www.envsciarch.com

ENVIRONMENTAL SCIENCE ARCHIVES
ISSN: 2583-5092
Volume V Issue 2, 2026

Received: 2026/08/01
Accepted: 2026/09/02
Published: 2026/09/06

check for updates

RESEARCH PAPER

OPEN ACCESS

Machine Learning-Based Estimation of Monthly Potential Evapotranspiration for Sustainable Water Resources Management using Random Forest and Support Vector Machine

Vivek Kumar[1], Azmat Raja[1], Nishikant Kumar[2], Raushan Anand[2], Ashish Anand[1], Raunak Kumar[3], Pusplata Puja[4] and Om Prakash Singh[1]

[1]Department of Civil Engineering, Supaul College of Engineering, Supaul, Bihar, India
[2]Department of Civil Engineering, BP Mandal College of Engineering, Madhepura, Bihar, India
[3] Department of Civil Engineering, Government Engineering College, West Champaran, Bettiah, Bihar, India
[4]Department of Civil Engineering, Saharsa College of Engineering, Saharsa, Bihar, India
*Correspondence for materials should be addressed to AR (email: azmatr.phd18.ce@nitp.ac.in)

Abstract
Accurate estimation of potential evapotranspiration (PET) is essential for irrigation planning, agricultural water allocation, drought assessment, and sustainable water resources management. Conventional PET methods require several meteorological variables that may be incomplete or unavailable in data-scarce regions. This study evaluates Random Forest (RF) and Support Vector Machine (SVM) for monthly PET estimation in Supaul District, Bihar, India, using monthly meteorological data from 2001–2022. Seven predictors—specific humidity at 2 m (QV2M), relative humidity at 2 m (RH2M), wind speed at 2 m (WS2M), maximum and minimum air temperature (T2M_MAX and T2M_MIN), ultraviolet radiation (UVB), and sunshine duration (SDDN)—were used, while Penman-derived PET was the target. Model performance was evaluated using R^2, RMSE, MAE, MBE, NSE, KGE, Willmott's index of agreement, residual diagnostics, Taylor diagrams, and Bland–Altman analysis. RF achieved $R^2 = 0.976$, RMSE = 6.30 mm month^{-1}, MAE = 4.30 mm month^{-1}, MBE = −0.21 mm month^{-1}, NSE = 0.976, KGE = 0.897, and d = 0.994. SVM achieved $R^2 = 0.828$, RMSE = 16.74 mm month^{-1}, MAE = 12.24 mm month^{-1}, and MBE = −1.94 mm month^{-1}. RF consistently showed closer agreement with the Penman-derived reference series and smaller residual dispersion. The results indicate that RF provides an efficient data-driven approach for reproducing monthly Penman-derived PET in the study region and may support irrigation and water-resource planning where complete operational calculations are inconvenient.

Keywords: Potential evapotranspiration; Machine learning; Water resources management; Irrigation planning; Sustainable civil engineering

Introduction
Potential evapotranspiration (PET) is a fundamental component of the hydrological cycle, representing atmospheric water demand under adequate moisture availability. It is widely used in irrigation scheduling, crop water requirement estimation, drought monitoring, watershed hydrology, groundwater recharge assessment, reservoir operation, and climate-change impact

studies. Reliable PET estimation is therefore essential for sustainable water resources management, particularly in agricultural regions facing increasing water demand and climatic variability (Allen et al., 1998; Fisher et al., 2017).

Several methods have been developed for PET estimation, ranging from empirical temperature-based equations to physically based approaches. Commonly used methods include Thornthwaite, Hargreaves–Samani, Priestley–Taylor, Penman, and FAO-56 Penman–Monteith. Among these, Penman and FAO-56 Penman–Monteith are widely regarded as robust reference methods because they incorporate multiple meteorological controls, including temperature, humidity, wind speed, and radiation (Penman, 1948; Allen et al., 1998; Djaman et al., 2018). However, their application requires multiple meteorological variables, which may be unavailable or incomplete in developing regions because of limited monitoring infrastructure, missing observations, and measurement uncertainties (Sentelhas et al., 2010; Martí et al., 2015).

Recent advances in artificial intelligence and machine learning (ML) have provided alternative approaches for modelling complex hydrological processes. Unlike conventional empirical and physically based methods, ML algorithms can learn nonlinear relationships from historical observations without explicitly specifying the underlying functional relationships. Their ability to process multidimensional datasets and represent complex interactions among climatic variables has supported applications in rainfall–runoff modelling, streamflow forecasting, groundwater prediction, drought assessment, soil moisture estimation, and evapotranspiration modelling (Breiman, 2001; Elbeltagi et al., 2022; Deb et al., 2025; Bhavya et al., 2025).

Among ML techniques, Random Forest (RF) and Support Vector Machine (SVM) are widely used in environmental and hydrological applications. RF employs an ensemble of decision trees generated through bootstrap sampling and random feature selection, thereby improving robustness and reducing prediction variance. SVM uses statistical learning theory and kernel functions to represent nonlinear relationships in a transformed feature space. Their performance, however, depends on climatic conditions, predictor selection, data quality, and regional hydro-meteorological characteristics (Breiman, 2001; Smola and Schölkopf, 2004; Kisi, 2015). Recent studies have reported promising performance of ML approaches for evapotranspiration estimation across different climatic conditions, with RF frequently demonstrating strong predictive capability, while SVM performance can be more sensitive to kernel selection and hyper parameter optimization. Recent studies have also demonstrated the applicability of soft-computing techniques for environmental prediction, including vegetation-dynamics modelling (Raj and Gopikrishnan, 2024).

Although previous studies have commonly evaluated ML models using statistical measures such as R^2, RMSE, and MAE, relatively fewer studies have combined conventional metrics with complementary diagnostics such as residual analysis, Q–Q plots, Taylor diagrams, Bland–Altman analysis (Willmott, 1981; Bland & Altman 1986; Taylor, 2001; Gupta et al., 2009). Such complementary evaluation can provide a broader assessment of predictive accuracy, agreement, error distribution, and model reliability.

Accordingly, this study develops and compares RF and SVM models for monthly PET estimation in Supaul District, Bihar, India, using NASA POWER-derived meteorological data from 2001–2022. The objectives are to:

- Estimate monthly PET using the Penman method from long-term NASA POWER-derived meteorological data (2001–2022).
- Develop RF and SVM models for monthly PET prediction.
- Compare model performance using R^2, RMSE, MAE, KGE and NSE.
- Assess model reliability using residual analysis, Q–Q plots, Taylor diagrams, Bland–Altman analysis.
- Identify the more suitable ML approach for monthly PET estimation in the humid subtropical agricultural environment of Supaul District.

Study Area

Supaul District is located in the northeastern part of Bihar, India, between 25°37′–26°25′ N latitude and 86°28′–87°10′ E longitude, covering an area of approximately 2,410 km² (Fig. 1). The district shares its northern boundary with Nepal and is bordered by the districts of Madhepura, Saharsa, Araria, and Madhubani. Situated within the middle Gangetic alluvial plains, the region is characterized by predominantly flat terrain with elevations ranging from approximately 40 to 70 m above mean sea level. The fertile alluvial deposits of the Kosi River and its tributaries make the district one of the major agricultural regions of eastern India.

Agriculture is the dominant economic activity in Supaul District, where paddy, wheat, maize, pulses, and oilseed crops constitute the principal cropping system. Because crop production relies heavily on irrigation, reliable estimation of PET is essential for irrigation scheduling, crop water requirement assessment, and efficient water allocation. In addition, the district is frequently affected by seasonal flooding associated with the Kosi River and periodic water stress during dry months, emphasizing the importance of effective water resources management.

