DOE Wizard – Multilevel Factorial Designs

Summary

The *Multilevel Factorial* designs are used to study the effects of *q* quantitative factors. The user begins by specifying a range over which each factor is to be varied and the number of different levels at which it will be studied. The program creates a datasheet containing all combinations of the different levels of the variables.

One important use of this type of design is to create a set of candidate runs from which an optimal subset can be selected using the *Select runs* feature of the DOE Wizard, which selects a subset of runs so as to achieve a D-optimal design.

Example

As an example, suppose the experimenter wished to study 3 factors over the region shown below:

Factor	**Low**	**High**	**Units**
temperature	160	180	degrees
flow rate	50	80	liters/min
concentration	20	40	%

Unfortunately, rather than designing a well-balanced experiment from the beginning, he decided to try a few runs that he thought might work well. The runs performed are shown below, together with the measured response:

run	*temperature*	*flow rate*	*concentration*	*yield*
1	160	50	20	18.7
2	180	75	40	32.9
3	165	80	25	22.1
4	175	80	35	29.0
5	170	55	25	22.8
6	180	75	35	29.9

After 6 runs and little progress, he decided to stop and design a good experiment. He did not, however, want to throw away the runs that had already been performed, each of which was costly and time-consuming.

This document will demonstrate how the DOE Wizard can be used to help select additional experiments to augment the runs that have already been performed.

