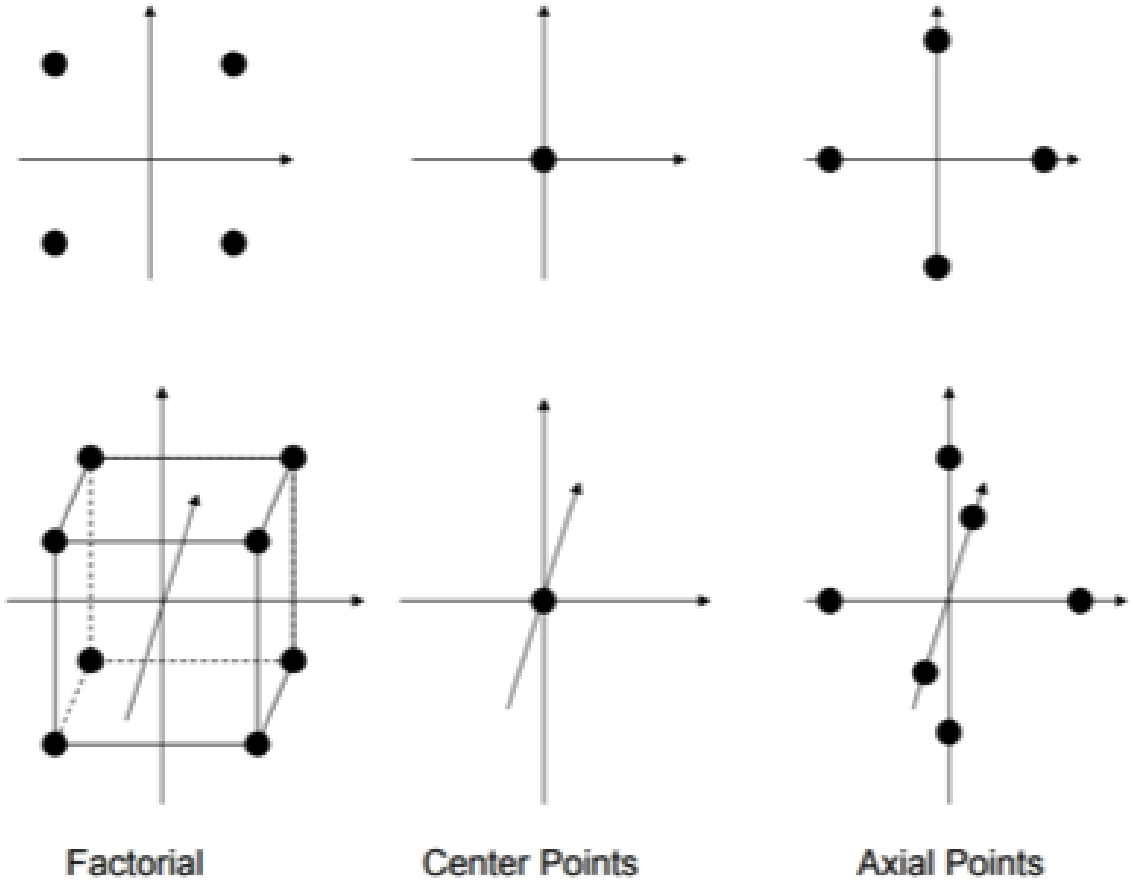
# Response Surface Methodology II

BIOE 498/598

3/8/2020

# The Central Composite Design (CCD)




1. A $2^k$ factorial or a Resolution V $2^{k-p}$ fractional design allow estimation of the FO and TWI terms.
2. The axial points estimate the pure quadratic (PQ) terms.
3. Replicated center points estimate the model's precision.

# Coding the CCD

You are optimizing protein expression levels in CHO cells by varying the serum concentration, casamino acid concentration, salt concentration, and the concentration of the induction molecule (IPTG).

