Mixture Design Tutorial

(Part 1/2- The Basics)

Introduction

In this tutorial you are introduced to mixture design. If you are in a hurry to learn about mixture design and analysis, bypass the sidebars. However, if/when you can circle back, takes advantage of these educational sidetracks.

☒ ***Explore more fundamental features of the software***: Mixture design is really a specialized form of response surface methods (RSM). To keep this tutorial to the point, we will not go back over features detailed in the RSM tutorials so you'd best either work through these first, or do so after completing this one on mixtures. Otherwise you will remain ignorant of many useful features and miss out on important nuances in interpreting outputs.

To gain a full working knowledge of this powerful tool, we recommend you attend our workshop on Mixture Design for Optimal Formulations. Call Stat-Ease or visit our website at www.statease.com for a schedule. For a free primer on mixture design, go to the Stat-Ease home page and follow the link that says "I'm a formulator." If you seek statistical details on this topic, see John Cornell's *Experiments with Mixtures*, 3rd edition, published by John Wiley and Sons, New York, in 2002.

This tutorial demonstrates only essential program functions. For more details, check our extensive Help system, accessible at any time by pressing F1. Its hypertext search capability makes it easy for you to find the information you need.

The Case Study – Formulating a Detergent

A detergent must be re-formulated to fine-tune two product attributes, which are measured as responses from a designed experiment:

- Y_1 - viscosity
- Y_2 - turbidity.

Three primary components vary as shown:

- 3% ≤ A (water) ≤ 8%
- 2% ≤ B (alcohol) ≤ 4%
- 2% ≤ C (urea) ≤ 4%

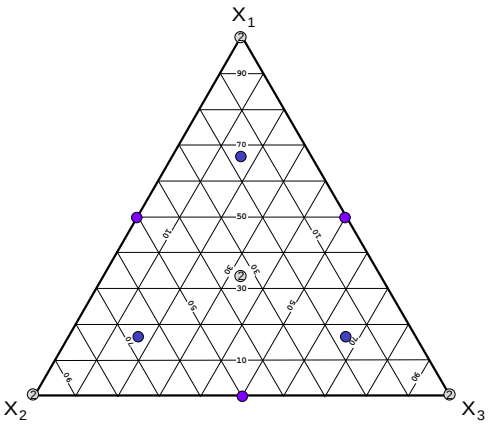
These components represent nine weight-percent of the total formulation, that is:

A + B + C = 9%

Other materials (held constant) make up the difference: 91 weight-percent of the detergent. For purposes of this experiment they are ignored.

Experimenters chose a standard mixture design called a simplex lattice. They augmented this design with axial check blends and the overall centroid. Vertices

and overall centroid were replicated, increasing the experiment size to 14 blends total.



Augmented simplex lattice (second degree)

This case study leads you through all the steps of design and analysis for mixtures. The next tutorial, Part 2, instructs how to simultaneously optimize the two responses.

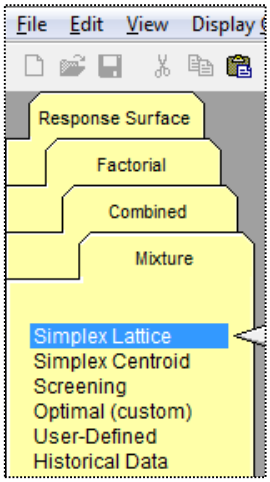
Design the Experiment

Start the program by finding and double clicking the Design-Expert software icon. Take the quickest route to initiating a new design by clicking the **New Design Button** on the opening screen as shown below.



Opening screen – Click New Design Button to begin

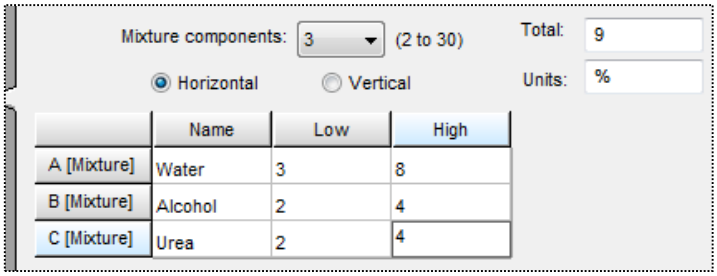
Click the **Mixture** folder tab. The design you want, a simplex lattice, comes up by default.



Choosing a mixture design

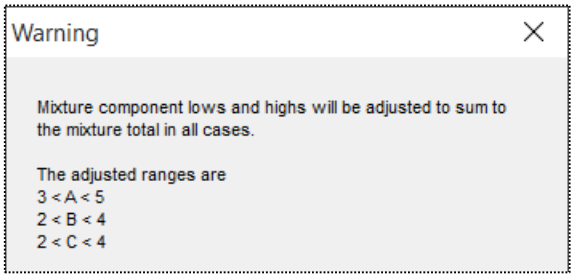
☒ ***Explore Screen Tips and Help on mixture design options***: Some on-screen details appear, but more are available by pressing the screen tips button (or via Tips, Screen Tips). Check this out! Close Tips by pressing **X** at its upper-right corner. Next explore **Help**, **Contents**. Double-click **Mixture Designs,** then **Mixture Design Choices**. After looking over all this helpful information, close Help by pressing **X**.

Now change the number of **Mixture Components** to **3**. Enter components and their limits as shown below in all **Name**, **Low,** and **High** fields, pressing the **Tab** key after each entry. Enter **9** in the **Total** field and **%** as your **Units**.



Entering components, limits, and total

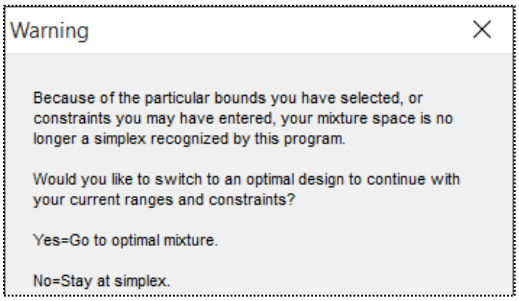
Press **Continue**. Immediately a warning appears.



Adjustment made to constraints

Press **OK**. Notice that although you entered the high limit for water as 8%, Design-Expert adjusts it to 5% — leaving room for 2% each of the other two ingredients within the 9% total. Otherwise, at 8%, water and the low levels of alcohol and urea would total 12%. Design-Expert recognizes that this does not compute. Very helpful!

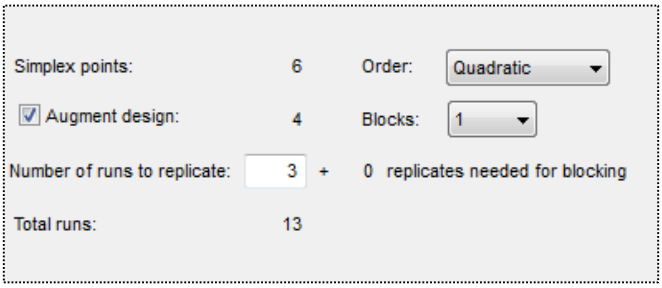
☒ ***Explore optimal design***: For many mixture designs you may be warned at this point that it cannot fit a simplex.



Warning that comes up if mixture region is not a simplex

You should then shift to the Optimal design choice. See this option detailed in the Multifactor RSM Optimal tutorial.

Press **Continue** and the software's adjustment lets you move on. Now you must choose the order of the model you expect is appropriate for the system being studied. In this case, assume that a quadratic polynomial, which includes second-order terms for curvature, will adequately model the responses. Therefore, leave the order at **Quadratic**. Keep the default check-mark at "Augment design" but change **Number of runs to replicate** to **3**. Press your **Tab** key to display the correct number of total runs.

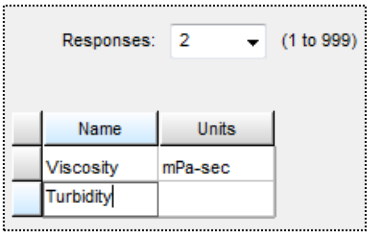


Simplex-lattice design form (after pressing Tab key)

By keeping (accepting) the "Augment design" check-mark, you allow Design-Expert to add the overall centroid and axial check blends to the design points.

☒ ***Explore why the experiment reduced the replicates***: The "Number of runs to replicate" field, which had defaulted to 4, causes the specified number of experiments to be duplicated. In this case, there are three points that are duplicated – the vertices of the triangular simplex. This makes the fourth replicate a bit awkward because it creates an imbalance in the design. Feel free to try this and see for yourself. Then rebuild the design saying "Yes" to Use previous design info" — thus preserving your typing of component names, etc.

Press **Continue** to proceed to the next step in the design process. In the **Responses** droplist, choose **2**. Then enter all response **Names** and **Units** as shown below.

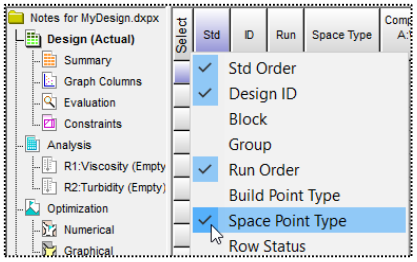


Response entries

Up to now you've been able to click the Back button at the lower right of your screen and move through the design forms to change requirements. When you press **Finish** on this page, Design-Expert completes the design setup for you.

Modify the Design

To top off this experiment design let's replicate the centroid. In the Design layout right click the **Select** column header at the upper-left corner and pick **Design ID**. Go back and also **Select** (display) the **Space Point Type** column. This is very helpful for insights about design geometry.



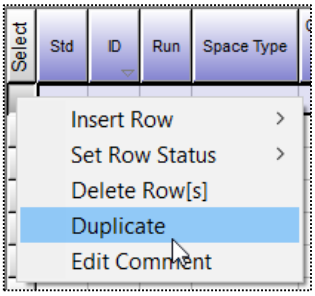
Adding columns to design layout via the Select option (right-click menu)

Next, double-click the column header labeled **Id**, to **Sort Ascending**. Now your screen should match that below except for the randomized run numbers.

Select	Std	ID	Run	Space Type	Component 1 A:Water %	Component 2 B:Alcohol %	Component 3 C:Urea %	Response 1 Viscosity mP-sec	Response 2 Turbidity
	10	0	2	Center	3.66667	2.66667	2.66667		
	11	1	4	Vertex	5	2	2		
	1	1	6	Vertex	5	2	2		
	2	2	3	Vertex	3	4	2		
	12	2	12	Vertex	3	4	2		
	13	3	7	Vertex	3	2	4		
	3	3	9	Vertex	3	2	4		
	4	4	1	CentEdge	4	3	2		
	5	5	5	CentEdge	4	2	3		
	6	6	10	CentEdge	3	3	3		
	7	7	13	AxialCB	4.33333	2.33333	2.33333		
	8	8	8	AxialCB	3.33333	3.33333	2.33333		
	9	9	11	AxialCB	3.33333	2.33333	3.33333		

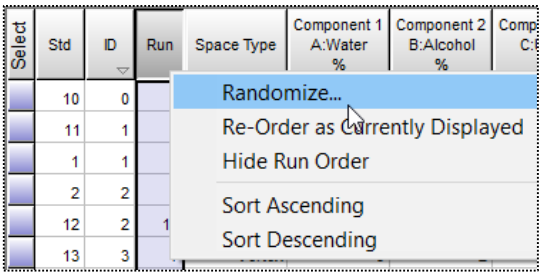
Initial design sorted by ID with point type shown (run order randomized)

The experimenters ran an additional centroid point, so in the box to the left of Id 0 (point type = "Center") right-click and select **Duplicate**.



Duplicating the centroid

Whenever you insert, delete, or duplicate rows, always right-click the **Run** column-header and choose **Randomize**.



Re-randomizing the run order

After randomization, the Run column is automatically sorted in ascending order.

☒ ***Explore what you would do next if this were your experiment***: Normally you'd now do a File, Print to produce a 'recipe' sheet to run your experiment. Go ahead if you wish or simply do a **File**, **Print Preview**.

Save Your Work

Because you've invested time into your design, it is prudent to save your work. Click **File** then **Save As**. The program displays a standard file dialog box. Use it to specify the name and destination of your data file. Enter a file name in the field with default extension **dxpx**. (We suggest **tut-mix.**) Click **Save**.

Analyze the Results

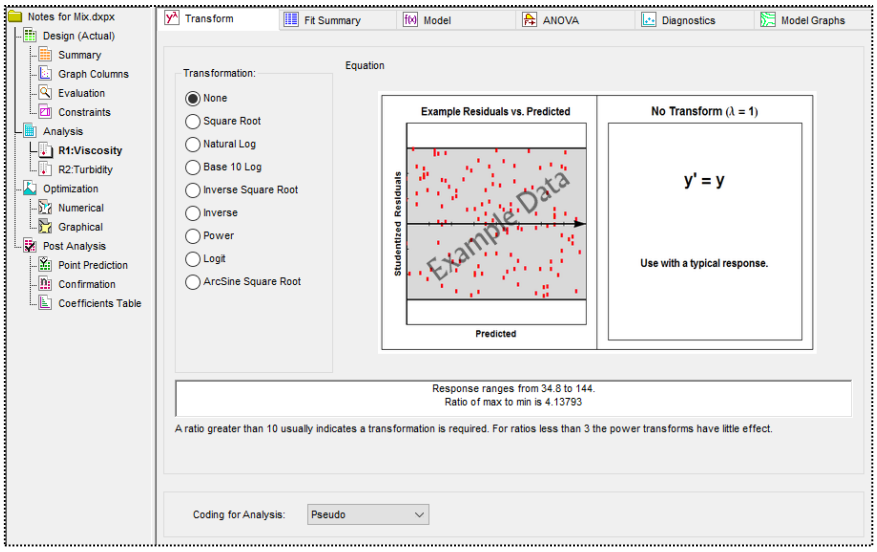
Assume your experiments are completed. You now need to enter responses into the Design-Expert software. For tutorial purposes, we see no benefit to making you type all the numbers. So to save time, read the response data in from a file that we've placed in your program's Data directory. Select **File**, **Open Design**. Click the file named **Mix.dxpx**. Press **OK**. You now should be displaying the response data shown below. (Note the design layout returns to the default selection, which we have not changed.)

There's no need for typing in this case, but normally you'd have invested much more work by this stage, so click **File** then **Save** to preserve all response data.

Run	Component 1 A:Water %	Component 2 B:Alcohol %	Component 3 C:Urea %	Response 1 Viscosity mPa-sec	Response 2 Turbidity
1	4.000	2.000	3.000	46.5	630
2	4.333	2.333	2.333	35.1	671
3	3.000	3.000	3.000	45.3	949
4	3.000	4.000	2.000	91.6	546
5	3.667	2.667	2.667	34.8	984
6	3.333	2.333	3.333	70.7	874
7	5.000	2.000	2.000	37.2	378
8	3.000	2.000	4.000	144	641
9	3.000	4.000	2.000	87.8	323
10	5.000	2.000	2.000	40.8	436
11	3.000	2.000	4.000	130	786
12	3.667	2.667	2.667	46	1122
13	4.000	3.000	2.000	67.9	436
14	3.333	3.333	2.333	51.7	730

Ready-made response data

Go to the **Analysis** branch of Design-Expert and click the **Viscosity** node. Let's progress through the tabs atop the window.



