Frequency Analysis of Extreme Mean Annual Rainfall Events

EM Masereka[a], FAO Otieno[b], GM Ochieng[c] J Snyman[d]

[a]D.Tech. student Civil Engineering Department, Tshwane University of Technology, Pretoria, South Africa

[b]Prof and Vice Chancellor, Masinde Muliro University of Science and Technology, Kakamega, Kenya

[c]Prof and Head of Dept; Civil Engineering and Building, Vaal University of Technology, Vanderbijlpark, South Africa

[d]Prof and Head of Dept. of Civil Engineering Tshwane University of Technology; Pretoria, South Africa.

Abstract

Frequency analysis of extreme mean annual rainfall events is important to water resource planners at catchment level because mean annual rainfall is an important parameter in determining mean annual runoff. Mean annual runoff is an important input in determining surface water available for water resource infrastructure development. The objective of this study was to carry out frequency analysis of extreme low mean annual rainfall events in 8 rainfall zones in the Sabie River catchment in South Africa.

Peaks Below Threshold (PBT) method was applied to extract extreme low mean annual rainfall events from 85 year record. Candidate Model Identification Criterion (CMIC) and Least Sum of Statistic Model Selection Criterion (LSSMSC) were applied to identify the best fit models for the frequency analysis. Parameters were estimated by maximum likelihood method. Quantile-Quantile (Q-Q) and Probability-Probability (P-P) plots were applied to evaluate the performance of the model selection criteria.

From the study, the quantiles at return periods of 5, 10, 25, 50, 100 and 200 years for each of 8 rainfall zones in Sabie River catchment were obtained.

Based on the results of the study, no single probability distribution function or model is the best fit for frequency analysis of extreme low mean annual rainfall events in all 8 rainfall zones Sabie River catchment.

Keywords: Best fit probability distribution function, candidate model identification criterion, least sum of statistics model selection criterion.

1. Introduction

1.1 Water resources situation in South Africa

South Africa is divided into 19 water management areas (WMA), 11 of these catchment areas are water deficit catchment where water resource demand is greater than the available water resource. Estimates carried out by the Department of Water Affairs and Forestry (DWAF) indicated that by 2025 two or more additional Water Management Areas will experience water deficit situation (IWMI, 2000). Although only 11 out of 19 water management areas are in situation of water resource deficit, on average the whole country can be classified as water stressed. The annual fresh water availability is estimated to be less than 1700m³ per capita. The 1700m³ per capita is taken as the threshold or index for water stress. Water deficit scenarios in water management areas are increased by frequent occurrences of extreme low mean annual rainfall events. Extreme Low mean annual events lead to low runoff, low stream flow and limited underground water recharging. Prolonged periods of low rainfall periods result into low production of agricultural products. This has serious financial implications to the nation. Frequency analysis of extreme mean annual rainfall is therefore important. The main focus of this study was to carry out frequency analysis of extreme low mean annual rainfall events in Sabie River catchment. Sabie River catchment is one of the water deficit catchments in South Africa. .

1.2 Frequency analysis of mean annual rainfall events

Mean annual rainfall events like any hydro meteorological variables change through time according to probabilistic laws. This process is called stochastic process. Stochastic frequency analysis therefore involves developing mathematical probability distribution functions (PDFs) or models that are applied to extrapolate and generate events based on sample data of those specific events. Numerous Stochastic models have been developed to extrapolate different aspects of hydro meteorological events including mean annual runoff, mean annual rainfall, temperature, stream flow, groundwater, soil moisture and wind. Shamir (2005) developed stochastic techniques to generate input data for modeling small to medium catchments.

