International Journal of Advanced Science and Technology
Vol.107 (2017), pp.21-32
http://dx.doi.org/10.14257/ijast.2017.107.02

Statistical Modelling and Analysis of Borehole Computer Experiment

Kazeem Adewale Osuolale[1], Waheed Babatunde Yahya[2] and Babatunde Lateef Adeleke[3]

Department of Statistics, University of Ilorin, P.M.B. 1515, Ilorin, Nigeria
whereisqosimadewale@gmail.com

Abstract

Computer experiments are popular techniques adopted in modern businesses, engineering, scientific and technological applications in the recent years. Its flexibility and wide applicability has made it more accepted than the classical physical experiments. The Design and Analysis of Computer Experiments (DACE) is fast growing in statistical experimental designs. In this work, an Orthogonal Array-based Latin Hypercube Design, that is, OA (N, k) LHD was applied for the development of borehole computer experiment. A computer experiment was conducted based on the OA (49, 8) LHD using a borehole computer model. The borehole computer model was used to simulate the real life borehole experiment. The Gaussian stochastic process (Gasp) model was employed to mimic the computer model in order to save time that may be required by a complex computer code and for the purpose of predictions of the flow rate of water at untried inputs. The Maximum Likelihood Estimation technique was used to estimate the parameters of the Gasp model. The results obtained using the Gasp model indicated that the radius (r_w) and the hydraulic conductivity (K_w) of the borehole were the most important factors that influenced the flow rate of water from an upper aquifer to a lower one. The fitted Gasp model was found to be very efficient since it yielded exact results on the test data cases. The model development and analysis were performed in MATLAB package.

Keywords: *Computer experiment, Gasp model, Orthogonal array-based Latin hypercube design, Borehole computer model, Space-filling design*

1. Introduction

An experimental design is the selection of inputs by which to compute the output of computer experiments in order to achieve specific aims. It has a matrix of input variable (***X***), where each column of ***X*** depicts a variable and each row is the combination of input variable values for a single experimental run. Conventional experimental designs originate from the theory of Design of Experiments (DOE) when physical experiments are performed [8] while space-filling designs are associated with computer experiments. Computer experiments are distinct from physical experiments because they have no random error and they deal with functions that are considered to have more complex behaviour. Properly designed experiments are essential for effective computer utilization. Deterministic computer experiments are more commonly employed in engineering, science and technology because the conventional physical experiments could require more time, money and some other resources to conduct. In some instances, the physical experiments could be difficult to perform. A computer experiment is an experiment performed using data obtained from a computer model instead of the physical process.

Received (June 10, 2017), Review Result (September 26, 2017), Accepted (October 8, 2017)

ISSN: 2005-4238 IJAST

The rapid growth in computer power has now made it possible to conduct deterministic experiments on simulators. The first computer experiment was reported to have been conducted by Enrico Fermi and colleagues [14] in Los Alamos in 1953 and since then, scientists in different disciplines have turned to computer experiments as a useful tool to understand their respective processes.

In this present work, a borehole computer experiment that utilizes a model which is a simple example of flow rate of water through a borehole from an upper aquifer to a lower aquifer that is separated by an impermeable rock layer was developed. The orthogonal array-based Latin hypercube design (OALHD) originally constructed by [12] is used to develop a borehole computer experiment and the Gasp model is employed to emulate a borehole computer model.

2. Material and Methods

The model development and analysis in this work were performed using MATLAB package. Orthogonal array Latin hypercube design (OALHD) was used to develop a borehole computer experiment through a model of flow rate of water and the Gasp model was subsequently used to emulate a borehole computer model. The results of OA (49, 8) LHD and its plot for bivariate projections among the eight input variables are provided in Table 1 and Figure 1 as constructed by the author in [13]. The approach employed by the author in [13] is different from the one employed in the construction of OALHD for three input variables computer experiment [10]. The borehole model has been investigated by several researchers including [16], [11], [7], [5] and [1] among others.

Table 1. OA (49, 8) LHD Constructed for Borehole Computer Experiment

OA (49, 8) LHD							
Design Points							
0.0102	0.0102	0.0102	0.0102	0.0102	0.0102	0.0102	0.0102
0.0306	0.1531	0.1531	0.2959	0.4388	0.7245	0.1531	0.8673
0.0510	0.2959	0.2959	0.5816	0.8673	0.4388	0.2959	0.7245
0.0714	0.4388	0.4388	0.8673	0.2959	0.1531	0.4388	0.5816
0.0918	0.5816	0.5816	0.1531	0.7245	0.8673	0.5816	0.4388
0.1122	0.7245	0.7245	0.4388	0.1531	0.5816	0.7245	0.2959
0.1327	0.8673	0.8673	0.7245	0.5816	0.2959	0.8673	0.1531
0.1531	0.0306	0.1735	0.1735	0.3163	0.4592	0.7449	0.1735
0.1735	0.1735	0.3163	0.4592	0.7449	0.1735	0.8878	0.0306
0.1939	0.3163	0.4592	0.7449	0.1735	0.8878	0.0306	0.8878
0.2143	0.4592	0.6020	0.0306	0.6020	0.6020	0.1735	0.7449
0.2347	0.6020	0.7449	0.3163	0.0306	0.3163	0.3163	0.6020
0.2551	0.7449	0.8878	0.6020	0.4592	0.0306	0.4592	0.4592
0.2755	0.8878	0.0306	0.8878	0.8878	0.7449	0.6020	0.3163
0.2959	0.0510	0.3367	0.3367	0.6224	0.9082	0.4796	0.3367
0.3163	0.1939	0.4796	0.6224	0.0510	0.6224	0.6224	0.1939
0.3367	0.3367	0.6224	0.9082	0.4796	0.3367	0.7653	0.0510
0.3571	0.4796	0.7653	0.1939	0.9082	0.0510	0.9082	0.9082
0.3776	0.6224	0.9082	0.4796	0.3367	0.7653	0.0510	0.7653
0.3980	0.7653	0.0510	0.7653	0.7653	0.4796	0.1939	0.6224
0.4184	0.9082	0.1939	0.0510	0.1939	0.1939	0.3367	0.4796
0.4388	0.0714	0.5000	0.5000	0.9286	0.3571	0.2143	0.5000
0.4592	0.2143	0.6429	0.7857	0.3571	0.0714	0.3571	0.3571
0.4796	0.3571	0.7857	0.0714	0.7857	0.7857	0.5000	0.2143
0.5000	0.5000	0.9286	0.3571	0.2143	0.5000	0.6429	0.0714