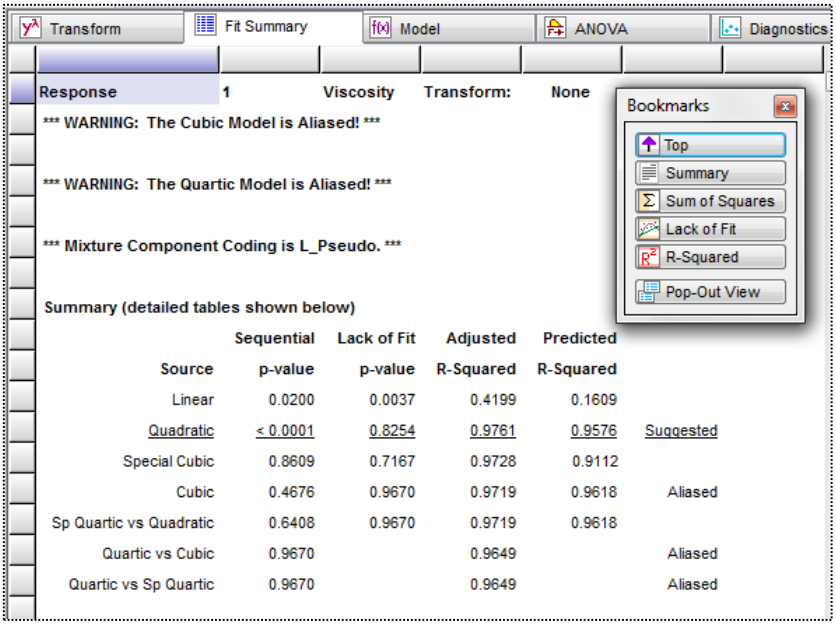
First step in analysis: Transformation options

First, consider doing a transformation on the response. In some cases this improves the statistical properties of the analysis. For example, when responses vary over several orders of magnitude, the log scale usually works best. In this case the ratio of maximum to minimum response is only a bit over 4, which isn't excessive (see detail at bottom of the screen), so leave the selection at its default, **None**, because no transformation is needed. Also, leave the coding for analysis as pseudo because this re-scales the actual component levels to 0 – 1.

☒ ***Explore details on coding for mixture models***: For complete details on pseudo and other coding for mixture models, see the textbook by Cornell mentioned at the outset of this tutorial. In the meantime, bring up **Help**, **Contents**. Then on

the Contents tab click **Mixture Designs** and **Mixture Design Details**. Select **Component Scaling in Mixture Designs**. After studying all the information you find here, close Help by pressing **X**.

Next click the **Fit Summary** tab. Here Design-Expert fits linear, quadratic, special cubic, and full cubic polynomials to the response.



Source	Sequential p-value	Lack of Fit p-value	Adjusted R-Squared	Predicted R-Squared	
Linear	0.0200	0.0037	0.4199	0.1609	
Quadratic	< 0.0001	0.8254	0.9761	0.9576	Suggested
Special Cubic	0.8609	0.7167	0.9728	0.9112	
Cubic	0.4676	0.9670	0.9719	0.9618	Aliased
Sp Quartic vs Quadratic	0.6408	0.9670	0.9719	0.9618	
Quartic vs Cubic	0.9670		0.9649		Aliased
Quartic vs Sp Quartic	0.9670		0.9649		Aliased

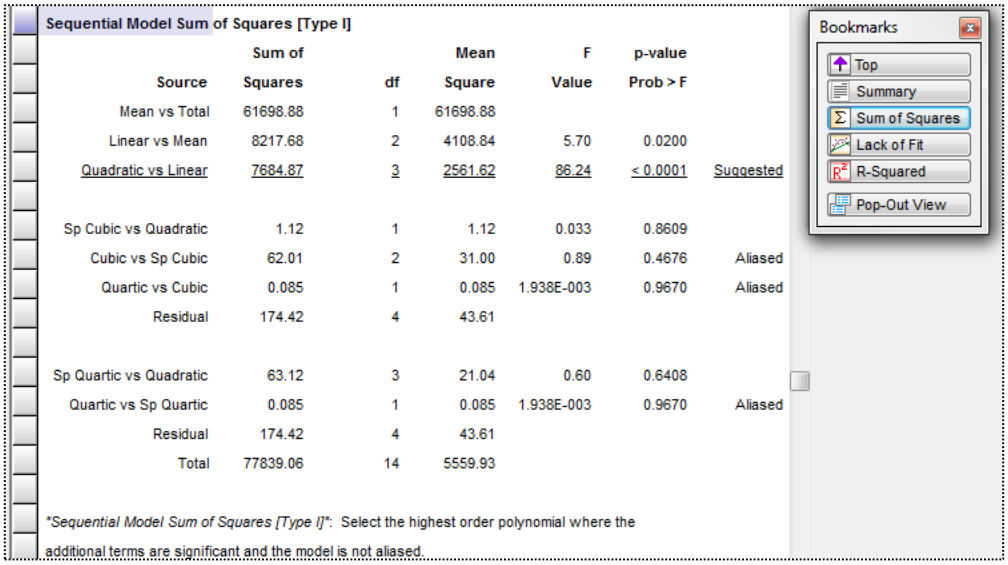
Fit summary reports

☒ ***Explore re-sizing column widths***: You may need to enlarge some columns horizontally so all outputs appear in full. If so, simply place the cursor at any column-header's right edge until the cursor changes to a double-ended arrow. Drag it to the right.

To begin your analysis, look for any warnings about aliasing. In this case, the full cubic model and beyond could not be estimated by the chosen design — an augmented simplex design. Remember, you chose only to fit a quadratic model, so this should be no surprise.

Next, pay heed to the model suggested by Design-Expert in the first table at the top, which re-caps what's detailed below.

Now on the floating **Bookmarks** press forward to the **Sum of Squares** breakdown.



Sequential Model Sum of Squares [Type I]

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Mean vs Total	61698.88	1	61698.88			
Linear vs Mean	8217.68	2	4108.84	5.70	0.0200	
Quadratic vs Linear	7684.87	3	2561.62	86.24	< 0.0001	Suggested
Sp Cubic vs Quadratic	1.12	1	1.12	0.033	0.8609	
Cubic vs Sp Cubic	62.01	2	31.00	0.89	0.4676	Aliased
Quartic vs Cubic	0.085	1	0.085	1.938E-003	0.9670	Aliased
Residual	174.42	4	43.61			
Sp Quartic vs Quadratic	63.12	3	21.04	0.60	0.6408	
Quartic vs Sp Quartic	0.085	1	0.085	1.938E-003	0.9670	Aliased
Residual	174.42	4	43.61			
Total	77839.06	14	5559.93			

Sequential Model Sum of Squares [Type I]: Select the highest order polynomial where the additional terms are significant and the model is not aliased.

Bookmarks: Top, Summary, Sum of Squares, Lack of Fit, R-Squared, Pop-Out View

Sequential models sum of squares table

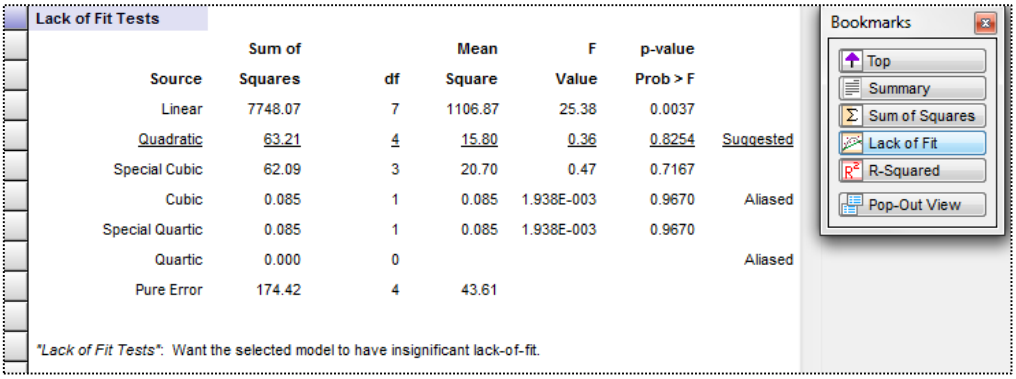
Analysis proceeds from a basis of the mean response. This is the default model if none of the factors causes a significant effect on the response. The output then shows the significance of each set of additional terms. Notice that below the level where the program says "Suggested" the p-values become insignificant (>0.05), thus there is no advantage to adding further terms.

☒ ***Explore more details on the breakdown of model sums of square***: Here's a line-by-line detailing:

- "Linear vs Mean": the significance of adding the linear terms after accounting for the mean. (Due to the constraint that the three components must sum to a fixed total, you will see only two degrees of freedom associated with the linear mixture model.)
- "Quadratic vs Linear": the significance of adding the quadratic terms to the linear terms already in the model.
- "Sp Cubic vs Quadratic": the contribution of the special cubic terms beyond the quadratic and linear terms.
- "Cubic vs Sp Cubic": the contribution of the full cubic terms beyond the special cubic, quadratic, and linear terms. (In this case, these terms are aliased.)
- And so on….

For each set of terms, probability ("Prob > F") should be examined to see if it falls below 0.05 (or whatever statistical significance level you choose). Adding terms up to quadratic significantly improves this particular model, but when you get to the special cubic level, there's no further improvement. The program automatically underlines at least one "Suggested" model. Always confirm this suggestion by reviewing all tables under Fit Summary.

On the floating **Bookmarks** tool click **Lack of Fit** to move on to the next table. This table compares residual error with pure error from replication. If residual error significantly exceeds pure error, then deviations remain in the residuals that can be removed using a more appropriate model. Residual error from the linear model shows significant lack of fit (this is bad), while quadratic, special cubic, and full cubic do not show significant lack of fit (this is good).



Lack of Fit Tests

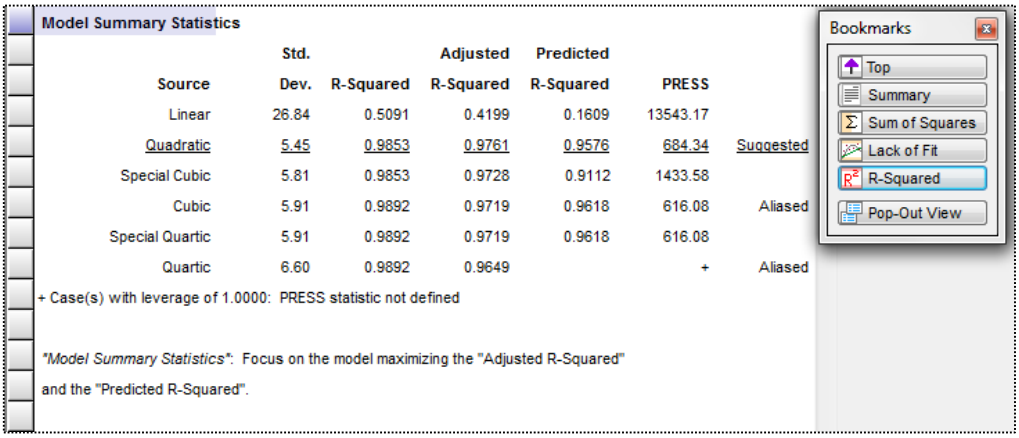
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Linear	7748.07	7	1106.87	25.38	0.0037	
Quadratic	63.21	4	15.80	0.36	0.8254	Suggested
Special Cubic	62.09	3	20.70	0.47	0.7167	
Cubic	0.085	1	0.085	1.938E-003	0.9670	Aliased
Special Quartic	0.085	1	0.085	1.938E-003	0.9670	
Quartic	0.000	0				Aliased
Pure Error	174.42	4	43.61			

*"Lack of Fit Tests": Want the selected model to have insignificant lack-of-fit.

Lack of fit table

At this point, the quadratic model statistically looks very good indeed.

Now on the floating **Bookmarks** tool click **R-Squared** to view the bottom table: "Model Summary Statistics." Here you see several comparative measures for model selection.



Model Summary Statistics

Source	Std. Dev.	R-Squared	Adjusted R-Squared	Predicted R-Squared	PRESS	
Linear	26.84	0.5091	0.4199	0.1609	13543.17	
Quadratic	5.45	0.9853	0.9761	0.9576	684.34	Suggested
Special Cubic	5.81	0.9853	0.9728	0.9112	1433.58	
Cubic	5.91	0.9892	0.9719	0.9618	616.08	Aliased
Special Quartic	5.91	0.9892	0.9719	0.9618	616.08	
Quartic	6.60	0.9892	0.9649		+	Aliased

+ Case(s) with leverage of 1.0000: PRESS statistic not defined

*"Model Summary Statistics": Focus on the model maximizing the "Adjusted R-Squared" and the "Predicted R-Squared".

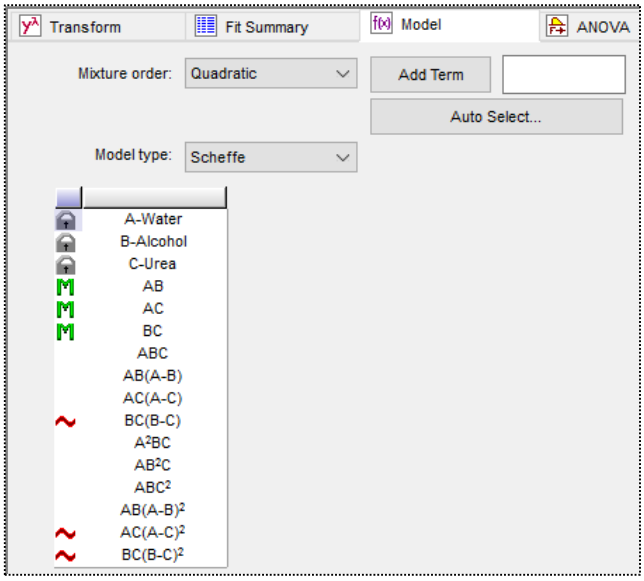
Summary statistics

Ignoring the aliased cubic model, the quadratic model comes out best: low standard deviation ("Std Dev"), high "R-Squared" statistics, and low "PRESS."

Before moving on, you may want to print the Fit Summary tables via File, Print. These tables, or any selected subset, can be cut and pasted into a word processor, spreadsheet, or any other Windows application. You're now ready to take an in-depth look at the quadratic model.

Model Selection and Analysis of Variance (ANOVA)

Click the **Model** tab atop the screen to see the model suggested by Design-Expert software.



