Strategies for Near Replicates in Response Surface Analysis

Thomas Bzik
Statistical Consultant
bziktj@verizon.net

Peter Parker
National Aeronautics and Space Administration
Hampton, VA 23681
peter.a.parker@nasa.gov

61st ASQ/ASA Fall Technical Conference, October 5-6, 2017, Philadelphia, PA

Pure-Error in Design of Experiments

Pure-error estimates are:

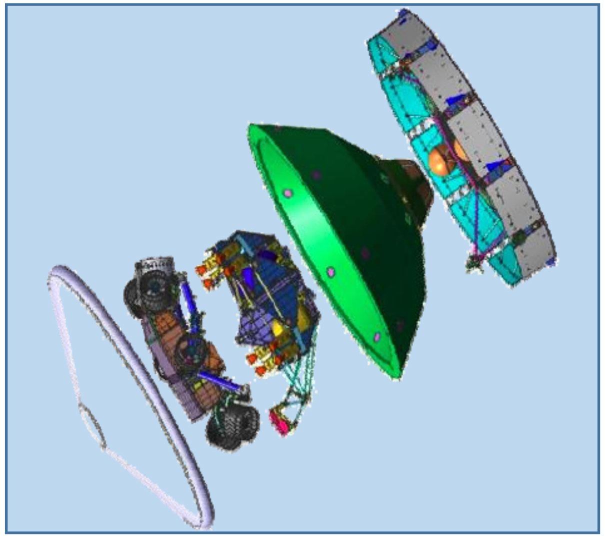
- Based on exact replicates - powerful and highly desirable in practice
  - Estimate of experimental variability, model-independent
  - Lack-of-Fit Testing
- Very easy to explain and approachable to non-statisticians

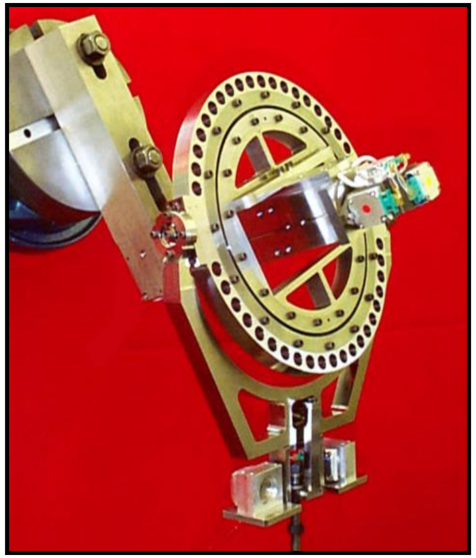
However,

- In practice, precise genuine replicates are often not achieved
  - Lack control precision/efficiency to set the factor levels (independent variable)
- Often, we can measure the factor levels of intended replicates, thereby resulting in “near replicates”

Relevant in Characterization Experiments

- High-precision measurement systems with accurate set-point measurements (feedback/monitor), but not controlled
- Lack of fit is important to isolate in the total uncertainty (residual)
- Guides strategic investment in improving calibration models
  - Invest in the apparatus
  - Explore high-order designs/models





Some Approaches in the Literature

- Montgomery, Martin, and Peck (1980), "**Interior Analysis of the Observations in Multiple Linear Regression**," Journal of Quality Technology, 12(3), pg. 165-173.
- Joglekar, Schuenemeyer, and LaRiccia (1989), "**Lack-of-Fit Testing When Replicates are not Available**," The American Statistician, 43(3), pg. 135-143.
- Su and Yang (2006), "**A Note on Lack-of-Fit Tests for Linear Models Without Replication**," Journal of the American Statistical Association, 101, pg. 205-210
- Christensen (1989) "**Lack-of-Fit Test Based on Near or Exact Replicates**," The Annals of Statistics, 17, pg. 673-683.
- Christensen and Sun (2010) "**Alternative Goodness-of-Fit Tests for Linear Models**," Journal of the American Statistical Association, 105, pg. 291-301.

From practice, none have been satisfactory solutions, nor are they widely used

Our Focus – What Can Be Done With Near Replicates?

Naïve Approach

Just pretend they are true replicates - Could lead to overestimation of PE depending on the transfer function from x to y. Could quantitatively evaluate propagation of error.

Correct Near Replicates into Estimated True Replicates

We will propose a methodology to correct the near replicates in a manner that allows PE to be estimated as if you had pure replicates. Then we will evaluate the performance of the proposed methodology.

