Quasi-Experimental based on Cross-Section Historical Data: An Initial Alternative Framework for Analyzing Causality in Manufacturing Process

Mochammad Arbi Hadiyat[1,2,*], Bertha Maya Sopha[1], Budhi Sholeh Wibowo[1]

[1] Industrial Engineering Program, Department of Industrial and Mechanical Engineering, Universitas Gadjah Mada, Indonesia; arbi@mail.ugm.ac.id, bertha_sopha@ugm.ac.id, budhi.sholehwibowo@ugm.ac.id

[2] Industrial Engineering Program, Universitas Surabaya, Indonesia; arbi@staff.ubaya.ac.id

* *Corresponding author*

Abstract

The classic design of experiment (DoE) requires randomization within the treatments to ensure statistical independence and fulfill mathematical assumptions for causality investigation purposes. For some cases with difficulties in randomizing the treatment, a quasi-DoE becomes an alternative with all its weaknesses. Meanwhile, a non-random cross-section historical data from such a smart manufacturing should be considered as one of the available resources, and there are big challenges to retrieving hidden information within. This paper proposes an alternative framework to select observations from historical data and treat it as a quasi-experimental that meets a type of classic DoE, followed by performing statistical analysis to build an evaluation and interpretation. As an initial validation of this framework, three factors historical data from a CNC milling process were recorded for a case study. The statistical analysis is successfully conducted by selecting an observational subset that matched a DoE design with satisfying its properties. It has been concluded that selecting the desired observations subset gives a similar interpretation to a classic DoE.

Keywords: Classic DoE, historical data, quasi-experiment, subset selection

Body of Knowledge: *Design & Manufacturing Engineering*

How to Cite:
Hadiyat, M. A., Sopha, B. M., & Wibowo, B. S. (2023). Quasi-Experimental based on Cross-Section Historical Data: An Initial Alternative Framework for Analyzing Causality in Manufacturing Process. *Prosiding Seminar Nasional Teknik Industri* (SENASTI) 2023, xxx-xxx.

1. Introduction

The classic DoE becomes the golden standard in investigating causalities among factors and their response in manufacturing (Montgomery, 2017). DoE ensures independence among parameter estimation in the model, accommodating interaction between factors, involves the defined covariates, and reduces the potential confounding (Voss et al., 2017). The existence of orthogonality in the DoE design also removes the dependencies between the factor level combination, so each included factor will perform its effect individually without any other factor influences unless for pre-defined inter-factor interaction (Roy, 2010). Moreover, since the classic DoE requires full randomization during the experiment, the treatments run as if they were a sampling from a certain population; thus, the inference and conclusion from DoE model satisfy the statistical generalization concept (see Tipton, 2014)

However, fulfilling all the classic DoE requirements becomes difficult for such a continuous manufacturing process. This kind of process has been running for a certain time before, and there is very little chance of interrupting the running process to accommodate extreme setting adjustment for classic DoE experiments (Tanco et al., 2009). No such factor setting adjustments were conducted except if some non-conform products were found, and the user needed to re-adjust the machine setting to optimize it. This situation obstructs the application of classic DoE ideally (see Tanco et al., 2010), but usually, the user still needs information on investigating factor-response causality, especially for optimization purposes. Nevertheless, performing a classic DoE experiment should be the main option for the user to analyze such a causality; and there are thousands of DoE successes in helping the user to investigate the influencing factors of a process or machine.

The following papers show that there is another option besides implementing DoE, i.e., performing DoE-like analysis based on historical data (see Loy et al., 2002). This option arises firstly in optimizing a manufacturing process, followed by (Sukthomya and Tannock, 2005) and (Chien et al., 2014). Some years before, a formal procedure in adopting historical data for DoE-like analysis was proposed by (Shainin and Shainin, 1988), who successfully implemented a DoE concept based on provided manufacturing data. It integrates many statistical tools, but some papers give critiques because of the lack of scientific bases (Tanco et al., 2008). Another approach in investigating the causality model is regression model and analysis based on provided historical data (see Draper and Smith, 1998); thus, since regression model uses all observational data without having full control of treatment, then the causality model gives weaker interpretation compared to DoE.