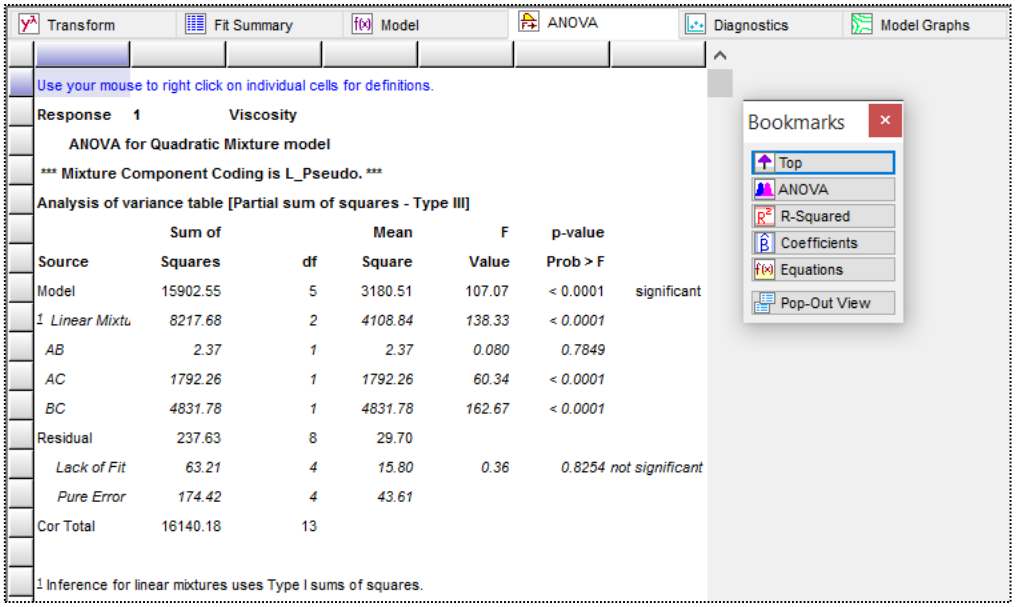
Choosing the model

☒ ***Explore how you can over-ride the model suggested by the program***: Press the screen tips button to see very helpful information about what you can do at this stage. You may select models other than this defaulting quadratic model from the pull down list. (Be sure to do this in the rare cases when Design-Expert suggests more than one model.) On the current screen you are allowed to manually reduce the model by clicking off terms that are not statistically significant. For example, in this case, you will see in a moment that the AB term is not statistically significant.

Also, as noted in the Tips screen, Design-Expert provides several automatic reduction algorithms as alternatives to "Manual" which can be accessed via the **Auto Select…** button. Click that button if you'd like to try one. You will see a recommendation pop up on what works best as a general rule. However, we recommend you not reduce mixture models unless you're sure, based on statistical and subject-matter knowledge, that this makes sense. If you really want to be competent on this, attend our Mixture Design for Optimal Formulations workshop.

Close Tips by clicking **X**.

Press the **ANOVA** button for details about the quadratic model.



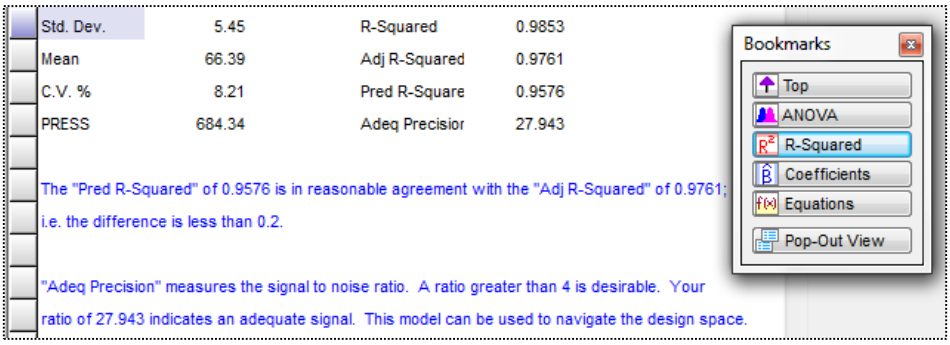
ANOVA report with annotations on

The statistics look very good.

☒ ***Explore details on the ANOVA results***: The model has a high F value and low probability values (Prob > F). That is good as you will infer from the annotation provided by Design-Expert. The probability values show the significance of each term.

P.S. Because the mixture model does not contain an intercept term, the main effect coefficients (linear terms) incorporate the overall average response and are tested together.

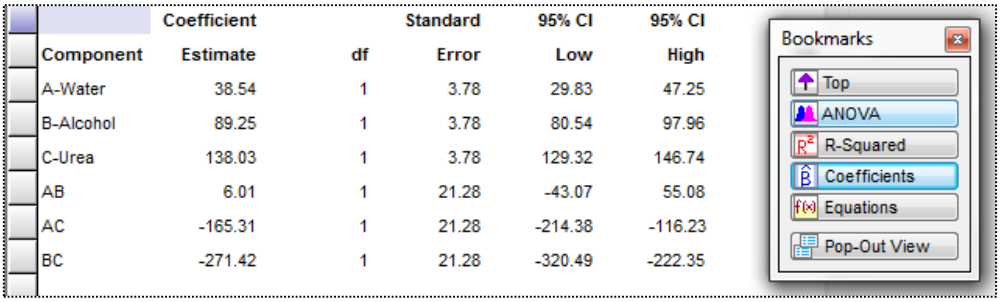
Use **Bookmarks** to jump to the next report – **R-Squared** statistics.



R-squared and other statistics after the ANOVA

These statistics, many of which you've already seen in the "Model Summary Statistics" table, all look good. Note the more than adequate precision (Adeq Precision) value of 27.943.

Next, view the coefficients and associated confidence intervals for the quadratic model.



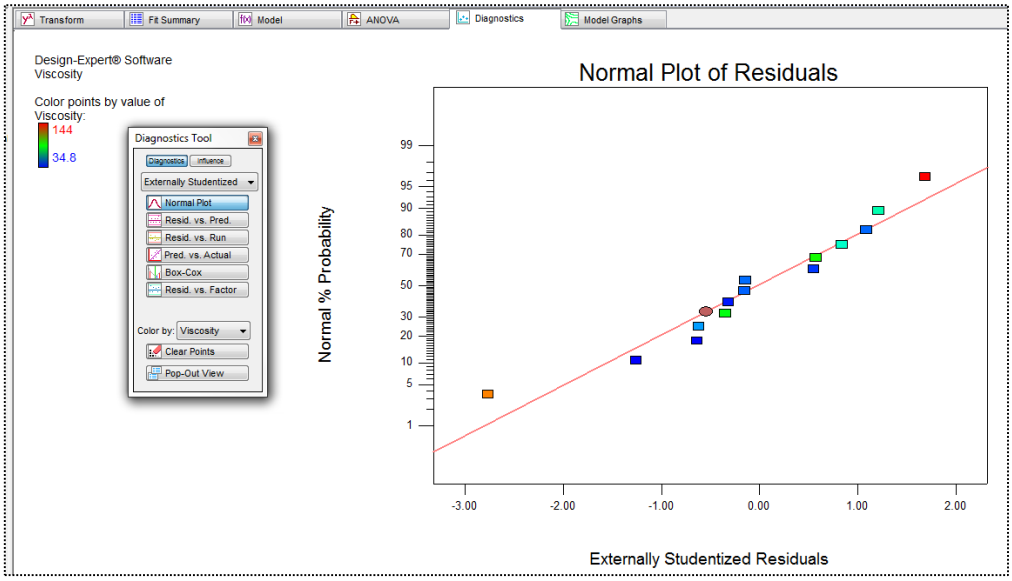
Coefficients for the quadratic model

☒ ***Explore alternative equations for response prediction***: Continue further to see several models that vary only by how components are coded. The annotations provide ideas on how they differ due to the coding.

Continue on to the next tab from ANOVA — **Diagnostics**.

Diagnose the Statistical Properties of the Model

The normal probability plot of the residuals, comes up by default.

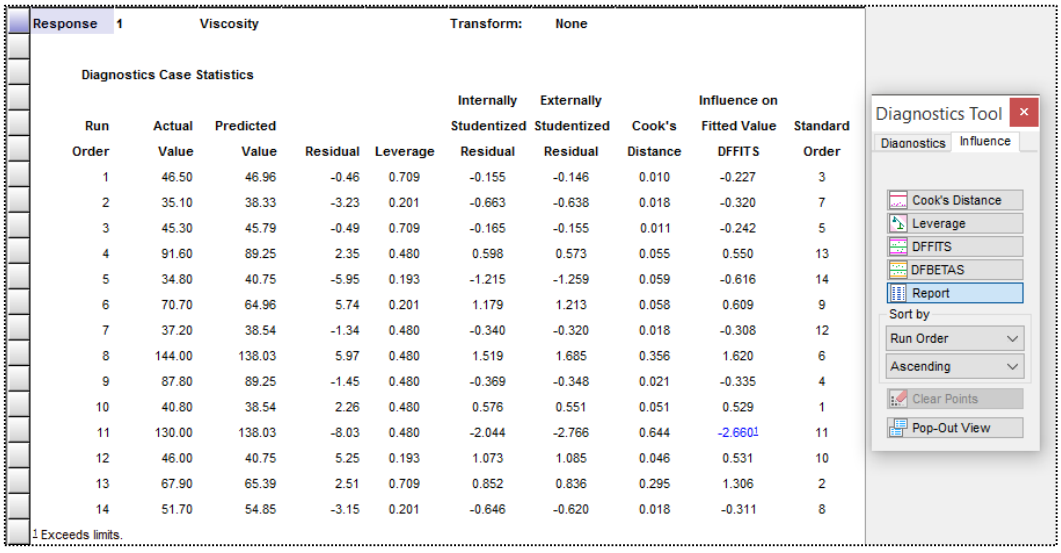


Normal Probability Plot of Residuals

The data points should be approximately linear. A non-linear pattern (such as an S-shaped curve) indicates non-normality in the error term, which may be corrected by a transformation. There are no signs of any problems in our data.

☒ ***Explore the form of residuals***: At the left of the screen you see the floating **Diagnostics Tool** palette. Be aware that residuals are externally studentized unless you elect otherwise (not advised). Studentization counteracts varying leverages due to design point locations. For example, center points carry little weight in the fit and thus exhibit low leverage. Externalizing the residuals isolates each one in comparison to the others so discrepant results stand out more.

Now click the **Influence** option. To bring up bring up case-by-case details on many of the statistics you can see on the graphs for diagnostic purposes: Press **Report**.



Response 1 Viscosity Transform: None

Diagnostics Case Statistics

Run Order	Actual Value	Predicted Value	Residual	Leverage	Internally Studentized Residual	Externally Studentized Residual	Cook's Distance	Influence on Fitted Value DFFITS	Standard Order
1	46.50	46.96	-0.46	0.709	-0.155	-0.146	0.010	-0.227	3
2	35.10	38.33	-3.23	0.201	-0.663	-0.638	0.018	-0.320	7
3	45.30	45.79	-0.49	0.709	-0.165	-0.155	0.011	-0.242	5
4	91.60	89.25	2.35	0.480	0.598	0.573	0.055	0.550	13
5	34.80	40.75	-5.95	0.193	-1.215	-1.259	0.059	-0.616	14
6	70.70	64.96	5.74	0.201	1.179	1.213	0.058	0.609	9
7	37.20	38.54	-1.34	0.480	-0.340	-0.320	0.018	-0.308	12
8	144.00	138.03	5.97	0.480	1.519	1.685	0.356	1.620	6
9	87.80	89.25	-1.45	0.480	-0.369	-0.348	0.021	-0.335	4
10	40.80	38.54	2.26	0.480	0.576	0.551	0.051	0.529	1
11	130.00	138.03	-8.03	0.480	-2.044	-2.766	0.644	-2.660	11
12	46.00	40.75	5.25	0.193	1.073	1.085	0.046	0.531	10
13	67.90	65.39	2.51	0.709	0.852	0.836	0.295	1.306	2
14	51.70	54.85	-3.15	0.201	-0.646	-0.620	0.018	-0.311	8

1 Exceeds limits.

Diagnostics tool and report

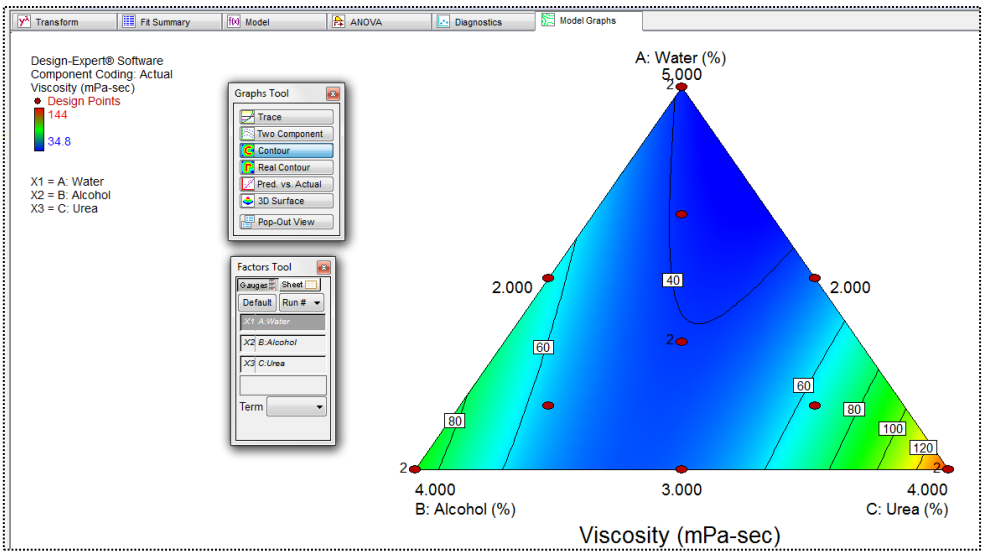
Notice that one value is flagged in blue (and with an asterisk) for exceeding suggested limits: DFFITs for standard order 11. As we discussed in the General One-Factor Tutorial (Part 2 – Advanced Features), this statistic stands for difference in fits. It measures change in each predicted value that occurs when that

response is deleted. Given that only this one diagnostic is flagged, it probably is not a cause for alarm. However, observe that it's one of the highest viscosity responses (Actual Value = 130.00), so the experimenter might want to double-check the accuracy of this response.

☒ ***Explore all the diagnostic graphs***: Check out all the graphs on either side of the tool if you like. Explanations for most of these graphs are addressed in earlier tutorials. Get more details via Screen Tips and Help. In this case, none of the graphs, excepting DFFITS, indicates any cause for alarm.

Examine Model Graphs

The residuals diagnosis reveals no statistical problems, so now let's generate response surface plots. Click the **Model Graphs** tab. The 2D contour plot comes up by default in graduated color shading.

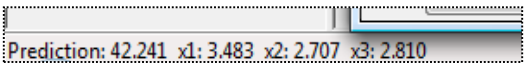


Response surface contour plot

Note that Design-Expert displays any actual point included in the design space. In this case you see a plot of viscosity as a function of the three mixture components. This slice includes two centroids as indicated by the red dot and the number "2" at the middle of the contour plot.

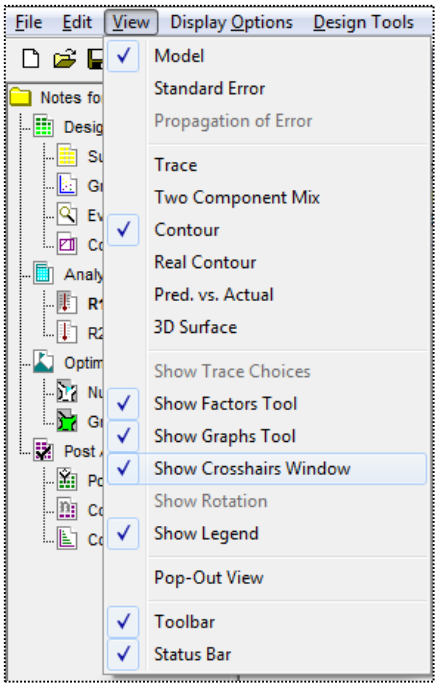
☒ ***Explore all the Factors Tool***: The Factors Tool displays along with the default plot. Move this floating tool as needed by clicking on the top border (title bar) and dragging it. The tool controls which factor(s) are plotted on the graph. The Gauges view is the default. Each component listed has either an axis label, indicating that it is currently appearing on the graph, or a red slider bar, which allows you to choose specific settings for those not currently plotted. This case study involves only three components, all of which fit on one mixture plot – a ternary diagram. Therefore, you do not see any red slider bars. If you did, they would default to the midpoint levels of the components not currently assigned to axes. You could then change a level by dragging the red slider bars left or right. If you'd like to see a demonstration of this feature, work through the Multifactor RSM Tutorial (Part 1 – The Basics).