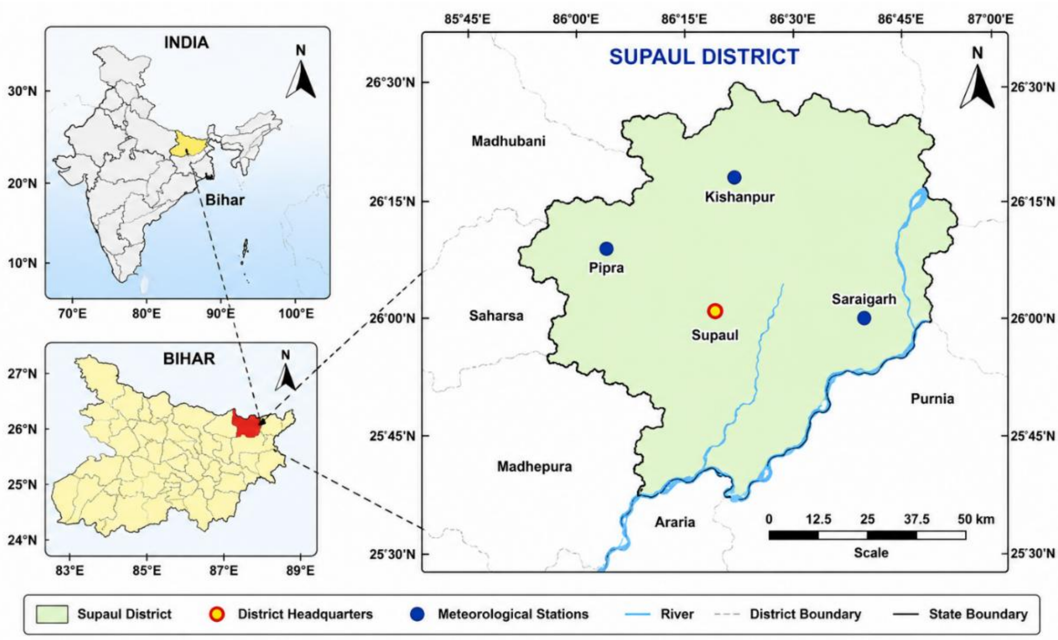


Fig. 1. Location of the study area showing the geographical position of Supaul District in Bihar, India, and the distribution of meteorological observation stations used in this study.

Materials and methods

Overall Methodological Framework

An integrated framework was adopted to estimate monthly potential evapotranspiration (PET) using the Penman method and two machine learning algorithms, namely Random Forest (RF) and Support Vector Machine (SVM). The methodology comprised six sequential stages. The complete workflow is illustrated in Fig. 2.

Meteorological Data Collection

Monthly NASA POWER-derived meteorological data from January 2001 to December 2022 were obtained from the NASA Prediction of Worldwide Energy Resources (NASA POWER) database (Sparks, 2018). The database provides globally consistent satellite-derived and reanalysis-based meteorological information and has been extensively applied in hydrological and agricultural studies.

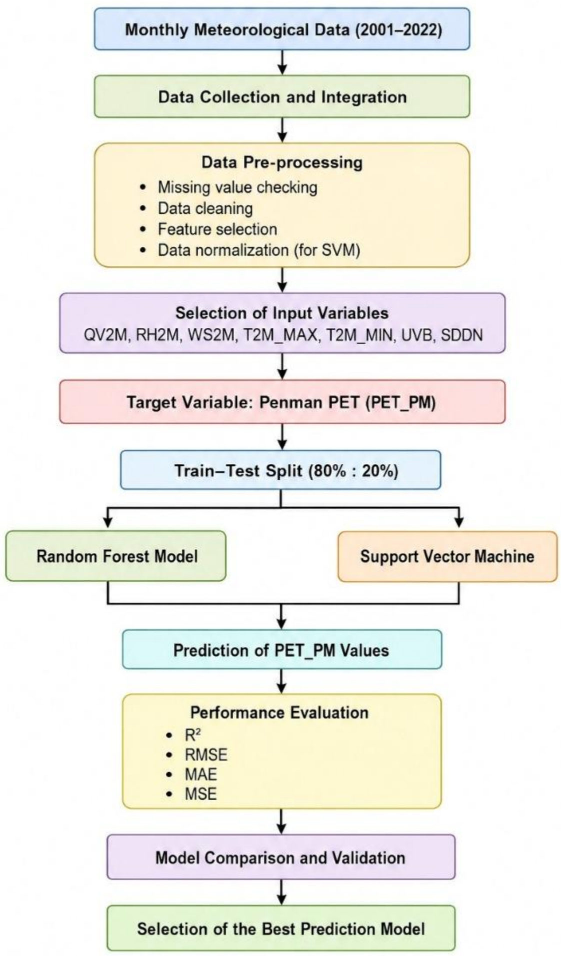


Fig 2. Overall research methodology adopted for monthly PET estimation using Random Forest and Support Vector Machine models. (Right)

Estimation of Potential Evapotranspiration

Monthly PET was computed using the Penman equation, which combines surface energy balance and aerodynamic transport processes by integrating net radiation, air temperature, atmospheric humidity, and wind speed (Penman, 1948; Allen et al., 1998). Owing to its physical basis and wide applicability across diverse climatic conditions, the Penman method was adopted as the reference approach in this study. The calculated monthly PET values served as target values for training and validating the machine learning models.

Data Preprocessing

Following quality control, the processed dataset was organized into a consistent monthly time series for subsequent model development. The predictor variables were examined to ensure consistent units, temporal alignment, and compatibility with the target PET series. The final dataset was then structured into input features and the corresponding target variable, ensuring that each observation contained the required meteorological predictors and associated PET value. This preprocessing step provided a standardized and reliable dataset for training, validation, and comparative evaluation of the Random Forest (RF) and Support Vector Machine (SVM) models.

Random Forest Model

Random Forest is an ensemble learning algorithm that constructs multiple regression trees using bootstrap samples and random feature selection (Breiman, 2001). The final PET prediction is obtained by averaging the outputs of all decision trees, thereby reducing prediction variance and improving model robustness (Fig. 3). The algorithm is well suited for hydrological applications because it effectively captures nonlinear relationships, is resistant to overfitting, and provides estimates of predictor importance (Elbeltagi et al., 2022). The Random Forest model was implemented using 50 decision trees. A fixed random seed of 123 was used to ensure reproducibility.

