Volume : 1 | Issue : 5 | May 2012 ISSN - 2250-1991

Research Paper Engineering

Study of Precipitation and Stream Flow Data- A Case Study of Kim Basin

* Prashant A. Ramani

* PG Student, Civil Engineering Department, GEC,SURAT-395001

ABSTRACT

Rainfall -runoff which is a non linear complex phenomenon and whose predictions are to be required as it is demanding and challenging, especially for the country having the outsized agricultural sector like India. These models are conventionally assigned to one of three broad categories: deterministic (physical), conceptual or parametric (also known as analytic or empirical) (Anderson and Burt, 1985; Watts, 1997). In this study Rainfall-runoff linear regression done by method of least squares, that can be used to provide reliable and accurate estimates of runoff. Also, the goodness of fit carried out of various distributions like Normal, Log Normal, Log Normal (3P) etc. It is done by using rainfall-runoff data of ten years (2001-2010) Kim basin.

Keywords : Linear Regression, least squares, goodness of fit

INTRODUCTION

The water resources systems are very complex due to the variety of objectives, conflicting nature of water uses, and their impact on socio-political environment because of the fact that the hydrological variables like precipitation, runoff, evaporation etc. The relationship of rainfall-runoff is known to be highly non-linear and complex. The rainfall-runoff relationship is one of the most complex hydrologic phenomena to understand because of various reasons such as uncertainty in the rainfall, uneven pattern of rainfall, variations with respect to space and time, etc. Rainfall -runoff which is a non linear complex phenomenon and whose predictions are to be required as it is demanding and challenging, especially for the country having the outsized agricultural sector like India. These models are conventionally assigned to one of three broad categories: deterministic (physical), conceptual or parametric (also known as analytic or empirical) (Anderson and Burt, 1985; Watts, 1997). For the better management of water it is always require estimating the quantities in advance. Kim River is West flowing River in Gujarat State. For its effective management it is very appropriate to predict the contributions due to rainfall in it. Therefore the present study was undertaken in order to develop rainfall-runoff linear regression that can be used to provide reliable and accurate estimates of runoff. Also, the goodness of fit carried out of various distributions like Normal, Log Normal, Log Normal (3P) etc.

2. STUDY AREA AND OBJECTIVES

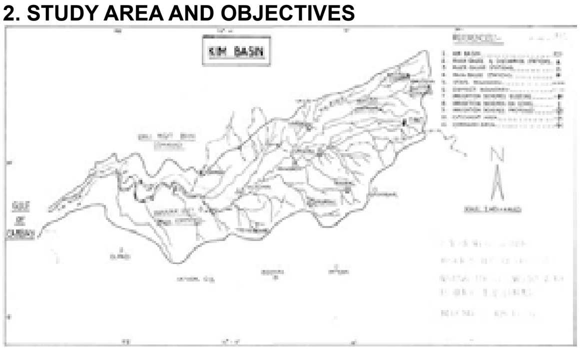


Figure 1: kim basin index map

Kim River is one of the west flowing rivers in Gujarat state. It originates from Saputara Hill ranges in Bharuch district and falls in Gulf of Khabhat near Village Kantiajal of Hansot taluka of Bharuch district after flowing south west direction for a length of 107 km. The river Kim, for the first 80 km, of its course passes through Rajpipala and Valia talukas. For the remaining the river flows in a western direction between Ankleshwar and Olpad taluka of Surat District. The main tributaries of Kim River are Ghanta River and Tokri River. The river basin extends over an area of 1286 sq km of which the catchment area up to the site is 117.9 sq km. The river basin lies between 21^0 33^0 40^0 North latitude and 73^0 12^0 14^0 West longitudes. The winter season is the most pleasant in the basin. The maximum, minimum temperature at site Dehali varies from 27^0 C to 44^0 C and 26^0 to 10^0 respectively

Objectives

I. To Suggest optimum density of gauges as per IS Standard and from Adequacy of Rain Gauge Station for Kim Basin.

II. To determine linear regression by method of least squares for Annual Rainfall-Annual Runoff and for Average Annual Rainfall and Average Annual Runoff of Kim Basin.

III. To Test the Goodness of fit of Normal, Log Normal and Log Normal (3P) Distribution through Easy-fit Software for Sminrov-Kolmogrov test and Chi-Squared test.

