DOE Wizard – Computer Generated Designs

Summary

The *Computer Generated* designs allow you to create experimental designs which have optimal properties with respect to the estimation of specific statistical models. Given the definition of an experimental region, a model to be estimated, and the number of experimental runs that can be performed, the program searches for a set of runs that maximizes a selected design optimality criteria.

Computer generated designs are included as an option in the Experimental Design Wizard when constructing several types of designs, including screening designs, response surface designs, and mixture experiments. In addition to lower and upper limits for each experimental factor, constraints may be specified based on linear combinations of the factors. The optimality criteria offered include D-optimality, A-optimality, G-optimality, and I-optimality.

The results of previous experimental runs may be included in the design by entering them into the experiment datasheet before requesting that the computer generate new runs. Additional runs will then be added to those already performed to maximize the specified optimality criterion. This allows the procedure to improve a partial design that has already been performed.

Example #1: Optimized Response Surface Experiment

In the book titled Optimal Design of Experiments: A Case Study Approach (2011), Goos and Jones describe an experiment in which the experimenters wished to model the effect of 4 experimental factors on the peel strength of a package. The factors of interest were:

Factor	**Type**	**Low**	**High**	**Units**
temperature	continuous	193	230	degrees C
pressure	continuous	2.2	3.2	bar
speed	continuous	32	50	cpm
supplier	categorical	1	3	

They wished to estimate a model containing:

1. Main effects of all 4 factors.
2. Two-factor interactions amongst all of the factors.
3. Quadratic effects for the 3 continuous factors.

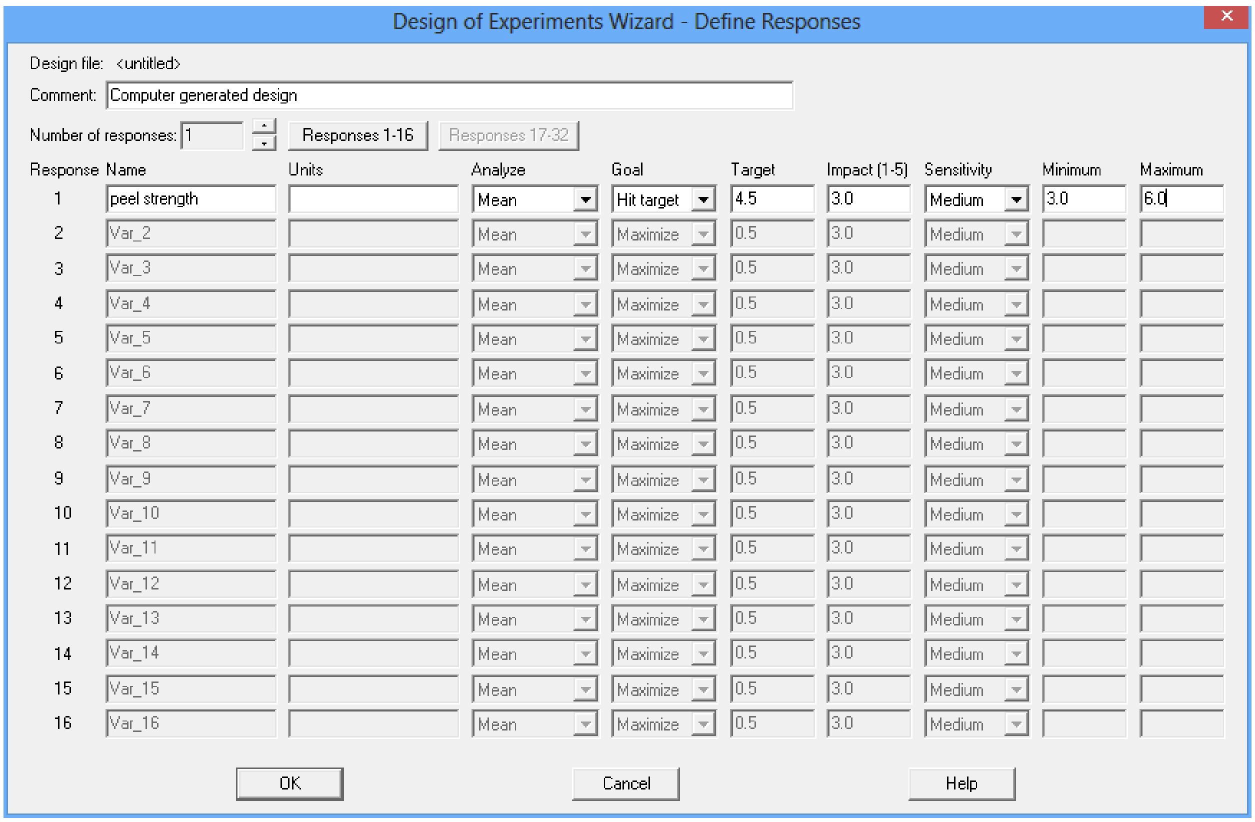
This document will demonstrate how the DOE Wizard can be used to select a good design for this situation.

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:



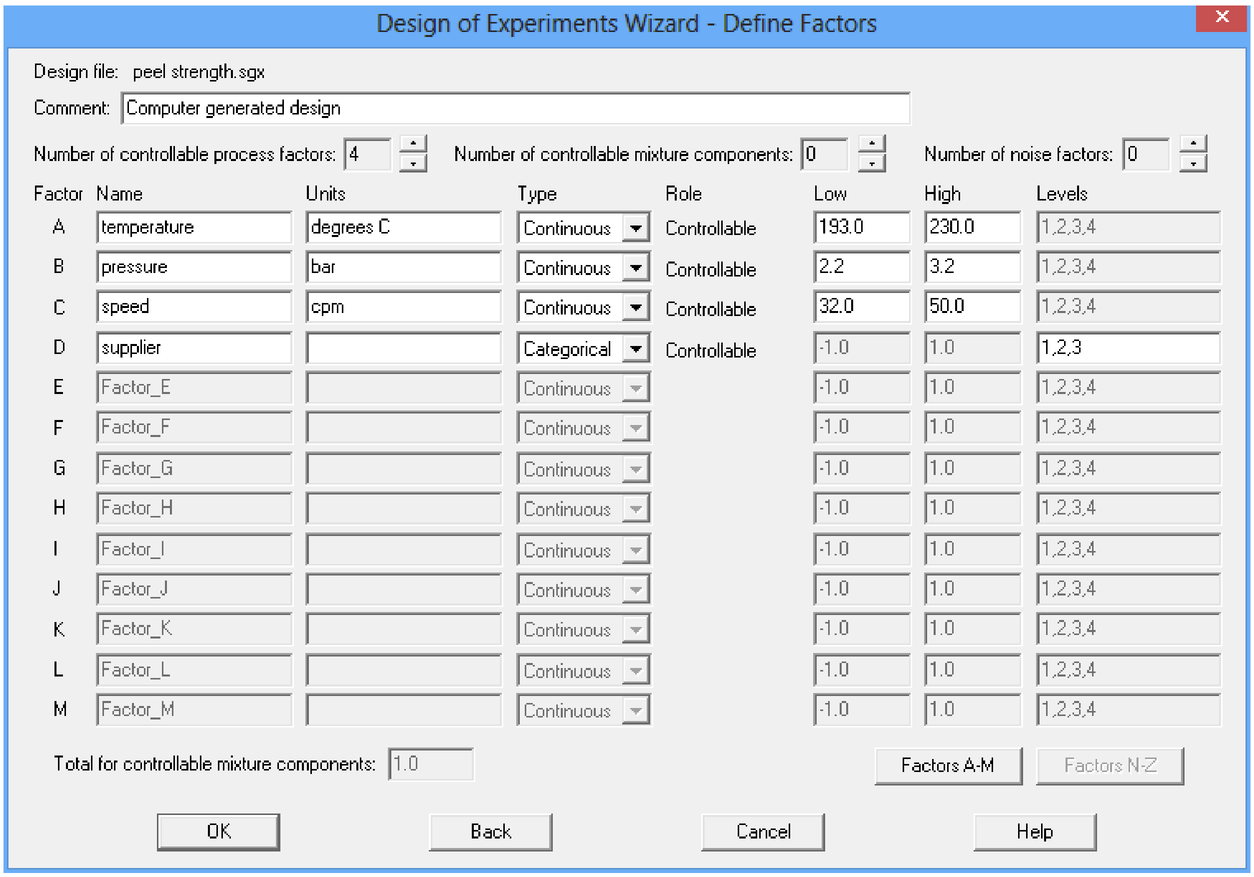
- **Name**: The name for the variable is *peel strength*.
- **Units:** The units are not relevant in this example.
- **Analyze**: The parameter of interest is the *mean* peel strength.
- **Goal**: The goal of the experiment is to hit a target peel strength close to 4.5.
- **Impact**: This is not relevant since there is only one response.

- **Sensitivity**: The importance of being close to the best desired value has little impact on the optimization result when there is only one response.

- **Minimum and Maximum**: The range of acceptable peel strength was 3.0 to 6.0.

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 4 controllable process factors:



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

- **Units** – Units are optional.

- **Type** – Set the type of the first 3 factors to *Continuous*, since they can be set at any value within a continuous interval. The fourth factor is categorical.

- **Role** - The four factors are all *Controllable*.

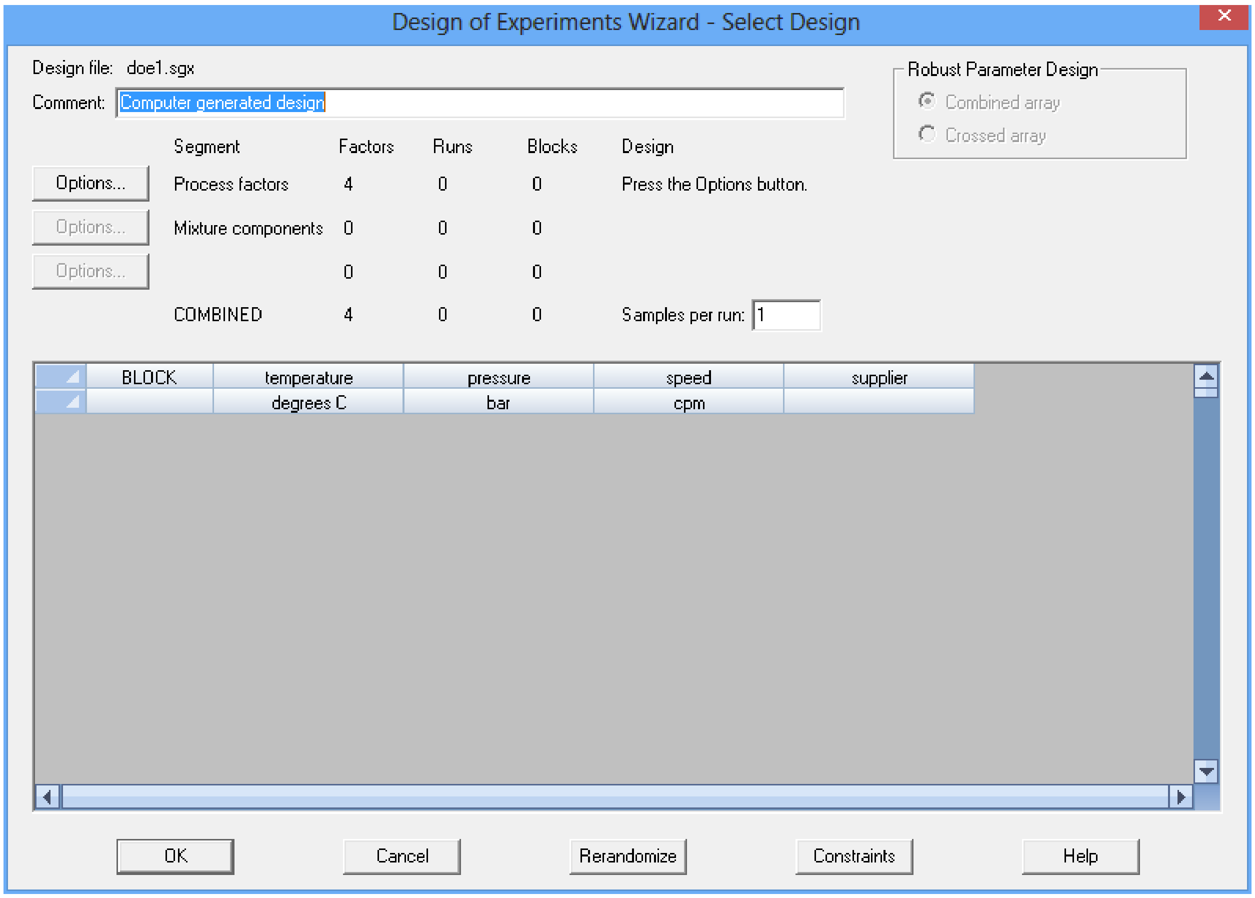
- **Low** - the lower level L_j for the continuous factors.

- **High** - the upper level U_j for the continuous factors.

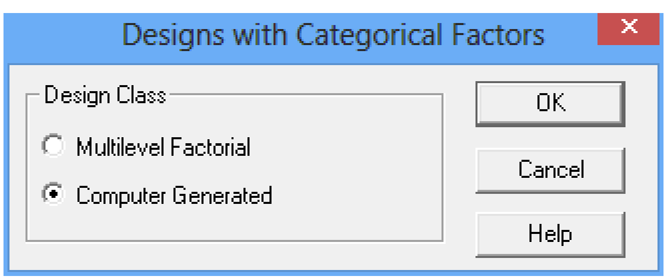
- **Levels** – the allowable levels of the categorical factor *supplier*.

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:



Since the experiment involves a categorical factor, two types of designs are offered:

1. *Multilevel Factorial* - designs involving all combinations of selected levels of each experimental factor. For example, a 3x3x3x3 factorial design would have 81 runs consisting of all combinations of 3 levels of each factor.

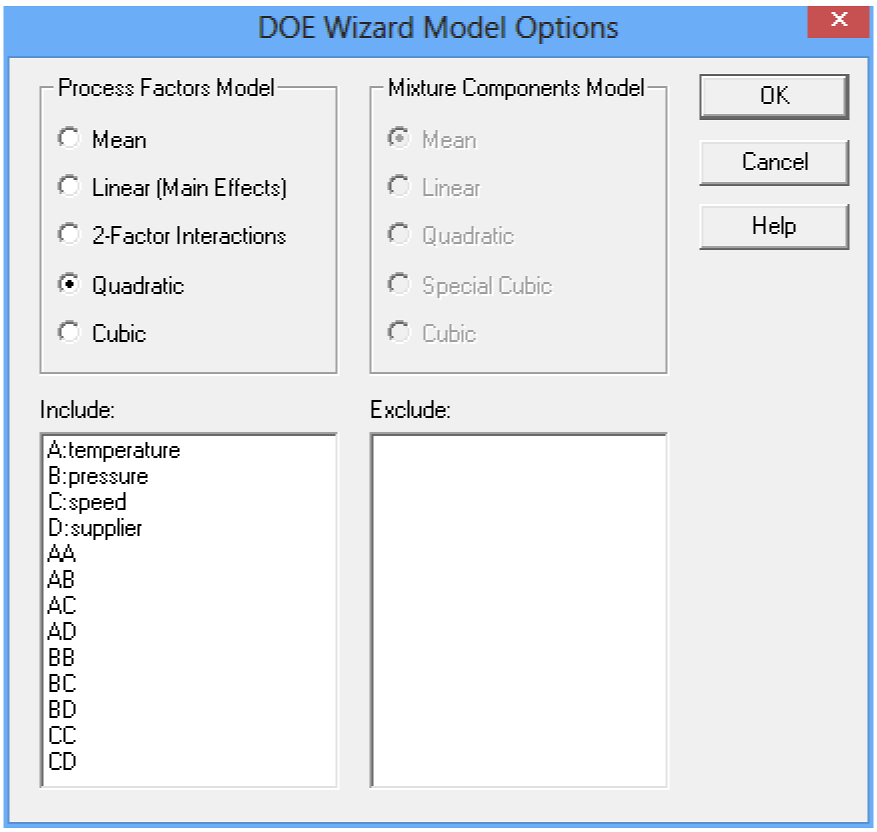
2. *Computer Generated* – If this type of design is selected, the computer will select a set of runs during Step 5 that are optimal for the model to be fit.

Select *Computer Generated* and press *OK*, which will return you to the *Select Design* dialog box:

If satisfied, press *OK* to save the design selection and return to the DOE Wizard's main window, which will now contain a summary of the design:

Step #4: Specify Model

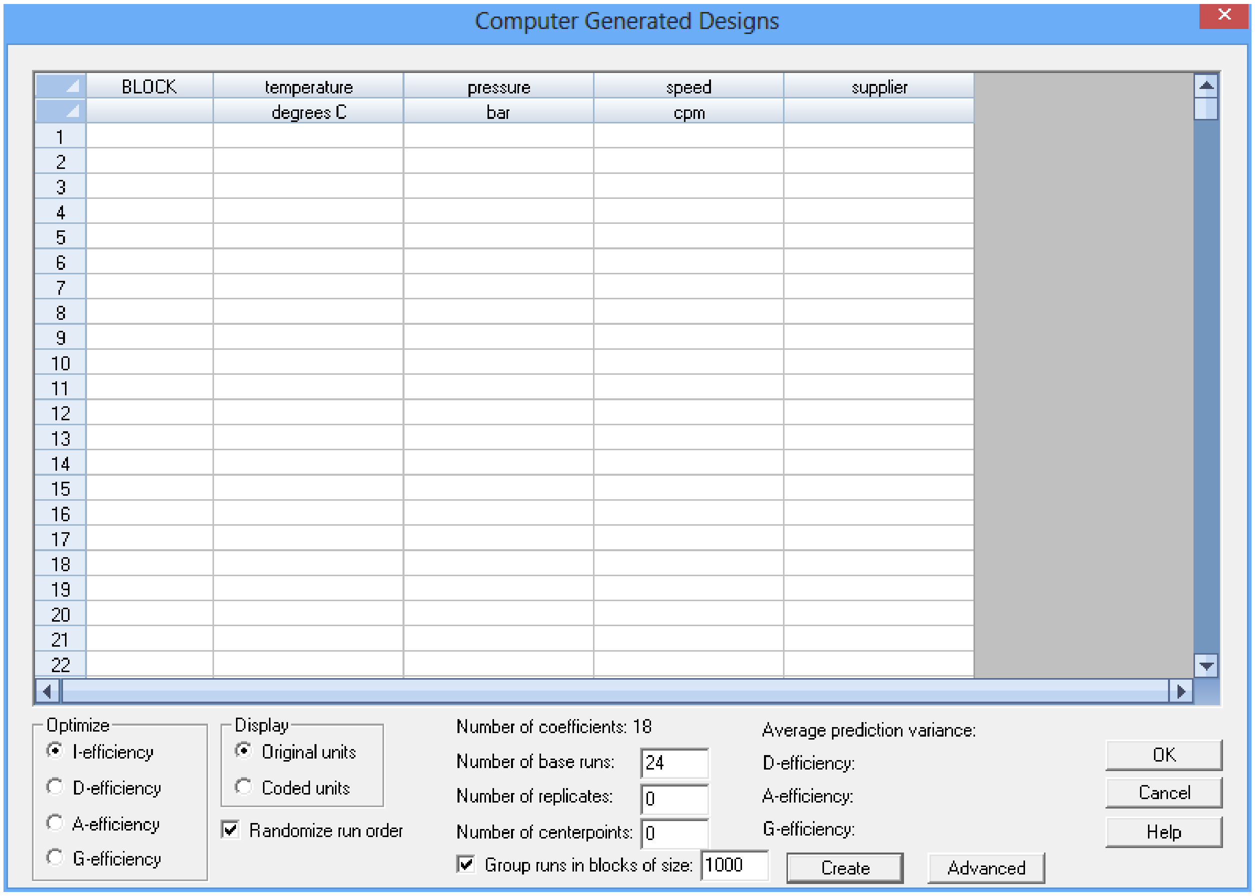
The next step in the design selection process specifies the model that will be fit to the response data. Pressing the fourth button on the DOE Wizard's toolbar displays a dialog box to make that choice:



The *quadratic* model includes main effects for each of the 4 experimental factors, 6 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 selects the runs to be performed. Press the *Step 5: Select Runs* button on the DOE Wizard toolbar to display the following dialog box:



The dialog box shows the number of coefficients in the model to be estimated. At least as many runs must be selected as there are coefficients. Typically, at least 3 additional runs will be added to that number in order to estimate the experimental error.

There are several important fields to be completed:

- *Optimize* - the criterion to be used to select the experimental runs. Briefly, D-efficiency measures the information generated by the experiment about the model parameters. A-efficiency measures the average variance of the estimates of the model parameters. G-efficiency measures the maximum variance of the predicted response at the design points. I-efficiency measures the average prediction variance over the design space. It is often recommended that D-efficiency be selected when creating a screening design and I-efficiency selected when creating a response surface design.

