



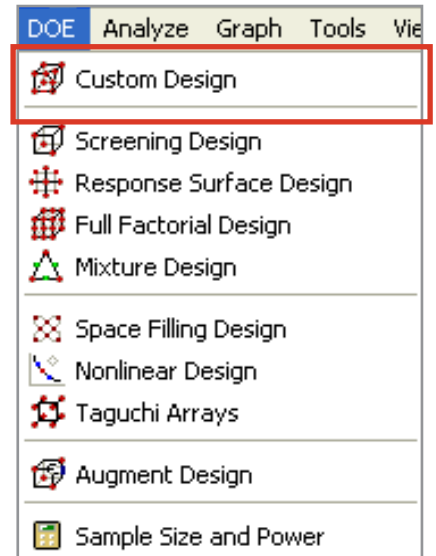
## JMP DESIGN OF EXPERIMENTS (DOE)



Fit the design to the problem,  
not the problem to the design

# The JMP DOE Difference: Tailor-Made Designs

**Don't be forced to fit your  
problem to the available designs**



For those whose problems can be addressed with classical designs, JMP offers classical design in a totally interactive environment that includes the full power of JMP—graphics, scripting, journaling, project collaboration and trusted analytics.

However, a pre-defined design rarely provides an exact match for your unique industrial process. That is why JMP, unlike most DOE software, goes well-beyond classical designs to give you the flexibility to tailor a design that fits your specific circumstances. With the Custom Designer in JMP, you describe the process variables and constraints, and then JMP constructs a design to match. You can have many different factor roles in the same experiment. Factors can take on continuous, categorical, blocking, mixture and covariate roles. Enter one or more

responses, enter factors, enter factor constraints, choose a model, modify sample size alternatives, choose the run order and add center points and replicates. You can use the Custom Designer for routine factor screening, response optimization, full factorial designs and mixture problems.

When creating designs with the Custom Designer, you can:

- Add factors with any role in any experiment.
- Specify any number of runs per block.
- Include categorical factors with any number of levels and mix factor roles as required by your engineering situation, such as having non-mixture factors in a mixture experiment.
- Create any design with continuous or categorical covariate factors (factors whose values are fixed in advance of the experiment).
- Disallow certain regions of the factor space by defining linear inequality constraints.

### Assess the quality of your design

After generating a design, JMP gives you diagnostic tools to assess the quality of the design. These tools include a Prediction Variance Profiler, a Fraction of Design Space Plot, a Prediction Variance Surface and a table of the Relative Variance of Coefficients that includes the signal-to-noise ratio and power calculations.

### Analyze experimental data

After you define the design you want, JMP saves the design in a data table of runs. JMP constructs scripts that analyze the data correctly and then saves them with the data table. Every design has a script called **Model** that launches the **Fit Model** dialog with the appropriate response and model effects showing in the dialog. Screening designs also have a script called **Screening** that automatically performs a screening analysis.

|   | X1 | X2 |
|---|----|----|
| 1 | 1  | -1 |
| 2 | -1 | 1  |
| 3 | 1  | 1  |
| 4 | 1  | 1  |
| 5 | 1  | 1  |

### How does the Custom Designer work?

The Custom Designer starts by creating a random design with each point inside the range of each factor. The computational method is an iterative algorithm called coordinate exchange. Each iteration of the algorithm tests every value of every factor in the design to determine if replacing that value improves the optimality criterion. If so, the new value replaces the old. Iteration continues until no replacement occurs in an entire iterate.

To help ensure the solution is a global (rather than local) optimum, the whole process is repeated several times using a different random start. The Custom Designer displays the best of these designs. If a design problem has several equivalent solutions (designs with equal precision for estimating the model coefficients as a group), the design algorithm may generate different, but equivalent, designs.

# Custom Designer: Capabilities

## **Cutting-edge technology makes the impossible become simple**

### **Supersaturated designs for factor screening**

Factor screening relies on the sparsity principle. The experimenter expects that only a few of many factors in a screening experiment are active. A supersaturated design can examine dozens of factors using fewer than half as many runs as factors, and can help identify the vital few factors that are driving the process.

### **Optimal designs with random effects**

Does your process change from day to day? If so, the runs you do each day comprise a random block. Allowing any factor to change between runs in a given day results in a randomized block design. Sometimes, though, one or more factors are hard to change between runs. In that case you would like to keep the hard-to-change factors constant for all the runs in a day. This results in a split-plot design. The Custom Designer supports randomized block, split-plot, split-split-plot and even strip-plot designs. The models for these designs have one or more random effects. The JMP Fit Model platform analyzes models with random effects using REML methodology.

### **Process variables in a mixture experiment**

If you have factors that are ingredients in a mixture, you can use either the classic Mixture Designer, or the Custom Designer. However, classic mixture designs are limited because they require all the factors to be mixture components. But, you might want to vary the process settings along with the percentages of the mixture ingredients. The Custom Designer can handle mixture ingredients and process variables in the same study.

### **Screening with flexible sample and block sizes**

When you create a design using the classic Screening Designer, the available block sizes for the listed designs are only powers of two. However, custom designs in JMP can have blocks of any size. After you complete a design with a blocking factor, the Custom Designer shows the appropriate number of blocks calculated as the sample size divided by the number of runs per block.

### **Response surface models with categorical factors**

The classic Response Surface Designer only deals with quantitative factors. You could use the Response Surface Designer to produce an RSM design with a qualitative factor by replicating the design for each level of the factor. The Custom Designer accommodates a categorical or blocking factor when creating an RSM model. Using the Custom Designer is simpler and more cost-effective because fewer runs are required.

### **Avoiding unworkable factor combinations**

Sometimes it is impossible to vary factors independently over their entire experimental range. Therefore, you have factors whose levels are constrained. JMP gives you the flexibility to disallow particular combinations of levels of factors using linear inequality constraints or using a special **Disallowed Combinations** command.

