JMP Academic Case Study 066

Chemical Process Improvement in Resin Production: Part 1

Process capability, process modeling, Design Space Profiler

Produced by

Frank Deruyck, HOGENT University of Applied Sciences
frank.deruyck@hogent.be

Volker Kraft, JMP Global Academic Team
volker.kraft@jmp.com

Chemical Process Improvement in Resin Production: Part 1

Process capability, process modeling, Design Space Profiler

Key ideas

A chemical company producing resins faces two problems in production: On the one hand, the quality of the produced resins is too low. On the other hand, the required offline measurements to detect defects are very inefficient. In this first of two case studies, we investigate process stability and capability to understand the baseline quality before building a process model to improve the process quality by controlling the critical parameters.

Background

BLX Chemicals is a leading producer of high-quality resin that serves as essential building blocks across numerous industries. From advanced manufacturing and coatings to adhesives, composites, and innovative material solutions, resins help customers create products that are stronger, more durable, and more sustainable.

A critical quality attribute (CQA) of a resin is purity, which describes how free it is from contaminants or any unreacted components. High purity ensures consistent performance, stability, and reliability in applications such as coatings, adhesives, and polymer production. Even small impurities can influence color, strength, or reactivity, making careful purification and quality control essential.

In recent months, the resin manufacturing process operated at the best achievable stability, which was still not optimal. As a result, purity did not meet its specification for too many batches, which led to a poor baseline capability of the process. The chemical engineering team wants to address this issue.

To ensure acceptable product quality, an additional purification step is required. This initial purification step is applied before the final purification takes place, which leads to an increased processing time. It is further required to measure purity on each product. Since sufficient data for online monitoring of purity are not available, an offline measurement of purity became necessary, adding even more time to the overall process.

The resin production runs 22.5 hours per day over two shifts and uses two reactors.

The task

The primary objective of our first project is the elimination of the additional purification step, aiming at significant time and cost savings. These are the suggested analysis steps:

1. Investigate the baseline quality of the current manufacturing process:
   - Explore the relationships between the critical process parameters and purity.
   - Use control charts to analyze stability of the process.
   - Check normality assumptions and compare process capability for both reactors.
2. Build and apply a process model for optimization:
   - Build a response surface model using the given experimental data.
   - Control the ranges of process parameters to improve the in spec portion of produced resins.
   - Simulate new data under the optimized conditions to estimate the expected capability.
   - Compare the baseline capability with the improved capability by simulating new data under optimized conditions.

The data

SPC Purity.jmp

Purity	Quality attribute of the chemical product in %
Shift	Production shift (1, 2)
Hour	Production time (1-20)
Reactor	Reactor (1, 2)