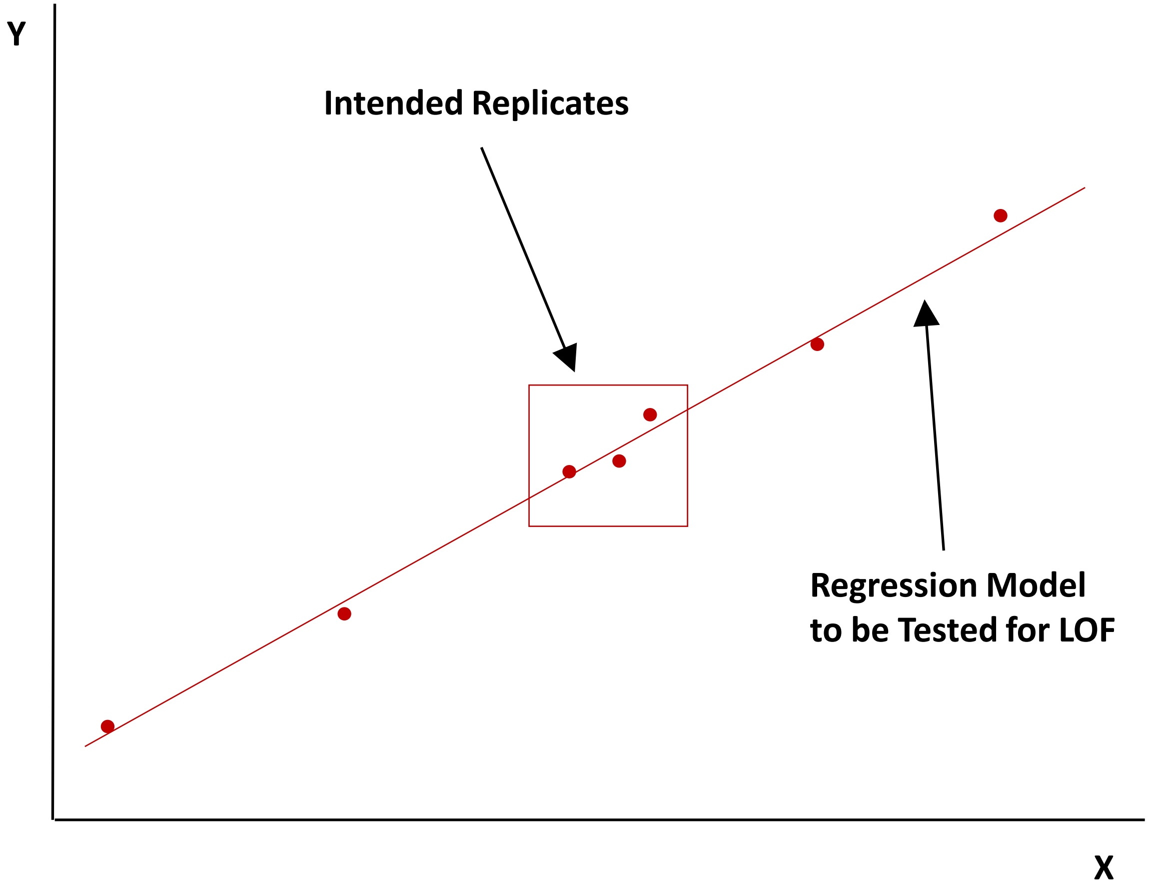
Presentation Outline

- Introduce our methodology to convert Near to Estimated replicates
- Simulation Study to test methodology performance and robustness
  - Models
    - Linear Model without Lack of Fit, then introduce Lack of Fit
    - Three-factor models (ME+2FI) without LoF, with LoF
  - "Near" definition and insights on practical ranges
  - Replication Strategy (Center, Edge)
  - Lack of Fit Structure Sensitivity
- Summary and Practitioner Guidance

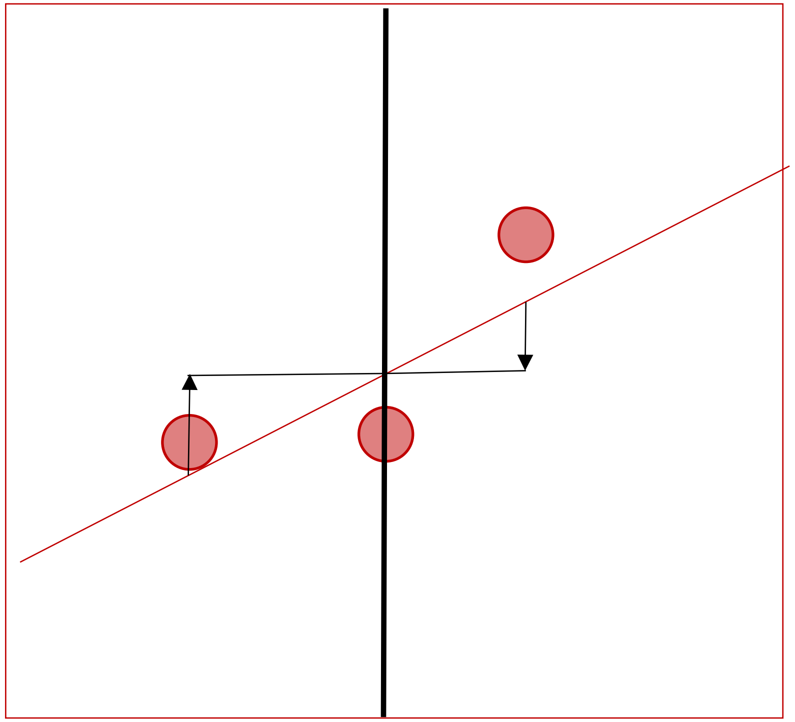
The Correction Methodology: Description

1. Fit the regression model that is intended to be checked for Model LOF
2. Identify your set(s) of near replicates (typically intended replicates)
3. For each set of near replicates:
   a) Calculate the independent variable center of mass for the set
   b) Use the regression model to correct each point's observed response to the center of mass (see example)
   c) Use these corrected values to estimate PE for each set
   d) Pool PE estimates if there are multiple sets of near replicates
   e) Corrected values are <u>only</u> used for PE estimation, nothing else