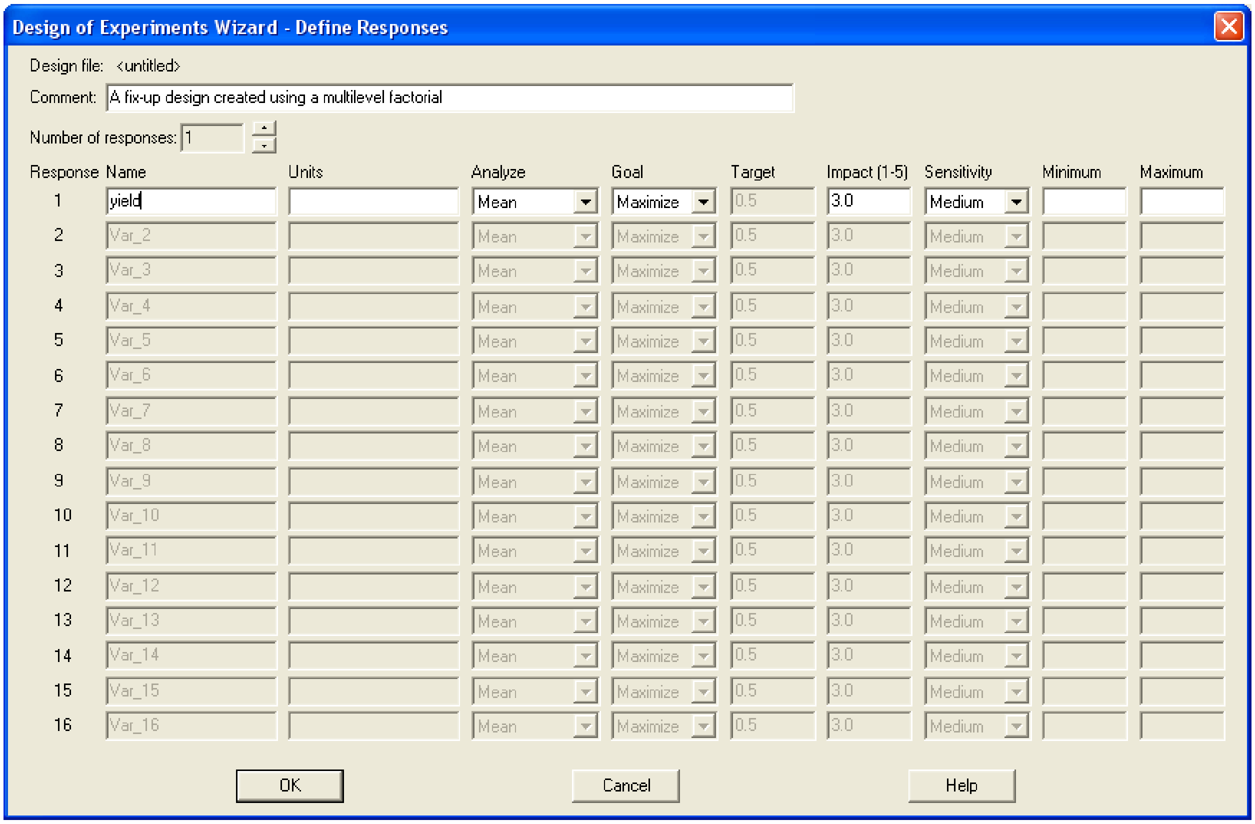
Sample StatFolio: doewiz multilevel.sgp

Design Creation

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

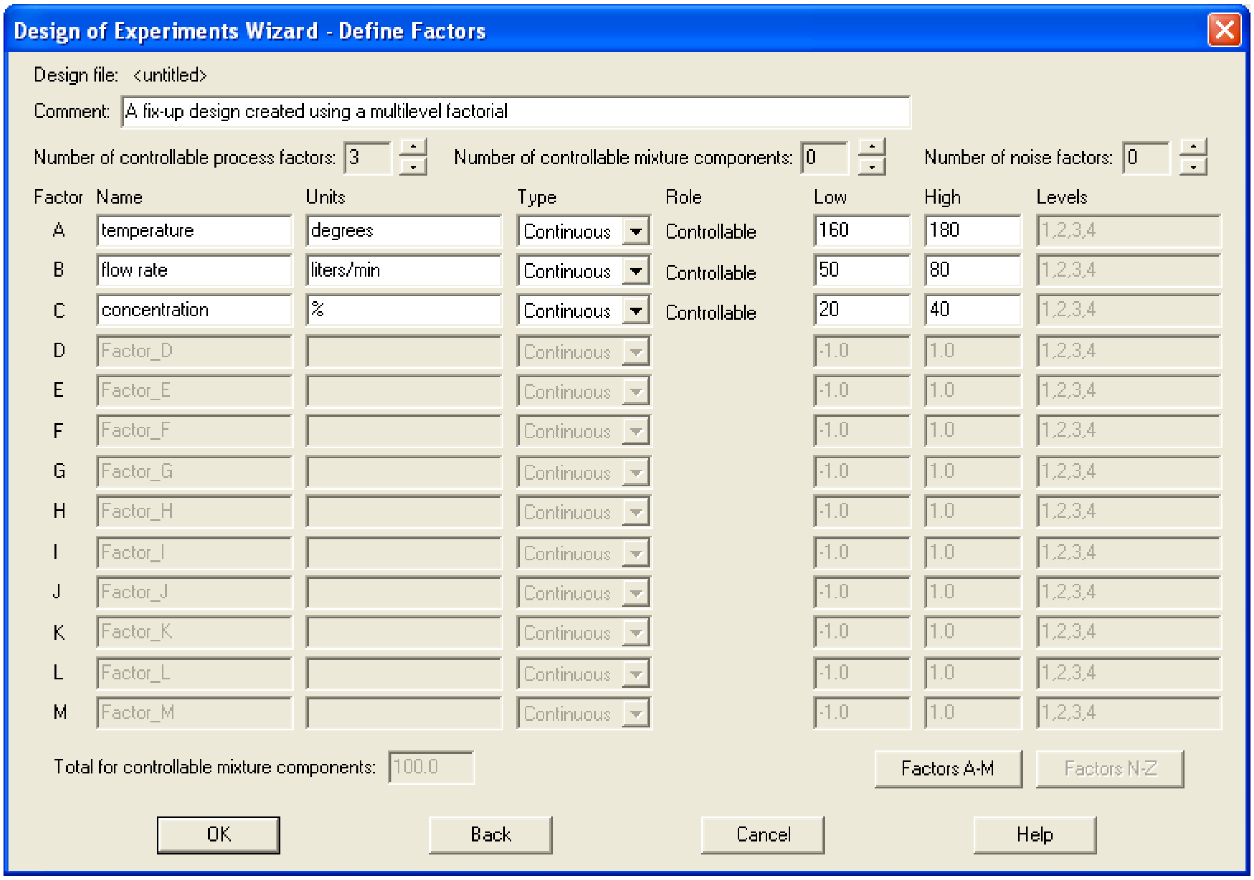
The first step of the design creation process displays a dialog box used to specify the response variables. For the current example, there is a single response variable:



- **Name**: The name for the variable is *yield*.
- **Units:** The units are not relevant in this example.
- **Analyze**: The parameter of interest is the *mean* yield.
- **Goal**: The goal of the experiment is to maximize the yield.
- **Impact**: The relative importance of each response (not relevant if only one response).
- **Sensitivity**: The importance of being close to the best desired value.
- **Minimum and Maximum**: Range of desirable values for the response.

Step #2 – Define Experimental Factors

The second step displays a dialog box on which to specify the factors that will be varied. In the current example, there are 3 factors:



- **Name** – Each factor must be assigned a unique name.

- **Units** – Units are optional.

- **Type** – Set the type of each factor to *Continuous*, since they can be set at any value within a continuous interval.

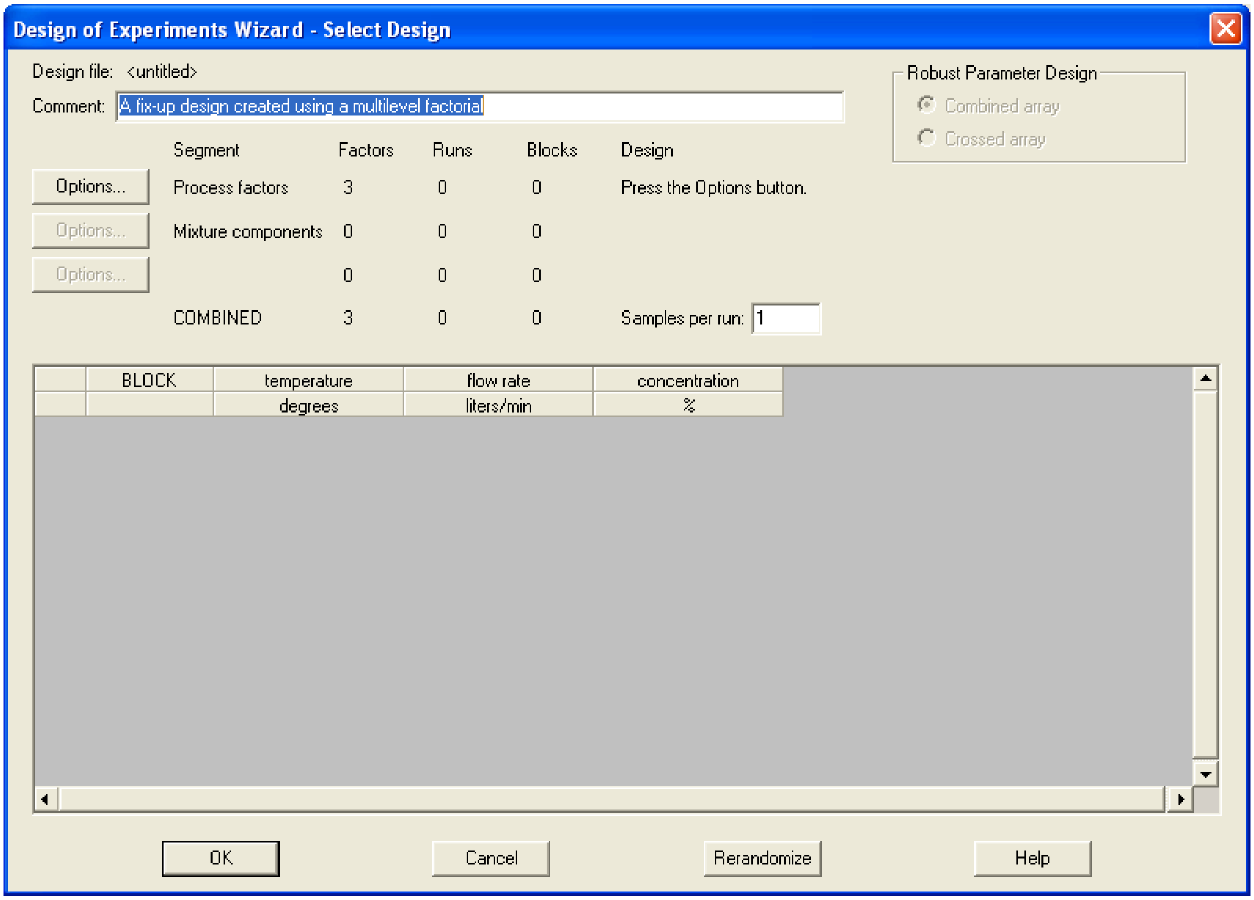
- The three factors are all *Controllable*.

