Confidence Intervals for Derringer and Suich Desirability Function Optimal Points [1]

Peter A. Calhoun, PhD

Beau A. Nunnally, PhD
Lance E. Champagne, PhD
Christine Schubert Kabban, PhD

Methods and Tools for T&E
DATAWorks 2023
Alexandria, VA

April 27, 2023

Distribution Statement A - Approved for Public Release

Introduction

The Derringer and Suich (DS) desirability function (DF) method is often used for multi-objective optimization of linear models in Response Surface Methodology (RSM). This optimization technique lacks an adequate inference method for constructing confidence intervals. Practitioners compensate for lack of adequate inference method by:

- Treating desirability as deterministic, reporting point estimate,
- Using inference method that ignores covariance between responses,
- Using robust estimation methods that are less affected by variability with no inference.

This research studies 1 existing and 7 novel inference methods for constructing confidence intervals for the optimal values of Derringer and Suich Desirability Functions.

Multiple Linear Regression (MLR) [2, 3, 4, 5]

Let n be the number of observations, $p = k + 1$ be the number of parameters

- Linear model of the form

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

  - $\boldsymbol{y}$, $n \times 1$ vector of responses
  - $\boldsymbol{X}$, $n \times p$ model matrix
  - $\boldsymbol{\beta}$, $p \times 1$ vector of model parameters
  - $\boldsymbol{\varepsilon}$, $n \times 1$ vector of random error with mean 0 and variance σ^2
- $\boldsymbol{\varepsilon}$ most commonly assumed $N(0, \sigma^2)$
- $\boldsymbol{\beta}$ estimated using OLS or MLE and are BLUE
- $MSE = \frac{1}{n-p}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$ is unbiased estimator for σ^2

Multivariate Multiple Linear Regression (MMLR) [6, 7]

Adding to previous linear model, let m be the number of responses

- Linear model of the form

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{\Xi}$$

  - $\boldsymbol{Y}$, $n \times m$ matrix of responses
  - $\boldsymbol{X}$, $n \times p$ model matrix
  - $\boldsymbol{B}$, $p \times m$ matrix of model parameters
  - $\boldsymbol{\Xi}$, $n \times m$ matrix of random error with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}$
- $\boldsymbol{\Xi}$ most commonly assumed $N_m(\mathbf{0}, \boldsymbol{\Sigma})$
- $\boldsymbol{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m)$ uses OLS estimates and are BLUE
- $\boldsymbol{S}_e = \frac{1}{n-p}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}})'(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}})$ is unbiased estimator for $\boldsymbol{\Sigma}$

Vectorized Matrix and Kronecker Product [8, 9, 10]

Vectorizing an $m \times p$ matrix by rearranging columns such that they are stacked into a $mp \times 1$ vector.

Example: Let $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ then

$$vec(\boldsymbol{A}) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{pmatrix}$$

The Kronecker product multiplies each index of the first $m \times n$ matrix by the entire second $p \times q$ matrix resulting in an $mp \times nq$ matrix.

Example: Let $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\boldsymbol{B}$ be a $p \times q$ matrix. The Kronecker product, $\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B}$, is

$$\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} \end{pmatrix}$$

Desirability Functions [2, 11]

- Additive Desirability Index – Weighted Arithmetic Mean

$$D_i^{add} = \sum_{r=1}^{m} w_r d_{ir} \in [0, 1]$$

- Multiplicative Desirability Index – Weighted Geometric Mean

$$D_i^{mult} = \prod_{r=1}^{m} d_{ir}^{w_r} \in [0, 1]$$

- $\boldsymbol{w} = (w_1, \ldots, w_m)$ is vector of weights, $\sum_{r=1}^{m} w_r = 1$
- d_{ir} is desirability function for observation i in response r

Derringer and Suich Method [12, 11]

Maximization

$$\mathrm{d}_{ir}^{max} = \begin{cases} 0, & y_{ir} < L_r \\ \left(\frac{y_{ir}-L_r}{T_r-L_r}\right)^{l_r}, & L_r \leq y_{ir} \leq T_r \\ 1, & y_{ir} > T_r \end{cases}$$

Minimization

$$\mathrm{d}_{ir}^{min} = \begin{cases} 1, & y_{ir} < T_r \\ \left(\frac{U_r-y_{ir}}{U_r-T_r}\right)^{l_r}, & T_r \leq y_{ir} \leq U_r \\ 0, & y_{ir} > U_r \end{cases}$$

Match Target

$$\mathrm{d}_{ir}^{tgt} = \begin{cases} 0, & y_{ir} < L_r \\ \left(\frac{y_{ir}-L_r}{T_r-L_r}\right)^{l_{1r}}, & L_r \leq y_{ir} \leq T_r \\ \left(\frac{U_r-y_{ir}}{U_r-T_r}\right)^{l_{2r}}, & T_r \leq y_{ir} \leq U_r \\ 0, & y_{ir} > U_r \end{cases}$$

- where L_r is a lower bound, U_r is an upper bound, T_r is the target value
- U_r, T_r, and L_r can be arbitrarily chosen according to optimization desire
- l_r places emphasis on a particular d_{ir} to be closer to T
  - $l_r > 1$ places additional emphasis
  - $0 < l_r < 1$ places less emphasis
  - $l_r = 1$ places equal emphasis
- Other methods available but focus is Derringer and Suich Method

Methods

The following 8 methods are considered, all but BW are novel for this application

- BW – Best/Worst Case Method
- **UG** – Univariate Generalized Method
- **MG** – Multivariate Generalized Method
- MVNS – MVN Surfaces Method
- MVtS – MVt Surfaces Method
- **MVNSSig** – MVN Surfaces Method with Covariance
- **MVtSSig** – MVt Surfaces Method with Covariance
- BSR – Multivariate Bootstrap Residuals Method

BW, UG, MVNS, and MVtS assume independence of the responses
MG, MVtSSig account for correlation with MVt distribution
MVNSSig account for correlation with MVN distribution
BSR accounts for correlation by sampling errors from m linear models simultaneously

Univariate Generalized Pivotal Quantity

A pivotal quantity for a predicted observation is used to construct $R_{y_{ir}}$

$$\hat{y}_{ir} \sim N(\boldsymbol{x}_i \boldsymbol{\beta}_r, \sigma_r^2 h_{ii})$$

$$Z = \frac{\hat{y}_{ir} - \boldsymbol{x}_i \boldsymbol{\beta}_r}{\sqrt{\sigma_r^2 h_{ii}}} \sim N(0,1)$$

$$E\left[y_{ir}\right] = \hat{y}_{ir} - Z\sqrt{\sigma_r^2 h_{ii}}$$

σ_r^2 needs to be replaced to remove nuisance parameters. Let

$$U = SSE_r/\sigma_r^2 \sim \chi^2_{n-p}$$

$$\sigma_r^2 = SSE_r/U$$

then the pivot becomes

$$R_{E[y_{ir}]} = \hat{y}_{ir} - \frac{Z}{\sqrt{U/(n-p)}} \cdot \sqrt{MSE_r h_{ii}}$$

$$= \hat{y}_{ir} - t \cdot \sqrt{MSE_r h_{ii}}$$

Multivariate Generalized Pivotal Quantity

If $\boldsymbol{Y} \sim N_m(\mathbf{0}, \boldsymbol{\Sigma})$ and $U \sim \chi^2_\nu$ then

$$\boldsymbol{\mu} + \frac{\boldsymbol{Y}}{\sqrt{U/\nu}} \sim t_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}_t), \;\; \boldsymbol{\Sigma}_t = \boldsymbol{\Sigma} \cdot \frac{\nu}{\nu - 2}$$

$$vec(\hat{\boldsymbol{Y}}) \sim N_m(vec(\boldsymbol{X}\boldsymbol{B}), \boldsymbol{\Sigma} \otimes \boldsymbol{H})$$

$$\boldsymbol{P} = vec(\hat{\boldsymbol{Y}}) - E[vec(\boldsymbol{Y})] \sim N_m(\mathbf{0}, \boldsymbol{\Sigma} \otimes \boldsymbol{H})$$

Let $U \sim \chi^2_{n-p}$ then

$$t = \boldsymbol{vec}(\hat{\boldsymbol{Y}}) + \frac{\boldsymbol{P}}{\sqrt{U/(n-p)}} \sim t_m(vec(\hat{\boldsymbol{Y}}), \boldsymbol{\Sigma}_t \otimes \boldsymbol{H}, n-p)$$

where t_m is a multivariate t-distribution with $n-p$ degrees of freedom and $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma} \cdot \frac{n-p}{n-p-2}$. Thus, the multivariate pivotal quantity is

$$R_{vec(\hat{\boldsymbol{Y}})} = t_m(vec(\hat{\boldsymbol{Y}}), \boldsymbol{\Sigma}_t \otimes \boldsymbol{H}, n-p)$$

Simulated Surfaces

Original MVNS and MVtS simulated surfaces methods do not incorporate covariance. MVNSSig and MVtSSig vectorize the parameter matrix $\boldsymbol{B}$ and use $\hat{\boldsymbol{B}}$ and $\boldsymbol{S}_e$ as unbiased estimates for the unknown parameters in the sampling distribution.

- MVNSSig:
$$vec(\hat{\boldsymbol{B}}) \sim N_m(vec(\boldsymbol{B}), \boldsymbol{\Sigma} \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1})$$
- MVtSSig:
$$vec(\hat{\boldsymbol{B}}) = t_m(vec(\boldsymbol{B}), (\boldsymbol{S}_e \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1}) \cdot \frac{n-p}{n-p-2})$$

These are resampled similar to parametric bootstraps and used to make simulated linear models of the form

$$\hat{\boldsymbol{Y}}^* = \boldsymbol{X}\hat{\boldsymbol{B}}^*$$

First Order Models [13] (1/2)

First order models were considered to determine effect of correlation between responses and angle between response surfaces. 7 correlation levels and 7 plane angles were each considered using two response surfaces

$$\begin{aligned}
\rho_1 &= -0.8, & \theta_1 &= 143, & \boldsymbol{\beta}_{11} &= (10, 9.5, 9.5), & \boldsymbol{\beta}_{12} &= (10, -7, 1) \\
\rho_2 &= -0.5, & \theta_2 &= 120, & \boldsymbol{\beta}_{21} &= (10, 2.5, 9.5), & \boldsymbol{\beta}_{22} &= (10, -5.5, -1.5) \\
\rho_3 &= -0.3, & \theta_3 &= 107, & \boldsymbol{\beta}_{31} &= (10, 6.5, 9.5), & \boldsymbol{\beta}_{32} &= (10, 5.5, -7) \\
\rho_4 &= 0, & \theta_4 &= 90, & \boldsymbol{\beta}_{41} &= (10, 10, 10), & \boldsymbol{\beta}_{42} &= (10, 10, -10) \\
\rho_5 &= 0.3, & \theta_5 &= 73, & \boldsymbol{\beta}_{51} &= (10, 6.5, 9.5), & \boldsymbol{\beta}_{52} &= (10, -5.5, 7) \\
\rho_6 &= 0.5, & \theta_6 &= 60, & \boldsymbol{\beta}_{61} &= (10, 2.5, 9.5), & \boldsymbol{\beta}_{62} &= (10, 5.5, 1.5) \\
\rho_7 &= 0.8, & \theta_7 &= 37, & \boldsymbol{\beta}_{71} &= (10, 9.5, 9.5), & \boldsymbol{\beta}_{72} &= (10, 7, 1)
\end{aligned}$$

$$\theta_2 = 120: \;\; y_1 = 10 + 2.5x_1 + 9.5x_2, \;\; y_2 = 10 - 5.5x_1 - 1.5x_2$$

First Order Models [13] (2/2)

Random samples of data from true surfaces created using a bivariate normal

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{\Xi}$$

where $\boldsymbol{\Xi} \sim N_2(\mathbf{0}, \boldsymbol{\Sigma}_k)$ and

$$\boldsymbol{\Sigma}_k = \begin{pmatrix} \sigma_{k1}^2 & \rho_k \sigma_{k1} \sigma_{k2} \\ \rho_k \sigma_{k1} \sigma_{k2} & \sigma_{k2}^2 \end{pmatrix}$$

For the FO models, unit variance is used and

$$\boldsymbol{\Sigma}_k = \begin{pmatrix} 1 & \rho_k \\ \rho_k & 1 \end{pmatrix}$$

Second Order Models (1/3)

The second order models are based on Myers et al. (2016) Chemical Process Optimization Problem

Chemical Process Optimization Problem

	Natural Variables		Coded Variables		Responses		
i	ξ_1(**Time**)	ξ_2(**Temp**)	x_1	x_2	y_1(**Yield**)	y_2(**Viscosity**)	y_3(**Molecular Weight**)
1	80	170	-1	-1	76.5	62	2940
2	80	180	-1	1	77.0	60	3470
3	90	170	1	-1	78.0	66	3680
4	90	180	1	1	79.5	59	3890
5	85	175	0	0	79.9	72	3480
6	85	175	0	0	80.3	69	3200
7	85	175	0	0	80.0	68	3410
8	85	175	0	0	79.7	70	3290
9	85	175	0	0	79.8	71	3500
10	92.07	175	1.414	0	78.4	68	3360
11	77.93	175	-1.414	0	75.6	71	3020
12	85	182.07	0	1.414	78.5	58	3630
13	85	167.93	0	-1.414	77.0	57	3150

