INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

R&D SHOWCASE 2021

REFERENCE EVAPOTRANSPIRATION USING MACHINE LEARNING MODELS OVER SEMI-ARID CLIMATES

[1]Adeeba Ayaz , [2]Maddu Rajesh, [3]Shailesh Kumar Singh, [4]Rehana. S.

International Institute of Information Technology, Hyderabad -500032

INTRODUCTION

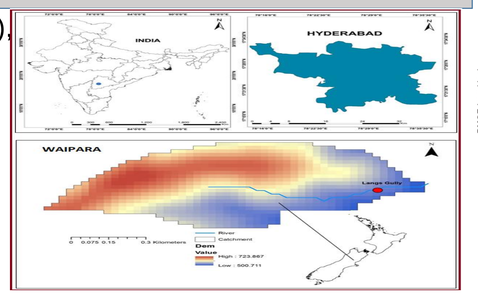
Reference Evapotranspiration (ET) is a complex hydrological variable defined by various climatic variables affecting water and energy balances and critical factors for crop water requirements and irrigation scheduling.

OBJECTIVES

- To develop different Machine Learning models, Support Vector Regression(SVR), Artificial Neural Network(ANN), Random Forest(RF) and Gradient Boosting Regression(GBR) for modelling ET in Hyderabad and Waipara Stations.
- To assess the performance and stability of these models with different input combinations over the two stations.
- To find an appropriate approach to boost the modelling performance under the limited input factors condition.

STUDY AREA & METHODOLOGY

The study areas are Hyderabad(Top), which is the capital southern Indian province of Telangana and Waipara(Bottom) situated in the South Island of New Zealand ,in the Waipara River. The study used everyday meteorological data from Jan(1965)-Dec(2015).



The standard equation used is FAO-Penman Monteith method, given as

$$ET = \frac{0.408D(R_n - G) + g(900/T+273)U_2(es - ea)}{D + g(1 + 0.34U_2)}$$

Where, ET = reference evapotranspiration (mm d^{-1}), D = slope vapor pressure curve [k pa°C^{-1}], R_n = net radiation (MJ m^{-2} d^{-1}), G = soil heat flux (MJ m^{-2} d^{-1}), U_2 = wind speed measured at 2 m height [m s^{-1}], (e_s - e_a) = pressure deficit for measurement at 2 m height [k Pa], T = average temperature at 2 m height (°C), 900 = coefficient for the reference crop [l J^{-1} Kg K d^{-1}], g = psychrometric constant [k pa°C^{-1}], 0.34 = wind coefficient for the reference crop [s m^{-1}].

Four ML models were implemented for modelling the ET relationship of the Hyderabad and Waipara stations, namely, ANN, GBR, SVR and RF regressor and a comparison was made between the models.

RESULTS & DISSCUSIONS

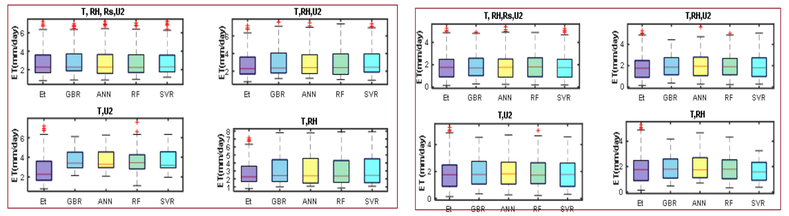


Fig. 1. Comparison of observed and estimated ET by different models (with varying parameters of input for the validation period at Hyderabad (left) and Waipara Stations (Right).

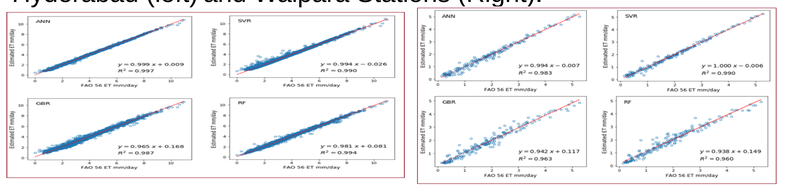


Fig.2. Scatter plots of the FAO 56 ET and those estimated by the ML models with all parameters of input for during the testing period at Hyderabad (Left) and Waipara Stations (Right).