- **Low** - the lower level L_j for the factor.

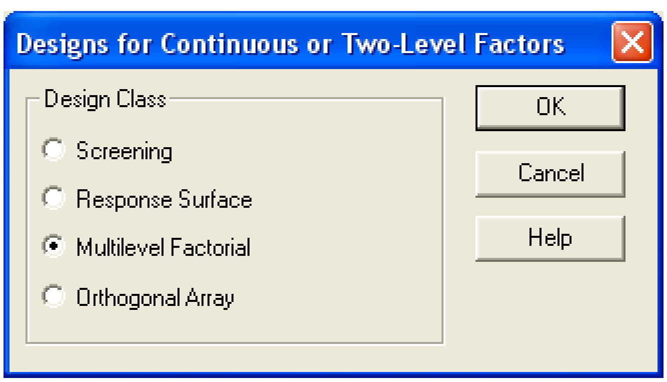
- **High** - the upper level U_j for the factor.

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:



Four general classes of designs are offered:

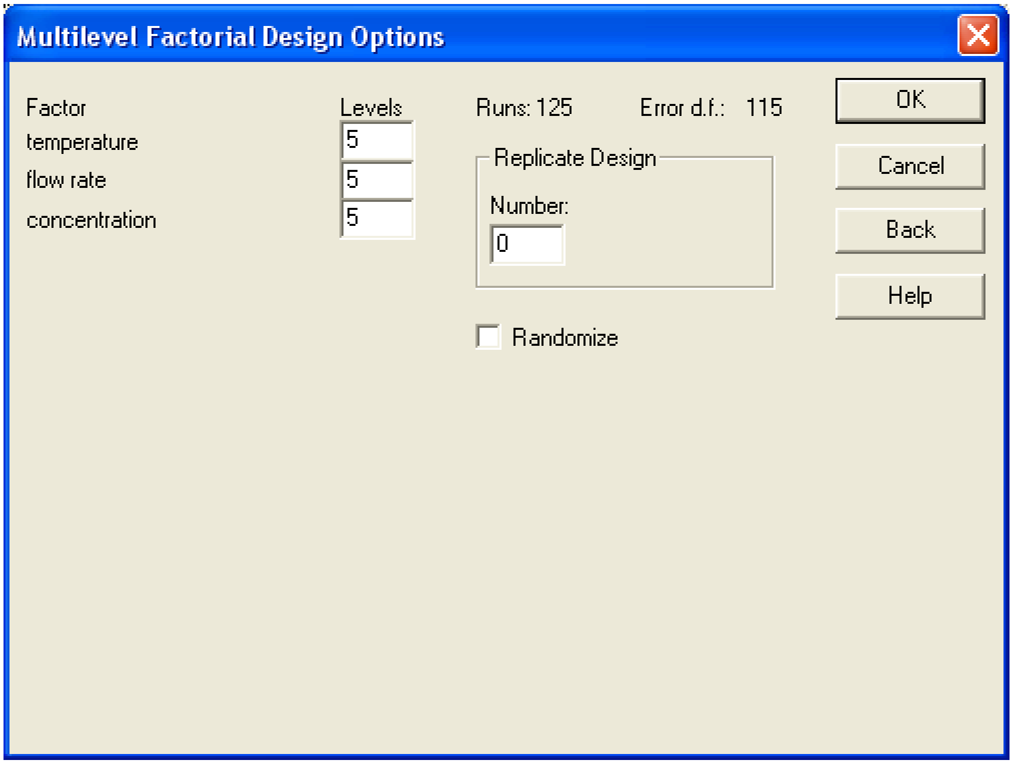
1. *Screening* - designs intended to select the most important factors affecting a response. Most of the designs involve only 2 levels of each factor. The factors may be quantitative or categorical.

2. *Response Surface* - designs intended to select the optimal settings of a set of experimental factors. The designs involve at least 3 levels of the experimental factors, which must be quantitative.

3. *Multilevel Factorial* - designs involving different numbers of levels for each experimental factor. The factors must be quantitative.

