

FUSION PRO SOFTWARE

10 Reasons Why **Fusion Pro** is the
World's Best Design of Experiments Software



S-MATRIX CORPORATION
www.smatrix.com

Contents

1.	Workflow Based Expert System Framework	3
2.	Supports Risk Assessment.....	4
4.	Simple Experiment Setup	6
5.	Automated Experimental Design Selection	7
6.	Automates Handling of Complex Response Data.....	8
7.	Automated Data Modeling	12
8.	Identifies Critical Process Parameters (CPPs)	13
9.	Identifies Best Performing and Most Robust Conditions.....	15
10.	Graphically Visualizes & Reports the Robust Design Space	18

1. Workflow Based Expert System Framework

Scientists and engineers spend an enormous amount of time learning abstract statistical experimental design and data modeling theories in order to use general statistics software correctly and effectively, and avoid the many mistakes common to non-statistician users.

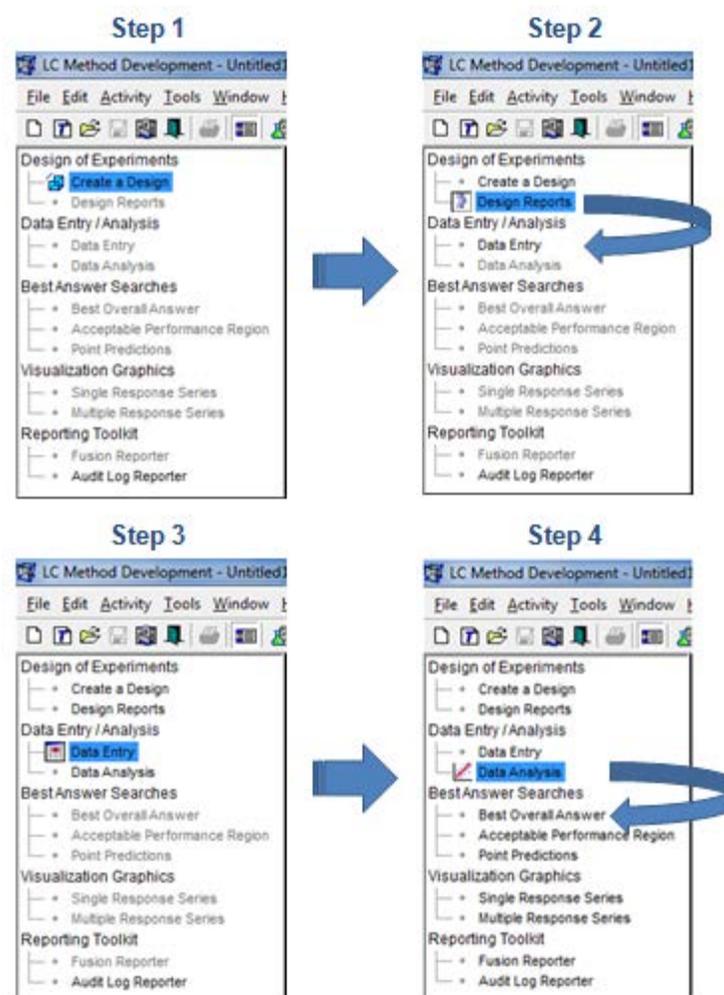
Fusion Pro is expert system software which is completely aligned with QbD principles and guidances, and designed for working scientists and engineers. The software provides a complete suite of Design of Experiments (DOE) technologies, including advanced modeling and simulation capabilities, to support formulation and process development R&D work within a simple workflow based work environment, as illustrated and described below.

Step 1 – you start in Create a Design. Here you define the experiment variables you want to study. Note that this is the only activated work area until your design is generated.

Step 2 – in the Automated mode, **Fusion Pro** generates the most statistically defensible experimental design. It then activates Design Reports, brings you there automatically, and also activates Data Entry where you enter your test results.

Step 3 – Once you enter your experiment test results, **Fusion Pro** activates the Data Analysis work area.

Step 4 – In the automated mode, **Fusion Pro** analyzes and models all your test results data sets with one click. **Fusion Pro** then activates the work areas where you use these models to search for best operating conditions and visualize your QbD Design Space and PARs.



2. Supports Risk Assessment

Fusion Pro integrates the four “*Supporting Statistical Tools*” shown in blue highlight below which are listed in the **ICH Q9 Quality Risk Management** guidance (refer to ICH Q9 Nov 9, 2005, Section I.9, reproduced below).

I.9 Supporting Statistical Tools

“Statistical tools can support and facilitate quality risk management. They can enable effective data assessment, aid in determining the significance of the data set(s), and facilitate more reliable decision making. A listing of some of the principal statistical tools commonly used in the pharmaceutical industry is provided:”

- Control Charts, for example: ...
- Design of Experiments (DOE)
- Histograms
- Pareto Charts
- Process Capability Analysis

As part of formulation and process development experimentation **Risk Assessment** includes identification of Critical Process Parameters (CPPs) and accurate characterization of all the types of effects that the CPPs have on the defined Critical Quality Attributes (CQAs). This quantitative characterization leads to knowledge of the level of control of each CPP required for a robust formulation or process. **Fusion Pro** enables the user to study multiple parameters in combination, can characterize their independent, interactive, and complex effects, and can illustrate these effects in its graphical reporting of the Design Space and Proven Acceptable Ranges for the study parameters (see Item 6). **Fusion Pro** therefore provides strategic tools and capabilities for statistical experimental design based studies which support Quality Risk Management (QRM) as defined in the **ICH Q9 Quality Risk Management** guidance (refer to ICH Q9 Section II.3, reproduced on the following page).