I-optimal DOE.jmp

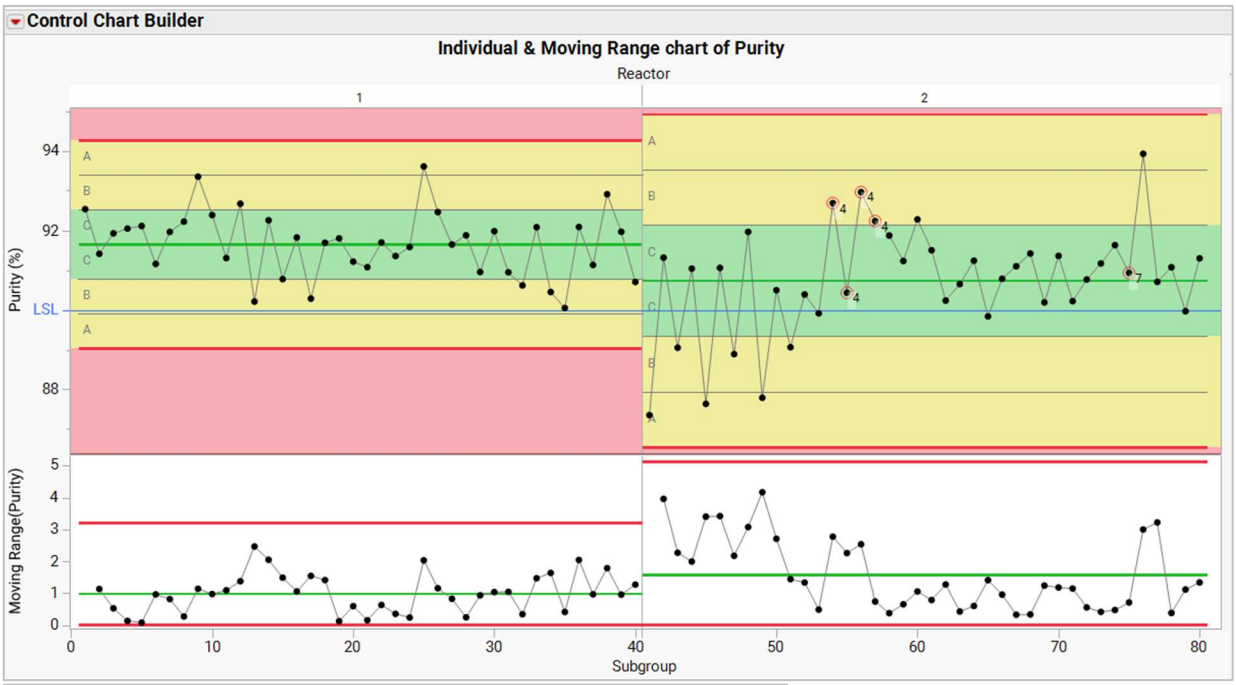
Run	Experiment run #
Random Block	Blocking factor for day (1, 2, 3, 4)
Temperature	Continuous factor #1 (135-145 °C)
Time	Continuous factor #2 (10-20 min)
Pressure	Continuous factor #3 (100-160 10^3Pa)
Flowrate	Continuous factor #4 (100-180 l/h)
Purity	Response variable with spec limit

Analysis

Investigate the baseline quality

We begin by investigating process stability for Purity using a control chart (data set: **SPC Purity.jmp**; Exhibit 1). The Individual & Moving Range chart shows random variation in Purity for both reactors. However, Reactor 2 displays warnings that may indicate special causes of instability. Although these problems should be fixed first, right now there are no obvious options to stabilize the process. Therefore, the quality team decides to move on to capability analysis and an overall improvement initiative by optimizing a process model.

Exhibit 1 Control chart of purity



Purity Limit Summaries

Points plotted	Reactor	LCL	Avg	UCL	Limits Sigma	Subgroup Size
Individual	1	89	92	94	Moving Range	1
Individual	2	87	91	95	Moving Range	1
Moving Range	1	0	1	3	Moving Range	1
Moving Range	2	0	2	5	Moving Range	1

(Analyze > Control Chart Builder. Select Purity > Y; select Reactor > Phase to create an Individual & MR chart. Right-click > Limits > Zones and Shade Zones. Right-click > Warnings > Tests > All Tests. Hour > Subgroup to change into XBar & R chart.)

Exploring our data and measuring the baseline capability leads us to the Distribution platform. As a first step (Exhibit 2), we click on the histograms and bar charts to explore relationships between Shift, Hour, and Reactor, as well as between these variables and Purity.

Exhibit 2 Distributions

(Analyze > Distribution. Assign all variables to Y, Columns. Deselect Create Process Capability and select Histograms Only.)

The product specification requires a Purity of 90% or higher, otherwise the produced resin would be classified as a defect (also called “out of spec”). This goal is called a specification limit and is already set as a column property on the output column. Either click the asterisk (*) beside the column name in the Columns section, or right-click the column header on the data table to check the spec limit. In this case, we have the lower spec limit (LSL) set to 90.

Since the distributions and capability analysis are shown by Reactor, we can easily compare both reactors side by side (Exhibit 3 shows Reactor 1 only). The Purity histograms relative to the LSL confirm that (the more stable) Reactor 1 produces better results than Reactor 2.

Running the goodness of fit tests after a continuous fit of normal distributions to our data, there is some minor evidence that only Reactor 2 deviates from the normal distribution. Although this is good to know while considering analyses that are based on the standard deviation σ, no further action is taken to address this finding.

