

## RESPONSE SURFACE METHODOLOGY VIA DESIRABILITY FUNCTION TECHNIQUES FOR OPTIMIZING CORRELATED RESPONSES OF ELECTRICAL CONDUCTIVITY AND TOTAL DISSOLVED SOLIDS OF SELECTED BOREHOLE WATER

\*<sup>1</sup>Yusuff K. M., <sup>2</sup>Lawal M., <sup>3</sup>Audu A. T. and <sup>1</sup>Wale-Orojo O. A.

<sup>1</sup>Department of Statistics, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria

<sup>2</sup>Department of Mathematical and Computer Science, Fountain University, Osogbo, Osun State, Nigeria

<sup>3</sup>Department of Statistics, Federal School of Statistics, Ibadan, Oyo State, Nigeria

Corresponding Author's email: [kehindemide@yahoo.com](mailto:kehindemide@yahoo.com)

### ABSTRACT

The health benefits in the description and observation of quantitative contents of quality parameters present or contained in any water source cannot be underestimated as they determine selection of best choice from available water sources for different intended uses as well as resource consumption. It also helps to compare the observed quantity of the quality with the acceptable standards or limits to get desired results. Physical parameters like pH, temperature, electrical conductivity (EC) and total dissolved solids (TDS) among others are determined by present of other chemical properties like Cations ( $Mg^{2+}$ ,  $Ca^{2+}$ ,  $Na^+$ , etc), Anions ( $Cl^-$ ,  $NO_3^-$ ,  $SO_4^{2-}$ , etc), heavy metals and other dissolved materials during the course of its formation in different proportions and amounts. This study observed EC and TDS of 20 selected boreholes as two close and correlated water quality parameters as well as two of the major water quality parameters that account for overall quality of any water source, despite their different quantitative contents and physical features, they are likely determined by the same set of cations and anions with similar constraint equations. In contrast to linear programming, multiple criteria optimization models were fitted for EC and TDS using Response Surface Methodology via desirability techniques, optimal values obtained in this case measured against several criteria are found to lie between acceptable standards limits for drinking water, other numerical values and descriptive features in the final results reflect that the response equations obtained were well fitted.

**Keywords:** Desirability Function, Electrical Conductivity, Optimal value, Response Surface Methodology, Total Dissolved Solids, Water Quality Parameters.

### INTRODUCTION

Product optimization is an essential process in the fields of Science, Engineering and Technology as optimization of yields and productivities has been a major goal in biological and physical components since the very beginning of decision processes on industrial and production systems. Optimization is a technique that explores possible and likely behaviors of systems with numerous input responses with goal of identifying the best possible outcome, Cornell (1996). In mathematical terms, the “outcome” is the value of some functions, and “best possible” often is the maximum or minimum of the function which is the point with highest or lowest possible value. The function itself is often called objective function and its arguments are called control variables or at times input variables observed from various input responses that determine or account for quality of interest, Cornell (1996). In practices, if one multiplies the

objective function by -1, the former maximum becomes minimum and vice versa. Thus, finding a maximum or a minimum is basically the same, and we talk generally about finding extremum or optimum. Raymond, *et al.* (2016) affirmed that the objective function is controlled by some conditions of input factors or variable as defined by set of constraint functions that define boundaries or limits that guarantee each response viability or acceptable standards in the system. Thus, the constraints are dictated by system component, technical and economical factor among others to have final equilibrium output from objective function. A complete optimization to optimize a response of interest involves solving response of interest involved and validating of objective and constraint equations. Summarily, the optimization task with several system components typically reads;

$$\begin{array}{l}
 \text{Maximize:} \quad f(X_1, X_2, \dots, X_h) \\
 \text{Subject to:} \quad f(X_1) \leq \text{Constant} \\
 \quad \quad \quad f(X_1, X_2) \geq \text{Constant} \\
 \quad \quad \quad f(X_1, X_2, X_3) = \text{Constant} \\
 \quad \quad \quad X_1 > 0, X_2 > 0, \dots, X_h > 0
 \end{array} \quad (1)$$

### Multi-response Design and Model

Multi-response experiment requires careful consideration of the multivariate nature of data observed together in a process. In fitting model for multi-response observation, Khuri and Cornell (1996) stated that response variables should not be investigated individually and independently of one another as interrelationships that may exist among them can render such investigation meaningless. Hill and Hunter (1966) cited several papers in which multiple response are investigated with desire to optimize several response functions simultaneously with a completely defined constraint equations with full information about the condition of constraints optimal solution has to satisfy.