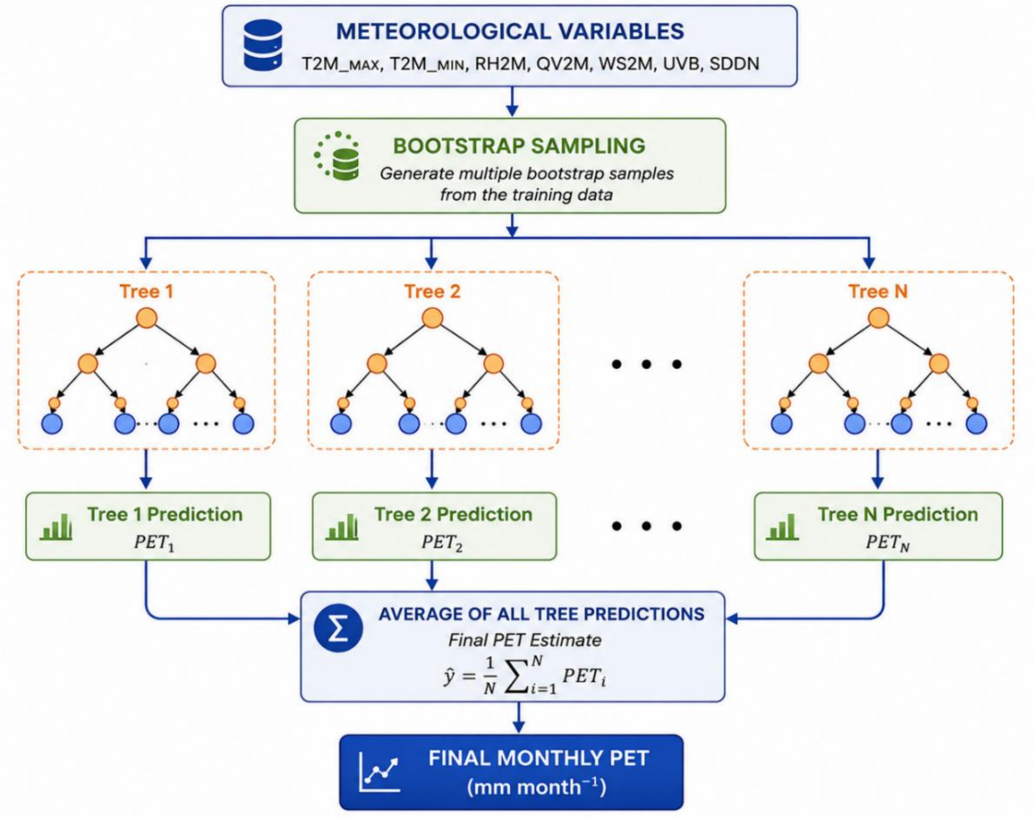


Fig 3. Workflow illustrating the development and validation of the Random Forest model.

Support Vector Machine Model

Support Vector Machine is a supervised machine learning algorithm that estimates nonlinear regression functions by mapping predictor variables into a higher-dimensional feature space through kernel functions Smola and Schölkopf (2004). Predictor variables were normalized before model training to improve optimization efficiency. The SVM workflow adopted in this study is shown in Fig. 4. SVM model using an epsilon-regression formulation with a radial basis function (RBF) kernel.

Model Training and Validation

The complete dataset was randomly divided into training (80%) and testing (20%) subsets using a fixed random seed to ensure reproducibility. The training dataset was used to develop both RF and SVM models, whereas the testing dataset was reserved exclusively for independent model evaluation, thereby minimizing overfitting and providing an unbiased assessment of predictive performance.

Model Performance Evaluation

Model performance was evaluated using four statistical indicators (Nash and Sutcliffe, 1970):

- Coefficient of Determination (R^2)
- Root Mean Square Error (RMSE)
- Mean Absolute Error (MAE)
- NSE

To further examine prediction reliability and uncertainty, additional diagnostic analyses were performed, including:

- Observed versus predicted scatter plots
- Residual plots
- Residual histograms
- Quantile–Quantile (Q–Q) plots

- Feature importance analysis
- Taylor diagrams
- Bland–Altman analysis

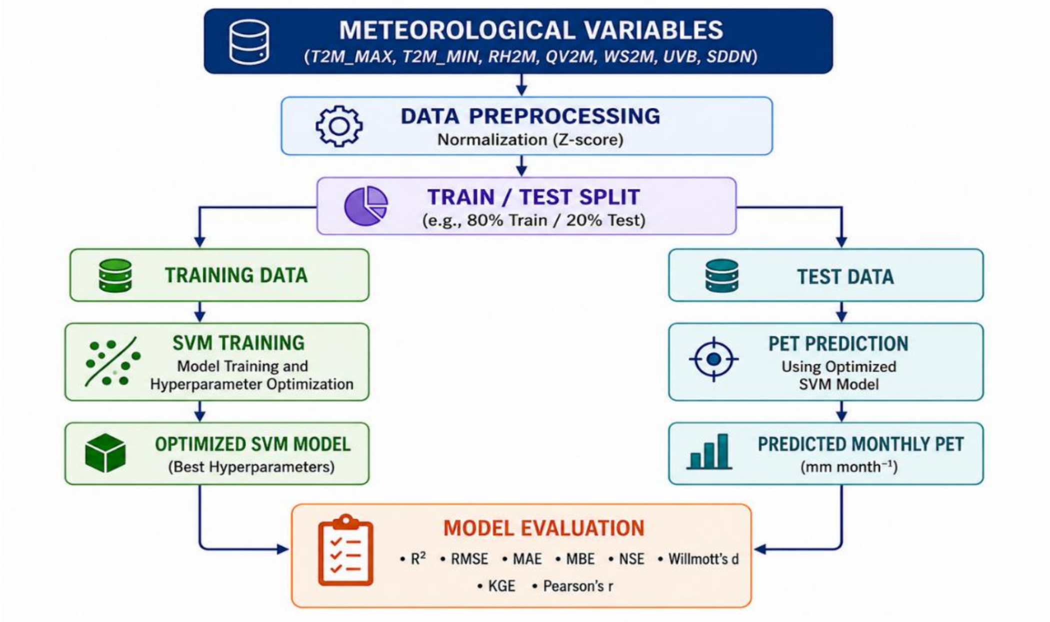


Fig 4. Workflow illustrating the development and validation of the Support Vector Machine model.

Results and Discussion

Temporal Variation of Monthly Potential Evapotranspiration

The temporal variation of Penman-derived monthly potential evapotranspiration (PET) during 2001–2022 is presented in Fig. 5. A pronounced seasonal pattern is evident, with PET increasing from winter to the pre-monsoon season, reaching its maximum during April–June, and subsequently declining during the southwest monsoon before attaining minimum values in winter. This seasonal behaviour reflects the combined influence of air temperature, solar radiation, atmospheric humidity, and wind speed on atmospheric evaporative demand.

The elevated PET observed during the pre-monsoon period is primarily associated with higher temperatures, greater incoming solar radiation, and lower relative humidity, which collectively enhance the vapour pressure deficit and increase evaporative demand. In contrast, PET decreases during the monsoon despite relatively high temperatures because increased cloud cover reduces available solar radiation while high humidity suppresses evaporation. During winter, lower temperatures and reduced solar radiation further decrease atmospheric water demand, resulting in the lowest PET values (Djaman et al., 2018).

The observed seasonal pattern is consistent with the climatic characteristics of humid subtropical regions and highlights the importance of accurate PET estimation for irrigation scheduling and agricultural water management. Reliable prediction of seasonal PET variations is essential for optimizing irrigation practices, improving water-use efficiency, and supporting climate-resilient agricultural planning.

