Desalination and Water Treatment
www.deswater.com
doi: 10.1080/19443994.2015.1112306

57 (2016) 21130–21142
September

Taylor & Francis
Taylor & Francis Group

Sensitivity analysis of model parameters on biochemical oxygen demand in integrated solar and hydraulic jump enhanced waste stabilization pond

N.M. Ogarekpe[a,*], J.C. Agunwamba[b], I.L. Nwaogazie[c]

[a]*Department of Civil Engineering, Cross River University of Technology, Calabar, Nigeria, email: nkpaogarekpe@yahoo.com*
[b]*Department of Civil Engineering, University of Nigeria, Nsukka, Nigeria, email: jonah.agunwamba@unn.edu.ng*
[c]*Department of Civil and Environmental Engineering, University of Port Harcourt, Port Harcourt, Nigeria, email: ifynwaogazie@yahoo.com*

Received 12 January 2015; Accepted 18 October 2015

ABSTRACT

The sensitivity analysis of selected parameters on biochemical oxygen demand (BOD) in the integrated solar and hydraulic jump enhanced waste stabilization pond (ISHJEWSP) was carried out by regression analysis approach. Based on empirical data, the linear relationship between the response variable, BOD, and the predictor variables of pH, temperature, algae concentration, dissolved oxygen (DO), inlet velocity, distance from inlet to the point of initiation of hydraulic jump, angle representing change in pond bed slope, and intensity of solar radiation was tested using various hypotheses at $\alpha = 5\%$ level of significance. The test of hypotheses revealed that not all the regression coefficients can be taken as zero and that the observed t-values of predictor variables X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), X_7 (x_{θ_s}), and the constant are statistically significant at $\alpha = 5\%$. The test on the equality of regression coefficients of the reduced model revealed that it is significant to infer that X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s}) do not have the same incremental effect in determining the concentration of BOD at $\alpha = 5\%$. Also, the test that a subset of regression coefficients equals to zero revealed that at $\alpha = 5\%$, it is significant to infer that the deletion of X_2 (x_{temp}), X_5 (x_{v}), and X_8 (x_{I}) does not adversely affect the explanatory strength of the model. The model developed gave a good multiple linear regression coefficient of correlation of 0.937 with a standard error of 5.17 at a significance level of 0.05. The good multiple linear regression coefficients of correlation obtained from the regression of algae concentration on temperature and solar radiation; pH on temperature and solar radiation revealed that temperature and solar radiation play a vital role in BOD degradation in the ISHJEWSP.

Keywords: Sensitivity analysis; BOD; Regression; Hypothesis; ISHJEWSP

1. Introduction

Engineering and scientific phenomena are often studied with the aid of mathematical models designed to simulate complex physical processes [1]. Sensitivity analysis has been used in scientific research to explore the validity of models [2]. The determination of the most sensitive parameters of a model is important especially in the time and cost savings associated with their practical utilization. The development of models as perhaps complementary or otherwise alternative methods to the laborious and time consuming

*Corresponding author.

experimental procedures cannot be over emphasized. This opinion perhaps held in the context of future needs where prevailing site and operational conditions are similar.

In the past, the sensitivity index approach has been used by Hoffman and Gardner [3] to account for all possible values when determining parameter sensitivity. Many methods of carrying out sensitivity analysis have been reviewed [4–11]. Hamby [1] stated that regression analysis provides the most comprehensive sensitivity measure and is commonly utilized to build response surfaces that approximate complex models.

Regression analysis is one of the most widely used statistical tools because it provides simple methods for establishing a functional relationship among variables [12]. Typically, regression analysis is used to investigate the relationships between a dependent variable (either categorical or continuous) and a set of independent variables based on a sample from a particular population [13]. Across behavioral science disciplines, multiple linear regression (MR) is a standard statistical technique in a researcher's toolbox [14].

In the past, standard theories of regression analysis have been discussed [15–19]. A general linear regression model can be formulated as shown in Eq. (1):

$$y_i = \beta_0 + \varepsilon_i + \sum_{i=1}^{n} \beta_p x_{ip} \quad i = 1, 2, \ldots, n \tag{1}$$

The regression model can be rewritten as shown in Eq. (2):

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{2}$$

Where y_i is the ith value of the response variable y; x_{i1}, x_{i2}, …, x_{ip} represent the ith independent variable value from total set of n observations; p represents the number of estimated regression coefficients; β_i represents the ith regression coefficient corresponding to x_{ip}; β_0 is the constant coefficient that represents the intercept; ε_i represents the error in the approximation of y_i; n is the number of observations.

To estimate the regression coefficients $\beta_0, \beta_1, \beta_2, \ldots \beta_p$, the least square method is used in order to minimize the sum of squares of the errors. From Eq. (1), the errors can be written as:

$$\varepsilon_i = y_i - \beta_0 - \sum_{i=1}^{n} \beta_p x_{ip} \quad i = 1, 2, \ldots, n \tag{3}$$

The sum of squares of these errors is given by Eq. (4):

$$S = \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_{i1} - \ldots - \beta_p x_{ip} \right)^2 \tag{4}$$

By a direct application of calculus, it can be shown that the least squares estimates $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$, which minimize $S\ (\beta_0, \beta_1, \ldots, \beta_p)$, are given by the solution of system of linear equations known as the normal equations [12]. For this study, the regression coefficients were obtained by methods of multiple linear regression using Microsoft Excel.