You want to vary [IPTG] linearly over the range of 0.05 to 0.2 mM. What are the levels you would use in a CCD design?

$$\alpha = \sqrt[4]{F}$$

$$\text{A} = \text{center(A)} + \frac{\text{range(A)}}{2\alpha}[\text{code}]$$

# Cement Workability Experiment

Table 10.1 *CCD in Coded and Actual Units for Cement Workability Experiment*

| Run | $x_1$ | $x_2$ | $x_3$ | Water/Cement | Black Liq. | SNF | $y$ |
|---|---|---|---|---|---|---|---|
| 1 | −1 | −1 | −1 | 0.330 | 0.120 | 0.080 | 109.5 |
| 2 | 1 | −1 | −1 | 0.350 | 0.120 | 0.080 | 120.0 |
| 3 | −1 | 1 | −1 | 0.330 | 0.180 | 0.080 | 110.5 |
| 4 | 1 | 1 | −1 | 0.350 | 0.180 | 0.080 | 124.5 |
| 5 | −1 | −1 | 1 | 0.330 | 0.120 | 0.120 | 117.0 |
| 6 | 1 | −1 | 1 | 0.350 | 0.120 | 0.120 | 130.0 |
| 7 | −1 | 1 | 1 | 0.330 | 0.180 | 0.120 | 121.0 |
| 8 | 1 | 1 | 1 | 0.350 | 0.180 | 0.120 | 132.0 |
| 9 | 0 | 0 | 0 | 0.340 | 0.150 | 0.100 | 117.0 |
| 10 | 0 | 0 | 0 | 0.340 | 0.150 | 0.100 | 117.0 |
| 11 | 0 | 0 | 0 | 0.340 | 0.150 | 0.100 | 115.0 |
| 12 | −1.68 | 0 | 0 | 0.323 | 0.150 | 0.100 | 109.5 |
| 13 | 1.68 | 0 | 0 | 0.357 | 0.150 | 0.100 | 132.0 |
| 14 | 0 | −1.68 | 0 | 0.340 | 0.100 | 0.100 | 120.0 |
| 15 | 0 | 1.68 | 0 | 0.340 | 0.200 | 0.100 | 121.0 |
| 16 | 0 | 0 | −1.68 | 0.340 | 0.150 | 0.066 | 115.0 |
| 17 | 0 | 0 | 1.68 | 0.340 | 0.150 | 0.134 | 127.0 |
| 18 | 0 | 0 | 0 | 0.340 | 0.150 | 0.100 | 116.0 |
| 19 | 0 | 0 | 0 | 0.340 | 0.150 | 0.100 | 117.0 |
| 20 | 0 | 0 | 0 | 0.340 | 0.150 | 0.100 | 117.0 |

# Correct number of center points (Box and Hunter 1957)

| factors ($k$) | 2 | 3 | 4 | 5 | $5-1$ | 6 |
|---|---|---|---|---|---|---|
| factorial points | 4 | 8 | 16 | 32 | 16 | 64 |
| axial points | 4 | 6 | 8 | 10 | 10 | 12 |
| center points | 5 | 6 | 7 | 10 | 6 | 15 |
| $\alpha$ | 1.414 | 1.682 | 2.000 | 2.378 | 2.000 | 2.828 |

| factors ($k$) | $6-1$ | 7 | $7-1$ | 8 | $8-1$ | $8-2$ |
|---|---|---|---|---|---|---|
| factorial points | 32 | 128 | 64 | 256 | 128 | 64 |
| axial points | 12 | 14 | 14 | 16 | 16 | 16 |
| center points | 9 | 21 | 14 | 28 | 20 | 13 |
| $\alpha$ | 2.378 | 3.364 | 2.828 | 4.000 | 3.364 | 2.828 |

# Two-stage RSM with a CCD

1. Run the factorial design and ~1/2 of the center points.
2. Fit a FO and TWI model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{i} \beta_{ij} x_i x_j$$

3. If the fit is “good enough", stop. Otherwise, you need a quadratic model.
4. Run the axial points and the rest of the center points.
5. Fit the full quadratic model with a blocking factor

$$y = \beta_0 + \text{FO}(x_1, \ldots, x_k) + \text{TWI}(x_1, \ldots, x_k) + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \beta_{\text{block}} \text{block}$$

6. Optimize

# Box-Behnken Design

- A CCD is the optimal RSM design (uniform precision, rotatable, can estimate curvature).
- However, it requires five levels per factor ($-\alpha$, $-1$, 0, 1, $\alpha$).
- Five levels can be prohibitive if each level requires a separate prototype or large changes to processes.
- The Box-Behnken Design requires only three levels per factor ($-1$, 0, 1).