- *Display* - whether the runs should be displayed in their original units or coded units.

- *Randomize run order* - whether the order of the runs should be randomized.

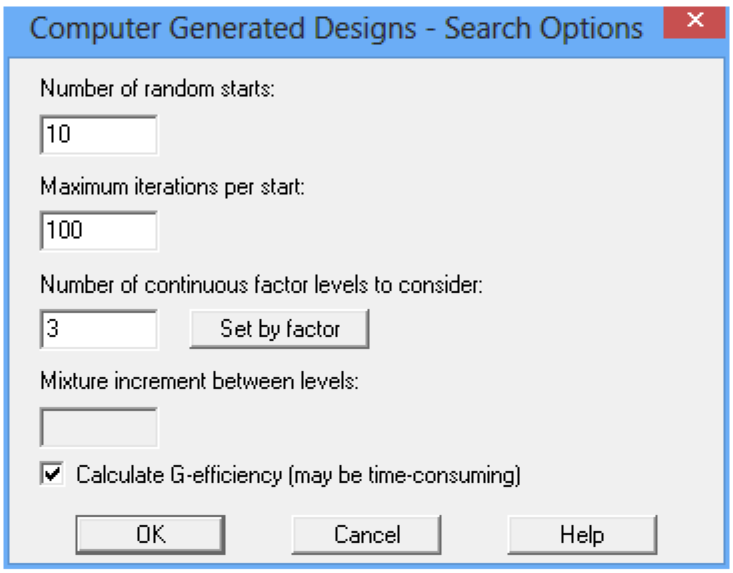
- *Number of base runs* - the number of different combinations of the factors that should be generated. This number must be greater than or equal to the number of coefficients in the model that will be fit to the data.

- *Number of replicates* - the number of additional runs to be added that are replicates of one or more of the base runs.

- *Number of centerpoints* - the number of additional runs to be added at the center of the experimental region.

- *Group runs in blocks of size* – If desired, the new runs may be placed in blocks of the specified size. Additional terms will be added to the model to allow for differences between the blocks. Selecting this option also places any experimental runs that have already been performed in a different block than the new runs.

- *Advanced* button – displays a dialog box for changing the search options:



Number of random starts – number of times the algorithm will search for an optimal design from a different random start.

Maximum iterations per start – maximum number of times that the algorithm will try exchanging runs before a solution is accepted.

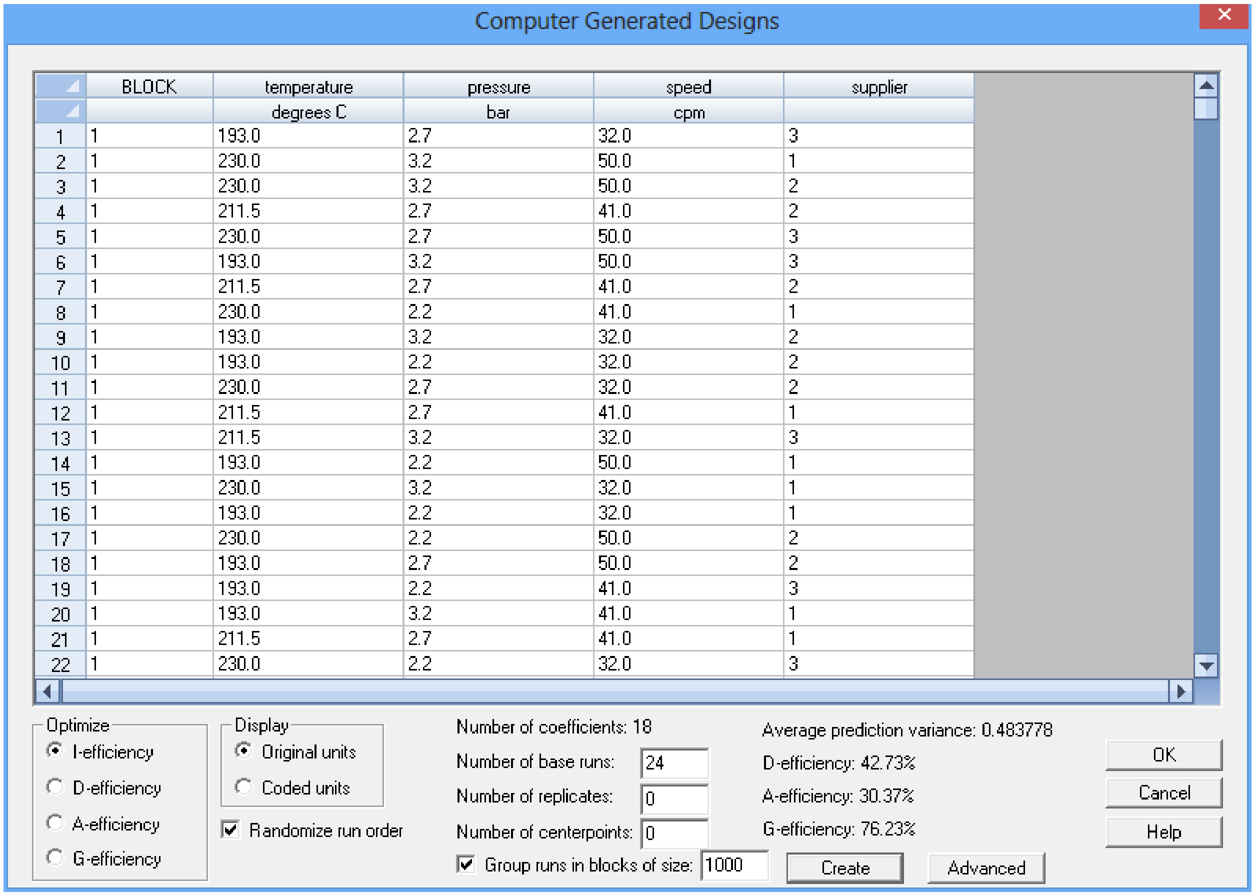
Number of factor levels to consider - number of levels at which experimental runs may be performed for continuous factors, ranging from the low level of the factor to the high level. Specifying a larger number increases the time required to create the design but may improve the design efficiency. If different numbers of levels are desired for different factors, they may be set by pressing the *Set by factor* button.

Mixture increment between levels – for experiments containing mixture components, the amount by which each component will be changed when attempting to find an optimal set of runs.

Calculate G-efficiency - whether the G-efficiency of the design should be calculated and displayed. This statistic requires calculating the prediction variance at every candidate point considered when constructing the design, which can be very large if there are many factors or factor levels to consider.

In this case, the experimenters decided that they could afford to perform 24 runs, all of which will be allocated to base runs (different combinations of the factors). They also specified that the experiment should include runs at only 3 levels of each continuous factor (the low level, the center level, and the high level).

To generate an I-optimal design, set *Optimize* to *I-efficiency* and press the *Create* button. When the algorithm is complete, the selected experimental runs will be added to the dialog box:

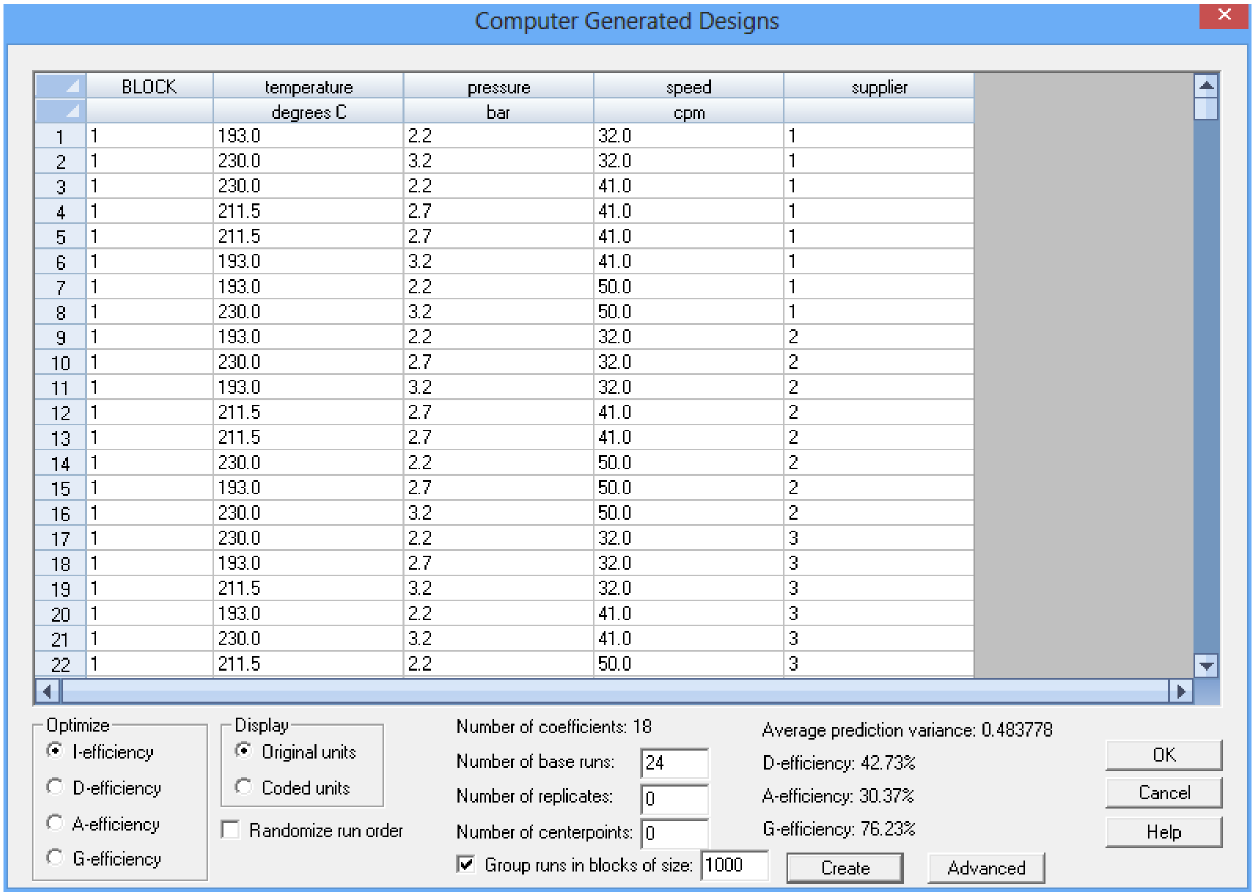


	BLOCK	temperature	pressure	speed	supplier
		degrees C	bar	cpm	
1	1	193.0	2.7	32.0	3
2	1	230.0	3.2	50.0	1
3	1	230.0	3.2	50.0	2
4	1	211.5	2.7	41.0	2
5	1	230.0	2.7	50.0	3
6	1	193.0	3.2	50.0	3
7	1	211.5	2.7	41.0	2
8	1	230.0	2.2	41.0	1
9	1	193.0	3.2	32.0	2
10	1	193.0	2.2	32.0	2
11	1	230.0	2.7	32.0	2
12	1	211.5	2.7	41.0	1
13	1	211.5	3.2	32.0	3
14	1	193.0	2.2	50.0	1
15	1	230.0	3.2	32.0	1
16	1	193.0	2.2	32.0	1
17	1	230.0	2.2	50.0	2
18	1	193.0	2.7	50.0	2
19	1	193.0	2.2	41.0	3
20	1	193.0	3.2	41.0	1
21	1	211.5	2.7	41.0	1
22	1	230.0	2.2	32.0	3

The efficiencies of the selected design and the average prediction variance are also displayed.

Note: The efficiency values should only be used to compare alternative experiments containing the same number of runs. Their absolute magnitude is usually not important.

If you uncheck the button labeled *Randomize run order*, the runs will be displayed in standard order:



Note that 8 runs are performed for each supplier.

You can change some settings and try again or press *OK* to accept the selection, at which point the selected runs will be placed in datasheet A:

C:\Data\Version17 Tests\doewiz compgen rsm.sgx

	BLOCK	temperature	pressure	speed	supplier	peel strength
		degrees C	bar	cpm		
1	1	193.0	2.7	32.0	3	
2	1	230.0	3.2	50.0	1	
3	1	230.0	3.2	50.0	2	
4	1	211.5	2.7	41.0	2	
5	1	230.0	2.7	50.0	3	
6	1	193.0	3.2	50.0	3	
7	1	211.5	2.7	41.0	2	
8	1	230.0	2.2	41.0	1	
9	1	193.0	3.2	32.0	2	
10	1	193.0	2.2	32.0	2	
11	1	230.0	2.7	32.0	2	
12	1	211.5	2.7	41.0	1	
13	1	211.5	3.2	32.0	3	
14	1	193.0	2.2	50.0	1	
15	1	230.0	3.2	32.0	1	
16	1	193.0	2.2	32.0	1	
17	1	230.0	2.2	50.0	2	
18	1	193.0	2.7	50.0	2	
19	1	193.0	2.2	41.0	3	
20	1	193.0	3.2	41.0	1	
21	1	211.5	2.7	41.0	1	
22	1	230.0	2.2	32.0	3	
23	1	230.0	3.2	41.0	3	
24	1	211.5	2.2	50.0	3	

A B C

The main DOE Wizard window will reflect the design:

Experimental Design Wizard

Step 1:Define responses	Step 3:Specify model	Step 5:Select runs	Step 7:Save experiment	Step 9:Optimize responses	Step 11:Augment design
Step 2:Define exp. factors	Step 4:Select design	Step 6:Evaluate design	Step 8:Analyze data	Step 10: Save results	Step 12:Extrapolate

Experimental Design Wizard

Step 1: Define the response variables to be measured

Name	*Units*	*Analyze*	*Goal*	*Target*	*Impact*	*Sensitivity*	*Low*	*High*
peel strength		Mean	Hit target	4.5	3.0	Medium	3.0	6.0

Step 2: Define the experimental factors to be varied

Name	*Units*	*Type*	*Role*	*Low*	*High*	*Levels*
A:temperature	degrees C	Continuous	Controllable	193.0	230.0	
B:pressure	bar	Continuous	Controllable	2.2	3.2	
C:speed	cpm	Continuous	Controllable	32.0	50.0	
D:supplier		Categorical	Controllable			1,2,3

Step 3: Specify the initial model to be fit to the experimental results

Factors	*Model*	*Coefficients*	*Excluded effects*
Process	quadratic	18	

Step 4: 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	Computer generated design							

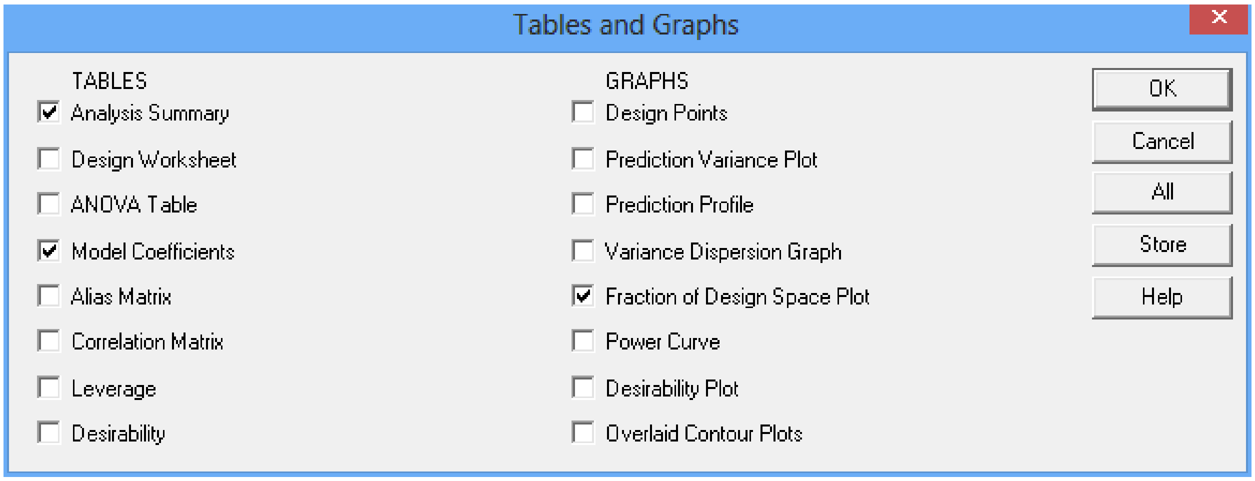
Number of samples per run: 1

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

If the selection is acceptable, press *Step 7: Save experiment* to save the design.

Evaluate Design

After the design has been created, press the button labeled *Step 6: Evaluate Design* on the DOE Wizard toolbar to display various design diagnostics:



The *Model Coefficients* table shows the relative standard error of each coefficient in the model to be estimated:

Model Coefficients

Coefficient	*Standard Error*	*VIF*	*Ri-Squared*	*Power at SN = 0.5*	*Power at SN = 1.0*	*Power at SN = 2.0*
constant	0.466817			7.41%	14.85%	43.68%
A	0.245327	1.08333	0.0769231	13.90%	40.32%	92.11%
B	0.259094	1.07407	0.0689655	12.96%	36.90%	89.23%
C	0.259094	1.07407	0.0689655	12.96%	36.90%	89.23%
D	0.311438	1.5519	0.355627	10.47%	27.32%	76.28%
D	0.296604	1.40759	0.289564	11.04%	29.58%	80.10%
AA	0.584816	1.53904	0.350246	6.53%	11.22%	30.27%
AB	0.280009	1.0846	0.0780028	11.79%	32.50%	84.30%
AC	0.280009	1.0846	0.0780028	11.79%	32.50%	84.30%
AD	0.350264	1.47222	0.320755	9.31%	22.65%	66.48%
AD	0.340207	1.38889	0.28	9.57%	23.71%	68.95%
BB	0.505853	1.36473	0.267254	7.05%	13.36%	38.37%
BC	0.312182	1.16949	0.144928	10.44%	27.21%	76.08%
BD	0.387896	1.50463	0.335385	8.50%	19.36%	57.89%
BD	0.350264	1.47222	0.320755	9.31%	22.65%	66.48%
CC	0.505853	1.36473	0.267254	7.05%	13.36%	38.37%
CD	0.360041	1.2963	0.228571	9.07%	21.69%	64.14%
CD	0.350264	1.22685	0.184906	9.31%	22.65%	66.48%

alpha = 5.0%, sigma estimated from total error with 6 d.f.

The standard error is relative in the sense that it is the multiple of the residual standard error, which is not known until the experiment has been performed. Of particular interest are the VIFs (Variance Inflation Factors), which show how much the variance of each coefficient has been increased relative to a perfectly orthogonal design. Since the largest VIF is less than 1.4, there has been relatively little variance inflation.

