

Research Article

A Simple Normal Approximation for Weibull Distribution with Application to Estimation of Upper Prediction Limit

H. V. Kulkarni and S. K. Powar

Department of Statistics, Shivaji University, Kolhapur 416 004, India

Correspondence should be addressed to H. V. Kulkarni, kulkarnih1@rediffmail.com

Received 30 May 2011; Accepted 24 September 2011

Academic Editor: A. Thavaneswaran

Copyright © 2011 H. V. Kulkarni and S. K. Powar. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a simple close-to-normal approximation to a Weibull random variable (r.v.) and consider the problem of estimation of upper prediction limit (UPL) that includes at least l out of m future observations from a Weibull distribution at each of r locations, based on the proposed approximation and the well-known Box-Cox normal approximation. A comparative study based on Monte Carlo simulations revealed that the normal approximation-based UPLs for Weibull distribution outperform those based on the existing generalized variable (GV) approach. The normal approximation-based UPLs have markedly larger coverage probabilities than GV approach, particularly for small unknown shape parameter where the distribution is highly skewed, and for small sample sizes which are commonly encountered in industrial applications. Results are illustrated with a real dataset for practitioners.

1. Introduction

Weibull distribution is widely used in reliability and survival analysis due to its flexible shape and ability to model a wide range of failure rates. It can be derived theoretically as a form of extreme value distribution, governing the time to occurrence of the “weakest link” of many competing failure processes. Its special case with shape parameter $b = 2$ is the Rayleigh distribution which is commonly used for modeling the magnitude of radial error when x and y coordinate errors are independent normal variables with zero mean and the same standard deviation while the case $b = 1$ corresponds to the widely used exponential distribution.

Let X follow a Weibull distribution with scale parameter a and shape parameter b . The pdf of X is given by

$$f_X(x; a, b) = \begin{cases} \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} \exp\left(-\left(\frac{x}{a}\right)^b\right), & x > 0, a > 0, b > 0, \\ 0, & \text{otherwise.} \end{cases} \quad (1.1)$$

If a random sample of size n is given, an important statistical problem is to construct a $(1-\alpha)100\%$ UPL such that l of m future sample values are below the UPL at each of r locations (or time periods). This problem for Weibull-distributed data is less attended in the literature. Such prediction limits are very much useful in monitoring and control problems during the operation of a production process in industries, particularly when the characteristic of interest is smaller the better type. UPLs are also useful in groundwater quality monitoring in the vicinity of hazardous waste management facilities (HWMF). For example, to monitor ground water quality, a series of m samples, that is, measurements of a pollutant like vinyl chloride from each of r monitoring wells in the vicinity of an HWMF are often compared with UPL based on a sample of n measurements obtained from one or more upgradient sampling locations of the facility. If at least l samples out of these m samples from each of r locations are less than the UPL, then the facility is considered to be within compliance. If this requirement is not met, then monitoring of contaminants like vinyl chloride is much more needful. Bhaumik and Gibbons [1] have discussed applications of UPLs in the fields like molecular genetics and industrial quality control. Davis and McNichols [2] obtained UPL assuming normality for the parent distribution. Bhaumik and Gibbons [1] and Krishnamoorthy et al. [3] proposed approximate methods for constructing UPL for Gamma distribution. Krishnamoorthy et al. [4] used the GV approach for constructing UPL for Weibull distribution.

In this paper, we propose a simple close to normal transformation for Weibull distribution when the shape parameter b is known. The transformation is based on two key features of the normal distribution, namely, symmetry and the tail behaviour. This transformation and the well-known Box-Cox transformation are used to obtain approximate UPLs for Weibull distribution when the shape parameter b is known. For unknown shape parameter b , replacing it by its maximum likelihood estimator (mle) gave equally good results. A simulation-based comparison of the proposed UPLs with existing ones revealed that the proposed UPLs outperform their competitors even for small sample sizes, and more prominently for small shape parameters which are frequently encountered in many real applications.