Using the estimated regression coefficients, the fitted least squares regression equation is given as follows:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \cdots + \hat{\beta}_p x_{ip}, \quad i = 1, 2, \ldots, n \tag{5}$$

The ordinary least squares residuals are given by Eq. (6):

$$e_i = y_i - \hat{y}_i, \quad i = 1, 2, 3, \ldots, n \tag{6}$$

The Error Sum of Square (SSE) is given by:

$$\mathrm{SSE} = \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2 \tag{7}$$

An unbiased estimate of σ^2 is given by:

$$\sigma^2 = \frac{\mathrm{SSE}}{\mathrm{d.f.}} \tag{8}$$

where:

$$\mathrm{d.f.} = n - p - 1 \tag{9}$$

The correlation coefficient between Y and $\hat{y}$, is given by:

$$\mathrm{Cor}\left(Y, \hat{Y}\right) = \frac{\sum (y_i - \bar{y}) \sum \left(\hat{y}_i - \bar{\hat{y}}\right)}{\sqrt{\sum (y_i - \bar{y})^2 \sum \left(\hat{y}_i - \bar{\hat{y}}\right)^2}} \tag{10}$$

where $\bar{y}$ is the mean of the response variable Y and $\bar{\hat{y}}$ is the mean of the fitted values. The coefficient of determination $R^2 = [\mathrm{Cor}\,(Y, \hat{Y})]^2$ is also given as follows:

$$R^2 = \frac{\text{SSR}}{\text{SST}} = 1 - \frac{\text{SSE}}{\text{SST}}$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (11)$$

The adjusted R^2, R_a^2 is defined as shown in Eq. (12):

$$R_a^2 = 1 - \frac{n-1}{n-p-1}\left(1 - R^2\right) \quad (12)$$

Using the properties of the least squares estimators, the statistical inference regarding the regression coefficients can be made [2]. Chatterjee and Hadi [2] presented that the statistic for testing H_0: $\beta_0 = \beta_j^0$ vs. H_1: $\beta_j \neq \beta_j^0$ where β_j^0 is a constant chosen by the investigator is:

$$t_j = \frac{\hat{\beta}_j - \beta_j^0}{s.e.\left(\hat{\beta}_j\right)} \quad (13)$$

which has a Student's t-distribution with $(n - p - 1)$ degrees of freedom.

If $\beta_j^0 = 0$, then we have Eq. (14):

$$t_j = \frac{\hat{\beta}_j}{s.e.\left(\hat{\beta}_j\right)} \quad (14)$$

Also

$$F = \frac{\left(R_p^2 - R_q^2\right)/(p-q)}{\left(1 - R_p^2\right)/(n-p-1)}, \quad \text{d.f.}\, p-q,\; n-p-1 \quad (15)$$

where F is the observed value of the F-test; d.f. is the degree of freedom; $p - q$, $n - p - 1$ are the degrees of freedom; SST denotes the Total Sum of Squared deviation in Y from its mean $\bar{y}$; SSR denotes the sum of squares due to regression; SSE denotes the Sum of Squares of the residuals (errors); SSE(FM) denotes the residual sum of squares from the full model; SSE(RM) denotes the residual sum of squares from the reduced model; $\bar{y}$ is the mean of the response variable Y; $\bar{\hat{y}}$ is the mean of the fitted values; β_j^0 is a constant chosen by the investigator; $s.e.(\hat{\beta}_j)$ is the standard error of $\hat{\beta}_j$; R_p^2 denotes the sample multiple correlation coefficient of determination that is obtained when the full model with all the p variables in it is fitted to the data; R_a^2 denotes the adjusted multiple correlation coefficient of determination of the fitted model; R_q^2 denotes the sample multiple correlation coefficient of determination when the model is fitted with q specific variables; σ^2 denotes the variance.

Empirical models have been developed to describe the kinetics of organic degradation in waste stabilization ponds [20–22]. Saqqar and Pescod [23] and Mayo [24] have all developed models for predicting fecal coliform removal in WSP based on environmental factors. In addition, researches have been conducted on the prediction of biochemical oxygen demand (BOD) removal in solar-enhanced WSP [25], sludge filtration equation [26]. Predictive equations for the dispersion number have been proposed [27–32], but some of these have then been criticized when evaluated by others. However, there is paucity of information in literature with respect to research in ISHJEWSP. This research is aimed at the determination of the most sensitive parameters in the BOD prediction model for the ISHJEWSP using regression analysis as sensitivity measure.

The integrated solar and hydraulic jump enhanced waste stabilization pond (ISHJEWSP) is introduced as a new technology that incorporates solar reflector and the introduction of hydraulic jump through change in pond bed slope of the conventional waste stabilization pond. The essence is for the purpose of increasing the treatment efficiency of the conventional WSP and consequently, the reduction in land area requirement [33].

2. Materials and methods

2.1. Description of area of study

Nsukka is a town and Local Government Area in southeast Nigeria in Enugu State. Nsukka urban is the home to the prestigious University of Nigeria. Located at the northeastern end of the University campus about 800 m from the junior staff quarters, the treatment plant at Nsukka consists of a screen (6 mm bar racks set at 12 mm centers) followed by two Imhoff tanks, each measuring about 6.667 m × 4.667 m × 10 m, and two facultative waste stabilization ponds. Sludge is discarded from the Imhoff tank once every 28 d onto the drying beds, so that the beds are loaded at 40-d interval. The beds have a total area of 417 m^2. Although its efficiency has deteriorated, its effluent is used for uncontrolled vegetable irrigation by some village dwellers. The poor effluent quality is also partly attributable to overloading because of population growth.