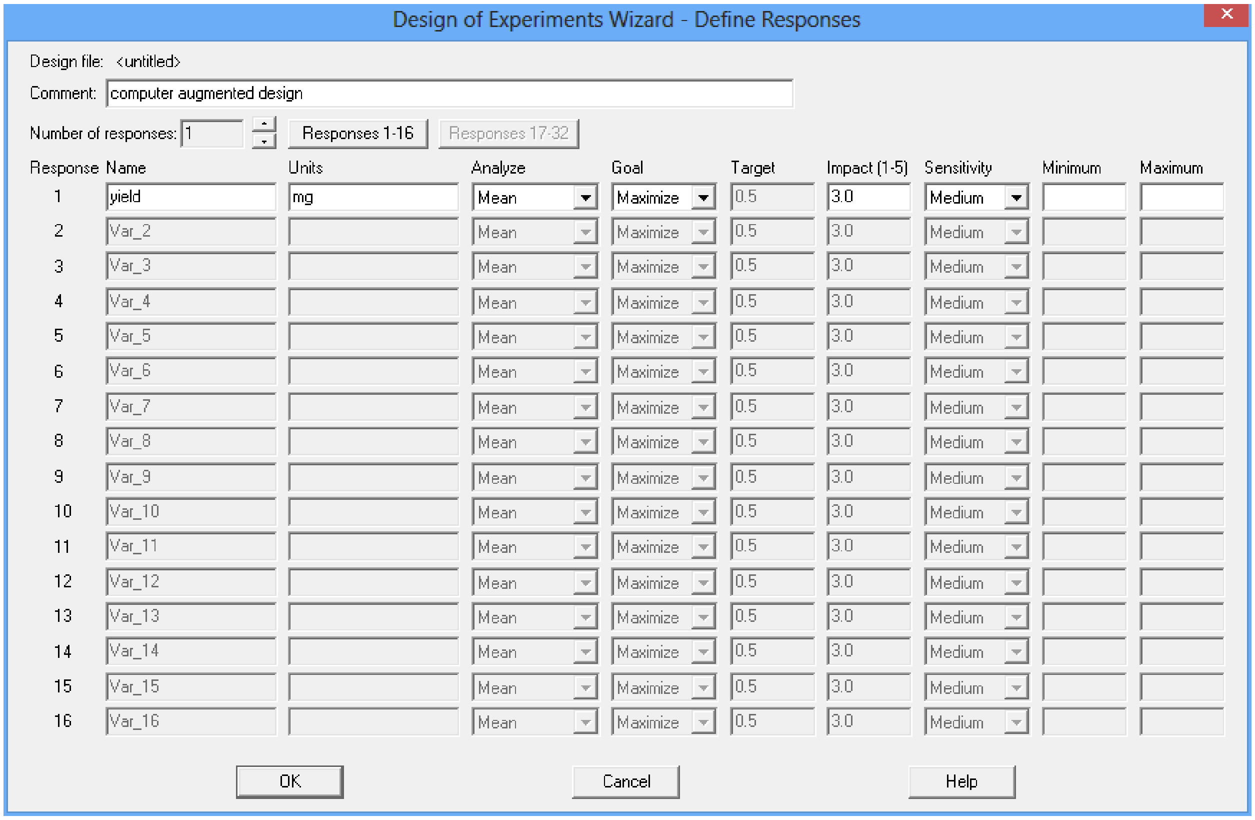
Example #2: Adding Runs to an Existing Experiment

After running a set of experimental runs, it is not uncommon to desire that additional runs be added to the original design. This may happen because of various reasons:

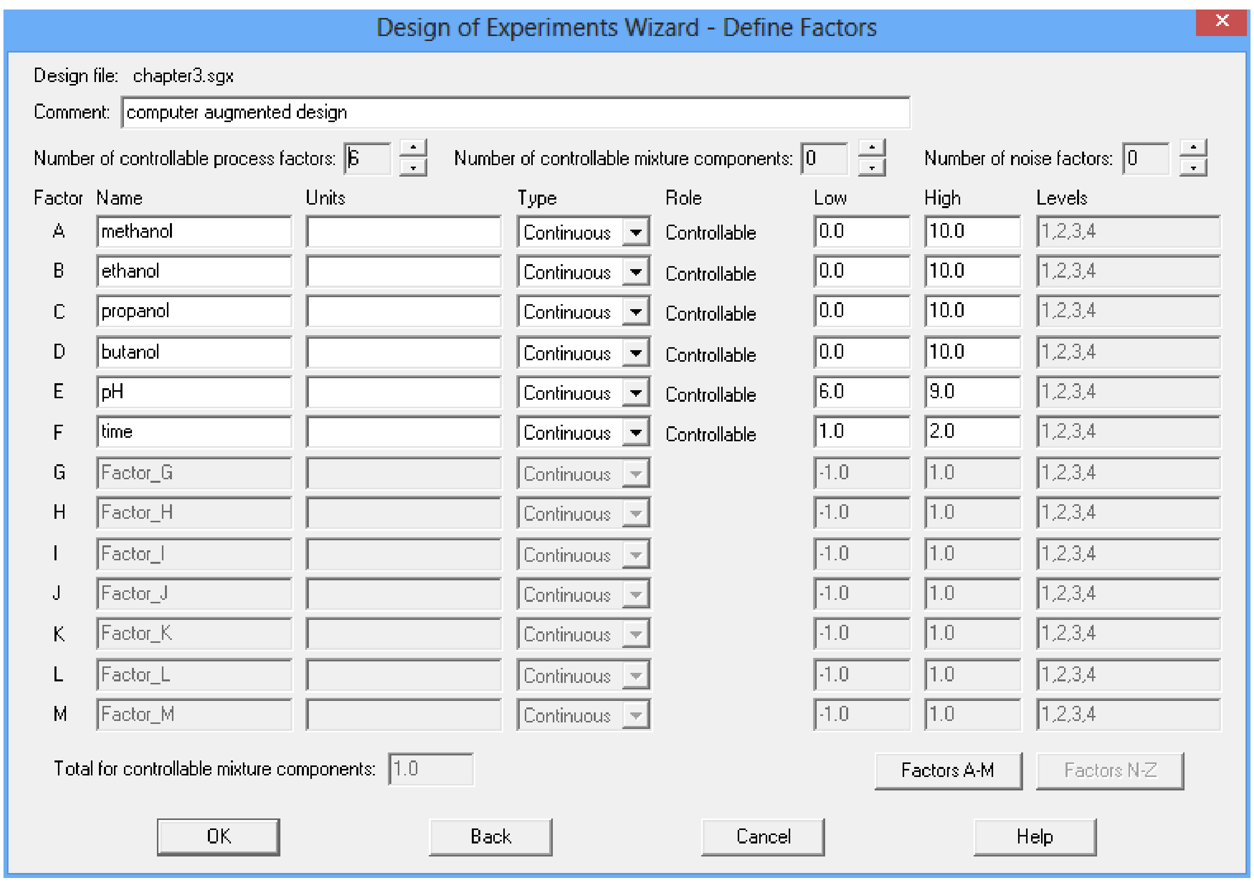
1. The initial design might have been constructed poorly. It is not uncommon to find experimenters who begin with a haphazard approach to factor level selection and only realize later that the interpretation of their results would be much easier with a designed experiment.
2. In other cases, unexpected effects may be observed in the first design and additional runs may be desired to follow up on the initial discoveries.
3. In the case of screening designs involving many factors, there may be considerable confounding amongst the effects in the initial design. Breaking apart specific confounding patterns then requires additional runs.
4. Conversion of a screening design to a response design may require additional runs to add higher-order terms to the statistical model.

Goos and Jones (2011) describe a situation where a 12-run experiment was performed involving 6 factors, each at 2 levels. The design was constructed to estimate a model involving all 6 main effects and 1 specific 2-factor interaction. After analyzing the results of the first experiment, they wished to augment the design so that it could estimate 5 additional interactions.

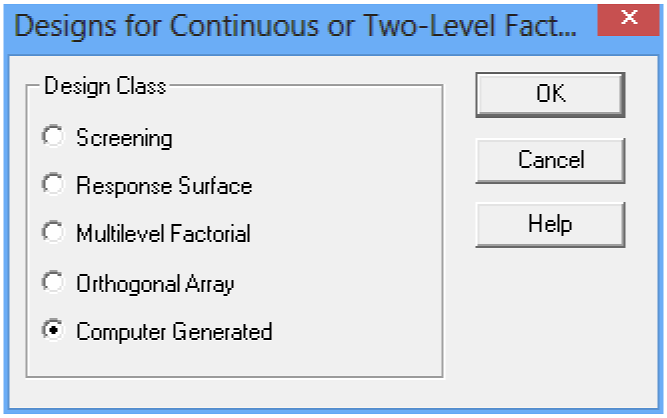
To reconstruct their experiment, start with a StatFolio and select *DOE – Experimental Design Wizard* from the main menu. In Step 1, enter the following information about the response variable:



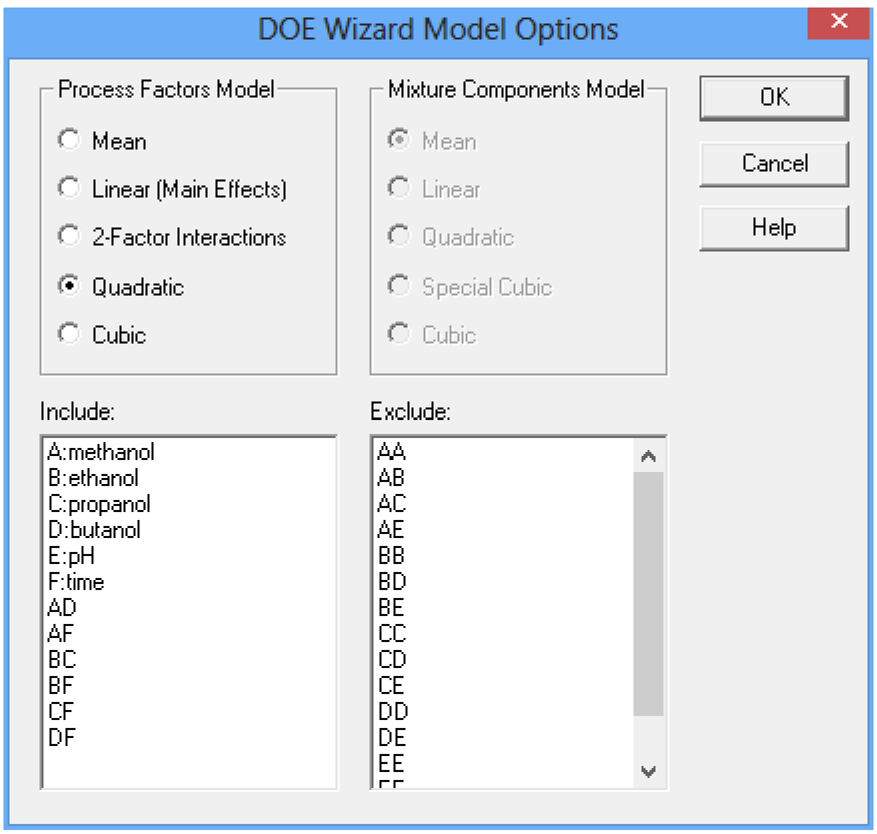
In step 2, enter the following information about the experimental factors:



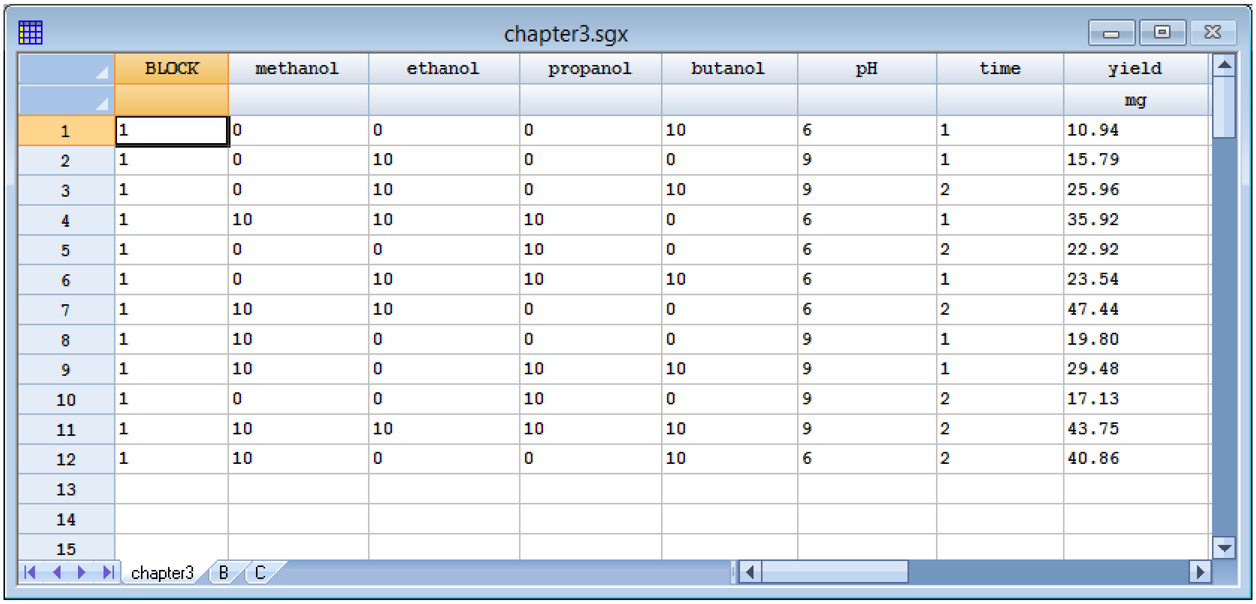
In Step 3, select a *Computer Generated* design:



In Step 4, select a model involving all 6 main effects and 6 two-factor interactions:



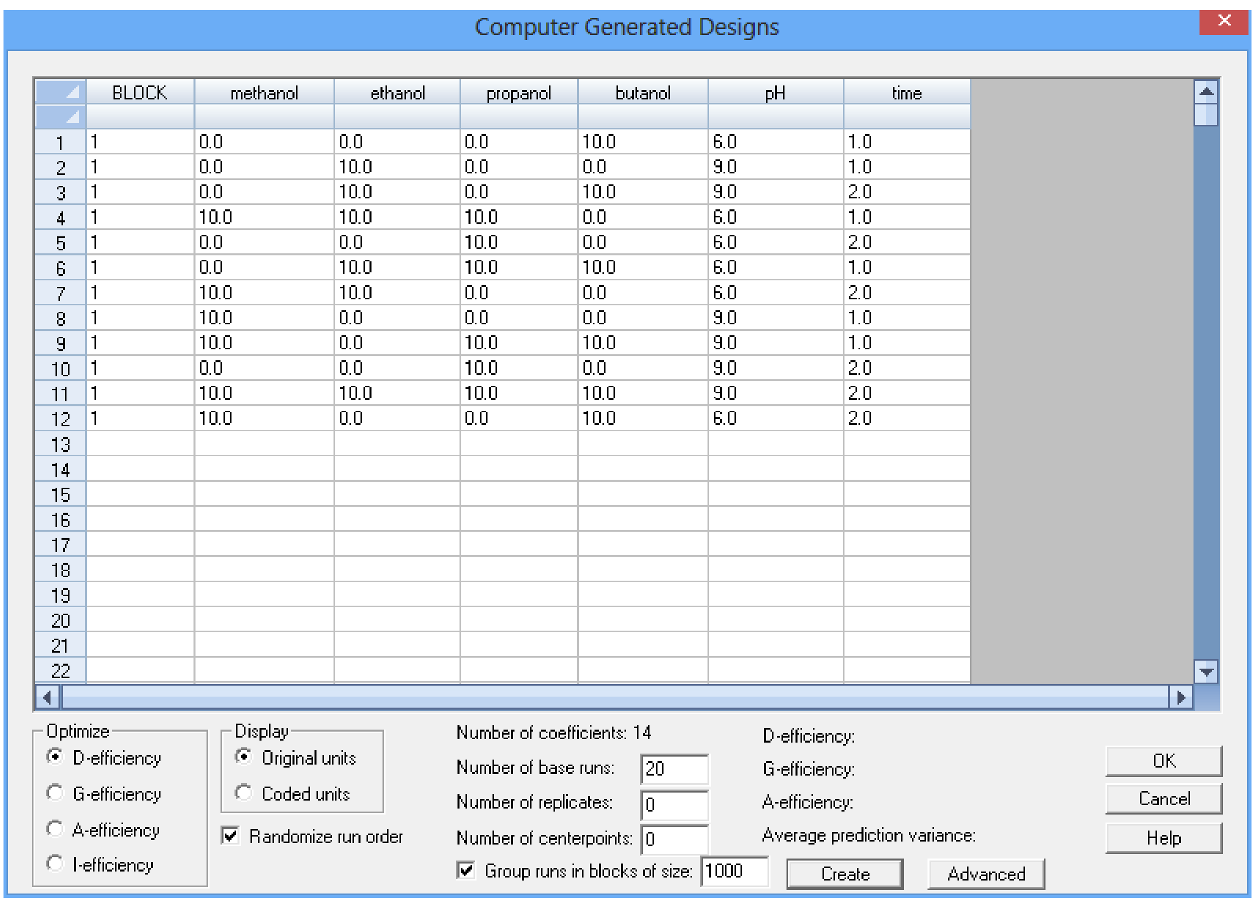
Before selecting the experimental runs, go to the DataBook and put in information about the 12 runs that have already been performed:



	BLOCK	methanol	ethanol	propanol	butanol	pH	time	yield
								mg
1	1	0	0	0	10	6	1	10.94
2	1	0	10	0	0	9	1	15.79
3	1	0	10	0	10	9	2	25.96
4	1	10	10	10	0	6	1	35.92
5	1	0	0	10	0	6	2	22.92
6	1	0	10	10	10	6	1	23.54
7	1	10	10	0	0	6	2	47.44
8	1	10	0	0	0	9	1	19.80
9	1	10	0	10	10	9	1	29.48
10	1	0	0	10	0	9	2	17.13
11	1	10	10	10	10	9	2	43.75
12	1	10	0	0	10	6	2	40.86
13								
14								
15								

Be sure to include the values for *Yield*, so that the algorithm will know that those runs have already been performed.

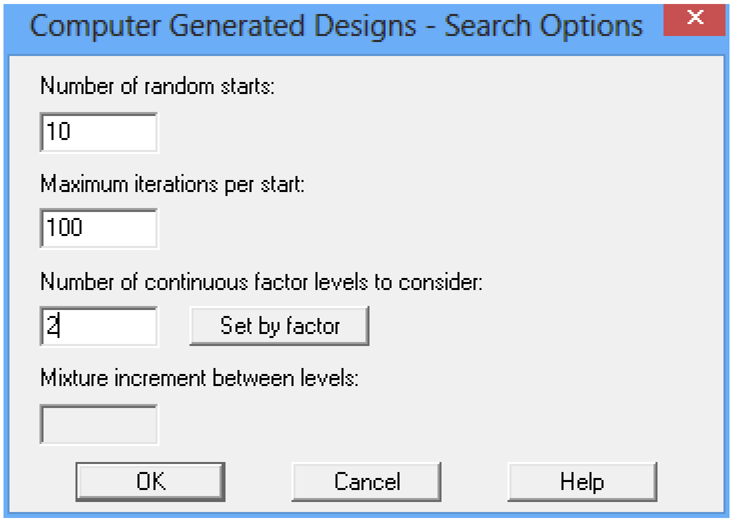
Now return to the DOE Wizard and press the button labeled *Step 5: Select runs*. The dialog box will show the 12 runs that have already been performed:



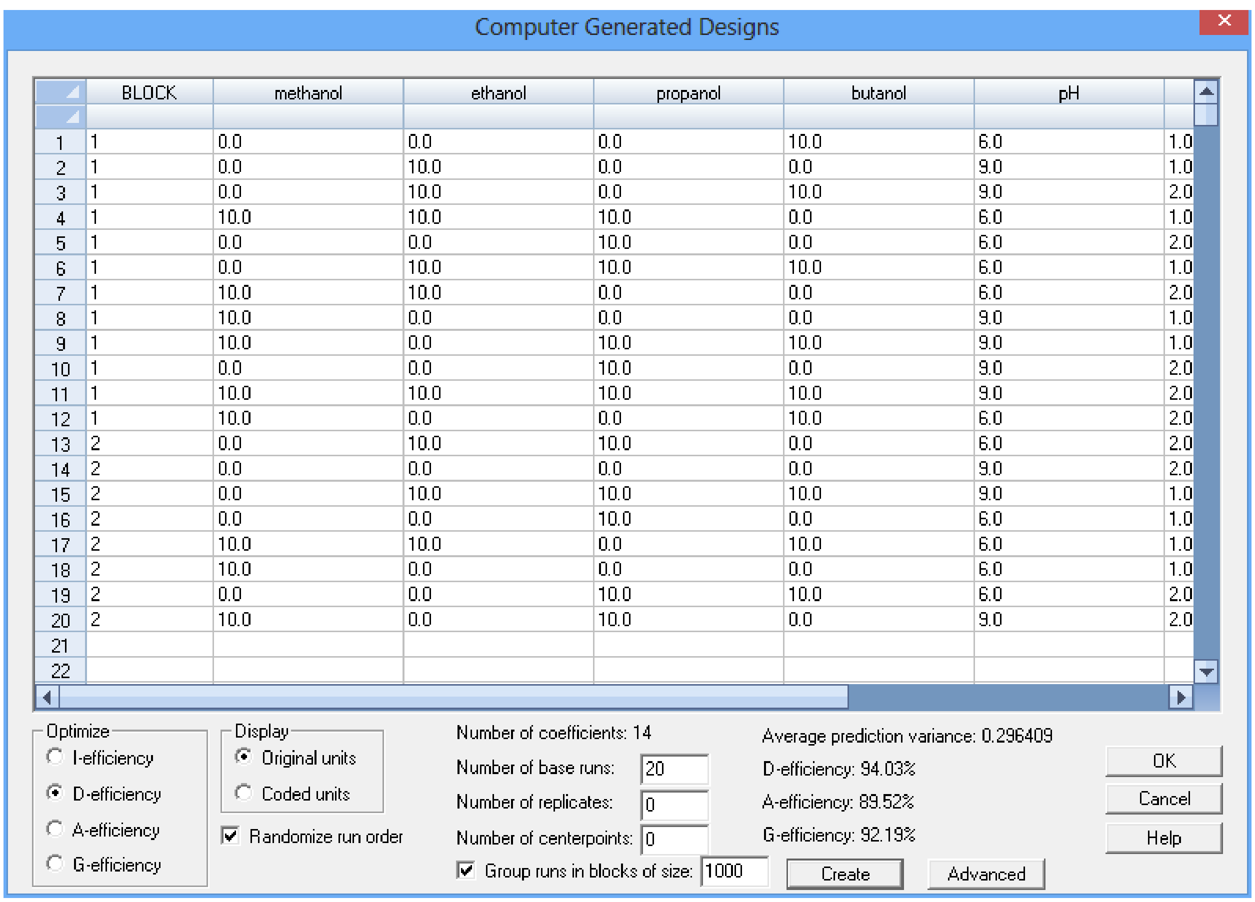
	BLOCK	methanol	ethanol	propanol	butanol	pH	time
1	1	0.0	0.0	0.0	10.0	6.0	1.0
2	1	0.0	10.0	0.0	0.0	9.0	1.0
3	1	0.0	10.0	0.0	10.0	9.0	2.0
4	1	10.0	10.0	10.0	0.0	6.0	1.0
5	1	0.0	0.0	10.0	0.0	6.0	2.0
6	1	0.0	10.0	10.0	10.0	6.0	1.0
7	1	10.0	10.0	0.0	0.0	6.0	2.0
8	1	10.0	0.0	0.0	0.0	9.0	1.0
9	1	10.0	0.0	10.0	10.0	9.0	1.0
10	1	0.0	0.0	10.0	0.0	9.0	2.0
11	1	10.0	10.0	10.0	10.0	9.0	2.0
12	1	10.0	0.0	0.0	10.0	6.0	2.0
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							