# Box-Behnken Design

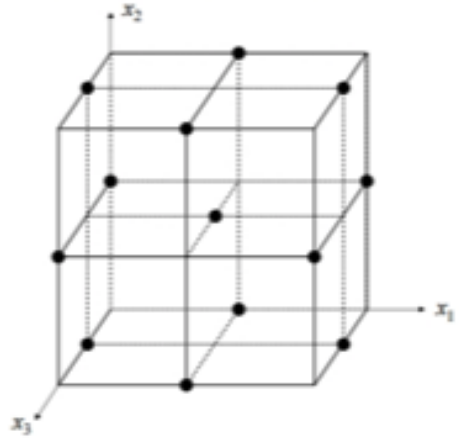
- A CCD is the optimal RSM design (uniform precision, rotatable, can estimate curvature).
- However, it requires five levels per factor ($-\alpha$, $-1$, 0, 1, $\alpha$).
- Five levels can be prohibitive if each level requires a separate prototype or large changes to processes.
- The Box-Behnken Design requires only three levels per factor ($-1$, 0, 1).

| run | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|
| 1 | −1 | −1 | 0 |
| 2 | 1 | −1 | 0 |
| 3 | −1 | 1 | 0 |
| 4 | 1 | 1 | 0 |
| 5 | −1 | 0 | −1 |
| 6 | 1 | 0 | −1 |
| 7 | −1 | 0 | 1 |
| 8 | 1 | 0 | 1 |
| 9 | 0 | −1 | −1 |
| 10 | 0 | 1 | −1 |
| 11 | 0 | −1 | 1 |
| 12 | 0 | 1 | 1 |
| 13 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 |