2.2. Description of experimental setup

Experimental research and design were adopted. The experimental setup consisted of one sewage storage tank (1.2 m × 1.2 m × 0.6 m) and an overhead storage tank (1.5 m × 1.5 m × 1.2 m) as shown in Tables 1–2 below. Three sets of experimental ponds with varying locations of change in pond bed slope were constructed using metallic tanks with each set consisting of eight experimental ponds (*A*, *B*, *C*, *D*, *E*, *F*, *G*, *H*) with varying widths. Six out of the eight ponds were constructed with tilt frames of size 1.0 m × 0.3 m, fixed at varying angles in accordance with the relative position of the sun per week. The tilt frames were made of a flat wooden board wrapped with aluminum foil paper to serve as solar reflectors. The foil paper was to act as solar reflector, with each of the six ponds having one reflector each at the outlet position (west facing). One out of the eight ponds was constructed without a change in slope and solar reflector to serve as control experiment, while the other though without change in slope however was fitted with solar reflector in order to investigate the effect of solar radiation on the conventional WSP. For each set studied, ponds *C*, *D*, *E*, *F*, *G*, and *H* were constructed with varying locations of point of initiation of hydraulic jump. Half-inches diameter inlet pipes were fitted centrally to the experimental ponds. The outlet pipes were centrally fitted to the experimental ponds. To control the inflow and outflow, valves were fitted at the inlet and outlet pipes of the experimental ponds. The two storage tanks were usually filled to supply the eight ponds with sewage effluent from the Imhoff tank of the University of Nigeria, Nsukka sewage treatment plant through a hose with the aid of an electromechanical water pump. The influent samples for the laboratory analysis were obtained from the storage tank immediately after being filled. Also, the experimental ponds were immediately filled and samples collected at the outlets after two days.

2.3. Data collection and analysis

Wastewater samples collected from the inlet and outlet for varying inlet velocities and varying locations of point of initiation of hydraulic jump were examined for physicochemical and biological characteristics for a period of nine months. The parameters examined were temperature, pH, detention time, dissolved oxygen (DO), total coliform count (TCC), total suspended solids (TSS), *E. coli*, and algae concentrations, and biochemical oxygen demand (BOD). All the analyses were carried out using appropriate water testing meters and in accordance with the standard methods [34]. Microsoft Excel was used to perform all the statistical analyses.

2.4. Model development

In developing the model for the prediction of BOD in the ISHJEWSP, it is assumed that a linear relationship exists between the response variable (y_{BOD_5}) and the predictor variables (x_{pH}, x_{temp}, x_{Algae}, x_{DO}, x_v, x_{HJL}, x_θ, x_I). y_{BOD_5} is directly proportional to x_{pH}, x_{temp}, x_{Algae}, x_{DO}, x_v, x_{HJL}, x_θ, x_I. The addition of these parameters duly factored with their regression coefficients to replace proportionality yields Eq. (16):

$$y_{\mathrm{BOD}_5} = \beta_0 + \beta_1 x_{\mathrm{pH}} + \beta_2 x_{\mathrm{temp}} + \beta_3 x_{\mathrm{Algae}} + \beta_4 x_{\mathrm{DO}} + \beta_5 x_v + \beta_6 x_{\mathrm{HJL}} + \beta_7 x_{\theta_s} + \beta_8 x_I \quad (16)$$

This is identical to the multiple linear regression model given by Eq. (1).

Where y_{BOD_5} is the dependent variable which is the concentration of BOD_5 of the ISHJEWSP (mg/l); x_{pH} is pH; x_{temp} is temperature of the ISHJEWSP (°C); x_{Algae} is algae concentration of the ISHJEWSP (μg chlorophyll a/l); x_{DO} is DO of the ISHJEWSP (mg/l); x_v is the inlet velocity (m/s); x_{HJL} is the distance from inlet to the point of initiation of hydraulic jump (m); x_{θ_s} is the angle representing change in pond bed slope (°); x_I represents the intensity of solar radiation (kW/m^2); β_1, β_2, β_3, β_4, β_5, β_6, β_7, and β_8 are corresponding regression coefficients.

Table 1
Detailed experimental characteristics of the various ponds due to variations in location of point of initiation of hydraulic jump [33]

Experimental setups	No. of solar ponds	Characteristics (location of point of initiation of hydraulic jump from the inlet)	Purpose
Set 1	8	0.5 m	Effect of location of point of initiation of hydraulic jump
Set 2	8	0.4 m	Effect of location of point of initiation of hydraulic jump
Set 3	8	0.3 m	Effect of location of point of initiation of hydraulic jump

Table 2
Detailed experimental characteristics of the various ponds due to varying inlet velocities [33]

Experimental setups	No. of experimental ponds	Characteristics (velocity) (m/s)	Purpose
Set 1	8	0.39	Effect of inlet velocity
	8	0.42	Effect of inlet velocity
	8	0.46	Effect of inlet velocity
Set 2	8	0.39	Effect of inlet velocity
	8	0.42	Effect of inlet velocity
	8	0.46	Effect of inlet velocity
Set 3	8	0.39	Effect of inlet velocity
	8	0.42	Effect of inlet velocity
	8	0.46	Effect of inlet velocity

