

Multiway Blocking of Designs of Experiments

Nam-Ky Nguyen¹, Tung-Dinh Pham², Mai Phuong Vuong³

¹*Vietnam Institute for Advanced Study in Mathematics, Hanoi, Vietnam*

²*VNU University of Science, Vietnam National University, Hanoi*

³*Hanoi University of Science & Technology, Vietnam*

Received: 07 January 2021; Revised: 05 February 2021; Accepted: 09 February 2021

Abstract

Fisher's three R's or three principles of designs of experiments are (i) Randomisation; (ii) Replications; and (iii) Local control or blocking (also called noise reduction). Of the three, blocking is the most difficult. Works on blocked 3-level designs are very limited. In addition, there might be more than one extraneous variations or blocking factors. As such, there is a need for a general method to do multiway blocking of experimental designs. This paper extends the idea of orthogonal blocking of Box and Hunter (1957) from one blocking factor to several blocking factors. It then presents a blocking algorithm which can impose several blocking/noise factors on popular experimental designs. Particular attention will be given to 2-level, 3-level and mixed-level screening designs such as those introduced by Jones and Nachtsheim (2013) and Nguyen *et al.* (2020).

Key words: Box-Behnken designs; Central composite designs; Definitive screening designs; Fractional factorial designs; Orthogonal blocking.

1. Introduction

An experiment is conducted to determine the relationship between input factors affecting a process and the output of that process. There are controllable input factors to be studied as well as nuisance (uncontrollable) ones to be eliminated. While the former can be modified to optimise the output, this is not the case of the latter. Examples of uncontrollable factors are: (i) different batches of raw material; (ii) different machine; (iii) different operators; (iv) different locations; (v) different times, *etc.* A well-designed experiment minimises the effects of these uncontrollable factors by partitioning the set of experimental runs into more homogeneous subsets. This noise reduction exercise is called local control or blocking. It makes experiments more sensitive in detecting significant effects and hence less experimentation may be required. Examples of the scenarios when designs of more than one blocking factors can be used are given below.

Example 1: A 2^5 factorial experiment to identify interaction effects for different additives in linear low-density polyethylene film (Hoang *et al.*, 2004, Mee, 2009, p. 79). The factors and levels are: (A) Antioxidant A (ppm), 0 and 400; (B) Antioxidant B (ppm), 0 and 400; (C) Acid Scavenger (ppm), 0 and 1000; (D) Anti-block agent (ppm), 0 and 2000; (E) Slip

additive (ppm), 0 and 800. As the full 2^5 factorial would take at least three days to complete, it was divided into four blocks. Let's assume the experimenter wishes to add to the model an additional block factor, *i.e.* times of the day (8AM and 2PM) and find a suitable design which can accommodate the new blocking factor.

Example 2: A 9-factor DSD in 21 runs was conducted to investigate the oxidation reactions in homogeneous Co^{2+} /PMS system (Zhang *et al.*, 2018). The objective is to evaluate the suitability of the DSD approach in optimising the operating parameters of Co^{2+} /PMS system and to identify the significant effects involved in the reaction system. See Jones and Nachtsheim (2011) for the use of DSD as a screening design. The nine factors in this experiment are: (1) NaCl, (2) NaH_2PO_4 , (3) NaHCO_3 , (4) NaNO_3 , (5) Na_2SO_4 , (6) HA, (7) PMS, (8) AO II, and (9) Co^{2+} . The first five factors were set at 0, 10, and 20mM, HA at 0, 20, 40mg dm^{-3} , PMS at 2, 6, 10mM. AO II at 50, 75, 100 mg dm^{-3} and Co^{2+} at 0, 0.68, and 1.36mM. Zhang *et al.* (2018) stated that they could not use response surface designs (RSDs), such as the Box-Behnken designs (BBD), the Doehlert design and the central composite design, as they could not afford the enormous number of runs required by these designs. At the same time, the popular Plackett-Burman design is unable to capture the quadratic and interaction effects. Let's assume the experiment was performed in two different reactors and three different days and the experimenters wish to add these two blocking factors to the model.

