ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

Estimation Methods for the Wakeby Distribution

Jose A. Raynal-Villasenor, Humberto R. Martinez-Avila (+), Maria E. Raynal-Gutierrez
Department of Civil and Environmental Engineering, Universidad de las Americas Puebla

Abstract— The probability weighted moments (PWM) and Iterative Linear Regression (ILR) methods for the estimation of the parameters and quantiles of the Wakeby distribution, were tested through distribution sampling experiments, towards its application in flood frequency analysis. The procedures to compute the parameters and quantiles are shown in the paper. The results showed that PWM method, in the estimation of parameters and quantiles, showed an overall better behavior that those provided by the ILR method. In the latter case, that better performance of the PWM method is quite obvious.

Index Terms— **Estimation of parameters, Wakeby distribution, probability weighted moments, iterative linear regression, flood frequency analysis.**

I. INTRODUCTION

A subject of paramount interest in planning and design of water works is that related with flood frequency analysis. Due to the characteristic that design values have, given that they are linked to a return period or to a non-exceedance probability, the use of mathematical models known as probability distribution functions is a must. Among the most widely used probability distribution functions for hydrological analysis, related with flood frequency analysis, are ([1]-[4]):

a) Two and three parameters Log-Normal (LN2 and LN3)
b) Pearson type III (PIII)
c) Log-Pearson Type III (LP-III)
d) Extreme Value Type I (EVI)
e) Extreme Value Type II (EVII)
f) General Extreme Value (GEV)
g) Wakeby (W)

With regard to the Wakeby distribution, [5] used the Wakeby distribution to characterize precipitation extremes for weather index-based insurance in the Zhujiang River Basin in China. Their results showed that maximum precipitation and 5-day-maximum precipitation are best described by the Wakeby distribution. The competing distributions with the Wakeby distribution were the Pearson III, Generalized Extreme Value and the Generalized Pareto distributions. Reference [6] compared the Wakeby distribution with Beta-Kappa and Beta-P distributions for representing annual extreme and partial duration rainfall series. They found that the Wakeby distribution provided better results than those by the Beta-Kappa and Beta-P distributions. [7] modelled the summer extreme rainfall in the Korean Peninsula using the Wakeby distribution. They obtained isopluvial maps of estimated design values corresponding to selected return periods. Reference [8] measured the applicability for the Wakeby distribution for flood frequency analysis in Eastern Australia. They developed regression equations to test if the Wakeby distribution fitted the annual maximum flood series at any given flow gauging station. Reference [9] introduced a numerical least squares method to estimate the parameters of the Wakeby distribution. The method was tested with flood data samples of Turkish rivers and the parameters obtained by such method were as good as those produced by the L-moments and the curve fitting method produced by MATLAB. Reference [10] simulated extreme precipitation in the Yangtse River Basin in China by using the Wakeby distribution. In such study, four probability distributions were used, namely General Extreme Value, Generalized Pareto, Generalized Logistic and Wakeby distributions. They observed that the Wakeby distribution can describe the probability distribution of the precipitation extremes calculated both from the observational and projected data.

II. THE WAKEBY DISTRIBUTION

The Wakeby distribution can be expressed in its inverse form as, [11] - [13]:

$$x = m + a\,[1 - (1 - F(x))^{b}\,] - c\,[\,1 - (1 - F(x))^{-d}] \qquad (1)$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

Where:

$F = F(x) = P\ (X \leq x)$

$x \geq m$

m, a, b, c, and *d* are the parameters of the Wakeby distribution. And the probability density function is:

$$f(x) = dF(x) / dx = [a\ b\ (1 - F(x))^{b-1} + c\ d\ (1 - F(x))^{-d-1}\ ]^{-1} \quad (2)$$

$f(x) \geq 0$ for all values of x, and $(a\ b + c\ d) \geq 0$. If $F(x) = 0$, then $x = m$ and $f(x) = 1 / (a\ b + c\ d)$. If $F(x) = 1$, the upper limit of x is ∞ or $(m + a - c)$. The value of parameter b has a strong influence on the shape of the Wakeby distribution. If 0 < b < 1, the Wakeby probability density distribution is a monotonically decreasing function, if $b > 1$ the probability density function will have a mode in some x value above its lower limit.
Another way to define the Wakeby distribution is, [14]-[15]:

$$x = -a\ (1-F)^{b} + c\ (1-F)^{-d} + e \quad (3)$$

where:

$$e = m + a - c \quad (4)$$

III. ESTIMATION PROCEDURES FOR THE WAKEBY DISTRIBUTION

In the Wakeby distribution, some traditional methods for the estimation of parameters, like those of moments and maximum likelihood, are not readily available given that a probability density function cannot be expressed in an explicit way. So, some other methods have to be applied to estimate the parameters of the Wakeby distribution. In the case of this paper, the methods that have been chosen are those of probability weighted moments (PWM) and iterative linear regression (ILR).