Exhibit 3 Distributions with capability analysis (Reactor 1 only)

Distributions

Reactor=1

Purity

Compare Distributions

Show	Distribution		AICc ^	BIC	-2*LogLikelihood
☑	Normal	——	101.21798	104.27141	96.893655

Fitted Normal Distribution

Parameter		Estimate	Std Error	Lower 95%	Upper 95%
Location	μ	91.666216	0.1300667	91.403131	91.929301
Dispersion	σ	0.8226143	0.0937456	0.6738536	1.056266

Measures	
-2*LogLikelihood	96.893655
AICc	101.21798
BIC	104.27141

Goodness-of-Fit Test

	W	Prob<W
Shapiro-Wilk	0.9853162	0.8739

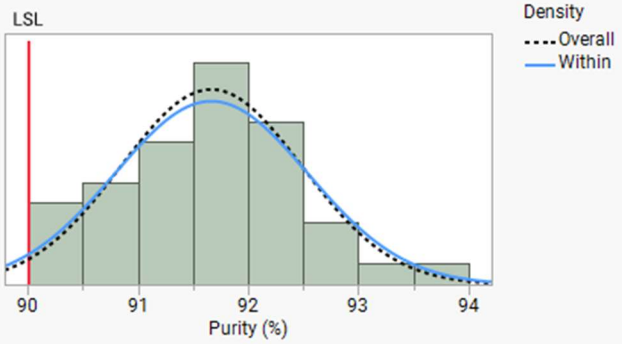
	A^2	Simulated p-Value
Anderson-Darling	0.209783	0.8460

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

Process Capability

Purity Capability

Histogram



Process Summary

LSL	90
N	40
Sample Mean	91.66622
Within Sigma	0.87535
Overall Sigma	0.822614
Stability Index	0.939755

Within sigma estimated by average moving range.

Within Sigma Capability

Index	Estimate	Lower 95%	Upper 95%
Cpk	0.634	0.427	0.838
Cpl	0.634	0.427	0.838

Overall Sigma Capability

Index	Estimate	Lower 95%	Upper 95%
Ppk	0.675	0.492	0.855
Ppl	0.675	0.492	0.855

Nonconformance

Portion	Observed %	Expected Within %	Expected Overall %
Below LSL	0.0000	2.8489	2.1407
Total Outside	0.0000	2.8489	2.1407

(Analyze > Distribution. Assign Purity to Y, Columns and *Reactor to By. Deselect Create Process Capability and select Histograms Only. Broadcast or repeat: Purity red triangle > Histogram Options > deselect Vertical. Select Continuous Fit > Fit Normal, add Goodness of Fit. Select Process Capability.)*

Measuring baseline capability (“How capable is our process meeting the specification?”), the Process Performance Index Ppk is reported for both reactors as part of the capability analysis. This index is calculated by dividing the distance between the mean μ and the LSL by 3σ, where σ is the overall (or

long-term) standard variation. An LSL at 3σ would result in Ppk = 1. A higher index is better. The Ppk for Reactor 1, and especially for Reactor 2, shows that the process can be clearly classified as “not capable.” While we have no defects for Reactor 1 in our observed data (see Nonconformance section), the expected defect rate – at the left tail of the fitted distribution – is above 2%. For Reactor 2, the expected defect rate is close to 30%. All these findings motivate our next step, an initiative to improve the process (both capability and stability) by optimizing a process model.

Improving the process

The quality team has decided to improve the process by optimizing the ranges (lower and upper limits) of the process settings. The factor settings with a significant impact on purity have already been determined by screening experiments; they are as follows:

- Temperature (135-145 °C)
- Reaction Time (10-20 min)
- Pressure (100-160 10^3Pa)
- Dosing Flowrate Reactant (100-180 l/h)

For optimization purposes, the data have already been collected designed to fit a Response Surface Model (RSM). The design steps of this experiment are not described here, but you can explore the settings in Custom Design by running the table scripts DOE Dialog or Scatter Plot Matrix. The Custom Designer was told that we want to maximize Purity as our response, based on four continuous factors fitting an RSM to the experimental data. A random blocking effect has been added to capture the day-by-day-variation. For convenience, the designer also prepared the analysis by adding a Fit Model script to the design table. The analysis estimates the model parameters using the REML[1] method, due to the random (blocking) effect in our model.