3. Results and discussion

3.1. Modeling the BOD concentration in the ISHJEWSP

The model was developed through the application of multiple linear regression analysis on the measured variables for Pond *D* using Microsoft Excel. The process involved estimating the model regression coefficients β_0, β_1, β_2, β_3, β_4, β_5, β_6, β_7, and β_8. The application of regression analysis using Microsoft Excel on the data revealed that a linear correlation existed between the parameters. The model developed gave a good multiple linear regression coefficient of correlation of 0.938 with a standard error of 5.224 at a significance level of 0.05. Eq. (17) below, is obtained as the full empirical multiple linear regression model for the prediction of the BOD in the ISHJEWSP with reference to the University of Nigeria, Nsukka treatment plant. The summary of output of the multiple linear regression model is shown in Table 3.

$$y_{\mathrm{BOD}_5} = -2.82\,x_{\mathrm{pH}} - 0.49\,x_{\mathrm{temp}} - 0.09\,x_{\mathrm{Algae}} + 16.68\,x_{\mathrm{DO}} + 12.69\,x_v + 105.52\,x_{\mathrm{HJL}} + 1.17\,x_{\theta_s} + 1.40\,x_I - 123.45 \quad (17)$$

where y_{BOD_5} represent the BOD_5 in the ISHJEWSP (mg/l); x_{temp} is temperature of the ISHJEWSP (°C); x_{Algae} is algae concentration in the ISHJEWSP (μg chlorophyll a/l); x_{DO} is DO in the ISHJEWSP (mg/l); x_v is the inlet velocity (m/s); x_{HJL} is the distance from inlet to the point of initiation of hydraulic jump (m); x_{θ_s} is the angle representing change in pond bed slope (°); and x_I represents the intensity of solar radiation ($\mathrm{kW/m^2}$).

The model in Eq. (17) was developed on the linearity assumption. There is therefore need for the determination of the key predictor parameters.

3.2. Sensitivity analysis of model parameters on BOD concentration prediction in ISHJEWSP

Based on empirical data, the relationship between the response variable, BOD, and the predictor variables of pH, temperature, algae concentration, DO, inlet velocity, distance from inlet to the point of initiation of hydraulic jump, angle representing change in pond bed slope, intensity of solar radiation obtained in Eq. (17) were tested using various hypotheses at $\alpha = 5\%$ level of significance.

3.3. Testing all regression coefficients equal to zero

The hypothesis that all predictor variables under consideration do not explain the response variable and that all their regression coefficients are zero was tested. In this case, the reduced model (RM) and full model (FM) become:

$$\mathrm{RM}: H_0: Y = \beta_0 + \varepsilon \quad (18)$$

$$\mathrm{FM}: H_1: Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_8 x_8 + \varepsilon \quad (19)$$

From the ANOVA table in Table 3, the sum of squares due to error in the full model is SSE (FM) = SSE = 2701.25. Under the null hypothesis, where all the β's are zero, the number of parameters estimated for the reduced model is therefore one (β_0). Consequently, the sum of squares of the residuals in the reduced model is SSE (RM) = SST = 22,450. The observed *F*-value for the full model is 90.47 as shown in Table 3. The *F*-critical value at 8 and 99 degrees of freedom and 5% level of significance is 2.03. Since the observed *F*-value is larger than the critical *F*-value, the

Table 3
Summary of output of regression analysis of Y (y_{BOD_5}) on X_1 (x_{pH}), X_2 (x_{temp}), X_3 (x_{Algae}), X_4 (x_{DO}), X_5 (x_v), X_6 (x_{HJL}), X_7 (x_{θ_s}), and X_8 (x_I)

Regression statistics								
Multiple R	0.937910973							
R^2	0.879676993							
Adjusted R^2	0.869953922							
Standard error	5.223539863							
Observations	108							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	8	19748.7485	2468.593562	90.4731612	4.68672E-42			
Residual	99	2701.251502	27.2853687					
Total	107	22,450						
	Coefficients	Standard error	t Stat	p-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	−123.4519203	54.6117476	−2.26053781	0.02597736	−231.813476	−15.090365	−231.813476	−15.090365
X variable 1	−2.818049995	1.642956327	−1.715231226	0.08943124	−6.07803179	0.441931799	−6.07803179	0.4419318
X variable 2	−0.491935063	0.674420079	−0.72941936	0.46746744	−1.83013082	0.846260691	−1.83013082	0.84626069
X variable 3	−0.087315434	0.014841388	−5.883239059	5.4865E-08	−0.11676397	−0.0578669	−0.11676397	−0.0578669
X variable 4	16.67759918	2.432534152	6.856059622	6.1399E-10	11.85092368	21.50427468	11.85092368	21.5042747
X variable 5	12.68836935	18.47712292	0.686706984	0.49387275	−23.9742512	49.35098986	−23.9742512	49.3509899
X variable 6	105.5221408	44.64433701	2.363617603	0.02005049	16.93809054	194.1061911	16.93809054	194.106191
X variable 7	1.165576683	0.454581467	2.564065561	0.01184818	0.263588431	2.067564935	0.263588431	2.06756493
X variable 8	1.397930019	8.568354356	0.163150351	0.87073264	−15.6035439	18.39940398	−15.6035439	18.399404

null hypothesis is rejected; not all the βs can be taken as zero.