### **Designing experiments with fixed covariate factors**

Sometimes you have fixed quantitative measurements (a covariate) on the experimental units before the experiment begins. If this variable affects the experimental response, the covariate should be a design factor. However, unlike other factor types, this covariate factor cannot be controlled in the study. The Custom Designer supplies a reasonable design option. You can specify a covariate factor and the Custom Designer uses the supplied covariate values to create an optimal design, giving the number of runs you specify.

# Custom Designer: Responses and Factors

## Express the opportunities and goals up front

When you choose **DOE > Custom Design**, the Custom Designer launch dialog appears with panels to enter responses and factors.

### Enter responses

The following Responses panel illustrates how to enter responses along with additional information about them.

Click to enter multiple responses at once.

Click to enter lower and upper limits and importance weights.

To add one response at a time, click and select a goal type **Maximize, Match Target**.

Double-click and type to change response name.

Click to change response goal **Maximize, Match Target, Minimize** or **None**.

| Response Name | Goal     | Lower Limit | Upper Limit | Importance |
|---------------|----------|-------------|-------------|------------|
| optional item | Maximize | .           | .           | .          |

### Enter factors

The following Factors panel illustrates how to enter different types of factors, their level values, and the level of difficulty involved in changing level settings. Indicating a factor is **Hard** or **Very Hard** to change yields split-plot or split-split plot designs.

To add one factor, click and select type of factor **Continuous, Categorical, Blocking, Covariate, Mixture, Constant** or **Uncontrolled**.

To add multiple factors, click and enter number of factors, then click **Add Factor** and select type of factor.

Click and select **Add Level** to increase levels of character variables.

Double-click and type to change response name.

Click and choose **Easy, Hard** or **Very Hard** to indicate the difficulty of changing factor settings.

Click and type to change factor values. To remove a character level, click, press delete, then return.

| Name | Role        | Changes | Values       |
|------|-------------|---------|--------------|
| X1   | Continuous  | Easy    | -1   1       |
| X2   | Categorical | Easy    | L1   L2   L3 |
| X3   | Mixture     | Easy    | 0   1        |

# Custom Designer: Design Diagnostics

## Get a feel for your design



After defining a custom design, but before you create the design table, you can preview the design and investigate details by looking at various plots and tables that serve as design diagnostic tools. Most designs display these diagnostics:

- Prediction Variance Profile plot.
- Fraction of Design Space plot.
- Prediction Variance Surface plot.

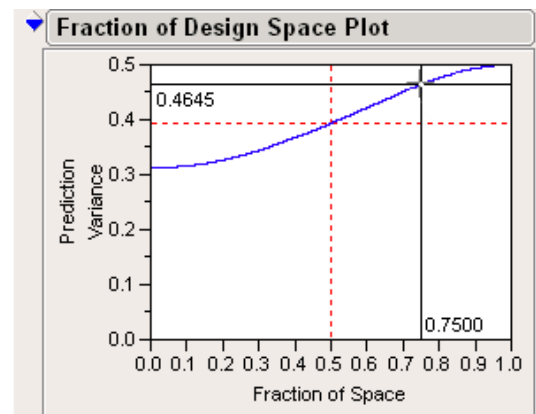
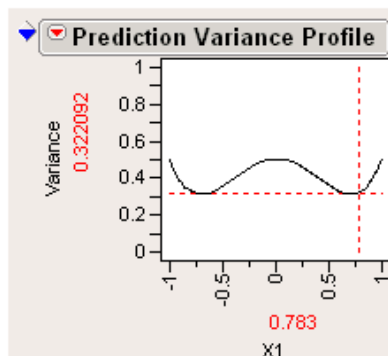
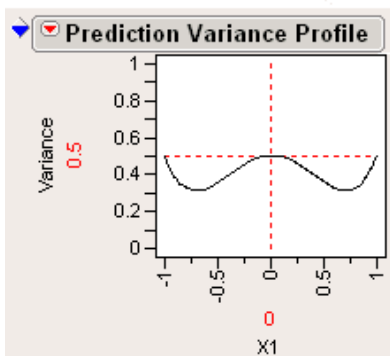
- Relative Variance of Coefficients and Power table.
- Alias matrix.

The prediction variance for any factor setting is the product of the error variance and a quantity that depends on the design and the factor setting. This ratio, called the relative variance of prediction, can be calculated before acquiring the data. It is ideal for the prediction variance to be small throughout the allowable regions of the factors.

In the Prediction Variance Profile, you can drag the vertical trace lines in the plot to change the factor settings to different points. Also, using the **Maximum Desirability** command on the Prediction Variance Profile title bar identifies the maximum prediction variance for a model. Comparing the prediction variance profilers for two designs side-by-side is one way to compare two designs.

The Fraction of Design Space plot is a way to see how much of the model prediction variance lies above (or below) a given value.

As an example, Prediction Variance plots and the Fraction of Design Space plot for a single-factor quadratic model are shown following. The Prediction Variance plot on the left shows that 100% of the values are smaller than 0.5. You can move the vertical trace and also see that all the values are above 0.322, as shown in the middle plot. The Fraction of Design Space plot on the right displays the same information. The X axis is the proportion or percentage of prediction variance values, ranging from 0 to 100%, and the Y axis is the range of prediction variance values. Using the crosshair tool shows that 75% of the possible factor settings have a relative predictive variance less than 0.4645.

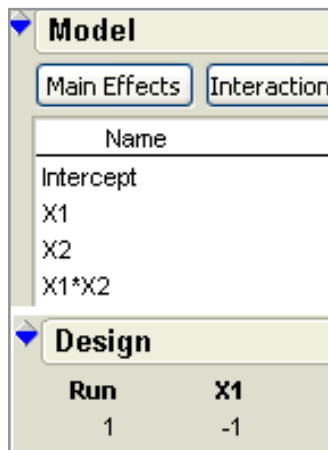


# Custom Designer: Bayesian D-Optimal

## Steer the design with possibilities as well as requirements

In screening designs, experimenters often add center points and other checkpoints to a design to help determine whether the assumed model is adequate. Although this is good practice, it is also ad hoc. The Custom Designer provides a way to improve on this ad hoc practice while supplying a theoretical foundation and an easy-to-use interface for choosing a design robust to the modeling assumptions.

