DOE Wizard – Definitive Screening Designs

statgraphics 18
centurion

Revised: 10/10/2017

Summary

The DOE Wizard can construct and analyze definitive screening designs (DSD) as described in the articles referenced as the end of this document. Definitive screening designs may be constructed for any combination of continuous and 2-level categorical factors where the total number of factors is between 4 and 16. Both blocked and unblocked designs are available.

Definitive screening designs are small designs capable of estimating models involving both linear and quadratic effects, although second-order interactions are partially confounded with themselves and with quadratic effects. In addition, designs for 6 or more factors collapse into designs which can estimate the full second-order model (including interactions) for any 3 factors.

Jones and Nachtsheim (2011) state that using DSD designs often makes it unnecessary to perform follow-up experiments. They list the following desirable properties of those designs:

1. The required number of runs is very small, usually between 1 and 3 more than twice the number of factors.

2. Main effects are independent of two-factor interactions.

3. Two-factor interactions are not perfectly confounded with other two-factor interactions, although they are correlated.

4. For continuous factors, all of the quadratic effects can be estimated.

5. Quadratic effects are orthogonal to linear main effects and only partially confounded with two-factor interactions.

6. For designs involving 6 through 12 factors, the full second-order model can be estimated for any 3 or less factors.

Designs Containing Only Continuous Factors

For designs containing *m* continuous factors run in a single block, the total number of runs to be performed is

$$n = 2m' + 1 \qquad (1)$$

where

$$m' = m + k. \qquad (2)$$

If *m* is even, *k* = 0, while if *m* is odd, *k* = 1. For designs runs in *B* blocks where *B* > 1, the total number of runs is

$$n = 2m' + B - k \qquad (3)$$

The designs are constructed using conference matrices as described by Xiao, Lin and Bai (2012).

As an example, a typical screening experiment will be constructed involving 5 factors and 1 response. The example, which involves a chemical reaction, is discussed in Chapter 12 of the well-known book by Box, Hunter and Hunter (2005). The factors that will be varied are:

X1: feed rate
X2: amount of catalyst
X3: agitation rate
X4: temperature
X5: concentration

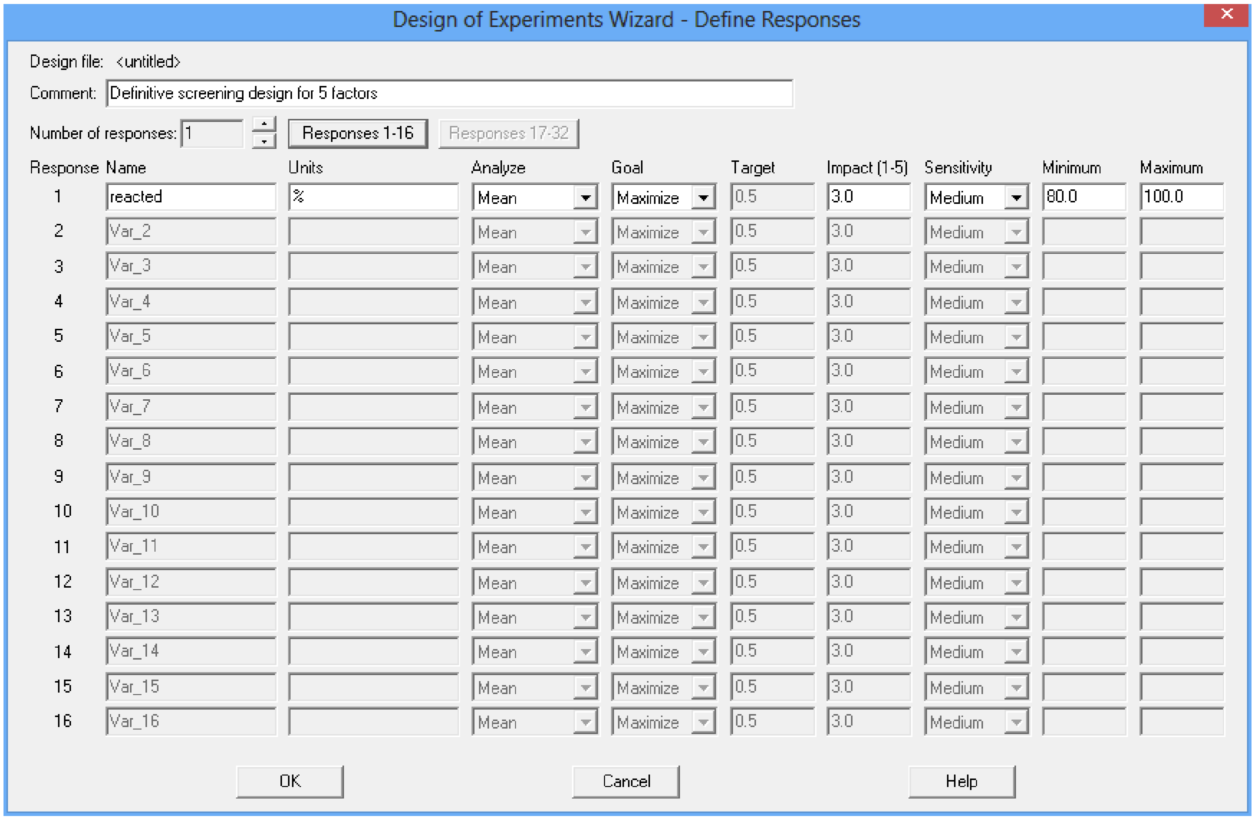
There is one response variable:

Y: percent reacted

To begin the design creation process, start with an empty StatFolio. Select *DOE – Experimental Design Wizard* to load the DOE Wizard's main window. Then push each button in sequence to create the design.