Now:

1. Set *Optimize* to *D-efficiency* to create a D-optimal design.
2. Set *Number of base runs* to 20 to add 8 additional runs to the original 12 runs.
3. Push the *Advanced* button and set *Number of factor levels to consider* to 2 to consider only runs involving the low and high levels of each factor:



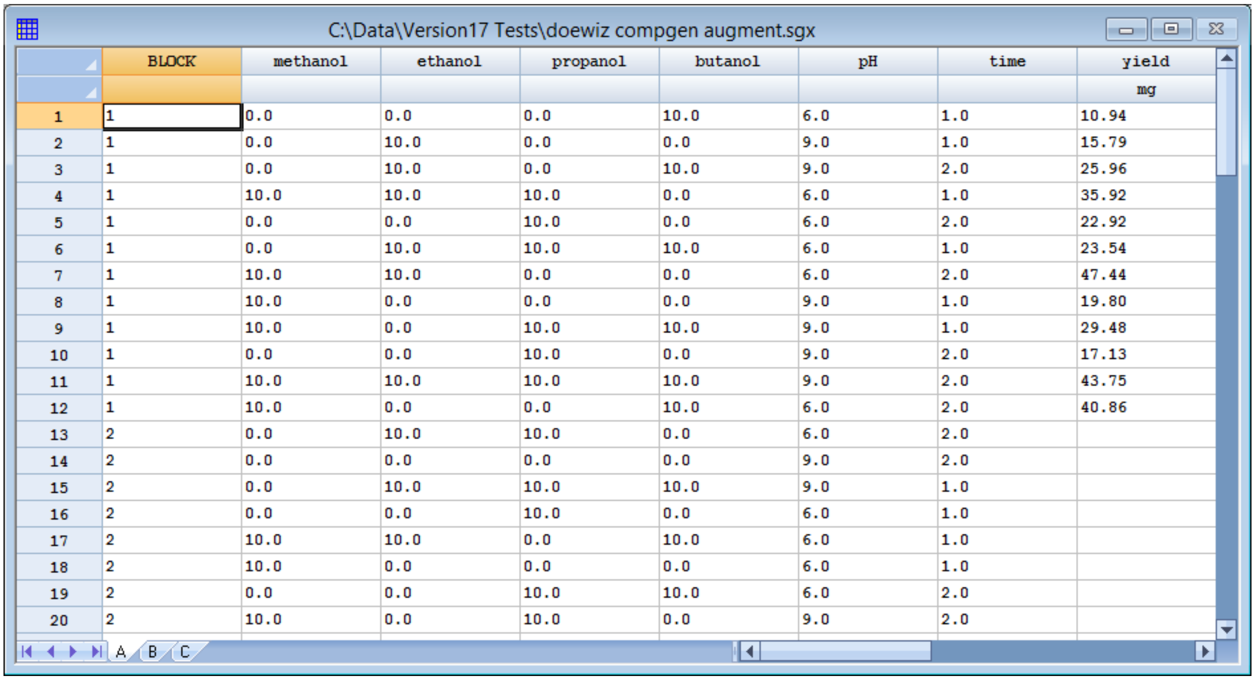
4. Push the *Create* button to initiate the design creation process. When the alogorithm is complete, 8 additional runs will be added to the dialog box:



	BLOCK	methanol	ethanol	propanol	butanol	pH	
1	1	0.0	0.0	0.0	10.0	6.0	1.0
2	1	0.0	10.0	0.0	0.0	9.0	1.0
3	1	0.0	10.0	0.0	10.0	9.0	2.0
4	1	10.0	10.0	10.0	0.0	6.0	1.0
5	1	0.0	0.0	10.0	0.0	6.0	2.0
6	1	0.0	10.0	10.0	10.0	6.0	1.0
7	1	10.0	10.0	0.0	0.0	6.0	2.0
8	1	10.0	0.0	0.0	0.0	9.0	1.0
9	1	10.0	0.0	10.0	10.0	9.0	1.0
10	1	0.0	0.0	10.0	0.0	9.0	2.0
11	1	10.0	10.0	10.0	10.0	9.0	2.0
12	1	10.0	0.0	0.0	10.0	6.0	2.0
13	2	0.0	10.0	10.0	0.0	6.0	2.0
14	2	0.0	0.0	0.0	0.0	9.0	2.0
15	2	0.0	10.0	10.0	10.0	9.0	1.0
16	2	0.0	0.0	10.0	0.0	6.0	1.0
17	2	10.0	10.0	0.0	10.0	6.0	1.0
18	2	10.0	0.0	0.0	0.0	6.0	1.0
19	2	0.0	0.0	10.0	10.0	6.0	2.0
20	2	10.0	0.0	10.0	0.0	9.0	2.0
21							
22							

5. Push *OK* to save the design.

The new runs have now been added to the DataBook:


C:\Data\Version17 Tests\doewiz compgen augment.sgx

	BLOCK	methanol	ethanol	propanol	butanol	pH	time	yield
								mg
1	1	0.0	0.0	0.0	10.0	6.0	1.0	10.94
2	1	0.0	10.0	0.0	0.0	9.0	1.0	15.79
3	1	0.0	10.0	0.0	10.0	9.0	2.0	25.96
4	1	10.0	10.0	10.0	0.0	6.0	1.0	35.92
5	1	0.0	0.0	10.0	0.0	6.0	2.0	22.92
6	1	0.0	10.0	10.0	10.0	6.0	1.0	23.54
7	1	10.0	10.0	0.0	0.0	6.0	2.0	47.44
8	1	10.0	0.0	0.0	0.0	9.0	1.0	19.80
9	1	10.0	0.0	10.0	10.0	9.0	1.0	29.48
10	1	0.0	0.0	10.0	0.0	9.0	2.0	17.13
11	1	10.0	10.0	10.0	10.0	9.0	2.0	43.75
12	1	10.0	0.0	0.0	10.0	6.0	2.0	40.86
13	2	0.0	10.0	10.0	0.0	6.0	2.0	
14	2	0.0	0.0	0.0	0.0	9.0	2.0	
15	2	0.0	10.0	10.0	10.0	9.0	1.0	
16	2	0.0	0.0	10.0	0.0	6.0	1.0	
17	2	10.0	10.0	0.0	10.0	6.0	1.0	
18	2	10.0	0.0	0.0	0.0	6.0	1.0	
19	2	0.0	0.0	10.0	10.0	6.0	2.0	
20	2	10.0	0.0	10.0	0.0	9.0	2.0	

Example #3: Incorporating Design Region Constraints

It is not infrequent to find when designing an experiment that the feasible region in which design points may be placed is not cuboidal. In other words, having set low and high ranges for each of the factors, certain combinations of factor levels falling between the lows and the highs may not correspond to acceptable locations at which to perform an experiment. Statgraphics allows you to specify constraints on linear combinations of the factors that limit the locations at which experiments may be run.

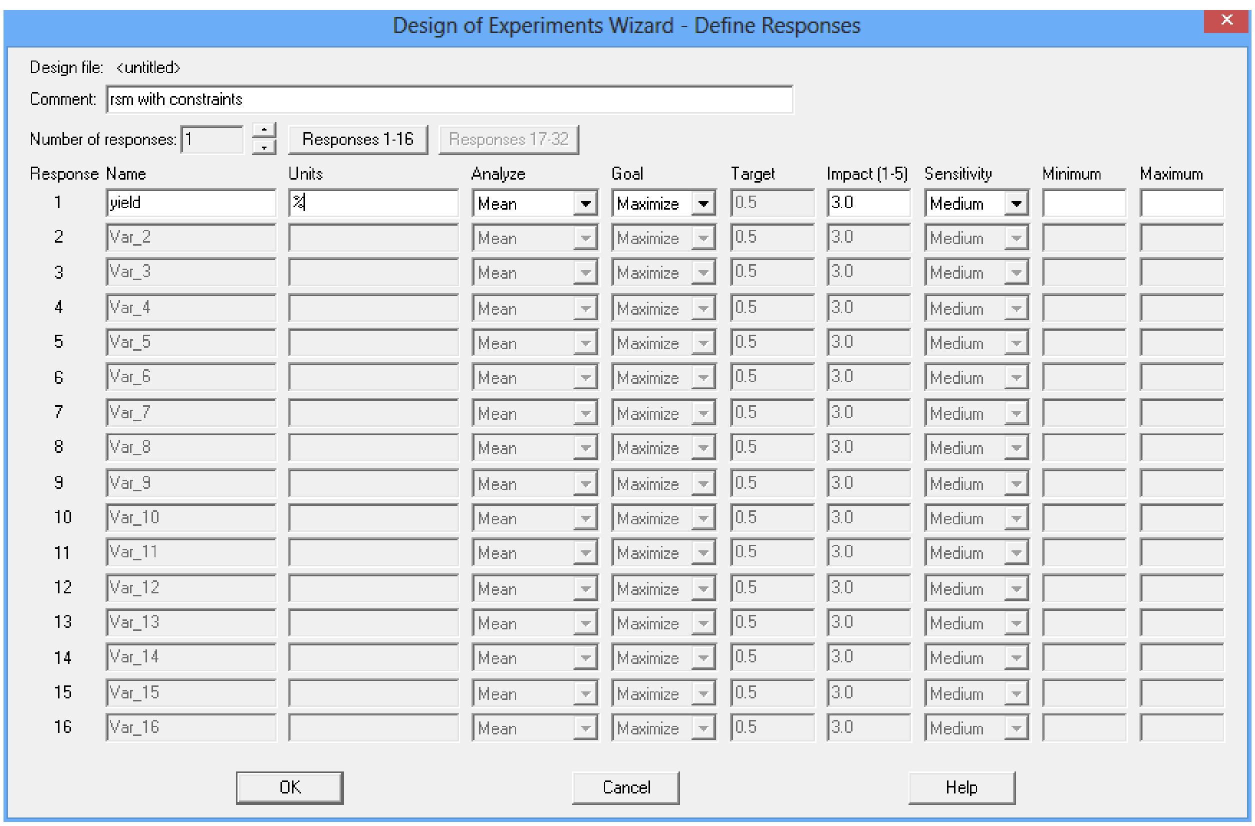
Goos and Jones (2011) describe an experiment intended to maximize the yield of a chemical process. The first factor (Time) was allowed to vary between 430 and 640 seconds. The second factor (Temperature) was allowed to vary between 500 and 550 degrees Kelvin. However, experiments could only be run in regions where two additional constraints were satisfied:

$$0.3 \text{ Time} + \text{Temperature} \geq 539.8$$

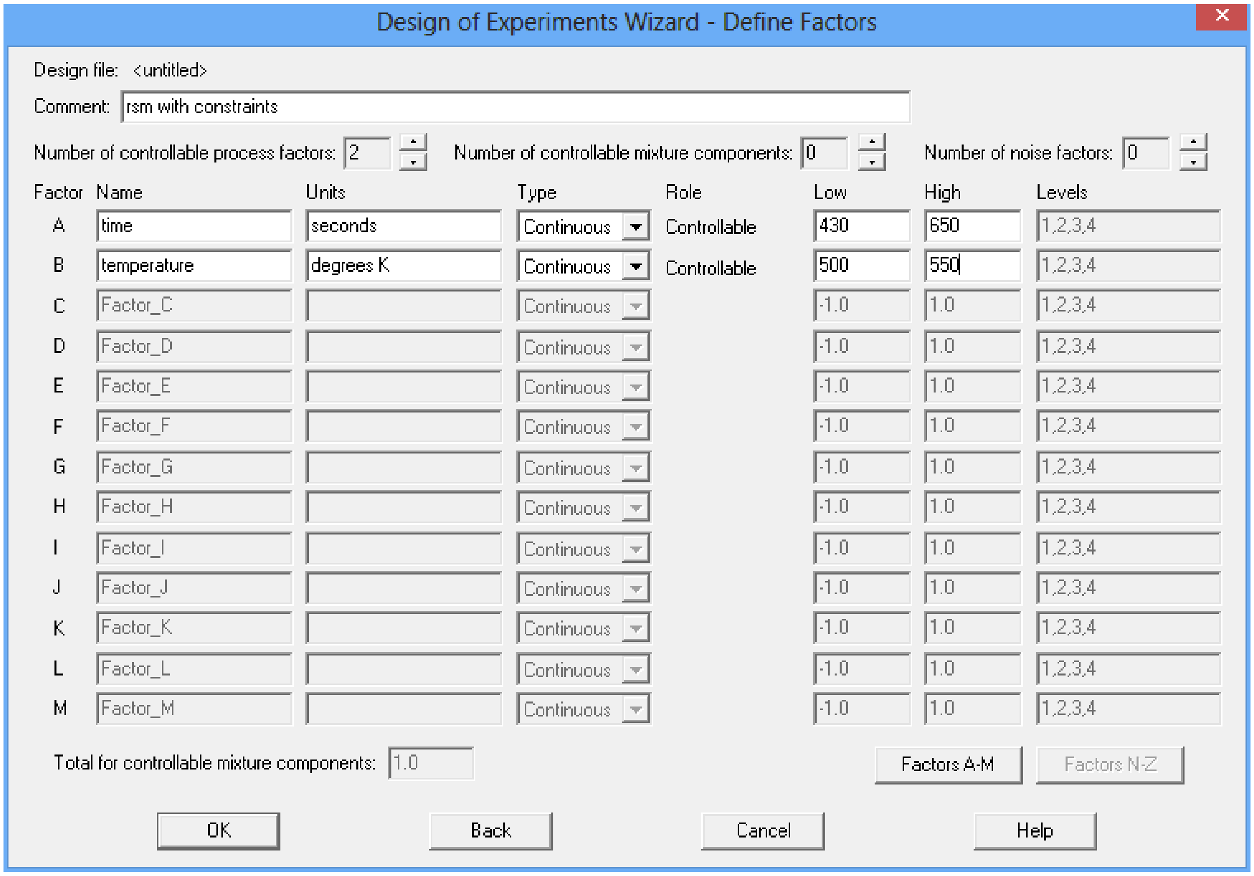
$$0.09 \text{ Time} + \text{Temperature} \leq 587.8$$

Outside that constrained region, it was thought that the process would behave much differently than within the region, so that any experimental points outside the region could potentially distort the estimated statistical model.

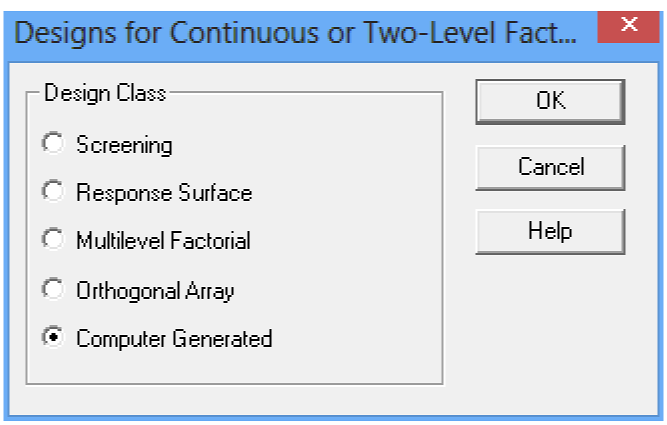
To reconstruct their experiment, start with an StatFolio and select *DOE – Experimental Design Wizard* from the main menu. In Step 1, enter the following information about the response variable:



In step 2, enter the following information about the experimental factors:

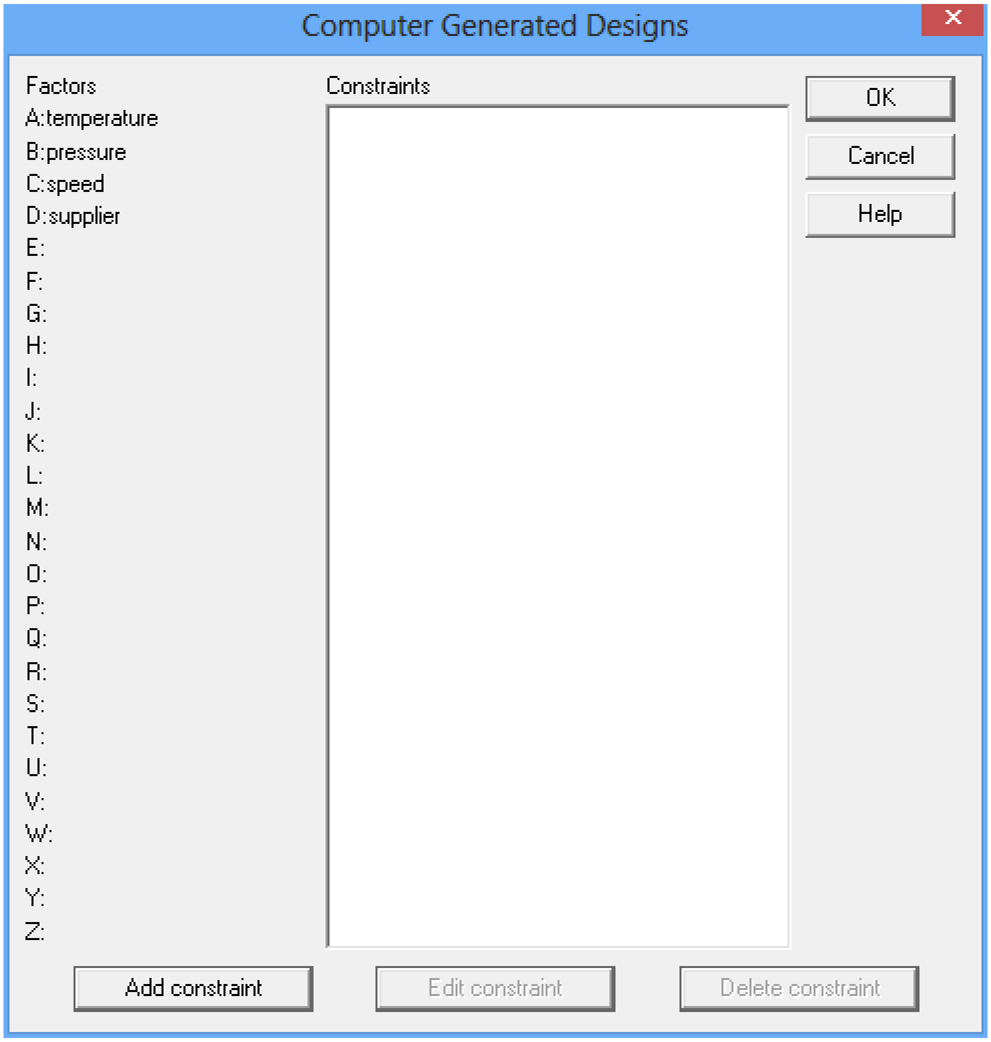


In Step 3, select a *Computer Generated* design on the first dialog box:



Press *OK* to return to the *Select Design* dialog box.