0.5204	0.6429	0.0714	0.6429	0.6429	0.2143	0.7857	0.9286
0.5408	0.7857	0.2143	0.9286	0.0714	0.9286	0.9286	0.7857
0.5612	0.9286	0.3571	0.2143	0.5000	0.6429	0.0714	0.6429
0.5816	0.0918	0.6633	0.6633	0.2347	0.8061	0.9490	0.6633
0.6020	0.2347	0.8061	0.9490	0.6633	0.5204	0.0918	0.5204
0.6224	0.3776	0.9490	0.2347	0.0918	0.2347	0.2347	0.3776
0.6429	0.5204	0.0918	0.5204	0.5204	0.9490	0.3776	0.2347
0.6633	0.6633	0.2347	0.8061	0.9490	0.6633	0.5204	0.0918
0.6837	0.8061	0.3776	0.0918	0.3776	0.3776	0.6633	0.9490
0.7041	0.9490	0.5204	0.3776	0.8061	0.0918	0.8061	0.8061
0.7245	0.1122	0.8265	0.8265	0.5408	0.2551	0.6837	0.8265
0.7449	0.2551	0.9694	0.1122	0.9694	0.9694	0.8265	0.6837
0.7653	0.3980	0.1122	0.3980	0.3980	0.6837	0.9694	0.5408
0.7857	0.5408	0.2551	0.6837	0.8265	0.3980	0.1122	0.3980
0.8061	0.6837	0.3980	0.9694	0.2551	0.1122	0.2551	0.2551
0.8265	0.8265	0.5408	0.2551	0.6837	0.8265	0.3980	0.1122
0.8469	0.9694	0.6837	0.5408	0.1122	0.5408	0.5408	0.9694
0.8673	0.1327	0.9898	0.9898	0.8469	0.7041	0.4184	0.9898
0.8878	0.2755	0.1327	0.2755	0.2755	0.4184	0.5612	0.8469
0.9082	0.4184	0.2755	0.5612	0.7041	0.1327	0.7041	0.7041
0.9286	0.5612	0.4184	0.8469	0.1327	0.8469	0.8469	0.5612
0.9490	0.7041	0.5612	0.1327	0.5612	0.5612	0.9898	0.4184
0.9694	0.8469	0.7041	0.4184	0.9898	0.2755	0.1327	0.2755
0.9898	0.9898	0.8469	0.7041	0.4184	0.9898	0.2755	0.1327

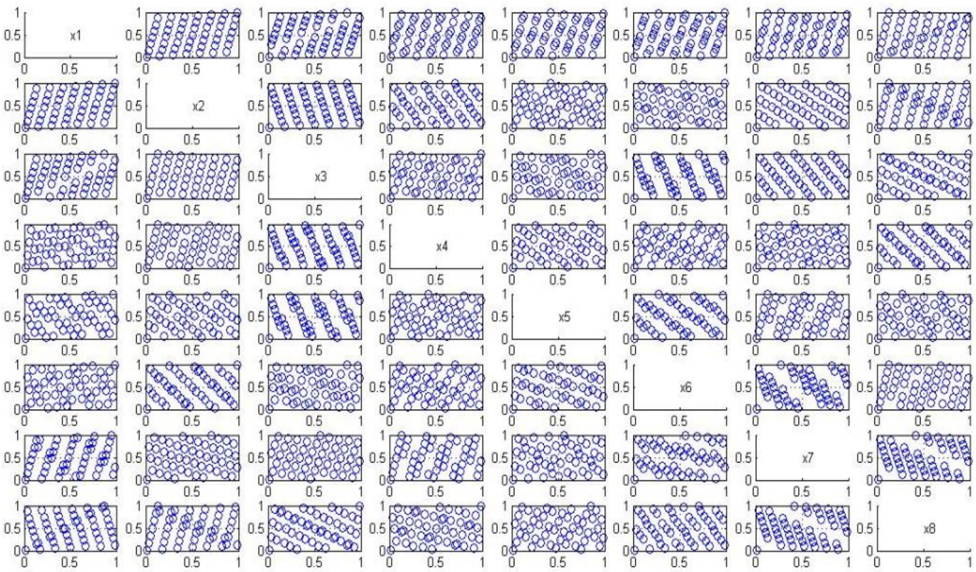

Figure 1. Projection Properties of OA (49, 8) LHD

The OA (49, 8) LHD given in Table 1 contained 49 runs and 8 input variables and was scaled according to the assumed range for design variables using Equation 1:

$$y_{OALHD} = \frac{y_{data} - y_{data(min)}}{y_{data(max)} - y_{data(min)}} \tag{1}$$

$$y_{data} = y_{OALHD}\left(y_{data(max)} - y_{data(min)}\right) + y_{data(min)}$$