Launching the model fit to our data, note that the model effects include the random block effect, as well as the (fixed) main effects, two-factor interactions, and quadratic terms. As the estimation method REML is pre-configured by the design tool, leave the options at their default settings (which is always a good idea when using JMP unless you have specific needs or you already know better).

Investigating the model report (Exhibit 4), we remove the four least significant effects one by one to simplify the model a bit. The Actual by Predicted Plot shows a reasonable model with high R-squared and low p-value. There are no strange predictions (outliers), but – as expected – the measured purity of some runs did not match with the specification.

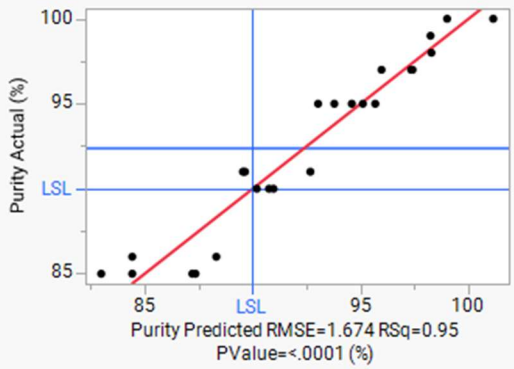
The Fixed Effect Tests show which effects are significant: Temperature, Time and Pressure, the Temperature*Time interaction, the quadratic effect for Time, and the Time*Flowrate interaction. Interestingly, Flowrate is not significant as a main effect, which is an exception from the heredity principle. By checking the REML Variance Component Estimates (less than 10% of the variation of the residual (model error) can be explained by the day-by-day-variation), the other part remains as noise. For that reason, we can just ignore the blocking factor during our next step: the process optimization.

[1] Restricted maximum likelihood: for more details, open jmp.com/help and search for REML.

Exhibit 4 Model output

Response Purity

Actual by Predicted Plot



Effect Summary

Source	Logworth	PValue
Temperature(135,145)	4.554	0.00003
Time*Time	3.904	0.00012
Time(10,20)	3.458	0.00035 ^
Temperature*Time	3.305	0.00050
Pressure(100,160)	3.229	0.00059
Time*Flowrate	1.755	0.01759
Flowrate*Flowrate	0.960	0.10962
Pressure*Pressure	0.703	0.19812
Temperature*Temperature	0.485	0.32723
Flowrate(100,180)	0.091	0.81078 ^

Summary of Fit

RSquare	0.94523
RSquare Adj	0.903099
Root Mean Square Error	1.673956
Mean of Response	92.41667
Observations (or Sum Wgts)	24

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Wald p-Value	Pct of Total
Random Block	0.1063252	0.297937	0.7735002	-1.218096	1.8139696	0.7001	9.611
Residual		2.8021289	1.2649074	1.3605253	8.7475317		90.389
Total		3.1000659	1.2238106	1.6236106	8.1078684		100.000

-2 Residual Log Likelihood = 77.014537603

Note: Total is the sum of the positive variance components.

Total including negative estimates = 3.1000659

Fixed Effect Tests

Source	Nparm	DF	DFDen	F Ratio	Prob > F
Temperature(135,145)	1	1	11.4	45.1639	<.0001*
Time(10,20)	1	1	9.8	28.4974	0.0003*
Pressure(100,160)	1	1	11.2	22.3281	0.0006*
Flowrate(100,180)	1	1	10.8	0.0602	0.8108
Temperature*Temperature	1	1	9.8	1.0631	0.3272
Temperature*Time	1	1	12.5	21.7090	0.0005*
Time*Time	1	1	9.8	37.2121	0.0001*
Pressure*Pressure	1	1	9.8	1.9045	0.1981
Time*Flowrate	1	1	10.8	7.8225	0.0176*
Flowrate*Flowrate	1	1	10.2	3.0721	0.1096