The Correction Methodology: Basic Visualization



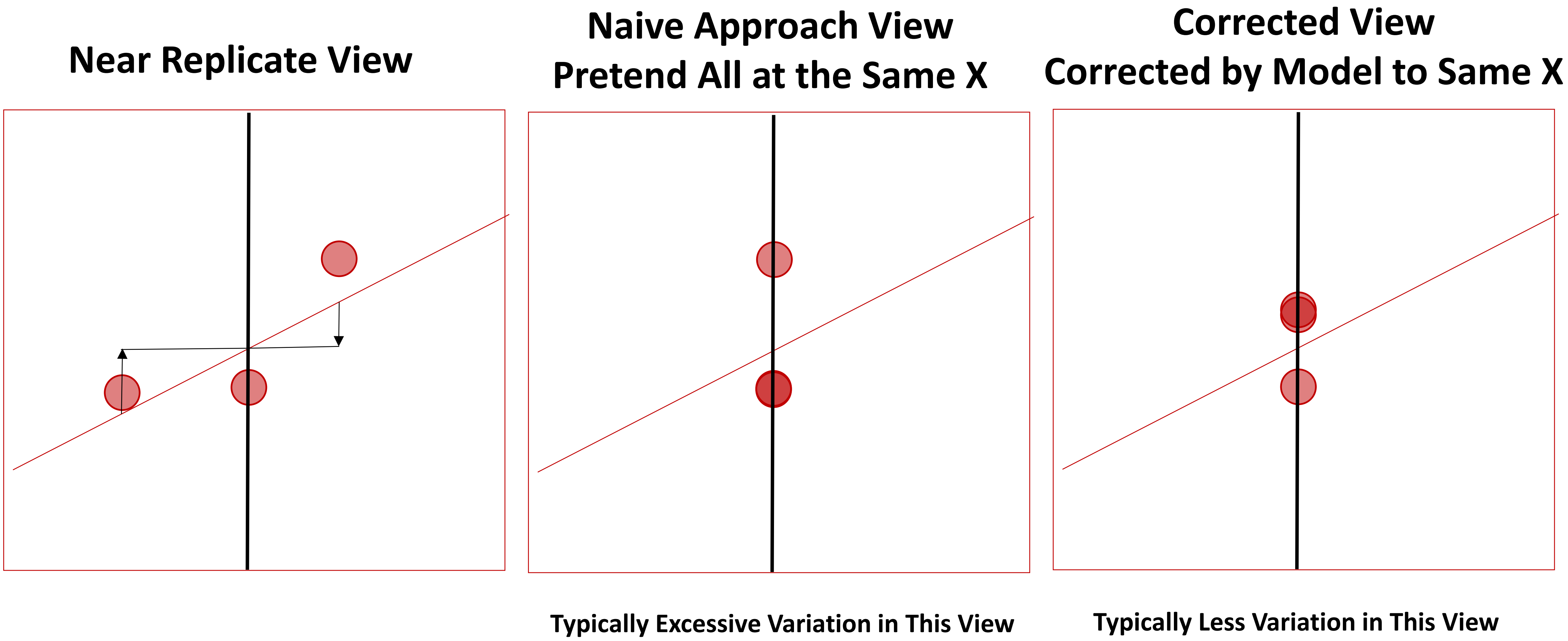
Point on left will be adjusted upward by impact of slope difference from the center of mass
Point on right will be adjusted downward by impact of slope difference from the center of mass
Midish point will be barely adjusted upward



Center of Mass

After adjustment, Y variation from estimated model due to nearness is removed

The Correction Methodology: Basic Visualization



Example of Correction for a Linear Model

X	Y
1.0	15.0
1.8	19.0
2.0	20.0
2.2	21.0
3.0	25.0

True Model: Y = 10 + 5X (without any simulated error)

Fitted Model: Y = 10 + 5X

The near replicates were three attempts to get X=2

Naïve estimate of PE variance = Variance (19, 20, 21) = 1

X	Y	X Offset	Adjust Amount	Adjusted Y
1.8	19.0	0.2	1	20.0
2.0	20.0	0.0	0	20.0
2.2	21.0	-0.2	-1	20.0

Center of mass - Near Replicates = Mean(1.8, 2.0, 2.2) = 2.0
(In practice, the center of mass will rarely be equal to the target)

X Offset = Center of Mass - X

Adjust Amount = X offset * est. slope = X offset * 5

Adjusted Y = Y + Adjust Amount

Corrected PE variance = Variance (20, 20, 20) = 0

Correction for more complex models uses all slope estimates including X in the model

The Correction Methodology: Intuitive Expectations

1. **Very robust if no Model LOF present (more exact)**
2. **Useful improvement if Model LOF present (more approximate)**