3.4. Test of hypothesis of regression coefficients

The t-values in Table 3 test the null hypothesis H_0 against an alternative hypothesis H_1:

$$H_0 : \beta_j = 0, \quad j = 0, 1, 2, ..., p \tag{20}$$

$$H_1 : \beta_j \neq 0, \quad j = 0, 1, 2, ..., p \tag{21}$$

From Table 3 it is seen that only the regression coefficients of X_4 (x_{DO}), X_5 (x_v), X_6 (x_{HJL}), X_7 (x_{θ_s}), and X_8 (x_I) are statistically significantly different from zero. However, applying the t statistics, the critical t-value at 99 degree of freedom and 5% level of significance is 1.66. From Table 3, the computed t-values of X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), X_7 (x_{θ_s}), and the constant are statistically significant as compared to the critical t-value. From Table 3, the p-values of X_3 (x_{pH}), X_4 (x_{DO}), X_6 (x_{HJL}), X_7 (x_{θ_s}), and the constant are statistically significant with p-values of 5.48E-08, 6.14E-10, 0.02, 0.01, and 0.03, respectively.

3.5. Testing a subset of regression coefficients equal to zero

In order to determine the most important predictors that describe the response, the variables which were statistically significant with regards to the t-test of the full model were tested. The aforementioned significant parameters were X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s}). The reduced model is therefore:

$$\text{RM: } Y = \beta_0 + \beta_1 x_1 + \beta_3 x_3 + \beta_4 x_4 + \beta_6 x_6 + \beta_7 x_7 + \varepsilon. \tag{22}$$

This corresponds to the null hypothesis:

$$H_0 : \beta_2 = \beta_5 = \beta_8 = 0 \tag{23}$$

The regression output from fitting this model is given in Table 4. Hence the F-test, as given in Eq. (15) is as follows:

$$F = \frac{[0.8797 - 0.8784]/3}{(1 - 0.8797)/99} = 0.36 \tag{24}$$

The F-critical value at 3 and 99 degrees of freedom and 5% level of significance is 2.7. The observed F-value is less than the critical F-value, and thus not statistically significant. Therefore, the null hypothesis is not rejected. The variables X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s}) together explain the variation in Y as adequately as the full set of eight variables. Hence, at $\alpha = 5\%$ it is significant to infer that the deletion of X_2 (x_{temp}), X_5 (x_v) and X_8 (x_I) does not adversely affect the explanatory strength of the model.

3.6. Testing the equality of regression coefficients

This test was carried out to determine whether the five regression coefficients in the reduced model, with summary of regression output in Table 4, are the same. The test is performed with the assumption that the regression coefficients of the full model for X_2 (x_{temp}), X_5 (x_v), and X_8 (x_I) are zero. The null hypothesis to be tested is:

$$H_0 : \beta_1 = \beta_3 = \beta_4 = \beta_6 = \beta_7| \quad (\beta_2 = \beta_5 = \beta_8 = 0) \tag{25}$$

If $\beta_1 = \beta_3 = \beta_4 = \beta_6 = \beta_7 = \beta_1'$, then, Eq. (22) which is the reduced model becomes:

$$Y = \beta_0' + \beta_1' (X_1 + X_3 + X_4 + X_6 + X_7) + \varepsilon \tag{26}$$

Let $X_1 + X_3 + X_4 + X_6 + X_7 = Z$, then Eq. (26) becomes:

$$Y = \beta_0' + \beta_1' Z + \varepsilon \tag{27}$$

The least squares estimates of β_0', β_1', and the simple linear regression correlation coefficient, as shown in Table 5 below, yields the fitted Eq. (28):

$$\hat{Y} = 89.54 - 0.11\,Z \tag{28}$$

Hence the F-test, as given in Eq. (15) is:

$$F = \frac{[0.8784 - 0.4624]/(5 - 1)}{(1 - 0.8797)/(108 - 5 - 1)} = 87.24 \tag{29}$$

The F-critical value at 4 and 102 degrees of freedom and 5% level of significance is 2.46. The observed F-value is greater than the critical F-value, and thus statistically significant. Therefore, the null hypothesis is rejected. Hence, at $\alpha = 5\%$, X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s}) do not have the same incremental effect in determining the concentration of BOD.

Table 4
Summary of output of regression analysis of Y (y_{BOD_5}) on X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s})

Regression statistics								
Multiple R	0.937242637							
R^2	0.878423761							
Adjusted R^2	0.872464142							
Standard error	5.172880409							
Observations	108							
ANOVA								
	df	SS	MS	*F*	Significance *F*			
Regression	5	19720.613	3944.123	147.3959	4.98946E-45			
Residual	102	2729.3866	26.75869					
Total	107	22,450						
	Coefficients	Standard error	*t* Stat	*p*-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	−133.621924	50.610878	−2.64018	0.009588	−234.008361	−33.23549	−234.00836	−33.2355
X variable 1	−3.24723118	1.4337553	−2.26484	0.025637	−6.09107796	−0.403384	−6.091078	−0.40338
X variable 3	−0.09139166	0.0113277	−8.06796	1.45E-12	−0.11386016	−0.068923	−0.1138602	−0.06892
X variable 4	16.84154933	2.2121855	7.613082	1.39E-11	12.45368987	21.22941	12.4536899	21.22941
X variable 6	111.1154547	42.448612	2.617646	0.010201	26.91883407	195.3121	26.9188341	195.3121
X variable 7	1.214541313	0.4315223	2.81455	0.005863	0.358618847	2.070464	0.35861885	2.070464

Table 5
Summary of output of regression analysis of Y (y_{BOD_5}) on Z ($X_1 + X_3 + X_4 + X_6 + X_7$)