# Box-Behnken Design — disadvantages

1. The BBD runs cannot be split into two blocks.

Figure 10.4 *Graphical Representation of Three-Factor Box-Behnken Design*

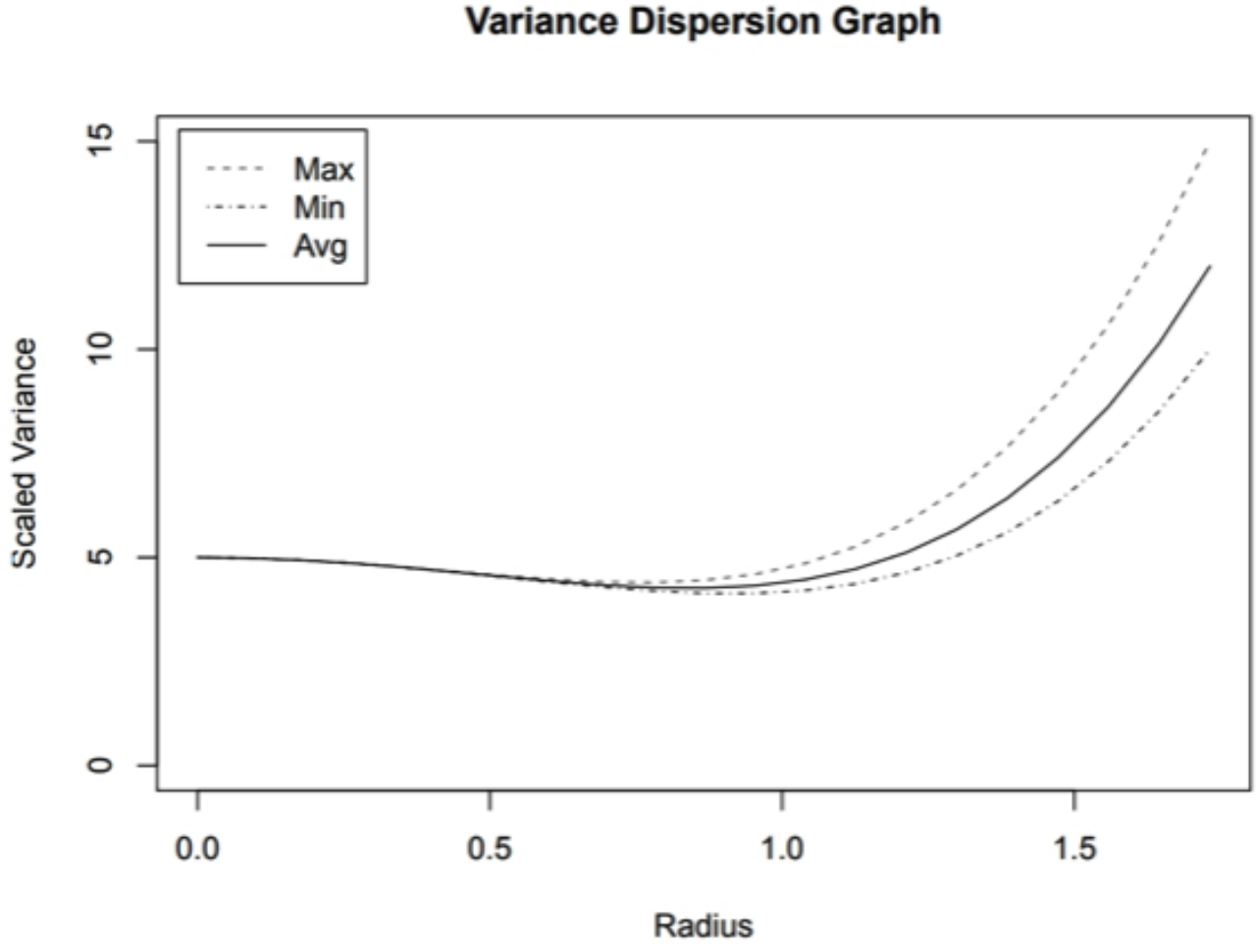



If you are sure your experiment requires quadratic terms, this is not a disadvantage.

# Box-Behnken Design — disadvantages

2. The BBD is not rotatable and barely misses uniform precision.

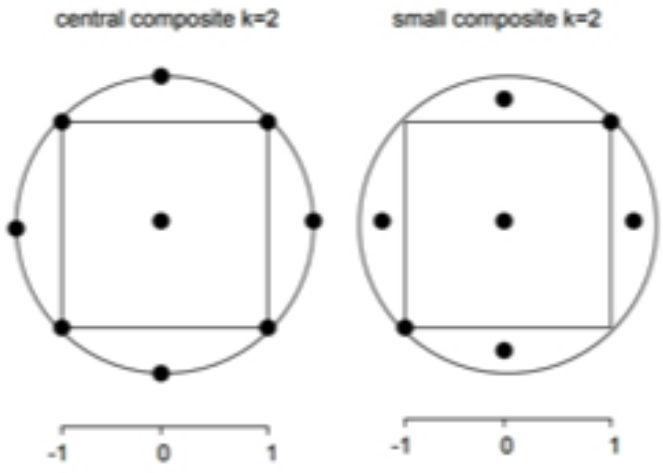
Figure 10.5 *Variance Dispersion Graph for Box-Behnken Design in Three Factors*



# Small Composite Designs — if runs *really* matter

A true CCD contains a $2^k$ factorial or resolution V $2^{k-p}$ fractional design.

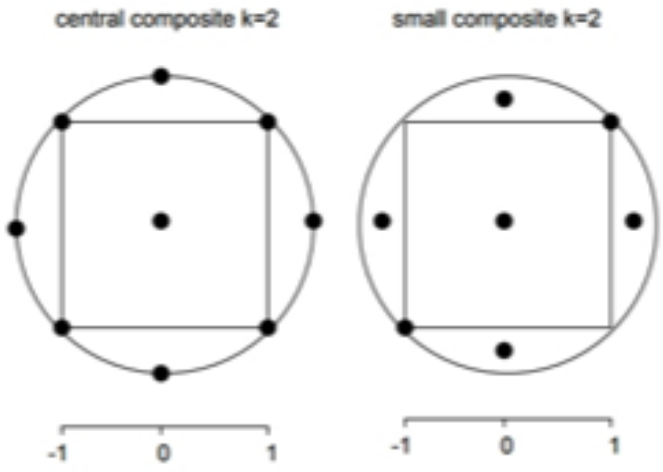
A Small Composite Design (SCD) replaces these with a resolution III $2^{k-p}$ design.



