BORNEO SCIENCE 11: 23-31, JUNE 2002

MODELLING BRIDGE PIER SCOUR EQUATION USING REGRESSION METHODS

Ahmad Shukri Yahaya[1], Aminuddin Ab Ghani[1] & Nor Azazi Zakaria[2]

[1]Unit Perkhidmatan Akademik
Universiti Sains Malaysia, Kampus Kejuruteraan, 14300 Nibong Tebal
Seberang Perai Selatan, Pulau Pinang, Malaysia.
[2] Pusat Penyelidikan Kejuruteraan Sungai Dan Saliran Bandar (REDAC)
Universiti Sains Malaysia, Kampus Kejuruteraan, 14300 Nibong Tebal
Seberang Perai Selatan, Pulau Pinang, Malaysia.

ABSTRACT. *There are a number of equations that have been developed to predict local scour depth at bridge piers using experimental data. A few researches have formed pier scour equations using available field data. This paper discusses one of the equations and provides an extended statistical regression analysis on the model. By determining outliers and influential observations in the data, a new regression model was obtained which provides a better prediction for pier scour depth.*

KEYWORDS. Pier scour, sediment transport, outliers, influential observations

INTRODUCTION

The presence of bridge piers causes an abrupt change in the direction of the approach flow resulting in the removal of bed material, thus the problem of local scour at piers is developed. Sediment movement in the approach flow determines whether the local scour is clear water (no sediment transport) or live bed (with sediment transport).

Kafi & Alam (1995) discusses various scour equations which have been developed in the literature. They found out that the accuracy of laboratory based equations were improved if the coefficients and exponents of these equations were derived using field data. Based on the suggestions by Kafi & Alam (1995), Ab. Ghani & Nalluri (1996) developed a number of new equations to predict pier scour using the available field data.

AB. GHANI & NALLURI'S (1996) EQUATIONS

This paper examines one of the equations given by Ab. Ghani & Nalluri (1996) and provides an extended regression analysis of the equation. Two new equations were developed utilising available field data from literature. Applications of dimensional analyses on these data results in the following equations:-

$$\frac{y_s}{d} = 1.46\left(\frac{a}{d}\right)^{0.97}\left(\frac{y_0}{a}\right)^{0.52}\left[\frac{Q}{ay_0\sqrt{gy_0}}\right]^{0.24} \tag{1}$$

with adjusted $r^2 = 0.95$ and

$$\frac{y_s}{d} = 1.46\left(\frac{a}{d}\right)^{0.44}\left(\frac{y_0}{d}\right)^{0.52}\left[\frac{Q}{ay_0\sqrt{gy_0}}\right]^{0.24} \tag{2}$$

with adjusted $r^2 = 0.95$ where a is the pier width, d the mean diameter of the bed material, y_0 the mean flow depth, Q the maximum flood discharge, and g the gravitational constant. Both equations can be approximately transformed to:

$$y_s = 1.46\left[aQ\sqrt{\frac{y_0}{g}}\right]^{1/4} \tag{3}$$

Equation (3) confirms the important effect of flow depth, pier width and flow discharge in computing scour depth.

THE GENERAL LINEAR REGRESSION MODEL

The general linear regression model can be written in matrix form as follows:

$$\underset{(n\times1)}{Y} = \underset{(n\times p)}{X}\,\underset{(p\times1)}{\beta} + \underset{(n\times1)}{\varepsilon} \tag{4}$$

where Y is a vector of responses
β is a vector of parameters
X is a matrix of constants
ε is a vector of errors

By using the least squares method, the values of β can be obtained. However, the general linear regression model must satisfy the following assumptions:

i) The mean of ε is zero.
ii) The variance of ε is constant.
iii) The probability distribution of ε is normal.
iv) The random errors are independent.

OUTLIERS

Outliers are extreme observations, that is, they are cases in the data that are well separated from the rest of the data. To identify outlying X observations, the hat matrix can be used. The hat matrix, H is defined as $\underset{(n\times n)}{H} = X(X'X)^{-1}X'$. Leverage values greater than $\frac{2p}{n}$ are outliers with regard to their X values.

For Y values, the studentised deleted residual method can be used to identify outliers. The studentised deleted residual is given by $d_i^* = e_i\left[\frac{n-p-1}{SSE(1-h_{ii})-e_i^2}\right]^{1/2}$. The t test with $(n-p)$ degrees of freedom can be used to identify the outlying Y values.

