Optimization of Z Scores in Multiple Response DOEs for Robust Design

Christopher Stanard
GE CRD Applied Statistics Program

2001 ASA Albany, NY Spring Meeting:
Applications in Statistics – What's Hot / What's Next III
March 24, 2001

stanard@crd.ge.com ph: (518) 387-4131 fax (518) 387-5714

Abstract

Designed Experiments in industrial settings typically have multiple responses with multiple constraints and objectives. Particularly in robust design, it is important to optimize the means and minimize the variances.

This talk will demonstrate new procedures for accomplishing these tasks using the Six Sigma Z score metric.

Outline

- General Methodology & Discussion
- Example Problem
  - Problem Overview
  - Problem Steps
  - Problem Solution
- Key Observations
- Bibliography

General Methodology & Discussion

Most Robust DOE Optimization Methods today employ methods to characterize the mean and standard deviation using factorial analysis, response surface regression equations, Taguchi loss functions, or the like. Overlaid contour plots techniques find the "sweet spot" satisfying operation constraints and numerical optimization techniques find the best points when considering some subset of the desired problem characteristics.

Optimizing the Six Sigma Z-score (a DPMO or PPM index) as a response, however, can take into account the factor and response means, standard deviations, and operating constraints.

General Methodology (continued)

Since we don't generally have equations for two-sided Z-scores, however, we optimize the Z-lowers and the Z-uppers. Z-lower is essentially 3*Cpk_lower, and Z-upper is essentially 3*Cpk_upper. Optimizing on the equivalent Cpk constraints instead of Z-scores is essentially the same process. Recall Cpk_lower = (Mean-LSL)/(3*Std Dev.), and Cpk_upper = (Upr Spec Limit-Mean)/(3*Std Dev.) . These are very tractable formulas and should pose no problem to use. After finding the optimal point, points, or region, the actual Z can be determined using a Six Sigma scorecard such as the ROSTAR scorecards used in this presentation.

If deviation from target or mean is also important, one can add Cpm as another response to be optimized.

Y-bar and Std Dev(Y) responses may still be useful to include for informational and illustrative purposes.

General Methodology (continued)

If desired, Taguchi responses such as difference from target (reminiscent of Cpm) can be used. Common Inner-Outer Array S/N functions for smaller the better, nominal the best, larger the better, and signed target can also be used as additional factors. I recommend using still using Z with these, though it isn’t a problem to just add the specifications to the Taguchi equations if desired (i.e. replace y by y-LSL and USL-y).

The simplest example for illustrating this technique is a replicated DOE with constant variance and no transformations. In this case, I recommend using the Root MSE as the estimate of the standard deviation. (A less conservative approach might be to use the SE of the Mean, and some might try the SE of Prediction in some cases.) The technique is also applicable for DOEs with (variance stabilizing, normal fit, etc.) transformed responses, factorial responses with standard deviations possibly varying at different design points, DOEs where some form of the standard deviation is used as a response, Taguchi designs, Mixture designs, other types of general designed experiments, and more.

General Methodology (continued)

DOEs generally deal with normally distributed data, but a transformed log response or lognormal equivalent is also fine. It would make sense to extend this further to other parametric regression equation types with different error structures such as Weibull, extreme value, logistic, and so forth.

In cases where there it is possible to get further control factor measurements such as GRR, or historical, or industry standard data, it is further possible to add dispersion parameters to the DOE X factors and use propogation of error or/and simulation to get an even more robust measure of the Z-score taking into account the variation throughout the system. This would be applicable not only for DOE, but also historical regression and related areas.

Again, the distributions involved do not have to be normal. The ROSTAR (RObust Statistics for All Responses) Six Sigma scorecards handle a number of parametric distributions as well as empirical distributions.

Example Problem Overview

- Several Input X Variables
- Several Response Y Variables
- Limited Operating Regions for Y Variables That Must Be Simultaneously Optimized
- Related Z-scores fromYs, Specs, and Standard Deviations

Assumptions

- iid ~ $N(0,\sigma)$, const. var., & other usual assumptions for Xs & errors
- Y variables are uncorrelated

Note: In industrial situations, we find these assumptions are often relaxed. In Robust Design, replicated points and Taguchi inner-outer arrays often help in modeling equations for non-constant variance. Equations for mean, standard deviation, and/or a Taguchi Signal Response are often modeled.