2. Conditions for Orthogonal Blocks

Consider the following model for an n -run design with m factors x_1, \dots, x_m , out of which m_3 factors are at 3-level and the rest are at 2-level, arranged in b blocks:

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\delta} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (1)$$

where \mathbf{y} is an $n \times 1$ column vector of response, \mathbf{Z} an $n \times b$ matrix containing b dummy variables, $\boldsymbol{\delta}$ a $b \times 1$ column vector representing block effects, \mathbf{X} is an $n \times p$ expanded design matrix, $\boldsymbol{\beta}$ a $p \times 1$ column vector of parameters to be estimated, and $\boldsymbol{\epsilon}$ an $n \times 1$ column vector of random errors assumed to have zero mean and variance σ^2 . \mathbf{X} includes a column of 1's, representing the intercept; m columns representing the main-effects (MEs); and depending on the model might also include $\binom{m}{2}$ columns representing the 2-factor interactions (2FIs) and m_3 columns representing the second-order effects (SOEs).

Following Box and Hunter 1957, Section 8 and Khuri and Cornell, 1996, Chapter 8, we scale the columns of \mathbf{Z} by subtracting the value of each column from the column mean. Equation (1) can then be written as:

$$\mathbf{y} = \tilde{\mathbf{Z}}\boldsymbol{\delta} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (2)$$

The condition for orthogonal blocks can be written as:

$$\tilde{\mathbf{Z}}'\mathbf{X} = \mathbf{0}. \quad (3)$$

As the row sum of $\tilde{\mathbf{Z}}$ in (2) is zero, which is an example of perfect multicollinearity, to avoid the singular data matrix, we drop the last column of $\tilde{\mathbf{Z}}$ and the last element of $\boldsymbol{\delta}$ in

(2). The least square solution for the unknown parameters δ and β in (2) is the solution of the following equation:

$$\begin{pmatrix} \tilde{\mathbf{Z}}' \\ \mathbf{X}' \end{pmatrix} \mathbf{y} = \begin{pmatrix} \tilde{\mathbf{Z}}'\tilde{\mathbf{Z}} & \tilde{\mathbf{Z}}'\mathbf{X} \\ \mathbf{X}'\tilde{\mathbf{Z}} & \mathbf{X}'\mathbf{X} \end{pmatrix} \begin{pmatrix} \hat{\delta} \\ \hat{\beta} \end{pmatrix} \quad (4)$$

When the orthogonal block condition in Equation (3) is satisfied, it can be seen that the solution for β from (4) will be the same as the one from the equation $\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{X}\hat{\beta}$, *i.e.* the equation for a model without blocking.

Remarks

1. When there are r blocking factors, the matrix $\tilde{\mathbf{Z}}$ in (2) can be partitioned as $(\tilde{\mathbf{Z}}_1 \dots \tilde{\mathbf{Z}}_r)$, where $\tilde{\mathbf{Z}}_l$ ($l = 1, \dots, r$) is matrix of size $n \times (b_l - 1)$ and b_l the settings of the blocking factor l .

2. Let \mathbf{x}'_i and \mathbf{x}'_u be two rows of \mathbf{X} . Let $\tilde{\mathbf{z}}'_i$ and $\tilde{\mathbf{z}}'_u$ be the corresponding row vectors of $\tilde{\mathbf{Z}}$. Swapping the i th and u th row of \mathbf{X} is the same as adding the following matrix to $\tilde{\mathbf{Z}}'\mathbf{X}$:

$$-(\tilde{\mathbf{z}}'_i - \tilde{\mathbf{z}}'_u)(\mathbf{x}_i - \mathbf{x}_u)'. \quad (5)$$

We use this matrix result to develop BLOCK, an algorithm for blocking various types of experimental designs, including DSDs and DSD-based mixed-level designs.

3. Two Steps of The BLOCK Algorithm

Here are two steps of BLOCK with r blocking factors using the results in Equation (5):

1. Allocate the n runs of the unblocked design to the blocking factors randomly. Calculate f , the sum of squares of the elements of $\tilde{\mathbf{Z}}'\mathbf{X}$.

