

Chapter 14

1. (a) The response when $A = 0$ and $B = 0$ is 1.68, which we represent as $(A, B, \text{response})$ or, in this case, $(0, 0, 1.68)$. For the first cycle, we increase A in steps of one until the response begins to decrease or until we reach a boundary, obtaining the following additional results:

$$(1, 0, 1.88), (2, 0, 2.00), (3, 0, 2.04), (4, 0, 2.00)$$

For the second cycle, we return to $(3, 0, 2.04)$ and increase B in steps of one, obtaining these results:

$$(3, 1, 2.56), (3, 2, 3.00), (3, 3, 3.36), (3, 4, 3.64), \\ (3, 5, 3.84), (3, 6, 3.96), (3, 7, 4.00), (3, 8, 3.96)$$

For the third cycle, we return to $(3, 7, 4.00)$ and increase A in steps of one, obtaining a result of $(4, 7, 3.96)$. Because this response is smaller than our current best response of 4.00, we try decreasing A by a step of one, which gives $(2, 7, 3.96)$. Having explored the response in all directions around $(3, 7, 4.00)$, we know that the optimum response is 4.00 at $A = 3$ and $B = 7$.

Figure SM14.1a shows the progress of the optimization as a three-dimensional scatterplot with the figure's floor showing a contour plot of the response surface. Figure SM14.1b shows a three-dimensional surface plot of the response surface.

- (b) The response when $A = 0$ and $B = 0$ is 4.00, which we represent as $(0, 0, 4.00)$. For the first cycle, we increase A in steps of one until the response begins to decrease or until we reach a boundary, obtaining a results of $(1, 0, 3.60)$; as this response is smaller than the initial step, this ends the first cycle.

At this point, our best response is 2.04 at $A = 3$ and at $B = 0$.

At this point, our best response is 4.00 at $A = 3$ and at $B = 7$.

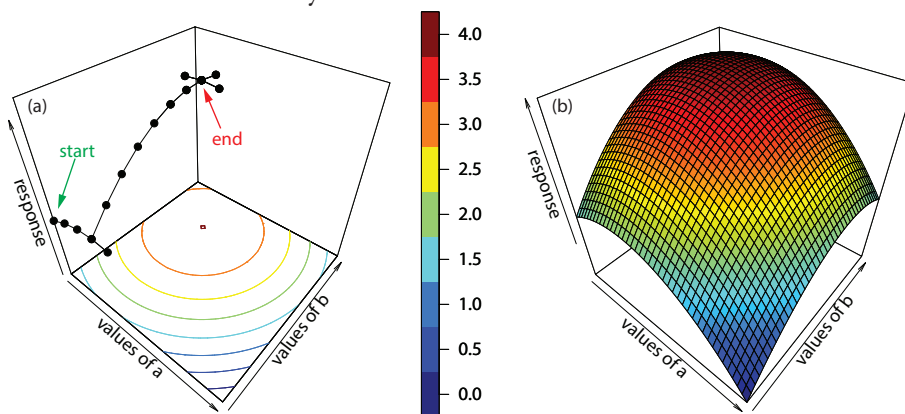


Figure SM14.1 The progress of a one-factor-at-a-time optimization for the equation in Problem 1a is shown in (a) as a scatterplot in three dimensions with a contour plot of the response surface on the figure's floor. The full response surface is shown in (b). The legend shows the colors used for the individual contour lines; the response surface provides for a greater resolution in the response by using gradations between these colors.

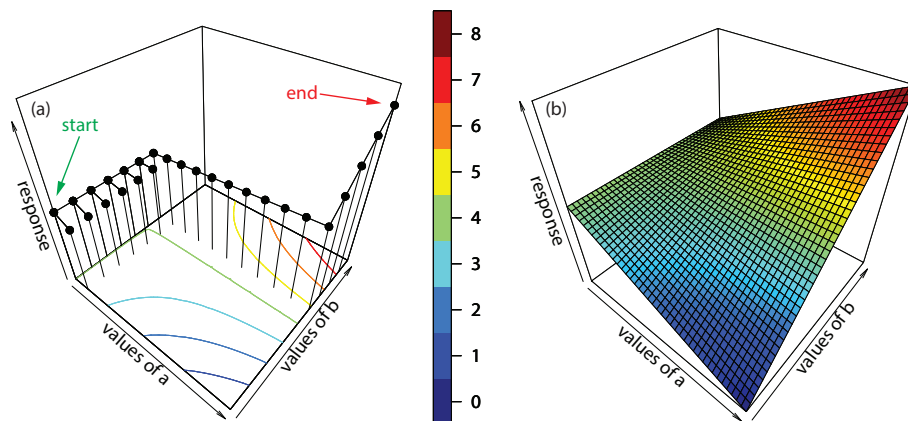


Figure SM14.2 The progress of a one-factor-at-a-time optimization for the equation in Problem 1b is shown in (a) as a scatterplot in three dimensions with a contour plot of the response surface on the figure's floor. The full response surface is shown in (b). The legend shows the colors used for the individual contour lines; the response surface provides for a greater resolution in the response by using gradations between these colors.

We begin the second cycle by returning to $(0, 0, 4.00)$ and increase the value of B by one, obtaining a result of $(0, 1, 4.00)$. Because the response did not increase, we end the second cycle and, for the third cycle, we increase the value of A , obtaining a result of $(1, 1, 3.68)$. Continuing in this fashion, the remainder of the steps are

$(0, 1, 4.00), (0, 2, 4.00), (1, 2, 3.76), (0, 2, 4.00), (0, 3, 4.00)$
 $(1, 3, 3.84), (0, 3, 4.00), (0, 4, 4.00), (1, 4, 3.92), (0, 4, 4.00)$
 $(0, 5, 4.00), (1, 5, 4.00), (0, 5, 4.00), (0, 6, 4.00), (1, 6, 4.08)$
 $(2, 6, 4.16), (3, 6, 4.24), (4, 6, 4.32), (5, 6, 4.40), (6, 6, 4.48)$
 $(7, 6, 4.56), (8, 6, 4.64), (9, 6, 4.72), (10, 6, 4.80), (10, 7, 5.60)$
 $(10, 8, 6.40), (10, 9, 7.20), (10, 10, 8.00)$

The optimum response is 8.00 at $A = 10$ and $B = 10$.

Figure SM14.2a shows the progress of the optimization as a three-dimensional scatterplot with the figure's floor showing a contour plot for the response surface. Figure SM14.2b shows a three-dimensional surface plot of the response surface.

