International Journal of Applied Engineering and Technology ISSN: 2277-212X (Online)
An Online International Journal Available at http://www.cibtech.org/jet.htm
2011 Vol. 1 (1) October-December, pp.98-105/N. Vivekanandan
Research Article

COMPARISON OF PWM ESTIMATORS OF WEIBULL DISTRIBUTIONS FOR LOW-FLOW FREQUENCY ANALYSIS

N. Vivekanandan
Central Water and Power Research Station, Pune 411024, Maharashtra, India
****Author for Correspondence***

ABSTRACT
Quantitative information on the low-flow regime of a stream is of utmost importance while making decisions on varied water resources management issues. The paper details a study on estimation of low-flow for different return periods for rivers Godavari at Pathagudam and Narmada at Mandleshwar sites using probability weighted moments of 2-parameter/ 3-parameter Weibull (WB2 and WB3) distributions. Kolmogorov-Smirnov is used for checking the adequacy of fitting of WB2 and WB3 to annual minimum d-day average flow data for different durations of 'd' such as 7-, 14- and 30-days. Diagnostic analysis involving correlation coefficient and root mean square error are used for selection of an appropriate distribution amongst WB2 and WB3 for estimation of low-flow. The analysis shows that WB3 distribution is more appropriate than WB2 for low-flow estimation for the rivers considered in the study. Low-flow frequency curves using WB2 and WB3 distributions are developed and presented in the paper.

Key Words: *Correlation coefficient, Kolmogorov-Smirnov, Low-flow, Mean square error, Weibull*

INTRODUCTION
Low-flow analysis is an important aspect of water quality management, reservoir storage design, determining minimum release policy and safe surface water withdrawals. Many studies have focused on flow duration analysis, low-flow frequency analysis, flow recession analysis and storage-yield analysis. There is, however, a need to classify clearly the definition of the low-flow and drought. Low-flow is a seasonal phenomenon and an integral component of flow regime of any river. Drought, on the other hand, is a natural event resulting from less than normal precipitation for an extended period of time (Nathan and McMahon, 1990). Thus, drought, which is a more general phenomenon, can be characterized by more factors than low-flow alone. Hydrological literature describes that there are many interlinking natural factors that contribute to low-flow which includes direct river withdrawals for human activity and artificial afforestation in the catchment (Vogel and Kroll, 1990).
Number of measures and indices used to characterize the low-flow, including mean annual runoff, mean daily flow, median flow, and absolute minimum flow. Among these, the low-flow frequency analysis is most commonly performed on the basis of a series of annual minimum d-day average flow for different duration of 'd' such as 7-, 14- and 30-days. An associated, annual event based, low-flow statistic Q(d,T) gives low-flow estimates, which is defined as the annual minimum d-day average flow that is expected to be occurred once in T year return period (Kernell, 1994). In most of the cases involving low-flow frequency analysis, the available flow data are insufficient to conduct an accurate analysis of an extreme low-flow event. Therefore, various statistical distributions are used to improve the accuracy of the low-flow frequency analysis. After finding the annual d-day minimum flow, a probability distribution should be selected to estimate the value of Q(d,T).
Research reports iterated that 2-parameter normal/ log-normal (N2 and LN2), Pearson Type-3 and log-Pearson Type-3, 2-parameter/ 3-parameter Weibull (WB2 and WB3) distributions are commonly available for low-flow frequency analysis (Durrans, 1996; O¨ no¨z and Bayazit, 2001). In low-flow analysis, a selected distribution must fit an annual minimum d-day flow and must have a finite lower limit of at least zero. Since WB2 and WB3 distributions satisfy these conditions, both distributions are considered in the present study for low-flow frequency analysis (Lee1 and Kim, 2008).
Number of methods like method of moments (MOM), maximum likelihood (MLM) and probability-weighted moments (PWM) are commonly used for low-flow modelling. Research reports indicated that there is no possibility of getting MLM estimates for the data set having small samples, if the number of parameters is large. On the other hand, MOM estimates are usually inferior in quality especially for distributions with three or more number of parameters because higher order moments

are more likely to be highly biased in relative small samples. Under these circumstances, PWM is considered to be more efficient method for determining the parameters of Weibull distributions (Sharad et al, 2007). For WB2, parameters are obtained by equating the first two moments of the distribution with the corresponding sample moments. For WB3, the parameters estimated by the way of setting of first three sample moments to the corresponding population moments.

