Petroleum and Coal

**Article** **Open Access**

# A Response Surface Model for Predicting Sand Production Rate for Oil Wells in the Niger Delta

*Anthony Ogbaegbe Chikwe, Onyebuchi Ivan Nwanwe*, Augustine Chinemerem Duru, Jude Emeka Odo*

*Department of Petroleum Engineering, Federal University of Technology, Owerri Nigeria*

Received July 30, 2023; Accepted November 13, 2023

**Abstract**

The petroleum industry is becoming increasingly concerned with sand production, which is common in the Niger Delta and presents technical, operational, and financial difficulties. Literature has devoted a lot of time to the creation of sanding prediction tools and efficient management techniques. However, the majority of the published theoretical models have been supported by data from other petroleum regions than the Niger Delta or evidence from laboratories. An easy-to-use machine learning model was created that takes into account the idea of dimensionless quantities related to sanding. The parameters taken into consideration included Reynold's Number, Loading factor, Gas-Liquid Ratio and water-cut. An equation for predicting sand production rate is the output from this work. Model validation was carried out on the developed model and the deviation was less than 5% demonstrating the validity of the proposed model. The developed model, when compared to existing models, forecasts superior outcomes, particularly when the boost factor and GLR are very high. The study has a wide range of practical implications, including reservoir management, general well completion design, and field development plans and economics.

***Keywords:*** *Sand production rate; Response surface methodology; Reynold's number; Loading factor; Water cut, Gas liquid ratio.*

## 1. Introduction

A significant proportion of the world's oil and gas reserves is contained in weakly consolidated sand stone reservoirs and hence is prone to sand production [1]. Sand produced has no commercial value. Not only does formation sand clog wells and lower recovery rates, but it also erodes equipment and collects in surface containers. The most common way to control formation sand is to either reduce production rate or employ a variety of different methods, both of which are expensive. Due to the above stated challenges, sand production is a major issue during oil and gas production especially in Pliocene, younger tertiary basins and even in more compact formations worldwide. Material degradation is a key process leading to sanding. However, these approaches have limitations as they can only predict catastrophic sand failure while neglecting the continuous changes and post failure characterization of the formation and the volume of the sand. Real-time, effective sand management is necessary because of the financial, operational, and safety ramifications of sand failures [2]. Passive preventive tactics and sand control measures are the two basic categories into which sand management techniques have been divided [3]. However, some researchers had earlier emphasized that creating a whole sand management strategy necessitates assessment of the formation strength, stress characterisation, failure modelling, sand exclusion studies, sand rate and size prediction, and utilization of field sand rate data [4]. The accurate assessment of the quantity and size of the produced sand is arguably the biggest problem in the sand management chain. This is necessary to ensure proper sand control facility design and to keep erosion limits for chokes and pipes within acceptable bounds. Theoretical models, laboratory experiments, and field observations are techniques for forecasting sanding rates. Theoretical models based on several sand failure processes have been published in which [5–7] are a few examples. An analytical model for forecasting the start of sanding from Canadian heavy oil sands was created in 1994 by Geilikman *et al.* [8]. In 1996 van der Hoek *et al.* drew on the theoretical and experimental research of Geilikman *et al.* [8] to arrive at their own conclusions [9]. Neural network technology was proposed and applied for modelling sand production [10]. In fact, other researchers developed a number of additional models for predicting the commencement of sand [1,11]. The majority of modern models, however, have relied on geo-mechanical concepts to forecast sand output after the initial onset.

Sand production has been found to occur in moderately consolidated and poorly cemented sandstones. A lot of Empirical, Analytical and Numerical Models for Predicting sand production rate is made available in the literature. A disadvantage of these Models is that they require humongous rock mechanical data as an input data which are not readily available in practical field. Additionally, these models call for laborious calculations that are impractical when prompt sand control choices are required. The difficulty to access the field input data is the major challenge in carrying out this project. This work developed an easy to use model that can be used to predict sand production rate from oil wells in the Niger Delta.

## 2. Literature review

Using geomechanics and fluid flow, Chin & Ramos [12] developed a sand production model to measure volume of sand output throughout the declination phase, the early drawdown phase and bean-up phase. Their proposed model was then tested by carrying out comparison with the data from full-scale laboratory perforation sanding tests and the simulated model from the proposed model. They came to the conclusion that flowrate, producing time, oil viscosity and rock strength were the key factors that regulate volumetric sand production. The primary issue with oil and gas fields that produce from elastic sediments is the control of formation sand [3]. Sand Sediments produced during the younger Tertiary eras are usually problematic, and we expect sand production problems in wells that are not consolidated when completion is about to be done. In older formations, Poor completion and production techniques consequently reduce in-situ rock strength and lead to sand failure [3]. The most fundamental elements of sand production were described as rate of fluid production, fluctuating in-situ stresses and strength of the formation. It is important to note that methods of predicting sand production include laboratory mechanical rock testing, analogy, production tests, well log analysis and laboratory tests. The paper came to the conclusion that the data needed to predict propensity of sanding as rock dynamic elastic constants, formation intrinsic strength, production test data and log data. Various methodologies have been proposed and are documented in the literature to assess the conditions of petroleum reservoirs with respect to onset sand production [13].