(c) The response when $A = 0$ and $B = 0$ is 3.267, which we represent as $(0, 0, 3.267)$. For the first cycle, we increase A in steps of one until the response begins to decrease or until we reach a boundary, obtaining the following additional results:

$(1, 0, 4.651), (2, 0, 5.736), (3, 0, 6.521),$
 $(4, 0, 7.004), (5, 0, 7.187), (6, 0, 7.068)$

For the second cycle, we return to $(5, 0, 7.187)$ and increase B in steps of one, obtaining these results:

Note that until we reach $A = 0$ and $B = 6$, we keep probing toward larger values of A without increasing the response, and then probing toward larger values of B , also without increasing the response. Once we reach $A = 0$ and $B = 6$, however, we find that an increase in A finally increases the response. Once we reach the boundary for A , we continue to increase B until we reach the optimum response at $A = 10$ and $B = 10$.

At this point, our best response is 7.187 at $A = 5$ and at $B = 0$.

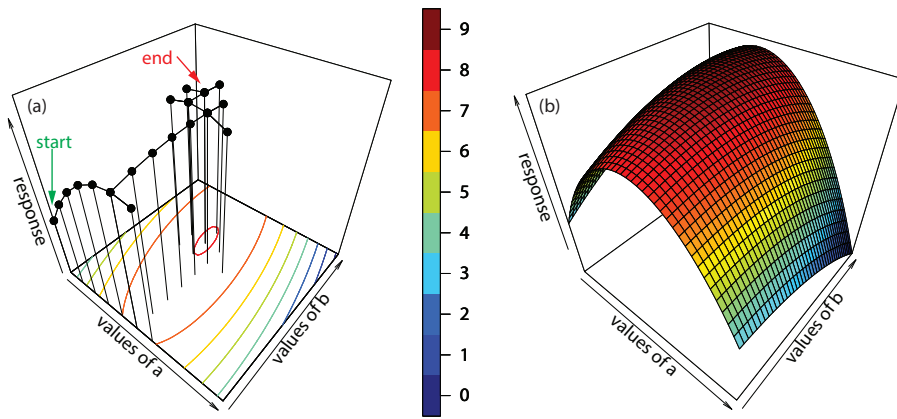


Figure SM14.3 The progress of a one-factor-at-a-time optimization for the equation in Problem 1c is shown in (a) as a scatterplot in three dimensions with a contour plot of the response surface on the figure's floor. The full response surface is shown in (b). The legend shows the colors used for the individual contour lines; the response surface provides for a greater resolution in the response by using gradations between these colors.

$$(5, 1, 7.436), (5, 2, 7.631), (5, 3, 7.772), \\ (5, 4, 7.858), (5, 5, 7.889), (5, 6, 7.865)$$

At this point, our best response is 7.889 at $A = 5$ and at $B = 5$.

For the next cycle, we return to $(5, 5, 7.889)$ and increase A in steps of one, obtaining a response for $(6, 5, 7.481)$ that is smaller; probing in the other direction gives $(4, 5, 7.996)$ and then $(3, 5, 7.801)$. Returning to $(4, 5, 7.966)$, we find our optimum response at $(4, 6, 8.003)$, with movement in all other directions giving a smaller response. Note that using a fixed step size of one prevents us from reaching the true optimum at $A = 3.91$ and $B = 6.22$.

Figure SM14.3a shows the progress of the optimization as a three-dimensional scatterplot with the figure's floor showing a contour plot for the response surface. Figure SM14.3b shows a three-dimensional surface plot of the response surface.

- Given a step size of 1.0 in both directions and $A = 0$ and $B = 0$ as the starting point for the first simplex, the other two vertices for the first simplex are at $A = 1$ and at $B = 0$, and at $A = 1.5$ and at $B = 0.87$. The responses for the first three vertices are $(0, 0, 3.264)$, $(1.0, 0, 4.651)$, and $(0.5, 0.87, 4.442)$, respectively. The vertex with the worst response is $(0, 0, 3.264)$; thus, we reject this vertex and replace it with coordinates of

$$A = 2\left(\frac{1 + 0.5}{2}\right) - 0 = 1.5 \\ B = 2\left(\frac{0.87 + 0}{2}\right) - 0 = 0.87$$

The following table summarizes all the steps in the simplex optimization. The column labeled “vertex” shows the 25 unique experiments along with their values for A , for B , and for the response. The column

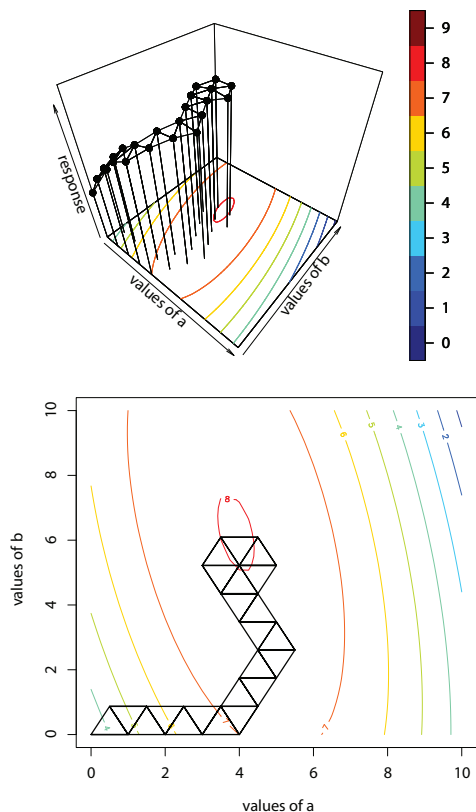


Figure SM14.4 Two views showing the progress of a simplex optimization of the equation in Problem 1c in (a) three dimensions and in (b) two dimensions. The legend shows the colors used for the individual contour lines. Figure SM14.3b shows the full response surface for this problem.

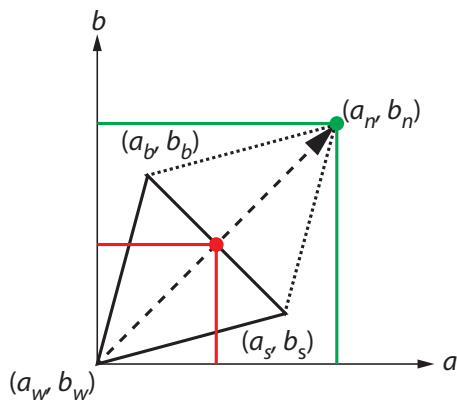


Figure SM14.5 Diagram showing vertices of original simplex and the reflection of the worst vertex across the midpoint (red circle) of the best and the next-best vertices to give the new vertex (green circle). See text for additional details.

