International Journal of Scientific & Engineering Research Volume 10, Issue 6, June-2019
ISSN 2229-5518

DEVELOPING STAGE - DISCHARGE MODEL USING LEAST SQUARE METHOD FOR BARO GAUGING STATION

NGBAA DORNU SIMEON
Federal University of Technology Owerri

Abstract— Developed Stage-Discharge Models for Baro is presented based on the statistical method of least squares to estimate the potential discharge volume from the sub basin. Missing records were determined by the mass curve analysis for periods that stage and discharge records were not available. Test on time homogeneity showed that the interaction of stream Stage and stream discharge exhibited common profile but non-stationary. The trend of the annual subset mean and the coefficient of variation for stage and discharge along the river decrease downstream. Six models were developed, calibrated and validated for suitability using six set of equations namely; simple linear, polynomial of second degree (quadratic), power (geometric), exponential, logarithmic and modified power. Comparison of the values of their coefficient of correlation (R), coefficient of determination (R2), standard error of estimate (SE), hypothesis test, calibration/validation test and error function test (SEF) favoured model (VI) for Baro gauging station. The application of the models showed that [Q = 1883.22(h - 180)0.8794] for Baro should be adopted. Thus, from the analysis it can be ascertained that stage-discharge model for Baro along river Niger followed a modified power model (model VI) of the form [Q = α*(h – h0)b] and the stage at zero discharge in the river "ho" is a hypothetical value that cannot be measured in the field. Thus, the regression model is robust enough to perform satisfactorily in a catchment such as the Niger River and can therefore be used as a complimentary tool in the estimation of the discharge potential. Recommendations are made that hydrological modeling be applied in water resources planning and management activities for environmental sustainability.

Index Terms— Coefficients of Determination, Correlation Coefficient, Data, Discharge, Error Estimate, Least square, Mass curve analysis, Model, Regression, Stage, and Standard Hydrology.

1 INTRODUCTION

FOR an effective and sustainable water resource planning and management, the need for complete and reliable hydrological data is indisputable. To get such data there is a need to develop and maintain hydrological stations in proper networks but since the availability of this data in sizes and number of years required for effective planning and management of water resources system is somehow difficult, there is need to develop a model whereby such data can be used to estimate these data that will enhance the management of water resources system.

Models can be expressed mathematically to represent a system or sets of data; Models are also seen as mathematical representations of sets of relationships between variables or parameters (Nwaogazie, 2006; Nwadike, 2008). In this study, we shall be looking at Mathematical models as representing a set of variables which establishes relationships between stage and discharge variables.

Hydrologic data are the building blocks for management of any water resources abstraction. Many sources of data may be accessed to support model development and verification, statistical analyses, and other studies (Viessman, and Lewis, 2008). The unavailability of long term hydrological data such as discharge measurements and invariably lack of discharge models has been the major difficulties encountered by engineers and hydrologists in design and planning of water resources structures in developing countries like Nigeria and other underdeveloped countries (Ojha et al., 2008; Viessman and Lewis, 2008; Raghunath, 2006 and Sonuga,) 1990. Since most of the river basins in Nigeria do not have data especially discharge in sizes and number of years required for hydrological study, the use of discrete statistical analysis to determine appropriate and suitable model that can be used to generate this data was necessary. Thus, measurements of river stage provide the best alternative of measuring discharge.

The study reported here emanated from global concern to generate data for water resources management and here in Nigeria one of the key imperatives of the federal ministry of water resources is; Establishing the means to acquire, collate, manage and disseminate hydrological, hydro meteorological and hydro geological information for each of the river basins in Nigeria. This research study is set to provide tool to achieve part of the derived objectives of the federal water road map.

The stage-discharge (H-Q) relationship is a fundamental technique employed in discharge calculation. Typically, the relationship is established from periodic measurements of stream discharge and corresponding water surface elevation, height or stage.

