MODELING INFLOWS INTO THE GARIEP DAM

S.J. Bekker, J.H. Pretorius & D.J. de Waal
Dept. Mathematical Statistics & Actuarial Sciences
University of the Free State
Bloemfontein

ABSTRACT

ESKOM has a hydropower station at the Gariep dam. Predictions of the inflows are necessary for ESKOM to manage the water level. In this paper we consider three possible distributions to model the annual and monthly inflows.

1. INTRODUCTION

Spillage at the Gariep Dam in the Orange river in South Africa causes financial losses to ESKOM, the main supplier of electricity in Southern Africa. De Waal and Verster (2009) have shown that if the loss due to spillage is taken as approximately R4,35/1000m^3 then the total financial loss from 1970 to 2006, of which 13 years recorded spillage, amounts to R 76 950 708. ESKOM manage the outlet of water through 4 hydro turbines each having an outlet capacity of 200 m^3/s (Department of water affairs website). The aim is to manage the outlet of water through the turbines such that the risk of spillage is minimized given the constraints specified by the Department of Water Affairs and Forestry (DWAF). The DWAF formulated a curve on the water level of the dam such that there is water available for irrigation purposes downstream. ESKOM is free to use water above the curve, but if it gets below the curve, restrictions are imposed on the amount of water let out (De Waal and Verster, 2009). The inflows into the dam are usually high in the summer months, October to April, when heavy rains can occur in the catchment areas, while the demand for electricity is usually high in the winter months, June to August. Furthermore, it has been shown by De Waal (2009) that the Southern Oscillation Index (SOI) has a significant correlation with the rainfall in the catchment areas of the Gariep dam and therefore also with the inflows into the Gariep dam.

To place the problem of spillage into perspective, consider that the Gariep dam has a total capacity of approximately 5 500 million m^3 (Department of water affairs website). Therefore, in the event of a flood as was recorded in 1988, when the inflows for February and March was 4147.5 million m^3 and 4886.7 million m^3 respectively, ESKOM stand to lose a significant amount of money in the form of lost electricity that could have been generated.

We will consider three possible distributions to model the inflow into the Gariep dam while incorporating the effect of the SOI. In Section 2 we will consider the LogNormal model as it was discussed by de Waal (2009) as a starting point in predicting the total inflow for each month. In Section 3 we will consider the effect of using the SOI of different months in predicting the inflow for each month separately. In Section 4 the Beta Type 2 distribution is considered as a possibility to model the inflow of separate months. Another approach, the Weibull distribution will be considered in Section 5.

2. MONTHLY PREDICTIONS USING THE LOGNORMAL MODEL

De Waal (2009) introduced the LogNormal model with posterior predictive density to predict the total inflow for a year given the observed value of the previous year's October SOI. The LogNormal distribution is used because by the Central Limit Theorem totals will always tend to a Normal distribution but the inflow data is always positive and skewed to the right.

De Waal (2009) showed that if we let Y be equal to the log of the annual inflows, expressed in millions, then Y follows a Normal distribution. Expressing the mean as a linear combination of the observed SOI for October of the previous year.

$\mu = \beta_0 + \beta_1 \times X$, X is used to represent the observed SOI value

Estimating $\underline{\beta}$ by linear regression: $\beta_0 = 8.63216$ & $\beta_1 = 0.033108$

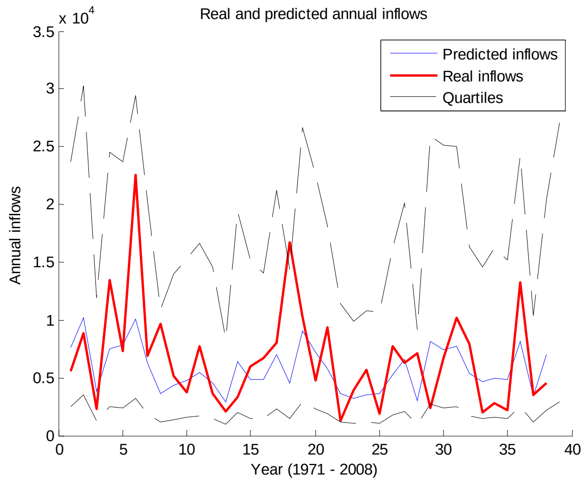

Figure 1: Real and predicted annual inflows with the 1st and 3rd quartiles shown

Using Y_0 to represent the expected inflow for 2010 and $\tilde{X}$ to represent the observed SOI value for October 2009 of -14.7, the posterior predictive density of $Y_0|\tilde{X}$ is $t_\nu(\hat{\mu}_0, \hat{\sigma}^2{}_0)$,where

