SAS® System Tools for

Design and Analysis of Experiments

Donna O. Fulenwider
Randall D. Tobias
Robert N. Rodriguez

SAS Institute Inc., Cary, NC

Abstract

SAS software provides a comprehensive set of tools for the design and analysis of experiments. Procedures such as FACTEX, OPTEX, GLM, RSREG, and NLP provide statisticians with powerful algorithms for construction, analysis, and optimization of experimental designs. The new ADX Interface provides a solution for engineers and researchers who prefer a point-and-click interface for harnessing the power of SAS Software for standard experimental designs. This paper presents the highlights of a two-hour SUGI workshop on the features of the experimental design solution provided by the new ADX Interface. It highlights the interactive graphical and modeling capabilities of the Interface as well as recently developed features for process optimization.

Experimental Design

An increasing number of engineers, researchers, and scientists are using designed experimentation to discover and exploit cause-and-effect relationships between important factors. Experimental design is a systematic approach for applying statistical methods to the experimental process. Based on the questions to be answered, a testing plan or design can be constructed with a certain number of factors at an appropriate number of levels. The factor levels are changed systematically and simultaneously according to the design in order to evaluate their effects on the process. These effects are studied using statistical models, and the models are then used for determining:

- the factors that are most important out of many initial factors (screening)
- the factor settings for optimal operating conditions
- the factor settings that reduce variability due to environmental conditions or noise (an approach known as quality engineering)

The end result is a maximum amount of information in a minimal amount of time at a known cost.

Experimental Design Tools

Experimental design involves five steps for which software support is essential:

- construction of the design
- analysis of the response data
- optimization of the responses in the design space
- organization and saving of information about the experiment as it evolves
- reporting of results.

Within the budget constraints of an experiment, various designs are typically available, but not all of them are equally efficient. With no loss of information, the number of runs required by a statistically optimal design can reduce the number of runs required by a "one-factor-at-a-time" design while dramatically increasing the amount and quality of model information. However, the construction of optimal designs requires powerful algorithms.

SAS software provides various tools for the design and analysis of experimental design data. The FACTEX and OPTEX procedures in SAS/QC® software provide powerful algorithms for the construction of designs such as factorial and optimal designs. The GLM, MIXED, and RSREG procedures in SAS/STAT® software provide model fitting. The RSREG procedure specifically provides specialized methods for fitting quadratic response surface models including techniques for response optimization. The NLP Procedure in SAS/OR® software provides the nonlinear optimization capabilities for complex response surfaces. Together these procedures comprise a rich set of experimental design tools for the statistical programmer easily integrated with all facets of the SAS System.

Experimental Design Solution

Beginning with Release 6.03 of the SAS System, SAS Institute recognized that customers wanted the power of experimental design tools in an easy to use menu-driven system. The ADX Menu System for Design of Experiments was created to provide prompt-driven access to SAS software tools for nonprogrammers and statisticians interested in these experimental design methods. During the past few years, the ADX development team has been at work on a completely new interface for the design and analysis of experiments. Prabhu, Rodriguez, and Tobias (1996) provided the first glimpse of the ADX Interface at SUGI 22. Extensive usability studies and feedback from customer demonstrations have provided valuable information, and the result of that effort is the experimental version of the new ADX Interface.

The New ADX Interface

Recently released as experimental software on desktop platforms for Release 6.12 and scheduled for production in Version 7, the ADX Interface is an exciting addition to SAS/QC Software. The Interface is being developed in response to requests from customers for features such as

- a more intuitive interface
- easier data entry
- greater flexibility for model fitting
- interactive graphics
- HTML output for intranet reporting

The Interface is not intended for use as a general statistical tool. The ADX Interface is a self-contained environment designed to bring powerful SAS tools to you as needed throughout the experimental design process. The following example provides an introduction to the new ADX Interface for the analysis of a two-factor design.

Example

This example comes from the automotive industry. Engineers in an engine manufacturing facility seek to reduce the emissions produced by the engine while maintaining or improving its fuel economy. The experiment includes five factors (inputs) that affect process performance. The responses (outputs) to be measured are fuel economy (miles per gallon), hydrocarbon emissions (gm/ml), and nitrogen oxide emissions (gm/ml).