The main objectives of this study was to establish a functional stage-discharge relationship between stage-discharge in order to generate the discharge potential from a given stage measurement.

The general objectives include:

1. to examine the relationship between stage and discharge for Baro gauging station along the river Niger;
2. to develop a model for Baro gauging station along river Niger using the least square method of regression analysis;
3. determine which model out of the six model's setup is best in terms of accuracy, simplicity, and applicability for educational and research purposes;

4. to validate the adopted model and determine if the data it produces is reasonable by comparing the results of the validation to the actual.

The relationship between stage (m) and the corresponding discharge (m3/s) is an exponential function (Ojha et al., 2008). Strong nonlinear bivariate and multivariate correlations are common in hydrology and various mathematical models are used to describe the relationships, such as parabolic, exponential, hyperbolic, power, and other forms that provide better graphical fits than straight lines. In this study, regression modeling was adopted to fit discharge data from Baro discharge gauging station along Niger River for fifty-three years duration (1960-2012). The approach was to utilize the original data in model calibration and using coefficient of correlation, coefficient of determination and standard error of estimate as evaluation parameters for good curve fitting.

2 MATERIAL AND MATHODS

2.1 Types of and Sources of Data

The data type for this research was basically secondary data. And these include daily archival data on stage (water level) and discharge; these data from the two stations were made available by the Nigeria Institute of Hydrological Services Agency (NIHSA) Abuja.

2.2 Method of Data Quality Analysis

In the preliminary treatment of the data, basics statistical like computation of the totals, mean, variance and standard deviation were employed for the analysis of stage height and discharge in each of the stations.

The analytical techniques employed include mass curve analysis for the determination of missing data, homogeneity and consistency analysis, product moment correlation coefficient test, coefficient of determination regression analysis, hypothesis test and error function test.

2.3 Determination of Missing Records using Mass Curve Analysis

The theory of the mass curve is based on the fact that a graph of the cumulative of one parameter is plotted against the time. This is one of the methods employed to estimate missing value's in hydrology from a set of data.
In this study, the cumulative depth of stage (or discharge) for any year starting from the first year of record was plotted as ordinate against time as abscissa. The plot was extrapolated to periods of no record and the cumulative depth at the extrapolated time read off from the graph. The depth of stage or discharge for year t_2 was obtained by subtracting the cumulative depth at time t1 from that at time t_2.

2.4 Homogeneity and Consistency Analysis

The annual field data collected for Jiderebode, Jebba, Baro, Lokoja and Onitsha hydrological stations were subjected to homogeneity and consistency analysis. In the consistency analysis, the subset mean of annual stage or discharge were compared to the long term mean by calculating the deviation of the subset mean from the long term mean and expressed in percentage.

$$\textbf{Deviation from Mean} = \left[\frac{\textbf{Subset mean} - \textbf{Longterm mean}}{\textbf{Longterm Mean}}\right] \textbf{x} \ \frac{\textbf{100}}{\textbf{1}} \qquad \textbf{1.0}$$

Also, as part of the consistency analysis, coefficient of variation which is a measure of variability was employed to determine the degree of spatial variations in annual stage or discharge in river Niger. The formula is;

$$\mathbf{CV} = \frac{S}{X}\ X\ \frac{\mathbf{100}}{\mathbf{1}} \qquad \mathbf{2.0}$$

Where; C.V = coefficient of variation, S = Standard variation (deviation) and X = Mean (stage or discharge)

2.5 Best-fit Evaluation criteria

The evaluation of Coefficient of determination (R2) is determine by the equation;
It describes the extent of best-fit and is expressed as:

$$R^2 = \frac{[\sum(X_i - \bar{X})(Y_i - \bar{Y})]}{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2} \qquad 3.0$$

(b) Coefficient of correlation (r):

