International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 3.45 (SJIF-2015), e-ISSN: 2455-2585
Volume 4, Issue 5, May-2018

Rainfall-Runoff relationship by Linear Regression Analysis

Mr. Ramani Prashant[1], Mr. Aankit Doni[2]

[1]*Civil engineering & gsfc university, prashant.ramani@gsfcuniversity.ac.in*

[2]*Civil engineering & gsfc university, aankit.doni@gsfcuniversity.ac.in*

Abstract*— The Rainfall-Runoff is a non-linear complex phenomenon. Predictions of rainfall-runoff is required as it's demand frequently raising since past few years due to change in hydrological condition, especially for the country like India having huge agricultural sector which depends upon monsoon cycle for large crop yields but same time it is challenging to do so. The current study focuses univariate linear regression and multivariate linear regression analysis of rainfall-runoff.

***Keywords*— Rainfall, Runoff, Kim River, Univariate Linear Regression, Multivariate Linear Regression**

I. INTRODUCTION

The water resources systems are very complex as it is depending upon the various hydrological variables viz. precipitation, runoff, evaporation, evapotranspiration, infiltration, v etc. 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 geographical location, climate change etc. For the better management of water, it is always requiring prediction of the rainfall-runoff trend in advance. The present study is done for Kim river basin for the rainfall-runoff data of year 2001-2010 of Kosamba, Kim, Tadkeswar rain gauge station of Kim river basin. Kim River is one of the west flowing rivers in Gujarat state. Therefore, the present study was undertaken in order to develop rainfall-runoff univariate linear regression & multivariate linear regression that can be used to provide reliable and accurate prediction of runoff.

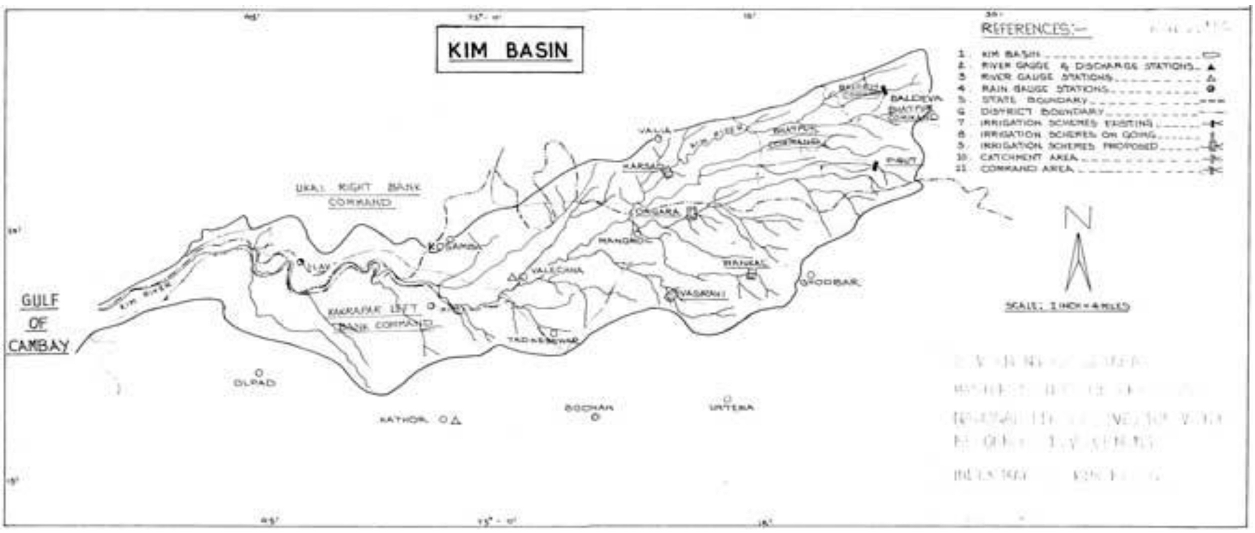

Figure 1: Kim River Basin (source: http://www.guj-nwrws.gujarat.gov.in)

II. LITERATURE REVIEW

A. Review from Journals

F.H.S. Chiew, T.C. Piechot, J.A. Dracup, T.A. McMahon (1998) studied about El Nino/Southern Oscillation and Australian rain- fall, stream flow and drought: Links and potential for forecasting and El Nino/Southern Oscillation (ENSO) have been linked to climate anomalies throughout the world.

Elias Masrya, Jan Mielniczuk (1999) studied Local linear regression estimation for time series with long-range dependence.

Renato Coppi, Pier- paolo 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.

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 ap- proximate 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) reviewed Variable selection in linear regression and several approaches based on normalized maximum likelihood.

Turkey Özlem TERZİ and Sadık Önal (2012) studied application of artificial neural networks and multiple linear regression to forecast monthly river flow in Turkey.

Guochang Wang, Nan Lin, Baoxue Zhang (2012) developed Functional linear regression after spline transformation.

B. **Univariate Linear regression analysis by Least Square Methods**

Two variables y (dependent) and x (independent) can be correlated by plotting them on x and y axis. If they plot on a straight line there is a close linear relationship; on another hand if the points depart appreciably (without a definite trend), the graph is called scatter diagram or plot.