This paper details a computer-aided procedure for evaluation of low-flow using Q(d,T) by adopting PWM of WB2 and WB3 distributions for rivers Godavari at Pathagudam and Narmada at Mandleshwar. The methodology can be used to evaluate the frequency and magnitude of the annual minimum d-day average flow at a gauged site for different return periods. The objective of the paper is to identify the suitable method amongst WB2 and WB3 for evaluation of low-flow using goodness-of-fit (GoF) and diagnostic analysis. GoF test involving Kolmogorov-Smirnov (KS) is used for checking the adequacy of fitting of WB2 and WB3 to the recorded annual minimum d-day average flow. Diagnostic analysis involving correlation coefficient (CC) and root mean square error (RMSE) are used for selection of an appropriate method for estimation of low-flow. The paper describes the methodology adopted in determining the PWM parameters of WB2 and WB3 distributions, GoF test and diagnostic analysis, and the results obtained thereof.

MATERIALS AND METHOD

The probability density function [f(q)] and cumulative distribution function [F(q)] of WB3 is given by:

$$f(q) = \left(\frac{b(q-m)^{b-1}}{a^{b}} \right) \mathrm{Exp}\left(-\left(\frac{q-m}{a} \right)^{b} \right); \; m \le q \le 0,\ a,b,m>0 \qquad (1)$$

$$F(q) = 1 - \mathrm{Exp}\left(-\left(\frac{q-m}{a} \right)^{b} \right) \qquad (2)$$

where a, b and m are scale, shape and location parameters of the distribution respectively (Cohn and Stedinger, 2001). The procedures involving determination of parameters of WB3 by PWM are given below:

$$a = \frac{a_0 - 2a_1}{\left(1-2^{-1/b}\right)\left(\Gamma\left(1/b\right)+1\right)} \qquad (3)$$

$$b = a_0 - \left(\Gamma\left(1/b\right)+1\right) \qquad (4)$$

$$m = 1/(7.859C + 2.9554C^2) \qquad (5)$$

where

$$m = \frac{a_0 - 2a_1}{a_0 - 3a_2} - \frac{\mathrm{Log}2}{\mathrm{Log}3} \quad \text{and} \quad a_r = \frac{1}{N}\sum_{i=1}^{N}\left[(N-i)C_r q_i / (N-1)C_r\right] \qquad (6)$$

If m=0 then WB3 reduces to WB2 for which the parameters can be obtained from Eqs. (3) and (4). The parameters are further used to estimate Q(d,T) for different return periods using Eq. (7) and is given by:

$$Q(d,T) = m - a\left[\mathrm{Log}\left(1-\left(1/T\right)\right)\right]^{1/b} \qquad (7)$$

Goodness-of-Fit Test

Kolmogorov-Smirnov (KS) is used for checking the adequacy of fitting of PWM of WB2 and WB3 to the recorded flow data. The KS statistic (K_C) is defined by:

$$K_C = \operatorname*{Max}_{i=1}^{N}\left(F_e(q_i) - F_D(q_i)\right) \quad \text{with} \quad F_e(q_i) = (i - 0.35)/N \tag{8}$$

where, $F_e(q_i)$ is the empirical cumulative distribution function (CDF) of q_i, $F_D(q_i)$ is the computed CDF of q_i (D'Agostino and Stephens, 1986). The rejection region of KS statistic at the desired significance level 'η' is $K_C \rangle K_{N,1-\eta}$.

If the computed values of KS of the distribution are less than that of theoretical value at the desired significance level 'η' then the selected distribution is accepted to be adequate than any distribution.

Diagnostic Analysis

Diagnostic analysis involving correlation coefficient (CC) and root mean square error (RMSE) are adopted for evaluating the applicability of an appropriate distribution for estimation of low-flow and are given by:

$$CC = \frac{\sum_{i=1}^{N}\left(Q_{oi}(d,T) - \overline{Q_{oi}(d,T)}\right)\left(Q_{ei}(d,T) - \overline{Q_{ei}(d,T)}\right)}{\sqrt{\left(\sum_{i=1}^{N}\left(Q_{oi}(d,T) - \overline{Q_{oi}(d,T)}\right)^2\right)\left(\sum_{i=1}^{N}\left(Q_{ei}(d,T) - \overline{Q_{ei}(d,T)}\right)^2\right)}} \tag{9}$$