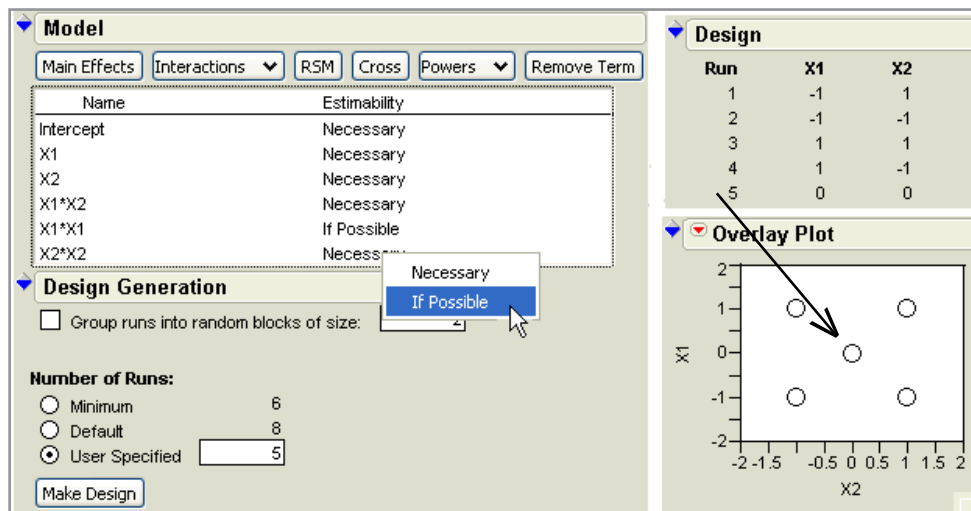
Suppose you have a model for a design with two main effects and an interaction. The minimum number of runs for this design is four (placed on the vertices of the design space), as shown in the illustration to the right. Also, suppose you can afford an extra run. You would like to use this additional point as a checkpoint for curvature.



If you go back to the design generation panel in the Custom Designer and specify five runs instead of the minimum four, the Custom Designer replicates one of the four vertices. Replicating any run is the optimal choice for improving the estimates of the terms in the model, but it provides no way to check for lack of fit.

To generate a five-run design, with the extra run serving as a checkpoint, use the **Powers** button on the Model panel and enter quadratic terms ( $X1^2$  and  $X2^2$ ) to the model. The assumed model consists of the primary terms designated **Necessary** in JMP. The additional higher-order terms are called potential terms and are designated **If Possible**. To take advantage of the benefits of the approach using **If Possible** model terms, the sample size should be larger than the number of primary terms but smaller than the sum of the primary and potential terms. That is,  $p < n < p + q$ .

The formal name of the approach using **If Possible** model terms is Bayesian D-Optimal design. This type of design allows the precise estimation of all of the primary terms while providing omnibus detectability (and some estimability) for the potential terms.



# Custom Designer: Optimality Choices

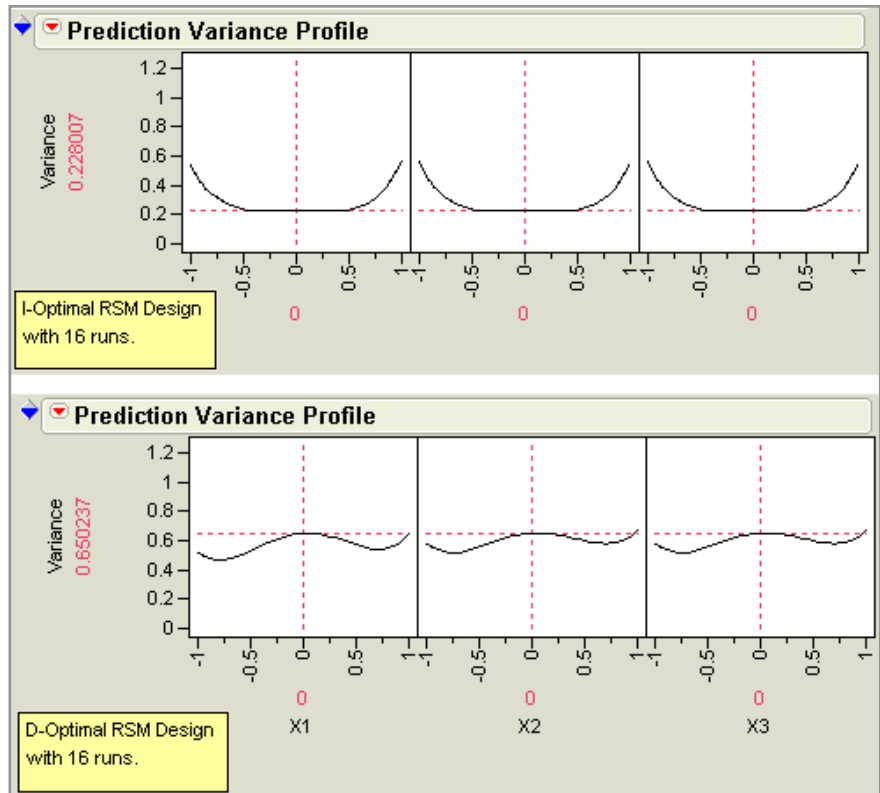
Select 'Best Inference Design' or 'Best Prediction Design'

The Custom Designer generates designs using a mathematical optimality criterion.

- D-Optimal designs are most appropriate for screening experiments because the optimality criterion focuses on precise estimates of the coefficients. If an experimenter has precise estimates of the factor effects, then it is easy to tell which factors' effects are important and which are negligible. So, D-Optimal designs are most efficient for designing experiments where the primary goal is inference.
- I-Optimal designs minimize the average prediction variance inside the region of the factors. This makes I-Optimal designs useful for prediction. As a result, I-optimality is the recommended criterion for JMP response surface designs.

