© 2018 IJSRSET | Volume 4 | Issue 11 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology
DOI : https://doi.org/10.32628/IJSRSET21841110

G-Optimal Design in Non-linear Models to Increase Silicon Oxide Purity Levels and Electrical Conductivity

Muklas Riva'i*[1], Bagus Sartono[2], Erfiani[3], Irzaman[4]
[1, 2, 3] Department of Statistics, Bogor Agricultural University, Bogor, West Java, Indonesia
[4] Department of Physics, Bogor Agricultural University, Bogor, West Java, Indonesia

ABSTRACT

Optimal design is a design which required in determining the points of variable factors that would be attempted to optimize the relevant information so that fulfilled the desired criteria. The optimal fulfillment criteria based on the information matrix of the selected model. The experimental design in silicon oxide with a purity rise silicon oxide obtained following the exponential decay distribution and approaches an asymptotic value which is a non-linear models. This research is aimed to obtain the best designs in determining the levels of silicon. The method used is the G-optimal criterion on non-linear models using the exchange algorithm. G-optimal is an optimal criteria in order to minimize the maximum variety of the estimation responses. The results of this study shows that the best G-optimal design for non-linear models to the respons of purification levels of SiO_2 (Y_1) and electrical conductivity (Y_2) is alternative 3 ($Y_1 = 0.75$ and $Y_2 = 0.25$) with temperature levels of 700°C, 705°C, 710°C, 735°C, 740°C,750°C, 835°C, 840°C, 860°C, 865°C, 870°C and 875°C with the G-efficiency value of Y_1 is 78.10% and the G-efficiency value of Y_2 is 67.48%.

Keywords : G-Optimal, Non-Linear Model, Optimal Design, Silicon Oxide

I. INTRODUCTION

The advances of science and technology demanding change in the various fields including the change in research and experiment. In order to provide clear answers and conducted with minimum costs, the research required an appropriate statistical method. One problem that can be solved by statistical methods is to determine the effects of several experimental factors. This can be accomplished through a design that is processed using optimal design theory [1].

One of the developing experiments is an experiment to obtain inexpensive raw materials of silicon oxide through the purification process to increase the selling value. In most areas in Indonesia, the biomass sector as materials for silicon oxide has a huge potential. Based on the selling power by Aldrich [2], the price of one gram silicon oxide with 9.99% purity rate is Rp. 4000. It means that the selling power obtained from 0.5376 tons of silicon oxide is Rp. 2,1504 billion per year. Additionally, 0.5376 tons of silicon per year can supply 1.8% of the national silicon oxide needs (30 tons per year).

Silicon oxide which has low purity can be improved through the purification process by setting the temperature, the silicon purity levels, and the different of rising temperature rates. The combination of each factor which affects the refining process through a series of experiments will produce different purity levels. According to Montgomery [3], the experiment was designed to obtain the conclusions with an error as small as possible (optimal design).

The main problem in optimal design lies in the selection of treatment or combination of treatments that must be tested in order to obtain the desired efficiency.

Optimal design is needed to determine the points of the factor variables that would be attempted to optimize the relevant information so that fulfilled the desired criteria. The required criteria to obtain the optimal experimental design matrix is the variance of the estimator response must be minimum. Hence, it needs an appropriate and more optimal design to produce an accurate statistical inference with minimum experimental costs. For this purpose, the optimal criteria and efficiency values of the design are used. Optimal design is needed to determine which point of variable X will be tried with to maximize the amount of relevant information. G-optimality design is an optimal criterion with the aim of minimizing the maximum variety of the alleged responses in the range of research interests [1].

Non-linear models have been widely used in the fields of phamacokinetics and chemistry, such as the exponensial models [4]. In the nonlinear models, the information matrix depends on unknown parameters. Overcoming this problem, some additional types of information are needed both from the initial value and distribution. Some researchers used non-linear models in optimal design [5-7]. The results of the selection of design points in this study are the optimum design with a more economical experiment cost. Therefore the researchers are interested to study the G-optimal design for non-linear models in increasing the purity levels of Silicon oxide.

II. MATERIAL AND METHOD

A. Models Used

The model in this study is a non-linear model obtained from the relationship between factors and responses. The used factors in the design are the temperature and the purity level of silicon with the increasing temperature per minute is 0.1°C. The temperature factor levels are at intervals of 700°C to 900°C. The levels of the silicon oxide purity factor are 70%, 80%, 90%, and 99.5%. There are two responses in this study i.e. the purity level of silicon oxide and the electrical conductivity. Model selection in this study is done subjectively by researchers by using the information from the experts in their fields.