labeled “simplex” shows the three vertices that make up each simplex. For each simplex, the vertex that we reject is shown in **bold** font; note that on two occasions, the rejected vertex, shown in ***bold-italic*** font, has the second-worst response (either because of a boundary condition or because the new vertex has the worst response)

vertex	A	B	response	simplex
1	0	0	3.264	—
2	1.0	0	4.651	—
3	0.5	0.87	4.442	1, 2, 3
4	1.5	0.87	5.627	2, 3, 4
5	2.0	0	5.736	2, 4, 5
6	2.5	0.87	6.512	4, 5, 6
7	3.0	0	6.521	5, 6, 7
8	3.5	0.87	7.096	6, 7, 8
9	4.0	0	7.004	7, 8, 9
10	4.5	0.87	7.378	8, 9, 10
11	4.0	1.74	7.504	8, 10, 11
12	5.0	1.74	7.586	10, 11, 12
13	4.5	2.61	7.745	11, 12, 13
14	5.5	2.61	7.626	12, 13, 14
15	5.0	3.48	7.820	13, 14, 15
16	4.0	3.48	7.839	13, 15, 16
17	4.5	4.35	7.947	15, 16, 17
18	3.5	4.35	7.866	16, 17, 18
19	4.0	5.22	8.008	17, 18, 19
20	5.0	5.22	7.888	17, 19, 20
21	4.5	6.09	7.983	19, 20, 21
22	3.5	6.09	8.002	19, 21, 22
23	3.0	5.22	7.826	19, 22, 23
24	3.5	4.35	7.866	19, 23, 24
25	4.5	4.35	7.947	19, 24, 25

Figure SM14.4 shows the progress of the simplex optimization in three dimensions and in two dimensions.

- To help us in the derivation, we will use the diagram shown in Figure SM14.5 where a and b are the coordinates of a vertex, and w , b , s , and n identify the vertex with, respectively, the worst response, the best response, the second-best response, and the new vertex. The **red** circle marks the midpoint between the best vertex and the second-best vertex; its coordinates are

$$a_{mp} = \frac{a_b + a_s}{2}$$

$$b_{mp} = \frac{b_b + b_s}{2}$$

The distance along the a -axis between the worst vertex's coordinate of a_w and the midpoint's coordinate of a_{mp} is

$$\frac{a_b + a_s}{2} - a_w$$

The distance along the a -axis between the worst vertex's coordinate and the new vertex's coordinate is twice that to the midpoint, which means the a coordinate for the new vertex is

$$a_n = 2\left(\frac{a_b + a_s}{2} - a_w\right) + a_w$$

which simplifies to equation 14.3

$$a_n = 2\left(\frac{a_b + a_s}{2}\right) - a_w$$

Using the same approach for coordinates relative to the b -axis yields equation 14.4

$$b_n = 2\left(\frac{b_b + b_s}{2}\right) - b_w$$

The value for the coordinate a_n is the value for the coordinate a_w plus the distance along the a -axis between the new vertex and the worst vertex.

4. In coded form, the values for b_0 , b_a , b_b , and b_{ab} are

$$b_0 = \frac{1}{4}(5.92 + 2.08 + 4.48 + 3.52) = 4.00$$

$$b_a = \frac{1}{4}(5.92 + 2.08 - 4.48 - 3.52) = 0$$

$$b_b = \frac{1}{4}(5.92 - 2.08 + 4.48 - 3.52) = 1.20$$

$$b_{ab} = \frac{1}{4}(5.92 - 2.08 - 4.48 + 3.52) = 0.72$$

which gives us the following equation for the response surface in coded form

$$R = 4.00 + 1.20B^* + 0.72A^*B^*$$

To convert this equation into its uncoded form, we first note the following relationships between coded and uncoded values for A and for B

$$A = 5 + 3A^* \quad B = 5 + 3B^*$$

$$A^* = \frac{A}{3} - \frac{5}{3} \quad B^* = \frac{B}{3} - \frac{5}{3}$$

Substituting these two equations back into the response surface's coded equation gives

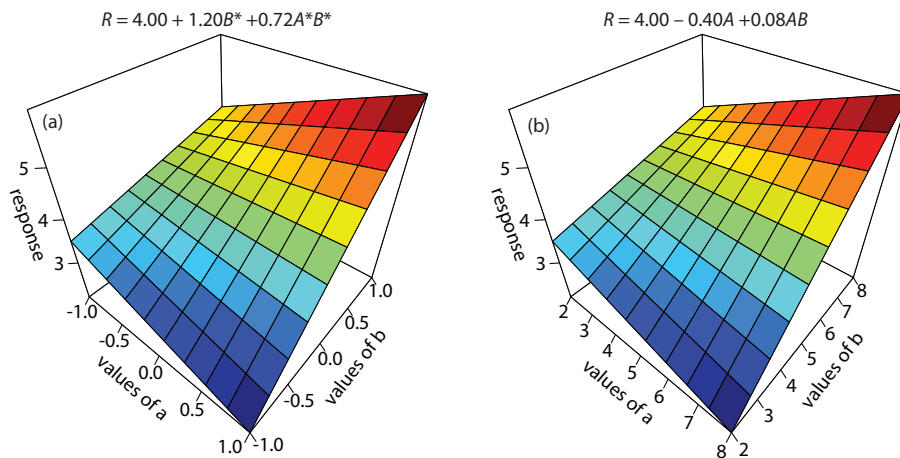


Figure SM14.6 Response surfaces based on the (a) coded and the (b) uncoded equations derived from the data in Problem 4. Note that the two response surfaces are identical even though their equations are very different.

$$R = 4.00 + 1.20\left(\frac{B}{3} - \frac{5}{3}\right) + 0.72\left(\frac{A}{3} - \frac{5}{3}\right)\left(\frac{B}{3} - \frac{5}{3}\right)$$

$$R = 4.00 + 0.40B - 2.00 + 0.08AB - 0.40A - 0.40B + 2.00$$

$$R = 4.00 - 0.40A + 0.08AB$$

When we examine carefully both equations, we see they convey the same information: that the system's response depends on the relative values of A and B (or A^* and B^*) and that the affect of A (or A^*) depends on the value of B (or B^*), with larger values of A (or more positive values of A^*) decreasing the response for smaller values of B (or more negative values of B^*).

Although the mathematical form of the equation is important, it is more important that we interpret what it tells us about how each factor affects the response.