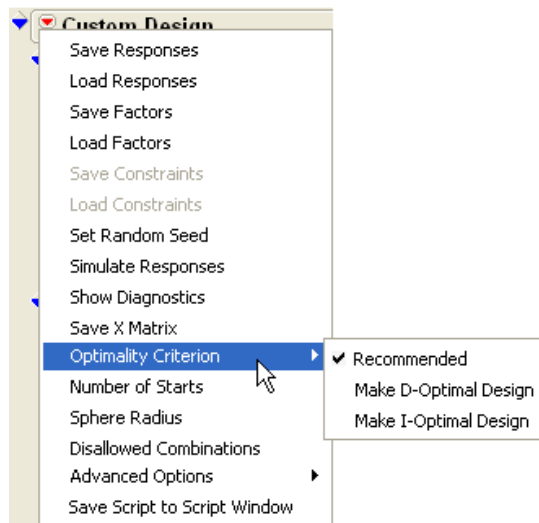
An I-Optimal design tends to place fewer runs at the extremes of the design space than does a D-Optimal design. As a result, D-Optimal designs often predict better at the extreme values of the factors.

The plots shown here compare 16-run response surface models for both criteria. The I-Optimal design has the lowest prediction variance at the center - there are two center points in this design. The D-Optimal design has no center points and its prediction variance at the center of the factor space is about three times the variance of the I-Optimal design.



To change the design criterion, click the red triangle icon on the Custom Design title bar and select **Optimality Criterion**, then choose **Make D-Optimal Design** or **Make I-Optimal Design**.

The default **Recommended** criterion is D-optimal for all design types unless you used the RSM button in the Model panel to add effects that make the model quadratic.



# Custom Designer: Supersaturated Designs

## Screen even when you have more factors than observations

Sometimes you cannot afford to do as many runs as there are factors. It may seem impossible to handle a problem like this. However, JMP has the ability to create and analyze supersaturated designs. Using Bayesian D-Optimality, these designs examine dozens of factors using fewer than half as many runs.

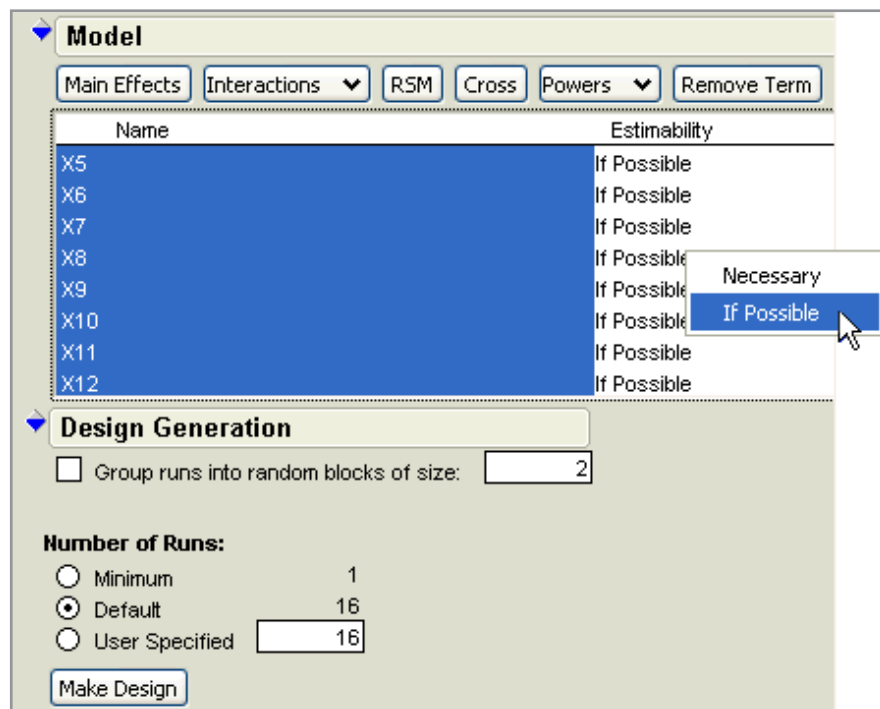
Many of the fractional factorial designs available using the Screening Designer are saturated with respect to a main effects model. In the analysis of a saturated design, you can fit the model, but there are no degrees of freedom for error or for lack of fit. Until recently, saturated designs represented the limit of efficiency in designs for screening.

In general, factor screening relies on the sparsity principle. The experimenter expects only a few of the factors in a screening experiment to be active. The problem is not knowing which are the vital few factors and which are the trivial many.

By default, all factors in the Factors panel show with their estimability set to **Necessary**. In that case, the least number of runs you can enter is the minimum given by the Custom Designer.

In a supersaturated design, all terms are designated **If Possible**. This approach provides the machinery for creating a supersaturated design. Now you can enter any number of runs and the make a design. The Custom

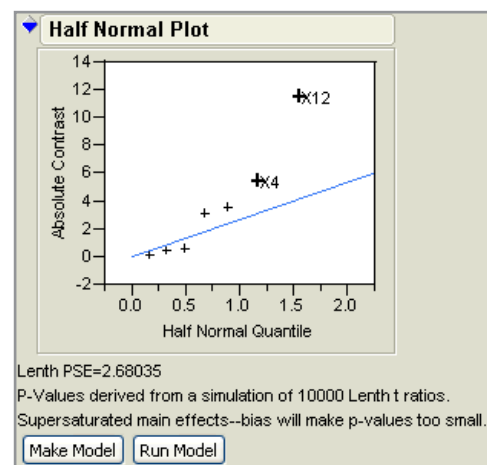
Designer judiciously creates a set of runs that maximizes the information about the factors and tends to produce lowest possible pairwise correlations on average.



## Analyze a supersaturated design

After gathering data, the Screening platform (**Analyze > Modeling > Screening**) offers a streamlined analysis, designed to identify active factors. Its purpose is to evaluate the main effects model. The results given by the Screening platform are concise and easily interpretable. The example shown here is an eight-run design that has twelve factors and simulated response data. The Screening platform clearly shows only two active factors, **X4** and **X12**.