Table 2. Input Variables for Borehole Model

Variable	*Variable name*	*Minimum*	*Maximum*
X_1 (r_w)	Radius of Borehole (metre)	0.05	0.1
X_2 (r)	Radius of Influence (metre)	100	25050
X_3 (T_u)	Transmissivity of Upper Aquifer (m^2/yr)	63070	89335
X_4(H_u)	Potentiometric Head of Upper Aquifer (metre)	990	1045
X_5(T_l)	Transmissivity of Lower Aquifer (m^2/yr)	63.1	89.55
X_6(H_l)	Potentiometric Head of Lower Aquifer(metre)	700	760
X_7(L)	Length of Borehole (metre)	1120	1400
X_8(K_w)	Hydraulic Conductivity of Borehole (metre/yr)	9855	10950
Y	Flow Rate of Water (m^3/yr)		

The scaled OALHD was used to develop a borehole computer experiment using the simulator in Equation 2:

$$y = \frac{2\pi T_u (H_u - H_l)}{1n(r/r_w)\left[1 + \frac{2LT_u}{1n(r/r_w)r_w^2 k_w} + \frac{T_u}{T_l}\right]} \quad (2)$$

where

r_w(m) = radius of borehole
r(m) = radius of influence
T_l(m^2/yr) = transmissivity of lower aquifer
T_u(m^2/yr) = transmissivity of upper aquifer
H_l (m) = potentiometric head of lower aquifer
H_u(m) = potentiometric head of upper aquifer
L (m) = length of borehole and
K_w(m/yr) = hydraulic conductivity of borehole

The scale input variables and the output from a borehole computer model constitute the experimental results for the training data sets as provided in Table 3.

Table 3. Experimental Data for Borehole Computer Experiment (Training Data Sets)

r_w	R	T_u	H_u	T_l	H_l	L	K_w	Y
0.055	1397.057	63922.849	991.247	63.556	700.833	1123.256	9957.012	24.111
0.056	2113.633	64202.654	991.357	63.561	700.834	1123.256	10086.946	25.194
0.056	2830.210	64482.458	991.468	63.557	700.834	1123.256	10216.880	26.309
0.057	3546.786	64762.263	991.320	63.562	700.833	1123.256	10346.814	27.433
0.058	4263.362	64062.752	991.431	63.558	700.834	1123.256	10476.748	28.613
0.059	4979.939	64342.556	991.284	63.563	700.834	1123.256	10606.682	29.801
0.060	5696.515	64622.361	991.394	63.560	700.833	1123.256	10736.616	31.047
0.061	2216.001	64082.738	991.289	63.557	700.833	1123.256	9975.574	29.671
0.061	2932.578	64362.542	991.399	63.562	700.834	1123.256	10105.508	30.911
0.062	3649.154	64642.347	991.252	63.559	700.834	1123.256	10235.442	32.156
0.063	4365.731	63942.835	991.362	63.563	700.834	1123.256	10365.376	33.464
0.064	5082.307	64222.640	991.473	63.560	700.833	1123.256	10495.310	34.806
0.065	5798.883	64502.445	991.326	63.556	700.834	1123.256	10625.244	36.152
0.066	1499.425	64782.249	991.436	63.561	700.834	1123.256	10755.178	37.583
0.067	3034.946	64242.626	991.331	63.559	700.834	1123.256	9994.136	35.816
0.067	3751.522	64522.431	991.441	63.563	700.833	1123.256	10124.070	37.219
0.068	4468.099	64802.235	991.294	63.560	700.834	1123.256	10254.004	38.624

