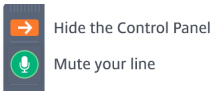
Dive into Diagnostics to Discover Data Discrepancies

Shari Kraber, MS Applied Statistics

shari@statease.com

August 2023

Making the most of this learning opportunity



To prevent audio disruptions, all attendees will be muted.

Questions can be posted in the Question area. If they are not addressed during the webinar, I will reply via email afterwards.

Questions may also be sent to stathelp@statease.com. Please provide your company name and, if you are using Design-Expert, the serial number (found under Help, About).

Note: The slides and a recording of this webinar will be posted on the Webinars page of the Stat-Ease website for your review.

Stat-Ease, Inc.

Design of Experiments Software and Training
www.statease.com

Stat-Ease 360

DESIGN EXPERT

YouTube Statistics Made Easy by Stat-Ease
@StatisticsMadeEasybyStatEase

Stat-Ease Training:
Sharpen Up Your DOE Skills

Modern DOE for Process Optimization

Mixture Design for Optimal Formulations

ESSENTIALS Series short courses

Individuals	Teams (6+ people)
Improve your DOE skills	Choose your date & time
Topics applicable to both novice and advanced practitioners	Add company case studies

Email list sign up:
www.statease.com/publications/signup/

Agenda

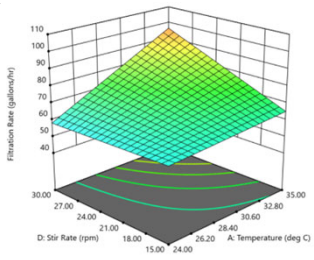
- **Analyzing DOE results – why check diagnostics?**
- Diagnostic basics
- Case Study – Well (bad diagnostics and a transformation)
- Wrap Up

Purpose of a Designed Experiment

Goal: Take DOE data (small sample) and → turn it into a predictive model (of a larger system).

Std	Run	Factor 1 A:Temperat... deg C	Factor 2 B:Pressure psig	Factor 3 C:Concentr... percent	Factor 4 D:Stir Rate rpm	Response 1 Filtration R... gallons/hr
11	1	24.00	15.00	2.00	30.00	45
9	2	24.00	10.00	2.00	30.00	43
12	3	35.00	15.00	2.00	30.00	104
13	4	24.00	10.00	4.00	30.00	75
15	5	24.00	15.00	4.00	30.00	70
14	6	35.00	10.00	4.00	30.00	86
2	7	35.00	10.00	2.00	15.00	71
3	8	24.00	15.00	2.00	15.00	48
7	9	24.00	15.00	4.00	15.00	80
10	10	35.00	10.00	2.00	30.00	100
8	11	35.00	15.00	4.00	15.00	65
1	12	24.00	10.00	2.00	15.00	45
4	13	35.00	15.00	2.00	15.00	65
6	14	35.00	10.00	4.00	15.00	60
16	15	35.00	15.00	4.00	30.00	96
5	16	24.00	10.00	4.00	15.00	68

Filtration Rate	=
+70.06	
+10.81	* A
+4.94	* C
+7.31	* D
-9.06	* AC
+8.31	* AD



The Statistical Tool - ANOVA

The statistical relationship between the factors and the responses is built using the statistical tool called analysis of variance (ANOVA).

ANOVA for selected factorial model

Response 1: Filtration Rate

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	5535.81	5	1107.16	56.74	< 0.0001	significant
A-Temperature	1870.56	1	1870.56	95.86	< 0.0001	
C-Concentration	390.06	1	390.06	19.99	0.0012	
D-Stir Rate	855.56	1	855.56	43.85	< 0.0001	
AC	1314.06	1	1314.06	67.34	< 0.0001	
AD	1105.56	1	1105.56	56.66	< 0.0001	
Residual	195.12	10	19.51			
Cor Total	5730.94	15				

Question: are the p-values, and related R^2 values, correct statistically?

ANOVA Assumptions

ANOVA validity assumes that the model is correct, and the residuals:

- Are (nearly) independent
- Have a mean = 0
- Have a constant variance
- Follow a "well-behaved" distribution (approx. normal)

➢ If the model is correct, then the residuals should look like a sample of random variables – no "signal" remaining.