3. LITRATURE REVIEW

F.H.S. Chiew, T.C. Piechot, J.A. Dracup, T.A. McMahon (1997) did study of El Nino/Southern Oscillation and Australian rainfall, stream flow and drought: Links and potential for forecasting. El Nino/Southern Oscillation (ENSO) have been linked to climate anomalies throughout the world. Renato Coppi, Pierpaolo D'Urso, Paolo Giordani, Adriana Santoro (2006) determined Least squares estimation of a linear regression model with LR fuzzy response. Pierpaolo D'Urso, Adriana Santoro (2006) tested Goodness of fit and variable selection in the fuzzy multiple linear regression. A. S. Korkhin (2009) developed **linear regression with non stationary variables and constraints on its parameters.** Masashi Sugiyam, Shinichi Nakajima (2009) studied **Pool-based active learning in approximate linear regression.** Guochang Wang, Nan Lin, Baoxue Zhang (2011) developed Functional linear regression after spline transformation. Ciprian Doru Giurcneanu, Sayed Alireza Razavi, Antti Liski (2011) **Review: Variable selection in linear regression: Several approaches based on nor-**

 ISSN - 2250-1991

malized maximum likelihood. Turkey Özlem TERZİ* and Sadık Önal (2011), Application of artificial neural networks and multiple linear regression to forecast monthly river flow in Turkey.

METHODOLOGY

Finding out optimum density of gauges as per IS Standard and from Adequacy of Rain Gauge Station for Kim Basin.

From practical consideration of Indian conditions, the Indian Standard (IS: 4987-1968) recommends the following densities as sufficient.

I. In plains: 1 station per 520 km²

II. In regions of average elevation 1000m: 1 station per 260-390 km²; and

III. In predominantly hilly areas with heavy rainfall: 1 station per 130 km²

$$N = \left(\frac{C_v}{\varepsilon}\right)^2$$

...Rain gauge Station from below Equation

Where N = optimal number of stations, ε = allowable degree of error in the estimate of the mean rain fall and C_v = coefficient of variation of the rain fall values at the existing rainfall values $P_1, P_2,...,P_i,...,P_m$ in a known time.

Determining linear regression by method of least squares for annual rainfall- runoff and for average annual rainfall-runoff of Kim Basin.

If the trend is a straight line, the relationship is linear and has the equation

$$y = a + bx$$

The least squares line equation may be obtained by solving for a and b the two normal equations

$$\sum y = na + b\sum x$$

$$\sum xy = a\sum x + b\sum x^2$$

Testing the Goodness of fit of Normal, Log Normal and Log Normal (3P) Distribution through Easy-fit Software for Sminrov Kolmogrov and Chi-Squared test.

RESULT AND DISCUSSION

Optimum Density of Rain gauge Station.

Year:2001 to 2010	Total Catchment area km²	No. of Rain gauge required as per Indian Standard (IS:4987-1968)	Actually provided Rain gauge	No. of Rain gauge require for 10% error
1	1286	2	3	1
2	1286	2	3	4
3	1286	2	3	2
4	1286	2	3	2
5	1286	2	3	5
6	1286	2	3	3
7	1286	2	3	3
8	1286	2	3	1
9	1286	2	3	2
10	1286	2	3	9

Table 1: No. of Optimum Stations for Various Years Kim Basin

5.2 Linear regression by method of least squares

1. for Annual Rainfall-Runoff.

...least square based on following equations.

$$y = a + bx \quad (1)$$

The least squares line equation (1) may be obtained by solving for a and b the two normal equations

$$\sum y = na + b\sum x \quad (2)$$

$$\sum xy = a\sum x + b\sum x^2 \quad (3)$$

Where n = number of pairs of observed values of x and y.

Solution

$$y = 2.86x - 2656$$

$$R = 2.86P - 2656$$

Where R= Runoff, P= rainfall, both r in mm.

Above equation determined from the below calculation.