Sand production can be anticipated when product of Shear modulus (G) and bulk modulus (Kb) known as elastic constants is less than 8E11 [3]. It is important to note that this value 8E11 is known as the threshold value below which sand production will occur. The elastic constants are derived from density and acoustic logs. Since sanding can only take place or sand can only be formed after the formation fracture or have failed. Stein and Hilchie [14] ignored the influence of other factors that affect sand output while correlating sand production from the reservoir with production rate. Operational characteristics that affect sand output, such as flowrate, bean-up pattern, and drawdown, are deemed to be overly numerous and vary from field to field. Unreliable predictions of sand production may emerge from using only a small subset of these characteristics and extrapolating the findings from one field to another. According to Willson *et al.* [15] , while determining the basic strength of a formation, U, the collapse pressure of a thick-walled cylinder test was utilized. The authors also noted that the largest effective tangential stress should be less than the formation's effective strength, U, adjacent to the hole in order to prevent sand production. Some researchers recommended making a wide range of analytical and numerical sanding onset prediction models available in literatures [16]. These models call for enormous input parameters related to rock mechanics that are infrequently present in field practice and necessitate laborious computations, such as those required by finite element models, which are impractical in situations requiring quick sand control decisions. To calculate the critical drawdown and flow rate at which sand formation commences, they constructed a straightforward and user-friendly mechanistic sanding onset prediction model.

According to Walton *et al.* [17], sandstone formations are used to drill and produce a sizable portion of the world's oil and gas wells. The in-situ stress and subsequent changes brought on by activities associated to petroleum, such as drilling and production, are what determine the mechanical rigidity and strength of these sandstone formations. The formation failure or breakdown that results when the mechanical strength is exceeded may cause wellbore instability and, as a result, sand production. Han *et al.* [18] laid emphasis on sand prediction and management and stated that they were placed mainly on the unconsolidated and poorly cemented sandstones thereby creating corresponding failure models, because of their high tendency to produce sand while neglecting more compact formations. These formations however, have been seen to fail at some point, producing sand in the course of getting the crude oil to the surface, hence, causing adverse problems that are very costly to remedy. For Niger-Delta formations subject to open-hole completions, Adeyanju & Oyekunle [19] developed a geo-mechanical model for a coupled reservoir that can be utilized to forecast volumetric sand production and related wellbore stability. The model is based on mixture theory, and for each of the relevant phases-solid matrix, fluidized solid, oil, water and gas phase- mechanics and laws of conservation equations are developed. Results indicate that flow velocity, confining pressure, pressure drawdown, and fluid viscosity have a significant impact on how much sand is produced.

Today's oil and gas sector relies on conventional sand prediction methods that are based on field observation, laboratory sand prediction trials, and theoretical or numerical modelling [20]. A link between sand production and operational field characteristics is typically attempted using methodologies based on field observation and experience. These correlations and models are often developed using a small sample from a large collection of variables that may have an impact on the production of sand. Azadbakht *et al.* [21] developed a numerical model that can predict the volumetric sand production rate in injector wells based on the hypothesis that sand production in injector wells is primarily related to the back-flow and cross-flow generated during shut-in as well as water-hammer pressure pulsing in wellbore as a result of rapid flow rate changes. Their model for explaining the parameters necessary for sanding initiation and propagation included a given set of criteria employing geomechanics principles and fundamental physics of sand formation. The numerical model considered the impacts of water weakening effect, cross-flow effect and rock strength on the sanding behaviour of injectors. Okereke *et al.* [22] accounted for gas-liquid ratio, Reynold's number, loading factor and water cut as the intrinsic parameters that affect sanding potential. They came to a conclusion that the proposed model gave a more accurate result when validated with boost factor that is the GLR considered to be significantly high. Every reservoir is considered to have a peculiar sand production rate correlation indices which represents its tendency to produce sand. Understanding a formation's mechanical strength is essential for predicting sand production and recommending sand control completion [23]. Since cores aren't always available, the model provided a method to measure rock strength without being restricted by core testing. They conducted triaxial and hydrostatic testing to establish the failure envelope. The study's findings showed that a single normalized failure envelope exists and that, with knowledge of the critical pressure, it can be used to characterize the failure envelope for a sandstone formation. There is a correlation between the critical pressure and compressional wave speed (at equivalent depths of burial). Other researchers carried out research on sand output forecasting in gas wells [7]. The suggested technology was evaluated on 13 fields along the US Gulf Coast, and it has since been widely used throughout the world by the now-defunct Arco. The rock's strength was assessed using core testing and log correlations, and the results were contrasted. The prediction procedure is distinct from the popular log-based sand prediction methodology.

## 3. Methodology

The data used for this work was obtained from literature [22] while an experimental design software was used to develop the model. The intrinsic parameters that affect sanding potential are gas-liquid ratio, water cut, loading factor and Reynolds number which are the input data to the model being developed in this work. The responses obtained from the field which is defined is the sand production rate is the output variable.