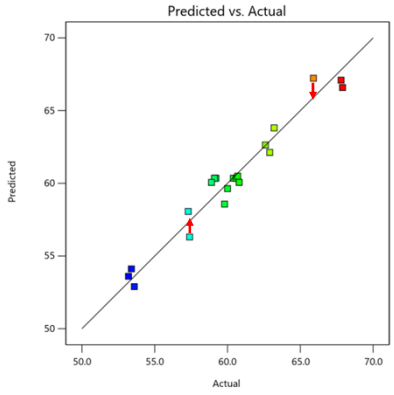
➢ If the assumptions are violated, then the p-values, R-squared values, predictions, etc, are wrong.

Agenda

- Analyzing DOE results – why check diagnostics?
- **Diagnostic basics**
- Case Study – Well (bad diagnostics and a transformation)
- Wrap Up

What is a residual?

Residual – difference between the actual data and the predicted value. $e_i = y_i - \hat{y}_i$



Reasons **not** to use the raw residuals to verify ANOVA:

- Sensitive to under-fitting or over-fitting of the model (too few or too many terms)
- Variance is not constant

Better residuals

Internally studentized residuals (standardized) – the raw residual divided by its estimated standard error. This helps with nonconstant variance but does not follow a t distribution.

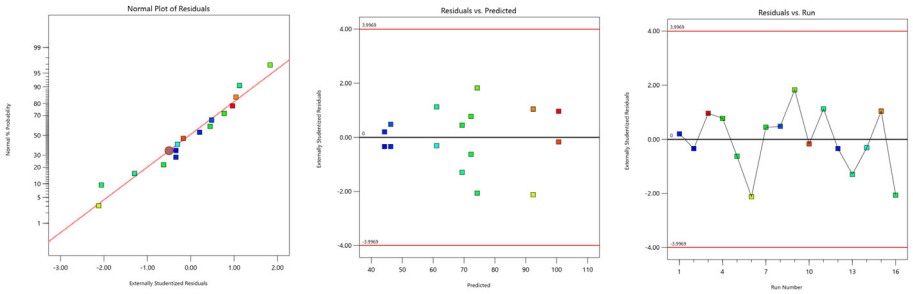
$$r_i = \frac{e_i}{\sqrt{MS_{res}(1 - h_{ii})}}$$

Externally studentized residuals (deleted) – adjusted to drop the *i*th data point. Corrects for nonconstant variance and follows a t distribution (which converges to the normal).

$$t_i = \frac{e_i}{\sqrt{MS_{res(-i)}(1 - h_{ii})}}$$

BEST for verifying assumptions

Diagnostic plots of Externally Std. Residuals



Normal plot of residuals – look for large violations of normality

Residuals vs Predicted – look for non-constant variance across range

Residuals vs Run – check for time-based trends or patterns

Agenda

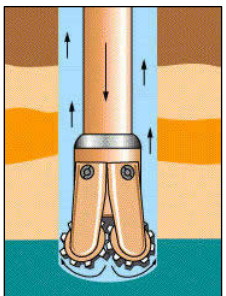
- Analyzing DOE results – why check diagnostics?
- Diagnostic basics
- **Case Study – Well (bad diagnostics and a transformation)**
- Wrap Up

Well-Drilling Experiment
2^4 factorial – Background

This is a case study with a twist (pun-intended). The goal is to increase the advance rate of a drilling operation.

The drill is driven by a pipe, which carries a water-based mud down-hole to:

- Carry off the pulverized rock and
- Lubricate the drill-bit.


Mud circulation in the hole

The mud carries the bits of rock up the annulus of the borehole. It then runs through gratings to shed off the rock-cuttings before being re-circulated.

Well-Drilling Experiment

Input factors and ranges to study:

Factor	Units	Low Level (−)	High Level (+)
Load	tons	50	70
Flow	gal/min	180	220
Speed	1/sec	45000	55000
Mud	base gum	guar	xanthan

Open Well and analyze.