It is the square root of the coefficient of determination (r^2):

$$R = \sqrt{R^2} \qquad 4.0$$

2.6 Tests for Goodness of Fit

The test for goodness of fit was conducted using equation (3.0) and equation (4.0) to test for the coefficient of determination (R^2) and coefficient of correlation (R) respectively. The coefficient of determination, R^2 measures the proportion of the total variation in Y that is explained by the linear model. That is, R2expresses the proportion of the total variation in the values of the variable Y that can be accounted for or explained by a linear relationship with the values of the random variable X. Thus; unlike the value of a test statistic, the coefficient of determination does not have a critical value that enables us to draw conclusions. In general, the higher the value of R^2, the better the model fits the data.
R2 = 1: Perfect match between the line and the data points. $R^2 = 0$; It means the relationship between x and y is zero.

2.7 Model Set-Up

In this research, six set of regression equations were considered to setup the model. These sets of regression equation are; linear, quadratic, power, exponential, Logarithmic, and modified power as given below;

I. $Q = a_0 + a_1 H$ 5.0

II. $Q = a_0 + a_1 H + a_2 H^2$ 6.0

III. $Q = aH^b$ 7.0

IV. $Q = Ae^{bH}$ 8.0

V. $Q = a_0 + a_1 \ln(H)$ 9.0

VI. $Q = a[h - h_0]^b$ 10.0

2.8 Error Function Test

The error functions are used to support interpretation of the results in single estimations. In this study, the judgment of the parameter fit is often obtained by a combination of visual inspection of the graphical display of the estimations and by the use of equation (3.30). The estimated discharge is an important guidance in model calibration.
The r^2 coefficient represented in equation (3.30) indicates the correlation between observed and estimated flows and gives a measure of the correctness of the model (Saelthurn, 1995).

$$r^2 = 1 - \left[\frac{\Sigma(Q_{obs} - Q_{est})^2}{\Sigma(Q_{obs} - Q_{obsaverage})^2}\right] \qquad 11.0$$

Where; Q_{obs} is the observed discharge in m³/s, Q_{est} is the estimated discharge in m³/s; and $Q_{obsaverage}$ is the average of the observed discharge values in m³/s.

2.9 Hypotheses Test

In order to establish an effective relation between stage and discharge, the following null hypothesis and alternative hypothesis are tested.

I. Null hypothesis (H_0): There is no significant relationship between stage (water level) and discharge in each of the study area.

II. Alternative hypothesis (H_a): There is a significant relationship between stage (water level) and discharge in each of the study area.

2.10 The Rejection Rules are (for significance level α)

In this research, the significance level of alpha (α) which is our acceptable risk (probability) of rejecting a true null hypothesis is 5% (0.05). Therefore, our confidence level would be 95% (0.95) which is 1-α.

In other to determine whether or not we can reject the null hypothesis in favour of the alternative hypothesis, the following condition must be ascertained;

1. We reject H_0: $\rho = 0$ in favour of H_a: $\rho \neq 0$ if;

$$h > t\frac{a}{2}, n-2 \qquad 12.0$$

2. Calculate the p-value, and reject H_0: $\rho = 0$ in favour of Ha: $\rho \neq 0$ if;

$$\text{P-value} \leq \alpha \qquad 13.0$$

3 RESULT AND ANALYSIS

3.1 Result of Mass Curve Analysis

Table 1 showed the estimated missing annual discharge data for Baro Hydrological station that were estimated using the mass curve analysis, the estimated annual discharges were marked with asterisk (*) indicated that for the given hydrological year (i.e. 2003 - 2005) there were no available record for stage and discharge data.

TABLE 1
MISSING DATA FOR STAGE AND DISCHARGE AT BARO

S/N	Year	Stage (m)	Discharge (m³/s)
1	2003	1116*	745033*
2	2004	861.7*	560146*
3	2005	820.3*	530562*

3.2 Result of Consistency Analysis

Consistency analysis of Mean Annual Stage and Discharge for Baro

Table 2 showed the result of consistency analysis of the mean annual stage and discharge at Baro, Lokoja station of River Niger. The deviation mean expressed in percentages are in bracket. The subsets of 25 years show that the stage and discharge at this location is time homogenous and time homogenous series are purely random and stationary. The subsets of 10 years are equally consistent and homogeneous for the long term mean for stage and discharge. It is clear that the interaction of stream Stage and stream discharge exhibited common profile i.e. the year of maximum stage depth is the same year of maximum discharge and the year of minimum stage is also the same year of minimum discharge.