What the method actually delivers will be explored in subsequent detail

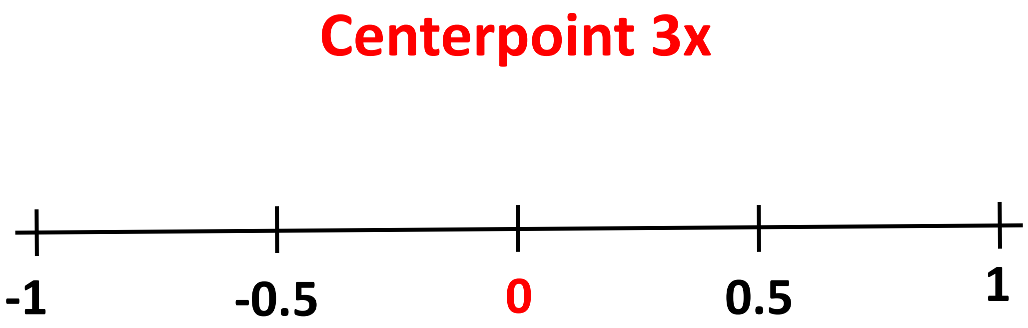
Testing the Correction Methodology

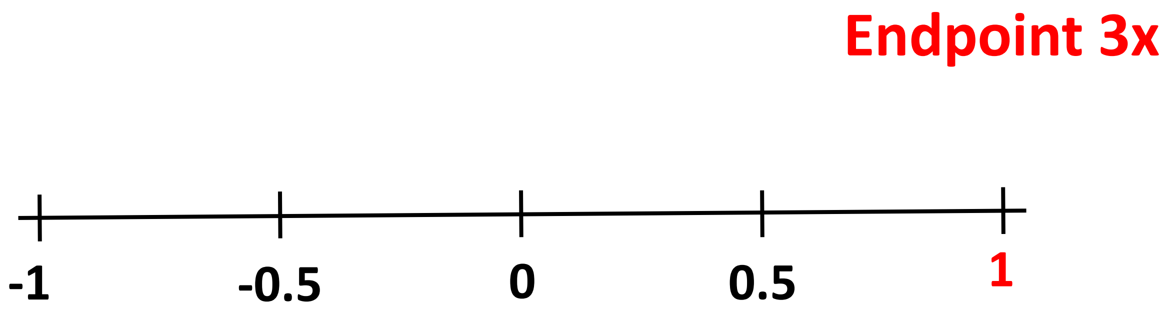
Simulation study will explore polynomial models of varying complexity and varying degrees of LOF to understand both benefits and potential weaknesses of the methodology.

Reported results are based on the average of many simulations of the entire process of correction method application.

We will start with lower dimensionality and increase this in later studies.

The Design Layouts Used for One-Dimensional Simulations
Replicated Experiment Target is in Red





Each experiment is run once except for the red colored experiment which is replicated three times. Note the purposeful consideration of method ability to handle relatively low PE (df =2) as well as only 3 df for LOF. This was part of stressing the correction methodology. Increasing any of the two df sources above made the method behave better than described.

Center Replicated, No Model LOF

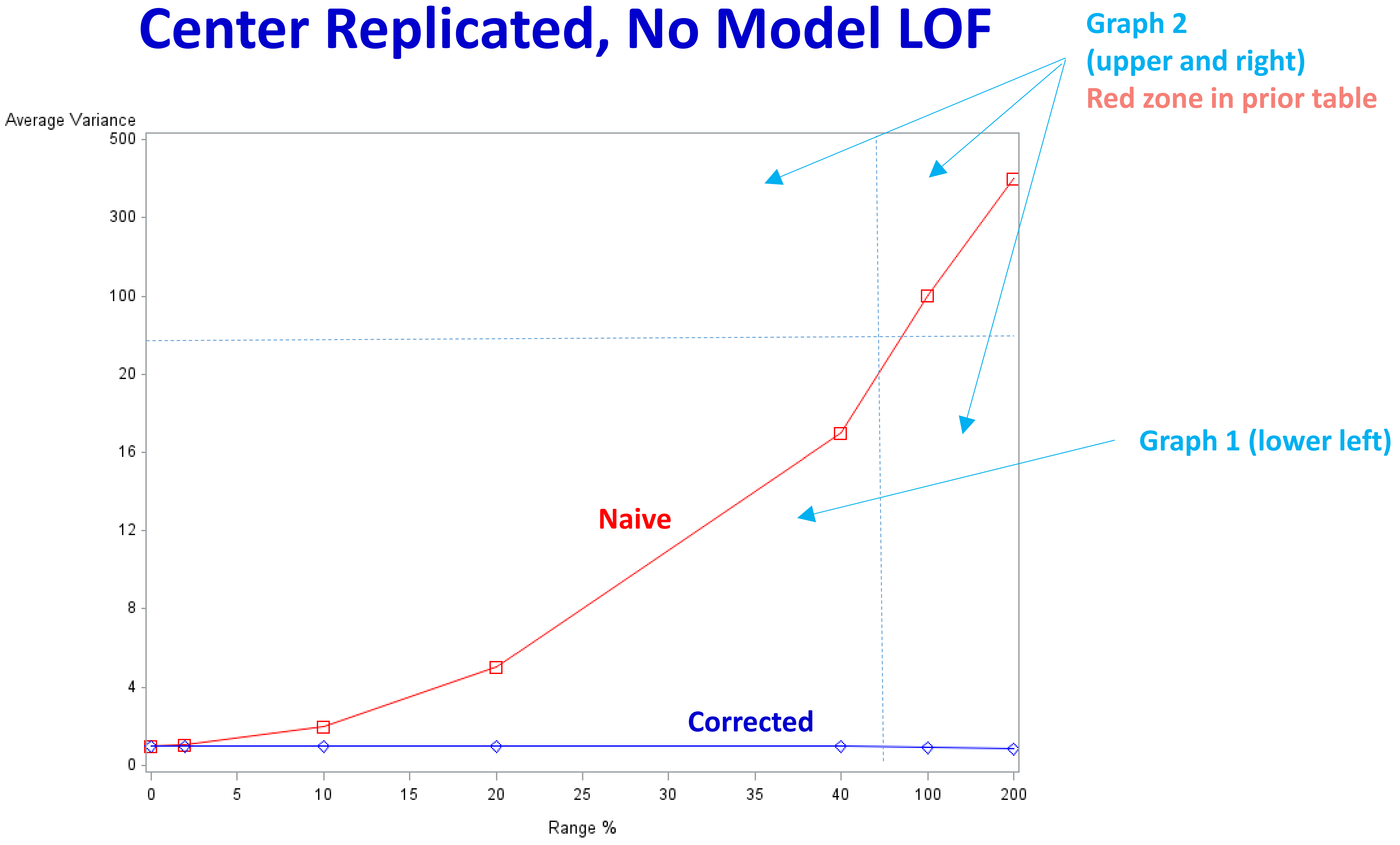
Linear Model, No Model LOF, Replication at Center		
Model: Y = 100 + 20*X + Ɛ where Ɛ ~ N(0,1), MSE=1		
Simulated Near Replicate Variation in X		
Range %	**Naïve Variance**	**Corrected Variance**
0%	0.9994	0.9994
2%	1.0427	1.0018
10%	2.0055	1.0009
20%	5.0073	0.9980
40%	17.0053	0.9870
100%	100.9540	0.9308
200%	400.7180	0.8365

Range % (variation in near replicates)
= 4 times the std. dev. of X divided by the DOE's range expressed as a percentage

Naïve Variance:
The variance computed by pretending the near replicates are true replicates. It should be close to one where the naïve approach works best.