Well-Drilling Experiment

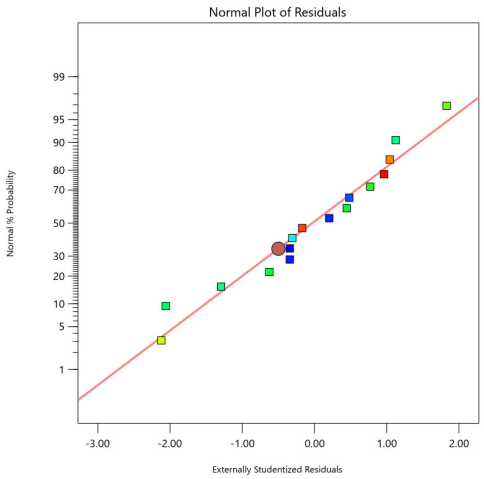
Fit a model and view ANOVA and Fit Statistics:

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	216.00	5	43.20	48.93	< 0.0001	significant
B-Flow	40.80	1	40.80	46.21	< 0.0001	
C-Speed	142.50	1	142.50	161.42	< 0.0001	
D-Mud	17.00	1	17.00	19.25	0.0014	
BC	8.48	1	8.48	9.61	0.0113	
CD	7.22	1	7.22	8.18	0.0169	
Residual	8.83	10	0.8828			
Cor Total	224.83	15				

Std. Dev.	0.9396		**R^2**	0.9607
Mean	5.92		**Adjusted R^2**	0.9411
C.V. %	15.86		**Predicted R^2**	0.8995
			Adeq Precision	19.5066

All looks good here!
Check diagnostics →

Normal plot of residuals – what it should be



The residuals should follow a well-behaved distribution.

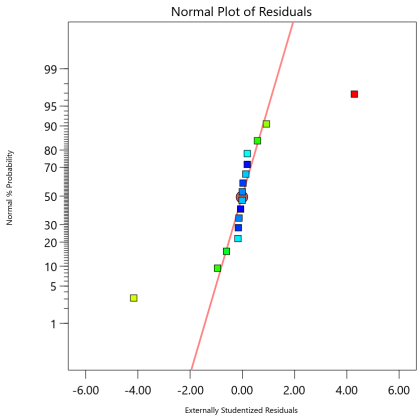
- Single peak
- Roughly symmetric
- Tails die rapidly

Roughly a normal distribution.

ANOVA is robust to moderate deviations from normality, so no need for a formal test (like Shapiro Wilks).

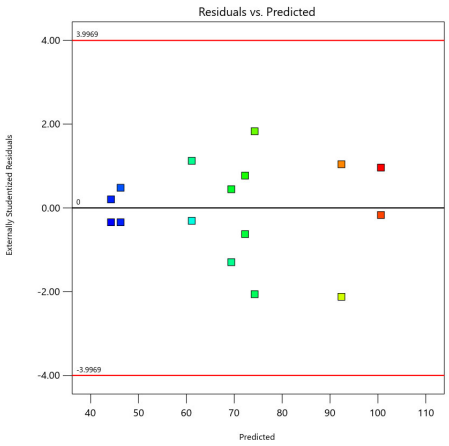
Well-Drilling Experiment
Diagnostic plots

Normal probability plot of residuals: want approximate normality



The residuals are not approximately normally distributed. This shows an S-shaped curve.

Residuals vs Predicted – what it should be

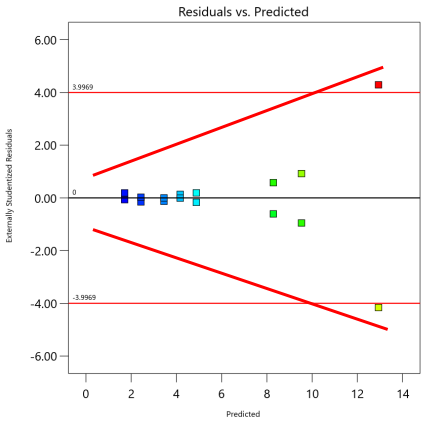


The plot should be a random scatter of points.

Look for systematic patterns or evidence of increasing variability as the predicted value increases.

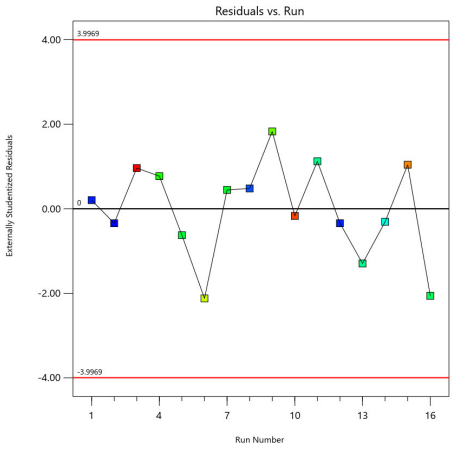
Well-Drilling Experiment
Diagnostic plots

Predicted vs Residuals: Look for constant variance left to right.