Second Order Models (2/3)

True response surfaces are based on rounded values from MLR models. The linear models estimated from the data are

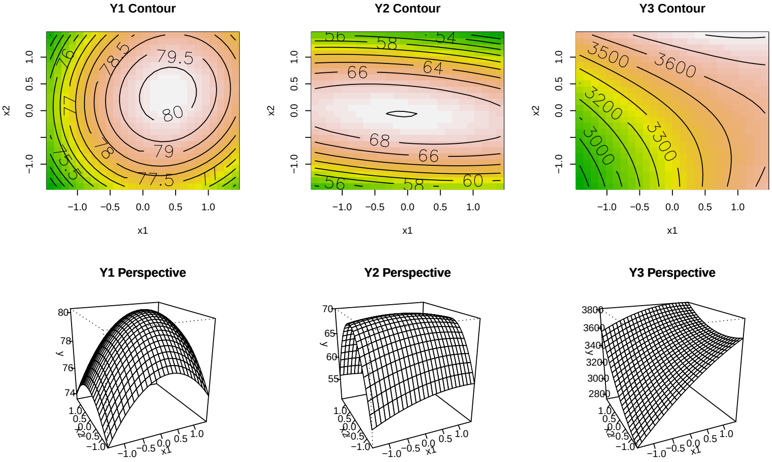
$$\hat{y}_1 = 79.9400 + 0.9951x_1 + 0.5152x_2 + 0.2500x_1x_2 - 1.3764x_1^2 - 1.0013x_2^2$$
$$\hat{y}_2 = 70.0002 - 0.1553x_1 - 0.9484x_2 - 1.2500x_1x_2 - 0.6873x_1^2 - 6.6891x_2^2$$
$$\hat{y}_3 = 3375.975 + 205.126x_1 + 177.367x_2 - 80x_1x_2 - 41.744x_1^2 + 58.286x_2^2$$

Random samples of data from true surfaces created using a multivariate normal, $\boldsymbol{\Xi} \sim N_3(\mathbf{0}, \boldsymbol{\Sigma}_{chem})$

$\boldsymbol{\Sigma}_{chem}$ replaced by $\boldsymbol{S}_e$ from original data as 'true' covariance

$$\boldsymbol{\Sigma}_{chem} = S_e = \begin{pmatrix} 0.07091 & -0.25988 & 13.46764 \\ -0.25988 & 5.17458 & 36.95738 \\ 13.46764 & 36.95738 & 29695.65911 \end{pmatrix}$$

Second Order Models (3/3)



Contour and Perspective Plots of Second Order Data

Model Scenarios (1/2)

Four optimization scenarios considered for **FO models**. In all cases, $\boldsymbol{l} = 1$ and $\boldsymbol{w} = (0.5, 0.5)$

- **Max/Min Three Correlation Levels with Differing Plane Angles**
- **Max/Min Three Plane Angles with Differing Correlation**
- Max/Min Matching Plane Angles and Correlation
- Max/Tgt Matching Plane Angles and Correlation

Three optimization scenarios considered for **SO models**. In each scenario, linearity and weights change.

- Max/Min Reduced Chemical Process with Y_1 and Y_3
- Max/Tgt Reduced Chemical Process with Y_1 and Y_2
- **Max/Tgt/Min Chemical Process with** Y_1, Y_2, Y_3

Each of these consider a total of 12 combinations.

Model Scenarios (2/2)

Max/Min or Max/Tgt Scenario

$$\boldsymbol{w}_1 = (w_1, w_2) = (0.5, 0.5)$$
$$\boldsymbol{w}_2 = (0.8, 0.2)$$
$$\boldsymbol{w}_3 = (0.2, 0.8)$$

Max/Tgt/Min Scenario

$$\boldsymbol{w}_1 = (w_1, w_2, w_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$$
$$\boldsymbol{w}_2 = (0.6, 0.2, 0.2)$$
$$\boldsymbol{w}_3 = (0.2, 0.6, 0.2)$$

Max/Min Scenario

$$\boldsymbol{l}_1 = (l_1, l_2) = (1, 1)$$
$$\boldsymbol{l}_2 = (0.1, 1)$$
$$\boldsymbol{l}_3 = (1, 10)$$
$$\boldsymbol{l}_4 = (0.1, 10)$$

Max/Tgt Scenario

$$\boldsymbol{l}_1 = (l_1, l_{12}, l_{22}) = (1, 1, 1)$$
$$\boldsymbol{l}_2 = (0.1, 1, 1)$$
$$\boldsymbol{l}_3 = (1, 10, 1)$$
$$\boldsymbol{l}_4 = (0.1, 10, 1)$$

Max/Tgt/Min Scenario

$$\boldsymbol{l}_1 = (l_1, l_{12}, l_{22}, l_3) = (1, 1, 1, 1)$$
$$\boldsymbol{l}_2 = (0.1, 1, 1, 1)$$
$$\boldsymbol{l}_3 = (1, 10, 1, 1)$$
$$\boldsymbol{l}_4 = (0.1, 10, 1, 1)$$

Simulation (1/2)

Original designs for data generation based on 9 and 13 run designed experiments. $N = 25$ candidate interpolated points considered

Interpolation grid of $\boldsymbol{X}$ from (-1,1)

		x_1				
		-1	-0.5	0	0.5	1
	-1	1	2	3	4	5
	-0.5	6	7	8	9	10
x_2	0	11	12	13	14	15
	0.5	16	17	18	19	20
	1	21	22	23	24	25

Reference of observation number to $\boldsymbol{X}$-space for (x_1, x_2)

Simulation (2/2)

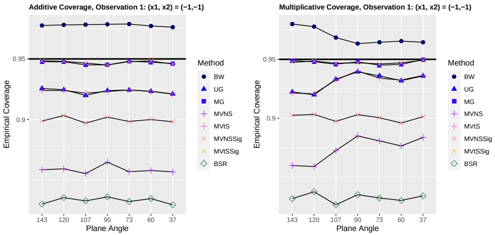
Goal is recording accurate coverage, width, symmetry of confidence intervals

- Number of simulations, $G = 10,000$
- Number of samples in resampling techniques, $B = 2,000$
- $\alpha/2$ and $1-\alpha/2$ quantiles used in B samples for each $i \in (1, G)$

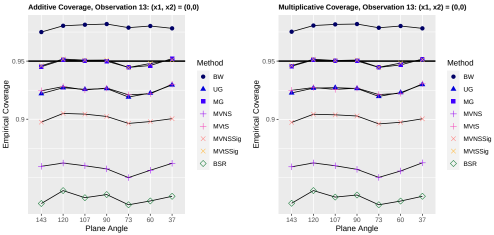
$$CP = \frac{\sum_{g=1}^{G} I_{\theta \in (L(\hat{\theta}_g)), U(\hat{\theta}_g))}}{G} \qquad AW = \frac{\sum_{g=1}^{G} U(\hat{\theta}_g) - L(\hat{\theta}_g)}{G}$$

$$SYM = \frac{\sum_{g=1}^{G} I_{\theta > U(\hat{\theta}_g)} - I_{\theta < L(\hat{\theta}_g)}}{G}$$

FO Max/Min Same Correlation, Differing Plane Angles (1/2)

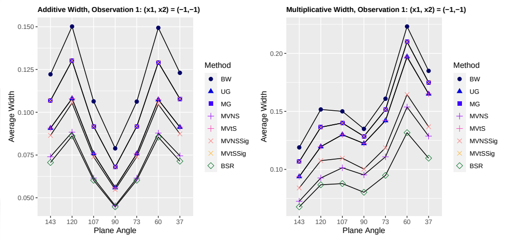


Additive and Multiplicative Plots for Coverage at Observation 1, Max/Min (Constant Correlation), $\rho = -0.5$

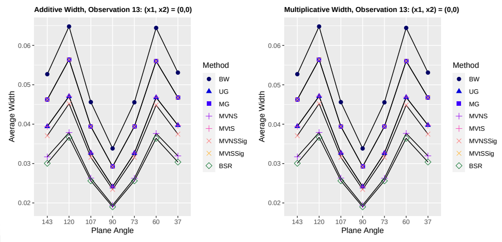


Additive and Multiplicative Plots for Coverage at Observation 13, Max/Min (Constant Correlation), $\rho = -0.5$

FO Max/Min Same Correlation, Differing Plane Angles (2/2)

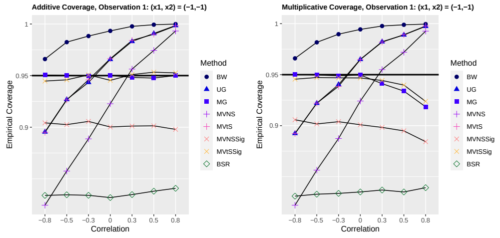


Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Correlation), $\rho = -0.5$

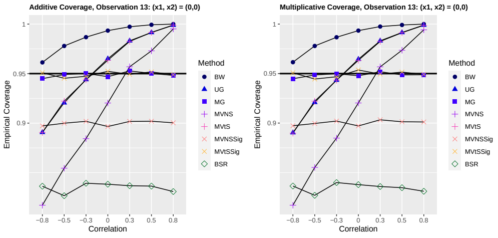


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Correlation), $\rho = -0.5$

FO Max/Min Same Plane Angles, Differing Correlation (1/2)

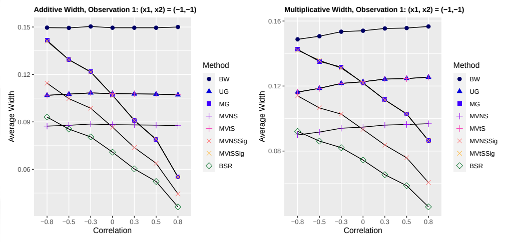


Additive and Multiplicative Plots for Coverage at Observation 1, Max/Min (Constant Plane Angles), $\theta = 120$

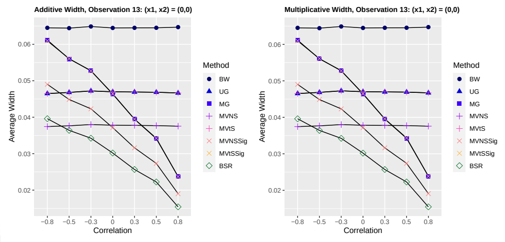


Additive and Multiplicative Plots for Coverage at Observation 13, Max/Min (Constant Plane Angles), $\theta = 120$

FO Max/Min Same Plane Angles, Differing Correlation (2/2)



Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Plane Angles), $\theta = 120$

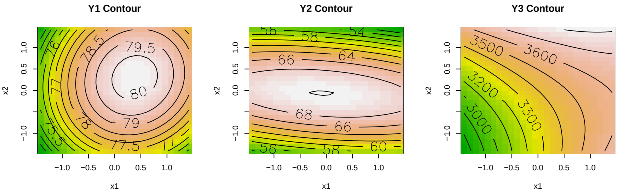


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Plane Angles), $\theta = 120$

SO Max/Tgt/Min Surface Models (1/6)

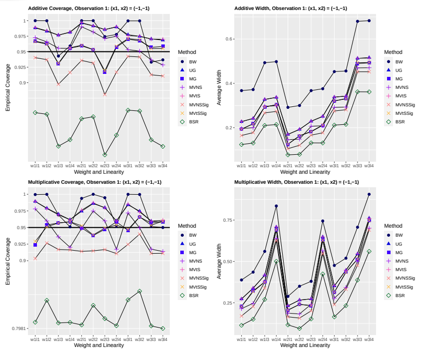
Optimal and Worst Solutions for each $\boldsymbol{w}_i\boldsymbol{l}_j$ Combination

w_i	l_j	Add Max	Mult Max	d_1Max	d_2Max	d_3Max	Add Min	Mult Min	d_1Min	d_2Min	d_3Min
1	1	19	3	19	4	1	21	1	1	13	25
1	2	3	2	19	4	1	25	25	1	13	25
1	3	19	19	19	20	1	21	25	1	25	25
1	4	6	6	19	20	1	25	25	1	25	25
2	1	3	3	19	4	1	1	1	1	13	25
2	2	4	3	19	4	1	23	23	1	13	25
2	3	19	19	19	20	1	21	24	1	25	25
2	4	19	7	19	20	1	24	24	1	25	25
3	1	4	4	19	4	1	21	21	1	13	25
3	2	4	2	19	4	1	24	24	1	13	25
3	3	19	19	19	20	1	25	25	1	25	25
3	4	20	8	19	20	1	25	25	1	25	25



SO Max/Tgt/Min Surface Models (2/6)

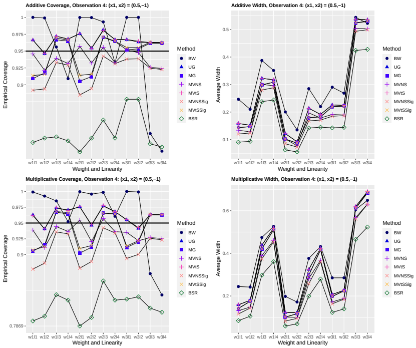
- Solution at location $\boldsymbol{x} = (-1, -1)$ favors Y_3
- $\boldsymbol{w}_1$ equal weights, $\boldsymbol{w}_2$ favors Y_1, $\boldsymbol{w}_3$ favors Y_2
- $\boldsymbol{l}_1$ equal linearity, $\boldsymbol{l}_2$ less emphasis on Y_1, $\boldsymbol{l}_3$ more emphasis lower half of Y_2, $\boldsymbol{l}_4$ less Y_1 and more Y_2 lower half
- Small overall desirability for w2l3 (add) and w1l1 (mult) resulting in less coverage
- $\boldsymbol{l}_3$ and $\boldsymbol{l}_4$ result in larger difference in values near optimal



Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt/Min Second Order

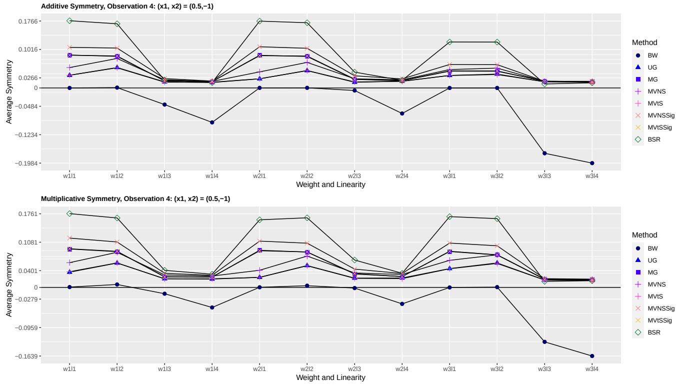
SO Max/Tgt/Min Surface Models (3/6)

- Solution at location $\boldsymbol{x} = (0.5, -1)$ favors Y_2
- $\boldsymbol{w}_1$ equal weights, $\boldsymbol{w}_2$ favors Y_1, $\boldsymbol{w}_3$ favors Y_2
- $\boldsymbol{l}_1$ equal linearity, $\boldsymbol{l}_2$ less emphasis on Y_1, $\boldsymbol{l}_3$ more emphasis lower half of Y_2, $\boldsymbol{l}_4$ less Y_1 and more Y_2 lower half
- Solution has minimum desirability for w2l1 (add) and w1l1, w2l2 (mult)



Additive and Multiplicative Plots for Coverage and Width at Observation 4, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (4/6)

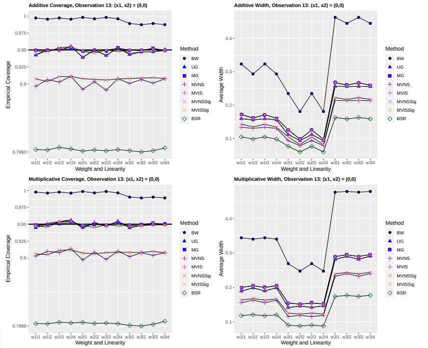


Additive and Multiplicative Plots for Symmetry at Observation 4, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (5/6)

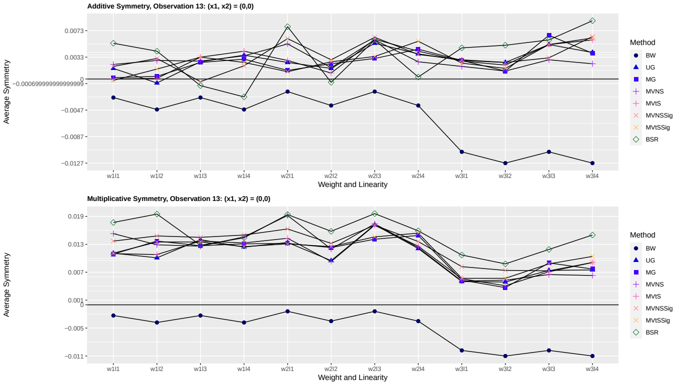
- Solution at location $\boldsymbol{x} = (0, 0)$ favors Y_1
- Low prediction variance makes coverage consistent for weights and linearity
- Width makes appropriate changes to accommodate coverage needs

$$\boldsymbol{C} = \begin{pmatrix} 1 & -0.42893 & 0.29331 \\ -0.42893 & 1 & 0.094275 \\ 0.29331 & 0.094275 & 1 \end{pmatrix}$$



Additive and Multiplicative Plots for Coverage and Width at Observation 13, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (6/6)



Additive and Multiplicative Plots for Symmetry at Observation 13, Max/Tgt/Min Second Order

Summary

- Previously only a single inference method used for DS DF optimal values
- Showcased 8 (7 novel) inference methods for DS DF optimal values
- Coverage not affected by plane angles; however, width changes d
- BW, UG, MVNS, MVtS coverage/width affected by missing covariance
- Coverage below limit when target objective optimal, true parameter is above interval; beneficial for decision maker
- Coverage decreases when solution has disagreeing optimal points; mitigate with contour plots
- Prediction variance affects coverage/width, replications recommended
- MG and MVtSSig are best performers for FO and SO models; BW conservative option

Algorithmic Steps

- Fit response surface model for m responses using linear regression,
- Calculate sample covariance $\boldsymbol{S}_e$ from MMLR,
- Determine, $\boldsymbol{X}^*$, the set of N candidate locations,
- For each response r, determine target and evasion values (T_r, L_r, U_r),
- For each response r, determine linearity parameters l_r, weights w_r, and DF type
- Calculate initial response surface, $\hat{\boldsymbol{Y}} = \boldsymbol{X}^* \hat{\boldsymbol{B}}$,
- For each observation in $\boldsymbol{X}^*$ using inference method of choice, sample B $\hat{\boldsymbol{y}}_i$ values, $\hat{\boldsymbol{y}}_{ib}$,
- For each observation in $\boldsymbol{X}^*$, calculate $\alpha/2$ and $1-\alpha/2$ quantiles from simulated random sample, $\hat{\boldsymbol{y}}_{ib}$,
- For each observation in $\boldsymbol{X}^*$, calculate applicable desirability function: d_{ir}^{max}, d_{ir}^{min}, or d_{ir}^{tgt},
- For each observation in $\boldsymbol{X}^*$, calculate applicable desirability index: D_i^{add} or D_i^{mult}

Algorithmic steps to compute CP and AW

Questions?

peter.calhoun.1@us.af.mil
peteracalhoun@gmail.com
https://www.linkedin.com/in/peteracalhoun/

References (1/3)

[1] P. A. Calhoun, "Statistical Inference on Desirability Function Optimal Points to Evaluate Multi-Objective Response Surface," 2020. *Theses and Dissertations*.

[2] R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook, *Response Surface Methodology*, 4th ed. Hoboken, NJ: Wiley, 2016.

[3] D. Montgomery, E. Peck, and G. Vining, *Introduction to Linear Regression Analysis*, ser. Wiley Series in Probability and Statistics. Wiley, 2015.

[4] D. C. Montgomery, *Design and Analysis of Experiments*. Hoboken, NJ: Wiley, 2017.

[5] M. H. Kutner, C. J. Nachtsheim, and J. Neter, *Applied Linear Regression Models*. New York, NY: McGraw-Hill/Irwin, 2004.

[6] A. Rencher and W. Christensen, *Methods of Multivariate Analysis*, ser. Wiley Series in Probability and Statistics. Wiley, 2012. [Online]. Available: https://books.google.com/books?id=fWMTV3wSpTcC

[7] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York, NY: Wiley, 1982.

[8] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," 2007.

[9] R. Schafer, *An Introduction to Nonassociative Algebras*, ser. Dover Books on Mathematics. Dover Publications, 2017. [Online]. Available: https://books.google.com/books?id=JnE9DwAAQBAJ

[10] E. W. Weisstein, "Kronecker product," https://mathworld.wolfram.com/KroneckerProduct.html, accessed: 2022-08-07.

[11] E. C. Harrington, "The desirability function," *Industrial quality control*, vol. 21, no. 10, pp. 494–498, 1965.

[12] G. Derringer and R. Suich, "Simultaneous optimization of several response variables," *Journal of quality technology*, vol. 12, no. 4, pp. 214–219, 1980.

[13] O. Bretscher, *Linear Algebra with Applications*. Pearson Prentice Hall, 2009.

[14] N. Helwig, "Lecture notes in multivariate linear regression," January 2017. [Online]. Available: http://users.stat.umn.edu/~helwig/notes/mvlr-Notes.pdf

References (2/3)

[15] C. A. Robertson and J. G. Fryer, "Some descriptive properties of normal mixtures," *Scandinavian Actuarial Journal*, vol. 1969, no. 3-4, pp. 137–146, 1969. [Online]. Available: https://doi.org/10.1080/03461238.1969.10404590

[16] J. A. Cornell, *Experiments with mixtures: Designs, models, and the analysis of mixture data*, 2011, cited by: 909.

[17] S. Nadarajah and T. K. Pogány, "On the distribution of the product of correlated normal random variables," *Comptes Rendus Mathematique*, vol. 354, no. 2, pp. 201–204, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.crma.2015.10.019

[18] R. E. Gaunt, "The basic distributional theory for the product of zero mean correlated normal random variables," *Statistica Neerlandica*, 2022.

[19] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions, volume 2*. John wiley & sons, 1995, vol. 289.

[20] S. Weerahandi, *Exact Statistical Methods for Data Analysis*. Springer Science & Business Media, 2003.

[21] B. A. Nunnally, "Statistical Inference to Evaluate and Compare the Performance of Correlated Multi-State Classification Systems," 2018. *Theses and Dissertations*.

[22] J. L. Chapman, L. Lu, and C. M. Anderson-Cook, "Incorporating response variability and estimation uncertainty into Pareto front optimization," *Computers and Industrial Engineering*, 2014b.

[23] P. A. Calhoun, "Characterizing Uncertainty in Correlated Response Variables for Pareto Front Optimization," 2020. *Theses and Dissertations*.

[24] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, 1st ed., ser. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1994.

[25] D. J. Eck, "Bootstrapping for multivariate linear regression models," *Statistics Probability Letters*, vol. 134, pp. 141–149, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167715217303450

[26] G. Casella and R. L. Berger, *Statistical Inference*. Belmont, CA: Duxbury, 2002.

[27] S. Kotz and S. Nadarajah, *Multivariate t-distributions and their applications*. Cambridge University Press, 2004.

References (3/3)

[28] P. Lin, "Some characterizations of the multivariate t distribution," *Journal of Multivariate Analysis*, vol. 2, no. 3, pp. 339–344, 1972. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0047259X72900218

[29] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Continuous multivariate distributions, Volume 1: Models and applications*. New York, NY: John Wiley & Sons, 2000.

[30] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Discrete Multivariate Distributions*. New York, NY: John Wiley & Sons, 1997.

β Normal Equations — Return to slide 5 (1/2)

The linear regression model can be rewritten to be in terms of the random error

$$\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}$$

The sum of the squared error can be expressed as

$$\begin{aligned} L = \sum_{i=1}^{n} \varepsilon_i^2 &= \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} \\ &= (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \\ &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} \\ &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} \end{aligned}$$

which is the least squares estimator

β Normal Equations — Return to slide 5 (2/2)

The least squares estimator, L, must satisfy

$$\left.\frac{\partial L}{\partial \beta}\right|_{\boldsymbol{b}} = -2\boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b} = 0$$

Dividing each term by 2 and rearranging the equation gives

$$\begin{aligned} \boldsymbol{X}'\boldsymbol{X}\boldsymbol{b} &= \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{b} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \end{aligned}$$

where $\boldsymbol{b}$ are the least squares estimator for $\boldsymbol{\beta}$

MMLR Expected Value and Covariance [14, 6] (1/2)

$\hat{\boldsymbol{Y}}$ can be vectorized and Kronecker product can be used to find expected value and covariance matrix.

$$\begin{aligned} E[vec(\hat{\boldsymbol{Y}})] &= E[vec(\boldsymbol{X}\hat{\boldsymbol{B}})] \\ &= E[vec(\boldsymbol{I}_n\boldsymbol{X}\hat{\boldsymbol{B}})] \\ &= E[(\hat{\boldsymbol{B}}' \otimes \boldsymbol{I}_n)vec(\boldsymbol{X})] \\ &= (\boldsymbol{B}' \otimes \boldsymbol{I}_n)vec(\boldsymbol{X}) \end{aligned}$$

$$\begin{aligned} V[vec(\hat{\boldsymbol{Y}})] &= V[vec(\boldsymbol{H}\boldsymbol{Y})] \\ &= V[(\boldsymbol{I}_m \otimes \boldsymbol{H})vec(\boldsymbol{Y})] \\ &= (\boldsymbol{I}_m \otimes \boldsymbol{H})V[vec(\boldsymbol{Y})](\boldsymbol{I}_m \otimes \boldsymbol{H})' \\ &= (\boldsymbol{I}_m \otimes \boldsymbol{H})(\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m)(\boldsymbol{I}_m \otimes \boldsymbol{H})' \\ &= (\boldsymbol{I}_m\boldsymbol{\Sigma}\boldsymbol{I}_m \otimes \boldsymbol{H}\boldsymbol{I}_m\boldsymbol{H}') \\ &= \boldsymbol{\Sigma} \otimes \boldsymbol{H} = \boldsymbol{\Sigma} \otimes \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \end{aligned}$$

MMLR Expected Value and Covariance [14, 6] (2/2)

Similarly, the expected value and covariance matrix for $\hat{\boldsymbol{B}}$ can be found. For the covariance, let $\boldsymbol{C} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$

$$
\begin{aligned}
E[\hat{\boldsymbol{B}}] &= E[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}] \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E[\boldsymbol{Y}] \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{B} \\
&= \boldsymbol{B}
\end{aligned}
$$