0.069	5184.675	64102.724	991.405	63.556	700.834	1123.256	10383.938	40.099
0.070	5901.251	64382.528	991.257	63.561	700.833	1123.256	10513.872	41.575
0.071	1601.793	64662.333	991.368	63.558	700.834	1123.256	10643.806	43.146
0.072	2318.369	63962.821	991.478	63.562	700.834	1123.256	10773.740	44.730
0.073	3853.890	64402.514	991.373	63.560	700.834	1123.256	10012.698	42.546
0.073	4570.467	64682.319	991.483	63.557	700.833	1123.256	10142.632	44.119
0.074	5287.043	63982.807	991.336	63.561	700.834	1123.256	10272.566	45.691
0.075	6003.619	64262.612	991.447	63.558	700.834	1123.256	10402.500	47.341
0.076	1704.161	64542.417	991.299	63.562	700.834	1123.256	10532.434	49.016
0.077	2420.737	64822.221	991.410	63.559	700.833	1123.256	10662.368	50.740
0.078	3137.314	64122.710	991.263	63.564	700.834	1123.256	10792.302	52.459
0.078	4672.835	64562.403	991.415	63.561	700.834	1123.256	10031.260	49.861
0.079	5389.411	64842.207	991.268	63.558	700.834	1123.256	10161.194	51.567
0.080	6105.988	64142.696	991.378	63.563	700.833	1123.256	10291.128	53.358
0.081	1806.529	64422.500	991.489	63.559	700.834	1123.256	10421.062	55.225
0.082	2523.106	64702.305	991.341	63.564	700.834	1123.256	10550.996	57.044
0.083	3239.682	64002.793	991.452	63.560	700.833	1123.256	10680.930	58.955
0.084	3956.258	64282.598	991.305	63.557	700.834	1123.256	10810.864	60.855
0.084	5491.779	64722.291	991.457	63.563	700.834	1123.256	10049.822	57.763
0.085	6208.356	64022.779	991.310	63.559	700.834	1123.256	10179.756	59.648
0.086	1908.897	64302.584	991.420	63.564	700.833	1123.256	10309.690	61.668
0.087	2625.474	64582.389	991.273	63.560	700.834	1123.256	10439.624	63.628
0.088	3342.050	64862.193	991.383	63.557	700.834	1123.256	10569.558	65.687
0.089	4058.626	64162.682	991.494	63.562	700.834	1123.256	10699.492	67.791
0.089	4775.203	64442.486	991.347	63.558	700.833	1123.256	10829.426	69.879
0.090	6310.724	64882.179	991.499	63.564	700.834	1123.256	10068.384	66.253
0.091	2011.265	64182.668	991.352	63.561	700.834	1123.256	10198.318	68.372
0.092	2727.842	64462.472	991.462	63.557	700.834	1123.256	10328.252	70.538
0.093	3444.418	64742.277	991.315	63.562	700.833	1123.256	10458.186	72.687
0.094	4160.994	64042.765	991.426	63.558	700.834	1123.256	10588.120	74.945
0.095	4877.571	64322.570	991.278	63.563	700.834	1123.256	10718.054	77.181
0.095	5594.147	64602.375	991.389	63.559	700.833	1123.256	10847.988	79.530

3. Modelling Borehole Computer Experiment

A borehole computer model is a model of a physical process to be emulated using a metamodel. A metamodel is referred to as a model used to emulate a borehole computer model. A wide variety of techniques have been discussed in the literature for creating the metamodels [14]. These techniques include response surface modelling [9], Radial Basis Functions ([2]; [3]), Multivariate Adaptive Regression Splines [4] and Support Vector Machine [15]. In this study, a Gaussian stochastic process (Gasp) model was investigated as an alternative technique for approximating a borehole computer model. The Gasp model allows a wide range of correlation functions $R(X,X')$ to be used. A Gaussian correlation function was chosen in this study. The output of the borehole computer experimental data was modelled using a Gaussian stochastic process model as described below:

$$Y(x) = \sum_{j=1}^{k} \beta_j f_j(x) + Z(x) \tag{3}$$

$$\beta_j f_j(x) = \beta_1 f_1(x) + \beta_2 f_2(x) + \cdots \beta_k f_k(x)$$
$$= [f_1(x) + f_2(x) + \cdots f_k(x)]\beta_j$$

$$= f(x)^T \beta_j \tag{4}$$

where $f_1(x), \ldots, f_k(x)$ are k known regression functions and β_1, …, β_k are their corresponding (unknown) parameters and $Z(x)$ is a stochastic process which is assumed to have mean zero and variance-covariance structure

$$Cov\left(Z(x), Z(x')\right) = \sigma_z^2 R(X, X') \tag{5}$$

σ_z^2 is the process variance and $R(X, X')$ is the Gaussian correlation function that can be tuned to the data . The Gaussian correlation function is given as

$$R(X, X') = \prod_{j=1}^{d} \exp(-\theta_j \mid x_j - x_j' \mid^2) \tag{6}$$

where $\theta_j \geq 0$. The parameter θ is important in the correlation structure of Z. When θ is large there is a small correlation between observations and therefore prediction is more difficult whereas there is a large correlation between observations and prediction is much simpler when θ is small. The selection of the correlation function is very useful in the prediction process. The author in [6] discussed the effects of θ on the prediction of output of a computer experiment. The correlation matrix, R is given as an (*n x n*) matrix given in Equation 7:

$$\boldsymbol{R} = \begin{bmatrix} R(x_1, x_1) & \cdots & R(x_1, x_n) \\ \vdots & \ddots & \vdots \\ R(x_n, x_1) & \cdots & R(x_n, x_n) \end{bmatrix} \tag{7}$$

The matrix R is symmetric since $R(x_i, x_j) = R(x_j, x_i)$ and the diagonal consists of all ones because $R(x_i, x_i) = 1$. The correlation between an unknown point x and the n known sample points is given by the vector:

$$r_x = \left[R(x_1, x), \ldots, R(x_n, x)\right]^T \tag{8}$$

The best Linear Unbiased Predictor (BLUP) is obtained by minimizing the mean square error of the predictions. The BLUP at an untried point x is therefore given as:

$$\hat{y}(x) = f_x^T \hat{\beta} + r^T(x) R^{-1}(y - F\hat{\beta}) \tag{9}$$

where F is the expanded design matrix *n x k* given by

$$F = \begin{pmatrix} f^T(x_1) \\ \ldots\ldots \\ f^T(x_n) \end{pmatrix} \tag{10}$$

$$f(x) = \left[f_1(x), \ldots, f_k(x)\right]^T \tag{11}$$

The Maximum Likelihood Estimation (MLE) method was used to estimate the Gasp model parameters (β, θ and σ) and it is an objective estimator that is most consistent with the observed data. The MLE assumes the residuals have a known probability distribution shape, that is, the Gaussian probability distribution. The correlation parameter θ was found using mlegp package in R software. The MLE estimation of β equals its least-squares estimate and is given by