After *Computer Generated* is selected, the *Constraints* button at the bottom of the *Select Design* dialog box will be enabled. Pressing that button displays a dialog box in which up to 20 constraints involving the experimental factors may be entered:



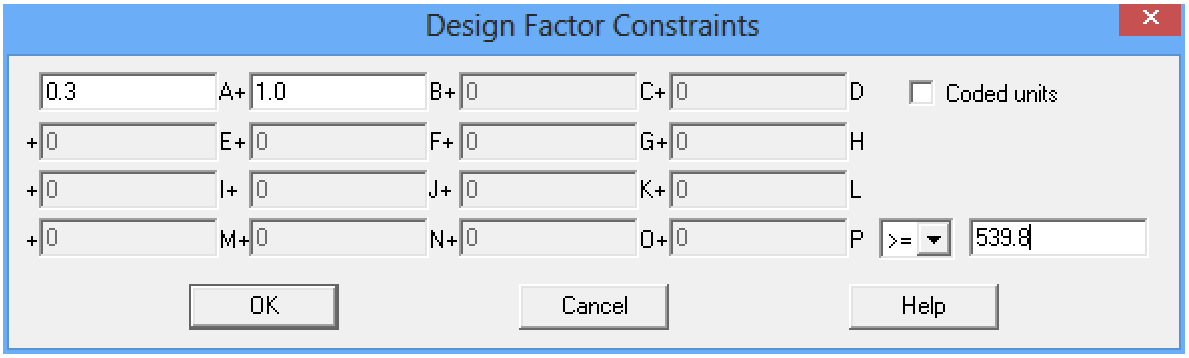
You may enter constraints of the form

$$C_1X_1 + C_2X_2 + C_3X_3 \leq d$$

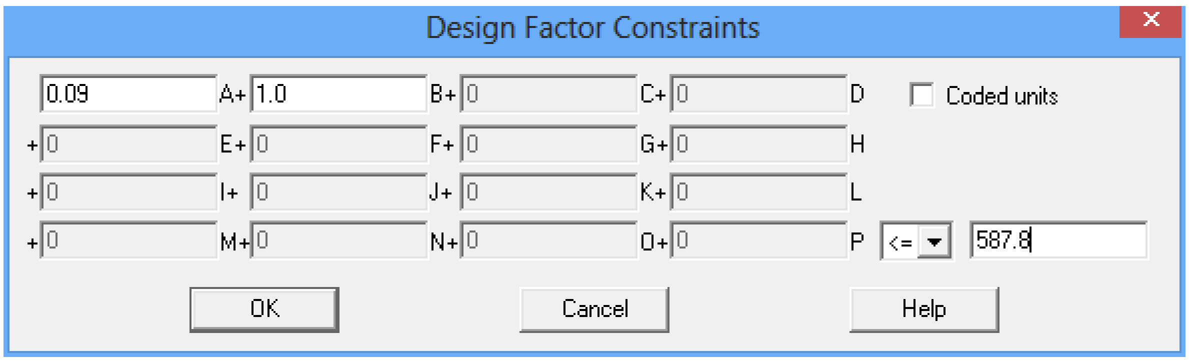
or

$$C_1X_1 + C_2X_2 + C_3X_3 \geq d$$

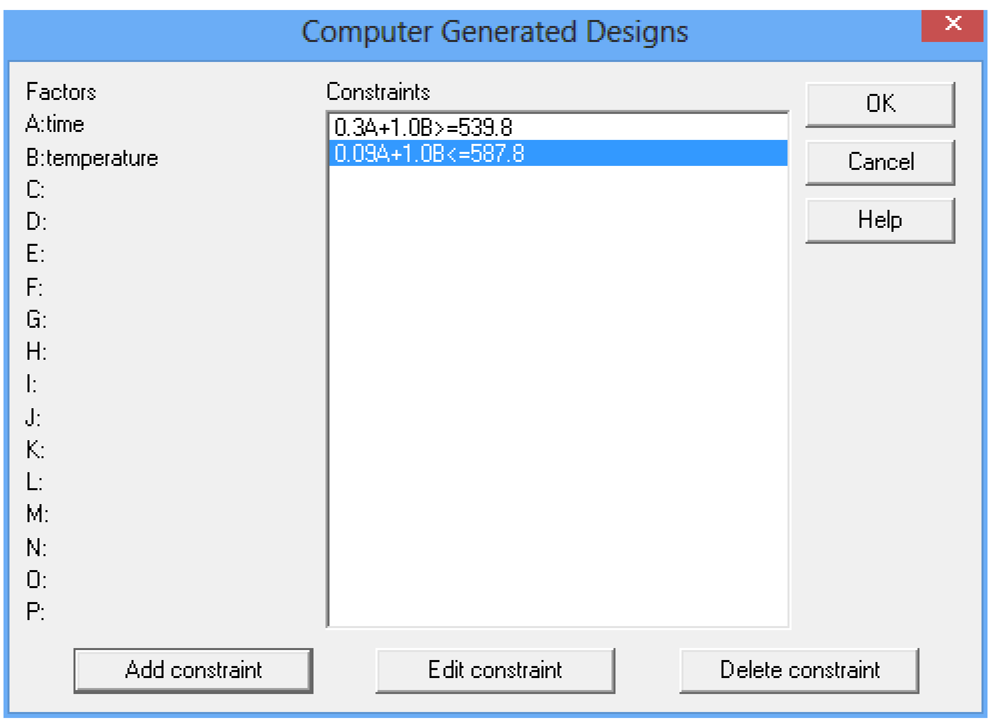
by pressing the *Add constraint* button and entering the values of C_1, C_2 and C_3. For the current example, the first constraint should be added by entering the following:



Press *OK* and then *Add constraint* again to add the second constraint:

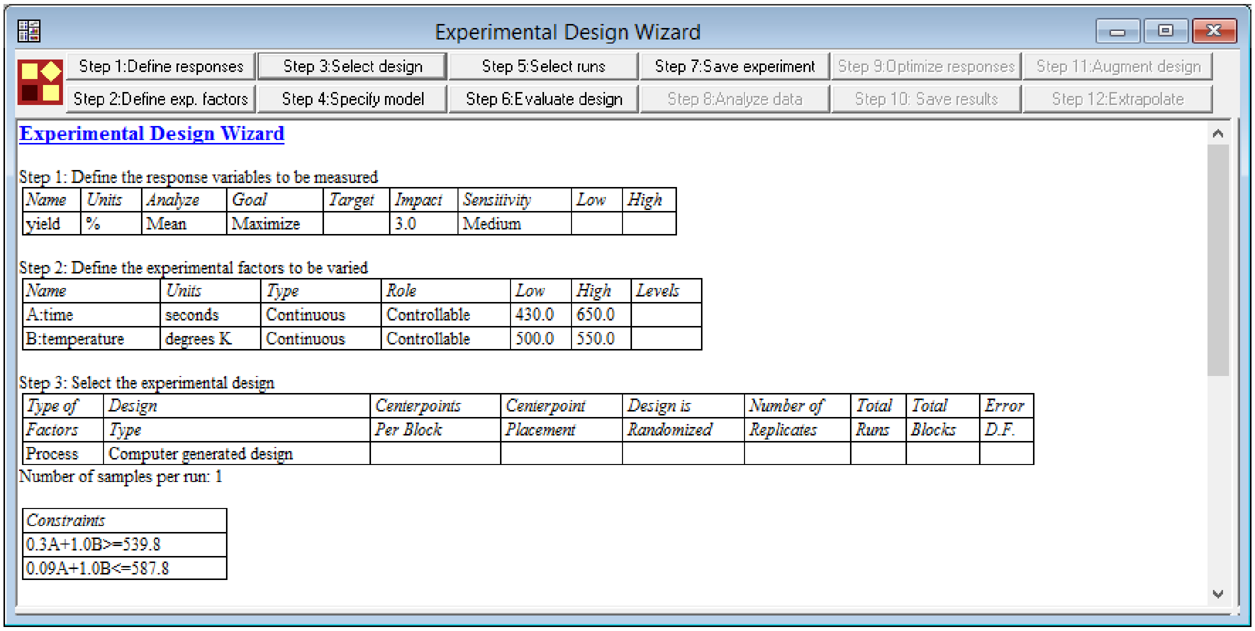


Press *OK* again and review the entered constraints:

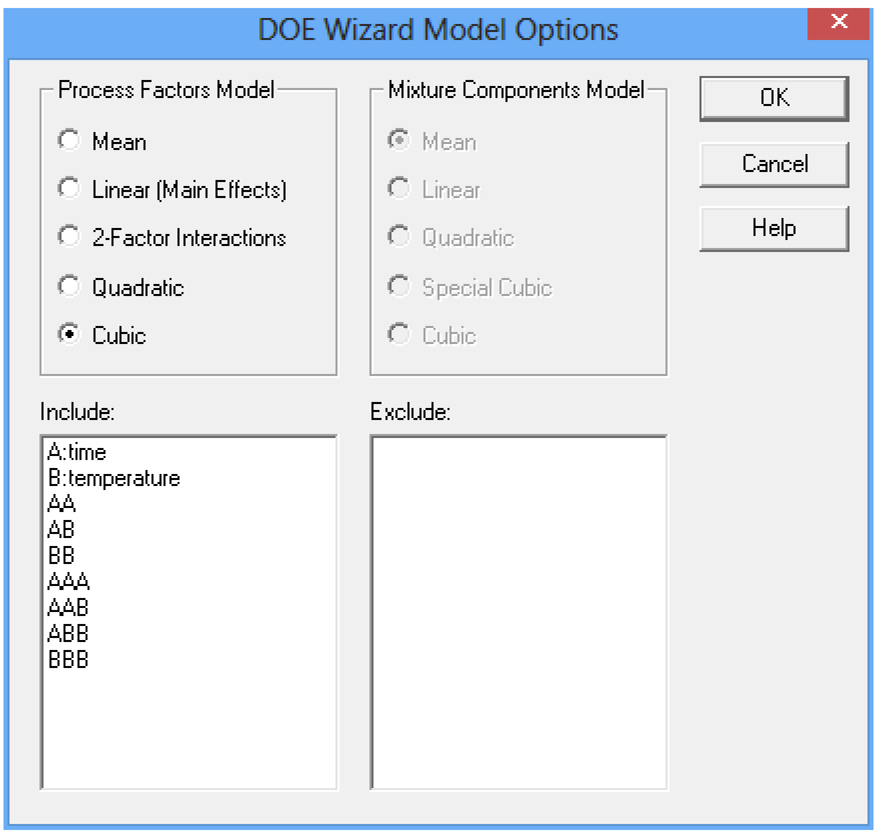


Note: if your design contains both process factors and mixture components, each constraint may involve only one type of factor.

Press *OK* twice to return to the main DOE Wizard window, which will display the entered constraints:

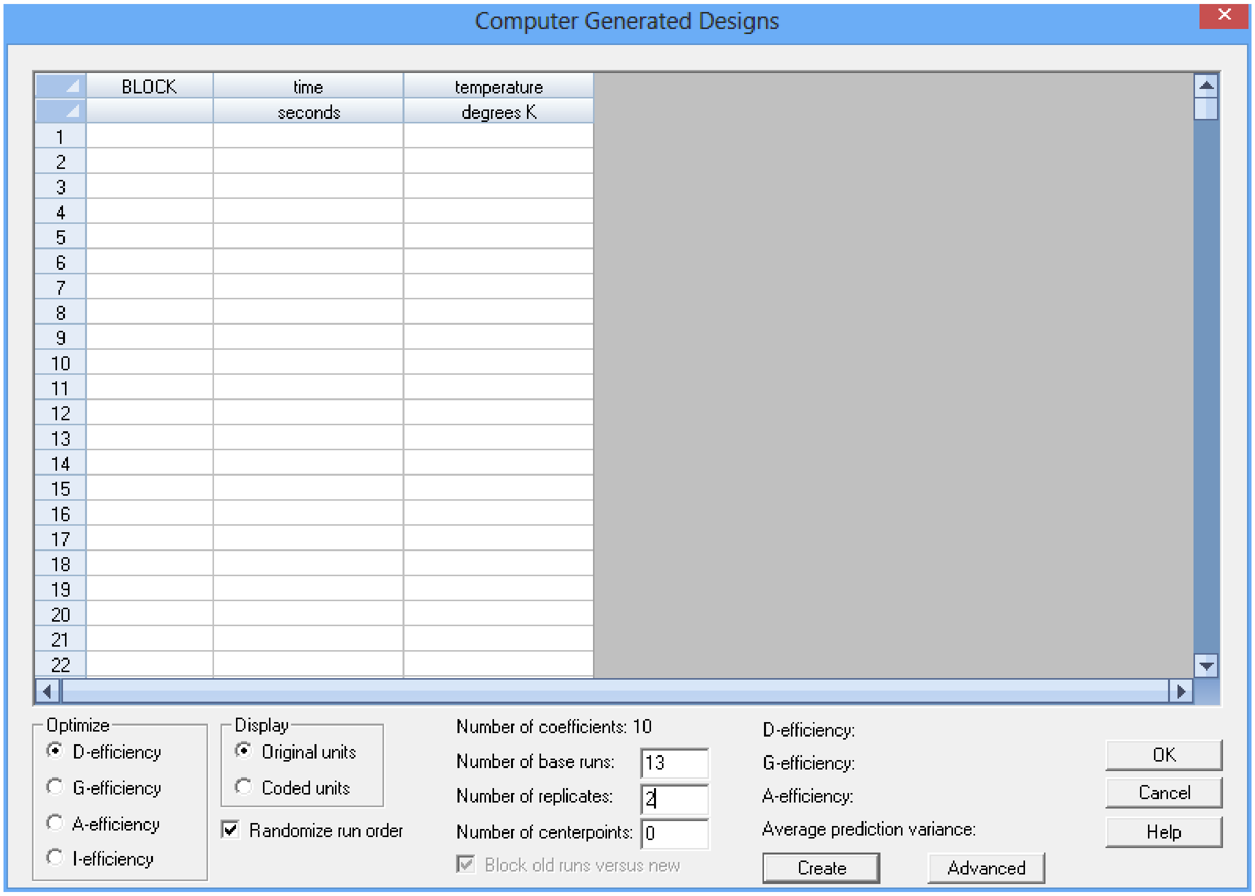


In Step 4, select a cubic model:

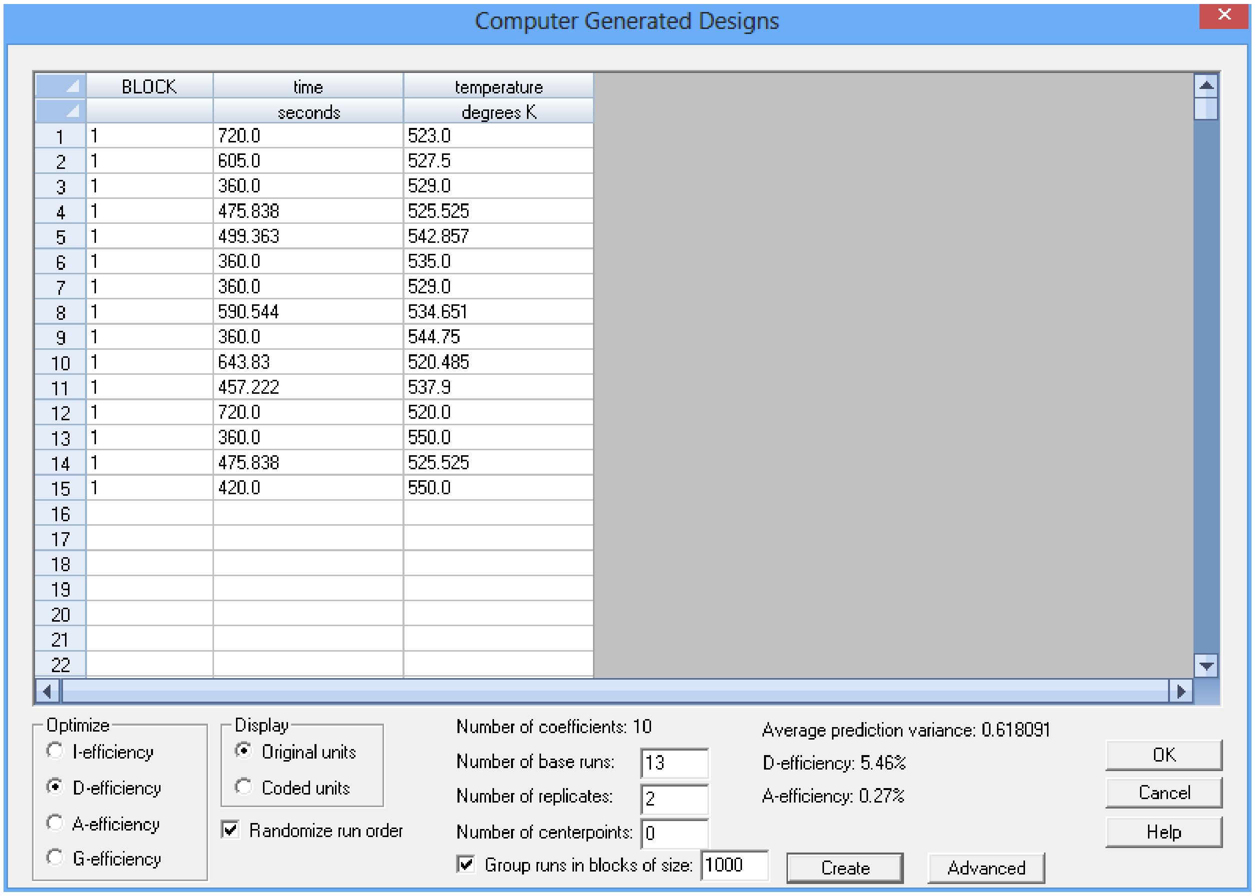


This model involves 2 main effects, 1 two-factor interaction, 2 quadratic effects, 2 cubic terms of the form x_j^3, and 2 mixed cubic terms of the form $x_i x_j^2$.

Now press the button labeled *Step 5: Select runs*. Complete the dialog box as shown below:



The settings request a D-optimal design with a total of 15 runs, 2 of which are to be replicates of 2 of the 13 base runs. Press the *Create* button to generate the design:



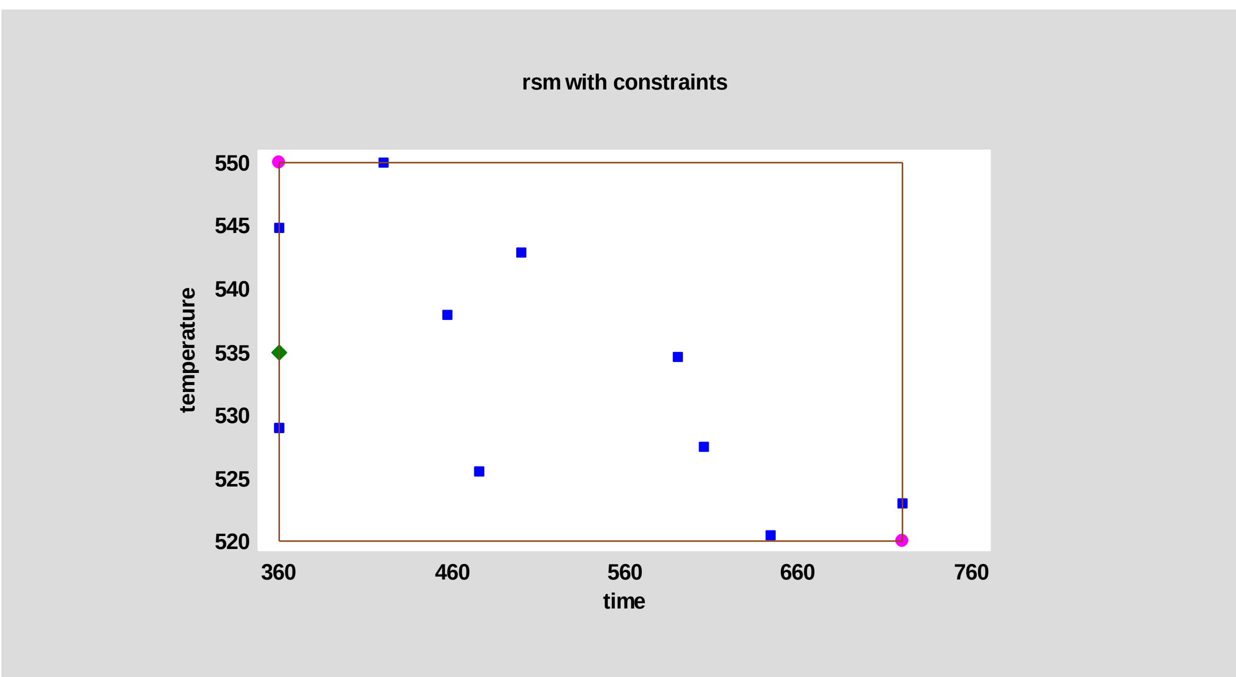
Push *OK* to save the design.

Once the design has been saved to the datasheet, you may round off the levels if desired:

C:\Data\Version17 Tests\doewiz compgen constraints.sgx

	BLOCK	time	temperature	yield
		seconds	degrees K	%
1	1	720.0	523.0	
2	1	605.0	527.5	
3	1	360.0	529.0	
4	1	476	526	
5	1	499	543	
6	1	360.0	535.0	
7	1	360.0	529.0	
8	1	591	535	
9	1	360.0	545	
10	1	644	520	
11	1	457	538	
12	1	720.0	520.0	
13	1	360.0	550.0	
14	1	476	526	
15	1	420.0	550.0	
16				

To display the design points, select *Step 6: Evaluate design* from the DOE Wizard toolbar. The *Design Points* graph displays the data as shown below:



Original vertices are shown as round point symbols.