II.3 Quality Risk Management as Part of development

To design a quality product and its manufacturing process to consistently deliver the intended performance of the product (see ICH Q8); (*see Items 3, 9, 10, and 11*)

To enhance knowledge of product performance over a wide range of material attributes (e.g., particle size distribution, moisture content, flow properties), processing options and process parameters; (*see Items 3, 4, 9, 10 and 11*)

To assess the critical attributes of raw materials, solvents, Active Pharmaceutical Ingredient (API) starting materials, APIs, excipients, or packaging materials; (*see Items 3, 7, and 8*)

To establish appropriate specifications, identify critical process parameters and establish manufacturing controls ...; (*see Items 9, 10, 11, and 12*)

To decrease variability of quality attributes:

- reduce product and material defects
- reduce manufacturing defects

(*See Item 10*)

To assess the need for additional studies (e.g., bioequivalence, stability) relating to scale up and technology transfer; (*see Items 3, 9, and 10*)

To make use of the “design space” concept (see ICH Q8). (*see Items 10, 11, and 12*)

4. Simple Experiment Setup

Fusion Pro has the ability to create statistically valid experiments for any combination of variables selected for study. The built-in Design Wizard does this by evaluating your level of study (screening or optimization), and the nature of your defined study variables (Continuous or Discrete Process, Categorical, or Mixture).

Figure 1 shows a part of the Experiment Setup workspace. The study defined in the figure is a combined Mixture-Process study which includes three *Mixture* variables (formulation components) and one Continuous numeric Process variable – Compaction Force. Note that the Mixture setup controls provide the flexibility to define the constant portion of the formulation – the “Bulking Agent” in this example - as well as the study components. Note also that the user can specify the exact nature of each study variable:

- Continuous: A variable with levels that exist on a continuous number scale in which levels intermediate to those defined for the variable have meaning and can be achieved in an experiment.
- Discrete Numeric: A variable with levels that exist on a continuous number scale, but have only specific levels that are desired or that can be achieved given current equipment, conditions, etc.
- Categorical: A *non*-numeric variable having levels that do not exist on a numerical scale.

Figure 1

Experiment Setup

Mixture Variable Settings

No. of Mixture Variables: 4

Units: % Mixture Amount: 100.0

Mixture Variable	State	Lower Bound	Upper Bound
Excipient 1	Variable	0.0	40.0
Excipient 2	Variable	0.0	40.0
Excipient 3	Variable	0.0	40.0
Bulking Agent	Constant	60.00	---

Process Variable Settings

No. of Process Variables: 1

Name	Units	Type	Lower Bound	Upper Bound
Compaction Force	kN	Continuous	9.68	20.35

State: Variable Constant

5. Automated Experimental Design Selection

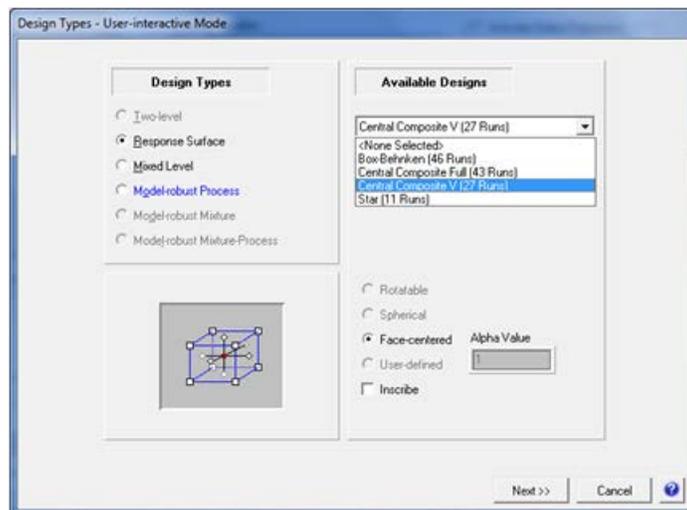
Fusion Pro has a built-in Design Wizard with both Automated and User-interactive Modes. The Automated Mode will automatically select the most efficient DOE design based on your stage of work (e.g. screening or optimization) and the types of variables you selected for study (see Item 4). The User-interactive mode gives the user complete control over the design selection and configuration settings while showing its statistically defensible recommendations, as shown in the wizard dialog pictured in Figure 2.

Fusion Pro has a comprehensive experimental design library which includes the statistical experimental design types listed below. This assures the user that **Fusion Pro** can generate a statistically valid and defensible DOE design for any combination of formulation and/or process variables you want to study. However, **Fusion Pro** will disable DOE design options which are not valid for your selected variables.

Fusion Pro has built-in knowledge of each experiment design type in its DOE Design library.