The residuals have small variation with small predictions and increasing variation with larger predictions.

Residuals vs Run Order – what it should be

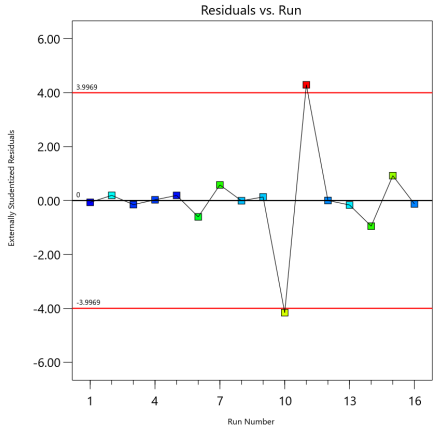


This time-ordered plot should show no patterns or trends.

- Watch for upward or downward trends, clumps of data.
- Watch for points not predicted well by the model, usually called "outliers".

Well-Drilling Experiment
Diagnostic plots

Residuals vs Run: want no time-based patterns/trends, no outliers

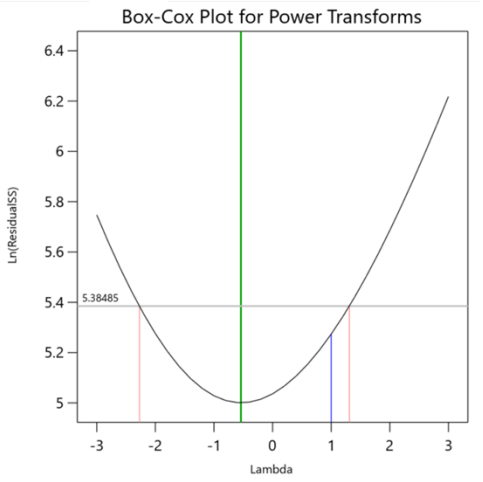


There are no particular trends, but there are 2 points that are not fit well by this model. "Outliers"??

Box Cox plot – what it should be

Current transform: None
Current Lambda = 1
Recommended transform: None

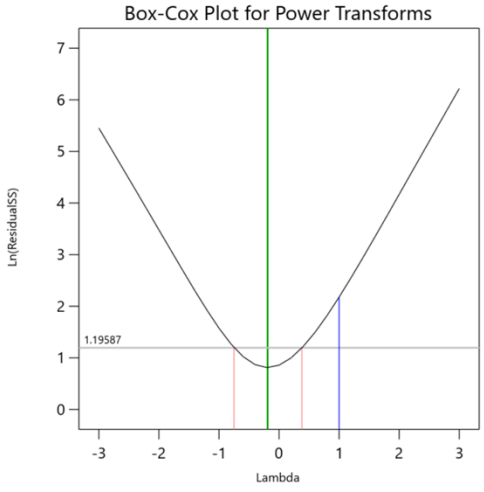
This plot will either recommend NO transformation or specify a particular legend.



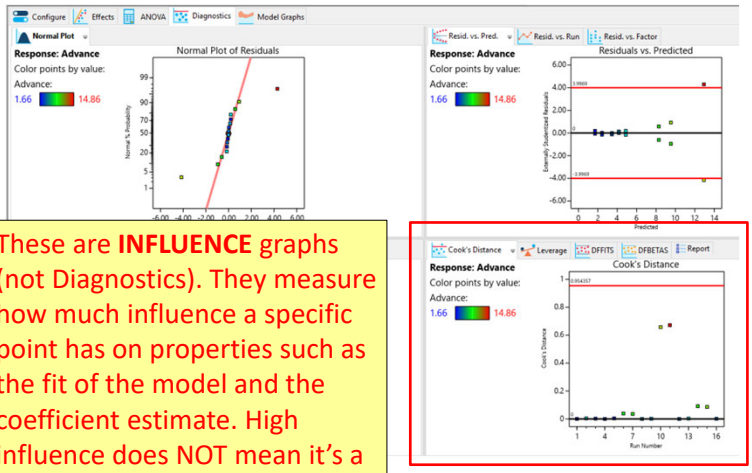
Well-Drilling Experiment
Box-Cox plot for transformations

Response: Advance
Current transform: None
Current Lambda = 1
Recommended transform: Log
(Lambda = 0)

The recommended transformation is a Log, so re-analyze the data using this scale.