Place your mouse cursor over the contour graph. Observe how it turns into a cross (+). Then notice in the lower-left corner of the screen that Design-Expert displays the predicted response and coordinates.



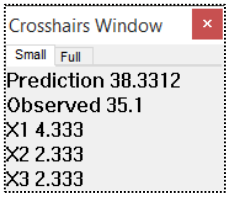
Coordinates display at lower-left corner of screen

To enable a handier tool for reading coordinates off contour plots, go to **View**, **Show Crosshairs Window**.



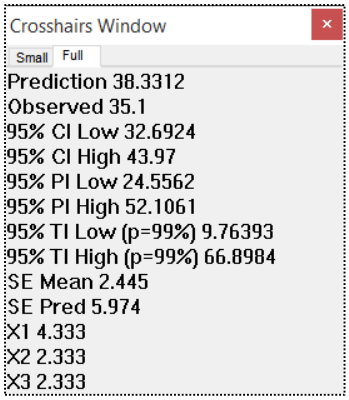
Showing crosshairs window

Now move your mouse over the contour plot and notice that Design-Expert generates the predicted response for specific values of the factors that correspond to that point. If you place the crosshair over an actual point, for example, the check-blend midway between the centroid and the upper vertex (corner labeled "A"), you also see the observed value (in this case: 35.100) as shown below.



Prediction where an actual run was performed

☒ ***Explore all the Crosshairs capability***: Press the **Full** button to see confidence and prediction intervals in addition to the coordinates and predicted response, as shown above right.

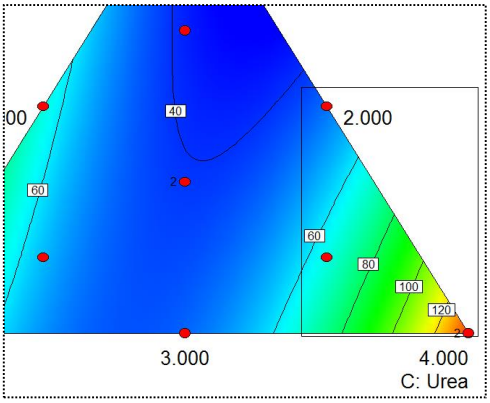


Full Crosshairs display

Close the crosshairs window by clicking **X**.

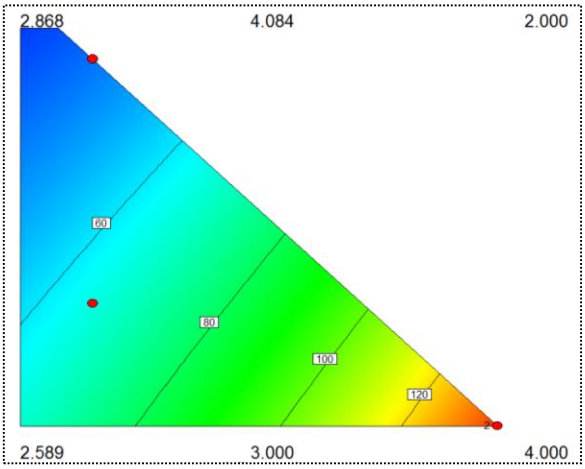
Zooming In and Out

Let's say you're interested in highest values for viscosity. With your left mouse button held down, drag over the lower right corner of the contour graph.



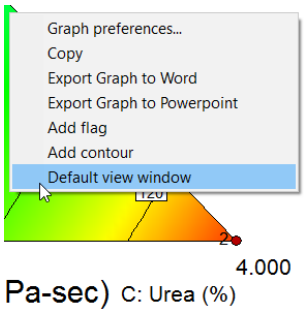
Corner identified for zoom

Now the area you chose is magnified.



Zoomed-in area on contour plot

To revert to the full triangle plot, right-click anywhere over the plot and select **Default View Window**.

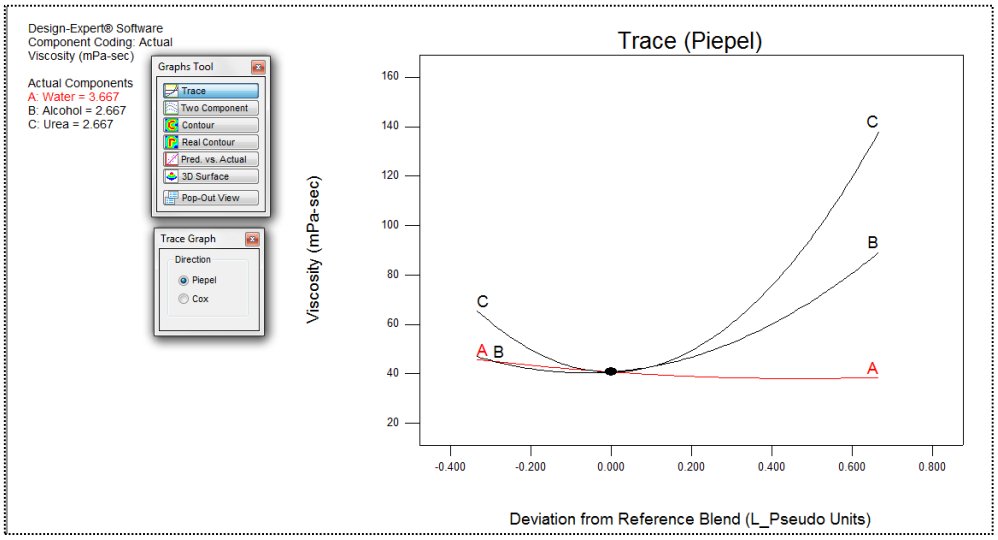


Default view window

☒ ***Explore how to make the contour plot more presentable***: Many more features are available to modify the look of contour plots. These are detailed in the Multifactor RSM Tutorial (Part 3 – Advanced Topics) Tips and Tricks for Making Response Graphs Most Presentable.

Trace Plot

Wouldn't it be handy to see all your factors on one response plot? You can do this with the trace plot, which provides silhouette views of the response surface. The real benefit from this plot is for selecting axes and constants in contour and 3D plots. From the floating Graphs Tool select **Trace**. Trace plots show the effects of changing each component along an imaginary line from the reference blend (defaulted to the overall centroid) to the vertex. For example, click on the curve for A and it changes color.



Trace plot – component A highlighted

Notice that viscosity (the response) is not very sensitive to this component.

☒ ***Explore options for trace directions***: In this case, where the experimental region forms a simplex, it matters little which direction you take. Check this out by going to the **Trace Graph** tool and pressing **Cox**. In the Cox direction, as the amount of any component increases, the amounts of all other components decrease, but their ratio to one another remains constant. Chemists may like this because it preserves the reaction stoichiometry. However, when plotted in this direction, traces for highly constrained mixture components (such as a catalyst for a chemical reaction) become

truncated. Thus, mixture-design experts argue that although it no longer holds actual ratios constant, Piepel's direction provides a more helpful plot by providing the broadest coverage of the experimental space. For this reason Piepel is the preferred plot in Design-Expert. For more detail, search in Help for "trace plot".

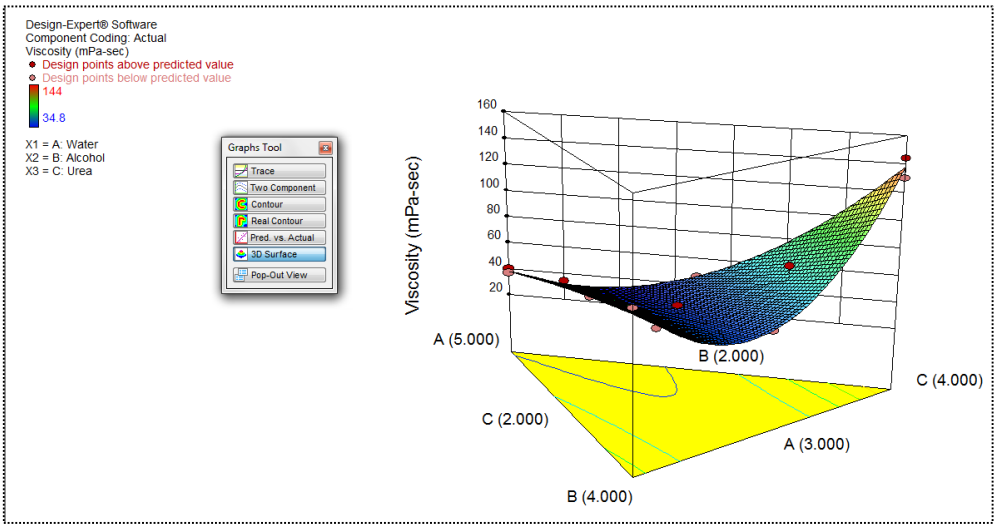
P.S. Trace plots depend greatly on where you place the starting point (by default the centroid). See for yourself by moving slide bars on the **Factors Tool**. When you are done, press the **Default**. Consider that the traces are one-dimensional only, and thus cannot provide a very useful view of a response surface. A 3D response plot provides a better picture of the surface, and ultimately provides the basis for numerical optimization. It's the ultimate tool for determining the most desirable mixture composition.

If you experiment on more than three mixture components, use the trace plot to find those components that most affect the response. Choose these influential components for the axes on the contour plots. Set as constants those components that create relatively small effects. Your 2D contour and 3D plots will then be sliced in ways that are most visually interesting.

☒ ***Explore this heads-up on how to deal with more than three components***: When you have more than three components to plot, Design-Expert software uses the composition at the optimum as the default for the remaining constant axes. For example, if you design for four components, the experimental space is a tetrahedron. Within this three-dimensional space you may find several optimums, which require multiple triangular "slices," one for each optimum.

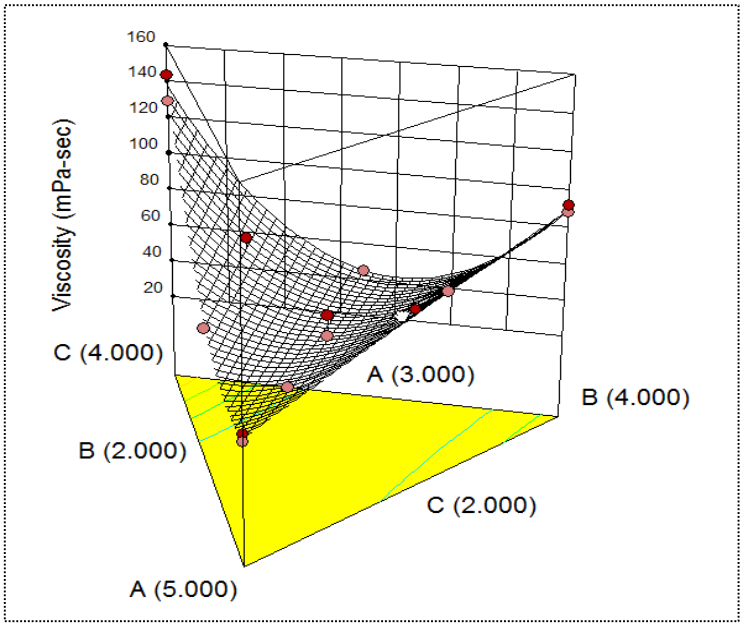
Generating a 3D View of the Response Surface

Now to really get a feel for how response varies as a function of the two factors chosen for display, select **View**, **3D Surface**. A three-dimensional display of the response surface appears. If coordinates encompass actual design points, these emerge.



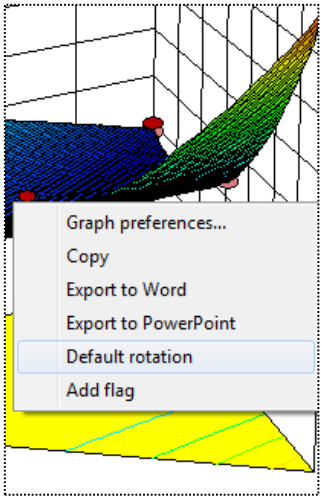
3D response surface plot

You can rotate the 3D plot directly by grabbing it with your mouse. It turns into a hand ✋ when placed over the graph. Then click and hold the left mouse-button and drag. Try it! What's really neat is how it becomes transparent so you can see hidden points falling below the surface.



Transparent view as surface is rotated, which allows points to be seen better

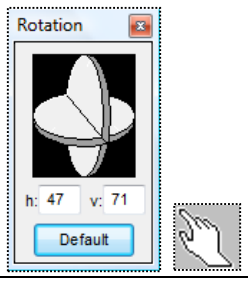
When you're done spinning the graph around, right click over it and select preferences **Default rotation**.



Right-click menu for 3D graph

The graph then re-sets to its original settings.

☒ ***Explore another way to rotate graphs***: Select **View**, **Show Rotation** for a tool that makes it easy to view 3D surface plots from any angle.



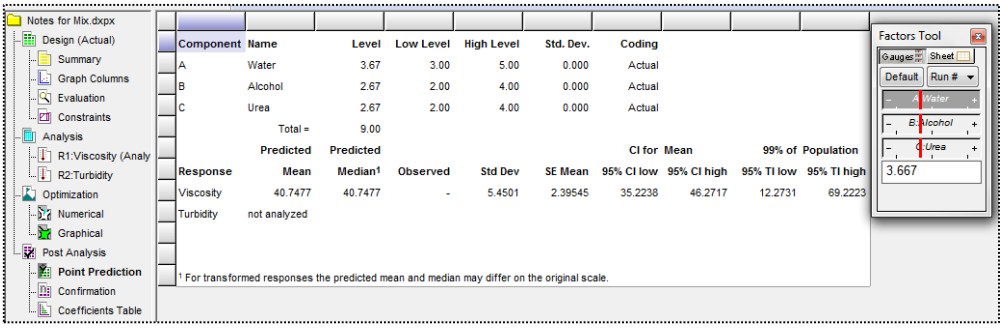
Control for rotating 3D plot

Move your cursor over the tool. The pointer changes to a hand. Now use the hand to rotate the vertical or horizontal wheel. Watch the 3D surface change. It's fun! Notice that you can specify precise horizontal ("h") and vertical ("v") coordinates. Give this a try, too. Then press **Default** and **X** off the view of Rotation tool.

Design-Expert offers many options for 3D graphs via its Graph Preferences, which come up with a right-click over the plot. For example, if you don't like graduated colors, go to the Surface Graphs tab and change 3D graph shading to wire frame view (a transparent look).

Response Prediction

Response prediction in Design-Expert software falls under the **Post Analysis** branch, which will be explored more fully in the next tutorial in this series. It allows you to generate predicted response(s) for any set of factors. To see how this works, click the **Point Prediction** node.



Point Prediction

You now see the predicted responses from this particular blend - the centroid. The **Factors Tool** opens along with the point prediction window. Move the floating tool as needed by clicking and dragging the top border. You can also drag the handy red sliders on the component gauges to view other blends. Note that in a mixture you can only vary two of the three components independently. Can you find a combination that produces viscosity of 43? (Hint: push Urea up a bit.) Don't try too hard, because in the next section of this tutorial you will make use of Design-Expert's optimization features to accomplish this objective.