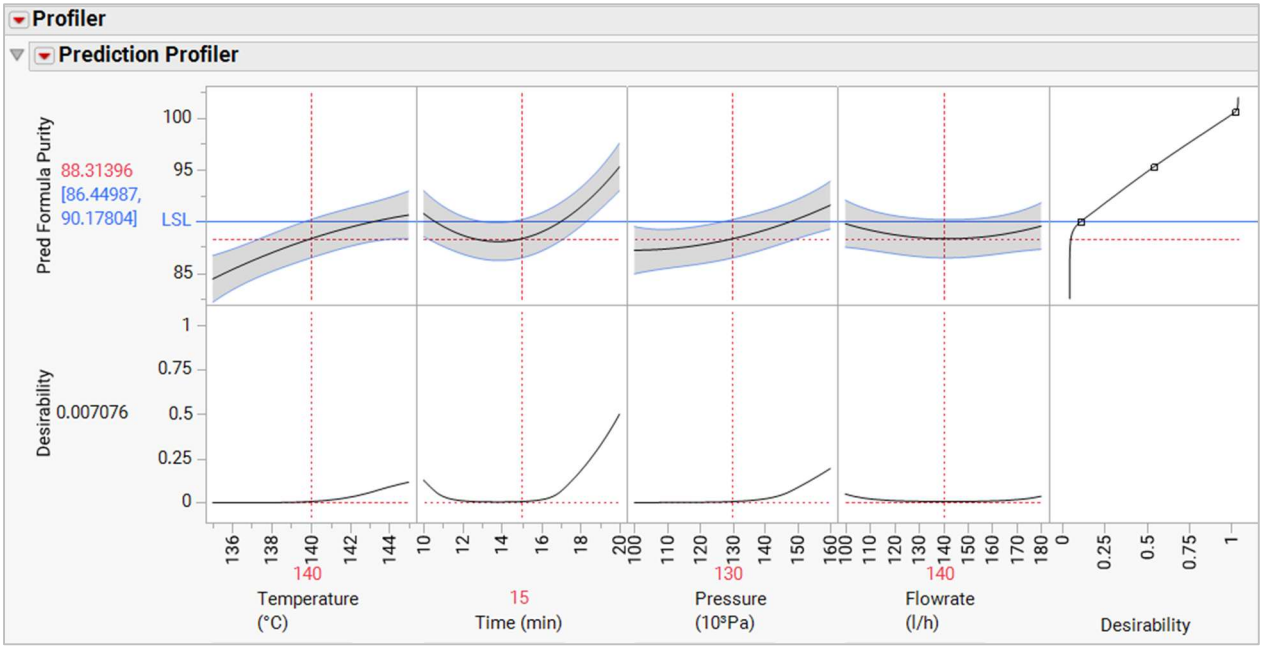
(Run script Fit Model > Run, Row Diagnostics > Plot Actual by Predicted, remove least significant model effects keeping only 10 effects. See Effect Summary; Save Columns > Prediction and Interval Formulas.)

The prediction formula can be saved from the model (instructions are below Exhibit 4) and loaded into the Profiler (Exhibit 5). The Prediction Profiler shows the relationships between Purity and the process settings still at their midpoints, with confidence intervals and desirability functions (we want to maximize the response, with 90% as the lower spec limit).

The quality team decides to follow a design space optimization restricting the factor settings ranges to smaller ranges, instead of maximizing the desirability in the Profiler and implementing (and fixing!) the best combination of factor settings. This optimization approach – also known as Quality by Design (or QbD) – is available as a Profiler extension called Design Space Profiler (DSP) on the red triangle. Behind the scenes, the DSP simulates 10,000 data points uniformly distributed over the design space. For each point, Purity is predicted based on the model, with green indicating InSpec and red indicating OutOfSpec results. Shrinking the factor ranges can help increase the InSpec Portion, while the Volume Portion decreases at the same time. Clicking Move Inward automatically takes the steepest path increasing the InSpec Portion.

We stop moving lower and upper limits as soon as the InSpec Portion exceeds 99%, allowing less than 1% defects occurring in the resulting design space. However, the Volume Portion nearly reached 1% (remember, less could affect the accuracy of the predicted portions), and the ranges for Pressure and especially Temperature are quite small. If they are too small, process engineers should be consulted if these tighter controls are technically possible.

Exhibit 5 Prediction Profiler (at midpoints)



(Graph > Profiler, use Pred Formula Purity as Y, Prediction Formula, OK.)

