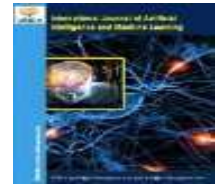


SvedbergOpen
DISSEMINATION OF KNOWLEDGEInternational Journal of Artificial
Intelligence and Machine LearningPublisher's Home Page: <https://www.svedbergopen.com/>

Research Paper

Open Access

Study of ET₀ by using Conventional and Soft Computing Techniques in the Eastern Gandak Project in Bihar, India – A Case Study

Abhinav Prakash Singh¹, L B Roy^{2*}¹Department of Civil Engineering, National Institute of Technology Patna, Patna, 800005, India. E-mail: Abhinavs.ph21.ce@nitp.ac.in,
Orcid Id: 0000-0001-5022-9469²Department of Civil Engineering, National Institute of Technology Patna, Patna, 800005, India. E-mail: lbroy@nitp.ac.in; Orcid : 0000-0003-4209-6007**ABSTRACT**

Reference evapotranspiration (ET₀) is a key parameter for irrigation planning, water resource management, and sustainable agricultural development. Because direct field measurement of evapotranspiration is difficult, time-consuming, and costly, ET₀ is commonly estimated using empirical and data-driven approaches based on meteorological variables. In this study, reference evapotranspiration was estimated for the Eastern Gandak Project region of Bihar, India, using both conventional empirical methods and soft computing techniques. Long-term climatic data recorded at the IARI regional station, Pusa, for the period 1981–2024 were used for analysis.

Four conventional methods—Hargreaves, Pan Evaporation, FAO-24 Modified Penman, and FAO-56 Penman–Monteith—were employed to compute ET₀, with the FAO-56 Penman–Monteith method adopted as the benchmark. In addition, three soft computing approaches, namely Artificial Neural Network (ANN), Support Vector Machine (SVM), and Gaussian Process Regression (GPR), were developed to estimate ET₀ using different combinations of meteorological inputs. Model performance was evaluated using the coefficient of determination (R²) for both training and testing datasets.

The results indicate that among the conventional methods, the FAO-24 Modified Penman method shows the strongest agreement with the FAO-56 Penman–Monteith method. Among the soft computing techniques, ANN demonstrates superior predictive capability on the testing dataset (R² up to 0.98) compared to SVM and GPR, indicating its robustness in capturing the nonlinear relationship between climatic variables and reference evapotranspiration. The findings highlight the potential of soft computing techniques, particularly ANN, as reliable alternatives for ET₀ estimation in data-rich and data-limited conditions.

Keywords: Reference Evapotranspiration; Penman–Monteith Method; Gaussian Process Regression; Artificial Neural Network; Coefficient of Determination.

1. Introduction

Rapid population growth, increasing food demand, and climate variability have intensified pressure on available freshwater resources, particularly in agrarian countries such as India. To meet rising agricultural water demands, extensive investments have been made in major and medium irrigation projects. However, several assessments have shown that the actual performance of irrigation systems often remains below their designed potential due to inefficiencies in storage, conveyance, distribution, and on-field application of water (Doorenbos and Pruitt, 1977; Pereira et al., 2015). These challenges highlight the importance of accurate estimation of crop water requirements for effective irrigation planning and sustainable agricultural management.

Crop water requirement is defined as the amount of water needed to support normal crop growth under prevailing agro-climatic conditions. Since the proportion of water used directly for plant metabolic processes is negligible compared to atmospheric losses, crop water requirement is commonly approximated by evapotranspiration. Direct measurement of evapotranspiration using lysimeters or detailed soil water balance techniques provides reliable results but is seldom feasible for routine applications due to high costs, operational complexity, and the need for skilled manpower (Allen et al., 1998). As a result, reference evapotranspiration (ET₀), which represents the evaporative demand of the atmosphere independent of crop type and management practices, has become a fundamental parameter in irrigation water management.

Reference evapotranspiration is governed primarily by climatic variables such as air temperature, humidity, solar radiation, and wind speed. Over the years, several empirical and semi-empirical methods have been developed to estimate ET₀ using different combinations of meteorological data. Methods recommended under FAO-24, including the Pan Evaporation, Blaney–Criddle, Radiation, and Modified Penman approaches, have been widely applied but often require regional calibration to achieve acceptable accuracy (Doorenbos and Pruitt, 1977). To overcome the limitations of regional dependency and improve consistency, the FAO-56 Penman–Monteith method has been proposed as a standardized approach and is currently regarded as the

most reliable method for ET_0 estimation under diverse climatic conditions (Allen et al., 1998; Pereira et al., 2015).

Despite its robustness, the FAO-56 Penman–Monteith method requires a complete set of meteorological inputs, which may not always be available, particularly in developing regions or historical datasets. This limitation has encouraged the development of alternative approaches capable of estimating ET_0 with limited climatic information. In this context, soft computing and machine learning techniques have gained increasing attention due to their ability to capture complex nonlinear relationships between climatic variables and evapotranspiration without explicit physical assumptions (Firat, 2008; Mehdizadeh et al., 2017).

Recent studies have also shown that deep learning–based models can further improve ET_0 estimation accuracy (Islam et al., 2021). Several comparative studies across different climatic regions have reported improved ET_0 estimation using ANN and SVM approaches (Traore et al., 2010; Tabari et al., 2012), with further advances achieved through hybrid frameworks (Gocic & Trajkovic, 2015).

Among machine learning approaches, Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Gaussian Process Regression (GPR) have been widely applied for ET_0 estimation across different climatic regions. ANN-based models have shown strong predictive capability in modeling evapotranspiration processes, often outperforming conventional empirical methods (Kumar et al., 2002; Kisi, 2009; Hassan Afzaal et al., 2020). SVM has been reported as a robust technique for regression problems involving limited data and high-dimensional feature spaces (Vapnik et al., 1997; Lu and Wang, 2005), while GPR offers the additional advantage of probabilistic prediction and uncertainty quantification, which is valuable in hydrological modeling (Wang et al., 2009; Abdullah et al., 2015).

Recent studies have further demonstrated the effectiveness of hybrid and data-driven models for ET_0 estimation under varying climatic conditions, particularly in South and Southeast Asia (Islam et al., 2019; Salam et al., 2020; Jerin et al., 2021). However, the performance of these models is strongly influenced by regional climate characteristics, input data availability, and model structure. Comparative evaluations of multiple soft computing techniques using long-term meteorological datasets remain limited for the eastern Indo-Gangetic plains, including Bihar, where agriculture is highly dependent on irrigation from canal systems.

The present study aims to estimate reference evapotranspiration for the Eastern Gandak Project region of Bihar using both conventional empirical methods and soft computing techniques. Four conventional methods—Hargreaves, Pan Evaporation, FAO-24 Modified Penman, and FAO-56 Penman–Monteith—are evaluated, with the FAO-56 Penman–Monteith method adopted as the reference standard. In addition, ANN, SVM, and GPR models are developed using different combinations of meteorological inputs to assess their suitability for ET_0 estimation under varying data availability conditions. The outcomes of this study are expected to support improved irrigation scheduling and sustainable water resource management in canal-command agricultural regions.

2. The Study Area

The Eastern Gandak Project plays a critical role in canal-based irrigation in north Bihar and is highly sensitive to seasonal variability in evapotranspiration demand (Pereira et al., 2015). Bhagwanpur Distributary is the tail end distributary of the Eastern Gandak Project which takes off from the Vaishali Branch Canal. The command area lies in Muzaffarpur and Vaishali districts of the Indian state of Bihar between longitude 85°7'30" E and 85°15'0" E and latitude 25°52'30" N and 26°3'0" N. The Bhagwanpur distributaries command area is covered by Saraiya block in Muzaffarpur district and Vaishali block in Vaishali district. It is ridge canal and runs through 9 and 17 villages of Muzaffarpur and Vaishali districts respectively. The Bhagwanpur Distributary has a gross command area of 2250 hectares and a total cultivable command area of 1841 hectares. The location map of the command area of the Bhagwanpur distributary is shown in Fig-1.