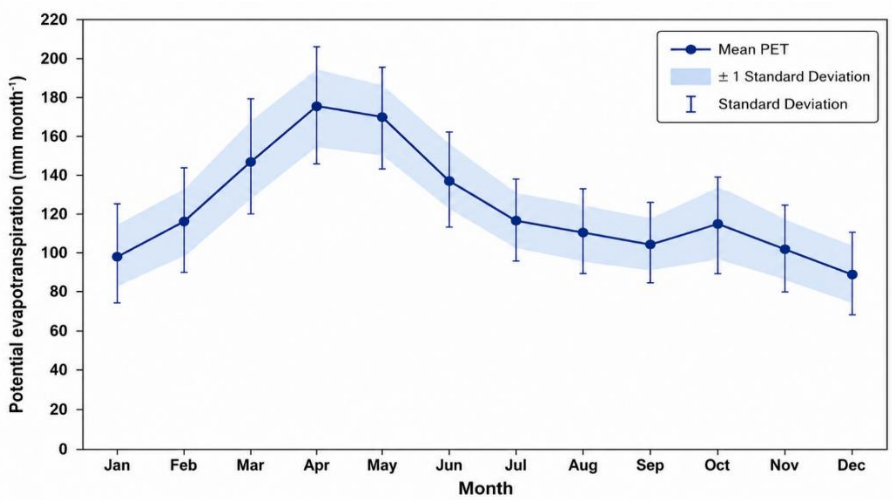
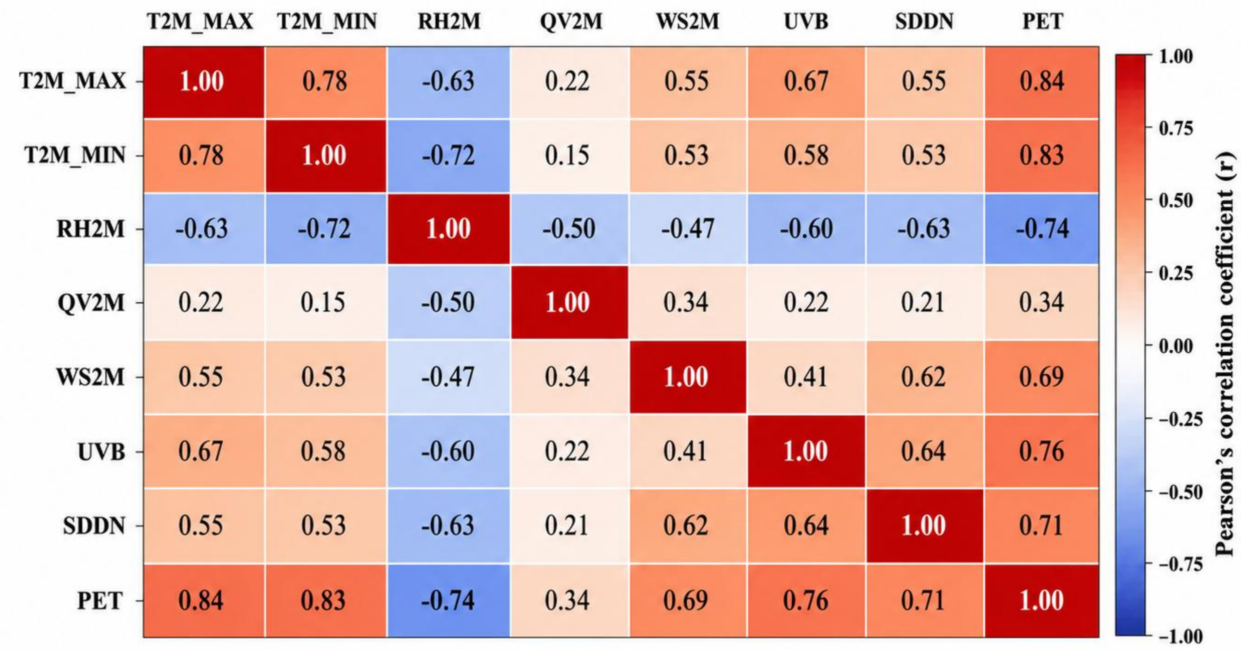


Fig 5. Mean monthly PET in Supaul District during the study period (2001–2022).

Relationship between Meteorological Variables and Potential Evapotranspiration

The correlation matrix presented in Fig. 6 illustrates the linear relationships among the meteorological variables used in this study and their association with monthly Penman-derived potential evapotranspiration (PET). Pearson's correlation coefficients reveal that PET is influenced by multiple climatic variables rather than by a single atmospheric parameter, highlighting the complexity of evapotranspiration processes in humid subtropical environments. The observed correlation patterns provide valuable insight into the physical mechanisms governing atmospheric evaporative demand and justify the application of nonlinear machine learning models for PET estimation. Among all predictor variables, maximum air temperature (T2M_MAX) exhibited the strongest positive correlation with PET, indicating that increasing air temperature substantially enhances atmospheric evaporative demand.



Variable	Description	Unit	Variable	Description	Unit
QV2M	Specific humidity at 2 m	kg kg^{-1}	T2M_MIN	Minimum air temperature	°C
RH2M	Relative humidity at 2 m	%	UVB	Ultraviolet radiation	W m^{-2}
WS2M	Wind speed at 2 m	m s^{-1}	SDDN	Sunshine duration	h
T2M_MAX	Maximum air temperature	°C	PET	Penman-derived monthly potential evapotranspiration	mm $month^{-1}$

Note: Warm colours (red) indicate positive correlations and cool colours (blue) indicate negative correlations. Values in each cell are Pearson correlation coefficients (r).

Fig. 6. Correlation matrix illustrating the relationships among meteorological variables and Penman-derived monthly PET.

Performance of the Random Forest and Support Vector Machine Model

The predictive performance of the Random Forest (RF) and Support Vector Machine (SVM) models is illustrated in Figs. 7 and 8, respectively, which compare the observed and predicted monthly potential evapotranspiration (PET) values for the testing dataset. The close agreement between the observed and predicted values, together with the strong clustering of data points around the 1:1 reference line, demonstrates the ability of the model to accurately reproduce the seasonal and interannual variability of monthly PET.

The RF model demonstrated excellent predictive performance, with close agreement between observed and predicted PET (R^2 = 0.976, RMSE = 6.30 mm $month^{-1}$, MAE = 4.30 mm $month^{-1}$, and MBE = −0.21 mm $month^{-1}$). Its superior performance can be attributed to the ensemble structure of RF, which effectively captures nonlinear relationships among meteorological variables while reducing prediction variance and overfitting. These results are consistent with previous studies reporting strong RF performance for evapotranspiration estimation under diverse climatic conditions (Bhavya et al., 2025). In comparison, SVM showed good but lower predictive accuracy (R^2 = 0.828, RMSE = 16.74 mm $month^{-1}$, MAE = 12.24 mm $month^{-1}$, and MBE = −1.94 mm $month^{-1}$), indicating slight underestimation. Its lower performance may reflect greater sensitivity to kernel selection, hyperparameter optimization, and complex nonlinear interactions under seasonal variability. Nevertheless, SVM reproduced the overall temporal variability of monthly PET, although RF provided more accurate and stable predictions.

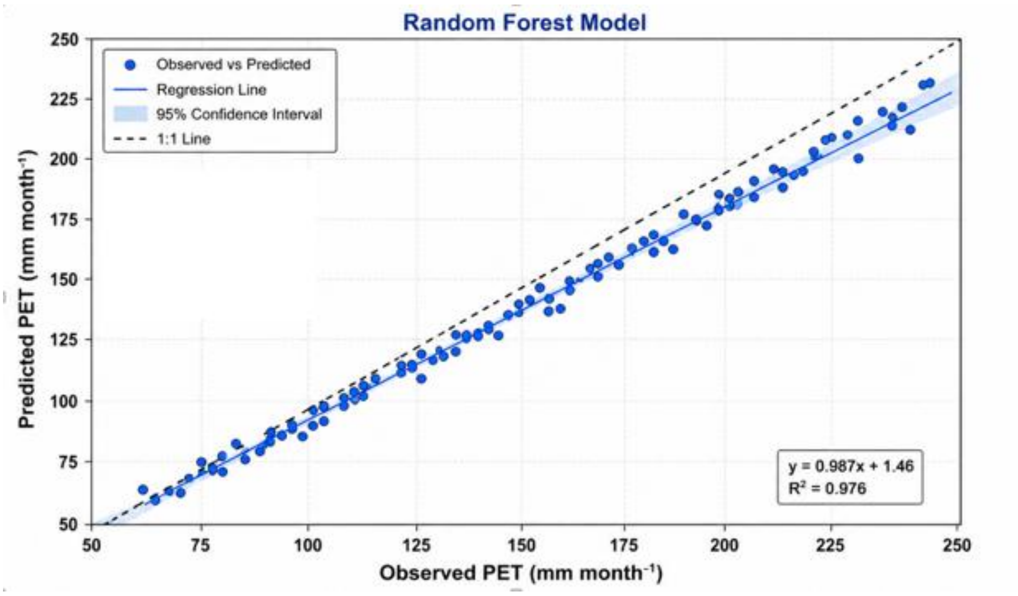


Fig. 7. Observed versus Random Forest-predicted monthly PET.

