



received: 10 February 2021
accepted: 1 August 2021

pages: 50-65

CRITICAL EVALUATION INTO THE PRACTICAL UTILITY OF THE DESIGN OF EXPERIMENTS

MITHUN SHARMA  SHILPI SHARMA 

ABSTRACT

The research aims to emphasise the relevance of the Design of Experiments (DOE) technique as a reliable method for ensuring efficient use of statistical methods in routine industrial processes. A case study approach with a deductive strategy was used to assess the effectiveness of different DOE methods to achieve the desired objectives. Screening, mid-resolution and high-resolution DOE methods helped identify, characterise, and optimise an experimental variable against the desired output response. A general framework for effective DOE is provided as part of DOE planning, including defining DOE objectives, selection criteria, noise reduction, and application across industries. Overall, various DOE models proved successful in identifying a complicated relationship between experimental variables and output response. However, when ideal DOE models may not be feasible, reducing test run by choosing lower resolution DOE or fewer replicates can still provide important insights into the experimental variables' impact on output responses.

Mithun Sharma

O.P. Jindal Global University, India
ORCID 0000-0002-0108-6331

Corresponding author:
e-mail: mithun.sharma@gmail.com

KEY WORDS

Six-Sigma, Design of Experiments, screening, Taguchi screening, full factorial

Shilpi Sharma

O.P. Jindal Global University, India
ORCID 0000-0003-4316-8551

10.2478/emj-2021-0021

INTRODUCTION

Presently, the world is in a state of pandemonium, disordered and frenzied, which makes the work of scientists and engineers formidable. Evolving new products, advancing old models, managing ongoing

processes, and repairing gadgets are some of the many routine technical tasks for engineers. Daily life is contingent on engineers' real-life technical activities, which are highly complex, multi-staged processes. The engineering field is continuously under

Sharma, M., & Sharma, S. (2021). Critical evaluation into the practical utility of the Design of Experiments. *Engineering Management in Production and Services*, 13(3), 50-65. doi: 10.2478/emj-2021-0021

pressure to re-engineer its process-control mechanisms to keep up with the exponentially increasing advancements in new technology to prevent lags in the process yield quality. To be a frontrunner and beat the business competition, engineers must step up from simple solutions to analysing complex interactions stemming from a thousand different factors. The success of any complex process is contingent on experimental testing, process control, and statistical analysis. Unfortunately, training in advanced statistical techniques is not a regular part of engineering qualification in most countries, which raises serious questions about the accuracy of the analysis most engineers are obliged to perform (Bisgaard, 1991). In this context, the Design of Experiments (DOE) is an efficient tool for applying statistics to routine experimental manufacturing tasks (Davim, 2016; Lye, 2005).

DOE is a toolkit, a set of guidelines that offers several statistical tests for managing resolute modifications to more than one input variable in an engineering process and assessing its impact on the resulting outputs (Montgomery, 20017). It is a cost-effective and time-saving quality improvement technique for managing and optimising processes across various industries. The DOE qualified as a distinguished class of its own and was an independent toolkit during the twentieth century but is now a part of the trendy Six-Sigma package (Brady & Allen, 2006). This paper emphasises the relevance of the DOE technique as a reliable method for ensuring efficient use of statistical methods in routine industrial processes. The following sections of the literature review present more information on the nature, stages, and types of DOE techniques, their benefits over the traditional experimental methods, their relevance to the Six-Sigma methodology, the wide applications of DOE across various industries, and the gaps in its current use. This is followed by results from multiple DOE categories across different sectors to gain practical insights into the process shared in the discussion section of the manuscript.

1. LITERATURE REVIEW

1.1. WHAT IS DOE?

The design of experiments (DOE) technique relies on several statistical methods to optimise the modelling of various factors and response variables simultaneously (Hecht et al., 2016). DOE's origin date

back to 1926, whereby Ronald E. Fisher used it to systematically eliminate environmental biases from a geometrically patterned agricultural field experiment (Fisher, 1926). The principles used in this first design construction have emerged as pillars of statistics. Many of his predecessors defined the importance of replication in understanding variation, including William's student Gosset. Fisher and his contemporaries, such as Stuart Chapin (1950, randomisation) and Wald and Tukey (1943 and 1947, respectively, blocking), elaborated on DOE applications in many other industries. Later in the 21st century, many other statisticians, such as Box et al. (1987) and Montgomery (2017), further strengthened the DOE technique's mathematical foundations, providing a reliable platform to engineers and other practitioners.

The application of DOE models is known to benefit a wide range of industries, such as Aviation (Yondo et al., 2018; Davis et al., 1996), Service (Antony et al., 2020; Setamanit, 2018), Chemical Engineering (Yoo, 2020; Durakovic & Torlak, 2017; Andersson et al., 1996), Engineering (Schlueter & Geyer, 2018), Environment (Okatia et al., 2016; Yang & Tsai, 1998), Food Technology (Yu et al., 2018; Chen, 1996), Manufacturing (Wesling & Emamjomeh, 1994), Materials (Paulo et al., 2017; Bucher & Loos, 1994), Medical (Yip et al., 2020; Mager, 1997), and Microbiology (Puente-Massaguer et al., 2020). Within the field of engineering and science, DOE's applications are frequently visible in the materials sector (30 %), mechanical engineering (18%), chemical engineering (17 %), and the industrial sector (13 %).

1.2. BASIC PRINCIPLES OF DESIGN OF EXPERIMENTS

Every process, manufacturing or service will inevitably experience variation resulting in rejects or failures. These variations could stem from different testing temperatures or even concentrations, types of raw materials, and adjustments of process parameters across different testing sessions. Controlling such variations is expensive and time-consuming, and some processes might be more sensitive than others. Thus, the aim is to identify the sources of such manufacturing variations and control them by setting appropriate process controls, reducing the process sensitivity, and improving output quality (Kackar & Shoemaker, 2021). To do that, an engineer needs three types of information: a) general, measurable metrics that may indicate functional characteristics

of the final product, b) process engineer also needs exact parameters for the process variables so that they can manipulate their standard settings, and c) uncontrollable external environmental sources of noise, such as the factory's humidity and temperature. Thus, the objective is to identify control parameter settings that minimise the noise factor effect on functional characteristics of the final output. The whole step-by-step model is a robust design, first discussed by Taguchi as a quality control strategy (Antony & Roy, 1999). The DOE strategy works on a similar model to identify the underlying causes of process problems. It is a disciplined approach that systematically manipulates settings of the control parameter variables to minimise the noise effect on the functional characteristics of the output (Gardner & Bieker, 2000).

While applying the DOE model in a process, engineers may have to deal with two kinds of process variables that they need to manipulate for reducing noise effects on the final output: quantitative and qualitative. Quantitative settings are measurable goals, controllable, and usually pre-determined across different levels of process settings. The term "level" refers to settings or process specifications for a factor in an experiment. Technical specifications of a particular manufacturing process constitute these variables. On the other hand, qualitative variables are external, disconnected, and may require many more levels than a quantitative factor. Type of supplier, characteristics of a catalyst, and the composition of raw materials are examples of qualitative factors in a manufacturing process. In a standard DOE, an engineer performs a pre-determined number of trials to assess the effects of different variable levels on an outcome variable.

Three frequently used stage-wise processes in experiments to enhance the output efficiency statistically and significantly and minimise bias are: a) randomisation is an essential building block of statistical methods in experimental designs. It is the process of randomly assigning factors to testing units to ensure an equal probability of each factor's allocation to any one test; b) replication enables an experimenter to estimate the experimental error by simple repetition of necessary tests; c) blocking further enhances an experimental design's efficiency by grouping homogeneous experiments and controlling any extraneous variation sources.

In essence, the DOE aims to detect optimum and efficient solutions to process problems through systematic planning, designing, and statistically analysing engineering experiments. The success of an

industrially designed experiment depends on efficient planning, accurate design choice, thorough statistical analysis of the data, and collaborative teamwork. Traditionally, engineers heavily relied on the one-factor-at-a-time (OFAT) technique for their experimental needs. The OFAT technique draws on a null hypothesis that no difference exists between two settings of a manufacturing process parameter, and it sequentially tests all the inherent factors, one at a time.

1.3. OFAT vs DOE

The most significant risk of using the OFAT technique is its failure to identify the real optimum value and ignoring the potential interactions between different factors. Another difficulty in using the OFAT technique is its inability to separate the process's inherent noise from the actual test improvement. Therefore, a viable alternative is the DOE technique, enabling simultaneous combinatorial testing of all factors in a full factorial design model. If there are "k" number of factors and "X" numbers of levels within a testing range, then a DOE strategy will allow for "Xk" number of experiments (Hecht et al., 2021). This strategy produces novel reaction conditions and provides for the systematic elimination of researcher bias. Other noted advantages of DOE over the traditional OFAT technique (Czitrom, 1999): a) it requires fewer resources, such as time, material, and the number of experiments for the amount and quality of results it produces, b) the estimated effects of each factor on the response variables are exact, and it can yield information on a larger region of the factor space. DOE has proven to offer substantial financial benefits by improving a product's yield and quality by its use at three different times in the life of a manufacturing process: a) to determine the optimal testing conditions in the early days of initiating a process, b) to establish the robustness of a methodology for formal applications to regulatory agencies, and c) for refining existing conditions of already known procedures.

The DOE technique further gains a significant edge over OFAT by carrying out multiple testing involving several factors and simultaneously assessing their impact on the product yield and quality. The DOE technique can also isolate the interdependent effects of a combination of factors, if any, by testing the results using appropriate statistical methods. The most commonly used statistical techniques are analysis of variance (ANOVA), half-normal and contour

plots, and response surface methodology. Nowadays, statisticians and technical professionals across many industries acknowledge the practical utility of DOE as an indispensable quality improvement tool. It is also a part of one of the most popular business world philosophies: the Six-Sigma methodology.

1.4. DESIGN OF EXPERIMENTS AS A KEY TOOL OF SIX-SIGMA

Six-Sigma is a quality improvement tool that consists of a systematic and organised set of scientifically proven statistical methods for defect reduction in a product and carrying out strategic improvements in a process thereby improving customer satisfaction. The stellar popularity of the Six-Sigma technique as a quality improvement methodology relies upon its scientifically organised and data-driven statistical approach (Hahn, 2005), which has proven to be a winning combination time and again. A general stepwise sequence of a Six-Sigma practitioner consists of five logical stages of Define, Measure, Analyse, Improve, and Control, popularly known as DMAIC. It starts with delineating the project scope and measurable goals, followed by identifying specific characteristics of the process. The following two stages of Measure and Analyse gauge the process attributes for an efficient implementation during the Improvement phase in a manner that leads to sustainable, long-term benefits, monitored under the final Control stage (Goh, 2002). The most critical and bulky steps for achieving the project's goals are the Analyse and Improve phases, which rely on experimentation, thus creating a perfect niche and fit for the DOE technique. The various stages of a DOE are further elaborated in the next section.

Stage 1. First and foremost, engineers agree on the broad aims, specific objectives, and the input and output variables for the intended experiments. According to Wu and Hamada's classification system, there are three ways of conducting this first stage. The first type is the traditional approach and mainly concerns the screening and characterisation stages in a process. The second type of DOE referred to as the response surface methodology (RSM), is intended to understand process behaviour and find the optimal performance point (Myers et al., 2004). The third type borrows from Taguchi's principles and focuses on detecting those controllable process settings that will effectively minimise the variability induced by noise factors, known as the robust parameter design (RPD) (Robinson et al., 2003).

Stage 2. The next stage delves into the specification of the different levels of the previously selected input variables for the manufacturing experiment. Often engineers rely on their expert knowledge and experience for agreeing on the number of factors and their levels during this stage.

Stage 3. A range of experimental design options is available for use in a DOE engineering model. The most popular (31 %) is Taguchi's orthogonal matrices, consisting of tables to display experiments for specific constituents of factors. The next common choice (16 %) among manufacturing practitioners is full factorial designs that produce an exhaustive list of all possible combinations for every single factor and their level. Another common (14 %) type of experimental design is fractional factorial, i.e., the design, which carefully selects and conducts only fractions of the full factorial design experiments. Finally, central composite designs (9 %) enhance several factorial design experiments with axial and centre points. Other less popular designs are Plackett-Burmann (4 %) and optimal designs (4 %). The Taguchi method's higher popularity amongst engineers is partly due to its solid statistical base that significantly enhances its practicality (Box, 1988; Nair, 1992). However, due to a lack of statistical training amongst most engineers, manufacturing industries still fail to capitalise on the robust methodology and often rely on other designs, such as orthogonal matrices, full or partial factorial designs.

Stage 4. After selecting the number of factors, their levels, and the number of design options, engineers contemplate and set the number of experiments they need. This decision is usually contingent on the available budget, access to resources, and the engineers' expertise combined with the previous variables of the type of design and factors.

Stage 5. The final decision relates to a real or simulated computer experiment. An actual, tangible experimental process using factual data usually occurs in a laboratory or manufacturing plant, and a simulated option can be through computer software. Computer simulated experiments for product and process development have recently gained tremendous popularity (Kennett & Steinberg, 2006). A simulation is cost-effective as it facilitates simultaneous testing of many factors and proves to be especially useful in case an experiment turns out to be defective.

Gaps in the current research. It is noticeable that despite a steep rise in the number of scientific articles using the DOE technique, there is a gap in the effec-

tive use of advanced statistical analyses. For example, between 2003 – 2007, there were about six publications per year, which increased by 145 %, to 14.7 publications/year for the next five-year period of 2008 – 2013. However, a significant proportion of engineering publications (at least 77 in the last five years) suggested low use of advanced statistical techniques. Furthermore, many scientific articles excluded relevant details of the pre-experimental steps, potentially reducing the educational value of the publications. Therefore, the scientific community should invest more tremendous efforts in conquering the valley between the engineering community and the statisticians. One of the ways is to provide a practical DOE framework with statistics aiding and not restricting the process.

This paper aims to provide the general framework with critical consideration to improve the effectiveness of the DOE in industrial applications. Furthermore, as the same personnel led the DOEs, it removed user variation and evaluated different DOEs objectively across industries. This is a novel idea as all the previous comparative DOE-related research publications have been independent or restricted to a single sector.

2. RESEARCH METHODS

We present a comparative evaluation of the utility of three distinct categories of DOE, conducted across a wide range of industries. These three types are also the most common forms of a DOE design:

- Screening,
- Mid-Level Resolution, and
- High Resolution or Response Surface Designs.

Many factors influence the creation of DOE, and hence the framework below was adopted and recommended to increase its effectiveness.

1. Define the problem. Explain the issue to be resolved or understood using the experimental designs.

2. Establish the objective. State the purpose of the DOE, i.e., to identify significant input variables, identify an interaction between variables, or characterise and optimise them.

3. Select the output. Decide how to measure the performance.

4. Select the input factors and their levels. Use the CNX (Controlled Noise eXperimental) diagram (Fig. 2) to help select experimental input factors. Controlled input factors remain constant and experi-

mental factors are changed as per DOE to evaluate their impact on the output. The effect of noise factors should be minimised during the DOE.

5. Select the experimental design & sample size. Below are some of the factors that should be considered while selecting the DOE.

- Aims and objectives of the experiment. An essential consideration in selecting a DOE is whether the intention is to use it to identify significant input variables only or characterise and optimise input variables.
- The number of input variables. Typically, in the case of input factors higher than six, a screening design is more appropriate; for a design involving less than six input variables, industries prefer a mid-level resolution design; and for processes inputting between two and four factors, the full factorial design is the optimal choice.
- The number of levels and replicates. The number of input factor levels helps characterise its relationship with the output, especially a non-linear relationship. Replicates can be used.
- Calculate pure error derived from the measurements. Ideally, the more the input factor levels and replicates, the better it is for DOE. Still, considering the DOE practicality, it may not always be possible to opt for the ideal design.
- Randomisation protects against unknown or uncontrolled input variables.

6. Collect the data: Once the test plan is created, perform the tests and collect the response/output data.

7. Analyse the data. Standard statistical software like Minitab can be used to analyse the DOE data in detail, helping in identifying, characterising and optimising significant input factors and their interactions.

8. Draw conclusions. Based on the analysis, make conclusions to help with the design and development of permanent corrective action.

9. Evaluate whether all objectives were achieved or not. If not, a follow-up DOE might be required.

Following the above framework, vital outcomes from various DOEs were present. They were conducted across different industries under three broad categories: Screening, Mid-Resolution and High-Resolution Designs. The utility of different DOEs was analysed based on their effectiveness and efficiency in meeting the desired objectives. The same personnel led these experimental designs eliminating the noise from operator variation due to skillset and methodology.

3. RESEARCH RESULTS

3.1. SCREENING DESIGNS

A screening DOE is used to gain insights between experimental inputs and outputs in cases when many experimental variables are present. To elucidate current research objectives, a couple of examples were taken from electronics wafer manufacturing and the wireless communication industry.

3.1.1. POOR LI-FI CONNECTION SPEED

An L36 (2³, 3³) Taguchi DOE was used in a warehouse logistic company to identify factors affecting connection speed between two devices communicating wirelessly via Light Fidelity (Li-Fi) technology. There were six experimental factors with multi-level input values. The example presented here was one of many screening DOEs conducted to gain insight into a large pool of potential Key Process Input Variables (KPIVs). Higher-resolution DOE would have required thrice the experimental runs, which was not practical considering the experiment’s effort. The process’s knowledge helped select a Taguchi design that evaluated potentially significant factors and interactions among the factors. The main

effects plot and response ranking for the mean connection speed is shown in Fig. 1 and Table 1, respectively.

Taguchi DOE was successful in identifying the key factors affecting the connection speed. The optimum settings, calculated using Minitab’s predict Taguchi results, yielded low connection speed when used during confirmation testing. For this reason, we recommend a follow-up using a screening DOE with a high resolution fractional or full factorial DOE. Results from the full factorial DOE are discussed below. We suggest that the top four factors from Taguchi DOE should be statistically significant in determining mean connection speed, but characterisation plots are likely to be different.

3.1.2. ELECTRONICS WAFER MANUFACTURING

It is an industry where conducting experimental design is very valuable but also very expensive. Each test run costs approx. GBP 500, which made it critical to keep the experimental runs as low as possible. Even before considering DOE, a detailed analysis of the existing data was conducted to reduce potential KPIVs from 40 to nine. However, nine was still too high to conduct a high resolution or more than 2-level DOE. For Taguchi or any DOE to be successful, care-

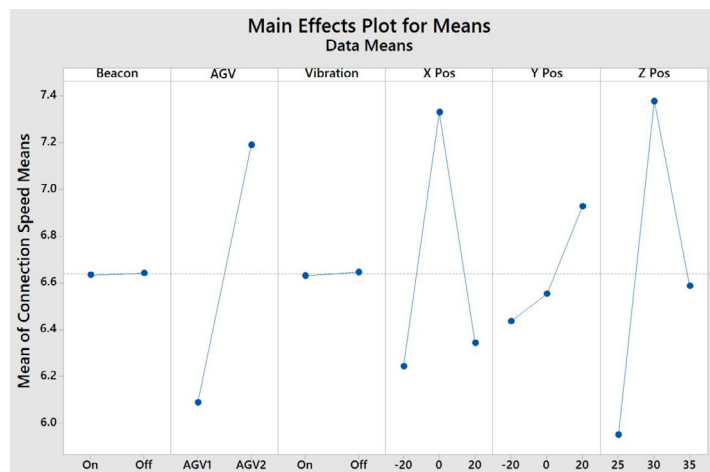


Fig. 1. Main effects plot for mean connection speed

Tab. 1. Response for mean connection speed

LEVEL	BEACON	AGV.	VIBRATION	X POS	Y POS	Z POS
1	6.634	6.088	6.632	6.243	6.437	5.951
2	6.642	7.190	6.646	7.329	6.553	7.378
3				6.344	6.927	6.587
Delta	0.008	1.102	0.014	1.087	0.490	1.426
Rank	6	2	5	3	4	1

ful selection of KPIVs levels, managing and controlling the noise variables are critical. A reasonable understanding of the KPIVs through the Six-Sigma methodology helped create a detailed CNX diagram shown in Fig. 2.

We used a Taguchi L12 design to rank the input factors in the order of their impact on TTV. Taguchi L12 implied that by performing 12 lapping batches, nine experimental factors at two different values were

evaluated for their effects on TTV, as shown in Fig. 3 and Table 2.

From Taguchi DOE, the top two ranking factors (the sun gear ratio and plate temperature) seemed to impact TTV significantly. Moreover, the bottom plate speed, with a ranking of three, did not appear to be significant but was still included in the further analysis; because Taguchi is only a screening DOE, so a follow-up with a more comprehensive DOE to

Tab. 2. Response for Means

Level	Sun Gear Ratio	Bottom Plate Speed	Exhaust Timer	Acceleration Timer	Plate Temp.	Recycle Slurry Status	Active agent concentration	Slurry Temp.	Slurry Mixing Time
1	1.2893	1.0338	1.0053	1.0092	1.1997	0.9940	1.0120	0.9957	1.0062
2	0.7123	0.9678	0.9963	0.9925	0.8020	1.0077	0.9897	1.0060	0.9955
Delta	0.5770	0.0660	0.0090	0.0167	0.3977	0.0137	0.0223	0.0103	0.0107
Rank	1	3	9	5	2	6	4	8	7

CNX Diagram for Taguchi DOE

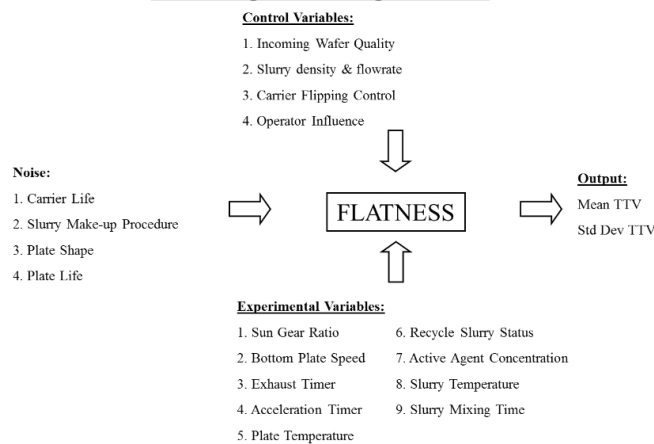


Fig. 2. CNX diagram for DOE

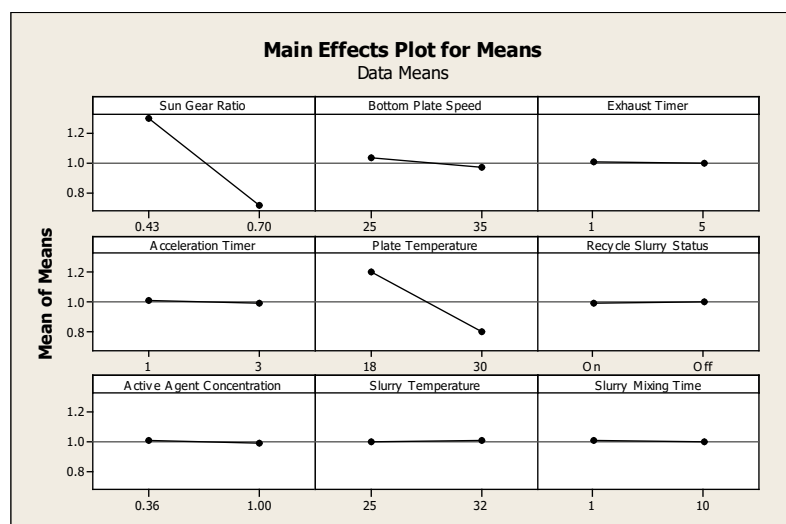


Fig. 3. Main effect plots for means

confirm the results of screening DOE is usually recommended.

3.2. MID-RESOLUTION DESIGNS

Mid-Resolution Designs are an optimal choice when the interaction between KPIVs is significant to the theoretical model while keeping the experimental runs to a minimum. In this section, a couple of examples are presented from the same industry.

3.2.1. THE PLUMBING INDUSTRY

This set of fractional factorial experimental designs was conducted in a UK-based, global plumbing company renowned for its bathroom products. A soon-to-be-launched mixer shower product was demonstrating a high reject rate for leak failures. The

product consisted of 12 sub-assemblies and components that went through five assembly stations and two leak tests at different assembly stages. At the final leak stage, approx. 12 % of the product was rejected for a high leak rate resulting in a high financial loss due to scrap and rework. After the initial root cause analysis, the team identified six potential factors with possible interactions. A quarter-factorial of two levels with resolution IV and a 20 test run design was deemed most suitable for the purpose. Centre points were also added to identify the possible non-linearity of the KPIVs. The output from the variance analysis and model summary is presented in Table 3, and the Pareto chart is shown in Fig. 4.

The response data were categorised based on visual inspection, which was an interesting attempt to evaluate the effectiveness of DOE against ordinal

Tab. 3. Analysis of variance and model summary

SOURCE	DF	ADJ SS	ADJ MS	F-VALUE	P-VALUE
Model	17	236.700	13.9235	26.52	0.037
Linear	6	116.200	19.3667	36.89	0.027
Housing	1	72.250	72.2500	137.62	0.007
Cover	1	0.450	0.4500	0.86	0.452
Energy	1	12.250	12.2500	23.33	0.040
Working Pressure	1	9.000	9.0000	17.14	0.054
Standstill Delay	1	6.250	6.2500	11.90	0.075
Amplitude	1	16.000	16.0000	30.48	0.031
2-Way Interactions	8	84.250	10.5312	20.06	0.048
Housing*Cover	1	0.250	0.2500	0.48	0.561
Housing*Energy	1	25.000	25.0000	47.62	0.020
Housing*Working Pressure	1	0.250	0.2500	0.48	0.561
Housing*Standstill Delay	1	25.000	25.0000	47.62	0.020
Housing*Amplitude	1	20.250	20.2500	38.57	0.025
Cover*Working Pressure	1	0.250	0.2500	0.48	0.561
Cover*Amplitude	1	12.250	12.2500	23.33	0.040
Energy*Standstill Delay	1	0.000	0.0000	0.00	1.000
3-Way Interactions	2	90.450	45.2250	86.14	0.011
Housing*Cover*Amplitude	1	36.000	36.0000	68.57	0.014
Cover*Energy*Standstill Delay	1	54.450	54.4500	103.71	0.010
Curvature	1	0.000	0.0000	0.00	1.000
Error	2	1.050	0.5250		
Total	19	237.750			
S	R-sq	R-sq(adj)	R-sq(pred)		
0.724569	99.56%	95.80%	24.06%		

response data. An excellent R-sq value suggested an ideal regression model. Significant KPIVs were identified from the model, and an optimum setting was established using the Minitab software. Confirmation testing validated the theoretical model.

3.2.2. ROBUSTNESS OF THE OPTIMUM SETTINGS

The optimum settings derived in the previous section need to work for different batches of three key components, which are housing, cover, and base. Although visual check (Attribute Data) indicates weld quality, burst strength (continuous data) is the correct measure that required destructive testing. Considering the total loss of a part due to destructive testing, a half-factorial, two-level, and resolution III was conducted with these three factors. The variance analysis results are shown in Table 4 and the Pareto chart in Fig. 5.

We obtained excellent experimental design results with a 100 % R-sq value and no factors exerting a statistically significant burst strength impact. However, further analysis revealed that the DOE could not calculate p-value and R-sq (adj), suggesting some missing data links. Visual inspection of the weld further revealed quality issues with certain

parts. As this was a compliance and safety issue, a full factorial DOE technique was performed with one replicate using the same input variables. The results are presented in section 3.3.1. It was also found that although it was a fractional factorial DOE, the resolution was only III, which had put it into the screening DOE category and hence less useful in characterising the factors.

3.3. HIGH-RESOLUTION FULL FACTORIAL DESIGNS

High-Resolution Full factorial designs are complete models with quadratic interactions for identification, characterisation, and optimisation of KPIVs.

This is the follow up full factorial DOE technique to the Taguchi screening discussed earlier in section 3.1.1. A four-factor and two-level full factorial DOE was performed with centre points using experimental factors identified from the screening DOE. The main effects plot and model summary generated using Minitab selecting stepwise options are shown in Fig. 6 and Table 5.

All four experimental variables and a two-level interaction were found to be statistically significant. A robust regression model, R-sq (Adj) of 90.61 %, was

Tab. 4. Analysis of variance and model summary

SOURCE	DF	ADJ SS.	ADJ MS	F-VALUE	P-VALUE
Model	3	6.68750	2.22917	*	*
Linear	3	6.68750	2.22917	*	*
Housing	1	1.56250	1.56250	*	*
Base	1	0.06250	0.06250	*	*
Cover	1	5.06250	5.06250	*	*
Error	0	*	*		
Total	3	6.68750			
S	R-sq	R-sq(adj)	R-sq(pred)		
	* 100.00%	*	*		

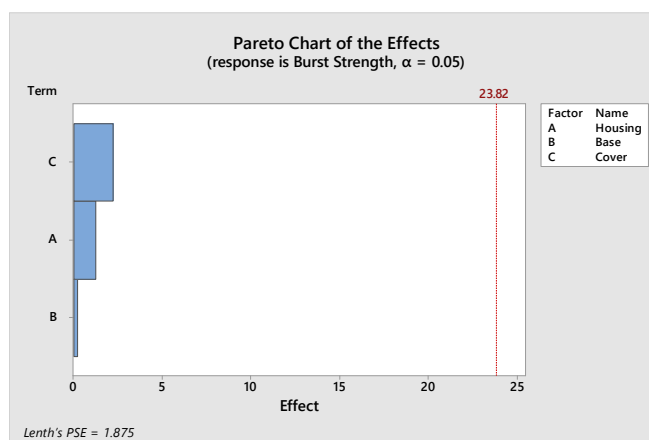


Fig. 5. Pareto chart of the effects

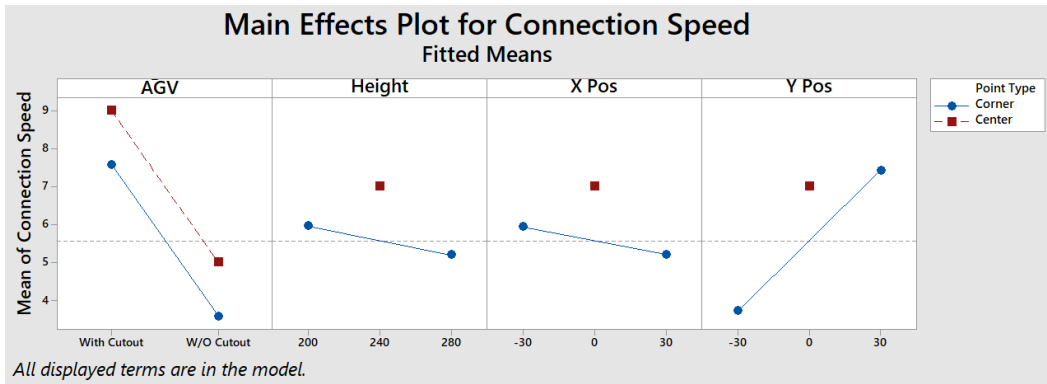


Fig. 6. Main effects plot for mean connection speed

Tab. 5. Model summary and coded co-efficient for connection speed

S	R-SQ	R-SQ(ADJ)		R-SQ(PRED)		
0.929010	91.56%	90.61%		88.99%		
TERM	EFFECT	COEF	SE COEF	T-VALUE	P-VALUE	VIF.
Constant		5.579	0.147	37.98	0.000	
AGV	-3.992	-1.996	0.120	-16.64	0.000	1.00
Height	-0.757	-0.378	0.147	-2.58	0.013	1.00
X Pos	-0.720	-0.360	0.147	-2.45	0.018	1.00
Y Pos	3.702	1.851	0.147	12.60	0.000	1.00
AGV*Y Pos	2.860	1.430	0.147	9.74	0.000	1.00
Ct Pt		1.436	0.254	5.64	0.000	1.00

achieved with a combination of four factors tested during the experiment suggesting that most of the process variation can be controlled via these factors. With the help of Minitab’s Response Optimiser, optimum settings were identified, including tolerances, for all the factors. It helped develop a robust solution that can deal with manufacturing, installation, and operational variations between the two AGVs without significantly affecting connection speed. It is essential to highlight that some of the obtained settings differed from what was identified during the screening phase, further highlighting the potential shortcomings of screening DOE. Pilot testing was conducted with the new set-up using single Li-Fi, but multiple AGVs. Achieved process capability (Cpk) of 3.18 was much higher than the industry standard of 1.67, which also confirmed the solution’s robustness.

This is the follow up full factorial DOE to the Taguchi screening discussed earlier in section 3.1.2. Three-factor and two-level full factorial DOE with centre points was designed to characterise and optimise the three experimental factors. The DOE successfully identified statistically significant factors and the effect of their interactions on the output (Fig. 7). The factors above the red line were statistically significant, and the factors below were non-significant.

The main effects plot shown in Fig. 8 indicated that an increase in the sun gear ratio, plate temperature and the bottom plate speed resulted in reduced TTV. Further, as centre points (marked red) for all the three factors were not on the line, the impact of these factors on TTV was non-linear. For example, the plate temperature increase to 24 °C did not reduce TTV; however, a plate temperature closer to 30 °C was most likely to reduce the TTV reject rate.

Results from Minitab’s Response Optimiser are shown in Fig. 9, suggesting that to achieve the minimum TTV (0.462), inputs should be set to the parameters highlighted in red.

The theoretical predictive DOE model was validated via a confirmation run using the recommended input values that yielded similar results to the DOE model.

This is the follow up full factorial DOE to the fractional factorial DOE discussed earlier in the plumbing industry section 3.2.2. A three-factor and two-level full factorial DOE with a replicate was conducted. The results of the Pareto chart of the standardised effects are shown in Fig. 10.

In contrast to fractional factorial DOE, the full factorial DOE was able to identify “housing” as a statistically significant factor. The same full facto-

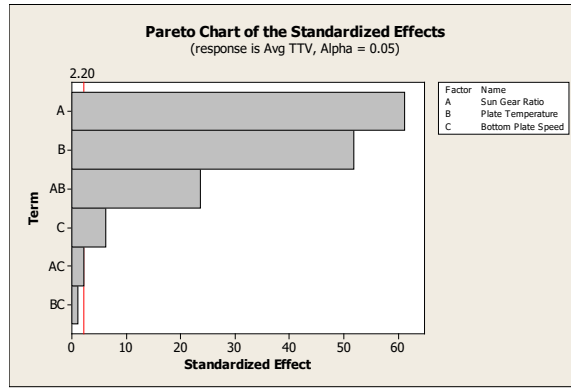


Fig. 7. Normal and standardised effects for Pareto charts

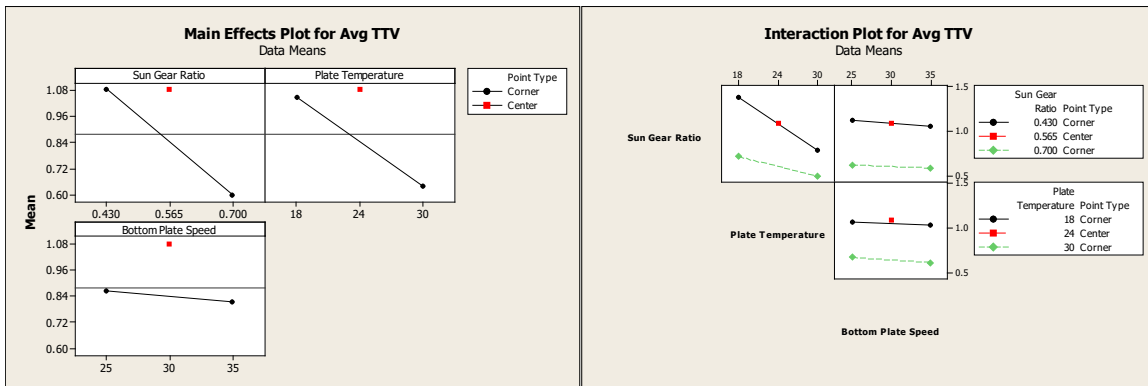


Fig. 8. Main effects and interaction plots for Avg TTV

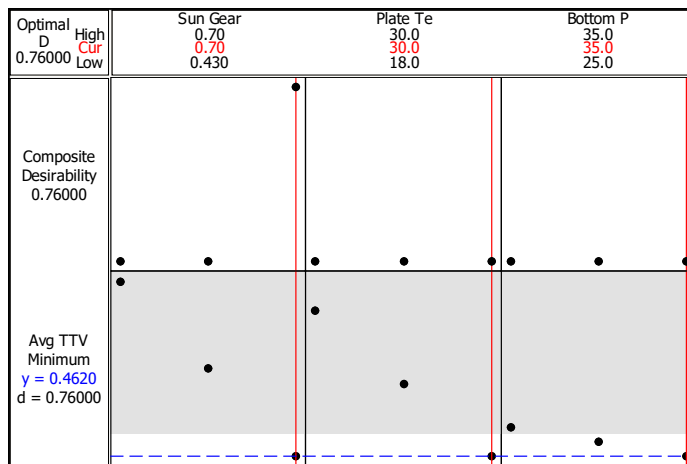


Fig. 9. Response optimiser result

rial DOE model without replicates was analysed as an academic exercise, presented in Fig. 11. It can be seen that without replicates, even the full factorial DOE model failed to detect any significant factors. Investigating it further revealed that high variation in the burst strength measurements made it challenging to identify process shifts. In such cases, it is recommended to have as many replicates as possible.

Based on the insights gained from the DOE, further investigation into two types of housing revealed that poor roundness of the low performing housing was responsible for reduced burst strength. Confirmation testing and the pilot study conducted using suitable housing provided similar results to the theoretical DOE model. Further data collected after the resolution of housing roundness also provided excellent results with no failures.

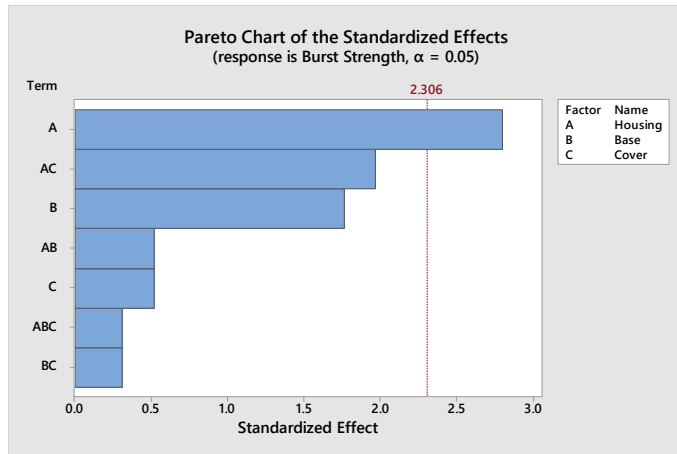


Fig. 10. Normal and standardised effects for Pareto charts (with a replicate)

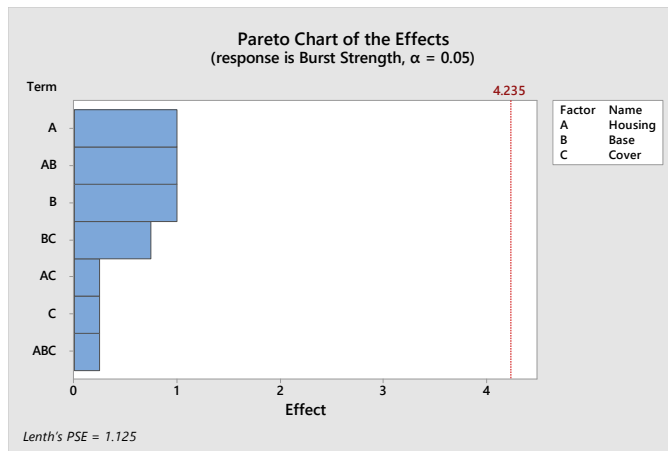


Fig. 11. Normal and standardised effects for Pareto charts (without a replicate)

3.4. SUMMARY OF VARIOUS DOE RESULTS

Comparison between multiple DOE techniques and their utility is summarised in Table 6.

Tab. 6. Comparative summary of DOE results

DOE TYPE	INPUT FACTORS	EFFORT / COST	IDENTIFYING KPIV & INTERACTIONS	CHARACTERISING KPIV & INTERACTIONS	OPTIMISING KPIV & INTERACTIONS	APPLICATION
Screening	6,9	Low	KPIVs only	Not reliable	Not reliable	At the start of development or problem solving, where the number of potential input variables is high
Mid-resolution	3,6	Medium	KPIV & some of 2-way interactions	KPIV & some of the 2-way interactions with low error	Not reliable	The number of potential input variables is <9 with no 4-way or higher interactions
High-resolution	3,4	High	KPIV and up to all interactions	KPIV & up to all the interactions with very low error	Reliable	This should be the preferred choice as long as it is practical

4. DISCUSSION OF THE RESULTS

In line with previous research (Ilzarbe et al., 2008; Antony et al., 1999; Hecht et al., 2016), various DOE models were highly influential in attaining the experiments' objectives. This research study found that defining the experimental designs' goals as part of DOE planning is vital. For example, a) Is DOE used for KPIVs identification only or to characterise and optimise KPIVs as well? b) Is the interaction between KPIVs significant? c) To what level must the KPIVs be characterised? However, DOE objectives may change if many tests are not feasible due to a large set of experimental factors and their levels. Every factor or level added in the full factorial design increases test runs significantly, calculated by FL where F is the number of factors, and L is the number of levels. For example, a 2-level and 3-factor will require 8 test runs compared to 27 test runs for a 3-level and 3-factor or 16-factors for a 2-level and 4-factor design. Similarly, a fraction factorial or screening DOE will use half or less than a full factorial design. As shown in Section 3, a less comprehensive DOE model must be chosen to achieve the initial objectives set out for experiments in such cases.

Consistent with previous research studies (Zheng et al., 2013; Robinson et al., 2003; Box, 1988; Box et al., 1988; Nair, 1992), it was found that screening DOEs were highly effective in identifying KPIVs while managing a high number of variables. In Section 3.1, both screening DOEs managed to identify statistically significant KPIVs, which remained the same even when conducted with higher resolution DOE. However, it was argued that screening DOE effectiveness was highly dependent on the user's skillset to choose the correct levels and screening model. Previous studies have highlighted the ineffectiveness of screening DOE in characterising the KPIVs. This paper confirmed this conclusion as the optimum setting derived from screening DOEs in both cases failed to give desired results during confirmation testing. As the name suggests, screening DOE can be good for identifying KPIVs, but not that useful for characterising and optimising them.

It is well established that the Design of Experiments is a powerful and versatile analytical tool that works across industries with different data types (e.g., Durakovic, 2017; Hecht et al., 2016; Ilzarbe et al., 2008). This research further elucidates that the DOE works with continuous variable data and ordinal response data. In Section 3.2.1, the response variable

was a visual inspection which categorised ultrasonic weld quality. However, as a general rule, before any data analysis, data validation should be performed. In this case, a Measurement System Analysis (MSA) was performed using the agreement analysis to identify reproducibility and repeatability error. In both cases, the agreement between and within visual inspectors was more remarkable than 90 %. The case studies presented in Sections 3.2.2 and 3.3 also stresses the relevance of MSAs. High variation in the burst strength measurement made it difficult for various DOEs to create an accurate regression model. In line with previous research studies (Sukthomya & Tannock, 2005), DOEs were found to remain effective when an independent variable was discrete or categorical. All the DOE cases discussed in this manuscript had either categorical or discrete data as an independent variable.

It is essential when performing a DOE that the noise factors are identified and minimised as much as possible (Kackar & Shoemaker, 2021; Gremyr et al., 2003). However, at times maybe inadvertently, noise can be introduced in the DOE. Previous research studies recommended randomising experimental runs to reduce the impact of noise factors on the DOE model (Hecht et al., 2016; Antony, 2014; Box, 1990). In Sections 3.1.2 & 3.3, a couple of factors (the slurry mixing time and the plate temperature) are time-dependent, meaning the longer a machine runs, the better they become. If a similar DOE without these experimental variables and a randomised test was administered, the first few test runs would have always resulted in a poor TTV regardless of the test set-up. Another way of reducing the impact of noise factors is by repeating test runs (Hecht et al., 2016; Antony, 2014). For example, in section 3.3, a measurement system variation for the process was acceptable (% Tolerance Variation: 19.88), but it was still too high for the DOE to accurately identify and characterise independent variables. In this case, replicates helped to form correct modelling.

In line with previous studies (Weissman & Anderson, 2015), it is highly recommended to undertake confirmation testing after any DOE. The larger the sample size, with normal process variation, the better it is for validation. With a couple of case studies presented in Sections 3.1.1 and 3.2.2, it was found that the characterisation of independent variables was way off the mark, and optimum settings did not give the desired results. Confirmation testing becomes even more critical when the understanding of a process is limited.

The availability of advanced statistical software has made the DOE application more accessible without great statistical expertise for users (Durakovic, 2017; Hibbert, 2012). Also, the Minitab software was found to be extremely helpful in selecting and analysing various DOEs. The stepwise function was repeatedly used to generate accurate regression models. The Response Optimiser was also very helpful in identifying optimum settings, including tolerances for the independent variable, especially significant interactions were present.

CONCLUSIONS

The current study aimed to emphasise the relevance of the DOE technique as a reliable method for ensuring efficient use of statistical methods in routine industrial processes. The cases presented in the manuscript led to the conclusion that DOE effectively achieves the desired objectives. The DOE model should be selected based on the goals, i.e., screening, identifying, characterising or optimising KPIVs and their interactions. DOE is a powerful and versatile analytical tool that works across industries with different data types. However, if noise factors are not appropriately managed, it can easily result in inaccurate regression models. Randomisation and replicates are a couple of acceptable practices to reduce noise. Further, to avoid following the incorrect DOE conclusions, confirmation testing after any DOE is highly recommended.

Overall, various DOE models proved successful in identifying a complicated relationship between experimental variables and output. The different cases presented in the research study were conducted by the same person with a decent knowledge of the experimental product or process. It will, therefore, be useful for future research to try evaluating the impact of personnel and lack of product or process knowledge on DOE effectiveness.

ACKNOWLEDGEMENTS

The publication of the article for 11th International Conference on Engineering, Project, and Production Management - EPPM2021 was financed in the framework of the contract no. DNK/SN/465770/2020 by the Ministry of Science and Higher Education within the "Excellent Science" programme.



LITERATURE

- Andersson, P. M., Lundstedt, T. & Abramo, L. (1996). Synthesis and optimisation of 1-pyrrolemethane sulfonate by means of experimental design. *Journal of Chemometrics*, 10, 379-384.
- Antony, J. (2014). Some Useful and Practical Tips for Making Your Industrial Experiments Successful. In J. Antony (Ed.), *Design of Experiments for Engineers and Scientists. Second Edition* (pp. 113-123). Elsevier. doi: 10.1016/B978-0-08-099417-8.00008-0
- Antony, J., & Roy, R. K. (1999). Improving the process quality using statistical Design of experiments: A case study. *Quality Assurance*, 6, 87-95.
- Arnoldsson, K. C., & Kaufmann, P. (1994). Lipid class analysis by normal phase high performance liquid chromatography, development and optimisation using multivariate methods. *Chromatographia*, 38, 317-324.
- Beijersten, I., & Westerlund, D. (1995). Derivatisation Of Dipeptides With 4-Fluoro-7-Nitro-2,1,3 Benzoxadiazole For Laser-Induced Fluorescence And Separation By Micellar Electrokinetic Chromatography. *Journal of Chromatography*, 716, 389-399.
- Bisgaard, S. (1991). Teaching Statistics to Engineers. *The American Statistician*, 45(4), 274-283.
- Bisgaard, S. (1992). Industrial use of statistically designed experiments: Case study references and some historical anecdotes. *Quality Engineering*, 4, 547-562.
- Box G. E. P., Bisgaard, S., & Fung, C. (1988). An explanation and critique of Taguchi's contributions to quality engineering. *Quality and Reliability Engineering International*, 4, 123-131.
- Box, G. E. P. (1988). Signal-to-noise ratios, performance criteria, and transformation. *Technometrics*, 30, 1-40.
- Box, G. E. P. (2001). Statistics for discovery. *Journal of Applied Statistics*, 28(3-4), 285-299.
- Box, G. E. P., Hunter, W. G. J., & Hunter, S. (1987). Statistics for experimenters: an introduction to Design.
- Brady, J. E., & Allen, T. T. (2006). Six Sigma literature: A review and agenda for future research. *Quality and Reliability Engineering International*, 22, 335-367.
- Bucher, R. A., & Loos, A. C. (1994). Parametric statistical analysis of electrostatic powder prepping. *Journal of Advanced Materials*, 25, 44-50.
- Bzik, T. J., Henderson, P. B., & Hobbs, J. P. (1998). Increasing the precision and accuracy of top loading balances: Application of experimental design. *Analytical Chemistry*, 70, 58-63.
- Carlson, A. D., Hofer, J. D., & Riggan, R. M. (1997). Development of an optimised peptide map for recombinant activated human protein c by means of an experimental design strategy. *Analytica Chimica Acta*, 352, 221-230.

- Chapin, S. F. (1950). Research note on randomisation in a social experiment. *Science*, 112, 760-761.
- Chen, H. C. (1996). Optimising the concentrations of carbon, nitrogen and phosphorus in a citric acid fermentation with response surface method. *Food Biotechnology*, 10, 13-27.
- Czitrom, V. (1999). One factor at a time versus Designed Experiments. *The American Statistician*, 53(2), 126-131.
- data analysis, and model building. New York: Wiley.
- Davim, J. P. (Ed.). (2016). *Design of Experiments in Production Engineering*. Springer International Publishing.
- Davis, B. L., Cavanagh, P. R., Sommer, H. J., & Wu, G. (1996). Ground reaction forces during locomotion in simulated microgravity. *Aviation Space and Environmental Medicine*, 67, 235-242.
- Durakovic, B. (2017). Design of Experiments Application, Concepts, Examples: State of the Art. *Periodicals of Engineering and Natural Sciences*, 5(3), 421-439. doi: 10.21533/pen
- Durakovic, B., & Torlak, M. (2017). Simulation and experimental validation of phase change material and water used as heat storage medium in window applications. *Journal of Materials and Environmental Science*, 8(5), 1837-1746.
- Fisher, R. A. (1926). The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain*, 33, 503-513.
- Gardner, R., Bieker, J., Elwell, S., Thalman, R., & Rivera, E. (2000). Solving tough semiconductor manufacturing problems using data mining. In Proceedings of IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop, 46-55.
- Goh, T. N. (2002). The role of statistical Design of Experiments in Six Sigma: Perspectives of a practitioner. *Quality Engineering*, 14(4), 659-671.
- Gremyr, I., Arvidsson, M., & Johansson, P. (2003). Robust Design Methodology: Status in the Swedish Manufacturing Industry. *Qualitative Reliability Engineering International*, 19, 285-293.
- Hahn, G. J. (2005). Six Sigma: 20 Key Lessons Learned. *Quality and Reliability Engineering International*, 21, 225-233.
- Hecht, E. S., Oberg, A. L., & Muddiman, D. C. (2016). Optimising Mass Spectrometry Analyses: A Tailored Review on the Utility of Design of Experiments. *Journal of American Society of Mass Spectrometry*, 27, 767-785.
- Hibbert, D. B. (2012). Experimental design in chromatography: a tutorial review. *Journal of Chromatography*, 910, 2-13.
- Ilzarbe, L., Álvarez, M. J., Viles, E., & Tanco, M. (2008). Practical applications of design of experiments in the field of engineering: a bibliographical review. *Qualitative Reliability Engineering International*, 24, 417-428.
- Kackar, R. N., & Shoemaker, A. C. (2021). Robust Design: A cost-effective method for improving manufacturing processes. *AT&T Technical Journal*, 65, 39-50.
- Kenett, R. S., & Steinberg, D. M. (2006). New frontiers in the Design of experiments. *Quality Progress*, 39(8), 61-65.
- Lye, L. M. (2005). *Tools and toys for teaching design of experiments methodology*. In 33rd Annual General Conference of the Canada, Toronto, Ontario, Canada.
- Mager, P. P. (1997). How design statistics concepts can improve experimentation in medicinal chemistry. *Medicinal Research Reviews*, 17, 453-475.
- Montgomery, D. C. (2017). *Design and Analysis of Experiments*. John Wiley & Sons, Inc.
- Myers, R. H., Montgomery, D. C., Vining, G. G., Borror, C. M., & Kowalski, S. M. (2004). Response surface methodology: A retrospective and literature survey. *Journal of Quality Technology*, 36(1), 53-77.
- Nair, V. N. (1992). Taguchi's parameter design: A panel discussion. *Technometrics*, 34, 127-161.
- Okatia, V., Behzadmehra, A., & Farsad, S. (2016). Analysis of a solar desalinator (humidification-dehumidification cycle) including a compound system consisting of a solar humidifier and subsurface condenser using DoE. *Desalination*, 397, 9-21.
- Paulo, F., & Santos, L. (2017). Design of experiments for microencapsulation applications: A review. *Materials Science and Engineering: C*, 77(August), 1327-1340.
- Puente-Massaguer, E., Lecina, M., & Gòdia, F. (2020). Integrating nanoparticle quantification and statistical design of experiments for efficient HIV-1 virus-like particle production in High Five cells. *Applied Microbiology and Biotechnology*, 104, 1569-1582. doi: 10.1007/s00253-019-10319-x
- Robinson, T. J., Borror, C. M., & Myers, R.H. (2004). Robust parameter design: A review. *Quality and Reliability Engineering International*, 20, 81-101.
- Schlueter, A., & Geyer, P. (2018). Linking BIM and Design of Experiments to balance architectural and technical design factors for energy performance. *Automation in Construction*, 86(February), 33-43.
- Setamanit, S. (2018). Evaluation of outsourcing transportation contract using simulation and design of experiment. *Polish Journal of Management Studies*, 18(2), 300-310.
- Sukthomya, W., & Tannock, J. (2005). The optimisation of neural network parameters using Taguchi's design of experiments approach: an application in manufacturing process modelling. *Neural Computing and Applications*, 14, 337-344. doi: 10.1007/s00521-005-0470-3
- Tukey, J. W. (1947). Non-parametric estimation II. Statistically equivalent blocks and tolerance regions – the continuous case. *Annals of Mathematical Statistics*, 18, 529-539.
- Wald, A. (1943). An extension of Wilks' method for setting tolerance limits. *Annals of Mathematical Statistics*, 14, 45-55.
- Wesling, P., & Emamjomeh, A. (1994). T.A.B. Inner-lead bond process characterisation for single point laser bonding. *IEEE Transactions on Components Packaging & amp. Manufacturing Technology Part A*, 17, 142-148.

- Yang, G. C. C., & Tsai, C. M. (1998). A study on heavy metal extractability and subsequent recovery by electrolysis for a municipal incinerator fly ash. *Journal of Hazardous Materials*, 58, 103-120.
- Yip, H. M., Wang, Z., Navarro-Alarcon, D., Li, P., Cheung, T. H., Greiffenhagen, Ch., & Liu, Y. (2020). A collaborative robotic uterine positioning system for laparoscopic hysterectomy: Design and experiments. *International Journal of Medical Robotics and Computer Assisted Surgery*, 16(4), e2103. doi: 10.1002/ircs.2103
- Yondo, R., Andrés, E., & Valero, E. (2018). A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses. *Progress in Aerospace Sciences*, 96, 23-61. doi: 10.1016/j.paerosci.2017.11.003
- Yoo, K. S. (2020). Application of Statistical Design of Experiments in the Field of Chemical Engineering: A Bibliographical Review. *The Korean Society of Industrial and Engineering Chemistry*, 31(2), 138-146. doi: 10.14478/ACE.2020.1018
- Yu, P., Low, M. Y., & Zhou, W. (2018). Design of experiments and regression modelling in food flavour and sensory analysis: A review. *Trends in Food Science & Technology*, 71(January), 202-215.
- Zheng, H., Clausen, M. R., Dalsgaard, T. K., Mortensen, G., & Bertram, H. C. (2013). Time-saving design of experiment protocol for optimisation of LC-MS data processing in metabolomic approaches. *Analytical Chemistry*, 85, 7109-7116.