2. Repeat searching for a pair of runs in different blocks such that the swap of the run positions results in the biggest reduction in f . If the search is successful, swap their positions, update f and $\tilde{\mathbf{Z}}'\mathbf{X}$. This step is repeated until $f=0$ or until f cannot be reduced further.

Each computer try has these two steps. Several tries are required for each design and the one with the smallest f will be chosen. For designs with the same f , the one with the smallest block factor (BF) will be chosen where BF is calculated as:

$$BF = (|\mathbf{X}'\mathbf{X}|/(|\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}}||\mathbf{X}'\mathbf{X}|))^{1/(p-v)} \quad (6)$$

Here $\mathbf{X} = (\tilde{\mathbf{Z}} \mathbf{X})$, p is the number of parameters to be estimated and $v = \sum_{l=1}^r (b_l - 1)$ is the degree of freedom associated with the r blocking factors or columns of $\tilde{\mathbf{Z}}$. Clearly, BF equals 1 means the design is orthogonally blocked.

Remarks

1. For a factorial or fractional factorial design (FFD), the orthogonality between block variables and MEs is considered more important than the one between the former and 2FIs. For a 3-level or a mixed-level screening design, the orthogonality between block variables and MEs is considered more important than the one between the former and SOEs. In these situations, partition \mathbf{X} as $(\mathbf{X}_1 \ \mathbf{X}_2)$ where \mathbf{X}_1 is associated with the more important effects and partition $\tilde{\mathbf{Z}}'\mathbf{X}$ as $(\tilde{\mathbf{Z}}'\mathbf{X}_1 \ \tilde{\mathbf{Z}}'\mathbf{X}_2)$. Let g be the sum of squares of the elements of $\tilde{\mathbf{Z}}'\mathbf{X}_1$ and f the sum of squares of the elements of $\tilde{\mathbf{Z}}'\mathbf{X}$ as defined previously. g and f will then be used as the primary and secondary objective functions respectively.

2. BLOCK can be implemented sequentially, *i.e.* at each step, a blocking factor is added to the design. We use this sequential approach to block very large designs with several blocking factors.

3. In a sense, our blocking algorithm is an extension of one of Nguyen (2001), which only works with one blocking factors, to several blocking factors. Ours is more general than the one of Gilmour and Trinca (2003), which only work with two blocking factors. Clearly, ours does not require matrix inversions and therefore is considered faster (and less prone to get trapped in the local optima) than the ones by other authors (See *e.g.* Cook and Nachtsheim, 1989; Gilmour and Trinca, 2003 and Jones and Nachtsheim, 2016).

4. Discussion

In the followings, we will show the solutions for the two designs problems mentioned in the Introduction.

Example 1: Table 1 shows how a 2^5 factorial can be blocked using two blocking factors (day and time of the day) and the interaction model which includes the MEs and 2fi's terms. Our constructed design is an orthogonally blocked one, meaning the factors **A**, **B**, **C**, **D** and **E** are orthogonal to both days and times of day (8AM and 2PM).

Example 2: Table 2 (a) shows the layout of a 9-factor DSD with two blocking factors, *i.e.* reactor and day, using the pure quadratic model which includes the SOE terms and MEs terms. Three centre runs have been added to the original 21-run DSD. For this design, $g = 0$, $f = 29$ and $BF = 0.807$. Now let us assume that the unblocked design for the above experiment is a 24-run mixed-level design with the first five factors at 3-level and the rest are at 2-level constructed from a Hadamard matrix of order 12 (see Nguyen *et al.*, 2020). Table 2 (b) shows the layout the blocked design. For this design, $g = 0$, $f = 7.11$ and $BF = 0.928$.

To visualise the confounding patterns of blocked designs we use the correlation cell plots (CCPs). These CCPs, proposed by Jones and Nachtsheim (2011), display the magnitude of the correlation between the blocking factors and the MEs, SOEs (of 3-level factors) and 2FIs of the designs under study. The colour of each cell in these plots ranges from white (no correlation) to dark (correlation of 1 or close to 1).

