Modeling Reference Evapotranspiration with Wind Speed and Relative Humidity Data

Rajesh Vijaykumar Kherde[*1]and Sumana Sahoo[#2].

*Sandip Institute of Technology & Research Center, Sandip Foundation, Nashik, Maharashtra, India.
[1] rkherde@gmail.com
#Transheight Consultants Pvt. Ltd. Bangalore, Karnataka, India
[2]sumanatrans@gmail.com

Abstract--**The objective of this research is to investigate whether evapotranspiration can be modelled using wind speed and relative humidity data only. A linear model based on wind speed and relative humidity has been formulated for estimation of pan coefficient (K_p). The equation is calibrated by using multivariate regression analysis. The coefficients of wind speed and relative humidity thus obtained were used for deciding the range of Monte Carlo simulations. Nash-Sutcliffe efficiency (NSE) is used as an objective function and the results of Monte Carlo simulations were used for sensitivity analysis. Sensitive model parameter shows differences in separation and form between the cumulative frequency distribution curves. Sensitivity analysis revealed that the simple linear model developed in this study is sensitive to relative humidity only; the wind velocity is an insensitive parameter. The impression after conducting this study is that the pan evaporation equation can be modelled using relative humidity and wind velocity data with the Nash-Sutcliff efficiency varying between 65 to 90 per cent which is acceptable.**

Keyword - Evapotranspiration, Pan Evaporation, Monte-Carlo Simulations, Sensitivity Analysis

I.   INTRODUCTION

Reference evapotranspiration is useful for water resource planning and irrigation scheduling. The meteorological data required for estimation of reference evapotranspiration (ET0) by standard FAO-56 PM equation may not be available in data sparse region. In such case temperature based empirical formulas are used for estimation of reference evapotranspiration. In absence of temperature data the two readily available climate data viz wind speed and relative humidity can be used for conversion of pan evaporation into ET0. The conversion relies on accurate evaluation of pan coefficient (Kp). Meteorological and lysimeter data are mostly not available. Measurement of ET0 using empirical equations thus becomes unfeasible. On the other hand pan evaporation data collected using class A pan is widely available. Conversion of pan evaporation into reference evapotranspiration is achieved through estimation of correction factor (Kp). Pan coefficients rely upon wind speed, relative humidity and upwind fetch distance. A number of empirical equations were proposed in past for conversion of pan evaporation into ET0. These equations include second degree equation (Frevert et al., 1983), nonlinear equation (Allen and Pruitt, 1991), the simple equation developed by Snyder (1992) and Orang (1998) equation. The best method for estimating pan coefficient is Snyder equation (Hadi Modaberi and Mostafa Assari, 2014). Researchers (Yi Li and Tusheng Ren, 2010) studied time scale effect on reference evaporation and noticed that Kp at the daily scale were generally larger than the Kp at monthly scale. Hossein Tabari et.al (2011) concluded that majority of mass transfer based equations underestimated ET0 in humid environment and among the pan evaporation based methods, the ET0 calculated by Snyder equation best matched the ETo estimates from the PMF-56 equation with the lowest error rates. Tim R. McVicaret.al. concluded that the influence of topography, especially aspect, is seen on the resultant ET0 and Kp, but not Epansurfaces.Khil-Ha Lee and Hong-Yeon Cho suggested equation that relies on daily wind speed and surface temperature data and obtained reasonable match. a reference evapotranspiration model based on fuzzy least square method presented by Sriram A.V, Rashmi C.N (2014) revelled that the fuzzy regression model could be successfully employed in estimating the daily Potential Evapotranspiration. Vassilis G. Aschonitiset. al. (2011) showed that Cuenca's equation provided more accurate daily estimations. Eric W. Harmsenet. al (2004) updated pan evaporation coefficient values for the seven university of Puerto Rico experimental substations and noticed significant difference between the mean Kp values calculated with pan evaporation data from 1960 to 1980 and those with data from 1981 to 2000.Chong-yuXuet. al. (2006) analysed spatial distribution and temporal trend of reference evapotranspiration and pan evaporation in changing catchment. Their study showed that the spatial distribution pattern of the pan coefficient is significantly influenced by wind speed and relative humidity in the region. Hamidreza Golkar HamzeeYazd et. al. (2015) concluded that the regions with a wider vegetation cover and a more various vegetation density had less evapotranspiration and as the range of vegetation reduces and the density increases, a stepped rise in evapotranspiration is expected.

II. DATA USED

The meteorological data was collected form hydrological data user group (HDUG) Nashik, Maharashtra, India. The data comprises maximum and minimum temperature, relative humidity, wind velocity (measured at 2m above ground) and pan evaporation. The data were sorted year wise and calculations were done for aerodynamic resistance (Ra), bulk surface resistance (Rs), saturation vapour pressure deficit, net radiation (Rn), soil heat flux (G) etc. Using measured and calculated data the year wise daily reference evapotranspiration values based on FAO-56 penman's equation were calculated for each station. For missing data the average value of preceding and succeeding years for the particular metrological parameter has been adopted.