# Small Composite Designs — if runs *really* matter

A true CCD contains a $2^k$ factorial or resolution V $2^{k-p}$ fractional design.

A Small Composite Design (SCD) replaces these with a resolution III $2^{k-p}$ design.



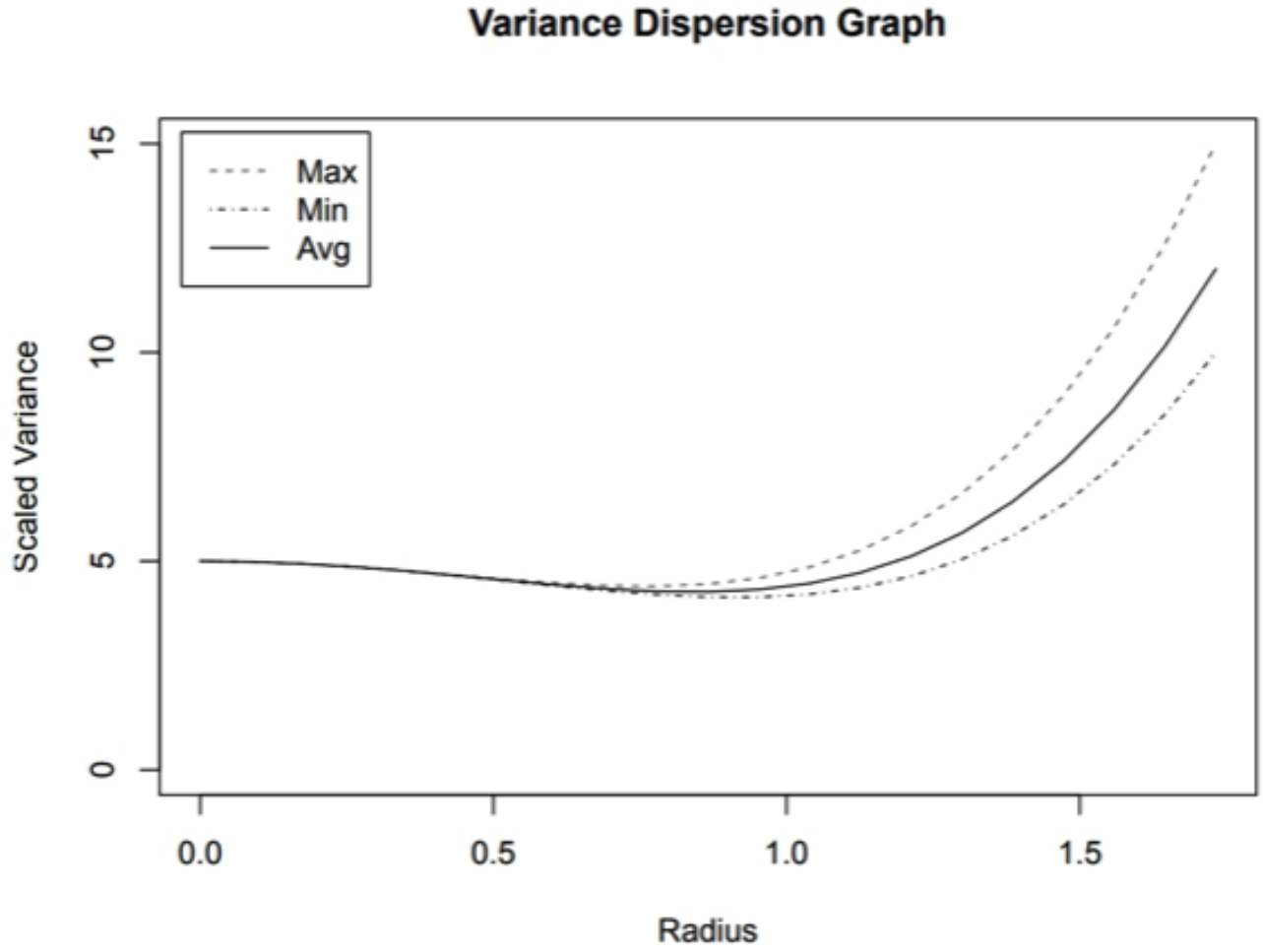

The axial and center points make up for the lost factorial points and allow estimation of the FO and TWI terms.