$$y_{ui} = f(X_u, \beta) + \varepsilon_{ui} \quad u = 1, 2, \dots, N$$

$$i = 1, 2, \dots, r \quad (2)$$

Where  $X_j$  is the vector  $(X_{j1}, X_{j2}, \dots, X_{jk})$  being  $u^{th}$  the level of the  $j^{th}$  coded variables ( $u = 1, 2, \dots, N; j = 1, 2, \dots, r$ ),  $\beta$  is a vector of unknown parameters,  $\varepsilon_{ui}$  is a random error, and  $f$  is a function of known form for the  $i^{th}$  response and is assumed to be continuous. Similarly, response variables can be represented by polynomial regression models in the values of  $x_i$  within a certain region of  $\mathfrak{R}$ . Hence, the  $i^{th}$  response model can be written in vector form as

$$Y_i = X_i \beta_i + \varepsilon_i \quad i = 1, 2, \dots, r \quad (3)$$

Where  $Y_i$  is an  $N \times 1$  vector of observations on the  $i^{th}$  response,  $X_i$  is an  $N \times P$  matrix of rank  $P$  of known function of the setting of coded variables,  $\beta$  is a  $P \times 1$  vector of unknown constant parameters and  $\varepsilon_i$  is a random error vector associated with  $i^{th}$  response ( $i = 1, 2, \dots, r$ ). The assumptions on  $\varepsilon_i$  are that

$$E(\varepsilon_i) = 0$$

$$V(\varepsilon_i) = \sigma_{ii} I_N \quad i = 1, 2, \dots, r$$

$$Cov(\varepsilon_i \varepsilon_j) = \sigma_{ij} I_N \quad i, j = 1, 2, \dots, r \quad i \neq j \quad (4)$$

The  $r \times r$  matrix whose  $(i, j)^{th}$  element is  $\sigma_{ij}$ , ( $i, j = 1, 2, \dots, r$ ) is denoted by  $\Sigma$ .

Also, the  $r$  equations in 3 can be represented by

$$Y = X \beta + \varepsilon \quad (5)$$

### Water Quality Parameters As A Multi-response Experiment

As opposite to experiments with single response variables which are referred to as single-response experiments. However, numerous experiments involve measurements associated with several response variables, in such cases number of responses are measured simultaneously for each setting of group of input variables which are referred to as multi-response experiment by Khuri and Cornell (1996). There are numerous number of multiresponse experiments where researchers' interest is to determine optimum combinations of various system components on the basis of acceptability, nutritional and economic value among other considerations. Like many other systems or processes that quality performance are determined by numerical contents or compositions of different or specific components of that system or process in which no single quality parameter can perform better in isolation.

Water is natural substance that during the process of its

### Representation of General Multi-response Model

In (Cornell (1996), Way Kuo, *et al.* (2001)), it was supported that the design problem in the multi-response case is more complex than in the case of a single response as each of the response values has different set of input variables, at times when set of input variables are the same for some responses, the linear relationship or contrast may differ. In a system process of different components, the constraint conditions on input variables are the same irrespective of response value. In matrix form,  $i^{th}$  response value at  $u^{th}$  experimental run is represented by

formation, its contents, components and quality is determined by different organic and inorganic matters that have contact with during and after its formation e.g. different types of soluble and insoluble rocks, soils, atmospheric and biological matters. Different sources of water account for unequal compositions and contents of quality parameters and as a result, there are needs for its different users to make decisions on selected responses and factors of choice to meet their desired limits of acceptance for optimal condition subject to different response requirement with respect to their specification limits as constraints and to analyze the feasibility of optimal values to give desired product quality.

### An Overview of Multi-Response Optimization Approaches

In the work of Raissi and Eslami (2009), this process is regarded as a multi-response optimization problem. Most of the common methods are incomplete in such a way that a response variable is selected as the primary one and is optimized by adhering to the other constraints set by the criteria. Among

methodologies developed to resolve the multi-response problems. Khuri and Cornell (1987) surveyed the multi-response problem using a response surface method. Tai, *et al.* (1992) assigned a weight for each response to resolve the problem. Pignatiello (1993) made use of a squared deviation-from-target and a variance to and from an expected loss function. Layne (1995) used a procedure capable of considering three functions: weighted loss function, desirability function and distance function. Myers and Montgomery (1997) referred to this as a popular approach to formulate and solve the problem as constrained optimization problem. Kim, *et al.* (2001) classified it as a priority based as is similar to bounded objective in the multi objective decision making problems where response with highest performance is chosen as the objective and the rest of the functions are considered as constraints. Myers and Carter (1973) first suggested the idea where he assumed two responses as a ‘Primary response’ and a ‘Constraint response’ with goal to find condition on a set of design variables which maximize the primary response function subject to the constraint response function. Biles (1975) considered multiple process responses and extended the Myers and Carter’s idea. Del Castillo and Montgomery (1993) studied the approach later year. On designs with multiple response, Logothetis and Haigh (1988) discussed a manufacturing process with five responses, one of the five response variables was selected as primary and optimized the objective function while ignoring the possible correlations among the responses and considered the constraints of other determining input variables. Derringer and Suich (1980) proposed that a particular response  $Y_i$  among other responses in the same experimental set up may be maximized or minimized, or assigned a target value. If  $L_i$ ,  $U_i$  and  $T_i$  be lower, upper and target value respectively, the desired value for  $Y_i$  with set of constraints to get its optimal value could be obtained via desirability function  $d_i(Y_i)$ .