III. METHODOLOGY

The Meteorological data was collected for the period of 2002 to 2008 for four Meteorological stations. However substantial amount of data (period of one month or more) was missing for the years 2000, 2001 and 2009. Hence, these years were excluded from the analysis. Multivariate analysis was performed for each year's data of every station. The results of multivariate analysis are tabulated in Table I.

Table I. Multivariate analysis for coefficients of evapotranspiration parameters.

STATION	YEAR	β1	β2	β0	R2	NSE
Dhaba	2002	-0.00267	-0.00042	1.269898	0.022174	0.623979
	2003	0.015891	-0.00032	0.493192	0.107817	0.750077
	2004	0.001178	-0.00021	0.931664	0.025333	0.300055
	2005	0.01108	0.000541	0.437437	0.08585	0.606397
	2006	0.001125	0.000141	1.077582	0.01429	0.729244
	2007	0.021864	0.002244	0.522432	0.19006	-0.26854
	2008	0.015614	0.001065	0.671637	0.116108	0.244223
	ALL	0.00791315	0.000724	0.699796	0.022174	0.612076
Hinganghat	2002	-0.08148	-0.00808	1.539044	0.768867	-39.2172
	2003	-0.05815	-0.00693	1.722595	0.605059	-12.694
	2004	-0.00856	-1.8E-05	1.071052	0.229482	-0.34011
	2005	-0.02337	-0.00577	2.081749	0.373966	-0.24135
	2006	0.002079	0.0008	0.842876	0.110281	0.558528
	2007	0.006024	0.001189	0.440664	0.031147	0.434907
	2008	-0.0417	-0.00223	0.801702	0.642625	-48.7588
	ALL	-0.0193118	-0.00199	1.265506	0.768867	-2.22403
Sirpur	2002	0.04462	0.003019	-0.33041	0.148702	-1.24542
	2003	0.02407	0.003314	0.403431	0.268951	0.033546
	2004	0.006207	0.000334	0.622644	0.079567	0.434252
	2005	-0.00195	-0.00069	1.462252	0.18407	-0.75664
	2006	0.015601	0.002676	0.493623	0.213782	0.73381
	2007	0.003905	0.000899	0.860124	0.110888	0.599399
	2008	0.003783	0.000254	1.958243	0.076583	0.286484
	ALL	0.006059	0.000939	0.825886	0.148702	0.564802

The intention behind carrying out the multivariate analysis was to find the coefficients of linear regression model with two variables, namely Relative humidity and wind velocity as given under,

$$K_p = \beta_0 + \beta_1 * RH_{mean} + \beta_2 * U_2$$

Here, RHmean corresponds to relative humidity (%), and U2 corresponds to the wind speed (m/s) measured at 2 m height. Once this value has been determined, reference evapotranspiration can be calculated as:

$$ET_0 = K_p * E_p$$

The primary goal is to determine the set of coefficients β_i, such that model predicts the values of ET0 as accurately as possible compared to those produced by FAO-56 Penman-Monteith equation. The purpose was to investigate whether a linear model with only two variables (RHmean& U2) is adequate to fit the standard FAO-56 Penman- Monteith ET0 results. We also wanted to investigate the sensitivity of the parameters β_i by ranking

the parameters according to Monte-Carlo realizations and examining the proximity of the yearly sample sets. However the results of multivariate analysis are not really encouraging except for some years, were NSE is above 60 per cent. But it served the purpose of finding parameter spaces for Monte Carlo simulations.

Every model has a set of parameters that cannot be measured directly, but these parameters can be construed from the calibration process. The calibration process is generally a trial-and-error method that adjusts the parameter values in such a manner that the input-output behaviour of the model matches with the real world system which it represents. In evapotranspiration modelling the model generated values of evapotranspiration (ETp) are matched with those calculated using FAO-56 PM equation.

Manual calibration procedures are labour-intensive and time consuming. These difficulties with the manual calibration method have led to the evolution of automatic calibration procedures which utilize the speed and power of computers. Many studies using manual calibration have reported the difficulties in finding the optimum parameter estimate. Multiple local optimum parameter sets have been noticed while employing optimization algorithms irrespective of the modelling methodology. It still remains typically difficult, if not impossible to find a unique "Best" parameter set whose performance measure differs significantly from other parameter sets.

IV. MONTE CARLO SIMULATION TECHNIQUE

Repeated random sampling is the basis of all computational algorithms that are classified as Monte Carlo methods and are used to obtain numerical results. In this technique the distribution of unknown probabilistic entity is obtained by running over the simulations many times. Monte Carlo methods vary, but tend to follow a particular pattern:

- Selecting imprecisely known model input parameters to be sampled.

- Assigning ranges and probability distributions for each of these parameters.

- Generating many sample sets (realizations) with random values of model parameters.

- Running the model for all realizations to estimate uncertainty in model outcomes.