Screening results also summarize the most active factors with a contrast value, a bar chart of the contrasts, and Lenth's t-Ratio with both individual and simultaneous p values. The Half Normal plot has buttons to launch or run the Fit Model dialog, completed with the most active factors from the screening results.



# Custom Designer: Technical Notes

## Using optimality in custom designs

The Custom Designer produces optimal designs using various criteria.

### D-optimality

- Default design type except when the model is a response surface design.
- Minimizes the variance of the model coefficient estimates—appropriate for first-order models and screening, where the experimental goal is to identify the active factors and estimate parameters.
- Depends on a pre-stated model—a limitation because in most real situations, the form of the pre-stated model is not known.
- Has runs whose purpose is to lower the variability of the coefficients of a pre-stated model; might not allow for checking that the model is correct; might not include center points when investigating a first-order model; might only have  $p$  distinct runs with no degrees of freedom for lack of fit.
- Maximizes  $D$  when  $D = \det[\mathbf{X}^T \mathbf{X}]$ .
- D-Optimal split plot designs maximize  $D$  when  $D = \det[\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}]$  where  $\mathbf{V}^{-1}$  is the block diagonal variance matrix of the responses.

### Bayesian D-optimality

- Modification of the D-Optimality criterion that estimates the coefficients in a model; can detect and estimate some higher-order terms; advantageous when there are interactions or curvature.

- Best when the sample size is larger than the number of primary terms but smaller than the sum of the primary and potential term ( $p + q > n > p$ ); allows precise estimation of all of the primary terms while providing omnibus detectability (and some estimability) for the potential terms.
- Uses the potential terms to force in runs that allow for detecting any inadequacy in the model containing only the primary terms. Let  $\mathbf{K}$  be the  $(p + q)$  by  $(p + q)$  diagonal matrix whose first  $p$  diagonal elements are equal to 0 and whose last  $q$  diagonal elements are the constant,  $k$ . The Bayesian D-Optimal design maximizes the determinant of  $(\mathbf{X}^T \mathbf{X} + \mathbf{K})$ . The difference between the criterion for D-Optimality and Bayesian D-Optimality is this constant added to the diagonal elements corresponding to the potential terms in the  $\mathbf{X}^T \mathbf{X}$  matrix.

### I-optimality

- Minimizes the average variance of prediction over the region of the data.
- Appropriate when the goal is to predict the response rather than the coefficients, such as in response surface design problems; can predict the response anywhere inside the region of data and therefore find the factor settings that produce the most desirable response value; best when precise estimation of the response takes precedence over precise estimation of the parameters.

- Maximizes this criterion: If  $f^T(x)$  denotes a row of the  $\mathbf{X}$  matrix corresponding to factor combinations  $x$ , then

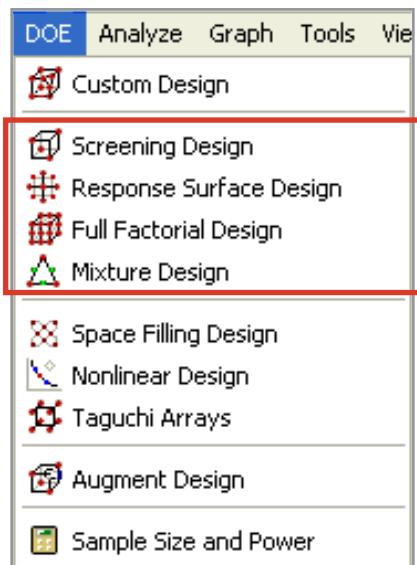
$$I = \int_R f^T(x)(X^T X)^{-1} f(x) dx = \text{Trace}[(X^T X^{-1})M]$$

$$\text{where } M = \int_R f(x)f(x)^T dx$$

is a moment matrix that is independent of the design and can be computed in advance.

# Classic Designs: Predefined Methods

## Oldies but goodies



### Screening designs

Screening experiments separate the group of factors having a significant influence on the response from the rest of the factors. JMP supplies a list of screening designs for an unlimited number of factors and designs that group the experimental runs into blocks of equal sizes where the size is a power of two.

### Response surface designs

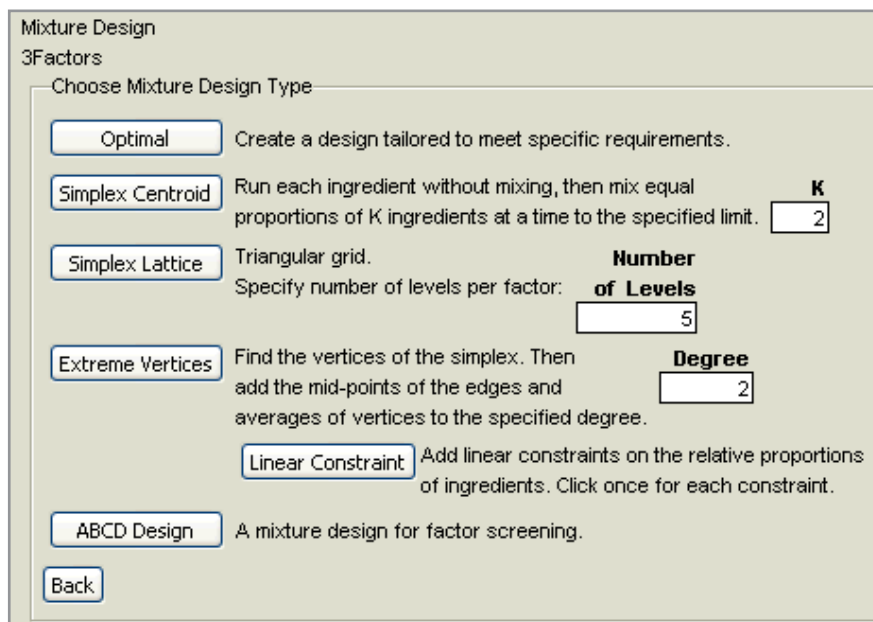
JMP offers standard response surface methodology (RSM) that helps you find the optimal response within the specified ranges of the factors. These designs are capable of fitting a second order prediction equation for the response. The Response Surface Designer offers a Box Behnken design and choice of central composite designs for different sample sizes.