**Optimizing Through the Desirability Function Approach**

An analytic technique for optimization of multi-response design based on the concept of utility or desirability of a property associated with a given response objective function introduced by Harrington in 1965. This approach uses an estimated response such as  $y_{i(x)}$  transformed to a scale free

The desirability function  $d_i$  for the three scenarios of optimization problems are illustrated as

$$\left. \begin{aligned}
 \text{STB} \quad d_i &= \left[ \frac{y_i^{\max} - y_{i(x)}}{y_i^{\max} - y_i^{\min}} \right]^r && y_i^{\min} \leq y_{i(x)} \leq y_i^{\max} \\
 \text{NTB} \quad d_i &= \begin{cases} \left[ \frac{y_{i(x)} - y_i^{\min}}{t_i - y_i^{\min}} \right]^{r_1} & y_i^{\min} \leq y_{i(x)} \leq t_i \\ \left[ \frac{y_i^{\max} - y_{i(x)}}{y_i^{\max} - t_i} \right]^{r_2} & t_i \leq y_{i(x)} \leq y_i^{\max} \end{cases} \\
 \text{LTB} \quad d_i &= \left[ \frac{y_{i(x)} - y_i^{\max}}{y_i^{\max} - y_i^{\min}} \right]^r && y_i^{\min} \leq y_{i(x)} \leq y_i^{\max}
 \end{aligned} \right\} \tag{10}$$

The *min* and *max* indexes on the  $y_i$  denote the lower and upper limits accepted for  $y_{i(x)}$  respectively. The

value  $d_i$  that is called desirability which ranges from 0 to 1 and also allows users to specify minimum and maximum acceptable values for each response. In Harrington (1965) desirability ( $D$ ) is also in the  $[0,1]$  interval is obtained by combining all desirabilities ( $d_i$ ). Derringer and Suich (1980) extended the idea and presented a method to construct an overall desirability. There are three scenerios as in the case of response surface work and any one serves as a specific goals for each of the input variables. As sugested by Taguchi (1987) an approach that provides information about the mean and variance of observations, the summary statistic is computed across for observations which is called Signal-to-Noise Ratio (SNR) and emphasis on variance reduction. Each of the three scenarios depends on the choice of experimenter and or the nature or goal of the variable to optimised which can be as;

**The Smaller The Better (STB)**, the experimenter wishes to minimize the response, in this case the SNR is given by

$$\text{SNR}_s = -10 \log \sum_{i=1}^n \frac{y_i^2}{n} \tag{6}$$

**The Larger The Better (LTB)**, the experimenter wishes to maximize the response, this case is treated in the same fashion as the STB case, but  $y_i$  in equation 6 is replaced by  $1/y_i$ .

Thus, we have

$$\text{SNR}_l = -10 \log \sum_{i=1}^n \frac{1/y_i^2}{n} \tag{7}$$

**Nominal The Better (NTB) or The Target is Best**, the experimenter wishes to achieve a particular value for the response, in this case we are attempting to determine value of  $x$  that achieves a target value for the response, the SNR used by Taguchi is given by

$$\text{SNR}_s = -10 \log s^2 \tag{8}$$

where

$$s^2 = \sum_{i=1}^n (y_i - \bar{y})^2 / (n - 1) \tag{9}$$

Thus  $s^2$  is the sample variance.

$r$ ,  $r_1$  and  $r_2$  are weights specified by user for their different specific situations and  $t$  is the target value. The advantages of

this approach over other approaches are that easy to use, understand and model.

The overall desirability  $D$  is maximized with respect to the controllable factors using a geometric mean function as

$$D = (d_1, \times d_2, \times \dots d_k)^{1/k}$$

$$D = (d_1^{w_1} d_2^{w_2} \dots d_j^{w_j})^{1/\sum w_j} \tag{11}$$

$w_j$  is the number of similar  $d_i$ 's  
 The approach is based on the idea that the quality of a product or process must meet all  $k$  quality characteristics. In adopting the approach for solving the problems of optimization of several responses is the use of a multicriteria methodology which is applied when various responses have to be considered at the same time and it is necessary to find optimal compromises between the total number of input variables taken into account at a time. The Derringer function or desirability function of Harrington, (1965) is the most important and most currently used multicriteria methodology in the optimization of analytical procedures by constructing a desirability for each individual response. In summary, the measured properties related to each response are transformed into dimensionless individual desirability ( $d_i$ ) scale. Through the individual functions, the analysis showcases the specifications that each response must fulfil in the measuring procedure. The scale of the individual desirability function ranges between  $d_i = 0$ , for a completely undesirable response, and  $d_i = 1$ , for a fully desired response, above which further improvement would have no importance. The benefit of this transformation is that it makes it possible to combine the results obtained for properties

measured on different orders of magnitude.  
 From equation 10, there is need to transform into a dimensionless individual desirability,  $d_i$  in the desirability function. There are two cases for transformation to consider: one-sided and two-sided transformations. One-sided transformations are used when the goal is to either maximize or minimize the response. Two-sided transformations are used when goal is for the response to achieve some specified target value. The goal on the  $i^{th}$  response for individual value of response gives corresponding individual desirability by equation 10.

