Scientific Research Journal *Vol. 4 No. 1, 39-48, 2007*

Low Flow Analysis for Drought Management

Ismail Atan[1], Jurina Jaafar & Aminuddin Mohd Baki

Institute for Infrastructure Engineering and
Sustainable Management (IIESM)
Faculty of Civil Engineering
Universiti Teknologi MARA (UiTM), Malaysia
[1]Email: ismail585@yahoo.com

ABSTRACT

A statistical method to fit the lower tail of the cumulative distribution function of streamflow is presented. The proposed model is distribution-free that is it does not require the specification of a parent distribution. The potential use of the proposed model is its ability to estimate small quantities of low flow distribution for drought management. The accurate estimation of low flow is very important for water resources planning and to illustrate the applicability of this model, the Langat River was selected.

Keywords*: flow distribution, drought management, water resources planning*

Introduction

The characteristic of streamflow is important for water supply studies and for the assessment of water quality parameters. Riggs [1] introduced the parametric methods to investigate low flow while MacMohan [2] presented a survey of low flow analysis. Loaciga [3] developed a non parameteric method to estimate low flow using the location, scale and the shape of the parameter of the cumulative density function of streamflows. The aim of this paper is to estimate the parameters of low flow using maximum likelihood function.

ISSN 1675-7009

Basic Assumption of the Low Flow Data

In analyzing the drought events, it is generally assumed that these events are independently distributed in time. The independence assumption may be considered to be physical plausible when the time separating drought events is relative long. Among the widely used test of independence is the Wald and Wolfowitz test [4], which is based on the motion of serial correlation.

Another assumption required is that all the elements of the sample should belong to the same population. In some cases this condition of homogeneity of annual low flow series is not met because the sample is composed of events of different origin (rainfall and evaporation). Checking the homogeneity of an annual low flow sample can be done with the help of the Mann and Whitney test.

The most important assumption with regard to low flow series is that no significant changes in the general characteristics of the series occur from one year to the other year. It is quite reasonable practice to treat low flow series as stationary within the time interval of the year.

The Distribution Free Model

Suppose that flow variates, *i.e.*, $X_1, X_2, ...X_N$, are ordered from smallest to largest and the sorted record is $X_{(1)} < X_{(2)} < .... < X_{(N)}$. The next step in the distribution free approach is to censor out the variates $X_{(m+2)}, ..., X_{(N)}$, and the trimmed subsample $X_1, X_2, ...X_{(m)}$ is used to fit the lower tail of the CDF. The purpose of censoring out $X_{(m+1)}, X_{(m+2)}, ..., X_{(N)}$, is to derive an approximation to the lower tail of the CDF that is not adversely affected by the flow values above a threshold flow $X_{(m)} = x_{(m)}$. Suppose that the cumulative density function,CDF is given as [5]:

$$\mathrm{P}\left(X \le x\right) = \mathrm{F}_x\left(x\right) \approx \left(\frac{x-a}{b}\right)^c \tag{1}$$

In which a = the lower bound to X; and b and c are positive constant. Equation 1 implies that the lower tail of the CDF of flows can be approximated by a power function that is defined by a location parameter a, and scale and shape parameters, b and c respectively. It is assumed that Equation 1 is valid only as X approaches the lower bound a from above.

The estimation of $\hat{a}, \hat{b}$ and $\hat{c}$ is done by method of the Maximum Likelihood and the steps to estimate the quantiles are as follows:

1. Plot (a) the interval $0 \le a < x_{(i)}$ to approximate $\hat{a}$;
2. Use a univariate search technique to locate $\hat{a}$ precisely; and
3. A FORTRAN programming (see Appendix 2) is used to compute the values of $\hat{b}$ and $\hat{c}$ where

$$\hat{b} = \left(\frac{N-m}{m} \right)^{1/c} \left[x_{(m)} - a \right] \tag{2}$$

$$\hat{c} = \frac{m}{\sum_{i=1}^{m} \ln \left[\frac{x_{(m)} - a}{x_{(i)} - a} \right]} \quad \text{‘(3)}$$

Having computed the values of $\hat{a}, \hat{b}$ and $\hat{c}$, maximum likelihood estimator of the *p*th quantile is given as:

$$\hat{x}_p = \hat{a} + \hat{b}(p)^{1/\hat{c}} \tag{4}$$

in which the probability of $p = P(X \le x)$ is close to zero.