Table 1. Factor Names and Levels

Factor Name	Low Level	High Level
Air-to fuel ratio	16	17
Engine gas recirculation (lb/min)	.18	.92
Spark advance (degrees btc)	20	40
Engine speed (rpm)	1000	1600
Engine Torque (lb-ft)	45	115

You can invoke the experimental version of the ADX Interface by issuing the *xadx* command (in a command bar or on the command line) within the SAS System. Upon invoking the ADX Interface, you enter the ADX desktop shown in Figure 1. The desktop is the place to

- organize and view designs
- construct new designs
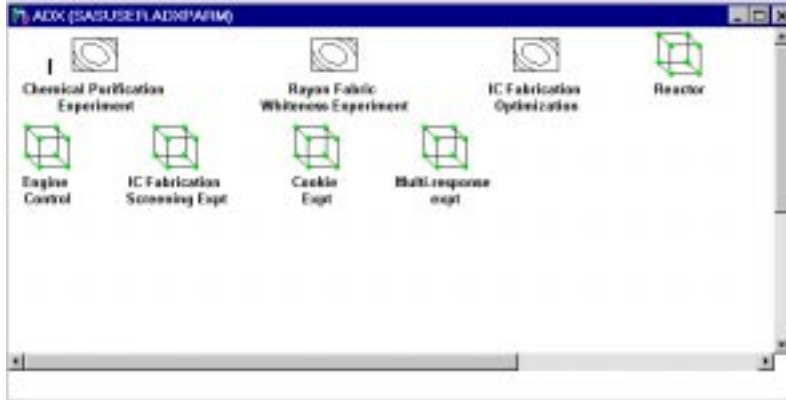- copy, augment, delete, and export existing designs

2

Figure 1. The ADX Desktop

To construct the design for the experiment, select File. The pop-up menu in Figure 2 appears with the choice highlighted to create a new two-level design.
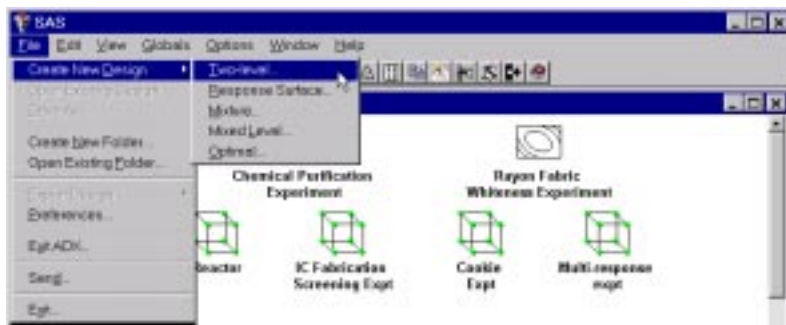

Figure 2. Create a New Design

The main design window for creating and working with designs is shown in Figure 3. Since this is a new design, the design window is empty. The buttons on the right of the window represent the logical steps in conducting the analysis of an experimental design, from defining the variables (factors, responses, and blocks) to reporting on the experiment. Buttons are successively ungrayed as their actions become relevant in the progression of the experiment.
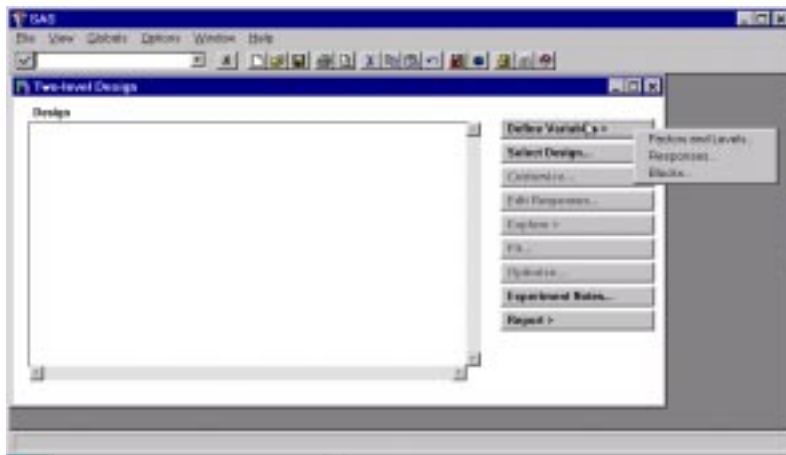
Figure 3. The Main Design Window