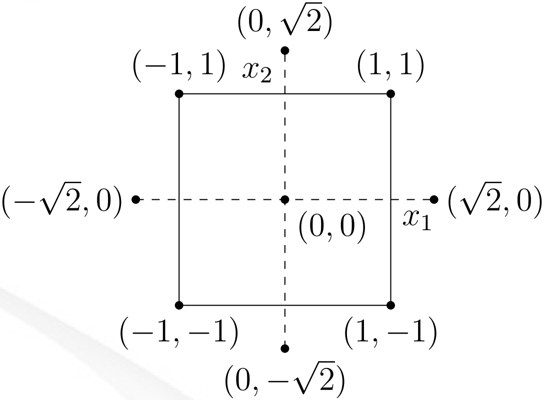
$$
\begin{aligned}
V[vec(\hat{\boldsymbol{B}})] &= V[vec(\boldsymbol{C}\boldsymbol{Y})] \\
&= V[vec(\boldsymbol{C}\boldsymbol{Y}\boldsymbol{I}_m)] \\
&= V[(\boldsymbol{I}_m \otimes \boldsymbol{C})vec(\boldsymbol{Y})] \\
&= (\boldsymbol{I}_m \otimes \boldsymbol{C})V[vec(\boldsymbol{Y})](\boldsymbol{I}_m \otimes \boldsymbol{C})' \\
&= (\boldsymbol{I}_m \otimes \boldsymbol{C})(\boldsymbol{\Sigma} \otimes \boldsymbol{I}_n)(\boldsymbol{I}_m \otimes \boldsymbol{C}') \\
&= (\boldsymbol{I}_m\boldsymbol{\Sigma}\boldsymbol{I}_m \otimes \boldsymbol{C}\boldsymbol{I}_n\boldsymbol{C}') \\
&= \boldsymbol{\Sigma} \otimes \boldsymbol{C}\boldsymbol{C}' \\
&= \boldsymbol{\Sigma} \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= \boldsymbol{\Sigma} \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1}
\end{aligned}
$$

Response Surface Methodology (RSM) [2, 4]

- Methodology of developing, improving, or optimizing a product or process.
- Sequentially uses designed experiments with gradient search to find optimal response
- Curvature often present resulting in second order models
- Estimated model fit using linear regression
- Multi-objective optimization methods include
  - Overlaying Contour Plots
  - Constrained Optimization
  - Pareto Fronts – $\boldsymbol{y}_1 = (y_{11}, y_{21}) = (10, 15)$, $\boldsymbol{y}_2 = (15, 15)$
  - Desirability Functions

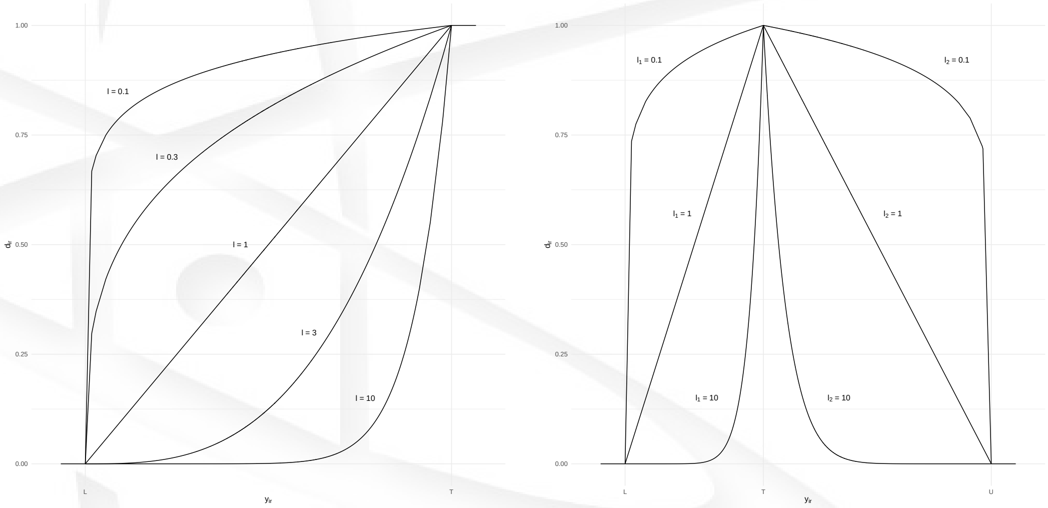
Central Composite Design [2, 4]

- 2^k factorial design augmented with n_c center points, $2k$ axial points
  - 2^k points provide an initial 'box' of high and low points
  - n_c center points reduce variability in the center
  - $2k$ axial points produce circular design
  - $\alpha = \sqrt[4]{2^k}$ axial distance allows rotatability
- Preferred designed experiment for second order models



Central Composite Design with 2 Factors and $\alpha = \sqrt{2}$

Desirability Function Plots



DF for max/minimization (left) and match target (right) for differing linearity parameter, l

Unknown Desirability Distribution

Recall the linear regression model,

$$\boldsymbol{y}_r = \boldsymbol{X}\boldsymbol{\beta}_r + \boldsymbol{\varepsilon}_r$$

- $E[\boldsymbol{\varepsilon}_r] = 0$
- $V[\boldsymbol{\varepsilon}_r] = \sigma_r^2$
- $\varepsilon_{ir} \perp\!\!\!\perp \varepsilon_{jr}$

When $\boldsymbol{\varepsilon}_r$ is assumed to be normally distributed then

$$\hat{\boldsymbol{y}}_r \sim N(\boldsymbol{X}\boldsymbol{\beta}_r, \sigma_r^2 \boldsymbol{H}) \implies \hat{y}_{ir} \sim N(\boldsymbol{x}_i\boldsymbol{\beta}_r, \sigma_r^2 h_{ii})$$

where $\boldsymbol{H}$ is the hat matrix of $\boldsymbol{X}$

Linear Additive Desirability Index [15, 16]

Recall the desirability functions, with linearity $l = 1$,

Maximization $\boldsymbol{y}_r \in (L_r, T_r)$

$$d_{ir}^{max} = \frac{\hat{y}_{ir} - L_r}{T_r - L_r}$$

$$d_{ir}^{max} \sim N\left(\frac{\boldsymbol{x}_i\boldsymbol{\beta}_r - L_r}{T_r - L_r}, \frac{\sigma_r^2 h_{ii}}{(T_r - L_r)^2}\right)$$

Minimization $\boldsymbol{y}_r \in (T_r, U_r)$

$$d_{ir}^{min} = \frac{U_r - \hat{y}_{ir}}{U_r - T_r}$$

$$d_{ir}^{min} \sim N\left(\frac{-(\boldsymbol{x}_i\boldsymbol{\beta}_r - U_r)}{U_r - T_r}, \frac{\sigma_r^2 h_{ii}}{(U_r - T_r)^2}\right)$$

When $m = 2$ the additive desirability index is

$$D_i^{add} = \sum_{r=1}^{2} w_r d_{ir} = w_1 d_{i1} + w_2 d_{i2}$$

thus when the weights sum to 1 and only when independence is assumed,

$$D_i^{max} \sim N\left(w_1\frac{\boldsymbol{x}_i\boldsymbol{\beta}_1 - L_1}{T_1 - L_1} - w_2\frac{(\boldsymbol{x}_i\boldsymbol{\beta}_2 - U_2)}{U_2 - T_2}, w_1^2\frac{\sigma_1^2}{(T_1 - L_1)^2} + w_2^2\frac{\sigma_2^2}{(U_2 - T_2)^2}\right)$$

Non-Linear and/or Multiplicative Desirability Index [17, 18, 19]

Maximization $\boldsymbol{y} \in (L, T)$

$$d_{ir}^{max} = \left(\frac{\hat{y}_{ir} - L_r}{T_r - L_r}\right)^l$$

Minimization $\boldsymbol{y} \in (T, U)$

$$d_{ir}^{min} = \left(\frac{U_r - \hat{y}_{ir}}{U_r - T_r}\right)^l$$

The multiplicative desirability function for m responses is

$$D_i^{mult} = \prod_{r=1}^{m} d_{ir}^{w_r}$$

Distribution of D_i becomes unknown and intractable except in special cases such as $l = 2$ with additive form.

Best/Worst Case Method (BW) [2, 3, 4, 5]

$$E[y_{ir}] = \hat{y}_{ir} \pm t_{1-\alpha/2,n-p}\sqrt{MSE_r \cdot h_{ii}}$$

- Assumes independence of responses
- Best (worst) case scenario occurs at upper (lower) bound for maximization and inverse is true for minimization
- Best case scenario occurs at target value, worst case scenario occurs at largest absolute deviation from target for match target
- Upper (lower) bound of D_i is calculated using the m best (worst) case values

Generalized Method [20, 21]

Generalized confidence intervals require a generalized pivotal quantity.

$$R = r(\boldsymbol{X}, \boldsymbol{x}, \zeta)$$

where r is a function of $\boldsymbol{X}$ (random sample), and possibly $\boldsymbol{x}$ (observed), $\zeta = (\theta, \boldsymbol{\delta})$

- The distribution of R is free of unknown parameters
- $r_{obs} = r(\boldsymbol{x}; \boldsymbol{x}, \zeta)$ does not depend on nuisance parameters, $\boldsymbol{\delta}$

Based on sufficient statistic and sampling distribution of nuisance parameter. Monte Carlo random samples used with generalized pivotal quantity to construct confidence intervals

Simulated Surface Methods [22, 23, 2]

The distribution for regression parameters is

$$\hat{\boldsymbol{\beta}}_r \sim N_m(\boldsymbol{\beta}_r, \sigma_r^2(\boldsymbol{X}'\boldsymbol{X})^{-1})$$

- Assumes independence of responses
- Use $\hat{\boldsymbol{\beta}}_r$ and MSE_r as unbiased estimators for $\boldsymbol{\beta}_r$ and σ_r^2
- Sample from $N_m(\hat{\boldsymbol{\beta}}_r, MSE_r((\boldsymbol{X}'\boldsymbol{X})^{-1}))$
- Sample from $t_m(\hat{\boldsymbol{\beta}}_r, MSE_r((\boldsymbol{X}'\boldsymbol{X})^{-1}) \cdot \frac{n-p}{n-p-2})$

Simulated surface model variability in individual responses and can be used to construct a parametric bootstrap confidence interval

Nonparametric Bootstrap Method [24, 25]

Two methods for nonparametric bootstrapping regression models, pairs and residuals. Only bootstrapping residuals were considered for final results
Let $\boldsymbol{C} = (\boldsymbol{Y}, \boldsymbol{X})$ be a dataset of m responses and p factors, a row is $\boldsymbol{c}_i = (\boldsymbol{y}_i, \boldsymbol{x}_i)$

- Bootstrap Residuals
  - Sample n residuals, e_i, from linear model of $\boldsymbol{C}$ with replacement
  - Add randomly sampled residuals to predicted values from model and fit new model with $\boldsymbol{C}_b^*$

$$\boldsymbol{C}_b^* = \{(\boldsymbol{x}_1\hat{\boldsymbol{\beta}} + e_{b1}, \boldsymbol{x}_1), \ldots, (\boldsymbol{x}_n\hat{\boldsymbol{\beta}} + e_{bn}, \boldsymbol{x}_n)\}$$

- The previous definition is normally performed for individual regression models. Eck (2018) provides a multivariate generalization by sampling error vectors from MMLR models to incorporate covariance.

Multivariate Distributions - MV Normal & MV t [6, 26, 27, 28]

- Let $\boldsymbol{X}$ be a vector of m random variables with mean vector, $\boldsymbol{\mu}$, and covariance matrix, $\boldsymbol{\Sigma}$, then the pdf of the MVN is defined as

$$f(\boldsymbol{x}|\boldsymbol{\mu}, m, \boldsymbol{\Sigma}) = \frac{1}{(\sqrt{2\pi})^m |\boldsymbol{\Sigma}|^{1/2}} e^{-(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})/2}$$

  where $\boldsymbol{x} \in \mathbb{R}$, $\boldsymbol{\mu} \in \mathbb{R}$, $m \in \mathbb{N}$, and $\boldsymbol{\Sigma}$ is positive definite

- Let $\boldsymbol{X}$ be a vector of m random variables with shift vector, $\boldsymbol{\mu}$, and correlation matrix, $\boldsymbol{R}$, then the pdf of the MVt used in this research is defined as

$$f(\boldsymbol{x}|\boldsymbol{\mu}, \nu, m, \mathbf{R}) = \frac{\Gamma((\nu+m)/2)}{(\nu\pi)^{m/2}\Gamma(\nu/2)|\mathbf{R}|^{1/2}} \left[1 + \frac{1}{\nu}(\boldsymbol{x}-\boldsymbol{\mu})'\mathbf{R}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^{-(\nu+m)/2}$$

  where $\boldsymbol{x} \in \mathbb{R}$, $\boldsymbol{\mu} \in \mathbb{R}$, $\nu \in \mathbb{N}$, $m \in \mathbb{N}$, and r_{ij} are entries of $\boldsymbol{R}$ where $-1 \leq r_{ij} \leq 1$

Multivariate Distributions – Dirichlet [29, 30]

- Let $\boldsymbol{X}$ be a vector of m random variables with parameter vector, $\boldsymbol{\alpha} = (\alpha_0, \dots \alpha_m)$, the pdf of the standard Dirichlet distribution defined as

$$f(\boldsymbol{x}|\boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{j=0}^{m} \alpha_j\right)}{\prod_{j=0}^{m} \Gamma(\alpha_j)} \left(1 - \sum_{j=1}^{m} x_j\right)^{\alpha_0 - 1} \prod_{j=1}^{m} x_j^{\alpha_j - 1}$$

  where $x_j > 0$, $j = 1, \dots, m$, and $\sum_{j=1}^{m} x_j \leq 1$.