The two CCPs in Figures 1 (a) and 1 (b) show the confounding patterns of the two designs in Tables 2 (a) and 2 (b). It can be seen the first blocking factor (reactor) is orthogonal to both MEs and SOEs in Figure 1 (b) but is only orthogonal to MEs in Figure

Table 1: A 2^5 factorial with two blocking factors: day and time of the day. The low and high levels are coded as -1 and 1

Day	Time	A	B	C	D	E	Day	Time	A	B	C	D	E
1	1	1	1	-1	1	-1	3	1	1	-1	-1	-1	-1
1	1	-1	1	1	-1	-1	3	1	-1	-1	1	1	-1
1	1	-1	-1	-1	-1	1	3	1	1	1	1	-1	1
1	1	1	-1	1	1	1	3	1	-1	1	-1	1	1
1	2	1	-1	1	-1	-1	3	2	1	1	1	1	-1
1	2	-1	1	1	1	1	3	2	-1	1	-1	-1	-1
1	2	1	1	-1	-1	1	3	2	-1	-1	1	-1	1
1	2	-1	-1	-1	1	-1	3	2	1	-1	-1	1	1
2	1	-1	-1	-1	-1	-1	4	1	-1	1	1	-1	1
2	1	-1	1	1	1	-1	4	1	1	1	-1	-1	-1
2	1	1	1	-1	1	1	4	1	-1	-1	-1	1	1
2	1	1	-1	1	-1	1	4	1	1	-1	1	1	-1
2	2	1	1	1	-1	-1	4	2	-1	-1	1	-1	-1
2	2	1	-1	-1	1	-1	4	2	1	-1	-1	-1	1
2	2	-1	1	-1	-1	1	4	2	-1	1	-1	1	-1
2	2	-1	-1	1	1	1	4	2	1	1	1	1	1

1 (a). The second blocking factor (day) is orthogonal to MEs in both Figures 1 (a) and 1 (b) but is only near-orthogonal to the SOEs in these figures. We say that for both designs, the MEs of both designs are clear of block effects but the SOEs are partially confounded (slightly correlated) with block effects.

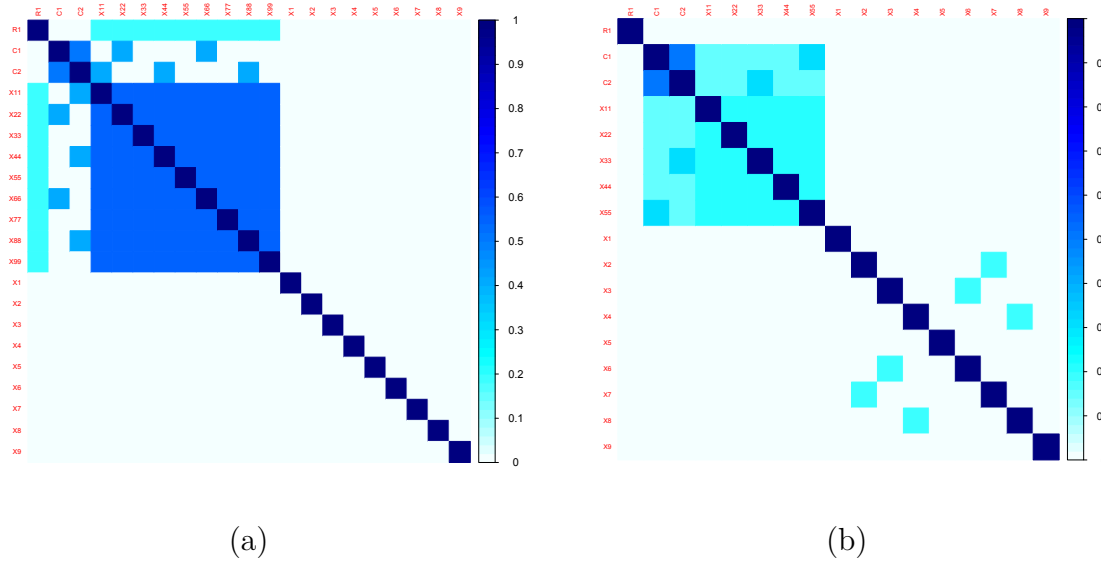


Figure 1: CCPs showing the confounding patterns of: (a) a blocked 9-factor DSD, (b) a blocked mixed-level screening designs in Table 2