This paper leads to a hypothesis that already provided historical data becomes an alternative to conducting classic DoE experiments. Of course, the historical data should contain information supporting causality model building, such as cross-sectional continuous data type, recorded factor, and response (Hadiyat et al., 2022). The main purpose of proposing this new framework is to apply historical or observational data or non-designed experiment data as an alternative to conducting a designed DoE, maximizing information gained from data as if it is from a quasi-experimental design.

Only few papers develop the multifactor or factorial design in engineering topics that implement the type of quasi-experiment, and almost all of the research covers social topics (see McKinley and Rose, 2019). As a fully controlled experiment, the classic DoE gives strong internal validity to study the effects of factors, ignoring the external validity that has been limited to include with the treatment. Meanwhile, as shown in **Figure 1**, the quasi-experiment provides higher external validity; instead of systems created or modified for the sake of study, the interventions being evaluated in quasi-experiments have been executed using real-life systems (see Geldsetzer and Fawzi, 2017). This paper proposes an alternative in implementing a quasi-DoE based on non-randomized and non-experimental historical data to gain more information within it rather than conducting new costly experimentation.

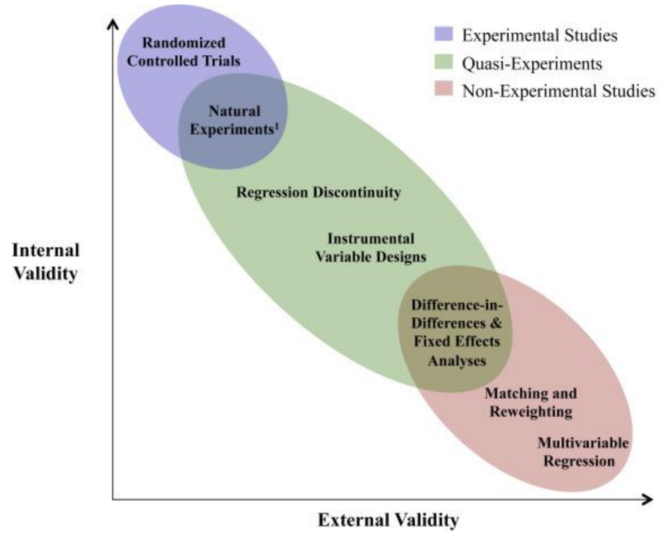


Figure 1. Trade off between internal and external validity (adopted from Geldsetzer and Fawzi, 2017)

The concept of instance selection starts with the need to select the most informative data to include in the machine learning model to improve its prediction accuracy (see García et al., 2015). The keyword in is gaining useful information from data and ignoring others that disturb the modelling process. The selection process runs iteratively based on certain criteria that optimize the information gained. In this paper, the instance selection concept is adopted to select the most appropriate observation within historical data that satisfies some properties of the classic DoE experiment, and the selected observation becomes similar to a quasi-DoE.

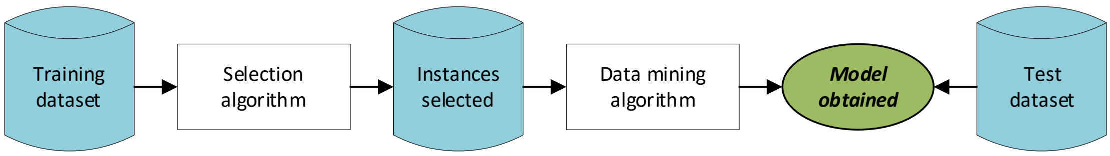


Figure 2. Instance selection process (adopted from García et al., 2015)