Factors can be defined in a data entry window, or factors can be obtained from a factor library. This feature can minimize the amount of time it takes to define factors. Likewise, a response library exists for response variable information. Figure 4 shows an example of a factor library.
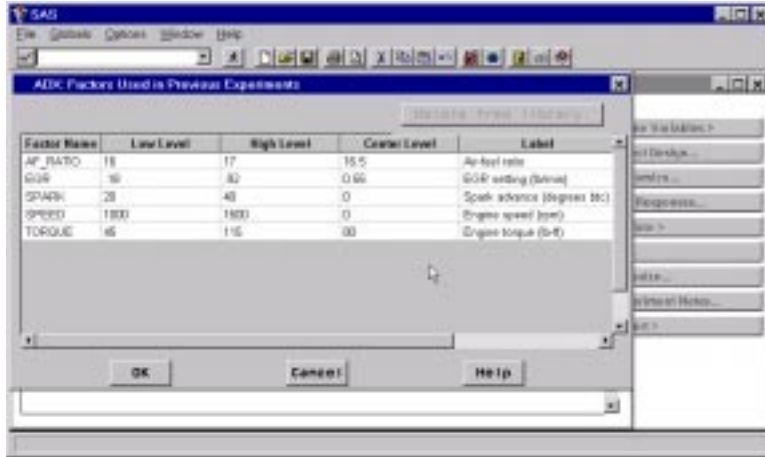

Figure 4. The Factor Library

Once the factors and responses have been defined for this experiment, the next step is to select the design. The design selection window shown in Figure 5 provides a list of available designs for the number of factors in this experiment: a quarter fraction or Resolution III design, a half fraction or Resolution V design, and the full factorial design. If blocking is selected, more designs are available for selection from the list. For simplicity, a full factorial design is chosen for this example.
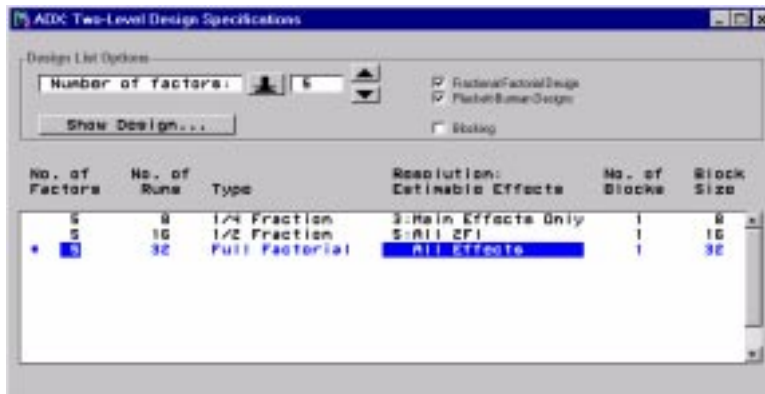

Figure 5. Design Selection List

When selecting a design, you may want to view the aliasing structure and the confounding rules. Selecting the design from the list and then selecting the *Show Design…* button displays this information, as shown in Figure 6.

In this experiment, you expect speed to be confounded with the three-way interaction of air-to-fuel ratio (AF_RATIO), engine gas recirculation (EGR), and engine spark (SPARK). Also, engine torque (TORQUE) is confounded with the two-way interaction of EGR and SPARK.