The optimization of the parameters of a model requires the use of an objective function. Objective function is a reference numerical quantity enabling the calibration to be improved. The choice of objective function to be used for a given model is a subjective decision which influences the values of the parameters and the performance of the model. The objective functions for hydrologic simulations used in the present study are given below,

The regression coefficient or the measure of the model efficiency discussed by Nash and Sutcliffe (1970), used in the present study to evaluate model performance is as given below,

$$\text{NSE } (\eta) = 1 - \frac{\sum_1^n (ET_0 - ET_p)^2}{\sum_1^n (ET_0 - ET)^2}$$

Where,

ET_0= Standard reference evapotranspiration given by FAO-56 PM equation.

ET_p = Model generated evapotranspiration.

ET = Average value of ET_0.

n = Total number of time steps in calibration period.

The value of NSE can be negative or positive, with a maximum absolute value of 1. A positive value indicate that the simulated values describe the trend of the measured data better than the mean of the observational values, while a negative value indicates that the corresponding model output is dis-similar to the behavior of the studied system. Obviously, a higher NSE value (close to 1) means a better fit of the predicted reference evapotranspiration to the observed one. Model simulations with negative NSE values are considered unacceptable. For best fit the other three criterions (performance indicators) should move close to zero. These criteria's are,

Sum of squared errors, $\text{SSE} = \sum_{i=1}^n (ET_0 - ET_p)^2$

Sum of squared log error, $\text{SLE} = \sum_{i=1}^n \{\log (ET_0) - \log (ET_p))\}^2$

Sum of absolute error, $\text{SAE} = \sum_{i=1}^n |(ET_0 - ET_p)|$

Rajesh Vijaykumar Kherde et al. / International Journal of Engineering and Technology (IJET)

Table II. Optimization of coefficients using Monte-Carlo Simulation (Range: β0 ~0.3- 2, β1~ 0.001- 0.4, β2 ~ 0.001-0.0035)

STATION	YEAR	β_0	β_1	β_2	K_P	NSE	SSE	SLE	SAE
DHABA	2002	1.4194	0.00493	0.00342	1.09624	0.7086	3.08342	0.12596	1.38827
	2003	1.31862	0.00300	0.00492	1.15671	0.90220	2.07359	0.14622	1.18799
	2004	1.49869	0.00301	0.00353	1.28166	0.79728	1.96951	0.09527	1.09259
	2005	1.32257	0.00300	0.00383	1.12072	0.72995	1.89998	0.06447	1.01474
	2006	1.32139	0.00310	0.00300	1.10622	0.72817	1.91253	0.06683	1.02576
	2007	1.34703	0.00308	0.00451	1.14097	0.73423	2.69189	0.09307	1.14721
	2008	1.27694	0.00306	0.00372	1.07495	0.74923	2.68809	0.11664	1.32650
HINGANGHAT	2002	1.49722	0.00456	0.00356	1.21138	0.65630	4.42394	0.08160	1.51567
	2003	1.49172	0.00480	0.00318	1.23204	0.68266	5.30104	0.08079	1.51538
	2004	1.48919	0.00300	0.00391	1.27308	0.79665	1.97557	0.09662	1.09790
	2005	1.49575	0.00472	0.00315	1.12157	0.74935	2.98818	0.06780	1.26541
	2006	1.48347	0.00498	0.00332	1.12378	0.74513	2.71817	0.09231	1.21479
	2007	1.24467	0.00310	0.00466	1.07705	0.67729	3.10379	0.17895	1.33329
	2008	1.04414	0.00301	0.00378	0.84666	0.50728	2.42750	0.14922	1.18784
SIRPUR	2002	1.11444	0.00307	0.00302	1.02321	0.68452	3.49088	0.25676	1.29324
	2003	1.19993	0.00306	0.00427	1.03372	0.79245	3.46713	0.15783	1.35759
	2004	1.28704	0.00302	0.00306	1.06873	0.56512	4.21631	0.14740	1.63698
	2005	1.35851	0.00301	0.00441	1.12796	0.88122	1.93426	0.11796	1.13358
	2006	1.23176	0.00348	0.00308	1.01694	0.65792	4.53437	0.14579	1.54312
	2007	1.23671	0.00364	0.00337	1.0116	0.65783	4.53554	0.147	1.54168
	2008	1.49903	0.00307	0.00381	1.30635	0.84052	3.29364	0.089737	1.465666
AVERAGE	AVG	1.34182	0.003515	0.003695	1.116747	0.6462	4.2681	0.1438	1.5182

Table III. Results of Monte-Carlo simulations on daily scale.

STATION	YEAR	K_P	NSE	SSE	SLE	SAE
DHABA	2002	1.0977	0.6951	3.2254	0.1193	1.4078
	2003	1.1042	0.8969	2.1843	0.1534	1.2264
	2004	1.1125	0.6784	3.1239	0.1521	1.4688
	2005	0.9095	0.6156	2.8063	0.1827	1.2185
	2006	1.1018	0.7254	1.9318	0.06872	1.0331
	2007	1.1226	0.7278	2.7565	0.1014	1.1762
	2008	1.1260	0.7429	2.7559	0.1094	1.3267
HINGANGHAT	2002	1.1299	0.5898	5.2792	0.07890	1.5351
	2003	1.1266	0.6198	6.3508	0.0834	1.5543
	2004	1.1125	0.6784	3.1239	0.1521	1.4688
	2005	1.0874	0.6949	3.6366	0.0662	1.2926
	2006	1.0003	0.6648	3.5748	0.1205	1.3195
	2007	1.1091	0.6507	3.3593	0.1646	1.4145
	2008	1.1232	0.6637	6.6801	0.2633	2.3654
SIRPUR	2002	1.1905	0.6892	5.0763	0.2129	1.6058
	2003	1.1262	0.7458	4.2460	0.1289	1.4489
	2004	1.1105	0.6782	4.2869	0.1405	1.6230