Regression statistics								
Multiple R	0.679985222							
R^2	0.462379903							
Adjusted R^2	0.457308015							
Standard error	10.6707019							
Observations	108							
ANOVA								
	df	SS	MS	*F*	Significance *F*			
Regression	1	10380.42882	10380.429	91.16524834	5.8399E-16			
Residual	106	12069.57118	113.86388					
Total	107	22,450						
	Coefficients	Standard error	*t* Stat	*p*-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	89.53824536	4.436001667	20.184448	4.20681E-38	80.7434406	98.33305012	80.74344061	98.33305012
Z variable	−0.108669721	0.011381353	−9.548049	5.83991E-16	−0.13123436	−0.08610508	−0.13123436	−0.08610508

3.7. Multiple linear regression model of BOD prediction in the ISHJEWSP

The test of hypotheses revealed that not all the regression coefficients can be taken as zero and that the computed t-values of predictor variables X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), X_7 (x_{θ_s}), and the constant are statistically significant at $\alpha = 5\%$. The test on the equality of regression coefficients of the reduced model revealed that it is significant to infer that X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s}) do not have the same incremental effect in determining the concentration of BOD at $\alpha = 5\%$. Also, the test that a subset of regression coefficients equals to zero revealed that at $\alpha = 5\%$, it is significant to infer that the deletion of X_2 (x_{temp}), X_5 (x_v), and X_8 (x_I) does not adversely affect the explanatory strength of the model.

The application of regression analysis using Microsoft Excel on the experimental data obtained for y_{BOD_5}, X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s}) revealed that a linear correlation existed between the parameters. The model developed gave a good multiple linear regression coefficient of correlation of 0.937 with a standard error of 5.173 at a significance level of 0.05. Therefore, Eq. (30) is presented as the degradation model of BOD in the ISHJEWSP considering the multiple linear regression approach with reference to the University of Nigeria, Nsukka treatment plant. The summary of output of the multiple linear regression model is shown in Table 4.

$$y_{\text{BOD}_5} = -3.25\,x_{\text{pH}} - 0.09\,x_{\text{Algae}} + 16.84\,x_{\text{DO}} + 111.12\,x_{\text{HJL}} + 1.21\,x_{\theta_s} - 133.62 \quad (30)$$

3.8. Effects of temperature, solar radiation, and inlet velocity on the degradation of BOD in the ISHJEWSP

Temperature has been found to be one of the most important variables affecting biological processes [35,36]. Algal activity is also retarded at low temperatures. Even under conditions of high solar radiation, intensity of algal growth is affected by low temperatures [37]. The main mechanism of oxygenation in pond systems is the oxygen provided by the algal population [38,39].

The results obtained from the ISHJEWSP (Pond D) revealed that the changes in temperature and intensity of solar ration resulted in the variation of such parameters as DO, pH, algae concentration, and biochemical oxygen demand. Subjecting algae concentration (response variable) and temperature (predictor variable) to simple linear regression, the results revealed that there exists a good linear relationship between the variables with a coefficient of correlation of 0.828 at $\alpha = 5\%$. Similarly, the simple linear regression of algae concentration (response variable) and intensity of solar radiation (predictor variable) revealed that there exists a good linear relationship between the variables with a coefficient of correlation of 0.796 at $\alpha = 5\%$. Also, the multiple linear regression of algae concentration (response variable) on temperature and intensity of solar radiation as predictor variables revealed that there exists a good linear relationship between the variables with a multiple linear regression coefficient of correlation of 0.839 at $\alpha = 5\%$. The summary of output of the multiple linear regression model is shown in Table 6 below. Also, the multiple linear regression of pH (response variable)

Table 6
Summary of output of regression analysis of Y (y_{Algae}) on X_1 (x_{temp}), and X_2 (x_I)

Regression statistics	
Multiple R	0.838963
R^2	0.703859
Adjusted R^2	0.698219
Standard error	47.56048
Observations	108

ANOVA

	df	SS	MS	*F*	Significance *F*
Regression	2	564507.3	282253.7	124.7806	1.79E-28
Residual	105	237509.9	2261.999		
Total	107	802017.2			

	Coefficients	Standard error	*t* Stat	*p*-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	−536.269	131.7873	−4.0692	9.15E-05	−797.579	−274.959	−797.579	−274.959
X variable 1	23.81498	4.752464	5.01108	2.2E-06	14.39172	33.23824	14.39172	33.23824
X variable 2	157.6166	61.23257	2.574064	0.011446	36.20368	279.0294	36.20368	279.0294

Table 7
Summary of output of regression analysis of Y (y_{pH}) on X_1 (x_{temp}), and X_2 (x_I)

Regression statistics								
Multiple R	0.88695965							
R^2	0.78669741							
Adjusted R^2	0.78263451							
Standard error	0.41100958							
Observations	108							
ANOVA								
	df	SS	MS	*F*	Significance *F*			
Regression	2	65.41913	32.70957	193.6292	5.92E-36			
Residual	105	17.73753	0.168929					
Total	107	83.15667						
	Coefficients	Standard error	*t* Stat	*p*-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.95646747	1.138883	0.839829	0.402912	−1.30173	3.214662	−1.30173	3.214662
X variable 1	0.22676593	0.04107	5.521451	2.45E-07	0.145332	0.3082	0.145332	0.3082
X variable 2	2.08980596	0.529161	3.949278	0.000142	1.040576	3.139035	1.040576	3.139035

on temperature and intensity of solar radiation as predictor variables revealed that there exists a good linear relationship between the variables with a multiple linear regression coefficient of correlation of 0.887 at $\alpha = 5\%$. The summary of output of the multiple linear regression model is shown in Table 7 below.

