JMP052: Optimization of Microbial Cultivation Process

Design of Experiments, Predictive Modeling

Produced by

Andreas Trautmann, Lonza AG
andreas.trautmann@lonza.com

Claire Baril, Lonza AG
claire.baril@lonza.com

Volker Kraft
volker.kraft@jmp.com

Muralidhara A
muralidhara.a@jmp.com

Lonza

jmp STATISTICAL DISCOVERY FROM SAS

Optimization of Microbial Cultivation Process

Design of Experiments, Predictive Modeling

Key ideas

This case study requires the use of design of experiments (DOE) to optimize the microbial cultivation process. An I-optimal custom design is generated and evaluated for diagnostic measures, followed by fitting a statistical model to experimental data. Model accuracy and diagnostics were analyzed, and optimal factor setting was ascertained using advanced prediction.

Background

Lonza is a Swiss multinational biotech and pharmaceutical company, headquartered in Basel, with major facilities in Europe, North America, and South Asia. Lonza was founded in 1987 and it's known for technological innovation with world class manufacturing and process excellence. The company provides product development services, early phase discovery and custom manufacturing of active pharmaceutical ingredients for the pharma, consumer health and nutrition sector. This also includes scientific innovation and manufacturing technology optimization.

Lonza has established a purpose to ***Enable a healthier world*** and the company believes that this is the reason for its existence. The vision of the company is ***Bring any therapy to life***. Lonza is involved in the manufacturing of biologics with several pharmaceutical companies and is recognized by Ethisphere as one of the world's most ethical companies. As of 2021, Lonza has more than 5,278 trademark fillings, 792 brands and 1,065 small and large molecules.

Lonza Group comprises two segments, namely Lonza Pharma, Biotech & Nutrition (LBPN) and Lonza Specialty Ingredients (LSI). LBPN focuses on delivering innovative solutions that prevent, treat or even cure disease. LSI focuses on microbial control for personal care, as well as for the protection of homes, schools, workplaces and the environment from mold and other potential pathogens. Lonza has divided its business into four divisions: Small Molecules, Biologics, Cell & Gene, and Capsules & Health Ingredients.

Bioreactors

0.25L scale | 2L scale | 30L scale | 1500L scale

A microbial cultivation process is a method of multiplying microorganisms by letting them reproduce in predetermined culture media under controlled laboratory conditions. Microbial cultures are used to determine the type of organism, its abundance in the sample being tested, or both. A bioreactor is a manufactured device or system that supports a biologically active environment. It refers to a device or system designed to grow cells or tissues in the context of cell culture. These devices are used for

biochemical/bioprocess engineering. The sizes of the bioreactor can vary over several orders of magnitudes. One of the microbial cultivation processes uses three different sizes: small scale (0.25 L), lab scale (2-30 L) and pre-pilot scale (1500 L).

The task

The Manufacturing team was tasked to speed development of this novel biomanufacturing processes to help products reach the market in a short time. The team focused on increasing the yield of the process, because a poor yield translates to higher manufacturing costs. An economically unprofitable product will not reach the market even if it were to be beneficial to customers.

The process involves four continuous process variables: feed rate, complex compounds, temperature, and pH. Instead of actual values, the standard convention of -1 and +1 are used to represent the low and high levels of factor settings.

The initial factor settings and corresponding yield are:

Feed rate	Complex compounds	Temperature	pH	Yield
0	-0.6	0	0	41%

The goal is to find out the key process variables and their optimal process settings that would maximize the yield. The team also wants to understand the relationship between input and output variables.

Design of experiments (DOE)

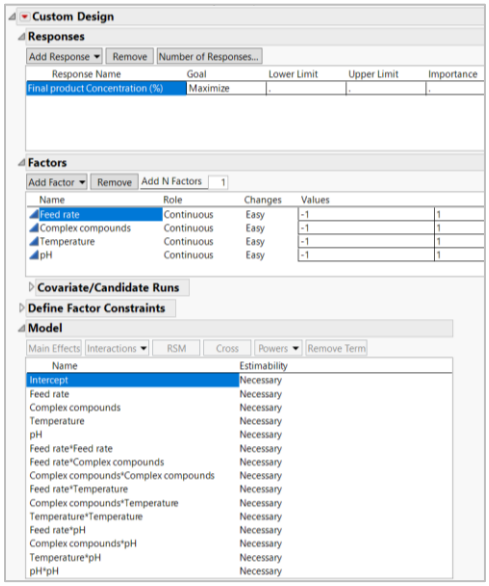
With an aim of identifying the key parameters affecting the yield, the team leveraged a statistically designed experiment to collect the right data. Knowing that there are many classical designs available and supporting a sequential approach, they had the option to identify the key factors in a screening experiment as a first step, followed by an optimization experiment. However, to minimize the overall resource usage, the team created a tailor-made design platform for a more efficient approach.