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} y \tag{12}$$

and the MLE of the process variance is also given by

$$\hat{\sigma}_z^2 = \frac{1}{n}\left(y - F\hat{\beta}\right)^T R^{-1}\left(y - F\hat{\beta}\right) \tag{13}$$

4. Analysis of Borehole Computer Experiment

The training datasets given in Table 3 showed that the eight input variables involved in the borehole computer experiment were of different scales. These variables were normalized by subtracting their means and multiplying by the reciprocal of their standard deviations before the analysis. This is required to lessen the dimension effect of each design variable and avert the Gasp model from being inconsistent in prediction. The normalized experimental data based on the borehole computer experiment and the estimated results are given in Table 4 and Table 5, respectively.

Table 4. Normalized Experimental Data for Borehole Computer Experiment

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	Y
-1.680	-1.680	-1.680	-1.680	-1.680	-1.680	-1.680	-1.680	-1.565
-1.610	-1.190	-0.700	-0.210	0.280	0.770	1.260	-1.190	-1.495
-1.540	-0.700	0.280	1.260	-1.190	-0.210	0.770	-0.700	-1.424
-1.470	-0.210	1.260	-0.700	0.770	-1.190	0.280	-0.210	-1.353
-1.400	0.280	-1.190	0.770	-0.700	1.260	-0.210	0.280	-1.277
-1.330	0.770	-0.210	-1.190	1.260	0.280	-0.700	0.770	-1.201
-1.260	1.260	0.770	0.280	-0.210	-0.700	-1.190	1.260	-1.122
-1.190	-1.120	-1.120	-1.120	-1.120	-1.120	-1.120	-1.610	-1.210
-1.120	-0.630	-0.140	0.350	0.840	1.330	-1.610	-1.120	-1.131
-1.050	-0.140	0.840	-1.610	-0.630	0.350	1.330	-0.630	-1.051
-0.980	0.350	-1.610	-0.140	1.330	-0.630	0.840	-0.140	-0.968
-0.910	0.840	-0.630	1.330	-0.140	-1.610	0.350	0.350	-0.882
-0.840	1.330	0.350	-0.630	-1.610	0.840	-0.140	0.840	-0.796
-0.770	-1.610	1.330	0.840	0.350	-0.140	-0.630	1.330	-0.705
-0.700	-0.560	-0.560	-0.560	-0.560	-0.560	-0.560	-1.540	-0.818
-0.630	-0.070	0.420	0.910	1.400	-1.540	-1.050	-1.050	-0.728
-0.560	0.420	1.400	-1.050	-0.070	0.910	-1.540	-0.560	-0.638
-0.490	0.910	-1.050	0.420	-1.540	-0.070	1.400	-0.070	-0.544
-0.420	1.400	-0.070	-1.540	0.420	-1.050	0.910	0.420	-0.450
-0.350	-1.540	0.910	-0.070	-1.050	1.400	0.420	0.910	-0.350
-0.280	-1.050	-1.540	1.400	0.910	0.420	-0.070	1.400	-0.249
-0.210	0.000	0.000	0.000	0.000	0.000	0.000	-1.470	-0.388
-0.140	0.490	0.980	1.470	-1.470	-0.980	-0.490	-0.980	-0.288
-0.070	0.980	-1.470	-0.490	0.490	1.470	-0.980	-0.490	-0.187
0.000	1.470	-0.490	0.980	-0.980	0.490	-1.470	0.000	-0.082
0.070	-1.470	0.490	-0.980	0.980	-0.490	1.470	0.490	0.025
0.140	-0.980	1.470	0.490	-0.490	-1.470	0.980	0.980	0.135
0.210	-0.490	-0.980	-1.470	1.470	0.980	0.490	1.470	0.244
0.280	0.560	0.560	0.560	0.560	0.560	0.560	-1.400	0.079
0.350	1.050	1.540	-1.400	-0.910	-0.420	0.070	-0.910	0.188
0.420	1.540	-0.910	0.070	1.050	-1.400	-0.420	-0.420	0.302
0.490	-1.400	0.070	1.540	-0.420	1.050	-0.910	0.070	0.421

0.560	-0.910	1.050	-0.420	1.540	0.070	-1.400	0.560	0.537
0.630	-0.420	-1.400	1.050	0.070	-0.910	1.540	1.050	0.659
0.700	0.070	-0.420	-0.910	-1.400	1.540	1.050	1.540	0.780
0.770	1.120	1.120	1.120	1.120	1.120	1.120	-1.330	0.583
0.840	1.610	-1.330	-0.840	-0.350	0.140	0.630	-0.840	0.703
0.910	-1.330	-0.350	0.630	1.610	-0.840	0.140	-0.350	0.832
0.980	-0.840	0.630	-1.330	0.140	1.610	-0.350	0.140	0.957
1.050	-0.350	1.610	0.140	-1.330	0.630	-0.840	0.630	1.089
1.120	0.140	-0.840	1.610	0.630	-0.350	-1.330	1.120	1.223
1.190	0.630	0.140	-0.350	-0.840	-1.330	1.610	1.610	1.356
1.260	1.680	1.680	1.680	1.680	1.680	1.680	-1.260	1.125
1.330	-1.260	-0.770	-0.280	0.210	0.700	1.190	-0.770	1.260
1.400	-0.770	0.210	1.190	-1.260	-0.280	0.700	-0.280	1.398
1.470	-0.280	1.190	-0.770	0.700	-1.260	0.210	0.210	1.535
1.540	0.210	-1.260	0.700	-0.770	1.190	-0.280	0.700	1.679
1.610	0.700	-0.280	-1.260	1.190	0.210	-0.770	1.190	1.822
1.680	1.190	0.700	0.210	-0.280	-0.770	-1.260	1.680	1.972