The article is organized as follows. Section 2 provides a brief review of the GV approach-based UPL developed by Krishnamoorthy et al. [4]. Section 3 describes the proposed normal approximation method and develops UPL for Weibull distribution based on the proposed and the Box-Cox transformations. Section 4 reports the comparison of the proposed UPLs with the GV approach-based UPL with respect to the simulated expected coverages and expected lengths. Section 5 illustrates the methods using a real dataset. Section 6 provides overall conclusions and recommendations.

2. GV Method for Obtaining UPL for Weibull (a, b) Distribution

Krishnamoorthy et al. [4] proposed $(1-\alpha)100\%$ UPL that includes at least l out of m future observations from a Weibull (a, b) distribution at each of r locations as

$$\exp\left(\hat{\eta} + u_{n,r,l,m}\hat{\beta}\right), \quad (2.1)$$

where $\hat{\eta} = \log(\hat{a})$ and $\hat{\beta} = 1/\hat{b}$. They computed value for $u_{n,r,l,m}$ using the following simulation study.

For the given values of n, r, l, m , and $1 - \alpha$, the following procedure is repeated N (say 100000) times.

- (I) n independent and identically distributed (i.i.d.) random variables $y_1^*, y_2^*, \dots, y_n^*$ are generated from the extreme-value (0,1) distribution and mles $\hat{\eta}^*, \hat{\beta}^*$ are computed.
- (II) For the given $\hat{\eta}^*, \hat{\beta}^*$ following procedure is repeated r times.
 - (i) m i.i.d. random variables $y_{i1}^*, y_{i2}^*, \dots, y_{im}^*$; $i = 1, 2, \dots, r$ are generated from the extreme-value (0,1) distribution.
 - (ii) l th order statistic based on these samples $y_{i(l)}^*$; $i = 1, 2, \dots, r$ are computed.
- (III) The quantities $y_{r,m,l}^* = \max\{y_{i(l)}^*; i = 1, 2, \dots, r\}$ and $u = (y_{r,m,l}^* - \hat{\eta}^*)/\hat{\beta}^*$ are computed.

Then $100(1 - \alpha)$ th percentile of the generated N values of u is the estimate of $u_{n,r,m,l}$. Note that the distribution of the pivotal quantity based on which UPL is developed does not depend on any unknown parameters, thus it is an exact method.

3. The Proposed Close-To-Normal Power Transformation-Based UPL

3.1. The Proposed Close-To-Normal Power Transformation

The proposed transformation is based on the two key features governing normality, namely, the symmetry and tail behaviour of the normal distribution.

Let X follow a two parameter Weibull (a, b) distribution and the shape parameter b is known. We consider a transformation $Y = X^p$ for the r.v. X where the power p is chosen so that the distribution of the transformed variable Y has very small deviation from symmetry and simultaneously has tail behaviour very close to that of the normal distribution with the same mean and variance. Straightforward calculations show that the skewness of the distribution of Y is given by

$$\gamma(p, b) = \gamma(\theta) = \frac{\Gamma(1 + (3\theta)) - 3\Gamma(1 + (\theta))\Gamma(1 + (2\theta)) + 2(\Gamma(1 + (\theta)))^3}{(\Gamma(1 + (2\theta)) - (\Gamma(1 + (\theta)))^2)^{3/2}}, \quad (3.1)$$

which is a function of the ratio $\theta = p/b$, and does not depend on the scale parameter a . Treating $\gamma(p, b)$ as a function of θ , a solution for $\gamma(\theta) = 0$ is $\theta_1 = 0.2776$, for which the distribution of the variable $Y = X^p$ where $p = b\theta$, is exactly symmetric. To achieve control over the tail behaviour, it is noted that the k th central moment of the transformed r.v. Y is $E(Y^k) = E(X^{kp}) = a^{kp}\Gamma(1 + k\theta)$ leading to the mean and standard deviation of Y given by

$$\mu_y = a^p\Gamma(1 + \theta), \quad \sigma_y = a^p \left(\Gamma(1 + 2\theta) - (\Gamma(1 + \theta))^2 \right)^{0.5}. \quad (3.2)$$