By using the Custom Design platform in JMP, if a predefined classical design doesn't perfectly fit the problem, one can construct an optimal design that is custom-built for the specific experimental situation at hand. The Custom Designer creates a wide array of design types capable of addressing an extensive range of experimental conditions and goals. Custom Design has a unique ability to configure factors, constraints and factor settings; information and other experimental conditions; and resource restrictions.

Exhibit 1 Response and Factors for Designing an Experiment

Response		**Goal**	
Final Product Concentration (%)		Maximize (100%)	
Factors	**Nature**	**Lower Limit**	**Upper Limit**
Feed rate	Continuous	-1	+1
Complex compounds	Continuous	-1	+1
Temperature	Continuous	-1	+1
pH	Continuous	-1	+1

Exhibit 2 Custom Design



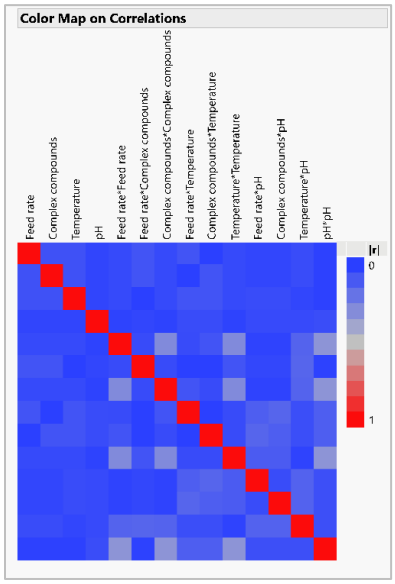
Run	Feed rate	Complex compounds	Temperature	pH
1	-0.1	1	1	-1
2	0	0	1	1
3	-1	0	0	0
4	0	1	-1	0
5	-1	-1	-1	1
6	1	-1	-1	0
7	1	1	1	0
8	0	0	0	0
9	0	-1	0	1
10	0	-1	-1	-1
11	1	-1	0	-1
12	1	0	1	-1
13	0	0	0	0
14	-1	1	1	1
15	-1	-1	1	0
16	1	1	-1	-1
17	-1	1	0	-1
18	-1	-1	1	-1
19	-1	0	-1	-1
20	0	0	0	0
21	-1	-1	-1	1
22	1	0	-1	1
23	1	1	0	1
24	1	-1	1	1

To create, DOE>Custom Design>Add response: Final Product Concentration (%). Add four continuous factors as shown and populate the lower and upper values. Under the Model options, choose RSM to add all the quadratic and two-way interaction effects for all the continuous variables selected. Enter 3 as the number of center points. Set the user-specified number of runs as 24. Click Make Design and Make Table. The above process will create a data table along with the details of the experimental runs with factor settings.

Evaluation of custom design

The Design Evaluation outline provides different ways to evaluate a custom design.

Exhibit 3 Color Map on Correlations for Custom Design



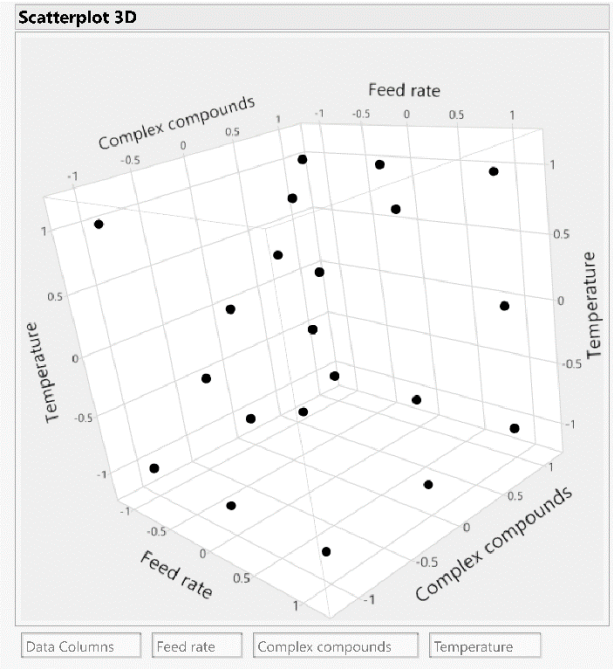
For this example, open Design Evaluation and examine the color map on correlations. (If the chart appears in black and white, use the red triangle next to Color Map on Correlations drop-down and select Use Blue to Red color theme for the color map.)