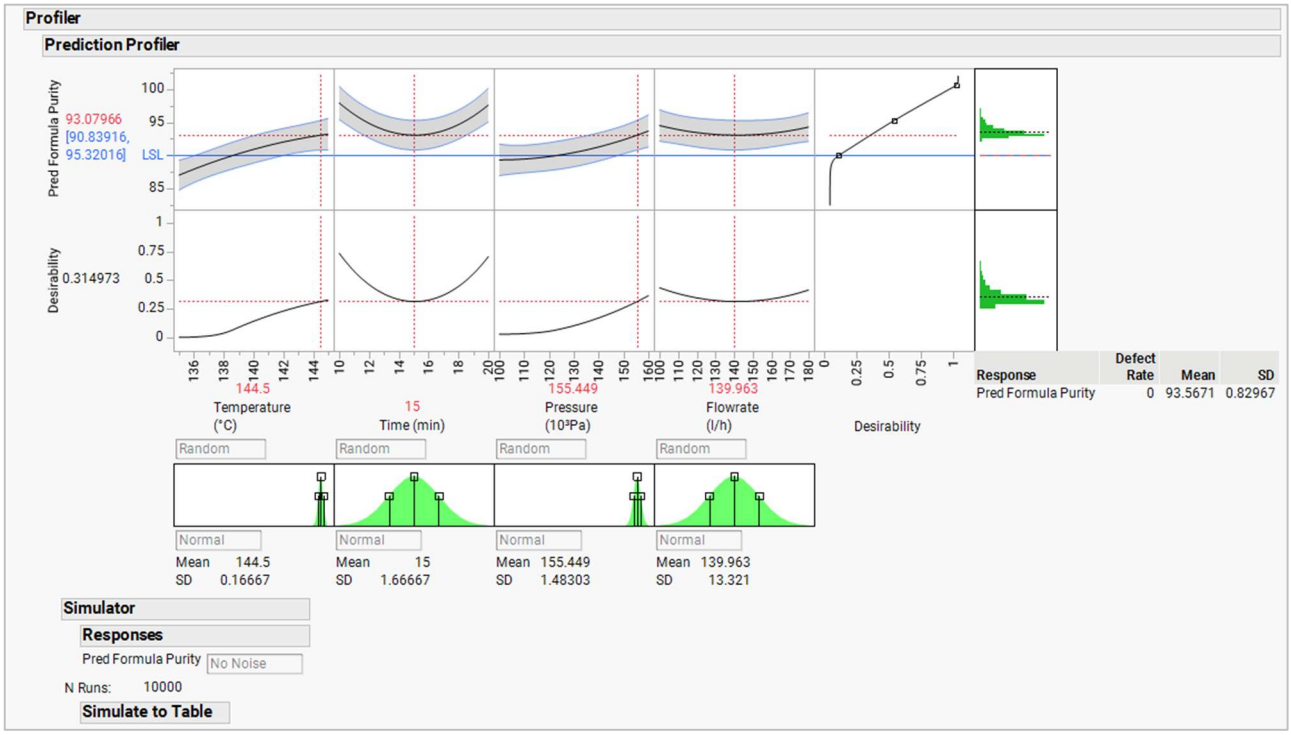
Exhibit 6 Design Space Profiler (optimized)

(Prediction Profiler red triangle > Design Space Profiler, Make and Connect Random Table > Embed Factor Space Scatterplots > OK, click Move Inward / - Outward to improve the InSpec Portion.)

Comparison with baseline quality

While the InSpec Portion at >99% already looks promising, we should consider that it is based on simulated data uniformly distributed over the factor space. Preferring a normal distribution around the mid settings can get us closer to the “real” process capability and the defect rates to be expected from the optimized process. Exhibit 7 shows the result after sending the final midpoints to the Profiler and the ranges to the Simulator. We are assuming normally distributed process settings, centered at the midpoints and 6 Sigma between the lower and upper limits.

Exhibit 7 Profiler and Simulator



(DSP > Send Midpoints to Profiler and Send Limits to Simulator with Normal at 3 Sigma. Add No Noise to Response Pred Formula. Click Simulate.)

For better comparability with our low baseline capability (Ppk=0,675!), we add no additional noise to the Monte-Carlo simulated response values. Each time Simulate is clicked, 10,000 response values are simulated with a distribution indicated by the green histogram on the response row in the Profiler. The “observed” (simulated) defect rate is already around zero, which is great news.