INFLUENTIAL OBSERVATIONS

After identifying outlying observations, these observations were checked to determine whether it is influential. An observation is said to be influential if its exclusion causes major changes in the fitted regression model.

To measure the influence of the combined impact of the i^{th} observation on all of the estimated regression coefficients, the Cook's distance measure can be used. The Cook's distance measure is given by:

$$D_i = \frac{e_i^{\,2}}{pMSE}\left[\frac{h_{ii}}{(1-h_{ii})^2}\right] \tag{5}$$

If $D_i = F_{(1-\alpha):(p,n-p)} > 50^{th}$ percentile, than the observation is influential.

The influence on the fitted values can be measured using $DFFITS$. The $DFFITS$ values are given by $(DFFITS)_i = e_i\left[\frac{n-p-1}{SSE(1-h_{ii})-e_i^{\,2}}\right]^{\frac{1}{2}}\left(\frac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}}$. A case is influential if the absolute value of $DFFITS$ exceeds 1 for small to medium size data sets and $2\sqrt{p/n}$ for large data sets.

DATA

The data was obtained at random from field data found in the literature (Kafi & Alam, 1995; Kothyari *et al*., 1992). The independent variables given in the data are flow characteristics, geometry of the piers, bed material size and fluid properties while the dependent variable is the scour depth at piers. The sample size is 40. Table 1 summarises ranges of field data available.

Table 1. Ranges of available field data

Variables	Range
$Q(m^3/s)$	68 - 63763
a(m)	0.92 – 11.27
d(m)	0.10 – 230
y_0(m)	0.62 – 18.36
y_s(m)	1.23 – 35.67

MODEL

Based on Equation (1) developed by Ab. Ghani & Nalluri (1996), the regression odel considered for determining the scour depth around bridge piers is as follows:

$$\frac{y_s}{d} = \beta_0 \left(\frac{a}{d}\right)^{\beta_1} \left(\frac{y_0}{a}\right)^{\beta_2} \left(\frac{Q}{a y_0 \sqrt{g y_0}}\right)^{\beta_3} \tag{6}$$

here Q is the flow discharge $\left(m^3/s\right)$

a is the pier width (m)

d is the sediment size (mm)

y_0 is the flow depth (m)

y_s is the scour depth (m)

g is the gravitational acceleration $\left(m/s^2\right)$

For this model, a transformation is required so that the assumptions of the linear gression model using least squares method is fulfilled. The transformed model is given as llows:

$$\ln\left(\frac{y_s}{d}\right) = \ln(\beta_0) + \beta_1 \ln\left(\frac{a}{d}\right) + \beta_2 \ln\left(\frac{y_0}{a}\right) + \beta_3 \ln\left(\frac{Q}{a y_0 \sqrt{g y_0}}\right) \tag{7}$$

r

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 \tag{7a}$$

RESULTS

The regression model obtained using the least squares method is given by :

$$\ln\left(\frac{y_s}{d}\right) = 0.166535 + 0.969302 \ln\left(\frac{a}{d}\right) + 0.524133 \ln\left(\frac{y_0}{a}\right) + 0.240191 \ln\left(\frac{Q}{a y_0 \sqrt{g y_0}}\right) \tag{8}$$

with the coefficient of determination, $r^2 = 0.955$. The analysis of variance test for model adequacy shows that the model is adequate $(F = 256.73064, p = 0.0)$.

Figures 1 and 2 show the normal probability plot for the standardised residuals and the residual plot against $\hat{y}$ respectively. Figure 1 shows that the residual is approximately normal. However in Figure 2, the standardised residual plot against $\hat{y}$, shows that there might be outliers in the data. Thus tests for outliers and influential observations were carried out.

By using the leverage values, there are three observations which are outliers namely cases 28, 29 and 36 (h_{ii} = 0.21163, 0.61453 and 0.28315 respectively). For detecting outlying y observations, observations 8, 36, 37 and 39 were found to be outliers (d_i^* = 1.95782, -4.05631, -1.92756 and -2.80944 respectively). Please refer to Figure 2 where observations 8, 36, 37 and 39 are shown.