Table 5. Results of the Estimated Model for Borehole Computer Experiment

Model Parameters	Estimated Values	P-values
$\hat{\beta}_0$	0.8150	
$\hat{\beta}_1$	0.9425	<0.001
$\hat{\beta}_2$	-4.6449e-04	0.9692
$\hat{\beta}_3$	-7.9773e-06	0.9401
$\hat{\beta}_4$	8.3103e-04	0.1442
$\hat{\beta}_5$	-1.1769e-05	0.9841
$\hat{\beta}_6$	-1.5467e-04	0.9475
$\hat{\beta}_7$	1.0115e-05	0.9281
$\hat{\beta}_8$	0.0727	0.0540
$\hat{\sigma}_z^2$	0.1018	

The estimated linear main effects $\hat{\beta}_i$, $i = 1, 2,3,4,5,6,7$ and 8 (corresponding to r_w , r, T_u, H_u,T_l, H_l, L *and* K_w, respectively) of the borehole computer model were presented in Table 5 with their respective *p-values* for the t-test for $i = 1,\ldots, 8$ and $\hat{\sigma}_z^2$. The linear main effects for r_w (0.9425) and K_w (0.0727) are relatively large while their respective *p-values* <0.001 and 0.0540 are quite small. This implies that, the radius of borehole (r_w) is highly significant and its hydraulic conductivity (K_w) is also significant with *p-value* = 0.0540. The estimated values for the radius (r_w) and the hydraulic conductivity of borehole (K_w) have the same signs, indicating that r_w and K_w have the same effects on the flow rate of water.

For the sake of prediction, additional simulations were performed to verify the accuracy of the Gasp model since the goodness of fit obtained from the training datasets may not be sufficient to assess the accuracy of newly predicted points. The additional simulations constitute the test datasets. The assumed range for the test data is given in Table 6 and the normalized experimental data is also given in Table 7.

Table 6. Input and Output Variables for Borehole Model (Test data)

Variable	Variable Name	Minimum	Maximum
X_1	Radius of Borehole (metre)	0.11	0.15
X_2	Radius of Influence (metre)	25051	50e3
X_3	Transmissivity of Upper Aquifer (m^2/yr.)	89336	115600
X_4	Potentiometric Head of Upper Aquifer (metre)	1046	1100
X_5	Transmissivity of Lower Aquifer (m^2/yr.)	89.56	116
X_6	Potentiometric Head of Lower Aquifer(metre)	761	820
X_7	Length of Borehole (metre)	1401	1680
X_8	Hydraulic Conductivity of Borehole (metre/yr.)	10951	12045
Y	Flow Rate of Water (m^3/yr.)	-	-

Table 7. Normalized Experimental Data for Borehole Computer Experiment

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
-1.680	-1.680	-1.680	-1.680	-1.680	-1.680	-1.680	-1.680	-1.725
-1.610	-1.190	-0.700	-0.210	0.280	0.770	1.260	-1.190	-1.611
-1.540	-0.700	0.280	1.260	-1.190	-0.210	0.770	-0.700	-1.496
-1.470	-0.210	1.260	-0.700	0.770	-1.190	0.280	-0.210	-1.383
-1.400	0.280	-1.190	0.770	-0.700	1.260	-0.210	0.280	-1.265
-1.330	0.770	-0.210	-1.190	1.260	0.280	-0.700	0.770	-1.149
-1.260	1.260	0.770	0.280	-0.210	-0.700	-1.190	1.260	-1.027
-1.190	-1.120	-1.120	-1.120	-1.120	-1.120	-1.120	-1.610	-1.316
-1.120	-0.630	-0.140	0.350	0.840	1.330	-1.610	-1.120	-1.195
-1.050	-0.140	0.840	-1.610	-0.630	0.350	1.330	-0.630	-1.077
-0.980	0.350	-1.610	-0.140	1.330	-0.630	0.840	-0.140	-0.953
-0.910	0.840	-0.630	1.330	-0.140	-1.610	0.350	0.350	-0.827
-0.840	1.330	0.350	-0.630	-1.610	0.840	-0.140	0.840	-0.704
-0.770	-1.610	1.330	0.840	0.350	-0.140	-0.630	1.330	-0.574
-0.700	-0.560	-0.560	-0.560	-0.560	-0.560	-0.560	-1.540	-0.889
-0.630	-0.070	0.420	0.910	1.400	-1.540	-1.050	-1.050	-0.761
-0.560	0.420	1.400	-1.050	-0.070	0.910	-1.540	-0.560	-0.636
-0.490	0.910	-1.050	0.420	-1.540	-0.070	1.400	-0.070	-0.504
-0.420	1.400	-0.070	-1.540	0.420	-1.050	0.910	0.420	-0.376
-0.350	-1.540	0.910	-0.070	-1.050	1.400	0.420	0.910	-0.240
-0.280	-1.050	-1.540	1.400	0.910	0.420	-0.070	1.400	-0.103
-0.210	0.000	0.000	0.000	0.000	0.000	0.000	-1.470	-0.445
-0.140	0.490	0.980	1.470	-1.470	-0.980	-0.490	-0.980	-0.310
-0.070	0.980	-1.470	-0.490	0.490	1.470	-0.980	-0.490	-0.177
0.000	1.470	-0.490	0.980	-0.980	0.490	-1.470	0.000	-0.038
0.070	-1.470	0.490	-0.980	0.980	-0.490	1.470	0.490	0.098
0.140	-0.980	1.470	0.490	-0.490	-1.470	0.980	0.980	0.242
0.210	-0.490	-0.980	-1.470	1.470	0.980	0.490	1.470	0.381
0.280	0.560	0.560	0.560	0.560	0.560	0.560	-1.400	0.016
0.350	1.050	1.540	-1.400	-0.910	-0.420	0.070	-0.910	0.154
0.420	1.540	-0.910	0.070	1.050	-1.400	-0.420	-0.420	0.299
0.490	-1.400	0.070	1.540	-0.420	1.050	-0.910	0.070	0.447
0.560	-0.910	1.050	-0.420	1.540	0.070	-1.400	0.560	0.590
0.630	-0.420	-1.400	1.050	0.070	-0.910	1.540	1.050	0.741
0.700	0.070	-0.420	-0.910	-1.400	1.540	1.050	1.540	0.888
0.770	1.120	1.120	1.120	1.120	1.120	1.120	-1.330	0.495
0.840	1.610	-1.330	-0.840	-0.350	0.140	0.630	-0.840	0.640
0.910	-1.330	-0.350	0.630	1.610	-0.840	0.140	-0.350	0.794
0.980	-0.840	0.630	-1.330	0.140	1.610	-0.350	0.140	0.943
1.050	-0.350	1.610	0.140	-1.330	0.630	-0.840	0.630	1.100
1.120	0.140	-0.840	1.610	0.630	-0.350	-1.330	1.120	1.259

