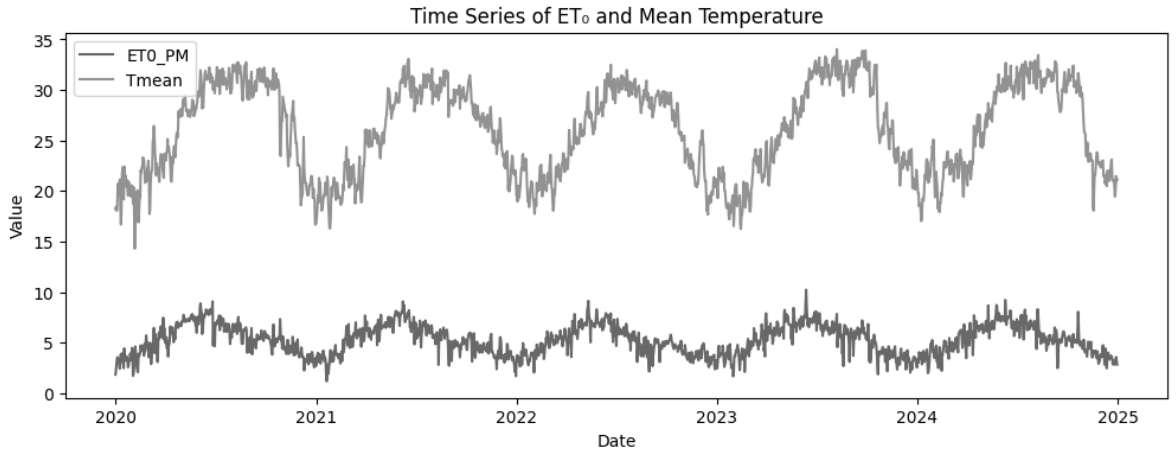

Figure 1. Time series of reference evapotranspiration and mean temperature.

This synchrony between Tmean and ET_o_PM validates the selection of mean temperature as a primary predictive variable in the AI models, a relationship widely recognized in agrometeorological research where temperature is a fundamental driver of evapotranspiration (Irmak et al., 2003).

Figure 2 presents the distribution of daily mean temperature over the study period, shown as a histogram with a kernel density estimate (KDE). A bimodal tendency can be observed, with local peaks around 22 °C and 30 °C, reflecting the seasonal cycle of the semi-arid climate. Approximately 10% of the days fall below 18 °C and 90% below 32 °C, while the distribution tails extend towards 15 °C and 35 °C, indicating winter cooling and occasional summer heat episodes. Beyond describing the climatic range, the relatively smooth distribution and absence of extreme outliers are relevant for model stability, as machine learning algorithms are less prone to overfitting when predictor variables follow physically consistent and bounded patterns (Irmak et al., 2003). This implies that mean temperature provides a robust and generalisable signal for ET_o estimation across seasons, supporting its role as a key predictor in the AI models.

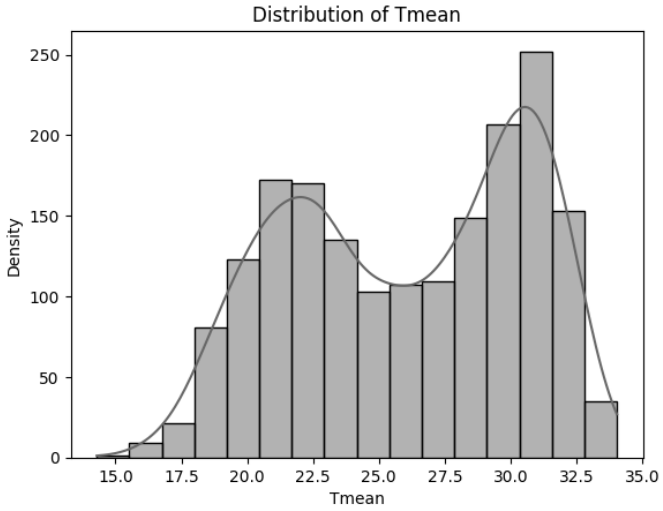

Figure 2. Distribution of Tmean.

As shown in Figure 3, the daily relative humidity at 2 m (RH2M) exhibits an approximately normal distribution centred around 60%, with values ranging from 37% to 88%. Beyond this descriptive range, RH is recognised as a secondary yet influential predictor of reference evapotranspiration, as it affects the vapor pressure deficit and the evaporative demand of the atmosphere. Agrometeorological guidelines (Allen et al., 1998) include RH

as an indispensable component in ET_0 estimation. Moreover, Córdova et al. (2015) quantified that estimating RH instead of using measured data can introduce up to 14% errors in ET_0 under highland conditions. This underscores that, although less dominant than radiation or temperature, RH contributes significantly to model fidelity by capturing moisture-driven variability in semi-arid climates.

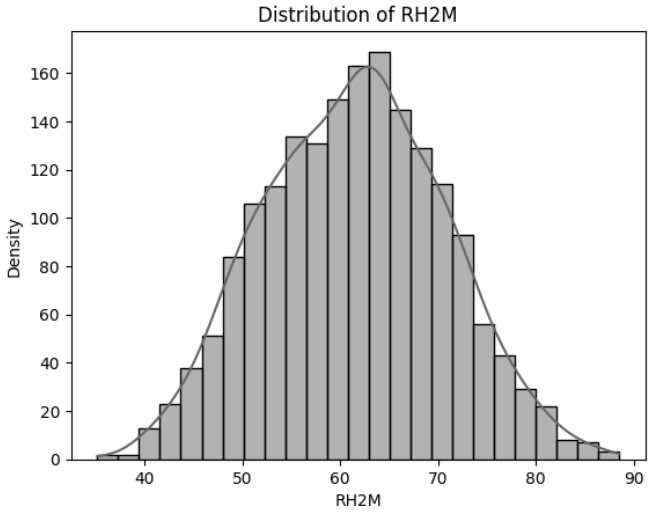

Figure 3. Distribution of RH2M.

The 10th percentile of the sample lies near 50%, while the 90th percentile is around 75%, indicating that most days maintain moderate to high humidity levels. The symmetry of the kernel density estimate (KDE) curve and the shape of the histogram reflect the stability of the atmospheric moisture regime in the region, supporting the inclusion of RH2M as a key predictor variable in the AI models.

Figure 4 shows the distribution of the daily vapour pressure deficit (VPD) estimated for the 2020–2024 period. The histogram with a kernel density estimate (KDE) curve reveals a modal peak near 1.0 kPa, indicating that most days exhibit a moderate VPD around that value.

The 10th percentile is approximately 0.5 kPa, while the 90th percentile is close to 2.0 kPa. Only 3.8% of the days (≈70 out of 1,827 records) presented VPD values above 2.0 kPa, corresponding to atypical heatwaves or very dry episodes. This long right tail, although infrequent, is critical in explaining the higher predictive capacity of the XGBoost model, as it successfully captures nonlinear responses under extreme vapour deficit conditions that tend

to be smoothed out by LSTM. This asymmetry confirms that, although the prevailing climate regime is predominantly of moderate deficit, there are seasonal events with significantly higher values—highlighting the relevance of VPD as a predictor variable in evapotranspiration models.