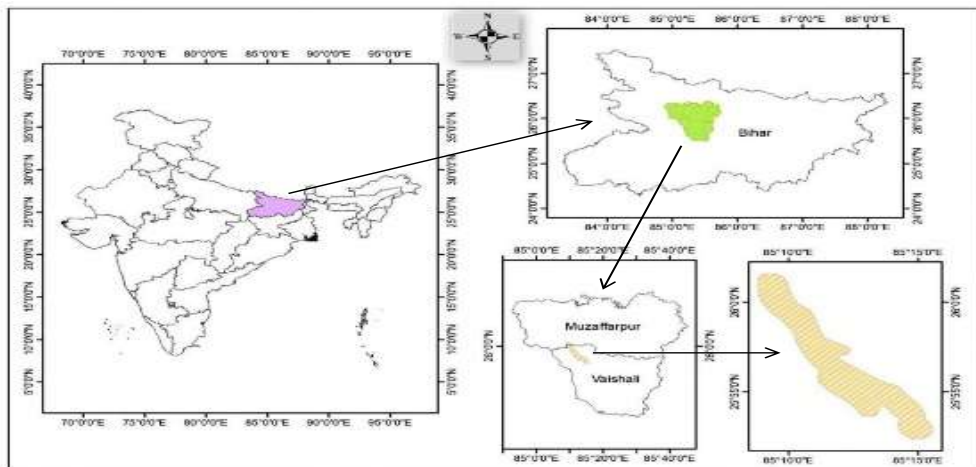


Figure 1. Bhagwanpur distributary

3. Methodology

Estimation of ET₀ Using Conventional Methods

Reference evapotranspiration (ET₀) for the IARI regional station at Pusa (Samastipur, Bihar) was estimated using four widely adopted conventional methods: Hargreaves method, Pan Evaporation method, FAO-24 Modified Penman method, and FAO-56 Penman–Monteith method. Monthly ET₀ values were computed for the period 1981–2024, except for the Pan Evaporation method, which was applied only for 2000–2024 due to the unavailability of earlier pan evaporation data. These methods were selected to represent temperature-based, radiation-based, and combination approaches with varying data requirements.

3.1. Pan Evaporation Method

The Pan Evaporation method estimates reference evapotranspiration based on measured evaporation from a standardized open water surface, typically a Class A evaporation pan. Since evaporation from a pan differs from evapotranspiration from a cropped surface, a pan coefficient is applied to account for differences in heat storage, aerodynamic conditions, and surface characteristics. Despite its simplicity, this method is sensitive to local environmental factors such as wind speed, humidity, and pan surroundings, which may introduce uncertainty if not properly calibrated (Doorenbos & Pruitt, 1977; Pereira et al., 2015).

The reference evapotranspiration is estimated using:

$$ET_0 = K_p \times E_{pan}$$

where K_p is the pan coefficient and E_{pan} is the measured pan evaporation (mm day⁻¹).

Although the Pan Evaporation method is easy to implement, several studies have reported its tendency to underestimate or overestimate ET₀ under varying climatic conditions, particularly during humid or windy periods (Allen et al., 1998; Islam et al., 2019).

3.2. Modified Penman Method

The FAO-24 Modified Penman method is a combination approach that integrates both energy balance and aerodynamic components to estimate reference evapotranspiration. It accounts for net radiation, wind speed, temperature, and vapor pressure deficit, thereby offering improved physical representation compared to purely empirical methods. However, this method has been reported to overestimate ET₀ in certain climatic regions, especially under humid conditions, unless region-specific adjustments are applied (Doorenbos & Pruitt, 1977; Pereira et al., 2015).

The FAO-24 Modified Penman equation is expressed as:

$$ET_0 = (WR_n + (1 - W)f(u)(e_s - e_a)) \times c$$

where W is the weighting factor, R_n is the net radiation expressed in equivalent evapotranspiration (mm day⁻¹), $f(u)$ is the wind function, e_s represents the vapor pressure deficit (mbar), and c is an adjustment factor accounting for day and night conditions. Due to its relatively higher data demand and sensitivity to climatic variability, the Modified Penman method is often used for comparison rather than routine application in operational irrigation scheduling.

3.3. Hargreaves Method

The Hargreaves method is a temperature-based empirical approach that estimates reference evapotranspiration using air temperature and extraterrestrial radiation. Its simplicity and minimal data requirement make it particularly useful in regions where comprehensive meteorological data are unavailable. The method has been shown to perform reasonably well in arid and semi-arid climates, although its accuracy may decline in humid regions or areas with strong wind effects (Hargreaves & Samani, 1985; Khoob, 2007).

The Hargreaves equation is given by:

$$ET_0 = 0.0023(T + 17.8)R_a(T_{max} - T_{min})^{0.5}$$

where T is the mean air temperature (°C), R_a is the extraterrestrial radiation (mm day⁻¹), and $(T_{max} - T_{min})$ represents the temperature range (°C). Although the Hargreaves method may overestimate ET₀ during hot periods, it remains a practical alternative when radiation, humidity, or wind data are missing (Valiantzas, 2013).

3.4. Penman – Monteith Method

The FAO-56 Penman–Monteith method is widely recognized as the standard approach for estimating reference evapotranspiration and serves as a benchmark for evaluating other ET₀ estimation methods. This method provides a physically based representation of evapotranspiration by combining energy balance and aerodynamic processes and is applicable across a wide range of climatic conditions without the need for local calibration (Allen et al., 1998; Pereira et al., 2015).

The FAO-56 Penman–Monteith equation is expressed as:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)}$$

where ET₀ is the reference evapotranspiration (mm day⁻¹), R_n is the net radiation (MJ m⁻² day⁻¹), G is the soil heat flux density (MJ m⁻² day⁻¹), T is the mean air temperature (°C), u_2 is the wind speed at 2 m height (m s⁻¹), e_s and e_a are the saturation and actual vapor pressures (kPa), Δ is the slope of the saturation vapor pressure

curve ($\text{kPa } ^\circ\text{C}^{-1}$), and γ is the psychrometric constant ($\text{kPa } ^\circ\text{C}^{-1}$).

In the present study, the FAO-56 Penman–Monteith method was adopted as the reference standard. ET_0 values estimated using other conventional methods were compared against Penman–Monteith estimates, and their performance was evaluated using the coefficient of determination (R^2), which quantifies the strength of linear agreement between methods

Estimation of ET_0 using Soft Computing Techniques

Soft computing techniques have gained wide acceptance in hydrological and agricultural modeling due to their ability to capture nonlinear relationships between climatic variables and hydrological responses without requiring explicit physical assumptions. In evapotranspiration modeling, these techniques are particularly useful when meteorological inputs are limited, noisy, or incomplete. In the present study, three soft computing approaches—Support Vector Machine (SVM), Gaussian Process Regression (GPR), and Artificial Neural Network (ANN)—were employed to estimate reference evapotranspiration (ET_0) using long-term climatic data from the IARI regional station, Pusa, Bihar.

Monthly meteorological data for the period 1981–2024 were used to develop the models. The dataset was divided into two subsets, with 75% of the data used for model training and the remaining 25% reserved for testing and validation. All input variables were normalized prior to model training to ensure numerical stability and improved convergence. Different combinations of input variables were considered to evaluate model performance under varying data availability conditions.

3.5. Support Vector Machine

Support Vector Machine is a supervised machine learning technique grounded in statistical learning theory and is widely used for regression and classification problems. In regression applications, known as Support Vector Regression (SVR), the objective is to identify a function that deviates from observed values by no more than a specified margin while maintaining maximum flatness. This principle enables SVM to generalize well even when training data are limited or high-dimensional (Vapnik et al., 1997; Vapnik, 1999).

SVM performs nonlinear regression by mapping input variables into a high-dimensional feature space through kernel functions, where linear regression is carried out. Commonly used kernels include linear, polynomial, radial basis function (RBF), and sigmoid kernels. The choice of kernel and associated hyperparameters plays a critical role in model accuracy. In evapotranspiration studies, SVM has demonstrated strong robustness and competitive performance compared to empirical and neural network models (Lu & Wang, 2005; Mehdizadeh et al., 2017).

In the present study, ϵ -insensitive SVM regression was implemented using MATLAB (version 9.4). Meteorological variables such as temperature, relative humidity, wind speed, and sunshine duration were used as inputs, while ET_0 estimated by the FAO-56 Penman–Monteith method was used as the target output. Model performance was evaluated using the coefficient of determination (R^2) for both training and testing datasets.

3.6. Gaussian Process Regression

Gaussian Process Regression is a non-parametric, probabilistic machine learning approach that defines a distribution over possible functions that fit the data. Unlike deterministic regression methods, GPR provides both mean predictions and associated uncertainty, making it particularly valuable for hydrological applications where data uncertainty is significant. GPR models are governed by covariance (kernel) functions that control smoothness, variability, and correlation structure of the predicted function (Wang et al., 2009; Abdullah et al., 2015).

One of the main advantages of GPR is its ability to automatically learn hyperparameters from the training data by maximizing the marginal likelihood. Commonly used kernel functions include Squared Exponential, Rational Quadratic, Matérn, and Exponential kernels. Previous studies have reported that GPR performs well in evapotranspiration estimation, especially when sufficient historical data are available and uncertainty quantification is required (Abdullah et al., 2015; Feng et al., 2017).