### 3.1. Reynold's number (Re)

Since Reynold's Number is a function of reservoir parameters such as size, density, perforation number, permeability, viscosity and flow rate per perforation, ,it is of great importance to consider that as a key parameter that influence sanding. Reynold's Number helps to define flow patterns in different fluid flow situations. Reynold's number can be evaluated using equation 1.

$$R_e = 131735 * 10^{-6} * \frac{K\beta\rho v}{\mu} \tag{1}$$

where β, the non-Darcy flow coefficient is given by equation 2

$$\beta = \frac{265 * 10^{-10}}{K\hat{}12} \tag{2}$$

where K= permeability in millidarcy (MD); β = non-Darcy flow coefficient ( ft-I); ρ = density ( Lb/ft^3); v = velocity of the fluid crossing the lateral perforation (inches/seconds); μ = fluid viscosity (cP).

### 3.2. Loading factor (LF)

Loading factor which is defined as the state stress acting on a perforation borehole or tunnel. Loading factor is function of in-situ stresses, well trajectory, reservoir pressure, drawdown and depletion, Thick wall cylinder strength (TWC). TWC is defined as the fundamental strength of the formation. Loading factor can be evaluated using the equation 3 and the parameter F in equation 3 can be evaluated using equation 4.

$$LF = \frac{3S_H - S_V - 2P_{wf} - F(P_R - P_{wf})}{3.1(TWC_{sp})} \tag{3}$$

$$F = \frac{(1-2v)(C_b - C_r)}{(1-v)C_b} \tag{4}$$

where $S_H = S_V$ = total stresses on a plane perpendicular to the axis of wellbore; $P_R$ = reservior pressure (Psia); $P_{wf}$ = wellbore flowing pressure (Psia); F= poro-elastic constant; V= Poisson's ratio; $C_b = C_r$ = bulk and grain rock compressibility; $TWC_{sp}$= thick walled-cylinder strength.

It is good to note the following conditions;

1. when Loading factor is less than 1, the formation hasn't failed
2. when Loading factor is greater than 1, the formation has failed and subsequently, sand may be produced.

### 3.3. Water-cut (WC)

The percentage of water generated in a well relative to all liquids produced is known as the "water-cut." Oil and water mix as they flow out the well and the field fills with water. Water cut refers to how much water is present in these wells. It is important to keep in mind that water cut increases the likelihood of sand formation, which is based on perforation strength and sanding depends on the mineralogical constitution of the sandstone and extent of residual water saturation. Water cut boosts sand production; hence its impact was taken into consideration when this work's model was being developed. Water cut can be calculated using the equation 5.

$$WC = \frac{water\ produced(\frac{bbl}{m})}{water\ produced(\frac{bbl}{m}) + oil\ produced(\frac{bbl}{m}) * 100\%} \tag{5}$$

### 3.4. Gas-liquid-ratio (GLR)

The ratio of a volume of gas to a volume of liquid at the same pressure and temperature is known as the gas-liquid ratio. This is very important for characterizing the behaviour of a reservoir. High GLR predicts better SPR results. GLR can be calculated using equation 6

$$GLR = \frac{Gas\ Produced(\frac{scf}{m})}{Gas\ produced(\frac{scf}{m}) + Liquid\ Produced} \tag{6}$$

where Liquid produced= water produced + oil produced (7)

## 4. Methods

### 4.1. Model development data set

The data used for this work as discussed in section 3.2 above comprises of input parameters and an output parameter. The input parameter comprises of water cut, gas liquid ratio, Reynolds number and loading factor. The output parameter which is Sand production Rate (SPR) considered as measured in pounds per month (lb/month) was generated using the mathematical model equation as developed by Okereke *et.al.* [22] which is dependent of the parameters listed above. The input and output data set used for this machine learning model can be summarized in Tables 1 and 2 respectively.

Table 1. Summary of Input data set used for model development.

| Input parameters | Mean | Median | Mode | Max | Min | No of counts |
|---|---|---|---|---|---|---|
| RE | 0.064032 | 0.0191 | 0.0161 | 0.36845 | 0.0018 | 72 |
| LF | 1.6567 | 1.715 | 1.45 | 2.54 | 0.54 | 72 |
| WC | 0.4 | 0.41 | 0.33 | 0.905 | -0.165 | 72 |
| GLR | 3081.833 | 442 | 107 | 22504 | -7340 | 72 |

Table 2. Summary of output data used for model development.

| Output parameter | Mean | Median | Mode | Max. | Min. | No of counts |
|---|---|---|---|---|---|---|
| (bbl/month) | 1116.819 | 86.488 | --------- | 7839.3 | 0.801 | 72 |

The input and output data were obtained from the research work conducted by Okereke *et al.* [22] for three different wells with the data being merged to produce a single model which can work for the 3 wells. This was done due to the non- availability of real field data. The equation was inserted into an excel data sheet with the corresponding input data and then the output data set was generated. The equations according to Okereke *et al.* [22] are shown in Equations 8, 9, and 10 for OMD64/D1, OMD64/D2, and OMF65/D3 respectively.

$$Q_{sand} = 7938.66 * R_e^{0.8957} * LF * \exp(0.5657w + 1.22E - 05G) \text{ for OMD 64/D1} \tag{8}$$

$$Q_{sand} = 196.12 * R_e^{1.2771} * LF * \exp(6.8081w + 3.28E - 05G) \text{ for OMD 64/D2} \tag{9}$$

$$Q_{sand} = 912.78 * R_e^{1.015} * LF * \exp(0.1407w + 5.27E - 04G) \text{ For OMD 65/D3} \tag{10}$$

This equation was merged together to develop a single machine learning model that can predict sand production rate measured in bbl/month for the wells. This data set was modelled using Design Expert Software which makes use of response surface model (RSM) as the modelling tool. The features of this software shall be discussed in section 4.2.1.