The first model used in this study is the relationship between temperature and the purity level of silicon oxide. According to Coniwanti et al. [8], the higher the combustion temperature of silicon oxide, the higher the purity of silicon oxide obtained. The increase in the purity of silicon is obtained by following the exponential decay distribution and approached in an asymptotic value. The exponential decay equation can be written as follows.

$$Y1(t, K, \theta) = [A_0 K]\{1 - \exp(-\theta t)\}$$

with:

Y1 : expected value of the response

A_0 : constants

K : purity levels

θ : parameters

t : temperature

The second model used in this study is the relationship between the temperature and the electrical conductivity.

$$Y2\ (\mathrm{t}, \mathrm{K}, \theta) = A_0 K e^{-\theta(1/t)}$$

with:

$Y2$: electrical conductivity of SiO_2

A_0 : constants

K : purity levels

θ : parameters

t : temperature

The two models used in this study are non-linear models, so an approach is needed to simplify the model. Taylor's approach is an approach used to approach non-linear equations through linear equations. The simplification is used to simplify the algorithm that will be done. In general, the Taylor approach can be written as

$$\boldsymbol{e^{x} \cong 1 + \frac{x}{1!} + \frac{x^{2}}{2!} + \frac{x^{3}}{3!} + \cdots}$$

B. Simulation of design point selection

The algorithm used is the Exchange Algorithm. For example, chosen r point from the candidate group contains m point to fit the model with q parameters with q≤r <m. The process of this algorithm is done by randomly selecting points as much as r, then choose one point from the remaining candidate groups so that the optimal criteria are the most optimum. There are (r + 1) selected point, then dispose one of those point which causes the value of the criteria become low [1].

The G-optimal criterion is a criterion to obtain X so that minimize the maximum variance of the estimated response in the range of research. Several studies show that design matrix X will reach the optimum criteria when the design points contain the extreme points on a specific range in each factor. According to Aguiar et al. [9], the variance function is as a measure of uncertainty of the estimated response. The variance of the estimated response for a single candidate x_i can be calculated using the following formula:

$$Var(\hat{y}(x)) = (x_i'(X'X)^{-1}x_i)\sigma^2$$

with x_i is a vector that describes a single experiment and x_i' is a transpose from x_i. In the G-optimal design, the selected candidates are candidates which have the smallest maximum estimated response variance, as follows:

$$\min_{x=1,2,\ldots,n} \max_{\xi_n} Var(\hat{y}(x)$$

$$= \min_{x=1,2,\ldots,n}\left(\left(\max_{\xi_n}\sum_{i=1}^{n} x_i'(X'X)^{-1}x_i\right)\sigma^2\right)$$

The G-optimal design can be said to be designed at the best condition if the maximum of prediction variance is the same as the parameter (p), where p is the number of parameters in the model. The indicator above is often called as G-efficiency. G-Efficiency is used to determine the goodness of the chosen design. Here's the formula of G-efficiency at the G-optimal design.

$$G - efficiency = 100\,x\left(\frac{p/N}{\max_{\xi_n}\sum_{i=1}^{n} x_i'(X'X)^{-1}x_i}\right)^{1/2}$$

The algorithm to determine the design point of the temperature factors is described as follows:

1. Create N design point of the candidate model from the temperature factor which is in the interval of 700°C to 900°C.
2. Based on the predetermined model, take 12 design points (x_i) so that the matrix $\mathbf{X}_{(0)}$ is obtained.
3. Calculate and determine the maximum variance of each estimated response (Y_1 and Y_2) at the matrix $\mathbf{X}_{(0)}$ with the following formula:
$$Max\,Var\left(\hat{y}_j(x)\right) = max\left(\sum_{i=1}^{12} x_i'(X'X)^{-1}\,x_i\right)\sigma^2$$
with $i = 1,2,\ldots,12\ dan\ j = 1,2$
4. Determine the maximum variance of combined responses based on the proportions on Y_1 in the amount of P_1% and Y_2 in the amount of P_2% with the following formula:
$Max\,Var\left(\hat{y}(x)\right) = P1\,Var\left(\hat{y}_1(x)\right) + P2\,Var\left(\hat{y}_2(x)\right)$
with $P_1 > P_2$.
5. Substitute x_i with a new design point (x_j) so that a new matrix $\mathbf{X}_{(1)}$ is obtained.
6. Calculate the maximum variance of the combined responses from the matrix $\mathbf{X}_{(1)}$
7. Compare the average variance of combined response matrix $\mathbf{X}_{(0)}$ and matrix $\mathbf{X}_{(1)}$, then select the matrix $\boldsymbol{X}$ with the smallest maximum variance combined response. The selected matrix is used as the initial matrix ($\mathbf{X}_{(0)}$).

