Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 17(1) 2024: 36-44
http://ejournal.undip.ac.id/index.php/media_statistika

A-OPTIMAL DESIGN IN NON-LINEAR MODELS TO INCREASE SILICON DIOXIDE PURITY LEVELS

Ghea Weisha[1], Erfiani[1], Utami Dyah Syafitri[1], Irzaman[2]
[1] Department of Statistics, IPB University, Bogor, Indonesia
[2] Department of Physics, IPB University, Bogor, Indonesia

e-mail: *utamids@apps.ipb.ac.id*

DOI: 10.14710/medstat.17.1.36-44

Article Info:
Received: 28 April 2022
Accepted: 21 September 2024
Available Online: 14 October 2024

Keywords:
Optimal Design; A-Optimal; Silicon Dioxide
.

Abstract: Silica is the most mineral found on earth and is widely used in industry. Silica used in industry is usually silicon dioxide with a purity ≥ 95% and its often sold at a higher cost. To obtain the silica at a lower cost, silica extraction from biomass such as rice husk can be conducted. The purity of silica extracted from biomass tends to be lower than that of mineral silica. Silica with low purity can be increased by adjusting the temperature and the rate of temperature rise. This research aims to obtain the best design to determine the purity of silicon dioxide. The design of this study was generated based on the A-optimality criterion using the DETMAX algorithm. The A-optimality criterion is minimizing the trace of the variance-covariance of the parameter estimation. The best design points obtained using A-optimal design consist of three temperature groups: the minimum temperature of 800°C, the middle temperature of 850°C, and the maximum temperature of 900°C, with varying rates of temperature rise. Points were repeated at the temperature of 850°C, with rates of temperature rise of 1.67°C/min and 3.34°C/min.

1. INTRODUCTION

Silicon dioxide (SiO2), or silica is a common mineral found on Earth. Utilizing silica is quite large, including in the production of windows, glasses, beverage bottles, and more. Silica is also a widely used material in semiconductor devices as the primary material for electronic components (Arul Prishya et al., 2023). High-purity silica (≥ 95%) is used in industrial applications such as microcomputer chips, solar cells, and others (Ghosh et al., 2017). Mineral silica is generally sold at a relatively high price. While mineral silica commands a premium, silica can also be extracted from biomass sources like rice husks, rice straw, and bamboo leaves, offering a more accessible alternative.

Previous studies have shown that rice husk ash contains the highest silica content among the agricultural biomass studied, ranging from 90-98% of its dry weight (Casnan et al., 2019). Bamboo leaf ash, with a silica content of 75.9% (Sa'diyah et al., 2016), comes in second place. Straw ash, on the other hand, contains approximately 75% silica (Har et al., 2019). Therefore, based on these findings, rice husk ash exhibits the highest silica content.

In most countries that grow rice, the rice husk is often thrown away or burned after the rice is processed. As agricultural waste, the annual output of rice husks is 120 million

tons worldwide (Mohd Kamal et al., 2014). Compared with other biomass, the main content of rice husk ash is silicon dioxide (Guo et al., 2021). Due to the extremely high selling price of mineral silica, rice husk waste can be used to produce silica and it has more economical alternative. The low purity of silica can be improved through an extraction process by adjusting the temperature factor (°C) and rate of temperature rise (°C/min), where the combination of these factors will affect the improvement of silica purity to varying degrees (Adli et al., 2018; Aminullah et al., 2018; Sintha et al., 2017).

To obtain the combination of these factors, an experimental design is needed. Experimental design is conducted to obtain results with minimal error (Montgomery, 2020). One of the challenges in conducting experimental design is that the combination of these factors can result in a large number of experiments. To investigate the impact of multiple experimental factors, researchers can employ optimal design theory (Atkinson et al., 2007).

The optimal design is used to identify the best test points to achieve the desired outcomes in an experiment. The optimal design uses several criteria, one of which is the optimal design with criteria based on parameter estimators, such as D-, A-, I-optimal criteria and others. This research was conducted collaboratively by a group of researchers. The members was tasked with investigating the relative merits of I-optimal (Aliu et al., 2024) and G-optimal (Wulandari et al., 2023). A previous study evaluated the performance of A-optimal and D-optimal designs, which are criteria used in screening experiments to assess parameter estimate variance, the A-optimal design gave better results than the D-optimal design (Jones et al., 2020).