General Steps

- Develop Equations
  - Design, Set Up, Perform, & Evaluate DOE to Get Response Surface Equations
  - Error Checking and Model Validation
  - Specify Targets and Constraints
  - Weight Relative Importance of Optimizing Each Y Variable
- Optimization (Response Mean, Z-score)
  - Numerical Multiple Response Optimization
  - Overlaid Contour Plots

Distance Metrics

- In order to Optimize Multiple Responses (Y's), They May Be Weighted and Combined Together in a Single "Distance Metric"
- Optimizing the Distance Metric Optimizes the Y's
- Weights of Individual Responses May Be Varied

For Literature Review, See "Multiple Response Optimization Using Designed Experiments – A Practical Example," The 1998 ASA Albany Mini-Conference, What Hot/What's Next III, March 28, 1998, GE CRD, Schenectady, NY 12309.

Derringer and Suich Distance Metrics

- Bounded Responses
- Goals for Each Response May Be:
  - Equals Target
  - Minimize
  - Maximize
- Units of Distance: Desire $\in [0,1]$
  - Higher Values Are More Desirable
  - Overall Combined Distance Is the Geometric Mean of the Individual Distances

Sample Problem

Simulated Plastics Manufacturing Example

Description: 5 factors, 3 responses

Uniform Precision Central Composite Design (Quadratic w/ 2 Fact. Interactn's, No Aliases)

1/2 Fraction, k = 5, α = 2, 2 replicates, 12 center + 32 corner + 20 axial points= 64 tot. points

G-Efficiency = 74.7%, Scaled D-Optimality Criterion = 1.517

(Based on 2^{5-1}_{V} design with 2 replicates, axial and center points added)

Y Responses:		**Index 1**
		Viscosity
		Gloss
X Factors:	**A**	**Size**
	B	**Linking**
	C	**Loading**
	D	**MolWt**
	E	**Graft**

(Designed & Analyzed Using Design Expert Software by Stat-Ease)

Sample Problem (continued)

Responses Mean Equations In Terms of 1, -1 Coded X Variables

Equation 1:	**Index 1**	**= 12.88 + 0.80 A - 1.5 C + 0.61 AC**, Std Dev = 0.8246
Equation 2:	**Viscosity**	**= 191.54 + 19.73 C - 71.29 E - 64.43 CE**, Std Dev = 10.86
Equation 3:	**Gloss**	**= 80 + 3.92 A - 2.31 B + 0.84 C - 0.80 AC - 0.23 BC**, Std Dev = 0.4359

For Optimization of Response Means:

Response	LSL Minimum	USL Maximum	Target Optimiz. Goal	Weight (1-5)
Index 1	13	15	14	3
Viscosity	120	190	160 (off ctr)	3
Gloss	72	84	74 (off ctr)	3

Note on Weighting: 1 is least impt. 3 is nominal, 5 is most impt (for this example)

LSL - Lower Specification Limit (for Product Quality)

USL - Upper Specification Limit (for Product Quality)

Sample Problem (continued)

<u>Z Score Capability Equations In Terms of 1, -1 Coded X Variables</u>

Index 1 Z Lwr	**= ((12.88 + 0.80 A - 1.5 C + 0.61 AC) - 13)/0.8246**
Index 1 Z Upr	**= (15 - (12.88 + 0.80 A - 1.5 C + 0.61 AC))/0.8246**
Viscosity Z Lwr	**= ((191.54 + 19.73 C - 71.29 E - 64.43 CE) - 120)/10.86**
Viscosity Z Upr	**= (190 - (191.54 + 19.73 C - 71.29 E - 64.43 CE))/10.86**
Gloss Z Lwr	**= (84 - (80 + 3.92 A - 2.31 B + 0.84 C - 0.80 AC - 0.23 BC))/0.4359**
Gloss Z Upr	**= ((80 + 3.92 A - 2.31 B + 0.84 C - 0.80 AC - 0.23 BC) - 74)/0.4359**