Table 1

α	0.010	0.025	0.050	0.095	0.975	0.990
$c(\alpha)$	-0.0283	-0.0063	0.0044	0.0009	0.0072	0.0170

Furthermore, the α th quantile of a normal distribution with mean μ_y and standard deviation σ_y is given by

$$\xi_\alpha = \mu_y + \sigma_y Z_\alpha = a^p \left(\Gamma(1 + \theta) + \left(\Gamma(1 + 2\theta) - (\Gamma(1 + \theta))^2 \right)^{0.5} Z_\alpha \right). \quad (3.3)$$

Similarly, if x_α is the α th quantile of the Weibull (a, b) distribution that is, $x_\alpha = a(-\log(1-\alpha))^{1/b}$, it easily follows that the α th quantile y_α of the distribution of Y is given by

$$y_\alpha = x_\alpha^p = a^p (-\log(1 - \alpha))^\theta. \quad (3.4)$$

To make the tail behavior of the distribution of Y very close to that of the normal distribution with same mean and standard deviation as that of Y , we solve the equation $\xi_\alpha - y_\alpha = 0$ for the commonly used choice of $\alpha = 0.025$ and $\alpha = 0.975$ for a two-sided interval leading to the solutions respectively $\theta_2 = 0.2698$ and $\theta_3 = 0.2994$. To control the symmetry and tail behaviour of the distribution of transformed r.v. Y simultaneously close to the normal distribution, we suggest taking $p = \theta b$ where $\theta = (\theta_1 + \theta_2 + \theta_3)/3 = 0.2823$ as the power of a Weibull r.v. X . From (3.3) and (3.4) it follows that for this choice of θ , the difference between the two quantiles ξ_α and y_α is given by $(a^b)^{0.2823} c(\alpha)$ where $c(\alpha) = (\Gamma(1 + \theta) + (\Gamma(1 + 2\theta) - (\Gamma(1 + \theta))^2)^{0.5} Z_\alpha - (-\log(1 - \alpha))^\theta$ is a constant depending on α . The values of $c(\alpha)$ for various commonly used choices of α are given in Table 1.

We note that the constant $c(\alpha)$ is considerably small for commonly used level of significance $\alpha = 0.05$, and further numerical study revealed that the accuracy of the proposed transformation is very good for small values of b (say $b < 4$) and for small to moderate values of a (say $a < 100$), which covers a reasonable subset of the parameter space and commonly encountered real situations. We recall that the choice of p is uniform for all $a > 0$ since γ is free from a . When b is unknown, we take $p = \theta \hat{b}$, replacing b by its mle \hat{b} . In the sequel we refer to this transformation as the close to normal power transformation (CNPT).

3.2. UPL Based on the CNPT

Let Y_1, Y_2, \dots, Y_n be a random sample of size n from a normal distribution with mean μ and standard deviation σ . Let \bar{Y} and S_y be the sample mean and sample standard deviation. Then Davis and McNichols [2] suggested UPL that includes at least l out of m future observations from the same normal distribution at each of r locations as

$$\bar{Y} + k_u S_y, \quad (3.5)$$

where the value of k_u for selected values of (n, r, l, m) and level α is the solution to the following equation

$$\int_0^1 \frac{N_t(\sqrt{nk_u}; n-1, \sqrt{n}\Phi^{-1}(x))r(I(x; l, m+1-l))^{r-1}x^{l-1}(1-x)^{m-l}}{\beta(l, m+1-l)} dx = \alpha, \quad (3.6)$$

where $N_t(x; v, \delta)$ is the cumulative distribution function (Cdf) of noncentral t r.v. with v df and noncentrality parameter δ , $\Phi^{-1}(x)$ is inverse Cdf of the standard normal distribution at x , $\beta(m, n)$ is the usual beta function and $I(x; m, n)$ is the Cdf of a beta distribution with parameters m and n .