### 4.2. Model development environment

In this project work, the SPR model was developed using the Design Expert software which works based on response surface method. The model developed using this software is a data driven cubic model which can predict sand production rate. This section can be summarised in this 2 sub-heading. 1. Introduction to machine learning; 2. Design Expert Software; 3. Response Surface Model (RSM).

#### *4.2.1. Design of experiment (design expert)*

A statistical software program used exclusively for experiment design is called Design-Expert (DOE). This provides characterisation, screening, optimization, comparison tests, mixture designs, robust parameter, and combined designs. Up to 50 factors can be screened using the test matrices provided by Design-Expert. The fact that design-expert offers analysis of variance (ANOVA) has statistical significance. Additionally, it offers graphical tools that show how each component affects the intended results and subsequently highlight anomalies in the data set. This offers a wide range of graphical and analytical techniques for model interpretation and fitting. There are 5 stages of methodology of Design of Experiment which includes: a. planning; b. screening; c. optimization; d. robustness testing; e. verification.

#### *4.2.2. Response surface method/model (RSM)*

A mathematical and statistical tool used for designing, optimizing, and upgrading the process is the response surface model or response surface method. When a number of independent factors, often referred to as predictor variables, have an impact on the dependent variable or response, this is crucial to the analysis of the issue. RSM is employed to strengthen a process in the face of unpredictable noise and even pursue numerous objectives. Additionally, it is employed to meet predetermined goals, lessen process variability, and maximize or minimize responses. The purpose of response surface modelling is summarised as follows:1. Analyse and rectify process problems and weak points; 2. Find optimal or improved settings; 3. Robust the process or product against the external influence.

#### *4.2.3. Model validation*

Due to the fact that models cannot be trusted, it is of great importance to carry out model validation. In the course of this project work, model validation was done using 2 method: 1. Error analysis; 2. Cross plots.

#### *4.2.4. Error analysis*

By isolating, observing, and diagnosing inaccurate machine learning predictions, error analysis helps us to identify where the model performs well and poorly. In this work, the error analysis was done using the following method: a. Mean absolute error (MAE);b. Mean square error (MSE); c. Root mean square error (RMSE); d. Coefficient of determination score ($R^2$ score).

**Mean absolute error**

MAE is a measure of errors between paired observations expressing same phenomenon. It is the amount of error that exist in a measurement, which shows the difference between true value and predicted values. Eqn. (11) shows the mathematical equation for MAE.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \quad (11)$$

where, $MAE$ = mean absolute error, $y_i$ = predicted value, $x_i$ = true value, $n$ = total No. of data points.

**Mean square error**

MSE measures the average of the squares of the errors, i.e. it measures the average of the squares of the errors. It accounts for the amount of error in mathematical and statistical models, it basic mathematical principle is shown in eqn. (12).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \quad (12)$$

where, $MSE$ = mean square error, $n$ = total No. of data points, $Y_i$ = true or observed values, $\hat{Y}_i$ = predicted values.

**Root mean square error**

This is the square root of the mean of the square of all of the error. It is considered an excellent general purpose error metric for numerical predictions. It is a frequently used measure of the differences between values predicted by the model and values observed. It is given by the formula in eqn. 13.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad (13)$$

where; $RMSE$ = root mean square error, $n$ = total No. of data points, $Y_i$ = true or observed values, $\hat{Y}_i$ = predicted values

**Coefficient of determination score**

This is known as the R-Squared score. It is the proportion of the variation in the dependent variable that is predicted from the independent variable. Best possible score is 1.00 and it can be negative (since some models can be arbitrarily worse). It can be represented by the formula in eqn. 3.9d

$$R^2 = 1 - \frac{RSS}{TSS} \tag{14}$$

where; $R^2$ = coefficient of determination, $RSS$ = sum of squares residuals, $TSS$ = total sum of squares.

## 5. Results and discussion

### 5.1. Results

A summary of the results from this work are presented in tables and plots. The results are grouped as 1. Machine learning model developed; 2. Comparison of predicted vs actual model; 3. Effects of input parameters on predicted SPR; 4. Statistical results;5. Results from error analysis carried out;6. Cross plots.

#### *5.1.1. Machine learning model developed*

The developed model is a function of Reynold's number, loading factor, gas-liquid ratio and water-cut as shown in Equation 1.

$$Q_{sand} = \mathrm{F}(\mathrm{LF}, \mathrm{RE}, \mathrm{GLR}, \mathrm{WC}) \tag{15}$$

The developed SPR model is given by equation 2

$$Q_{sand} = 379.939 + -9{,}416.62 * Re + -480.763 * LF + -997.747 * WC + 0.00235646 * GLR + 12{,}765 * Re * LF + 16{,}959.7 * Re * WC + 929.591 * LF * WC \tag{16}$$

where $Q_{sand}$ = sand production rate in bbl/month, Re= Reynold's number, LF= loading factor, WC = water cut, GLR= gas liquid ratio in scf/bbl.