☒ ***Explore a more precise way to specify component levels***: Click the **Sheet** button to get a convenient entry form for specific component values. Be careful though because the ingredients must add up to the fixed total you specified earlier: 9 wt %. Design-Expert makes adjustments as you go — perhaps in ways you do not anticipate. Don't worry: If you get too far off, simply press **Default** to return to the centroid.

Analyze the Data for the Second Response

This last step is a BIG one. Analyze the data for the second response, turbidity (Y_2). Be sure you find the appropriate polynomial to fit the data, examine the residuals, and plot the response surface. *(Hint: The correct model is special cubic.)*

Before you quit, do a **File**, **Save** to preserve your analysis. Design-Expert saves your models. To leave Design-Expert, use the File, Exit menu selection. The program warns you to save again if you've modified any files.

This tutorial gives you a good start using Design-Expert software for mixtures. We suggest you now go on to the Mixture Optimization Tutorial. You also may want to

work the tutorials about using response surface methods (RSM) for process variables. To learn more about mixture design, attend Mixture Design for Optimal Formulations, an extensive, trainer lead, workshop presented by Stat-Ease. Call or visit our web site at (www.statease.com/training/workshops.html) for a schedule.

Mixture Design Tutorial
(Part 2/2 – Optimization)

Introduction

This tutorial demonstrates the use of Design-Expert® software for optimization of mixture experiments. It's based on the data from the preceding tutorial (Part 1 – The Basics). You should go back to that section if you've not already completed it. Much of what's detailed in this Mixture Design Tutorial (Part 2 – Optimization) is a repeat of the Multifactor RSM Tutorial (Part 2 – Optimization). If you've already completed that RSM tutorial, simply skip over the areas in this tutorial that you find redundant.

☒ ***Explore how to get an in-depth knowledge of optimization tools***: For details about optimization, use the software's extensive on-screen program Help. Also, Stat-Ease provides in-depth training in its workshop titled Mixture Designs for Optimal Formulations. Call for information on content and schedules, or better yet, visit our web site at www.statease.com.

Start the program by finding and double clicking the Design-Expert software icon. To ensure being on the same page for this tutorial, go to **File** and **Open Design** file **Mix-a.dxpx**, which contains the experimental data as well as the response models.

The file you just loaded includes analyzed models as well as raw data for each response. Recall that the formulators chose a three-component simplex lattice design to study their detergent formulation. The components are water, alcohol, and urea. The experimenters held all other ingredients constant. They measured two responses: viscosity and turbidity. You will now optimize this mixture using their analyzed models.

☒ ***Explore the modeling embedded in the data file***: To see a description of the file contents, click the **Summary** node under the Design branch at the left of your screen. Drag the left border and open the window to see the report better. You can also re-size columns with your mouse. Now look at the bottom lines on responses R1 (viscosity) and R2 (turbidity). Observe under the "Model" column that these are fitted to quadratic and special cubic; respectively. That is good to know.

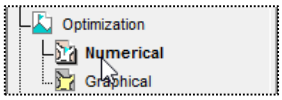
P.S. For complete details on the models fitted, go down to the bottom of the tree and click the **Coefficients Table** node under the Post Analysis branch.

Numerical Optimization

Design-Expert software's numerical optimization maximizes, minimizes, or targets:

- A single response
- A single response, subject to upper and/or lower boundaries on other responses
- Combinations of two or more responses.

We lead you through the last above case: a multiple-response optimization. Under the Optimization branch of the program, click the **Numerical** node to start the process.



Starting the Numerical Optimization

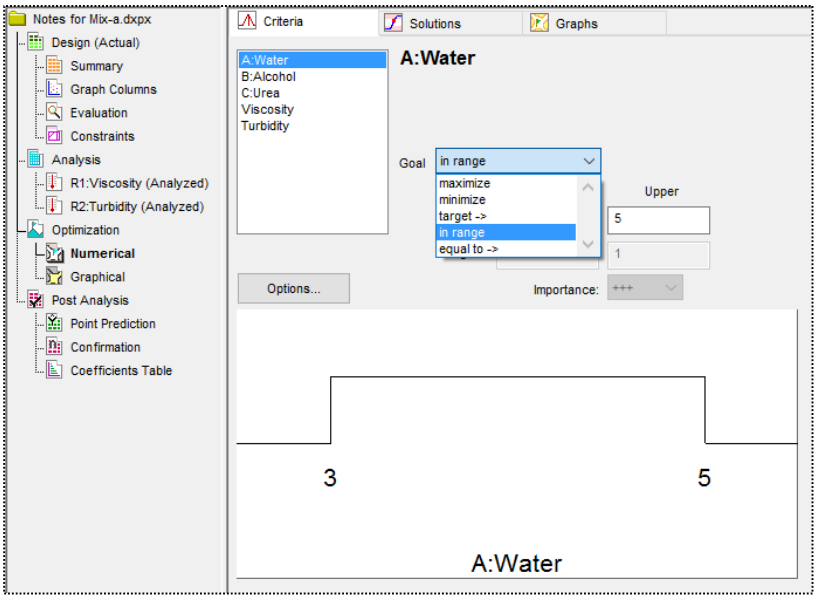
Setting the Optimization Criteria

Design-Expert allows you to set criteria for all variables, including components and propagation of error (POE). (We will get to POE later.) The limits for the responses default to the observed extremes.

☒ ***Explore the options for setting goals***: Now you reach the crucial phase of numerical optimization: assigning "Optimization Parameters." The program provides five possibilities for a "Goal" to construct desirability indices (d_i): none (to disregard any given response), maximize, minimize, target, in range (simple constraint) and equal to (components only).

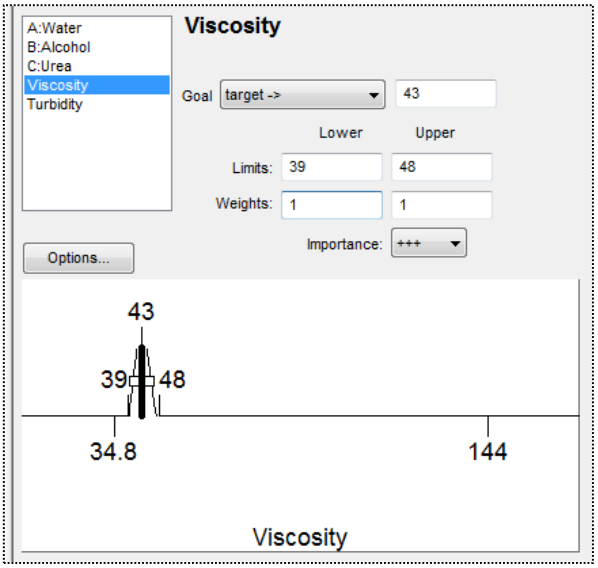
Desirabilities range from zero to one for any given response. The program combines individual desirabilities into a single number and then searches for the greatest overall desirability. A value of one represents the case where all goals are met perfectly. A zero indicates that one or more responses fall outside desirable limits. Design-Expert uses an optimization method developed by Derringer and Suich, described by Myers, Montgomery and Anderson-Cook in *Response Surface Methodology*, 3rd edition, John Wiley and Sons, New York, 2009.

In this case, components are allowed to range within their pre-established constraints, but be aware they can be set to desired goals. For example, because water is cheap, you could set its goal to maximize.



Options for goals on components

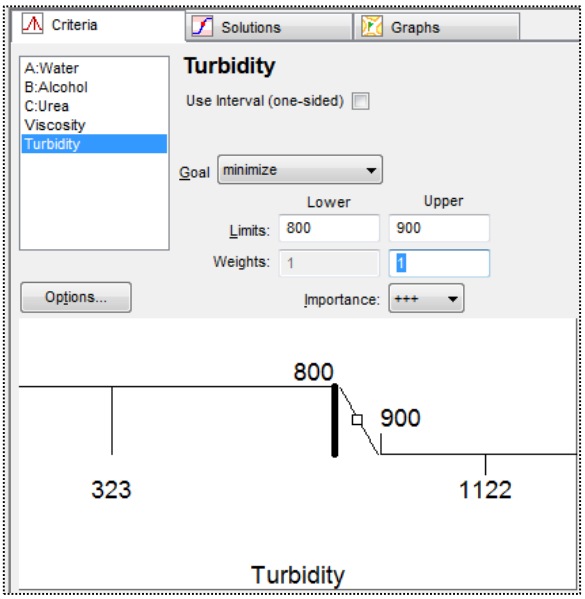
Notice that components can be set equal to specified levels. Leave water at its "in range" default and click the first response – **Viscosity**. Set its **Goal** to **target->** of **43**. Enter **Limits** as **Lower** of **39** and **Upper** of **48**. Press **Tab** to set your entries.



Setting Target for first response of viscosity

These limits indicate it is most desirable to achieve the targeted value of 43, but values in the range of 39-48 are acceptable. Values outside that range have no (zero) desirability.

Now click the second response — **Turbidity**. Select its **Goal** to **minimize**, with **Limits** set at **Lower** of **800** and **Upper** of **900**. Press **Tab** to set your entries. You must provide both these thresholds to get the desirability equation to work properly. By default they are set at the observed response range, in this case 323 to 1122. However, evidently in this case there's no advantage to getting the detergent's turbidity below 800 – it already appears as clear as can be to the consumer's eye. On the other hand, when turbidity exceeds 900, it looks as bad as it gets (too cloudy).



Aiming for minimum on second response of turbidity

These settings create the following desirability functions:

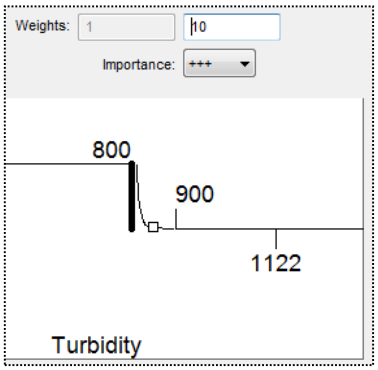
1. Viscosity:
   - if less than 39, desirability (d_i) equals zero
   - from 39 to 43, d_i ramps up from zero to one
   - from 43 to 48, d_i ramps back down to zero
   - if greater than 48, d_i equals zero.

2. Turbidity:
   - if less than 800, d_i equals one
   - from 800 to 900, d_i ramps down from one to zero
   - if over 900, d_i equals zero.

☒ ***Explore Screen Tips on numerical optimization***: Do not forget that at your fingertips you will find advice about using the sophisticated features of Design-Expert software: Press the screen tips icon for an overview about Numerical Optimization. Close out Tips by pressing **X** at the upper-right corner of its screen.

Changing Desirability Weights and the (Relative) Importance of Variables

The user can select additional parameters, called "weights," for each response. Weights give added emphasis to upper or lower bounds, or emphasize a target value. With a weight of 1, d_i varies from 0 to 1 in linear fashion. Weights greater than 1 (maximum weight is 10) give more emphasis to goals. Weights less than 1 (minimum weight is 0.1) give less emphasis to goals. Weights can be quickly changed by 'grabbing' (via left mouse-click and drag) the handles (the squares ▫) on the desirability ramps. Try pulling the handle on the ramp down as shown below.



Weights change by grabbing handle with mouse

Notice that Weight now reads 10. You've made it much more desirable to get near the turbidity goal of 800. Before moving on, re-enter **Upper Weights** to its default value of **1** and press the **Tab** key. This straightens the desirability ramp.

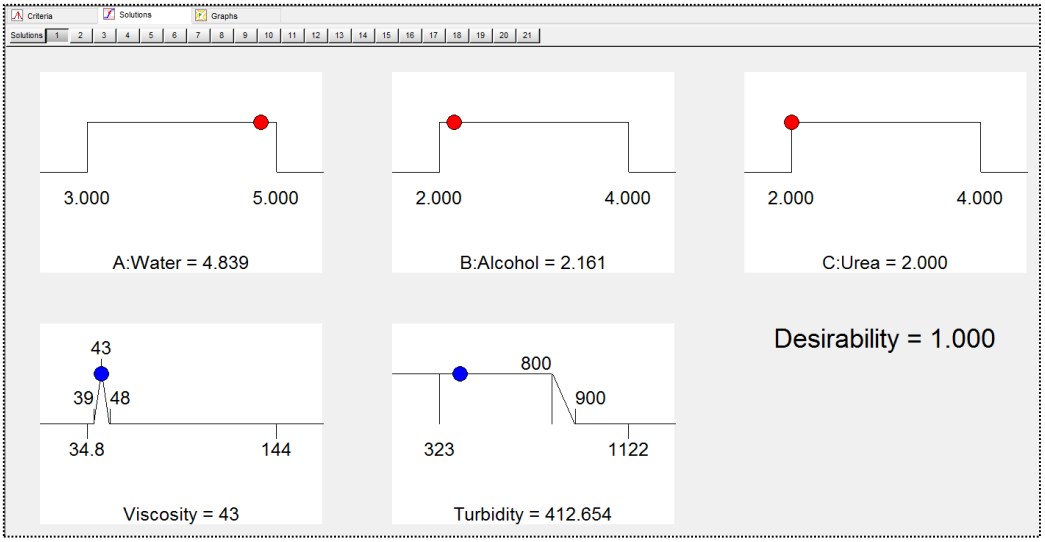
"Importance" is a tool for changing relative priorities for achieving goals you establish for some or all of the variables. If you want to emphasize one variable over the rest, set its importance higher. Design-Expert offers five levels of importance ranging from 1 plus (+) to 5 pluses (+++++). For this study, leave **Importance** at **+++**, a medium setting. By leaving all importance criteria at their defaults, none of the goals is favored over any other.

☒ ***Explore Help on numerical optimization***: For an in-depth explanation of constructing desirability functions, and formulas for weights and importance, select **Help** from the main menu. Then go to **Contents** and select **Optimization**

Node, then expand into the two **Numerical Optimization** topics — **"Desirability Details"** and **"Desirability Function"**. When you finish viewing Help, close the screen by pressing **X** at the upper-right corner of its screen.

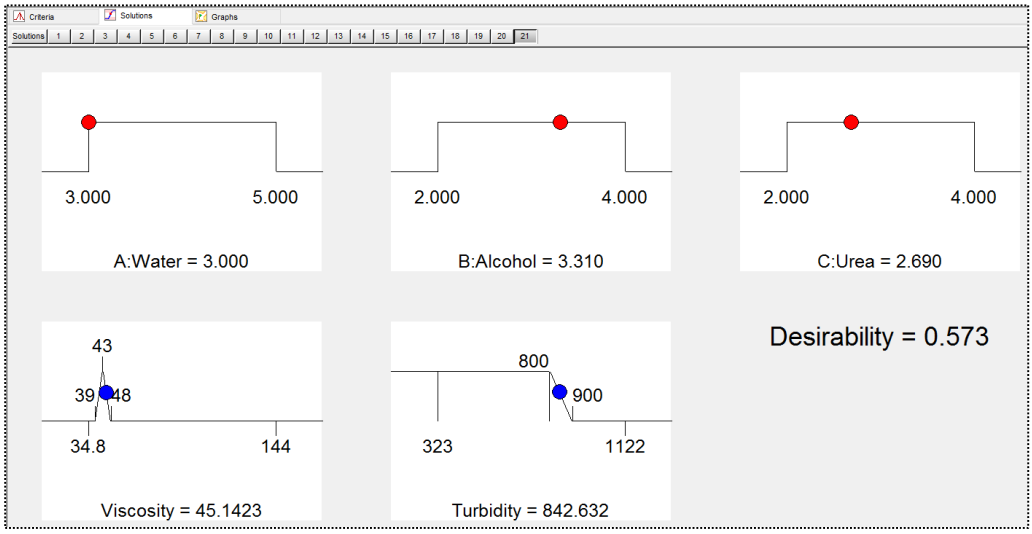
Running the optimization

Start the optimization by clicking the **Solutions** tab. Design-Expert brings up the Ramps view by default.