Table 3: Regression line for the Annual Rainfall-Runoff Data for River Kim (2001-2010)

Rainfall in mm	Runoff in mm	$\bar{x}$	$\bar{y}$	$x^2 \times 10^4$	$xy \times 10^4$	$\Delta x = x - \bar{x}$	$\Delta y = y - \bar{y}$	$(\Delta x)^2 \times 10^4$	$(\Delta y)^2 \times 10^4$	$\Delta x . \Delta y \times 10^4$		Working
x	y											
2813.6	830	3633.4	5041.5	791.63	233.52	-819.83	-4211.5	67.21	1773.67	345.27		$\bar{x}$ = Σx/n = 36334.34/10 = 3633.43 mm
2618.54	865			685.67	226.5	-1014.89	-4176.5	103	1744.31	423.86		
4811	585			2314.57	281.44	1177.57	-4456.5	138.66	1986.03	-524.78		$\bar{y}$ = Σy/n = 50415/10 = 5041.50
3751	11268			1407	4226.62	117.57	6226.5	1.38	3876.93	73.2		
5172	5879			2674.95	3040.61	1538.57	837.5	236.71	70.14	128.85		Normal Equation
5273.2	13217			2780.66	6969.58	1639.77	8175.5	268.88	6683.88	1340.59		50415 = 10a + 36334b....(i)
2604.7	5735			678.46	1493.79	-1028.73	693.5	105.82	48.09	-71.34		20715 × 10^4 = 36334a + 14333.56 × 10^4
3996	8034			1596.8	3210.38	362.57	2992.5	13.14	895.5	108.49		
2755.5	774			759.27	213.27	-877.93	-4267.5	77.07	1821.15	374.65		Dividing throughout by 10^4
2538.8	3228			644.55	819.52	-1094.63	-1813.5	119.82	328.87	198.51		20715 = 3.63a + 14333b.......(ii)
36334.34	50415			14333.56	20715.23			1131.69	19228.57	2397.3		solving (i) and (ii)
												a = -2656, b = 2.118
		correlation coefficient, r = Σ(Δx. Δy) / (Σ(Δx)² .(Δy)²)½										Regression line is
												y = 2.86x - 2656
		r = 2397.3 × 10^4 / (1131.69 × 19228.57 × 10^4 × 10^4)½										R = 2.86P - 2656
												Where, R and P are in mm
		r = 0.514										

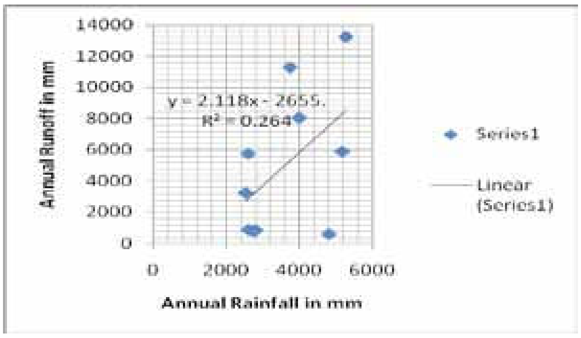


Figure 2: Linear Regression of Annual Rainfall-Runoff

Standard error of estimate for Annual Rainfall-Runoff.

$$S_{y.x} = \sigma_y\sqrt{1-r^2}$$

$$\sigma_y = \sqrt{\frac{1922.86\times10^4}{10-1}} = \sqrt{\frac{\Sigma(\Delta y)^2}{n-1}} = 1462\text{mm}$$

$$S_{y.x} = 1462\sqrt{1-(0.514)^2} = 1254\text{mm}$$

2. for Average Annual Rainfall-Runoff.

$$y = 0.504x - 274.7$$

$$R = 0.504P - 274.7$$

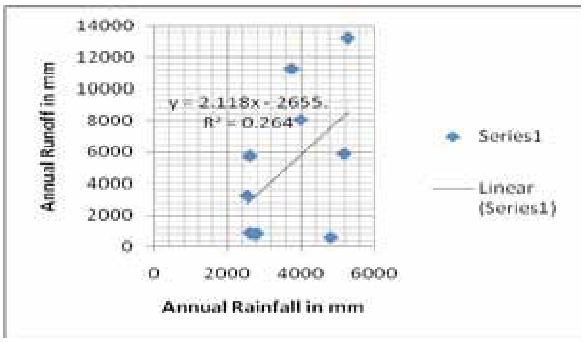
Where R= Runoff, P= rainfall, both r in mm.