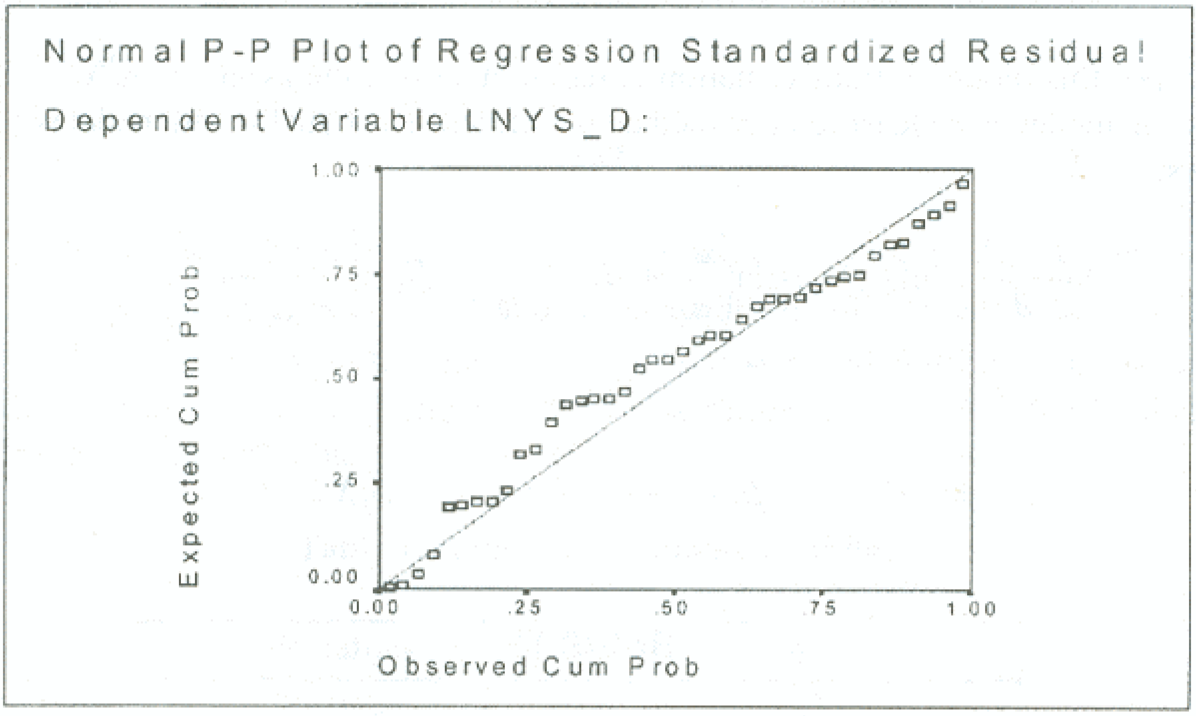

Figure 1. Normal probability plot for standardised residual (Equation 8)