  - Derived from $j+1$ $\chi^2_{\nu_j}$ random variables
  - Alternatively known as the Multivariate Beta Distribution

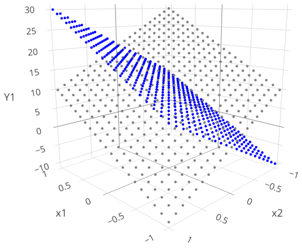
Multivariate Distributions – Wishart [29, 7]

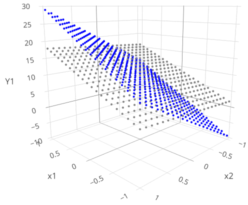
- If $\boldsymbol{Z}$ is an $n \times m$ distributed $N_m(0, I_n \otimes \Sigma)$ then $\boldsymbol{X} = \boldsymbol{Z}'\boldsymbol{Z}$ has a Wishart Distribution with n degrees of freedom and covariance matrix Σ. The pdf of the Wishart is defined as

$$f(\boldsymbol{x}|m, n, \boldsymbol{\Sigma}) = \frac{exp(tr\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\right))\,(det(\boldsymbol{x}))^{(n-m-1)/2}}{2^{mn/2}\Gamma_m\left(\frac{1}{2}n\right)(det(\boldsymbol{\Sigma}))^{n/2}}$$

  where $\boldsymbol{X} > 0$, $\Gamma_m(a)$ is the multivariate gamma function, $tr(.)$ is the trace, and $det(.)$ is the determinant
  - Alternatively known as Multivariate Chi-Square Distribution
  - Analogous properties to Chi-square distribution

Example of FO Model Planes





Examples of True Response Surface Planes with 90° angle (left) and 37° angle (right)

Simulation

L_r, T_r, U_r for d_{ir} calculated from $100 \cdot (1 - \alpha/2)$ prediction interval on true value y_{ir} for outer bounds

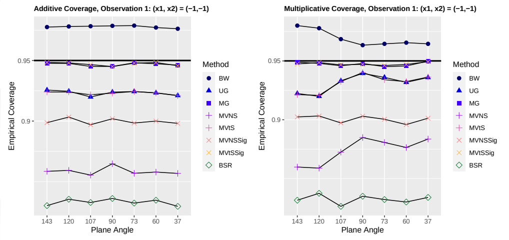
$$y_{ir} \pm t_{1-\alpha/2,n-p} \cdot \sqrt{\sigma_i^2(1 + h_{ii})}$$

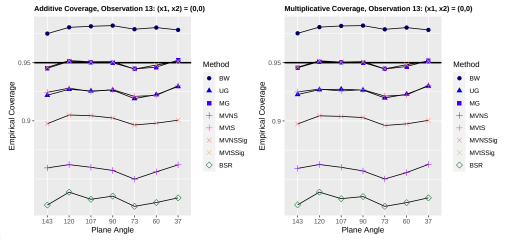
FO Model target based arbitrarily on 60th percentile of true response of Y_2 for each plane angle.

$$T_k = (11.6, 11.35, 11.7, 12, 11.7, 11.35, 11.6)$$

SO Model target is 65 from the original problem set.

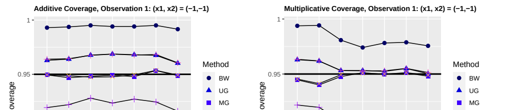
FO Max/Min Same Correlation with Differing Plane Angles (1/5)

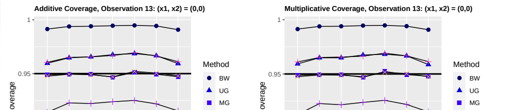




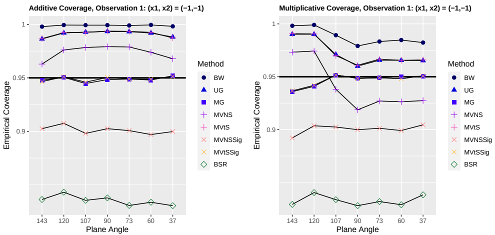
Additive and Multiplicative Plots for Coverage at Observation 1, Max/Min (Constant Correlation), $\rho = -0.5$

Additive and Multiplicative Plots for Coverage at Observation 13, Max/Min (Constant Correlation), $\rho = -0.5$

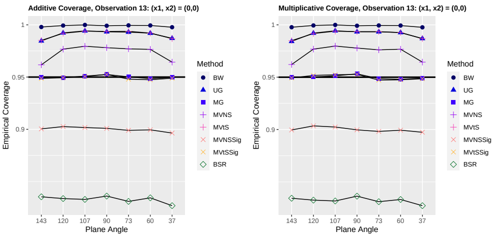




FO Max/Min Same Correlation with Differing Plane Angles (2/5)

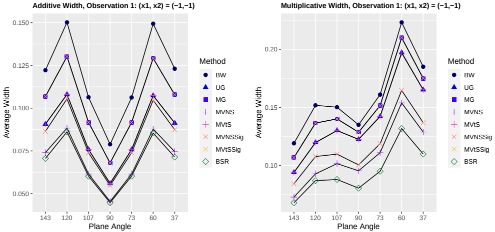


Additive and Multiplicative Plots for Coverage at Observation 1, Max/Min (Constant Correlation), $\rho = 0.5$

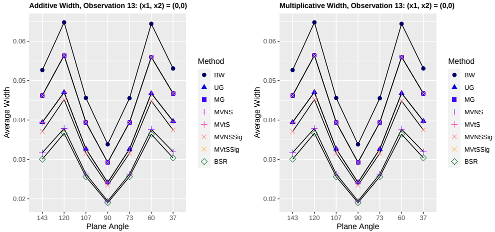


Additive and Multiplicative Plots for Coverage at Observation 13, Max/Min (Constant Correlation), $\rho = 0.5$

FO Max/Min Same Correlation with Differing Plane Angles (3/5)

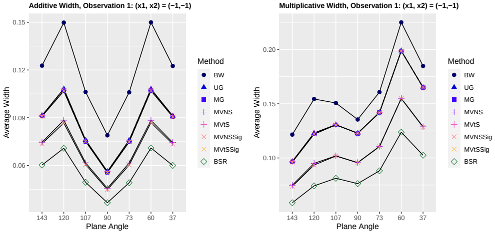


Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Correlation), $\rho = -0.5$

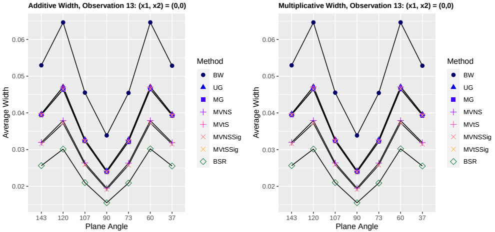


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Correlation), $\rho = -0.5$

FO Max/Min Same Correlation with Differing Plane Angles (4/5)

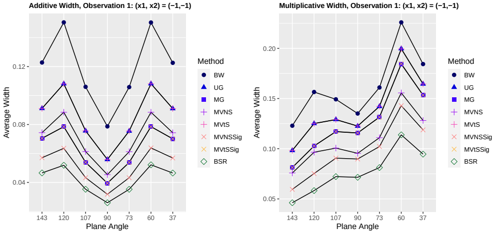


Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Correlation), $\rho = 0$

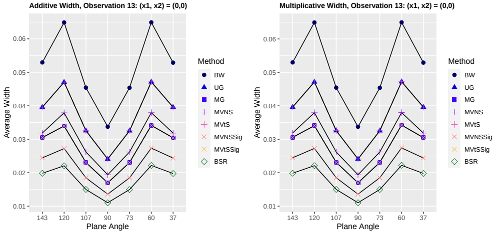


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Correlation), $\rho = 0$

FO Max/Min Same Correlation with Differing Plane Angles (5/5)

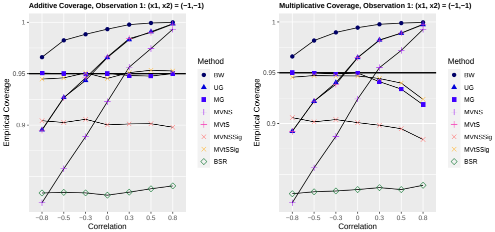


Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Correlation), $\rho = 0.5$

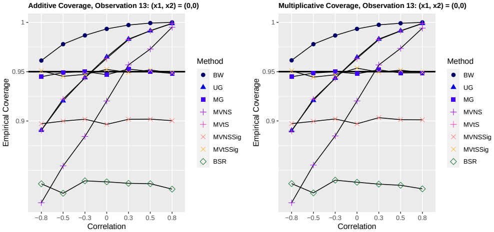


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Correlation), $\rho = 0.5$

FO Max/Min Same Plane Angles with Differing Correlation (1/6)

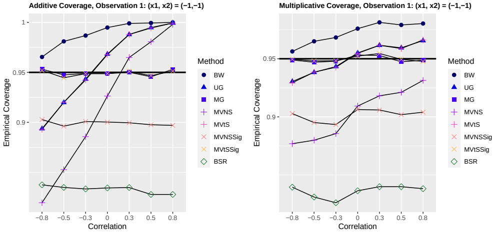


Additive and Multiplicative Plots for Coverage at Observation 1, Max/Min (Constant Plane Angles), $\theta = 120$

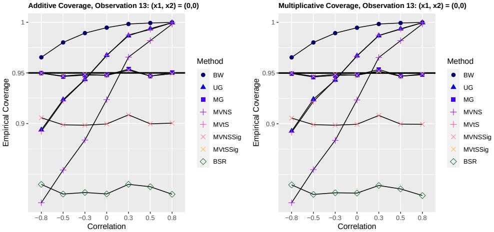


Additive and Multiplicative Plots for Coverage at Observation 13, Max/Min (Constant Plane Angles), $\theta = 120$

FO Max/Min Same Plane Angles with Differing Correlation (2/6)

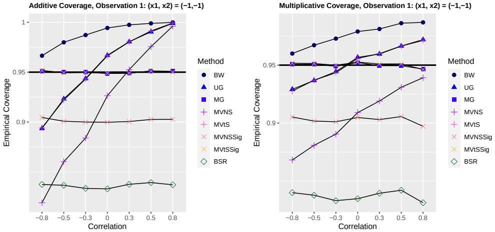


Additive and Multiplicative Plots for Coverage at Observation 1, Max/Min (Constant Plane Angles), $\theta = 90$

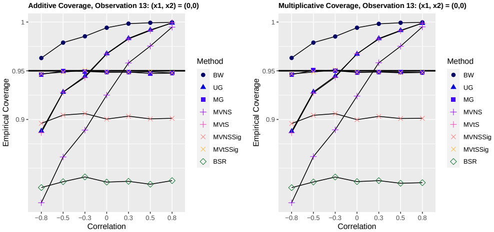


Additive and Multiplicative Plots for Coverage at Observation 13, Max/Min (Constant Plane Angles), $\theta = 90$

FO Max/Min Same Plane Angles with Differing Correlation (3/6)

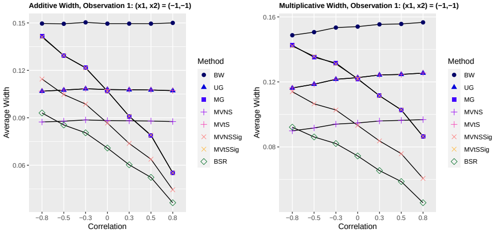


Additive and Multiplicative Plots for Coverage at Observation 1, Max/Min (Constant Plane Angles), $\theta = 60$

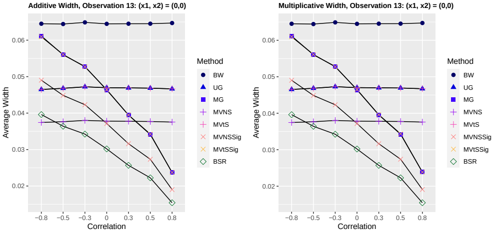


Additive and Multiplicative Plots for Coverage at Observation 13, Max/Min (Constant Plane Angles), $\theta = 60$

FO Max/Min Same Plane Angles with Differing Correlation (4/6)

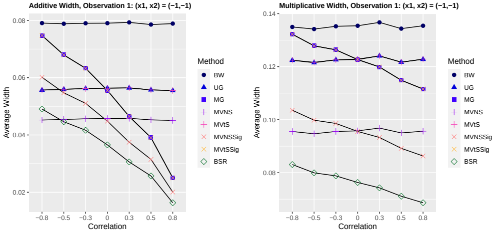


Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Plane Angles), $\theta = 120$

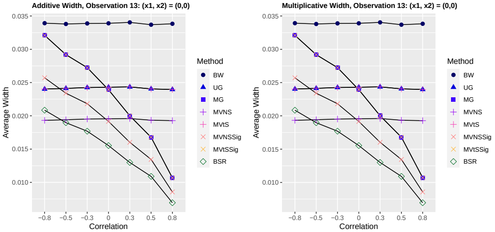


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Plane Angles), $\theta = 120$