The color map on correlations shows the absolute value of the correlation between any two effects that appear either in the Model or in the Alias Terms outline. The main effects are represented by the seven terms in the upper-left corner of the map. The blue corresponding to the correlation between two effects indicates correlations of 0. The red represents a correlation of 1.

Red indicates absolute correlations of 1, here just reflecting that each term is perfectly correlated with itself. It follows that no main effect is confounded (totally correlated) with any two-way interaction. A closer look at this diagram will reveal that collinearity is negligible and all main effects are nearly orthogonal and can be estimated nearly independently.

One can also visualize the factor settings using a 3D scatter plot.

Exhibit 4 Scatter Plot



To create, Graph>Scatterplot 3D>Select three continuous factors.

The data Custom Design Results.jmp

The results from the 24-run randomized experiment are shown in Exhibit 5

Exhibit 5 Custom Design Results

	Feed rate (-)	Complex compounds (-)	Temperature (-)	pH (-)	Final product conc. (%)
1	-0.1	1	1	-1	109
2	0	0	1	1	32
3	-1	0	0	0	44
4	0	1	-1	0	21
5	-1	-1	-1	1	23
6	1	-1	-1	0	38
7	1	1	1	0	67
8	0	0	0	0	32
9	0	-1	0	1	34
10	0	-1	-1	-1	33
11	1	-1	0	-1	61
12	1	0	1	-1	•
13	0	0	0	0	38
14	-1	1	1	1	•
15	-1	-1	1	0	53
16	1	1	-1	-1	47
17	-1	1	0	-1	68
18	-1	-1	1	-1	41
19	-1	0	-1	-1	29
20	0	0	0	0	35
21	-1	1	-1	1	22
22	1	0	-1	1	38
23	1	1	0	1	35
24	1	-1	1	1	•

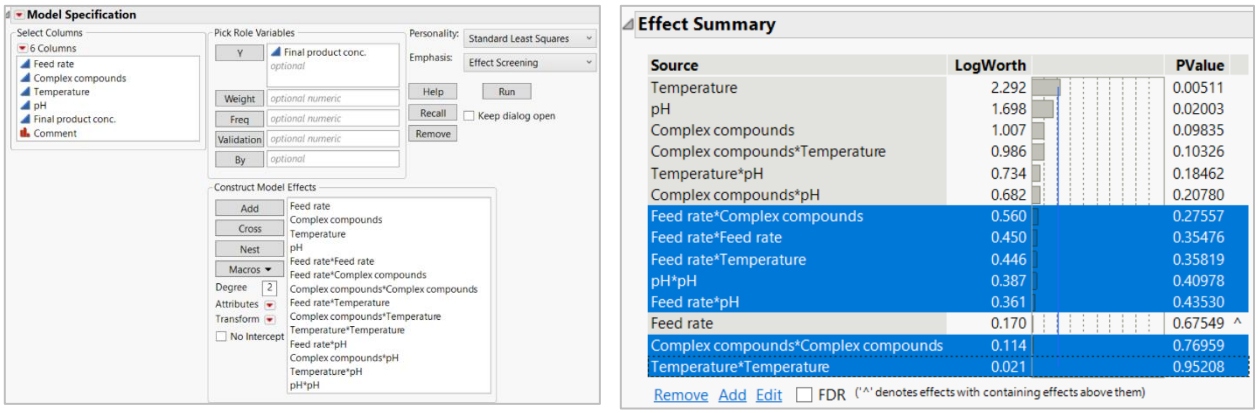
There were three runs where the final product concentration was unable to populate due to technical challenges. The variables in the data set are **Feed rate, Complex compounds, Temperature, pH**, and **Final Product Concentration (%)**. All the variables are continuous.