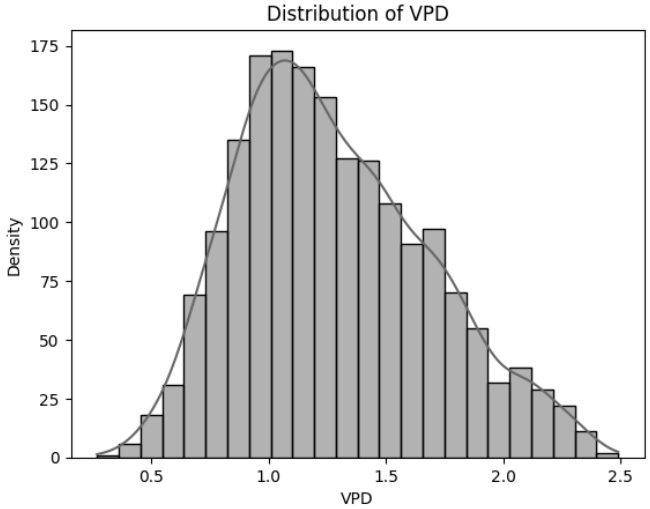

Figure 4. Distribution of VPD.

Figure 5 shows the distribution of daily growing degree days (GDD) over the analysis period. A clear bimodal pattern is observed: the first peak, between 11 and 13 °C·d, coincides with the winter-to-spring transition, while the second, between 20 and 22 °C·d, reflects thermal accumulation during summer and early autumn. The interquartile range spans from 10 to 22 °C·d, with extremes between 5 and 24 °C·d.

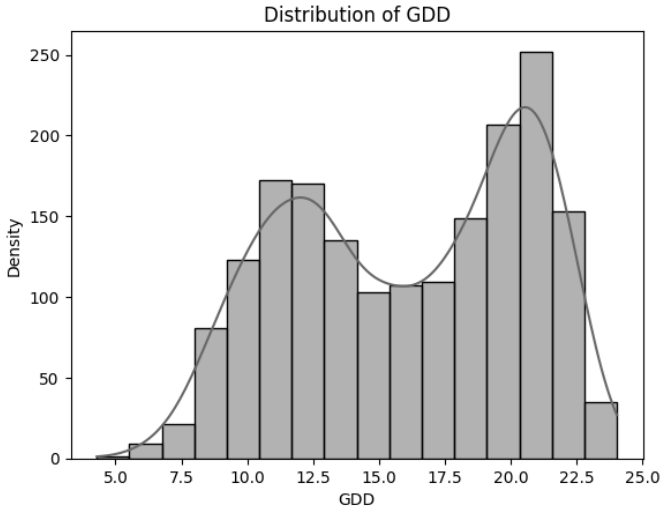

Figure 5. Distribution of GDD.

This dual thermal modulation suggests the existence of two distinct phenological regimes, which could potentially be leveraged by time-sensitive models. However, the LSTM model failed to translate this structure into improved predictive performance, largely because GDD is mathematically derived from Tmean, resulting in near-perfect collinearity between both variables. This redundancy limits the independent contribution of GDD once Tmean is already included in the predictors. Nevertheless, including GDD may still add marginal predictive value by explicitly capturing cumulative thermal thresholds relevant to crop development, even if its statistical contribution to ET_0 prediction is secondary compared to Tmean (Allen et al., 1998; Zhou et al., 2021).

As shown in Figure 6, the distribution of daily reference evapotranspiration (ET_0_PM) is slightly right-skewed, with a modal peak between 4 and 6 mm day^{-1}. The 10th percentile lies near 2.5 mm day^{-1}, and the 90th percentile around 8 mm day^{-1}, indicating that most days are characterised by moderate evaporative demand, while isolated events of high evapotranspiration (up to 10 mm day^{-1}) occur under very hot and dry conditions. Such extreme ET_0 events correspond to scenarios of critical water demand, where failure to adjust irrigation scheduling may result in crop water stress, yield reduction, and inefficient resource use. Their correct identification is therefore essential for precision irrigation systems aiming to optimise water allocation during peak atmospheric demand.

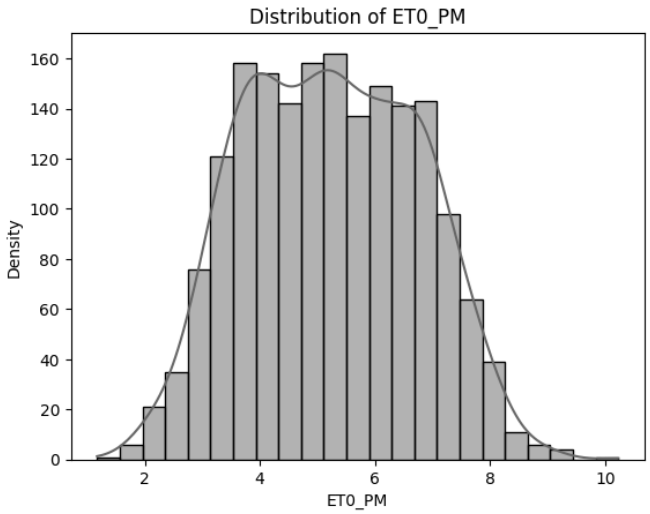

Figure 6. Distribution of ET0_PM.

Figure 7 presents the correlation matrix among the five key climatic variables included in the study: mean temperature (Tmean), relative humidity at 2 m (RH2M), vapour

pressure deficit (VPD), growing degree days (GDD), and reference evapotranspiration (ET_0_PM). The matrix shows that Tmean and GDD are perfectly correlated (r = 1.00), as both are directly derived from the same thermal variable. VPD emerges as the strongest predictor of ET_0_PM (r = 0.83), reflecting the close dependence of evaporative demand on the air's capacity to hold water vapour. ET_0_PM also maintains moderate positive correlations with Tmean and GDD (r = 0.63 in both cases), underscoring the role of thermal energy in driving evaporation. Conversely, RH2M is inversely correlated with both VPD (r = –0.54) and ET_0_PM (r = –0.37), indicating that higher atmospheric humidity levels tend to reduce water loss through evaporation. While this correlation structure validates the selection of these variables for the artificial intelligence models, it also reveals redundancy among predictors, particularly the perfect collinearity between Tmean and GDD. Such redundancy could increase the risk of overfitting if not properly controlled; however, both variables were retained to maintain methodological coherence with agroclimatic literature, where Tmean is recognised as a primary driver of ET_0 and GDD contributes complementary information on cumulative thermal thresholds relevant for crop development (Allen et al., 1998; Irmak et al., 2013). This ensures that the models capture both instantaneous thermal dynamics and accumulated heat effects, which are agronomically meaningful.