Example #4: Designs with Mixture Variables

Many experiments involve determining tradeoffs between components which must sum to a fixed value. A typical example is a formulation which consists of 3 components. While the experimenter may be able to tradeoff one component for another, the sum of the percentages of each component must equal 100%. Such “mixture” experiments may or may not include additional process factors.

A typical example of a mixture problem, discussed by Myers and Montgomery (2002), involves the formulation of a rocket propellant. The propellant is a mixture of three components: a fuel, an oxidizer, and a binder. The researcher wished to find a combination of these three components which achieved a satisfactory burn rate. Since an inert component made up 10% of the propellant, the factors were constrained by:

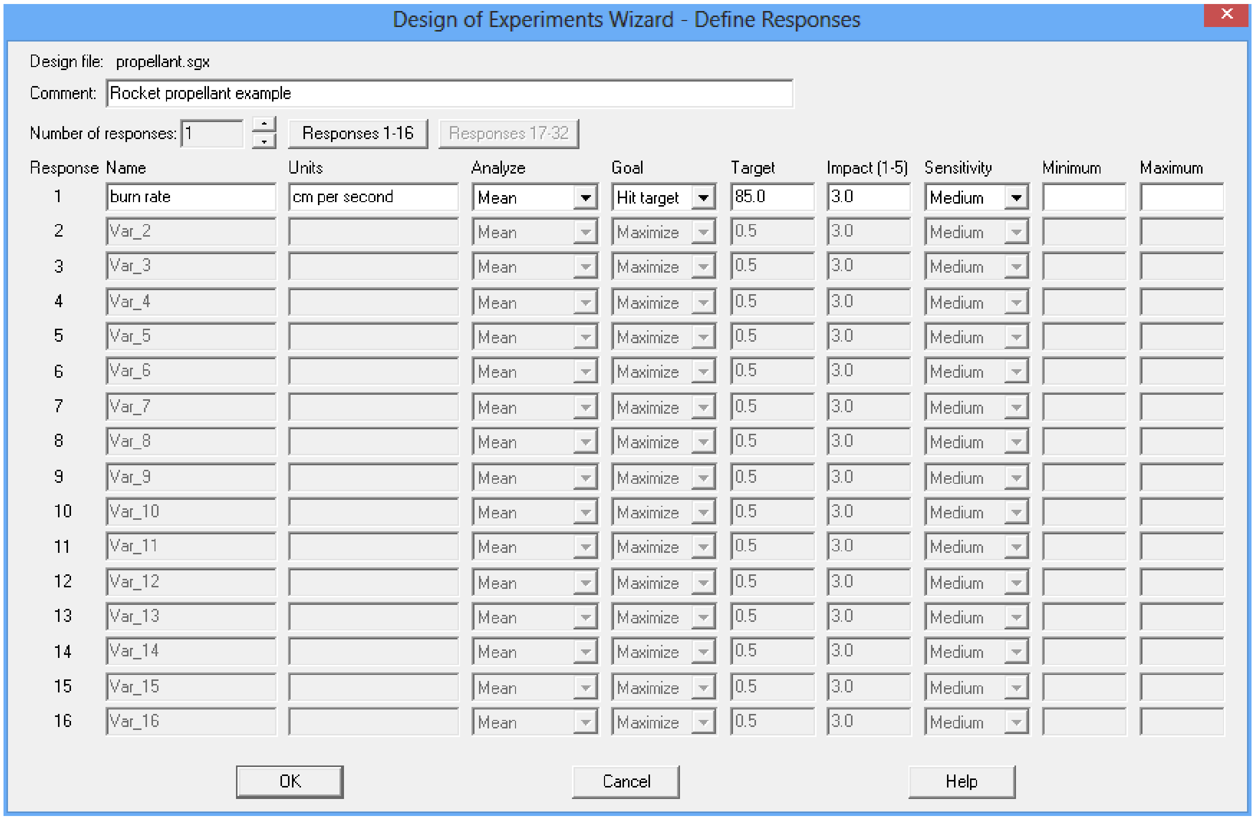
$$fuel + oxidizer + binder = 90\%.$$

In addition, there were lower bounds for each of the three components:

$$30\% \le fuel$$
$$20\% \le oxidizer$$
$$20\% \le binder$$

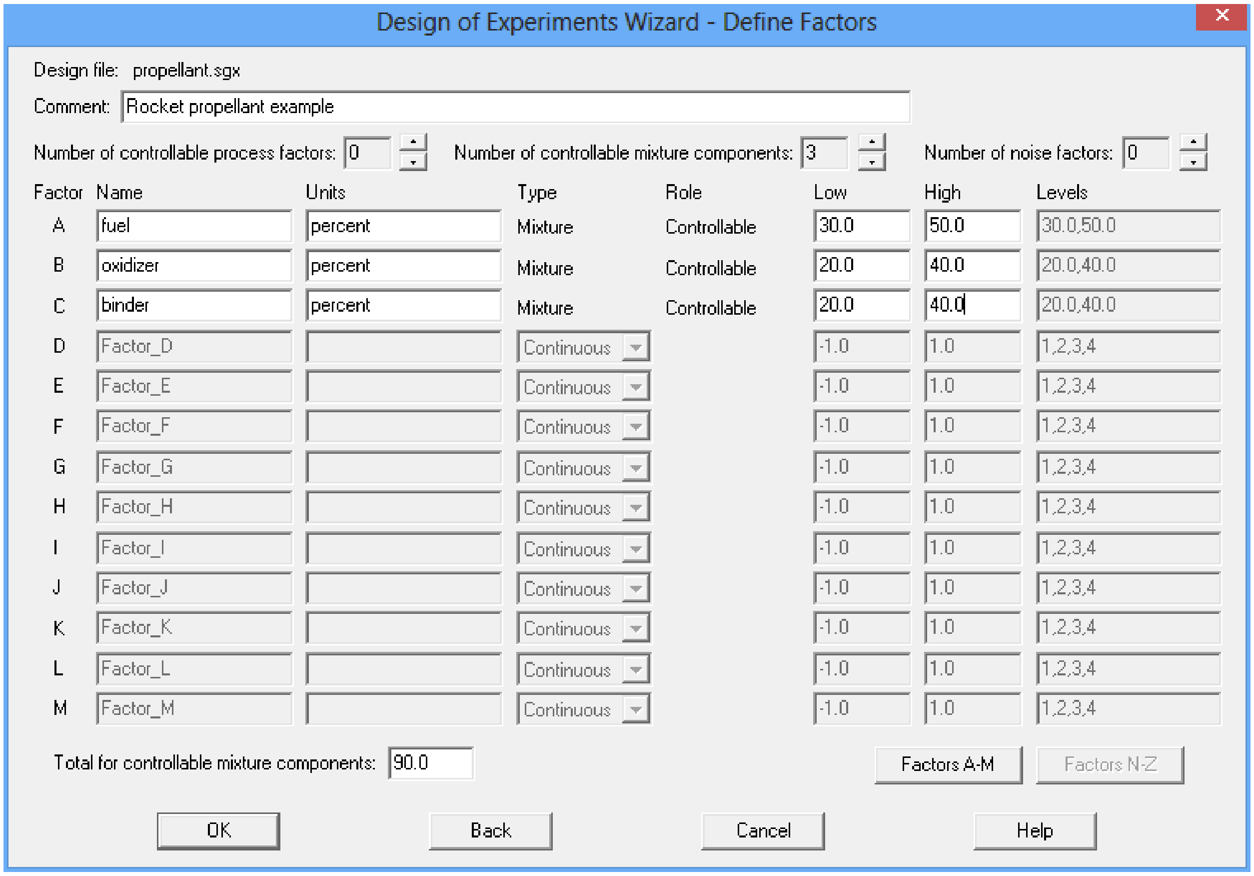
Given the constraints, there is a remaining 20% of the mixture that can be any combination of *fuel*, *oxidizer*, and/or *binder*.

To reconstruct their experiment, start with an StatFolio and select *DOE – Experimental Design Wizard* from the main menu. In Step 1, enter the following information about the response variable:



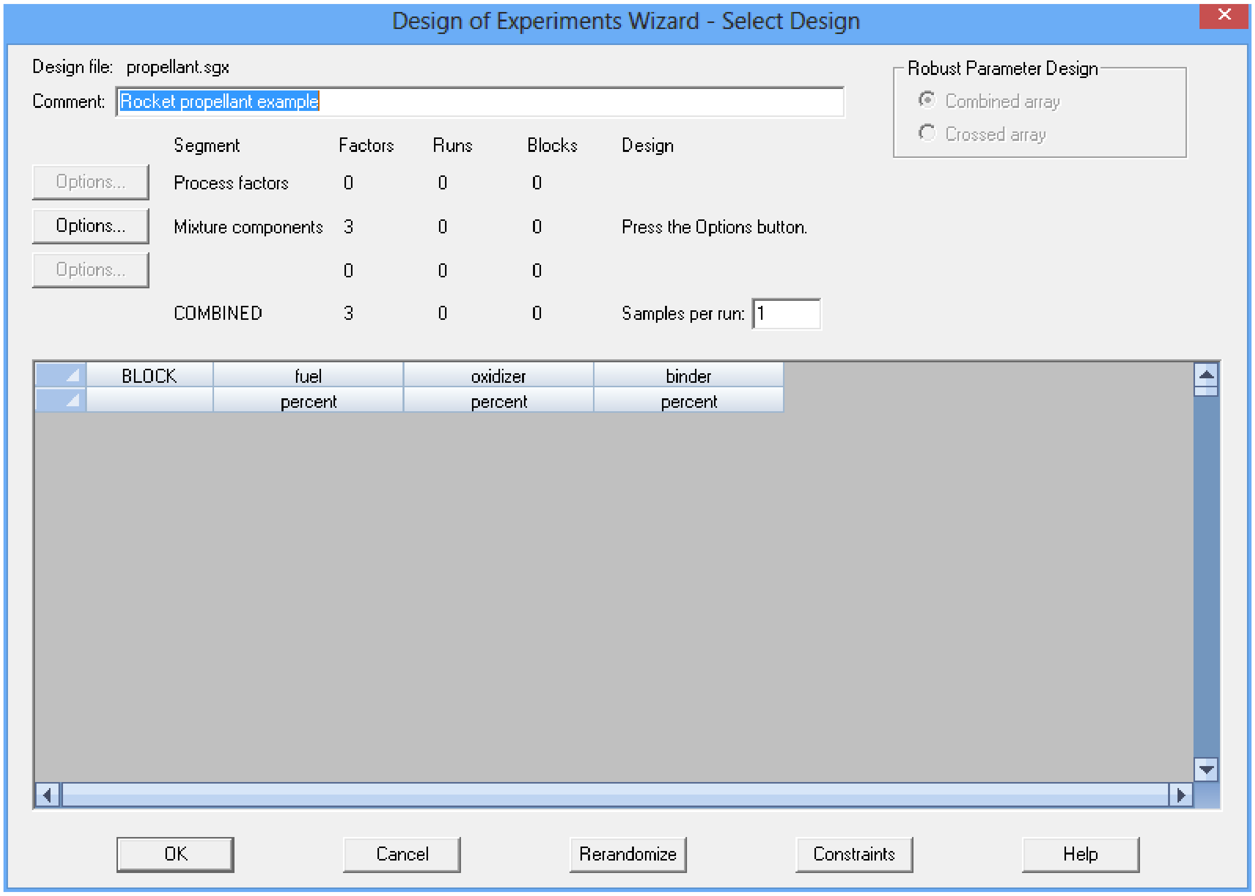
The goal of the experiment was to achieve a burn rate close to 85 cm per second.

In step 2, enter the following information about the experimental factors:

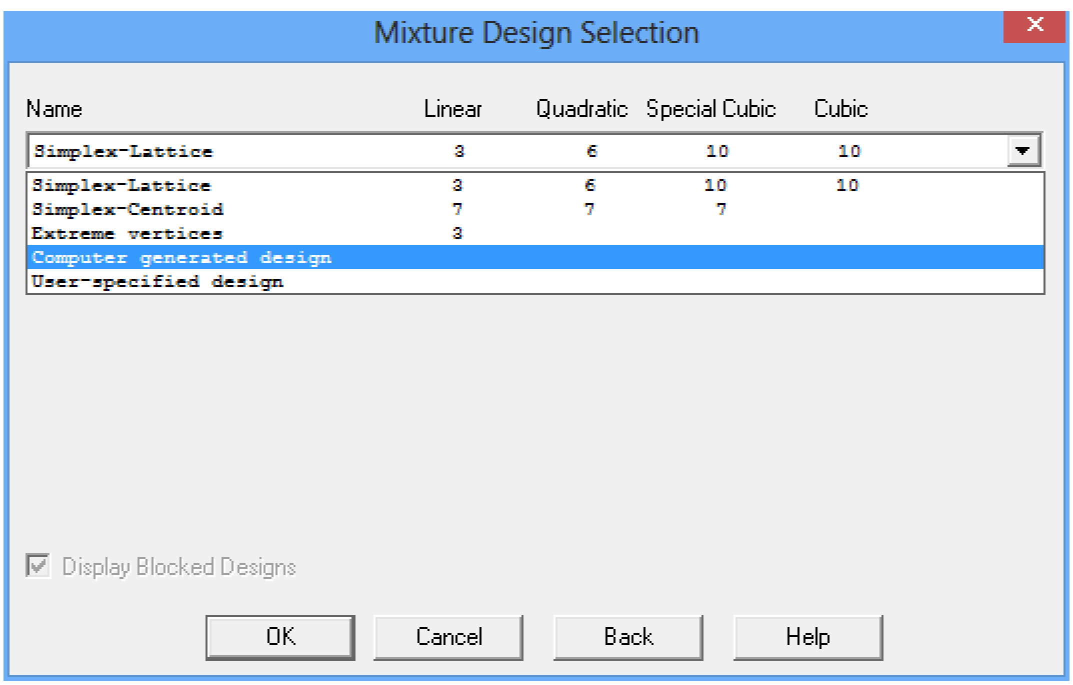


There are 3 mixture components which may sum to 90%.

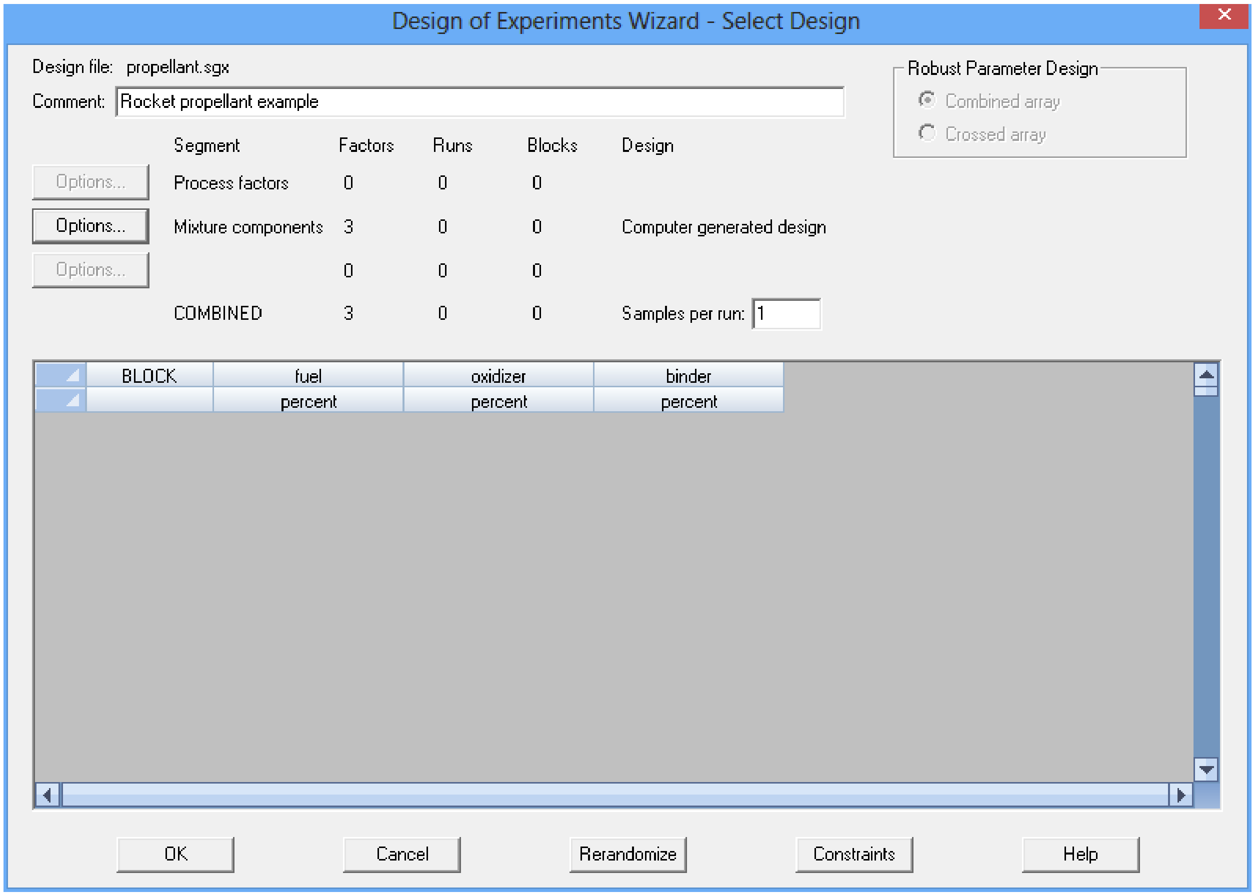
Pushing the button labeled *Step 3: Select design* displays the dialog box shown below:



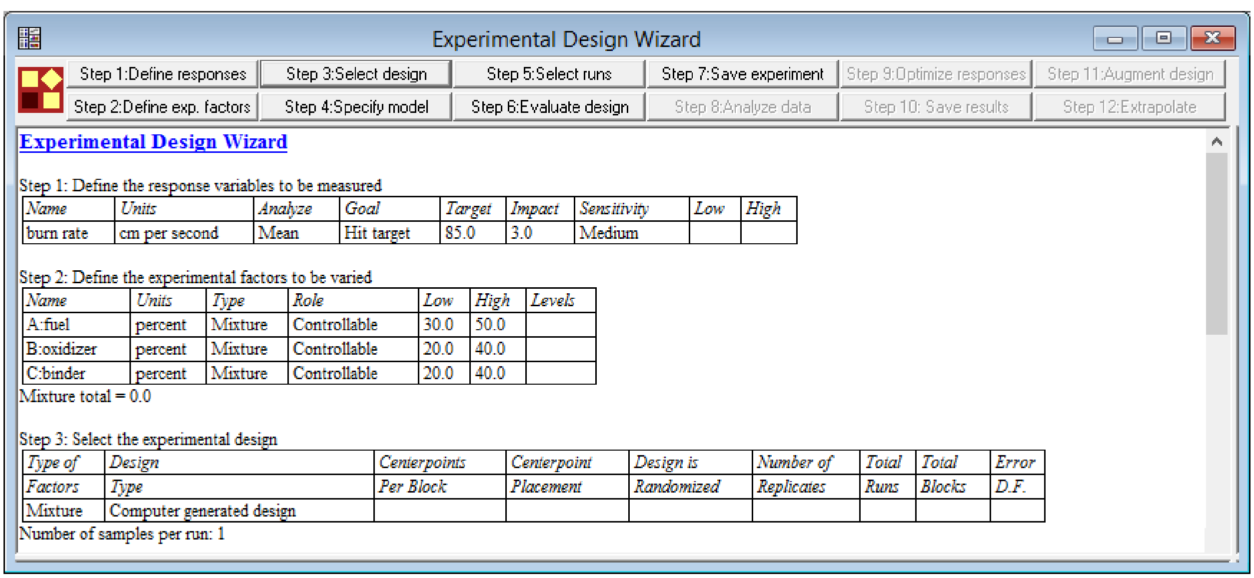
Press the *Options* button for the mixture components to display a list of designs appropriate for 3 mixture components:



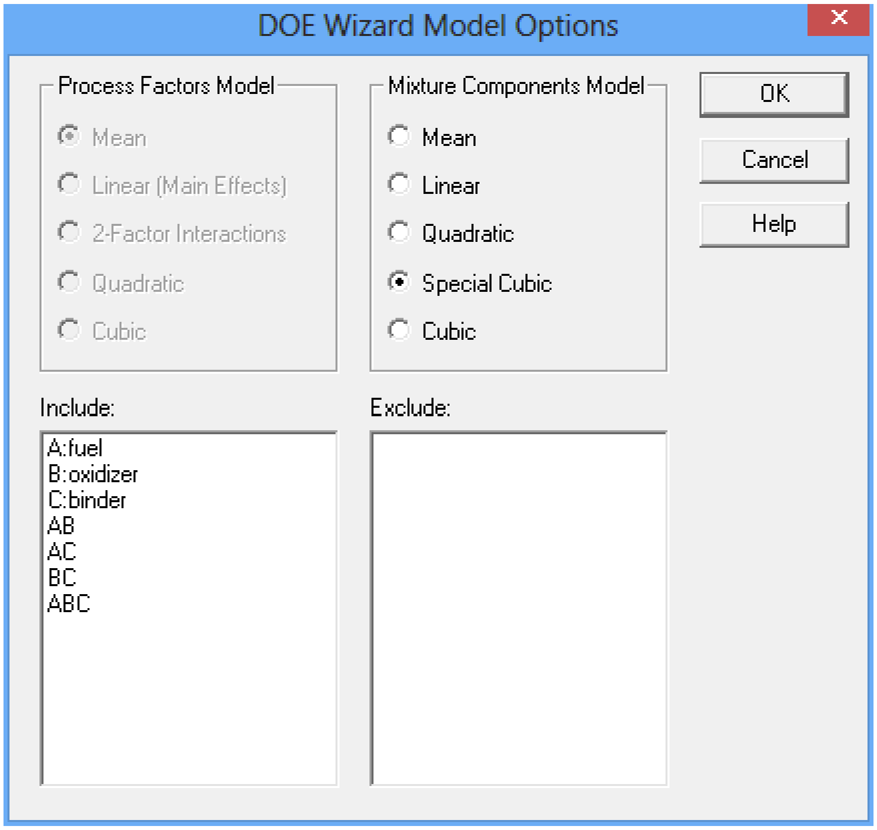
Select *Computer generated design* and press *OK* to return to the previous dialog box:



Press *OK* again to return to the main DOE Wizard window:

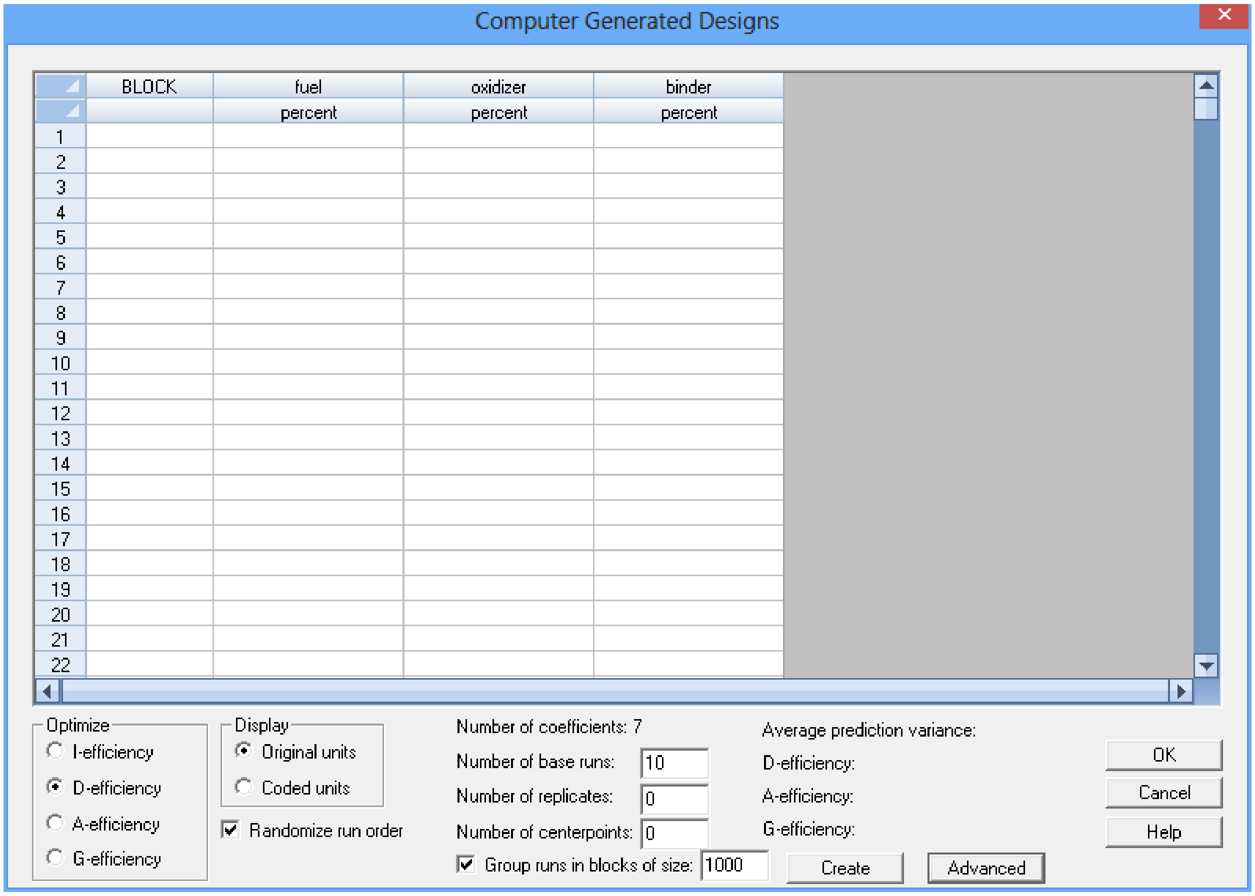


Press the *Step 4* button to select the desired model:



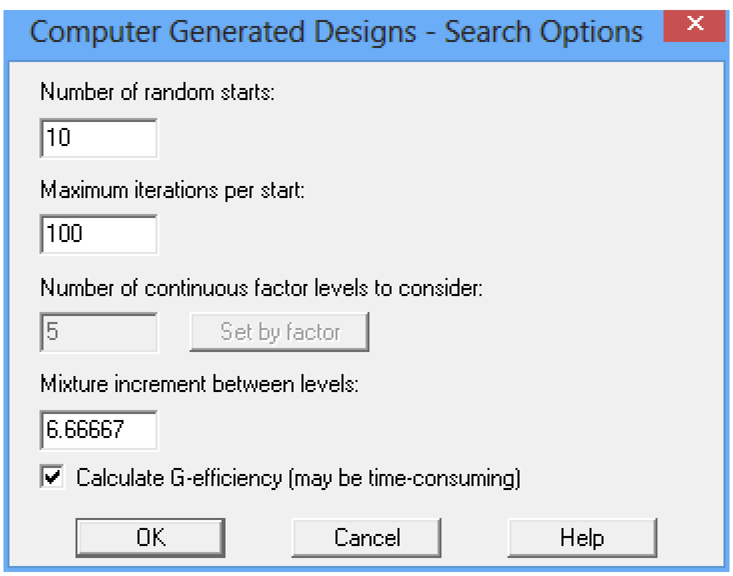
In the example, the researchers choose the popular *Special Cubic* model which has 7 coefficients that must be estimated.

Now press the button labeled *Step 5: Select runs*. Complete the dialog box as shown below:



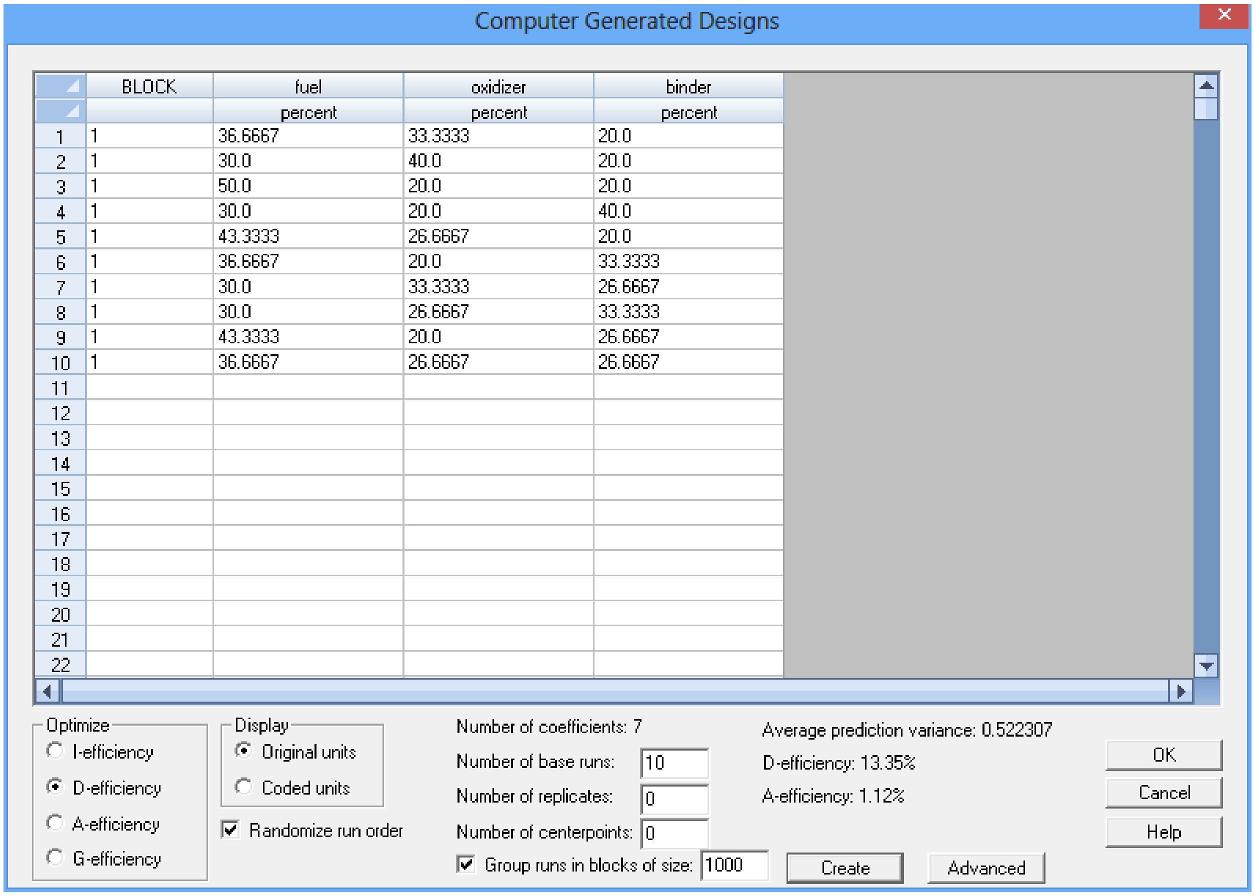
The settings request a D-optimal design with a total of 10 runs.

Before creating the design, press the *Advanced* button and set the *Mixture increment* field to 6.6667:

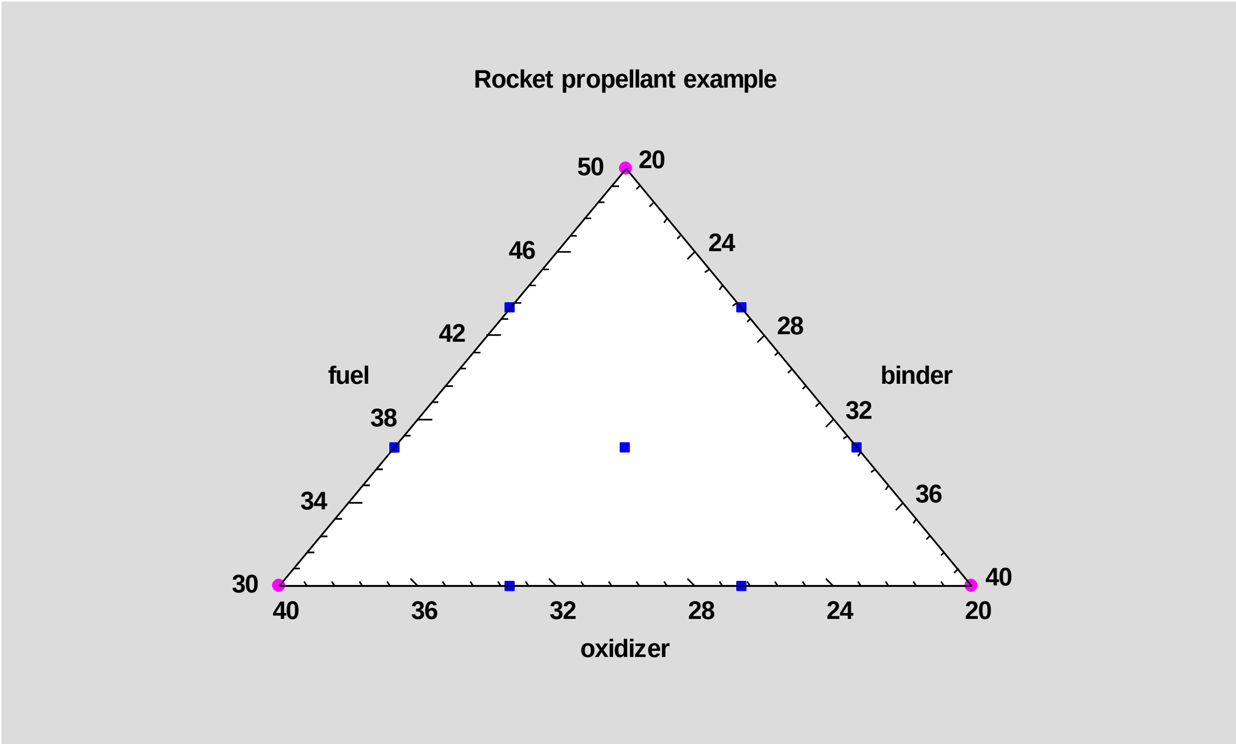


This tells the program to consider experimental runs spaced at increments of 6 2/3% between the low and high levels of each component.

Press the *Create* button to generate the design:



To display the design points, select *Step 6: Evaluate design* from the DOE Wizard toolbar. The *Design Points* graph displays the data as shown below:



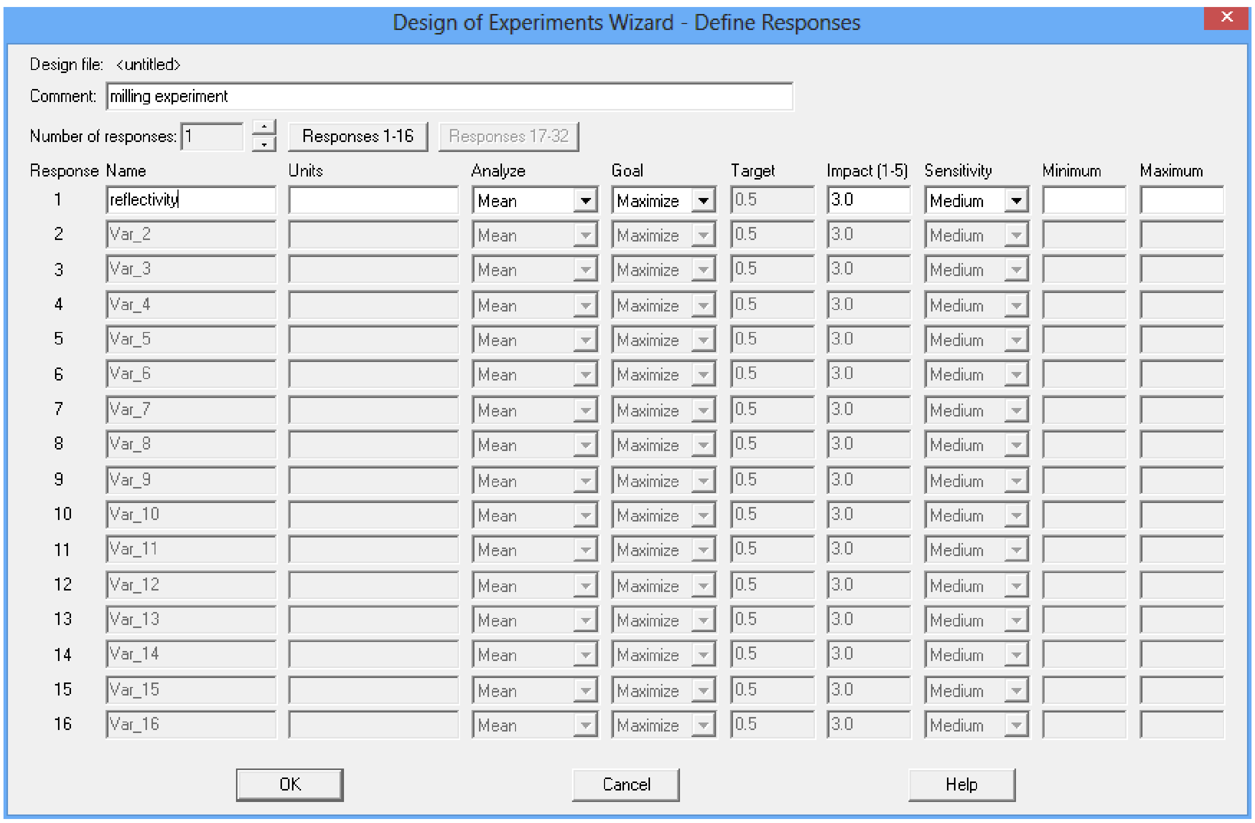
The selected runs include the 3 vertices, 2 points on each face, and the centerpoint.

Example #5: Designs with Both Process and Mixture Variables

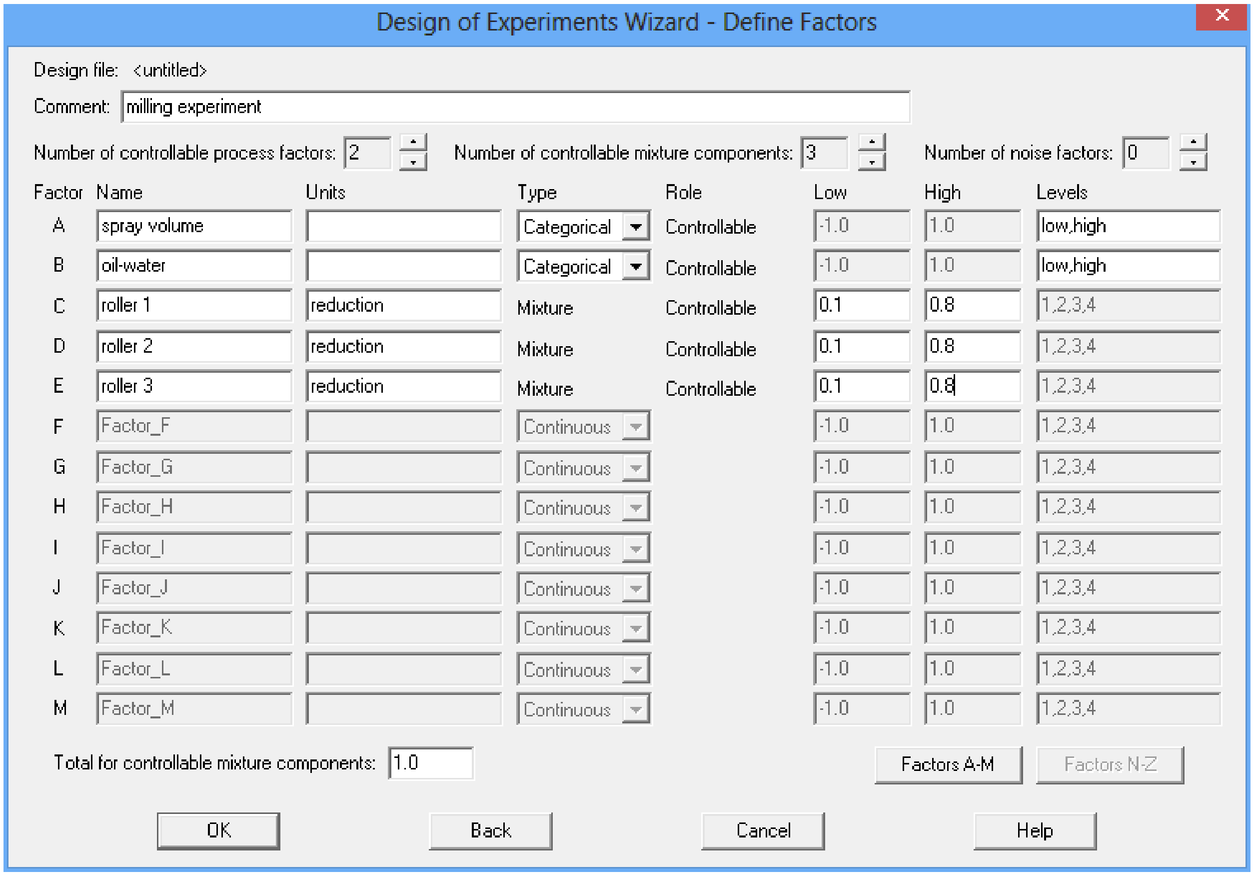
Many experiments involve determining tradeoffs between components which must sum to a fixed value. A typical example is a formulation which consists of 3 components. While the experimenter may be able to tradeoff one component for another, the sum of the percentages of each component must equal 100%. Such "mixture" experiments may or may not include additional process factors.