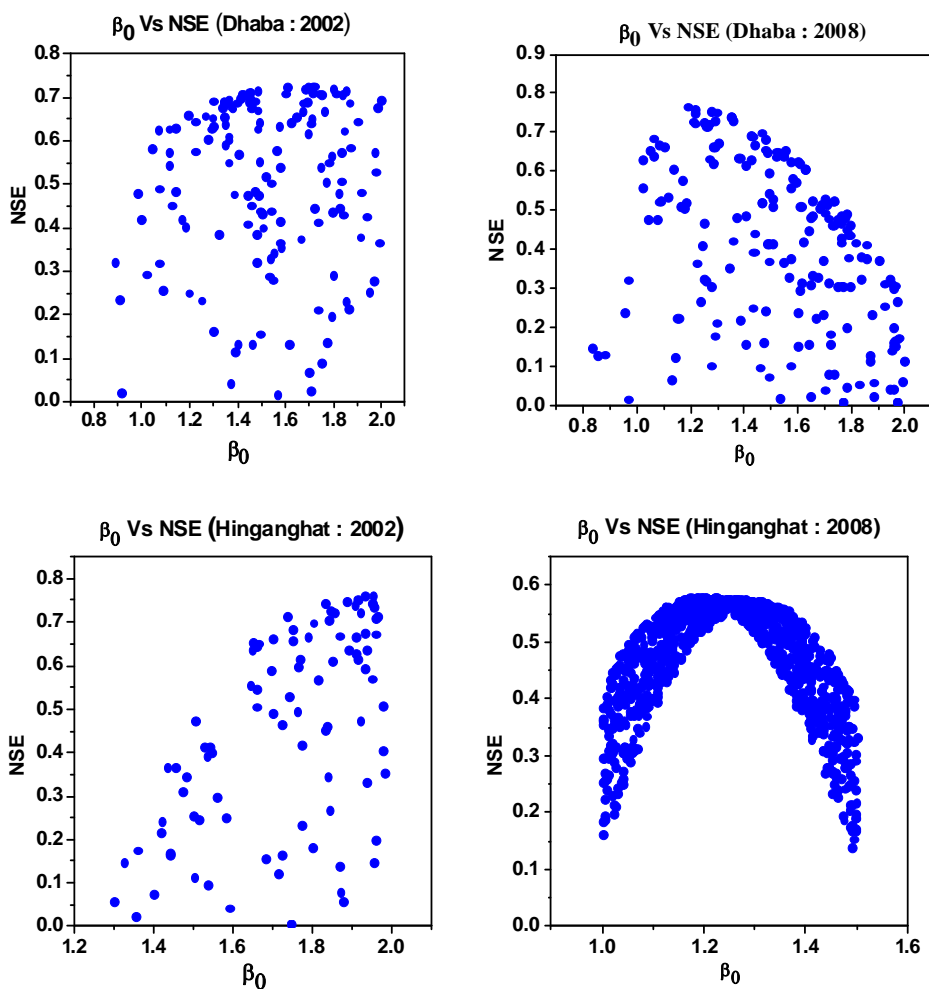|					
	2005	1.1444	0.6981	7.2994	0.2372	1.7907
	2006	1.0941	0.8715	2.0918	0.1382	1.1933
	2007	1.1383	0.6267	4.9618	0.1455	1.6682
	2008	1.1491	0.7158	5.8690	0.1601	2.0723

V.   THE DOTTY PLOTS

Monte Carlo simulations were carried out for each year's data of every station. The dotty plots represent various run of model for Monte Carlo realizations. They are basically scattering diagrams of parameter value verses some objective function value. While each dot represents one run of the model for different randomly chosen parameter value, they essentially represent projection of sample of points on goodness of fit response surface.