Let X_1, X_2, \dots, X_n be a random sample of size n from a Weibull (a, b) distribution. Let U be the normal based UPL obtained using (3.5), based on $Y_i = X_i^p$; where $p = 0.2823b$; $i = 1, 2, \dots, n$. Then $(U)^{1/p}$ is the proposed UPL that is expected to include at least l out of m future observations from the Weibull (a, b) distribution at each of r locations for known shape parameter b , with probability $1 - \alpha$. When the shape parameter b is unknown, we suggest to replace it by its mle \hat{b} and the proposed UPL is $(U)^{1/\hat{p}}$ where $\hat{p} = 0.2823\hat{b}$. It is noted that \hat{p} is a consistent estimator for p . Hence for large samples $(U)^{1/\hat{p}}$ is expected to be close to $(U)^{1/p}$. Small sample behavior of $(U)^{1/\hat{p}}$ is studied through simulation.

3.3. Box-Cox Transformation and Kullback-Leibler Information-(BCKL-) Based UPL

Hernández and Johnson [5] proposed the transformation

$$Y = \begin{cases} \frac{X^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \log(X) & \text{if } \lambda = 0, \end{cases} \quad (3.7)$$

to Weibull r.v. X for approximating distribution of Y to normal and used a solution $\lambda = 0.2654b$ for known shape parameter b that minimizes the Kullback-Leibler information between the distribution of Y and the normal distribution with the same mean and variance as that of Y . This transformation was used by Yang et al. [6] for obtaining prediction interval for a single future observation from Weibull (a, b) distribution. Using this transformation for the UPL problem under consideration, a $(1 - \alpha)100\%$ UPL for Weibull distribution is $(1 + \lambda U)^{1/\lambda}$ where U is the normal based UPL obtained using (3.5). As before an unknown value of b will be replaced by its mle \hat{b} . This also enjoys the large sample properties mentioned above and its small sample behavior is studied through simulation.

4. Comparison

In this section we compare the performance of above two proposed UPLs with the UPL based on GV method-based on a simulation study with respect to expected lengths and expected coverages of the UPLs.

For fixed values of the parameters a , b , and sample size n , we generate n random numbers X_i ; $i = 1, 2, \dots, n$ from the Weibull (a, b) distribution, and set $Y_{1i} = X_i^p$,

Table 2: Simulated expected lengths of 95% UPL that contain at least l of m future observations at each of r locations from Weibull (a, b) distribution for sample sizes $n = 6, 10, 20$ using GV, CNPT, and BCKL methods*.