<u>For Optimization of Z Score Indices:</u>

<u>Z Score Index</u>	<u>Minimum</u>	<u>Weight (1-5)</u>
Index 1 Z Lwr	1	5
Index 1 Z Upr	1	5
Viscosity Z Lwr	2	5
Viscosity Z Upr	2	5
Gloss Z Lwr	4.5	5
Gloss Z Uprb	4.5	5

Note on Weighting: 1 is least impt. 3 is nominal, 5 is most impt (for this example)

Note on Z Scores: No “Z-shift” is used for this example. Zlwr = 3 Cpk_lwr & Zupr = 3 Cpk_upr

Sample Problem (continued)

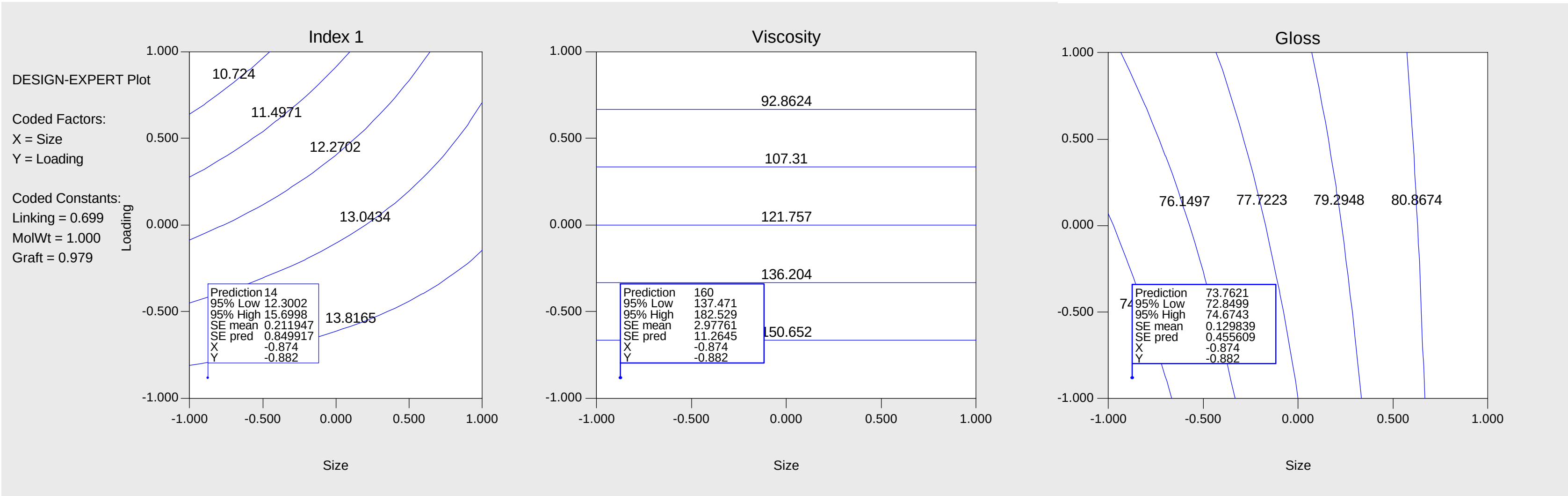
Response Optimization -- Not including Z-Scores

Numerical Optimization Using Distance Metrics Found Several Points to Fit the Criteria. Here's one of them:

X Variables					**Predicted Response Values**			
A	**B**	**C**	**D**	**E**	**Y1**	**Y2**	**Y3**	**Distance**
Size	**Linking**	**Loading**	**MolWt**	**Graft**	**Index 1**	**Visc.**	**Gloss**	**Desirability**
-0.874	0.699	-0.882	*1.000*	0.979	14	160	73.76	0.961

Contour Plots For the Individual Responses: (A Prelude to Overlaid Cntr Plots)