It is therefore significant to infer that though Eq. (30) does not include temperature and intensity of solar radiation as key predictors, it is worthy of note that their direct effect on the pH and algal bloom with its subsequent ripple effect, imply that they play a vital role in BOD degradation in the ISHJEWSP. This assertion is buttressed by the comparative analysis of the conventional waste stabilization pond and the ISHJEWSP which revealed that the latter performed better in fecal coliform bacteria removal [40]. The verification of the conventional model gave good average coefficients of correlation of $R = 0.800 \pm 0.173$ between the measured and calculated N_e/N_o and $R = 0.924 \pm 0.034$ for the ISHJEWSP, respectively [40]. Air bubble entrainment in a hydraulic jump starts for $Fr_1 > 1$ to 1.3 [31–33]. The high observed t-value ($p < 0.05$) for the DO in Table 4 is indicative of air entrainment associated with the occurrence of hydraulic jump in the ISHJEWSP. The inflow Froude numbers of 1.1, 1.2, and 1.3 corresponding to inlet velocities 0.39, 0.42, and 0.46 have contributed its quota to the DO in the ISHJEWSP in addition to the contributions by algae and other pond oxygenators.

4. Conclusions

Sensitivity analysis was carried out on the relationship between the response variable (BOD) and the predictor variables of pH, temperature, algae concentration, DO, inlet velocity, distance from inlet to the point of initiation of hydraulic jump, angle representing change in pond bed slope, and intensity of solar radiation using regression approach.

Testing all regression coefficients equal to zero on the linear model revealed that the null hypothesis is rejected; implying that not all the βs can be taken as zero. The test of hypothesis of regression coefficients revealed that the observed t-values of predictor variables X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), X_7 (x_{θ_s}), and the constant are statistically significant at $\alpha = 5\%$ level of significance.

Furthermore, the test that a subset of regression coefficients equals to zero revealed that at $\alpha = 5\%$, it is significant to infer that the deletion of X_2 (x_{temp}), X_5 (x_v), and X_8 (x_I) does not adversely affect the explanatory strength of the model. Also, the test on the equality of regression coefficients of the reduced model revealed that at significance level $\alpha = 5\%$, it is significant to infer that X_1 (x_{pH}), X_3 (x_{Algae}), X_4 (x_{DO}), X_6 (x_{HJL}), and X_7 (x_{θ_s}) do not have the same incremental effect in determining concentration of BOD.

Eq. (30) is presented as a multiple linear regression model for the prediction of the biochemical oxygen demand (BOD) concentration in the ISHJEWSP with reference to the University of Nigeria, Nsukka treatment plant. Though Eq. (30) does not include temperature and intensity of solar radiation as key predictors, it is worthy of note that their direct effect on the pH and algal bloom with its ripple effect, imply that they play a vital role in BOD degradation in the ISHJEWSP.

Similarly, the air entrainment associated with hydraulic jump, suggests its positive impact on the DO concentration of the ISHJEWSP.

Acknowledgments

Part of the results presented in this paper rely on TRODAN data collected and managed by the Centre for Atmospheric Research, National Space Research and Development Agency, Federal Ministry of Science and Technology, Anyigba, Nigeria. We thank the Centre for Atmospheric Research and their partners for promoting high standards of atmospheric observatory practice as well as the Federal Government of Nigeria for continuous funding of the Nigerian Space program (www.carnasrda.com).