5. (a) Letting a represent Ca and letting b represent Al, the values for b_0 , b_a , b_b , and b_{ab} in coded form are

$$b_0 = \frac{1}{4}(54.29 + 98.44 + 19.18 + 38.53) = 52.61$$

$$b_a = \frac{1}{4}(54.29 + 98.44 - 19.18 - 38.53) = 23.755$$

$$b_b = \frac{1}{4}(54.29 - 98.44 + 19.18 - 38.53) = -15.875$$

$$b_{ab} = \frac{1}{4}(54.29 - 98.44 - 19.18 + 38.53) = -6.20$$

which gives us the following equation for the response surface in coded form

$$R = 52.610 + 23.755Ca^* - 15.875Al^* - 6.20Ca^*Al^*$$

(b) The original data shows that a larger concentration of Al suppresses the signal for Ca; thus, we want to find the maximum concentration of Al that results in a decrease in the response of less than

5%. First, we determine the response for a solution that is 6.00 ppm in Ca and that has no Al. The following equations relate the actual concentrations of each species to its coded form

$$Ca = 7 + 3Ca^* \quad Al = 80 + 80Al^*$$

Substituting in 6.00 ppm for Ca and 0.00 ppm for Al gives $-1/3$ for Ca^* and -1 for Al^* . Substituting these values back into the response surface's coded equation

$$R = 52.610 + 23.755\left(\frac{-1}{3}\right) - 15.875(-1) - 6.20\left(\frac{-1}{3}\right)(-1)$$

gives the response as 58.50. Decreasing this response by 5% leaves us with a response of 55.58. Substituting this response into the response surface's coded equation, along with the coded value of $-1/3$ for Ca^* , and solving for Al^* gives

$$55.58 = 52.610 + 23.755\left(\frac{-1}{3}\right) - 15.875Al^* - 6.20\left(\frac{-1}{3}\right)Al^*$$

$$10.88 = -13.81Al^*$$

$$Al^* = -0.789$$

The maximum allowed concentration of aluminum, therefore, is

$$Al = 80 + 80(-0.789) = 16.9 \text{ ppm Al}$$

6. (a) The values for b_0 , b_x , b_y , b_z , b_{xy} , b_{xz} , b_{yz} , and b_{xyz} in coded form are

$$b_0 = \frac{1}{8} \left(\frac{28 + 17 + 41 + 34 + 56 + 51 + 42 + 36}{8} \right) = 38.125 \approx 38.1$$

$$b_x = \frac{1}{8} \left(\frac{-28 + 17 - 41 + 34 - 56 + 51 - 42 + 36}{8} \right) = -3.625 \approx -3.6$$

$$b_y = \frac{1}{8} \left(\frac{-28 - 17 + 41 + 34 - 56 - 51 + 42 + 36}{8} \right) = 0.125 \approx 0.1$$

$$b_z = \frac{1}{8} \left(\frac{-28 - 17 - 41 - 34 + 56 + 51 + 42 + 36}{8} \right) = 8.125 \approx 8.1$$

$$b_{xy} = \frac{1}{8} \left(\frac{28 - 17 - 41 + 34 + 56 - 51 - 42 + 36}{8} \right) = 0.375 \approx 0.4$$

$$b_{xz} = \frac{1}{8} \left(\frac{28 - 17 + 41 - 34 - 56 + 51 - 42 + 36}{8} \right) = 0.875 \approx 0.9$$

$$b_{yz} = \frac{1}{8} \left(\frac{28 + 17 - 41 - 34 - 56 - 51 + 42 + 36}{8} \right) = -7.375 \approx -7.4$$

$$b_{xyz} = \frac{1}{8} \left(\frac{-28 + 17 + 41 - 34 + 56 - 51 - 42 + 36}{8} \right) = -0.625 \approx -0.6$$

The coded equation for the response surface, therefore, is

$$R = 38.1 - 3.6X^* + 0.1Y^* + 8.1Z^* + 0.4X^*Y^* + 0.9X^*Z^* - 7.4Y^*Z^* - 0.6X^*Y^*Z^*$$

(b) The important effects are the temperature (X^*) and the reactant's concentration (Z^*), and an interaction between the reactant's concentration and the type of catalyst (Y^*Z^*), which leave us with

$$R = 38.1 - 3.6X^* + 8.1Z^* - 7.4Y^*Z^*$$

(c) Because the catalyst is a categorical variable, not a numerical variable, we cannot transform its coded value (Y^*) into a number.

(d) The response surface's simple coded equation shows us that the effect of the catalyst depends on the reactant's concentration as it appears only in the interaction term Y^*Z^* . For smaller concentrations of reactant—when Z^* is less than 0 or the reactant's concentration is less than 0.375 M—catalyst B is the best choice because the term $-7.4Y^*Z^*$ is positive; the opposite is true for larger concentrations of reactant—when Z^* is greater than 0 or the reactant's concentration is greater than 0.375 M—where catalyst A is the best choice.

(e) For the temperature and the concentration of reactant, the following equations relate a coded value to its actual value

$$X = 130 + 10X^* \quad Z = 0.375 + 0.125Z^*$$

Substituting in the desired temperature and concentration, and solving for X^* and for Z^* gives

$$125 = 130 + 10X^* \quad 0.45 = 0.375 + 0.125Z^*$$

$$-5 = 10X^* \quad 0.075 = 0.125Z^*$$

$$X^* = -0.5 \quad Z^* = 0.6$$

Because Z^* is greater than zero, we know that the best catalyst is type A, for which Y^* is -1 . Substituting these values into the response surface's coded equation gives the percent yield as

$$R = 38.1 - 3.6(-0.5) + 8.1(0.6) - 7.4(-1)(0.6) = 49.2\%$$

7. (a) The values for $b_0, b_x, b_y, b_z, b_{xy}, b_{xz}, b_{yz}$, and b_{xyz} in coded form are

$$b_0 = \frac{1}{8} \left(\begin{array}{c} 1.55 + 5.40 + 3.50 + 6.75 + \\ 2.45 + 3.60 + 3.05 + 7.10 \end{array} \right) = 4.175 \approx 4.18$$

$$b_x = \frac{1}{8} \left(\begin{array}{c} -1.55 + 5.40 - 3.50 + 6.75 - \\ 2.45 + 3.60 - 3.05 + 7.10 \end{array} \right) = 1.538 \approx 1.54$$

$$b_y = \frac{1}{8} \left(\begin{array}{c} -1.55 - 5.40 + 3.50 + 6.75 - \\ 2.45 - 3.60 + 3.05 + 7.10 \end{array} \right) = 0.925 \approx 0.92$$

$$b_z = \frac{1}{8} \left(\begin{array}{c} -1.55 - 5.40 - 3.50 - 6.75 + \\ 2.45 + 3.60 + 3.05 + 7.10 \end{array} \right) = -0.125 \approx -0.12$$

$$b_{xy} = \frac{1}{8} \left(\begin{array}{c} 1.55 - 5.40 - 3.50 + 6.75 + \\ 2.45 - 3.60 - 3.05 + 7.10 \end{array} \right) = 0.288 \approx 0.29$$

$$b_{xz} = \frac{1}{8} \left(\begin{array}{c} 1.55 - 5.40 + 3.50 - 6.75 - \\ 2.45 + 3.60 - 3.05 + 7.10 \end{array} \right) = -0.238 \approx -0.24$$

$$b_{yz} = \frac{1}{8} \left(\begin{array}{c} 1.55 + 5.40 - 3.50 - 6.75 - \\ 2.45 - 3.60 + 3.05 + 7.10 \end{array} \right) = 0.100 \approx 0.10$$

$$b_{xyz} = \frac{1}{8} \left(\begin{array}{c} -1.55 + 5.40 + 3.50 - 6.75 + \\ 2.45 - 3.60 - 3.05 + 7.10 \end{array} \right) = 0.438 \approx 0.44$$