Corrected Variance
The variance after applying the correction methodology. It should be close to one where the corrected method works.

Center Replicated, No Model LOF



What is a Possible Rational Range of Application?

Std Dev X	Range %	Rational?
0	0%	NA
0.01	2%	Yes
0.05	10%	Yes
0.1	20%	Maybe
0.2	40%	Stretch
0.5	100%	No
1	200%	No

Higher dimensional models will only be examined in a more rational range

In the one control dimension cases the simulations range far beyond any rational construction of near replicates. The only reason why is to probe the method beyond rational limits of application to better understand its limitations. Judge method performance in its realm of likely application.

End Replicated (X = 1), No Model LOF

Linear Model, No Model LOF, Replication at End (1)		
Model: Y = 100 + 20*X + Ɛ where Ɛ ~ N(0,1), MSE=1		
Simulated Near Replicate Variation in X		
Range %	**Naïve Variance**	**Corrected Variance**
0%	1.0018	1.0018
2%	1.0427	1.0018
10%	2.0055	1.0012
20%	5.0073	0.9993
40%	17.0053	0.9919
100%	100.9540	0.9483
200%	400.7180	0.8553

Analogously good results when compared to centerpoint replicated example

Center Replicated (X = 0), Model LOF

Nonlinear Model, Model LOF, Replication at Center		
Model: $Y = 100+20*((x+2)**1.3) + \varepsilon$ where $\varepsilon \sim N(0,1)$, MSE=1		
Fit Model: $Y = a + bX$		
Simulated Near Replicate Variation in X		
Range %	Naïve Variance	Corrected Variance
0%	1.0018	1.0018
2%	1.1057	1.0018
10%	3.5682	1.0012
20%	11.2497	1.0002
40%	41.9150	1.0102
100%	254.2128	1.4557
200%	919.8681	4.1044

Worked well, but eventual bias in correction outside rational application range

End Replicated (X = 1), Model LOF

Nonlinear Model, Model LOF, Replication at End (X=1)		
Model: Y = 100+20*((x+2)**1.3) + Ɛ where Ɛ ~ N(0,1), MSE=1		
Fit Model: Y = a + bX		
Simulated Near Replicate Variation in X		
Range %	Naïve Variance	Corrected Variance
0%	1.0018	1.0018
2%	1.1341	1.0036
10%	4.2741	1.0410
20%	14.0727	1.1535
40%	53.2253	1.5599
100%	326.0179	3.2791
200%	1271.3100	6.0018

Works passably with only mild bias inside rational application range
Eventual under correction is form of bias and is more rapid than observed at the centerpoint

Bias in Testing – Be Mean to the Correction Methodology

The prior testing was performed under moderately harsh test conditions (high model R-square, e.g., 99.5%). Increasing the simulated model MSE (lowered R-square) improves the relative method performance for all illustrated scenarios.

Higher df also helps method performance in all tested scenarios.

Our goal is to probe the methodology's behavior. Good, bad or approximate can be judged by yourselves depending on application.

We feel pretty comfortable with the linear case results.

Jumping to Three Dimensions: The Design Layouts

2^3 factorial with replicated center (3x)

X1	X2	X3
-1	-1	-1
-1	-1	1
-1	1	-1
-1	1	1
1	-1	-1
1	-1	1
1	1	-1
1	1	1
0	0	0
0	0	0
0	0	0

2^3 factorial with (1, 1, 1) replicated corner (3x) and center

X1	X2	X3
-1	-1	-1
-1	-1	1
-1	1	-1
-1	1	1
1	-1	-1
1	-1	1
1	1	-1
0	0	0
1	1	1
1	1	1
1	1	1

Center Replicated, No Model LOF

Case: True Model is Multivariate, No Model LOF, Replication at Center				
Simulated Model: Y = 50 + 20x1 + 5x2 - 5x3 - 5x1*x2 + Ɛ where Ɛ ~ N(0,1), MSE=1				
Fitted Model: Y = a + bx1 + cx2 + dx3 + ex1*x2				
Simulated Near Replicate Variation in X1				
Range %	**Naive Variance**	**Corrected Variance**	**Naive Power**	**Corrected Power**
0%	1.0018	1.0018	0.050	0.050
2%	1.0427	1.0018	0.048	0.050
5%	1.2539	1.0017	0.044	0.050
10%	2.0055	1.0015	0.031	0.050
15%	3.2567	1.0011	0.020	0.050
20%	5.0073	1.0005	0.014	0.050
40%	17.0053	0.9968	0.004	0.050

- **With no LOF the methodology remains robust in the presence of more dimensionality**
- **Some results match some prior results due to model choice, near replication var. and use of same seed**
- **Match is interesting in that increased dimensionality was irrelevant in method impact (narrow).**