This research aims to find the best design for estimating parameters in a non-linear model using the A-optimality criterion. Non-linear models are more complex than linear models, making it harder to find the optimal design. Determining the optimal design in the non-linear model is relatively more difficult than in the linear model due to the need for additional information in the information matrix. The goal of this study is to find the best design for determining the purity of silica levels.

2. LITERATURE REVIEW

2.1. Non-linier Model

A non-linear model is a relationship between the response variable and the explanatory variable that is not linear in the parameters. In general, the non-linear model can be written as follows:

$$y_i = \eta(t, \theta) + \varepsilon_i \quad (i = 1,2, \dots, n) \tag{1}$$

The non-linear model in this study is a form of relationship between temperature (factor) and purity level. The relationship between temperature and the silica purity levels follows an exponential distribution (Rivai et al., 2018). The non-linear model widely used in the study of pharmacokinetics and chemical kinetics is the exponential decay model (Atkinson et al., 2007). The exponential decay model with single factor is as follows:

$$f(t) = [A_0]\{1 - \exp(-\theta t)\} \tag{2}$$

with $f(t)$ = expectation value; A_0 = constants; θ = parameter; t = factor.

Estimating the parameters of non-linear models usually cannot be solved analytically, so a numerical method is needed to obtain the estimated parameters. One of the numerical methods that can be used is the Taylor approach (Kouki & Griffiths, 2019). A

model is approximated using the Taylor approach to get results closer to the initial solution. The nth-order Taylor polynomial approach used for the two variables is as follows:

$$f(t,r) = \sum_{i=0}^{n} \sum_{j=0}^{n-i} \frac{\frac{d^{(i+j)}}{\partial t^i \partial r^j} f(a,b)}{i!j!} (t-a)^i (r-b)^j \quad (3)$$

where $f(t,r)$ are a function of t and r, a, and b are constants.

2.2. Optimal Design

The experimental design generally arranges the possibilities that arise in an experiment. If an experiment has a factor consisting of several levels, the number of experiments carried out will be enormous. The more factors are used, the more experiments are carried out, so it is increasingly challenging to choose design points that produce optimal experiments.

Optimal design is one part of the experimental design which aims to obtain n optimal design points that can produce a model coefficient with the smallest standard error (Aguiar et al., 1995). The A-optimality criterion in the design is expected to obtain parameter estimates with a minimum variance and increase the accuracy of predicting the results of an experiment.

2.3. A-Optimality Criteria

An A-optimal design minimizes the sum of the diagonal elements of the ordinary least-squares estimator's variance-covariance matrix (Jones et al., 2020). The sum is called the trace of the variance-covariance matrix, an A-optimal design minimizes the trace$(\mathbf{X}^T\mathbf{X})^{-1}$. The form of matrix $\mathbf{X}$ can be explained as follows (Atkinson et al., 2007):

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

$\mathbf{X}$ is the n x p matrix, where n is the number of experiments and p is the number of parameters.

A measure of the design goodness can be calculated with the efficiency values, for A-optimality criteria is called A-efficiency value. The A-efficiency value can be written as follows (Atkinson et al., 2007):

$$A_{eff} = 100 \times \left(\frac{p/N}{trace(\mathbf{X}^T\mathbf{X})^{-1}} \right) \quad (4)$$

where p is the number of parameters, N is the number of design points.

2.4. Design Points Selection

An algorithm forms the design point candidate selection. One of them is the point-exchange algorithm. This algorithm aims to remove or add design matrix points by looking at the effect of the modification. There are several ways of finding an optimal solution. The simplest one is a randomized selective (Aguiar et al., 1995). Previous research has indicated that the randomized selective algorithm can be less accurate and more unstable, whereas deterministic algorithms have demonstrated higher accuracy, faster execution times, and simpler implementation (Broadbent et al., 2010). This suggests that deterministic algorithms like DETMAX algorithm might be a more suitable choice for applications that require reliable and efficient performance.

The DETMAX algorithm is often used for non-linear models, where it can be challenging to find optimal design points using other methods. The DETMAX algorithm is

an iterative algorithm used in experimental design to find optimal design points. It works by adjusting the design points in a systematic way to minimize a specified objective function, such as the average variance of the parameter estimates The steps of the DETMAX algorithm are as follows (Mitchell, 2000):

1. Create a list of possible experimental setups that explore the full range of design. These points will form the candidate set (N).
2. Starting with a randomly chosen set of n design points (initial design).
3. Calculate the A-optimal value, trace$(\mathbf{X}^T\mathbf{X})^{-1}$**,** and A-efficiency based on Equation (4).
4. Add or Remove design points at the initial design to minimize the trace.
5. Repeat the process until the trace of the inverse information matrix reaches a minimum or stopping criterion is met. The Algorithm should stop when $A_{eff(i)} > A_{eff(i-1)}$.