Ramps report on numerical optimization (Your results may differ)

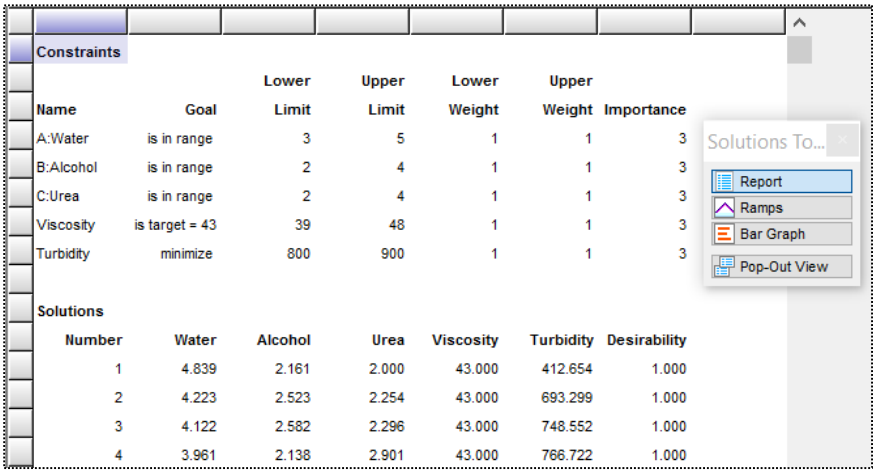
The ramp display combines individual graphs for easier interpretation. The dot on each ramp reflects the factor setting or response prediction for that solution. The height of the dot shows how desirable it is. Press the different solution buttons (1, 2, 3,…) and watch the dots. The red ones representing the component levels move around quite a bit, but do the responses remain within their goals (desirability of 1)? Near the graph's top, click the last solution (solution 21 in this case) on your screen. Does your solution look something like the one below?



Sub-optimum solution that ranks least desirable

If your search also uncovered the above local optimum, note that viscosity falls off target and turbidity becomes excessive, thus making it less desirable than the option for higher temperature.

Go back a step up on the **Solutions Tool** by pressing for the **Report**.



Constraints

Name	Goal	Lower Limit	Upper Limit	Lower Weight	Upper Weight	Importance
A:Water	is in range	3	5	1	1	3
B:Alcohol	is in range	2	4	1	1	3
C:Urea	is in range	2	4	1	1	3
Viscosity	is target = 43	39	48	1	1	3
Turbidity	minimize	800	900	1	1	3

Solutions To...
Report
Ramps
Bar Graph
Pop-Out View

Solutions

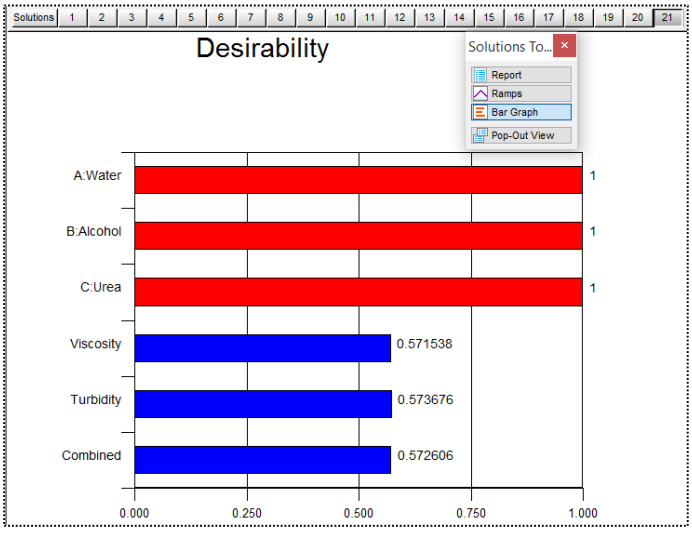
Number	Water	Alcohol	Urea	Viscosity	Turbidity	Desirability
1	4.839	2.161	2.000	43.000	412.654	1.000
2	4.223	2.523	2.254	43.000	693.299	1.000
3	4.122	2.582	2.296	43.000	748.552	1.000
4	3.961	2.138	2.901	43.000	766.722	1.000

Numerical Optimization Report on Solutions (Your results may differ)

The report starts with a recap of your optimization specifications. Then it lists solutions in order of desirability. It ends with detailing of the starting points for the search.

☒ ***Explore the Report***: Scroll down this report to see how where the program starts its empirical searches for desirable results. Multiple cycles improve the odds of finding multiple local optimums, some of which will be higher in desirability than others. In this case Design-Expert grinds through 110 cycles of optimization, starting from the 10 design points plus 100 more at random.

Go back now to the **Solutions Tool** and select the **Bar Graph**.

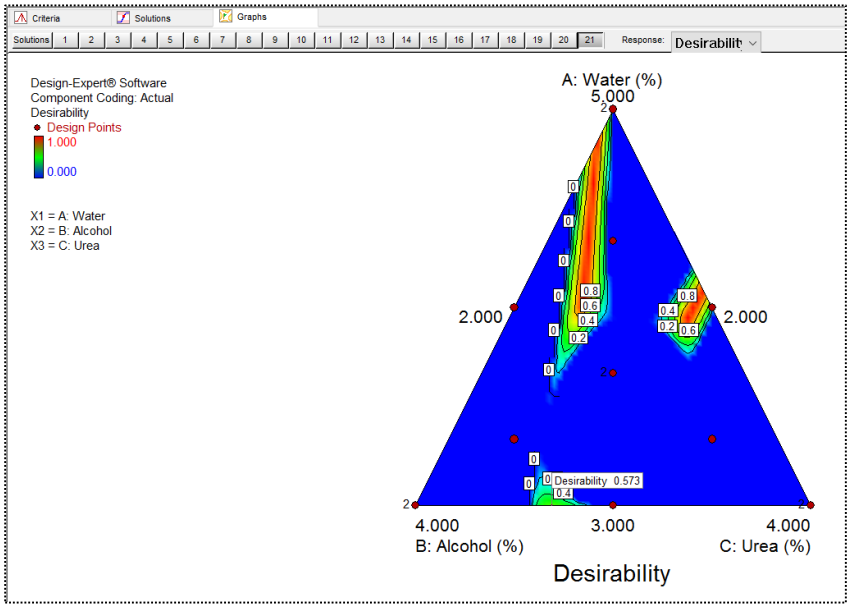


Solution to multiple-response optimization – desirability bar graph

The above bar graph shows how well each variable satisfies the criteria and the overall combined desirability: Values near one are good. This is not the best solution!

Optimization Graphs

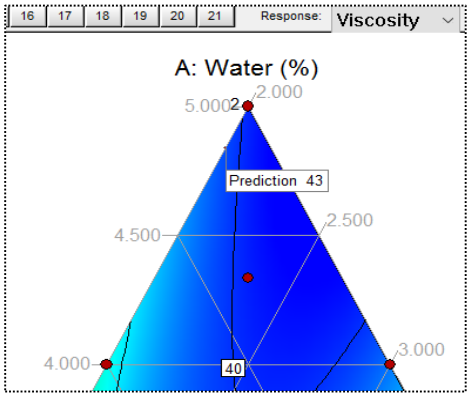
Pressing the **Graphs** tab brings up the graphs of "All Responses", including the desirability function. Select **Desirability** from the droplist to view a contour graph of overall desirability. It now becomes obvious that at least somewhat desirable formulations fall with three distinct 'sweet spots' as indicated by the three graduated color areas within the blue background.



Desirability contour graph

The screen shot above came from a graph done showing graduated colors — cool blue for lower desirability and warm yellow for higher. Design-Expert software sets a flag at the optimal point for solution 21 (or whichever one is your worst).

Now click back through the numbered Solutions choices atop your screen (or use the left arrow once you've clicked) until the flag relocates to the largest sweet spot (the one with the largest area) at the top of the triangular mixture space. To view the responses associated with this desirability (sweet spot), press the droplist arrow for **Response** and select **Viscosity**.

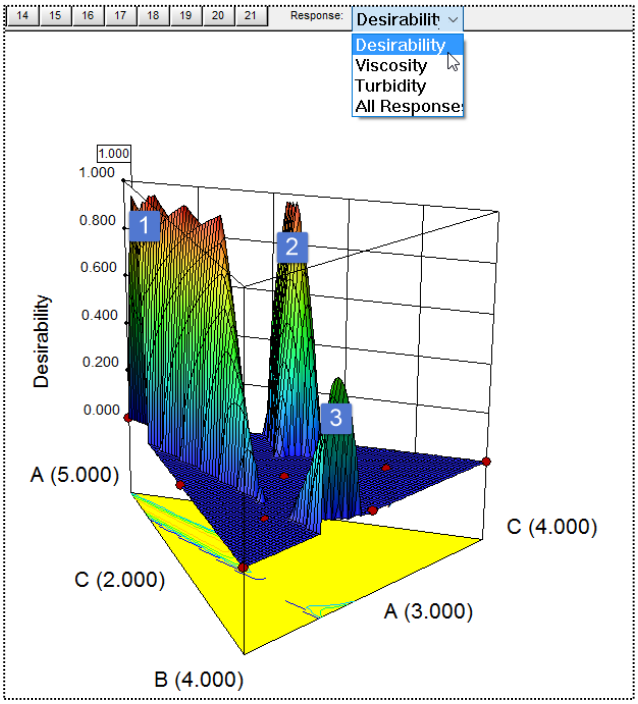


Most desirable point flagged (grid lines added — see sidebar to explore this)

☒ ***Explore how to add grid lines***: Right-click the graph and select **Graph preferences**, go to the **Surface Graphs** tab and check on **Show contour grid lines**. The gridlines appear in the plot above. There are many other options on this and other Graph preferences tabs. Look them over if you like and then press **OK** to see how options specified by this tutorial affect your contour plot. If you like, look at the optimal turbidity response as well.

P.S. For tutorial purposes, go back and press **Default** on all Graph Preference tabs to re-set the original layouts.

To view the desirability surface in three dimensions, again click **Response** and choose **Desirability**. Then from the floating Graphs Tool select **3D Surface**.



3D view of desirability at default resolution in color

Now you can see one high ridge (1) where desirability can be maintained at a maximum level over a range of compositions. Another high point (2) can be achieved, but it requires sharp control of the composition. The other peak (3) is less desirable (lower).

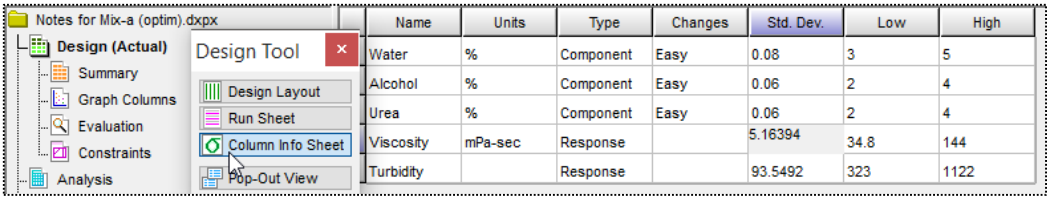
☒ ***Explore going to very high graph resolution***: So as not to tax user's computers, Design-Expert defaults short of maximum resolution. Try smoothing out the 3D desirability surface via a right-click over the graph, selecting **Graph Preferences** and then on the **Surface Graphs** tab changing the **Graph resolution** to **Very High**. Press **OK** for the new graph preferences. The go back and re-set things to the **Default**.

Adding Propagation of Error (POE) to the Optimization

If you have prior knowledge of the variation in your component amounts, this information can be fed into Design-Expert software. Then you can generate propagation of error (POE) plots showing how that error transmits to the response. Look for compositions that minimize transmitted variation, thus creating a formula that's robust to slight variations in the measured amounts.

Start by clicking the **Design** node on the left side of the screen to get back to the design layout. Select **Column Info Sheet** from the floating **Design Tool** palette.

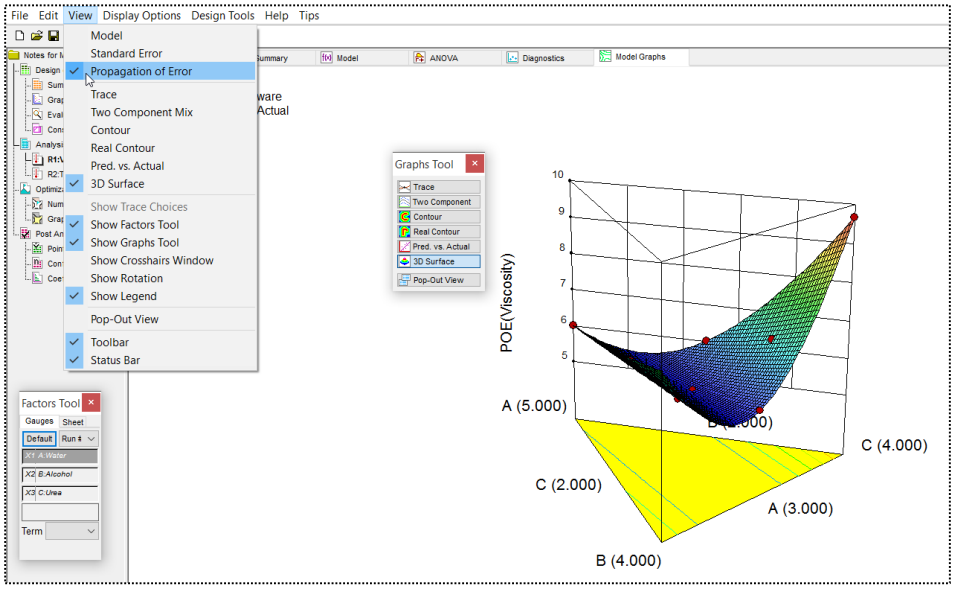
Enter the following information into the **Std. Dev.** column: Water: **0.08**, Alcohol: **0.06**, Urea: **0.06**, as shown on the screen below.



Name	Units	Type	Changes	Std. Dev.	Low	High
Water	%	Component	Easy	0.08	3	5
Alcohol	%	Component	Easy	0.06	2	4
Urea	%	Component	Easy	0.06	2	4
Viscosity	mPa-sec	Response		5.16394	34.8	144
Turbidity		Response		93.5492	323	1122

Column Info Sheet with standard deviations filled in

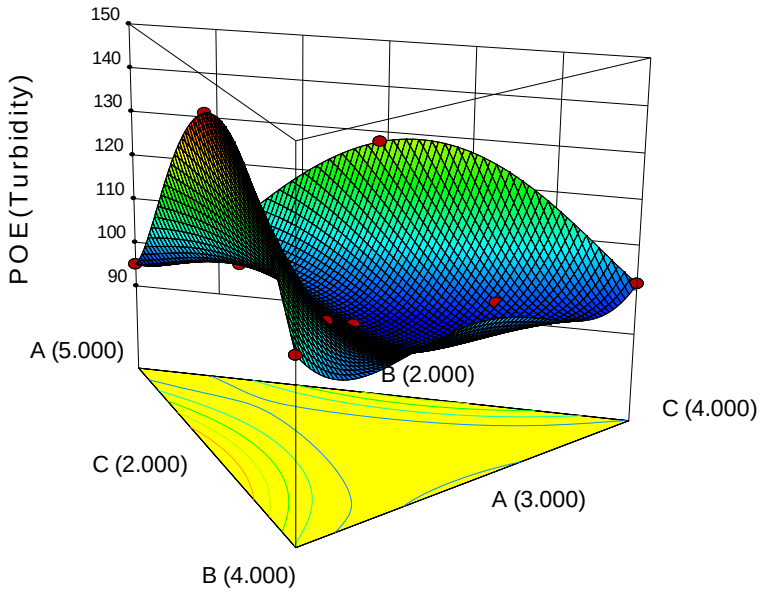
Now you can calculate propagation of error by generating graphs for each response. First, click the **Viscosity** analysis node and press the **Model Graphs** tab. Next, select **View**, **Propagation of Error**, which previously was grayed out. Also choose **3D Surface** view. Now your screen should match what's shown below.