TABLE 2
COMPARISON OF SUBSET MEANS WITH LONG TERM MEAN OF STAGE AND DISCHARGE FOR BARO

Term	Stage	Discharge	CV for Stage	CV for Discharge
Long Term	1010.5	693907.6	28.22%	34.19%
1st (10 Years)	1254.4	850142.3		
	(24.14%)	(22.52%)		
2nd (10 Years)	980.8	678172.5		
	(-2.94%)	(-2.27%)		
3rd (10 Years)	789.5	564700		
	(-21.87%)	(-18.62%)		
4th (10 Years)	965.3	675410.5		
	(-4.47%)	(-2.67%)		
5th (10 Years)	1035.8	689063.5		

	(2.50%)	(-0.70%)
1st (25 Years)	1037.1	713868.2
	2.63%	2.88%
2nd (25 Years)	973.2	669127.4
	(-3.69%)	(-3.57)

3.3 Model Set-Up for Baro

Table 3 showed the result of regression analysis, standard error of estimate (SE) and test of hypothesis using student's t distribution for each of model 1, 2, 3, 4, 5 and 6 at Baro gauging station. The result of the regression analysis showed that model 6 has the highest value of coefficient of correlation (R) and coefficient of determination (R^2) (0.9778 and 0.9561 respectively). The result of the standard error of estimate of discharge on stage also, showed that model 6 has the least value of standard error of estimate of 0.0367 (3.67%) as compared to the other models. Again, the result of hypothesis test, for 5% significance level, model 6 has highest t statistic value of 33.31 which was greater than the T-critical (table value of t) of 2.008. Consequently, the null hypothesis was rejected in favour of the alternative hypothesis. Also, model 6 has the smallest P-value (2.82E-36) as compared to other models which is less than 5% significance. Thus, the null hypothesis was rejected in favour of the alternative hypothesis with 95% confidence. Established on the assertions from the analysis, Model 6 (modified power model) should be adopted for prediction of annual discharge at Baro station and the cumulative annual stage at zero discharge is estimated to be 180m which implied that the daily stage at zero discharge is about 0.5m. The observation made from this analysis suggested that "h_0 is a hypothetical value that cannot be measured in the field but estimated by trial and error optimization

The deviation in the coefficient of variation may not be unconnected with other factor such as slope of the river or flow from another nearby river and the river geometry.

TABLE 3
MODEL PERFORMANCE INDICATOR FOR BARO

M	Model's Type	Goodness of Fit (G_F)	Standard Error (S_E)	t-stat	T-Critical	P-Value
1	$Q = 785.62h - 99976$	R = 0.9445 R^2 = 0.8921	79437.80	20.53	2.008	2.58E-26
2	$Q = 0.1778h^2 - 425.69h - 67712$	R = 0.9485 R^2 = 0.8996	77372.45	1.94	2.008	0.058
3	$Q = 182.74h^{1.190}$	R = 0.9709 R^2 = 0.9426	0.042	28.94	2.008	2.60E-33
4	$Q = 17804e^{0.0013h}$	R = 0.9238 R^2 = 0.8534	0.1542	17.23	2.008	6.53E-23
5	$Q = 664069 \ln(h) -4E+06$	R = 0.9030 R^2 = 0.8154	103888.81	15.011	2.008	2.37E-20
6	$Q = 1883.22(h - 180)^{0.8794}$	R = 0.9778 R^2 = 0.9561	0.0367	33.31	2.008	2.82E-36