Model Application

The model developed is illustrated with a series of annual flow volumes in the Langat River which originate in the western slope of the mountain ridge penetrating the Malay Peninsula. The river collects water in the mountain areas and generally flows southwestward in the mountainous terrain. It gradually changes direction and flows westward into the Straits of Malacca. The total catchments area of the Langat Basin is 1,988 km^2. Table 1 contains 25-year record of flow volumes. The annual low flow series is first tested for independence and stationary distribution using MINITAB macro (Appendix 1), and the parameters are determined (Appendix 2, FORTRAN programming) and subsequently the low quantiles are estimated.

In Figure 1, it is quite obvious that variation of annual low flow has no significant changes in level change, trend and cyclic movement throughout the years. This indicates that the data may be stationary (the data falls within the boundary limits of $\mu \pm 3\sigma$). The Mann-Whitney test was conducted and the null hypotheses of stationary distribution are not

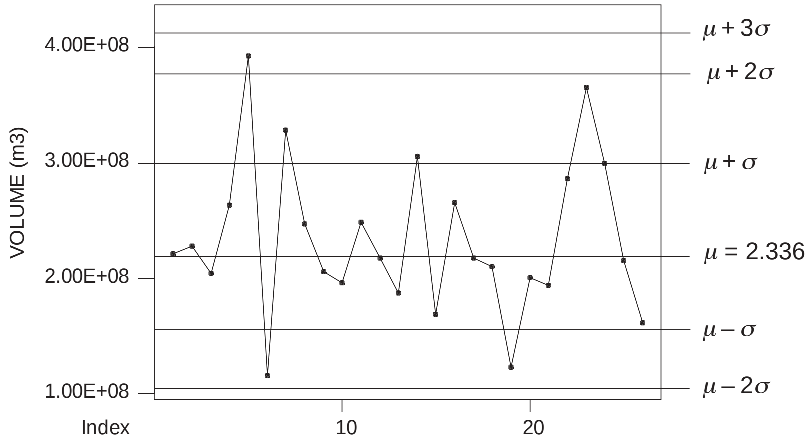

Figure 1: The Time Series Plot of Langat River

rejected. In testing the independence, there is an equal number of flows are above and below the median and therefore the null hypotheses is also not be rejected that is the null hypotheses of independence of successive flows is validated by the data.

The first step in this modeling process is to estimate m, the size of trimmed sample. In this study, the m value for Langat River was estimated as 7, 9, 10 and 12 and the wide range of m values tested was to acquire a clear understanding of the model performance and the quantiles are estimated as follows:

For quantiles, $\hat{x}_p{}^b$ calculation of low quantile of the Langat River is, by taking $m = 7$ and using equation 4, then

$$\hat{x}_p{}^b = 0.000 + 240.043\left(\frac{1}{26}\right)^{1/4.098}$$

$$= 125 \times 10^6 \text{ m}^3$$

$$\hat{x}_p{}^b = 0.000 + 240.043\left(\frac{2}{26}\right)^{1/4.098}$$

$$= 143 \times 10^6 \text{ m}^3 \text{ etc.}$$

The detail results of the low flow quantiles are given in Table 2.