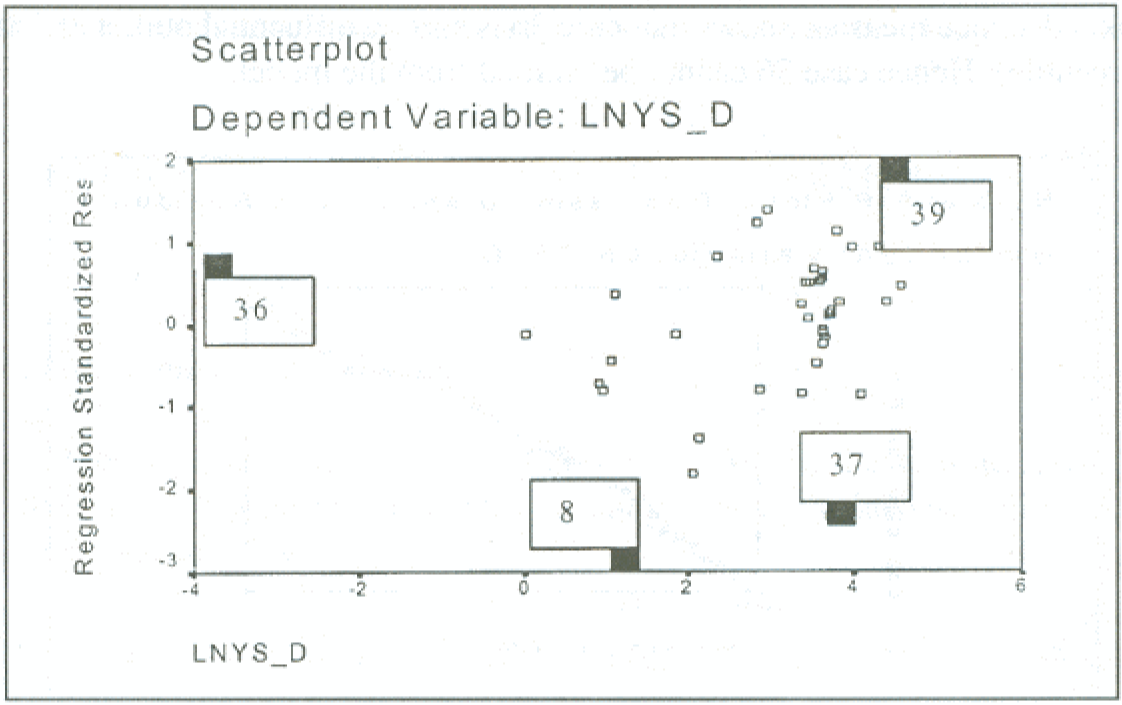

Figure 2. Standardised residual plot against $\hat{y}$ (Equation 8)

Among all the outliers given above, the Cook's distance measure show that case 36 to be an influential outlier ($D_i = 1.28186 = 70^{th}$ percentile). None of the fitted values are influential.

Thus an improved regression was obtained by omitting cases 8, 37 and 39. Thus the model obtained is given by:

$$\ln\left(\frac{y_s}{d}\right) = 0.030075 + 0.971976\ln\left(\frac{a}{d}\right) + 0.602400\ln\left(\frac{y_0}{a}\right) + 0.284267\ln\left(\frac{Q}{ay_0\sqrt{gy_0}}\right) \quad (9)$$

with the coefficient of determination, $r^2 = 0.971$. The analysis of variance test for model adequacy shows that the model is adequate $(F = 367.86316, p = 0.0)$.

Figure 3 and 4 show the normal probability plot for the standardised residuals and the residual plot against $\hat{y}$ respectively for the new regression equation. Figure 3 shows that the residual is approximately normal. Figure 4 below shows that the residuals have approximately zero mean and constant variance.

The Cook's distance measure shows that case 36 is still an influential outlier (1.78865 = 84.5th percentile). Hence case 36 cannot be omitted from the model.

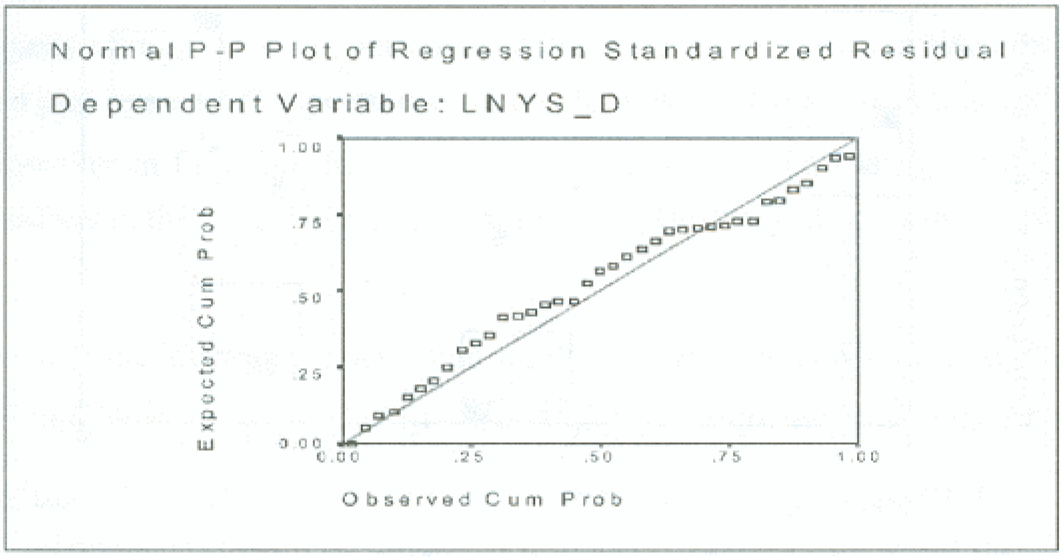

Figure 3. Normal probability plot for standardised residual (Equation 9)