Analysis of experimental output

Let us analyze the experimental results using a multiple regression model to find out the statistically significant model effects. Use the data set Custom Design Results.jmp

JMP facilitates the task of data analysis by saving a Model script to the design table that it creates. You can either run this script after you conduct your experiment and enter your data or build a model using analyze>Fit Model option. The script opens a Fit Model window containing the effects that you specified in the Model outline of the Design window.

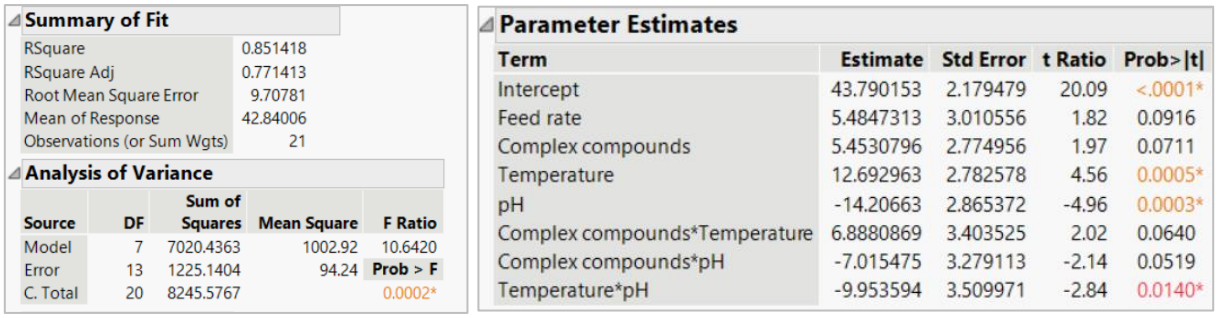
Exhibit 6 Building Initial Regression Model and Selecting Effects



To create, Analyze>Fit Model>select Final product concentration % as Y variable and all other variables as Model Effects. To add the quadratic and interaction effects, choose Response Surface option under the Macro drop-down. Click Run. One can also explore the Stepwise regression by changing the Personality.

From the results, scroll down to Effect Summary and confirm that many effects are statistically not significant (P-value is higher than 0.05), and remove the least significant effects one by one. The selected effects are removed to fine-tune the model and will rebuild the regression model each time. The resulting regression output is shown in Exhibit 7. Ensure that the final model has all the four main effects and the three additional interaction effects.

Exhibit 7 Summary of Fit, ANOVA and Parameter Estimates of the Fine-tuned Model



The Summary of Fit report provides details on RSquare. An RSquare closer to 1 indicates a better fit to the data. The value of 0.851 indicates that the model explains most of the variation observed in the response. The P-value in the ANOVA table tells us that the model is statistically significant overall. However, the ANOVA doesn't tell us which predictors are significant. For this, use the information reported in the Parameter Estimates table (a more detailed version of the Effect Summary).

The Parameter Estimates show that pH and Temperature are statistically significant. The interaction Time and pH is also statistically significant.

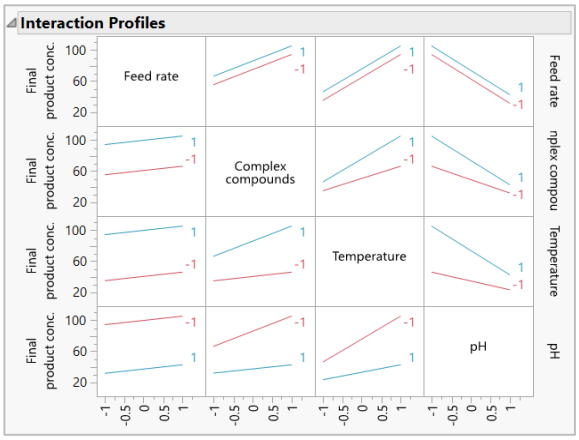
Process optimization using Prediction Profiler

The profilers in JMP provide a highly interactive cross-sectional views of any response surface. The most optimized process parameters or even the most robust factor settings can be identified using the JMP

Profiler. The Prediction Profiler displays the response surface traces between the response and each factor. One can explore all dimensions of an opportunity space in one plot. Each trace shows the predicted response as one factor is changed while the other factors are held constant at the given values. The Prediction Profiler recomputes the profiles and predicted response (in real-time) as one varies the value of a factor.