Table 2: The Estimation of Low Flow Quantiles of Langat River at Rantau Panjang, Selangor

	Parameters			Quantiles, $\hat{x}_p{}^b$							
m^a (1)	$\hat{a}$ (2)	$\hat{b}$ (3)	$\hat{c}$ (4)	1/26 (5)	2/26 (6)	3/26 (7)	4/26 (8)	5/26 (9)	6/26 (10)	7/26 (11)	8/26 (12)
7	0.00	240.04	4.09	125	143	156	165	172	179	184	190
9	0.00	230.96	5.26	124	141	153	165	168	175	180	184
10	0.00	223.26	5.70	126	142	152	160	167	173	177	181
12	0.00	222.05	5.31	120	137	147	156	163	168	173	178

m^a Defined the value $x_{(m)}$ at which the sample is trimmed.
$\hat{x}_p{}^b$ The quantiles are defined in terms of the Weibull plotting positions ($i/N + 1 = 1/26$) units are in m^3.
The observed values of the quantiles from the flow volume record in Table 2 are 115.7, 123.3, 169.0, 187.0, 194.2, 196.4, 204.6 and 200.5.

Figure 2, 2(a), 2(b), 2(c) and 2(d) show a plot in logarithm scales of flow quantiles corresponding to $p = P(X \leq x)$ and it can be seen that the approximation of the lower quantiles are excellent for all *m* values.

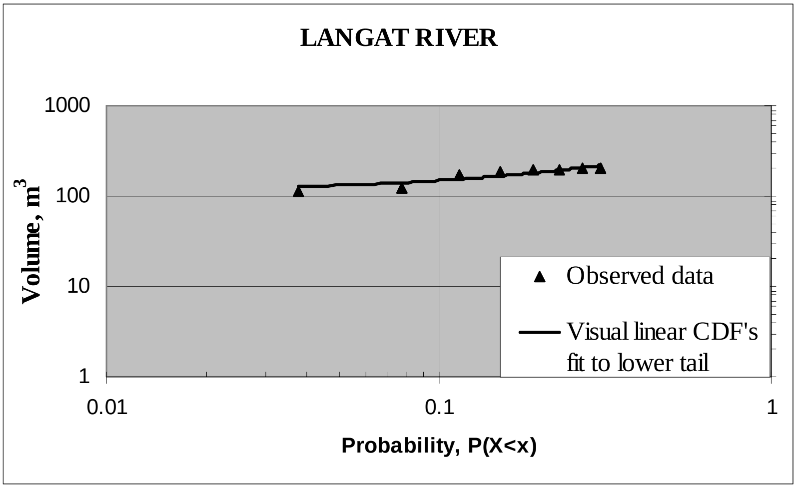

Figure 2: Observed Data for Langat River Corresponding to Probability, $P(X \leq x)$