Upon exiting the design selection window, the design table is populated with a design matrix.   The design matrix can be shown with coded or uncoded levels. Note that the Customize, Edit Responses, and Explore buttons have become active.
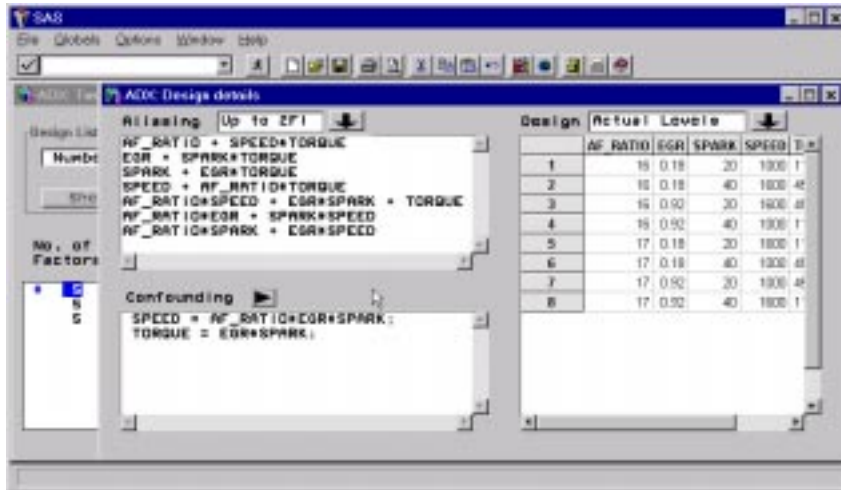
Figure 6. Design Details

At this point, you may want to add center points, replicate runs, replicate the design, or define outer arrays. These are all done within the Customize selection. The engine control example does not require any customization.

Entering the responses can be done in spreadsheet fashion through the Edit Responses window. The ADX Interface makes it easy to enter response values in a number of ways. You can export the design to another spreadsheet package for data entry, or you can cut and paste the design from ADX into another spreadsheet such as Excel. Once response values are entered in the other spreadsheet, you can cut and paste the response values into the Edit Responses window. However, factor and response variables must be defined within the ADX Interface to assure proper bookkeeping and analysis.

Assume that all response values are entered into the design through the Edit Responses window. The design window shows the three responses added to the design and the Fit and Optimize buttons become active (Figure 7).
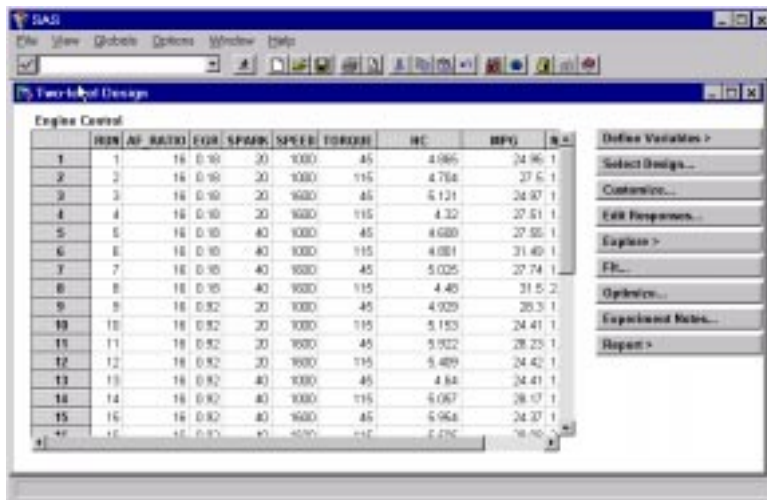

Figure 7. Main Design Window after Response Values are Added

The ADX Interface provides many plots for exploring your data. At this point in the experiment, no models have been fit, so that you are simply looking for patterns and relationships within the data. Selecting the Explore button invokes a window with a tab object, enabling you to switch easily from one graphical display to another. Main effects and interaction plots, and scatter, cube, and factorial plots are available. The engine control example looks at the interaction plot to investigate the relationships between factors for the three response variables. Figure 8 shows the interaction plot for Hydrocarbons.
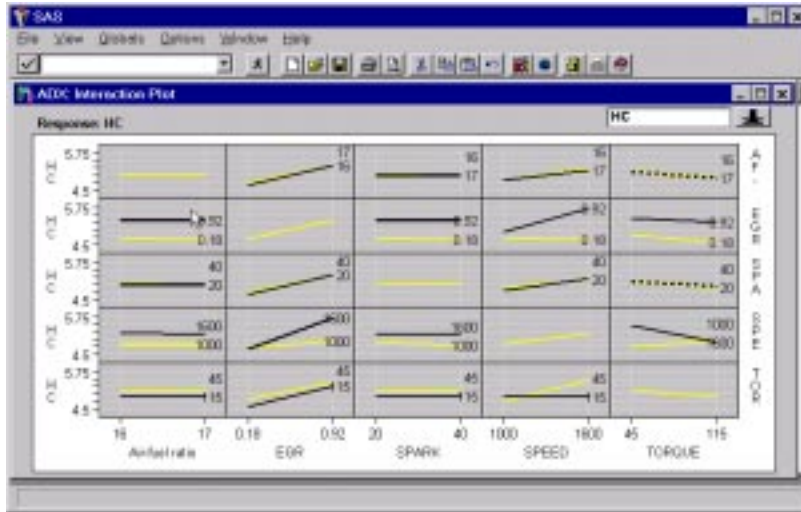
5