3D view of the POE graph

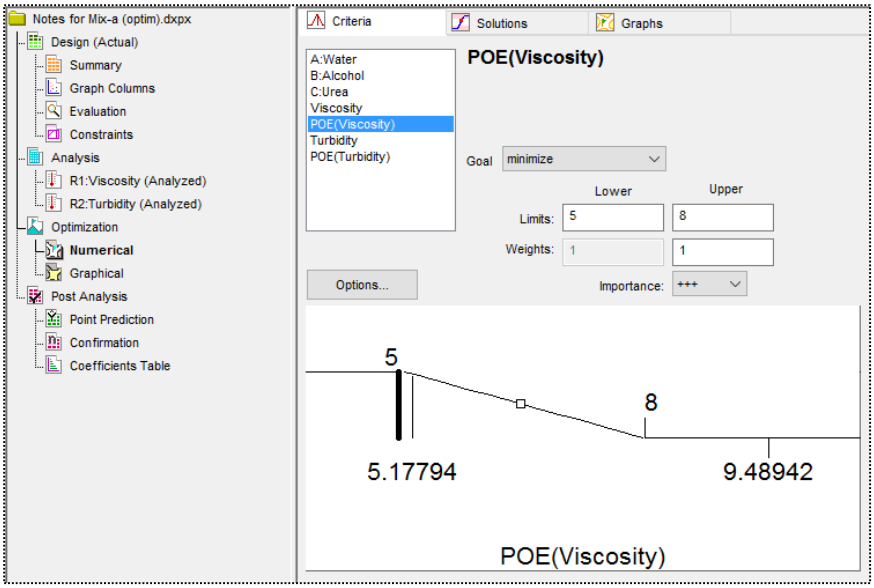
The surface reaches a minimum where the least amount of error is transmitted (propagated) to the viscosity response. These minima occur at flat regions on model graphs where formulations are most robust to varying amounts of components.

Click the **Turbidity** node, press the ***Model Graphs*** button and select **View**, **Propagation of Error** and look at its **3D Surface**. Rotate it so you can see the surface best.



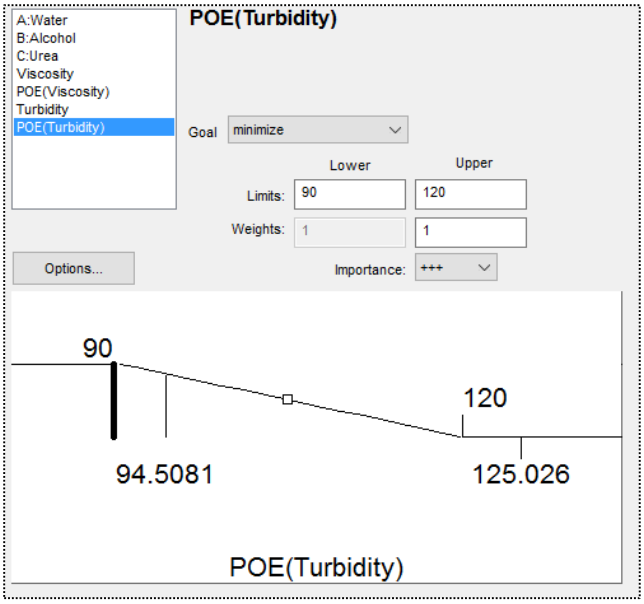
POE surface for turbidity

Now that you've found optimum conditions for the two responses, let's go back and add criteria for the propagation of error. Click the **Numerical** optimization node. Select **POE (Viscosity)** and establish a **Goal** to **minimize** with **Limits** of **Lower** at **5** and **Upper** of **8**.



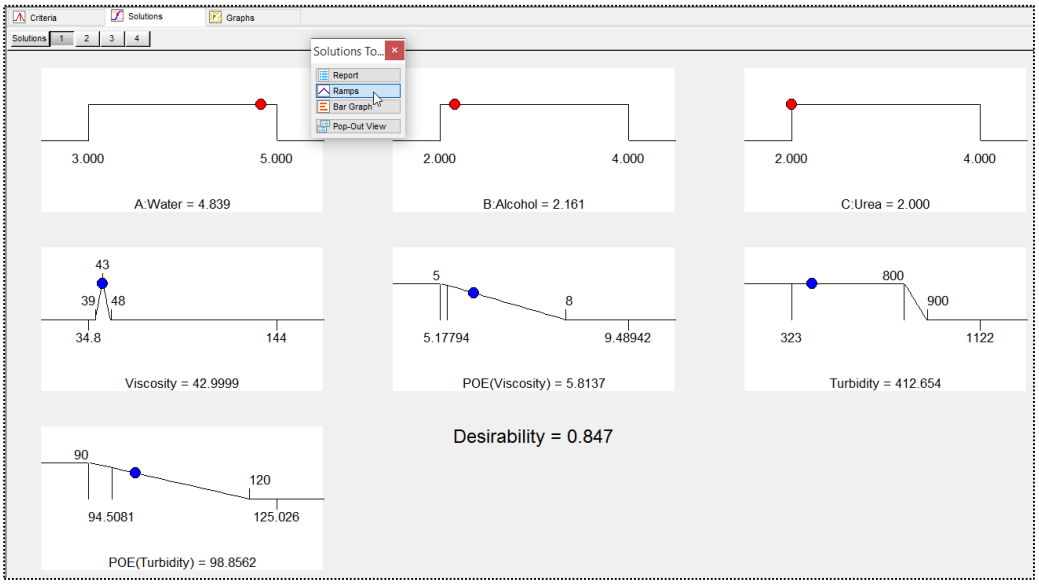
Set goal and limits for POE (Viscosity)

Select **POE (Turbidity)** and set its **Goal** also to **minimize** with **Limits** of **Lower** at **90** and **Upper** of **120**.



Criteria for POE (Turbidity)

Now click the **Solutions** tab to generate new solutions with the additional criteria. (You may need to press Ramps on the Solutions Tool to get the view shown below.)



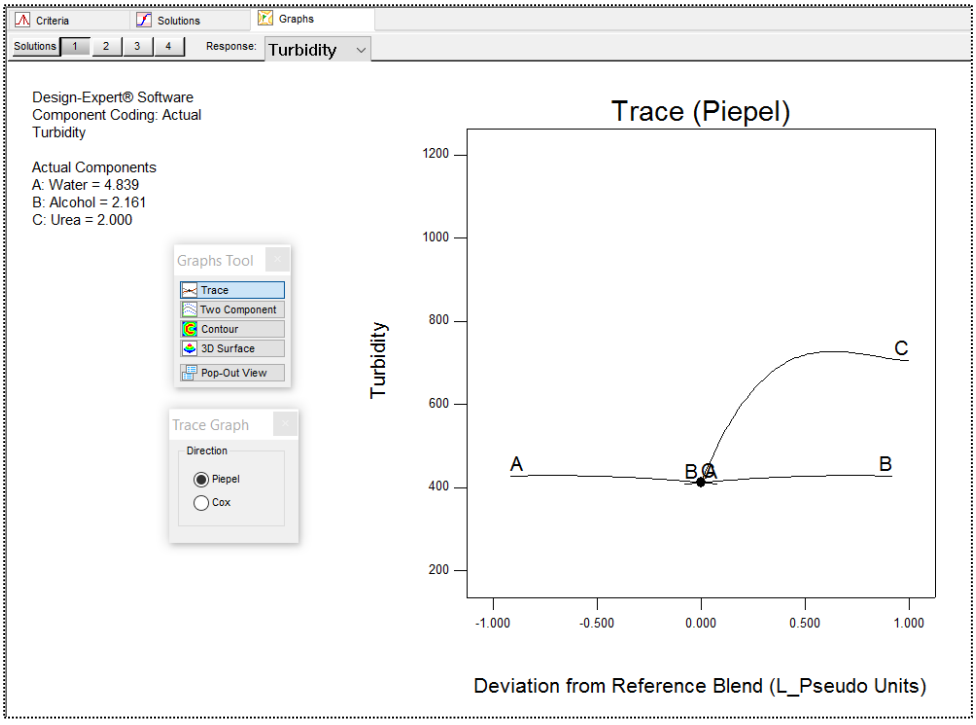
Solutions Generated with Added POE Criteria (Your results may differ)

The number 1 solution represents the formulation that best achieves the target value of 43 for viscosity and minimizes turbidity, while at the same time finds the spot with the minimum POE (most robust to slight variations in the component amounts).

☒ ***Explore alternative solutions***: If you can take the time, review the alternative solutions, which may be nearly as good based on the criteria you entered. There may be some alternative solutions that make better tradeoffs among the mutual goals.

Viewing Trace Plots from Optimal Point

Continue on to the numerical optimization **Graphs** to look at the "All Responses" plots. Here, you get a bird's eye view of all of the responses and how the solution was arrived at. Note that the Desirability plot does not look much different from before because adding POE criteria had only a small impact on the result. However, this is a good time to get a feel for the sensitivity of responses around the optimum point. Observe this by changing **Response** to **Turbidity**. Then select **Trace** from the **Graphs Tool** palette.



Trace plot viewed from optimal point (remember, your optimum may differ slightly)

Now you see that changing component A (water) and B (Alcohol) makes little difference on this response, whereas its very dependent on C (urea).

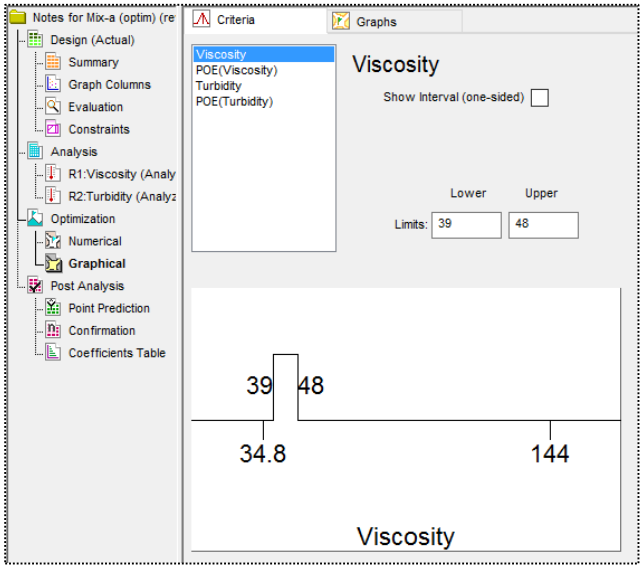
Take a look at the trace for the other response — turbidity. It looks even more interesting!

Graphical Optimization

By shading out regions that fall outside of specified contours, you can identify desirable sweet spots for each response – windows of opportunity where all specifications can be met. In this case, response specifications are:

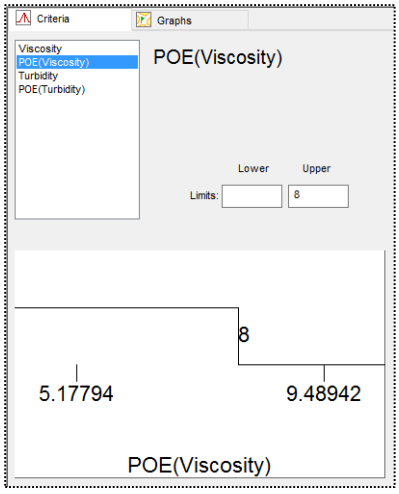
- 39 < Viscosity < 48
- POE (Viscosity) < 8
- Turbidity < 900
- POE (Turbidity) < 120

To overlay plots of all these responses, click the **Graphical** optimization node. For the **Viscosity** response, if the following values are not already pre-set, enter a **Lower** limit of **39** and an **Upper** limit of **48**.



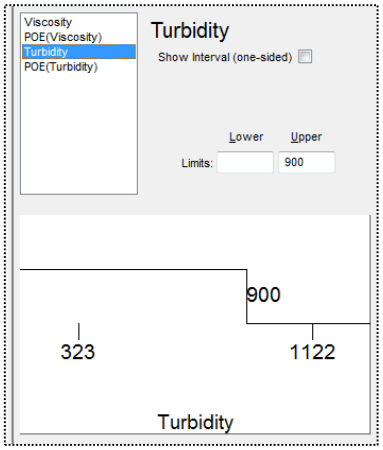
Setting criteria for Graphical optimization: Viscosity response

Click the **POE(Viscosity)** response. If the following value is not already pre-set, enter an **Upper** limit of **8**. Do not enter a lower limit — it will not be needed for the graphical optimization when simply minimizing.



Graphical criterion for POE of viscosity

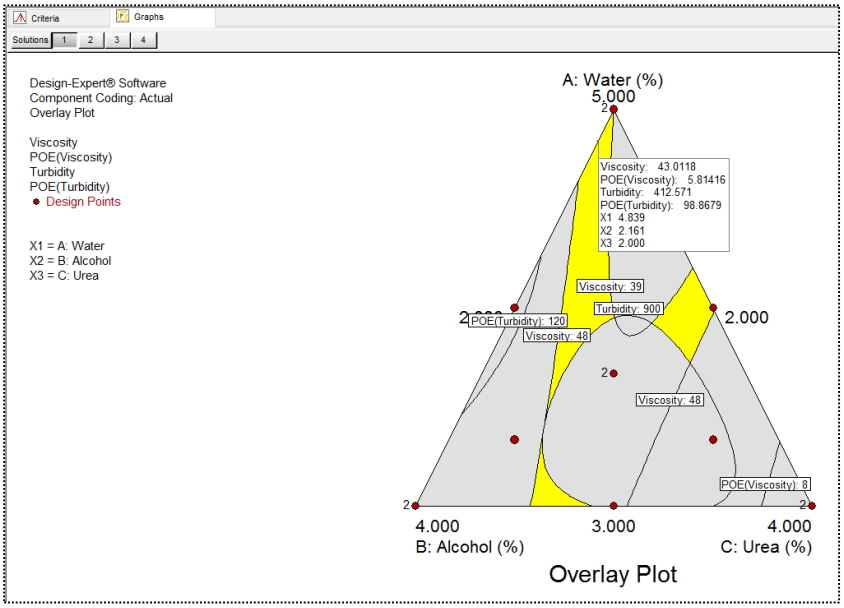
Press forward to the **Turbidity** response and, if the following value is not already pre-set, enter an **Upper** limit of **900**. This again is a minimization, so don't enter a lower limit.



Setting criteria for turbidity

Click the **POE(Turbidity)** response and, if the following value is not already pre-set, enter an **Upper** limit of **120**.

Press the **Graphs** tab to produce the "overlay" plot.