You can also use a PB design or a custom fractional design for even fewer runs.

# Small Composite Designs — disadvantages

2. The SCD is not rotatable and has uniform precision only to radius 0.8.

Figure 10.5 *Variance Dispersion Graph for Box-Behnken Design in Three Factors*

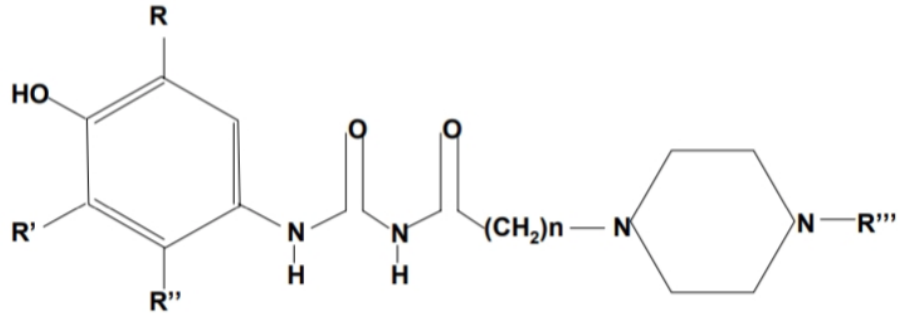



Estimates can be terrible beyond radius 1.

# RSM design example 1: QSAR

We want to perform a Quantitative Structure Activity Relation (QSAR) analysis on hydroxyphenylureas.

Figure 10.12 *General Structure of Hydroxyphenylureas*




The response is oxygen free radical scavenging quantified by the binding constant. Factors are:

- Hydration energy (HE)
- Molecular dipole moment ($DM_z$)
- Symmetry index (S0K)

# RSM design example 1: QSAR

Table 10.6 *Library of Substituted Hydroxyphenylurea Compounds*

| Compound | R | R′ | R″ | R‴ | HE | DMz | S0K |
|---|---|---|---|---|---|---|---|
| 1 | H | H | H | $CH_3$ | −12.221 | −0.162 | 64.138 |
| 2 | H | H | H | $CH_2Ph$ | −14.015 | −0.068 | 88.547 |
| 3 | H | H | H | Ph | −14.502 | 0.372 | 85.567 |
| 4 | H | H | H | $2CH_3OC_6H_4$ | −14.893 | 1.035 | 96.053 |
| 5 | H | $OCH_3$ | H | $CH_3$ | −12.855 | 1.091 | 74.124 |
| 6 | H | $OCH_3$ | H | $CH_2Ph$ | −14.628 | 1.115 | 99.002 |
| 7 | H | $OCH_3$ | H | Ph | −15.123 | 1.554 | 96.053 |
| 8 | H | $OCH_3$ | H | $2CH_3OC_6H_4$ | −15.492 | 2.221 | 106.607 |
| 9 | H | $OC_2H_5$ | H | $CH_3$ | −11.813 | 1.219 | 77.02 |
| 10 | H | $OC_2H_5$ | H | $CH_2Ph$ | −13.593 | 1.188 | 101.978 |
| 11 | H | $OC_2H_5$ | H | Ph | −14.088 | 1.621 | 99.002 |
| 12 | $CH_3$ | $OC_2H_5$ | H | $2CH_3OC_6H_4$ | −14.46 | 2.266 | 109.535 |
| 13 | $CH_3$ | H | $CH_3$ | $CH_3$ | −8.519 | −0.56 | 71.949 |
| 14 | $CH_3$ | H | $CH_3$ | $CH_2Ph$ | −10.287 | −0.675 | 96.6 |
| 15 | $CH_3$ | H | $CH_3$ | Ph | −10.798 | −0.134 | 96.62 |
| 16 | $CH_3$ | H | $CH_3$ | $2CH_3OC_6H_4$ | −11.167 | 0.418 | 104.047 |