4. *Orthogonal Array* – a general class of designs developed by Genichi Taguchi. The factors may be quantitative or categorical.

If *Multilevel Factorial* is selected, a third dialog box will be displayed on which to indicate the number of levels at which each factor should be run:



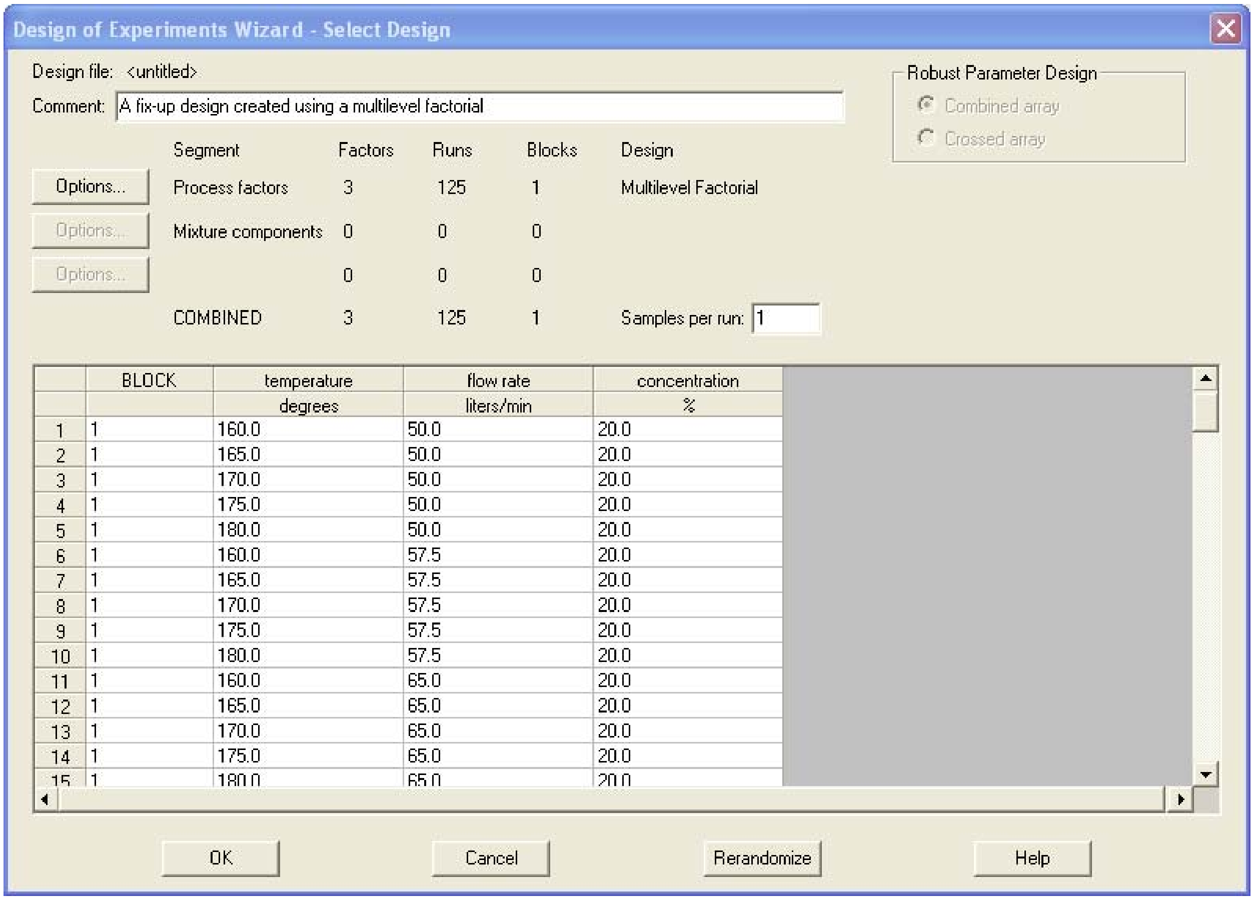
- **Levels** – the number of levels at which each factor should be run.

- **Replicate design** - if a number other than 0 is entered, the entire design will be repeated the indicated number of times.