$$\nu = n - 2 = 38 - 2 = 36 \qquad ; \qquad \hat{\sigma}^2{}_0 = \frac{\nu}{\nu - 2} S^2\left(1 - \tilde{X}M^{-1}\tilde{X}^T\right)^{-1}$$

$$S^2 = \frac{\left(Y - X\hat{\beta}\right)^T\left(Y - X\hat{\beta}\right)}{\nu} \qquad ; \qquad M = X^T X + \tilde{X}^T \tilde{X}$$

$\mu = \beta_0 + \beta_1 \times -14.7 = 8.14547$, so that the prediction for 2010 can be obtained as

$$E(Inflow_{2010}) = 3447.733715\ million\ m^3$$

By adjusting a standard t-distribution the 1st and 3rd quartiles can be obtained:

$$Q_1 = 2275.465782\ million\ m^3$$
$$Q_3 = 5223.902203\ million\ m^3$$

By using this model a correlation of approximately 0.54 can be obtained.

However, ESKOM needs the predictions to be made for each month separately and not the year as a whole. In order to do this we fitted an 11th degree polynomial to the monthly inflow data of 1971 – 2008 and calculated the weight given to each month. By using these weights, given in Table 1, the expected annual inflow can then be subdivided into the expected inflow for each month. Figure 2 shows the best fit 11th degree curve that can be obtained by using the data.

Table 1: Weights assigned to each month and the expected inflow for 2010

Month	**January**	**February**	**March**	**April**
Weight	0.126049	0.180328	0.160997	0.088928
Inflow (m^3)	434.581977	621.722299	555.074134	306.599189
Month	**May**	**June**	**July**	**August**
Weight	0.046610	0.027112	0.014139	0.026830
Inflow (m^3)	160.697673	93.475253	48.748080	92.501124
Month	**September**	**October**	**November**	**December**
Weight	0.043741	0.079946	0.101756	0.103566
Inflow (m^3)	150.806877	275.631342	350.828080	357.067688

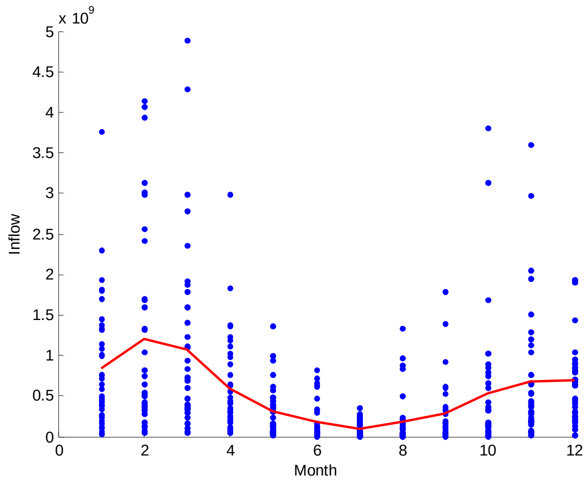

Figure 2: Best fit 11th degree polynomial

3. USING THE SOI OF DIFFERENT MONTHS

Table 2 shows the correlation of the observed SOI of different months of the previous year with the inflow of each month. It can be seen that even though the SOI for October is the best to use when predicting the total inflow for a given year it is not necessarily the best when predicting the inflow for each month separately.