Graphical optimization

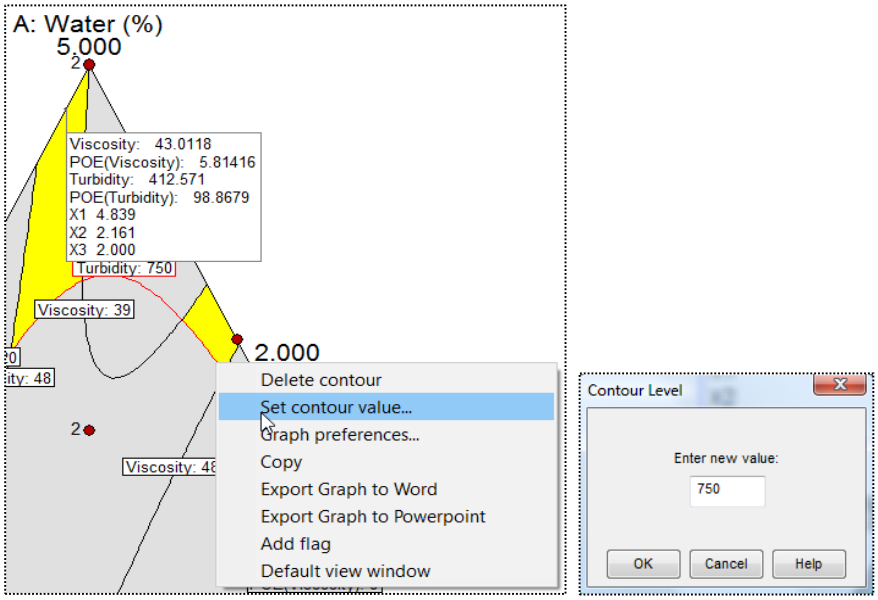
Notice that regions not meeting your specifications are grayed out, leaving (hopefully!) an operating window or "sweet spot."

Notice the flag remains planted at the optimum. That's handy! This Design-Expert display may not look as fancy as 3D desirability, but it is very useful to show windows of operability where requirements simultaneously meet critical properties. Grayed areas on the graphical optimization plot do <u>not</u> meet selection criteria. The clear "window" shows where you can set factors to satisfy requirements for both responses.

The lines that mark the high or low boundaries on the responses can be identified with a mouse-click. Notice that the contour and its label change color for easy

identification. Click outside the graph to reset the contour and its label to the original color.

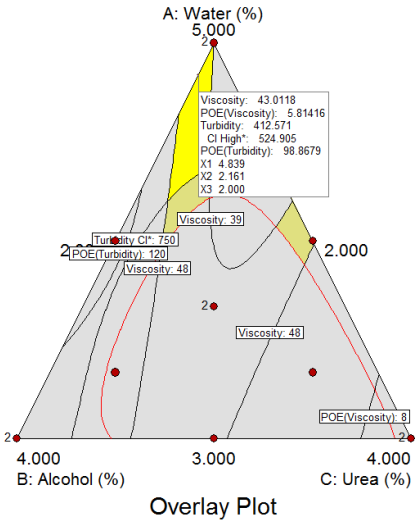
Let's say someone wonders whether the 900 maximum for turbidity can be decreased. What will this do to the operating window? Find out by clicking the 900 turbidity contour line – you know you've got it when it turns red. Then drag the contour until it reaches a value of approximately **750**. Finally right-click over this contour, select **Set contour value** and enter **750**.



Setting the turbidity contour value

Press **OK** to get the 750 contour level. Notice the smaller sweet spot has disappeared and the medium one considerably reduced in area. To reset the original sweet spot, go back to **Criteria** and reset **Turbidity** to an **Upper** limit of **900**.

☒ ***Explore adding uncertainty intervals around your window of operability***: For **Turbidity** click **Show Interval (one-sided)** and then press forward to **Graphs**. This pushes in the boundary by the confidence interval, thus accounting for uncertainty in the mean prediction based on the model derived from this mixture experiment.



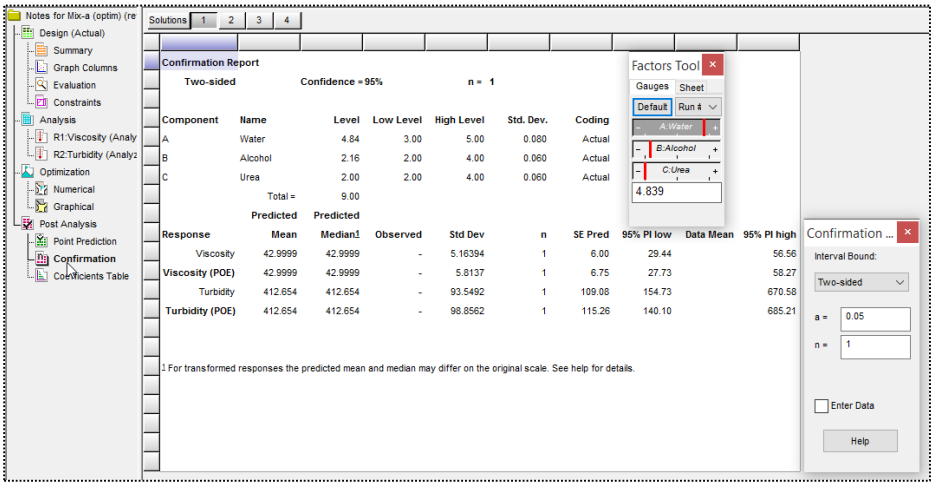
Confidence intervals (CI) superimposed on turbidity

If you are subject to FDA regulation and participate in their quality by design (QBD) initiative, the CI-bounded window provides a relatively safe operating region — a functional design space — for any particular unit operation. Manufacturing design space requires tolerance intervals. This tutorial experiment provided too few runs to support adding the CI for viscosity, much less the imposition of TIs. Learn how to size designs properly for manufacturing QBD by attending our Designed Experiments for Pharma workshop.

Graphical optimization works great for three components, but as the number increases, it becomes more and more tedious. Once you find solutions much more quickly by using the numerical optimization feature, return to the graphical optimization and produce outputs for presentation purposes.

Response Prediction at the Optimum

Click the **Confirmation** node (near bottom left on your screen). Notice it defaults to your first solution.



Confirmation set to Solutions 1 (yours may be different)

This defaults to the prediction interval (PI) for a single point.

☒ ***Explore the Confirmation feature***: You had best replicate the optimal formulation six or so times and then click the Enter Data option to type these into Design-Expert. It then computes the Data Mean and puts this in the middle of the PI values for evaluation. Try entering some numbers for yourself and see what happens.

Save the Data to a File

Now that you've invested all this time into setting up the optimization for this design, it is wise to save your work. Click the **File** menu item and select **Save As**. You can now specify the **File name** (we suggest **tut-MIX-opt**) to **Save as** type *.dxpx" in the Data folder for Design-Expert (or wherever you want to Save in).

Final Comments

We feel that numerical optimization provides powerful insights when combined with graphical analysis. Numerical optimization becomes essential when investigating many components with many responses. However, computerized optimization does not work very well in the absence of subject-matter knowledge.

For example, a naive user may define impossible optimization criteria. The result will be zero desirability everywhere! To avoid this, try setting broad, acceptable ranges. Narrow down the ranges as you gain knowledge about how changing factor levels affect responses. Often, you will need to make more than one pass to find the "best" factor levels that satisfy constraints on several responses simultaneously.

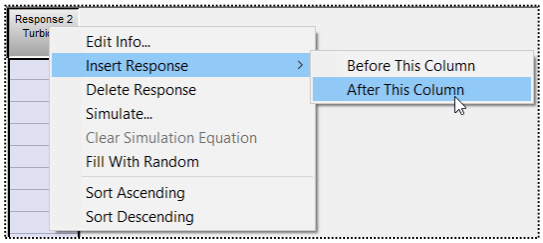
Using Design-Expert software allows you to explore the impact of changing multiple components on multiple responses — and to find maximally desirable solutions quickly via numerical optimization. For your final report, finish up with a graphical overlay plot at the optimum "slice." (Don't forget you can set goals on the components themselves. For example, in this case it might be wise to try maximizing the amount of cheap water.)

Learn more about mixture design methods at our workshop titled Mixture Designs for Optimal Formulations. To get the latest class schedule, go to the Training link at www.statease.com. Also, we appreciate your questions and comments on Design-Expert software. Address these to stathelp@statease.com.

Postscript: Adding a Cost Equation

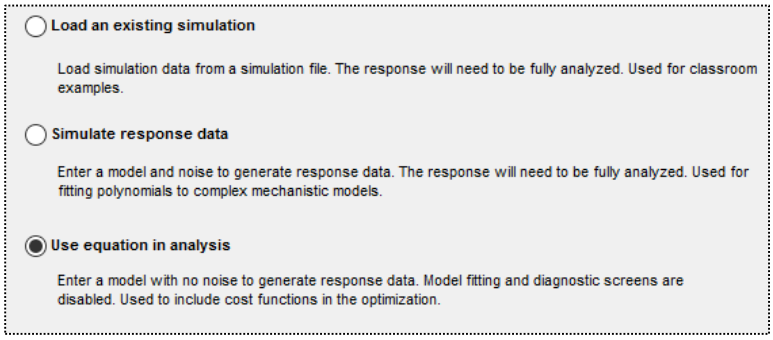
In the comments above, we suggested you consider maximizing the cheapest ingredient – water in this case. Conversely, you may have an incredibly expensive material in your formulation that obviously needs to be minimized. With only a small amount of effort, you can set up cost as a response to be included in Design-Expert's numerical optimization.

Re-**open** the **Mix-a.dxpx** file. In the **Design** branch, right-click the last response column. From the menu, select **Insert Response**, **After This Column**.



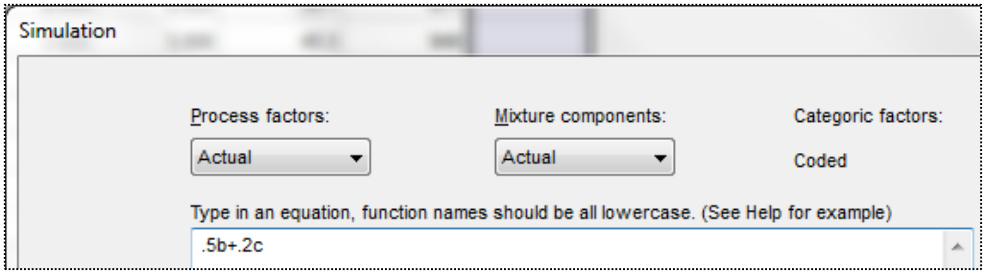
Inserting a new response

Next, right-click the new untitled response header and select **Simulate**. Then choose **Use equation in analysis**.



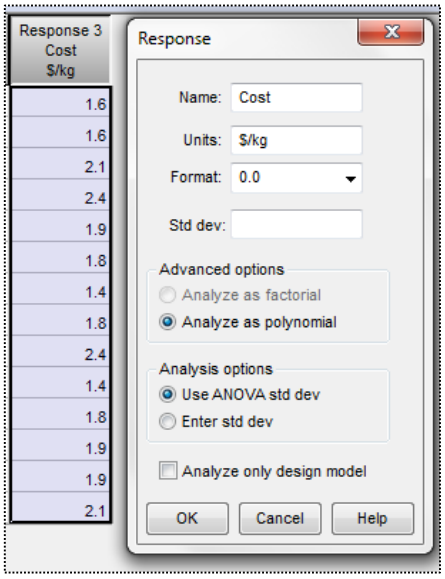
Preparing to enter cost equation

Press **Next** and enter into the area provided **.5b+.2c** (alcohol at $0.50 per kilo and urea at $0.20 cents – assume water costs practically nothing). Also, enter the **Response Name** as **Cost** and **Response units** in **$/kg**.



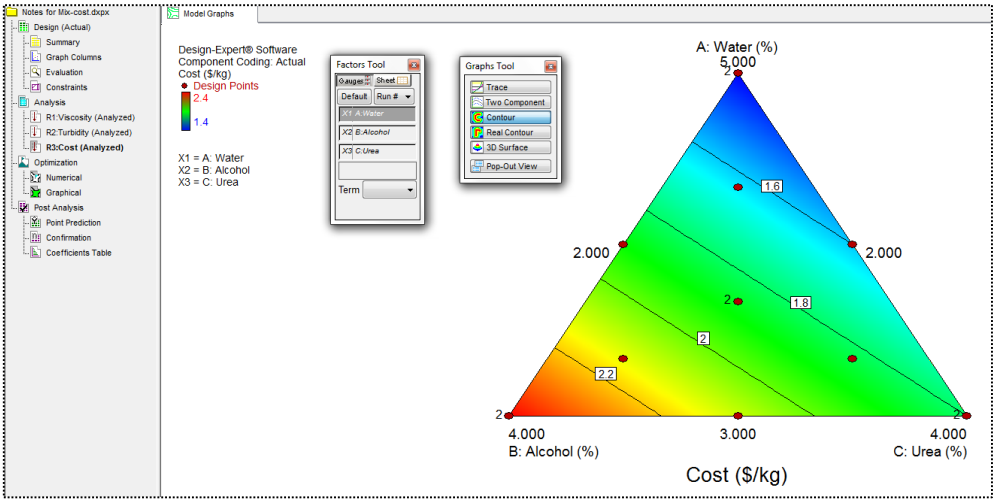
Entering the cost equation

Press **Finish** to accept the equation and calculate costs for all formulations in this mixture design. To make these more presentable, right-click the **Response 3** column header, select **Edit Info,** and in the droplist change **Format** to **0.0**. Press **OK**.



Costs formatted

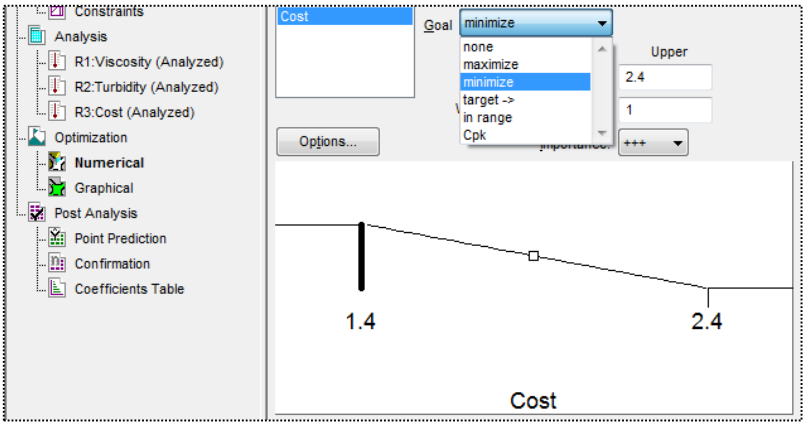
Now, under the Analysis branch, click the **Cost** node to bring up the model graph directly — no modeling is necessary because you already entered the deterministic equation.



Contour plot of cost

The water shows blue due to it being so cheap.

This sets the stage to include cost in your multiple response optimization. As pictured below, go to **Optimization** node **Numerical**, select **Cost** and set its **Goal** to **minimize**.



Minimizing cost

Pressing Solutions at this stage only tells you what you already know: The lowest cost formula is at the greatest amount of water within the specified constraints. Re-enter the goals for viscosity and turbidity and their POEs if you like, but it really isn't necessary now. Wait until you do your own mixture design and then make use of this postscript tip to take costs into account.