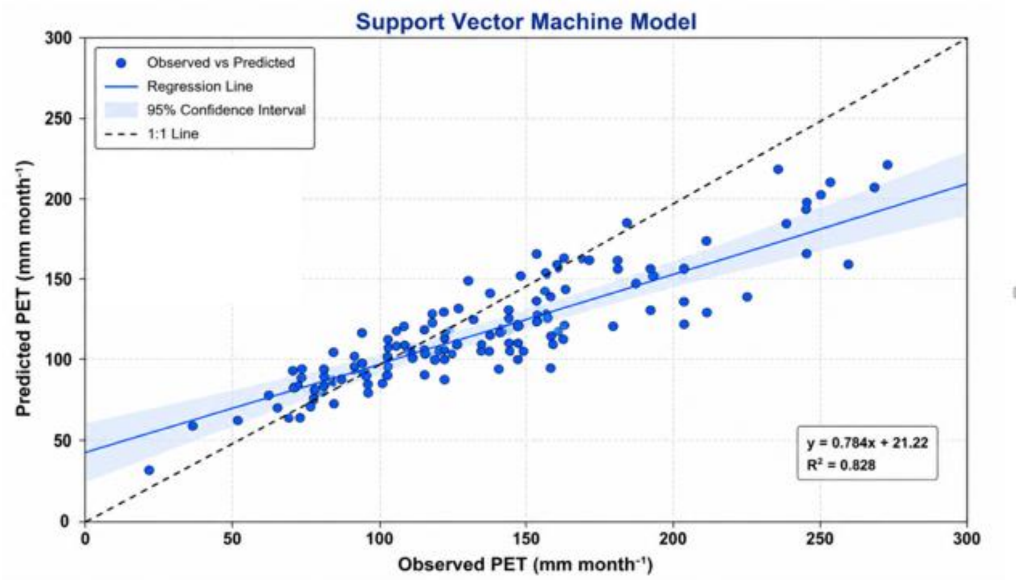


Fig. 8. Observed versus SVM-predicted monthly PET.

Residual Analysis and Error Distribution

Residual analysis provides valuable insight into the reliability, stability, and uncertainty of predictive models beyond conventional performance statistics. While indicators such as the coefficient of determination (R^2), RMSE, and MAE quantify the overall predictive capability of a model, residual diagnostics reveal whether prediction errors are randomly distributed, normally distributed, and free from systematic bias. In the present study, residual plots, residual frequency distributions, and Quantile–Quantile (Q–Q) plots were employed to comprehensively evaluate the performance of the Random Forest and Support Vector Machine models (Fig. 9–11).

Residual diagnostics showed that both RF and SVM errors were generally distributed around the zero-error line without obvious systematic trends or substantial heteroscedasticity. However, RF exhibited a narrower and more uniform residual distribution, indicating lower prediction uncertainty and greater stability than SVM. The residual histograms (Fig. 10) similarly showed approximately symmetric distributions centred near zero, with RF producing smaller errors, consistent with its lower RMSE (6.30 mm month^{-1}) and MAE (4.30 mm month^{-1}) compared with SVM (16.74 and 12.24 mm month^{-1}, respectively). Q–Q plots (Fig. 11) further indicated that RF residuals more closely followed the theoretical distribution, whereas SVM showed greater deviations at the tails, indicating less conformity to the assumed residual distribution.

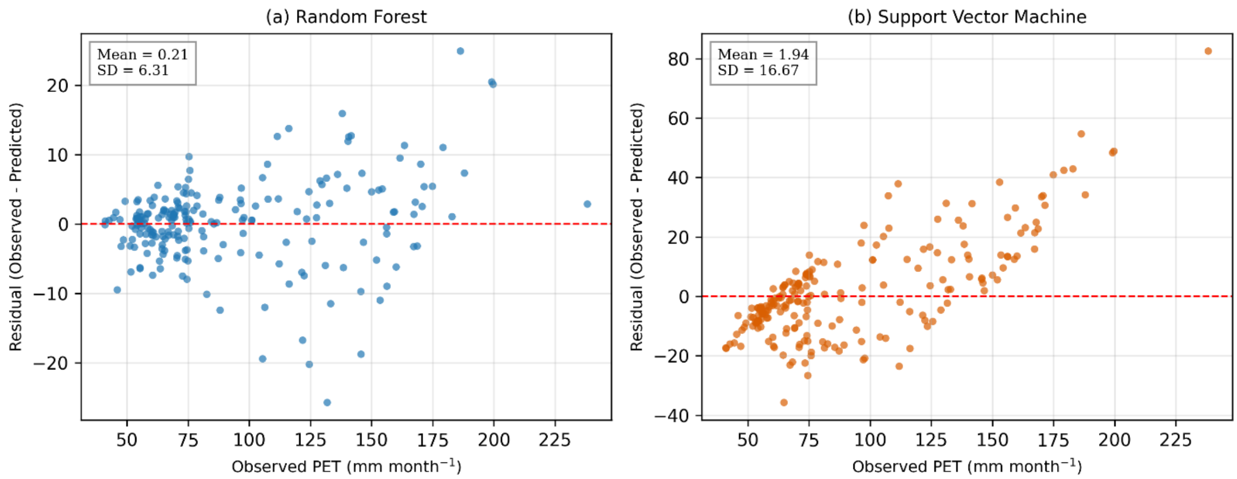


Fig. 9. Residual plots for the Random Forest (a) and Support Vector Machine (b) models.

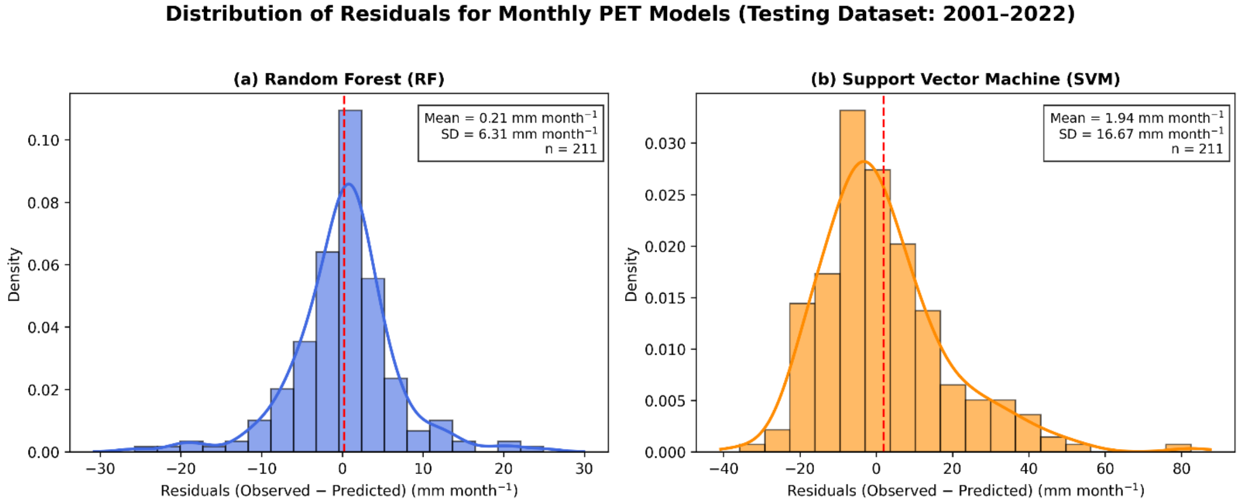


Fig. 10. Histograms of residuals for the Random Forest (a) and Support Vector Machine (SVM) (b) models.