8. Repeat steps 5, 6, and 7 by iterating process until the design point is selected with the smallest maximum variance combined response.
9. Calculate the G-efficiency value in the selected design.

III.RESULTS AND DISCUSSION

A. Non-linear Model Approach

The Taylor series expansion approach is one way to simplify non-linear equations to linear equations [1]. The principle used is Taylor through the k-order Polynomial equation. The higher the order chosen the smaller the resulting error will be. Yet, this is inversely proportional to the model obtained since the higher the order chosen is the more complex the model. The selection of order in this study considers both the error and the model. The following table shows the results from the Taylor approach for the two models used.

Table 1 Taylor approach for the purification level of SiO_2 (Y_1) and the electrical conductivity (Y_2) in each model

Model	Average error of Y_1	Average error of Y_2
Order-1	0.20774	0.16340
Order-2	0.05349	0.03659
Order-3	0.01018	0.00607
Order-4	0.00154	0.00080

Based on the results, the chosen order is the 3rd order since it has a small error and the model is not complex. At the purification level SiO_2 (Y_1) with the third order, Taylor approach produces an average error of 0.01018 while the electrical conductivity (Y_2) produces an average error of 0.00607. The approach model produced in the 3rd order is the 3rd order polynomial. The selection of the order will show a $n \times p$ of matrix $\boldsymbol{X}$ with p is the number of parameters and n is the number of design points taken. Therefore the 3rd order matrix is sized 12x4.

B. Design for The Purification Levels of Silicon Oxide

This design was made when the observations only see the purification level of silicon oxide (Y_1) without prioritizing the electrical conductivity response (Y_2). Modifications are formed by making the proportion for Y_1 equal to 100% while Y_2 is not given a proportion. The designs are presented in Table 2.

Table 2 Design point and G-efficiency of the model

No	Temperature (°C) Alternative 1 Y_1=1.00,Y_2=0
1	720
2	725
3	730
4	735
5	765
6	775
7	835
8	840
9	880
10	885
11	890
12	895
G-efficiency of Y_1	86.17%
G-efficiency of Y_2	60.66%

The best design with the G-optimal criteria is found in Table 2. The design is when the study only considered the response of the purification level of SiO_2. Based on the Table 2, the G-efficiency value of Y_1 is 86.17%, which means that the design can explain the parameters of the relationship between temperature and the purification level of SiO_2. Then if the model is used to electrical conductivity response (Y_2), the G-efficiency is 60.66%. Therefore the design is good to used for the response of purification level of SiO_2.

C. Design for electrical conductivity

This design was made when the observations made only wanted to see the response of the electrical conductivity (Y_2) without prioritizing the purification level of silicon oxide (Y_1). Modifications are formed by making the proportion for Y_2 equal to 100% while Y_1 is not given a proportion. The following designs are presented in Table 3.

Table 3 Design points and G-efficiency of the model

No	Temperature (°C) Alternative 2 Y_1=0,Y_2=1.00
1	700
2	720
3	730
4	735
5	750
6	755
7	760
8	845
9	850
10	855
11	860
12	885
G-efficiency of Y_1	59.69%
G-efficiency of Y_2	72.58%

When the design only considers the electrical conductivity respons, the best design with the G-optimal criteria is in Table 3. Based on Table 3, the G-efficiency value of the electrical conductivity response is 72.58%. Then if the model is used to the purification level of SiO_2 response, the G-efficiency is 59.69%.

D. Design for The Purification Levels of SiO_2 and The Electrical Conductivity

The design which consider both responses between the purification levels of SiO_2 (Y_1) and the electrical conductivity (Y_2) could be formed. The modification process is done by making the various proportion for Y_1 and Y_2. The proportion for Y_1 is greater than Y_2 since the response from Y_2 is more prioritized than the response Y_1. The proportions formed by 4 alternatives are alternative 3 (Y_1 = 75, Y_2= 25), alternative 4 (Y_1= 70, Y_2 = 30), alternative 5 (Y_1 = 65, Y_2 = 35) and alternative 6 (Y_1= 60, Y_2 = 40). The results are presented in Table 4.