The coded equation for the response surface, therefore, is

$$R = 4.18 + 1.54X^* + 0.92Y^* - 0.12Z^* + \\ 0.29X^*Y^* - 0.24X^*Z^* + 0.1Y^*Z^* + 0.44X^*Y^*Z^*$$

(b) The important effects are the presence or absence of benzocaine (X^*) and the temperature (Y^*), which leave us with

$$R = 4.18 + 1.54X^* + 0.92Y^*$$

8. (a) The values for b_0 , b_x , b_y , b_z , b_{xy} , b_{xz} , b_{yz} , and b_{xyz} in coded form are

$$b_0 = \frac{1}{8}(2 + 6 + 4 + 8 + 10 + 18 + 8 + 12) = 8.5$$

$$b_x = \frac{1}{8}(-2 + 6 - 4 + 8 - 10 + 18 - 8 + 12) = 2.5$$

$$b_y = \frac{1}{8}(-2 - 6 + 4 + 8 - 10 - 18 + 8 + 12) = -0.5$$

$$b_z = \frac{1}{8}(-2 - 6 - 4 - 8 + 10 + 18 + 8 + 12) = 3.5$$

$$b_{xy} = \frac{1}{8}(2 - 6 - 4 + 8 + 10 - 18 - 8 + 12) = -0.5$$

$$b_{xz} = \frac{1}{8}(2 - 6 + 4 - 8 - 10 + 18 - 8 + 12) = 0.5$$

$$b_{yz} = \frac{1}{8}(2 + 6 - 4 - 8 - 10 - 18 + 8 + 12) = -1.5$$

$$b_{xyz} = \frac{1}{8}(-2 + 6 + 4 - 8 + 10 - 18 - 8 + 12) = -0.5$$

The coded equation for the response surface, therefore, is

$$R = 8.5 + 2.5X^* - 0.5Y^* + 3.5Z^* - \\ 0.5X^*Y^* + 0.5X^*Z^* - 1.5Y^*Z^* - 0.5X^*Y^*Z^*$$

(b) The important effects are the temperature (X^*), the pressure (Y^*), and the interaction between the pressure and the residence time (Y^*Z^*), which leave us with

$$R = 8.5 + 2.5X^* + 3.5Z^* - 1.5Y^*Z^*$$

(c) The mean response is an 8.6% yield for the three trials at the center of the experimental design, with a standard deviation of 0.529%. A 95% confidence interval for the mean response is

$$\mu = \bar{X} \pm \frac{ts}{\sqrt{n}} = 8.60\% \pm \frac{(4.303)(0.529\%)}{\sqrt{3}} = 8.60\% \pm 1.31\%$$

The average response for the eight trials in the experimental design is given by b_0 and is equal to 8.5; as this falls within the confidence interval, there is no evidence, at $\alpha = 0.05$, of curvature in the data and a first-order model is a reasonable choice.

9. (a) When considering the response in terms of ΔE , the values for b_0 , b_x , b_y , b_z , b_{xy} , b_{xz} , b_{yz} , and b_{xyz} in coded form are

$$b_0 = \frac{1}{8} \begin{pmatrix} 37.45 + 31.70 + 32.10 + 27.20 + \\ 39.85 + 32.85 + 35.00 + 32.15 \end{pmatrix} = 33.54$$

$$b_x = \frac{1}{8} \begin{pmatrix} -37.45 + 31.70 - 32.10 + 27.20 - \\ 39.85 + 32.85 - 35.00 + 32.15 \end{pmatrix} = -2.56$$

$$b_y = \frac{1}{8} \begin{pmatrix} -37.45 - 31.70 + 32.10 + 27.20 - \\ 39.85 - 32.85 + 35.00 + 32.15 \end{pmatrix} = -1.92$$

$$b_z = \frac{1}{8} \begin{pmatrix} -37.45 - 31.70 - 32.10 - 27.20 + \\ 39.85 + 32.85 + 35.00 + 32.15 \end{pmatrix} = 1.42$$

$$b_{xy} = \frac{1}{8} \begin{pmatrix} 37.45 - 31.70 - 32.10 + 27.20 + \\ 39.85 - 32.85 - 35.00 + 32.15 \end{pmatrix} = 0.62$$

$$b_{xz} = \frac{1}{8} \begin{pmatrix} 37.45 - 31.70 + 32.10 - 27.20 - \\ 39.85 + 32.85 - 35.00 + 32.15 \end{pmatrix} = 0.10$$

$$b_{yz} = \frac{1}{8} \begin{pmatrix} 37.45 + 31.70 - 32.10 - 27.20 - \\ 39.85 - 32.85 + 35.00 + 32.15 \end{pmatrix} = 0.54$$

$$b_{xyz} = \frac{1}{8} \begin{pmatrix} -37.45 + 31.70 + 32.10 - 27.20 + \\ 39.85 - 32.85 - 35.00 + 32.15 \end{pmatrix} = 0.41$$