A. *Method of Probability Weighted Moments*

The probability weighted moments (PWM) of a probability distribution function can be defined as, [11]:

$$M_{l,j,k} = E\left[x^l F^j (1-F)^k = \int_0^1 [x(F)]^l F^j (1-F)^k dF\right] \quad (5)$$

where *l*, *j* and *k* are real numbers.

For the Wakeby distribution, when $m \neq 0$, eq. (5) becomes:

$$M_{l,0,0} = \sum_{s=0}^{l} \binom{l}{s} (m + a - c)^{l-s} \sum_{t=0}^{s} \binom{s}{t} (-a)^{s-t} c^t [b(s-t) - dt + 1]^{-1} \quad (6)$$

Using the notation $M_{(k)} = M_{1,0,k}$, then:

$$M_{(k)} = \frac{m}{1+k} + \frac{(a-c)}{1+k} - \frac{a}{1+k+b} + \frac{c}{1+k-d} \quad (7)$$

The parameters are obtained through the use of the following expressions, ([3], [11]):

$$\hat{b} = \frac{(N_3 C_1 - N_1 C_3) \pm [(N_1 C_3 - N_3 C_1)^2 - 4(N_1 C_2 - N_2 C_1)(N_2 C_3 - N_3 C_2]^{1/2}}{2(N_2 C_3 - N_3 C_2)} \quad (8)$$

$$\hat{d} = \frac{(N_1 + b N_2)}{(N_2 + b N_3)} \quad (9)$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

$$\hat{m} = \frac{[\{3\}-\{2\}-\{1\}+\{0\}]}{4} \quad (10)$$

$$\hat{a} = \frac{(b+1)(b+2)}{b(b+d)}\left[\frac{\{1\}}{2+b} - \frac{\{0\}}{1+b} - m\right] \quad (11)$$

$$\hat{c} = \frac{(1-d)(2-d)}{d(b+d)}\left[\frac{-\{1\}}{2-d} + \frac{\{0\}}{1-d} + m\right] \quad (12)$$

$$N_{4-j} = (4)^j M_3 - (3)^{1+j} M_2 + 3(2)^j (M)_1 - M_0 \quad (13)$$
$j = 1, 2, 3$

$$C_{4-j} = (5)^j M_4 - 3(4)^j M_3 + (3)^{1+j} M_2 - (2)^j M_1 \quad (14)$$
$j = 1, 2, 3$

$$\{k\} = (k+1)(k+1+b)(k+1-d)M_k \quad (15)$$
$k = 0, 1, 2, 3$

When $m = 0$, then eqs. (13) and (14) become, [3]:

$$N_{4-j} = -(3)^j M_2 - (2)^{1+j} M_1 - M_0 \quad (16)$$
$j = 1, 2, 3$

$$C_{4-j} = -(4)^j M_3 - 2(3)^j M_2 - (2)^j M_1 \quad (17)$$
$j = 1, 2, 3$

Equation (8) produces two values and [11] recommended to take the larger value of $\hat{b}$ without loss of generality. The Wakeby parameter estimation may not succeed if any of the following conditions hold, [13]:

1. Imaginary value of $\hat{b}$ or $\hat{b} < 0.3$ or $\hat{b} > 50$
2. $\hat{d} \geq 1$; the mean does not exist
3. Invalid probability density function: $f(m) = \frac{1}{\hat{a}\hat{b}+\hat{c}\hat{d}} < 0$
4. Improperly defined probability distribution function: $F(x_1) > F(x_2)$ for $x_1 < x_2$ for a combination of signs of parameters
5. Same a point 4 but for combination of parameters

The following procedure have been suggested, [13], for the estimation of the parameters of the Wakeby distribution, when using the method of probability weighted moments:

1. Assume $m \neq 0$ and determine the parameters by using the previous equations. Test for error conditions; if none hold, accept the parameters; if any error condition holds go to step 2
2. Assume $m = 0$, then fit a Wakeby distribution of four parameters. If this proves to be successful, then accept the parameters; otherwise go to step 3
3. Assume $m \neq 0$. Fix a number of allowed iterations. Set b equal to $b_{max} = 50$ and choose Δb. Go to step 4
4. If the allowed number of iterations is exceeded go to step 6. If not, set $\hat{b}$ to $(\hat{b} - \Delta b)$. If $\hat{b}$ falls below $b_{min} = 0.3$ go to step 6. Determine $\hat{d}, \hat{a}, \hat{c}$ and $\hat{m}$ from the previous equations and test for error conditions 1 to 3; if none hold go to step 5. If either conditions 1 or 2 hold go to beginning of step 4 to start the next iteration. If only condition 3 holds, decrease the value of Δb by half and start a new iteration at beginning of step 4
5. Test for error conditions 4 and 5. If neither hold, accept the parameters, otherwise go to step 6
6. Assume $m = 0$ proceed as in step 5, except the appropriate equations (for $m = 0$) are used to calculate $\hat{d}, \hat{a}$, and $\hat{c}$. If number of allowed iterations is exceeded or if $\hat{b}$ falls below $b_{min} = 0.3$ then the algorithm fails; otherwise go to step 7
7. $m = 0$. If a solution is obtained for which error conditions 1 to 3 do not hold, test for error conditions 4 and 5. If either condition 4 or 5 holds, the algorithm fails; otherwise the algorithm succeeds

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

B. Method of Iterative Linear Regression

The method of iterative linear regression (ILR) was initially proposed by [14]. The method consists in two phases, the first one the Wakeby distribution is analyzed just on its right hand side; in the second phase the left hand side of the distribution is analyzed.
The procedure consists in taking a value of Fc, which represents the separation point between the right and left tails of the distribution. The corresponding right hand side of the Wakeby distribution is analyzed when $Fc < F$ and the values of two of the parameters are assumed as: $a = 0$ and $b = 1$. Value for e are being explored and the values of parameters c and d are obtained by linear regression by the following expression:

$$log[x_k - e + a(1 - F_k)^b] = \log(c) - d\ log(1 - F_k)] \tag{18}$$

For any x_k when $F_k > F$. The values of the parameters selected in the phase one are those that produce the least sum of the squares obtained by linear regression.
In the phase two, with the computed values of the parameters c, d and e, the parameters a and b are calculated by linear regression by using the following equation:

$$log[-x_k + e + c(1 - F_k)^{-d}] = log(a) + b\ log(1 - F_k) \tag{19}$$

For any x_k when $F_k < F$. With the new values of a and b, the phase one is being repeated and then phase two is repeated and then the values are chosen as the final ones. The values F_c always have to be within the following limits: $0 < F_c < 1$. With sample sizes of 60, [14] has varied F_c for every sample size between 5 and 55.

IV. DISCUSSION OF RESULTS

In order to perform distributional sampling experiments, the same six Wakeby distributions used by [12]-[13], when they applied the probability weighted moments method, were utilized with the following values of the parameters and skewness coefficient for the Wakeby distribution.

Table 1. Parameters and skewness coefficient (γ) for the six Wakeby distributions used in the study, [12]-[13]

Distribution Name	Parameters				Skewness Coefficient
	a	b	c	d	γ
W1	1	16	4	0.20	4.14
W2	1	7.5	5	0.12	2.00
W3	1	1	5	0.12	1.91
W4	1	16	10	0.04	1.10
W5	1	1	10	0.04	1.11
W6	1	2.5	10	0.02	0.0

With the previous values of the parameters of the Wakeby distribution and using eq. (1), several samples were constructed with sizes of n = 11, 31, 51 and 101. Then, 101000 Wakeby distributed random number were generated, and the number of samples considered for each sample size were 9181, 3258, 1980 and 1000 samples, respectively. The parameters and quantiles of the Wakeby distribution were then estimated for each combination of sample size and distribution option by both selected methods of estimation of parameters, namely PWM an ILR. The statistical measures used to assess the suitability of each combination of sample size, number of samples and method of estimation of parameters and quantiles in the distribution sampling experiments, were the bias, variance and the mean squared error (MSE). The definitions of such statistical measures are as follows:

$$\widehat{\sigma^2}(\hat{\theta}) = \frac{\sum_{i=1}^{N}(\hat{\theta}_i - \hat{\mu}[\hat{\theta}])^2}{(N-1)} \tag{20}$$

$$Bias = \theta - \hat{\mu}(\hat{\theta}) \tag{21}$$

$$MSE = \widehat{\sigma^2}(\hat{\theta}) + [\theta - \hat{\mu}(\hat{\theta})]^2 \tag{22}$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

The results for the six distributions shown in table 1, for parameters a, b, c, d, and m, with regard to bias are shown in figs.1-5. The results for the six distributions shown in table 1, for parameters a, b, c, d, and m, with regard to variance are shown in figs.6-10. The results for the six distributions shown in table 1, for parameters a, b, c, d, and m, with regard to mean squared error are shown in figs.10-15.