1.190	0.630	0.140	-0.350	-0.840	-1.330	1.610	1.610	1.414
1.260	1.680	1.680	1.680	1.680	1.680	1.680	-1.260	0.991
1.330	-1.260	-0.770	-0.280	0.210	0.700	1.190	-0.770	1.145
1.400	-0.770	0.210	1.190	-1.260	-0.280	0.700	-0.280	1.306
1.470	-0.280	1.190	-0.770	0.700	-1.260	0.210	0.210	1.463
1.540	0.210	-1.260	0.700	-0.770	1.190	-0.280	0.700	1.628
1.610	0.700	-0.280	-1.260	1.190	0.210	-0.770	1.190	1.788
1.680	1.190	0.700	0.210	-0.280	-0.770	-1.260	1.680	1.958

Based on the data in Table 7, the predicted flow rate of water, y is calculated using Equation 9 and the Gasp model interpolates the test data, that is, the predicted values at untried inputs gave the same results as the simulated values. This quality makes Gasp model approximately an exact interpolator. The graph of the predicted versus simulated output of the borehole model is given in Figure 2 as shown below:

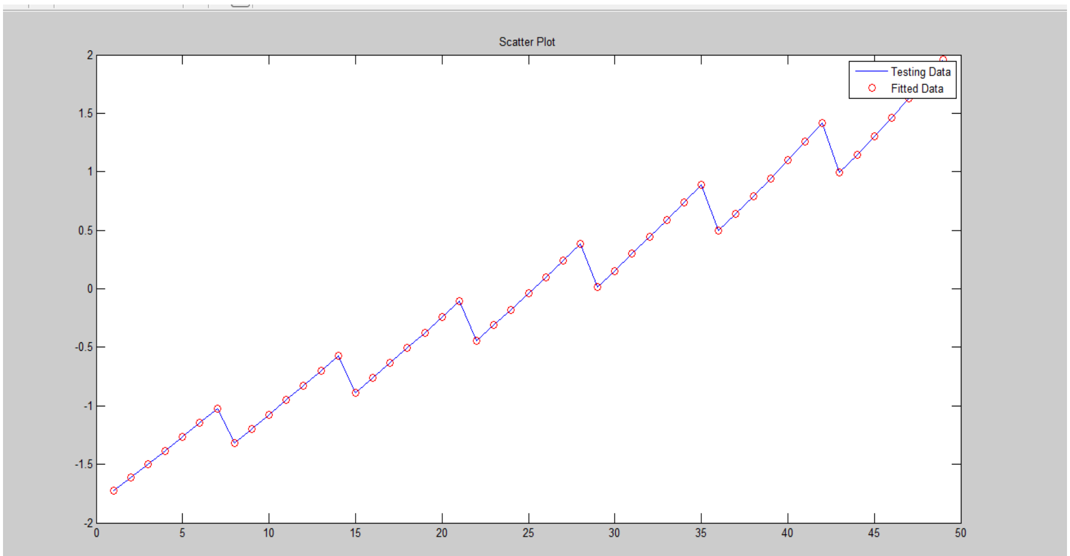

Figure 2. Graph of the Predicted y (Flow Rate of Water- Dotted Points) against Simulated Output (Solid Blue Line) over 49 Experimental Runs

5. Discussion of Results

The constructed OA(49, 8) LHD was used to develop a borehole computer experiment by scaling OA(49, 8) LHD and then simulate the experimental output using Equations 1 and 2, respectively. The scaled input variables and the output simulated from a computer model form the experimental results for the training datasets. The training datasets were used to fit a Gasp model. The fitted Gasp model showed that the radius (r_w) and Hydraulic Conductivity of Borehole (K_w) are important variables in modelling borehole experiment as given in Table 5.