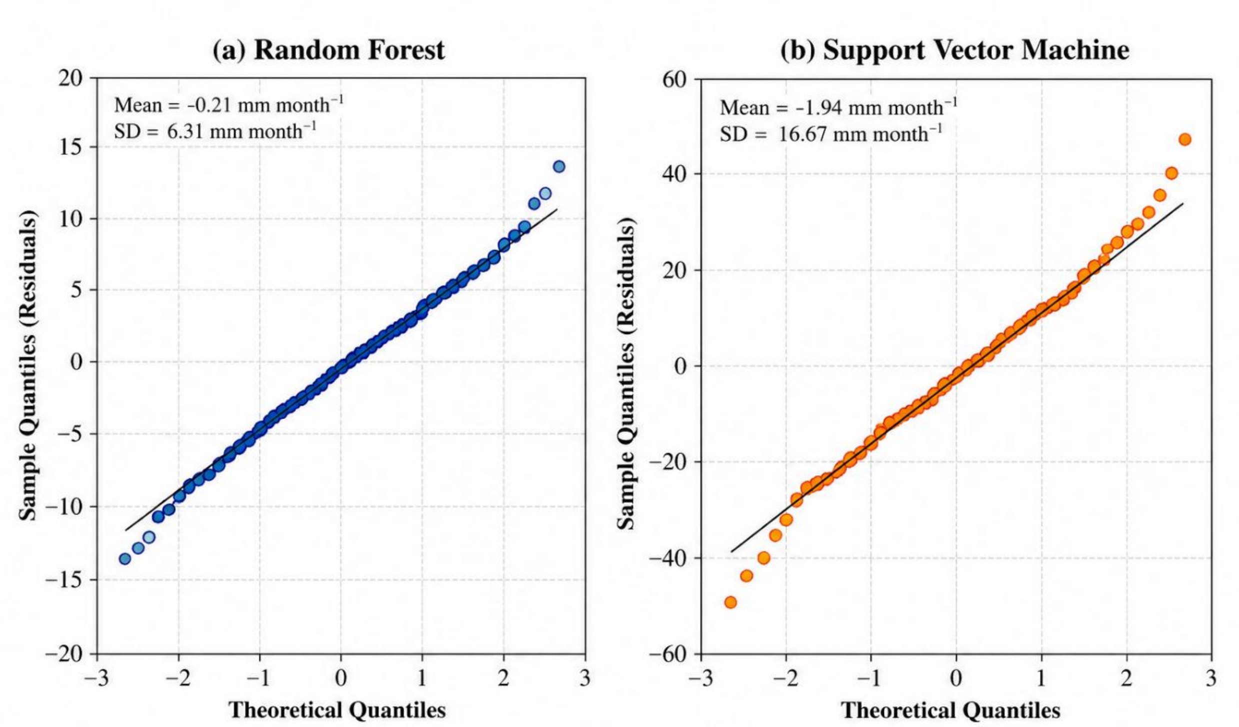


Fig. 11. Quantile–Quantile (Q–Q) plots for the Random Forest (a) and Support Vector Machine (b) models

Comparative Performance Evaluation of Machine Learning Models

Fig. 12 demonstrates the overall superiority of the RF model over SVM across multiple statistical and hydrological metrics. RF achieved higher R^2 (0.9756), NSE (0.9756), KGE (0.897), and Willmott's agreement index (0.9937), while producing substantially lower RMSE (6.30 mm month^{-1}), MAE (4.30 mm month^{-1}), and MBE (−0.21 mm month^{-1}) than SVM (R^2 and NSE = 0.8276; KGE = 0.443; d = 0.9421; RMSE = 16.74 mm month^{-1}; MAE = 12.24 mm month^{-1}; MBE = −1.94 mm month^{-1}). These results indicate that RF provides more accurate predictions, lower systematic bias, and better representation of the variability and temporal behaviour of monthly PET.

The findings are consistent with previous studies reporting superior performance of ensemble learning approaches for evapotranspiration estimation compared with kernel-based methods (Elbeltagi et al., 2022).

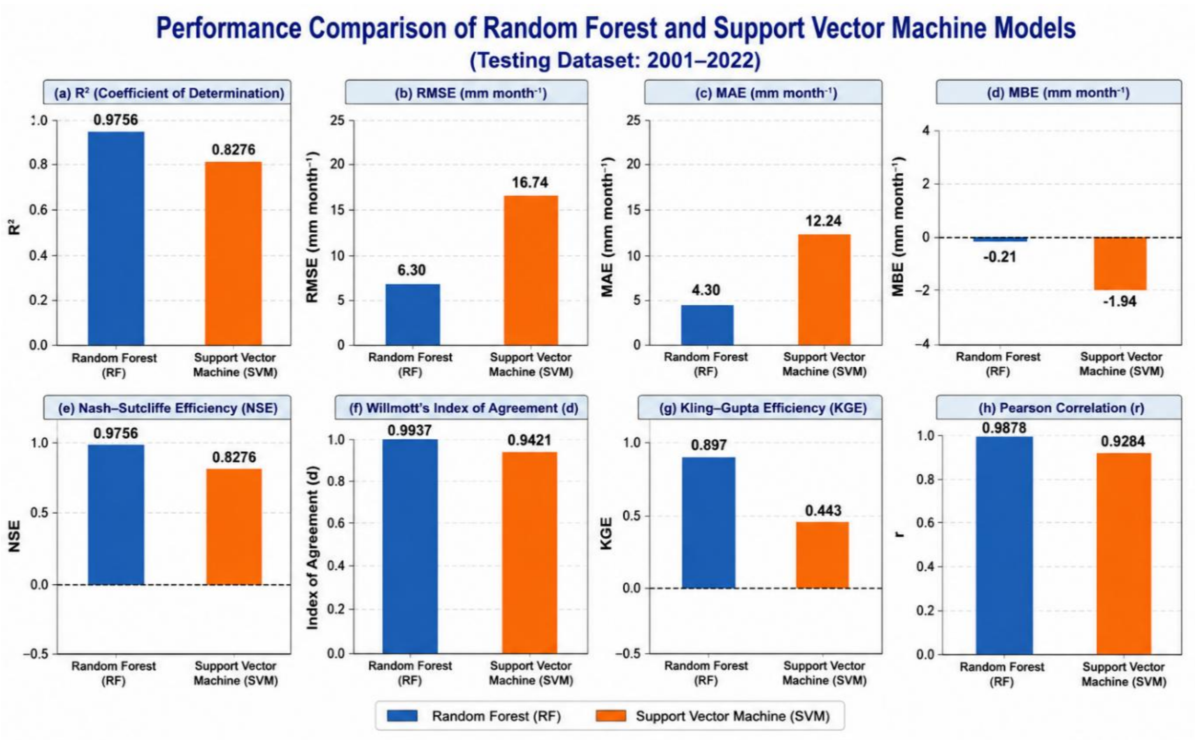


Fig. 12. Comparative performance of the RF and SVM models.

Time-Series Comparison of Observed and Predicted Potential Evapotranspiration

Fig. 13 shows that both RF and SVM reproduced the seasonal variability of monthly PET, with higher values during the pre-monsoon period and lower values during winter. However, RF showed a closer agreement with the observed PET throughout the study period, capturing both seasonal peaks and troughs with relatively small deviations. In contrast, SVM followed the general seasonal pattern but showed greater deviations, particularly during periods of high PET.

The close temporal agreement of RF is consistent with its higher R^2 (0.976) and lower RMSE and MAE, indicating better predictive accuracy and stability. This superior performance may be attributed to the ability of the RF ensemble to capture nonlinear interactions among meteorological variables and represent seasonal and interannual PET variability more effectively (Djaman et al., 2018). Overall, the time-series comparison confirms RF as the more reliable model for monthly PET estimation and its potential application in water resources planning.

Taylor Diagram Analysis

Fig. 14 further evaluates model performance using a Taylor diagram, which simultaneously represents the correlation coefficient, standard deviation, and centered RMSE, providing a comprehensive comparison of model agreement with observed PET (Gupta et al., 2009). RF was positioned closer to the reference point than SVM, indicating higher correlation, lower centered RMSE, and a standard deviation more consistent with the observed series. Although SVM showed satisfactory performance, its greater distance from the reference point indicated comparatively lower accuracy.