$$RMSE = \left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{Q_{oi}(d,T) - Q_{ei}(d,T)}{Q_{oi}(d,T)}\right)^2\right)^{1/2} \tag{10}$$

where $Q_{oi}(d,T)$ and $Q_{ei}(d,T)$ are the observed and estimated low-flow respectively for i^{th} record for a given return period (T) and desired value of 'd'. $\overline{Q_{oi}(d,T)}$ and $\overline{Q_{ei}(d,T)}$ are the average value of observed and estimated low-flow respectively. The distribution having the good correlation and minimum RMSE is considered to be a good choice for low-flow modelling (Chen and Adams, 2006).

RESULTS AND DISCUSSIONS

Study Area and Data Used

An attempt has been made to estimate the low-flow using Q(d, T) for different return periods from 1.01 to 100-yr using annual minimum d-day average flow by adopting WB2 and WB3. Data in respect of rivers Godavari at Pathagudam and Narmada at Mandleshwar sites for the period 1975 to 2004 are used.

Estimation of Q(d, T) using WB2 and WB3 Distributions

By adopting the procedures, as described above, a computer program was developed and used to determine the annual minimum d-day average flow for different durations of 'd' such as 7-, 14- and 30-days. The program computes the PWM parameters of WB2 and WB3, low-flow estimates for different return periods and GoF test statistics. Table 1 gives the parameters of WB2 and WB3 determined by PWM for the series of annual minimum d-day average flow for different durations of 'd' such as 7-, 14- and 30-days for rivers Godavari at Pathagudam and Narmada at Mandleshwar sites. Tables 2 and 3 give the values of Q(d,T) for different duration of 'd' such as 7-, 14- and 30-days for different return periods from 1.01 to 100-yr obtained using PWM of WB2 and WB3 distributions for the rivers under study.

Table 1: PWM estimators of WB2 and WB3 distributions

Parameters	WB2			WB3		
	d=7	d=14	d=30	d=7	d=14	d=30
Godavari at Pathagudam						
a	11.099	12.340	14.303	5.920	8.475	10.738
b	3.062	3.422	3.316	1.474	2.193	2.362
m	-	-	-	4.564	3.585	3.317
Narmada at Mandleshwar						
a	39.837	41.692	43.642	36.172	37.845	37.171
b	2.765	2.925	2.986	2.466	2.608	2.468
m	-	-	-	3.373	3.572	5.993

Table 2: Low-flow estimates for different return periods using WB2 and WB3 distributions for river Godavari at Pathagudam site

Return period (yr)	Estimated low-flow (m^3/s)					
	WB2			WB3		
	d=7	d=14	d=30	d=7	d=14	d=30
2	9.9	11.1	12.8	9.2	10.8	12.5
5	6.8	8.0	9.1	6.7	7.9	9.0
10	5.3	6.4	7.3	5.9	6.6	7.5
20	4.2	5.2	5.8	5.4	5.8	6.4
50	3.1	4.0	4.4	5.0	5.0	5.4
100	2.5	3.2	3.6	4.6	4.8	4.9

Table 3: Low-flow estimates for different return periods using WB2 and WB3 distributions for river Narmada at Mandleshwar site

Return period (yr)	Estimated low-flow (m^3/s)					
	WB2			WB3		
	d=7	d=14	d=30	d=7	d=14	d=30
2	34.9	36.8	38.6	34.5	36.5	38.0
5	23.2	25.0	26.4	23.1	24.9	26.2
10	17.7	19.3	20.5	17.9	19.5	20.9
20	13.6	15.1	16.1	14.2	15.7	17.2
50	9.7	11.0	11.8	10.8	12.0	13.6
100	7.5	8.7	9.4	9.0	10.1	11.8

CDF and Low-flow Frequency Curves

Parameters of WB2 and WB3 were used to develop CDF curves for the sites under study and delineated in Figures 1 and 2. From CDF curves, it can be seen that the fitted curve using WB3 shows very good convergence for the entire set of data when compared with WB2. The estimated values of Q(d,T) for different durations of 'd' such as 7-, 14- and 30-days for different return periods from 1.01 to 100-yr were used to develop low-flow frequency curves and shown in Figures 1 and 2.