3.4 Model Calibration and Validation

Model Calibration, Validation and Error Function Test for Baro Gauging Station

Table 4 showed the result of calibration, validation and error function test for River Niger at Baro gauging station with six different models using annual observed/estimated discharge versus annual stage, the coefficient of determination (R^2) which measures the correctness of the model had high and positive values for model 6 [$Q = a(h - h_0)^b$] was 0.9561 this implied that 95.61% of the total estimated variance is explained. The standard error function test (SE.F) for model 6, which measured of the variation between the observed discharge and the estimated discharge had small values of 0.0439 indicates that the regression model fit the data with an error margin of 4.39%. Therefore, the result of these calibration, validation and error function test further confirm that model 6 (modified power model) should be adopted for the prediction of discharge potential for Baro gauging station.

TABLE 4
MODEL CALIBRATION, VALIDATION AND ERROR FUNCTION TEST (USING ANNUAL OBSERVED AND ESTIMATED DISCHARGE VERSUS ANNUAL STAGE CURVE)

Loc.	Terms	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Baro	R^2	0.8921	0.8996	0.9426	0.8534	0.8154	0.9561
	S_{EF}	0.1079	0.1004	0.0574	0.1466	0.1846	0.0439

4 CONCLUSIONS

The stage-discharge models in this research were developed from the annual discharge and annual stage from Baro gauging stations along Niger River.

Given the main focus of the study, Hydrological (stage and discharge) data for Baro gauging stations were collected and analyzed using regression model to estimate the potential discharge volume from the sub basin. The data included daily series of stage, and discharge for some periods in which data were available for five hydrological stations along the River Niger system.

The test for homogeneity and consistency of the stage and discharge records showed that the interaction of stream Stage and stream discharge exhibited common profile and are non-stationary.

The application of the six regression models namely; linear, polynomial of second degree (quadratic), power (geometric), exponential, logarithmic and modified power showed that the modified power model should be used in the predictions of annual discharge at Baro gauging stations. This can be justified, since it was proven with conclusion drawn by other researchers (Boiten, 2000, Herschy, 1999, Rantz, 1982, WMO, 1994, Schmidt, 2002) who studied the relationship between stage and discharge and established that the stage discharge relationship follows a power and modified power model.

The result of calibration, validation and error function test showed that the coefficient of determination has a very high and positive value of 0.0561 which indicated that 95.61% of the total estimated variance is explained.

The standard error function has a small value ranged of 0.0439 indicating an error margin of 4.39. Thus, the regression model is robust enough to perform satisfactorily in a catchment such as the Niger River and can therefore be used as a complimentary tool in the estimation of the discharge potential.

Therefore, the research will play an important role in ensuring that future allocation of water resources in the localities are scientifically based and efficiently used so as to satisfy the needs of both human and natural systems.

Acknowledgment

A study of this nature and scope necessarily relies heavily on the work of others. Some indication of the extent of my indebtedness is given by the references. Several people who were involved in the work gave documentation, textbooks and discussed their experience willingly. To all of them I say a big thank you.

References

[1] [2] Arora, k R. (2007) Irrigation, Water Power and Water Resources Engineering. 4th edition. Delhi: Standard Publisher Distributors.

[2] [3] Barlow, P. M. and Leake, S. A. (2012) ―Streamflow depletion by wells―Understanding and managing the effects of groundwater pumping on streamflow ‖. U.S. Geological Survey Circular 1376, 84

[3] [4] Harshburger, B. J., Von P. W., Karen S. H., Brandon C. M., Troy R. B., and Albert R. (2012) ―Generation of Ensemble Streamflow Forecasts Using an Enhanced Version of the Snowmelt Runoff Model‖. Journal of the American Water Resources Association (JAWRA) 48(4): 643-655.

[4] [5] Bronstert, A., Niehoff, D. and Burger, G. (2002). ―Effects of climate and land-use change on storm runoff generation: present knowledge and modeling capabilities‖. Hydrol. Process. 16, pp.509–529.