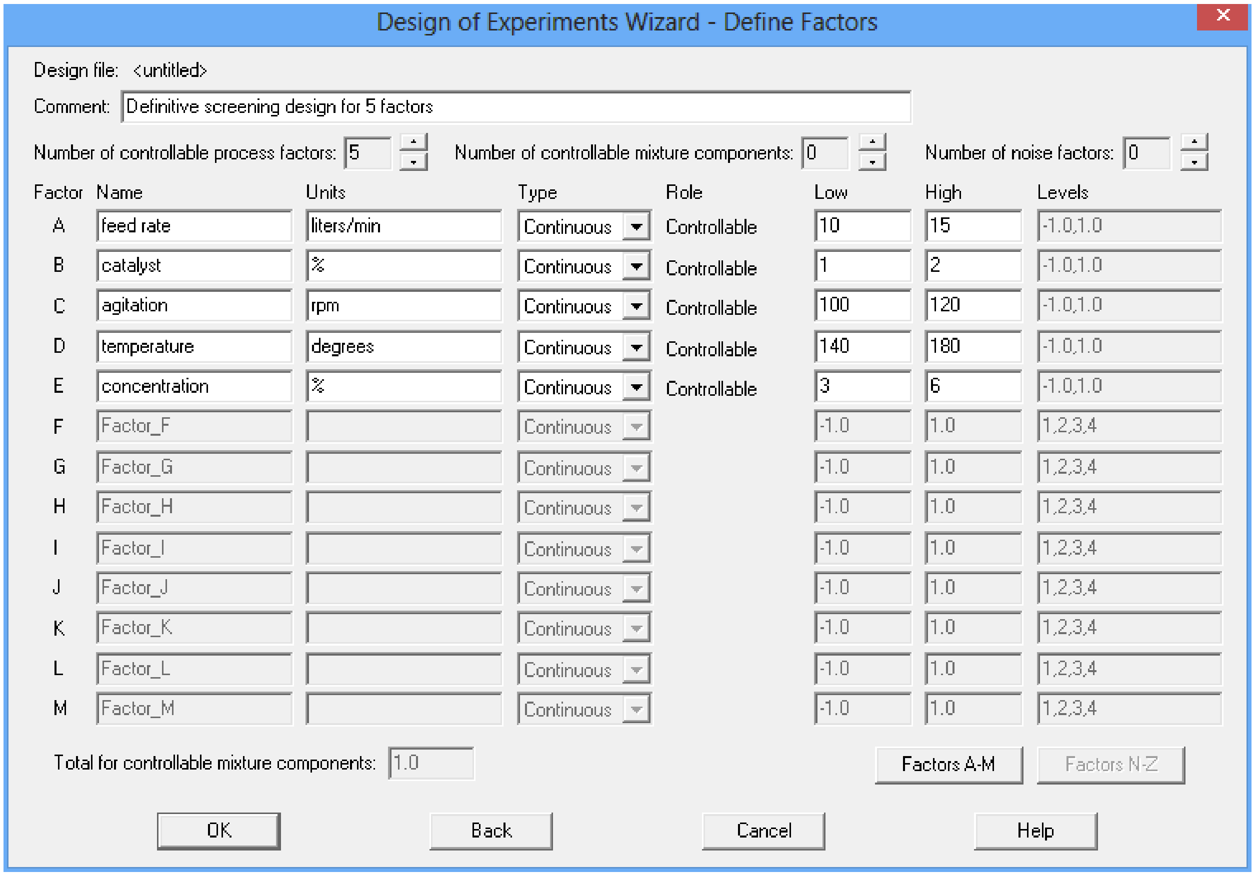
Step #1 – Define Responses

The first step of the design creation process displays a dialog box used to specify the response variables. For the current example, there is a single response variable:



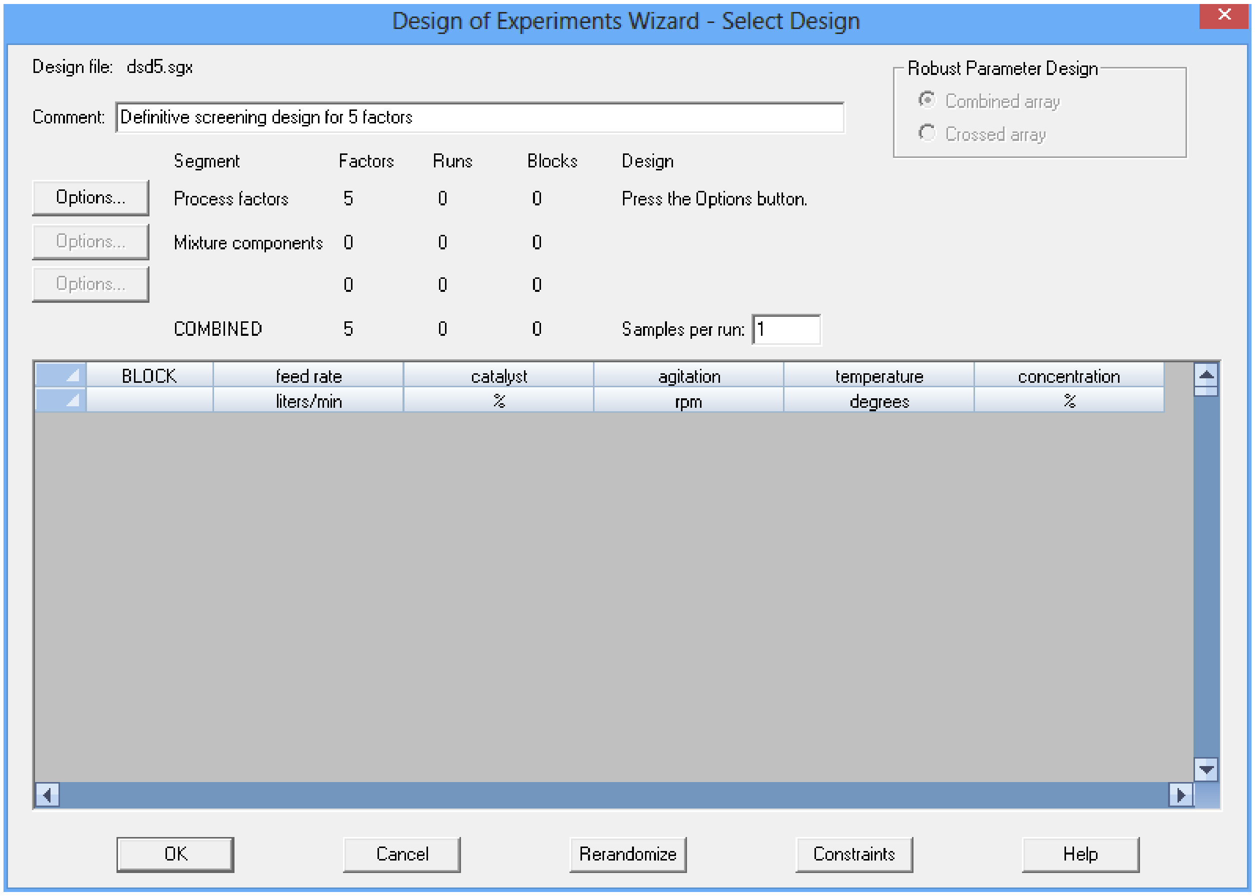
Step #2 – Define Experimental Factors

The second step displays a dialog box on which to specify the factors that will be varied. In the chemical reaction example, there are 5 factors:

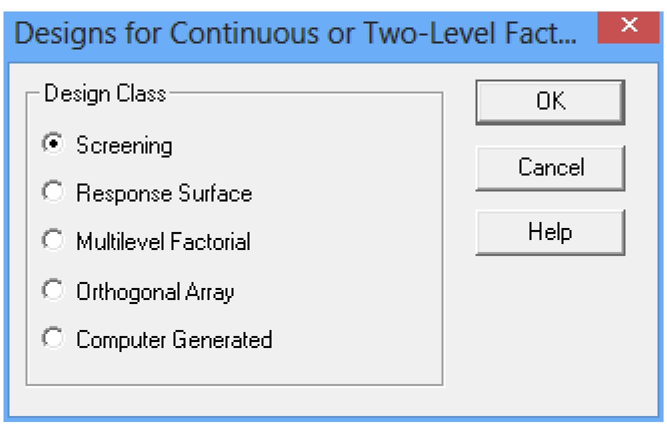


Step #3 – Select Design

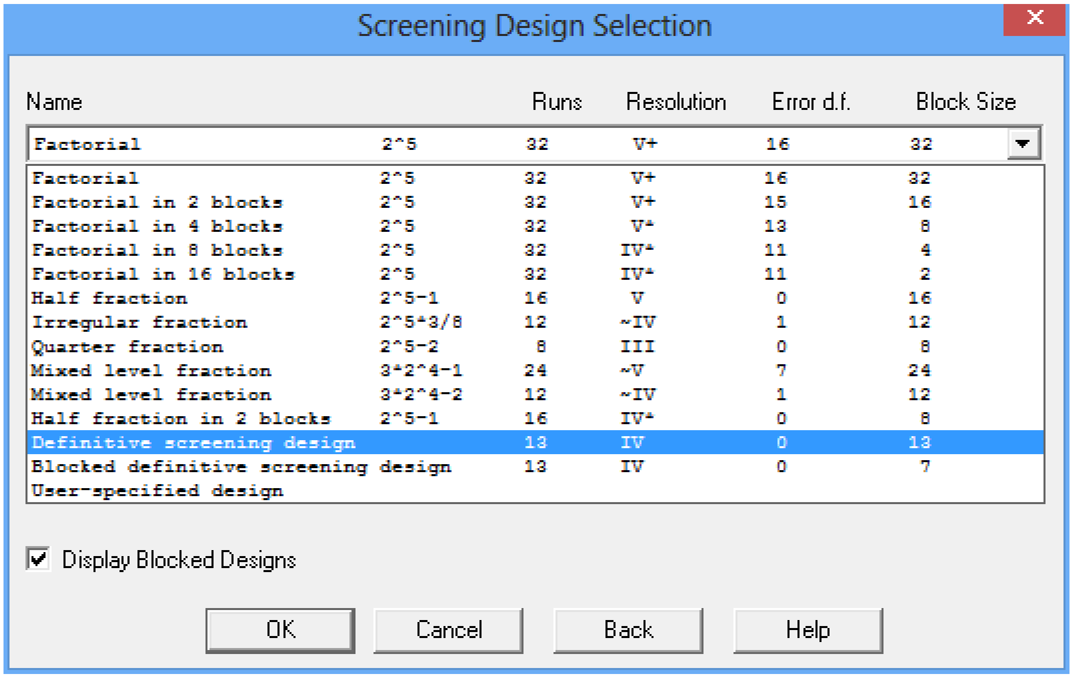
The third step begins by displaying the dialog box shown below:



Since all of the factors are controllable process factors, only one *Options* button is enabled. Pressing that button displays a second dialog box:



Select *Screening* and press *OK*. This will display a third dialog box listing all of the screening designs available for 5 experimental factors:

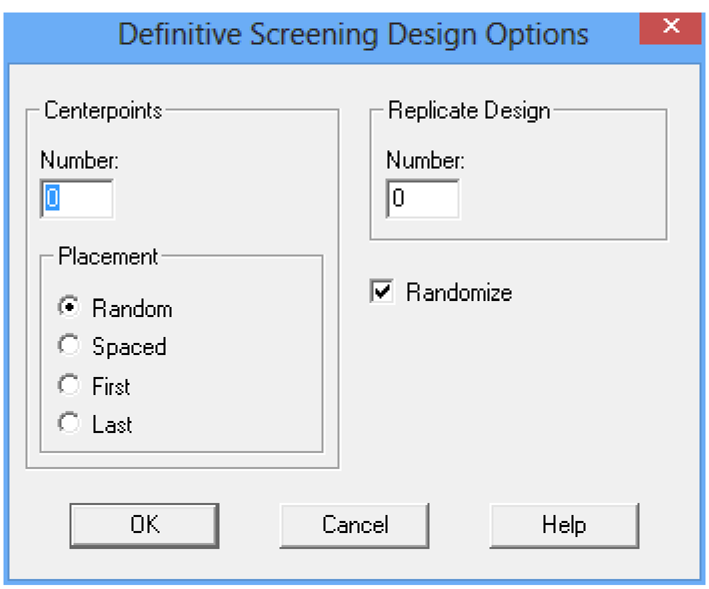


Two options for creating a DSD are listed:

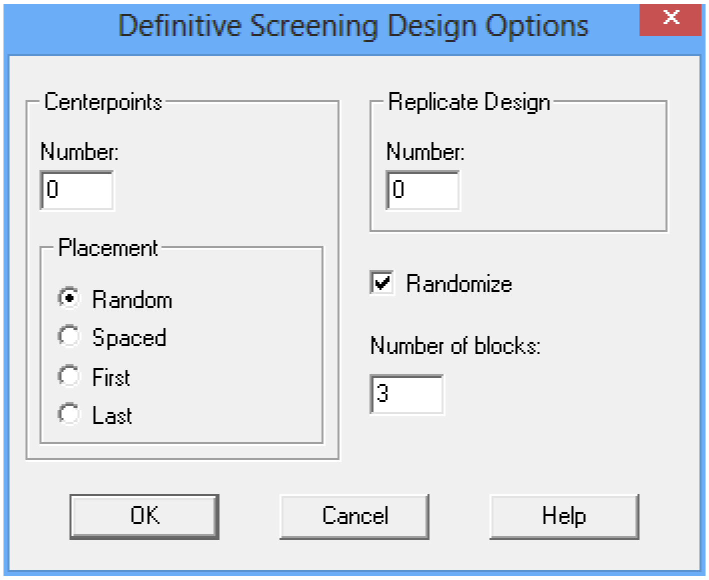
1. *Definitive screening design* – creates a design with n = 13 runs, all in a single block.

2. *Blocked definitive screening design* – creates a DSD in 2 or more blocks. The default design has n = 13 runs divided into 2 blocks, with a maximum of 7 runs in any block. Note that the number of blocks may be increased on the next dialog box, which will increase the total number of runs.