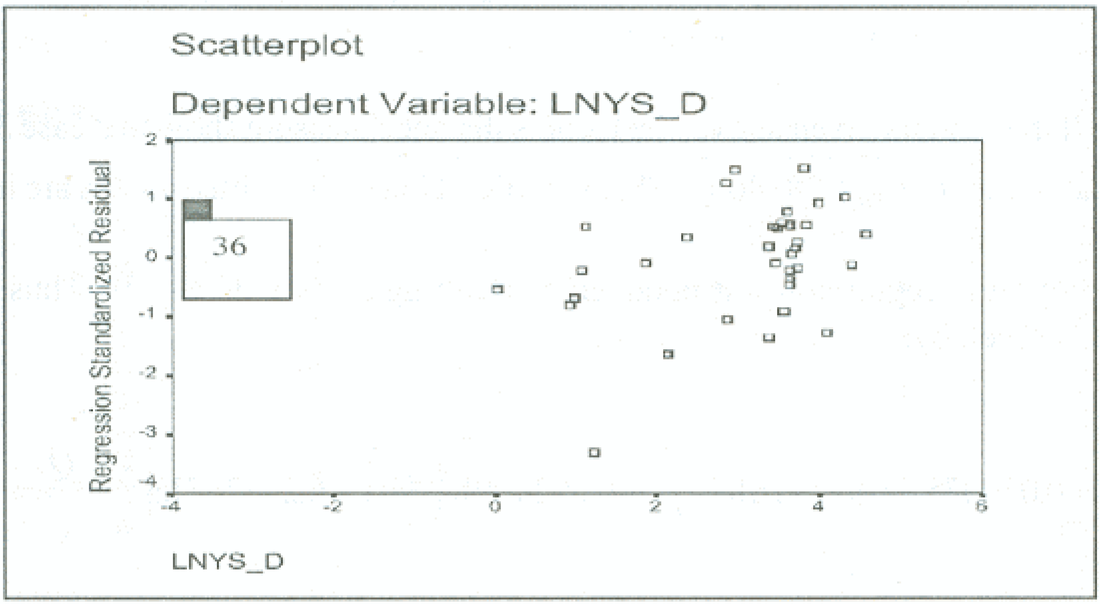

Figure 4. Standardised residual plot against (Equation 9)

CONCLUSION

This paper discusses the linear regression method to predict bridge pier scour for field data using the least squares method. The dependent variables used in predicting the pier scour are flow depth, pier width and maximum flood discharge.

By considering all available field data, the linear regression model (Equation 8) is given as follows:

$$\ln\left(\frac{y_s}{d}\right) = 0.166535 + 0.969302\ln\left(\frac{a}{d}\right) + 0.524133\ln\left(\frac{y_0}{a}\right) + 0.240191\ln\left(\frac{Q}{ay_0\sqrt{gy_0}}\right) \quad (8)$$

with the coefficient of determination,

However, there exist outliers in cases 8, 36,37 and 39. Among these outliers, only case 36 is influential. Thus a new regression model (Equation 9) was obtained by omitting cases 8, 37 and 39 and is given below:

$$\ln\left(\frac{y_s}{d}\right) = 0.030075 + 0.971976\ln\left(\frac{a}{d}\right) + 0.602400\ln\left(\frac{y_0}{a}\right) + 0.284267\ln\left(\frac{Q}{ay_0\sqrt{gy_0}}\right) \quad (9)$$

with the coefficient of determination, This model has improved the predicting power of the model since the value for has increased from 0.955 to 0.971. Thus if cases 8, 37 and 39 can be considered as outliers, this new regression model (Equation 9) can be used to predict pier scour better than the previous model (Equation 8).

REFERENCES

Ab. Ghani, A. & Nallurim, C. 1996. Development of Pier Scour Equations Using Field Data. *10th Congress Of Asia-Pacific Division Of International Association For Hydraulic Research, 26-29 August, Langkawi, Malaysia,* **1:** 295 – 302.

Kafi, M. & Alam, J. 1995. Modification Of Local Scour Equations. *Journal Of The Institution Of Engineers (India).* **76(5)**: 25-28.

Kothyari, U. C., Garde, R. J. & Ranga Raju, K. G. 1992. Live Bed Scour Around Cylindrical Bridge Piers. *Journal Of Hydraulic Research.* **30(5)**: 701 - 715.

Mendenhall, N. & Sincich, J. 1994. *Statistics For Engineering And the Sciences.* New York: Prentice Hall.

Norusis, M. J. 1997. *SPSS 7.5 Guide To Data Analysis.* North Carolina: Prentice Hall.