From Figures 1 and 2, it can be seen that the estimates of Q(d, T) for different durations of 'd' such as 7-, 14- and 30-days given by WB3 are consistently higher for return periods above 5-yr when compared with the corresponding values of WB2 for river Godavari at Pathagudam and Narmada at Mandleshwar.

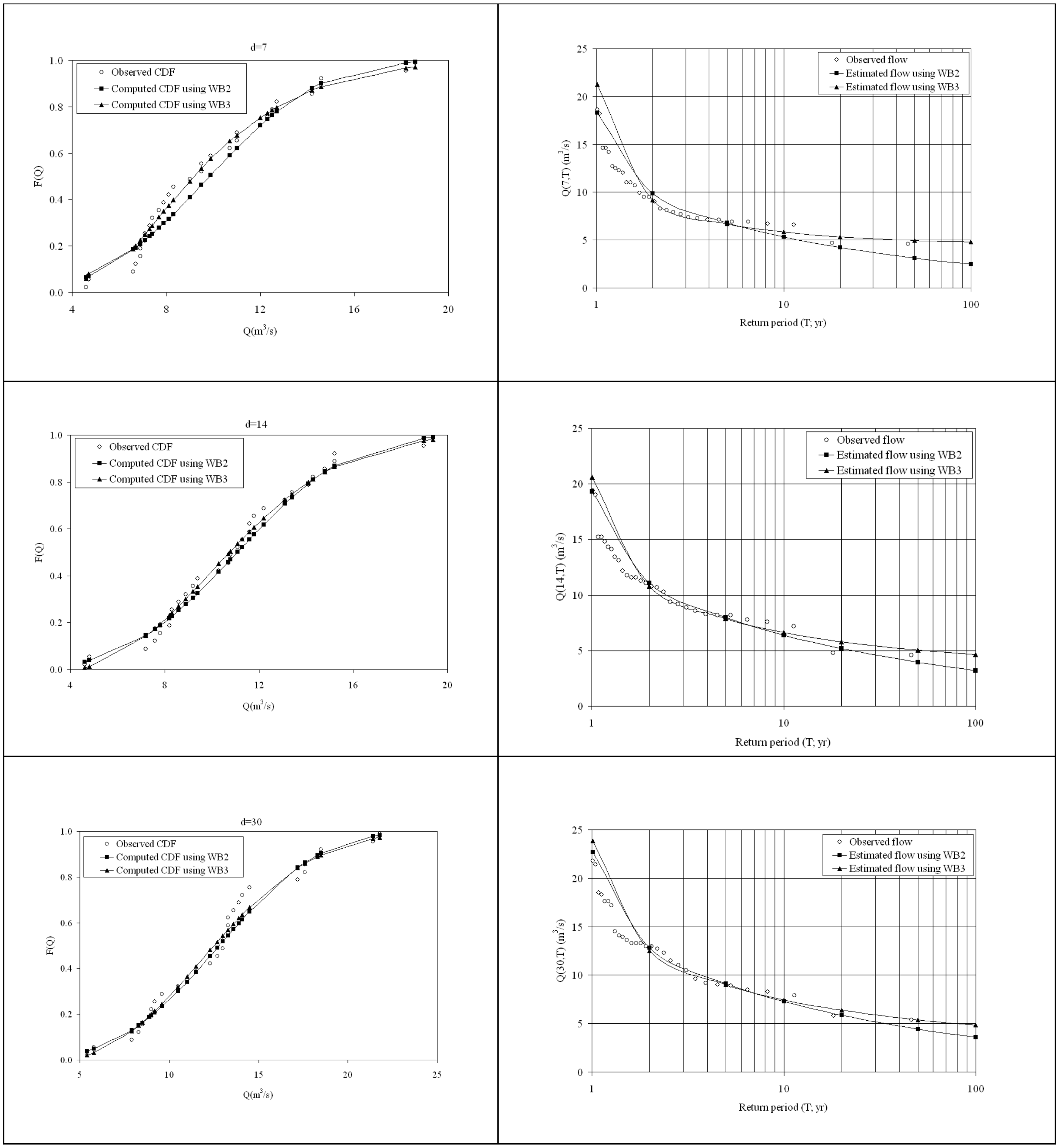

Figure 1: CDF and low-flow frequency curves for river Godavari at Pathagudam site