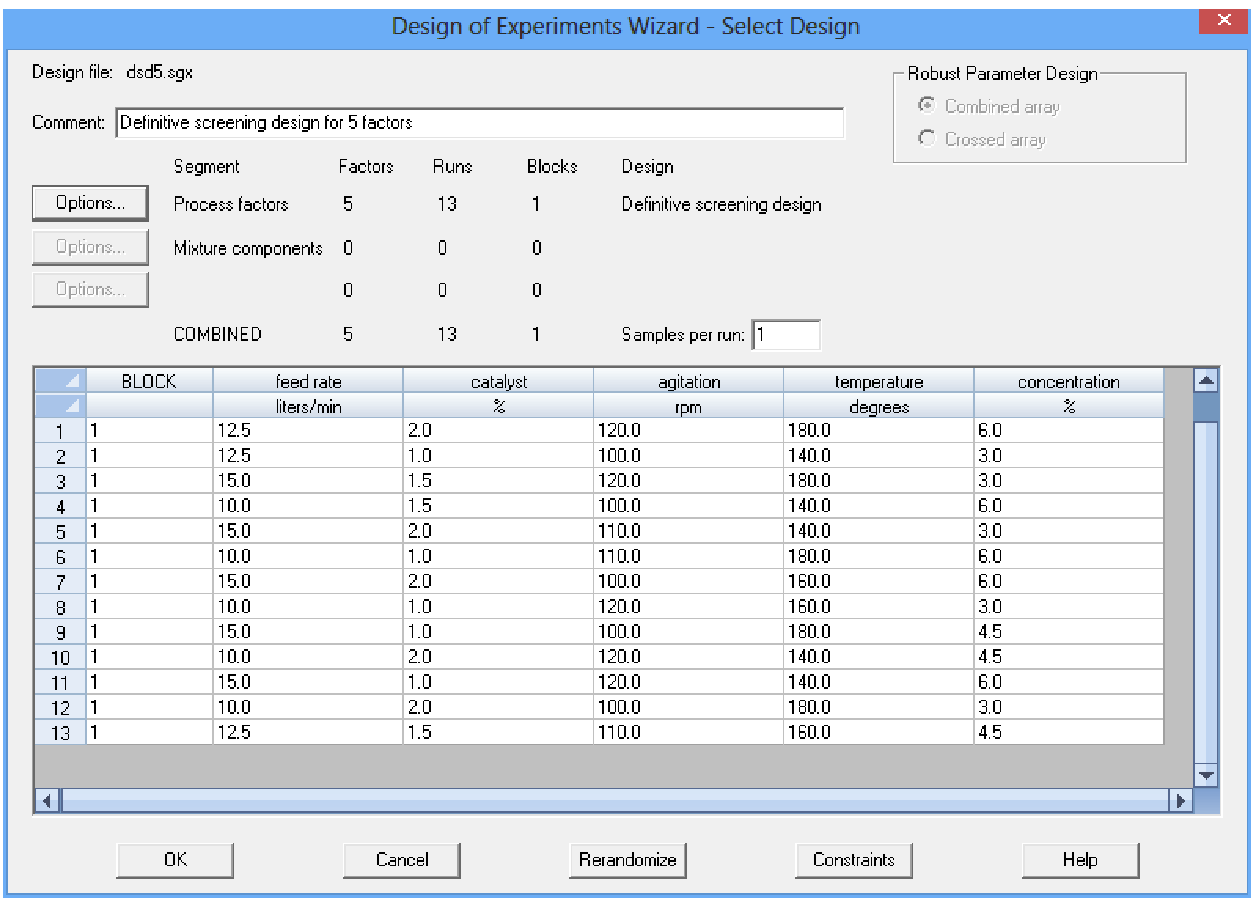
The final dialog box allows the analyst to add additional runs to the design and to specify the order in which the runs will be performed. If the unblocked DSD is chosen, the dialog box takes the following form:



If a blocked DSD is chosen, the dialog box also contains a field for specifying the number of blocks:

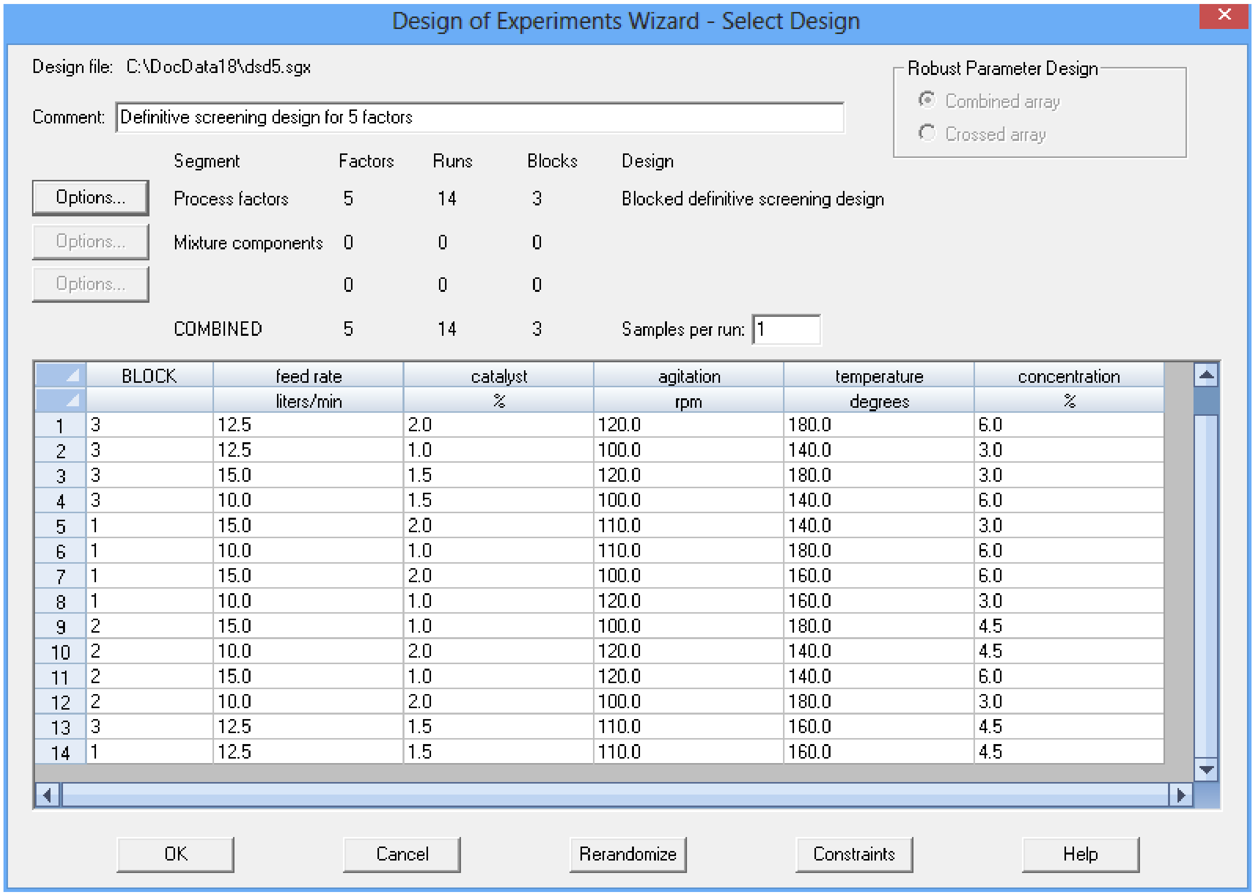


The unblocked DSD for 5 factors in shown below in non-random order:



Note that each factor is run at 3 levels: at its low level, at its high level, and halfway between the low and high levels.