Above equation determined from the below calculation.

Table 4: Regression line for the Average Annual Rainfall-Runoff Data for River Kim (2001-2010)

Rainfall in mm x	Runoff in mm y	$\bar{x}$	$\bar{y}$	$x^2 \times 10^4$	$xy \times 10^4$	$\Delta x = x - \bar{x}$	$\Delta y = y - \bar{y}$	$(\Delta x)^2 \times 10^4$	$(\Delta y)^2 \times 10^4$	$\Delta x.\Delta y \times 10^4$	Working
937.86	69.16	1211.1	335.65	87.95	6.48	-273.281	-266.491	7.46	7.1	7.28	$\bar{x} = \Sigma x/n = 12111.41/10 = 1211.1$ mm
872.84	72.12			76.18	6.29	-338.301	-263.531	11.44	6.94	8.91	
1603.66	48.77			257.17	7.82	392.519	-286.881	15.4	8.23	-11.26	$\bar{y} = \Sigma y/n = 3356.51/10 = 335.65$
1250.33	93.9			156.33	11.74	39.189	-241.751	0.15	5.8	-0.94	
1724	489.96			297.21	84.46	512.859	154.309	26.3	2.38	7.91	Normal Equation
1757.73	1101.48			308.96	193.61	546.589	765.829	29.87	58.64	41.85	3356 = 10a + 12111b....(i)
868.23	477.94			75.38	41.49	-342.911	142.289	11.75	2.02	-4.87	$469.76 \times 10^4 = 12111a + 1592 \times 10^4$
1332	669.57			177.42	89.18	120.859	333.919	1.46	11.15	4.03	
918.5	64.54			84.36	5.92	-292.641	-271.111	8.56	7.35	7.93	Dividing throughout by 10^4
846.26	269.07			71.61	22.77	-364.881	-66.581	13.31	0.44	2.42	470 = 1.21a + 1592b.......(ii)
12111.41	3356.51			1592.57	469.76			125.7	110.05	63.26	solving (i) and (ii)
											a = -274.7, b = 0.504
		correlation coefficient, $r = \Sigma(\Delta x.\Delta y) / (\Sigma(\Delta x)^2.(\Delta y)^2)^{1/2}$									Regression line is
											y = 0.504x - 274.7
		$r = 63.26 \times 10^4 / (125.7 \times 110.05 \times 10^4 \times 10^4)^{1/2}$									R = 0.504P - 274.7
											Where, R and P are in mm
		r = 0.537									

Figure 3: Linear Regression of Average Annual Rainfall-Runoff



Standard error of estimate for Average Annual Rainfall-Runoff.

$$\sigma_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n-1}} = \sqrt{\frac{\Sigma(\Delta y)^2}{n-1}}$$

$$= \sqrt{\frac{110.11\times10^4}{10-1}} = 349\text{mm}$$

$$S_{y.x} = 349\sqrt{1-(0.537)^2} = 294.41\text{mm}$$

The Goodness of fit of Normal, Log Normal and Log Normal (3P) Distribution by Easy-fit Software for Sminrov-Kolmogrov test at 10 % Significance Level.

For Annual Runoff.

Sr. No.	Distribution	Parameters	Remarks
1	Lognormal	s=1.1616 m=7.9731	Where, s, m, g parameters of Distribution known as Standard Deviation, Mean, Kurtosis Co-efficient.
2	Lognormal (3p)	s=5.8187 m=5.9677 g=585.0	
3	Normal	s=4622.2 m=5041.5	

Table 5: Parameters of Different Distribution

Sr. no.	Distribution	Smirnov Kolmogorov		Remarks
		Statistic Δ	Rank	
1	Lognormal	0.25129	2	Critical Value Δ_o= 0.37 for, n = 10 and α = 0.10 from table A in Appendix Hence $\Delta < \Delta_o$,
2	Lognormal (3P)	0.35035	3	Critical Value Δ_o = 0.37 for, n = 10 and α = 0.10 from table A in Appendix
3	Normal	0.21689	1	Critical Value Δ_o= 0.37 for, n = 10 and α= 0.10 from table A in Appendix Hence $\Delta < \Delta_o$,

Table 6: Smirnov Kolmogorov Test for Annual Runoff For Average Annual Runoff.