Monte-Carlo simulation results are plotted for positive values of Nash-Sutcliffe efficiency (NSE). The dotty plots of the year 2002 to 2008 drawn for all the three stations show randomness of the parameters β_0, β_1 and β_2. No trend however is observed, one parameter set seems to be optimum for every year of observation at each station giving highest NSE. These optimum parameter sets belong to different parameter spaces. The NSE for the entire observation period using average parameter set is 0.6462 and values of other performance indicators viz. SSE, SLE and SAE are 4.2681, 0.1438 and 1.5182 respectively.
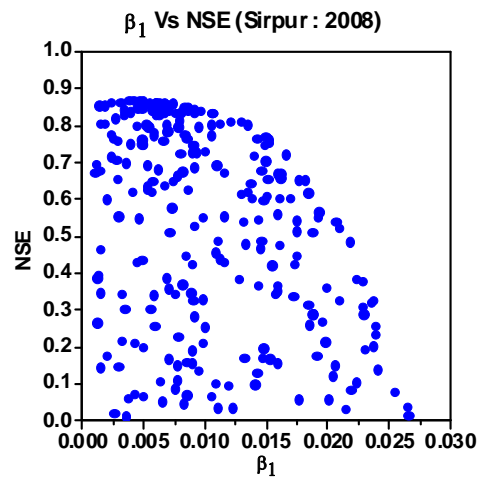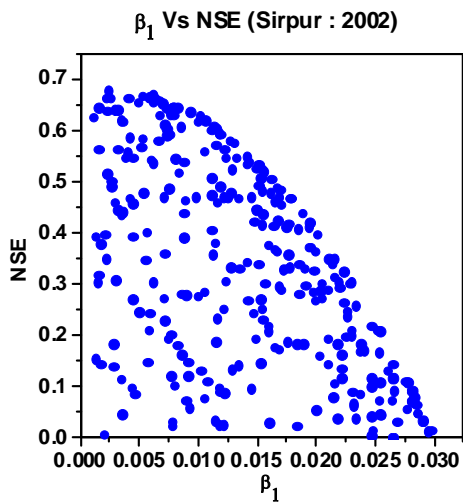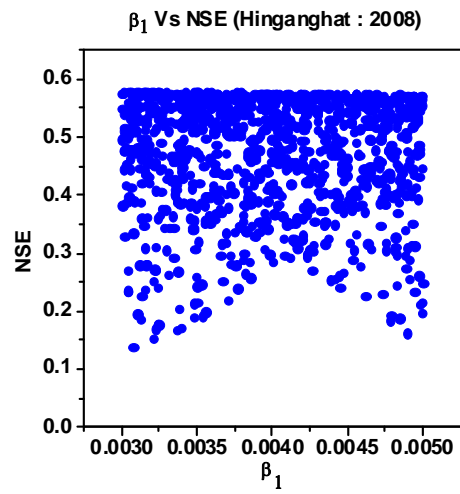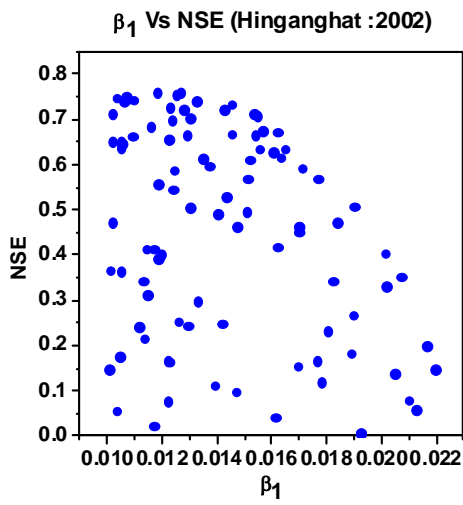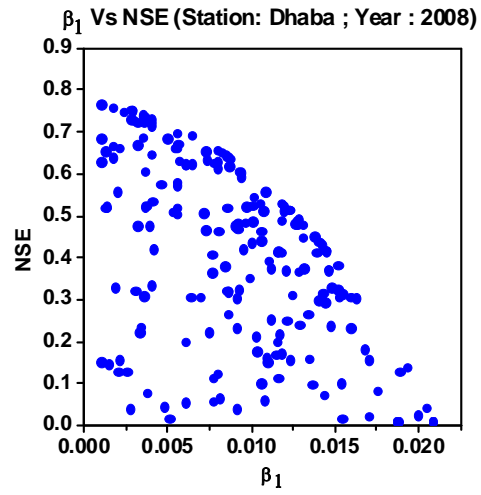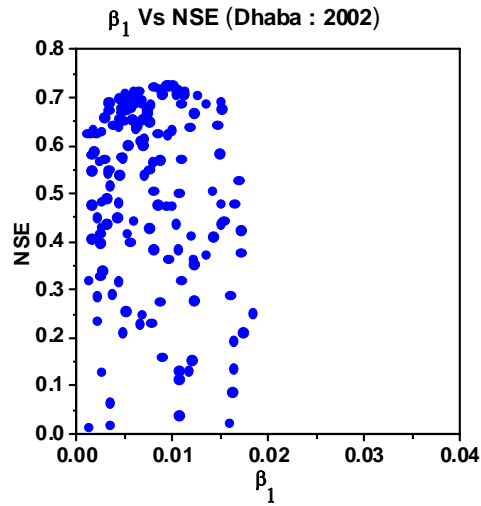
The final form of the equation is as follows:

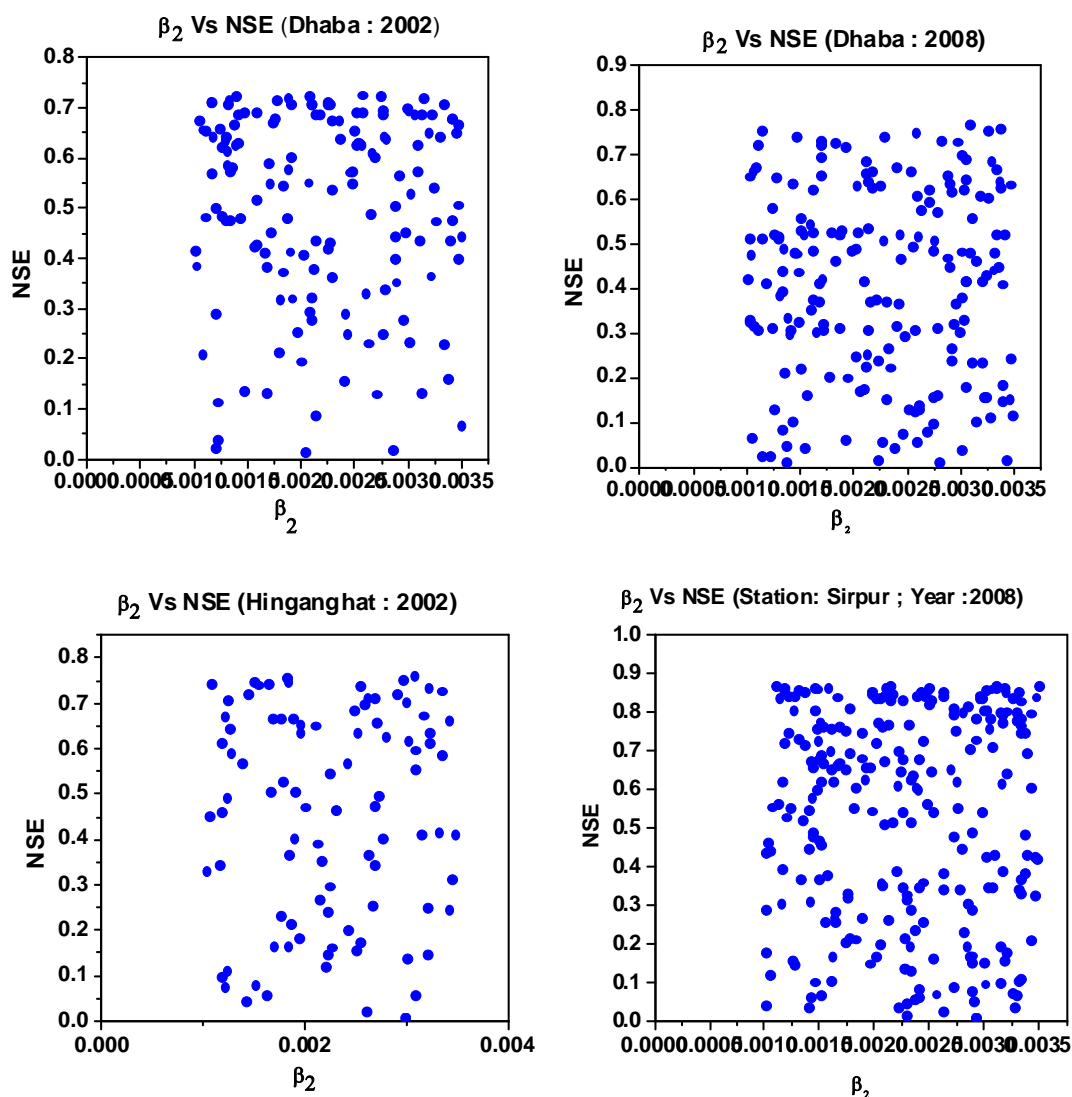**$K_p = 1.34182 + 0.003515\ RHmean + 0.003695\ U_2$**

The averaging of parameter values worked well except for some years of observation (Hinganghat 2008, Sirpur 2002, 2004 & 2005). This is due to data error, for rest of the year's NSE is acceptable if not exceptional.

(a)

(b)

Rajesh Vijaykumar Kherde et al. / International Journal of Engineering and Technology (IJET)

(C)

Figure 1. Dotty plots showing distribution of βi verses NSE.