3. MATERIAL AND METHOD

3.1. Model Used

An A-optimal design approach was employed to determine the optimal experimental conditions for increasing the silica purity levels. The best design results from the A-optimal design in non-linear models using the DETMAX algorithm with the help of SAS PROC OPTEX (SAS Institute Inc., 2018), we generated an A-optimal design with n experimental runs.

The model in this study is a non-linear model obtained from the relationship between temperature (°C) and rate of temperature rise (°C /min) as factors with the silica purity levels (%) as the response (Rivai et al., 2018). The used factors in this design are the temperature levels at intervals of 800°C to 900 °C and the rate of temperature rise at intervals of 1.67 °C/min to 5 °C/min with the increasing 0.5 °C/min.

The relationship between temperature and the silica purity levels of silica follows an exponential distribution. The non-linear model that is widely used in the study of pharmacokinetics and chemical kinetics is the exponential decay model (Atkinson et al., 2007). Based on the Equation (2), the exponential decay model used in this study is as follows:

$$f(t,r) = [A_0]\{1 - e^{-\theta_1 t + \theta_2 r}\} \quad (5)$$

where $f(t,r)$ = expected value of the response; A_0 = constants; θ_1, θ_2 = parameters; t = temperature; r = rate of temperature rise.

3.2. Steps of Point Exchange Algorithm

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

1. Create N design points (candidate set) from the temperatures levels at intervals 800 °C to 900 °C and the rate of temperature rise at intervals 1.67 °C/min to 5 °C/min with the increasing 0.5 °C/min.
2. Determine n choice sets which are chosen randomly from the candidate set so that the matrix, $\mathbf{X_{(0)}}$, is obtained.
3. Calculate the A-optimal value and the A-efficiency value based on Equation (4).
4. Exchange point on the choice set matrix with the candidate set. This step is called the improvement process. The exchanges process uses the DETMAX algorithm, which adds and reduces a current choice set from the candidate set.
5. Repeat steps 2, 3, and 4 by the iterating process until the design point is selected with the best A-efficiency value (the highest value), when $A_{eff(i)} > A_{eff(i-1)}$.

4. RESULTS AND DISCUSSION

4.1 Non-linear Model Approach

The Taylor method is a useful technique for simplifying non-linear equations by approximating them with linear equations. The approach adopted here focuses on finding an initial solution that minimizes error. Specifically, this method applies the Taylor expansion up to the nth-order polynomial equation. The selection of the appropriate order in this study was guided by the Mean Square Error (MSE) values, with the order that resulted in the smallest MSE being chosen. Table 1 summarizes the selected orders and their corresponding MSE values.

Table 1. MSE Values of the Silicon Dioxide Purity Level

Taylor Polynomial	MSE
First Order	7.90E-08
Second Order	5.08E-10
Third Order	1.95E-12

As shown in Table 1, the second-order Taylor polynomial was chosen for this study because it provides a lower MSE (5.08E-10) compared to the first-order polynomial, while maintaining a less complex model than the third-order alternative. This balance between model simplicity and prediction accuracy makes the second-order approach more suitable for the analysis of the Silicon Dioxide purity level. Using Equation (3) with θ_1= 0,005 and θ_2= 0,005, the second-order Taylor approximation employed in this study is detailed as follows:

$$\begin{aligned} f(t,r) = {} & 0.79413127 + 0.00037952t - 0.00037952r \\ & -1.813010525\ 10^{-7}t^2 + 3.6260211\ 10^{-7}tr \\ & -1.8130105\ 10^{-7}r^2 \end{aligned} \tag{6}$$

4.2 A-Optimal Design on Silicon Dioxide Purity Levels

A-optimal design was used to find optimal design points from the predictor variables, namely temperature, and rate of temperature rise on the response variable of the silicon dioxide purity levels. In this study, three alternatives will be carried out. The first alternative takes six design points, the second alternative takes nine design points, and the third alternative takes 12 design points. The efficiency values of the three options can be seen in Table 2.