Center Replicated, Model LOF

Case: True Model is Multivariate, Model LOF, Replication at Centerpoint				
Simulated Model: Y = 50 + 20x1 + 5x2 - 5x3 - 5x1*x2 + 5x1*x1 + Ɛ where Ɛ ~ N(0,1), MSE=1				
Fitted Model: Y = a + bx1 + cx2 + dx3 + ex1*x2 (Incomplete Model)				
Simulated Near Replicate Variation in X1				
Range %	Naive Variance	Corrected Variance	Naive Power	Corrected Power
0%	0.9998	0.9998	0.521	0.521
2%	1.0378	0.9974	0.508	0.521
5%	1.2453	0.9940	0.448	0.525
10%	1.9959	0.9966	0.316	0.527
15%	3.2509	1.0059	0.210	0.526
20%	4.9967	1.0025	0.146	0.533
40%	17.0095	1.1106	0.053	0.538

- **With LOF the center point replicates still worked extremely well**
- **Slightly theoretically imperfect in last line, but still very useful**

Corrected Power – Useful Results

The prior results used the following in its simulation, *how important is this?*

Simulated Model: $Y = 50 + 20x1 + 5x2 - 5x3 - 5x1*x2 + 5x1*x1 + \varepsilon$ where $\varepsilon \sim N(0,1)$, MSE=1

Changing coefficients in black does not meaningfully impact *corrected power* even if all are set to 0, even with sign and coefficient magnitude changes. Once the model being fit is of specified form, the variance named, the DOE named and the lack of fit coefficient(s) set, power appears stable within likely simulation error (these results are only fully relevant to the Range % = 0 line).

Corrected power is distinctly a function of LOF df, PE df, variance, the DOE and the point(s) being replicated and their frequency. More df = more power, lower variance = more power, bigger absolute LOF coefficient = more power.

(1, 1, 1) Replicated, No Model LOF

Case: True Model is Multivariate, No Model LOF, Replication at (1, 1, 1)				
Simulated Model: Y = 50 + 20x1 + 5x2 - 5x3 - 5x1*x2 + Ɛ where Ɛ ~ N(0,1), MSE=1				
Fitted Model: Y = a + bx1 + cx2 + dx3 + ex1*x2				
Simulated Near Replicate Variation in X1				
Range %	**Naive Variance**	**Corrected Variance**	**Naive Power**	**Corrected Power**
0%	1.0024	1.0024	0.049	0.049
2%	1.0252	1.0045	0.049	0.050
5%	1.1433	1.0200	0.045	0.048
10%	1.5600	1.0673	0.038	0.047
15%	2.2716	1.1423	0.028	0.047
20%	3.2489	1.2530	0.020	0.045
40%	9.9371	2.0005	0.007	0.031

- Gradual theoretical imperfection but still useful

(1, 1, 1) Replicated, Model LOF

Case: True Model is Multivariate, Model LOF, Replication at (1, 1, 1)				
Simulated Model: Y = 50 + 20x1 + 5x2 - 5x3 - 5x1*x2 + 5x1*x1 + Ɛ where Ɛ ~ N(0,1), MSE=1				
Fitted Model: Y = a + bx1 + cx2 + dx3 + ex1*x2 (Incomplete Model)				
Simulated Near Replicate Variation in X1				
Range %	**Naive Variance**	**Corrected Variance**	**Naive Power**	**Corrected Power**
0%	1.0003	1.0003	0.284	0.284
2%	1.0622	1.0030	0.271	0.284
5%	1.3925	1.0190	0.217	0.282
10%	2.5634	1.0626	0.126	0.276
15%	4.5251	1.1379	0.076	0.270
20%	7.2733	1.2431	0.051	0.266
40%	26.0577	1.9844	0.019	0.232

- Centerpoint worked better in terms of power, but not bias in this example.
- Gradual theoretical imperfection but method still useful.
- This is the core piece of the method, even if you eventually pollute the correction with differential LOF, it still yields a much better test than the naïve alternative.

Let’s Up the Method Stress Another Way

The next simulation will examine a tougher near replicate scenario. We’ll let the near replication extend over two of the variables, X1 and X2 rather than just X1. This will necessarily increase the differential impact of any LOF on the methodology.