FO Max/Min Same Plane Angles with Differing Correlation (5/6)

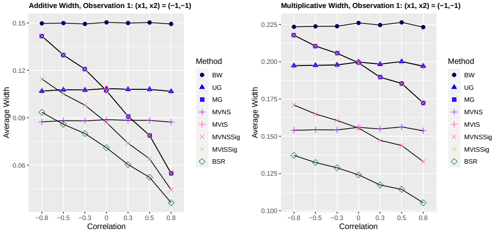


Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Plane Angles), $\theta = 90$

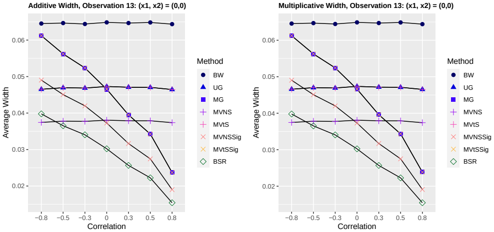


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Plane Angles), $\theta = 90$

FO Max/Min Same Plane Angles with Differing Correlation (6/6)

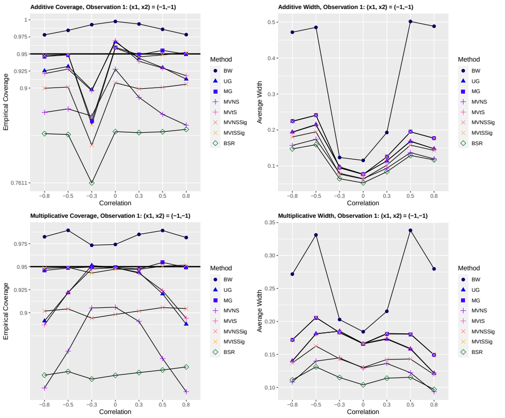


Additive and Multiplicative Plots for Width at Observation 1, Max/Min (Constant Plane Angles), $\theta = 60$

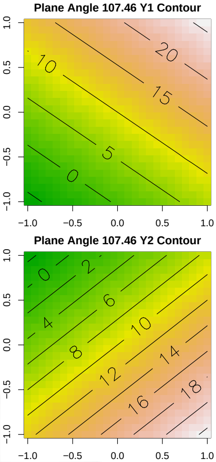


Additive and Multiplicative Plots for Width at Observation 13, Max/Min (Constant Plane Angles), $\theta = 60$

FO Max/Tgt Matching Plane Angle and Correlation (1/4)

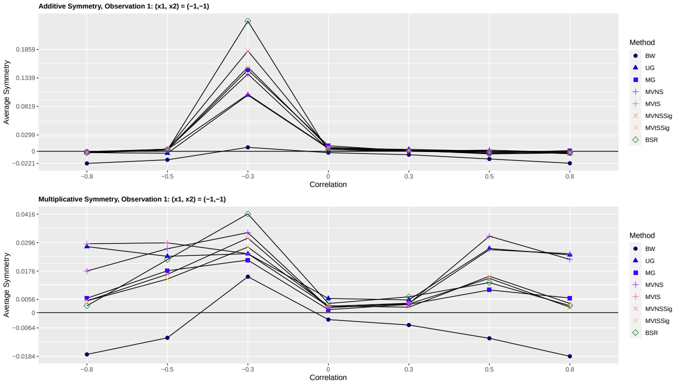


Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt



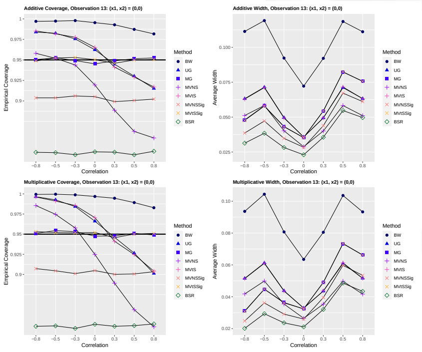
Contour plot for $\theta = 107$ corresponding to $\rho = -0.3$

FO Max/Tgt Matching Plane Angle and Correlation (2/4)

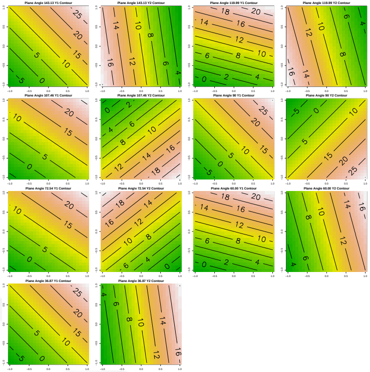


Additive and Multiplicative Plots for Symmetry at Observation 1, Max/Tgt

FO Max/Tgt Matching Plane Angle and Correlation (3/4)

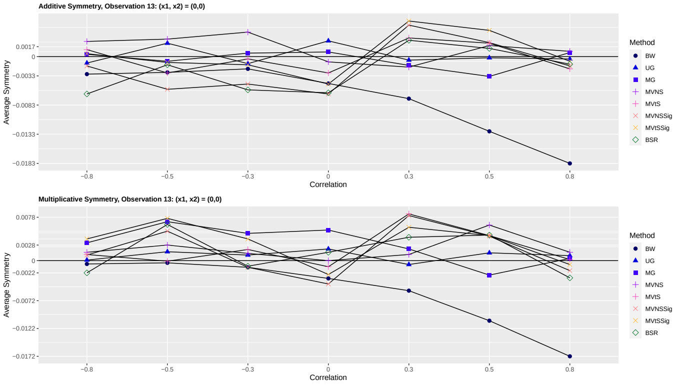


Additive and Multiplicative Plots for Coverage and Width at Observation 13, Max/Tgt



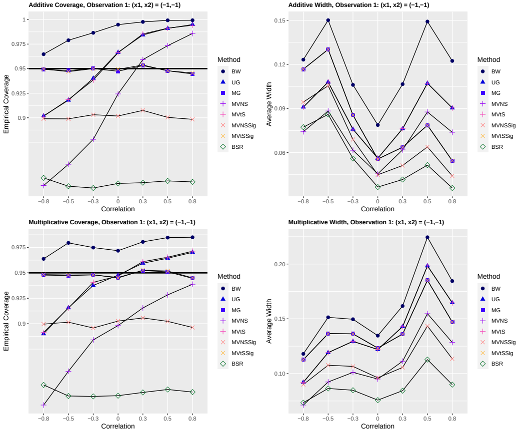
Contour plots for all FO Models

FO Max/Tgt Matching Plane Angle and Correlation (4/4)



Additive and Multiplicative Plots for Symmetry at Observation 13, Max/Tgt

FO Max/Min Matching Plane Angle and Correlation (1/2)



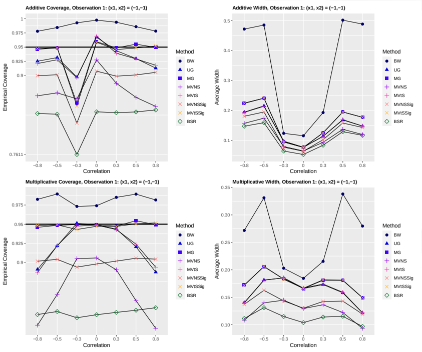
Additive and Multiplicative Plots for Coverage and Width at Observation 1 (Max/Min)

FO Max/Min Matching Plane Angle and Correlation (2/2)

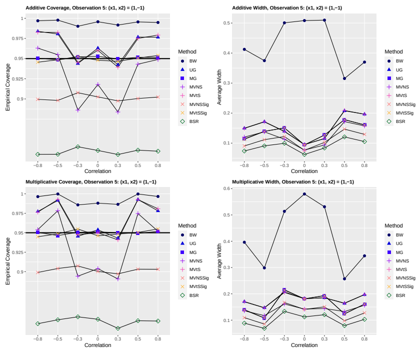
Coverage								
ρ	BW	UG	MG	MVNS	MVtS	MVNSSig	MVtSSig	BSR
-0.8	0.9647	0.9019	0.9493	0.831	0.9016	0.899	0.9496	0.8388
-0.5	0.9789	0.9179	0.948	0.8526	0.9185	0.8989	0.947	0.8303
-0.3	0.9865	0.9402	0.95	0.878	0.9384	0.9031	0.95	0.8285
0	0.9948	0.9666	0.9474	0.9243	0.9662	0.9018	0.9493	0.8331
0.3	0.9976	0.9843	0.9531	0.9592	0.9853	0.9075	0.9537	0.8339
0.5	0.9992	0.991	0.9479	0.9734	0.9912	0.9004	0.9478	0.8357
0.8	0.9993	0.9949	0.9446	0.9858	0.9942	0.8984	0.9456	0.8347
Width								
ρ	BW	UG	MG	MVNS	MVtS	MVNSSig	MVtSSig	BSR
-0.8	0.12323	0.09096	0.11676	0.07421	0.09094	0.09437	0.11676	0.07729
-0.5	0.15006	0.10801	0.13024	0.08831	0.108	0.10548	0.13015	0.08598
-0.3	0.10607	0.07587	0.0855	0.06155	0.0759	0.06891	0.08551	0.05597
0	0.07877	0.05614	0.05561	0.04556	0.05616	0.04489	0.05562	0.03654
0.3	0.10658	0.07625	0.06365	0.06187	0.07627	0.05122	0.06365	0.04168
0.5	0.14916	0.10716	0.07863	0.08764	0.10714	0.06387	0.07861	0.0515
0.8	0.12238	0.09036	0.05448	0.0738	0.09038	0.04412	0.05449	0.03591

Additive Desirability Inference Observation 1, $(x_1, x_2) =$ (-1,-1) Coverage and Width

FO Max/Tgt Matching Plane Angle and Correlation (1/3)

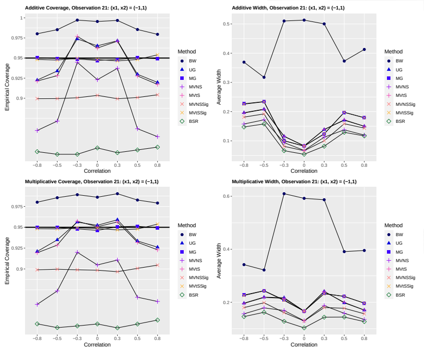


Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt

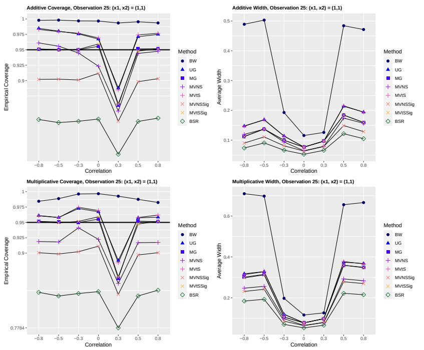


Additive and Multiplicative Plots for Coverage and Width at Observation 5, Max/Tgt

FO Max/Tgt Matching Plane Angle and Correlation (2/3)

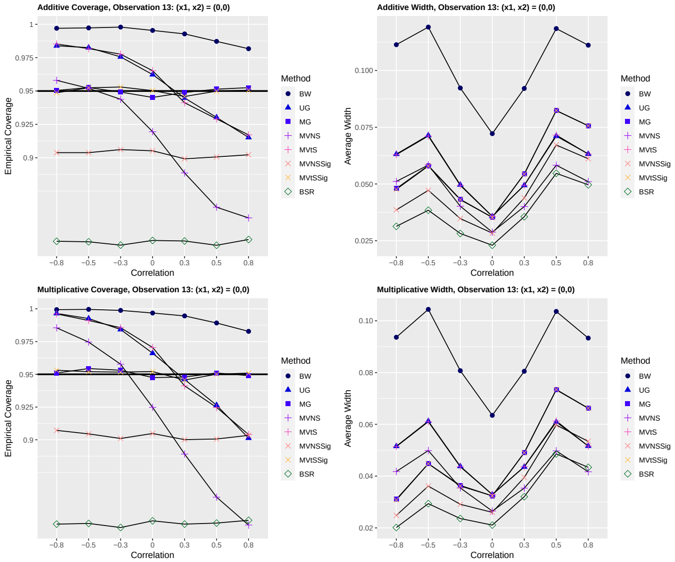


Additive and Multiplicative Plots for Coverage and Width at Observation 21, Max/Tgt

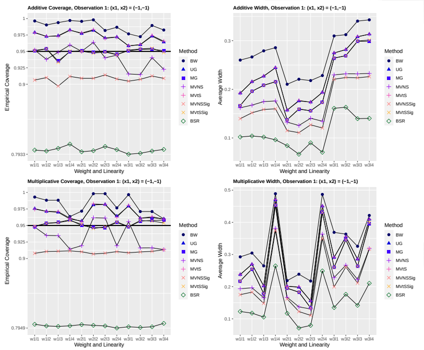


Additive and Multiplicative Plots for Coverage and Width at Observation 25, Max/Tgt

FO Max/Tgt Matching Plane Angle and Correlation (3/3)



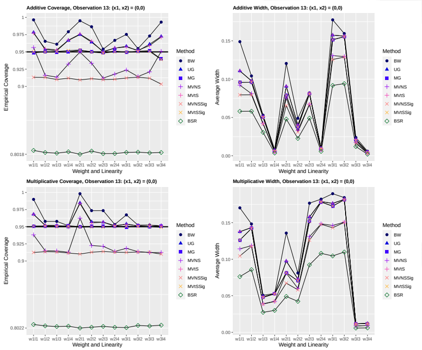
SO Max/Min Surface Models (1/2)