Since our model has interaction effects, the Interaction Plots option under the profiler provides a visual output showing a matrix of interaction plots. Each cell of the matrix contains a plot whose horizontal axis is scaled for the effect displayed in the column in which the plot appears. Line segments are plotted for the interaction of that effect with the effect displayed in the corresponding row. So, an interaction plot shows the interaction of the row effect with the column effect. An interaction is more significant, if the two lines are less parallel.

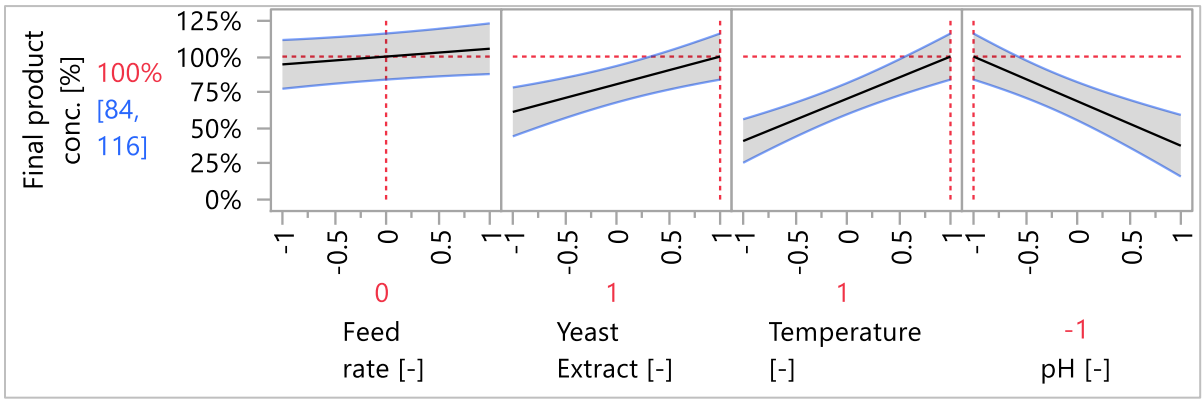
Exhibit 8 Interaction Profiles



To create, go to the red triangle next to response Final Product Conc>Factor Profiling>Profiler to get to the Prediction Profiler. Under the red triangle of Prediction Profiler, choose the option Interaction Profiles.

The Prediction Profiler has an additional built-in feature called Maximize Desirability which determines one combination of factor settings that results in a predicted response that optimizes a given desirability. Our desirability is to maximize yield, i.e., Final Product Concentration. Please note, there can be different factor level combinations that maximize the desirability to the same extent.

Exhibit 9 Profiler with Optimized Factor Settings



To change the factor settings and explore the changes in the response, go to the red triangle and choose Maximize Desirability to get one combination of optimized factor settings.

With a goal of maximizing the yield, the process team explored multiple factor settings and finalized one of them which is displayed visually in Exhibit 9.

Summary

Statistical insights

Organisations always strive to improve the processes by predicting the most significant factors affecting the responses. The process team leveraged smarter experimentation to understand the existing process. The unique custom design platform helped the team to extract maximum information from the minimum number of experiments. The goal of predicting the key variables affecting the yield and their factor settings for the optimum value was achieved.

Managerial implications

Achieving manufacturing excellence through a robust process is the goal of every organization. To reach that goal, it is important that the quality be imbibed throughout the process, rather than as an end-product inspection. Finding out the important factors affecting the output and understanding them practically helps in ascertaining the optimum process settings. By doing smarter experimentation using flexible and efficient designs, LONZA saved time and was able to extract maximum information from minimum resources. Finding the optimum factor settings helped increase the yield and better meet the time-to-market expectations.

JMP features and hints

This case used the DOE platform to create tailor-made design to suit experimental requirements. A 3D scatter plot was leveraged to visualize the design. Fit Model was used to identify the significant predictors and generate the predictive values of the response variables. Profiler was used to visualize the response-factor interactions and Prediction Profiler was used to arrive at the optimal process settings.