2. Method

This research focuses on developing an initial new framework for adopting observational data for quasi-DoE purposes. The stages start with identifying a classic DoE's main properties as the basis for how the historical data should be adopted as a quasi-DoE, followed by developing a framework to adopt the historical data as a quasi-DoE. Implementing the proposed framework also becomes part of this paper (Figure 3) and is completed with performance evaluation. As tools-development research, the methodology in this paper was designed to consider some alternatives to be adopted in building a quasi-DoE approach based on historical data. The main idea of this methodology covers how the provided historical data will be treated to meet or satisfy a condition as if it is a designed DoE experiment. Although not all of DoE properties will be satisfied, the aim of avoiding high-cost experiments will at least be achieved; of course, it will not be as ideal as conducting real experiments. As an implementation, historical data from the CNC milling process becomes a study case and will be compared with a real experiment based on DoE. The spread of experiment points is then evaluated for its distribution among the space of available factor operating conditions.

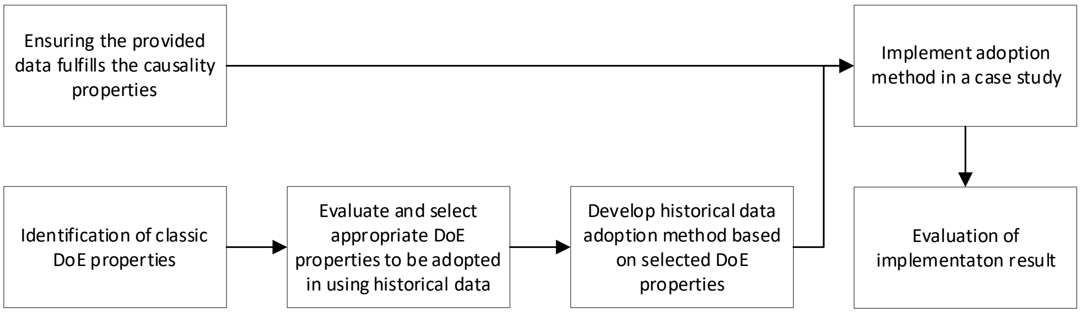


Figure 3. Methodology in developing quasi-DoE

3. Results and discussion

The classic DoE has some properties in its implementation (see Table 1), which become the basis for adopting historical data. Among these properties, the most adaptable one is the orthogonality among factors. The possibility of adopting DoE properties is only satisfied by the orthogonality and the spread of experiment points. Since not all DoE properties cannot be adopted, the historical data is impossible to replace real experiment data; that is why adopting historical data for DoE becomes one of the quasi-DoE techniques.

Table 1. Opportunities for adopting classic DoE properties in historical data

Main DoE properties	purpose	Adoptable opportunity in quasi-DoE	Adopted in quasi-DoE
Cross-sectional data involving factors and response (independent and (dependent vars)	Capture the influence of factors on the response	High possibility for cross-section historical data	Yes
randomization	Ensure no intervention or pattern in the response. It refers to the concept of a random sample in common statistics.	Not possible because historical data is observed without following any pre-determined treatments.	No
orthogonality	Ensure the independence between factors	High possibility to adopt. Historical data has the potential to be independent of factors	yes
Full control of factor level	Ensure the influence of a factor on the response is treated by changing its level/setting intentionally across the experimental area (operating condition)	Not possible; this is the main difference between classic DoE and Quasi-DoE, where the user uses already provided data instead of experimenting and interrupting the ongoing production process.	No
The experiment points representing the treatments, spread along/across the experimental area (operating condition of each factor)	Accommodate any possible combination factor levels as treatments to investigate their influence on the response	High possibility if the historical data contains and observes many combinations of factor level	yes
Statistical analysis using ANOVA technique	Determine which factors are statistically significant in influencing the response	High possibility, since the observation numbers exceed the degrees of freedom of ANOVA term analysis, standard ANOVA could be implemented	yes