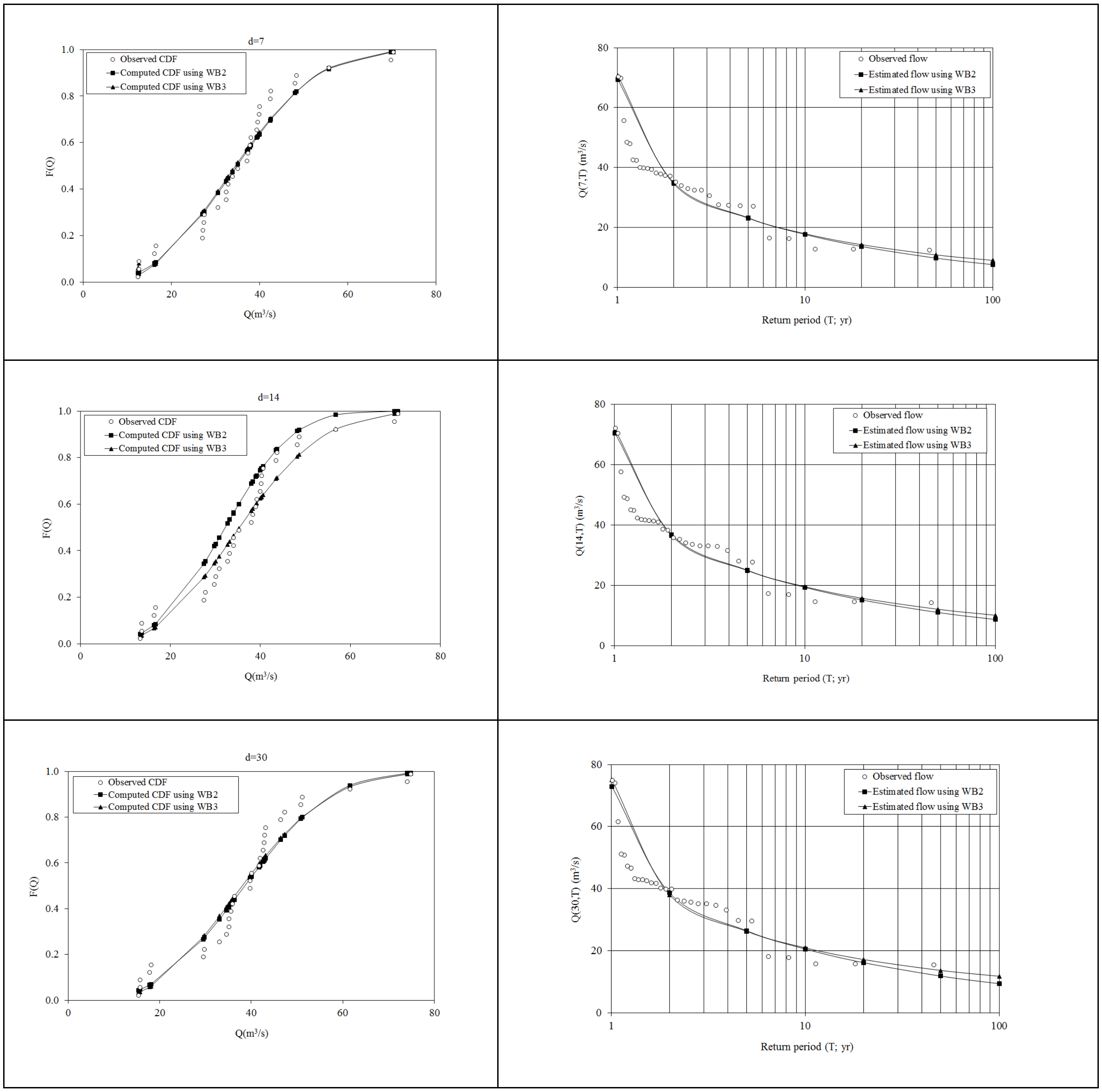

Figure 2: CDF and low-flow frequency curves for river Narmada at Mandleshwar site

Analysis Based on GoF Test

KS test statistics values for different durations of 'd' such as 7-, 14- and 30-days were computed by using Eq. (8) and given in Table 4.

Table 4: Computed values of KS test statistics using PWM of WB2 and WB3 distributions

Site	WB2			WB3		
	d=7	d=14	d=30	d=7	d=14	d=30
Pathagudam	0.130	0.091	0.119	0.121	0.077	0.099
Mandleshwar	0.122	0.110	0.134	0.116	0.104	0.122

From Table 4, it may be observed that the computed values of KS statistic obtained using PWM of WB2 and WB3 are less than the theoretical value of 0.242 at five percent level, and at this level, both distributions are considered to be significant for modelling of low-flow for different durations of 'd' such as 7-, 14- and 30-days for rivers Godavari and Narmada at the respective sites.