Sr. no.	Distribution	Parameters	Remarks
1	Lognormal	s=1.0961 m=5.2581	Where, s, m, g are parameters of Distribution known as Standard Deviation, Mean, Kurtosis Co-efficient.
2	Lognormal (3P)	s=6.0052 m=3.9542 g=48.77	
3	Normal	s=349.79 m=335.65	

Table 7: Parameters of Different Distribution

Sr. no.	Distribution	Smirnov Kolmogorov		Remarks
		Statistic Δ	Rank	
1	Lognormal	0.24316	1	Critical Value Δ_o = 0.37 for, n = 10 and α = 0.10 from table A in Appendix Hence $\Delta < \Delta_o$,
2	Lognormal (3P)	0.32106	3	Critical Value Δ_o = 0.37 for, n = 10 and α = 0.10 from table A in Appendix Hence $\Delta < \Delta_o$,
3	Normal	0.25526	2	Critical Value Δ_o = 0.37 for, n = 10 and α = 0.10 from table A in Appendix Hence $\Delta < \Delta_o$

Table 8: Smirnov Kolmogorov Test for Average Annual Runoff

The Goodness of fit of Normal, Log Normal and Log Normal (3P) Distribution by Easy fit Software for Chi-Squared test at 10 % Significance Level, for 8 degrees of freedom. For Annual Runoff

Sr. no.	Distribution	Chi-Squared		Remarks
		Statistic χ^2	Rank	
1	Lognormal	1.8838	2	Critical Value χ^2_o = 13.36 for, n = 10, α = 0.10 and V = 8 from table B Hence $\chi^2 < \chi^2_o$
2	Lognormal (3P)	N/A		N/A
3	Normal	0.77384	1	Critical Value χ^2_o = 13.36 for, n = 10, α = 0.10 and = 8 from table from table B Hence $\chi^2 < \chi^2_o$

Table 9: Chi-Squared Test For Average Annual Runoff.

Sr. no.	Distribution	Chi-Squared		Remarks
		Statistic χ^2	Rank	
1	Lognormal	1.5814	2	Critical Value χ^2_o = 13.36 for, n = 10, α = 0.10 and V = 8 from table B Hence $\chi^2 < \chi^2_o$
2	Lognormal (3P)	N/A		N/A
3	Normal	0.40822	1	Critical Value χ^2_o = 13.36 for, n = 10, α = 0.10 and V = 8 from table from table B Hence $\chi^2 < \chi^2_o$

Table 10: Chi-Squared Test for Average Runoff

CONCLUSION

Minimum 1 Rain gauge Station and Maximum 9 Rain gauge Station Required for Kim Basin.

Coefficient of correlation, r = 0.514 for Annual Rainfall and Annual Runoff r→1, indicates a close linear relationship. Coefficient of correlation, r = 0.537 for Average Annual Rainfall and Average Annual Runoff. r→1, indicates a linear relationship.

For Annual Runoff, Critical Valu Δ_0 = 0.37, Δ = 0.2519, $\Delta < \Delta_0$, Critical Value Δ_0= 0.37, Δ = 0.35035, $\Delta < \Delta_0$, Critical Value Δ_0 = 0.37, Δ = 0.21689, $\Delta < \Delta_0$ with respect to Log normal, Log normal (3p) and Normal distribution ranked 2, 3 and 1 respectively, hence accept the hypothesis for each three distribution by Sminrov-Kolmogorov Test.

For Average Annual Runoff, Critical Value Δ_0= 0.37, Δ = 0.24316, $\Delta < \Delta_0$, Critical Value Δ_0 = 0.37, Δ = 0.32106, $\Delta < \Delta_0$, Critical Value Δ_0 =0.37, Δ = 0.25526, $\Delta < \Delta_0$ with respect to Log normal, Log normal (3p) and Normal distribution ranked 1, 3 and 2 respectively, hence accept the hypothesis for each three distribution by Sminrov-Kolmogorov Test.