**Second Order Response Surface Model**

If all factors represent quantitative variables, the most informative model to assist analysis of the yields or response is a function of input variables, i.e.

$$Y = f(X_1, X_2, \dots, X_h) \tag{12}$$

the ordinary polynomials, the second order in particular have been extensively employed in exploring response surfaces. This is because it is generally accepted for its simple computation, easy to work with, easy to locate the optimum response. However, they exhibit the undesirable problems of unboundness, symmetry about the optimum. These polynomial models have been used in many biometry researches. In applying the response surface methodology, the dependent is viewed as a surface to which a mathematical model is fitted. For the development of regression equation related to various quality characteristics, the second order response surface may be assumed as

$$Y = B_0 + \sum_{i=1}^h B_i X_i + \sum_{i=1}^h B_{ii} X_i^2 + \sum_{i < j = 2}^h B_{ij} X_i X_j + e_r \tag{13}$$

Where  $e_r$  is a random error, the parameter  $B$ 's are called regression coefficients which are to be estimated and obtained by the design technique. The assumed surface  $Y$  contains linear, squared and cross-product terms of variables  $X_i$ 's. In order to estimate the regression coefficients, a number of experimental design techniques are available. Box and Hunter (1957) proposed that the scheme based on central composite rotatable, design fit second-order response surfaces very accurately. Many literatures on multi-response experiments utilized a second-order models. Also when restricting the response surface problem to response optimization, to select a design that will provide a good fitted model to the data, and in particular provide reliable parameter estimates, which can be used for precise prediction, second-order models are primarily used for these purposes.

In matrix form, equation 13 can be written as

$$Y = X\beta + E \tag{14}$$

Where  $Y$  is defined to be a  $h \times 1$  vector of coefficients of measured values or vector of observations,  $X$  is a  $h \times N$  matrix of known coefficients,  $\beta$  and  $E$  are vectors of  $N \times 1$  of unknown parameters and  $h \times 1$  of errors respectively. In

general matrix model form, equation 14 can be also be written as  $(Y, X\beta, \sigma^2 I_N)$ . Where  $I_N$  is the identity matrix and  $\sigma^2$  is a constant error (variance) which is fixed but unknown population constant and property which can only be estimated, for a well behaved model its characteristics depend on the error terms, under the model we assume that  $e_r \sim N(0, \sigma^2)$  that is  $e_r \sim N(0, \sigma^2)$  that  $E(e_r) = 0$  does not mean that  $e_r = 0$  for every  $r$ , that is the errors are not consistent. The solution of equation 14 can be obtained by least square matrix approach as

$$\beta = (X^T X)^{-1} X^T Y \tag{15}$$

$(X^T X)$  is a moment matrix and variance-covariance matrix of  $\beta$  is

$$\text{var}(\beta) = (X^T X)^{-1} \sigma^2 \tag{16}$$

In this study, estimation case is a straightforward non linear regression techniques in which researchers and beneficiaries of quality water can compute to compare existing acceptable water standards limits for different intended users of the resources with the results from the fitted observed data. The techniques used showcase a contribution to estimate and

monitor quality of water parameters as it recognizes existing commercial limits as adherence to quality product is becoming National priority in health sector.

#### MATERIALS AND METHODS

Borehole water samples were collected from 20 different locations across Akinyele Local Government, Oyo State, Nigeria. Each sample collected was analysed with the aid of Atomic Absorption Spectrophotometer and various quality parameters were measured to obtain different response quality values. Samples were analysed by appropriate certified and acceptable standard methods (APAH, 1998). The parameters of interest for the work are Electrical Conductivity (EC), Total Dissolve solutes (TDS), Chloride, Calcium and Magnesium ion contents. EC and TDS are taken as primary responses while Chloride, Calcium and Magnesium ions are set as constraint variables for each of primary responses. The Software application Design-Expert® 8 (Stat-Ease, Inc, Minneapolis, MN) was used in data analysis; it showcased the values of  $R^2$ , adjusted- $R^2$  and predicted- $R^2$  among others parameters estimated for fitted models.

#### Electrical Conductivity (EC) and Total Dissolve solutes (TDS)

Electrical conductivity of water is the measurement of its ability to carry an electric current and can be regarded as a crude indicator of water quality for primary purposes. It reflects the extent of solubility of mobile cations and anions and is related to the sum of ionised solutes or total dissolved solids which is the sum of cations and anions as well as organic and inorganic substances in water that can pass through a 2 micron filter. The relationship between EC and TDS is directly proportional and anyone can be estimated fairly accurately from other via a linear correlation and regression equation. However, as a rough approximation, the relationship between EC and TDS commonly used is

$$\text{TDS}(\text{mg/l}) = \text{EC}(\mu\text{Scm}^{-1}) \times 0.67 \quad (17)$$

The value  $k = 0.67$  in the above equation is for drinking water sources and varies for other sources for different uses and ranges from 0.54 to 0.96 irrespective of water sources and uses