Table 2. A-Efficiency Value Using DETMAX Algorithm

Design Number	A-efficiency		
	Alternative 1	Alternative 2	Alternative 3
1	24.9383	31.1997	32.7587
2	24.8916	31.1997	32.7587
3	24.8860	31.1997	32.7587
4	24.8860	31.1997	32.7587
5	24.8837	31.1997	32.7587
6	24.8837	31.1997	32.7587
7	24.8782	31.1997	32.7587
8	24.8616	31.1997	32.7587
9	24.8576	31.1997	32.7587
10	24.8497	31.1997	32.7587

Table 2 shows alternative 2 and alternative 3 have consistently higher A-efficiency values than alternative 1 across all design numbers. Alternative 3 consistently achieves the

highest A-efficiency values across all experimental runs. This indicates that Alternative 3 is the most efficient design for estimating the model parameters in this study. The best A-efficiency value of each alternative is shown in design number 1. The best design points of each choice are presented in Table 3, Table 4, and Table 5.

Table 3. A-Optimal Design with 6 Design Points

No	Temperature	Rate
1	800	1.67
2	800	4.17
3	850	3.00
4	850	5.00
5	900	1.67
6	900	4.17

Table 4. A-Optimal Design with 9 Design Points

No	Temperature	Rate
1	800	1.67
2	800	3.50
3	800	5.00
4	850	3.00
5	850	3.17
6	850	5.00
7	900	1.67
8	900	3.5
9	900	5.00

Table 5. A-Optimal Design with 12 Design Points

No	Temperature	Rate
1	800	1.67
2	800	3.17
3	800	5.00
4	850	1.67
5	850	1.67
6	850	3.34
7	850	3.34
8	850	3.34
9	850	5.00
10	900	1.67
11	900	3.17
12	900	5.00

Table 3, Table 4, and Table 5 present the optimal design points for a process based on the A-optimal criterion. The three tables correspond to different numbers of design points: 6, 9, and 12. The A-optimal design with six design points at Table 3 offers a balance between efficiency and experimental cost. However, if higher precision is desired, the A-optimal design with nine design points at Table 4 or twelve design points at Table 5 can be considered. The pattern of best A-optimal designs as shown in Figure 1.

The A-optimal designs in Figure 1, for alternative 1, 2, and 3 are categorized into three temperature groups: minimum, middle, and maximum. In all three alternatives, the first group consists with the minimum temperature 800°C, the second group consists with the middle temperature 850°C, and the third group consists with the minimum temperature

900°C. In alternatives 1 and 2, there were no repeated design points. However, in Alternative 3, there was a repetition of points at a temperature of 850°C with the rate of temperature rise of 1.67°C/min and 3.34°C/min. Based on Figure 1, each design suggests that as the number of design points increases, the frequency of repeated points within the design area also increases.