Table 2: Two blocked 24-run designs for Example 2: (a) a blocked 9-factor DSD; (b) a blocked mixed-level screening designs with five 3-level factors and four 2-level factors. The low, mid- and high levels are coded as $-1, 0$ and 1

Reactor	Day	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	1	1	1	-1	-1	1	-1	1	0
1	1	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0
1	1	-1	-1	1	-1	1	-1	0	1	1
1	2	0	-1	-1	-1	-1	-1	-1	-1	-1
1	2	-1	1	0	1	-1	-1	1	1	-1
1	2	1	-1	-1	0	1	1	1	1	-1
1	2	0	0	0	0	0	0	0	0	0
1	3	1	1	-1	1	1	-1	-1	0	1
1	3	1	-1	1	1	-1	0	1	-1	1
1	3	-1	1	-1	-1	0	1	1	-1	1
1	3	-1	0	1	1	1	1	-1	-1	-1
2	1	-1	-1	-1	1	-1	1	-1	1	1
2	1	-1	-1	-1	1	1	-1	1	-1	0
2	1	1	1	-1	1	-1	1	0	-1	-1
2	1	1	1	1	-1	1	-1	1	-1	-1
2	2	0	0	0	0	0	0	0	0	0
2	2	-1	1	1	0	-1	-1	-1	-1	1
2	2	1	-1	0	-1	1	1	-1	-1	1
2	2	0	1	1	1	1	1	1	1	1
2	3	-1	-1	1	-1	-1	1	1	0	-1
2	3	1	-1	1	1	0	-1	-1	1	-1
2	3	-1	1	-1	-1	1	0	-1	1	-1
2	3	1	0	-1	-1	-1	-1	1	1	1

(a)

Reactor	Day	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	-1	1	-1	0	-1	-1	-1	1	1
1	1	1	-1	1	0	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	2	-1	-1	1	1	0	1	-1	1	-1
1	2	1	0	-1	1	1	-1	1	1	-1
1	2	-1	0	1	-1	-1	1	-1	-1	1
1	2	1	1	-1	-1	0	-1	1	-1	1
1	3	1	-1	0	1	-1	-1	-1	1	1
1	3	1	1	1	1	-1	1	-1	-1	-1
1	3	-1	1	0	-1	1	1	1	-1	-1
1	3	-1	-1	-1	-1	1	-1	1	1	1
2	1	-1	-1	0	1	-1	-1	1	-1	-1
2	1	-1	0	1	-1	1	-1	-1	1	-1
2	1	1	0	-1	1	-1	1	1	-1	1
2	1	1	1	0	-1	1	1	-1	1	1
2	2	0	1	1	1	1	1	1	1	1
2	2	0	-1	-1	-1	-1	-1	-1	-1	-1
2	2	-1	1	-1	0	-1	1	1	1	-1
2	2	1	-1	1	0	1	-1	-1	-1	1
2	3	1	1	1	-1	-1	-1	1	1	-1
2	3	-1	1	1	1	0	-1	1	-1	1
2	3	-1	-1	-1	1	1	1	-1	-1	1
2	3	1	-1	-1	-1	0	1	-1	1	-1

(b)

Table 3: A resolution $V^{2^{6-1}}$ FFD arranged in eight blocks by using the BLOCK algorithm