VI. SENSITIVITY ANALYSIS

The primary objective of sensitivity analysis is to identify whether the perturbation of parameter significantly affect the model response i.e. the variable of interest. In case it is observed that impact of particular parameter is small, the relevant parameters may be replaced by constants or eliminated altogether. This strategy will not only help model construction but also model calibration on parameter estimation. In Pan evaporation modelling the main aim of carrying out sensitivity analysis is to investigate sensitivity of Kp to Relative humidity (RHmean) and wind velocity (U2).
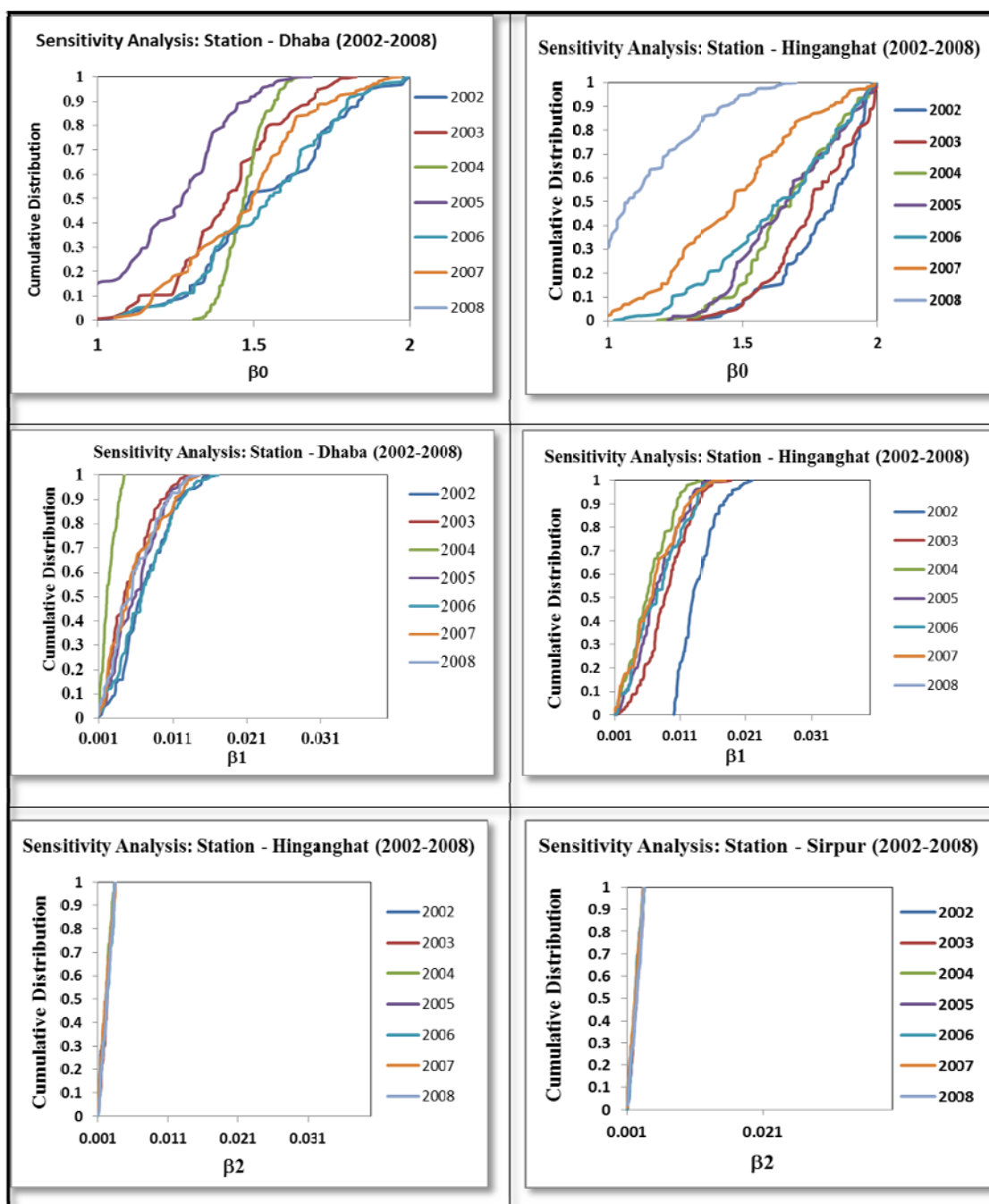
Figure 2: Parameter sensitivity Plots (2002-2008)

The parameter sensitivity plot shows the cumulative distributions of parameter values grouped according to the ranking of each Monte Carlo run for a given likelihood measure or predicted variable. Strong differences between the cumulative distributions for a given likelihood measure or variable indicates sensitivity to that parameter. Distributions plotting close together indicate a lack of sensitivity.

First, for each parameter, the Monte Carlo simulations are ranked and equally divided into ten bins based on the value of the objective function. Thus the first bin contained the best 10% of the simulations, the second bin the next best 10%, and so forth. Next, the values of the objective function in each bin are normalized so they ranged from 0 to 1. Finally, these normalized objective function values are plotted as a cumulative distribution function of the parameter value. Thus, for each panel in Figure 3 there are ten curves, each corresponding to a single bin. In general, an insensitive parameter will produce a straight one-to-one line whereas a sensitive model parameter will show differences in separation and form between the cumulative frequency distribution curves.

VII.CONCLUSIONS

The multivariate analysis was done to find the optimum values of βi, the coefficients of the evapotranspiration equation modeled using mean relative humidity and wind velocity. However the results of multivariate analysis did not yield acceptable results. Hence it can be concluded that multivariate analysis is not useful for optimization of models of pan evaporation studies. However it can used to find parameter spaces and hence the range of the dependent variables to be optimized using more advanced techniques. Monte Carlo simulation technique is used in this study to find the optimum parameter set for each year of observation. The Nash-Sutcliff efficiency was observed to vary between 65 to 90 percent which is acceptable. Other performance indicators (SSE, SLE & SAE) are in the acceptable range. The value of pan coefficient Kp on daily scale were observed to be greater than one as mentioned in the earlier literature. The averaging of the optimum parameter sets obtained for different years of analysis using MCS run worked well and the unique coefficients for the pan evaporation equation were obtained. Sensitivity analysis showed that Pan Evaporation equation is insensitive to the coefficient of wind velocity (β_2), However it is sensitive to the constant of the equation (β_0) and coefficient of relative humidity (β_1). The impression after conducting this study is that the pan evaporation equation can be modeled using relative humidity and wind velocity data based on the averaging of coefficients (β_1, β_2) and a constant of equation (β_0) of yearly analysis.