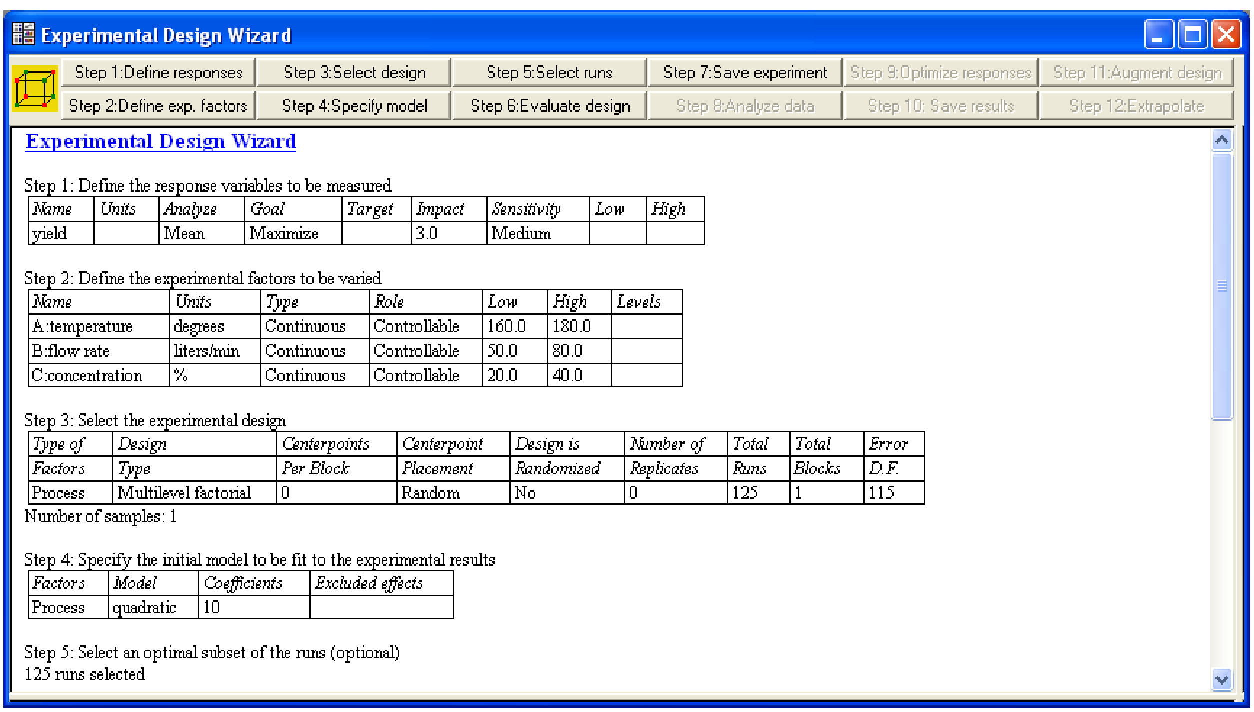
- **Randomize** - check this box to randomly order the runs in the experiment. Randomization is generally a good idea, since it can reduce the effect of lurking variables such as trends over time. However, when replicating the examples in this documentation, **do not** randomize the designs.

For the current example, select 5 for each factor, resulting in a total of 5 x 5 x 5 = 125 runs at each combination of the factors.

The tentatively selected design is displayed in the *Select Design* dialog box:

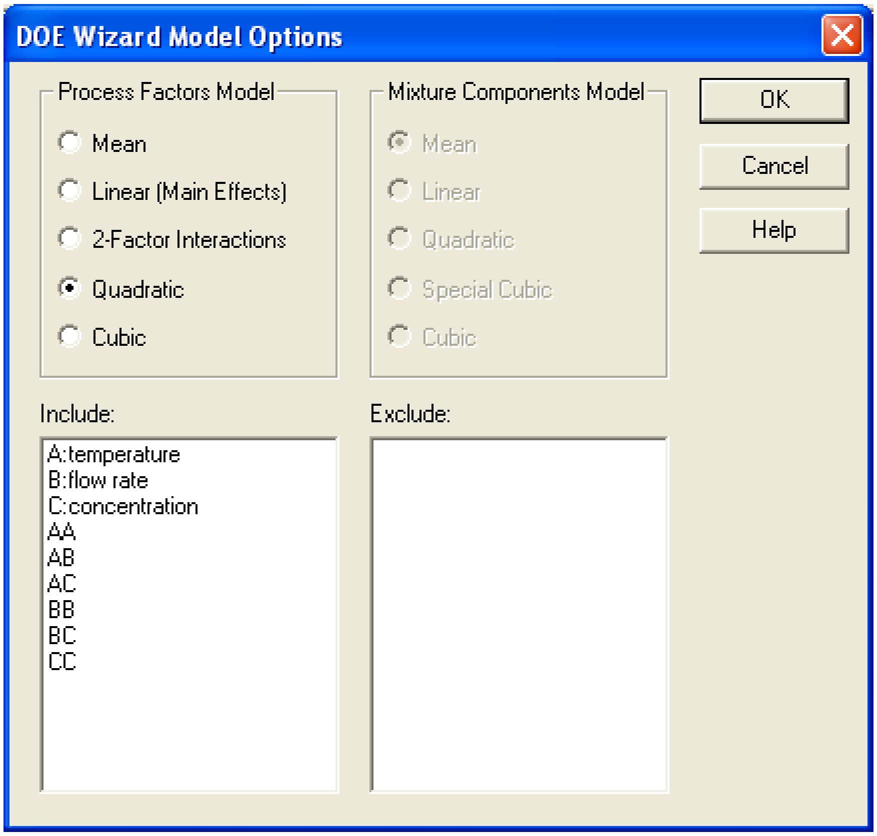


If the design is acceptable, press OK to save it to the STATGRAPHICS DataBook and return to the DOE Wizard's main window, which should now contain a summary of the design:



Step #4: Specify Model

The next step in the design selection process is to specify the model that will be fit to the response data. Pressing the fourth button on the DOE Wizard's toolbar displays a dialog box to make that choice:



The default *quadratic* model includes main effects for each of the 3 experimental factors, 3 two-factor interactions (shown as two-letter combinations with different factors), and 3 quadratic terms (shown as two-letter combinations with identical factors). Selected terms could be excluded by double-clicking on them with the left mouse button.

Step #5: Select runs

The next step is to select a subset of the 125 runs to add to the 6 runs already performed. Before pressing the button for step 5, the original 6 runs need to be added to the datasheet. Click on the DataBook and scroll to the 126th row. Then add the information for the original 6 runs, including the block number and measured response values:

C:\DocData16\multilevel2.sgx

	BLOCK	temperature	flow rate	concentration	yield	C
		degrees	liters/min	%		
120	1	180.0	72.5	40.0		
121	1	160.0	80.0	40.0		
122	1	165.0	80.0	40.0		
123	1	170.0	80.0	40.0		
124	1	175.0	80.0	40.0		
125	1	180.0	80.0	40.0		
126	1	160	50	20	18.7	
127	1	180	75	40	32.9	
128	1	165	80	25	22.1	
129	1	175	80	35	29.0	
130	1	170	55	25	22.8	
131	1	180	75	35	29.9	
132						
133						

multilevel2 / B / C / D / E / F / G / H / I / J / K

After entering the original 6 runs, return to the DOE Wizard and press the button labeled *Step 5*. The following dialog box will be displayed:

Design of Experiments Wizard - Select Runs

	BLOCK	temperature	flow rate	concentration	yield
		degrees	liters/min	%	
1	1	160.0	50.0	20.0	
2	1	165.0	50.0	20.0	
3	1	170.0	50.0	20.0	
4	1	175.0	50.0	20.0	
5	1	180.0	50.0	20.0	
6	1	160.0	57.5	20.0	
7	1	165.0	57.5	20.0	
8	1	170.0	57.5	20.0	
9	1	175.0	57.5	20.0	
10	1	180.0	57.5	20.0	
11	1	160.0	65.0	20.0	
12	1	165.0	65.0	20.0	
13	1	170.0	65.0	20.0	
14	1	175.0	65.0	20.0	
15	1	180.0	65.0	20.0	
16	1	160.0	72.5	20.0	
17	1	165.0	72.5	20.0	
18	1	170.0	72.5	20.0	
19	1	175.0	72.5	20.0	
20	1	180.0	72.5	20.0	
21	1	160.0	80.0	20.0	
22	1	165.0	80.0	20.0	

Number of runs desired: 13

Model coefficients: 10

Select runs using forward algorithm

Select runs using backward algorithm

☑ Apply exchange algorithm at end

D-efficiency:

A-efficiency:

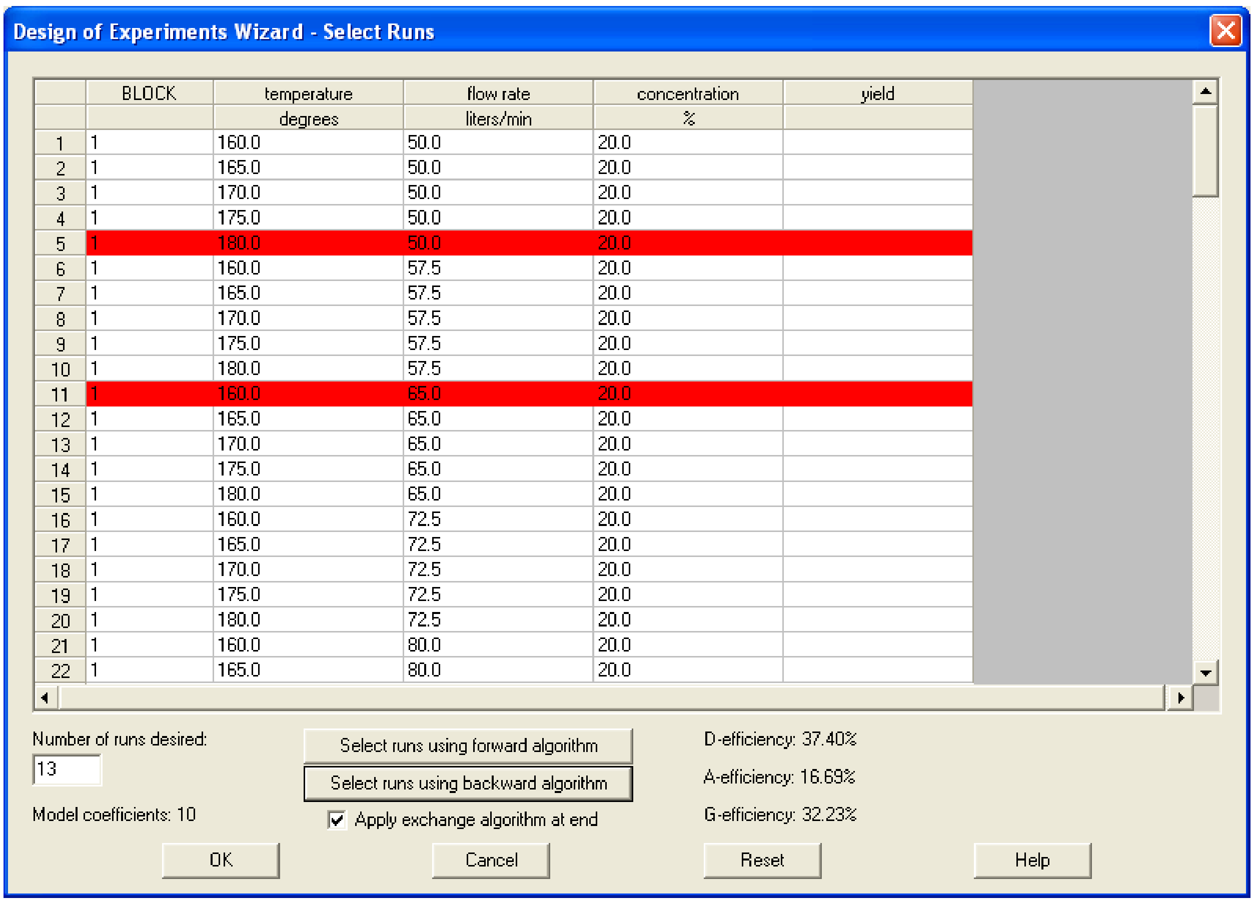
G-efficiency:

OK | Cancel | Reset | Help

In the bottom left is a field where the desired number of runs should be specified. It is usually a good idea to select at least 3 more runs than there are coefficients in the selected model.