The values obtained for the six distributions shown in table 1, for quantile 0.90 and 0.99, with regard to bias are shown in figs.16-17. The corresponding results for the six distributions shown in table 1, for quantile 0.90 and 0.99, with regard to variance are shown in figs.18-19. Finally, the results for the six distributions shown in table 1, for quantile 0.90 and 0.99, with regard to mean squared error are shown in figs.20-21.

The best overall performance was given by the PWM method, in both the estimation of parameters and quantiles. In the latter case, the much better performance of the PWM method is quite obvious.

V. CONCLUSION

In general terms and for the six Wakeby distributions analyzed, both methods presented similar reduction behaviors in the parameters and quantiles with regard to bias, variance and mean squared error as the sample size increased, as it was expected.

The W-3 and W-5 distributions (with parameter $b = 1$), were those with the largest convergence failures for both the probability weighted moments and the iterative linear regression methods.
The method of PWM showed to have the best overall performance, in both cases of estimation of parameters and quantiles, with regard to the cases explored in this study. In the latter case, the much better performance of the PWM method is quite obvious.

ACKNOWLEDGMENT

The authors wish to express their recognition to the Universidad of the Americas Puebla for the support granted to enable the publication of this paper.

REFERENCES

[1] G. W. Kite, Flood and Risk Analyses in Hydrology, Water Resources Publications, Littleton. 1988.

[2] J. D. Salas and R. Smith, Computer Programs of Distribution Functions in Hydrology, Colorado State University, Fort Collins. 1980.

[3] R. Rao and K. H. Hamed, Flood Frequency Analysis, CRC Press, Boca Raton. 2000.

[4] J. A. Raynal-Villasenor Frequency Analysis of Hydrologic Extremes, Lulu.com, http://www.lulu.com/spotlight/flodro4dot0atgmaildotcom . 2010.

[5] Fischer, T., Su, B., Luo, Y., et al (2012) Probability Distribution of Precipitation Extremes for Weather Index-Based Insurance in the Zhujiang River Basin, South China, Journal of Hydrometeorology, Vol. 13, No. 3, pp 1023-1037, Jun 2013.

[6] Oztekin, T., Wakeby Distribution for Representing Annual Extreme and Partial Duration Rainfall Series, Meteorological Application, Vol. 14, No. 4, pp 381-387, Dec. 2007.

[7] Park, J. S., Jung, H. S., Kim, R. S., et al, Modelling Summer Extreme Rainfall over the Korean Peninsula using Wakeby Distribution, International Journal of Climatology, Vol.21, No. 11, pp 1371-1384, Sep. 2011.

[8] Rahman, A., Zamman, M., Haddad, K., et al, Applicability of Wakeby Distribution in Flood Frequency Analysis: A Case Study for Eastern Australia, Hydrological Processes, Vol. 29, No. 4, pp 602-614, Feb. 2015.

[9] Oztekin, T., Estimation of the Parameters of Wakeby Distribution by Numerical Lest Squares Method and Applying it to the Annual Peak Flows of Turkish Rivers, Water Resources Management, Vol. 25, No. 5, pp 1299-1313, Mar. 2011.

[10] Su, B. Kundzwicz, Z. W., Jiang. T., et al, Simulation of Extreme Precipitation over the Yangtse River basin using Wakeby Distribution, Theoretical and Applied Climatology, Vol. 96, No. 3-4, pp 209-219, May 2009.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

[11] Greenwood, J. A., Maciunas Landwher, J., Matalas, N. C., Wallis, J. R., Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form, Water Resources Research, 15(5), 1049-1054. (1979).

[12] Landwher, J. M., Matalas, N. C., Wallis, J. R., Estimation of Parameters and Quantiles of Wakeby Distributions: 1. Known Lower Bounds, Water Resources Research, Vol. 15, No. 6, pp 1361-1372, 1979a.