1.3 Recommended (PDFs) for analysis of extreme hydro meteorological events in South Africa.

Log-Pearson 3 (LP3) probability distribution function has been recommended for the design of hydro meteorological events mostly flood and drought in South Africa (Alexander, 1990, 2001). Gorgens et al (2007) reported that both the LP3 and General Extreme Value distributions were suitable for frequency analysis of extreme hydro meteorological events in South Africa. Mkhandi et al. (2000) also reported that the Pearson Type 3 probability distribution function fitted with parameters by Method of Probability Weighted Moments (PWM) was found to be the most appropriate distribution for frequency analysis of hydro meteorological events in twelve (12) of the fifteen (15) relatively homogenous catchments in South Africa. Masereka et al. (2015) developed Candidate Model Identification Criterion

(CMIC) and Least Sum of Statistic Model Selection Criterion (LSSMSC) for identifying the best fit probability distribution functions for frequency analysis of extreme hydro meteorological events specifically extreme mean annual rainfall events in water deficit catchments in South Africa.

1.4 The objectives of the study

The main objective of the study was to carry out frequency analysis of extreme low mean annual rainfall events in Sabie River catchment.

The specific objectives of the study were:

1. To identify the best fit PDFs for frequency analysis of extreme low mean annual rainfall events in each of the 8 rainfall zones in Sabie River catchment.
2. To establish Quantile-Return period (Q_T-T) models for frequency analysis of the extreme low rainfall events in each of the 8 rainfall zones in Sabie River catchment based on the identified best it PDFs of the corresponding rainfall zone and
3 To apply the established Q_T-T model to carry out the frequency analysis of extreme low mean annual rainfall events in each rainfall zone.

2. Materials and Methods

The Sabie River catchment was case study catchment (Figure1). It is divided into of three (3) tertiary catchments. The 3 tertiary catchments are subdivided into twenty five (25) quaternary catchments. The quaternaries in each of the eight (8) rainfall zones are indicated in Table 1