Table 2: Correlation of previous year's SOI with monthly inflows

		Monthly Inflows											
		J	F	M	A	M	J	J	A	S	O	N	D
Monthly SOI	J	-0.10786	-0.01634	0.097161	0.019078	-0.00842	0.130112	0.30067	0.119562	0.183825	-0.03067	0.088766	0.056822
	F	0.007331	-0.13746	0.052967	0.068914	-0.02025	0.000724	-0.0455	-0.25085	-0.06901	-0.14719	0.07961	-0.01268
	M	0.207012	0.162363	0.259955	0.092075	0.284699	0.219934	0.294779	0.103223	0.071286	0.132212	0.231581	-0.05689
	A	0.211532	0.009135	0.121671	0.153531	0.33601	0.148348	0.28642	-0.03721	0.032671	0.182905	0.305574	-0.01435
	M	0.203586	0.018229	-0.02571	-0.02899	0.195836	0.054142	0.269736	-0.01513	-0.19326	0.131004	0.063541	-0.15315
	J	0.277537	0.072216	0.114244	0.098009	0.360758	0.27153	0.296696	0.268812	0.072998	0.345438	0.005456	-0.22738
	J	0.397783	0.23228	0.196007	0.106126	0.350775	0.330896	0.409552	0.14253	-0.05865	0.26848	0.199176	-0.12969
	A	0.475379	0.292549	0.248289	0.165798	0.465124	0.350596	0.22769	0.048502	-0.10292	0.18176	0.303956	-0.07177
	S	0.494792	0.302105	0.239705	0.252096	0.579682	0.39841	0.385287	0.130748	-0.04473	0.141961	0.254266	-0.01699
	O	0.450802	0.385429	0.39836	0.247993	0.61092	0.4062	0.398565	0.131292	0.123508	0.2377	0.247888	-0.08567
	N	0.446409	0.379765	0.295524	0.146357	0.516379	0.356877	0.194494	0.190871	0.01055	-0.01851	0.206714	0.03237
	D	0.526294	0.352744	0.305237	0.261966	0.573484	0.35358	0.230601	0.017379	-0.16432	0.097686	0.160524	-0.00404

However, the extra difficulty in predicting all the extra SOI values is considerable and the results obtained when this was done were not worthwhile, so for this paper we only consider the use of the October SOI values in all the models discussed.

4. THE BETA TYPE 2 DISTRIBUTION

If the monthly inflow data is assumed to follow a Beta Type 2 distribution it can be transformed to a Beta Type 1 distribution by performing the following transformation:

$$Y = \frac{X}{1-X} \sim Beta\ Type\ 1\ , \qquad where\ X \equiv Monthly\ Inflows \sim Beta\ Type\ 2$$

A QQ-plot and histogram with a pdf overlay can then be drawn to assess the fit of the data to the Beta Type 1 distribution. This has been done for each month and the resulting graphs for January are shown in Figures 3 and 4. From the histogram (Figure 3) it can clearly be seen that the pdf of the Beta Type 1 distribution does not fit the transformed data very well. The QQ-plot (Figure 4) indicates that the Beta Type 1 distribution has problems in the upper extremes. Due to the bad fit, the Beta Type 2 distribution cannot be used to model the monthly inflows and it will not be discussed further in this paper.

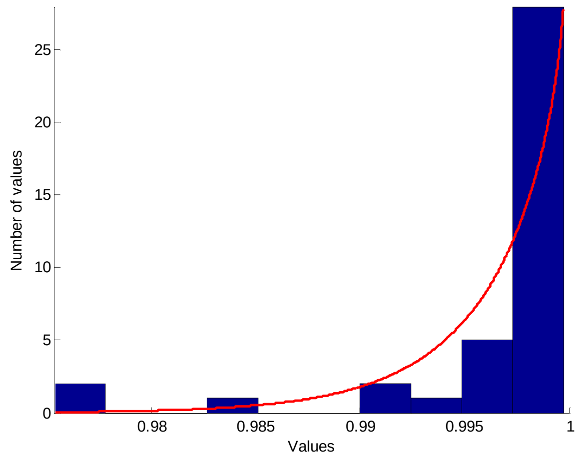

Figure 3: Histogram of transformed inflow data for January plotted with the pdf of the Beta Type 1 distribution.

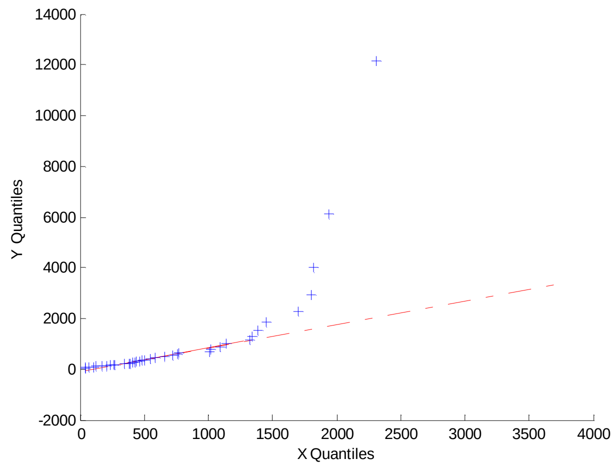

Figure 4: QQ-plot of the inflow data for January and the Beta Type 1 distribution.