In this study, GPR models were developed using the MATLAB Machine Learning Toolbox. Multiple kernel functions were tested, and the optimal model was selected based on performance during the validation phase. Similar to the SVM approach, FAO-56 Penman–Monteith ET_0 values were used as reference outputs for training and testing.

3.7. Artificial Neural Network

Artificial Neural Networks are data-driven computational models inspired by the structure and functioning of biological neural systems. ANNs consist of interconnected processing units (neurons) organized into an input layer, one or more hidden layers, and an output layer. Through a learning process, the network adjusts connection weights to minimize the difference between predicted and observed outputs. This adaptive learning capability enables ANNs to effectively model complex nonlinear relationships between climatic variables and evapotranspiration (Kumar et al., 2002; Kisi, 2009).

ANNs have been extensively applied for ET_0 estimation across different climatic regions and have often been found to outperform conventional empirical methods and other machine learning approaches when adequate training data are available (Khoob, 2008; Adamala et al., 2013; Hassan Afzaal et al., 2020). Variants such as backpropagation neural networks, genetic algorithm–optimized ANNs, extreme learning machines, and wavelet neural networks have further enhanced prediction accuracy in recent studies (Feng et al., 2016).

In the present study, a feed-forward ANN architecture was employed using the Levenberg–Marquardt

backpropagation algorithm for network training. A network with ten hidden neurons was selected based on trial-and-error optimization to balance accuracy and computational efficiency. The same sets of meteorological input combinations used for SVM and GPR models were applied to the ANN framework. Model performance was assessed using the coefficient of determination (R^2) for both training and testing datasets.

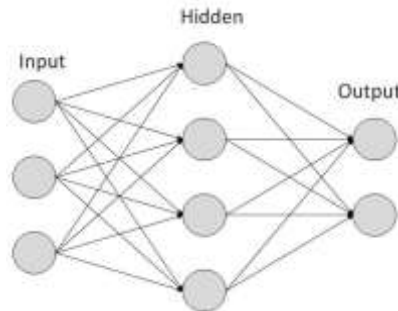


Figure 2. A Representation of ANN

3.8. Input Scenarios and Model Evaluation

To examine the effect of data availability on ET_0 estimation, four different input scenarios were considered, ranging from complete meteorological datasets to temperature-only inputs. This approach enables evaluation of model robustness under both data-rich and data-scarce conditions. The ET_0 values estimated using soft computing techniques were compared with those obtained from the FAO-56 Penman–Monteith method, and model accuracy was quantified using the coefficient of determination (R^2).

Table 1. Different Cases of Input Dataset Considered for Estimation of ET_0 Using Soft Computing Techniques

Case	Inputs	Output
1	T_{max} , T_{min} , RH_{max} , RH_{min} , Wind Speed, Sunshine Hours	ET_0
2	T_{max} , T_{min} , RH_{max} , RH_{min} , Wind Speed	ET_0
3	T_{max} , T_{min} , RH_{max} , RH_{min}	ET_0
4	T_{max} , T_{min}	ET_0

4. Result And Discussion

Monthly meteorological data for the period 1981–2024, including minimum and maximum temperature, relative humidity, wind speed, sunshine duration, and rainfall, were used to estimate reference evapotranspiration (ET_0) at the IARI regional station, Pusa (Samastipur, Bihar). The performance of both conventional and soft computing methods was evaluated using the FAO-56 Penman–Monteith method as the reference standard. Sample climatological data and ET_0 estimates for the year 2016 are presented in Tables 2 and 3, respectively.

Table2 Monthly Climatological Data for Pusa Station (2016)

Year	Month	Min Temp. (oC)	Max Temp. (oC)	Min Humidity (%)	Max Humidity (%)	Wind Speed (Km/day)	Sunshine Hours	Rainfall (mm)
2016	Jan	8.00	22.00	89.20	57.60	1.90	2.80	2.20
2016	Feb	12.10	26.70	87.90	52.30	2.90	5.70	2.80
2016	Mar	17.00	32.60	82.80	40.60	3.50	7.90	5.60
2016	Apr	21.70	38.30	70.50	29.50	5.70	8.10	3.20
2016	May	23.80	34.60	83.40	53.00	7.50	8.10	126.80
2016	Jun	26.30	35.20	83.80	62.90	7.30	7.00	105.10
2016	Jul	26.00	32.00	91.00	79.00	7.00	3.50	304.10
2016	Aug	26.20	33.70	86.20	68.70	6.70	7.00	110.80
2016	Sep	25.00	31.40	92.80	80.00	5.10	3.60	319.20
2016	Oct	24.00	32.10	90.60	69.00	3.70	5.00	34.60
2016	Nov	12.50	29.00	86.30	44.20	1.70	6.20	0.00

2016	Dec	11.50	21.20	90.30	64.40	2.70	1.80	0.00
------	-----	-------	-------	-------	-------	------	------	------

Table 3 Reference Evapotranspiration at IARI, Pusa Using Different Methods

Year	Month	Hargreaves Method	Pan Evaporation Method	Modified Penman Method	FAO 56 Penman Monteith Method
2016	January	2.65	0.7	1.71	1.31
2016	February	3.84	1.67	2.62	1.95
2016	March	5.21	2.85	3.88	3.12
2016	April	7.02	5.34	5.32	4.07
2016	May	5.82	4.1	4.78	4.56
2016	June	5.84	3.54	4.68	4.51
2016	July	4.38	2.3	3.47	3.4
2016	August	4.7	2.84	4.28	4.27
2016	September	3.95	2.05	2.99	2.96
2016	October	3.95	2.01	3.36	2.88
2016	November	3.69	1.49	2.51	2.01
2016	December	2.16	0.57	1.41	1.44

Comparison of Conventional Methods

Figure 3 illustrates the variation of mean monthly ET_0 estimated using the Hargreaves, Pan Evaporation, FAO-24 Modified Penman, and FAO-56 Penman–Monteith methods. The results indicate that the Hargreaves and Modified Penman methods tend to overestimate ET_0 compared to the Penman–Monteith method, particularly during the summer months. This overestimation can be attributed to the temperature-dominant nature of the Hargreaves method and the sensitivity of the Modified Penman method to radiation and wind speed inputs, which may amplify evaporative demand under hot climatic conditions.

The Pan Evaporation method shows reasonable agreement with the Penman–Monteith estimates during the pre-monsoon period (February to May) but underestimates ET_0 during monsoon and winter months. Similar seasonal biases have been reported in previous studies, where pan-based estimates were found to be strongly influenced by local pan exposure, wind conditions, and humidity variations (Allen et al., 1998; Islam et al., 2019).

Regression analysis between ET_0 values estimated by conventional methods and those obtained using the Penman–Monteith method is presented in Figures 4–6. The FAO-24 Modified Penman method exhibits the strongest correlation with the Penman–Monteith method, with a coefficient of determination (R^2) of 0.87. In contrast, the Pan Evaporation method shows the weakest correlation ($R^2 = 0.62$), indicating limited reliability without site-specific calibration. These findings are consistent with earlier comparative studies that identified Modified Penman as a relatively robust alternative among empirical approaches when sufficient climatic data are available (Pereira et al., 2015; Valiantzas, 2013).

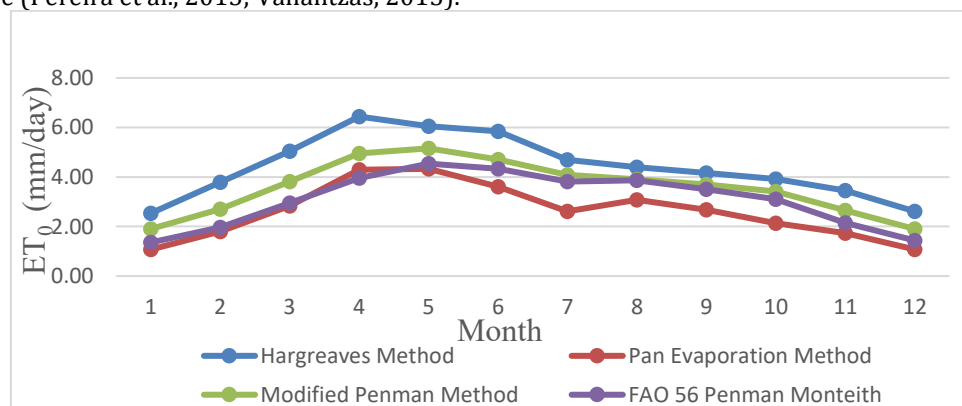


Figure 3 Comparisons between ET_0 Estimated Using Different Methods
Correlation between Penman Monteith Method and Other Methods

Figures 4 to 6 show the relationship between ET₀ calculated using the Penman Monteith method and ET₀ calculated using other methods.

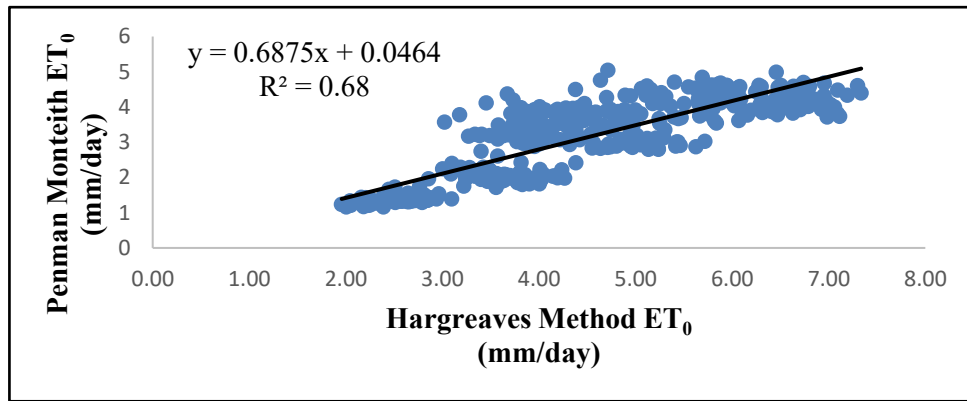


Figure 4 Correlation between Hargreaves Method ET₀ and FAO 56 Penman Monteith ET₀

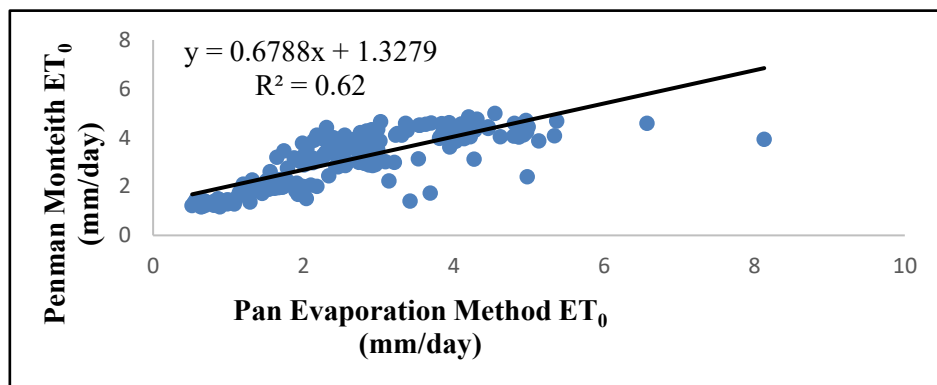


Figure 5 Correlation between Pan Evaporation Method ET₀ and FAO 56 Penman Monteith ET₀

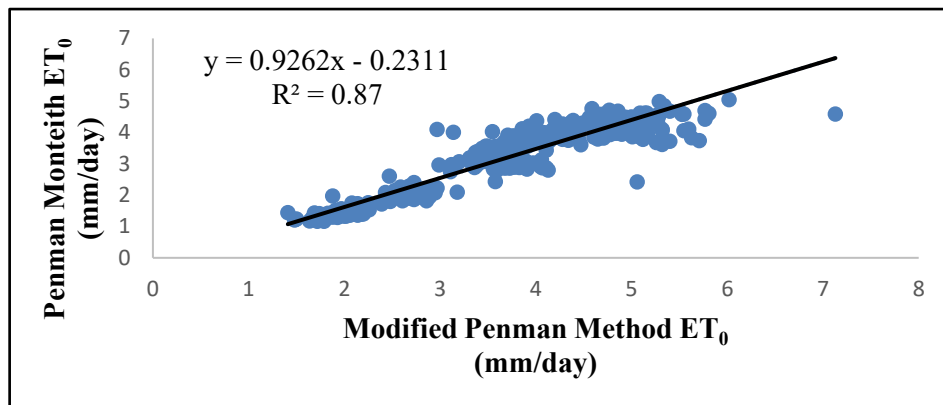


Figure 6 Correlations between Modified Penman Method ET₀ and FAO 56 Penman Monteith ET₀

Equations 7, 8, and 9 show the correlation equations for the Hargreaves technique vs. Penman Monteith method, Pan Evaporation method vs. Penman Monteith method, and Modified Penman method vs. Penman Monteith method.

$$y = 0.6875x + 0.0464; R^2 = 0.68 \quad (6)$$

$$y = 0.6788x + 1.3279; R^2 = 0.62 \quad (7)$$

$$y = 0.9262x - 0.2311; R^2 = 0.87 \quad (8)$$

$$y = \text{EToPM}, x = \text{EToH}$$

Figure 4, 5, and 6 shows that the Modified Penman Method has the best correlation with the Penman Monteith Method, with an R² of 0.87, while the Pan Evaporation Method has the weakest correlation with the Penman Monteith Method, with an R² of 0.62.

4.2. Estimation of ET₀ using Soft Computing Techniques

The performance of the soft computing models under different input scenarios is summarized in Tables 3 and 4 for training and testing datasets, respectively. Overall, all three techniques—SVM, GPR, and ANN—demonstrate strong predictive capability, with R² values exceeding 0.90 in most cases, highlighting their

suitability for ET_0 estimation.

For the training dataset, GPR yields the highest R^2 values, reaching 0.99 for input Cases 1 and 2. This superior performance may be attributed to the probabilistic nature of GPR and its ability to adaptively learn kernel hyperparameters from the data. However, during the testing phase, ANN consistently outperforms both SVM and GPR across all input scenarios. The ANN achieves the highest R^2 value of 0.98 for Case 1 and maintains stable performance even when the number of input variables is reduced.

The improved generalization capability of ANN in the testing phase suggests that neural networks are particularly effective in capturing the nonlinear and dynamic relationships between climatic variables and reference evapotranspiration. Similar observations have been reported by Kumar et al. (2002), Kisi (2009), and Hassan Afzaal et al. (2020), who found ANN models to outperform other machine learning techniques in ET_0 prediction under diverse climatic conditions.

SVM models show competitive performance, especially when complete meteorological inputs are available, but their accuracy declines slightly under reduced input scenarios. This behavior reflects the sensitivity of SVM to kernel selection and parameter tuning, as noted in previous hydrological modeling studies (Mehdizadeh et al., 2017). GPR models maintain relatively stable performance across all cases, demonstrating robustness under varying data availability, although their computational complexity increases with larger datasets.

Figures 7–24 illustrate the correlation between ET_0 values estimated using soft computing techniques and those obtained from the Penman–Monteith method for different cases and datasets. The clustering of data points around the 1:1 line further confirms the high predictive accuracy of the models, particularly for ANN-based estimates.

Table 3 Performance of Soft Computing Techniques in Training

Method	Coefficient of Determination (R^2)			
	Case 1	Case 2	Case 3	Case 4
SVM	0.98	0.96	0.93	0.91
GPR	0.99	0.99	0.97	0.95
ANN	0.98	0.96	0.91	0.92

Table 4 Performance of Soft Computing Techniques in Testing

Method	Coefficient of Determination (R^2)			
	Case 1	Case 2	Case 3	Case 4
SVM	0.97	0.92	0.95	0.93
GPR	0.97	0.92	0.94	0.94
ANN	0.98	0.96	0.95	0.95

Tables 3 and 4 shows that GPR outperforms all other methods for training data sets, with R^2 values of 0.99 for Cases 1 and 2, 0.97 and 0.95 for Cases 3 and 4, respectively. However, ANN performs best on the testing data set, with R^2 equal to 0.98 for Case 1, 0.96 for Case 2, and 0.95 for Cases 3 and 4.

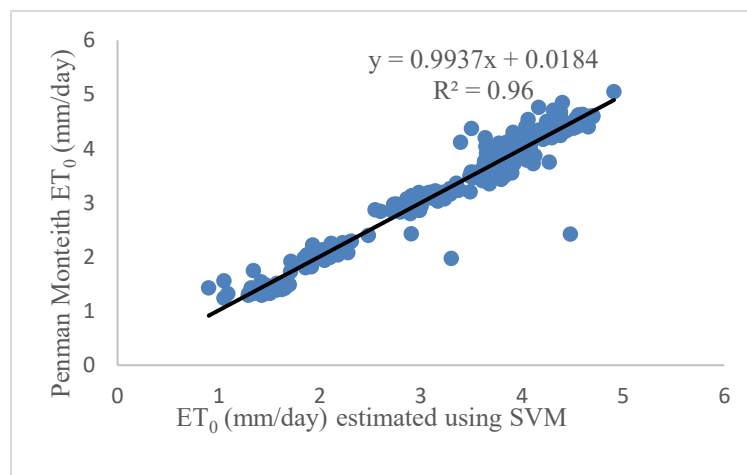


Figure 7 Performance of SVM in Training for Case 2

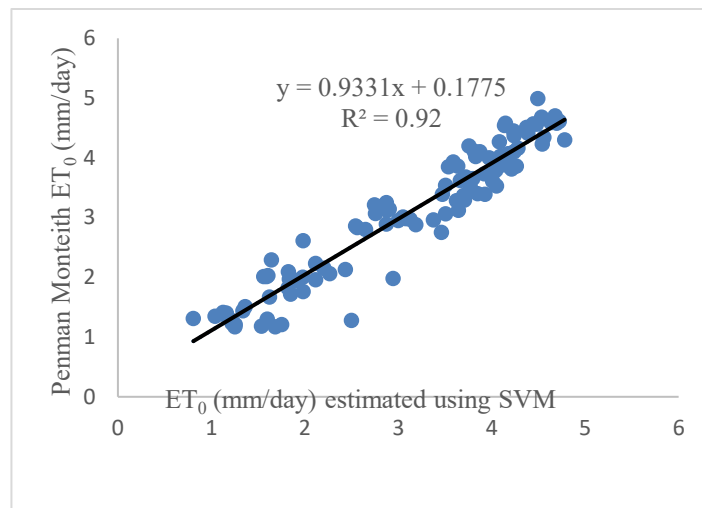


Figure 8 Performance of SVM in Testing for Case 2

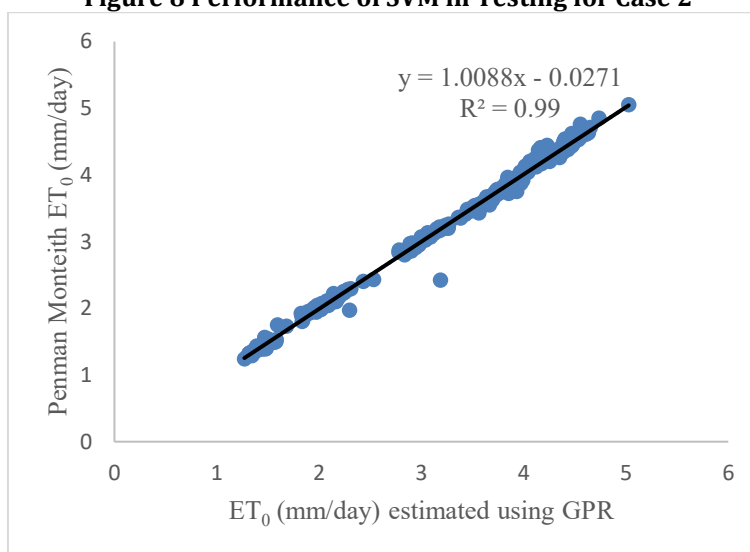


Figure 9 Performance of GPR in Training for Case 2

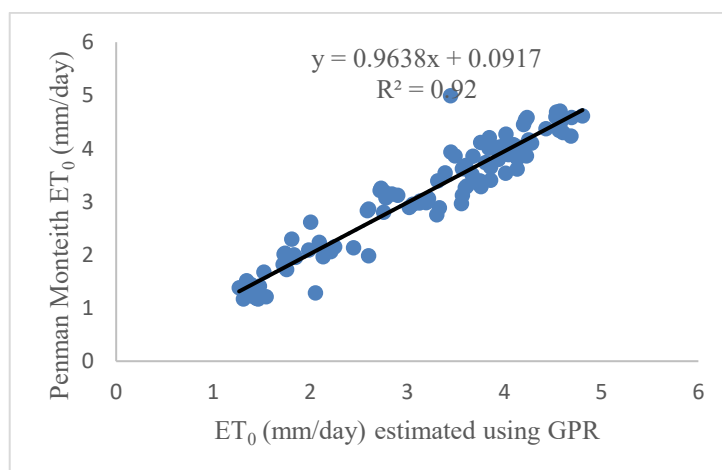


Figure 10 Performance of GPR in Testing for Case 2

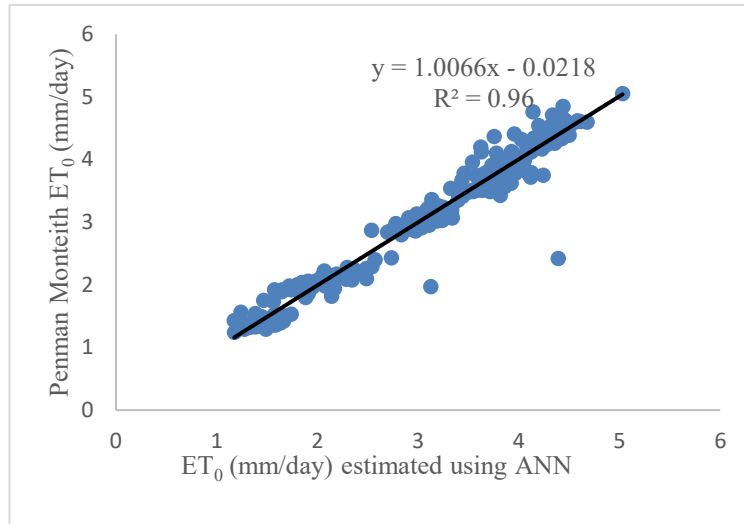


Figure 11 Performance of ANN in Training for Case 2

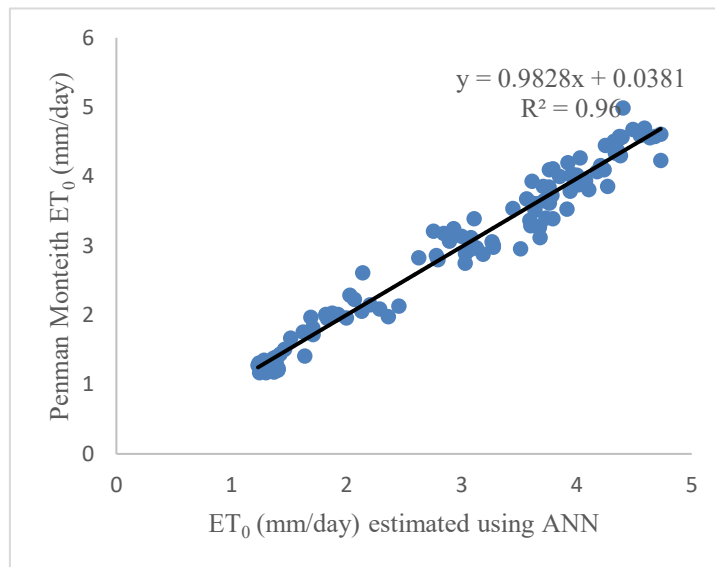


Figure 12 Performance of ANN in Testing for Case 2

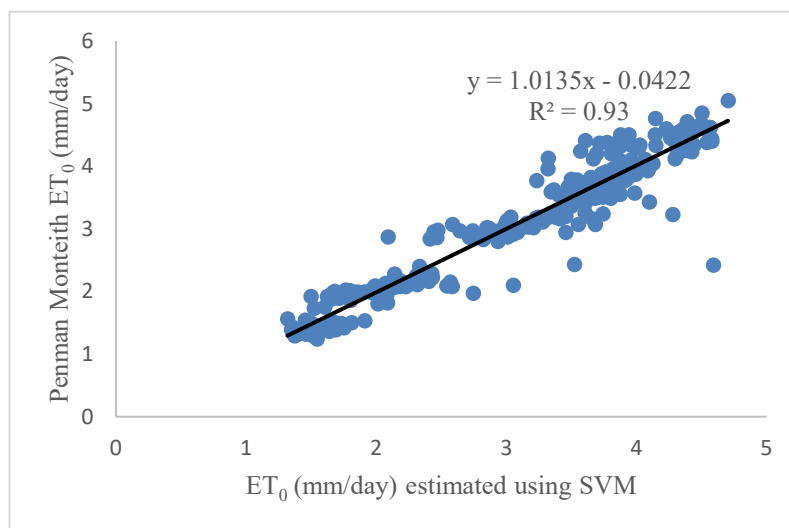


Figure 13 Performance of SVM in Training for Case 3

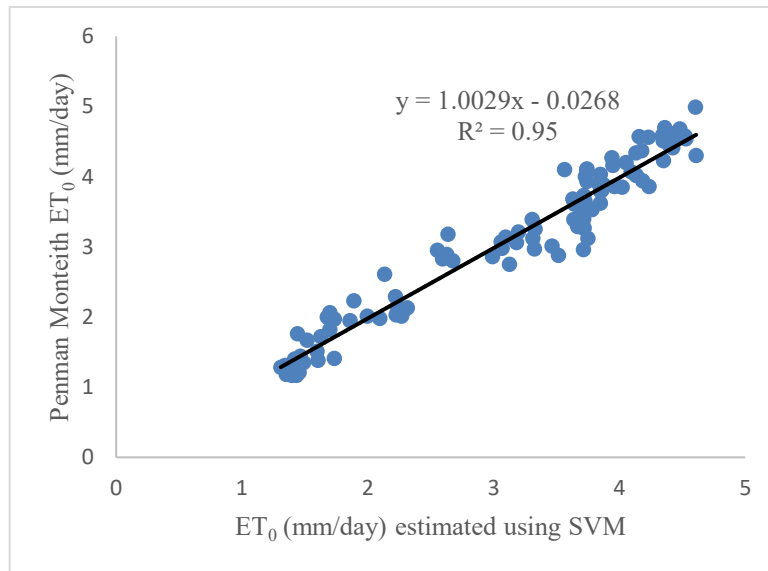


Figure 14 Performance of SVM in Testing for Case 3

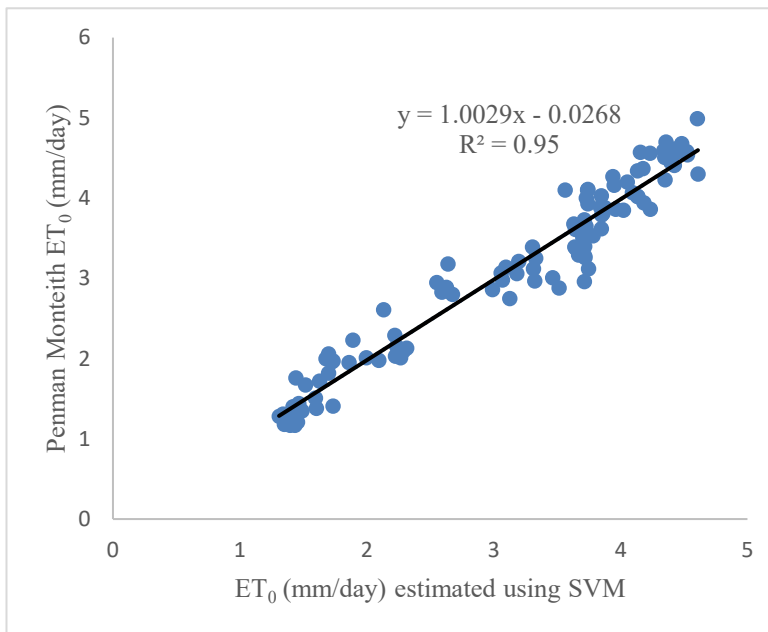


Figure 15 Performance of GPR in Training for Case 3

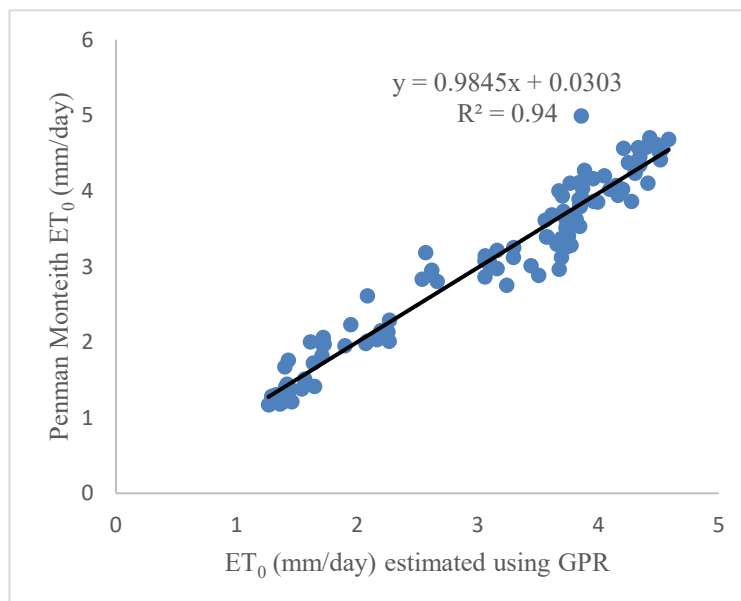


Figure 16 Performance of GPR in Testing for Case 3

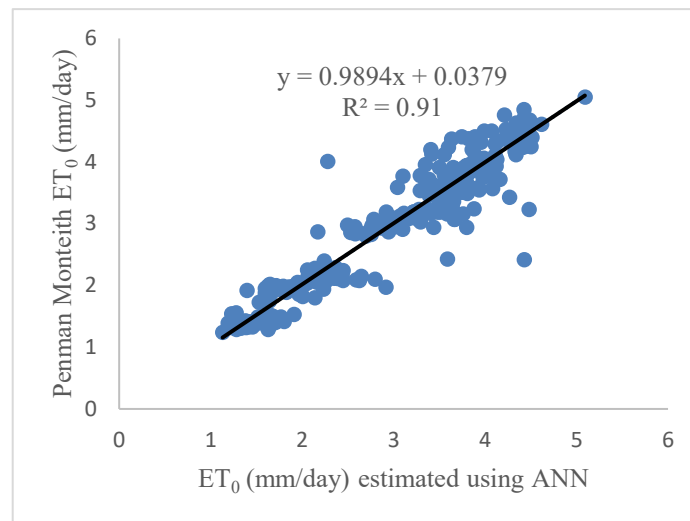


Figure 17 Performance of ANN in Training for Case 3

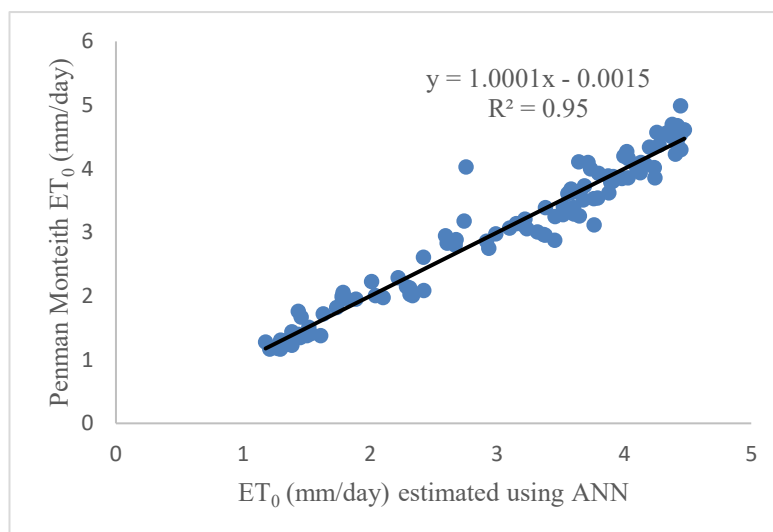


Figure 18 Performance of ANN in Testing for Case 3

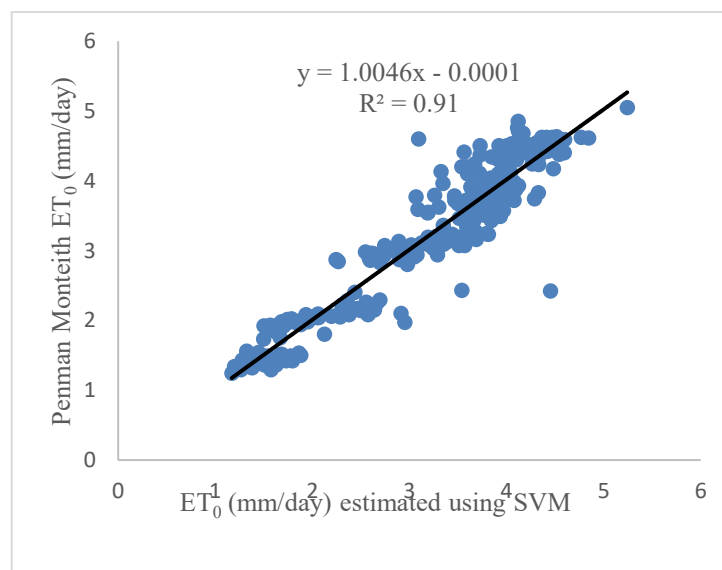


Figure 19 Performance of SVM in Training for Case 4

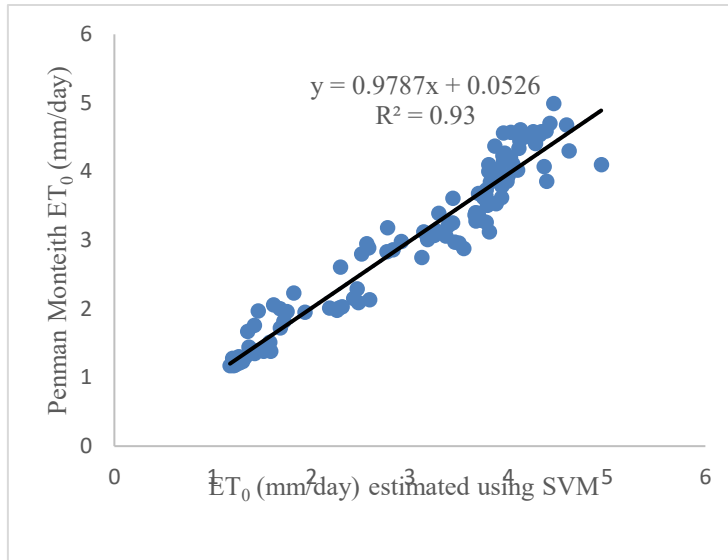


Figure 20 Performance of SVM in Testing for Case 4

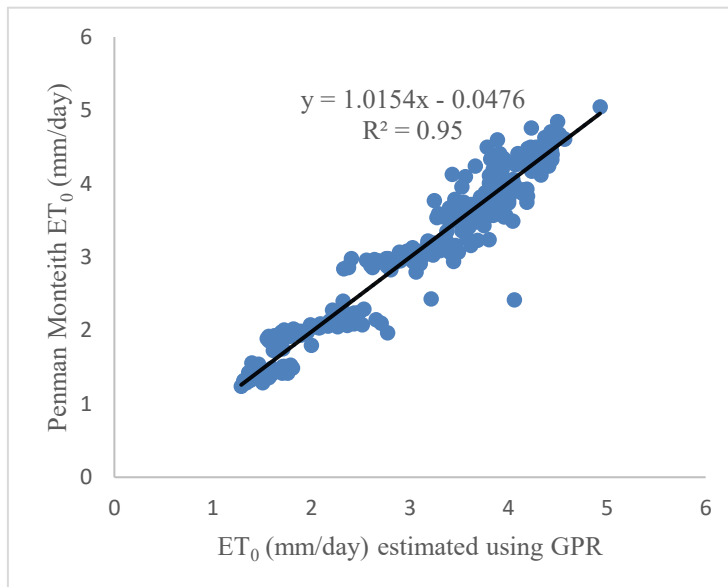


Figure 21 Performance of GPR in Training for Case 4

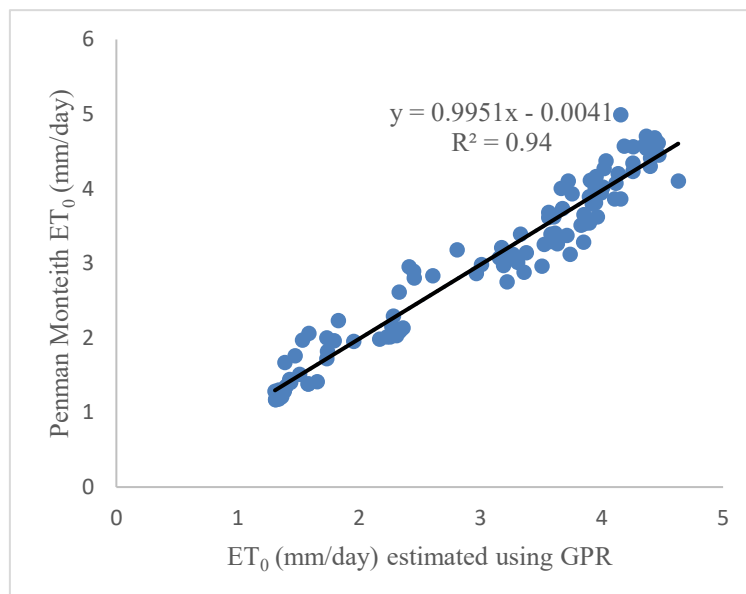


Figure 22 Performance of GPR in Testing for Case 4

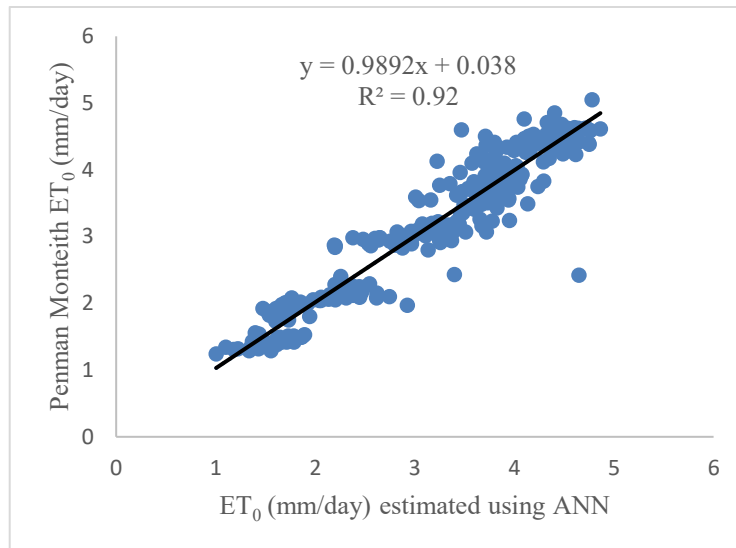


Figure 23. Performance of ANN in Training for Case 4

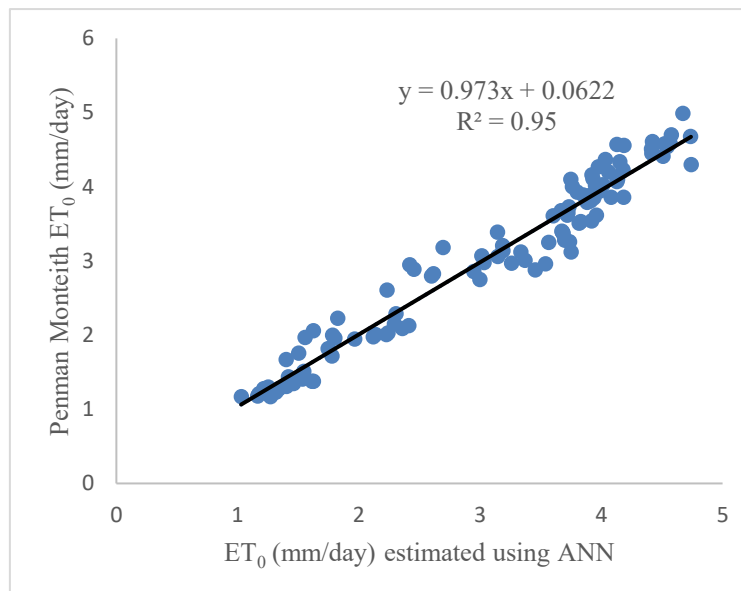


Figure 24 Performance of ANN in Training for Case 4

The performance of the soft computing models under different input scenarios is summarized in Tables 3 and 4 for training and testing datasets, respectively. Overall, all three techniques—SVM, GPR, and ANN—demonstrate strong predictive capability, with R^2 values exceeding 0.90 in most cases, highlighting their suitability for ET₀ estimation.

For the training dataset, GPR yields the highest R^2 values, reaching 0.99 for input Cases 1 and 2. This superior performance may be attributed to the probabilistic nature of GPR and its ability to adaptively learn kernel hyperparameters from the data. However, during the testing phase, ANN consistently outperforms both SVM and GPR across all input scenarios. The ANN achieves the highest R^2 value of 0.98 for Case 1 and maintains stable performance even when the number of input variables is reduced.

The improved generalization capability of ANN in the testing phase suggests that neural networks are particularly effective in capturing the nonlinear and dynamic relationships between climatic variables and reference evapotranspiration. Similar observations have been reported by Kumar et al. (2002), Kisi (2009), and Hassan Afzaal et al. (2020), who found ANN models to outperform other machine learning techniques in ET₀ prediction under diverse climatic conditions. The superior performance of ANN observed in this study is consistent with earlier findings reported for different hydro-climatic regions (Adamowski & Chan, 2011; Kisi & Cimen, 2011; Feng et al., 2019).

SVM models show competitive performance, especially when complete meteorological inputs are available, but their accuracy declines slightly under reduced input scenarios. This behavior reflects the sensitivity of SVM to kernel selection and parameter tuning, as noted in previous hydrological modeling studies (Mehdizadeh et al., 2017). GPR models maintain relatively stable performance across all cases, demonstrating robustness under varying data availability, although their computational complexity increases with larger datasets.

Figures 7–24 illustrate the correlation between ET₀ values estimated using soft computing techniques and those obtained from the Penman–Monteith method for different cases and datasets. The clustering of data

points around the 1:1 line further confirms the high predictive accuracy of the models, particularly for ANN-based estimates.

Implications for Irrigation Water Management

An analysis of monthly ET_0 and rainfall patterns indicates that atmospheric water demand exceeds rainfall from October to early June, highlighting the necessity for supplemental irrigation during these periods. Accurate estimation of ET_0 is therefore crucial for irrigation scheduling and water allocation planning in the Eastern Gandak Project command area. The superior performance of ANN models, even under limited data scenarios, suggests that soft computing techniques can serve as reliable tools for ET_0 estimation in regions where complete meteorological datasets are unavailable.

5. Conclusions

In this study, reference evapotranspiration (ET_0) was estimated for the Eastern Gandak Project region of Bihar using both conventional empirical methods and soft computing techniques based on long-term meteorological data from 1981 to 2024. The FAO-56 Penman–Monteith method was adopted as the reference standard to evaluate the performance of alternative approaches.

Among the conventional methods examined, the FAO-24 Modified Penman method exhibited the strongest agreement with the Penman–Monteith estimates, indicating its relative reliability when sufficient climatic data are available. In contrast, the Pan Evaporation method showed weaker correlation, emphasizing the influence of local pan exposure and the need for site-specific calibration when this method is applied. Temperature-based approaches such as the Hargreaves method were found to overestimate ET_0 during warmer periods, particularly under humid climatic conditions.

The soft computing techniques demonstrated superior predictive capability compared to conventional empirical methods. While Gaussian Process Regression achieved excellent performance during model training, Artificial Neural Networks consistently provided the most accurate and stable estimates during the testing phase across all input scenarios. This highlights the strong generalization ability of ANN models in capturing the nonlinear relationships between climatic variables and reference evapotranspiration. Support Vector Machine models also produced satisfactory results but showed greater sensitivity to reduced input datasets. Analysis of the seasonal relationship between ET_0 and rainfall revealed that atmospheric water demand exceeds effective rainfall from October to early June, indicating a clear requirement for supplemental irrigation during these periods. The findings suggest that soft computing models, particularly ANN-based approaches, can serve as reliable alternatives for ET_0 estimation in regions with limited or incomplete meteorological data. Overall, the results demonstrate that integrating data-driven techniques with conventional evapotranspiration methods can enhance irrigation water management and support sustainable agricultural planning in canal-command areas. Future research may focus on incorporating climate change projections, hybrid modeling frameworks, and uncertainty analysis to further improve reference evapotranspiration estimation under evolving climatic conditions.

References

1. Abdullah, S. S., Malek, M. A., Abdullah, N. S., Kisi, O., & Yap, K. S. (2015). Extreme learning machines: A new approach for prediction of reference evapotranspiration. *Journal of Hydrology*, 527, 184–195. <https://doi.org/10.1016/j.jhydrol.2015.04.073>
2. Adamala, S., Raghuvanshi, N. S., Mishra, A., & Tiwari, M. K. (2013). Evapotranspiration modeling using second-order neural networks. *Journal of Hydrologic Engineering*, 19(6), 1131–1140. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000871](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000871)
3. Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration: Guidelines for computing crop water requirements (FAO Irrigation and Drainage Paper No. 56). FAO.
4. Blaney, H. F., & Criddle, W. D. (1950). Determining water requirements in irrigated areas from climatological and irrigation data (USDA SCS TP-96). USDA.
5. Doorenbos, J., & Pruitt, W. O. (1977). Guidelines for predicting crop water requirements (FAO Irrigation and Drainage Paper No. 24). FAO.
6. Feng, Y., Cui, N., Zhao, L., Hu, X., & Gong, D. (2016). Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid regions of Southwest China. *Journal of Hydrology*, 536, 376–383. <https://doi.org/10.1016/j.jhydrol.2016.02.053>
7. Feng, Y., Cui, N., Zhao, L., Gong, D., & Zhang, K. (2017). Spatiotemporal variation of reference evapotranspiration during 1954–2013 in Southwest China. *Quaternary International*, 441, 129–139. <https://doi.org/10.1016/j.quaint.2016.10.030>
8. Firat, M. (2008). Comparison of artificial intelligence techniques for river flow forecasting. *Hydrology and Earth System Sciences*, 12(1), 123–139. <https://doi.org/10.5194/hess-12-123-2008>
9. Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, 1(2), 96–99.
10. Hassan Afzaal, A., Farooque, A. A., Abbas, F., Acharya, B., & Esau, T. (2020). Computation of evapotranspiration with artificial intelligence for precision water resource management. *Applied Sciences*, 10(5), 1621. <https://doi.org/10.3390/app10051621>

11. Islam, A. R. M. T., Shen, S., Yang, S., Hu, Z., & Chu, R. (2019). Assessing recent impacts of climate change on design water requirement of Boro rice season in Bangladesh. *Theoretical and Applied Climatology*, 138(1–2), 97–113. <https://doi.org/10.1007/s00704-019-02818-8>
12. Jerin, J. N., Islam, H. M. T., Islam, A. R. M. T., Shahid, S., & Hu, Z. (2021). Spatiotemporal trends in reference evapotranspiration and its driving factors in Bangladesh. *Theoretical and Applied Climatology*, 144(3–4), 793–808. <https://doi.org/10.1007/s00704-021-03566-4>
13. Kalteh, A. M. (2013). Monthly river flow forecasting using artificial neural networks and support vector regression models coupled with wavelet transform. *Computers & Geosciences*, 54, 1–8.
14. Khoob, A. R. (2007). Comparative study of Hargreaves and artificial neural network methodologies in estimating reference evapotranspiration in a semiarid environment. *Irrigation Science*, 26(3), 253–259.
15. Khoob, A. R. (2008). Artificial neural network estimation of reference evapotranspiration from pan evaporation in a semi-arid environment. *Irrigation Science*, 27(1), 35–39.
16. Kisi, O. (2009). Fuzzy genetic approach for modeling reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, 136(3), 175–183.
17. Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W., & Pruitt, W. O. (2002). Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128(4), 224–233.
18. Kumar, M., Raghuwanshi, N. S., & Singh, R. (2011). Artificial neural networks approach in evapotranspiration modeling: A review. *Irrigation Science*, 29(1), 11–25.
19. Lu, W. Z., & Wang, W. J. (2005). Potential assessment of the support vector machine method in forecasting ambient air pollutant trends. *Chemosphere*, 59(5), 693–701.
20. Mehdizadeh, S., Behmanesh, J., & Khalili, K. (2017). Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. *Computers and Electronics in Agriculture*, 139, 103–114.
21. Pereira, L. S., Allen, R. G., Smith, M., & Raes, D. (2015). Crop evapotranspiration estimation with FAO56: Past and future. *Agricultural Water Management*, 147, 4–20.
22. Salam, R., Islam, A. R. M. T., Pham, Q. B., Dehghani, M., Al-Ansari, N., & Linh, N. T. T. (2020). The optimal alternative for quantifying reference evapotranspiration in climatic sub-regions of Bangladesh. *Scientific Reports*, 10, 20171.
23. Valiantzas, J. D. (2013). Simplified forms for the standardized FAO-56 Penman–Monteith reference evapotranspiration using limited weather data. *Journal of Hydrology*, 505, 13–23.
24. Vapnik, V. N. (1999). *The nature of statistical learning theory* (2nd ed.). Springer.
25. Vapnik, V. N., Golowich, S. E., & Smola, A. J. (1997). Support vector method for function approximation, regression estimation, and signal processing. *Advances in Neural Information Processing Systems*, 9, 281–287.
26. Wang, S. D., Wu, K. P., & Wang, C. J. (2009). Kernel-based Gaussian process regression for nonlinear system modeling. *Pattern Recognition*, 42(5), 710–717.
27. Kisi, O., & Cimen, M. (2011). A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *Journal of Hydrology*, 399(1–2), 132–140.
28. Adamowski, J., & Chan, H. F. (2011). A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*, 407(1–4), 28–40.
29. Traore, S., Wang, Y. M., & Kerh, T. (2010). Artificial neural network for modeling reference evapotranspiration. *Agricultural Water Management*, 97(10), 1467–1477.
30. Tabari, H., Kisi, O., Ezani, A., & Talaei, P. H. (2012). SVM, ANFIS, and ANN modeling of reference evapotranspiration in different climates of Iran. *Water Resources Management*, 26, 303–317.
31. Gocic, M., & Trajkovic, S. (2015). Software for estimating reference evapotranspiration using limited weather data. *Computers and Electronics in Agriculture*, 113, 158–165.
32. Feng, Y., Gong, D., Zhang, Q., & Cui, N. (2019). Evaluating machine learning approaches for reference evapotranspiration estimation. *Hydrological Processes*, 33(15), 2163–2177.
33. Islam, A. R. M. T., Pham, Q. B., & Elbeltagi, A. (2021). Deep learning approaches for reference evapotranspiration estimation. *Journal of Hydrology*, 603, 127011.