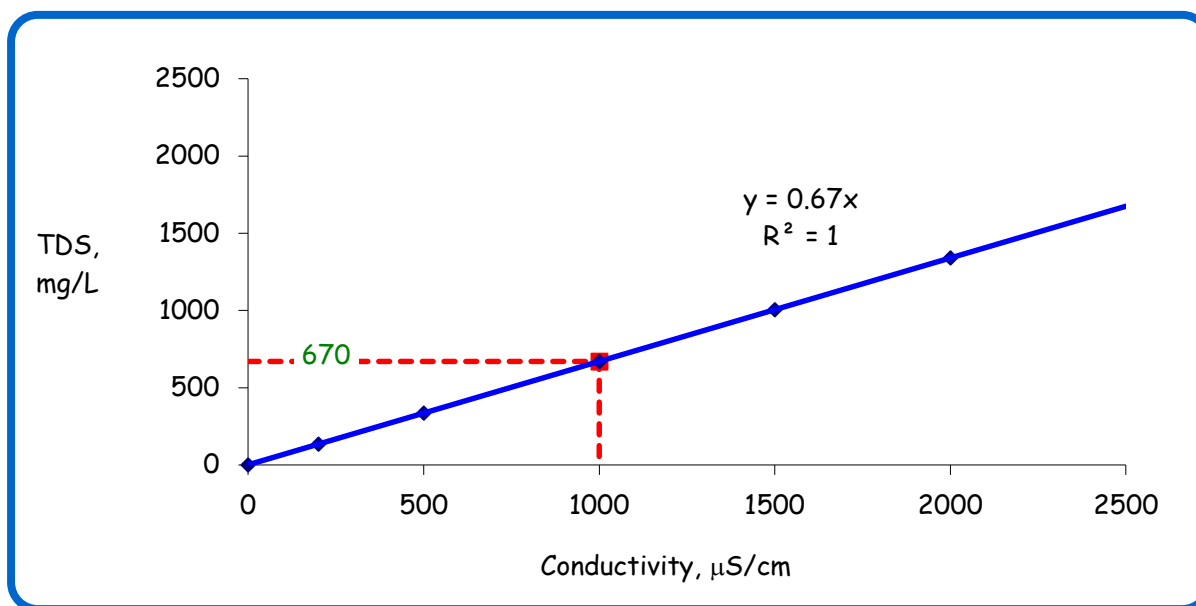


Figure 1: Showing relationship between Electrical Conductivity and Total Dissolved Solids of drinking water

High content of any of the parameters in drinking water poses serious health dangers, this is the main reason why they are used to monitor quality in drinking water through their acceptable limits guideline for use. For borehole water, EC value greater than  $500 \mu\text{Scm}^{-1}$  indicate that the water may be polluted, although, values as high as  $2000 \mu\text{Scm}^{-1}$  may be acceptable for farming, but for drinking water EC should not be more than  $500 \mu\text{Scm}^{-1}$  as water with higher value may have quality problem and be unpleasant to drink. For TDS, water with TDS greater than  $1200\text{mg/l}$  is very unusable.

Table 1: Some Physical Properties of Parameters of Borehole Water Samples

Borehole	Calcium, $Ca^{2+}$ (mg/l) $X_1$	Magnesium, $Mg^{2+}$ (mg/l) $X_2$	Chloride, $Cl^{-}$ (mg/l) $X_3$	Electrical Conductivity, EC ( $\mu S/cm$ ) $Y_1$	Total Dissolved Solid, TDS (mg/l) $Y_2$
1	27.2	16.6	18.0	180.0	200.0
2	53.6	53.7	30.0	355.0	370.0
3	27.1	15.7	60.0	245.0	260.0
4	80.0	34.2	90.0	490.0	510.0
5	92.0	12.2	49.0	315.0	340.0
6	67.6	30.0	46.0	350.0	375.0
7	57.6	19.6	50.0	280.0	290.0
8	0.8	8.30	30.0	110.0	110.0
9	14.0	23.4	30.0	115.0	120.0
10	51.2	9.77	10.0	195.0	210.0
11	40.8	37.2	48.0	330.0	350.0
12	29.6	18.0	35.0	180.0	200.0
13	62.4	23.0	80.0	305.0	325.0
14	76.0	20.0	88.0	340.0	360.0
15	27.2	20.0	26.0	180.0	200.0
16	36.0	25.2	80.0	295.0	310.0
17	68.0	57.6	122.0	715.0	750.0
18	50.4	16.6	58.0	280.0	296.0
19	35.2	17.1	12.0	185.0	190.0
20	51.2	10.2	12.2	416.0	440.0

Table 2: Descriptive Statistic of Parameters

Parameters	Minimum	Maximum	Mean	Standard Deviation	Acceptable limits		
					Upper	Lower	Penalty
EC	110.0	715.0	293.0	136	100	0	Maximum Allowable
TDS	110.0	750.0	310.0	142.0	500	0	Maximum Allowable
$Ca^{2+}$	14.0	92.0	47.4	22.7	75.0	0	Maximum Allowable
$Mg^{2+}$	9.77	57.6	23.4	13.1	50.0	0	Maximum Allowable
$Cl^{-}$	10.0	122.0	54.2	32.5	250	0	Maximum Allowable

RESULTS AND DISCUSSION

The second order response surface models fitted for each response are significant, the model incorporates interactive terms between three input factors  $X_1$ ,  $X_2$  and  $X_3$  that estimated each of responses  $Y_1$  and  $Y_2$  fitted as:

Effects on Electrical Conductivity

$$Y_{1(EC)} = 80.30 + 2.34X_1 - 2.09X_2 + 0.814X_3 + 0.0889X_1X_2 - 0.0380X_1X_3 + 0.0255X_2X_3 + 0.0190X_3^2 \tag{18}$$

Summary of ANOVA Table 3 shows that the model is significant at 0.05 with all single factors found significant ( $p < 0.05$ ), two factor interaction were also found significant except for one while quadratic term was significant were retained. The three  $R^2$  statistics values in Table 5 also explain the significant of the fitted model as predicted  $R^2$  of 82.0% is in reasonable agreement with adjusted  $R^2$  of 96.5%.