### Full factorial designs

These designs contain all possible combinations of a set of factors. JMP supports both continuous and categorical factors with up to nine levels.

### Mixture designs

JMP offers a way for you to design a set of factors that are ingredients in a mixture. You choose among several mixture design approaches, such as optimal, simplex centroid, simplex lattice, and extreme vertices and ABCE designs. You can specify number of replicates for mixture designs.

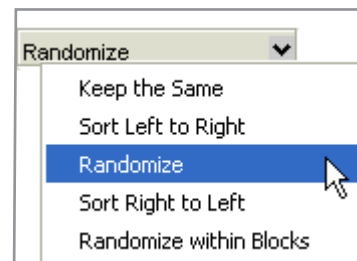


### Design options

Screening, response surface and full factorial designs let you add center points to the design and specify a number of replicates.

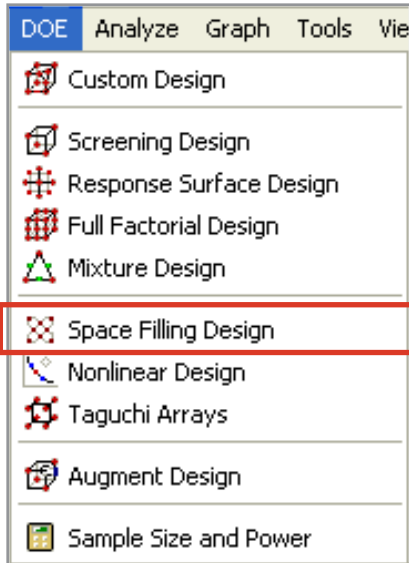
All designs offer options to order the runs in the output JMP data table.

Also, all designs let you preview the runs in the designs, and use a Back button to create a new design and see the effect of using different design options.



# Space Filling Designs: Computer Simulations

**Controlling bias is always a good thing**

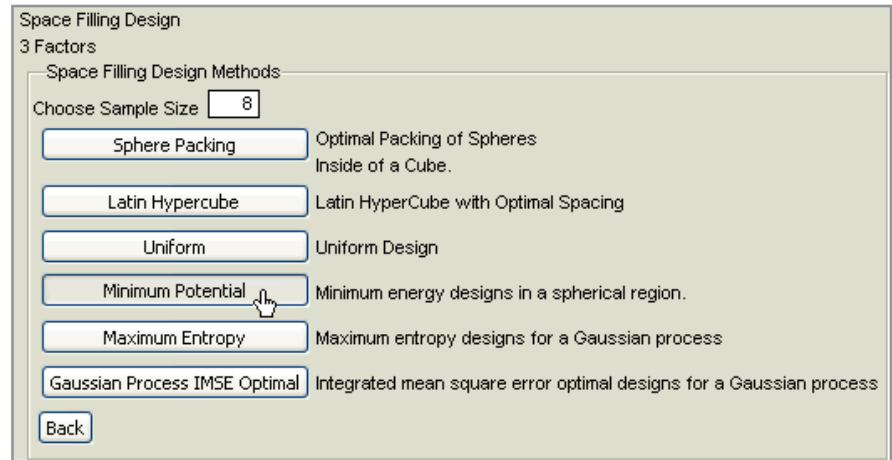


Space filling designs are useful for modeling systems that are deterministic or near-deterministic. One example of a deterministic system is a computer simulation. A goal of designed experiments on these systems is to find a simpler empirical model that adequately predicts the behavior of the system over limited ranges of the factors.

In experiments on deterministic systems, there is no variance but there is bias. Bias is the difference between the approximation model and the true mathematical function. Space filling designs attempt to bound the bias two possible ways:

- Spread the design points as far from each other as possible consistent with staying inside the experimental boundaries.
- Space the points evenly over the region of interest.

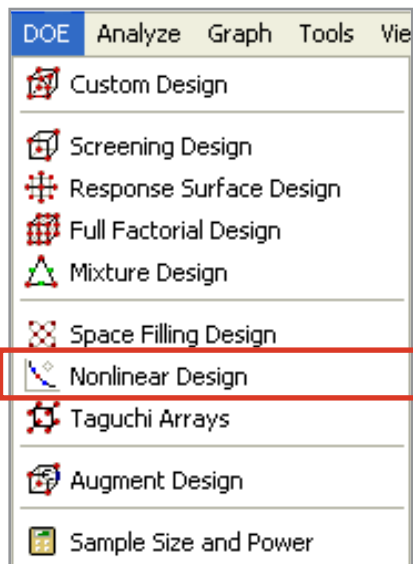
After you launch the Space Filling Designer and enter responses and factors, the menu gives the available space filling methods.



- **Sphere Packing** maximizes the minimum distance between pairs of design points.
- **Latin Hypercube** maximizes the minimum distance between design points but requires even spacing of the levels of each factor. This method produces designs that mimic the uniform distribution. The Latin Hypercube method is a compromise between the Sphere-Packing method and the Uniform design method.
- **Uniform** minimizes the discrepancy between the design points (which have an empirical uniform distribution) and a theoretical uniform distribution.
- **Minimum Potential** spreads points inside a sphere around the center.
- **Maximum Entropy** measures the amount of information contained in the distribution of a set of data.
- **Gaussian Process IMSE Optimal** creates a design that minimizes the integrated mean squared error of the Gaussian process over the experimental region.

# Nonlinear Designs: Fitting Models that are Nonlinear in their Parameters

**Don't let nonlinear models  
throw you a curve**



The Nonlinear Designer generates optimal designs and can optimally augment data for fitting models that are nonlinear in their parameters.