The Simulator also supports writing all data into a new data table. Running the Distribution script reveals a Ppk=1,543 (Exhibit 8), which is a significant improvement compared to the baseline capability.

Please note: Due to the random nature of all simulations (Design Space Profiler, Simulator), expect minor variation of the exact results.

Exhibit 8 Capability for simulated purity data

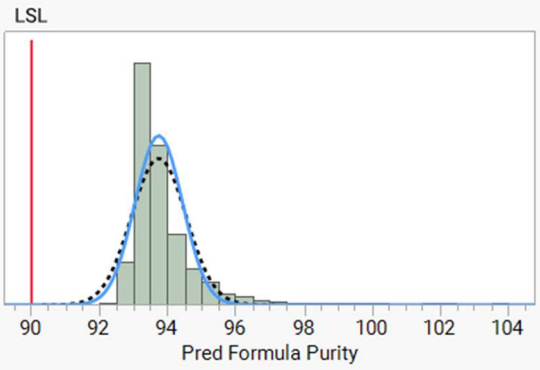
Distributions Student Edition

Pred Formula Purity

Process Capability

Pred Formula Purity Capability

Histogram



Density
Overall
Within

Process Summary

LSL	90
N	10000
Sample Mean	93.74819
Within Sigma	0.693876
Overall Sigma	0.803196
Stability Index	1.15755

Within sigma estimated by average moving range.

Within Sigma Capability

Index	Estimate	Lower 95%	Upper 95%
Cpk	1.801	1.768	1.833
Cpl	1.801	1.768	1.833

Overall Sigma Capability

Index	Estimate	Lower 95%	Upper 95%
Ppk	1.556	1.533	1.578
Ppl	1.556	1.533	1.578

Nonconformance

Portion	Observed %	Expected Within %	Expected Overall %
Below LSL	0.0000	0.0000	0.0002
Total Outside	0.0000	0.0000	0.0002

(Prediction Profiler > Simulator > Make Table without adding noise to the response. On the new data table, click the green triangle to run the Distribution script.)

What's next?

A mandatory next step would be confirming the improvement of process performance by some validation runs, measured under real-world conditions (e.g., keeping all settings in new ranges with random variation of process settings as expected in the production environment).

Summary

Statistical insights

Understanding the quality of a production process is key. Visual inspection of process data, such as using the Control Chart Builder or Distribution platform in JMP, can help to easily diagnose the health of a process and detect stability and capability issues. Standard quality metrics like the Process Performance Index Ppk help to better compare alternative processes or track even minor changes in one process.

The process improvement was based on a process model fitted to experimental data. Neither the design of the experiment (here, an I-optimal design with random blocking) nor the modeling step (here, using a REML mixed model) were discussed on a high level. It cannot be emphasized enough that both of these steps have to be conducted carefully to allow a significant step forward when using the model for process

optimization. JMP supports both DOE and modelling by many alternative methods, as well as using helpful diagnostics. Refer to other case studies to learn more about DOE or modeling.

A new approach has been chosen by the quality team to apply the model and optimize the process: The model prediction formula was loaded into the Prediction Profiler for a better process understanding. From here, the Design Space Profiler was used to change the lower and upper limits of process settings, focusing on a better InSpec Portion. After finding an acceptable solution, the results can be sent back to the Profiler and Simulator. The simulated data can also be sent into a capability analysis to retrieve the standard performance measurements.

Implications and further study

This project has been a successful step toward improving both the stability and capability of the resin production. A follow-up project (Part 2 of this case study) addresses our second goal: removing the time-consuming offline purity measurement by implementing a new online monitoring step using spectral characteristics of

Exercises

1. Compare the analysis done so far with alternative prediction formulas:

   a) Use the same modeling method, but remove all non-significant effects from the model before saving the prediction formula to the data table.

   b) Save the Conditional Pred Formula instead, which also takes the random effect into account.

2. Lock the range of Time to 12-15. Will this change much? How can you know just by exploring the Design Space Profiler?

3. How does adding random noise to the simulated response data change the expected capability?