- Full Factorial Designs
- Fractional Factorial Designs
- Plackett-Burman Designs
- Central Composite Designs
- Box-Behnken Designs
- Star Designs
- Mixed-level Designs
- Model Robust Designs (D-optimal, A+G Optimal)

Figure 2



6. Automates Handling of Complex Response Data

Many response measurement protocols either require or naturally provide multiple test results for each experiment run. But regression modeling requires a single response value for each run. This is why **Fusion Product Development** and **Fusion Pro** have powerful Response Data Handler Toolsets which make it simple and automated to collect and manage multiple test results for any response measurement, and convert repeat test data into analysis-ready data sets.

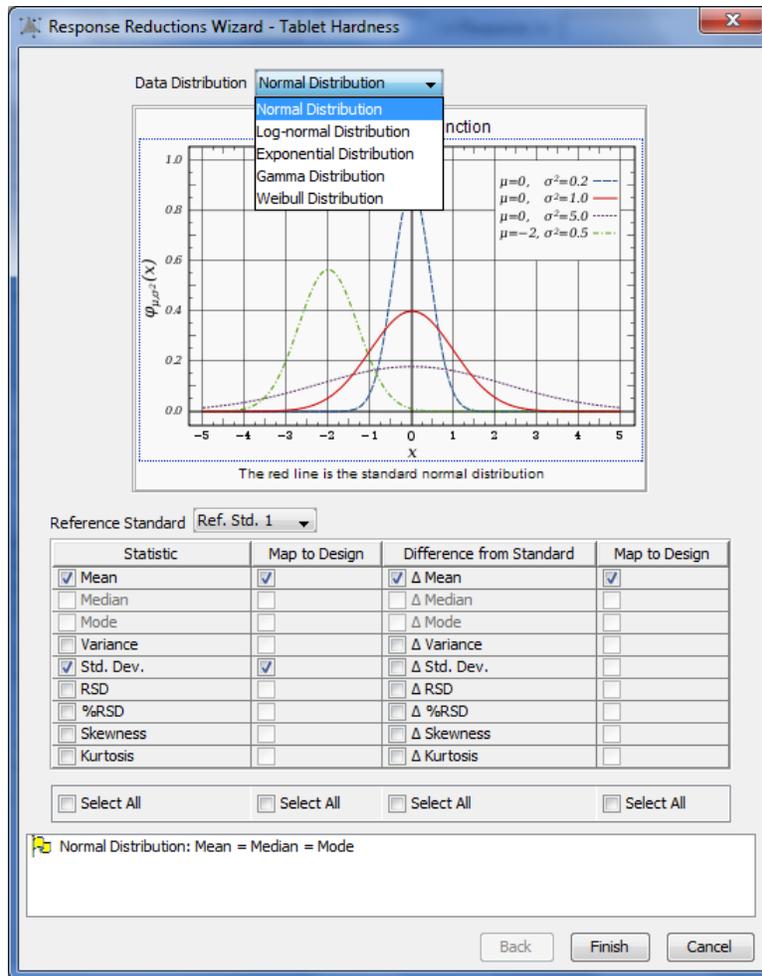
In addition, when your response measurement process includes one or more reference standards, the toolsets can compute the difference between the analysis-ready results computed from your experiment run repeat test data and the same result computed from the reference standard data. For example, a mean value can be computed from the test repeat data in each experiment run and also from the reference standard test repeat data. This enables the toolset to create two reduced responses for each experiment run – the mean response for the run and the difference of the mean from the reference standard mean response.

Descriptive Statistics Toolset – Built In

This toolset is included with **Fusion Product Development** and **Fusion Pro**. It is meant for response data sets which consist of multiple test repeats for each experiment run. Since linear regression analysis requires a single scalar response value for each experiment run, multiple test repeat data for each run must be reduced to one or more representative scalar response values. The most obvious reduction of multiple test repeat data values is computing the mean value (\bar{X}). As shown in Figure 4, the Descriptive Statistics toolset enables you to automatically compute the Mean, Variance, Standard Deviation (S.D.), Relative S.D., and % RSD from multiple test repeat data sets, and automatically map these analysis-ready responses to your experiment design.

In addition, Descriptive Statistics can be computed for test repeat data which follow one of the five data distributions: Normal, Log-normal, Exponential, Gamma, and Weibull. Specifying the distribution of your data assures that the descriptive statistics reduction options you select will be computed correctly.