[5] [6] Croke, B. F. W., Meritt, W.S. and Jakeman, A.J. (2004). ―A dynamic model for predicting hydrologic response to land cover changes in gauged and ungauged catchments‖. J. Hydrol. 291, pp.115–131.

[6] [7] Ezemonye, M. N. and Emeribe, C. N. (2011) ―Flood characteristics and management adaptations in parts of the Imo River system‖. Ethiopian Journal of Environmental Studies and Management, Vol. 4, No. 3

[7] [8] Isik, S., Kalin, L., Schoonover, J. E., Srivastava, P. and Lockaby, B. G. (2012) ―Modeling effects of changing land use/cover on daily streamflow: An Artificial Neural Network and curve number based hybrid approach‖ Journal of Hydrology, pp. 1-10.

[8] [9] Kalin, L. and Hantush, M. M., (2009) ―An auxiliary method to reduce potential adverse impacts of projected land developments: subwatershed prioritization‖. Environ. Manag. 43 (2), pp.311–325.

[9] [10] Lin, Y. P., Hong, N. M., Wu, P.J., Wu, C.F. and Verburg, P.H., (2007). ―Impact of land use change scenarios on hydrology and land use patterns in the Wu–Tu watershed in Northern Taiwan‖. Landsc. Urban Plann. 82, pp.111–126.

[10] [11] Matthews, O. P. & St. Germain, D. (2007) ―Boundaries and Transboundary Water Conflicts‖.

[11] [12] Journal of Water Resources Planning and Management, 133 (5), pp. 386-396.

[12] [13] Moore, D. S. and McCabe, G. P. (1993) Introduction to the Practice of Statistics. 2nd edition, New York: W.H. Freeman and company.

[13] [14] Nagy, C., Lockaby, B.G., Helms, B., Kalin, L. and Stoeckel, D. (2011) ―Water resources and land use and cover in a humid region: the Southeastern United States‖. J. Environ. Qual. 40, pp. 867–878.

[14] [15] Ojha, C. S. P., Berndtsson, R. and Bhunya, P. (2008) Engineering Hydrology. India: Oxford University Press.

[15] [16] Raghunath, H. M. (2006) Hydrology .2nd edition, New Delhi: New Age International (P) Ltd. Publishers.

[16] [17] Reddy, R. J. (2008) A Text Book of Hydrology. 2nd edition, New Delhi: University Science Press.

[17] [18] Siriwardena, L., Finlayson, B.L., McMahon, T.A., (2006) ‗The impact of land use change on

[18] [19] catchment hydrology in large catchments: the Comet River, Central Queensland, Australia‖. J. Hydrol. 326, 199–214.

[19] [20] Sonuga, J. O. (1990). ―Hydrological Decision Making for Design of Water Resources Systems‖ Proceedings of Nigerian National Committee for International Hydrological Programme (IHP), pp.30-44.

[20] [21] Suresh, R. (2008) Land and water management principles. Delhi: A. K. Jain.

[21] [22] van Dijk, A. I., Pena-Arancibia, J. L. and Bruijnzeel, L. A. (2011) ―Top-down analysis of collated streamflow from heterogeneous catchments leads to underestimation of land cover influence‖. Hydrol. Earth Sci. 8, pp. 4121–4150.

[22] [23] Viessman, W. and Lewis, L. G.(2008) Introduction To Hydrology 5th edition, New Delhi: Prentice-Hall Private Limited.

[23] [24] Wolf, A. T., Stahl, K. & Macomber, M. F. (2003a), Conflict and cooperation within international

[24] [25] river basins: The importance of institutional capacity, Water Resources Update, 125.

[25] [26] Zhao, F., Francis H. S. C., Lu Z., Jai V., Jean-Michel P. and Ming L. (2012) ―Application of a Macroscale Hydrologic Model to Estimate Streamflow across Southeast Australia‖. J. Hydrometeor, 13, pp. 1233–1250