Well-Drilling Experiment
What about the lower right graphs?



These are **INFLUENCE** graphs (not Diagnostics). They measure how much influence a specific point has on properties such as the fit of the model and the coefficient estimate. High influence does NOT mean it's a "bad" point!

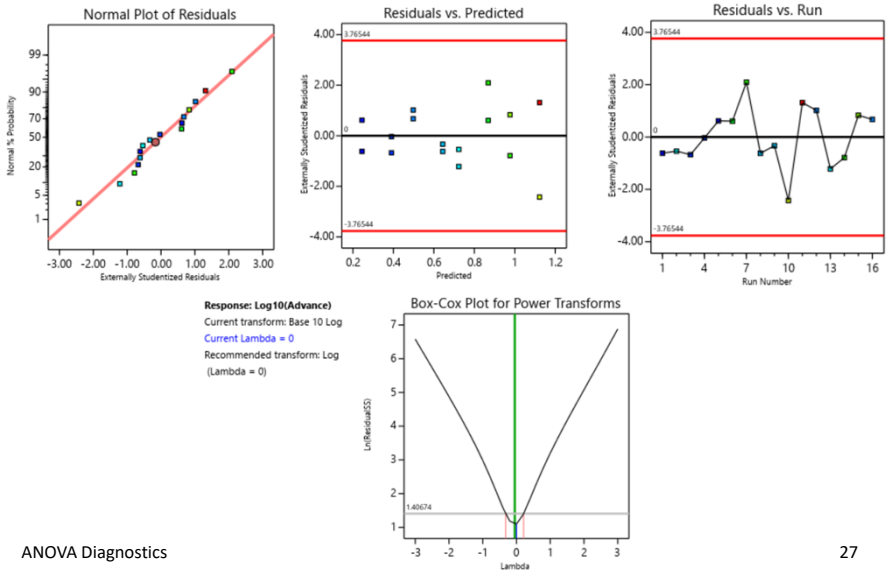
Well-Drilling Experiment
Re-analyze with log transform

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	1.26	3	0.4191	206.68	< 0.0001	significant
B-Flow	0.2568	1	0.2568	126.65	< 0.0001	
C-Speed	0.9155	1	0.9155	451.51	< 0.0001	
D-Mud	0.0849	1	0.0849	41.87	< 0.0001	
Residual	0.0243	12	0.0020			
Cor Total	1.28	15				

Std. Dev.	0.0450		**R^2**	0.9810
Mean	0.6835		**Adjusted R^2**	0.9763
C.V. %	6.59		**Predicted R^2**	0.9662
			Adeq Precision	38.9736

Fewer terms in the model, but it predicts better, explaining more variation.

Well-Drilling Experiment
Diagnostics after log transform



Agenda

- Analyzing DOE results – why check diagnostics?
- Diagnostic basics
- Case Study – Well (bad diagnostics and a transformation)
- **Wrap Up**

Why check diagnostics?
Satisfy ANOVA Assumptions

ANOVA validity assumes that the model is correct,
and the residuals:

- Are (nearly) independent (Resid vs Run Order)
- Have a mean = 0 (given by math)
- Have a constant variance (Resid vs Predicted)
- Follow a "well-behaved" distribution (Normal probability plot)

➢ If the model is correct, then the residuals should look like a sample of random variables – no "signal" remaining.

What if the residuals remain messy?

- Check the science – could there be any additional variables to consider?
- This is a small sample – did you have some "bad luck" and it just happens to have some unknown random effects?
- Do confirmation runs – use the model as is (even if a bit messy), pick some plausible/desirable conditions, and run them. Compare results to the predicted values from Point Prediction.
- Do a replicate or collect more data to enhance the information.

The researcher must weigh the risks of using an imperfect model versus spending more money to get more data.

Self-study options for learning more

YouTube Channel: www.youtube.com/c/StatisticsMadeEasybyStatEase

Playlist: New to DOE?

A collection of webinars on basic to intermediate-level topics.

Stat-Ease Academy: www.statease.com/training/academy/

Self-paced online courses covering the basics of factorial and fractional-factorial designs.

StatEase®
statistics made easy®

Make the most from every experiment!℠

DOE Questions? Email Stathelp@statease.com

Sales Questions? Email Sales@statease.com