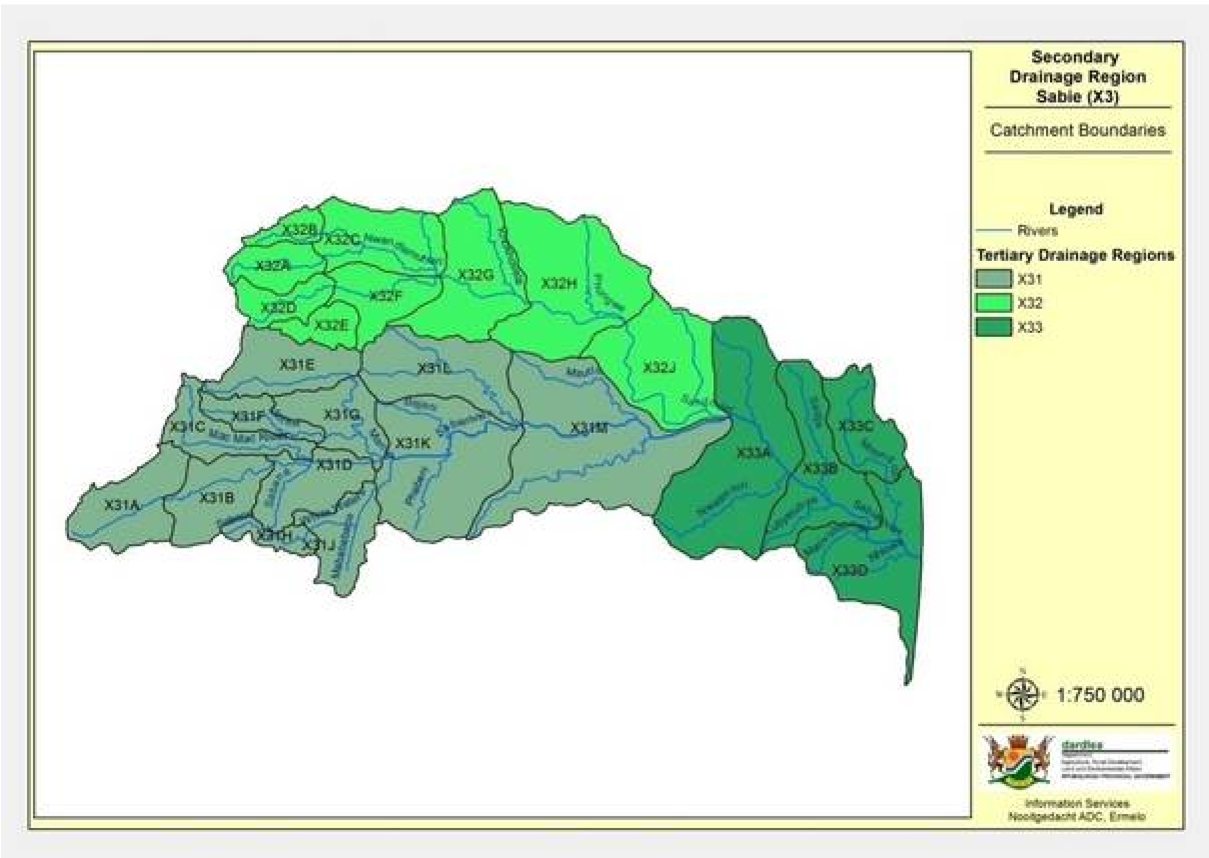

Figure 1 Tertiary and quaternary catchments of Sabie River catchment (Source: Dept. Agric. Mpumalanga 2010)

Table 1 Quaternary catchments and, rainfall zones Sabie River catchment

Quaternary	**Catchment area**		**S-pan evaporation**		**Rainfall**	
Catchment	**Gross**	**Net**	**evap**	**MAE**	**Rainfall**	**MAP**
	(km2)	(km2)	zone	(mm)	zone	(mm)
X31A	229.7	229.7	5A	1418.6	X3A1	1218
X31B	198.0	198.0	5A	1425.6	X3A1	1262
X31C	153.5	153.5	5A	1421.8	X3A1	1295
X31D	189.4	189.4	5A	1442.6	X3A2	960
X31E	212.7	212.7	5A	1430.8	X3B	1256
X31F	91.4	91.4	5A	1424.7	X3B	1329
X31G	167.7	167.7	5A	1454.2	X3B	953
X31H	61.3	61.3	5A	1420.8	X3A1	1164
X31J	153.7	153.7	5A	1432.9	X3A2	883
X31K	490.9	490.9	5A	1476.3	X3C	672
X31L	295.0	295.0	5A	1478.0	X3C	735
X31M	713.9	713.9	5A	1483.9	X3E	569
Tertiary	***2957.2***	***2957.2***		***1456***		***706***
X32A	110.5	110.5	5A	1440.6	X3D1	1092
X32B	54.4	54.4	5A	1460.4	X3D1	960
X32C	230.7	230.7	5A	1491.0	X3D2	766
X32D	98.0	98.0	5A	1450.0	X3D1	1139
X32E	79.6	79.6	5A	1454.5	X3D1	906
X32F	163.3	163.3	5A	1493.2	X3D2	728
X32G	339.5	339.5	5A	1493.0	X3D2	662
X32H	482.0	482.0	5A	1488.4	X3E	639
X32J	351.3	351.3	5A	1483.4	X3E	593
Tertiary	***1909.2***	***1909.2***		***1482***		***730***
X33A	601.6	601.6	5A	1469.8	X3F	540
X33B	316.9	316.9	5A	1461.9	X3F	525
X33C	178.5	178.5	5A	1452.7	X3F	484
X33D	311.1	311.1	5A	1453.5	X3F	469
Tertiary	***1408.1***	***1408.1***		***1462***		***514***

2.2 Frequency analysis processes

Frequency analysis of extreme low mean annual rainfall events in 8 rainfall zones in Sabie River catchment was carried out in five (6) steps described as follows:

1. Identification of extreme low mean annual events
2. Application of CMIC to identify candidate models
3. Application of LSSMSC to identify the best fit PDFs from the candidate models
4. Developing Q_T-T models for each extreme low mean rainfall events
5. Calculating quantiles (cm) of return periods (years) of 5, 10, 25, 50, 100 and 200 for each of the eight (8) rainfall zones and
6. Evaluation of the performance of CMIC and LSSMSC by Q-Q and P-P plots.