Figure 8. Interaction Plot for Hydrocarbon

The Properties selection on the pop-up menu within any graphics region or from the *View* pmenu provides the ability to customize colors, fonts, lines, and so forth for each individual graph or for broadcasting to all graphics within the ADX Interface. Also from the *View* pmenu you can include any graph in a final report (Figure 9). For HTML reports, graphics are displayed on the web page. For text reports, graphs are made available for inclusion in a word processor. Each time you choose to include a graph in a report, the ADX Interface saves the graph in a gif file and notifies you of the location of the file.

Figure 9. View Pmenu

Another plot that provides information about the relationships between factors for each response variable is the cube plot. Figure 10 displays the cube plot for hydrocarbon emissions (HC).
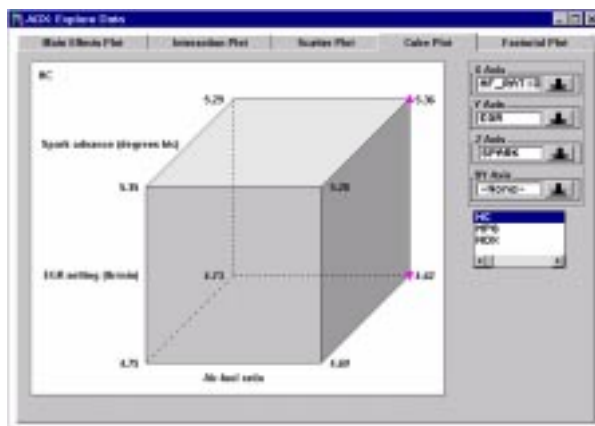

Figure 10. Cube Plot for Hydrocarbon

The value of hydrocarbon emission printed at each vertex is the mean response at that combination of factor levels. Symbols are used to denote the vertices with the largest and smallest mean response levels of hydrocarbon.   In the

cube plot tab, you can select individual cube plots for levels of other factors (BY variable) and also select multiple response variables. Scroll bars enable you to view all of the cube plots.

The next step in the experiment is to fit models for each of the individual responses. Assumption checking is provided for you before entering the Fit window. Stoplight indicators are used to indicate assumptions that may be violated. The Interface checks to see if a transformation of the response variable is needed. Also, residuals are checked for normality, as well as investigation of outlier and influential observations.

Once assumptions are checked, the ADX Interface presents views for effect selection. The default view is a table of effect estimates and p values. The highlighted effects in Figure 11 are those recommended for inclusion in a reduced model. The reduced model is then used for optimization. The default selection of effects may be based on different criteria. The selected effects for this example are based on a default α-level of 0.05. The table is interactive and enables you to select or deselect effects by simply selecting the row. For this experiment, the main effects of EGR, SPEED, and TORQUE are significant, as well as the interactions of SPEED with EGR and SPEED with TORQUE.
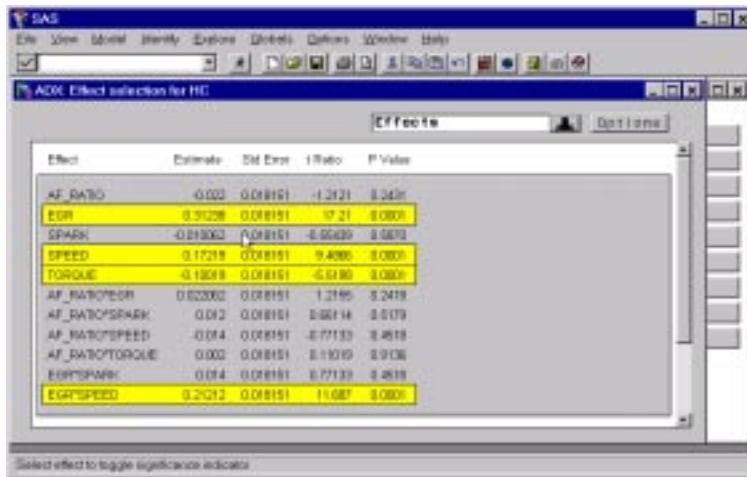