Figure 4



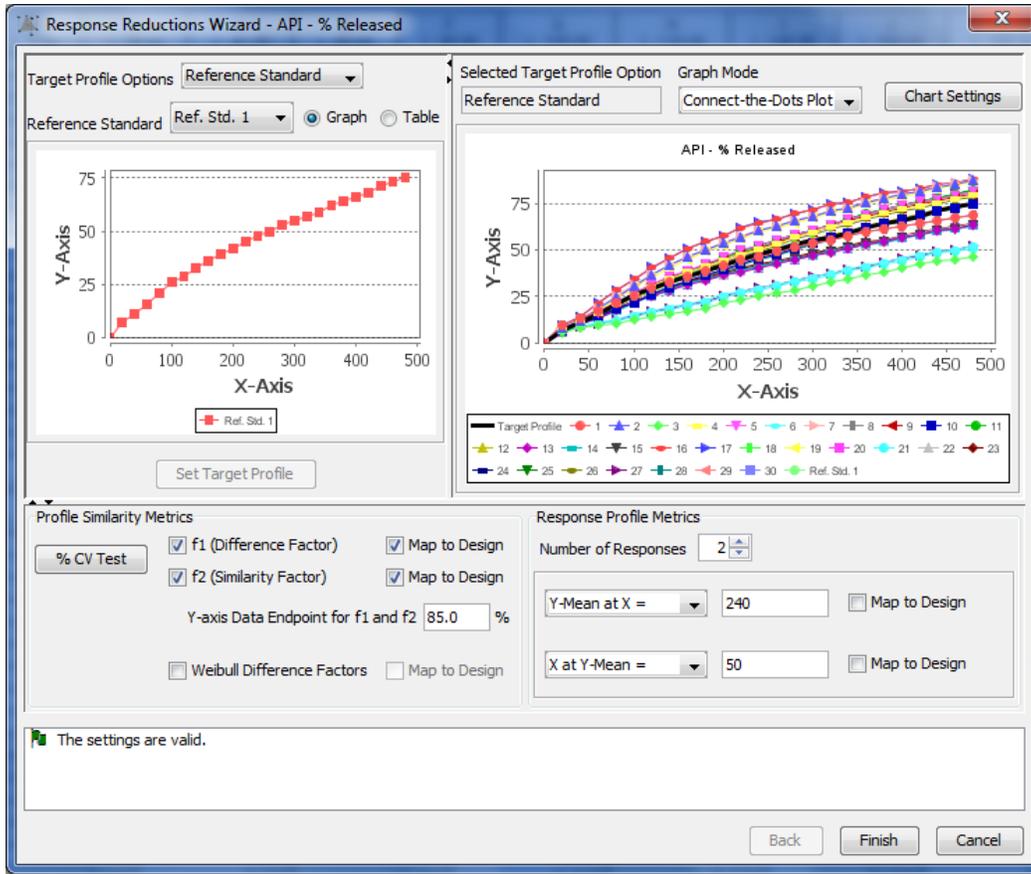
Time Series Toolset – Built In

This toolset automatically handles response data sets which consist of multiple test repeats obtained at different time points for each experiment run. This type of time series based response data would be obtained from dissolution testing and synthesis reaction testing. In these cases the response data set represents the trend, or profile, of how the response changes incrementally over time, as well as specific response values at different critical time points.

The Time Series toolset can automatically create testing plans for managing time-series response data. FPD can also export these testing plans to your Chromatography Data Software (CDS) as ready-to-run sequences, and automatically retrieve all chromatogram results data in preparation for computing the final analysis-ready results you need for your study (See Reason 6). As shown in Figure 5 the Time Series Toolset enables you to easily and automatically:

- Graph the time-series test results obtained for each experiment run.
- Compute mean response profiles from test repeat data at each time point.
- Compute f1 & f2 curve fit metrics
- Perform regulatory guidance specified data quality checks for f1 & f2 calcs.
- Compute sensitive Weibull curve fit metrics.
- Compute additional profile response metrics such as the mean response at a given test time point, or the time point associated with the given test result.
- Map all computed responses to the experimental design for automated data analysis.

Figure 5



7. Automated Data Modeling

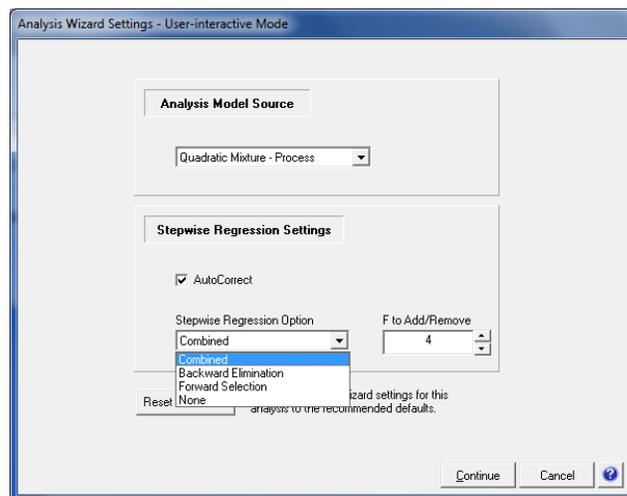
Fusion Pro has a built-in Analysis Wizard with both Automated and User-interactive Modes. The Automated Mode will automatically derive the most statistically defensible prediction equation (model) from the experimental data for each response included in the study. The Automated mode is a state-of-the-art Multiple Linear Regression (MLR) technology which contains all the relevant component analysis operations required of a fully automated MLR routine, implemented in the correct operational sequence. The User-interactive mode gives the user complete control over the MLR configuration settings while showing its statistically defensible recommendations, as shown in the wizard dialog pictured in Figure 5.

The correctness of **Fusion Pro**'s automated MLR routine was rigorously verified by Douglas C. Montgomery, Ph.D., Regents' Professor of Industrial Engineering & Statistics at Arizona State University. Dr. Montgomery is the acknowledged world leader in the field of applied industrial statistics with special expertise in statistical experimental design, also referred to as Design of Experiments (DOE).

