You are a chemist and are tasked with optimizing a chemical reaction:

$$R-X \; + \; R' \diagup\!\!\diagdown \xrightarrow[\text{Sodium Acetate}]{\text{Pd catalyst}} \; R' \diagup\!\!\diagdown\!\!\diagup R$$

You will try to *maximize* the **Yield** first by varying factors in an ad hoc fashion (traditional), and then by using a designed experiment (DOE).

You will find optimal factor settings for five factors, with High and Low settings based on reasonable starting points from prior experience.

Factor	Low	High
Volume	1.5	10
Catalyst	1	5
Temperature	50	120
Sodium Acetate	1.5	4
Time	4	24

Based on your subject matter knowledge, you suspect that – for these factor ranges – **Catalyst** and **Sodium Acetate** might not affect **Yield**. **Volume**, **Temperature**, and **Time** might have a curvilinear (polynomial) relationship with **Yield**, and there might be an interaction between **Temperature** and **Time**.

Traditional Optimization

For this part of the homework, you will vary the factor settings within the low and high limits based on your subject matter knowledge. You will simulate running the experiment to find the factor settings that produce the highest **Yield**. (*Note that experimental error is built into the simulation and entering the same factor settings will lead to different results.*)

1. Use the **Traditional Optimization Homework** table in the journal. For reference, two runs are included that use either all the highest settings or all the lowest settings.

2. Add a new row of data by entering settings for each factor at or between the high and low settings. You decide how you want to change them, based on your subject matter expertise!

3. Run the **Perform One Run** script in the table panel to get a result for **Yield**. (*If you don't get a result, check to make sure your settings are within the correct range.*)

4. Continue adding rows with different combinations of factor settings and running the experiment until you have a total of 16 runs.

5. What were the settings that produced the highest **Yield**, and what was the **Yield**?

DOE Optimization

For this part of the homework, an experiment was designed to estimate the main effects and a subset of all the possible higher order terms based on the experimenter's subject matter knowledge described above.

1. Use the **DOE Optimization Homework** table in the journal. The results of the 16-run experiment are in the table.

2. Run the **Profiler** script in the table and explore how each factor affects the **Yield** by dragging the vertical, dashed red lines to the left and right. Can you detect an interaction between any factors? (*Note: this is the same profiler graph you can see in the Fit Model platform. It is presented separately here using the Profiler platform for demonstration purposes.*)

3. Run the **Fit Model** script in the table to perform the analysis of the experiment. Which effects are most important? Are you able to detect higher order effects (interactions and polynomials) with this experiment?

4. Use the Prediction Profiler to find the optimum factor settings. What is the predicted **Yield** at these settings?

5. Do the results from the designed experiment match the results you got from your ad hoc experiment?

Solutions

You are a chemist and are tasked with optimizing a chemical reaction.

You will try to *maximize* the **Yield** first by varying factors in an ad hoc fashion (traditional), and then by using a designed experiment (DOE).

You will find optimal factor settings for five factors, with High and Low settings based on reasonable starting points from prior experience.

Based on your subject matter knowledge, you suspect that – for these factor ranges – **Catalyst** and **Sodium Acetate** might not affect **Yield**. **Volume**, **Temperature**, and **Time** might have a curvilinear (polynomial) relationship with **Yield**, and there might be an interaction between **Temperature** and **Time**.

Traditional Optimization

For this part of the homework, you will vary the factor settings within the low and high limits based on your subject matter knowledge. You will simulate running the experiment to find the factor settings that produce the highest **Yield**. (*Note that experimental error is built into the simulation and entering the same factor settings will lead to different results.*)

1. Use the **Traditional Optimization Homework** table in the journal. For reference, two runs are included that use either all the highest settings or all the lowest settings.

    a. Select **Traditional Optimization Homework** from the **Homework** section of the journal.

2. Add a new row of data by entering settings for each factor at or between the high and low settings. You decide how you want to change them, based on your subject matter expertise!

    a. Click in the first empty row of the data table and add values for each factor. Make sure your entries are not outside of the factor ranges indicated in rows 1 and 2.

3. Run the **Perform One Run** script in the table panel to get a result for **Yield**. (*If you don't get a result, check to make sure your settings are within the correct range.*)

    a. Click the green play button next to the **Perform One Run** script in the table panel.

    b. Check that a value has populated in the **Yield** column.

4. Continue adding rows with different combinations of factor settings and running the experiment until you have a total of 16 runs.

    a. Repeat the previous steps.

5. What were the settings that produced the highest **Yield**, and what was the **Yield**?

    a. Answers will vary.