5. THE WEIBULL DISTRIBUTION

Assuming that the monthly inflow data follows a Weibull distribution instead of a Beta Type 2 distribution as in Section 4, a histogram and a QQ-plot can be drawn for each month to assess the fit. From Figures 5 and 6 it can be seen that the Weibull distribution is a good candidate when trying to model the monthly inflows.

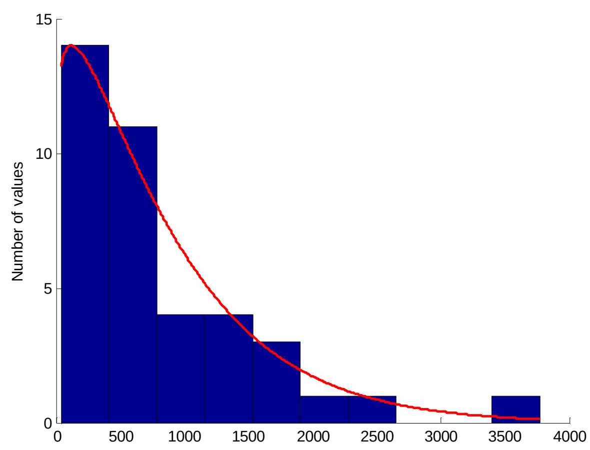

Figure 5: Weibull pdf plotted on the histogram on the monthly inflows for January

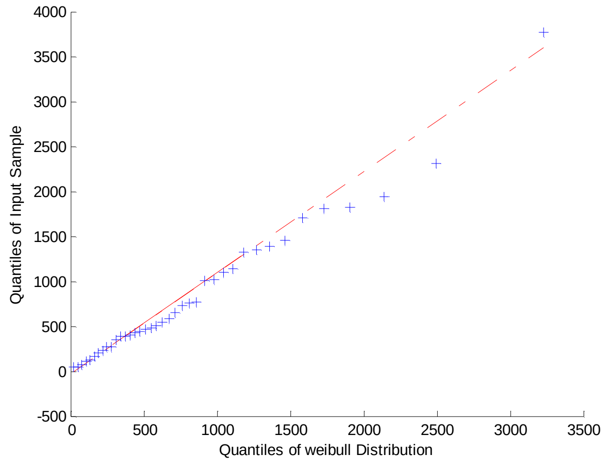

Figure 6: QQ-plot of the Weibull distribution against the monthly inflow data for January.

Let $Y = -\log(Monthly\ Inflows)$ then $Y \sim GBG(\mu, \sigma, \kappa, \xi)$, where GBG stands for the Generalized Burr Gamma distribution (Beirlant, *et. al.*, 1999). If $\kappa = 1$ and $\xi = 0$ then $Y \sim Weibull(\mu, \sigma)$ where μ denotes the mean of Y and σ denotes the standard deviation of Y, which is assumed to stay constant. Therefore μ and σ can easily be determined. In order to incorporate the impact of SOI it is necessary to express μ as a linear function of the observed SOI value of the previous October: $\mu = \beta_0 + \beta_1 \times SOI$. Now let $Z = \frac{Y-\mu}{\sigma}$ then $W = \psi(1) - Z \times \sqrt{\psi'(1)} \sim Exp(1)$ where $\psi(x)$, also known as the digamma function, denotes the logarithmic derivative of the gamma function. Therefore $\psi(1) = -0.57722$ and $\psi'(1) = \frac{\pi^2}{6}$. Thus $f(w) = \exp(-w)$ and using Jeffrey's prior $\left(\frac{1}{\sigma}\right)$ the posterior of Y can be determined as $\exp(-\Sigma w) \times \left(\frac{1}{\sigma}\right)^{n+1} \times \psi'(1)^{\frac{n}{2}}$, where n is the length of Y.

The Gibbs sampling method can then be used to estimate the values of β_0 and β_1 so that the expected inflow for any month can be determined as $\exp\left[-(\beta_0 + \beta_1 \times SOI)\right]$ for a given October SOI value. This process was repeated for each month. The expected annual inflow can then be obtained by summing over all the months. Figure 7 shows the expected annual inflow for different values of the SOI.