“It is especially noteworthy that about a decade ago a major international corporation spent over \$1,000,000.00 attempting to develop an automated MLR analysis capability. They were unsuccessful. This validation exercise therefore demonstrates that Fusion QbD’s automated MLR software is at once both correctly implemented and a major advance in the field of MLR analysis.”

Douglas C. Montgomery, Ph.D.

Figure 5



8. Identifies Critical Process Parameters (CPPs)

Fusion Pro automatically generates a full statistical reporting of each equation derived from the experimental data. One of these reports is the Mean Effects Pareto Ranking Table and Plot report shown in Figure 6 (report tables) and Figure 7 (report graph). This report was generated for the Tablet Hardness response data obtained from a tablet coater process optimization experiment. The effects ranking is based on the magnitude of the change in the response as the variables are changed across their individual experimental ranges. The ranking is applied to all effects, including interaction effects [e.g. $(X_i * X_j)$], curvature effects [e.g. $(X_i)^2$], and complex effects [e.g. $(X_i)^2 * X_j$] as well as independent additive effects (X_i).

For a given response, an experiment variable is a Critical Process Parameter (CPP) if one or more of the *statistically* significant effects involving the variable are of *practical* significance to the user. For example, Gun Distance is a CPP for this response, since its independent additive effect (D) and its interaction effects [(C*D) and (B*D)] are statistically significant *and* practically significant. However, Atomizing Air Pressure (A) is not a CPP for this response, since its only significant effect - the (A*D) interaction - only changes the Tablet Hardness by - 0.06 units, and therefore is not *practically* significant.

Figure 6

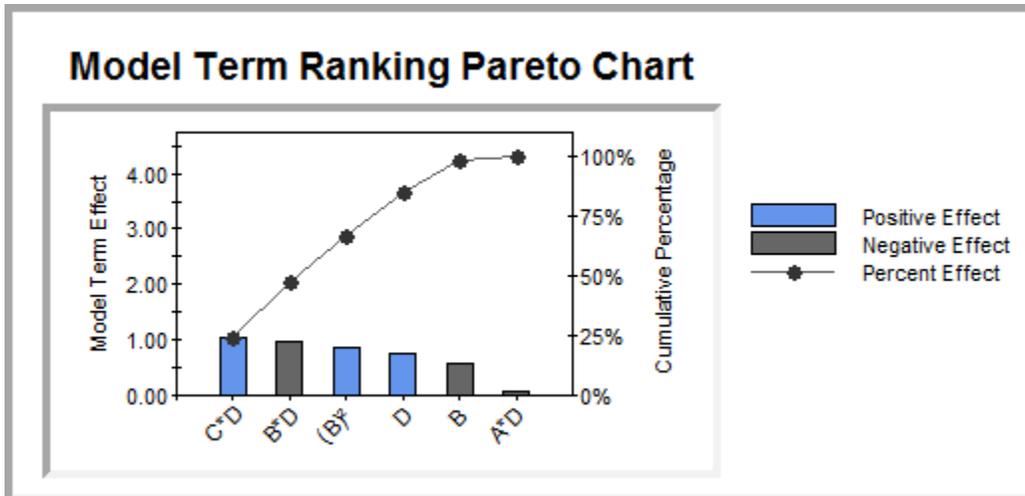
Study Variable Code Name Key

Study Variable	Code Name
Atomizing Air Pressure	A
Pattern Air Pressure	B
Spray Rate	C
Gun Distance	D

Mean Effects Ranking Table

Model Term	Model Term Range	Coefficient Value	Model Term Effect
C*D	2.0000	0.5248	1.0497
B*D	2.0000	-0.4936	-0.9872
(B)^2	1.0000	0.8525	0.8525
D	2.0000	0.3848	0.7695
B	2.0000	-0.2957	-0.5914
A*D	2.0000	-0.0310	-0.0619

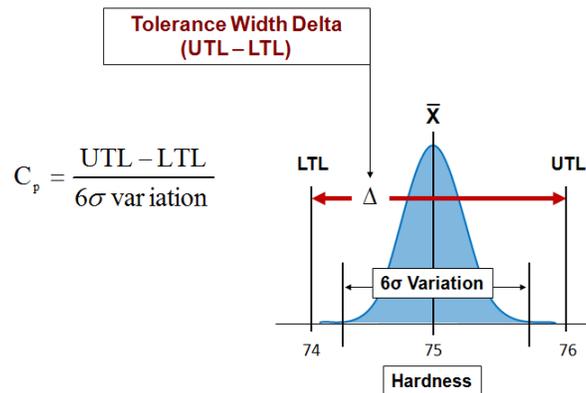
Figure 7



9. Identifies Best Performing and Most Robust Conditions

The robustness approach implemented in **Fusion Pro** uses Process Capability indices (e.g. C_p , C_{pk}) to quantify system robustness. C_p is a standard Statistical Process Control (SPC) metric widely used to quantify and evaluate variation in critical quality and performance characteristics. For a given critical performance characteristic, C_p is the ratio of the performance characteristic's allowable tolerance width to its inherent variation, and is computed as shown in Figure 8.