DOE Optimization

For this part of the homework, an experiment was designed to estimate the main effects and a subset of all the possible higher order terms based on the experimenter's subject matter knowledge described above.

1. Use the **DOE Optimization Homework** table in the journal. The results of the 16-run experiment are in the table.

    a. Select **DOE Optimization Homework** from the **Homework** section of the journal.

2. Run the **Profiler** script in the table and explore how each factor affects the **Yield** by dragging the vertical, dashed red lines to the left and right. Can you detect an interaction between any factors? (*Note: this is the same profiler graph you can see in the Fit Model platform. It is presented separately here using the Profiler platform for demonstration purposes.*)

    a. Click the green play button next to the **Profiler** script in the table panel.
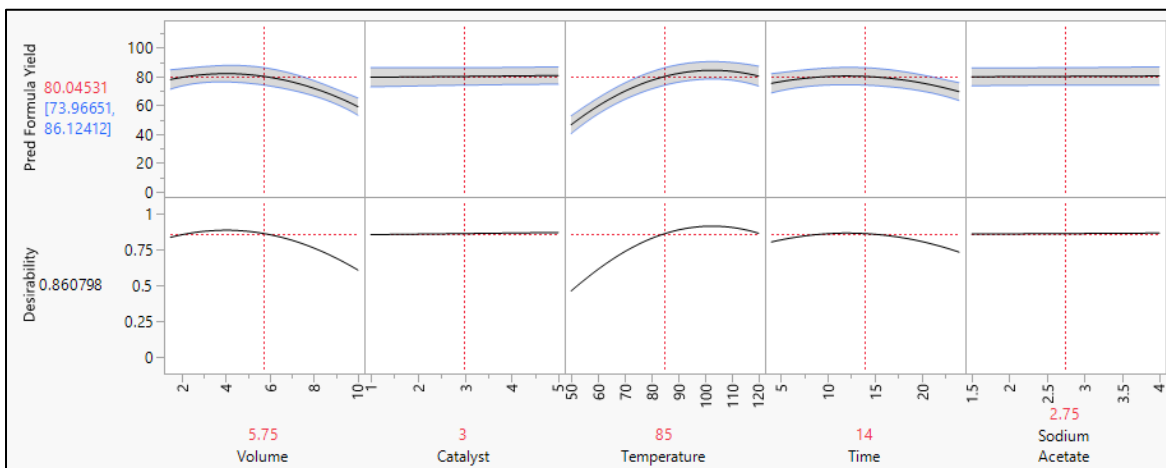

Volume, **Temperature**, and **Time** all appear to have an effect on the predicted value of **Yield**. There is curvature in the profiles of **Temperature**, and **Time**. **Catalyst** and **Sodium Acetate** don't appear to have a relationship with **Yield**.

b. Click and hold to drag the vertical, dashed red lines to new settings for the factors.

Changing the settings for **Volume**, **Catalyst**, and **Sodium Acetate** does not appear to change the shape of the relationship between the other factors and the predicted value of **Yield**. Changing the settings for either **Temperature** or **Time** does change the shape of the relationship of the other; this suggests an interaction between these two factors.
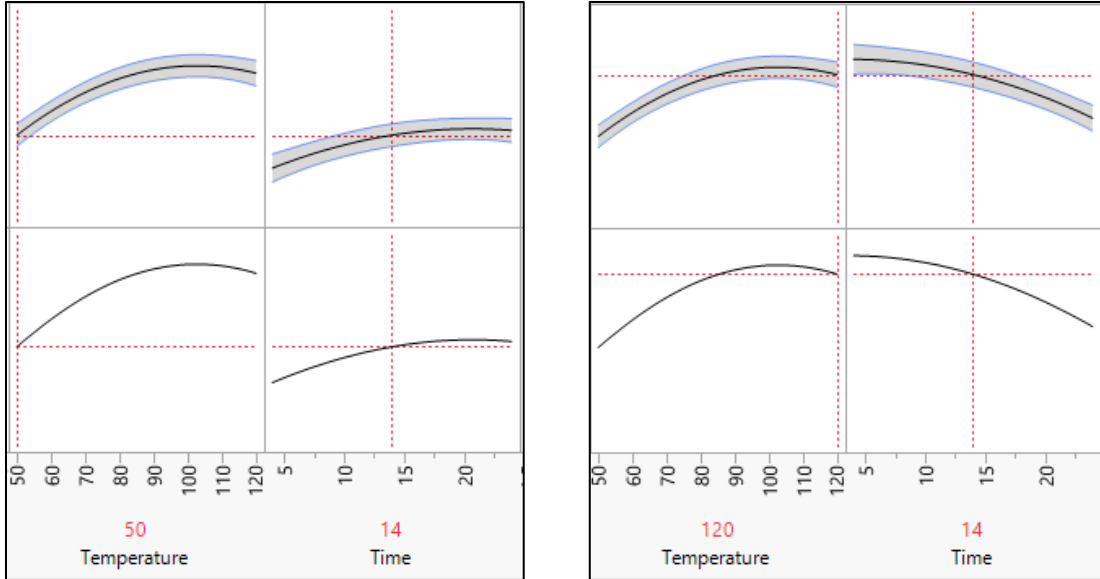
3. Run the **Fit Model** script in the table to perform the analysis of the experiment. Which effects are most important? Are you able to detect higher order effects (interactions and polynomials) with this experiment?