#### *5.1.2. Comparison of predicted SPR vs actual SPR*

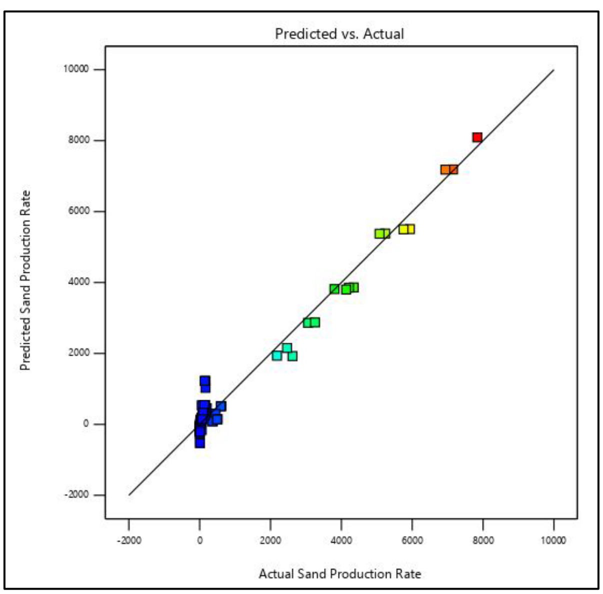


Figure 1. Plot of predicted SPR vs actual SPR.

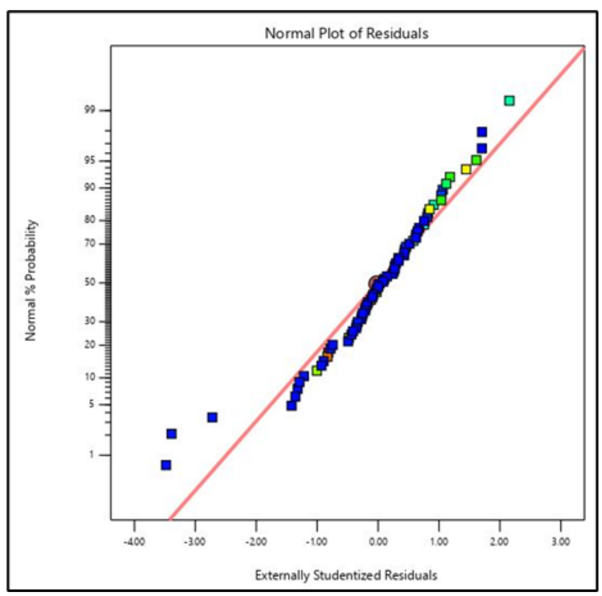


Figure 2. Normal probability plot of the residuals,

The percentage or rate of deviation of the predicted SPR from the actual SPR is provided in the appendix section of this work. The good thing to note is that the average deviation of the residual is 2.27 which is a good result. The plot of predicted vs actual SPR is shown in Figure 1. From the Figure 1 you will notice that the points are close to the fitted line with narrow confidence bands which is an indication of a good fit. This means that the residuals are i.e the difference between the actual and the predicted values isn't of much significant value. This can be better illustrated using the normal probability plot of the residuals as shown in the Figure 2. From Figure 2, the normal probability plot of the residuals is approximately linear thereby supporting the fact or condition that the error terms are normally distributed. The straight diagonal line here shows a normally distributed data.

### *5.1.3. Effects of input parameters on predicted SPR*

This section summarises the effects of Re, LF, GLR and WC on predicted SPR and the results are presented in Figure 3.