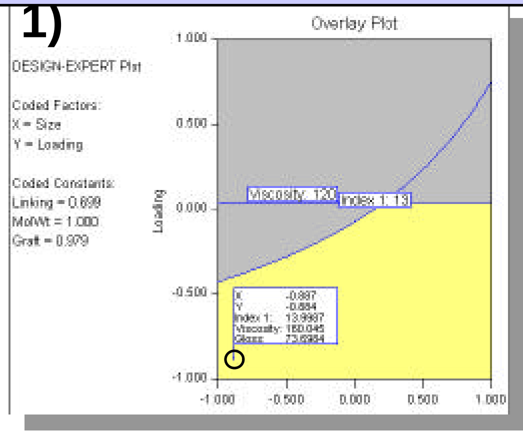
Index 1 **Viscosity** **Gloss**



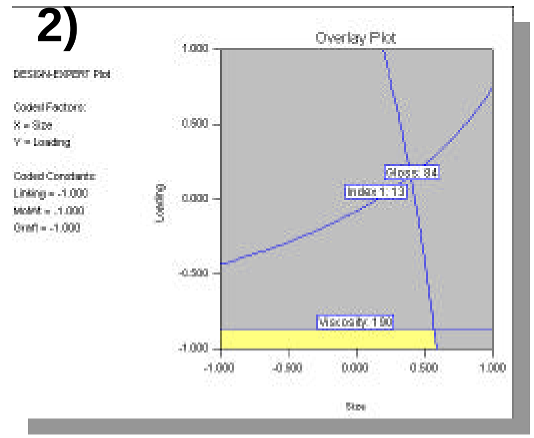
Sample Problem (continued)

Response Optimization -- Not including Z-Scores

1)



2)



Multiple Response Overlaid (or Overlay) Contour Plots:

X Axis: Loading, Y Axis: Size

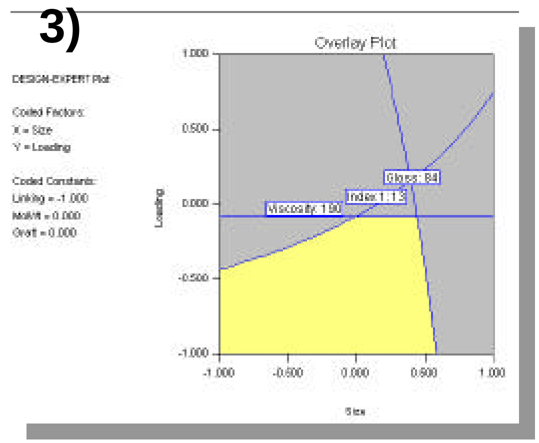
1) Optimized Point

2) Factors B, D, & E at -1 Level

3) Factor B at -1 Level,

Factors D & E at 0 Level

3)



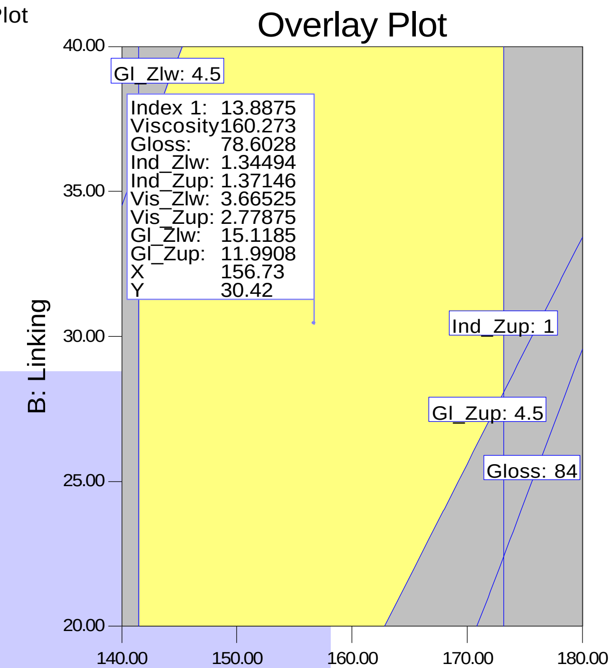
Sample Problem (continued)