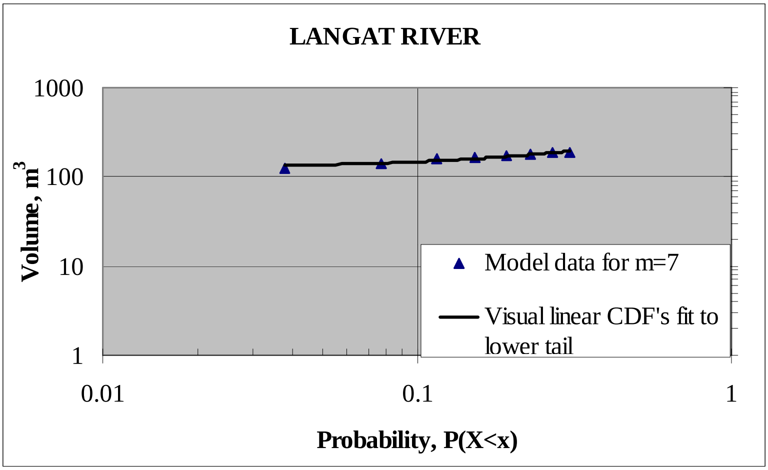

Figure 2(a): Flows Quantiles Corresponding to Probability, $P(X \leq x)$

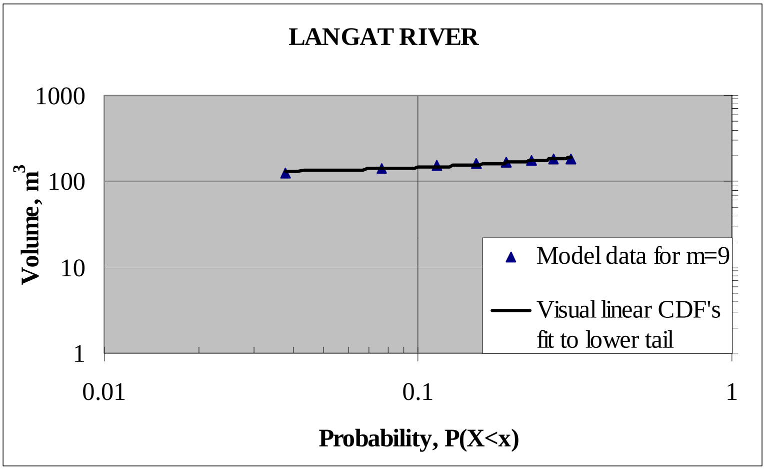

Figure 2(b): Flows Quantiles Corresponding to Probability, $P(X \leq x)$

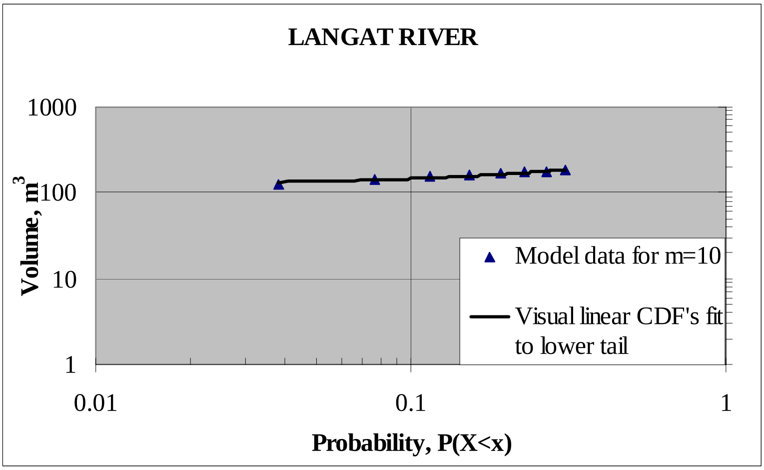

Figure 2(c): Flows Quantiles Corresponding to Probability, $P(X \leq x)$

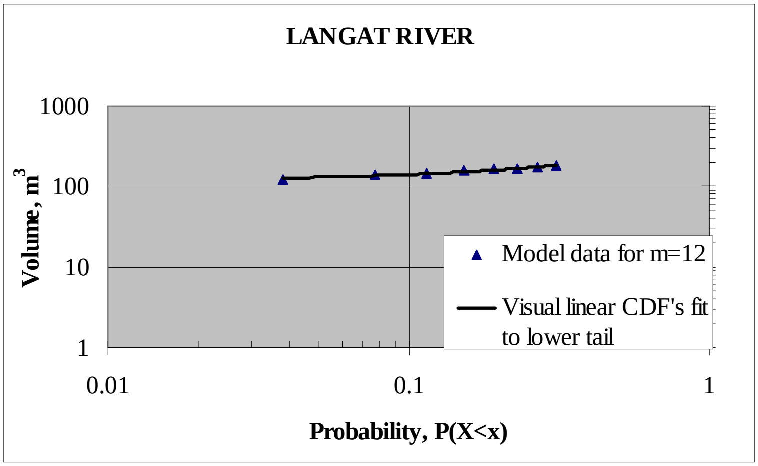

Figure 2(d): Flows Quantiles Corresponding to Probability, $P(X \leq x)$

Conclusion

There is evidence in the research literature that points to the severe difficulty to estimate accurately low quantiles and the parameteric models based on common distributions models usually fail to provide a good estimate. It is apparent that this model is suitable to estimate the lower tail of the cumulative density functions of streamflows and it is only based on the assumptions that the flow are stationary and independent.