[13] Landwher, J. M., Matalas, N. C., Wallis, J. R., Estimation of Parameters and Quantiles of Wakeby Distributions: 1. Unknown Lower Bounds, Water Resources Research, Vol. 15, No. 6, pp 1373-1379, 1979b.

[14] Houghton, J. C., Birth of a Parent: The Wakeby Distribution for Modeling Flood Flows, Water Resources Research, Vol. 14, No. 6, pp 1105-1109, 1978a.

[15] Houghton, J. C., The Incomplete Means Estimation Procedure Applied to Flood Frequency Analysis, Water Resources Research, Vol. 14, No. 6, pp 1111-1115, 1978b.

AUTHOR BIOGRAPHY

Jose A. Raynal-Villasenor, obtained his Ph. D. at Colorado State University in 1985, has published 43 research papers, 33 book chapters, 113 conference papers, 2 books and 6 edited books. He belongs to 15 professional societies and received several awards some of them coming from such societies and he is member of Mexico´s National System of Researchers since year 1984.

Humberto R. Martinez-Avila, obtained his Master in Water Resources Engineering at the Universidad Nacional Autonoma de Mexico. He was part of the faculty at the Colegio de Postgraduados at Chapingo, Mexico. He passed away several years ago.

Maria E. Raynal-Gutierrez, obtained her Ph. D. at Colorado State University in 2007, has published 10 research papers, 2 book chapters and 5 conference papers. She belongs to 2 professional societies and she is member of Mexico´s National System of Researchers since year 2014.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

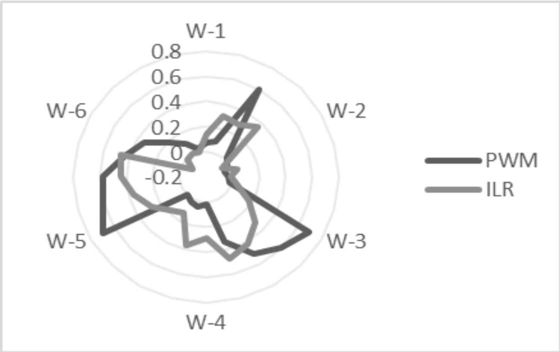

Fig 1. Bias for the six Wakeby distributions, parameter a, n = {11, 31, 51 and 101}

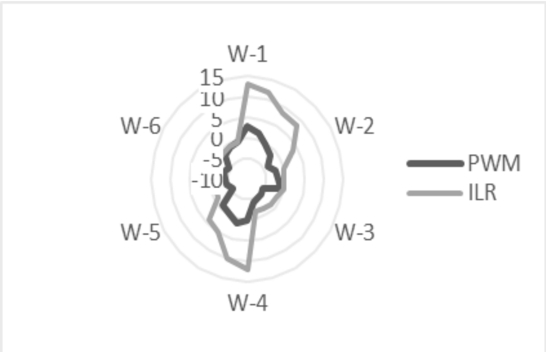

Fig 2. Bias for the six Wakeby distributions, parameter b, n = {11, 31, 51 and 101}

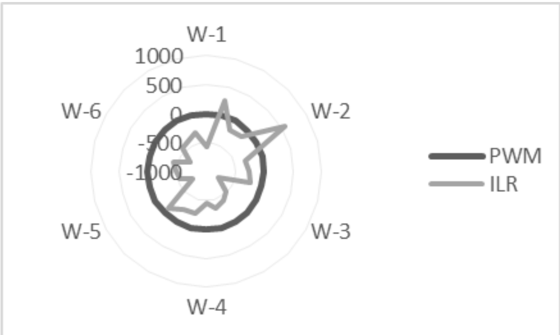

Fig 3. Bias for the six Wakeby distributions, parameter c, n = {11, 31, 51 and 101}

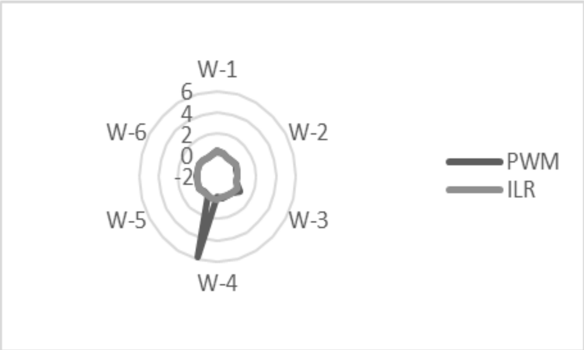

Fig 4. Bias for the six Wakeby distributions, parameter d, n = {11, 31, 51 and 101}