The main idea is that quasi-DoE is an alternative, not a replacement to the classic DoE, since there are limitations to conducting real experiments, but relevant historical data was provided. In other words, quasi-DoE wants to optimize the information gathering from historical data, so a quasi-DoE will have properties similar to a classic DoE. Historical data contains many level combinations of each factor, and it is different from classic DoE in terms of a balanced combination of factor levels as in common factorial design in classic DoE. Then, the challenge in adopting historical data for quasi-DoE becomes three options (see Table 2): adopting all the observations, selecting a subset of them, or modifying the linear model to capture the causality of factor and response. Based on Table 1 and **Error! Not a valid bookmark self-reference.**, a new framework for quasi-DoE is then developed according to those potential opportunities. Moreover, since the main idea in this paper focuses on adopting historical data for quasi-DoE, the developed framework also focuses on this. Considering that satisfying DoE properties is the main priority in this paper, the most reasonable choice is selecting a subset of historical observation and treating it as if it is a DoE and fulfils the requirements to be categorized as quasi-DoE.

Since there are many combinations of factor levels, a strategy to select the best observation subset is proposed based on maximizing the similarity of it to a classic DoE. In other words, the selected subset should be as close as possible to a DoE design by calculating orthogonality and points spread (variance) within the available operating condition (experiment area); see formulas (1) And (2)

Table 2. Strategies for adopting historical data for quasi-DoE

Historical data Adopting strategy	Scientific basis	weakness
Use all observation	Treat all observations as in regression analysis to capture the relationship between factor and response.	Regression is not designed as a DoE but is a fully happenstance observations-based model.
Select observation subset	Finding a DoE-like observation to imitate a type of classic DoE and treat it as a quasi-DoE with less orthogonality as in a D-optimal DoE design	Limitations of factor level combination within historical data make it difficult to find a subset that imitates a DoE.
Modify the causality model.	A linear model can be estimated to capture causality as in RSM analysis (advanced DoE). Alternatively, many advanced models were adopted, such as machine learning and nonparametric model approach.	Balanced Anova cannot be calculated, so the analysis should adopt a more complex one, such as a generalized linear model and another advanced model.

$$VIF = (X'X)^{-1}_{jj} \quad (1)$$

$$Var = \frac{\sum(Y_i - \bar{y})}{n} \quad (2)$$

Subset selection works by iteratively select observations that optimize these three criteria using multiobjective optimization technique with the same weight scalarization (Collette and Siarry, 2004).

$$\textit{Minimize}\ \ (\det(X'X)^{-1},\ \ \textstyle\sum_i^k VIF_i\ ,\ \ -\sum_i^k Var_i)$$
subject to each factor level available experiment area

An algorithm to select this subset is also developed based on these criteria. Assuming a selected observation by 1 (one) and an unselected observation by 0 (zero), combining selected and unselected observations forms a sequence similar to a DNA sequence in the Genetic Algorithm. Also, based on the concept of instance selection in data mining topics, this algorithm aims to optimize information gathering from complete historical data and ignore uninformative points. Thus, modification of this algorithm for selecting a subset is shown as in Algorithm 1; based on it, a framework for implementing quasi-DoE considers the iterative steps in finding the subset by maximizing three criteria in **Table *3***. The biggest constraint (as in Table 1) is the smaller number of factor level combinations within the historical data that obstruct the algorithm from finding the widest spread of data points, see Figure *4*.

Table 3. Evaluation of selected subset

Criteria	Purpose	Direction of optimization	Reference
Determinant of $(X'X)^{-1}$	Ensure orthogonality	Minimized	(Goos and Jones, 2011)
Variance Inflation Factor (VIF) of each factor	Ensure orthogonality	Minimized	(Hair et al., 2009)
factor level variance (Var) for each factor	Spread the experiment points within the experiment area	Maximized (Minimized the minus function)	(Montgomery, 2017)

Algorithm 1. Genetic algorithm pseudo code for subset selection

```
DEFINE
  Data X : contains factors and levels,  N : represents number of all observation
  n : number of selected observations for the subset, n<N (considering degrees of freedom)
  specify model term (linear, quadratic, interaction)
INITIALIZE
  Genetic algorithm properties (number of population, parents, offspring, mutation rate)
  Gene code 1: selected for subset, code 0: not selected for subset
  Generate initial population chromosomes represents selected observations for subset
  Involve the model term in the subset
WHILE termination criteria is not satisfied
  SELECT parents chromosomes from population
  CROSSOVER pairs of parents chromosomes to produce offsprings
  GENERATE MUTATION chromosomes from population
  COMBINE the offsprings with mutated chromosomes
  EVALUATING the fitness function (refer to criteria in Table 3)
  SELECT best chromosomes for next parents generation
ENDWHILE
```

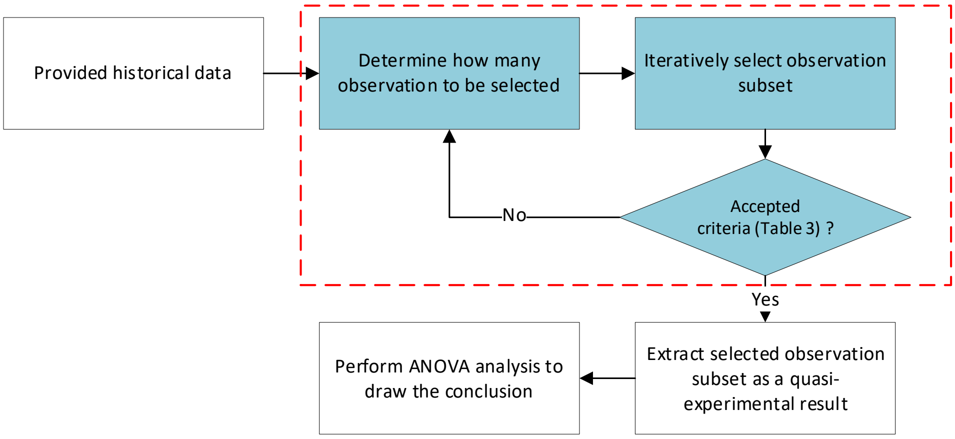
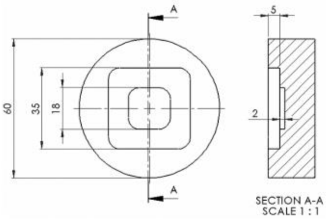


Figure 4. Developed framework for quasi-DoE

In order to implement the proposed frameworks, historical data from a CNC milling process was provided; it contains three factors and a single response. The data provides written records based on operator experience and judgment ini the CnC process to find the best but not optimal setting of factor levels without such designed experiment or scientific consideration in changing the factor setting. Any setting changes follow the operator's intuition in getting suitable equipment settings. As a comparison, the second data were provided by a classic DoE to investigate the influence factors to the response scientifically. The Algorithm 1 selects the subset as quasi-DoE from observational data and then compares the result with a real DoE experiment.

Table 4 case study for quasi-DoE (using CnC milling roland modela MDX-40)

Equipment/ machine and Specimen (in mm)			
Workpiece material		Polycarbonate thermoplastic	
Historical data	Number of obs	53 observations, no DoE design	
	Factors (with random level)	X1: Feed rate (mm/s)	
		X2: Step over (mm)	
		X3: Depth of Cut (mm)	
	Response	Y: Material Removal Rate	
DoE experiments	Number of experiments run	20 data from DoE experiment (CCD: Central Composite Design)	
	Factors (with pre-determined levels)	X1: Feed rate (mm/s)	Low: 0,1, middle: 0.235, high: 0.37
		X2: Step over (mm)	Low: 12, middle: 14.5, high: 17
		X3: Depth of Cut (mm)	Low: 0,3, middle: 0.65, high: 1
	Response	Y: Material Removal Rate, calculated using equation (3)	

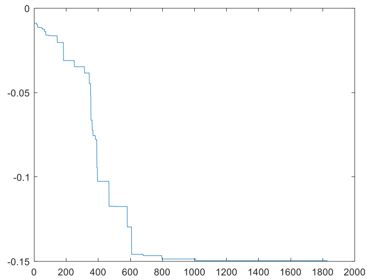


Figure 5. Iteration of algorithm 1, convergence is reached after 1000 epoch

A graphic from the complete observation is shown in **Error! Reference source not found.**. It Looks like all the points spread within the operating condition, but the orthogonality cannot be accepted since the designed DoE did not produce this data. A small correlation exists among the factors by looking at the linear pattern of spreading points; of course, it won't satisfy the DoE properties. Thus, Algorithm 1 should be implemented in order to select a subset that fulfils orthogonality, involving potential model terms (quadratic and/or interaction) to accommodate any second-order influence of factors (quadratic and interaction).

$$MRR = \frac{initial\ weight\ of\ workpiece - final\ weight\ of\ workpiece}{timing\ of\ machining} \quad (3)$$

Based on Algorithm 1, an observation subset is found after more than 500 iterations (see Figure 5). **Figure 6** shows the result of subset selection by algorithm 1; 20 observations from complete data (**Figure 6Error! Reference source not found.**) have been selected with satisfying orthogonality and high variance of points spreading. It means that the algorithm has successfully treated the historical data into a quasi-DoE that has similarity with real DoE experiments in the absence of randomization and balanced treatments.

Table 5. ANOVA comparison between subset data and classic DoE experiment.

Source	Degrees of freedom	ANOVA for subset data				ANOVA for classic DoE experiment			
		Sums of Square	Mean Square	F-Value	P-Value	Sums of Square	Mean Square	F-Value	P-Value
Regression	9	14.295	1.588	41.680	0.000	13.541	1.505	128.310	0.000
X1	1	0.091	0.091	2.390	0.154	0.243	0.243	20.720	0.001
X2	1	4.603	4.603	120.800	0.000	5.541	5.541	472.580	0.000
X3	1	1.357	1.357	35.600	0.000	6.587	6.587	561.720	0.000
$X1^2$	1	0.005	0.005	0.130	0.730	0.000	0.000	0.000	0.953
$X2^2$	1	0.002	0.002	0.050	0.822	0.005	0.005	0.440	0.523
$X3^2$	1	7.579	7.579	198.900	0.000	0.000	0.000	0.020	0.891
X1 X2	1	0.206	0.206	5.390	0.043	0.036	0.035	3.020	0.113
X1 X3	1	0.113	0.113	2.970	0.115	0.041	0.041	3.520	0.090
X2 X3	1	0.100	0.100	2.630	0.136	1.088	1.088	92.760	0.000
Error	10	0.381	MSE = 0.03811			0.117	MSE = 0.01173		
Total	19	14.676				13.659			

Note: green highlight shows the same significant/nonsignificant factors between both

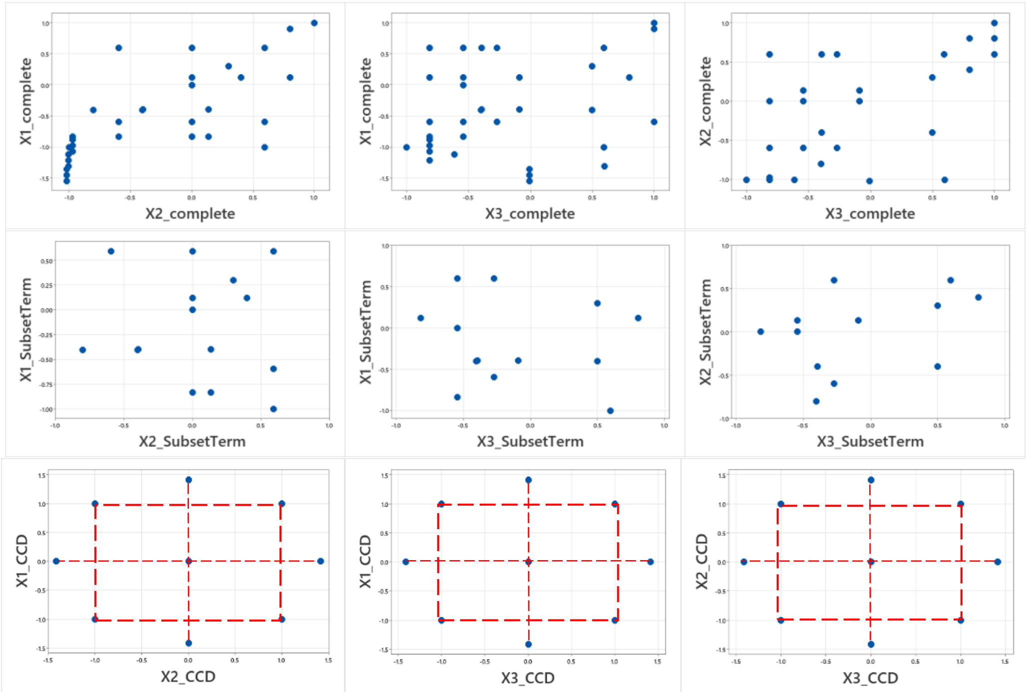


Figure 6. Data points from (a) complete data; (b) selected subset; (c) classic DoE experiment

As a comparison, a classic DoE experiment was also conducted with the same factors and response, and the points are shown in Figure 6. Of course, the real DoE experiment perfectly provides orthogonal conditions, high points spread, and is full of randomization and control. For this comparison, an ANOVA analysis is calculated for both the selected subset and classicDoE (**Table 5**). With the same factors and response as mentioned in Figure 4, though there are differences in the number of significant factors, and the classic DoE gives smaller MSE, there is still a similar result for certain factors. This result shows that the main effect terms of X_2, X_3 remain significant for both subset and DoE. It means there are potential methods to use historical data as an alternative to classic DoE experiments. For a more detailed comparison, the performance of both data is shown in
Table **6**

Table 6. Performance comparison between subset data and classic DoE experiment.

Properties	Real DoE	subset
Orthogonality	Perfect orthogonality	Moderate orthogonality
Determinant of $(X'X)^{-1}$	0.0000187	0.0005337
VIF (sum of)	9.06	15.51
factor level variance (Var)	0.01173	0.03811
Number of associated significant/ not-significant factors	5 associated factors from 9	
Experiment points and level	Fully controlled	Depending on the variation of historical data factor level
Randomization (measured using Runs Test)	Fully randomized, with Run Test Statistics 10.1	Not random, with Run Test Statistics 11.0

Finally, the developed algorithm should become part of the initial procedure/framework in adopting historical data as alternatives to DoE, or it can be assumed as quasi-DoE since some classic DoE properties cannot be satisfied. Thus, the opportunities for improving the proposed framework are shown in Table 7, with some limitations or constraints.

Table 7. Opportunities for quasi-DoE improvement

Frameworks weaknesses	Opportunity for improvement	Barriers
Observation points were not accommodated in all experiment areas	Historical data recording should involve all interested area	The data recording process cannot be fully controlled and depends on operator judgments
Cannot satisfy perfect DoE criteria	Develop an algorithm to find the most criteria-satisfying subset or initiate additional real experiment points to improve orthogonality.	The algorithm only moves within the recorded point.
Lack of randomization	Develop a new approach for quasi-DoE	Mathematical proving of Quasi DoE associated with a classic DoE

4. Conclusion

Classic DoE with real experiments still becomes the first choice with all the robustness of its analysis. However, for the provided historical data, there is an opportunity to use it as a less-informative alternative, not replacement, rather than conducting costly experiments; this becomes a type of quasi-DoE. The proposed framework starts with identifying classic DoE properties that can be adopted to the historical data, continuing by iteratively selecting a subset of observations using a genetic algorithm that satisfies certain DoE criteria, such as maximizing orthogonality and point spreading along the experiment area.

Historical data from a milling CNC process was recorded as a case study, and the result was then compared with real experiments using classic DoE. Although the comparison results show differences in both data regarding orthogonality, randomization, and point spreading, information of similar significance factors still provides useful information based on this historical data. Of course, some procedure improvements should be addressed to increase its performance.

References

Chien, C. F., Chang, K. H., & Wang, W. C. (2014). An empirical study of design-of-experiment data mining for yield-loss diagnosis for semiconductor manufacturing. *Journal of Intelligent Manufacturing*, *25*(5), 961–972.

Collette, Y., & Siarry, P. (2004). *Multiobjective Optimization: Principles and Case Studies*. Springer Berlin Heidelberg.

Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). John Wiley & Sons.

García, S., Luengo, J., & Herrera, F. (2015). Data Preprocessing in Data Mining. In *Intelligent Systems Reference Library* (Vol. 72). Springer International Publishing.

Geldsetzer, P., & Fawzi, W. (2017). Quasi-experimental study designs series—paper 2: complementary approaches to advancing global health knowledge. In *Journal of Clinical Epidemiology* (Vol. 89). Elsevier Inc.

Goos, P., & Jones, B. (2011). Optimal Design of Experiments. In *Optimal Design of Experiments*. Wiley.

Hadiyat, M. A., Sopha, B. M., & Wibowo, B. S. (2022). Response Surface Methodology Using

Observational Data: A Systematic Literature Review. *Applied Sciences (Switzerland)*, *12*(20).
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysis* (7th ed.). Pearson.
Liu, H., & Motoda, H. (2001). *Instance Selection and Construction for Data Mining* (H. Liu & H. Motorda (eds.)). Springer Science.
Loy, C., Goh, T. N., & Xie, M. (2002). Retrospective factorial fitting and reverse design of experiments. *Total Quality Management*, *13*(5), 589–602.
McKinley, J., & Rose, H. (2019). *The Routledge Handbook of Research Methods in Applied Linguistics*. Routledge.
Montgomery, D. C. (2017). Design and Analysis of Experiments. In *Design and Analysis of Experiments* (9th ed.). Wiley.
Roy, R. K. (2010). *A Primer on The Taguchi Method*. Society of Manufacturing Engineers (SME).
Shainin, D., & Shainin, P. (1988). Better than Taguchi orthogonal tables. *Quality and Reliability Engineering International*, *4*(2), 143–149.
Sukthomya, W., & Tannock, J. D. T. (2005). Taguchi experimental design for manufacturing process optimisation using historical data and a neural network process model. *International Journal of Quality and Reliability Management*, *22*(5), 485–502.
Tanco, M., Viles, E., Ilzarbe, L., & Alvarez, M. J. (2009). Barriers faced by engineers when applying design of experiments. *TQM Journal*, *21*(6), 565–575.
Tanco, M., Viles, E., Jesus Álvarez, M., & Ilzarbe, L. (2010). Why is not design of experiments widely used by engineers in Europe? *Journal of Applied Statistics*, *37*(12), 1961–1977.
Tanco, M., Viles, E., & Pozueta, L. (2008). Are All Designs of Experiments Approaches Suitable for Your Company? *Lecture Notes in Engineering and Computer Science*, *2171*(1), 1202–1207.
Tipton, E. (2014). How Generalizable Is Your Experiment? An Index for Comparing Experimental Samples and Populations. *Journal of Educational and Behavioral Statistics*, *39*(6), 478–501.
Voss, D., Dean, A., & Draguljic, D. (2017). Design and Analysis of Experiments Springer Texts in Statistics. In *Design and Analysis of Experiments Springer Texts in Statistics*.