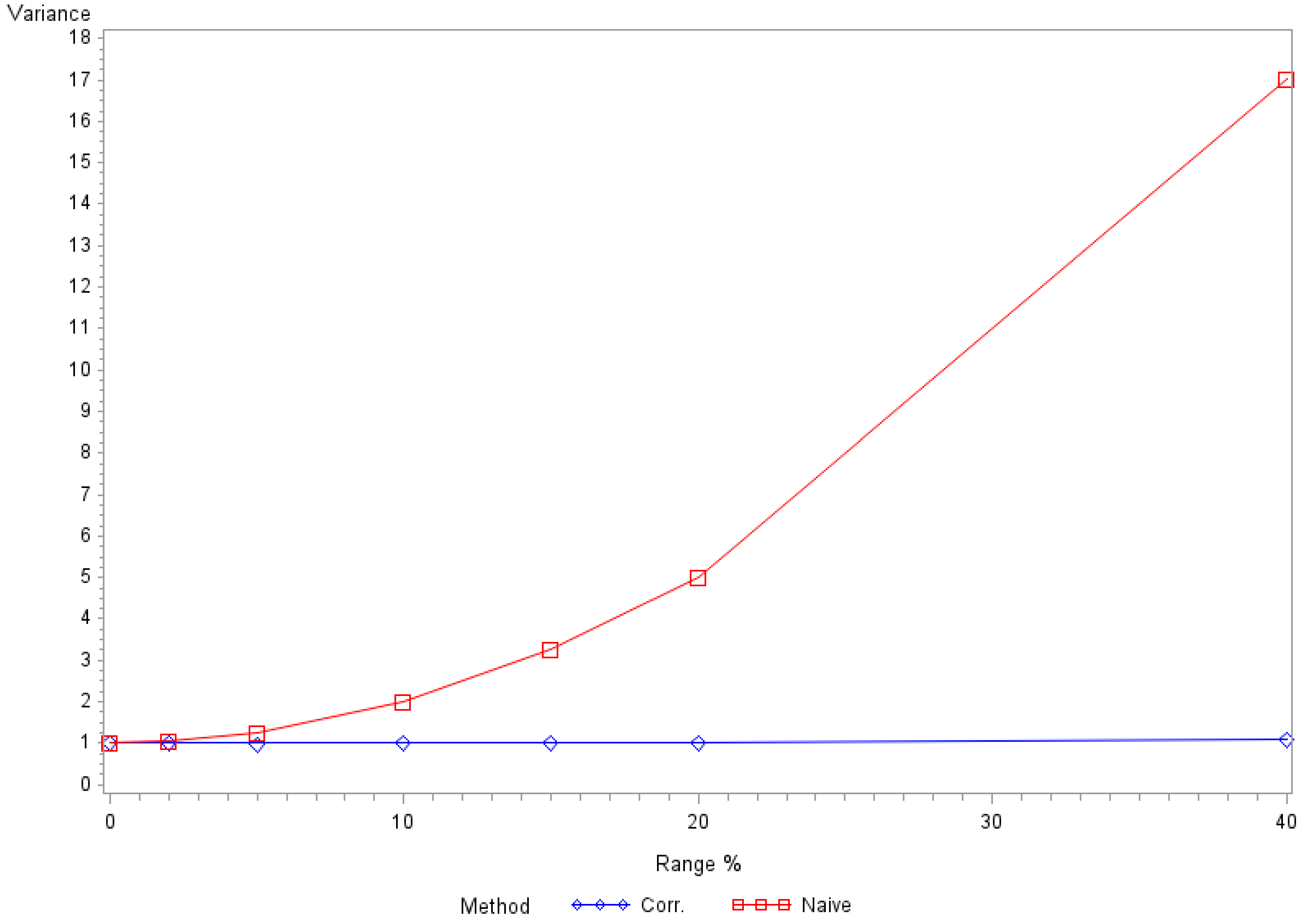
Next tables will narrow our focus to cases where the model has LOF rather than when it does not.

Center Replicated, Model LOF, Variation in X1 and X2

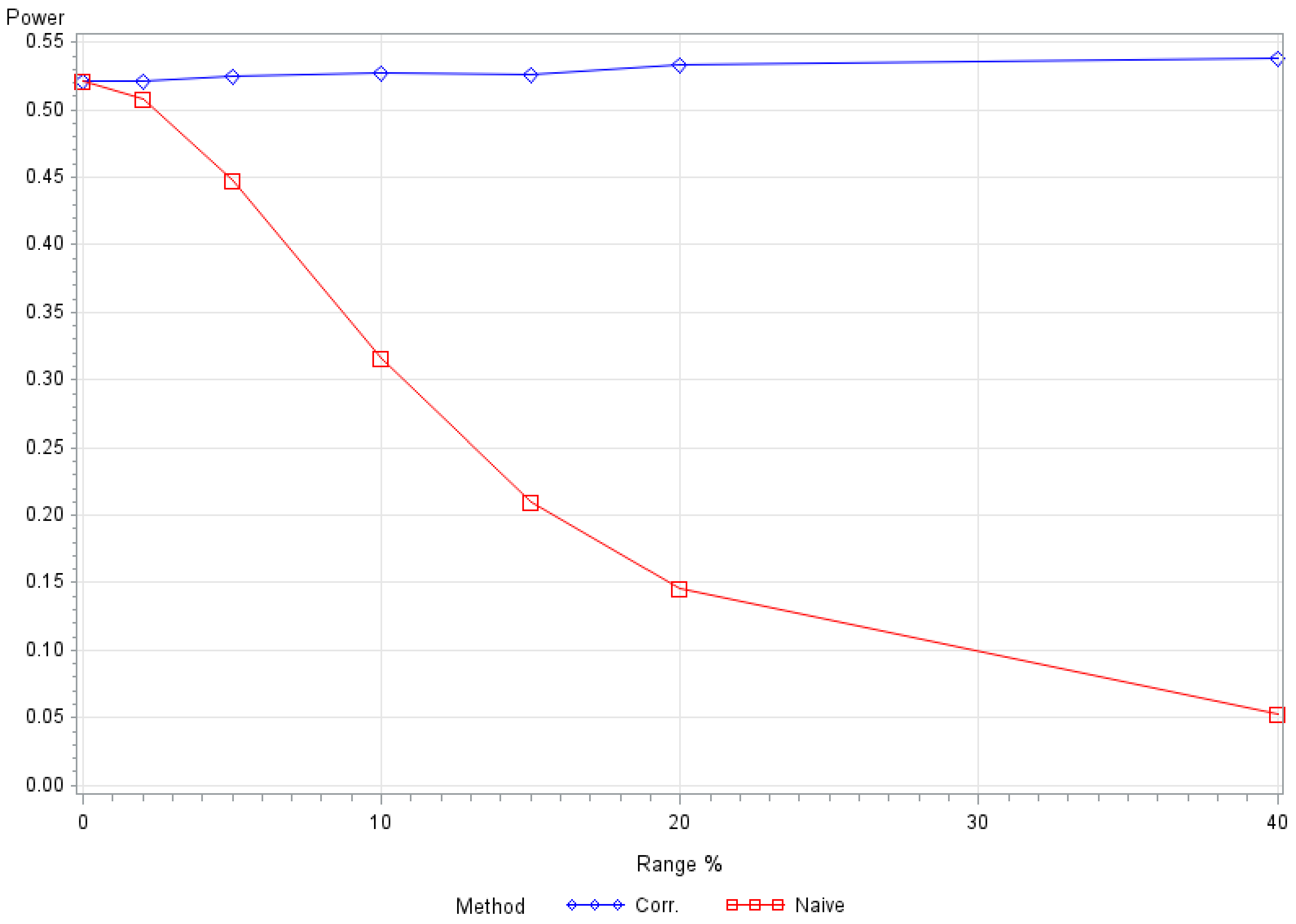
Case: True Model is Multivariate, Model LOF, Replication at Centerpoint				
Simulated Model: Y = 50 + 20x1 + 5x2 - 5x3 - 5x1*x2 + 5x1*x1 + Ɛ where Ɛ ~ N(0,1), MSE=1				
Fitted Model: Y = a + bx1 + cx2 + dx3 + ex1*x2 (Incomplete Model)				
Simulated Near Replicate Variation in both X1 and X2				
Range %	Naive Variance	Corrected Variance	Naive Power	Corrected Power
0%	0.9994	0.9994	0.521	0.521
2%	1.0474	1.0046	0.505	0.520
5%	1.2632	0.9973	0.446	0.524
10%	2.0532	0,9957	0.305	0.527
15%	3.3803	1.0047	0.205	0.527
20%	5.2624	1.0110	0.141	0.526
40%	18.4622	1.2069	0.050	0.507