	a	b	$n = 6$			$n = 10$			$n = 20$			
			GV	CNPT	BCKL	GV	CNPT	BCKL	GV	CNPT	BCKL	
$r = 4,$ $l = 2,$ $m = 6$	0.1	0.5	0.38	0.37	0.37	0.25	0.25	0.25	0.19	0.18	0.18	
		1	0.18	0.18	0.18	0.15	0.15	0.15	0.13	0.13	0.13	
		3	0.12	0.12	0.12	0.11	0.11	0.11	0.11	0.11	0.11	
	1	0.5	3.80	3.76	3.77	2.50	2.47	2.46	1.85	1.79	1.78	
		1	1.79	1.78	1.78	1.51	1.50	1.49	1.33	1.31	1.30	
		3	1.19	1.19	1.19	1.13	1.13	1.13	1.09	1.09	1.09	
	5	0.5	19.10	18.63	18.71	12.53	12.31	12.28	9.25	9.00	8.93	
		1	9.00	8.90	8.91	7.55	7.48	7.46	6.64	6.54	6.51	
		3	5.96	5.94	5.94	5.67	5.65	5.65	5.46	5.44	5.43	
	50	0.5	190.16	186.02	186.79	124.82	123.47	123.08	92.75	89.95	89.25	
		1	90.03	88.84	89.00	75.45	74.85	74.73	66.36	65.44	65.18	
		3	59.64	59.40	59.44	56.70	56.54	56.51	54.62	54.39	54.32	
	100	0.5	383.79	374.67	376.24	251.24	247.09	246.32	184.91	180.11	178.70	
		1	179.98	178.41	178.72	151.07	149.92	149.66	132.80	130.95	130.43	
		3	119.27	118.77	118.83	113.40	113.11	113.04	109.24	108.74	108.59	
	$r = 8,$ $l = 1,$ $m = 5$	0.1	0.5	0.27	0.27	0.26	0.18	0.18	0.18	0.14	0.13	0.13
			1	0.15	0.15	0.15	0.13	0.13	0.13	0.11	0.11	0.11
			3	0.11	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.10
1		0.5	2.66	2.64	2.63	1.82	1.79	1.78	1.38	1.34	1.32	
		1	1.51	1.50	1.50	1.29	1.27	1.27	1.14	1.13	1.12	
		3	1.13	1.12	1.12	1.08	1.07	1.07	1.04	1.03	1.03	
5		0.5	13.34	13.20	13.16	9.11	8.96	8.88	6.91	6.68	6.60	
		1	7.55	7.50	7.49	6.43	6.39	6.36	5.74	5.63	5.60	
		3	5.63	5.61	5.61	5.38	5.36	5.36	5.20	5.17	5.16	
50		0.5	133.95	132.56	130.91	91.13	89.64	88.90	69.09	66.84	66.04	
		1	75.44	75.21	74.90	64.36	63.77	63.47	57.36	56.35	56.00	
		3	56.31	56.16	56.11	53.78	53.59	53.56	51.98	51.72	51.63	
100		0.5	266.35	263.65	264.21	182.16	178.70	177.11	138.26	133.44	132.09	
		1	151.03	149.96	149.92	128.62	127.63	127.06	114.74	112.72	112.11	
		3	112.63	112.38	112.23	107.49	107.24	107.02	104.02	103.40	103.22	
$r = 16,$ $l = 2,$ $m = 5$		0.1	0.5	1.65	1.41	1.46	0.75	0.71	0.73	0.49	0.47	0.47
			1	0.36	0.33	0.33	0.26	0.25	0.25	0.22	0.21	0.21
			3	0.15	0.14	0.15	0.14	0.13	0.13	0.13	0.13	0.13
	1	0.5	16.54	14.00	14.60	7.55	7.13	7.26	4.87	4.64	4.67	
		1	3.56	3.28	3.34	2.59	2.51	2.54	2.15	2.10	2.11	
		3	1.48	1.44	1.45	1.35	1.34	1.34	1.28	1.27	1.27	
	5	0.5	82.85	70.44	73.11	37.77	35.69	36.25	24.36	23.18	23.40	
		1	17.79	16.52	16.75	12.93	12.54	12.67	10.73	10.47	10.53	
		3	7.40	7.22	7.26	6.77	6.70	6.73	6.41	6.36	6.37	
	50	0.5	830.35	710.92	733.32	376.12	355.27	362.96	242.46	232.30	234.39	
		1	177.94	164.84	166.79	129.32	125.54	127.11	107.39	104.72	105.50	
		3	74.12	72.18	72.67	67.71	67.05	67.21	64.11	63.63	63.68	

Table 2: Continued.

a	b	$n = 6$			$n = 10$			$n = 20$		
		GV	CNPT	BCKL	GV	CNPT	BCKL	GV	CNPT	BCKL
100	0.5	1650.20	1408.03	1456.67	754.87	713.90	725.43	485.52	464.88	467.82
	1	356.26	328.69	334.60	258.73	251.05	252.98	214.54	209.88	210.49
	3	148.20	144.36	145.19	135.32	134.11	134.56	128.23	127.21	127.43

* Results are obtained assuming unknown scale and shape parameters a, b .

$Y_{2i} = X_i^\lambda - 1/\lambda$; $i = 1, 2, \dots, n$ where $p = 0.2823b$ and $\lambda = 0.2654b$. Normal based UPLs, U_1 and U_2 are obtained using (3.5) based on the transformed samples Y_{1i} and Y_{2i} , $i = 1, 2, \dots, n$ respectively. Then $(U_1)^{1/p}$ and $(U_2)^{1/\lambda}$ are the UPLs based on the proposed CNPT and the BCKL transformation for Weibull (a, b) distribution with known shape parameter b . When the shape parameter b is unknown, we suggest to replace it by its mle \hat{b} . For the same sample X_1, X_2, \dots, X_n , the GV method based UPL is obtained using (2.1) and the procedure described in Section 2. Next we generate r sets of m random numbers say X_{ij} ; $i = 1, 2, \dots, r$; $j = 1, 2, \dots, m$ from Weibull (a, b) distribution, and set $X^* = \max\{X_{1(l)}, X_{2(l)}, \dots, X_{r(l)}\}$ where $X_{i(l)}$ is the l th order statistic from X_{ij} for $i = 1, 2, \dots, r$. This procedure is repeated 100000 times. Then the proportions of events $X^* < U_1^{1/p}$, $X^* < U_2^{1/\lambda}$ and $X^* < \exp(\hat{\eta} + u_{n,r,l,m}\hat{\beta})$, in these 100000 repetitions are the simulated coverage probabilities of the UPLs based on normal approximation, BCKL transformation, and GV method, respectively. Average of 100000 UPLs based on each of the three approaches discussed above are the simulated expected lengths of the corresponding UPLs. The simulated expected lengths and expected coverages for $n = 6, 10, 20$, $a = 0.1, 1, 5, 50, 100$, and $b = 0.5, 1, 3$ are reported in Tables 2 and 3 respectively. The combinations $(r = 4, l = 2, m = 6)$, $(r = 8, l = 1, m = 5)$, and $(r = 16, l = 2, m = 5)$ are chosen. The values of k_u for these combinations are computed using (3.6) for $\alpha = 0.05$.

4.1. Results of the Simulation Study

Following prominent facts are clearly visible from Table 2.

- (1) CNPT-based UPLs have uniformly excellent coverage probabilities even for small sample sizes as small as $n = 6$, for all $a > 0, b > 0$ and for all examined combinations of r, l , and m . The coverages are uniformly a little larger than those based on GV method, and the expected lengths are a little shorter than the same.
- (2) As mentioned in Section 2, GV method is exact and this fact is reflected in simulation study since its coverages are very close to the nominal coverage probability.
- (3) BCKL transformation-based UPLs have close to nominal coverage probabilities.

Based on these observations we recommend the proposed CNPT- and BCKL transformation-based UPLs that include at least l out of m future observations from a Weibull (a, b) distribution at each of r locations.

5. Illustrative Example

Nowadays vinyl chloride is one of the fifty most produced chemicals in the world. Its production almost doubled in the last 20 years and currently estimated to be about 27

Table 3: Percentage simulated coverage probabilities of 95% UPL that contain at least l of m future observations at each of r locations from Weibull (a, b) distribution for sample sizes $n = 6, 10, 20$ using GV, CNPT, and BCKL methods*.

	a	b	$n = 6$			$n = 10$			$n = 20$			
			GV	CNPT	BCKL	GV	CNPT	BCKL	GV	CNPT	BCKL	
$r = 4,$ $l = 2,$ $m = 6$	0.1	0.5	94.89	96.10	94.72	95.04	96.39	94.95	94.89	95.91	94.80	
		1	95.01	96.23	94.87	94.95	96.03	94.53	94.90	95.79	95.14	
		3	94.92	95.77	96.04	95.12	95.05	94.68	94.95	95.92	94.71	
	1	0.5	95.00	96.09	94.75	95.07	96.43	94.94	94.95	95.61	94.47	
		1	94.97	96.01	94.57	95.09	96.33	94.82	94.93	95.55	94.94	
		3	94.86	96.66	94.44	95.01	95.71	94.41	94.87	95.38	95.37	
	5	0.5	94.94	96.36	94.88	94.92	95.73	94.53	95.02	95.42	94.60	
		1	94.94	96.13	94.79	95.00	96.41	94.91	94.86	95.63	94.61	
		3	95.01	97.01	94.69	94.96	95.58	94.68	94.86	95.23	94.57	
	50	0.5	94.86	96.76	95.76	94.91	95.61	95.02	95.07	95.84	94.88	
		1	95.01	96.09	95.24	94.94	95.64	94.49	94.86	95.44	95.43	
		3	94.93	96.46	95.71	95.08	95.76	94.83	94.77	96.17	96.05	
	100	0.5	95.04	95.96	95.30	94.95	94.75	94.81	95.06	96.69	95.06	
		1	94.96	96.32	95.97	95.03	95.06	94.44	95.03	96.45	95.54	
		3	95.09	96.83	96.46	95.05	95.07	94.99	94.87	97.03	95.38	
	$r = 8,$ $l = 1,$ $m = 5$	0.1	0.5	95.08	96.14	94.68	94.89	96.39	94.90	95.09	95.76	94.92
			1	94.99	95.91	94.48	94.87	97.12	95.60	94.92	94.91	94.58
			3	94.92	95.29	94.44	94.90	94.98	94.87	94.89	95.72	95.62
1		0.5	94.92	96.52	95.08	94.98	96.31	94.79	94.97	95.85	94.48	
		1	95.01	96.02	94.65	95.15	96.92	95.40	95.03	95.91	94.76	
		3	95.03	95.24	94.47	95.09	94.63	94.52	94.93	96.60	95.99	
5		0.5	94.98	96.34	94.88	94.95	96.35	94.88	94.99	95.84	94.57	
		1	95.14	96.40	95.00	94.96	96.37	94.83	95.10	95.55	94.42	
		3	95.03	94.47	94.52	95.06	95.37	94.91	94.95	95.58	94.61	
50		0.5	95.02	96.13	94.52	94.96	95.19	95.10	94.96	95.97	94.78	
		1	94.99	95.22	94.61	94.94	95.52	94.70	95.10	95.47	94.61	
		3	95.10	95.31	94.50	94.95	95.27	94.68	95.16	95.13	94.58	
100		0.5	95.09	95.23	94.72	94.86	95.56	94.55	95.13	95.55	94.45	
		1	94.98	95.17	94.68	94.92	94.90	94.64	95.15	95.84	94.60	
		3	95.04	95.41	94.73	95.04	95.77	95.05	94.93	95.99	94.46	
$r = 16,$ $l = 2,$ $m = 5$		0.1	0.5	94.95	95.73	94.51	95.18	96.34	94.97	95.07	95.99	94.73
			1	94.92	95.41	94.19	95.01	96.31	95.05	95.18	95.76	94.72
			3	95.03	95.51	94.21	95.13	96.46	94.85	95.01	95.66	94.53
	1	0.5	94.90	95.36	94.16	95.04	96.00	94.68	95.20	95.89	94.54	
		1	94.94	95.66	94.49	95.16	96.11	94.64	95.09	95.53	94.61	
		3	94.93	95.68	94.86	95.07	95.87	94.41	95.08	95.18	94.64	
	5	0.5	95.02	95.57	94.29	95.18	95.64	94.42	95.02	95.23	94.55	
		1	94.96	95.01	94.95	95.05	95.99	94.75	95.09	95.46	94.66	
		3	94.84	94.71	94.06	95.01	95.21	94.14	95.18	95.49	94.59	
	50	0.5	95.00	95.11	94.30	94.94	95.68	94.82	95.13	95.52	94.68	
		1	94.97	95.91	94.50	94.85	95.86	94.55	95.13	95.07	94.54	
		3	94.86	95.86	94.81	95.19	95.78	95.25	95.13	95.62	94.64	

Table 3: Continued.

<i>a</i>	<i>b</i>	<i>n</i> = 6			<i>n</i> = 10			<i>n</i> = 20		
		GV	CNPT	BCKL	GV	CNPT	BCKL	GV	CNPT	BCKL
100	0.5	94.89	95.53	94.60	94.97	95.41	94.69	95.09	95.27	94.50
	1	94.95	95.43	94.37	95.02	95.57	94.52	95.05	95.84	94.47
	3	95.06	95.93	94.41	95.09	95.62	94.85	95.12	95.93	94.60

* Results are obtained assuming unknown scale and shape parameters *a*, *b*

Table 4: 95% Upper prediction limits for the vinyl chloride data.

<i>l</i>	<i>m</i>	<i>r</i>	<i>k_u</i>	GV	CNPT	BCKL
1	2	10	1.577	5.483	5.298	5.336
2	2	10	3.022	14.066	13.371	13.801
1	3	10	1.033	3.618	3.469	3.466
2	3	10	1.879	6.797	6.566	6.646
3	3	10	3.177	15.149	14.574	15.084

million tons/year worldwide. A high concentration of vinyl chloride in water can cause cancer and and liver damage. Therefore being toxic and carcinogenic to humans, more attention has to be given to vinyl chloride as a groundwater contaminant. In this section we illustrate the methods discussed in Sections 2 and 3 with a real dataset.

The data used here are vinyl chloride concentrations collected from clean upgradient monitoring wells. Krishnamoorthy et al. [4] showed an excellent fit of these data to a Weibull distribution. We computed various Weibull UPLs and compared them with those given in Krishnamoorthy et al. [4]. The a dataset representing the vinyl chloride concentration in micrograms per liter of water ($\mu\text{g}/L$), that is, number of parts per billion (ppb), from clean upgradient monitoring wells is 5.1, 2.4, 0.4, 0.5, 2.5, 0.1, 6.8, 1.2, 0.5, 0.6, 5.3, 2.3, 1.8, 1.2, 1.3, 1.1, 0.9, 3.2, 1.0, 0.9, 0.4, 0.6, 8.0, 0.4, 2.7, 0.2, 2.0, 0.2, 0.5, 0.8, 2.0, 2.9, 0.1, 4.0.

The Kolmogorov-Smirnov test to above dataset for fitting Weibull distribution resulted respective *P* value (2-tail) 0.94 indicating that Weibull is a good model for above dataset. Here $\hat{a} = 1.89, \hat{b} = 1.01$, sample mean is 1.88 and sample standard deviation is 1.95. The mle \hat{b} indicates that the above dataset is moderately skewed. In order to compare the proposed 95% UPLs with those of Krishnamoorthy et al. [4], we chose the various combinations of *r*, *l*, and *m* and are given in Table 4.

From Table 4, it seems that the proposed UPLs are little less than those of Krishna-moorthy et al. [4]. We also notice that all the UPLs are well above the nominal range of vinyl chloride concentration (2.0–2.4) suggested by US Environmental Protection Agency (USEPA) indicating that future vinyl chloride concentrations are likely to be larger than the nominal level and hence monitoring of these wells is necessary.

6. Overall Conclusion

The proposed normal approximation exhibits markedly well performance even for small sample sizes for almost all parameter combinations for estimation of UPL that includes atleast *l* out of *m* future observations from Weibull distribution at each of *r* locations. The superiority of normal approximation is much more strong for small shape parameters and small sample

sizes which are commonly observed in real situations. It has an added advantage of being computationally simple, which is important from practitioners point of view.

References

- [1] D. K. Bhaumik and R. D. Gibbons, "One-sided approximate prediction intervals for at least p of m observations from a gamma population at each of r locations," *Technometrics*, vol. 48, no. 1, pp. 112–119, 2006.
- [2] C. B. Davis and R. J. McNichols, "One-sided intervals for at least p of m observations from a normal population on each of r future occasions," *Technometrics*, vol. 29, no. 3, pp. 359–370, 1987.
- [3] K. Krishnamoorthy, T. Mathew, and S. Mukherjee, "Normal-based methods for a gamma distribution: prediction and tolerance intervals and stress-strength reliability," *Technometrics*, vol. 50, no. 1, pp. 69–78, 2008.
- [4] K. Krishnamoorthy, Y. Lin, and Y. Xia, "Confidence limits and prediction limits for a Weibull distribution based on the generalized variable approach," *Journal of Statistical Planning and Inference*, vol. 139, no. 8, pp. 2675–2684, 2009.
- [5] F. Hernández and R. A. Johnson, "The large-sample behavior of transformations to normality," *Journal of the American Statistical Association*, vol. 75, no. 372, pp. 855–861, 1980.
- [6] Z. Yang, S. P. See, and M. Xie, "Transformation approaches for the construction of Weibull prediction interval," *Computational Statistics & Data Analysis*, vol. 43, no. 3, pp. 357–368, 2003.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