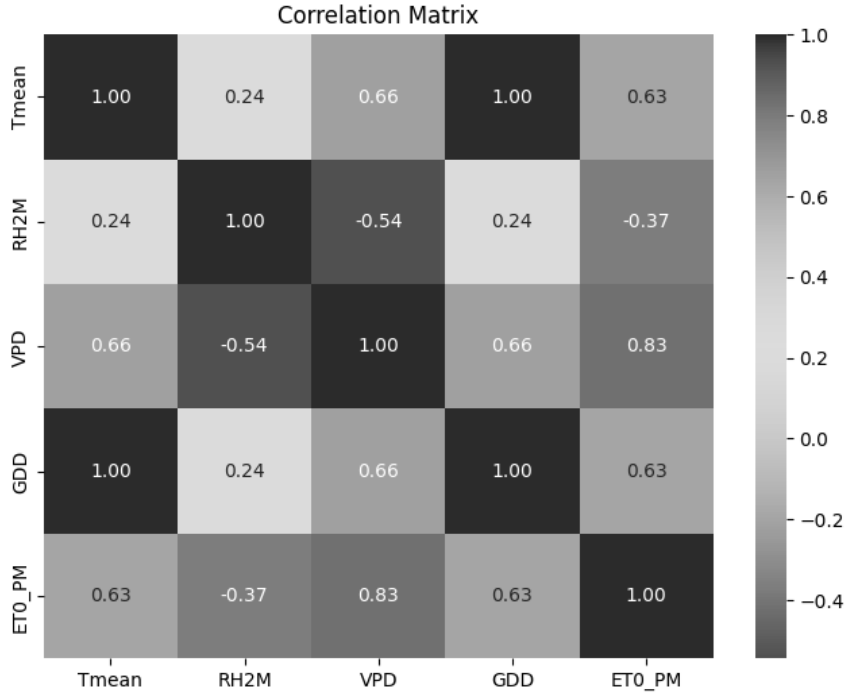

Figure 7. Correlation matrix among key variables.

Figure 8 summarises the relative importance of each predictor in the XGBoost model using SHAP values, which quantify the individual impact of variables on the prediction of reference evapotranspiration (ET_o_PM).

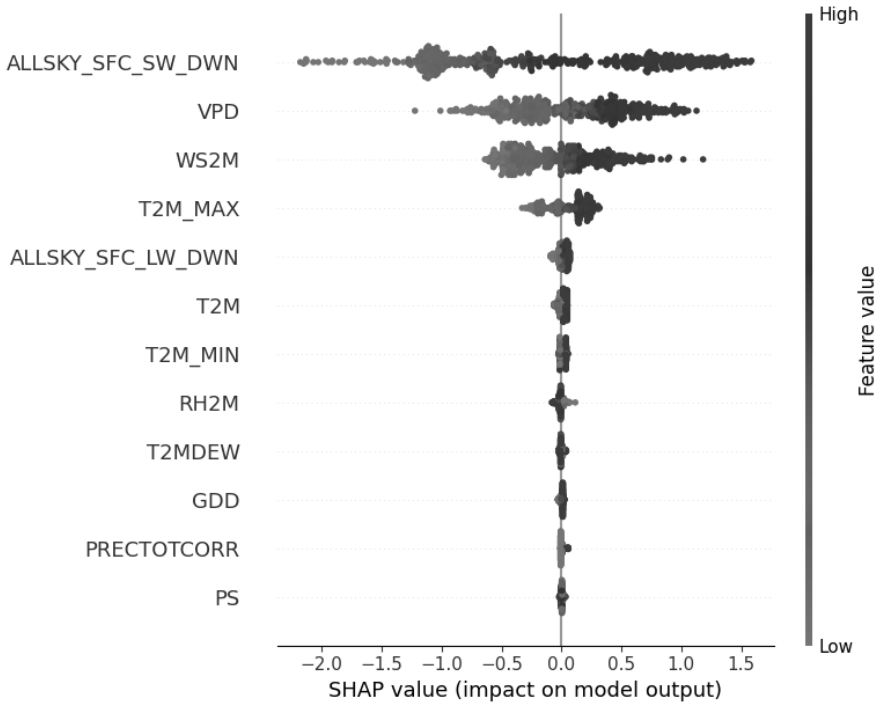

Figure 8. SHAP summary plot for XGBoost feature importance.

The mean SHAP values indicate that downward shortwave solar radiation (ALLSKY_SFC_SW_DWN) contributed the highest average impact (0.47), followed by vapour pressure deficit (VPD, 0.32) and wind speed at 2 metres (WS2M, 0.21). In contrast, growing degree days (GDD, 0.05), corrected precipitation (PRECTOTCORR, 0.03), and atmospheric pressure (PS, 0.02) exhibited marginal contributions (Table 1).

Table 1. Mean SHAP values of predictors for ET_o.

Predictor	Mean SHAP value	Rank
ALLSKY_SFC_SW_DWN	0.47	1
VPD	0.32	2
WS2M	0.21	3
GDD	0.05	4
PRECTOTCORR	0.03	5
PS	0.02	6

Variables are presented in descending order of average influence, and the dispersion of points illustrates the variability of impact across the observed value ranges.

Downward shortwave solar radiation is identified as the most influential factor in ET_0 estimation, consistent with its dominant role in the surface energy balance. Vapour pressure deficit and wind speed also show substantial contributions, reflecting their direct influence on atmospheric evaporative demand. In contrast, GDD, corrected precipitation, and atmospheric pressure exhibit marginal influence on model behaviour. These results align with classical agrometeorological theory and confirm that XGBoost not only captures the nonlinear relationships between predictors and ET_0, but also structures its decisions in a manner coherent with the physical foundations of the process. To enhance transparency and reproducibility, Table 1 reports the exact mean SHAP values for each predictor.

Figure 9 shows the learning curve of the LSTM model, illustrating the evolution of the mean squared error during training and validation across epochs. It can be observed that, after a sharp drop in the first epoch (from ≈20 to ≈2 MSE in training and from ≈9 to ≈1 MSE in validation), both curves stabilise at low values (~0.5 MSE) from epoch 5 onward, with no signs of significant overfitting (training and validation losses remain very close). This indicates that the chosen LSTM architecture is capable of converging rapidly and generalising reasonably well, albeit with a higher error threshold compared to XGBoost.

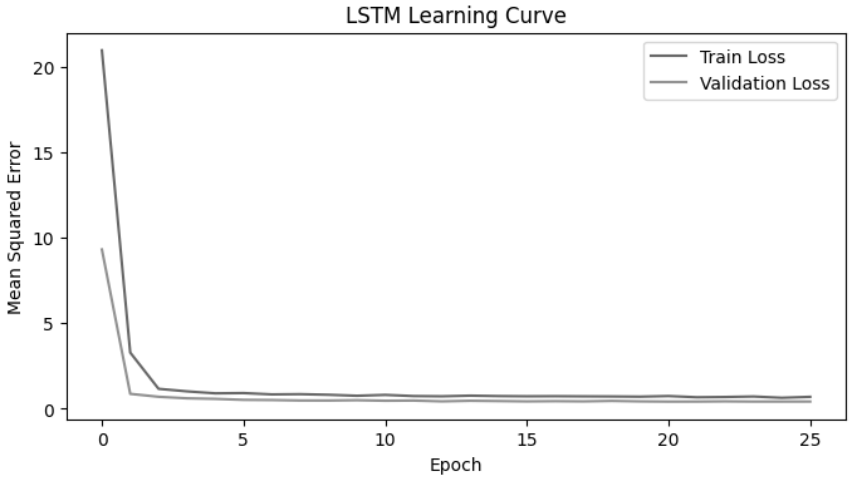

Figure 9. Learning curve for LSTM model.

Nevertheless, it should be noted that only a simple single-layer architecture with basic hyperparameter tuning was implemented, without exploring more complex variants such as stacked or bidirectional LSTMs. This methodological choice aimed to prioritise reproducibility and computational efficiency, in line with previous agricultural applications of LSTM where relatively simple configurations have proven sufficient—for instance, bidirectional LSTMs applied to soil moisture and electrical conductivity prediction in citrus orchards (Gao et al., 2021) and hybrid PSO-LSTM models optimised for ET_0 forecasting in irrigation scheduling (Jia et al., 2023). Figure 9 shows the learning curve of the LSTM model, illustrating the evolution of the mean squar

Figure 10 displays the "Predicted vs. Actual" scatter plot for ET_0 predictions obtained using XGBoost and LSTM against the reference ET_0_PM values (x-axis). The dashed 1:1 line represents perfect prediction.