- **Holds up well, but can see start of gradual fade from variance of 1 in the last line in the table.**
- **Even there it provides a useful test of LOF, although not a theoretically perfect one**
- **Compare powers to next table, better here.**

Center Replicated, Model LOF, Variation in X1 and X2



Center Replicated, Model LOF, Variation in X1 and X2



(1, 1, 1) Replicated, Model LOF, Variation in X1 and X2

Case: True Model is Multivariate, Model LOF, Replication at (1, 1, 1)				
Simulated Model: Y = 50 + 20x1 + 5x2 - 5x3 - 5x1*x2 + 5x1*x1 + Ɛ where Ɛ ~ N(0,1), MSE=1				
Fitted Model: Y = a + bx1 + cx2 + dx3 + ex1*x2 (Incomplete Model)				
Simulated Near Replicate Variation in Both X1 and X2				
Range %	**Naive Variance**	**Corrected Variance**	**Naive Power**	**Corrected Power**
0%	1.0014	1.0014	0.284	0.284
2%	1.0641	1.0072	0.269	0.283
5%	1.3921	1.0321	0.216	0.278
10%	2.5679	1.1260	0.126	0.264
15%	4.5266	1.2853	0.078	0.248
20%	7.2454	1.5006	0.051	0.226
40%	26.1820	3.0972	0.019	0.166

- Center point replication worked better in this example.
- Always useful vs. naïve, but a price for strong enough differential local LOF. Still better with correction vs. missing LOF!
- Gradually imperfect but still useful.

Pure Error Estimation with LOF – The Big Picture

Naïve: PE estimates inflated by how near replicate variation is transmitted

1. Transmission through uncorrected model (often substantial)
2. Transmission of locally differential LOF (often smaller)

Corrected: PE estimates inflated by how near replicate variation is transmitted

2. Transmission of locally differential LOF (often smaller)

The correction method operates in the *head space of regression analysis*, the unexplained variation left over after modeling (MSE). It performs best when the locally induced LOF (bias) is small relative to the total amount of unexplained variation. Near replicate's localized impact of LOF is smaller in that the impact of LOF is over the whole model. The correction methodology benefits from this.

Adaptive Analysis Possibilities

Consider a scenario in which near replicates, by being possibly too far apart or more weakly located, may have introduced too much differential LOF variation in the corrected result.

1. **Only use "better" sets of near replicates (closeness and / or location)**
2. **Use only subsets of the sets of near replicates (closeness)**
3. **Partition larger sets of near replicates into two or more internally nearer subgroups**

In all cases, if variance estimates consistently decline with adaptive analysis, especially in case 3 where all the data was still used, prefer the adaptive version of the analysis.

Unintended Replicate Possibility (Typically not RSM)

You have observational data without any replicated points, but you notice one or more clusters of points that are relatively close together.

Apply the correction methodology to such clusters of nearby points.

In this manner, an approximate LOF test can be constructed for observational data regression models! If it finds LOF, you probably have it. If it does not, it is not as certain a result, but is better than not trying to find LOF.

This realm of application may be more valuable than the RSM context.

Summary

Two witnessed past practitioner approaches to near replicates:

1) **Treat near replicates as replicates for the purpose of computing PE**
2) **Set near replicates to their target values prior to modeling and proceed as usual, even riskier than 1)**

Issues with 1): Inflated estimate of PE and lower power to detect LOF

Correction Method Philosophy: Find the "best" fix for near replicates when faced with LOF of unknown form

Summary

Demonstrated Correction Methodology

1) Is extremely robust when no model LOF is present
2) Fairly robust results when moderate LOF is present
3) Fairly robust to moderate variation in near replicate values
4) Even when imperfect, it still pulls us closer to a correct LOF judgement

We recommend that this correction methodology be added to your DOE toolkit.

Thank you for listening and questions are highly encouraged.