The coded equation for the response surface, therefore, is

$$R = 33.54 - 2.56X^* - 1.92Y^* + 1.42Z^* + \\ 0.62X^*Y^* + 0.10X^*Z^* + 0.54Y^*Z^* + 0.41X^*Y^*Z^*$$

(b) When considering the response in terms of samples per hour, the values for b_0 , b_x , b_y , b_z , b_{xy} , b_{xz} , b_{yz} , and b_{xyz} in coded form are

$$b_0 = \frac{1}{8} \begin{pmatrix} 21.5 + 26.0 + 30.0 + 33.0 + \\ 21.0 + 19.5 + 30.0 + 34.0 \end{pmatrix} = 26.9$$

$$b_x = \frac{1}{8} \begin{pmatrix} -21.5 + 26.0 - 30.0 + 33.0 - \\ 21.0 + 19.5 - 30.0 + 34.0 \end{pmatrix} = 1.2$$

$$b_y = \frac{1}{8} \left(\begin{array}{c} -21.5 - 26.0 + 30.0 + 33.0 - \\ 21.0 - 19.5 + 30.0 + 34.0 \end{array} \right) = 4.9$$

$$b_z = \frac{1}{8} \left(\begin{array}{c} -21.5 - 26.0 - 30.0 - 33.0 + \\ 21.0 + 19.5 + 30.0 + 34.0 \end{array} \right) = -0.8$$

$$b_{xy} = \frac{1}{8} \left(\begin{array}{c} 21.5 - 26.0 - 30.0 + 33.0 + \\ 21.0 - 19.5 - 30.0 + 34.0 \end{array} \right) = 0.5$$

$$b_{xz} = \frac{1}{8} \left(\begin{array}{c} 21.5 - 26.0 + 30.0 - 33.0 - \\ 21.0 + 19.5 - 30.0 + 34.0 \end{array} \right) = -0.6$$

$$b_{yz} = \frac{1}{8} \left(\begin{array}{c} 21.5 + 26.0 - 30.0 - 33.0 - \\ 21.0 - 19.5 + 30.0 + 34.0 \end{array} \right) = 1.0$$

$$b_{xyz} = \frac{1}{8} \left(\begin{array}{c} -21.5 + 26.0 + 30.0 - 33.0 + \\ 21.0 - 19.5 - 30.0 + 34.0 \end{array} \right) = 0.9$$

The coded equation for the response surface, therefore, is

$$R = 26.9 + 1.2X^* + 4.9Y^* - 0.8Z^* + 0.5X^*Y^* - 0.6X^*Z^* + Y^*Z^* + 0.9X^*Y^*Z^*$$

(c) To help us compare the response surfaces, let's gather the values for each term into a table; thus

parameter	ΔE	sample/h
b_0	33.54	26.9
b_x	-2.56	1.2
b_y	-1.92	4.9
b_z	1.42	-0.8
b_{xy}	0.62	0.5
b_{xz}	0.10	-0.6
b_{yz}	0.54	1.0
b_{xyz}	0.41	0.9

Looking at the main effects (b_x , b_y , and b_z), we see from the signs that the parameters that favor a high sampling rate (a smaller volume of sample, a shorter reactor length, and a faster carrier flow rate) result in smaller values for ΔE ; thus, the conditions that favor sensitivity do not favor the sampling rate.

(d) One way to answer this question is to look at the original data and see if for any individual experiment, the sensitivity and the sampling rate both exceed their mean values as given by their respective values for b_0 : 33.54 for ΔE and 26.9 sample/h for the sampling rate. Of the original experiments, this is the case only for run 7; thus, a reactor length of 1.5 cm ($X^* = -1$), a carrier flow rate of 2.2 mL/min

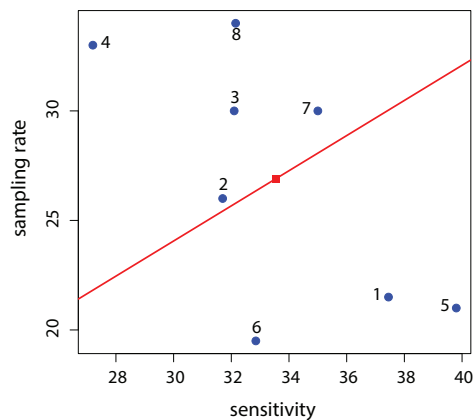


Figure SM14.7 Plot of sampling rate vs. sensitivity for the data in Problem 9. The **blue** dots are the results for the experimental runs used to model the response surface, the **red** square shows the mean sensitivity and mean sampling rate for the experimental data, and the **red** line shows equal percentage changes in sensitivity and sampling rate relative to their respective mean values. See text for further details.

($Y^* = +1$), and a sample volume of 150 μL provides the best compromise between sensitivity and sampling rate.

Another approach is to plot the sampling rate versus the sensitivity for each experimental run, as shown in Figure SM14.7 where the **blue** dots are the results for the eight experiments, the **red** square is the average sensitivity and the average rate, and the **red** line shows conditions that result in an equal percentage change in the sensitivity and the sampling rate relative to their mean values. The best experimental run is the one that lies closest to the **red** line and furthest to the upper-right corner. Again, the seventh experiment provides the best compromise between sampling rate and sensitivity.

10. (a) There are a total of 32 terms to calculate: one average (b_0), five main effects ($b_a, b_b, b_c, b_d,$ and b_e), 10 binary interactions ($b_{ab}, b_{ac}, b_{ad}, b_{ae}, b_{bc}, b_{bd}, b_{be}, b_{cd}, b_{ce},$ and b_{de}), 10 ternary interactions ($b_{abc}, b_{abd}, b_{abe}, b_{acd}, b_{ace}, b_{ade}, b_{bcd}, b_{bce}, b_{bde},$ and b_{cde}), five quaternary interactions ($b_{abcd}, b_{abce}, b_{abde}, b_{acde},$ and b_{bdce}), and one quinary interaction (b_{abcde}). We will not show here the equations for all 32 terms; instead, we provide the equation for one term in each set and summarize the results in a table.

$$b_0 = \frac{1}{32} \sum_{i=1}^{32} R_i \quad b_a = \frac{1}{32} \sum_{i=1}^{32} A_i^* R_i$$

$$b_{ab} = \frac{1}{32} \sum_{i=1}^{32} A_i^* B_i^* R_i \quad b_{abc} = \frac{1}{32} \sum_{i=1}^{32} A_i^* B_i^* C_i^* R_i$$

$$b_{abcd} = \frac{1}{32} \sum_{i=1}^{32} A_i^* B_i^* C_i^* D_i^* R_i \quad b_{abcde} = \frac{1}{32} \sum_{i=1}^{32} A_i^* B_i^* C_i^* D_i^* E_i^* R_i$$

term	value	term	value	term	value
b_0	0.49	b_{bd}	-0.008	b_{bcd}	0.001
b_a	0.050	b_{be}	0.008	b_{bce}	0
b_b	-0.071	b_{cd}	-0.021	b_{bde}	0.006
b_c	0.039	b_{ce}	-0.12	b_{cde}	0.025
b_d	0.074	b_{de}	-0.007	b_{abcd}	0.006
b_e	-0.15	b_{abc}	0.003	b_{abce}	0.007
b_{ab}	0.001	b_{abd}	0.005	b_{abde}	0.004
b_{ac}	-0.007	b_{abe}	-0.004	b_{acde}	0.009
b_{ad}	0.013	b_{acd}	0.003	b_{bdce}	0.005
b_{ae}	0.009	b_{ace}	0.049	b_{abcde}	-0.14
b_{bc}	0.014	b_{ade}	0.019		