The test data was also simulated using a 49-run experimental design in order to assess the accuracy of the Gasp model. These test data were used on the fitted Gasp model to predict the flow rate of water of a borehole. The predicted results interpolated the test data as shown in Figure 2. This, therefore, shows that the fitted Gasp model is very efficient.

6. Conclusion

A method for the statistical modelling and analysis of borehole computer experiment has been presented in this work. It is an exposition of the current trend in the design and analysis of experiments where a computer model was used to mimic a borehole experiment and results were obtained within the shortest possible time without having to wait for physical experimental results. This is simply called a computer experimentation approach.

The fitted Gasp model emulated the borehole computer model perfectly well and it also interpolated the test data. This work showed that the radius (r_w) and hydraulic conductivity of borehole (K_w) are important variables in modeling and predicting the flow rate of water of a borehole using a 49-run experimental design and Gasp model as an emulator of a borehole computer model. The estimated values for the radius (r_w) and the hydraulic conductivity of borehole (K_w) have the same signs, indicating that r_w and K_w have the same effects on the flow rate of water of a borehole. This study also showed that the *p-values* for the estimated parameters for r(m), T_l (m^2/yr), T_u (m^2/yr), H_l (m), H_u (m) and L (m) are very large as shown in Table 5. These six factors do not have significant linear main effects on the flow rate of water through the borehole.

The computer experimentation approach has opened up a new area in Design and Analysis of Experiments which is referred to as Design and Analysis of Computer Experiments. This area of study has permitted both greater complexity and more extensive use of mathematical models as computer models in scientific and engineering experimentations as well as in industrial processes. A Gasp model with correlation functions different from the Gaussian correlation function employed in this work could also be considered in order to compare the performance of Gasp model under different correlation functions and predict the output of the borehole computer experiment at untried inputs for future research.

References

[1] J. An and A. B. Owen, "Quasi-regression", Journal of Complexity, vol. 17, **(2001)**, pp. 588–607.

[2] N. Dyn, D. Levin and S. Rippa, "Numerical Procedures for Surface Fitting of Scattered Data by Radial Basis Functions", SIAM Journal of Scientific and Statistical Computing, vol. 7, no. 2, **(1986)**, 639-659.

[3] H. Fang and M. F. Horstemeyer, "Global Response Approximation with Radial Basis Functions", Journal of Engineering Optimization, vol. 38, no. 4, **(2006)**, pp. 407-424.

[4] J. H. Friedman, "Multivariate Adaptive Regressive Splines", The Annals of Statistics, vol. 19, no. 1, **(1991)**, pp. 1-67.

[5] W. M. Ho and Z.Q. Xu, "Applications of Uniform Design to Computer Experiments", Journal of Chinese Statistical Association, vol. 38, **(2000)**, pp. 395–410.

[6] J. R. Koehler and A. B. Owen, "Computer experiments", Handbook of Statistics, vol. 13, **(1996)**, pp. 261–308.

[7] M. D. Morris, T. J. Mitchell and D.Ylvisaker, "Bayesian Design and Analysis of Computer Experiments: Use of Derivatives in Surface Prediction", Technometrics, *vol.* 35, **(1993)**, pp. 243–255.

[8] D.C. Montgomery, "Design and Analysis of Experiments", 5th Edition, John Willey and Sons, Inco., New York, **(2001)**.

[9] R. H. Myers and D. C. Montgomery, "Response Surface Methodology: Process and Product Optimization using Designed Experiments", Wiley Series in Probability and Statistics, John Wiley and Sons, New York, **(1995)**.

[10] K.A. Osuolale , W.B. Yahya and B.L. Adeleke, "Construction of Space-filling Designs for Three Input Variables Computer Experiments", International Journal of Computer, Control, Quantum and Information Engineering, WASET, vol. 8, no. 9, **(2014)**, pp. 1624-1628.

[11] K.A. Osuolale and W.B. Yahya, "Statistical Design and Modelling of Borehole Computer Experiment", Proceedings of the 3rd U6 Consortium International Conference, Ilorin, Nigeria, (2015a) 6th September – 9th September.

[12] K.A. Osuolale, W.B. Yahya and B.L. Adeleke, "Construction of Orthogonal Array-Based Latin Hypercube Design for Deterministic Computer Experiments", Annals, Computer Science Series, vol. 13, no. 1, **(2015b)**, pp. 24-29.

[13] T. W. Simpson, T. M. Mauery, J.J. Korte and F.Mistree, "Kriging Models for Global Approximation in Simulation-Based Multidisciplinary Design Optimization", AIAA Journal, vol. 39, **(2001)**, pp. 2233-2241.

[14] S. Strogatz, "The Real Scientific Hero of 1953", New York Times, **(2003)**.

[15] G.G. Wang and S. Shan, "Review of Metamodeling Techniques in Support of Engineering Design Optimization" ASME Transactions, Journal of Mechanical Design, in press, **(2006)**.

[16] B.A. Worley, "Deterministic Uncertainty Analysis, *in* 'ORNL—0628" National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161, USA, **(1987)**.