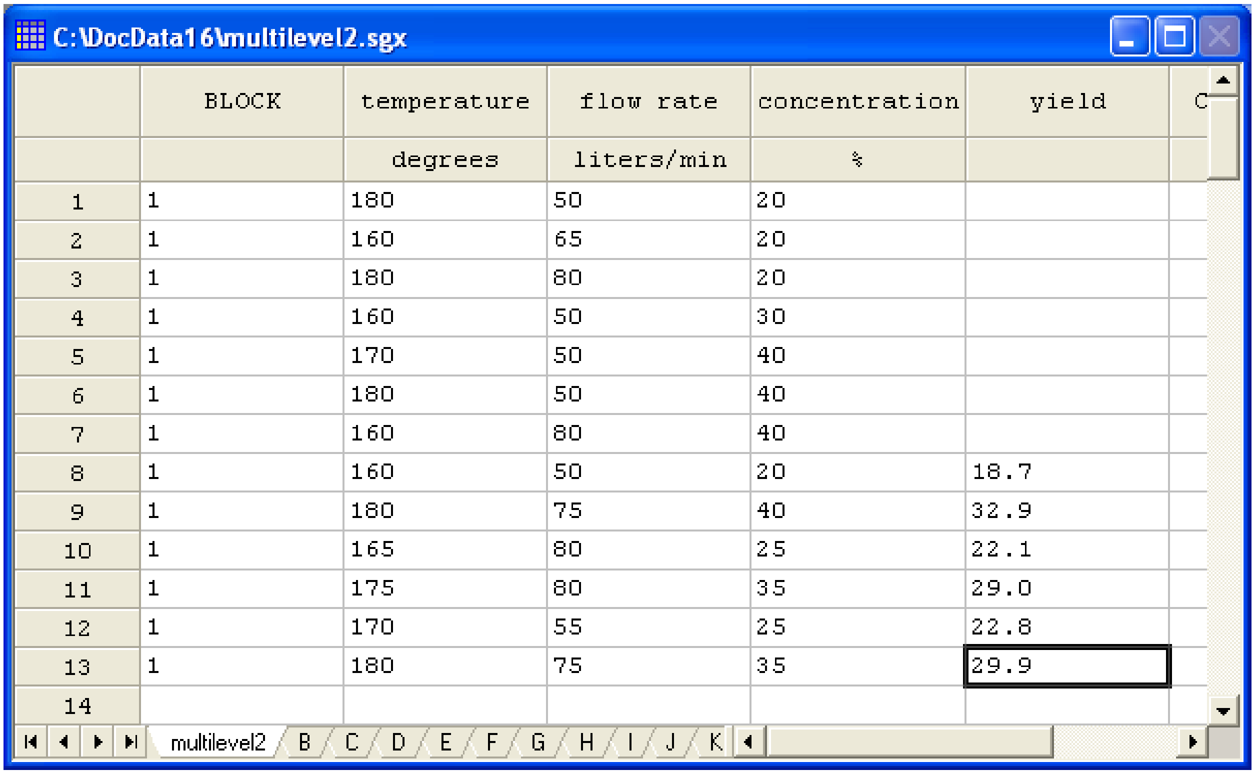
To select the runs, press either of the 2 buttons on the dialog box. Since the number of ways of choosing subsets of the candidate runs is too large to check all possibilities, STATGRAPHICS (like other programs) uses a selection algorithm to choose a subset. The *Forward* method begins with the runs that have already been performed (if any) and adds runs one at a time, adding at each step the run that adds the most to the D-efficiency of the experiment. The *Backward* method begins with all of the candidate runs and removes runs one at a time, removing at each step the run that adds the least to the D-efficiency of the experiment. In either case, once the desired number of runs has been selected, an *exchange* algorithm can be performed. This algorithm tests all pairs of runs consisting of one that has been selected and one that has not, making any exchanges that would increase the efficiency of the experiment. Exchanges continue until no further improvements can be made by switching one run that's been selected with one run that has not been selected.

When the algorithm is complete, the selected rows will be highlighted in red:



	BLOCK	temperature	flow rate	concentration	yield
		degrees	liters/min	%	
1	1	160.0	50.0	20.0	
2	1	165.0	50.0	20.0	
3	1	170.0	50.0	20.0	
4	1	175.0	50.0	20.0	
5	1	180.0	50.0	20.0	
6	1	160.0	57.5	20.0	
7	1	165.0	57.5	20.0	
8	1	170.0	57.5	20.0	
9	1	175.0	57.5	20.0	
10	1	180.0	57.5	20.0	
11	1	160.0	65.0	20.0	
12	1	165.0	65.0	20.0	
13	1	170.0	65.0	20.0	
14	1	175.0	65.0	20.0	
15	1	180.0	65.0	20.0	
16	1	160.0	72.5	20.0	
17	1	165.0	72.5	20.0	
18	1	170.0	72.5	20.0	
19	1	175.0	72.5	20.0	
20	1	180.0	72.5	20.0	
21	1	160.0	80.0	20.0	
22	1	165.0	80.0	20.0	