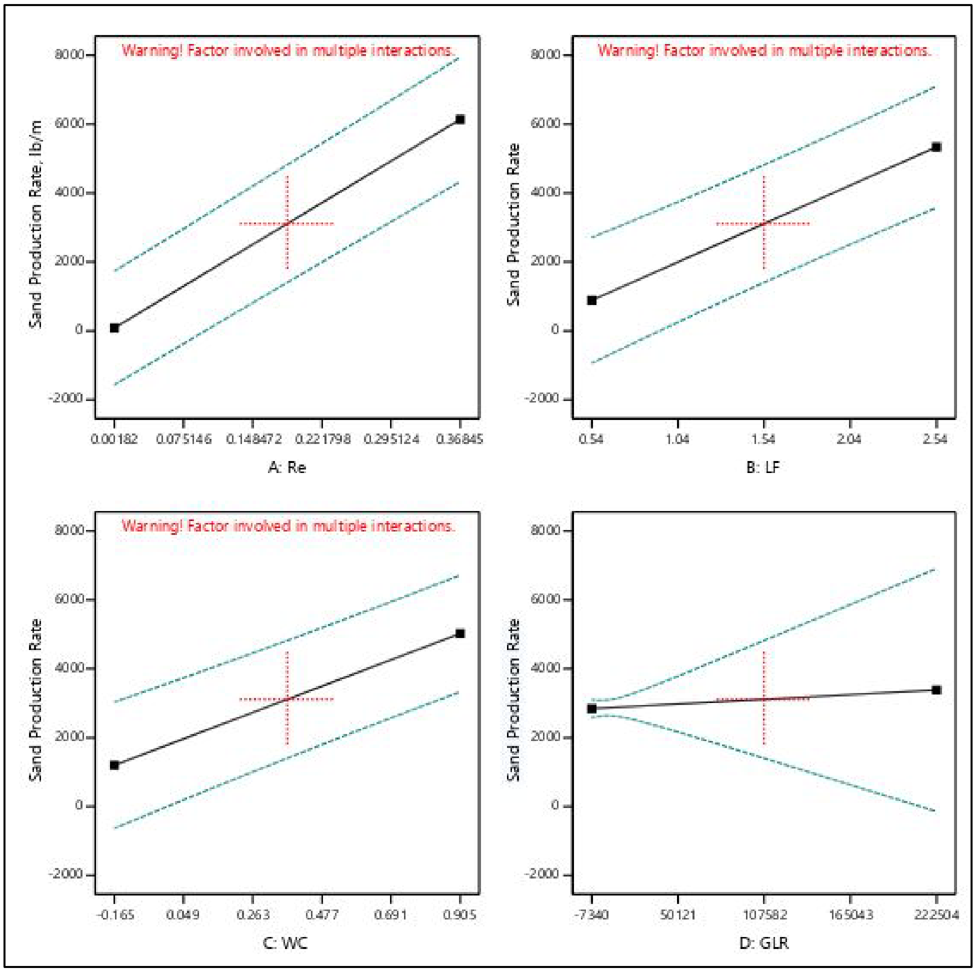


Figure 3. Effect of Re, LF, WC, and GLR on sand production rate.

### *Effect of water cut on predicted SPR*

This is presented in the bottom left picture of Figure 3 which illustrates how water-cut affects the predicted sand production rate. The response of water-cut thus determined show that the sand production rate increases with water cut for all values of loading factor and Gas liquid ratio. The cross-like lines in the plot indicates factors with multiple interactions during the modelling.

### *Effect of GLR on predicted SPR*

It can be deduced that GLR increases with increasing water-cut and vice versa (Bottom right). In order to avoid sanding, the Gas-liquid ratio should be kept fairly low. However, Models predicts more accurately when GLR is relatively high. The GLR versus SPR trend shows that GLR is directly proportional to SPR as such means that SPR increases as GLR increases.

***Effect of Reynolds number on predicted SPR***

As a standard, it can be deduced that for values of Re<0.1, sanding occurs. This indicates that sand production rate is dominated by the loading factor in this case. The top left picture of Figure 3 shows the effect of increasing or decreasing Reynold's number on sand production rate.

***Effect of loading factor on predicted SPR***

As earlier stated, LF>1 indicates propensity of the reservoir to produce sand. Loading factor increases with depth and this increase initiates sand production. This is better described in the Figure 3 (top left picture).

### *5.1.4. Statistical results*

This section covers for the statistical analysis carried out on the developed model to account for errors. This comprises of ANOVA and fit statistics.

***Analysis of variance (ANOVA)***

ANOVA is a statistical technique that isolates observed variance data into various components for use in further or additional tests. This seeks to learn more about how the dependent and independent variables are related.

Table 3. Analysis of variance on developed model.

| Source | Sum of squares | df | Mean Square | F-Value | P-Value | |
|---|---|---|---|---|---|---|
| **Model** | 2.932E+08 | 10 | 2.932E+07 | 243.57 | <0.0001 | significant |
| A-Re | 9.175E+06 | 1 | 9.175E+06 | 76.22 | <0.0001 | |
| B-LF | 4.130E+06 | 1 | 4.130E+06 | 34.30 | <0.0001 | |
| C-WC | 3.410E+06 | 1 | 3.410E+06 | 28.32 | <0.0001 | |
| D-GLR | 2.825E+05 | 1 | 2.825E+05 | 2.35 | 0.1307 | |
| AB | 4.755E+06 | 1 | 4.755E+06 | 39.50 | <0.0001 | |
| AC | 3.917E+06 | 1 | 3.917E+06 | 32.54 | <0.0001 | |
| AD | 2.496E+05 | 1 | 2.496E+05 | 2.07 | 0.1550 | |
| BC | 3.466E+05 | 1 | 3.466E+05 | 2.86 | 0.0958 | |
| BD | 3178.55 | 1 | 3178.55 | 0.0264 | 0.8715 | |
| CD | 91955.11 | 1 | 91955.11 | 0.7639 | 0.3855 | |
| **Residual** | 7.343E+06 | 61 | 1.204E+05 | | | |
| **Cor Total** | 3.005E+08 | 71 | | | | |

The following can be summarized from Table 3:

a. The **Model F-value** of 243.57 implies the model is significant. This is the ratio of explained variance to unexplained variance. There is only a 0.01% chance that an F-value this large could occur due to noise.

b. **P-values** less than 0.0500 indicate model terms are significant. P value describes how likely the data would have occurred under the null hypothesis of the statistical test. In this case A, B, C, AB, AC are significant model terms.

c. Values greater than 0.1000 indicate the model terms are not significant. If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

***Fit Statistics***

This covers for the mean, standard deviation, coefficient of variance, and adequate precision of the developed model. The statistics is provided in Table 2

Table 4. Fit statistics of developed model.

| | | | |
|---|---|---|---|
| Standard deviation | 346.96 | Adjusted $R^2$ | 0.9716 |
| Mean | 1116.82 | Predicted $R^2$ | 0.9679 |
| C.V % | 31.07 | Adequate Precision | 63.6447 |
| $R^2$ | 0.9756 | | |

The predicted $R^2$ of 0.9679 is in reasonable agreement with the adjusted $R^2$ of 0.9716; i.e. the difference is less than 0.2. Adequate precision measures the signal to noise ratio. A ratio greater than 4 is desirable. In this work, an adequacy in precision of 63.645 was obtained indicating an adequate signal and can be used to navigate the design space.