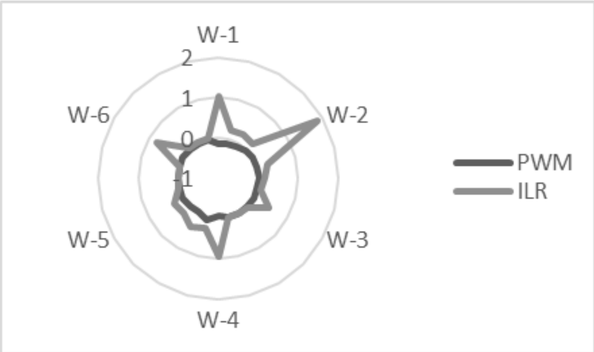

Fig 5. Bias for the six Wakeby distributions, parameter m, n = {11, 31, 51 and 101}

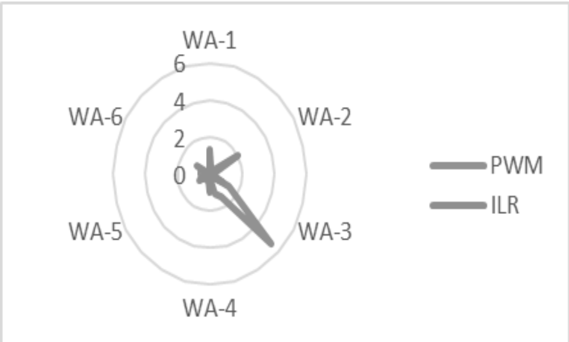

Fig 6. Variance for the Wakeby distributions, parameter a, n = {11, 31, 51 and 101}

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

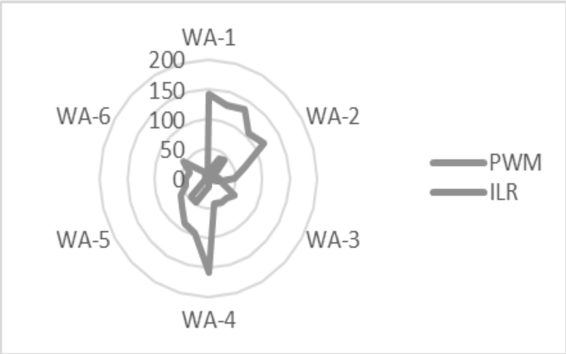

Fig 7. Variance for the Wakeby distributions, parameter *b, n = {11, 31, 51 and 101}*

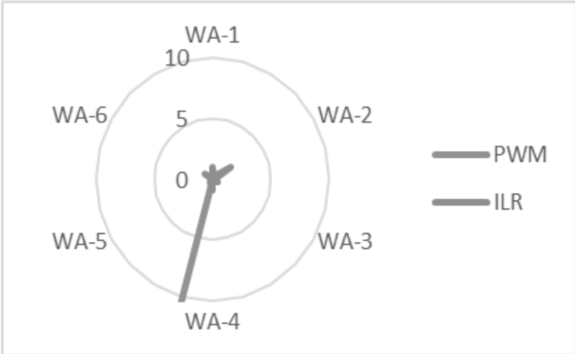

Fig 10. Variance for the Wakeby distributions, parameter m, *n = {11, 31, 51 and 101}*

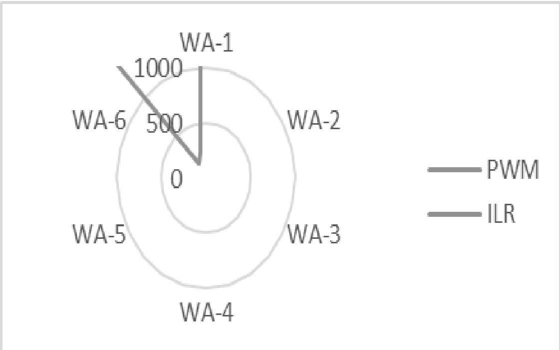

Fig 8. Variance for the Wakeby distributions, parameter *c, n = {11, 31, 51 and 101}*

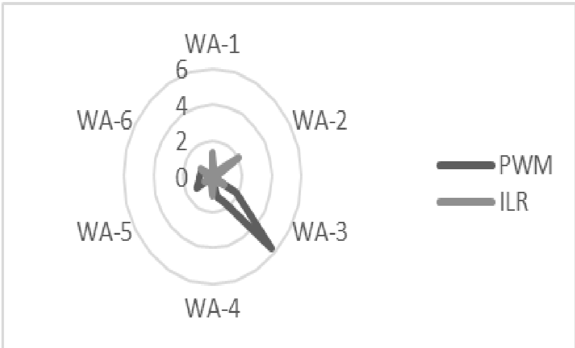

Fig 11. MSE for the Wakeby distributions, parameter *a, n = {11, 31, 51 and 101}*