Additive and Multiplicative Plots for Coverage and Width at Observation 1 (Max/Min Second Order)

- Solution location favors Y_3
- $\boldsymbol{w}_1$ equal weights, $\boldsymbol{w}_2$ favors Y_1, $\boldsymbol{w}_3$ favors Y_3
- $\boldsymbol{l}_1$ equal linearity, $\boldsymbol{l}_2$ less emphasis on Y_1, $\boldsymbol{l}_3$ more emphasis Y_3, $\boldsymbol{l}_4$ less Y_1 and more Y_3
- MG and MVtSSig maintain coverage
- Multiplicative width becomes large when $\boldsymbol{l}_4$ active because Y_3 is near optimal and Y_1 is not

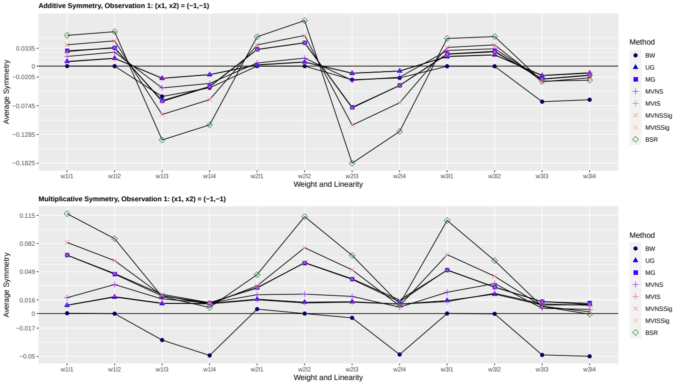
SO Max/Min Surface Models (2/2)



Additive and Multiplicative Plots for Coverage and Width at Observation 13 (Max/Min Second Order)

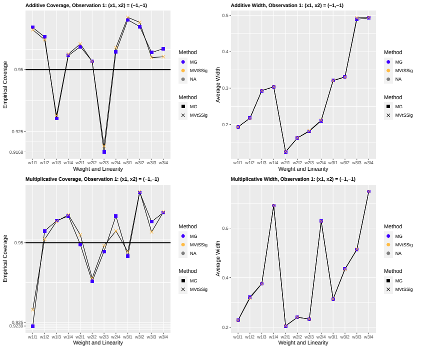
- Solution location favors Y_1
- $\boldsymbol{w}_1$ equal weights, $\boldsymbol{w}_2$ favors Y_1, $\boldsymbol{w}_3$ favors Y_3
- $\boldsymbol{l}_1$ equal linearity, $\boldsymbol{l}_2$ less emphasis on Y_1, $\boldsymbol{l}_3$ more emphasis Y_3, $\boldsymbol{l}_4$ less Y_1 and more Y_3
- MG and MVtSSig maintain coverage
- Widths roughly half size of corner point where prediction variance is lower

SO Max/Tgt/Min Surface Models (1/14)

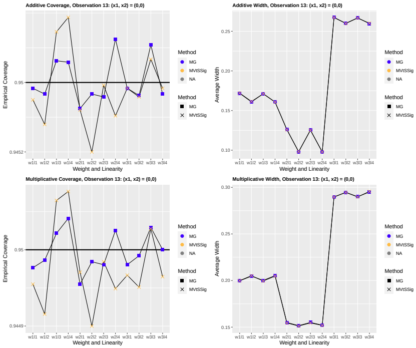


Additive and Multiplicative Plots for Symmetry at Observation 1, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (2/14)

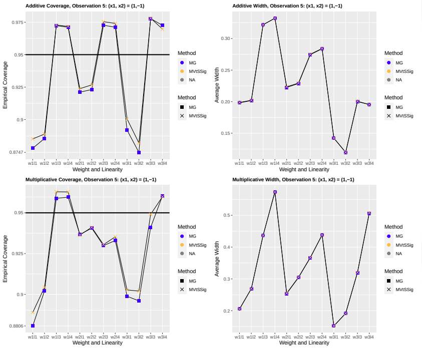


Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt/Min Second Order

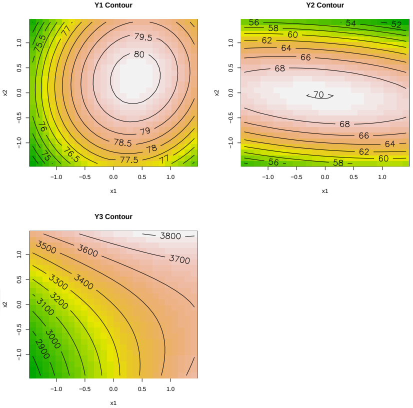


Additive and Multiplicative Plots for Coverage and Width at Observation 13, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (3/14)

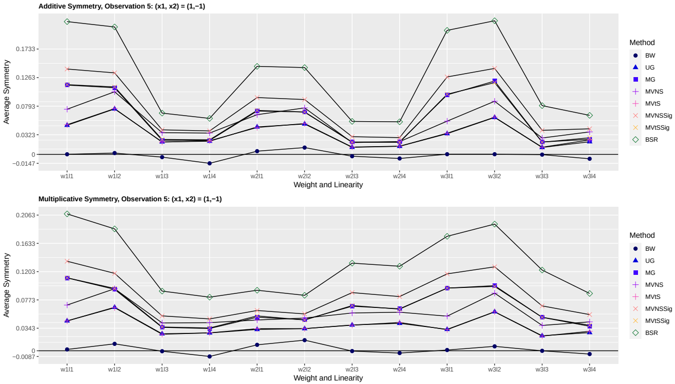


Additive and Multiplicative Plots for Coverage and Width at Observation 5, Max/Tgt/Min Second Order



Contour Plot for Second Order Models

SO Max/Tgt/Min Surface Models (4/14)

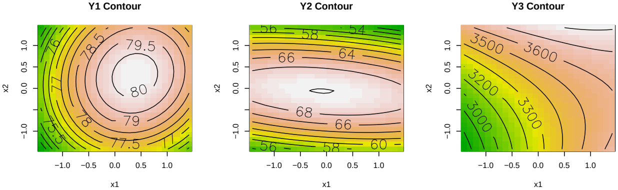


Additive and Multiplicative Plots for Symmetry at Observation 5, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (5/14)

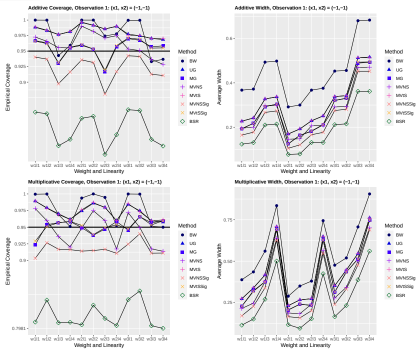
Optimal and Worst Solutions for each $\boldsymbol{w}_i \boldsymbol{l}_j$ Combination

w_i	l_j	Add Max	Mult Max	$d_1 Max$	$d_2 Max$	$d_3 Max$	Add Min	Mult Min	$d_1 Min$	$d_2 Min$	$d_3 Min$
1	1	19	3	19	4	1	21	1	1	13	25
1	2	3	2	19	4	1	25	25	1	13	25
1	3	19	19	19	20	1	21	25	1	25	25
1	4	6	6	19	20	1	25	25	1	25	25
2	1	3	3	19	4	1	1	1	1	13	25
2	2	4	3	19	4	1	23	23	1	13	25
2	3	19	19	19	20	1	21	24	1	25	25
2	4	19	7	19	20	1	24	24	1	25	25
3	1	4	4	19	4	1	21	21	1	13	25
3	2	4	2	19	4	1	24	24	1	13	25
3	3	19	19	19	20	1	25	25	1	25	25
3	4	20	8	19	20	1	25	25	1	25	25



SO Max/Tgt/Min Surface Models (6/14)

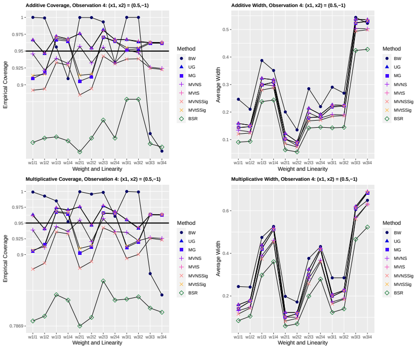
- Solution at location $\boldsymbol{x} = (-1, -1)$ favors Y_3
- $\boldsymbol{w}_1$ equal weights, $\boldsymbol{w}_2$ favors Y_1, $\boldsymbol{w}_3$ favors Y_2
- $\boldsymbol{l}_1$ equal linearity, $\boldsymbol{l}_2$ less emphasis on Y_1, $\boldsymbol{l}_3$ more emphasis lower half of Y_2, $\boldsymbol{l}_4$ less Y_1 and more Y_2 lower half
- Small overall desirability for w2l3 (add) and w1l1 (mult) resulting in less coverage
- l_4 is large emphasis on Y_2, results in large difference in values near optimal



Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt/Min Second Order

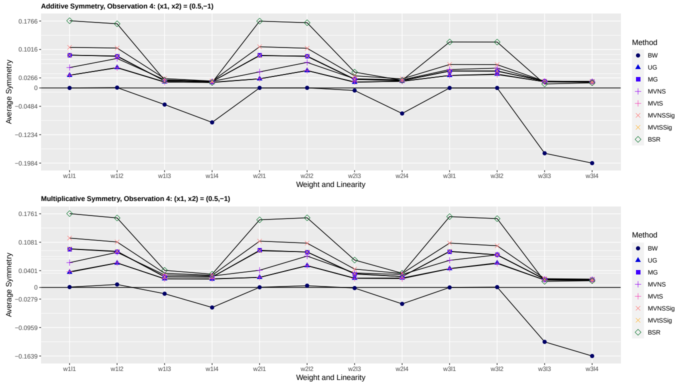
SO Max/Tgt/Min Surface Models (7/14)

- Solution at location $\boldsymbol{x} = (0.5, -1)$ favors Y_2
- $\boldsymbol{w}_1$ equal weights, $\boldsymbol{w}_2$ favors Y_1, $\boldsymbol{w}_3$ favors Y_2
- $\boldsymbol{l}_1$ equal linearity, $\boldsymbol{l}_2$ less emphasis on Y_1, $\boldsymbol{l}_3$ more emphasis lower half of Y_2, $\boldsymbol{l}_4$ less Y_1 and more Y_2 lower half
- Solution has minimum desirability for w2l1 (add) and w1l1, w2l2 (mult)



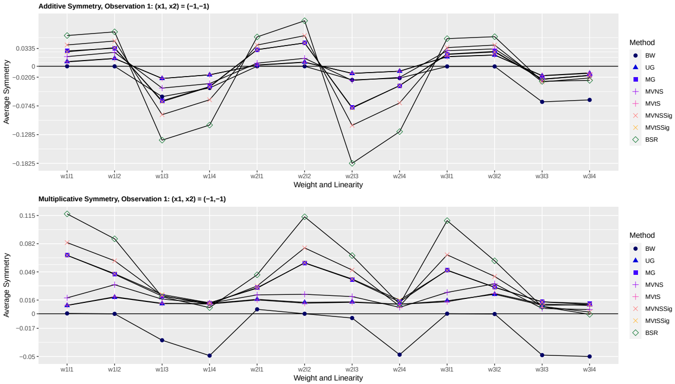
Additive and Multiplicative Plots for Coverage and Width at Observation 4, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (8/14)



Additive and Multiplicative Plots for Symmetry at Observation 4, Max/Tgt/Min Second Order

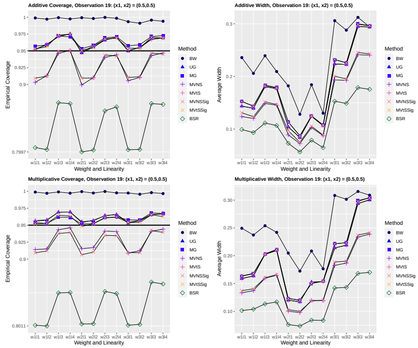
SO Max/Tgt/Min Surface Models (9/14)



Additive and Multiplicative Plots for Symmetry at Observation 1, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (10/14)

- Solution at location $\boldsymbol{x} = (0.5, 0.5)$ favors Y_1
- $\boldsymbol{w}_1$ equal weights, $\boldsymbol{w}_2$ favors Y_1, $\boldsymbol{w}_3$ favors Y_2
- $\boldsymbol{l}_1$ equal linearity, $\boldsymbol{l}_2$ less emphasis on Y_1, $\boldsymbol{l}_3$ more emphasis lower half of Y_2, $\boldsymbol{l}_4$ less Y_1 and more Y_2 lower half

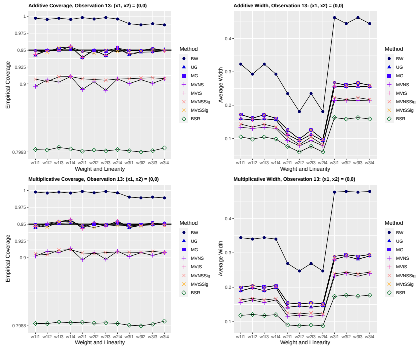


Additive and Multiplicative Plots for Coverage and Width at Observation 19, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (11/14)

