

## Selecting a Most Useful Predictive Model



Mark J. Anderson, PE, CQE, MBA  
Engineering Consultant  
Stat-Ease, Inc., Minneapolis, MN  
[mark@statease.com](mailto:mark@statease.com)

1

### Maximizing this educational opportunity



Welcome everyone! To make the most from this webinar:

- Attendees on mute
- Chat addressed afterward
- Send further questions to [mark@statease.com](mailto:mark@statease.com)

PS Presentation posted to [www.statease.com/webinars/](http://www.statease.com/webinars/)

 *Please press the raise-hand button if you are with me.*

Selecting a Most Useful Predictive Model

2

2

## This Webinar: What's In It for You



- How to achieve a statistically solid predictive model from experimental results, particularly from response surface designs on process factors (*main focus*) and/or mixture components
- Tips on deploying algorithmic model reduction to glean a most 'parsimonious' (i.e., simple) equation that produces a useful maps of future performance
- Insights on key statistics for assessing your model fit
- Briefing on top three diagnostics of normality, outlier identification and possible advantage by a transformation

Selecting a Most Useful Predictive Model

3

3

## RSM: Process Flowchart



Subject Matter Knowledge  
(Plus, Factorial Screening)

Vital Few Factors ( $x$ 's)

Process

Measured Response(s) ( $y(s)$ )



Polynomial Model

Response Surface



*"All models are wrong, but some are useful."* - George Box

Selecting a Most Useful Predictive Model

4

4

## RSM Case Study (Program Tutorial)



A chemist studied three process factors:

- A. Time (minutes)
- B. Temperature (degrees C)
- C. Catalyst (percent)

StatEase.360  
DESIGN EXPERT

To optimize two key responses:

1. Conversion (%) => Maximize (80% or better)
2. Activity => Target 63 ( $\pm 3$  allowable)

For convenience, the experiment is run in two blocks via a “central composite design” (CCD):

1. Two-level factorial with center points.
2. Axial runs (star points\*) plus more center points.

\*Will show this in RSM tutorial picture of CCD.



RSM  
Rebuild—Show CCD layout

Selecting a Most Useful Predictive Model

5

5

## Algorithmic Model Reduction Criterion and Selection Methods



Stat-Ease software options (**default in bold**):

Criterion	Selection method
<b>AICc</b>	<b>Forward</b> , Backward
BIC	Forward, Backward
p-value	Forward, Backward, Stepwise
Adj R-Squared	All Hierarchical

See detail on criterion in notes below or refer to the *Handbook for Experimenters*. For nearly orthogonal designs such as regular CCDs, all approaches work well. To guard against high collinearity, go backward with AICc. **(Consider also the more traditional p-value backward selection.)** If the models agree, then you’re done. Otherwise, go with the one that produces the best fit statistics, keeping in mind that:

*All model reduction must be guided by subject matter knowledge!*

Selecting a Most Useful Predictive Model

6

6

## Final Tips for Good Modeling



- Always try reducing the 'design-for' model, if need-be overriding the one chosen by the software based on the best combination of sequential p-values, lack-of-fit results and fit statistics—adjusted and predicted  $R^2$  (mind any gap  $> 0.2!$ ). It's worth a try, particularly for big models where many insignificant terms can mask one that would be worth including.
- Avoid “paralysis by analysis”—it may be better to press ahead, imperfections, e.g., LOF—presuming you have tried everything!). You be the judge: Is the model useful?
- Get a 2nd opinion from the Stat-Ease Consulting Team. Show us the data! Email your SE software file to [stathelp@statease.com](mailto:stathelp@statease.com) with some background information on what you hope to accomplish.

Selecting a Most Useful Predictive Model

7

7

## This Webinar: What's In It for You



- How to achieve a statistically solid predictive model from experimental results, particularly from response surface designs on process factors (main focus) and/or mixture components
- Tips on deploying algorithmic model reduction to glean a most 'parsimonious' (i.e., simple) equation that produces a useful maps of future performance
- Insights on key statistics for assessing your model fit
- Briefing on top three diagnostics of normality, outlier identification and possible advantage by a transformation

*Mission accomplished?*

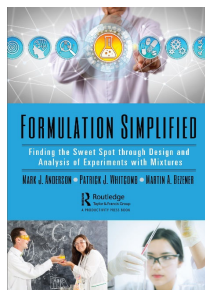
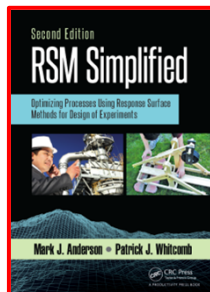
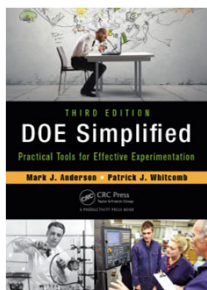
Selecting a Most Useful Predictive Model

8

8

## References

DOE/RSM/Formulation Simplified Series\*



\*Anderson, et al, Taylor & Francis, Productivity Press, New York, NY.




Selecting a Most Useful Predictive Model

9

9

## Stat-Ease Training: Sharpen Up Your DOE Skills



-  [Modern DOE for Process Optimization \(public or private\)](#)
-  [Mixture Design for Optimal Formulations \(public or private\)](#)
-  [Designed Experiments for Specific Industries \(private only\)](#)

Individuals	Teams (6+ people)
Improve your DOE skills	Choose your own date & time
Ideal for novice to advanced	Customize via select case studies



Learn more & then register:

[www.statease.com](http://www.statease.com)

Contact:

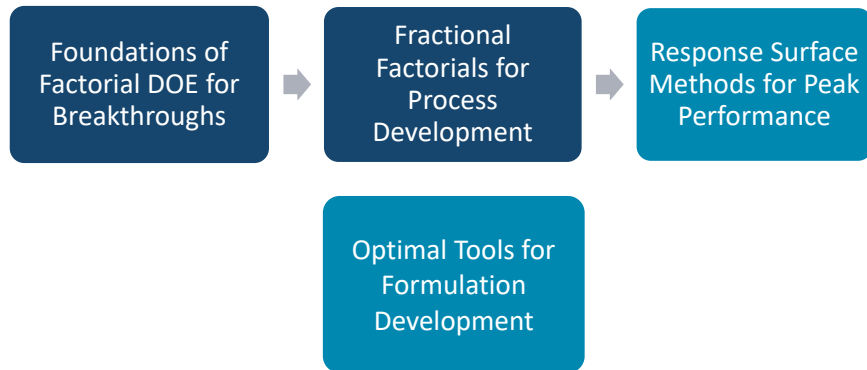
[workshops@statease.com](mailto:workshops@statease.com)

Selecting a Most Useful Predictive Model

10

10

Stat-Ease Training:  
*New=>* ESSENTIALS Series



*Two half-days of intense hands-on learning*

[www.statease.com/training/workshops](http://www.statease.com/training/workshops)

Selecting a Most Useful Predictive Model

11

11

**StatEase**  
statistics made easy

*Make the most from every experiment!<sup>SM</sup>  
via well-designed screening DOEs.*

Mark J. Anderson, Engineering Consultant  
Stat-Ease, Inc., Minneapolis, MN  
[mark@statease.com](mailto:mark@statease.com)  
[www.linkedin.com/in/markstat/](https://www.linkedin.com/in/markstat/)

Stay on for  
some chat  
if you like.

12