Block	A	B	C	D	E	F	Block	A	B	C	D	E	F
1	-1	1	-1	-1	1	-1	5	-1	1	1	1	1	-1
1	1	1	1	1	-1	-1	5	-1	1	-1	-1	-1	1
1	1	-1	-1	-1	-1	1	5	1	-1	1	-1	1	1
1	-1	-1	1	1	1	1	5	1	-1	-1	1	-1	-1
2	1	1	-1	1	1	-1	6	-1	-1	1	-1	1	-1
2	-1	1	1	1	-1	1	6	-1	-1	-1	1	-1	1
2	-1	-1	-1	-1	1	1	6	1	1	1	1	1	1
2	1	-1	1	-1	-1	-1	6	1	1	-1	-1	-1	-1
3	1	-1	1	1	-1	1	7	1	-1	-1	1	1	1
3	-1	1	1	-1	-1	-1	7	-1	1	-1	1	-1	-1
3	-1	-1	-1	1	1	-1	7	1	1	1	-1	1	-1
3	1	1	-1	-1	1	1	7	-1	-1	1	-1	-1	1
4	1	1	1	-1	-1	1	8	1	-1	-1	-1	1	-1
4	-1	-1	-1	-1	-1	-1	8	-1	1	1	-1	1	1
4	1	-1	1	1	1	-1	8	1	1	-1	1	-1	1
4	-1	1	-1	1	1	1	8	-1	-1	1	1	-1	-1

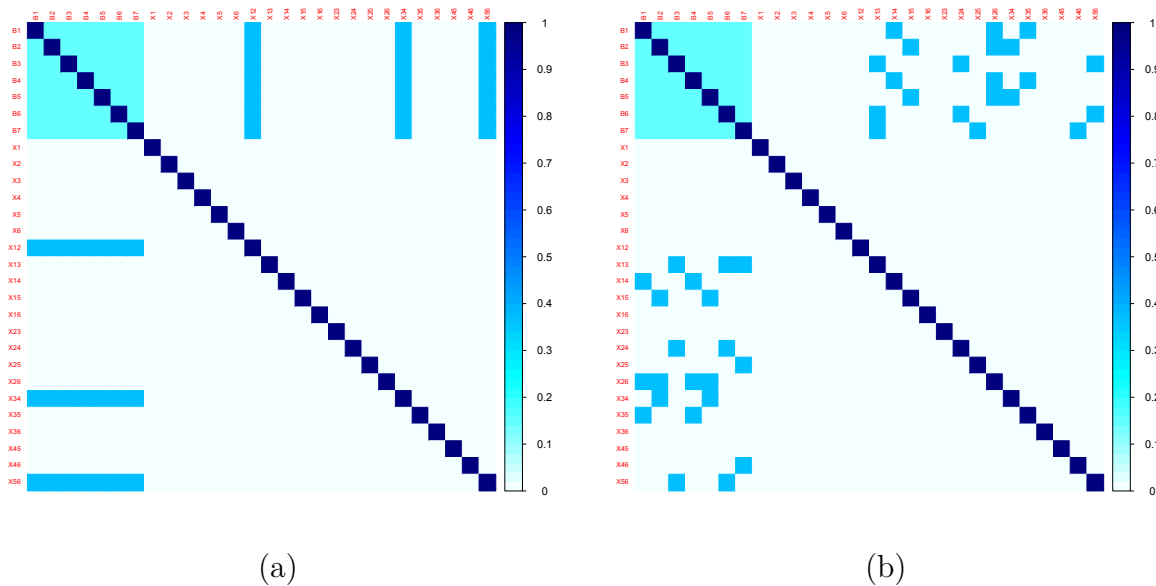


Figure 2: CCPs showing the confounding patterns of a resolution $V^{2^{6-1}}$ FFD arranged in eight blocks by: (a) using the block generators ACE, BCE and ADE, (b) using the BLOCK algorithm

It is reasonable to compare a block design available in a catalog and that constructed by BLOCK. Let's block a resolution $V^{2^{6-1}}$ FFD generated by the design generator $\mathbf{F}=\mathbf{ABCDE}$ in eight blocks. To construct this block design, we can use the block generators **ACE**, **BCE** and **ADE** (See Table 5B.3 of Wu and Hamada, 2009). The resulting design has clear MEs and 2FIs, except **AB**, **BC** and **CD**. In other words, these 2fi's are fully confounded with blocks and cannot be estimated. In contrast, our block design in Table 3 has clear MEs and

some clear 2FIs. Most 2FIs are however, partially confounded with blocks but can still be estimated.

Figures 2 (a) and 2 (b) show the confounding patterns of the blocked resolution $V 2^{6-1}$ FFDs in Tables 5B.3 of Wu and Hamada, 2009, and in Table 3, respectively. It can be seen from the two CCPs that the main effects of both designs are clear of block effects. However, for the CCP in Figure 2 (a), the 2fi's **AB**, **BC** and **CD** are confounded with block effects. At the same time, for the CCP in Figure 2 (b) the 2fi's are either clear or partially confounded with the block effects.

The runtime for BLOCK is minimal. BLOCK requires a couple of seconds to find a design solution for each design problem in this paper on a Mac mini M1 (each design parameter requires 1,000 tries).

5. Conclusion

Most block designs in the literature are cataloged designs and as such they are not flexible enough. The 2-level factorials and FFDs are only available in 2^q blocks, but not available in five, six or seven blocks. The 4-factor BBD, for example is available in three blocks but not in two blocks. Besides, catalogs do not offer designs having more than one blocking factor. BLOCK was developed with the philosophy “Design for experiment, not experiment for the design” in mind. We hope it could offer efficient alternatives to the existing catalog of block designs. In the pre-computer age, the constructed block design strived for simplicity in the analysis. Nowadays, it is not simplicity in the analysis but the design efficiency and the saving of experimental resources that counts.

The link to the supplemental material contains the Java implementation of the BLOCK algorithm in Section 3 is https://drive.google.com/drive/folders/1lju4hWxI0tY4zA_sxT5qVtgtYCYdH4LV. It also contains additional examples of blocked factorial and FFDs, blocked RSDs, blocked mixture designs, and blocked DSDs and DSD-based mixed-level designs (Jones and Nachtsheim, 2013) and Hadamard design-based mixed-level designs (Nguyen *et al.*, 2020).

References

- Box, G. E. P. and Hunter, J. S. (1957). Multi-factor experimental designs for exploring response surfaces. *The Annals of Mathematical Statistics*, **28**, 159-241.
- Cook, R. D. and Nachtsheim, C. J. (1989). Computer-aided blocking of factorial and response surface designs. *Technometrics*, **31**, 339-346.
- Gilmour, S. G. and Trinca L. A. (2003). Row-column response surface designs. *Journal of Quality Technology*, **35**, 184-193.
- Hoang, E. M., Liauw, C. M., Allen, N. S., Fontan, E. and Lafuente, P. (2004). Effects of additive interactions on the thermo-oxidative stabilization of a film grade metallocene LLDPE. *Journal of Vinyl and Additive Technology*, **10**, 148-156.
- Jones, B. and Nachtsheim, C. J. (2011). A class of three levels designs for definitive screening in the presence of second-order effects. *Journal of Quality Technology*, **43**, 1-15.
- Jones, B. and Nachtsheim, C. J. (2013). Definitive screening designs with added two-level categorical factors. *Journal of Quality Technology*, **45**, 121-129.

- Jones, B. and Nachtsheim, C. J. (2016). Blocking schemes for definitive screening designs, *Technometrics* **58**, 74-83.
- Khuri, A. I. and Cornell, J. A. (1996). Response Surfaces, Designs and Analyses. 2nd ed. New York, Marcel Dekker.
- Mee, R. W. (2009). *A Comprehensive Guide to Factorial Two-level Experimentation*. New York, NY: Springer.
- Nguyen, N-K (2001). Cutting experimental designs into blocks. *Australia and New Zealand Journal of Statistics*, **43**, 367-374.
- Nguyen, N-K., Pham, D-T. and Vuong, M. P. (2020). Constructing D-efficient mixed-level foldover designs using Hadamard matrices. *Technometrics*, **62**, 48-56,
- Wu, C. F. J and Hamada, M. (2009). Experiments: Planning, Analysis and Parameter Design Optimisation. New York: John Wiley and Sons, Inc.
- Zhang, C, Wei Chen, W, Xiana, J. and Degang Fub, D. (2018). Application of a novel definitive screening design to in situ chemical oxidation of acid orange-II dye by a Co^{2+} /PMS system. *RSC Advances*, **8**, 3934-3940.