2.2.1 Identification of extreme mean rainfall events by Peaks Below Threshold (PBT) method

The mean of the mean annual rainfall events was calculated for each zone for the period 1920-2004. The calculated mean of mean annual rainfall events was then chosen as the threshold. PBT method was then applied to identify extreme low mean annual rainfall events for each of the 8 rainfall zones. The rainfall events less than the threshold formed the extreme low mean annual rainfall events.

2.2.2 Application of CMIC to identify candidate PDFs

Masereka et al. (2015) have comprehensively described the development and the application of CMIC to identify candidate probability distribution functions for frequency analysis of extreme annual rainfall events. The development and the application of CMIC was based on the classification of continuous probability distribution functions (models) and bound characteristics of upper and lower tail events of the distributions of sample data. CMIC was adopted in this study to identify the candidate probability functions for frequency analysis of extreme low mean annual events in Sabie River catchment.

2.2.3 Application of LSSMSC to select best fit PDFs

Masereka et al. (2015) have in detail described the development and the application of LSSMSC. LSSMSC was developed based on statistics of the three (3) commonly applied goodness-of-fit tests. The three (3) goodness-of-fit tests are: Kolmogorov-Smirnov, Anderson-Darling and Chi-Square.

The definition of the developed LSSMSC was:

$$LSSMSC = Dn + A^2{}_n + abs.\chi^2{}_{red} \mid Least\ i \quad (i = 1..m) \tag{1.1}$$

where Dn = Kolomogrov-Smirnov statistic , $A^2{}_n$ = Anderson-Darling statistic

$abs.\chi^2{}_{red}$ = Absolute reduced Chi-Square , m = number of models.

In this study, LSSMSC was adopted to select the best fit PDFs from the candidate PDFs which had been identified by the application of CMIC. In each rainfall zone, the PDF with the least LSSMSC was taken as the best fit PDF for frequency analysis of extreme low annual rainfall events in that specific rainfall zone.

2.2.4 Development of Quantile-Return period (Q_T-T) models

Based on the identified best fit model for frequency analysis of extreme low mean annual rainfall events for each of the 8 rainfall zones, Q_T-T models were developed. The developed Q_T-T models were applied to extrapolate and estimate extreme low mean annual rainfall events for return periods of 5, 10, 25, 50, 100 and 200 years for each rainfall zone.

2.2.5 Evaluation of the performance of CMIC and LSSMSC by Q-Q and P-P plots

The evaluation of CMIC and LSSMSC was carried out by plotting theoretical quantile Vs empirical quantile and theoretical probability Vs empirical probability for each of the best fit PDF in each of the 8 rainfall zone.

The application of CMIC, LSSMSC, parameter estimations by methods of moments and maximum likelihood and plotting of Q-Q and P-P carried out in Easy Fit V5.5 software.

3. Results and Discussion

3.1 The candidate PDFs identified for frequency analysis of extreme low mean annual rainfall events

The results showed that the candidate PDFs for frequency analysis of extreme low mean annual rainfall events were: Generalised Pareto (GP), Wakeby (Wak), Generalised Logistic (GL), Log-Pearson 3 (LP3), and Generalised Extreme Value (GEV). The CDFs and PDFs of the identified candidate models are presented in Table 1.

3.2 Results of identification of best-fit PDFs for frequency analysis of extreme low mean annual rainfall events in Sabie River catchment

The best fit PDFs for frequency analysis of extreme low mean annual rainfall events in 8 rainfall zones in Sabie River catchment are presented in Tables 3. The Q_T-T models to estimate extreme low mean annual rainfall events of return periods of: 5, 10, 25, 50, 100 and 200 years are also included in Table 3. Table 2 is showing the performance of each PDF.

Log-Pearson 3 and Generalised Logistic models each was the best fit PDFs for frequency analysis of extreme low mean annual rainfall events in three (3) rainfall zones. The results show that there no single PDF that is the best fit for frequency analysis in all rain fall zones in Sabie River catchment.