#### *5.1.5. Model validation (error analysis)*

Table 5 summarises the values gotten when error analysis was carried out on the developed. The percentage error plots are shown in the appendix section of this work. The statistical error metrics which includes mean square error, mean absolute error and root square error. From the table and plots, it is observed that the model has minimal error since the $R^2$ value is high. This simply confirms and concludes the fact that the developed model is reliable and can be deployed.

Table 5. Error analysis result.

| | | | |
|---|---|---|---|
| Mean absolute error | 0.453 | Score | 0.9756 |
| Mean square error | 1.035 | Adjusted score | 0.9716 |
| Root mean square error | 1.017 | Adequate precision | 63.6447 |

## 6. Conclusion

The petroleum industry is becoming increasingly concerned with sand production, which is common in the Niger Delta and presents technical, operational, and financial difficulties. Literature has devoted a lot of time to the creation of sanding prediction tools and efficient management techniques. However, the majority of the published theoretical models have been supported by data from other petroleum regions than the Niger Delta or evidence from laboratories. Data analytics has gained a more relevance in the oil and gas industry recently as the industry seeks to analytically use data for optimal productivity and reduction of cost/time management purposes. As data is growing in this industry, the ability of the industry to use these data alongside intelligent models to solve problems is a very important and as well useful predicting sand production rate. In reservoir engineering and other disciplines in the oil and gas industry, history matching is often done and futuristic predictions are made towards the prediction of either the optimal production rate or the oil in place in a particular reserve. Same is applied in the use of intelligent models to make predictions towards the propensity of a reservoir to produce sand or not with respect to existing dataset recorded.

A machine learning model for estimating the rate of sand production in Niger Delta oil fields was developed. The developed model is very reliable since the average deviation of the residual is 2.27 which is fair enough. The proposed model's predictions will help in the planning of capital investments connected to sanding, such as delaying the installation of sand control mechanisms, digging disposal wells, and increasing the capacity of sand handling facilities.

The minimal deviation between the observed SPR and predicted SPR is an indicator that gas-liquid ratio, water cut, loading factor and Reynold's number plays a very important role in predicting sanding tendencies especially in loosely packed sedimentary basins like Niger-Delta.

SPR Models tends to predict accurately when the GLR is significantly high. It can be deduced that the input parameters vary directly to the rate of sand production. Hereby, any slight increase in any of the parameter will cause an increase in the rate of produced sand. To this effect, the parameters should be kept as low as possible. Unconsolidated formations are more likely to produce sand due to the fact that they are loosely arranged and unstratified.

## 7. Recommendation

For better results, it is recommended that the developed SPR model should be validated using additional data from diverse depositional conditions throughout the Niger Delta. Future research should take into account the variable rock mechanical characteristics of the examined reservoir. The major challenge from this work is the need for humongous data set which is not readily available for practical use. It is recommended that petrophysical data are made available in literatures in order to ease future predictions.

From the results, it is evident that machine learning models can be used for the predictions of sand production rate. Although, these models have not been deployed in real time predictions, as such more research can be done in this regard and deployment executed in nearest future for sand production prediction.

***Nomenclature***

| | |
|---|---|
| *ANOVA* | *Analysis of variance* |
| *CV* | *Coefficient of variance* |
| *GLR* | *Gas liquid ratio* |
| *KB* | *Bulk modulus* |
| *LF* | *Loading factor* |
| *LB/M* | *Pounds per month* |
| *RE* | *Reynold's number* |
| *RSM* | *Response surface model* |
| *SPR* | *Sand production rate* |
| *MAE* | *Mean absolute error* |
| *MSE* | *Mean square error* |
| *RMSE* | *Root mean square error* |
| *PPTB* | *Pounds per thousand barrels* |
| *TWC* | *Thick wall cylinder* |
| *UCS* | *Uniaxial compressive strength* |
| *SCF/BBL* | *Standard cubic feet/ barrels* |

**References**

[1] Nouri A, Vaziri H, Belhaj H, Islam R. Sand-production prediction: A new set of criteria for modeling based on large-scale transient experiments and numerical investigation. SPE J., 2006; 11: 227–237. https://doi.org/10.2118/90273-PA
[2] Oluyemi GF, Oyeneyin MB. Analytical Critical drawdown (CDD) Failure Model for real time Sanding Potential based on Hoek and Brown Failure Criterion. J. Pet. Gas Eng., 2010; 12: 16–27.
[3] Osisanya SO. Practical guidelines for predicting sand production. Soc. Pet. Eng.- Niger. Annu. Int. Conf. Exhib. 2010, NAICE, 2010; 1: 284–291. https://doi.org/10.2118/136980-ms
[4] Burton R, Chin L, Davis E, Enderlin M, Fuh G-F, Hodge R, Ramos R, van DeVerg P, Werner M, Mathews W, Petersen S. North Slope Heavy-Oil Sand-Control Strategy: Detailed Case Study of Sand Production Predictions and Field Measurements for Alaskan Heavy-Oil Multilateral Field Developments. In: SPE Annual Technical Conference and Exhibition (2005).
[5] Coates GR, Denoo SA. Mechanical Properties Program Using Borehole Analysis And Mohr's Circle, (1981).
[6] Bratli RK, Risnes R. Stability and Failure of Sand Arches. SPE J., 1961; 236–248
[7] Weingarten JS, Perkins TK. Prediction of sand production in gas wells: methods and Gulf of Mexico case studies. J. Pet. Technol., 1995; 47: 596–600. https://doi.org/10.2118/24797-PA
[8] Geilikman MB, Dusseault MB, Dullien FA. Sand production as a viscoplastic granular flow. Proc.-SPE Int. Symp. Form. Damage Control., 1994; 41–50. https://doi.org/10.2523/27343-ms
[9] van den Hoek PJ, Geilikman MB. Prediction of Sand Production Rate in Oil and Gas Reservoirs. Proc. - SPE Annu. Tech. Conf. Exhib., 2003; 3911–3919. https://doi.org/10.2118/84496-ms
[10] Kanj MY, Abousleiman Y. Realistic sanding predictions: a neural approach. Proc. - SPE Annu. Tech. Conf. Exhib. 2, (1999). https://doi.org/10.2118/56631-ms