REFERENCES

[1] Allen, R. G, W.O. Pruitt, "FAO 24 reference evapotranspiration factors",Journal of Irrigation and Draining Engineering, 1991Sept, Doi no: 10.1061/(ASCE)/0733- 9437/117:5(758)/758-773.
[2] Conceicao, Marco AntônioFonseca,"Reference evapotranspiration based on class A pan evaporation"Scientia Agricola, 2002Jul, Doi no: 10.1590/59(3)/417-420.
[3] Frevert, D.K., W. Hiller., and B.C. Braaten,"Estimation of FAO evapotranspiration coefficients", Journal of Irrigation and Draining Engineering, 1983 June, Doi no: 10.1061/(ASCE)/0733-9437/109:2(265)/265-270.
[4] HamidrezaGolkarHamzee Yazd, JavadBasirat, MojtabaTavousi and Abbas Alvani, "Study on SEBAL Model and Suggesting a Modified Model for Measuring Evapotranspiration", Indian Journal of Science and Technology, 2015 Apr/Vol 8(S8)/191–195.
[5] Lee, K. and Cho, H., "Simple Method for Estimating Pan Coefficients: Conversion of Pan Evaporation to Reference Evapotranspiration" Journal of Irrigation and Draining Engineering, 2012 Jan, Doi no: 10.1061/(ASCE)/IR.1943-4774.0000379/98-103.
[6] McMahon, T. A. and Peel, M. C. and Lowe, L. and Srikanthan, R. and McVicar, T. R. "Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis", Hydrology and Earth System Sciences, 2013 Apr /17(4)/ 1331—1363.
[7] Modaberi, Hadi; Assari, Mostafa,"Estimation of reference Evapotranspiration, the best pan coefficient and rice plant coefficient using the pan evaporation data in rice growth period (Case study: Mordab plain region in Guilan province, Iran)",International Journal of Engineering Sciences, 2014 Sept /3(9)/75-84
[8] Snyder, R., "Equation for Evaporation Pan to Evapotranspiration Conversions." Journal of Irrigation and Draining Engineering, 1992 Nov /10.1061/(ASCE)/0733-9437/118:6(977)/977-980.
[9] Sriram A.V, Rashmi C.N, "A Model for Evapotranspiration by Fuzzy Least Squares Method", IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE), 2014 Jan/e-ISSN: 2278-1684/p-ISSN: 2320-334X/11(1 Ver. II)/ PP43-49.
[10] Tabari H, HosseinzadehTalaeeP,"Local calibration of the Hargreaves and Priestley–Taylor equations for estimating refer-ence evapotranspiration in arid and cold climates of Iran based on the Penman-Monteithmodel",Journal of Hydrologic Engineering, 2011 Oct/ 10.1061/(ASCE)/HE.1943-5584.0000366/837-845.
[11] Yi Li, Robert Horton, TushengRen, and Chunyan Chen, "Investigating Time-ScaleEffects on Reference Evapotranspiration from Epan Data in North China",Journal of Applied Meteorology and Climatology, 2010 May/49/10.1175/2009JAMC2130.1/867–878.
[12] G. Aschonitis, V. , Z. Antonopoulos, V. and M. Papamichail, D,"Evaluation of pan coefficient equations in a semi-arid Mediterranean environment using the ASCE-standardized Penman-Monteithmethod",Agricultural Sciences, 2012 Jan/ DOI: 10.4236/as.2012.310083/ 58-65.
[13] Eric W. Harmsen , Antonio González-Pérez and Amos Winte, Re-evaluation of pan evaporation coefficients at seven locations in Puerto Ricor. J. Agric. Univ., 2013 Oct/ doi: 10.1111/ejss.12058/P.R. 88(3-4):109-122.
[14] Chong-yuXu, Lebing Gong, Tong Jiang, Deliang Chen, V.P. Singh, "Analysis of spatial distribution and temporal trend of reference evapotranspiration and pan evaporation in Changjiang (Yangtze River) catchment", Journal of Hydrology, 2006 Jul/ 327(1–2), / 81-93.

AUTHOR PROFILE

Dr. Rajesh Kherde is Professor & HOD of civil engineering department at Sandip Institute of Technology and Research Centre, Nashik. He is doctorate in civil engineering with specialization in water resources management. His area of research interest is rainfall runoff modeling, modeling evapotranspiration, GIS based hydrological modeling and flood inundation mapping.

Sumana Sahoo is graduate in civil engineering. Currently working as junior Engineer at Transheight Consultants Pvt. Ltd. Her present area of work is Highway Design, Estimation and Detailed Project Reports for Hyderabad Karnataka Regional Development Board.