a. Click the green play button next to the **Fit Model** script in the table panel.
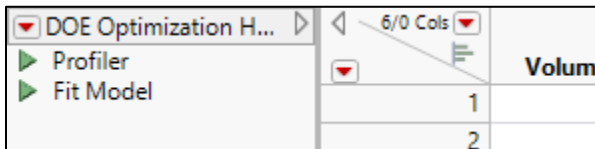

b. Examine the Effect Summary report to determine which effects are most important.

Effect Summary

Source	Logworth		PValue
Temperature(50,120)	5.896		0.00000
Temperature*Time	5.117		0.00001
Volume(1.5,10)	4.577		0.00003
Temperature*Temperature	3.932		0.00012
Volume*Volume	2.393		0.00405
Time*Time	2.084		0.00823
Time(4,24)	1.783		0.01648 ^
Catalyst(1,5)	0.323		0.47554
Sodium Acetate(1.5,4)	0.143		0.71864

Remove Add Edit Undo ☐ FDR ('^' denotes effects with containing effects above them)

Temperature, **Temperature*Time**, **Volume**, and **Temperature*Temperature** are the most important effects. The experiment was able to detect higher order terms.

4. Use the Prediction Profiler to find the optimum factor settings. What is the predicted **Yield** at these settings?

   a. Click the red triangle next to **Prediction Profiler** and select **Optimization and Desirability** > **Maximize Desirability**.
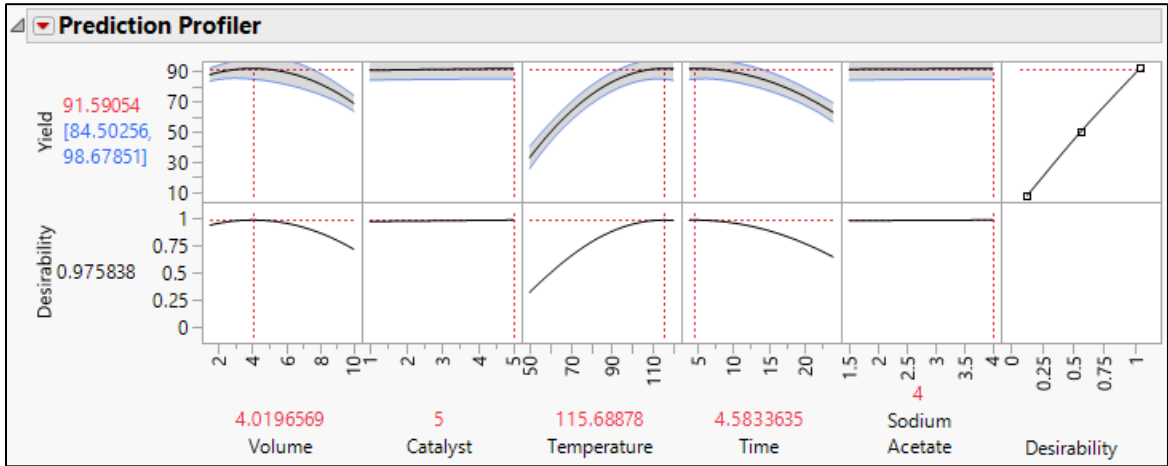
   At the optimum factor settings shown, the predicted **Yield** is 91.59%.

5. Do the results from the designed experiment match the results you got from your ad hoc experiment?

   a. Probably not! Especially because of the presence of interactions and polynomial effects, and the number of factors, it is very difficult to successfully change factor settings just based on subject matter knowledge and get a good understanding of the system. The designed experiment ensures that the higher order terms can be estimated, and an optimized response predicted. Of course, the next step would be to perform several runs at the chosen factor settings to confirm that the **Yield** is similar to what is predicted by this model.

   b. **Extra Credit**: Try to fit the full model (main effects, 2-factor interactions, and 2nd order polynomials) to the data you collected in your ad hoc experiment. Are you able to estimate the higher order terms?