**Table 3: Summary of ANOVA results for response surface model of Electrical Conductivity on Calcium, Magnesium and Chloride**

Source	Sum of Squares	df	Mean Square	F- Value	p-value	
Model	3.63E+005	7	5.18E+004	75.9	<	0.0001
X <sub>1</sub> -Calcium	1.18E+004	1	1.18E+004	17.4	0.00131	
X <sub>2</sub> -Magnesium	5.18E+004	1	5.18E+004	75.9	<	0.0001
X <sub>3</sub> -Chloride	4.36E+004	1	4.36E+004	63.9	<	0.0001
X <sub>1</sub> X <sub>2</sub>	3.41E+003	1	3.41E+003	5.01	0.0450	
X <sub>1</sub> X <sub>3</sub>	3.75E+003	1	3.75E+003	5.50	0.0370	
X <sub>2</sub> X <sub>3</sub>	2.51E+003	1	2.51E+003	3.68	0.0791	
X <sub>3</sub> <sup>2</sup>	7.11E+003	1	7.11E+003	10.4	0.00723	
Residual	8.18E+003	12	682.0			
Total	3.71E+005	19				

*Effects on Total Dissolved Solids*

$$Y_{2(TDS)} = 392 + 112X_1 + 106X_2 + 133X_3 + 78.9X_1X_2 - 68.4X_1X_3 + 52.1X_2X_3 \tag{19}$$

From equation 19, it may be observed that from ANOVA table 4 that all the single terms are significant while all interaction terms not significant, no quadratic term found significant as the model has only 0.01% chance that its value could due to noise, the three R<sup>2</sup> statistics in Table 5 indicate suitability of the model.

**Table 4: Summary of ANOVA results for response surface model of Total Dissolved Solids on Calcium, Magnesium and Chloride**

Source	Sum of Squares	df	Mean Square	F-Value	p-value
Model	3.85E+005	6	6.42E+004	46.5	<0.0001
X <sub>1</sub> -Calcium	1.22E+004	1	1.22E+004	8.80	0.0109
X <sub>2</sub> -Magnesium	5.55E+004	1	5.55E+004	40.2	<0.0001
X <sub>3</sub> -Chloride	5.31E+004	1	5.31E+004	38.4	<0.0001
X <sub>1</sub> X <sub>2</sub>	3.09E+003	1	3.09E+003	2.24	0.159
X <sub>1</sub> X <sub>3</sub>	2.61E+003	1	2.61E+003	1.89	0.192
X <sub>2</sub> X <sub>3</sub>	6.14E+003	1	6.14E+003	4.44	0.0550
Residual	1.80E+004	13	1.38E+003		
Total	4.03E+005	19			

**Table 5: Parameters of the fitted Response Surface Model EC and TDS**

Response	F-Value	P-Value	R <sup>2</sup>	Adj. R <sup>2</sup>	Pred. R <sup>2</sup>	Adeq. Pred.	COV
Y <sub>1(EC)</sub>	75.9	<0.0001	0.978	0.965	0.820	36.4	8.91%
Y <sub>2(TDS)</sub>	46.5	<0.0001	0.955	0.935	0.774	27.2	12%

**Optimization Results and Validation**

During the optimization stage, the desirability function approach was used to obtain the best compromise with respect to each response acceptable limits as constraints. The second order polynomial model was fitted to each response observed

data to obtain optimal values shown in Table 6. The goal was to minimize i.e. Smaller The Better (STB) scenario which reflects the optimum condition values for EC and TDS containing minimum amount of Calcium, Magnesium and Chloride.

**Table 6: Criteria and Output for Numerical Optimization of Selected Water Quality Parameters**

Criteria	Goal	Observed limits	Acceptable limits	Output
Electrical Conductivity, Y <sub>1</sub>	Minimized	110 – 715	0 – 100	80.3
Total Dissolved Solids, Y <sub>2</sub>	Minimized	110 – 750	0 – 500	210
Calcium, X <sub>1</sub>	Minimized	14 – 92	0 – 75	51.2
Magnesium, X <sub>2</sub>	Minimized	9.77 – 57.6	0 – 50	9.77
Chloride, X <sub>3</sub>	Minimized	10 – 122	0 – 250	10
Desirability, D				0.975

**Numerical Optimization Objective and Constraint Functions**

As the process includes both dependent and independent variables, the optimal solutions are characterized by objective function represented or obtained by fitted response model and set of constraint functions represented or obtained by acceptable conditions of the process or process equilibrium state. The optimization task in this process components typically reads.