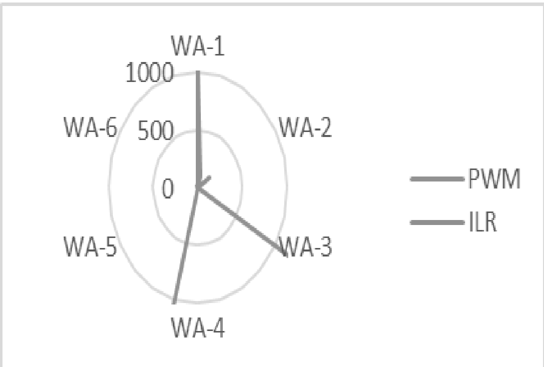

Fig 9. Variance for the Wakeby distributions, parameter *d, n = {11, 31, 51 and 101}*

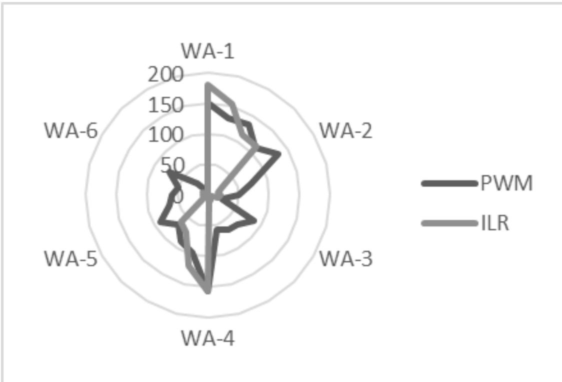

Fig 12. MSE for the Wakeby distributions, parameter *b, n = {11, 31, 51 and 101}*

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

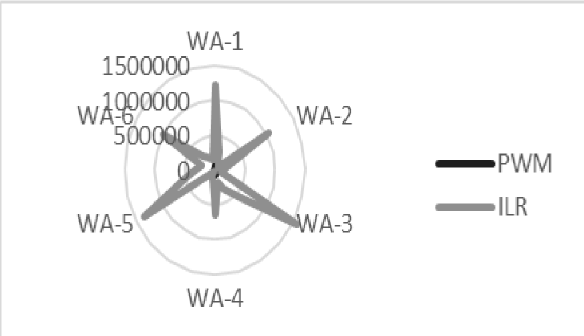

Fig 13. MSE for the Wakeby distributions, parameter *c*, *n* = {11, 31, 51 and 101}

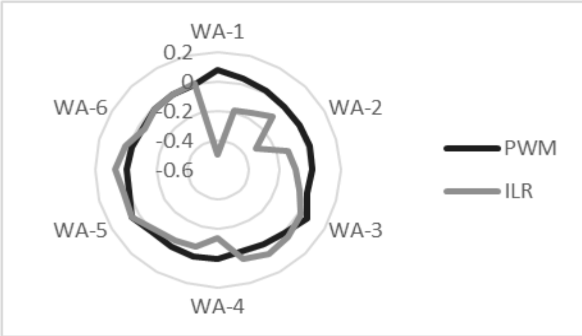

Fig 16. Bias for the six Wakeby distributions, quantile 0.90, *n* = {11, 31, 51 and 101}

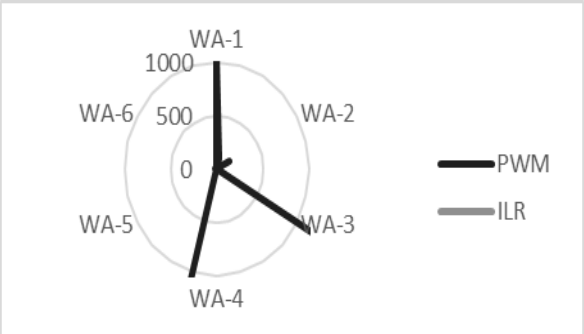

Fig14. MSE for the Wakeby distributions, parameter *d*, *n* = {11, 31, 51 and 101}