If a DSD with 3 blocks is requested, the design takes the following form:

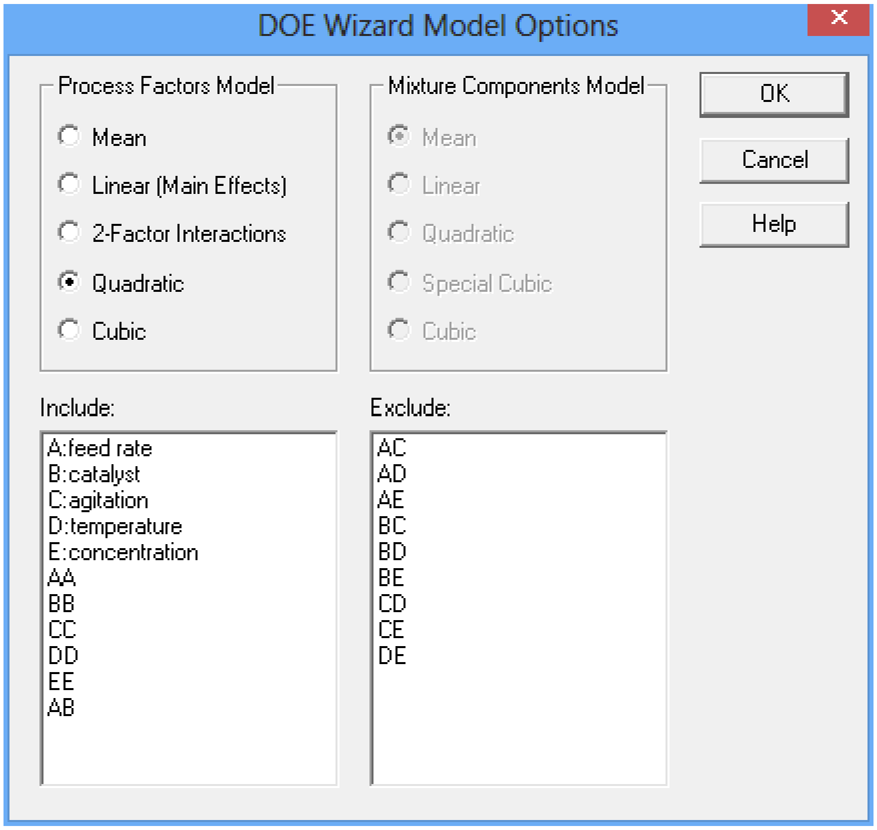


	BLOCK	feed rate	catalyst	agitation	temperature	concentration
		liters/min	%	rpm	degrees	%
1	3	12.5	2.0	120.0	180.0	6.0
2	3	12.5	1.0	100.0	140.0	3.0
3	3	15.0	1.5	120.0	180.0	3.0
4	3	10.0	1.5	100.0	140.0	6.0
5	1	15.0	2.0	110.0	140.0	3.0
6	1	10.0	1.0	110.0	180.0	6.0
7	1	15.0	2.0	100.0	160.0	6.0
8	1	10.0	1.0	120.0	160.0	3.0
9	2	15.0	1.0	100.0	180.0	4.5
10	2	10.0	2.0	120.0	140.0	4.5
11	2	15.0	1.0	120.0	140.0	6.0
12	2	10.0	2.0	100.0	180.0	3.0
13	3	12.5	1.5	110.0	160.0	4.5
14	1	12.5	1.5	110.0	160.0	4.5

There are a total of n = 14 runs with each factor at 3 levels. The blocking pattern has been selected following the methods of Jones and Nachtsheim (2012) which uses the D criterion to pick an optimal design.

Step #4 – Specify model

The fourth step in the DOE Wizard selects the model to be fit to the data once the experiment is performed. The default model is shown below:



It contains 5 main effects, 5 quadratic terms, and 1 two-factor interaction. The selected interaction is completely arbitrary and is partially confounded with all of the other two-factor interactions.

To examine the properties of the design, it is helpful to push the *Evaluate design* button and then select *Correlation Matrix*:

Correlation Matrix

	A	B	C	D	E	AA	BB	CC	DD	EE	AB
A	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
B	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
D	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
E	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AA	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.1333	0.1333	0.1333	0.1333	0.0000
BB	0.0000	0.0000	0.0000	0.0000	0.0000	0.1333	1.0000	0.1333	0.1333	0.1333	0.0000
CC	0.0000	0.0000	0.0000	0.0000	0.0000	0.1333	0.1333	1.0000	0.1333	0.1333	-0.4655
DD	0.0000	0.0000	0.0000	0.0000	0.0000	0.1333	0.1333	0.1333	1.0000	0.1333	-0.4655
EE	0.0000	0.0000	0.0000	0.0000	0.0000	0.1333	0.1333	0.1333	0.1333	1.0000	0.4655
AB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.4655	-0.4655	0.4655	1.0000

Notice the following properties:

1. The linear main effects (A, B, C, D, and E) are uncorrelated with any of the other terms.

2. The quadratic effects (AA, BB, CC, DD, and EE) are uncorrelated with the main effects but have small correlations with each other. They are also correlated with the AB interaction.

It is also helpful to display the *Alias Matrix*, which shows the confounding pattern for effects that are not included in the model:

Alias Matrix

Effect	AC	AD	AE	BC	BD	BE	CD	CE	DE
constant									
A									
B									
C									
D									
E									
AA				-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
BB	-1.0000	-1.0000	1.0000				-1.0000	1.0000	1.0000
CC	-1.0000	2.0000		1.0000		2.0000	1.0000	-1.0000	
DD		1.0000	-2.0000	2.0000	-1.0000		1.0000		1.0000
EE	2.0000	-2.0000	1.0000	-2.0000	2.0000	-1.0000		1.0000	-1.0000
AB	-1.0000	1.0000	-1.0000	1.0000	-1.0000	1.0000	1.0000	-1.0000	1.0000

While the main effects are clear of the omitted two-factor interactions, the other terms are not.