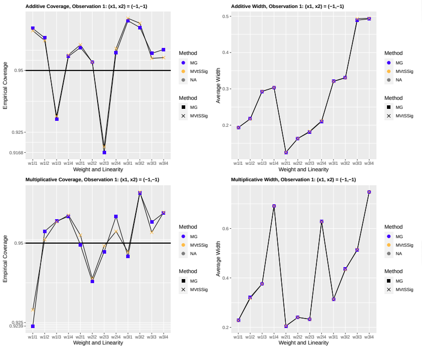
- Solution at location $\boldsymbol{x} = (0, 0)$ favors Y_1
- Low prediction variance makes coverage consistent for weights and linearity
- Width makes appropriate changes to accommodate coverage needs

$$\boldsymbol{C} = \begin{pmatrix} 1 & -0.42893 & 0.29331 \\ -0.42893 & 1 & 0.094275 \\ 0.29331 & 0.094275 & 1 \end{pmatrix}$$

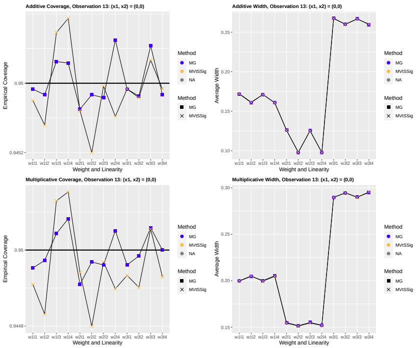


Additive and Multiplicative Plots for Coverage and Width at Observation 13, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (12/14)

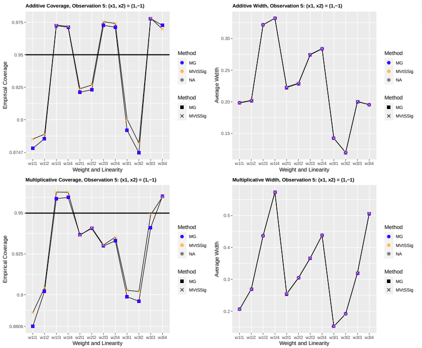


Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt/Min Second Order

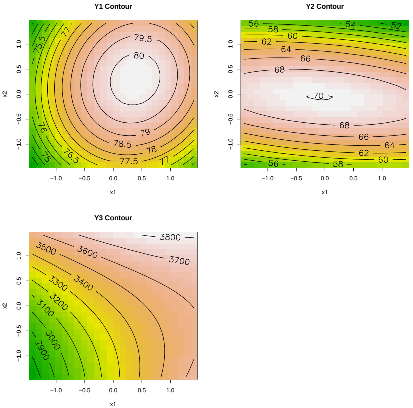


Additive and Multiplicative Plots for Coverage and Width at Observation 13, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models (13/14)

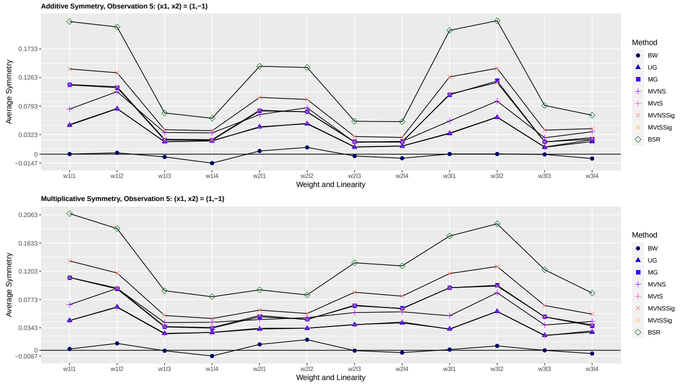


Additive and Multiplicative Plots for Coverage and Width at Observation 5, Max/Tgt/Min Second Order



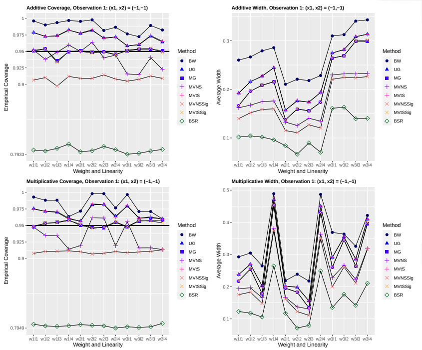
Contour Plot for Second Order Models

SO Max/Tgt/Min Surface Models (14/14)

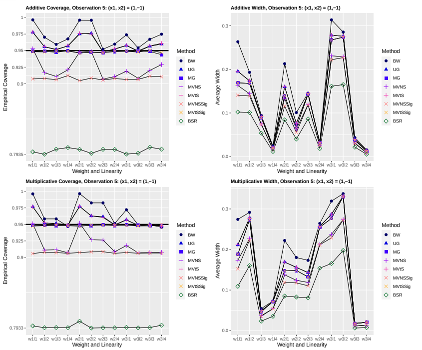


Additive and Multiplicative Plots for Symmetry at Observation 5, Max/Tgt/Min Second Order

SO Max/Min Surface Models Corner Points (1/2)

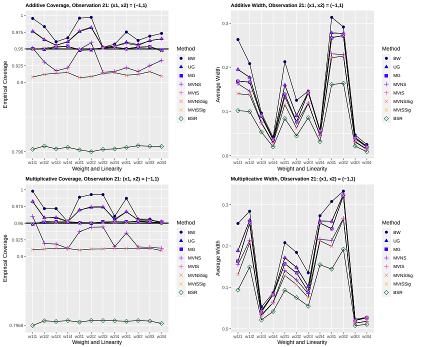


Additive and Multiplicative Plots for Coverage and Width at Observation 1, (Max/Min Second Order)

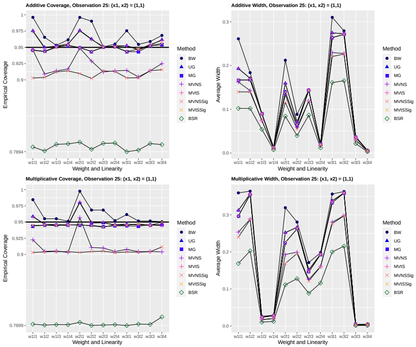


Additive and Multiplicative Plots for Coverage and Width at Observation 5, (Max/Min Second Order)

SO Max/Min Surface Models Corner Points (2/2)

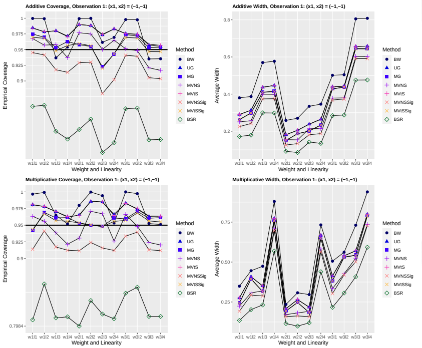


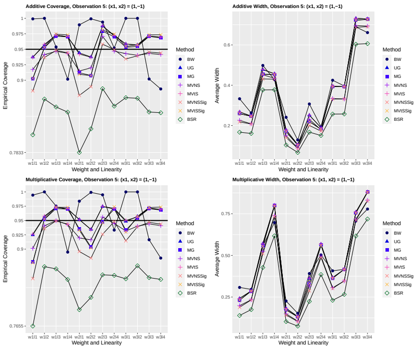
Additive and Multiplicative Plots for Coverage and Width at Observation 21, (Max/Min Second Order)



Additive and Multiplicative Plots for Coverage and Width at Observation 25, (Max/Min Second Order)

SO Max/Tgt Surface Models (1/3)

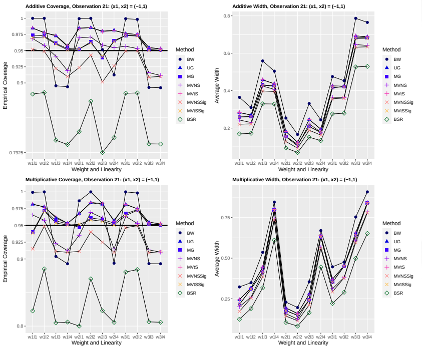


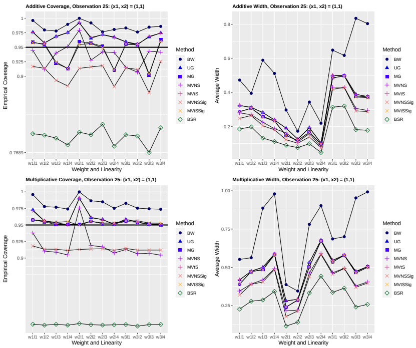


Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt Second Order

Additive and Multiplicative Plots for Coverage and Width at Observation 5, Max/Tgt Second Order

SO Max/Tgt Surface Models (2/3)

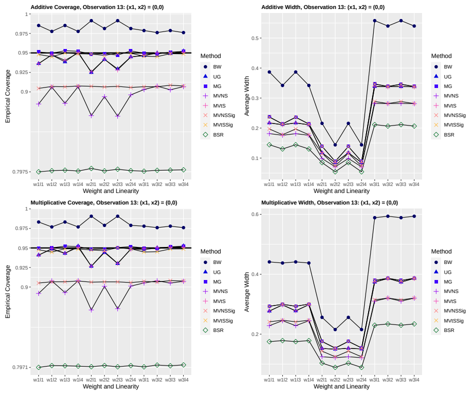




Additive and Multiplicative Plots for Coverage and Width at Observation 21, Max/Tgt Second Order

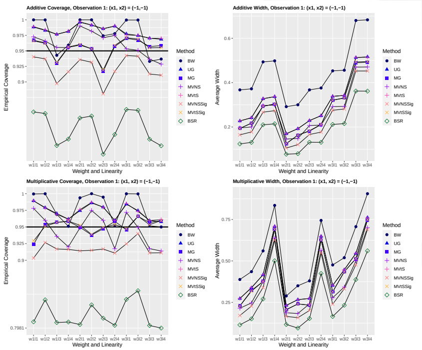
Additive and Multiplicative Plots for Coverage and Width at Observation 25, Max/Tgt Second Order

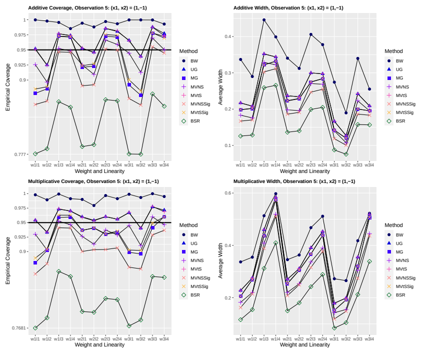
SO Max/Tgt Surface Models (3/3)



Additive and Multiplicative Plots for Coverage and Width at Observation 13, Max/Tgt Second Order

SO Max/Tgt/Min Surface Models Corner Points (1/2)

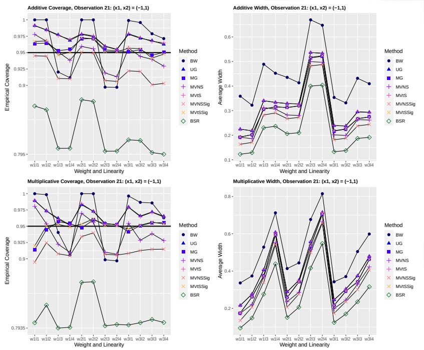


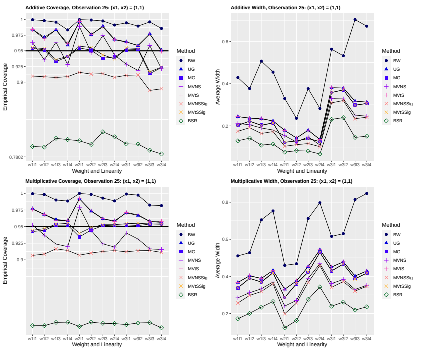


Additive and Multiplicative Plots for Coverage and Width at Observation 1, Max/Tgt/Min Second Order

Additive and Multiplicative Plots for Coverage and Width at Observation 5, Max/Tgt/Min Second Order

SO Max/Tgt/Min Surface Models Corner Points (2/2)

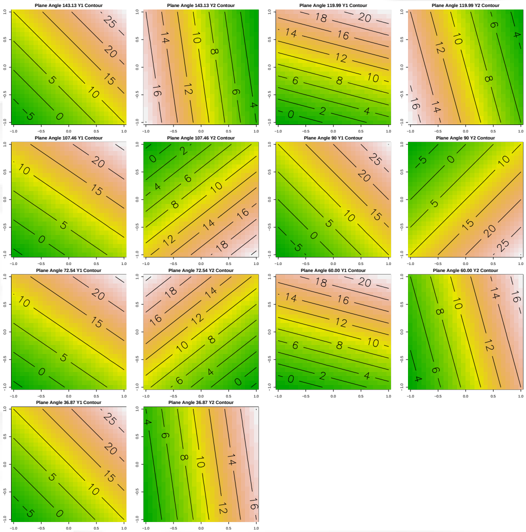




Additive and Multiplicative Plots for Coverage and Width at Observation 21, Max/Tgt/Min Second Order

Additive and Multiplicative Plots for Coverage and Width at Observation 25, Max/Tgt/Min Second Order

FO Contours



Contour Plots for First Order Models of Differing Plane Angles