The Taylor diagram therefore supports the results of the scatter plots, residual analysis, and time-series comparison, confirming the overall superior performance and reliability of RF for monthly PET estimation.

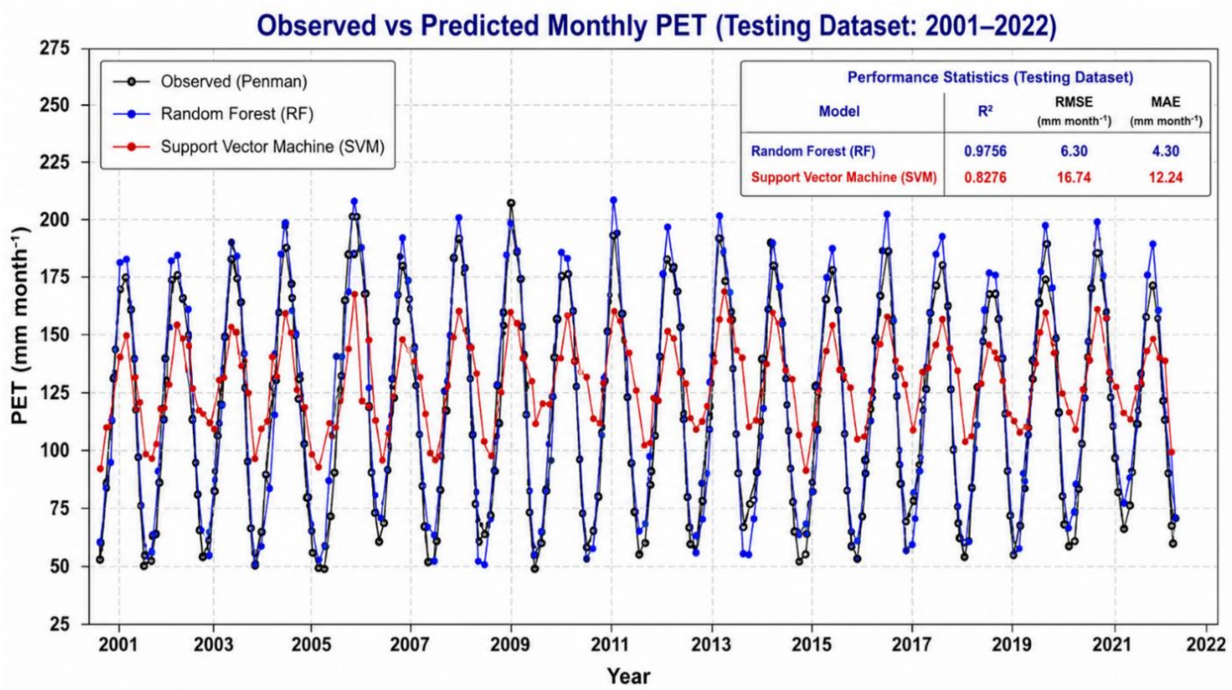


Fig.13. Time-series comparison of observed PET and values predicted by the RF and SVM models during the study period (2001–2022).

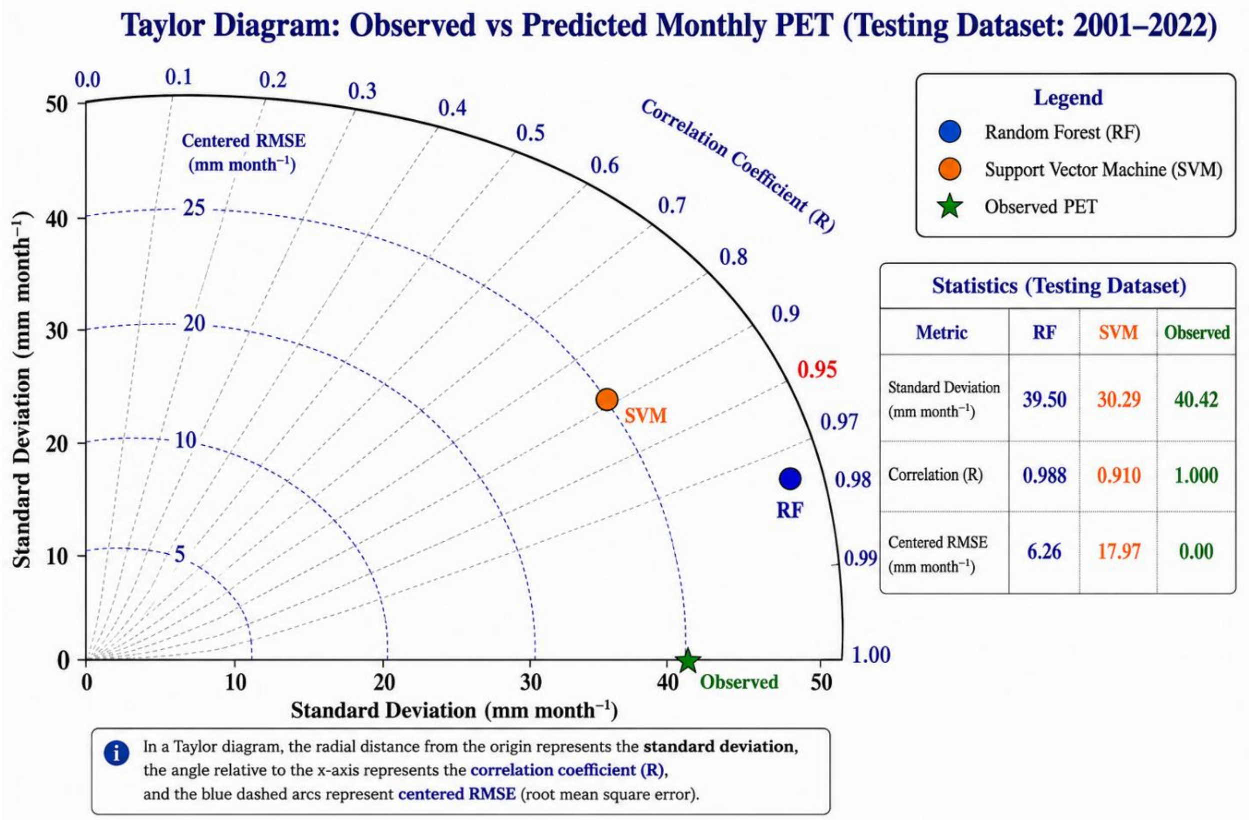


Fig. 14. Taylor diagram comparing the predictive performance of RF and SVM models against the observed PET.

Bland–Altman Analysis of Model Agreement

Fig. 15 evaluates the agreement between observed and predicted PET using Bland–Altman analysis, which assesses systematic bias and the 95% limits of agreement (Giavarina, 2015). RF showed a mean difference close to zero with relatively narrow limits of agreement and most observations within the 95% limits, indicating low prediction variability and good agreement with observed PET. In comparison, SVM exhibited wider limits and greater dispersion, particularly at higher PET values, indicating comparatively higher prediction uncertainty. These results are consistent with the lower RMSE, MAE, and MBE of RF and corroborate the findings from the residual analysis and Taylor diagram. Similar applications of comprehensive model-agreement assessment have been reported by Deb et al. (2025) further supporting the reliability of RF for monthly PET estimation under the study conditions.