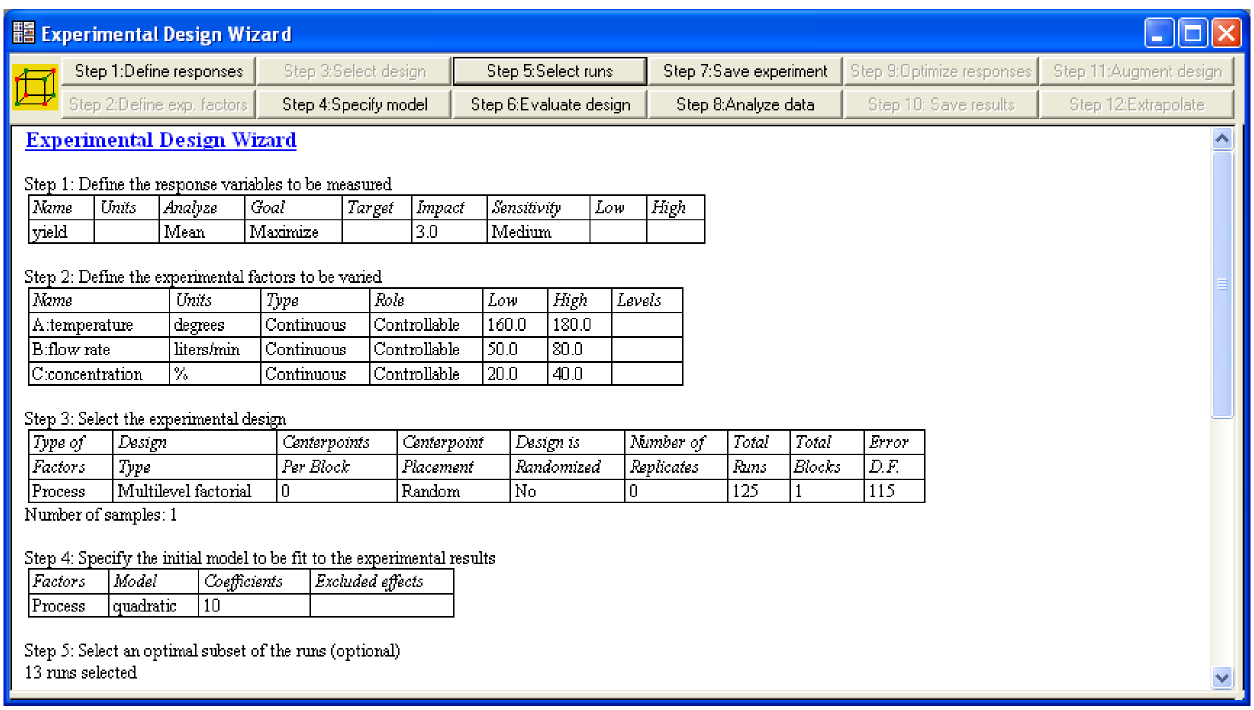
The efficiencies of the selected design will also be displayed. You can try another algorithm or press *OK* to accept the selection, at which point the rows of the datasheet will be reduced to the selected runs:



C:\DocData16\multilevel2.sgx

	BLOCK	temperature	flow rate	concentration	yield
		degrees	liters/min	%	
1	1	180	50	20	
2	1	160	65	20	
3	1	180	80	20	
4	1	160	50	30	
5	1	170	50	40	
6	1	180	50	40	
7	1	160	80	40	
8	1	160	50	20	18.7
9	1	180	75	40	32.9
10	1	165	80	25	22.1
11	1	175	80	35	29.0
12	1	170	55	25	22.8
13	1	180	75	35	29.9
14					

The main DOE Wizard window will reflect the design:



Experimental Design Wizard

Step 1: Define the response variables to be measured

Name	Units	Analyze	Goal	Target	Impact	Sensitivity	Low	High
yield		Mean	Maximize		3.0	Medium		

Step 2: Define the experimental factors to be varied

Name	Units	Type	Role	Low	High	Levels
A:temperature	degrees	Continuous	Controllable	160.0	180.0	
B:flow rate	liters/min	Continuous	Controllable	50.0	80.0	
C:concentration	%	Continuous	Controllable	20.0	40.0	

Step 3: Select the experimental design

Type of Factors	Design Type	Centerpoints Per Block	Centerpoint Placement	Design is Randomized	Number of Replicates	Total Runs	Total Blocks	Error D.F.
Process	Multilevel factorial	0	Random	No	0	125	1	115

Number of samples: 1

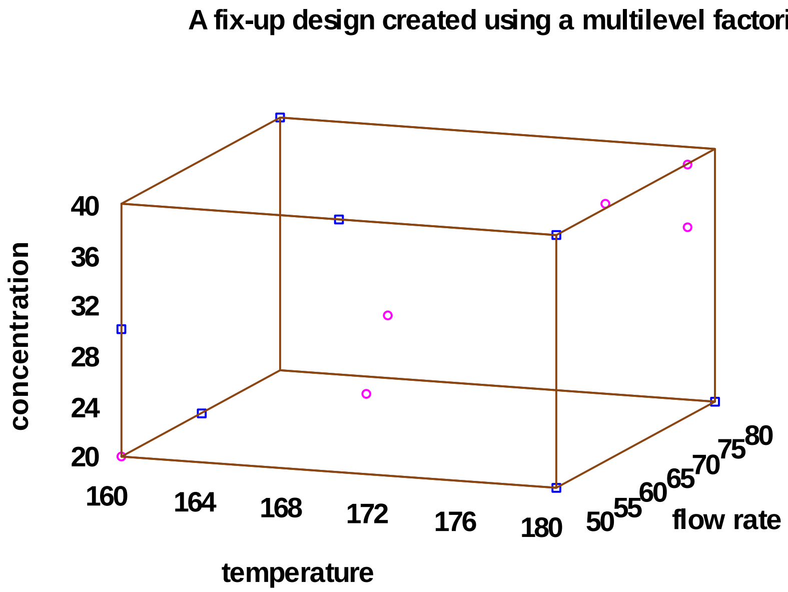
Step 4: Specify the initial model to be fit to the experimental results

Factors	Model	Coefficients	Excluded effects
Process	quadratic	10	

Step 5: Select an optimal subset of the runs (optional)
13 runs selected

If the selection is acceptable, press *Step 7: Save experiment* to save the reduced number of runs.

You can also use the *Design Plot* to display the final design:



The original runs are shown as red asterisks, while the added runs are shown as blue squares.