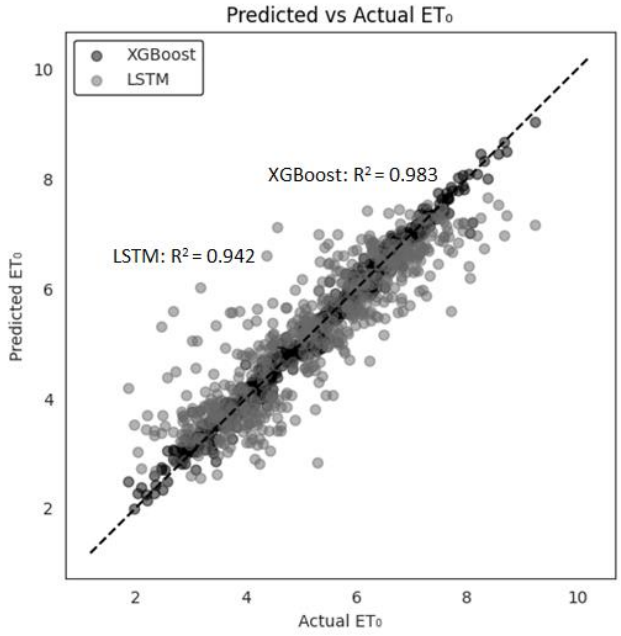

Figure 10. Predicted vs Actual for XGBoost and LSTM.

It can be seen that the XGBoost points cluster very closely around the line, indicating reduced errors and high precision, whereas the LSTM points exhibit greater dispersion around the diagonal, revealing slightly higher bias and lower overall accuracy. The correlation coefficients further support this visual impression: XGBoost achieved $R^2 = 0.983$,

while LSTM reached $R^2 = 0.942$. These values have also been included directly in the graph to provide a clearer quantitative comparison. This plot visually summarises the superiority of XGBoost in its ability to faithfully replicate daily reference evapotranspiration.

Figure 11 shows the time series of actual reference evapotranspiration (blue line) versus the predictions generated by XGBoost (orange line) and LSTM (green line) during the validation period (July 2023–January 2025). It is evident that XGBoost accurately reproduces both the peaks and sharp drops in ET_0, aligning almost point-by-point with the actual curve, whereas LSTM tends to smooth out fluctuations and exhibits a slight lag during extreme events (e.g., heatwaves or sudden drops). This behaviour reinforces the superiority of XGBoost in capturing the full temporal dynamics of evaporative demand under real-world precision irrigation conditions.

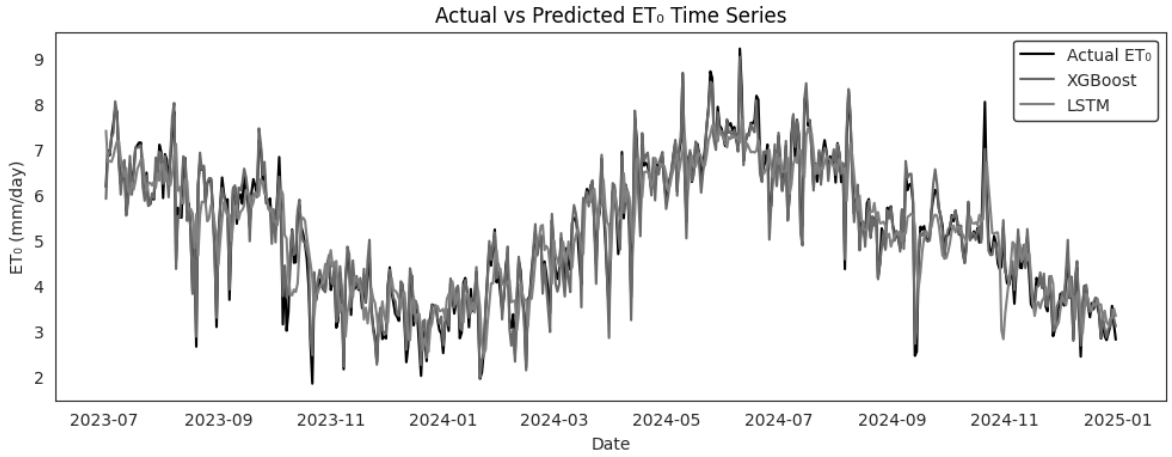

Figure 11. Time series of actual vs XGBoost and LSTM predictions.

Figure 12 shows the distribution of residuals (prediction – actual value) for XGBoost and LSTM, with their respective KDE curves superimposed.

The XGBoost residuals are more tightly clustered around zero, forming a narrow bell-shaped curve, while the LSTM residuals display a flatter distribution with longer tails in both directions.

This qualitative contrast is quantitatively confirmed by the normality diagnostics reported in Table 2: both models deviate significantly from normality (Shapiro–Wilk and Jarque–Bera tests, $p < 0.001$), yet XGBoost presents lower skewness (–0.43 vs. 0.13) and higher excess kurtosis (2.87 vs. 1.64) compared to LSTM. These results indicate that XGBoost not only

reduces mean prediction bias but also concentrates errors more symmetrically around zero, leading to lower variance and more stable performance.

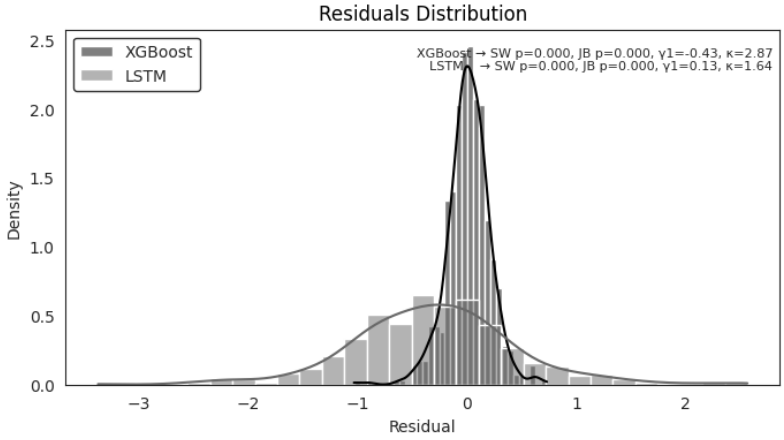

Figure 12. Residuals distribution for both models.

Table 2. Residual normality diagnostics for XGBoost and LSTM.

Model	W (Shapiro–Wilk)	p (SW)	JB (Jarque–Bera)	p (JB)	Skewness (γ_1)	Excess Kurtosis
XGBoost	0.96	0.00	200.08	0.00	-0.43	2.87
LSTM	0.98	0.00	61.12	0.00	0.13	1.64

Figure 13 presents the monthly distribution of XGBoost residuals (prediction – actual value) throughout the year. In each month, the median error lies very close to zero, indicating the absence of systematic bias, and the interquartile ranges (IQR) are relatively narrow, reflecting high consistency in predictions.