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

$$y = ax + b \text{ ....(1)}$$

Where, y= Runoff (dependent variable), x= Rainfall (independent variable), a & b = constant

Number of lines can be obtained depending on the values of a and b. the method of least squares is used to select the line that fits the data best. The principal of least squares states that the best line for fitting a series of observation is the one for which the sum of the squares of the departures is minimum. A departure is the difference between the observed value and the line. Since x is the independent variable, the departures of y are used.

$$\sum y = na + b\sum x \text{...........(2)}$$

$$\sum xy = a\sum x + b\sum x^2 \text{..........(3)}$$

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

The most commonly used statistical parameter for measuring the degree of association of two linearly dependent variables x and y, is the correlation coefficient.

$$r = \frac{\sum(\Delta x.\Delta y)}{\sqrt{\sum(\Delta x)^2.(\Delta y)^2}} \text{.........(4)}$$

$$r = \frac{\sum xy - n\overline{x\ y})}{(n-1)\sigma_x\sigma_y} \text{.........(5)}$$

Where, $\Delta x = x - \overline{x}, \Delta y = y - \overline{y}$

$\sigma_x\sigma_y$ = standard deviations of x and y, respectively,

x, y= middle of each class interval, respectively,

If r = 1, the correlation is perfect giving a straight line plot (regression line),

r = 0, no relation exists between x and y (scatter plot),

r→1 indicates a close linear relationship.

C. **Multivariate Linear regression analysis**

Multivariate linear regression analysis consist one dependent variable and multiple independent variable,

$$y = f(p_1, p_2, \dots p_n) \ldots\ldots\ldots(6)$$

Multivariate regression analysis can be done by given equation

$$y = a + b_1p_1 + b_2p_2 + b_3p_3 \ldots\ldots b_np_n \ldots\ldots\ldots(7)$$

Where, y = dependent variable, a = intercept, b_1 to b_n =regression coefficient p_1 to p_n

III. RESULTS

A. **Univariate Linear regression analysis**

Figure 2 & 3 shows the univariate linear regression analysis for annual data & seasonal data (June-Sep.) of Kim river basin.

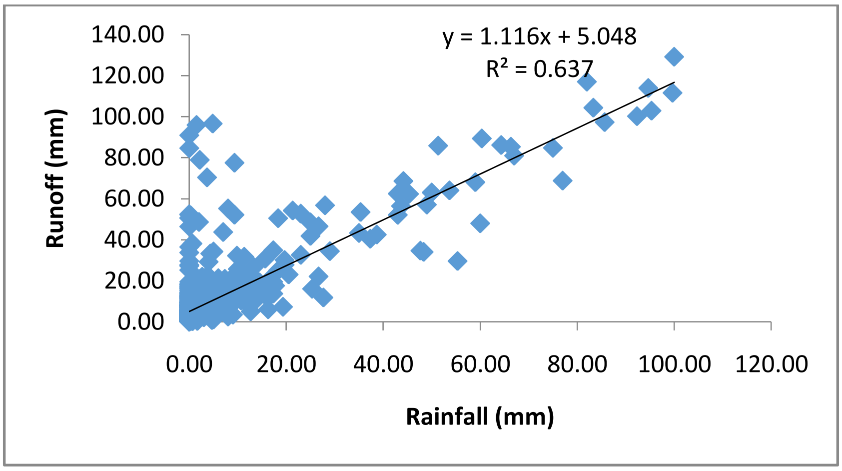

Figure 2: Scatter Diagram Rainfall-Runoff Linear for Annual data of Kim River Basin

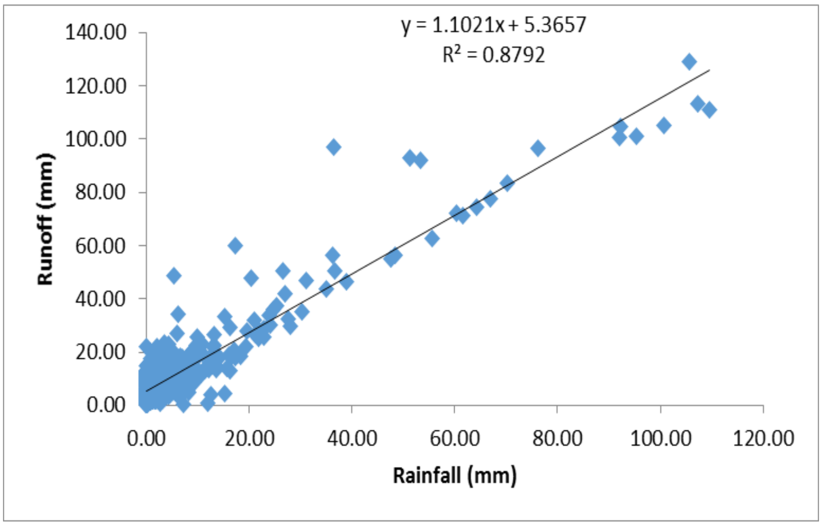

Figure 3: Scatter Diagram Rainfall-Runoff Seasonal (June-Sep.) data of Kim River Basin

B. Multivariate Linear regression analysis

Multiple Linear Regression analysis has been done for Kim Basin. Table 1 and 2 shows Rainfall-Runoff Relation and also the statistics of different variables for Annual Data & Seasonal (June-Sep.) for Kim Basin.