3.3 Results of frequency analysis of extreme low mean annual rainfall events in 8 rainfall zones in Sabie River catchment

The results of frequency analysis of extreme low mean annual rainfall events in the 8 zones of Sabie River catchment are presented in Table 3.

Table 1 Candidate models for frequency analysis of extreme low mean annual rainfall events

Distribution	CDF or PDF	Domain
Generalised Extreme Value	$F(x) = \begin{cases} \exp\left(-\left(1+kz\right)^{-1/k}\right) & k \neq 0 \\ \exp\left(-\exp(-z)\right) & k = 0 \end{cases}$	$\left. \begin{matrix} 1 + k\frac{(x-\mu)}{\sigma} > 0 & for & k \neq 0 \\ -\infty < x < +\infty & for & k = 0 \end{matrix} \right\}$
Generalised Pareto	$F(x) = \begin{cases} 1 - (1 + k\frac{(x-\mu)}{\sigma})^{-1/k} & k \neq 0 \\ 1 - \exp\left(-\frac{(x-\mu)}{\sigma}\right) & k = 0 \end{cases}$	$\left. \begin{matrix} \mu \leq x < +\infty & for & k \neq 0 \\ \mu \leq x \leq \mu - \sigma/k & for & k = 0 \end{matrix} \right\}$
Generalised Logistics	$F(x) = \begin{cases} \frac{1}{1+(1+kz)^{-1/k}} & k \neq 0 \\ \frac{1}{1+\exp(-z)} & k = 0 \end{cases}$	$\left. \begin{matrix} 1 + k\frac{(x-\mu)}{\sigma} > 0 & for & k \neq 0 \\ -\infty < x < +\infty & for & k = 0 \end{matrix} \right\}$
Log-Pearson 3	$F(x) = \left. \frac{\Gamma(\ln(x)-\gamma)/\beta^{(\alpha)}}{\Gamma(\alpha)} \right\}$	$\left. \begin{matrix} 0 < x \leq e^{\gamma} & \beta < 0 \\ e^{\gamma} \leq x < +\infty & \beta > 0 \end{matrix} \right\}$
Wakeby	$x(F) = \xi + \frac{\alpha}{\beta}\left(1 - (1-F)^{\beta}\right) - \frac{\gamma}{\delta}(1 - (1-F)^{-\delta}$	$\xi \leq x\ \infty$ if $\delta \geq 0$ and $\gamma > 0$ $\xi \leq x \leq \xi + \alpha/\beta - \gamma/\delta$ if $\delta < 0$ *or* $\gamma = 0$

Table 2 Performances of the best fit PDFS

Best fit model	Rainfall zones	%
Log-Pearson 3 (LP3)	3	37.5
Generalised Logistics (GL)	3	37.5
Generalised Pareto (GP)	0	0
Generalised Extreme Value (GEV)	2	25
Totals	8	100

Log-Pearson 3 was the best fit PDF in 3 out of the 8 rainfall zones. Generalised Logistic was the best fit PDF in 3 out of 8 rainfall zones and Generalised Extreme Value was best fit in 2 out of the 8 rainfall zones. The results indicate that not a single PDF is best fit PDF for frequency analysis of extreme mean annual rainfall events in Sabie River catchment.

Table 3 Recurrence intervals in years of extreme low mean annual rainfall events (cm) in Sabie River catchment.

Rainfall zone	Best-fit	Mathematical model	Recurrence interval in years					
			5	10	25	50	100	200
X3a1	LP3	$X_T = 85.49 - 9.845K_T$	77.06	73.18	69.84	67.76	66.24	64.76
X3a2	GL	$X_T = 100.23 - 8.797\{1-(T-1)^{-0.0964}\}$	99.12	98.55	97.90	97.49	97.08	96.71
X3b	GL	$X_T = 100.7 - 29.65\{1-(T-1)^{-0.1967}\}$	93.62	90.30	86.91	84.84	83.06	81.52
X3c	LP3	$X_T = 82.65 - 13.16\ K_T$	71.67	68.23	66.78	64.15	62.90	61.54
X3d1	GL	$X_T = 100.63 - 39.52\{1-(T-1)^{-0.1638}\}$	92.60	88.91	84.59	82.00	79.73	77.72
X3d2	GEV	$X_T = 100.86 - 22.67\left\{1-\left[-\ln\left(1-\frac{1}{T}\right)\right]^{0.5921}\right\}$	87.51	84.17	81.60	80.44	79.68	78.99
X3e	GEV	$X_T = 99.84 - 21.43\left\{1-\left[-\ln\left(1-\frac{1}{T}\right)\right]^{0.669}\right\}$	86.27	83.17	80.93	79.99	79.40	79.03
X3f	LP3	$X_T = 82.06 - 12.74k_T$	71.20	68.01	63.07	60.45	59.12	57.82