Figure 11. Effects Selection for HC

Other interactive graphics that are available for selecting effects to comprise the reduced model are
- normal plot
- half-normal plot,
- Lenth plot
- Pareto plot
- Bayes plot

The ANOVA table for the full model and suggested reduced model are also available (Figure 12).

Figure 12. ANOVA Table for HC

After fitting models for all three responses, you are interested in finding an optimal region of operation for the factors based on the predictive models. The Optimize selection from the main design window enables you to do just that.

The response optimization window appears as a tab object containing a prediction profiler and a contour optimizer and plot.   The default view is the prediction profile plot (Figure 13). The vertical axis in the plot shows predicted values (rows) for all three responses. The horizontal axis provides columns for the five factors. To aid in this graphical optimization, you can elect to use the desirability function. For more information on the use of the desirability function for optimizing multiple responses, refer to Myers and Montgomery (1995).

The ADX Interface enables you to interact with the graph in order to set the individual desirability values for each response. For HC and NOX emissions, large values are not desirable and would have a desirability of 0 whereas low emissions would have a desirability of 1. Conversely for MPG, high values are desirable (1) and low values are not (0). The function enables you also to set target values and to create a nonlinear desirability relationship.

After you set the desirability value for the individual responses and graphically optimize the overall desirability, values of the factors that may produce optimal predicted responses are found (Figure 13).
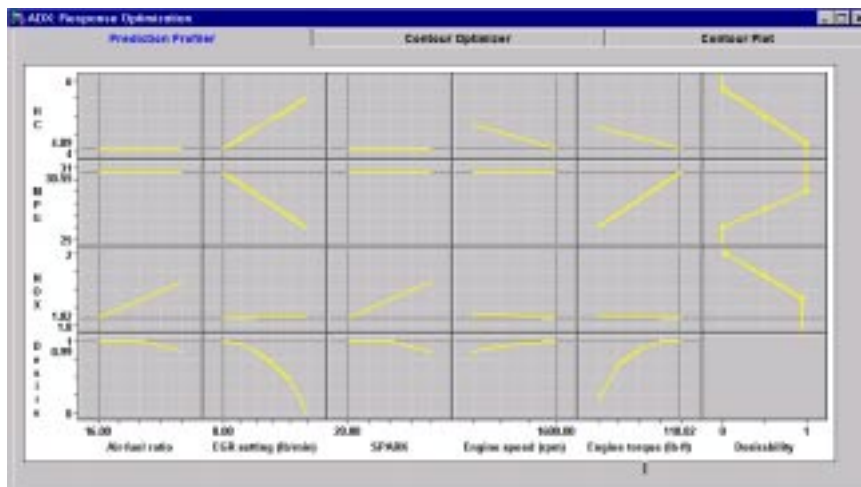

Figure 13. Prediction Profiler

You can choose to select the Contour Optimizer for further investigation. The Contour Optimizer enables you to

- overlay responses
- underlay the standard error of a single response
- set limits on the response values
- change the vertical and horizontal factors
- set fixed values for the nondisplay factors

Upon selecting the Contour Optimizer tab, you can also apply the settings from the prediction profiler.   As shown in Figure 14, the plots in the Optimization step can be linked by the transfer of settings from one plot to the other.
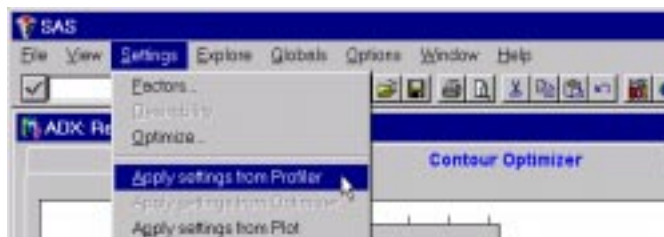