If we ignore any term with an absolute value less than 0.03, then the coded equation for the response surface is

$$R = 0.49 + 0.50A^* - 0.071B^* + 0.039C^* \\ + 0.074D^* - 0.15E^* - 0.12C^*E^* + 0.049A^*C^*E^*$$

(b) The coded equation suggests that the most desirable values for A^* and for D^* are positive as they appear only in terms with positive coefficients, and that the most desirable values for B^* are negative as it appears only in a term with a negative coefficient. Because E^* is held at its high, or +1 level, the most desirable value for C^* is negative as this will make $-0.12C^*E^*$ more positive than the term $0.049A^*C^*E^*$ is negative. This is consistent with the results from the simplex optimization as the flow rate (A) of 2278 mL/min is greater than its average factor level of 1421 mL/min (A^*), the amount of SiH_4 used (B) of 9.90 ppm is less than its average factor level of 16.1 ppm (B^*), the $\text{O}_2 + \text{N}_2$ flow rate (C) of 260.6 mL/min is greater its average factor level C^* of 232.5 mL/min, and the O_2/N_2 ratio (D) of 1.71 is greater than its average factor level (D^*) of 1.275.

11. Substituting in values of $X_1 = 10$ and $X_2 = 0$ gives a response of 519.7, or an absorbance of 0.520. Repeating using values of $X_1 = 0$ and $X_2 = 10$ gives a response of 637.5, or an absorbance of 0.638. Finally, letting $X_1 = 0$ and $X_2 = 0$ gives a response of 835.9, or an absorbance of 0.836.

These values are not reasonable as both H_2O_2 and H_2SO_4 are required reagents if the reaction is to develop color. Although the empirical model works well within the limit $8 \leq X_1 \leq 22$ and the limit $8 \leq X_2 \leq 22$, we cannot extend the model outside this range without introducing error.

This is, of course, the inherent danger of extrapolation.

12. The mean and the standard deviation for the 10 trials are 1.355 ppm and 0.1183 ppm, respectively. The relative standard deviation of

$$s_{rel} = \frac{0.1183 \text{ ppm}}{1.355 \text{ ppm}} \times 100 = 8.73\%$$

and the bias of

$$\frac{1.355 \text{ ppm} - 1.30 \text{ ppm}}{1.30 \text{ ppm}} \times 100 = 4.23\%$$

are within the prescribed limits; thus, the single operator characteristics are acceptable.

13. The following calculations show the effect of a change in each factor's level

$$E_A = \frac{98.9 + 98.5 + 97.7 + 97.0}{4} \\ - \frac{98.8 + 98.5 + 97.7 + 97.3}{4} = -0.05$$

$$E_B = \frac{98.9 + 98.5 + 98.8 + 98.5}{4} - \frac{97.7 + 97.0 + 97.7 + 97.3}{4} = 1.25$$

$$E_C = \frac{98.9 + 97.7 + 98.8 + 97.7}{4} - \frac{98.5 + 97.0 + 98.5 + 97.3}{4} = 0.45$$

$$E_D = \frac{98.9 + 98.5 + 97.7 + 97.3}{4} - \frac{97.7 + 97.0 + 98.8 + 98.5}{4} = 0.10$$

$$E_E = \frac{98.9 + 97.7 + 98.5 + 97.3}{4} - \frac{98.5 + 97.0 + 98.8 + 97.7}{4} = 0.10$$

$$E_F = \frac{98.9 + 97.0 + 98.8 + 97.3}{4} - \frac{98.5 + 97.7 + 98.5 + 97.7}{4} = -0.10$$

$$E_G = \frac{98.9 + 97.0 + 98.5 + 97.7}{4} - \frac{98.5 + 97.7 + 98.8 + 97.3}{4} = -0.05$$

The only significant factors are pH (factor B) and the digestion time (factor C). Both have a positive factor effect, which indicates that each factor's high level produces a more favorable recovery. The method's estimated standard deviation is

$$s = \sqrt{\frac{2}{7} \left\{ (-0.05)^2 + (1.25)^2 + (0.45)^2 + (0.10)^2 + (0.10)^2 + (-0.10)^2 + (-0.05)^2 \right\}} = 0.72$$

14. (a) The most accurate analyst is the one whose results are closest to the true mean values, which is indicated by the **red** star; thus, analyst 2 has the most accurate results.
- (b) The most precise analyst is the one whose results are closest to the diagonal line that represents no indeterminate error; thus, analyst 8 has the most precise results.
- (c) The least accurate analyst is the one whose results are furthest from the true mean values, which is indicated by the **red** star; thus, analyst 8 has the most accurate results.
- (d) The least precise analyst is the one whose results are furthest from the diagonal line that represents no indeterminate error; thus, analysts 1 and 10 have the least precise results.

Note that the results for analyst 8 remind us that accuracy and precision are not related, and that it is possible for work to be very precise and yet wholly inaccurate (or very accurate and very imprecise).

15. Figure SM14.8 shows the two sample plot where the mean for the first sample is 1.38 and the mean for the second sample is 1.50. A casual examination of the plot shows that six of the eight points are in the (+,+) or the (-,-) quadrants and that the distribution of the points is more elliptical than spherical; both suggest that systematic errors are present.

To estimate values for σ_{rand} and for σ_{sys} , we first calculate the differences, D_i , and the totals, T_i , for each analyst; thus

analyst	D_i	T_i
1	-0.22	2.92
2	0.02	2.68
3	-0.13	2.81
4	-0.10	3.10
5	-0.10	3.14
6	-0.13	2.91
7	-0.06	2.66
8	-0.21	2.85

To calculate the experimental standard deviations for the differences and the totals, we use equation 14.18 and equation 14.20, respectively, and are easy to calculate if first we find the regular standard deviation and then we divide it by $\sqrt{2}$; thus

$$s_D = 0.1232 \quad s_T = 0.0549$$

To determine if the systematic errors are significant, we use the following null hypothesis and one-tailed alternative hypothesis

$$H_0: s_T = s_D \quad H_A: s_T > s_D$$

Because the value of F_{exp}

$$F_{\text{exp}} = \frac{(s_T)^2}{(s_D)^2} = \frac{(0.1232)^2}{(0.0549)^2} = 5.04$$

exceeds the critical value of $F(0.05, 7, 7)$ of 3.787; thus, we reject the null hypothesis and accept the alternative hypothesis, finding evidence at $\alpha = 0.5$ that systematic errors are present in the data. The estimated precision for a single analyst is

$$\sigma_{\text{rand}} = s_D = 0.055$$

and the estimated standard deviation due to systematic differences between the analysts is

$$\sigma_{\text{sys}} = \sqrt{\frac{\sigma_T^2 - \sigma_D^2}{2}} = \sqrt{\frac{(0.1232)^2 - (0.0549)^2}{2}} = 0.078$$