For Annual Runoff, Critical Value χ_0^2 = 13.36, χ^2 = 1.8838, $\chi^2 < \chi_0^2$, Critical Value χ_0^2 = 13.36, χ_0^2 = 0.77384, $\chi^2 < \chi_0^2$, with respect to Log normal, and Normal distribution ranked 2 and1 respectively, hence accept the hypothesis for each two distribution by Chi-Squared Test. Log Normal (3P) distribution is not applicable in Chi-squared test because Kurtosis Co-efficient (□□ is based on 4th moment which cannot be estimated reliably unless the sample size is Large.

For Average Annual Runoff, Critical Value χ_0^2 = 13.36, χ^2 = 1.5814, $\chi_0^2 < \chi_0^2$, Critical Value χ_0^2 = 13.36, χ^2 = 0.40822, $\chi_0^2 < \chi_0^2$, with respect to Log normal, and Normal distribution ranked 2 and1 respectively, hence accept the hypothesis for each two distribution by Chi-Squared Test. Log Normal (3P) distribution is not applicable in Chi-squared test because Kurtosis Co-efficient (□□ is based on 4th moment which cannot be estimated reliably unless the sample size is Large.

Smirnov-Kolmogorov test has advantage over the Chi-Square test in that it does not lump the data and compare only discrete categories, but it rather compares all the data in an unaltered form as well as more convenient when the size of sample is small.

ACKNOWLEDGEMENT

Authors are thankful to Dr. Surat, B.M.Vadher, Principal, GEC, Suart, Dr. Akshi Agrawal, Director, Gujarat Technological University who have permitted us to work together.

REFERENCES

PAPERS : F.H.S. Chiew, T.C. Piechot, J.A. Dracup, T.A. McMahon (1998) done study of El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting. Journal of Hydrology, Volume 204, Issues 1–4, 30 January 1998, Pages 138-149. ~ Elias Masrya, Jan Mielniczuk (1999), Local linear regression estimation for time series with long-range dependence, Journal of Stochastic Processes and their Applications, Volume 82, 1999, February Pages 173-193. ~ Renato Coppi, Pierpaolo D'Urso, Paolo Giordani, Adriana Santoro (2006) determined Least squares estimation of a linear regression model with LR fuzzy response. Journal Computational Statistics & Data Analysis, Volume 51, Issue 1, November, 2006. ~ A. S. Korkhin (2009) developed Linear regression with nonstationary variables and constraints on its parameters, Journal Cybernetics and Systems Analysis, Volume 45 Issue 3, May 2009. ~ Masashi Sugiyam, Shinichi Nakajima (2009) studied Pool-based active learning in approximate linear regression, Journal Machine Learning, Volume 75 Issue 3, June 2009. ~ Ciprian Doru Giurcneanu,Sayed Alireza Razavi, Antti Liski (2011) Review: Variable selection in linear regression: Several approaches based on normalized maximum likelihood, Journal Signal Processing, Volume 91 Issue 8, August, 2011. ~ Turkey Özlem TERZİ* and Sadık Önal (2012) Application of artificial neural networks and multiple linear regression to forecast monthly river flow in Turkey, African Journal of Agricultural Research Vol. 7(8), pp. 1317-1323, 26 February, 2012.
ARTICLES : Pierpaolo D'Urso, Adriana Santoro (2006) tested Goodness of fit and variable selection in the fuzzy multiple linear regression, Fuzzy Sets and Systems, Volume 157, Issue 19, 1 October 2006, Pages 2627-2647. ~ Guochang Wang, Nan Lin, Baoxue Zhang (2012) developed Functional linear regression after spline transformation, Computational Statistics & Data Analysis, Volume 56, Issue 3, 1 March 2012, Pages 587-601.
WEB SITE : http://www.google.com ~ http://www.guj-nwrws.gujarat.gov.in ~ http://www.nwda.gov.in ~ http://www.Sciencedirect.com ~ http://www.mathwave.com ~ http://www.academicjournals.org ~ http://www.dl.acm.org
BOOKS : K Subramanya, Engineering Hydrology, 2nd addition, Tata McGraw Hill Publishing Company, 2004, pp 15,28-29. ~ Jayarami Reddy P., Stochastic Hydrology, Laxmi Publication, 2003, pp 103-104, 107-108, 253-255. ~H. M. Raghunath, Hydrology: principles. Analysis, design, revised Edition New Age International Publication, 2007, pp 341-346.