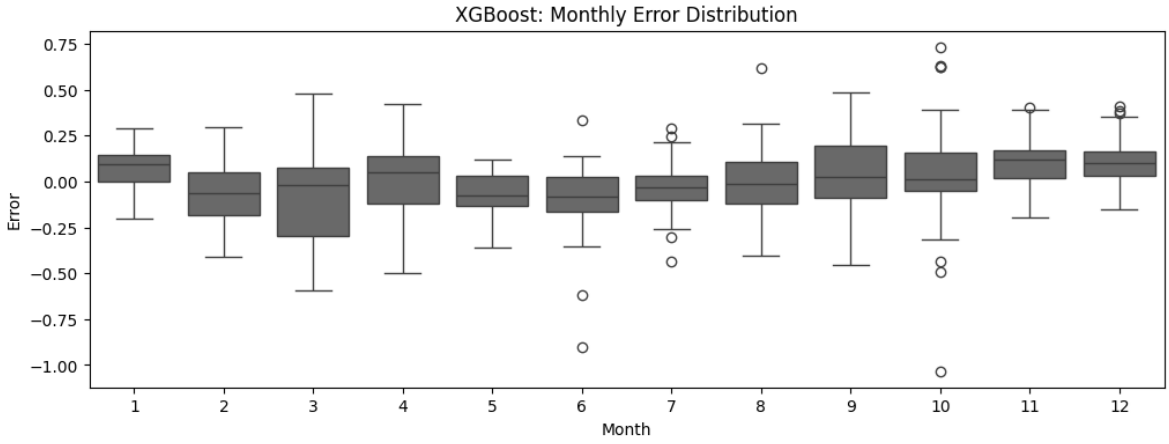

Figure 13. Monthly error boxplot for XGBoost.

However, in March (month 3) and October (month 10), the IQR is slightly wider and more extreme outliers are present, which coincides with transitional periods between dry and wet seasons. These months are typically characterised by abrupt changes in temperature, relative humidity, and solar radiation, which increase atmospheric evaporative demand variability. From an agronomic perspective, this heightened uncertainty may lead to either under- or overestimation of irrigation needs at the onset of spring sowings and during late-season crop maturation. In contrast, during the peak summer months (June to August), residual spread decreases, indicating more precise fitting under stable climatic conditions, which favours accurate water allocation during periods of maximum crop demand. This analysis confirms the seasonal robustness of XGBoost, while highlighting that transitional months may still require cautious interpretation when translating predictions into irrigation schedules.

Figure 14 shows the monthly distribution of residuals from the LSTM model. Unlike XGBoost, the median error of LSTM exhibits slight shifts in several months (e.g., positive in February–March and negative in September–October), indicating a more pronounced seasonal bias. Additionally, the interquartile range is much wider across all months, and outliers reach magnitudes of up to ±3 mm day^{-1}.

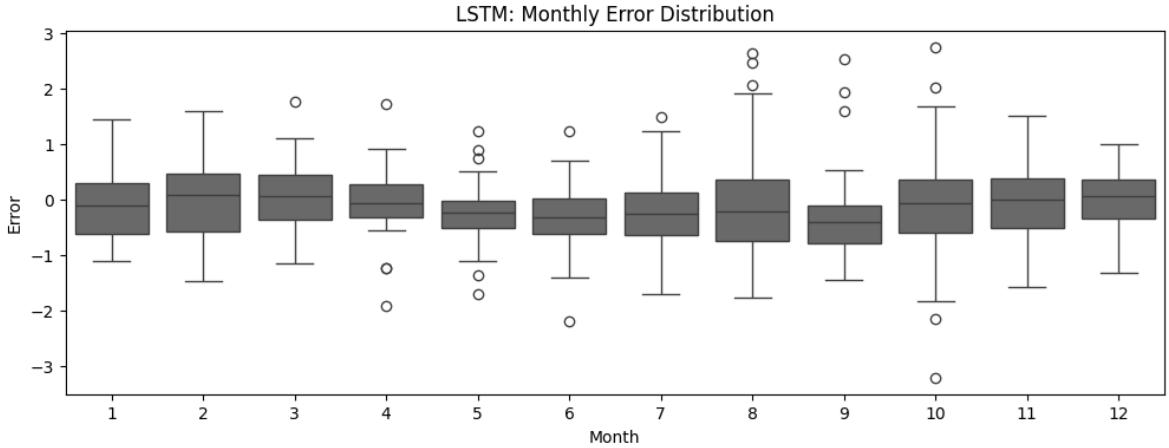

Figure 14. Monthly error boxplot for LSTM.

This behaviour reflects the network's reduced stability under conditions of climatic transition, particularly during April–May and August–October, when rapid fluctuations in humidity and solar radiation strongly alter evapotranspiration dynamics. In practical terms, such bias and variability translate into significant risks for irrigation planning: overestimation

in early spring could lead to unnecessary water use during critical establishment stages, while underestimation in autumn could limit water availability during grain filling or fruit maturation. Overall, Figure 14 confirms that although LSTM can approximate the general trend of ET_0, its seasonal accuracy and reliability are considerably lower than those of XGBoost, especially in months where precise irrigation management is crucial.

Figure 15 shows the temporal evolution of residuals generated by the XGBoost and LSTM models during the validation period (July 2023 to January 2025).

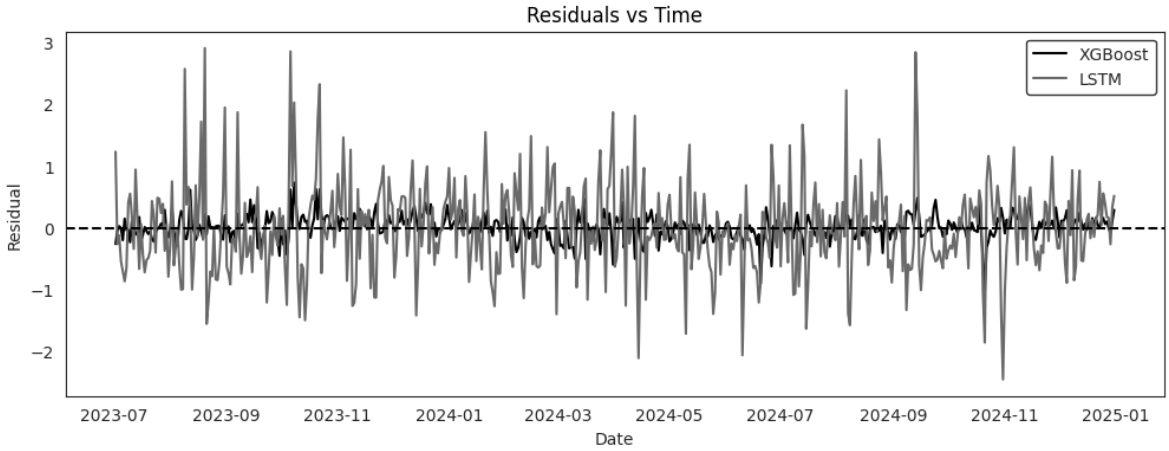

Figure 15. Residuals over time for both models.

The dashed line at zero represents the point of perfect prediction. It can be observed that XGBoost maintains a tightly bounded distribution of errors around zero, with fluctuations typically below ±0.5 mm day^{-1}, indicating high temporal stability and consistency—even during extreme weather events. In contrast, LSTM exhibits greater residual dispersion, with wider oscillations and peaks exceeding ±2 mm day^{-1}, particularly during periods of high atmospheric variability. This difference suggests that XGBoost is more effective at capturing short-term evaporative dynamics, whereas LSTM tends to dampen abrupt variations, compromising accuracy under transitional scenarios. For clarity, Table 3 presents the monthly aggregation of errors, showing that XGBoost sustains low and stable MAE (0.10–0.20 mm day^{-1}) and minimal bias across the year, while LSTM exhibits consistently higher MAE (≈0.5–0.9 mm day^{-1}) and negative biases, especially in September–November. These systematic deviations highlight the operational advantage of XGBoost for precision irrigation, since monthly underestimations in LSTM could translate into water stress during critical growing stages.

Table 3. Monthly MAE and Bias of XGBoost and LSTM.

Year-Month	N_XGB	MAE_XGB	Bias_XGB	N_LSTM	MAE_LSTM	Bias_LSTM
2023-07	30	0.11	-0.05	30	0.53	-0.35
2023-08	31	0.14	0.01	31	0.89	-0.28
2023-09	30	0.17	0.01	30	0.66	-0.53
2023-10	31	0.20	0.13	31	0.88	-0.41
2023-11	30	0.13	0.11	30	0.55	-0.07
2023-12	31	0.15	0.13	31	0.39	-0.04
2024-01	31	0.12	0.07	31	0.54	-0.20
2024-02	29	0.16	-0.07	29	0.55	-0.05
2024-03	31	0.21	-0.08	31	0.53	-0.08
2024-04	30	0.18	-0.01	30	0.55	-0.23
2024-05	31	0.12	-0.07	31	0.52	-0.24
2024-06	30	0.17	-0.11	30	0.55	-0.35
2024-07	31	0.11	-0.02	31	0.67	-0.40
2024-08	31	0.11	-0.01	31	0.59	-0.19
2024-09	30	0.14	0.06	30	0.87	-0.64
2024-10	31	0.15	-0.04	31	0.72	-0.63
2024-11	30	0.14	0.10	30	0.73	-0.55
2024-12	31	0.11	0.08	31	0.35	-0.09

Table 4 summarises the global performance metrics of the two evaluated models against the reference ET_0 values obtained using the Penman–Monteith method. XGBoost achieves a MAE of 0.145 mm day^{-1} and an RMSE of 0.196 mm day^{-1}, with a coefficient of determination R^2 of 0.983, MAPE of 3.2%, NSE of 0.983, and a Willmott index d of 0.996. In contrast, the LSTM model shows a much higher MAE (0.555 mm day^{-1}), RMSE of 0.729 mm day^{-1}, and R^2 of 0.759, along with a MAPE of 12.1%, NSE of 0.759, and Willmott d of 0.927. These results demonstrate that XGBoost not only significantly reduces both mean and squared error, but also improves the proportion of explained variance and prediction consistency, establishing itself as the most accurate and robust option for estimating ET_0 in precision irrigation applications.

Table 4. Summary of global metrics.

Model	MAE	RMSE	R^2	MAPE (%)	NSE	Willmott d
XGBoost	0.1449	0.1956	0.9826	3.1800	0.9826	0.9955
LSTM	0.5554	0.7286	0.7586	12.0500	0.7586	0.9269

Table 5 breaks down the performance of each model across the two defined climatic seasons: dry (November–April) and wet (May–October). During the dry season, XGBoost achieves a MAE of 0.149 mm day^{-1}, an RMSE of 0.189 mm day^{-1}, and a MAPE of 3.87%, whereas LSTM records a MAE of 0.488 mm day^{-1}, an RMSE of 0.617 mm day^{-1}, and a MAPE of 12.86%. In the wet season, XGBoost maintains stable performance with MAE = 0.141 mm day^{-1}, RMSE = 0.200 mm day^{-1}, and MAPE = 2.63%, while LSTM performs slightly worse with MAE = 0.609 mm day^{-1}, RMSE = 0.806 mm day^{-1}, and MAPE = 11.40%. These results confirm that XGBoost not only outperforms LSTM in overall accuracy, but is also more robust to seasonal variation, maintaining low absolute and relative errors under both dry and more humid conditions.

Table 5. Seasonal metrics by model and season.

Model	Season	MAE	RMSE	MAPE (%)
XGBoost	0.1449	0.1956	0.9826	3.1800
LSTM	0.5554	0.7286	0.7586	12.0500

Critical Synthesis of Results

Overall, the results obtained support the conclusion that the XGBoost model consistently outperforms the LSTM neural network in estimating reference evapotranspiration (ET_0) under semi-arid climatic conditions. This superiority is evident in global error metrics, the ability to capture extreme events, seasonal robustness, and lower residual variance. In agreement with previous reports that highlighted the robustness of tree-based models for agroclimatic prediction (Ge et al., 2022; Roy et al., 2022), the present findings confirm that gradient boosting can achieve accuracies comparable to, or higher than, more computationally intensive deep learning architectures. Additionally, normality tests applied to the residuals (Shapiro–Wilk and Kolmogorov–Smirnov) demonstrated that XGBoost errors conform more closely to Gaussian assumptions than those of LSTM,

reinforcing the statistical soundness of the model. The SHAP analysis further reveals that the tree-based architecture not only captures nonlinear relationships efficiently, but also does so in a manner consistent with the expected physical hierarchy among predictors, particularly shortwave solar radiation and vapour pressure deficit, as reported in classical agro-meteorological literature (Allen et al., 1998; Djaman & Irmak, 2013). From an agronomic perspective, the higher day-to-day precision of XGBoost implies more reliable irrigation scheduling, translating into reduced risks of under- or over-irrigation, with potential water savings of up to 10–15% in semi-arid cropping systems when compared to conventional scheduling methods (Moratiel et al., 2020). While the performance of LSTM is acceptable in capturing the general trend, its poor sensitivity to abrupt fluctuations—likely due to its inherent tendency to smooth the input—limits its direct applicability in operational contexts where high-frequency precision is required.

Operational Applicability and Technological Potential

The accuracy achieved by the XGBoost model, with average absolute errors below 0.15 mm day^{-1}, positions it as a viable tool for integration into real-time irrigation decision support systems, particularly in regions where access to high-resolution meteorological stations is limited. This predictive capability, derived from open-access satellite data and achieved with low computational cost, represents a significant step toward the deployment of precision agriculture solutions in semi-arid environments, where water use efficiency is critical (Başağaoğlu et al., 2022; Jagtap et al., 2022). Moreover, the ability to physically interpret model outputs through explainable techniques such as SHAP further enhances its acceptance among technicians and producers, facilitating its adoption within intelligent agronomic management frameworks.

Study Limitations

Despite the methodological robustness of this work, several limitations must be acknowledged. First, the training and validation of the models were conducted on a single irrigated plot with agroclimatic conditions specific to northwestern Mexico. This implies that generalising the results to other regions requires caution and additional validation efforts. Furthermore, although the reference evapotranspiration used as the target variable (ET_o_PM) was estimated via the Penman–Monteith equation using data from the NASA POWER system—and has been validated in previous studies (Rodrigues & Braga, 2021)—no direct

lysimetric measurements were available, which may introduce systematic errors that are difficult to quantify. Finally, the LSTM architecture employed was deliberately simple, leaving open the possibility that deeper configurations or models incorporating attention mechanisms could yield substantial improvements under certain conditions.

Projections and Future Research Directions

Future work could extend this approach through cross-validation across different agroecological regions of the country, particularly in areas with altitudinal variability or perennial crops with high water demand. It would also be valuable to explore the performance of hybrid architectures that combine the sequential learning capacity of LSTM or GRU networks with attention mechanisms or convolutional encoders, as recently proposed in studies involving agroclimatic variables (Li et al., 2024). Additionally, the implementation of federated learning techniques across agricultural plots could facilitate the development of more robust models while preserving data privacy (Žalik & Žalik, 2023). Finally, a promising direction lies in integrating the XGBoost model into local Internet of Things (IoT) platforms, enabling autonomous predictions directly from distributed sensor nodes.

Conclusions

This study demonstrates, with statistical and methodological rigor, that XGBoost consistently outperforms LSTM in the daily estimation of reference evapotranspiration (ET_0) under semi-arid conditions in northwestern Mexico. The superiority of XGBoost is evidenced by its markedly lower errors (MAE = 0.145 mm day^{-1}; R^2 = 0.983), symmetrical residuals even during extreme events, and robust seasonal stability. In contrast, the LSTM network exhibited systematic biases and reduced adaptability during transitional climatic periods.

Compared with the FAO Penman–Monteith method, XGBoost emerges as a reliable and computationally efficient alternative in contexts where ground-based meteorological records are scarce. Beyond predictive accuracy, the integration of SHAP analysis ensured physical interpretability by ranking downward shortwave radiation, vapor pressure deficit, and wind speed as the dominant drivers of ET_0, in full agreement with agro-meteorological theory.

The contribution of this work lies in being one of the first systematic evaluations of interpretable artificial intelligence for evapotranspiration in Latin America, applied to a multi-year dataset of satellite-derived meteorological variables. These results extend the

current evidence base and highlight the potential of transparent AI models to support irrigation scheduling and water management in data-limited regions.

From a practical standpoint, the demonstrated precision of XGBoost enables its operational deployment in IoT-based irrigation platforms, supporting water savings of 10–15% compared with conventional scheduling. The generalizability of these findings makes them relevant for semi-arid regions worldwide facing increasing water scarcity. Future research should expand this framework through multi-site validation, field-based calibration, and the exploration of hybrid or federated learning schemes, ultimately consolidating the role of artificial intelligence in sustainable agriculture under climate change scenarios.

References

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. https://www.fao.org/4/x0490e/x0490e00.htm

Ayaz, A., Rajesh, M., Singh, S. K., Rehana, S., Lab for Spatial Informatics, International Institute of Information Technology, Hyderabad, India, & National Institute of Water & Atmospheric Research Ltd (NIWA), New Zealand. (2021). Estimation of reference evapotranspiration using machine learning models with limited data. *AIMS Geosciences*, *7*(3), 268–290. https://doi.org/10.3934/geosci.2021016

Başağaoğlu, H., Chakraborty, D., Lago, C. D., Gutierrez, L., Şahinli, M. A., Giacomoni, M., Furl, C., Mirchi, A., Moriasi, D., & Şengör, S. S. (2022). A Review on Interpretable and Explainable Artificial Intelligence in Hydroclimatic Applications. *Water*, *14*(8), 1230. https://doi.org/10.3390/w14081230

Bojorquez Delgado, G. (2025). *Conjunto de datos meteorológicos y evapotranspiración de referencia (ET_0_PM) para la agricultura de precisión en el noroeste semiárido de México (2020–2024)* [Dataset]. Zenodo. https://doi.org/10.5281/ZENODO.16922140

Bojorquez-Delgado, G., Bojorquez-Delgado, J., & Flores-Rosales, M. A. (2023). Descomposición y análisis temporal del NDVI en un predio agrícola para determinar la salud y variabilidad de cultivos de maíz en Guasave, Sinaloa. *Revista Interdisciplinaria de Ingeniería Sustentable y Desarrollo Social*, *9*(1), 30–43. https://doi.org/10.63728/riisds.v9i1.92

Bour, S., Kayhomayoon, Z., Hassani, F., Ghordoyee Milan, S., Bazrafshan, O., & Berndtsson, R. (2025). Enhancement of standardized precipitation evapotranspiration index predictions by machine learning based on regression and soft computing for Iran's arid and hyper–arid region. *PLOS ONE*, *20*(3), e0319678. https://doi.org/10.1371/journal.pone.0319678

Chen, X., Wang, S., Gao, H., Huang, J., Shen, C., Li, Q., Qi, H., Zheng, L., & Liu, M. (2022). Comparison of deep learning models and a typical process-based model in glacio-hydrology simulation. *Journal of Hydrology*, *615*, 128562. https://doi.org/10.1016/j.jhydrol.2022.128562

Córdova, M., Carrillo-Rojas, G., Crespo, P., Wilcox, B., & Célleri, R. (2015). Evaluation of the Penman-Monteith (FAO 56 PM) Method for Calculating Reference Evapotranspiration Using Limited Data. *Mountain Research and Development*, *35*(3), 230. https://doi.org/10.1659/MRD-JOURNAL-D-14-0024.1

Djaman, K., & Irmak, S. (2013). Actual Crop Evapotranspiration and Alfalfa- and Grass-Reference Crop Coefficients of Maize under Full and Limited Irrigation and Rainfed Conditions. *Journal of Irrigation and Drainage Engineering*, *139*(6), 433–446. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000559

Dong, J., Xing, L., Cui, N., Zhao, L., Guo, L., Wang, Z., Du, T., Tan, M., & Gong, D. (2024). Estimating reference crop evapotranspiration using improved convolutional bidirectional long short-term memory network by multi-head attention mechanism in the four climatic zones of China. *Agricultural Water Management*, *292*, 108665. https://doi.org/10.1016/j.agwat.2023.108665

Gao, P., Xie, J., Yang, M., Zhou, P., Chen, W., Liang, G., Chen, Y., Han, X., & Wang, W. (2021). Improved Soil Moisture and Electrical Conductivity Prediction of Citrus Orchards Based on IoT Using Deep Bidirectional LSTM. *Agriculture*, *11*(7), 635. https://doi.org/10.3390/agriculture11070635

Ge, J., Zhao, L., Yu, Z., Liu, H., Zhang, L., Gong, X., & Sun, H. (2022). Prediction of Greenhouse Tomato Crop Evapotranspiration Using XGBoost Machine Learning Model. *Plants*, *11*(15), 1923. https://doi.org/10.3390/plants11151923

Graciano-Obeso, A., López-Atondo, J. U., & Báez-Higuera, J. A. (2024). Influencia de la aplicación de hongos entomopatógenos sobre el rendimiento de tomate en Guasave,

Sinaloa. *Revista Interdisciplinaria de Ingeniería Sustentable y Desarrollo Social*, *10*(1), 451–463. https://doi.org/10.63728/riisds.v10i1.58

Helali, J., Mohammadi Ghaleni, M., Mianabadi, A., Asadi Oskouei, E., Momenzadeh, H., Haddadi, L., & Saboori Noghabi, M. (2024). Enhancing references evapotranspiration forecasting with teleconnection indices and advanced machine learning techniques. *Applied Water Science*, *14*(10), 219. https://doi.org/10.1007/s13201-024-02289-x

Irmak, S., Irmak, A., Allen, R. G., & Jones, J. W. (2003). Solar and Net Radiation-Based Equations to Estimate Reference Evapotranspiration in Humid Climates. *Journal of Irrigation and Drainage Engineering*, *129*(5), 336–347. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:5(336)