Overlaid Contour Plot in the "Sweet Spot" with Optimized Z Scores Included

DESIGN-EXPERT Plot

Overlay Plot
X = A: Size
Y = B: Linking

Actual Factors
C: Loading = 5.75
D: MolWt = 97.00
E: Graft = 230.00

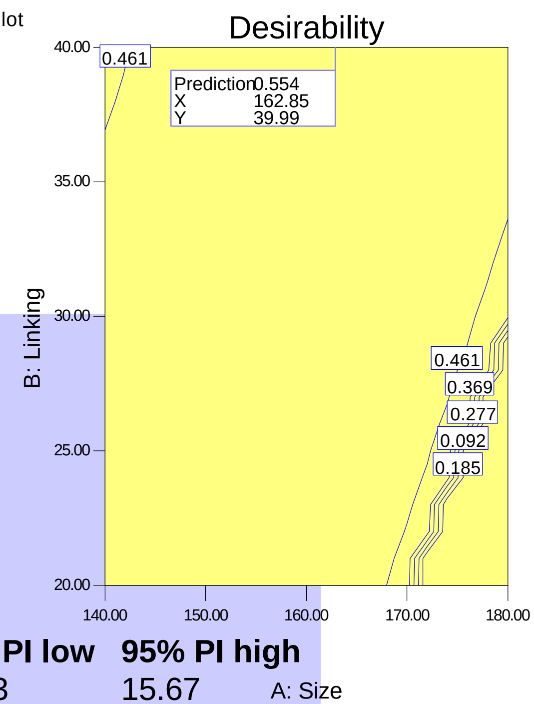


Name	**Level**	**Low Level**	**High Level**	**Std. Dev.**	
A	Size	156.00	140.00	180.00	0.000
B	Linking	30.00	20.00	40.00	0.000
C	Loading	5.75	5.00	10.00	0.000
D	MolWt	97.00	50.00	100.00	0.000
E	Graft	230.00	148.75	246.25	0.000

	Prediction	**SE Mean**	**95% CI low**	**95% CI high**	**SE Pred**	**95% PI low**	**95% PI high**
Index 1	13.87	0.14	13.60	14.15	0.83	12.21	15.54
Viscosity	160.27	2.22	155.82	164.72	11.09	138.09	182.45
Gloss	78.53	0.072	78.38	78.67	0.44	77.64	79.41
Ind_Zlw	1.33						
Ind_Zup	1.39						
Vis_Zlw	3.67						
Vis_Zup	2.78						
Gl_Zlw	14.95						
Gl_Zup	12.07						

Sample Problem (continued)

Similar Point, also in the "Sweet Spot," Identified Directly Through Numerical Optimization Technique



Factor	Name	Level	Low Level	High Level	Std. Dev.
A	Size	162.85	140.00	180.00	0.000
B	Linking	39.99	20.00	40.00	0.000
C	Loading	5.75	5.00	10.00	0.000
D	MolWt	94.43	50.00	100.00	0.000
E	Graft	230.46	148.75	246.25	0.000

	Prediction	SE Mean	95% CI low	95% CI high	SE Pred	95% PI low	95% PI high
Index 1	14.00	0.13	13.73	14.27	0.83	12.33	15.67
Viscosity	160.03	2.24	155.55	164.50	11.09	137.84	182.21
Gloss	77.92	0.11	77.70	78.14	0.45	77.02	78.82
Ind_Zlw	1.48						
Ind_Zup	1.23						
Vis_Zlw	3.64						
Vis_Zup	2.80						
Gl_Zlw	13.54						
Gl_Zup	14.35						

Sample Problem (continued)

Multi-Distribution Score Card

ANALYZED BY: Christopher Stanard
EXPERIMENT PERFORMED BY:
DESCRIPTION: Z & DOE Optimization Example
DATE:

Note: Only Enter Data in Green Cells

Enter Overall Z-shift (Zst-Zlt) **0**
(Typical M. Harry Shift Factor is 1.5)