# RSM design example 1: QSAR

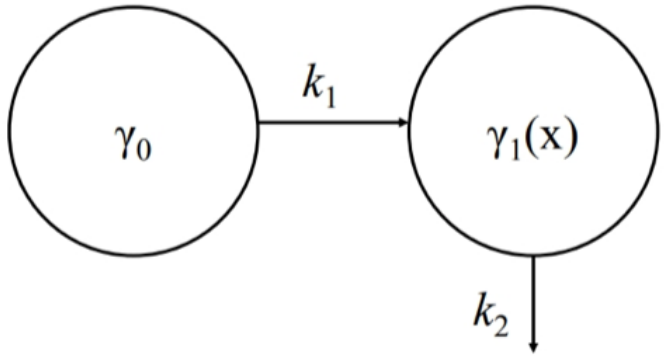
Table 10.6 *Library of Substituted Hydroxyphenylurea Compounds*

| Compound | R | R′ | R″ | R‴ | HE | DMz | S0K |
|---|---|---|---|---|---|---|---|
| 1 | H | H | H | $CH_3$ | −12.221 | −0.162 | 64.138 |
| 2 | H | H | H | $CH_2Ph$ | −14.015 | −0.068 | 88.547 |
| 3 | H | H | H | Ph | −14.502 | 0.372 | 85.567 |
| 4 | H | H | H | $2CH_3OC_6H_4$ | −14.893 | 1.035 | 96.053 |
| 5 | H | $OCH_3$ | H | $CH_3$ | −12.855 | 1.091 | 74.124 |
| 6 | H | $OCH_3$ | H | $CH_2Ph$ | −14.628 | 1.115 | 99.002 |
| 7 | H | $OCH_3$ | H | Ph | −15.123 | 1.554 | 96.053 |
| 8 | H | $OCH_3$ | H | $2CH_3OC_6H_4$ | −15.492 | 2.221 | 106.607 |
| 9 | H | $OC_2H_5$ | H | $CH_3$ | −11.813 | 1.219 | 77.02 |
| 10 | H | $OC_2H_5$ | H | $CH_2Ph$ | −13.593 | 1.188 | 101.978 |
| 11 | H | $OC_2H_5$ | H | Ph | −14.088 | 1.621 | 99.002 |
| 12 | $CH_3$ | $OC_2H_5$ | H | $2CH_3OC_6H_4$ | −14.46 | 2.266 | 109.535 |
| 13 | $CH_3$ | H | $CH_3$ | $CH_3$ | −8.519 | −0.56 | 71.949 |
| 14 | $CH_3$ | H | $CH_3$ | $CH_2Ph$ | −10.287 | −0.675 | 96.6 |
| 15 | $CH_3$ | H | $CH_3$ | Ph | −10.798 | −0.134 | 96.62 |
| 16 | $CH_3$ | H | $CH_3$ | $2CH_3OC_6H_4$ | −11.167 | 0.418 | 104.047 |

structures $\rightarrow$ physical parameters $\rightarrow$ response
$\rightarrow$ optimal parameters $\rightarrow$ optimal structures

# RSM design example 2: Tetracycline metabolism

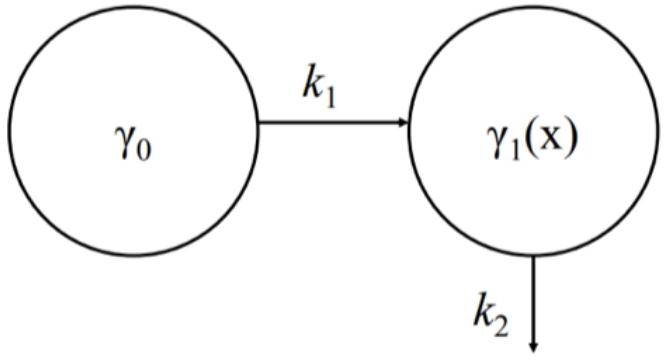
We want to build a kinetic model of tetracycline metabolism with four parameters.




$$y = \gamma_1(x) = \gamma_0[e^{-k_1(x-t_0)} - e^{-k_2(x-t_0)}]$$

# RSM design example 2: Tetracycline metabolism

We want to build a kinetic model of tetracycline metabolism with four parameters.




$$y = \gamma_1(x) = \gamma_0[e^{-k_1(x-t_0)} - e^{-k_2(x-t_0)}]$$

The input is the time ($x$) since the addition of the drug; however, we make the design optimal for the parameters.

We will come back to this example after break.