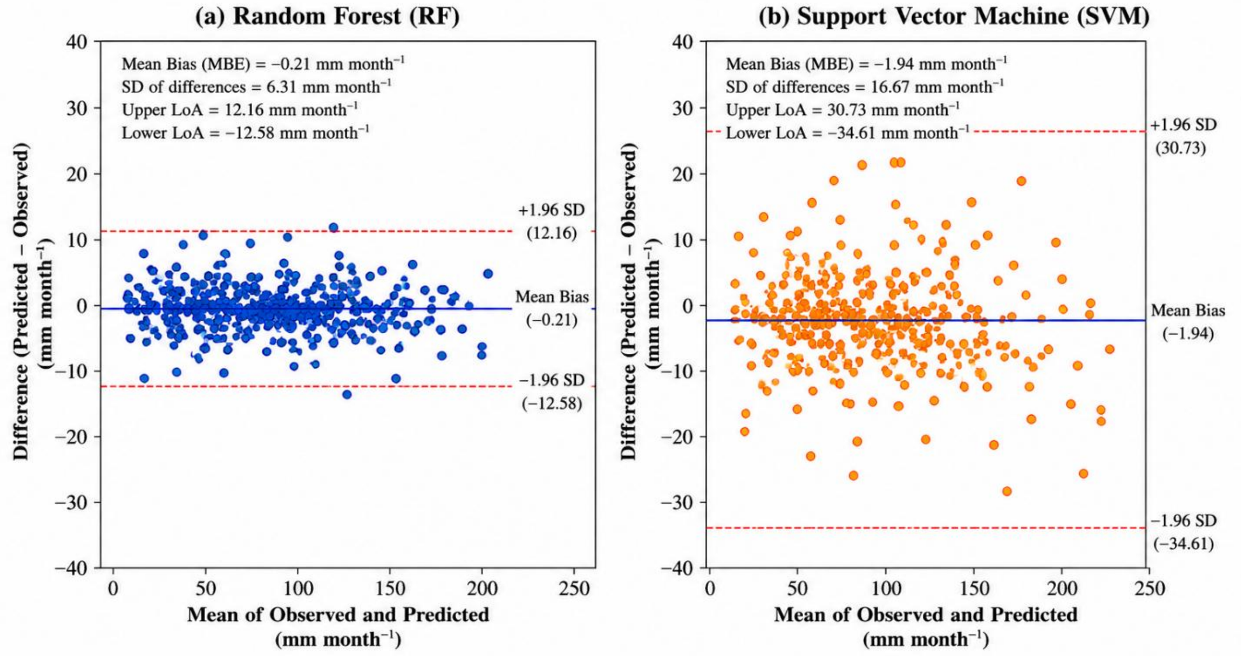


Fig. 15. Bland–Altman plots of observed and predicted PET for RF and SVM models. The solid horizontal lines represent the mean bias, while the dashed lines denote the 95% limits of agreement (mean bias ± 1.96 SD).

Conclusion

This study evaluated Random Forest (RF) and Support Vector Machine (SVM) models for monthly potential evapotranspiration (PET) estimation in Supaul District, Bihar, India, using NASA POWER-derived meteorological data from 2001–2022 and Penman-derived PET as the reference. Both models reproduced the seasonal variability of PET, but RF consistently outperformed SVM across the evaluated statistical and hydrological metrics. RF achieved $R^2 = 0.976$ compared with 0.828 for SVM, while RMSE was 6.30 mm month^{-1} for RF and 16.74 mm month^{-1} for SVM. RF achieved $R^2 = 0.976$ compared with 0.828 for SVM, while RMSE was 6.30 mm month^{-1} for RF and 16.74 mm month^{-1} for SVM. RF achieved $R^2 = 0.976$, RMSE = 6.30 mm month^{-1}, MAE = 4.30 mm month^{-1}, MBE = −0.21 mm month^{-1}, NSE = 0.976, KGE = 0.897, and Willmott's d = 0.994, indicating close agreement with the reference PET. Residual diagnostics, Q–Q plots, Taylor analysis, and Bland–Altman assessment further supported the comparatively better stability and agreement of RF. Overall, the findings indicate that RF provides an effective and computationally efficient approach for monthly PET estimation under the study conditions. The model has potential to support irrigation scheduling, agricultural water allocation, reservoir operation, and sustainable water-resources management, particularly where complete meteorological observations are limited.

References

Allen RG, et al. (1998) Crop evapotranspiration: guidelines for computing crop water requirements (FAO Irrigation and Drainage Paper No. 56). Food and Agriculture Organization of the United Nations.

Bhavya TR, et al. (2025) Estimating reference evapotranspiration using a machine learning approach. Water Science and Technology 92(6):819–842. DOI: https://doi.org/10.2166/wst.2025.078

Bland JM and Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet 327(8476):307–310.

Breiman L (2001) Random forests. Machine Learning 45(1):5–32.

Deb P, et al. (2025) Which machine learning algorithm is best suited for estimating reference evapotranspiration in humid subtropical climate? CLEAN – Soil, Air, Water 53(1):e202300441. DOI: https://doi.org/10.1002/clen.202300441

Djaman K, et al. (2018) Evaluation of evapotranspiration estimation methods under humid climatic conditions. Agricultural Water Management 193:134–145.

Elbeltagi A, et al. (2022) Machine learning approaches for reference evapotranspiration estimation: a comprehensive review. Agricultural Water Management 262:107407.

Fisher JB, et al. (2017) The future of evapotranspiration: global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resources Research 53:2618–2626.

Giavarina D (2015) Understanding Bland–Altman analysis. Biochemia Medica 25(2):141–151.

Gupta HV, et al. (2009) Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. Journal of Hydrology 377(1–2):80–91.

Kisi O (2015) Modeling reference evapotranspiration using machine learning techniques. Journal of Hydrology 528:312–320.

Martí P, Zarzo M and Ferrer J (2015) Evaluation of reference evapotranspiration estimation methods in Mediterranean climates. Agricultural Water Management 147:1–10.

Nash JE and Sutcliffe JV (1970) River flow forecasting through conceptual models. Part I: a discussion of principles. Journal of Hydrology 10(3):282–290.

Penman HL (1948) Natural evaporation from open water, bare soil and grass. Proceedings of the Royal Society A 193(1032):120–145.

Raj DK and Gopikrishnan T (2024) Soft computing techniques for predicting vegetation dynamics in Delhi. Journal of Earth System Science 133:241. DOI: https://doi.org/10.1007/s12040-024-02448-3

Sentelhas PC, Gillespie TJ and Santos EA (2010) Evaluation of methods for estimating reference evapotranspiration under humid conditions. Agricultural Water Management 98:155–162.

Smola AJ and Schölkopf B (2004) A tutorial on support vector regression. Statistics and Computing 14(3):199–222.

Sparks AH (2018) NASAPower: a NASA POWER global meteorology package for R. Journal of Open Source Software 3(30):1035.

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research: Atmospheres 106(D7):7183–7192.

Willmott CJ (1981) On the validation of models. Physical Geography 2(2):184–194.

Author Contributions
Vivek Kumar: Conceptualization, data curation, software, formal analysis. **Azmat Raja:** Conceptualization, methodology, supervision, validation, manuscript review and editing, interpretation of results, and project administration. **Nishikant Kumar:** Investigation, data collection, and manuscript review. **Raushan Anand:** Data curation, visualization, and manuscript editing. **Ashish Anand:** Methodology, validation, and technical review. **Raunak Kumar:** Investigation and data acquisition. **Pusplata Puja:** Data analysis, literature review, and manuscript editing. **Om Prakash Singh:** Supervision, technical guidance, and final manuscript review. **VK, AR, NK, RA, AA, RK, PP** and **OPS** conceived the concept, wrote and approved the manuscript.

Acknowledgements
Not applicable.

Funding
Not applicable.

Availability of data and materials
The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interest
The authors declare no competing interests.

Ethics approval
Not applicable.

Citation: Kumar V, Raja A, Kumar N, Anand R, Anand A, Kumar R, Puja P and Singh OP (2026) Machine Learning-Based Estimation of Monthly Potential Evapotranspiration for Sustainable Water Resources Management using Random Forest and Support Vector Machine. Environmental Science Archives 5(2): 466-478.