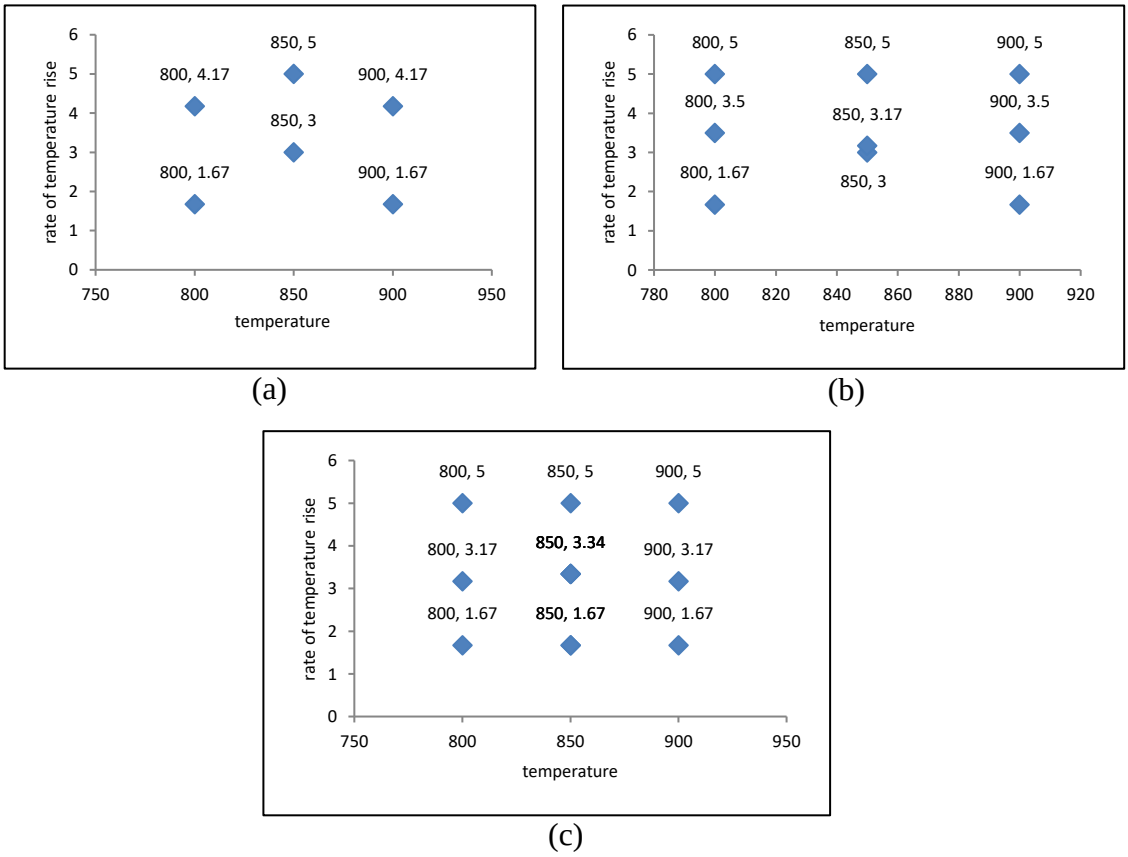

Figure 1. Design points for: (a) Alternative 1, (b) Alternative 2, (c) Alternative 3

A quadratic pattern of design points is often used in experiments to capture non-linear relationships between variables. In this case, the quadratic pattern in Alternative 3 ensures that the design points are distributed across the entire temperature range, allowing for a more accurate estimation of the model parameters and a better understanding of the non-linear relationship between temperature and rate of temperature rise.

The A-efficiency of each design points will be compared with those from previous studies. The first study yielded an A-efficiency of 21.98872 (Wulandari et al., 2023), while the second study obtained an A-efficiency of 18.52071 (Aliu et al., 2024). Based on the results, the design points obtained in this research have an A-efficiency of 32.7587, which is more powerful than those previous studies.

5. CONCLUSION

This study obtained the results using A-optimal design with DETMAX algorithm, which were significantly better than previous studies. The A-optimal design with an A-efficiency of 32.7587 was selected for this study, as it outperformed other designs. While the six point design offers a good balance between efficiency and cost, the twelve point design is better for achieving the optimal accuracy.

REFERENCES

Adli, M. Z., Sari, Y. W., & Irzaman. (2018). Extraction Silicon Dioxide (SiO2) from Charcoal of Baggase (Saccharum officinarum L). *IOP Conference Series: Earth and Environmental Science*, *187*(1). https://doi.org/10.1088/1755-1315/187/1/012004

Aguiar, P. F., Bourguignon, B., Khots, M. S., Massart, D. L., & Phan-Than-Luu, R. (1995). D-optimal designs. *Chemometrics and Intelligent Laboratory Systems*, *30*(2), 199–210. https://doi.org/10.1016/0169-7439(94)00076-X

Aliu, M. A., Syafitri, U. D., & Fitrianto, A. (2024). *The Comparison A-Optimal and I-Optimal Design in Non-Linear Models to Increase Purity Levels Silicon Dioxide The Comparison A-Optimal and I-Optimal Design in Non-Linear Models*. *6*(2), 187–194. https://doi.org/https://doi.org/10.37905/jjom.v6i2.26253

Aminullah, Rohaeti, E., Yuliarto, B., & Irzaman. (2018). Reduction of Silicon Dioxide from Bamboo Leaves and Its Analysis Using Energy Dispersive X-Ray and Fourier Transform-Infrared. *IOP Conference Series: Earth and Environmental Science*, *209*(1). https://doi.org/10.1088/1755-1315/209/1/012048

Arul Prishya, A. S., Chopra, L., & Manikanika. (2023). Comprehensive Review on Uses of Silicon Dioxide in Solar Cell. *Materials Today: Proceedings*, *72*, 1471–1478. https://doi.org/10.1016/j.matpr.2022.09.348

Atkinson, A. C., Donev, A. N., & Tobias, R. D. (2007). *Optimum Experimental Design, with SAS*. Oxford: Oxford University Press.