The results of the study show that the probability of annual rainfall being lower that the national average of 450mm is very low. However, very high figures of mean annual evaporation lead to water deficit in the catchment.

4. Conclusions

Based on the results of the study, no single probability distribution function or model is the best fit for frequency analysis of extreme low mean annual rainfall events in all 8 rainfall zones in Sabie River catchment. Log-Pearson 3 and Generalised Logistic each is best fit in 3 out of 8 rainfall zones. Generalised Extreme Value is best fit in 2 rainfall zones out of 8. From the results, it can be concluded that recommending a single model for frequency analysis of extreme hydro meteorological events as the best fit for a whole region or country is not advisable. It is therefore recommended that to carry out frequency analysis of extreme hydro meteorological events, the best fit model should be determined first by applying suitable model selection criteria rather than adopting a recommended model which has not been tested as a best fit model for the specific extreme hydro meteorological events.

Probability-Probability (P-P) and Quantile-Quantile (Q-Q) plots were applied to evaluate the performance of CMIC and LSSMSC. The plots showed that CMIC and LSSMSC performed fairly well. Although the P-P and Q-Q plots results cannot be considered completely conclusive, CMIC and LSSMSC criteria make useful tools as model selection method for frequency analysis of extreme hydrological events like mean annual rainfall events.

Acknowledgements

The study has been supported by the Provincial Dept. of Agriculture, Rural Development, Land and Environmental Affairs: Mpumalanga Provincial Government, South Africa.

References

Alexander, W.J.R., 1990. Flood Hydrology for Southern Africa. SANCOLD, Pretoria, RSA.

Alexander, W.J.R., 2001. Flood Risk Reduction Measures. University of Pretoria, Pretoria, RSA.

Dept. Agric. Mpumalanga, 2010. Research Report on Natural Resources of the Province : Internal Report.

Gorgens, A., H., M., Joint Peak- Volume (JPV) Design Flood Hydrographs for South Africa. WRC Report 1420/3/07 Water Research Commission, Pretoria, South Africa.

Gorgens, A., H., M., Lysons,S., Hayes, L., Makhabana, M., and D. Maluleke, 2007. Modernised South African Design Floo Pro actce in the context of Dam safety, WRC Report 1420 / 2 / 07, Water Research omission, Pretoria, South Africa

Masereka, E. M., Otieno, F.A.O., Ochieng G. M. and J. Snyman 2015. Best Fit and Selection of Probability Distribution Models for Frequency Analysis of Extreme Mean Annual Rainfall Events International Journal of Engineering Research and Development *e-ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.comVolume 11, Issue 04 (April 2015), PP.34-53*

Mathwave(2011), http://www.mathwave.com.

Mkhandi, S.H., Kachroo, R.K., Gunasekara, T.A.G., 2000. Flood frequency analysis of Southern Africa:II. Identification of regional distributions. Hydrological Sciences Journal, 45(3):449-464.

IWMI (International Water Management Institute), 2000. World water supply and demand: 1995 to 2025. Colombo, Sri Lanka: International Water Management Institute. Article [Online]. Available from: www.iwmi.cgiar.org/.../IWMI

Shamir,E., 2005. Application of temporal streamflow descriptors in hydrologic model parameter estimation. Water Resource Research 41 doi: 10 : 1029/2004 WR 003409 SSN 0043-1397