Figure 14. Transferring of settings in Optimization

For this experiment, you are interested in the overlaid contours of HC and MPG, using the standard error of NOX as the underlay selection.   The factors that you are interested in viewing are SPEED and TORQUE. The annotation selection within the contour pop-up menu enables you to traverse the contour and tag a location on the contour with the values of the factors at that location. This highly interactive feature can benefit many engineers responsible for process optimization. Figure 15 shows the contour plot with annotation.
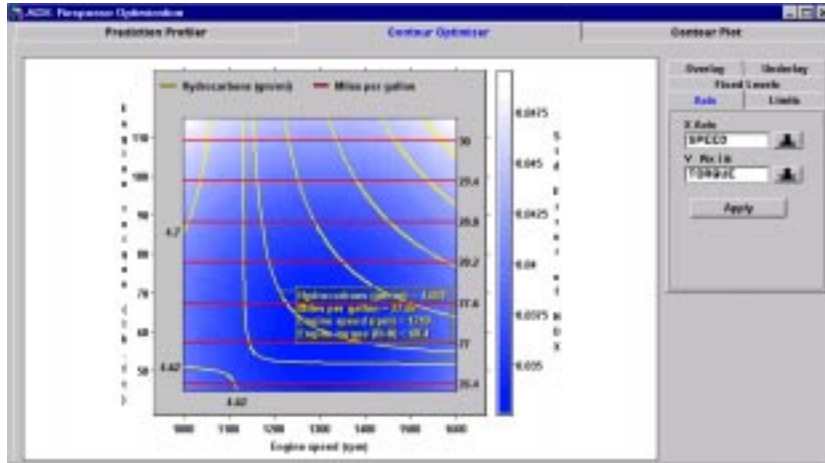

Figure 15. Contour Optimizer with Annotation

The organization and reporting of results is key to designed experimentation. The ADX Interface enables you to keep experimental notes for organizing and documenting the experiment. Notes can also be entered for each factor, response, and block variable. Reports are available in either text or HTML formats. You can also customize the report by selecting specific information and responses to be included in the report.(Figure 16).
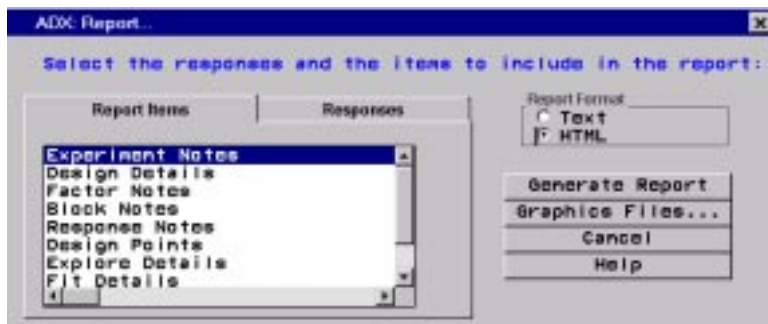

Figure 16. Report Facility

Also, for HTML output you can elect to review the graphics that you have chosen to include in your report. The list of graphics can be managed within the report facility. You can delete certain graphics from the report or change the graphics label for use in the report. Figure 17 shows an example of the graphics management facility.