Goos and Jones (2011) describe an experiment intended to study the effects on *reflectivity* in a rolling mill producing sheet aluminum. The experiment involved two process factors (*spray volume* and *oil-water ratio*) and three mixture components (percentage of thickness reduction assigned to each of 3 *rollers*).

To reconstruct their experiment, start with a StatFolio and select *DOE – Experimental Design Wizard* from the main menu. In Step 1, enter the following information about the response variable:

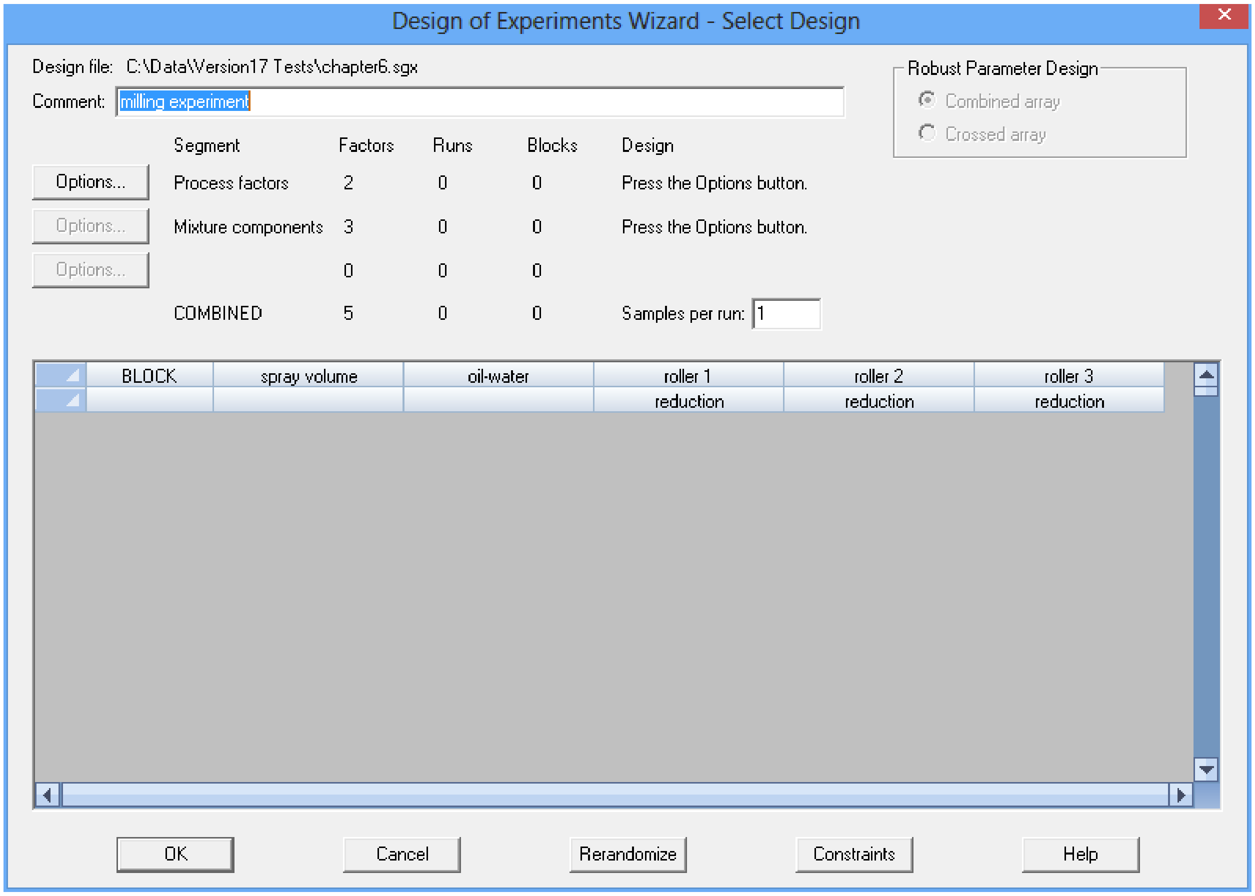


In step 2, enter the following information about the experimental factors:

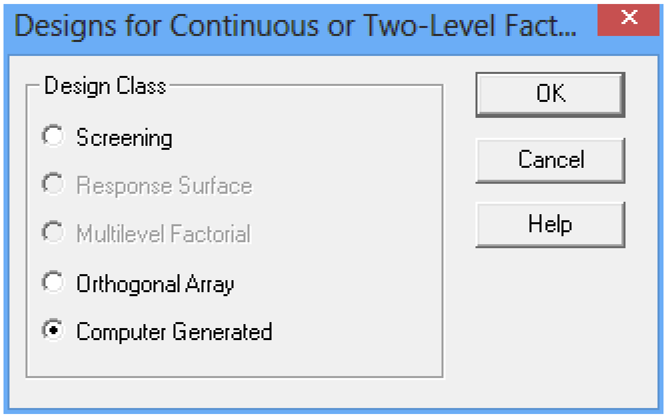


The two process variables are treated as categorical factors with 2 levels each. The 3 mixture components, which represent the proportion of thickness reduction performed by each roller, are each constrained to range between 0.1 and 0.8, summing to 1.

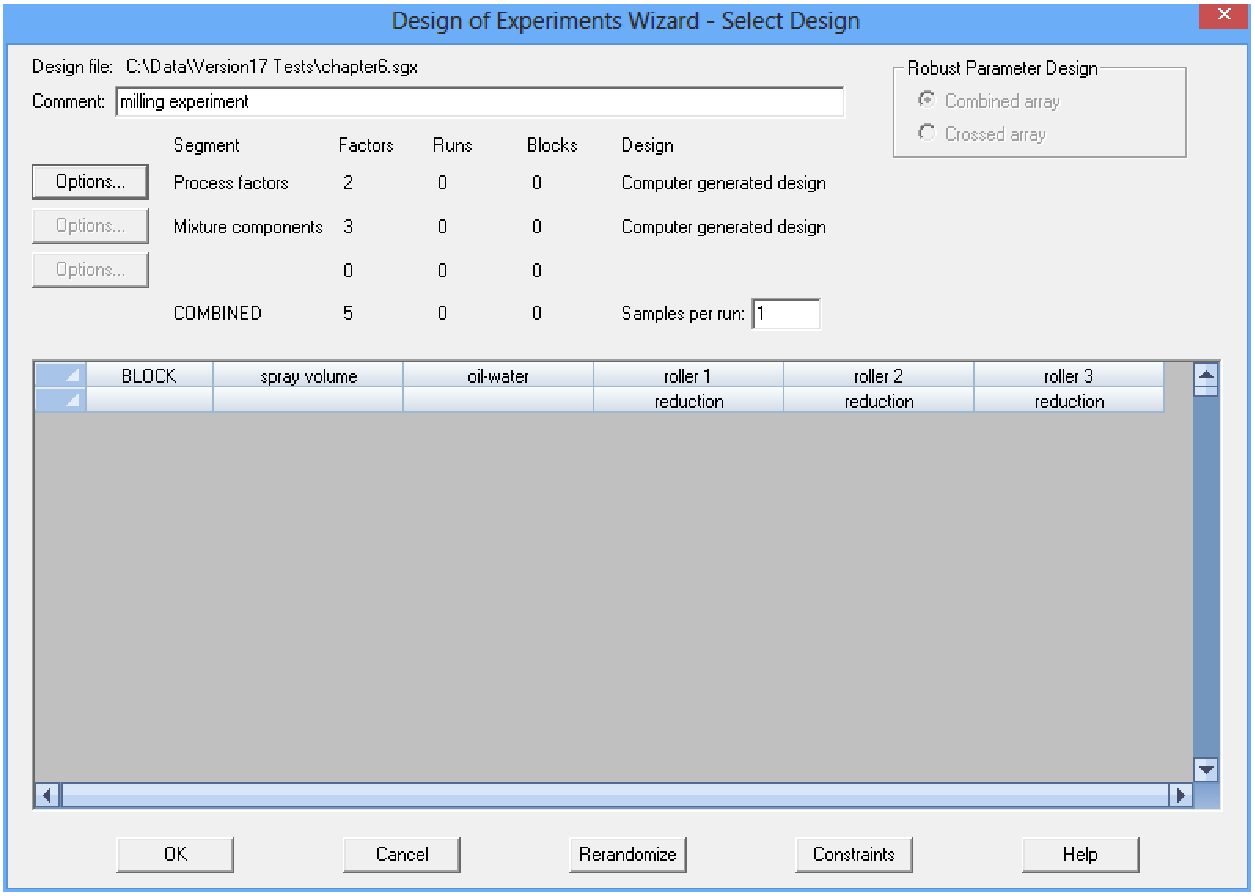
Pushing the button labeled *Step 3: Select design* displays the dialog box shown below:



Press the *Options* button for the process factors to display the following choices:

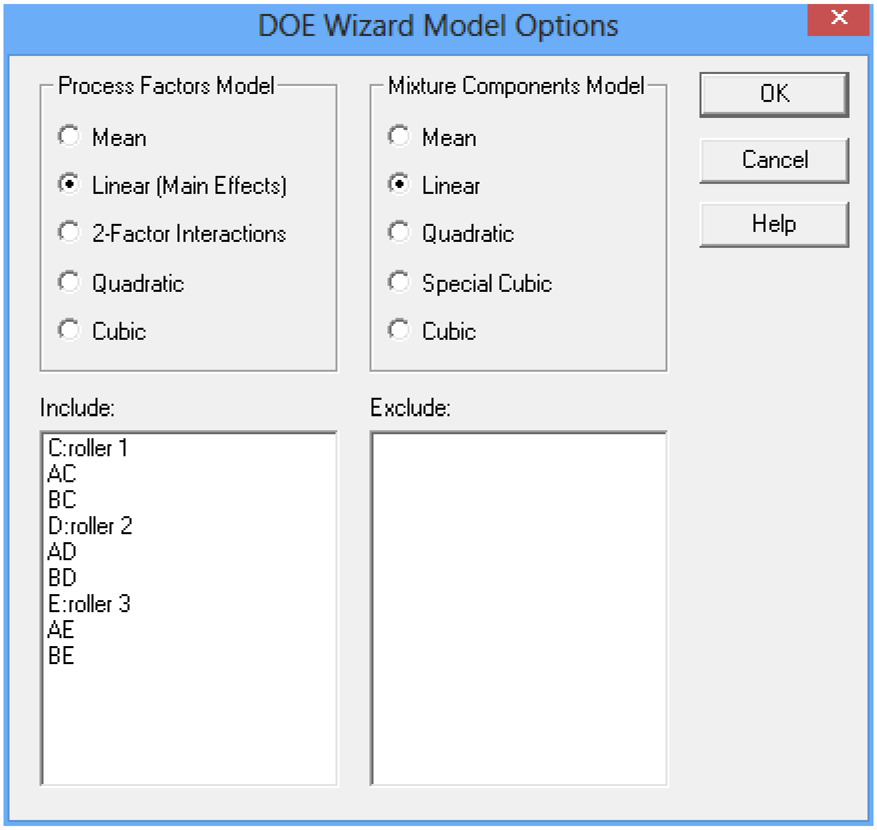


Select *Computer Generated* and press *OK* to return to the *Select Design* dialog box:



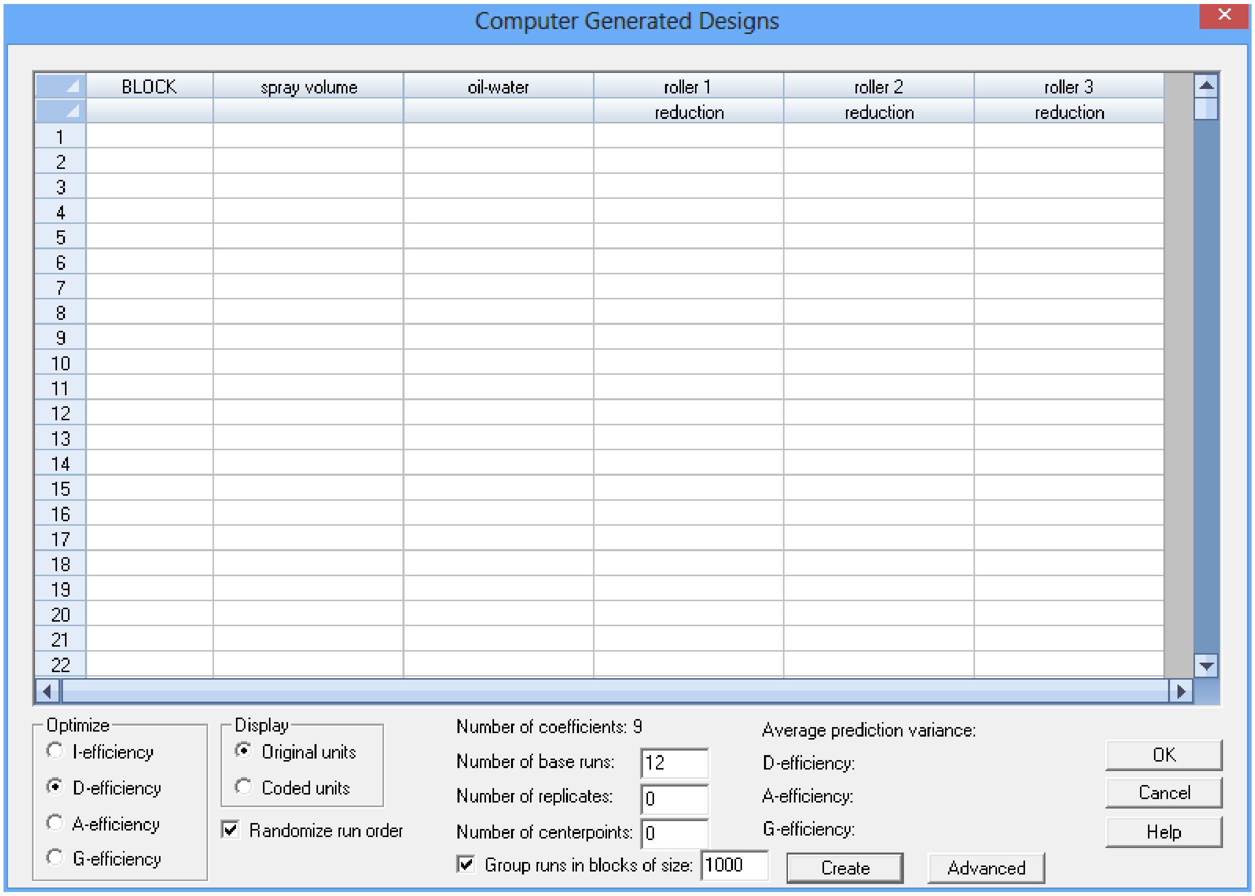
Notice that the *Design* for the mixture components has also been set to *Computer Generated Design*, since a single design will be generated for all 5 factors.

Next, push the button labeled *Step 4: Specify model* to display the dialog box shown below:



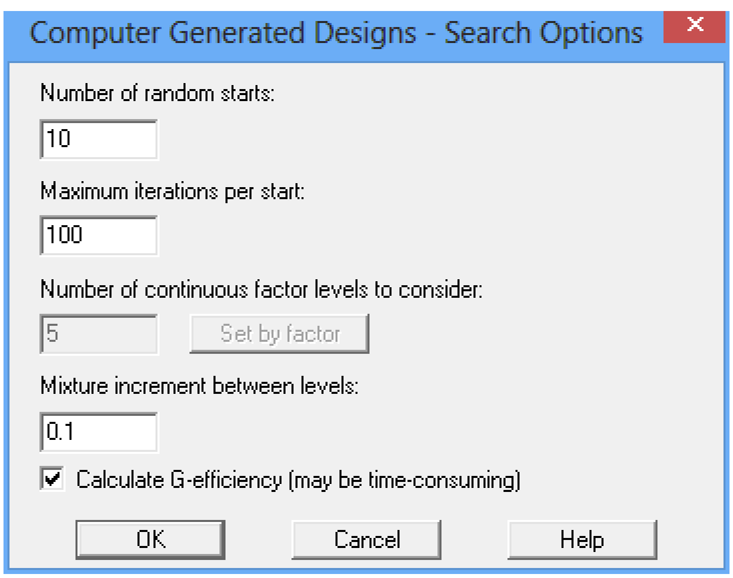
Selecting a linear model for both the processd factors and the mixture components results in a model contraining 9 coefficients that must be estimated.

After saving the choice of models, press the button labeled Step 5: Select runs to display the design creation dialog box:

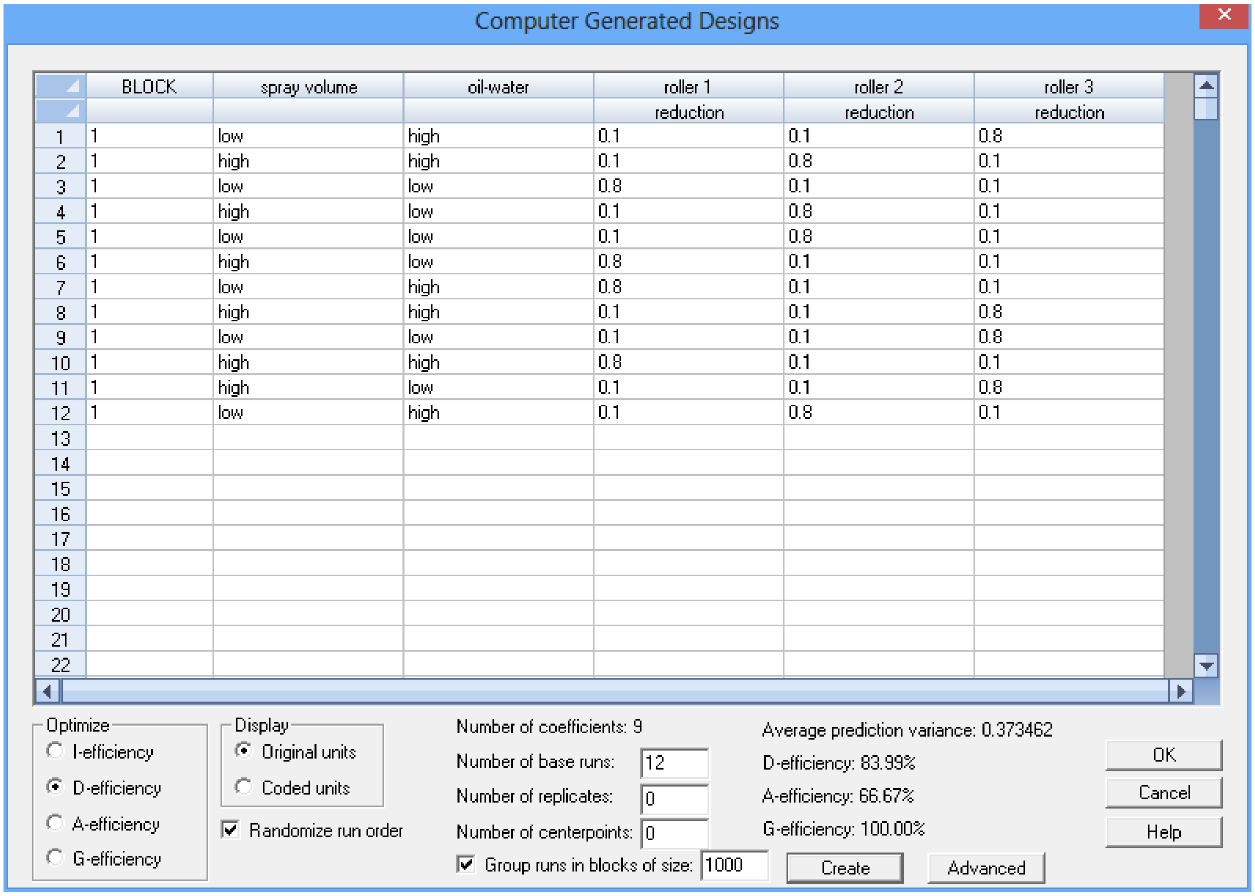


Be sure that "D-efficiency" is checked and that the requested number of base runs is 12.

Press the Advanced button and set the *Mixture increment* field to 0.1:



Press *OK* and then *Create* to create the design as shown below:



	BLOCK	spray volume	oil-water	roller 1 reduction	roller 2 reduction	roller 3 reduction
1	1	low	high	0.1	0.1	0.8
2	1	high	high	0.1	0.8	0.1
3	1	low	low	0.8	0.1	0.1
4	1	high	low	0.1	0.8	0.1
5	1	low	low	0.1	0.8	0.1
6	1	high	low	0.8	0.1	0.1
7	1	low	high	0.8	0.1	0.1
8	1	high	high	0.1	0.1	0.8
9	1	low	low	0.1	0.1	0.8
10	1	high	high	0.8	0.1	0.1
11	1	high	low	0.1	0.1	0.8
12	1	low	high	0.1	0.8	0.1
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						

To display the design points, select *Step 6: Evaluate design* from the DOE Wizard toolbar. The *Design Points* graph displays the experimental points as shown below