To use the Nonlinear Designer, you need a data table that has:

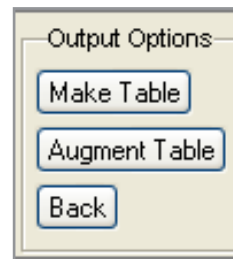
- One column for each factor.
- One column for the response.
- One column that contains a formula showing the functional relationship between the factor(s) and the response.

This is the same format for a table needed to use the nonlinear platform for modeling.

## Create a nonlinear design when there is prior data

If you have experimental data in a JMP data table, that table can be used to create a nonlinear design for improving the estimates of the model's parameters.

- Start by using the nonlinear fitting platform to get parameter estimates.
- Use options to save the parameter estimates in the data table and also generate confidence limits in the nonlinear report.
- Launch the Nonlinear Designer and specify the same model as used in the nonlinear fitting platform.
- Use the confidence limits from the analysis report as parameter ranges in the Nonlinear Designer dialog.
- Now the Nonlinear Designer can make the design. You have the option to make a new table of runs or augment the original table with additional runs.

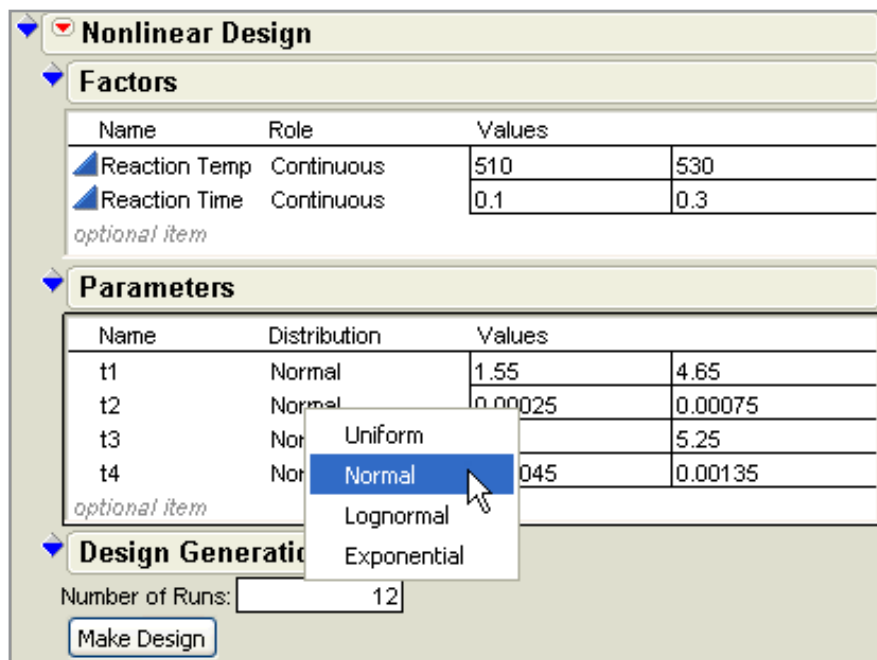


## Create a nonlinear design when there is no prior data

Nonlinear Designer can also be used when you have not yet collected data, but have a guess as to the values of the unknown parameters. You begin with a table template, which is a table that has columns with properties and formulas, but no observations.

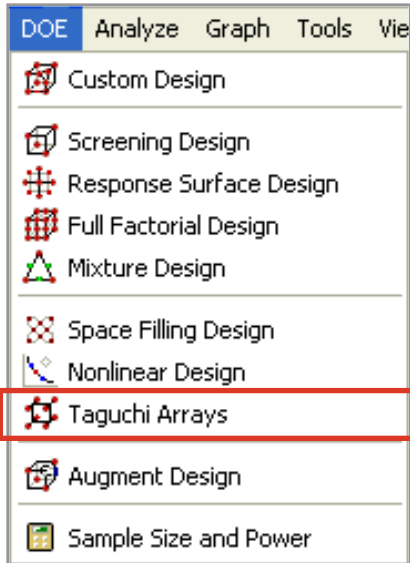
Enter the Y response and the predictor formula column into the Nonlinear Designer launch dialog, with your best guess for factor and parameter values, and supply a number of runs. Click **Make Design**, and then make or augment a design table.

Note that you can specify distributions for the parameters.



# Taguchi Arrays: Industrial Experimentation

**Quality was the watchword of the 1980's**

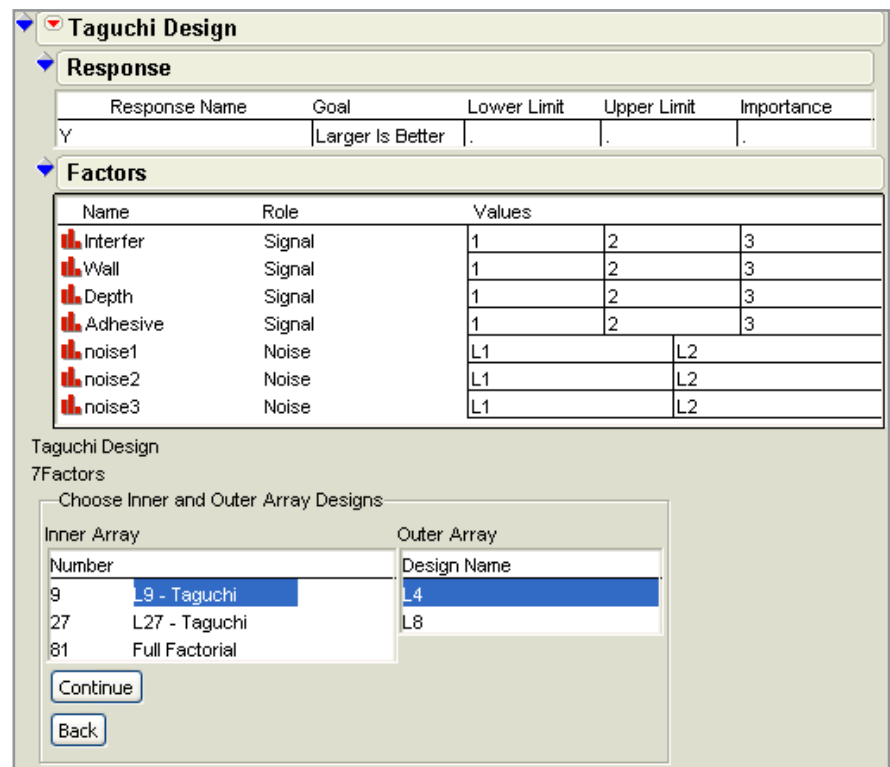


Genichi Taguchi was a leader in the growth of quality consciousness. One of Taguchi's technical contributions to the field of quality control was a new approach to industrial experimentation. The purpose of the Taguchi method was to develop products that worked well in spite of natural variation in materials, operators, suppliers and environmental change. Dividing system variables according to their signal and noise factors is a key ingredient in robust engineering.