In the case of the blocked design, the correlation matrix appears as follows:

Correlation Matrix

	block	block	A	B	C	D	E	AA	BB	CC	DD	EE
block	1.0000	0.5292	0.0000	0.0000	0.0000	0.0000	0.0000	0.1414	0.1414	0.5374	0.5374	-0.2546
block	0.5292	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.3742	-0.3742	0.3742	0.3742	0.0000
A	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
B	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
D	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
E	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AA	0.1414	-0.3742	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.3000	0.3000	0.3000	0.3000
BB	0.1414	-0.3742	0.0000	0.0000	0.0000	0.0000	0.0000	0.3000	1.0000	0.3000	0.3000	0.3000
CC	0.5374	0.3742	0.0000	0.0000	0.0000	0.0000	0.0000	0.3000	0.3000	1.0000	0.3000	0.3000
DD	0.5374	0.3742	0.0000	0.0000	0.0000	0.0000	0.0000	0.3000	0.3000	0.3000	1.0000	0.3000
EE	-0.2546	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.3000	0.3000	0.3000	0.3000	1.0000

While the linear effects are orthogonal to the blocks, the quadratic effects are not.

Designs Containing Categorical Factors

Definitive screening designs may also contain two-level categorical factors. Suppose the experimenter wishes to study m continuous factors and c categorical factors. If all runs are to be performed in a single block, then a DSD can be constructed containing

$$n = 2m' + 2 \tag{4}$$

runs where

$$m' = m + c + k. \tag{5}$$

If $m+c$ is even, $k = 0$, while if $m+c$ is odd, $k = 1$. For designs runs in B blocks where $B > 1$, the total number of runs is

