Received: 19 February 2020 | Revised: 19 May 2020 | Accepted: 15 June 2020
DOI: 10.1002/qre.2722

RESEARCH ARTICLE

WILEY

Three-level designs: Evaluation and comparison for screening purposes

Mohammed Alomair[1,2] | **Stelios Georgiou**[1] | **Stella Stylianou**[1]

[1]School of Science, Royal Melbourne Institute of Technology University, Melbourne, Victoria, Australia

[2]Department of Mathematics, Al Imam Muhammad Bin Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

Correspondence
Georgiou Stelios, School of Science, Royal Melbourne Institute of Technology University, Melbourne, Vic 3000, Australia.
Email: stelios.georgiou@rmit.edu.au

Abstract

Since their introduction by Box and Hunter, resolution criteria have been widely used when comparing regular fractional factorials designs. In this article, we investigate how a generalized resolution criterion can be used to assess some recently developed three-level screening designs, such as definitive screening designs (DSDs) and screening designs from weighing matrices. The aim of this paper is to capture the projection properties of those three-level screening designs, complementing the work of Deng and Tang, who used generalized resolution and minimum aberration criteria for ranking different two-level designs, particularly Plackett-Burman and other nonregular factorial designs. An advantage of generalized resolution, extended here to work on three-level designs, is that it offers a useful criterion for ranking three-level screening designs, whereas the Deng and Tang resolution is used mainly for the assessment of two-level designs. In addition, we applied a projection estimation capacity (PEC) criterion to select three-level screening designs with desirable properties. Practical examples and the best projections of the designs are presented in tables.

KEYWORDS

factorial designs, generalized resolution, linear models, projection estimation capacity, screening

1 | INTRODUCTION

Full factorial designs provide independent estimation for all factorial effects. The high cost of performing a full factorial is the reason that fractional factorial designs, which are subsets of full factorial designs, are used. Thus, fractional factorial designs are widely used in industrial settings to identify the most active factors that affect responses or processes. Orthogonality of two-level designs occurs when the number of +'s and −'s in each design column is equal and the four-level combinations (+ +), (− −), (+ −), and (− +), for each pair of factors, have the same frequency.

In this paper, we capture the projection properties of fractional factorial designs with three levels, where we use − as low, 0 as intermediate, and + as high levels, respectively, according to Montgomery.[1] Orthogonal fractional factorial designs and screening designs are commonly used in industrial research, bio-medical engineering, drug discovery, computer simulation experiments, and machine learning (see Dean and Lewis[2]). One of the desirable properties of screening designs is that the number of runs is limited because screening designs aim in identifying the most important factors from a large number of factors that may affect a response. Thus, it is important to investigate the projection properties of these designs. The projected subdesigns will have a small subset of factors and will be generated and studied because it is not known which factors will be active. When the active factors are identified, designs with good projections can reveal additional valuable information after the set of non-active factors is deleted.

Fractional factorial designs are categorized as either regular or nonregular. Defining relations are used for a regular fractional factorial design that has a simple aliasing structure so any two effects are orthogonal or fully aliased. A nonregular fractional factorial design has a more complicated aliasing structure, which means it is difficult to interpret the significance of some effects that are neither orthogonal nor fully aliased. The Plackett-Burman[3]design is an example of a nonregular fractional factorial design that is extensively used in screening experiments based on its flexible and economic run size (see Wu and Hamada[4]). Consequently, nonregular factorials have received more attention in the past decade. More work on two-level fractional factorial designs and their projection properties can be found on Hamada and Wu[5]; Lin and Draper[6]; Wang and Wu[7] and Cheng.[8,9] Deng and Tang[10] and Tang and Deng[11] studied Plackett-Burman and other nonregular factorial designs. They provided two criteria —generalized resolution and generalized minimum aberration —for ranking two-level nonregular designs.

There is less literature on studying the projection properties of designs with more than two levels. Only few researchers have investigated the projection properties of designs having more than two levels. For instance, Wang and Wu[7] and Cheng and Wu[12] worked on the hidden projection properties when the designs were projected onto three and four factors using the OA (18, 3^7, 2) design. Cheng and Wu[12] also included the orthogonal array OA (36, 3^{12}, 2) and OA (27, 3^8, 2) in their projection investigation. For more studies on projection properties of three-level designs, we refer to Xu, Cheng, and Wu[13]; Tsai, Gilmour, and Mead[14,15]; Evangelaras, Koukouvinos, and Lappas[16];and Dey. [17]

The problem of assessing three-level factorial designs, especially screening designs, should now get more attention as more screening designs with three levels appear in the literature. Deng and Tang[10] used generalized resolution and minimum aberration criteria for comparing and ranking screening designs. The criteria we use in this paper are natural generalizations of the criteria they applied to two-level non-regular factorials.

Following Deng and Tang,[10] a factorial design, regular or nonregular, is denoted by D and is regarded as a set of m columns $D = \{d_1, \ldots\ldots, d_m\}$ or as an $n \times m$ matrix $D = (d_{ij})$, depending on our preference. For $1 \le r \le m$ and any r-subset $T = \{d_{j_1}, \ldots, d_{j_r}\}$ of D, define

$$J_r(T) = J_r(d_{j_1}, \ldots, d_{j_r}) = \left|\sum_{i=1}^{n} d_{ij_1}, \ldots d_{ij_r}\right|. \tag{1}$$

Clearly, Deng and Tang[10] illustrated that for two-level orthogonal designs, $J_1(T) = J_2(T) = 0$. The value of $J_r(T)$ can be used to develop the generalized resolution and minimum aberrationcriteria.

The paper is organized as follows. In Section 2, we recall the construction of some definitive screening designs (DSDs) that Jones and Nachtsheim[18] suggested. These are presented in Table 1. In the same section, we present some alternative three-level designs that were recently appear in the literature. These use weighing matrices and a fold-over structure for their constructions. In Section 3, we introduce generalized resolution, and some examples

TABLE 1 Jones and Nachtsheim definitive screening designs (DSDs) structure for m factors

Run	x_1	x_2	...	x_m
1	0	±1	...	±1
2	0	∓1	...	∓1
3	±1	0	...	±1
4	∓1	0	...	∓1
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
$2m-1$	±1		...	0
$2m$	∓1		...	0
$2m+1$	0		...	0

of how this can be applied for design comparisons. The results on designs evaluation are presented and discussed. In Section 4, we define the confounding frequency vector (CFV) of a design based on the generalized minimum aberration criterion to discriminate between designs. In Section 6, we introduce the projection estimation capacity (PEC) criterion and we use some examples to select three-level screening designs with desirable properties.

2 | CONSTRUCTION OF THREE-LEVEL SCREENING DESIGNS

One of the advantages of DSDs is their small number of runs. For example, a design with four or more factors ($m \geq 4$) requires just twice as many runs (n) plus one (i.e., $n = 2m + 1$) (see Jones and Naschtsheim[18]).

In this article, the focus is on three-level screening designs, and we include the DSDs, WSD-2, and WSD-3 as their constructions are described below. DSDs can be constructed as suggested by Jones and Nachtsheim[18]:

$$D = \begin{bmatrix} C \\ 0's \\ -C \end{bmatrix},$$

where C is an $m \times m$ matrix and 0 is a $1 \times m$ zero matrix. D will have $2m + s$ runs. The construction of DSDs was introduced with $s = 1$ and can be considered as a special case of designs constructed from weighing matrices with 1 zero per row and column (conference matrix). W(6,5) is an example of a conference matrix (see Xiao et al.[19]) and can be used to generate a DSD of $m = 6$ factors.

$$DSD = WSD-1 = W(6,5) = \begin{pmatrix} 1 & 1 & -1 & -1 & 0 & 1 \\ 1 & -1 & -1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 & 1 & 1 \\ -1 & -1 & 0 & -1 & -1 & 1 \\ 0 & -1 & -1 & -1 & 1 & -1 \\ -1 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 & -1 \\ -1 & 1 & 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 & -1 & -1 \\ 1 & 1 & 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 & -1 & -1 \end{pmatrix}.$$

Moreover, a weighing matrix $W = W(m,w)$ is a square matrix of order m with entries from the set {0, ±1} satisfies $W W^T = W^T W = wI_n$, where m is the number of nonzero entries per row and column. Parameter w is called the weight of W. WSD-s for $s = 1$,2, and 3 used by Georgiou et al.[20] and were constructed by using a fold-over were structure with a $W(m,m-s)$ and by adding a selected number of center points. For example, the WSD-2 and WSD-3 can be constructed as

$$WSD-2 = \begin{bmatrix} W(m,m-2) \\ 0's \\ -W(m-2) \end{bmatrix}, WSD-3 = \begin{bmatrix} W(m,m-3) \\ 0's \\ -W(m-3) \end{bmatrix}.$$

The W(8,6) and W(8,5) are used to construct a WSD-2 and WSD-3, respectively:

$$WSD-2=\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & -1 & 0 & -1 & 1 & -1 \\ -1 & 0 & 1 & 1 & -1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 & 1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 & 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -1 & -1 & 1 & 0 & 1 \\ -1 & -1 & -1 & 0 & 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & -1 & 0 & -1 & -1 \\ 1 & -1 & 0 & 1 & 0 & 1 & -1 & 1 \\ 1 & 0 & -1 & -1 & 1 & -1 & -1 & 0 \\ 0 & -1 & 1 & -1 & -1 & -1 & 0 & 1 \\ 1 & 0 & -1 & 1 & -1 & -1 & 1 & 0 \\ 0 & -1 & 1 & 1 & 1 & -1 & 0 & -1 \\ 1 & 1 & 1 & 0 & -1 & 0 & -1 & -1 \\ 1 & -1 & 0 & -1 & 0 & 1 & 1 & -1 \end{pmatrix},$$

$$WSD-3=\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ -1 & 0 & 0 & -1 & 0 & -1 & 1 & -1 \\ -1 & 0 & 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & -1 & -1 & 0 & 0 & 1 \\ -1 & -1 & -1 & 0 & 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & -1 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & 0 & 1 & -1 & 1 \\ 1 & 0 & 0 & -1 & 1 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 & -1 & -1 & 0 & 1 \\ 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 1 & 0 & -1 & 0 & 0 & -1 \\ 1 & -1 & 0 & -1 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

For more details on the constructions and properties of WSD_s, we referred to Georgiou et al.[20]

3 | A GENERALIZED RESOLUTION CRITERION FOR THREE-LEVEL SCREENING DESIGNS

There are two equivalent methods for performing the resolution of a regular fractional factorial design. From the projection viewpoint, if all the possible level combinations 2^{r-1} in the projection design onto any $(r-1)$ factors occur with the same frequency, then we say that this regular factorial has resolution r. From the estimability viewpoint, there is the assumption that the interaction effects including $(r+1)/2$ or more factors are negligible, a regular factorial has resolution r if (r odd), and the interaction effects including $(r-1)/2$ or fewer factors can be estimated. In contrast, under the assumption that the interaction effects including $(r+2)/2$ or more factors are negligible, a regular factorial has resolution r if (r even), and the interaction effects including $(r-2)/2$ or fewer factors are estimable (for more information, see Box and Hunter[21]).

By using the estimability viewpoint, the resolution can be generalized to any factorial design. Webb[22] gave the definition of resolution that was later applied by Rechtschaffner[23] and Srivastava and Chopra[24] to construct useful designs. However, Deng and Tang[10] noted that this method can only be used as a classification rule, so it is not useful for ranking different designs. For instance, a resolution V design may be less efficient for estimating the main effects than a resolution III design when the experimental error is substantial.

Deng and Tang[10] went beyond that and defined generalized resolution and generalized minimum aberration for ranking designs. However, their method is only applicable to two-level factorial designs. In this paper, we extend the Deng and Tang[10] method, and we apply it to rank three-level screening designs.

3.1 | Calculating generalized resolution for screening designs

In this section, we apply the generalized resolution criterion to several three-level designs to evaluate and compare their performance under different models. We evaluate the designs using models that include just the main effects or main effects and interactions:

Model 1 (main effects)

$$y=\beta_0+\sum_{i=1}^{k}\beta_i x_i+\epsilon, \tag{2}$$

Model 2 (main effects and interactions)

$$y=\beta_0+\sum_{i=1}^{k}\beta_i x_i+\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\beta_{ij}x_{ij}+\epsilon, \tag{3}$$

where y is the response vector and x_i and $x_{ij}=x_i x_j$ are the columns that correspond to the main effects and two-factor interactions, respectively. β_0, β_1, and β_{ij} are unknown constant coefficients corresponding to the intercept, main effects, and two-factor interactions, respectively, whereas ϵ is the error vector with components ϵ_j being i.i.d. $N(0,\sigma^2)$.

For a screening designs D, r will be the smallest integer that achieves $max_{|T|=r}\, J_r(T)>0$, where J_r is defined as mentioned previously (1), and the maximization will be for all possible r subsets of the design columns.

The generalized resolution criterion will be taken from Deng and Tang[10]:

$$R(D)=r+[1-max\; J_r(T)/n]. \tag{4}$$

Clearly, $r\le R(D)<r+1$. Designs with higher generalized resolution are better than those with lower generalized resolution.

Example 1 In this example, we present the three-level screening designs DSD, WSD-2, and WSD-3 with 12 factors and 25 runs that are constructed from weighing matrices $W(12,11)$, $W(12,10)$, and $W(12,9)$, respectively, as these are shown on Table 2.

The generalized resolution of various examples are presented in Tables 3, 5, 7, and 8. In these tables, we have five columns where n is the number of runs, Design is the type of design, max J_r is the maximum value of the j characteristics among all subsets T of r distinct from columns of D, $R(D)$ is the generalized resolution, and Rep column include some representatives of the columns that constitute the projections that give the corresponding results.

TABLE 2 Orthogonal designs used for constructing 12 factors screening designs definitive screening design (DSD), WSD-2, and WSD-3

DSD

$$\left\{\begin{array}{cccccccccccc}
-1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 0 & -1 & 1 \\
-1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 0 & -1 & 1 & -1 \\
-1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 0 \\
-1 & -1 & 1 & 0 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 \\
-1 & -1 & 1 & 1 & 1 & -1 & 1 & 0 & -1 & 1 & -1 & -1 \\
-1 & 0 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 \\
-1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 0 & -1 \\
-1 & 1 & -1 & 1 & 0 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\
-1 & 1 & 0 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\
-1 & 1 & 1 & -1 & 1 & 0 & -1 & 1 & -1 & -1 & -1 & 1 \\
-1 & 1 & 1 & 1 & -1 & 1 & 0 & -1 & 1 & -1 & -1 & -1 \\
0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1
\end{array}\right\}$$

WSD-2

$$\left\{\begin{array}{cccccccccccc}
1 & 0 & 1 & -1 & 1 & 1 & -1 & 0 & -1 & -1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & -1 & 0 & -1 & -1 & 1 & 1 & -1 \\
0 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 0 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & 0 & 1 & -1 & 1 & -1 & 0 & -1 & -1 \\
-1 & -1 & 1 & 1 & 1 & 0 & 1 & -1 & -1 & -1 & -1 & 0 \\
-1 & 1 & -1 & 0 & 1 & 1 & -1 & -1 & 1 & -1 & 0 & -1 \\
1 & 0 & 1 & 1 & -1 & 1 & 1 & 0 & 1 & -1 & 1 & -1 \\
0 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & 0 & 1 & -1 & -1 \\
1 & 1 & 0 & 1 & 1 & -1 & 0 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & -1 & 0 & 1 & 1 & 1 & -1 & 1 & 1 & 0 & 1 \\
-1 & -1 & 1 & 1 & 1 & 0 & -1 & 1 & 1 & 1 & 1 & 0 \\
-1 & 1 & -1 & 1 & 0 & 1 & 1 & 1 & -1 & 0 & 1 & 1
\end{array}\right\}$$

WSD-3

$$\left\{\begin{array}{cccccccccccc}
1 & 0 & 1 & -1 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & -1 \\
1 & 1 & 0 & -1 & 1 & -1 & -1 & 1 & 0 & -1 & -1 & 0 \\
0 & 1 & 1 & 1 & -1 & -1 & 1 & 0 & -1 & -1 & 0 & -1 \\
1 & 1 & -1 & 1 & 0 & 1 & 1 & 0 & 1 & -1 & 0 & 1 \\
1 & -1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & -1 \\
-1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & -1 & 0 \\
0 & 1 & -1 & -1 & 0 & -1 & 1 & 0 & 1 & 1 & 1 & -1 \\
1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 0 & 1 & -1 & 1 \\
-1 & 0 & 1 & -1 & -1 & 0 & 0 & 1 & 1 & -1 & 1 & 1 \\
0 & 1 & 1 & 1 & 0 & -1 & -1 & -1 & 1 & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & 0 \\
1 & 0 & 1 & -1 & 1 & 0 & 1 & -1 & -1 & 0 & 1 & 1
\end{array}\right\}$$

TABLE 3 Generalized resolution for $m = 12$ factors with $n = 25$ runs

n	Design	max J_r	$R(D)$	Rep
				(1,3,4,9),(3,7,9,11),(6,7,9,10),(2,4,7,9),
				(6,8,10,11),(4,5,8,10),(4,5,7,9),(5,7,9,11),
				(7,10,11,12),(1,2,9,11),(1,4,6,12),(3,5,7,8),
25	DSD	8	4.6800	(1,4,5,11),(4,5,9,12),(3,8,9,11),(1,5,6,7),
				(1,2,8,9),(1,2,5,8),(2,6,10,11),(2,3,6,11),
				(1,5,6,11),(3,5,7,12),(5,8,9,11),(3,7,8,11),
				(5,6,7,8),(3,9,10,12),(2,6,8,9),(2,5,7,10),
				(4,5,7,8),(1,7,9,12),......
	WSD-2	16	4,36	(1,2,8,9),(4,5,10,11),(3,6,7,12)
				(3,4,8,11),(2,5,7,10),(1,7,9,12),(1,4,9,12),
				(1,3,7,10),(2,4,8,10),(2,6,9,12),(1,6,10,12),
	WSD-3	10	4.6	(2,4,7,11),(1,6,9,11),(1,6,8,12),(3,5,6,10),
				(5,7,10,12),(2,6,8,11),(3,5,7,11),(3,4,7,10),
				(5,6,9,12),(1,3,6,9),(2,4,9,11),(2,5,8,11),
				(3,5,7,9),(1,4,8,11),(3,5,8,10),(2,4,8,12)

For these three-level screening designs, $J_1(T) = J_2(T) = J_3(T) = 0$. The value of $R(D)$ can be a useful criterion to determine the projection properties when projecting onto r dimensions. Desirable designs have better projection properties when $R(D)$ is close to $r + 1$. To clarify, as shown in Table 3, generalized resolution is useful for assessing the screening designs based on the value of $max\ J_r$ and $R(D)$. Looking at the results, we could say that DSD with $max_{|T|=4}\ J_4(\mathrm{T}) = 8$ has better generalized resolution $R(D) = 4.68$ than the others. Designs WSD-2 and WSD-3 have $max_{|T|=4}\ J_4(T) = 16$ and 10, respectively, and lower generalized resolution $R(D)$, which is 4.36 and 4.6, respectively. However, in some cases, generalized resolution cannot discriminate between designs. When there are two or more designs with the same $R(D)$, as we illustrate in Example 2, then the minimum aberration criterion is used to distinguish the quality of the designs.

Example 2 In this example, we generate and study the DSD, WSD-2, and WSD-3 as their construction was illustrated in the previous section. These designs have at three levels, eight factors, and 17 runs. The orthogonal designs that are used, in the fold-over structure, to construct them are shown in Table 4.

In Table 5, all designs have the same four-dimensional projection properties based on j characteristics ($max_{|T|=4}$ $J_4(T) = 8$) and generalized resolution ($R(D) = 4.5294$). However, the projection properties of those designs are different. In Section 4, the above designs can be distinguished by using the minimum aberration criterion.

TABLE 4 Orthogonal designs used for constructing eight factors designs definitive screening design (DSD), WSD-2, and WSD-3

DSD	WSD-2	WSD-3
$\begin{bmatrix} -1 & -1 & -1 & 1 & -1 & 1 & 1 & 0 \\ -1 & -1 & 1 & -1 & 1 & 1 & 0 & -1 \\ -1 & -1 & 1 & 1 & 0 & -1 & -1 & 1 \\ -1 & 0 & -1 & -1 & 1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 & 1 & 0 & -1 & -1 \\ -1 & 1 & 0 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 0 & -1 & -1 & 1 & -1 \\ 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & -1 & 0 & -1 & 1 & -1 \\ -1 & 0 & 1 & 1 & -1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 & 1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 & 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -1 & -1 & 1 & 0 & 1 \\ -1 & -1 & -1 & 0 & 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 & 0 & -1 & -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ -1 & 0 & 0 & -1 & 0 & -1 & 1 & -1 \\ -1 & 0 & 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & -1 & -1 & 0 & 0 & 1 \\ -1 & -1 & -1 & 0 & 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 & 0 & -1 & -1 & 0 \end{bmatrix}$

TABLE 5 Generalized resolution for the designs with $m = 8$ factors and $n = 17$ runs

n	Design	max J_r	$R(D)$	Rep
				(2,4,6,8),(1,2,3,6),(1,4,6,7),(2,4,5,7),
				(2,4,5,6),(1,3,4,5),(1,4,6,8),(3,4,7,8),
				(2,3,5,8),(2,6,7,8),(2,3,5,7),(1,5,6,7),
17	DSD	8	4.5294	(3,4,5,6),(1,3,7,8),(1,2,4,7),(1,3,6,8),
				(1,2,3,4),(1,2,7,8),(1,2,5,6),(5,6,7,8),
				(2,3,6,7),(1,4,5,8),(1,2,5,8),(1,3,5,7),
				(3,5,6,8),(3,4,6,7),(4,5,7,8),(2,3,4,8)
	WSD-2	8	4.5294	(3,4,7,8),(3,4,5,6),(1,2,3,4),(1,2,7,8),
				(1,2,5,6),(5,6,7,8)
	WSD-3	8	4.5294	(1,2,3,4),(5,6,7,8)

Abbreviation: DSD, definitive screening design.

4 | MINIMUM ABERRATION CRITERION FOR THREE-LEVEL SCREENING DESIGNS

The minimum aberration criterion (MA) was introduced by Fries and Hunter[25] for ranking regular two-level designs. The definition of MA is only suitable for regular designs. The generators of regular design are used to calculate the MA criterion and that is the main reason that MA is not suitable for nonregular designs. Deng and Tang[10] and Tang and Deng[11]improved the definition of MA to meet these requirements. Clearly, the collection of subsets T of columns such that $J_r(T) = n$ for $r = \{4, \ldots, m\}$ is the defining relation of a regular factorial design D. In the defining relation, the r columns in T form a word of length r if $J_r(T) = n$. The number of words of length r is defined as $A_r(D)$, so the vector $W(D) = (A_3(D), \ldots\ldots, A_m(D))$ is a word length pattern for design D.

TABLE 6 Results for $m = 8$ factors with $n = 2m + 1 = 17$ runs

n	Design	r	J_r	Freq	CFV
		1	0	8	
		2	0	28	
	DSD	3	0	56	$[(0, 0, 0, 8)_1, (0, 0, 0, 28)_2, (0, 0, 0, 56)_3, (28, 0, 0, 42)_4]$
		4	0	42	
			8	28	
17	WSD-2	1	0	8	$[(0, 0, 0, 8)_1, (0, 0, 0, 28)_2, (0, 0, 0, 56)_3, (6, 24, 0, 40)_4]$
		2	0	28	
		3	0	56	
			0	40	
		4	4	24	
			8	6	
	WSD-3	1	0	8	$[(0, 0, 0, 8)_1,(0, 0, 0, 28)_2,(0, 0, 0, 56)_3,(2, 0, 32, 36)_4]$
		2	0	28	
		3	0	56	
			0	36	
		4	2	32	
			8	2	

Abbreviations: CFV, confounding frequency vector; DSD, definitive screening design.

The MA criterion can be applied to compare two-regular designs with the same resolution. Assume the word length patterns of two-regular designs D_1 and D_2, respectively are $W(D_1) = (A_3(D_1), ... A_m(D_1)$, $W(D_2) = (A_3(D_2), ... A_m(D_2)$.

When both D_1 and D_2 have the same resolution r, then D_1 is a better design than D_2 when it has a smaller number of words of length r. However, if $A_r(D_1) = A_r(D_2)$, then the $A_{r+1}(D_1)$ and $A_{r+1}(D_2)$ are used to check for differences between designs; otherwise, the process will continue until a different value appear or $r = m$. For more information about the results of MA of two-level designs, see Fries and Hunter[25]; Chen and Wu[26]; Tang and Wu[27]; and Cheng, Steinberg and Sun.[28]

Deng and Tang[10] applied the same idea to two-level nonregular factorial designs. They assumed that D_1 and D_2 are nonregular designs with the same generalized resolution $R(D_1) = R(D_2)$. The same generalized resolution of designs means the value of $maxJ_r$ are equal. Thus, if the frequency of J_r for deign D_1 is less than the frequency of J_r for design D_2, then D_1 is preferred. When the frequency of j characteristics of designs are the same, the r is selected to calculate the new J_r. If D_1 has lower value of the new J_r than D_2, then D_1 is preferred. Otherwise, the process is continue until the two designs can be distinguished or we decide they have exactly the same properties. Deng and Tang called this idea confounding frequency vectors (CFVs):

$$F = [(f_{31}, ..., f_{3t}); (f_{41} ...., f_{4t}); .....; (f_{m1}, ..., f_{mt})].$$

TABLE 7 Generalized resolution for designs with different number of runs ($n = 2m + 1$ and $r = 4$)

n	Design	max J_r	$R(D)$	Rep
13	DSD	4	4.6923	(2,3,4,6),(1,3,5,6),(2,3,5,6),(2,4,5,6),
				(1,2,4,5),(3,4,5,6),(2,3,4,5),(1,2,3,6),
				(1,2,4,6),(1,3,4,5),(1,4,5,6),(1,2,3,4),
				(1,2,5,6),(1,2,3,5),(1,3,4,6)
	WSD-2	4	4.6923	(2,3,5,6), (1,2,4,5), (1,3,4,6)
				(2,4,5,7),(2,5,8,9),(1,4,7,8),(3,4,7,10),
				(3,6,8,10),(1,3,9,10),(3,7,8,9),(4,5,6,8)
21	DSD	12	4.4286	,(1,6,7,10),(1,6,8,9),(1,2,8,10),(1,2,3,7),
				(3,4,5,9),(1,3,4,6),(2,6,7,8),(1,4,5,10),
				(1,3,5,8),(2,7,9,10),(2,4,6,10),(4,6,7,9),....
	WSD-2	12	4.4286	(3,5,8,10),(1,3,6,8),(2,4,7,9),(2,5,7,10),
				(1,4,6,9)
				(1,7,11,13),(3,8,10,12),(3,5,8,10),(7,10,12,13),
				(7,8,9,10),(1,6,7,9),(2,7,10,11),(1,2,3,9),
29	DSD	12	4.5862	(4,7,13,14),(3,4,10,13),(2,7,12,13),(6,9,11,12),
				(3,8,9,13),(2,6,8,12),(2,6,11,14),(6,7,9,13),
				(1,9,12,14),(1,4,10,11),(3,6,8,12),(8,11,12,14),...
				(3,4,10,11),(1,7,13,14),(1,2,4,7),(13,14,15,16),
				(4,6,11,15),(1,2,13,16),(1,2,9,12),(3,6,10,15),
33	DSD	24	4.2727	(6,8,10,12),(3,6,11,14),(3,7,10,14),(1,5,7,11),
				(3,9,13,14),(10,12,15,16),(3,9,10,12),(5,10,11,12),
				(2,5,12,13),(1,8,9,16),(10,12,13,14),(3,7,11,15),...
				(3,7,10,14),(3,7,11,15),(1,5,12,16),
	WSD-2	24	4.2727	(9,12,13,16),(1,5,9,13),(2,4,6,8),
				(10,11,14,15)
	WSD-3	24	4.2727	(13,14,15,16)

Abbreviation: DSD, definitive screening design.

Our criterion is a natural generalization of the criterion Deng and Tang[10] applied to two-level nonregular factorials. Let D be an orthogonal factorial design $n \times m$, with $n = 4t$. The frequency of r column combinations is f_{rj}, which gives $j_r = 4(t+1-j)$ for $j = 1, ..., t, t+1$. We have that $\sum_{j=1}^{t+1} f_{rj} = \binom{m}{r}$ and as $f_{1j} = f_{2j} = f_{3j} = 0$ for orthogonal screening designs, this helps in reducing the number of the f_{rj} we need to calculate for $r \geq 4$. The notation we use to define the CFV of D is

$$F = [(f_{41}, ..., f_{4t}); (f_{51} ...., f_{5t}); .....; (f_{m1}, ..., f_{mt})].$$

Let $f_i(D_1)$ and $f_i(D_2)$ be the i_{th} entries of the CFV of two designs D_1 and D_2, $i = 1, ..., (m-2)t$, and let i be the smallest integer such that $f_i(D_1) \neq f_i(D_2)$. If $f_i(D_1) < f_i(D_2)$, then D_1 has a less generalized aberration and hence is preferred.

The results are presented in Tables 6, 9, and 10. In these tables, we have six columns where n is the number of runs, Design is the type of design, J_r are the J characteristics calculated from Equation (1) using any subset T of r distinct columns of D, Freq is the number of different subsets of r distinct columns of D that give the same J_r, and CFV is the confounding frequency vectors. You can generate the CFV by looking at J_r and Freq.

Example 3 The three screening designs that are discussed in Example 2 have the same generalized resolution $R(D_1) = R(D_2) = R(D_3) = 4.5294$. The J_r for those three designs has four possible values (8, 4, 2, 0) listed in decreasing order. For all three designs, we have $J_1(D_i) = J_2(D_i) = J_3(D_i) = 0$. Counting the frequencies of J_r for $r = 1$, 2, and 3, we have $F_1(D_1)$

TABLE 8 Generalized resolution for designs with different number of runs ($n = 2m + 1$and $r = 4$)

n	Design	max J_r	$R(D)$	Rep
				(4,10,12,15),(6,10,17,18),(3,4,7,18),(4,9,12,17),
				(4,12,15,18),(1,5,7,14),(4,8,10,11),(1,6,12,17),
37	DSD	12	4.6757	(2,11,13,16),(5,7,17,18),(2,5,6,7),(3,7,14,16),
				(8,9,10,12),(1,4,6,16),(5,11,15,17),(9,12,13,15),
				(10,12,15,17),(3,6,12,17),(1,2,8,18),(2,6,13,18),...
				(2,6,11,15),(5,9,14,18),(1,5,10,14),
	WSD-2	28	4.2432	(4,9,13,18),(1,6,10,15),(3,8,12,17),
				(4,8,13,17),(3,7,12,16)
				(1,8,10,12),(5,11,15,19),(2,12,14,20),(4,13,15,20),
				(3,7,13,18),(1,7,10,19),(12,14,16,18),(9,11,16,18),
41	DSD	24	4.4146	(6,8,12,14),(1,7,13,20),(1,7,8,9),(3,5,8,15),
				(10,11,15,18),(1,7,14,17),(5,15,16,20),(3,7,14,15),
				(1,4,6,8),(3,4,5,12),(2,3,7,11),(2,6,19,20),...
	WSD-2	24	4.4146	(8,9,16,17),(2,8,15,18),(1,7,13,20),
				(4,5,14,15),(4,10,12,16),(2,8,12,17),
				(1,3,17,19),(3,9,15,19),(6,10,14,15),
				(5,6,13,16),(5,6,14,19)
				(12,14,16,18),(1,7,13,20),(4,5,14,15),
				(1,7,14,17),(8,12,13,16),(2,5,16,18),
	WSD-3	16	4.4839	(7,15,19,20),(5,6,15,17),(1,2,13,19),
				(10,14,17,18),(5,6,13,16),(2,4,12,18),
				(4,5,11,17)
45	WSD-2	24	4.4667	(4,5,16,17),(6,7,18,19),(8,9,20,21),
				(10,11,12,22),(3,4,15,16),(1,11,12,13),
				(7,8,19,20),(1,2,13,14),(5,6,17,18),
				(9,10,21,22)

Abbreviation: DSD, definitive screening design.

$= (0,0,0,8)$, $F_2(D_1) = (0,0,0,28)$, $F_3(D_1) = (0,0,0,56)$. As shown in Table 6, all designs have the same F_1,F_2,F_3, and the difference appears in F_4 where D_1 has $maxJ_r = 8$ with 28 frequency so the CFV of D_1 is

$$F(D_1) = [(0,0,0,8)_1, (0,0,0,28)_2, (0,0,0,56)_3, (28,0,0,42)_4].$$

$J_r(D_2)$ and $J_r(D_3)$ are calculated in a similar way. The CFV of D_2 and D_3 is

$$F(D_2) = [(0,0,0,8)_1, (0,0,0,28)_2, (0,0,0,56)_3, (6,24,0,40)_4],$$

$$F(D_3) = [(0,0,0,8)_1, (0,0,0,28)_2, (0,0,0,56)_3, (2,0,32,36)_4].$$

In Table 6, the comparison of the three designs shows that the WSD-3 is preferred as it has less F_4 in its CFV.

5 | RESULTS

Generalized resolution for each design is presented in Tables 7 and 8. We observe that the designs with the same number of runs have the same resolution for all cases with runs up to 33. This is expected because the three designs have the same $maxJ_r$. Thus, we need to distinguish designs by using CFV criterion, as shown in Tables 9 and 10. Note that it is not always possible to test all three designs because their existence is subject to the parameters. For example, when $n = 29$ and 45, some designs cannot exist. In this situation, no comparison is needed because there is only one design.

TABLE 9 Results for designs with $n = 2m + 1$, for $n = 13$ and $n = 21$

n	Design	r	J_r	Freq	CFV
	DSD	1	0	6	$[(0, 6)_1, (0, 15)_2, (0, 20)_3, (15, 0)_4]$
		2	0	15	
		3	0	20	
		4	4	15	
13		1	0	6	$[(0, 6)_1, (0, 15)_2, (0, 20)_3, (3, 12)_4]$
		2	0	15	
	WSD-2	3	0	20	
		4	0	12	
			4	3	
21		1	0	10	$[(0, 0, 0, 10)_1, (0, 0, 0, 45)_2, (0, 0, 0, 120)_3, (30, 0, 180, 0)_4]$
		2	0	45	
	DSD	3	0	120	
		4	4	180	
			12	30	
21		1	0	10	$[(0, 0, 0, 10)_1, (0, 0, 0, 45)_2, (0, 0, 0, 120)_3, (5, 20, 125, 60)_4]$
		2	0	45	
		3	0	120	
	WSD-2	4	0	60	
			4	125	
			8	20	
			12	5	

Abbreviations: CFV, confounding frequency vector; DSD, definitive screening design.

Moreover, the generalized resolution criterion is able to rank designs in only few cases. For example, the DSD has better generalized resolution ($R(D) = 4.67$) than the WSD-2 when $n = 37$. When $n = 41$, we can see that the WSD-3 has slightly better generalized resolution than the others.

6 | PROJECTION ESTIMATION CAPACITY

The generalized resolution and confounding frequency vector approaches are standard ways to compare two-level screening designs, and the extension to three-level designs is useful. However, it should be noted that none of these diagnostics can tell the practitioner what models are identifiable. Loeppky et al.[29] introduced projection estimation capacity (PEC) criterion that is closely related to the estimation capacity (EC) criterion initiated by Sun.[30] The aim of both criteria is to count the maximum number of estimable models of a specific design. The model space in EC criterion includes models with all k-main effects and some selected two-factors interactions. In PEC criterion, the model space includes all models that have some of the k-main effects and all their two-level interactions. To investigate the

TABLE 10 Results for designs with $n = 2m + 1 = 33$ runs

Design	r	J_r	Freq	CFV
DSD	1	0	16	$[(0, 0, 0, 0, 0, 0, 0, 0, 0, 16)_1, \ldots, (56, 0, 0, 168, 0, 0, 504, 0, 0, 1092)_4]$
	2	0	120	
	3	0	560	
	4	0	1092	
		8	504	
		16	168	
		24	56	
WSD-2	1	0	16	$[(0, 0, 0, 0, 0, 0, 0, 0, 0, 16)_1, \ldots, (7, 24, 0, 116, 0, 48, 381, 264, 0, 980)_4]$
	2	0	120	
	3	0	560	
		0	980	
		4	264	
		8	381	
	4	12	48	
		16	116	
		20	24	
		24	7	
WSD-3	1	0	16	$[(0, 0, 0, 0, 0, 0, 0, 0, 0, 16)_1, \ldots, (1, 0, 12, 54, 12, 48, 307, 528, 168, 690)_4]$
	2	0	120	
	3	0	560	
		0	690	
		2	168	
		4	528	
		8	307	
	4	12	48	
		14	12	
		16	54	
		18	12	
		24	1	

Abbreviations: CFV, confounding frequency vector; DSD, definitive screening design.

estimable number of models, we define D to be an $n \times m$ design matrix, we need to investigate all possible models that include k-main effects and all the 2-fi of these k effects. To clarify,

$$P_k(D) = \frac{\rho_k(D)}{\binom{m}{k}},$$

where$\rho_k(D)$ is the number of estimable models that contain k main effects with their associated two-level interactions. $\binom{m}{k}$ is the number of all possible of selections of k out of the m columns.

$(p_1, p_2, \ldots, p_m)$ is known as the PEC sequence of D.

Loeppky et al.[29] listed some key features for PEC criterion that we summarize in the following points:

(1) PEC criterion can be applied to both regular and nonregular, designs.
(2) PEC provides sufficient information about main effects and their two-factor interactions prior to the experiment.
(3) PEC can be used to compare designs with different number of factors and run sizes.

From the definition of PEC, it is easy to see that a design is desirable to have a large $P_k(D)$. Thus, a design with the highest values in the PEC sequence is preferred because it has the maximum projection estimation capacity (MPEC). Any model is not estimable if it has more parameters than degrees of freedom (trivial case). Cheng et al.[28] observed that under the EC criterion, the minimum aberration criterion can perform very well and that provides a statistical justification for MA in terms of the model robustness.

The projection estimation capacity (PEC) makes a step towards the evaluation and comparison of the designs based on the number of models they can identify. However, as it was noted by a reviewer, it is rare that all the interactions of a subset of factors are active. So, the PEC is perhaps a bit conservative.

In the next example, we evaluate the designs by using the PEC criterion.

Example 4 Suppose an experimenter want to investigate a range of effects for six factors. He applied DSD and WSD-2 to design the experiment. The aim was to estimate the main effects and their two-level interactions. As shown in Table 11, all models of the DSD and the WSD-2 are estimable when $k = 1$ and $k = 2$ factors. Thus, both designs can provide the same information if $k = 1$ or 2. However, DSD is preferable to WSD-2, in terms of PEC, because all three main effects and their two-levels interactions ($k = 3$) are estimable whether WSD-2 does not allow the estimation of all models that include three factors and all their corresponded two-factors interactions. Thus, DSD is the MPEC.

TABLE 11 Projection estimation capacity DSD and WSD-2 with $m = 6$ factors

Design	k	$\rho_k(D)$	$\binom{m}{k}$	$P_k(D)$	(PEC) sequence
DSD	1	6	6	1	(1,1,1,0.93,0,0)
	2	15	15	1	
	3	20	20	1	
	4	14	15	0.933	
	5	0	6	0	
	6	0	1	0	
WSD-2	1	6	6	1	(1,1,0.4,0,0,0)
	2	15	15	1	
	3	8	20	0.4	
	4	0	15	0	
	5	0	6	0	
	6	0	1	0	

Abbreviations: DSD, definitive screening design; PEC, projection estimation capacity.

For eight factors, we can construct all the three designs under consideration (DSD, WSD-2, WSD-3). The results of this comparison are shown in Table 12. For $k = 1$, $k = 2$ and $k = 3$ factors, we have the same PEC. However, if $k > 3$, then DSD is the MPEC based on the PEC sequence. In addition, we observe that WSD-3 is much better than WSD-2 with respect to the PEC sequence. Tables 13 and 14 show that DSD is preferred designs, but WSD-3 gives more information than WSD-2 when $m = 12$. It is obvious that there is a positive relationship between design size and PEC sequence because the variation between designs on PEC sequence becomes small when m is increasing (for $m = 6$, 8, 10, 12, 16, and 20, see Tables 11 to 16, respectively).

TABLE 12 Projection estimation capacity for $m = 8$ factors

Design	k	$\rho_k(D)$	$\binom{m}{k}$	$P_k(D)$	(PEC) sequence
DSD	1	8	8	1	(1,1,1,1,0,0)
	2	28	28	1	
	3	56	56	1	
	4	70	70	1	
	5	0	56	0	
	6	0	28	0	
WSD-2	1	8	8	1	(1,1,1,0.74,0,0)
	2	28	28	1	
	3	56	56	1	
	4	52	70	0.742857	
	5	0	56	0	
	6	0	28	0	
WSD-3	1	8	8	1	(1,1,1,0.81,0,0)
	2	28	28	1	
	3	56	56	1	
	4	57	70	0.814286	
	5	0	56	0	
	6	0	28	0	

Abbreviations: DSD, definitive screening design; PEC, projection estimation capacity.

TABLE 13 Projection estimation capacity for $m = 10$ factors

Design	k	$\rho_k(D)$	$\binom{m}{k}$	$P_k(D)$	(PEC) sequence
DSD	1	10	10	1	(1,1,1,1,0.94,0)
	2	45	45	1	
	3	120	120	1	
	4	210	210	1	
	5	238	252	0.944444	
	6	0	210	0	
WSD-2	1	10	10	1	(1,1,1,0.95,0.68,0)
	2	45	45	1	
	3	120	120	1	
	4	200	210	0.952381	
	5	173	252	0.686508	
	6	0	210	0	

Abbreviations: DSD, definitive screening design; PEC, projection estimation capacity.

TABLE 14 Projection estimation capacity for $m = 12$ factors

Design	k	$\rho_k(D)$	$\binom{m}{k}$	$P_k(D)$	(PEC) sequence
DSD	1	12	12	1	(1,1,1,1,1,0)
	2	66	66	1	
	3	220	220	1	
	4	495	495	1	
	5	792	792	1	
	6	0	924	0	
WSD-2	1	12	12	1	(1,1,1,0.99,0.91,0)
	2	66	66	1	
	3	220	220	1	
	4	492	495	0.993939	
	5	722	792	0.911616	
	6	0	924	0	
WSD-3	1	12	12	1	(1,1,1,1,0.97,0)
	2	66	66	1	
	3	220	220	1	
	4	495	495	1	
	5	774	792	0.977273	
	6	0	924	0	

Abbreviations: DSD, definitive screening design; PEC, projection estimation capacity.

TABLE 15 Projection estimation capacity for $m = 16$ factors

Design	k	$\rho_k(D)$	$\binom{m}{k}$	$P_k(D)$	(PEC) sequence
DSD	1	16	16	1	(1,1,1,1,1,0.93)
	2	120	120	1	
	3	560	560	1	
	4	1820	1820	1	
	5	4368	4368	1	
	6	7489	8008	0.93519	
WSD-2	1	16	16	1	(1,1,1,0.99,0.97,0.9)
	2	120	120	1	
	3	560	560	1	
	4	1813	1820	0.996154	
	5	4272	4368	0.978022	
	6	7253	8008	0.905719	
WSD-3	1	16	16	1	(1,1,1,0.99,0.99,0.96)
	2	120	120	1	
	3	560	560	1	
	4	1819	1820	0.999451	
	5	4353	4368	0.996566	
	6	7763	8008	0.969406	

Abbreviations: DSD, definitive screening design; PEC, projection estimation capacity.

TABLE 16 Projection estimation capacity for $m = 20$ factors

Design	k	$\rho_k(D)$	$\binom{m}{k}$	$P_k(D)$	(PEC) sequence
DSD	1	20	20	1	(1,1,1,1,1,0.999)
	2	190	190	1	
	3	1140	1140	1	
	4	4845	4845	1	
	5	15504	15504	1	
	6	38736	38760	0.999	
WSD-2	1	20	20	1	(1,1,1,1,1,0.999)
	2	190	190	1	
	3	1140	1140	1	
	4	4845	4845	1	
	5	15 504	15 504	1	
	6	38 758	38 760	0.999	
WSD-3	1	20	20	1	(1,1,1,1,1,0.999)
	2	190	190	1	
	3	1140	1140	1	
	4	4845	4845	1	
	5	15 504	15 504	1	
	6	38 759	38 760	0.999	

Abbreviations: DSD, definitive screening design; PEC, projection estimation capacity.

7 | CONCLUSION

One problem for three-level screening designs is to identify the best design with a prespecified criterion and a given number of runs. Generalized resolution and minimum aberration criteria are widely used for comparing designs since they were introduced by Box and Hunter.[21] Deng and Tang[10] extended this work to also work on two-level nonregular designs. We go beyond these criteria, while using them as a guide for assessing three-level screening designs. The examples show that there are multiple ways to evaluate three-level designs.

In this paper, we used the generalized resolution criterion and the minimum aberration criterion when the designs have the same generalized resolution. Further, we used the PEC criterion that was described by Loeppky et al.[29] The PEC criterion has several advantages that make it desirable for evaluating designs. However, two models in different projections may both be estimable but confounded with each other. This is especially problematic as the number of factors in each subset becomes large. Note that if two models are confounded, then there is no way to discriminate between them even though the PEC will include those two models in the number of estimable models.

Three different type of designs from the recent literature were evaluated with respect to the above criteria and for specific run sizes of the design under consideration. All the results and the examples presented here indicate that in all the tested cases, with any number of main effects and two factor interactions, the DSDs are superior of the WSD-2 and WSD-3. Practically, we could suggest the use of DSD over WSD-2 and WSD-3 in all those cases.

The WSD designs and their ability to estimate quadratic effects were discussed in detail in a recent paper of the authors but by using the alphabetic optimality criteria D, G, and A with models including quadratic effects (see Mohammed et al.[31]). Because in this paper, the focus is in developing the new criteria for three level designs and used them to compare the efficiency of the designs mainly for screening purposes (with also active interactions), the inclusion of quadratic effects was omitted. Note that there are cases where the WSDs are preferable with respect to the alphabetical optimality criteria, when quadratic effects are included (see Mohammed et al.[31]).

ACKNOWLEDGMENTS

The authors thank the referees and the associate editor for their valuable comments and suggestions that helped in improving the quality and presentation of this paper.

ORCID

Stelios Georgiou https://orcid.org/0000-0001-5096-0067
Stella Stylianou https://orcid.org/0000-0003-3677-2830

REFERENCES

1. Montgomery DC. *Design and Analysis of Experiments*. 9th Edition. New York: John Wiley & Sons; 2017.
2. Dean A, Lewis S. *Screening: Methods for Experimentation in Industry, Drug Discovery, and Genetics*. Springer-Verlag New York: Springer Science & Business Media; 2006. https://www.springer.com/gp/book/9780387280134
3. Plackett RL, Burman JP. The design of optimum multifactorial experiments. *Biometrika*. 1946;33(4):305-325.
4. Hamada M, Wu J. *Experiments: Planning, Analysis, and Parameter Design Optimization*. New York: Wiley; 2000.
5. Hamada M, Wu CFJ. Analysis of designed experiments with complex aliasing. *J Qual Technol*. 1992;24(3):130-137.
6. Lin DKJ, Draper NR. Projection properties of plackett and burrnan designs. *Technometrics*. 1992;34(4):423-428.
7. Wang JC, Wu CFJ. A hidden projection property of plackett-burman and related designs. *Stat Sin*. 1995;5:235-250.
8. Cheng C-S. Some projection properties of orthogonal arrays. *Annals Stat*. 1995;23(4):1223-1233.
9. Cheng C-S. Some hidden projection properties of orthogonal arrays with strength three. *Biometrika*. 1998;85(2):491-495.
10. Deng L-Y, Tang B. Generalized resolution and minimum aberration criteria for plackett-burman and other nonregular factorial designs. *Stat Sin*. 1999;9(4):1071-1082.
11. Tang B, Deng L-Y. Minimum g2-aberration for nonregular fractional factorial designs. *Annals Stat*. 1999;27(6):1914-1926.
12. Cheng S-W, Wu CFJ. Factor screening and response surface exploration. *Stat Sin*. 2001;11:553-580.
13. Xu H, Cheng S-W, Wu CFJ. Optimal projective three-level designs for factor screening and interaction detection. *Technometrics*. 2004; 46(3):280-292.
14. Tsai P-W, Gilmour SG, Mead R. Projective three-level main effects designs robust to model uncertainty. *Biometrika*. 2000;87(2):467-475.
15. Tsai P-W, Gilmour SG, Mead R. Some new three-level orthogonal main effects plans robust to model uncertainty. *Stat Sin*. 2004;14: 1075-1084.
16. Evangelaras H, Koukouvinos C, Lappas E. 18-run nonisomorphic three level orthogonal arrays. *Metrika*. 2007;66(1):31-37.
17. Dey A. Projection properties of some orthogonal arrays. *Stat Probab Lett*. 2005;75(4):298-306.
18. Jones B, Nachtsheim CJ. A class of three-level designs for definitive screening in the presence of second-order effects. *J Qual Technol*. 2011;43(1):1-15.
19. Xiao L, Lin DKJ, Bai F. Constructing definitive screening designs using conference matrices. *J Qual Technol*. 2012;44(1):2-8.
20. Georgiou SD, Stylianou S, Aggarwal M. Efficient three-level screening designs using weighing matrices. *Statistics*. 2014;48(4):815-833.
21. Box GEP, Hunter JS. The 2 k–p fractional factorial designs. *Technometrics*. 1961;3(3):311-351.
22. Webb SR. Characterization of non-orthogonal incomplete factorial designs. *Appendix B ARL*. 1964:65-116.
23. Rechtschaffner RL. Saturated fractions of 2 n and 3 n factorial designs. *Technometrics*. 1967;9(4):569-575.
24. Srivastava JN, Chopra DV. Balanced optimal 2m fractional factorial designs of resolution v, m ≤ 6. *Technometrics*. 1971;13(2):257-269.
25. Fries A, Hunter WG. Minimum aberration 2 k–p designs. *Technometrics*. 1980;22(4):601-608.
26. Chen J, Wu CFJ. Some results on s^{n-k} fractional factorial designs with minimum aberration or optimal moments. *Annals Stat*. 1991; 19(2):1028-1041.
27. Tang B, Wu CFJ. Characterization of minimum aberration 2^{n-k} designs in terms of their complementary designs. *Annals Stat*. 1996; 24(6):2549-2559.
28. Cheng C-S, Steinberg DM, Sun DX. Minimum aberration and model robustness for two-level fractional factorial designs. *J Royal Stat Soc Ser B (Stat Methodol)*. 1999;61(1):85-93.
29. Loeppky JL, Sitter RR, Tang B. Nonregular designs with desirable projection properties. *Technometrics*. 2007;49(4):454-467.
30. Sun DX. Estimation capacity and related topics in experimental design. university of waterloo. *Ph.D. Thesis*; 1993.
31. Mohammed A, Georgiou S, Aggarwal M. Projection properties of three-level screening designs. *Aust N Z J Stat*. to appear.

AUTHOR BIOGRAPHIES

Mohammed Alomair received the BS degree in mathematics from King Faisal University, Hofuf, Saudi Arabia, in 2008 and the MS degree in applied statistics from University of Northern Colorado, Colorado, USA, in 2015. He is current pursuing the PhD degree in statistics at RMIT University, Melbourne, VIC, Australia. From 2010 to 2012, he was a Research Assistant with the Imam Muhammad ibn Saud Islamic University, Saudi Arabia. His research interest includes screening, weighing designs, and projection properties.

Stelios Georgiou is currently an Associate Professor with the School of Science, RMIT University, Melbourne, Australia. He received the PhD degree in combinatorics and Statistics from the NTUA, Athens, Greece in 2003. He has authored over 80 papers in top journals and prestigious conferences. His research interests include statistical experimental designs, biostatistics, data analysis, statistical quality control, and combinatorial designs.

Stella Stylianou is currently an Assistant Professor with the School of Science, RMIT University, Melbourne, Australia. She received her PhD degree in Statistics from the NTUA, Athens, Greece in 2005. She has authored over 30 papers in top journals and prestigious conferences. Her research interests include statistical experimental designs, biostatistics, statistical quality control, and data analysis.

How to cite this article: Alomair M, Georgiou S, Stylianou S. Three-level designs: Evaluation and comparison for screening purposes. *Qual Reliab Engng Int*. 2021;37:90–107. https://doi.org/10.1002/qre.2722