For EC

Minimise:

$$Y_{1(EC)} = 80.30 + 2.34X_1 - 2.09X_2 + 0.814X_3 + 0.0889X_1X_2 - 0.0380X_1X_3 + 0.0255X_2X_3 + 0.0190X_3^2$$

Subject to:

$$0 < X_1 < 75$$

$$0 < X_2 < 50$$

$$0 < X_3 < 250$$

$$X_1, X_2, X_3 > 0$$

For TDS

Minimise:

$$Y_{2(TDS)} = 392 + 112X_1 + 106X_2 + 133X_3 + 78.9X_1X_2 - 68.4X_1X_3 + 52.1X_2X_3$$

Subject to:

$$0 < X_1 < 75$$

$$0 < X_2 < 50$$

$$0 < X_3 < 250$$

$$X_1, X_2, X_3 > 0$$

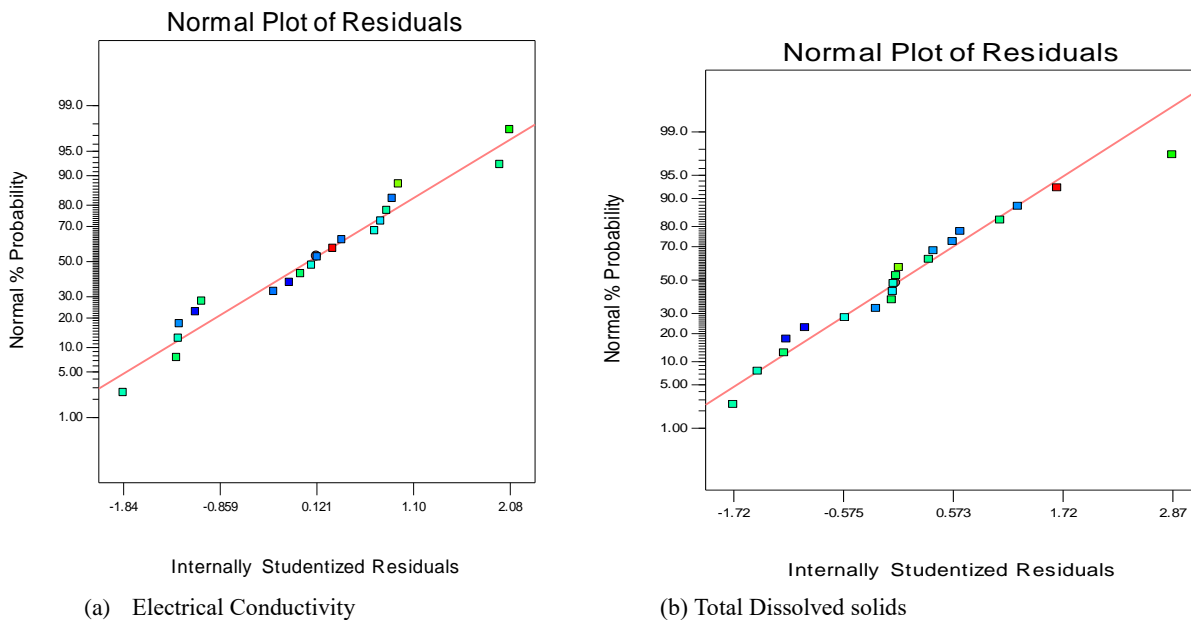


Figure 2: Normal Percentage Probability Plots of (a) Electrical Conductivity and (b) Total Dissolved Solids of drinking water



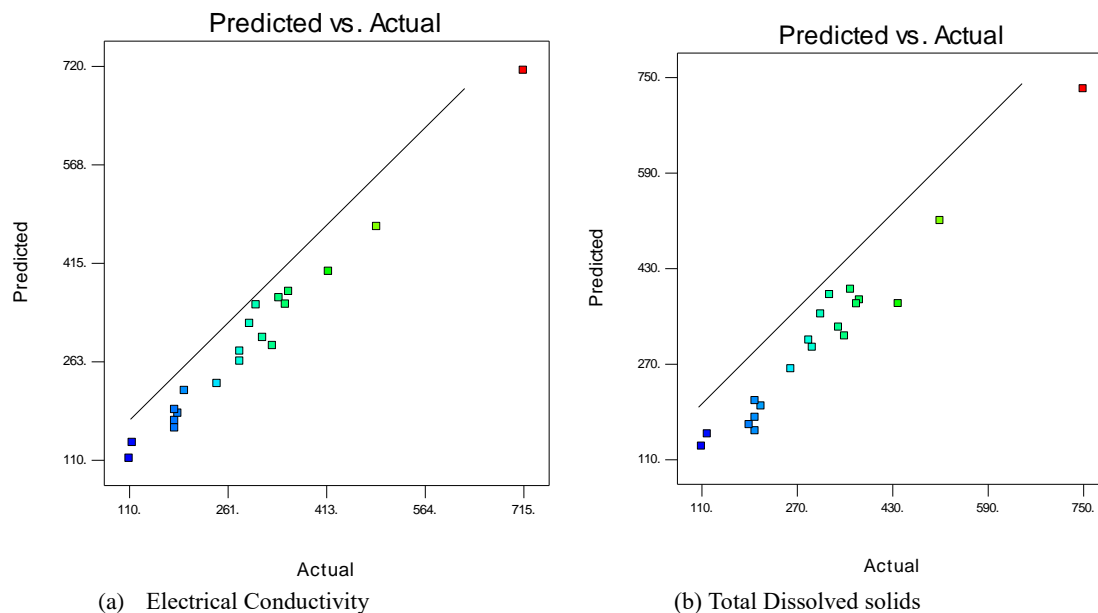


Figure 3: Graphs of predicted and actual values of (a) Electrical Conductivity and (b) Total Dissolved Solids of drinking water on Calcium, Magnesium and Chloride