[11] Rahmati H, Jafarpour M, Azadbakht S, Nouri A, Vaziri H, Chan D, Xiao Y. Review of Sand Production Prediction Models. J. Pet. Eng., 2013: 1–16. https://doi.org/10.1155/2013/864981

[12] Chin LY, Ramos GG. Predicting Volumetric Sand Production in Weak Reservoirs. Proc. SPE/ISRM Rock Mech. Pet. Eng. Conf., 2002; 161–170. https://doi.org/10.2523/78169-ms

[13] Ghiasi MM, Ghasemi MF, Heidaripour V, Mohammadi AH. Distinct Methodologies to Assess the Conditions of Petroleum Reservoirs with Respect to Onset of Sand Production. Pet. Coal. 61, 339–350 (2019). https://doi.org/10.1021/ac00060a011

[14] Stein N, Hilchie DW. Estimating the Maximum Production Rate Possible from Friable Sandstones Without Using Sand Control. J. Pet. Technol., 1972; 24: 1157–1160. https://doi.org/10.2118/3499-pa

[15] Willson SM, Moschovidis ZA, Cameron JR, Palmer ID. New Model for Predicting the Rate of Sand Production. Proc. SPE/ISRM Rock Mech. Pet. Eng. Conf., 2002; 152–160. https://doi.org/10.2523/78168-ms

[16] Yi X, Valkó PP, Russell JE. Predicting Critical Drawdown for the Onset of Sand Production. Proc. - SPE Int. Symp. Form. Damage Control.,2004; 807–818. https://doi.org/10.2523/86555-ms

[17] Walton IC, Atwood DC, Halleck PM, Bianco LCB. Perforating Unconsolidated Sands: An Experimental and Theoretical Investigation. Proc. - SPE Annu. Tech. Conf. Exhib. 2001;1151–1164. https://doi.org/10.2118/71458-ms

[18] Han K, Shepstone G, Harmawan I, Er U, Jusoh H, Sue Lin L, Pringle D, Koya R, Carney S, Barker L, Morita N, Papamichos E, Cerasi P, Sayers C, Heiland J, Bruno M, Diessl J. A comprehensive study of sanding rate from a gas field: From reservoir to completion, production, and surface facilities. SPE J., 2011; 16: 463–481. https://doi.org/10.2118/123478-PA

[19] Adeyanju OA, Oyekunle LO. Prediction of volumetric sand production and stability of wellbore in a Niger-Delta formation. Soc. Pet. Eng. - Niger. Annu. Int. Conf. Exhib. 2010, NAICE. 2010; 1: 134–153. https://doi.org/10.2118/136965-ms

[20] Veeken CAM, Davies DR, Kenter CJ, Kooijman AP. Sand production prediction review. Developing an integrated approach. Proc. - SPE Annu. Tech. Conf. Exhib. Pi, 1991; 335–346. https://doi.org/10.2523/22792-ms

[21] Azadbakht S, Jafarpour M, Rahmati H, Nouri A, Vaziri H, Chan D. A numerical model for predicting the rate of sand production in injector wells. Soc. Pet. Eng.-SPE Deep. Drill. Complet. Conf. 2012; 636–643. https://doi.org/10.2118/156394-ms

[22] Okereke N, Ogbuka V, Izuwa N, Kara F, Nwogu N, Nwanwe O, Baba Y, Kanshio S, Odo J Oguama I. Advanced Mathematical Model for Prediction of Sand Production Rate: A Niger-Delta Case-Study. Paper presented at the SPE Nigeria Annual International Conference and Exhibition, Virtual, August 2020.Paper Number: SPE-203692-MS. https://doi.org/10.2118/203692-MS

[23] Zhang JJ, Rai CS, Sondergeld CH. Mechanical strength of reservoir materials: Key information for sand prediction. SPE Reserv. Eval. Eng., 2000; 3: 127–131. https://doi.org/10.2118/62499-PA

*To whom correspondence should be addressed Onyebuchi Ivan Nwanwe, Department of Petroleum Engineering, Federal University of Technology, Owerri Nigeria, E-mail: onyebuchi.nwanwe@futo.edu.ng*
*ORCID: 0000-0002-9886-205X*