After selecting the report items and responses for your report, select Generate Report. This invokes the default browser on your desktop operating system for the HTML report. Output for this example is shown in Figure 18.
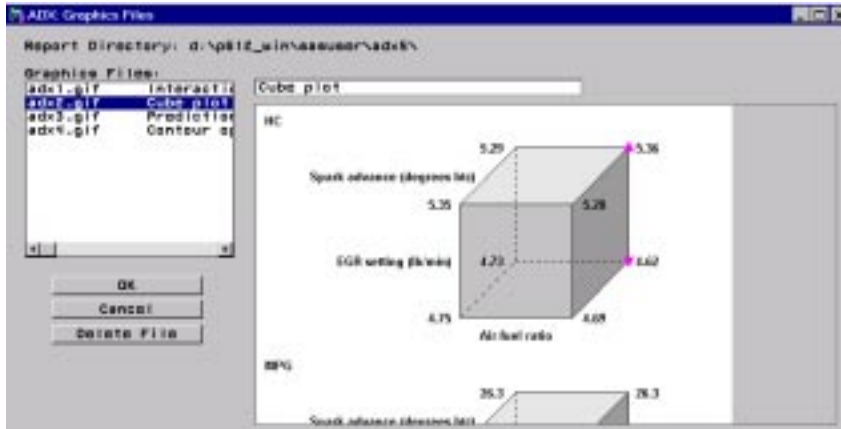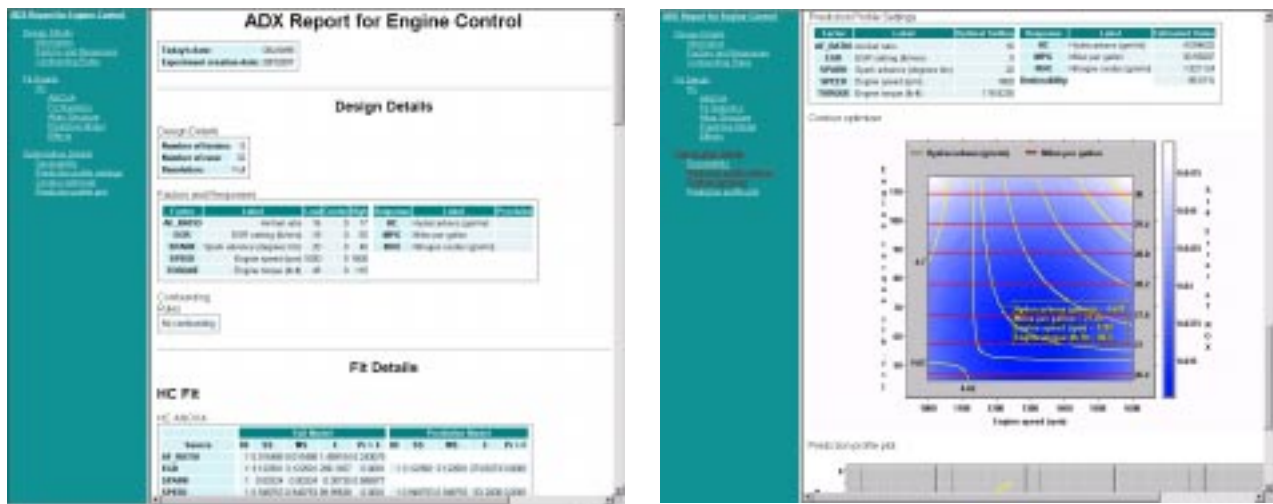
Figure 17. Graphics Management Facility

Figure 18. HTML Output

Conclusion

An increasing number of engineers, researchers, and scientists are using designed experimentation to discover and exploit cause-and-effect relationships between important factors. Software support for successful experimentation is essential for

- the construction of the design
- the analysis of the response data
- the optimization of the responses in the design space
- the organization and saving of information about the experiment as it evolves
- the reporting of results.

The new ADX Interface provides this support in a self-contained environment that brings the powerful tools of SAS software to the fingertips of engineers, researchers, and scientists.   Developed in response to customer requests, the Interface provides the guidance and analytical tools necessary to support designed experimentation.

For updates on the progress of the new ADX Interface for the Design and Analysis of Experiments, visit us at www.sas.com/rnd/app/qc/qcnew.html

References

Myers, Raymond H., and Douglas C. Montgomery (1995), *Response Surface Methodology,* New York: John Wiley & Sons.

Prabhu, Sharad W., Robert N. Rodriguez, and Randall D. Tobias (1996), "The New ADX Interface for Design and Analysis of Experiments," *Proceedings of the Twenty-Second Annual SAS Users Group International Conference, 22,1258-1270. I*

SAS, SAS/OR, SAS/QC, and SAS/STAT are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Acknowledgements

Contact

Donna O. Fulenwider
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
Phone: (919) 677-8000, x-4356
Fax: (919) 677-4444
Email: sasdof@wnt.sas.com