Broadbent, M., Brown, M., & Penner, K. (2010). Subset Selection Algorithms: Randomized vs. Deterministic. *SIAM Undergraduate Research Online*, *3*(5), 50–71. https://doi.org/10.1137/09s010435

Casnan, Noor, E., Hardjomidjojo, H., Irzaman, & Rohaeti, E. (2019). Scaling Up of the Pirolysis Procces to Produce Silica from Rice Husk. *Journal of Engineering and Technological Sciences*, *51*(6), 747–761. https://doi.org/10.5614/j.eng.technol.sci.2019.51.6.1

Ghosh, A., Johnson, J. E., Nuss, J. G., Stark, B. A., Hawkins, A. R., Tolley, L. T., Iverson, B. D., Tolley, H. D., & Lee, M. L. (2017). Extending the Upper Temperature Range of Gas Chromatography with All-Silicon Microchip Columns using a Heater/Clamp Assembly. *Journal of Chromatography A*, *1517*, 134–141. https://doi.org/10.1016/j.chroma.2017.08.036

Guo, W., Li, G., Zheng, Y., & Li, K. (2021). Nano-silica Extracted from Rice Husk and its Application in Acetic Acid Steam Reforming. *RSC Advances*, *11*(55), 34915–34922. https://doi.org/10.1039/d1ra05255a

Har, N. P., Irzaman, & Irmansyah. (2019). Crystallinity and Electrical Properties of Silicon Dioxide (SiO2) from Rice Straw. *AIP Conference Proceedings*, *2202*, 020–028. https://doi.org/10.1063/1.5141641

Jones, B., Allen-Moyer, K., & Goos, P. (2020). A-optimal versus D-optimal Design of Screening Experiments. *Journal of Quality Technology*, *53*(4), 369–382. https://doi.org/10.1080/00224065.2020.1757391

Kouki, R., & Griffiths, B. J. (2019). Introducing Taylor Series and Local Approximations using a Historical and Semiotic Approach. *International Electronic Journal of Mathematics Education*, *15*(2), 1–8. https://doi.org/10.29333/iejme/6293

Mitchell, T. J. (2000). An algorithm for the Construction of "D-Optimal" Experimental Designs. *Technometrics*, *42*(1), 48–54. https://doi.org/10.1080/00401706.2000.10485978

Mohd Kamal, N. L., Beddu, S., Nuruddin, M. F., Shafiq, N., & Muda, Z. C. (2014). Microwave Incinerated Rice Husk Ash (MIRHA) and Used Engine Oil (UEO): Towards Sustainable Concrete Production. *Applied Mechanics and Materials*, *567*, 434–439. https://doi.org/10.4028/www.scientific.net/AMM.567.434

Montgomery, D. C. (2020). Design and Analysis of Experiments, 10th Edition, Wiley. In *Wiley* (pp. 1–682).

Rivai, M., Sartono, B., Erfiani, & Irzaman. (2018). G Optimal Design in Non linear Models to Increase Silicon Oxide Purity Levels and Electrical Conductivity. *International Journal of Scientific Research in Science, Engineering and Technology*, *4*(11), 150–155. https://doi.org/10.32628/ijsrset21841110

Sa'diyah, H., Nurhimawan, S., Fatoni, S. A., Irmansyah, I., & Irzaman, I. (2016). *Ektraksi Silikon Dioksida Dari Daun Bambu*. *V*, SNF2016-BMP-13-SNF2016-BMP-16. https://doi.org/10.21009/0305020303

SAS Institute Inc. (2018). *SAS/QC ® 15.1 User's Guide The OPTEX Procedure*. SAS Institute Inc.

Sintha, I., Dahrul, M., Ismawati, S. S., Kurniati, M., Irmansyah, & Irzaman. (2017). Optimalization of Silicon Extraction from Husk Ashes by Excessive Magnesium Addition on Increasing Rate of Temperature Reduction. *IOP Conference Series: Earth and Environmental Science*, *65*(1). https://doi.org/10.1088/1755-1315/65/1/012032

Wulandari, N., Erfiani, E., Irzaman, I., & Syafitri, U. D. (2023). G-Optimal Design of Non-Linear Model to Increase Purity Levels of Silicon Dioxide. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *17*(2), 0659–0666. https://doi.org/10.30598/barekengvol17iss2pp0659-0666