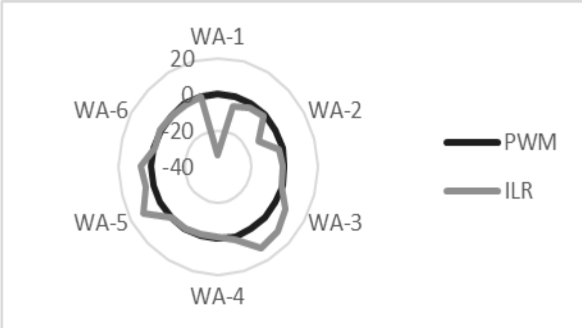

Fig 17. Bias for the six Wakeby distributions, quantile 0.99, *n* = {11, 31, 51 and 101}

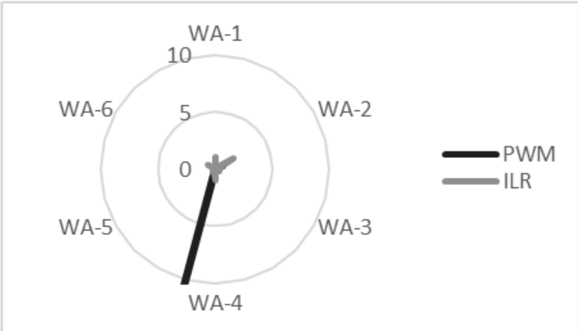

Fig 15. MSE for the Wakeby distributions, parameter *m*, *n* = {11, 31, 51 and 101}

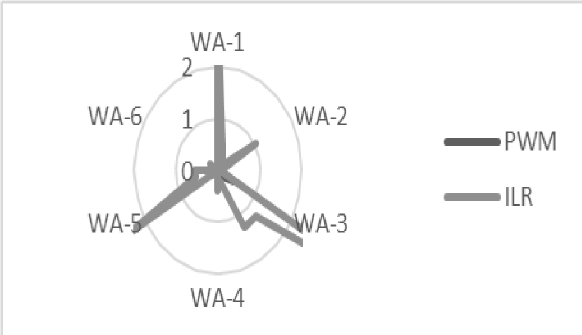

Fig 18. Variance for the six Wakeby distributions, quantile 0.90, *n* = {11, 31, 51 and 101}

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 2, March 2016

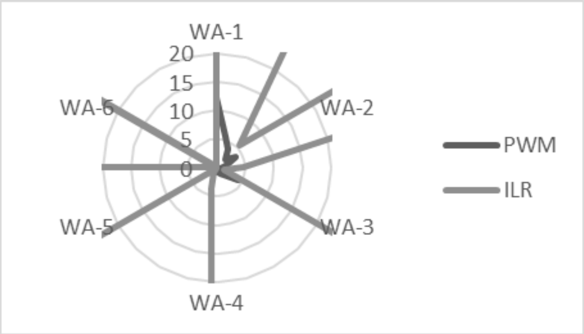

Fig 19. Variance for the six Wakeby distributions, quantile 0.99, n = {11, 31, 51 and 101}

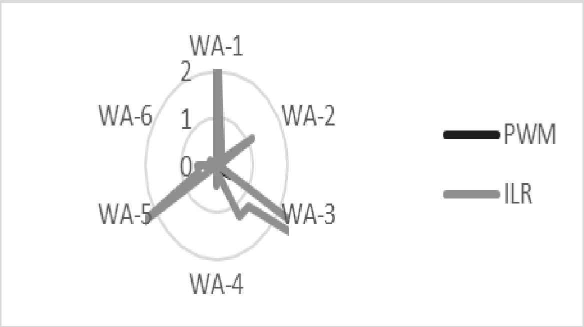

Fig 20. MSE for the six Wakeby distributions, quantile 0.90, n = {11, 31, 51 and 101}

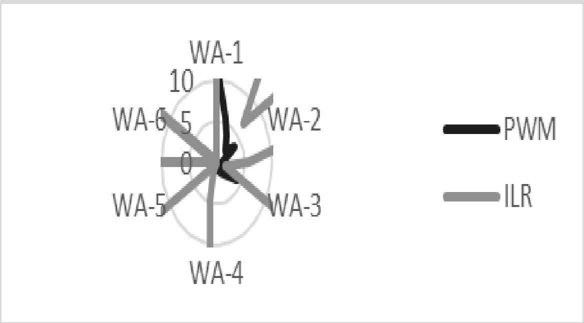

Fig 21. MSE for the six Wakeby distributions, quantile 0.99, n = {11, 31, 51 and 101}