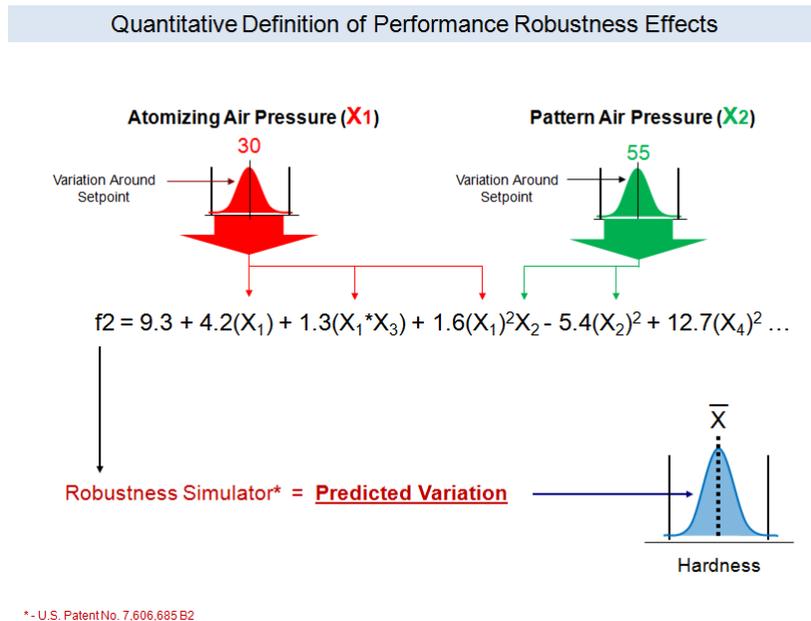
Figure 8



where **UTL** and **LTL** are the upper and lower response tolerance limits, and **6 σ variation** is the amount of observed response variation about the mean response value bounded by the $\pm 3\sigma$ confidence interval limits. C_p is therefore a direct measure of inherent process variation relative to specified tolerances. Figure 10 illustrates the C_p calculation elements for a Tablet Coater process optimization experiment result – the Hardness of the coated tablet (Hardness) – given a mean Hardness (\bar{X}) of 75.0 and tolerance limits of ± 1.0 . In classical SPC a process is deemed *capable* when its measured C_p is ≥ 1.33 . The value of 1.33 means that the inherent process variation, as defined by the 6σ variation limits, is equal to 75% of the tolerance limits ($4/3 = 1.33$). Conversely, a process is deemed *not capable* when its measured C_p is ≤ 1.00 , as the value of 1.00 means that the 6σ variation limits are located at the tolerance limits.

Figure 9 illustrates how **Fusion Pro** uses a critical CQA response prediction model in a Monte Carlo simulation mode to obtain the prediction of the response variation for given candidate variable level setting combination. First, 10,000 level settings are defined for each study variable using the normal distribution, with mean value equal to the variable's target level setting and the 6σ variation (the ±3.0σ limits) the user has defined for the variable. Next, the 10,000 combinations of the variable's variation distribution level settings are input into the CQA response prediction model, one combination at a time, which provides 10,000 response predictions. The distribution of these predictions correctly reflects the cumulative variation resulting from the combined variations of the individual study variables around their setpoints during a process run. Statistically characterizing this distribution then provides the 6σ variation value required in the Cp calculation.

Figure 9



Fusion Pro equations derived from the experiment data characterize the effects of the experiment variables on each analyzed response (CQA). **Fusion Pro** can combine these equations with the Monte Carlo Simulation equations for predicting robustness into its global Best Answer Search utility to find the best overall conditions from the current experiment.

The Best Answer Search operates based on the performance goals the user sets for the included responses. The search engine uses the equations to identify the level setting combination of the study variables which results in best overall performance in terms of meeting all goals simultaneously. For example, the dialog pictured in Figure 10 shows the goal settings for each response data set included in the study. Figure 11 shows part of the answer report from the optimization search.

Figure 10

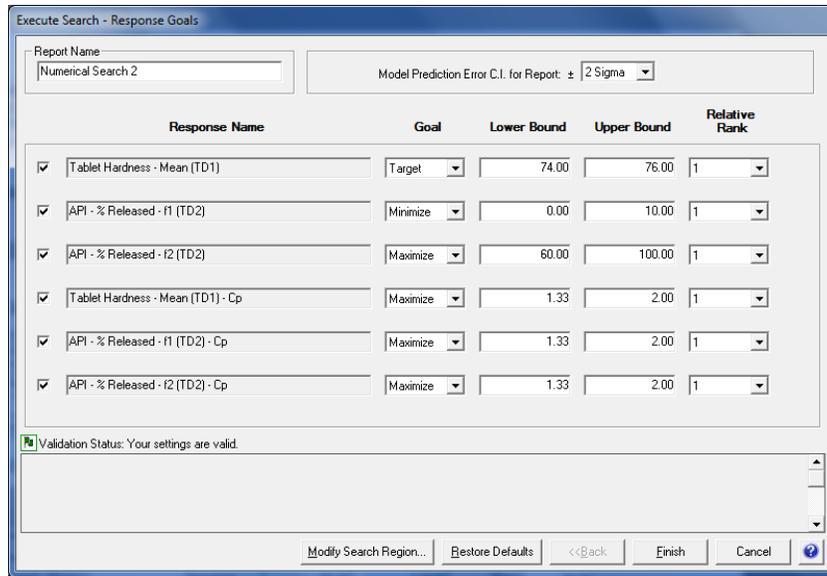


Figure 11

Answer #1

Variable Settings

Variable	Level Setting
Atomizing Air Pressure	38.7
Pattern Air Pressure	53.2
Spray Rate	40.0
Gun Distance	7.6