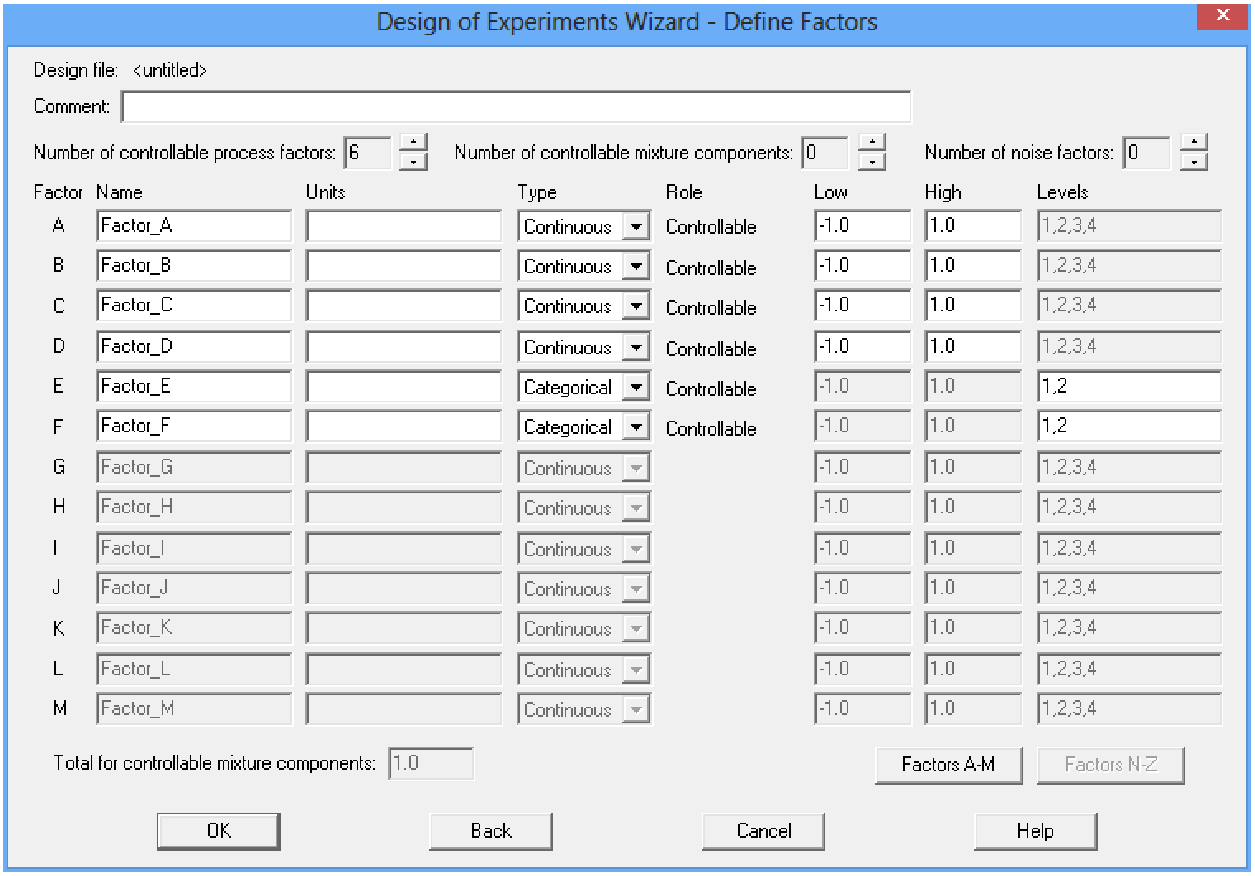
$$n = 2m' + B - k \tag{6}$$

if $(2m' + B - k)$ is even and

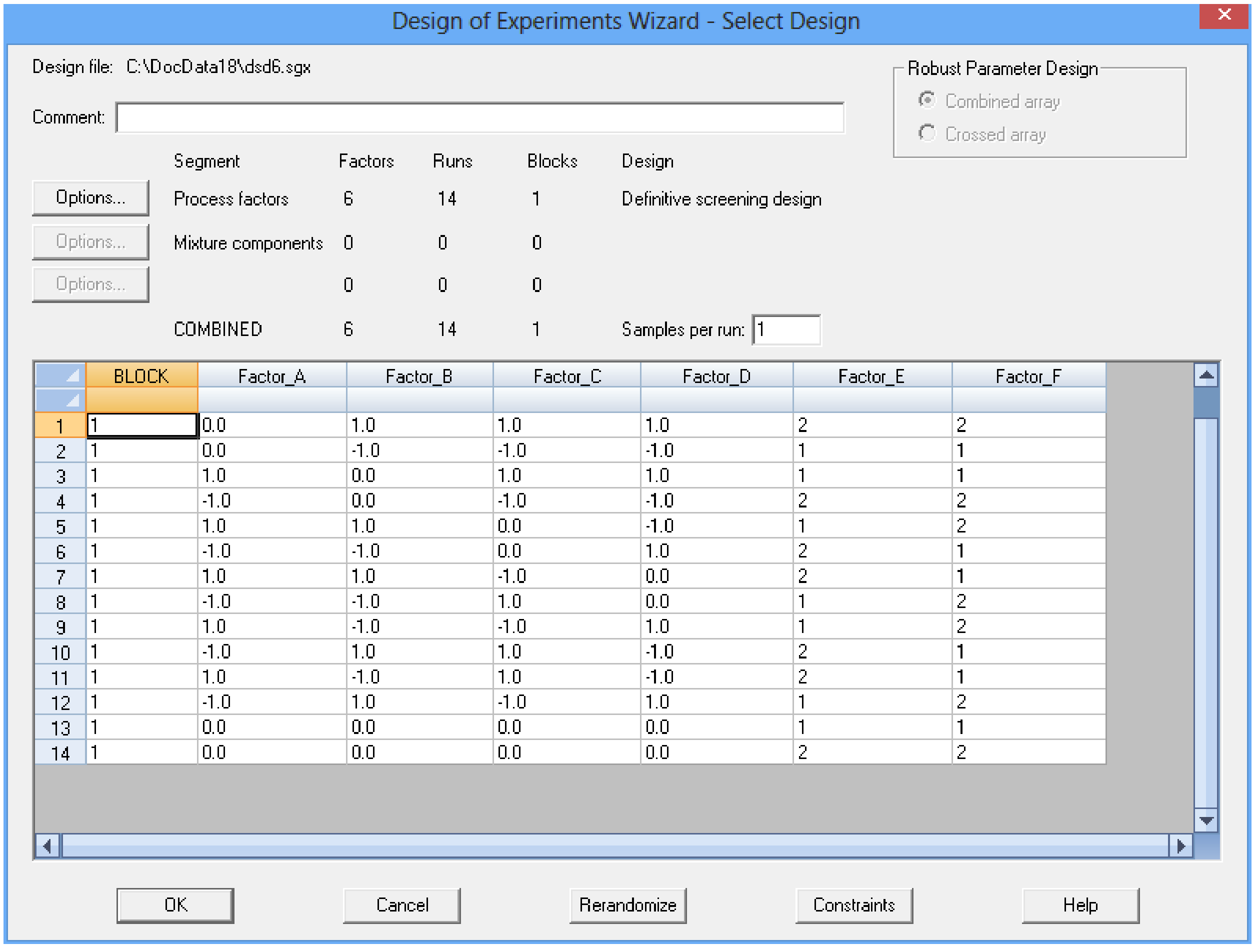
$$n = 2m' + B - k + 1 \tag{7}$$

if $(2m' + B - k)$ is odd.

As an example, consider a design with 4 continuous factors and 2 categorical factors:

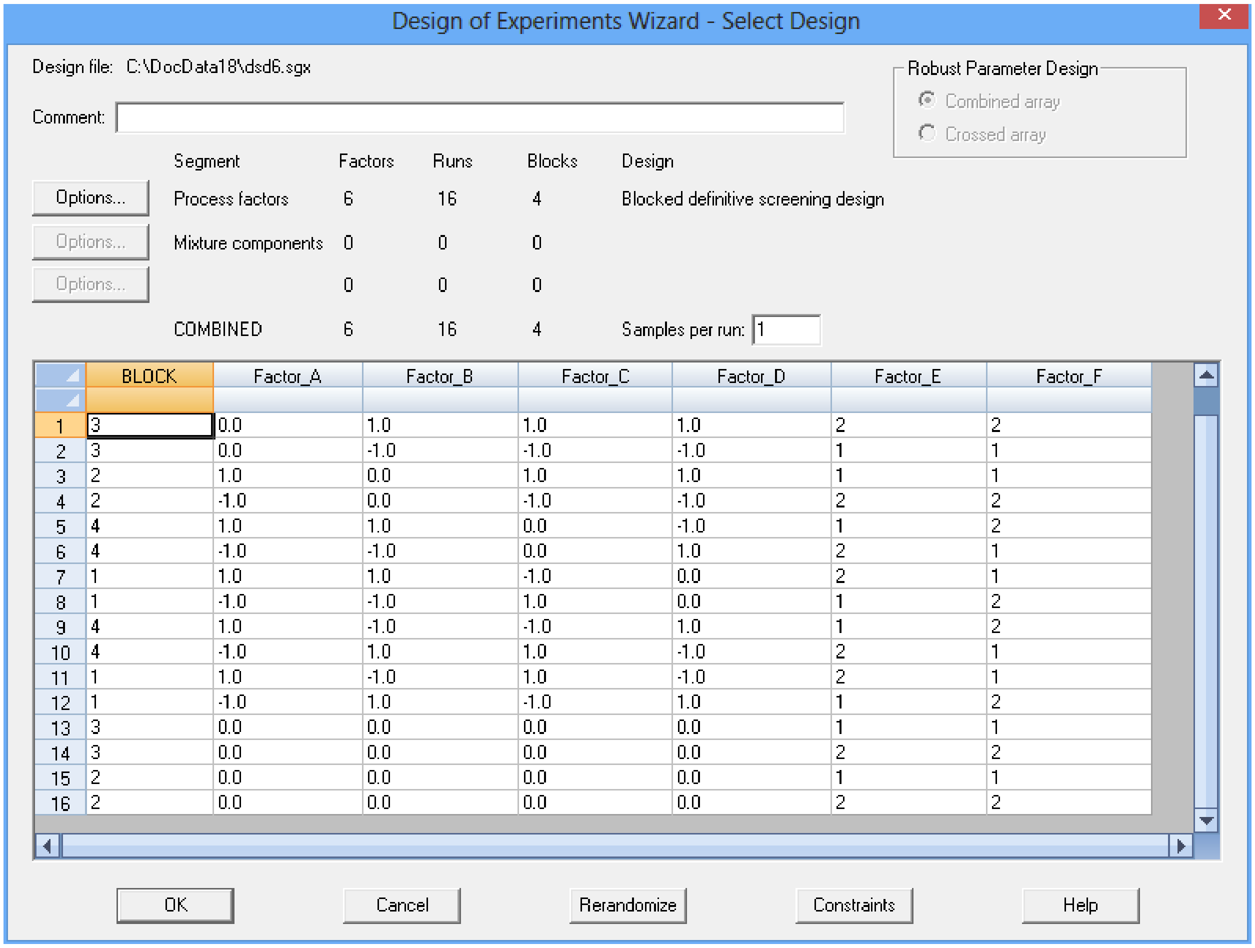


An unblocked DSD consists of the following 14 runs:



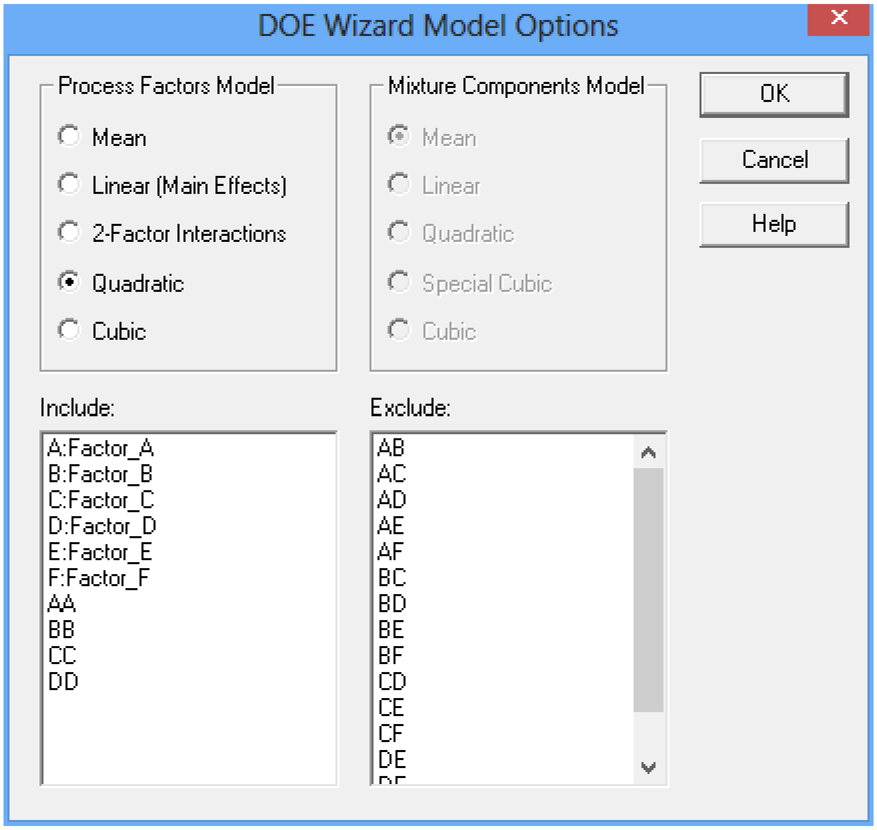
	BLOCK	Factor_A	Factor_B	Factor_C	Factor_D	Factor_E	Factor_F
1	1	0.0	1.0	1.0	1.0	2	2
2	1	0.0	-1.0	-1.0	-1.0	1	1
3	1	1.0	0.0	1.0	1.0	1	1
4	1	-1.0	0.0	-1.0	-1.0	2	2
5	1	1.0	1.0	0.0	-1.0	1	2
6	1	-1.0	-1.0	0.0	1.0	2	1
7	1	1.0	1.0	-1.0	0.0	2	1
8	1	-1.0	-1.0	1.0	0.0	1	2
9	1	1.0	-1.0	-1.0	1.0	1	2
10	1	-1.0	1.0	1.0	-1.0	2	1
11	1	1.0	-1.0	1.0	-1.0	2	1
12	1	-1.0	1.0	-1.0	1.0	1	2
13	1	0.0	0.0	0.0	0.0	1	1
14	1	0.0	0.0	0.0	0.0	2	2

A blocked DSD with $B = 4$ blocks has 16 runs:



	BLOCK	Factor_A	Factor_B	Factor_C	Factor_D	Factor_E	Factor_F
1	3	0.0	1.0	1.0	1.0	2	2
2	3	0.0	-1.0	-1.0	-1.0	1	1
3	2	1.0	0.0	1.0	1.0	1	1
4	2	-1.0	0.0	-1.0	-1.0	2	2
5	4	1.0	1.0	0.0	-1.0	1	2
6	4	-1.0	-1.0	0.0	1.0	2	1
7	1	1.0	1.0	-1.0	0.0	2	1
8	1	-1.0	-1.0	1.0	0.0	1	2
9	4	1.0	-1.0	-1.0	1.0	1	2
10	4	-1.0	1.0	1.0	-1.0	2	1
11	1	1.0	-1.0	1.0	-1.0	2	1
12	1	-1.0	1.0	-1.0	1.0	1	2
13	3	0.0	0.0	0.0	0.0	1	1
14	3	0.0	0.0	0.0	0.0	2	2
15	2	0.0	0.0	0.0	0.0	1	1
16	2	0.0	0.0	0.0	0.0	2	2

The default model has main effects for all of the factors and quadratic effects for the continuous factors:



As can be seen from the correlation matrix, there are small correlations between the main effects of the categorical factors and the main effects of the other factors, although the main effects of the continuous factors are orthogonal to each other:

Correlation Matrix

	A	B	C	D	E	F	AA	BB	CC	DD
A	1.0000	0.0000	0.0000	0.0000	-0.1690	-0.1690	0.0000	0.0000	0.0000	0.0000
B	0.0000	1.0000	0.0000	0.0000	0.1690	0.1690	0.0000	0.0000	0.0000	0.0000
C	0.0000	0.0000	1.0000	0.0000	0.1690	-0.1690	0.0000	0.0000	0.0000	0.0000
D	0.0000	0.0000	0.0000	1.0000	-0.1690	0.1690	0.0000	0.0000	0.0000	0.0000
E	-0.1690	0.1690	0.1690	-0.1690	1.0000	-0.1429	0.0000	0.0000	0.0000	0.0000
F	-0.1690	0.1690	-0.1690	0.1690	-0.1429	1.0000	0.0000	0.0000	0.0000	0.0000
AA	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.3000	0.3000	0.3000
BB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.3000	1.0000	0.3000	0.3000
CC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.3000	0.3000	1.0000	0.3000
DD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.3000	0.3000	0.3000	1.0000

References

Box, G. E. P., Hunter, W. G. and Hunter, J. S. (2005). Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building, 2nd edition. New York: John Wiley and Sons.

Jones, B. and Nachtsheim, C.J. (2011) "A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects", Journal of Quality Technology 43(1), pp. 1-15.

Jones, B. and Nachtsheim, C.J. (2013) "Definitive Screening Designs with Added Two-Level Categorical Factors", Journal of Quality Technology 45(2), pp. 121-129.

Jones, B. and Nachtsheim, C.J. (2016) "Blocking Schemes for Definitive Screening Designs", Technometrics 58(1), pp. 74-83.

Xiao, L., Lin, D.K.J. and Bai, F. (2012) "Constructing Definitive Screening Using Conference Matrices", Journal of Quality Technology 44(1), pp. 1-7.