Diagnostic Analysis

Diagnostic analysis was employed for selection of a suitable distribution amongst WB2 and WB3 for estimation of Q(d,T). The values of CC and RMSE of WB2 and WB3 distributions were computed by using Eqs. (9) and (10), and given in Table 5.

Table 5: Indices of CC and RMSE using PWM of WB2 and WB3 distributions

Site	WB2						WB3					
	d=7		d=14		d=30		d=7		d=14		d=30	
	A	B	A	B	A	B	A	B	A	B	A	B
Pathagudam	0.975	0.799	0.986	0.586	0.987	0.683	0.987	0.605	0.989	0.531	0.988	0.668
Mandleshwar	0.972	3.342	0.970	3.375	0.966	3.777	0.972	3.303	0.972	3.334	0.968	3.693

(A: Correlation Coefficient; B: Root Mean Square Error)

From Table 5, it can be seen that the RMSE values of WB3 are to be minimum when compared with the corresponding indices of WB2 for different durations of 'd' such as 7-, 14- and 30-days though there is good correlation between the observed and estimated values using WB2 and WB3 distributions for the rivers under study. From the results of GoF tests and Diagnostic analysis, WB3 distribution adjudged as best method and recommended for estimation of Q(d,T) for different return periods from 1.01 to 100-yr for different durations of 'd' such as 7-, 14- and 30-days for rivers Godavari at Pathagudam and Narmada at Mandleshwar sites.

Conclusions

The paper described a procedure for estimating the parameters of WB2 and WB3 distributions by PWM. The paper also described the methodology involved in evaluating the frequency and magnitude of annual minimum d-day average low-flow at a gauged site for different recurrence intervals with 'd' taking the values 7-, 14-, and 30- days. The results indicated that WB3 provides better results though there is a good correlation between the observed and estimated low-flow events by WB2 and WB3. From the results of GoF and diagnostic analysis, it is suggested that WB3 distribution could be used for determination of PWM estimators for low-flow modelling for rivers Godavari and Narmada at the respective sites. The results presented in the paper could be beneficial to the stakeholders for making any decisions on varied water resources management issues at the respective sites of the rivers Godavari and Narmada while modelling the low-flow events. Cumulative frequency distribution and low-flow frequency curves are also developed and presented in the paper.

REFERENCES

Chen J and Adams B J (2006). Integration of artificial neural networks with nonceptual models in rainfall-runoff modelling, *Journal of Hydrology*, **318**, 232-249.

Cohn T A, Lane W L and Stedinger J R (2001). Confidence intervals for expected moments algorithm flood quantile estimates, *Water Resources Research*, **37**, 1695–1706.

D'Agostino B. Ralph and Stephens A. Michael (1986). Goodness-of-Fit Techniques, Marcel Dekkar, Inc., 270 Madison Avenue, New York 10016, USA, 145-148.

Durrans S R (1996). Low-flow analysis with a conditional Weibull tail model, Water Resources Research, **32**, 1749–1760.

Kernell G. Ries III (1994). Estimation of low-flow duration discharges in Massachusetts, United States Geological Survey Water-Supply Paper **2418**.

Lee1 K S and Kim S U (2008). Identification of uncertainty in low-flow frequency analysis using Bayesian MCMC method, *Journal of Hydrological Processes*, **22**, 1949–1964.

Nathan R J and McMahon T A (1990). Practical aspects of low-flow frequency analysis, Water Resources Research, **26**, 2135-2141.
O¨ no¨z B and Bayazit M (2001). Power distribution for low-flow, Journal of Hydrologic Engineering, **6**, 429–435.
Sharad K. Jain, Pushpendra K. Agarwal and Vijay P. Singh (2007). Hydrology and Water Resources in India, Water Science & Technology Library, M/s Springer Link Publications, Netherlands, ISSN No. 0921-092X, Vol. 57, Part 3, 513-559, 597-639, 641-699.
Vogel R M and Kroll C N (1990). Generalised low-flow frequency relationships for ungauged sites in Massachusetts, Water Resources Bulletin, **26**, 241-253.