Predicted Results

Response	Goal	Predicted Result	Desirability	-2 Sigma Conf. Limit	+2 Sigma Conf. Limit
Tablet Hardness - Mean (TD1)	75.00	75.01	0.9943	74.87	75.14
API - % Released - f1 (TD2)	Minimize	0.38	0.9617	0.16	0.89
API - % Released - f2 (TD2)	Maximize	85.04	0.6260	82.41	87.67
Tablet Hardness - Mean (TD1) - Cp	Maximize	13.31	1.0000	13.31	13.31
API - % Released - f1 (TD2) - Cp	Maximize	3.08	1.0000	1.57	6.47
API - % Released - f2 (TD2) - Cp	Maximize	2.00	0.9949	1.98	2.02

Cumulative Desirability Target = 1.0000
 Cumulative Desirability Result = 0.9191

10. Graphically Visualizes & Reports the Robust Design Space

The Mean Performance and C_p metrics are applied directly to all CQAs to determine the relative robustness of a candidate level setting combination of the variables in terms of each CQA. This enables you in a single plot to see the region where the graphed variable settings will yield acceptable results for all responses simultaneously, including robustness responses – i.e., the combined **Robust Design Space**. Such an overlay graph is presented in Figure 12.a below, while Figure 12.b below presents the same graph after scribing the **Proven Acceptable Ranges** (PARs) onto the design space within the graph.

Figure 12.a. Design Space

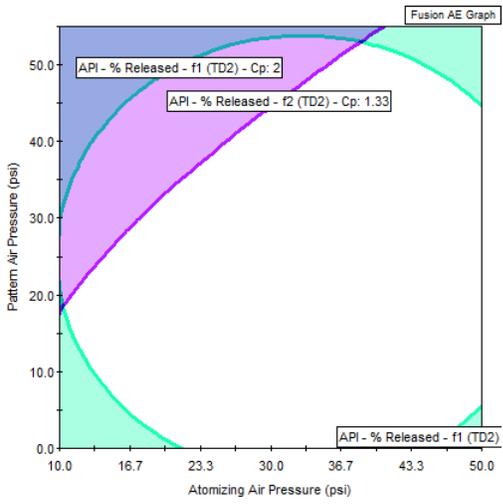
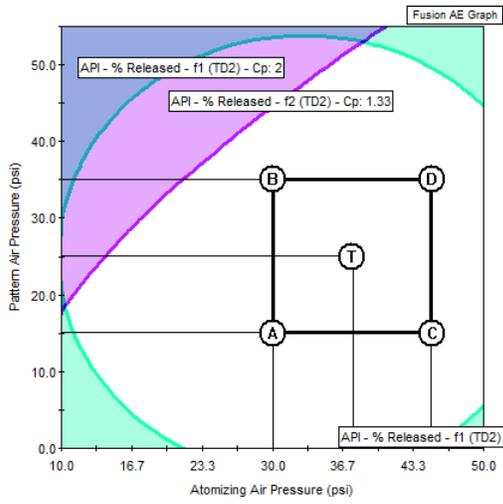
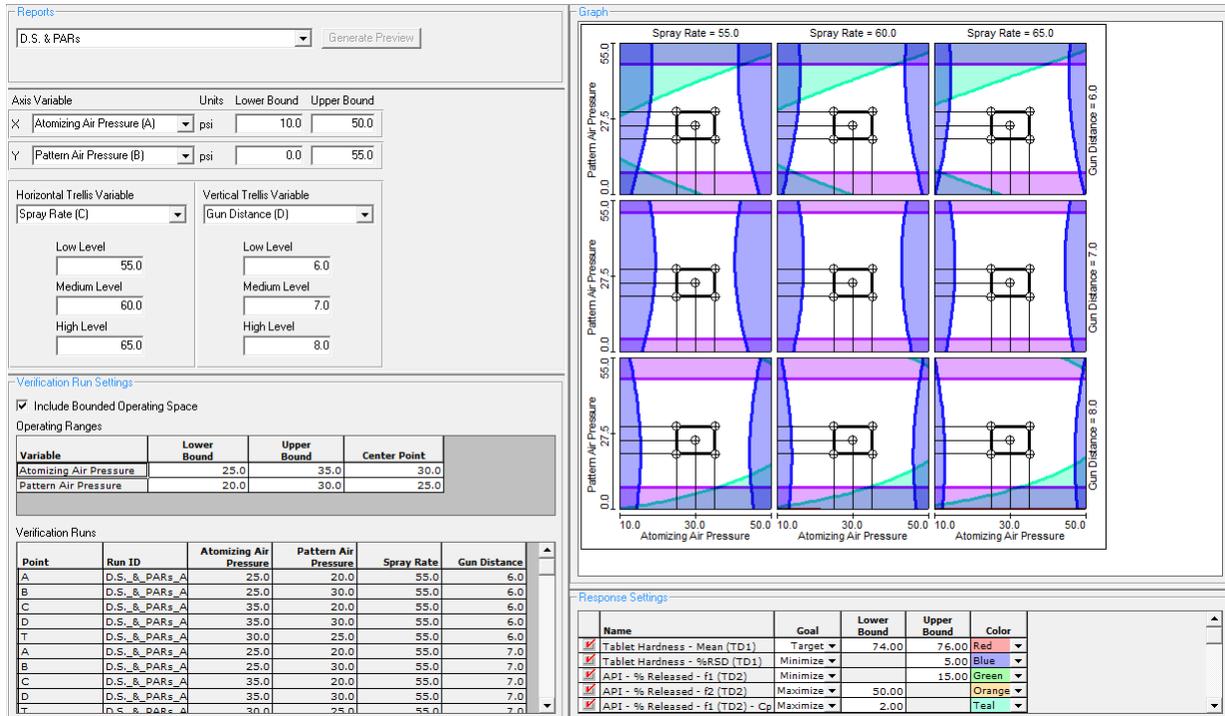