References

[1] Riggs, H. C. 1972. *Low-flow investigations. U.S. Geological Survey techniques for water resources investigations,* Book 4, Chapter B1, U.S. Geological Survey, Washington, D.C.

[2] MacMahon, T. A. 1972. *Low flow analyses of streams* - *Details of computational procedures and annotated bibliography.* Research Report No. 5, Department of Civil Engineering, Monash Unversity, Clayton, Australia.

[3] Hugo, A. L. and Miguel, A. M. 1988. *Fitting minima of Flows via Maximum Likelihood. ASCE*, pp. 78-89.

[4] Wald, A. and Wolfowitz, J. 1940. On a test whether two samples are from the same population. *Ann. Math. Statistics*, 11, pp. 147-162.

[5] Rockette, H., Antle, C. and Klimbo, L. A. 1974. Maximum estimation with the Weibull model. *Journal of American Statist. Assoc.*, 69 (345), pp. 246-249.

[6] David, H. A. 1981. *Order statistics*. John Wiley and Sons, Inc., New York, N. Y. 2nd Ed.

Appendix 1

Time Stationary Distribution

MINITAB MACRO:

```
note Test for time stationary
note ========================
name c1 'flow'
let k1=n(c1)
set c2
1:k1
end
name c2 'time'
rank c1 c3
name c3 'rank'
let c4=(c3-c2)**2
name c4 'D'
let c5=sum(c4)
name c5 'Dbar'
let c6=(k1*((k1*k1)-1))/6
name c6 'Ep'
let k3=(k1*k1*(k1+1)*(k1+1)*(k1-1))/36
let c7=sqrt(k3)
name c7 'sigmap'
let c8=(sum(c4)-c6)/c7
note Z=(Dbar-Ep)/sigmap
name c8 'Z'
prin c8
note use normal table to find prob(Z)
note if prob(Z) is between -1.28 and 1.28
note Do not reject null hypothesis
note The sample is time-stationary
note
```

Independence Flow

```
Program by Minitab;
note Test for independence
note =====================
let k4=median(c1)
runs k4 c1
note if p >alpha, do not reject null hypothesis
note The sample is independence
```

Appendix 2

Parameter Estimation using FOTRAN Program

This program is used to find the parameter of the $\hat{a}, \hat{b}$ and $\hat{c}$:

```
      PROGRAM LOWFLOW
      DOUBLE PRECISION X(1000),B, C, SUM, A, D
      INTEGER I, J, N, M
      OPEN (UNIT=1, FILE='SgLangat.dat')
      PRINT*,'ENTER N'
      READ*,N
      PRINT*,'ENTER M'
      READ*,M
      DO 10 I=1, N
      READ (1,*) X (I)
10    CONTINUE
C       SORT X FROM SMALLEST TO LARGEST
            DO 20 J=1, N-1
            XTEMP=X(J)
                  DO 30 I=J+1,N
                        IF (X(I).LT. X=TEMP) THEN
                        XTEMP2=XTEMP
                        XTEMP=X(I)
                        X(I)=XTEMP2
                        ENDIF
30                CONTINUE
                        X(J)=XTEMP
20          CONTINUE
            SUM=0
            A=0
      DO 40 I=1, M
            SUM=SUM+LOG ((X (M)-A)/(X (I)-A)
40    CONTINUE
            C=M/SUM
            D=1/DFLOAT(C)
            B=(((N-M)/DFOLAT (M))**D)*(X (M)-A)
      PRINT*,'                    VALUES OF A, B, C,'
      PRINT*,'                    ================,'
      PRINT 15, A, B, C
15    FORMAT (2X, 3F15.8)
      STOP
      END
```