References

[1] D.M. Hamby, A review of techniques for parameter sensitivity analysis of environmental models, Environ. Monit. Assess. 32 (1994) 135–154.
[2] J.R. Williams, F.R. Burton, R.F. Paige, F.A.C. Polack, Sensitivity analysis in model-driven engineering, in: R.B. France, J. Kazmeier, R. Breu, C. Atkinson (Eds.), MODELS, Springer-Verlag, Berlin Heidelberg, 2012, pp. 743–758.
[3] F.O. Hoffman, R.H. Gardner, Evaluation of uncertainties in environmental radiological assessment models, in: J.E. Till, H.R. Meyer (Eds.), Radiological Assessments: A Textbook on Environmental Dose Assessment, US Nuclear Regulatory Commission, Washington, DC, Report No. NUREG/CR-3332, 1983.
[4] R.H. Gardner, D.D. Huff, R.V. O'Neill, J.B. Mankin, J. Carney, J. Jones, Application of error analysis in a marsh hydrology model, Water Resour. Res. 16 (1980) 659–664.
[5] M. Demiralp, H. Rabitz, Chemical kinetic functional sensitivity analysis: Elementary sensitivities, J. Chem. Phys. 74 (1981) 3362–3375.
[6] R.H. Gardner, R.V. O'Neill, J.B. Mankin, J.H. Carney, A comparison of sensitivity analysis and error analysis based on a stream ecosystem model, Ecol. Modell. 12 (1981) 173–190.
[7] R.L. Iman, J.C. Helton, J.E. Campbell, An approach to sensitivity analysis of computer models: Part II—Ranking of input variables, response surface validation, distribution effect and technique synopsis, J Qual. Technol. 13 (1981) 232–240.
[8] D.J. Downing, R.H. Gardner, F.O. Hoffman, An examination of response-surface methodologies for uncertainties analysis in assessment models, Technometrics 27 (1985) 151–163.
[9] J.C. Helton, R.L. Iman, J.B. Brown, Sensitivity analysis of the asymptotic behavior of a model for the environmental movement of radionuclides, Ecol. Modell. 28 (1985) 243–278.
[10] M.J. Crick, M.D. Hill, D. Charles, The role of sensitivity analysis in assessing uncertainty, in: Proceeding of an NEA workshop on uncertainty analysis for performance assessments of radioactive waste disposal systems, OECD, Paris, 1987, pp. 1–258.
[11] R.L. Iman, J.C. Helton, The repeatability of uncertainty and sensitivity analyses for complex probabilistic risk assessments, Risk Anal. 11 (1991) 591–606.
[12] S. Chatterjee, A.S. Hadi, Regression Analysis by Example, fourth ed., Wiley-Interscience, A John Wiley & Sons Inc, Publication, Hoboken, New Jersey, 2006.
[13] C.S. Ding, Using regression mixture analysis in educational research, Pract. Assess. Res. Eval. 11 (2006) 1–11.
[14] L.L. Nathans, F.L. Oswald, K. Nimon, Interpreting multiple linear regression: A guidebook of variable importance, Pract. Assess. Res. Eval. 17 (2012) 1–19.
[15] N.R. Draper, H. Smith, Applied Regression Analysis, third ed., A Wiley–Interscience, John Wiley & Sons, Inc., Publication, Hoboken, New Jersey, 1998.
[16] F.A. Graybill, H.K. Iyer, Regression Analysis: Concepts and Applications, Duxbury Press, An Imprint of Wadsworth Publishing Company Belmont, California, 1994.
[17] A. Sen, M.S. Srivastava, Regression Analysis: Theory, Methods, and Applications, New York Inc., Springer-Verlag, 1990.
[18] G.A.F. Seber, Linear Regression Analysis, John Wiley & Sons, New York, NY, 1971.
[19] C.R. Rao, Linear Statistical Inference and Its Applications, second ed., John Wiley & Sons, New York, NY, 1973.
[20] M.G. McGarry, M.B. Pescod, Stabilization pond design criteria for tropical Asia, 2nd International Symposium For Waste Treatment Lagoons, Missouri Basin Eng. Health Council, Kanses City, 114, 1970.
[21] T.B. Larson, A dimensional design equation for sewage lagoons, Dissertation University of New Mexico Albuquerque, 1974.
[22] E.F. Gloyna, Facultative waste stabilization pond design, in: E.F. Gloyna, J.F. Malina, Jr., E.M. David (Eds.), Ponds as a Wastewater Treatment Alternative, University of Texas, College of Engineering, Center for Research in Water Resources, Austin TX, 1976, p. 143.
[23] M.M. Saqqar, M.B. Pescod, Modeling coliform reduction in waste water stabilization pond, Water Sci. Technol. 26 (1992) 1667–1677.
[24] A.W. Mayo, Modelling coliform mortality in waste stabilization ponds, J. Environ. Eng. 121 (1995) 140–152.
[25] J.T. Utsev, N.M. Ogarekpe, T. Tivde, Empirical regression model for biochemical oxygen demand removal in solar enhanced waste stabilization ponds, Energy Sci. Technol. 5 (2013) 31–35.
[26] C.B. Afangideh, J.O. Ademilyi, C.C. Nnaji, LMT Sludge filtration equation using dimensional analysis, Pac. J. Sci. Technol. 12 (2011) 512–519.
[27] J.C. Agunwamba, N. Egbuniwe, J.O. Ademiluyi, Prediction of the dispersion number in waste stabilization ponds, Water Res. 26 (1992) 85–89.
[28] C. Polprasert, K.K. Bhattarai, Dispersion model for waste stabilization ponds, J. Environ. Eng. 111 (1985) 45–59.
[29] T. Nameche, J. Vasel, Hydrodynamic studies and modelization for aerated lagoons and waste stabilization ponds, Water Res. 32 (1998) 3039–3045.

[30] S. Arceivala, Wastewater Treatment and Disposal, Marcel Dekker Inc., New York NY, 1981.
[31] R.A. Ferrara, D.R.F. Harlmann, Dynamic nutrient cycle model for waste stabilization ponds, J. Environ. Eng. Div. ASCE 106 (1980) 37–54.
[32] J. Agunwamba, A new method for dispersion number determination in waste stabilization pond, Water Air Soil Pollut. 63 (1992) 361–369.
[33] N.M. Ogarekpe, Integrated solar and hydraulic jump enhanced waste stabilization pond, Ph.D. thesis, Civil Engineering, University of Nigeria, Nsukka, 2015.
[34] APHA, Standard Methods for the Examination of Water and Wastewater, twentieth ed., American Public Health Association, USA, 1998.
[35] S.L. Klemetson, Solar heating of wastewater stabilization ponds, Colorado Water Resources Research Institue, Colorado, CO, 1983.
[36] S. Kayombo, T.S.A. Mbwette, J.Y.H. Katima, S.E. Ladegaard, S.E. Jorgensen, Waste stabilization ponds and constructed wetlands design manual, UNEP, Osaka, Japan, 2005.
[37] G.V.R. Marvis, Dynamic behavior of oxidation ponds, Second International Symposium for Waste Treatment Lagoons, Kansas City, Mo, 1970.
[38] A. Shilton, J. Harrison, Guidelines for the hydraulic design of waste stabilization ponds, Palmerston North, Institute of Technology and Engineering, New Zealand, 2003.
[39] D. Mara, H. Pearson, Design manual for waste stabilization ponds in Mediterranean countries, Lagoon Technology International, Leeds, England, 1998.
[40] N.M. Ogarekpe, J.C. Agunwamba, A new model for the prediction of the performance of integrated solar and hydraulic jump enhanced waste stabilization pond, Desalin. Water Treat. (2015), doi: 10.1080/19443994.2015.1053535.