Figure 12.b. Design Space & PARs



Fusion Pro also has a 4D trellis graphing capability shown in Figure 13. This lets you visualize and report a graphical representation of the 4-variable study variable combinations which result in **acceptable** performance (*unshaded* regions) and **unacceptable** performance (*shaded* regions). The 4D trellis graphics also operate based on the performance goals the user sets for the included responses. Note that a specific color is associated with each response to denote where the performance for that specific response is unacceptable. This enables the 4D trellis feature to graphically display the 4-variable Design Space and PARs in terms of meeting all user-defined goals simultaneously.

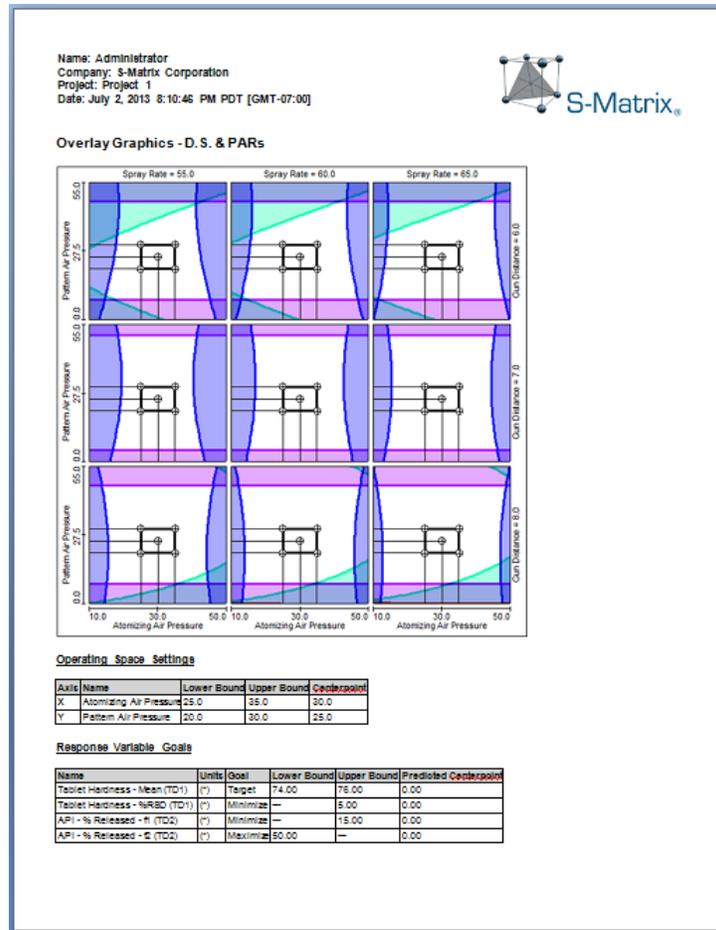
Figure 13



Generates Submittal Quality Reports

Fusion Pro enables the user to create final reports of the Design Space and proven acceptable ranges (PARs – the “operating space”) which conform to FDA/ICH Guidances. Figure 14 is an example of one of these **Fusion Pro** reports, which can be compared directly to the guidelines presented in the ICH Q8(R2) Guidance. This is an ideal format for reporting QbD experiment results.

Figure 14



Fusion Pro's comprehensive documentation also includes detailed statistical reports of the DOE design and the analysis results associated with each experiment. Any or all of these detail reports can be compiled into an internal report or an external report, depending on the target audience and the desired level of reporting detail.

For each response included in the experiment **Fusion Pro** generates a detail report of each of the interrelated analyses associated with the model (equation) it derives from the data. This includes an experimental error report, a transformation analysis report, a regression analysis report, a report of the statistical properties of the final model (equation), and a Pareto Ranking of Variable Effects report. In addition to these detail reports, **Fusion Pro** creates an Analysis Summary Report which contains a table and graph summarizing the key analysis results (Figure 15). The Analysis Summary Report defines the statistical quality of the model (model sufficiency), and is often used by customers as part of their analysis documentation.

Figure 15