Part/Component/Factor Description	Shrt Trm or Long Term	Distribution	Param 1	Param 2	Param 3	LTL	UTL	Row Zshift	# of Times Applied	Certainty (% Inside Spec Lims)	DPUlt	Zst
			mean	std dev	——-							
Index 1		Normal	14	0.13		13	15		1	100	1.4655E-14	7.60
			mean	std dev	——-							
Viscosity		Normal	160.03	2.24		120	190		1	100	##########	7+
			mean	std dev	——-							
Gloss		Normal	77.92	0.11		72	84		1	100	##########	7+

Note: User Inputs in Green Cells

Total DPUlt	1.46549E-14
Total Opportunities	3
Total DPMOlt	4.88498E-09
Zst	7.74

 Version 1.03 February 2001

Note: Using Std Error of Predicted Mean Rather Than Root MSE

DFSS ROSTAR: RObust Statistical Tools for All Responses

I recommend using the Root MSE rather than the std. error.

Six Sigma Score Card for Responses on Point Identified Directly Through Numerical Optimization Technique (Liberal)

Sample Problem (continued)

Multi-Distribution Score Card

Note: Only Enter Data in Green Cells

Enter Overall Z-shift (Zst-Zlt) **0**

(Typical M. Harry Shift Factor is 1.5)

ANALYZED BY: Christopher Stanard

EXPERIMENT PERFORMED BY:

DESCRIPTION: Z & DOE Optimization Example

DATE:

Part/Component/Factor Description	Shrt Trm or Long Term	Distribution	Distribution Parameters Param 1	Param 2	Param 3	LTL	UTL	Row Zshift	# of Times Applied	Certainty (% Inside Spec Lims)	DPUlt	Zst
			mean	std dev	——-							
Index 1		Normal	14	0.8246		13	15		1	77.475885	2.2524E-01	0.75
			mean	std dev	——-							
Viscosity		Normal	160.03	10.86		120	190		1	99.6993009	3.0070E-03	2.75
			mean	std dev	——-							
Gloss		Normal	77.92	0.4359		72	84		1	100	#########	7+

Note: User Inputs in Green Cells

Total DPUlt	0.228248141
Total Opportunities	3
Total DPMOlt	76082.71371
Zst	1.43

 Version 1.03 February 2001

Note: Using Root MSE

DESS ROSTAR: ROhust Statistical Tools for All Responses

Six Sigma Score Card for Responses on Point Identified Directly Through Numerical Optimization Technique (Conservative)

Key Observations

- Not Only Are Single Response DOEs Are Often Inadequate, but Multiple Optimization on Response Means (and Standard Deviations) Is Also Often Inadequate.
- Optimizing Z-scores (a DPMO or PPM Index) Is a Great Enhancement -- And Doable!!
- Multiple Response Requirements
  - Can Be ***Optimized*** Using
    Distance Measure Based Numerical Optimization
  - Can Be ***Satisficed*** Using
    Overlaid Contour Plots (Showing Feasible Regions)
- These These Tools Are Complementary

Bibliography

Derringer and Suich (1980). Journal of Quality Technology

Khuri, A. I. and Conlon, M. (1981). "Simultaneous Optimization of Multiple Responses by Polynomial Regression Functions," Technometrics Vol. 23, No. 4, pp. 363-375.

Montgomery, D. C. (1991). Design and Analysis of Experiments, Third Edition. New York: John Wiley & Sons.

Porter, Roper, Mason, Rossini, Banks, & Weiderholt (1991). Forecasting and Management of Technology. New York: John Wiley & Sons.

Rustagi, Jagdish S. (1994). Optimization Techniques in Statistics. San Diego: Academic Press, Inc.

Stanard, Christopher L (1998)."Multiple Response Optimization Using Designed Experiments – A Practical Example," The 1998 ASA Albany Mini-Conference, What Hot/What's Next III, March 28, 1998, GE CRD, Schenectady, NY 12309.

Wurl, Robin (1997). "Multiresponse Optimization Techniques: Sensitivity to Parameter Selection," Presented at the American Society for Quality Fall Technical Conference, October 1997 in Baltimore, MD.