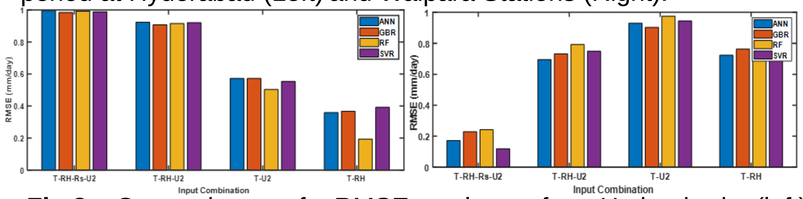


Fig.3 Comparison of RMSE values for Hyderabad (left) and Waipara station (Right) for different input combinations.

Parameters	Model	Training (Calibration)			Validation		
		R^2 (best is 1.0)	RMSE (best is 0.0)	MAE (best is 0.0)	R^2 (best is 1.0)	RMSE (best is 0.0)	MAE (best is 0.0)
All parameters	ANN	0.997	0.062	0.05	0.997	0.067	0.05
	GBR	0.99	0.167	0.12	0.987	0.183	0.13
	RF	0.998	0.057	0.031	0.994	0.123	0.079
	SVR	0.991	0.161	0.109	0.990	0.155	0.114
Temperature, Wind Speed, Relative Humidity	ANN	0.93	0.468	0.28	0.924	0.401	0.22
	GBR	0.932	0.457	0.28	0.908	0.457	0.29
	RF	0.989	0.183	0.105	0.916	0.436	0.261
	SVR	0.929	0.469	0.273	0.921	0.422	0.261
Temperature and Wind Speed	ANN	0.722	0.933	0.62	0.588	0.968	0.62
	GBR	0.734	0.908	0.61	0.572	0.989	0.63
	RF	0.88	0.608	0.397	0.503	1.065	0.68
	SVR	0.717	0.936	0.621	0.554	1.008	0.644
Temperature and Relative Humidity	ANN	0.533	1.202	0.90	0.359	1.222	0.94
	GBR	0.545	1.187	0.89	0.367	1.201	0.94
	RF	0.727	0.92	0.669	0.194	1.36	1.035
	SVR	0.529	1.209	0.899	0.392	1.177	0.908

Parameters	Model	Training (Calibration)			Validation		
		R^2 (best is 1.0)	RMSE (best is 0.0)	MAE (best is 0.0)	R^2 (best is 1.0)	RMSE (best is 0.0)	MAE (best is 0.0)
All parameters	ANN	0.983	0.151	0.11	0.983	0.172	0.12
	GBR	0.988	0.126	0.09	0.963	0.229	0.16
	RF	0.996	0.07	0.041	0.960	0.242	0.163
	SVR	0.994	0.086	0.067	0.990	0.119	0.085
Temperature, Wind Speed, Relative Humidity	ANN	0.764	0.575	0.42	0.65	0.695	0.51
	GBR	0.785	0.545	0.41	0.614	0.732	0.56
	RF	0.95	0.261	0.182	0.548	0.792	0.574
	SVR	0.74	0.599	0.417	0.596	0.749	0.543
Temperature and Wind Speed	ANN	0.512	0.805	0.60	0.378	0.93	0.75
	GBR	0.579	0.763	0.57	0.413	0.903	0.73
	RF	0.77	0.564	0.387	0.311	0.975	0.771
	SVR	0.459	0.865	0.603	0.359	0.944	0.744
Temperature and Relative Humidity	ANN	0.731	0.607	0.46	0.621	0.723	0.55
	GBR	0.769	0.564	0.42	0.582	0.762	0.57
	RF	0.867	0.426	0.294	0.49	0.833	0.592
	SVR	0.71	0.633	0.452	0.618	0.728	0.541

Table 2&3 Performance of different ML models for training and validation periods for Hyderabad(Left) and Waipara(Right) Stations.

CONCLUSION

- The study investigated that the best performance was when all input variables were used, the study, however, finds that even three input variable combination (Temperature, Wind Speed and Relative Humidity values) or two combination input variables (Temperature and Relative Humidity, Temperature and Wind Speed) also can provide practically identical results as using all data.
- The ANN model offered the most remarkable performance among four tested models regardless of under which station or input combination, trailed by SVR and GBR models, which could likewise accomplish moderately good performance.
- This study also concludes that even if not all parameter information is available in a particular station, we can use the three combination parameters which are Temperature, Relative Humidity, and Wind Speed or the two combinations, which are Temperature and Relative Humidity or Temperature and Wind Speed values, to estimate reference ET.

[1,2]Research Scholar , Lab for Spatial Informatics, IIITH,, [3]Hydrologist, National Institute of Water & Atmospheric Research Ltd (NIWA), New Zealand, [4]Assistant Professor, Lab for Spatial Informatics, IIITH