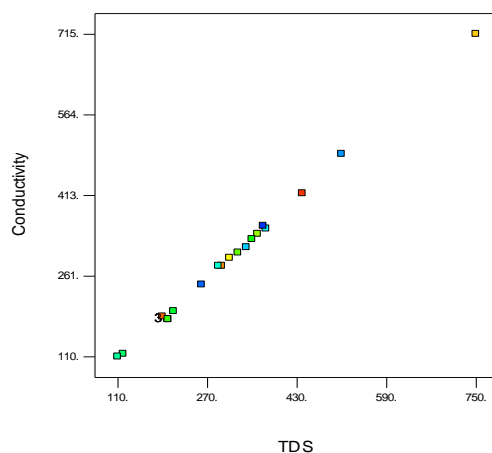


Figure 4: Graph of relationship between observed values of Electrical Conductivity and Total Dissolved Solids of drinking water

#### CONCLUSION AND RECOMMENDATION

RSM utilizes regression techniques to study experimental products and process, however, optimization of multiple response design depends too heavily on the assumptions of well estimated models fitted for the responses of interest. It can be seen that optimal values obtained for responses in Table 6 are within the acceptable standards with other model parameters which make the method to be reliable. The optimal values of  $80.3 \mu\text{Scm}^{-1}$  and  $210 \text{ mg/l}$  of EC and TDS respectively are obtained by the same optimal values of their same set of input factors of  $51.2 \text{ mg/l}$ ,  $9.77 \text{ mg/l}$  and  $10 \text{ mg/l}$  for Calcium, Magnesium and Chloride respectively. The result will be beneficial to water users most especially for drinking to improve the quality of the product for health reasons and benefits.

#### REFERENCES

- Biles W. E. (1975), A Response Surface Method for Experimental Optimization of Multiresponse Processes, *Industrial and Engineering Chemistry—Process Design and Development*, 14, 152–158.
- Box G. E. P. and Hunter J. S. (1957). Multifactor Experimental Design. *J. of Ann Math. Stat*, 28.
- Cornell A. I. (1996). *Response Surfaces; Designs And Analyses* (2nd ed.). New York, USA: Marcel Dekker, Inc.
- Del Castillo E. and Montgomery D. C. (1993), A Nonlinear Programming Solution to the Dual Response Problem, *Journal of Quality Technology*, 25, 3.

- Derringer G. and Suich R. (1980). Simultaneous Optimization of Several Response Variables. *Journal of Quality Technology*, 12, 214-219.
- Harrington E. C. Jn. (1965). The Desirability Function. *Industrial Quality Control*, 494-498.
- Hill W. J. and Hunter W. G. (1966). A Review of Response Surface Methodology: A Literature Review. *Technometrics*, 571-590.
- Khuri A. I. and Cornell J. A. (1987). *Response Surfaces*. New York: Dekker.
- Khuri A. I. and Cornell J. A. (1996). *Response Surface: Design and Analysis*. New York: Marcel Dekker, Inc.,
- Layne K. L. (1995). Method to Determine Optimum Factor Level for Multiple Responses in the Designed Experimentation. *Quality Engineering*, Vol. 7, No. 4, 649-656.
- Logothetis N. and Haigh A. (1988). Characterizing and Optimizing Multi-response Processes by Taguchi Method. *Quality and Reliability Engineering International* 4, 159-169.
- Myers R. H. and Carter W. H. (1973). Response Surface Techniques for Dual Response Systems. *Technometrics*, 15, 301-317.
- Myers R. H. and Montgomery D. C. (1997), A Tutorial on Generalized Linear Models, *Journal of Quality Technology*, 29, 274-291.
- Pignatiello J. J. Jr. (1993). Strategies for Robust Multi-response Quality Engineering. *IIE Transactions* 25, 5-15.
- Raymond H. Myers, Douglas C. Montgomery and Christine M. Anderson-Cook. (2016). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* (4th ed.). New Jersey: John Wiley & Sons, Inc.
- Raissi S. and R- Eslami Farsani. (2009). Statistical Process Optimization Through Multi-Response Surface Methodology. *World Academic of Science, Engineering and Technology* 27, 267-271.
- Taguchi G. (1987). *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*, UNIPUB/Kraus International, White Plains, NY
- Tai C. Y., Chen T. S. and Wu M. C. (1992). An Enhanced Taguchi Method for Optimizing SMT Process. *Journal of Electronics Manufacturing*, Vol. 2, 91-100.
- Way Kuo, V. Rajendra Prasad, Frank A. Tillman and Ching-Lai Hwang. (2001). *Optimal Reliability Design; Fundamental And Applications*. Cambridge, Cambridge, United Kindom: University Press, Cambridge.



©2020 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via <https://creativecommons.org/licenses/by/4.0/> which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited