Journal of Agrometeorology

(A publication of Association of Agrometeorologists)

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 27 (4) : 464-469 (December - 2025)
https://doi.org/10.54386/jam.v27i4.2909
https://journal.agrimetassociation.org/index.php/jam

Research Paper

Machine learning modeling of reference evapotranspiration in Central Luzon, Philippines

LEA S. CAGUIAT*, RONALDO B. SALUDES, MARION LUX Y. CASTRO and RUBENITO M. LAMPAYAN

Agrometeorology, Bio-Structures and Environment Engineering Division, Institute of Agricultural and Biosystems Engineering, College of Engineering and Agro-Industrial Technology, University of the Philippines, Los Baños, Laguna 4031 Philippines
** Corresponding Author: caguiatlea07@gmail.com*

ABSTRACT

Reference evapotranspiration (ET_o) is crucial for calculating irrigation requirements. Instruments that directly measure ET_o are still costly and limited while the empirical models are data intensive. Meteorological data of Central Luzon, Philippines (1985-2019) were used to estimate ET_o using the FAO Penman-Monteith method. The performances of machine learning algorithms in estimating ET_o were analyzed using ground-based weather data. Optimal models were determined using decision thresholds (RMSE<0.39 mm day^{-1}, R^2>0.75, MSE<0.15 mm day^{-1}, MAE<0.30 mm day^{-1}). The models were further assessed using principal component analysis for finding relevant variables (σ^2=0.95) and the Wilcoxon test for comparing two samples (α=0.05). Results show that optimal model required only two or three weather variables depending on the station. In general, the algorithms can be ranked as follows: Gaussian progress regression, Neural network, Support vector machines, Ensemble of trees, Regression trees, and Linear regression. The study reveals that machine learning can accurately predict ET_o using ground-based weather data, and it can be a good alternative to data-intensive empirical models.

Keywords: Reference evapotranspiration, FAO Penman-Monteith, Machine learning, Central Luzon, Limited weather data

Evapotranspiration (ET) is a combination of two processes where water is lost from soil and crop surfaces by evaporation and transpiration. One of the most common ET concepts is reference evapotranspiration (ET_o) which represents ET from a standardized vegetated surface, not short of water, at any given climatic condition (Allen *et al.,* 1998). Determining ET can be done through direct or indirect methods. Direct methods include lysimeters, the most accurate method of quantifying ET. However, lysimeters are still limited because they are difficult to construct and require high cost and maintenance. Indirect methods include climatological methods, where FAO Penman-Monteith is the most conventional. Its main drawback is that it requires numerous meteorological parameters, and many places do not have such data due to incomplete instruments in weather stations. New methods to estimate ET_o with fewer inputs are crucial for agricultural managers in irrigation system design and water management in rice farming regions like Central Luzon. With the rise of machine learning, advanced algorithms may now be explored for ET_o estimation based on historical data and trends.

Numerous studies reported the ability of machine learning to estimate ET in data-sparse regions (Valipour *et al.,* 2019; Granata, 2019). It has also been used for cross-station modeling where ET_o was estimated using the data from nearby stations (Karimi *et al.,* 2017). Efforts were also employed to overcome the lack of meteorological data using remote sensing (El-Shirbeny, 2016). Some machine learning algorithms include Linear regression, Instance-based models (IBM), Regression trees, Neural networks, Ensemble learning, Support vector machines (SVM), and Gaussian progress regression (GPR). Wu *et al.,* (2019) tested eight different models to estimate monthly mean daily ET_o in Jiangxi, China. Results show that the multivariate adaptive regression splines models were slightly superior to the others, notably when temperature data from adjacent stations were used. Feng and Tian (2021) proved that the K-nearest neighbor, the most common IBM, is consistent with the FAO Penman-Monteith method in ET_o modeling in semi-arid environments. ET_o estimation in China reported that tree-based models such as random forest and Gradient Boosting Decision Trees exhibited higher estimation accuracy than the other models in

Article info - DOI: *https://doi.org/10.54386/jam.v27i4.2909*
Received:28 January 2025; Accepted: 6 September 2025; Published online : 1 December 2025