Taguchi orthogonal arrays are two-, three- and mixed-level fractional factorial designs. The unique aspects of this approach are the use of signal and noise factors, inner and outer arrays and signal-to-noise ratios.

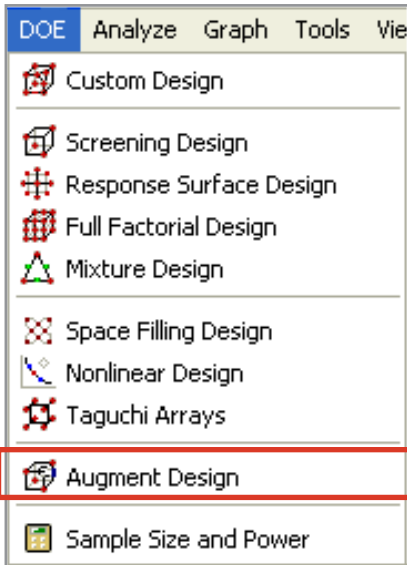
An inner design constructed over the control factors finds optimum settings. An outer design over the noise factors looks at how the response behaves for a wide range of noise conditions. The experiment is performed on all combinations of the inner and outer design runs. A performance statistic is calculated across the outer runs for each inner run. This becomes the response for a fit across the inner design runs.

Taguchi Designer supports all the features of the Taguchi method. You choose from inner and outer array designs, which use the traditional Taguchi orthogonal arrays, such as L4, L8, and L16, as shown in the following completed dialog.



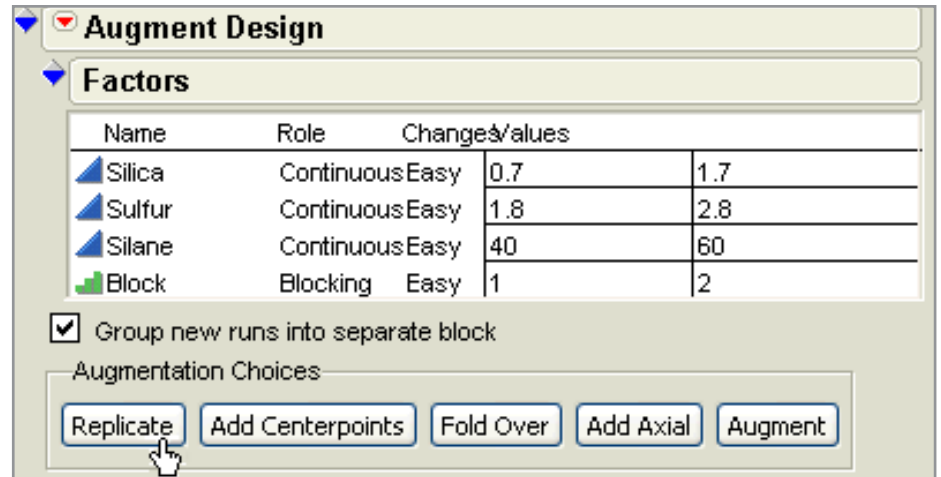
# Augment Designs: Add Runs to an Existing Design

Well-chosen runs can make a good design even better



JMP can add new runs to any existing design table. If you treat experimentation as an iterative process, you can master the temptation to assume that one successful screening experiment has optimized your process. You can also avoid disappointment if a screening experiment leaves behind some ambiguities. The Augment Designer modifies an existing JMP design data table, supporting your iterative process, using any of the following five choices.

- **Replicate** the design a specified number of times. Replication provides a direct check on the assumption that the error variance is constant. It also reduces the variability of the regression coefficients in the presence of large process or measurement variability.



- **Add Centerpoints**, which is useful to check for curvature and reduce the prediction error in the center of the factor region. Center points are usually replicated points that allow for an independent estimate of pure error, which can be used in a lack-of-fit test.
- **Fold Over** creates a foldover design which removes the confounding of two-factor interactions and main effects. This is especially useful as a follow-up to saturated or near-saturated fractional factorial or Plackett-Burman designs.
- **Add Axial** adds axial points together with center points, which transforms a screening design to a response surface design.

- **Augment** adds runs to the design using a model that can have more terms than the original model. Adding runs to a design is powerful because you can achieve the objectives of response surface methodology by changing a linear model to a full quadratic model and adding the necessary number of runs. For example, suppose you start with a two-factor, two-level, four-run design. If you add quadratic terms and five new points to the model, JMP generates the three-by-three full factorial as the optimal augmented design.

D-Optimal augmentation is a powerful tool for sequential design. Using this feature you can add terms to the original model and find optimal new test runs with respect to this expanded model. Note the **Group new runs into separate blocks** check box shown in the preceding Augment dialog. Check this box to group sets of experimental runs into separate blocks, which optimally blocks the second set with respect to the first.



JMP WORLD HEADQUARTERS SAS INSTITUTE INC. +1 919 677 8000 U.S. & CANADA SALES 800 727 0025 [www.jmp.com](http://www.jmp.com)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. -  
Other brand and product names are trademarks of their respective companies. Copyright © 2007 SAS Institute Inc. All rights reserved. 103044\_445055.0907