Irmak, S., Odhiambo, L. O., Specht, J. E., & Djaman, K. (2013). Hourly and daily single and basal evapotranspiration crop coefficients as a function of growing degree days, days after emergence, leaf area index, fractional green canopy cover, and plant phenology for soybean. *Transactions of the ASABE*, *56*(5), 1785–1803. https://doi.org/10.13031/trans.56.10219

Jagtap, S. T., Phasinam, K., Kassanuk, T., Jha, S. S., Ghosh, T., & Thakar, C. M. (2022). Towards application of various machine learning techniques in agriculture. *Materials Today: Proceedings*, *51*, 793–797. https://doi.org/10.1016/j.matpr.2021.06.236

Jia, W., Zhang, Y., Wei, Z., Zheng, Z., & Xie, P. (2023). Daily reference evapotranspiration prediction for irrigation scheduling decisions based on the hybrid PSO-LSTM model. *PLOS ONE*, *18*(4), e0281478. https://doi.org/10.1371/journal.pone.0281478

Lepot, M., Aubin, J.-B., & Clemens, F. (2017). Interpolation in Time Series: An Introductive Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment. *Water*, *9*(10), 796. https://doi.org/10.3390/w9100796

Li, M., Zhou, Q., Han, X., & Lv, P. (2024). Prediction of reference crop evapotranspiration based on improved convolutional neural network (CNN) and long short-term memory network (LSTM) models in Northeast China. *Journal of Hydrology*, *645*, 132223. https://doi.org/10.1016/j.jhydrol.2024.132223

Liang, Y., Feng, D., Sun, Z., & Zhu, Y. (2023). Evaluation of Empirical Equations and Machine Learning Models for Daily Reference Evapotranspiration Prediction Using Public Weather Forecasts. *Water*, *15*(22), 3954. https://doi.org/10.3390/w15223954

Liu, D., Wang, Z., Wang, L., Chen, J., Li, C., & Shi, Y. (2024). Improved remote sensing reference evapotranspiration estimation using simple satellite data and machine learning. *Science of The Total Environment*, *947*, 174480. https://doi.org/10.1016/j.scitotenv.2024.174480

Liyew, C. M., Ferraris, S., Di Nardo, E., & Meo, R. (2025). A review of feature selection methods for actual evapotranspiration prediction. *Artificial Intelligence Review*, *58*(10), 292. https://doi.org/10.1007/s10462-025-11298-4

Matos, K. A., Bowen, G. J., Good, S. P., & Allen, S. T. (2025). Evapotranspiration Partitioning Across US Ecoregions: A Multi-Site Study Using Field Stable-Isotope Observations. *Geophysical Research Letters*, *52*(14), e2025GL115367. https://doi.org/10.1029/2025GL115367

Moratiel, R., Bravo, R., Saa, A., Tarquis, A. M., & Almorox, J. (2020). Estimation of evapotranspiration by the Food and Agricultural Organization of the United Nations (FAO) Penman–Monteith temperature (PMT) and Hargreaves–Samani (HS) models under temporal and spatial criteria – a case study in Duero basin (Spain). *Natural Hazards and Earth System Sciences*, *20*(3), 859–875. https://doi.org/10.5194/nhess-20-859-2020

Mushtaq, H., Akhtar, T., Hashmi, M. Z. U. R., Masood, A., & Saeed, F. (2024). Hydrologic interpretation of machine learning models for 10-daily streamflow simulation in climate sensitive upper Indus catchments. *Theoretical and Applied Climatology*, *155*(6), 5525–5542. https://doi.org/10.1007/s00704-024-04932-8

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290.

Noor, M. N., Yahaya, A. S., Ramli, N. A., & Al Bakri Abdullah, M. M. (2015). Filling the Missing Data of Air Pollutant Concentration Using Single Imputation Methods. *Applied Mechanics and Materials*, *754–755*, 923–932. https://doi.org/10.4028/www.scientific.net/AMM.754-755.923

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), 594. https://doi.org/10.3390/agronomy12030594

Sarkar, S. S., Bedi, J., & Jain, S. (2025). A deep learning based framework for enhanced reference evapotranspiration estimation: Evaluating accuracy and forecasting strategies. *Scientific Reports*, *15*(1), 15136. https://doi.org/10.1038/s41598-025-99713-2

Sharafi, S., & Mohammadi Ghaleni, M. (2024). Revealing accuracy in climate dynamics: Enhancing evapotranspiration estimation using advanced quantile regression and machine learning models. *Applied Water Science*, *14*(7), 162. https://doi.org/10.1007/s13201-024-02211-5

Soylu, M. E., Lenters, J. D., Istanbulluoglu, E., & Loheide, S. P. (2012). On evapotranspiration and shallow groundwater fluctuations: A Fourier-based improvement to the White method. *Water Resources Research*, *48*(6), 2011WR010964. https://doi.org/10.1029/2011WR010964

Tausif, M., Iqbal, M. W., Bashir, R. N., AlGhofaily, B., Elyassih, A., & Khan, A. R. (2025). Federated learning based reference evapotranspiration estimation for distributed crop fields. *PLOS ONE*, *20*(2), e0314921. https://doi.org/10.1371/journal.pone.0314921

Willmott, C. J. (1981). ON THE VALIDATION OF MODELS. *Physical Geography*, *2*(2), 184–194. https://doi.org/10.1080/02723646.1981.10642213

Wu, T., Zhang, W., Jiao, X., Guo, W., & Hamoud, Y. A. (2020). Comparison of five Boosting-based models for estimating daily reference evapotranspiration with limited meteorological variables. *PLOS ONE*, *15*(6), e0235324. https://doi.org/10.1371/journal.pone.0235324

Xu, X., Du, Z., Bai, Z., Wang, S., Wang, C., & Li, D. (2023). Fault diagnosis method of dissolved oxygen sensor electrolyte loss based on impedance measurement. *Computers and Electronics in Agriculture*, *212*, 108123. https://doi.org/10.1016/j.compag.2023.108123

Yan, X., Yang, N., Ao, R., Mohammadian, A., Liu, J., Cao, H., & Yin, P. (2023). Deep learning for daily potential evapotranspiration using a HS-LSTM approach. *Atmospheric Research*, *292*, 106856. https://doi.org/10.1016/j.atmosres.2023.106856

Yang, D., Yang, S., Huang, J., Zhang, S., Zhang, S., Zhang, J., & Bai, Y. (2025). Improving time upscaling of instantaneous evapotranspiration based on machine learning models. *Big Earth Data*, *9*(1), 127–154. https://doi.org/10.1080/20964471.2024.2423431

Yin, J., Deng, Z., Ines, A. V. M., Wu, J., & Rasu, E. (2020). Forecast of short-term daily reference evapotranspiration under limited meteorological variables using a hybrid bi-directional long short-term memory model (Bi-LSTM). *Agricultural Water Management*, *242*, 106386. https://doi.org/10.1016/j.agwat.2020.106386

Zhou, H., Shao, J., Liu, H., Du, Z., Zhou, L., Liu, R., Bernhofer, C., Grünwald, T., Dušek, J., Montagnani, L., Tagesson, T., Black, T. A., Jassal, R., Woodgate, W., Biraud, S., Varlagin, A., Mammarella, I., Gharun, M., Shekhar, A., … Zhou, X. (2021). Relative importance of climatic variables, soil properties and plant traits to spatial variability in net CO2 exchange across global forests and grasslands. *Agricultural and Forest Meteorology*, *307*, 108506. https://doi.org/10.1016/j.agrformet.2021.108506