Table 4 Alternative designs with various proportions

No	Temperature (°C)			
	Alternative 3 Y_1=0.75,Y_2=0.25	Alternative 4 Y_1=0.70,Y_2=0.30	Alternative 5 Y_1=0.65,Y_2=0.35	Alternative 6 Y_1=0.60,Y_2=0.40
1	700	725	700	715
2	705	730	705	720
3	710	735	710	725
4	735	765	770	755
5	740	770	775	765
6	750	835	780	770
7	835	840	825	830
8	840	855	830	835
9	860	860	875	865
10	865	865	885	870
11	870	870	890	875
12	875	875	895	880
G-efficiency of Y_1	78.10%	76.14%	76.10%	75.86%
G-efficiency of Y_2	67.48%	67.23%	64.28%	67.78%

The design of each alternative has the same characteristics, i.e. there is no temperature lied in the average candidate. The average candidate for this design is 800°C with minimum value of 700 °C and maximum value of 900°C. When viewed from the G-efficiency value of Y_1, the highest value is alternative 3 of 78.10% and the lowest value is alternative 6 with 75.86 %. Then for the G-efficiency value of Y_2, the highest value is alternative 6 with 67.78% and the lowest value is alternative 5 with 64.28%. The G-efficiency difference value in each alternative for Y_1 is relatively small compared to Y_2. Therefore the appropriate alternatives to used when prioritize both responses is alternative 3.

E. The Comparison of Designs in Each Alternative

The determination of the best design is done by comparing the results of the previous design above, namely the design by prioritizing the purification level of SiO_2 responses (alternative 1), design by prioritizing electrical conductivity responses (alternative 2) and design by prioritizing both responses (alternative 3). The following is a plot of the design points.

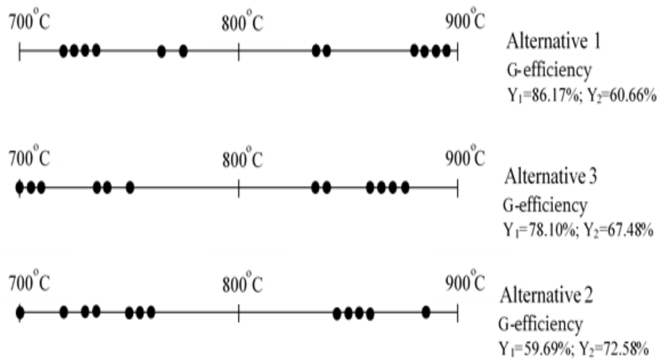

Figure 1. Design points for each alternative

Figure 1 shows a plot of the G-optimal design point and G-efficiency for each alternative. The plot of the three alternatives forms 4 groups of temperature. The shift in plot shape occurs in alternatives 2 which are close together so that they tend to form 2 groups of temperature. Then based on G-efficiency, alternative 3 has a G-efficiency of Y_1 and Y_2 which is high enough and balanced for both responses, which are 78.10% and 67.48% respectively. While the other two alternatives only have one high G-efficiency while the other one is low. Therefore the alternative design chosen is alternative 3.

IV.CONCLUSION

The best G-optimal design for non-linear models of the purity level response of SiO_2 (Y_1) and the electrical conductivity (Y_2) is alternative 3 (with Y_1= 0.75 and Y_2 = 0.25). The alternative design points 3 are 700°C, 705°C, 710°C, 735°C, 740°C, 750°C, 835°C, 840°C, 860°C, 865°C, 870°C and 875°C with the G-efficiency value of Y_1 is 78.10% and the G-efficiency value of Y_2 is 67.48%.

V. REFERENCES

[1] Atkinson AC, Donev AN, Tobias RD. 2007. Optimum Experimental Designs, with SAS. New York: Oxford university press

[2] Aldrich S. 2006. Advancing Science. Singapore; Sigma-Aldrich Pte Ltd.

[3] Montgomery DC. 2001. Design and Analysis of Experiments. New York (US): John Wiley and Sons, Inc.

[4] Widiharih T, Haryatmi S, Gunardi. 2014. Rancangan D-optimal untuk model eksponensial general. Media Statistika. 7(2): 71:76.

[5] Yang M, Stufken J. 2012. Identifying locally optimal designs for nonlinear models: a simple extension with profound consequences. *The Annals of Statistics*. 40(3): 1665–1681.

[6] Rodríguez C, Ortiz I, Martínez I. 2015. Locally and maximin optimal designs for multifactor nonlinear models. *Statistics*. 49(5): 1157-1168.

[7] Belmiro PM, Duarte, Wong KW, Nuno MC, Oliveira. 2016. Model-based optimal design of experiments-Semidefinite and nonlinear programming formulations. *Chemometrics and Intelligent Laboratory Systems*. 151: 153–163.

[8] Coniwanti P, Srikandhy R, Apriliyanni. 2008. Pengaruh proses pengeringan, normalitas hcl, dan temperatur pembakaran pada pembuatan silika dari sekam padi. *Jurnal Teknik Kimia*. 15 (1): 5-10

[9] Aguiar DPF et al. 1995. Tutorial D-optimal designs. *Journal of Chemometrics and Intelligent Laboratory System*. 30: 199-210.