Table 1: Synoptic weather stations in Central Luzon, Philippines

Station name	Latitude (°N)	Longitude (°E)	Elevation (m)	Duration
Baler (Radar), Aurora	15.75	121.63	173	1995-2018
Casiguran, Aurora	16.27	122.13	4	1985-2018
Clark Airport (DMIA), Pampanga	15.17	120.56	152	1998-2018
Cubi Point, Subic Bay Olongapo City	14.79	120.27	19	1995-2018

ET_o was computed using the FAO Penman-Monteith method. Due to the lack of solar radiation data, it was estimated using the differences in maximum and minimum temperature which is closely related to the existing solar radiation at a given location (Darshana *et al.,* 2013). The FAO ET_o calculator was used to facilitate the computations using the equation (Allen *et al.,* 1998):

Table 2: Input combinations used in machine learning modeling

Machine learning algorithms	Regression model options	
Linear Regression	(i) Linear (ii) Interactions linear	(iii) Robust linear (iv) Stepwise linear
Regression tree	Fine tree Medium tree	Coarse tree Optimizable tree
Support vector machine (SVM)	Linear SVM Quadratic SVM Cubic SVM Fine Gaussian SVM	Medium Gaussian SVM Coarse Gaussian SVM Optimizable SVM
Gaussian progress regression (GPR)	(i) Rational quadratic (ii) Squared exponential (iii) Matern 5/2	(iv) Exponential (v) Optimizable GPR
Ensembles of trees	Boosted trees Bagged trees	Optimizable ensemble
Neural network	Narrow neural network Medium neural network	Wide neural network Bilayered neural network Trilayered neural network

Table 3: Decision thresholds in evaluating the optimal machine learning model

Parameter	Range	Threshold
RMSE	0 to +∞	RMSE < 0.39
R^2	0 to 1	$R^2 > 0.75$
MSE	0 to +∞	MSE < 0.15
MAE	0 to +∞	MAE < 0.30
PCA	0.01 to 1	$\sigma^2 = 0.95$
Wilcoxon Test	0 to 1	$\alpha = 0.05$

the local application (Wu *et al.,* 2019). For cross-station scenarios, neural networks presented a good performance in the Mediterranean region of Turkey.

The general objective of the study is to conduct ET_o modeling in Central Luzon, Philippines. It specifically aimed to: (1) evaluate the performance of various machine learning algorithms and input combinations in estimating ET_o using ground-based weather data; and, (2) determine the optimal machine learning algorithm and input combination using the established decision thresholds.

MATERIALS AND METHODS

Data description

Daily meteorological data on maximum air temperatures (Tmax), minimum air temperatures (Tmin), relative humidity (RH), and wind speed from 1985 to 2018 were collected from four synoptic weather stations of Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA, 2020), Central Luzon, Philippines (Table 1).

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

where ET_o is the reference ET (mm day^{-1}); Rn is the net radiation at the crop surface (MJ/m^2/day); G is the soil heat flux density (MJ m^{-2} day^{-1}); T is the mean daily air temperature at 2 m (°C); u_2 is the wind speed at 2 m (m s^{-1}); e_s is the saturation vapor pressure at the temperature of air (kPa); e_a is the actual vapor pressure (kPa); Δ is the slope of vapor pressure curve (kPa °C^{-1}); γ is the psychrometric constant (kPa °C^{-1}).

Machine learning models

Six variants of the models for each scenario were built by changing the applied algorithms: (a) Linear regression; (b) Regression tree; (c) Support vector machine; (d) Gaussian progress regression; (e) Ensemble of trees; and (f) Neural network. Each algorithm comprises several model options presented in Table 2. The dataset was divided into training (60%), validation (20%), and test sets (20%). The training set is the sample of data used to fit the models. The validation set is used to provide an unbiased evaluation of a model fit on the training dataset while tuning model hyperparameters (Shah, 2017). Models that differed in input weather

Table 4: Input combinations used in evaluating the optimal machine learning model

Model	Station	Input combination
1 (One weather variable)	Baler	Tmax
	Casiguran	Tmax
	Clark	Tmax
	Cubi Point	RH
2 (Two weather variables)	Baler	Tmax, wind speed
	Casiguran	Tmax, RH
	Clark	Tmax, RH
	Cubi Point	RH, Tmax
3 (Three weather variables)	Baler	Tmax, wind speed, RH
	Casiguran	Tmax, RH, Tmin
	Clark	Tmax, RH, wind speed
	Cubi Point	RH, Tmax, wind speed
4 (Four weather variables)	Baler	Tmax, wind speed, RH, Tmin
	Casiguran	Tmax, RH, Tmin, wind speed
	Clark	Tmax, RH, wind speed, Tmin
	Cubi Point	RH, Tmax, wind speed, Tmin

variables (Tmax, Tmin, RH, and wind speed) per location were developed and tested. The models were built by adding one different variable into the input combination based on the result of feature selection and principal component analysis (PCA). The statistics and machine learning Toolbox™ of MATLAB R2021a were used to implement the regression algorithms. Steps in training regression models involved: (1) data selection; (2) training and validation; (3) model performance assessment; (4) choosing best model options; and (5) testing.

Statistical indicators

Statistical indices for measuring model performance were assessed using root mean square error (RMSE), coefficient of determination (R^2), mean square error (MSE), and mean absolute error (MAE). Higher R^2 values closer to 1 indicate high simulation accuracy whereas smaller values of RMSE, MSE, and MAE suggest better model performance. RMSE, MSE, and MAE match the unit of the response variable.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (ET_{o_m} - ET_{o_e})^2}$$

$$R^2 = \frac{[\sum_{i=1}^{n} \left(ET_{o_m} - \underline{ET_o}_m\right)\left(ET_{o_e} - \underline{ET_o}_e\right)]^2}{\sum_{i=1}^{n} (ET_{o_m} - \underline{ET_o}_m)^2 \sum_{i=1}^{n} (ET_{o_e} - ET_{o_e})^2}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (ET_{o_m} - ET_{o_e})^2$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |ET_{o_m} - ET_{o_e}|$$

where ET_{o_m}, ET_{o_e}, $\underline{ET_o}_m$ and ET_{o_e} are the measured, estimated, mean of measured, and mean of estimated ET_o by the FAO Penman-Monteith method and machine learning modeling, respectively; n is the number of observations.

Model selection

The decision thresholds for evaluating the best model for both scenarios are presented in Table 3. The thresholds for RMSE, MSE, and MAE were computed from a 10% mean deviation from the FAO Penman-Monteith method which is within the suggested limit of Hargreaves and Allen (2003). The threshold for R^2 can be described as a strong correlation in academic research (Mooi and Sarstedt, 2011). The models were further assessed using PCA to find relevant variables and reduce the dimensionality of datasets by determining the variance (σ^2). PCA is an analysis available in MATLAB for finding the relevant variables that explain at least 95% variance (MathWorks, 2021). In addition, the Wilcoxon test with a 5% significance level was implemented to compare two samples (Addinsoft, 2020).

RESULTS AND DISCUSSION

The input combination used for every station is presented in Table 4. Results show that GPR obtained the highest estimation accuracy for all stations in every input combination obtaining the highest R^2 (0.71-0.95) and lowest error values (RMSE=0.27-0.64, MSE=0.08-0.41, MAE=0.21-0.50). Out of all the GPR models, the optimizable GPR and Matern 5/2 GPR provided the highest accuracy. Very similar performances were shown by neural network and SVM which resulted to a high R^2 (0.71-0.95) and low error values (RMSE=0.28-0.64, MSE=0.08-0.41, MAE=0.20-0.51). Linear regression obtained the least predictive capabilities but based on the statistical indicators, it still produced results suitable for modeling daily ET_o. In general, the algorithms can be ranked as follows: GPR, neural network, SVM, ensemble of trees, regression tree, and linear regression.

The performance of all the models improves as the number of input variables increases. The use of four input weather variables (Tmax, Tmin, RH, wind speed) yielded the highest estimation accuracy with an 18% increase in R^2 and a decrease in error values (RMSE=0.22, MSE=0.19, MAE=0.17) relative to the model with one variable. Model 2 (RMSE=0.39, R^2=0.86, MSE=0.15, MAE=0.30) also obtained high accuracies using less input variables. The summary of the statistical indicators for Scenario 1 models is presented in Table 5.

In general, all models and combinations passed the Wilcoxon test. Based on the thresholds, the use of three variables (Model 3) in estimating ET_o in Baler and Casiguran gained enough variance and accuracy to reach the decision thresholds. Moreover, estimating ET_o in Clark and Cubi Point requires only two input weather variables (Model 2) to satisfy the thresholds. The results were consistent with both the rank correlation coefficient and the sensitivity analysis conducted in the study area, where the relevant weather variables were the same ones that showed significant correlation and the greatest sensitivity to ET_o (Caguiat *et al.,* 2022). Machine learning was also able to capture the most useful information stored in the available data without depending on any complex underlying knowledge about the specific area of interest (Khan *et al.,* 2023; Dou and Yang, 2018). Bijlwan *et al.,* (2024) were also able to determine the most important weather variable affecting ETo in Pantnagar, Uttarakhand using machine learning and

Table 5: Summary of the statistical indicators for the machine learning models (Scenario 1)

Station	Input combination*	Algorithm	Model	PCA	Wilcoxon test**	Statistical indicators RMSE (mm day^{-1})	R^2	MSE (mm day^{-1})	MAE (mm day^{-1})
Baler	Tmax	Neural network	Bilayered neural network	σ^2=0.85	P<0.05	0.50	0.71	0.25	0.39
	Tmax, Wind speed	Neural network	Bilayered neural network	σ^2=0.92	P<0.05	0.41	0.80	0.17	0.33
	Tmax, Wind speed, RH	Regression tree	Optimizable tree	σ^2=0.98	P<0.05	0.35	0.86	0.12	0.29
	Tmax, Wind speed, RH, Tmin	GPR	Optimizable GPR	σ^2=1.00	P<0.05	0.34	0.86	0.12	0.28
Casiguran	Tmax	Neural Network	Bilayered neural network	σ^2=0.69	P<0.05	0.42	0.81	0.18	0.33
	Tmax, RH	GPR	Optimizable GPR	σ^2=0.91	P<0.05	0.37	0.86	0.13	0.29
	Tmax, RH, Tmin	GPR	Optimizable GPR	σ^2=0.97	P<0.05	0.33	0.88	0.11	0.26
	Tmax, RH, Tmin, Wind speed	GPR	Optimizable GPR	σ^2=1.00	P<0.05	0.31	0.90	0.10	0.25
Clark	Tmax	GPR	Rational quadratic GPR	σ^2=0.94	P<0.05	0.53	0.73	0.29	0.42
	Tmax, RH	GPR	Matern 5/2 GPR	σ^2=0.98	P<0.05	0.38	0.87	0.15	0.30
	Tmax, RH, Wind speed	GPR	Optimizable GPR	σ^2=0.99	P<0.05	0.31	0.91	0.10	0.25
	Tmax, RH, Wind speed, Tmin	GPR	Optimizable GPR	σ^2=1.00	P<0.05	0.29	0.92	0.09	0.22
Cubi Point	RH	SVM	Medium Gaussian SVM	σ^2=0.93	P<0.05	0.64	0.71	0.41	0.50
	RH, Tmax	GPR	Optimizable GPR	σ^2=0.97	P<0.05	0.39	0.89	0.15	0.30
	RH, Tmax, Wind speed	Neural network	Trilayered neural network	σ^2=0.99	P<0.05	0.29	0.94	0.09	0.23
	RH, Tmax, Wind speed, Tmin	GPR	Optimizable GPR	σ^2=1.00	P<0.05	0.27	0.95	0.08	0.21

* The input combinations that passed the threshold are marked in bold; ** P<0.05 means that the result is significant at a 5% significance level (α = 0.05)

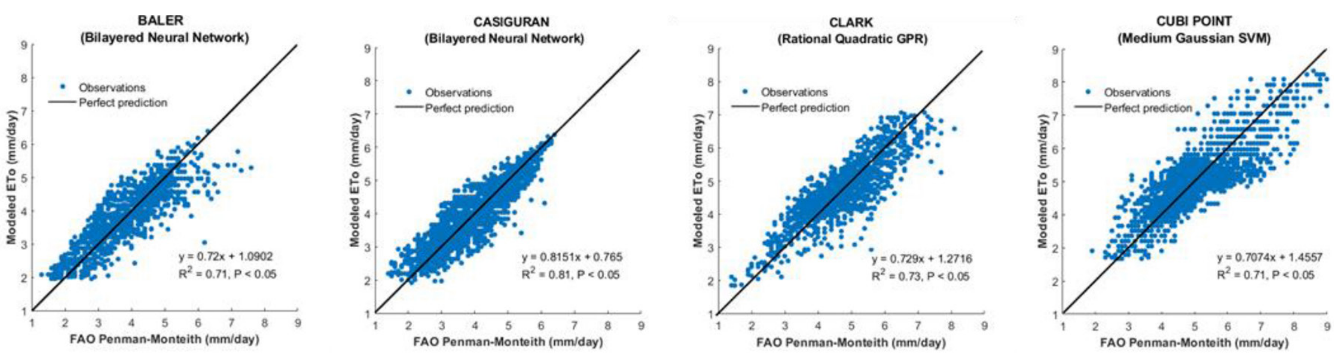

Fig. 1: Comparison between ETo of Model 1 and the FAO Penman-Monteith method

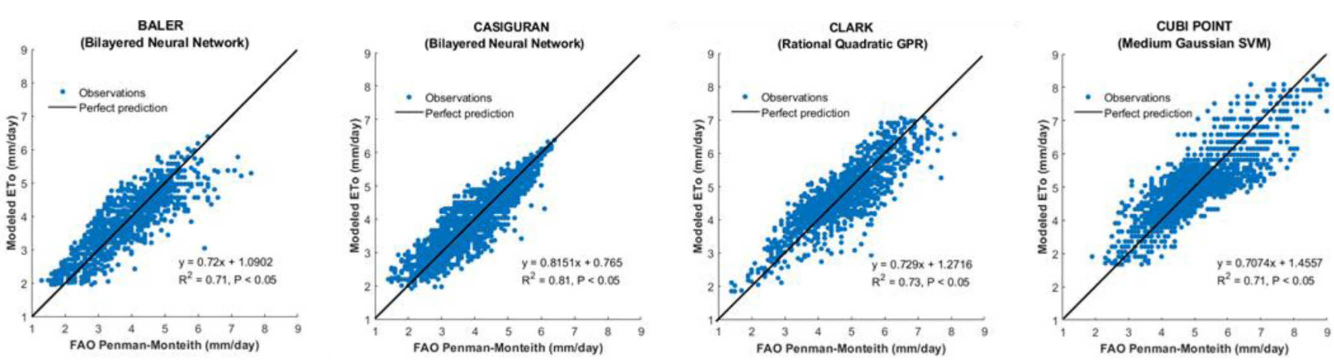

Fig. 2: Comparison between ETo of Model 2 and the FAO Penman-Monteith method

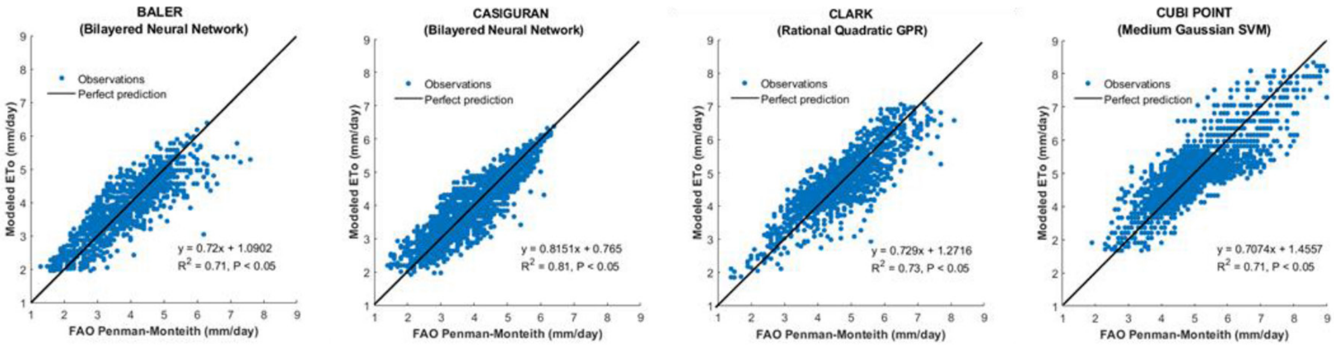

Fig. 3: Comparison between ETo of Model 3 and the FAO Penman-Monteith method

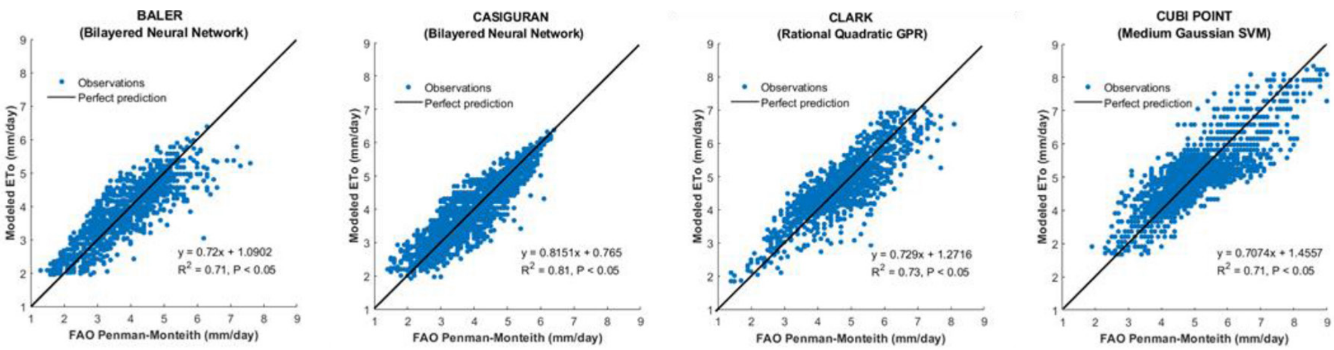

Fig. 4: Comparison between ETo of Model 4 and the FAO Penman-Monteith method

feature importance analysis. Results revealed that when there are data limitations, two or three input variables are enough to build a machine learning model that can estimate ET_o accurately. In this case, the ET_o in Central Luzon can be estimated using Tmax and RH (or wind speed) only. Fig. 1 to 4 show the relationship between ETo estimated by machine learning models and the FAO Penman-Monteith method.

CONCLUSION

GPR provided better ET_o estimates than the other machine learning algorithms. In general, the algorithms can be ranked as follows: GPR, Neural network, SVM, Ensemble of trees, Regression tree, and Linear regression. When ground-based data is used, the optimal combination requires only two or three input variables to satisfy the thresholds. This study proved that when there is limited ground-based weather data, machine learning is a powerful tool that can provide accurate ET_o predictions and can be a good alternative to data-intensive ET_o equations and empirical models.

ACKNOWLEDGMENT

The authors extend their deepest gratitude to the Agrometeorology, Biostructures, and Environment Engineering Division of the University of the Philippines-Los Baños for providing the resources used in the study, along with PAGASA, PhilRice, and TAU for the weather data needed for the study.

Funding: The study was funded by the Department of Science and Technology-Engineering Research and Development for Technology scholarship program.

Conflict of interest: The authors declare that there is no conflict of interest related to this article

Data availability: Data may be provided by the corresponding author upon request.

Authors contribution: **L. Caguiat**: Conceptualization, data acquisition, Analysis and interpretation, Drafting, Writing and editing; **R.B. Saludes:** Conceptualization, Analysis and interpretation, Review and editing; **M.L.Y. Castro:** Conceptualization, Review and editing; **R.M. Lampayan:** Conceptualization, Analysis and interpretation, Review and editing.

***Disclaimer**:* The contents, opinions, and views expressed in the research article published in the Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

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

REFERENCES

Addinsoft. (2020). XLSTAT support center: Wilcoxon signed rank test on two paired samples in excel tutorial. https://help.xlstat.com/s/article/wilcoxon-signed-rank-test-on-two-paired-samples-in-excel-tutorial?language=en_US

Allen, R., Pereira, L., Raes, D., and Smith, M. (1998). Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. Food and Agriculture Organization of the United Nations. https://

www.fao.org/4/X0490E/x0490e00.htm

Bijlwan, Amit, Shweta Pokhriyal, Ranjan, R., R.K Singh, and Ankita Jha. (2024). Machine learning methods for estimating reference evapotranspiration. *J. Agrometeorol.,* 26(1): 63-68. https://doi.org/10.54386/jam.v26i1.2462

Caguiat, L. S., Saludes, R. B., Castro, M. L. Y., and Lampayan, R. M. (2022). Spatio-Temporal Trend Analysis of Reference Evapotranspiration in Central Luzon, Philippines. *Philippine J. Sci.,* 152 (S1): 33-46. https://philjournalsci.dost.gov.ph/images/pdf/pjs_pdf/vol152_S1_Meteorology/spatio-temporal_trend_analysis_of_reference_evapotranspiration_.pdf

Darshana, Pandey, A., and Pandey, R. P. (2013). Analysing trends in reference evapotranspiration and weather variables in the Tons River Basin in Central India. *Stoch. Environ. Res. Risk Assess.,* 27(6):1407-1421. https://doi.org/10.1007/s00477-012-0677-7

Dou, X., and Yang, Y. (2018). Modeling Evapotranspiration Response to Climatic Forcings Using Data-Driven Techniques in Grassland Ecosystems. *Advan. Meteorol.,* 1-18. https://doi.org/10.1155/2018/1824317

El-Shirbeny, M. A. (2016). Evaluation Of Hargreaves Based on Remote Sensing Method To Estimate Potential Crop Evapotranspiration. *Intern. J. Geom.,* 11(23): 2143-2149. https://doi.org/10.21660/2016.23.1122

Feng, K., and Tian, J. (2021). RETRACTED ARTICLE: Forecasting reference evapotranspiration using data mining and limited climatic data. *Europ. J. Remote Sens.,* 54(sup2):363-371. https://doi.org/10.1080/22797254.2020.1801355

Granata, F. (2019). Evapotranspiration evaluation models based on machine learning algorithms—A comparative study. *Agric. Water Manag.,* 217: 303-315. https://doi.org/10.1016/j.agwat.2019.03.015

Hargreaves, G. H., and Allen, R. G. (2003). History and Evaluation of Hargreaves Evapotranspiration Equation. *J. Irrig. Drain. Engg.,* 129(1):53-63. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:1(53)

Karimi, S., Kisi, O., Kim, S., Nazemi, A. H., and Shiri, J. (2017). Modelling daily reference evapotranspiration in humid locations of South Korea using local and cross-station data management scenarios. *Intern. J. Climatol.,* 37(7): 3238-3246. https://doi.org/10.1002/joc.4911

Khan, Y., Kumar, V., Setiya, P., and Satpathi, A. (2023). Comparison of phenological weather indices based statistical, machine learning and hybrid models for soybean yield forecasting in Uttarakhand. *J. Agrometeorol.,* 25(3): 425-431. https://doi.org/10.54386/jam.v25i3.2232

MathWorks. (2021). Principal component analysis of raw data (Matlab R2021a documentation). https://www.mathworks.com/help/stats/pca.html;jsessionid=54b1f99cfeb0c309c7272a02abf8

Mooi, E., and Sarstedt, M. (2011). A Concise Guide to Market Research. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12541-6

PAGASA. (2020). Philippine Atmospheric, Geophysical and Astronomical Services Administration. Climatological data from DOST-PAGASA. http://bagong.pagasa.dost.gov.ph/climate/climate-data

Shah, T. (2017). Train, validation, and test sets. https://tarangshah.com/blog/2017-12-03/train-validation-and-test-sets/

Valipour, M., Gholami Sefidkouhi, M. A., Raeini-Sarjaz, M., and Guzman, S. M. (2019). A Hybrid Data-Driven Machine Learning Technique for Evapotranspiration Modeling in Various Climates. *Atmos.,* 10(6): 311. https://doi.org/10.3390/atmos10060311

Wu, L., Peng, Y., Fan, J., and Wang, Y. (2019). Machine learning models for the estimation of monthly mean daily reference evapotranspiration based on cross-station and synthetic data. *Hydrol. Res.,* 50(6): 1730-1750. https://doi.org/10.2166/nh.2019.060