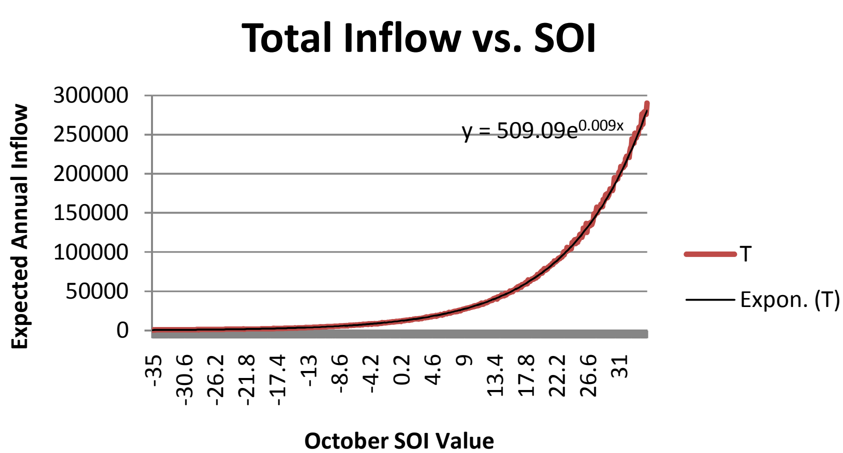

Figure 7: Expected annual inflow at a given previous October SOI value with an exponential curve fitted

Using the Weibull model, a table has been constructed of the expected inflow for each month at all possible values of the October SOI, a part of which is shown in Table 3. It is also possible to fit an exponential curve to the predictions that will make it easier to obtain a rough estimate of the prediction at a given SOI value, this curve is also shown in Figure 7.

Table 3: Expected inflows for specific October SOI of the previous year

SOI	EXPECTED INFLOW								
	J	**F**	**...**	**J**	**A**	**...**	**N**	**D**	**TOTAL**
-35	38.39	42.58	...	45.77	43.94	...	42.57	46.44	259.70
-34.9	36.65	47.64	...	46.23	48.40	...	47.98	47.33	274.22
-34.8	45.77	43.00	...	47.38	43.18	...	42.07	43.25	264.65
...			...			...			
-0.1	952.31	916.01	...	908.24	1081.94	...	964.03	931.89	5754.41
0	1031.93	1040.30	...	975.54	935.61	...	958.59	889.55	5831.53
0.1	1014.21	903.00	...	938.38	963.24	...	941.31	1031.50	5791.65
...			...			...			
34.8	21700.20	25940.24	...	22042.40	25159.04	...	22265.70	21779.07	277235.38
34.9	23275.42	24241.28	...	21955.24	20884.89	...	21937.41	22808.52	275684.61
35	23199.61	23837.65	...	24545.93	23899.46	...	25499.38	22564.28	290128.31

6. CONCLUSION AND RECOMMENDATIONS

Using the LogNormal model discussed in Section 2 a correlation of 0.5408, between the real and expected annual inflows, can be obtained. The Weibull model discussed in Section 5 can be used to predict the monthly and annual inflows given the observed value of the previous October's SOI, using this method a correlation of 0.5735 has been obtained. It may also be of interest to consider the effect of other ecological covariates on the inflows and to build a model that takes this into account.

6. REFERENCES

Beirlant, J., De Waal, D.J., Teugels, J.L., 1999. The generalized Burr-Gamma family of distributions with application in extreme value analysis. In: I. Berkes, E. Csáki, M. Csörgö, eds. 2002. *Limit Theorems in Probability and Statistics I.* Budapest, pp. 113-132.

Department of Water Affairs, *Gariep Dam* URL: http://www.dwa.gov.za/orange/mid_orange/gariep.aspx [Accessed on 20 August 2010]

De Waal, D.J., 2009. Posterior predictions on river discharges. In: M.J. Kallen & S.P. Kuniewski, eds. 2009. *Risk and Decision Analysis in Maintenance.* Amsterdam: IOS Press BV, pp.48-52.

De Waal, D.J. & Verster, A., 2009. Losses due to Spillage at Gariep Dam. In: Van Gelder, Proske & Vrijling, *7th International Probabilistic Workshop.* Delft, 25-26 November 2009, TU Delft: Delft.