TABLE 1

MULTIVARIATE LINEAR REGRESSION OF ANNUAL DATA

	Value	**Standard error**
Intercept	a = 11.16001	2.68023
Kosamba **(P1)**	b1 = 0.04078	0.12177
Kim **(P2)**	b2 = 0.27668	0.16547
Tadkeshwar **(P3)**	b3 = 0.02161	0.12563
Previous day Runoff **(Qt-1)**	b4 = 0.55175	0.02383
Hence the Equation, Y = 11.16 + 0.04P1 + 0.28P2 + 0.027 P3 + 0.55 Qt-1		
R-Square **(Coefficient of Determination)**	0.6127	

TABLE 2

MULTIVARIATE LINEAR REGRESSION OF SEASONAL (JUNE-SEP.) DATA

	Value	**Standard error**
Intercept	a = 2.03708	0.73944
Kosamba **(P1)**	b1 =0.51838	0.0617
Kim **(P2)**	b2 = -0.32218	0.0866
Tadkeshwar **(P3)**	b3 = 1.05403	0.06657
Previous day Runoff (Qt-1)	b4 = 0.45084	0.01351
Hence the Equation, Y = 2.037+ 0.51P1 – 0.32 P2 + 1.054 P3 + 0.45 Qt-1		
R-Square **(Coefficient of Determination)**	0.8404	

IV. RESULT DISCUSSION & CONCLUSIONS

1. A univariate linear regression analysis done with average daily rainfall as independent and daily runoff as dependent parameter and the value of the coefficient of determination R^2 is 0.63 & 0.87 respectively for annual data and seasonal (June-Sep.) data.
2. A multivariate regression analysis is carried out with four independent variables: precipitation at Kim, Kosamba and Tadkeshwar rain gauge stations are used as three independent variables and the runoff of the previous day is used as the fourth independent variable. The value of the coefficient of determination R^2 is equal to 0.61 & 0.84 respectively for annual data and seasonal (June-Sep.) data.
3. Seasonal (June- Sep.) data for the monsoon season taken separately gives a higher value coefficient of determination for both univariate linear regression and multivariate linear regression analysis than the annual data.

4. SCOPE OF THE WORK

1. Univariate Linear regression analysis can be done by other methods viz. the method of moments and the method of maximum likelihood.
2. Multivariate Linear Regression analysis can be done by obtaining more independent variable (if data is available) viz. evaporation of water from stream, infiltration rate etc.

REFERENCES

[1] F.H.S. Chiew, T.C. Piechot, J.A. Dracup, T.A. McMahon, *El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting*, Journal of Hydrology, vol. 204, issues 1–4, pp. 138-149, January 1998.

[2] Elias Masrya, Jan Mielniczuk, *Local linear regression estimation for time series with long-range dependence, Journal of Stochastic Processes and their Applications*, vol. 82, pp. 173-193, February 1999.

[3] Renato Coppi, Pierpaolo D'Urso, Paolo Giordani, Adriana Santoro, *Least squares estimation of a linear regression model with LR fuzzy response. Journal Computational Statistics & Data Analysis*, vol. 51, issue 1, November, 2006.

[4] Pierpaolo D'Urso, Adriana Santoro, *Goodness of fit and variable selection in the fuzzy multiple linear regression, Fuzzy Sets and Systems*, vol. 157, issue 19, pp. 2627-2647October 2006,.

[5] A. S. Korkhin, *Linear regression with nonstationary variables and constraints on its parameters, Journal Cybernetics and Systems Analysis*, vol. 45, issue 3, May 2009.

[6] Masashi Sugiyam, Shinichi Nakajima, *Pool-based active learning in approximate linear regression, Journal Machine Learning*, vol. 75 issue 3, June 2009.

[7] Ciprian Doru Giurcneanu,Sayed Alireza Razavi, Antti Liski, *Review: Variable selection in linear regression: Several approaches based on normalized maximum likelihood*, Journal Signal Processing, vol. 91, issue 8, August 2011.

[8] Turkey Özlem TERZİ* and Sadık Önal, *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, February 2012.

[9] Guochang Wang, Nan Lin, Baoxue Zhang, *Functional linear regression after spline transformation, Computational Statistics & Data Analysis*, vol. 56, issue 3, pp. 587-601, March 2012,.

[10] K Subramanya, *Engineering Hydrology*, 2004, 2nd addition, Tata McGraw Hill Publishing Company.

[11] Jayarami Reddy P., *Stochastic Hydrology*, Laxmi Publication, 2003.

[12] H. M. Raghunath, *Hydrology: principles. Analysis, Design*, revised Edition New Age International Publication, 2007.

[13] http://www.mathwave.com.