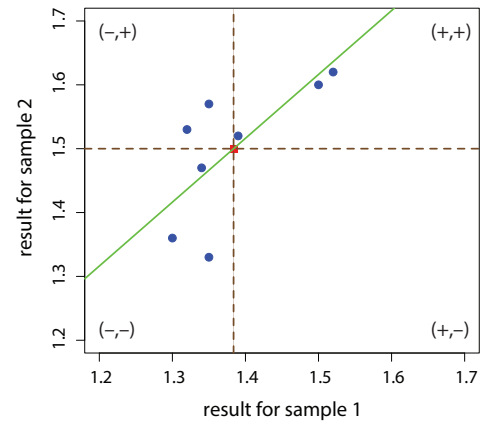


Figure SM14.8 Two-sample plot for the data in Problem 15. The blue dots are the results for each analyst, the red square is the average results for the two samples, the dashed brown lines divide the plot into four quadrants where the results for both samples exceeds the mean (+,+), where both samples are below the mean (-,-), and where one sample is above the mean and one below the mean, (+,-) and (-,+). The solid green line shows results with identical systematic errors.

Here we use a one-tailed alternative hypothesis because we are interested only in whether s_T is significantly greater than s_D .

16. (a) For an analysis of variance, we begin by calculating the global mean and the global variance for all 35 measurements using equation 14.22 and equation 14.23, respectively, obtaining values of $\bar{X} = 3.542$ and $\bar{s}^2 = 1.989$. Next, we calculate the mean value for each of the seven labs, obtaining results of

$$\bar{X}_A = 2.40 \quad \bar{X}_B = 3.60 \quad \bar{X}_C = 2.00$$

$$\bar{X}_D = 2.60 \quad \bar{X}_E = 4.80 \quad \bar{X}_F = 5.00 \quad \bar{X}_G = 4.40$$

To calculate the variance within the labs and the variance between the labs, we use the equations from Table 14.7; thus, the total sum-of-squares is

$$SS_t = \bar{s}^2(N - 1) = (1.989)(35 - 1) = 67.626$$

and the between lab sum-of-squares is

$$\begin{aligned} SS_b &= \sum_{i=1}^b n_i(X_i - \bar{X})^2 = (5)(2.40 - 3.542)^2 \\ &+ (5)(3.60 - 3.542)^2 + (5)(2.00 - 3.542)^2 \\ &+ (5)(2.60 - 3.542)^2 + (5)(4.80 - 3.542)^2 \\ &+ (5)(5.00 - 3.542)^2 + (5)(4.40 - 3.542)^2 = 45.086 \end{aligned}$$

and the within lab sum-of-squares is

$$SS_w = SS_t - SS_b = 67.626 - 45.086 = 22.540$$

The between lab variance, s_b^2 , and the within lab variance, s_w^2 , are

$$\begin{aligned} s_b^2 &= \frac{SS_b}{b - 1} = \frac{45.086}{7 - 1} = 7.514 \\ s_w^2 &= \frac{SS_w}{N - b} = \frac{22.540}{35 - 7} = 0.805 \end{aligned}$$

To determine if there is evidence that the differences between the labs is significant, we use an F -test of the following null hypothesis and one-tailed alternative hypothesis

$$H_0: s_b^2 = s_w^2 \quad H_A: s_b^2 > s_w^2$$

Because the value of F_{exp}

$$F_{\text{exp}} = \frac{s_b^2}{s_w^2} = \frac{(7.514)^2}{(0.805)^2} = 87.13$$

exceeds the critical value for $F(0.05, 6, 28)$, which is between 2.099 and 2.599, we reject the null hypothesis and accept the alternative hypothesis, finding evidence at $\alpha = 0.5$ that there are systematic differences between the results of the seven labs.

To evaluate the source(s) of this systematic difference, we use equation 14.27 to calculate t_{exp} for the difference between mean values, comparing t_{exp} to a critical value of 1.705 for a one-tailed t -test with

Here we use a one-tailed alternative hypothesis because we are interested only in whether s_b is significantly greater than s_w .

Here we use a one-tailed alternative hypothesis because we are interested only in whether the result for one lab is greater than the result for another lab.

28 degrees of freedom. For example, when comparing lab A to lab C, the two labs with the smallest mean values, we find

$$t_{\text{exp}} = \frac{|\bar{X}_A - \bar{X}_C|}{\sqrt{s_w^2}} \times \sqrt{\frac{n_A n_C}{n_A + n_C}} = \frac{|2.40 - 2.00|}{\sqrt{0.805}} \times \sqrt{\frac{5 \times 5}{5 + 5}} = 0.705$$

no evidence for a systematic difference at $\alpha = 0.05$ between lab A and lab C. The table below summarizes results for all seven labs

lab	C	A	D	B	G	E	F
\bar{X}	2.00	2.40	2.60	3.60	4.40	4.80	5.00
t_{exp}	0.705		1.762		0.705		
		0.352		1.410		0.352	

Note that the labs are organized from the lab with the smallest mean value (lab C) to the lab with the largest mean value (lab F) and that we compare mean values for adjacent labs only.

where there is no evidence of a significant difference between the results for labs C, A, and D (as shown by the **green** bar), where there is no evidence of a significant difference between the results for labs G, E, and F (as shown by the **blue** bar), and where there is no significant difference between the results for labs B and G (as shown by the **red** bar).

(b) The estimated values for σ_{rand}^2 and for σ_{sys}^2 are

$$\sigma_{\text{rand}}^2 \approx s_w^2 = 0.805$$

$$\sigma_{\text{sys}}^2 = \frac{s_b^2 - \sigma_{\text{rand}}^2}{n} = \frac{7.514 - 0.805}{5} = 1.34$$

17. First, let's write out the three sum-of-squares terms that appear in equation 14.23 (SS_p), equation 14.24 (SS_w), and equation 14.25 (SS_b)

$$SS_t = \sum_{i=1}^b \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$$

$$SS_w = \sum_{i=1}^b \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

$$SS_b = \sum_{i=1}^b n_i (\bar{X}_i - \bar{X})^2$$

so that we have them in front of us. Looking at the equation for SS_p , let's pull out the term within the parentheses,

$$X_{ij} - \bar{X}$$

and then subtract and add the term \bar{X}_i to it, grouping together parts of the equation using parentheses

$$(X_{ij} - \bar{X}) = (X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X})$$

Next, let's square both sides of the equation

$$(X_{ij} - \bar{X})^2 = \{(X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X})\}^2$$

$$(X_{ij} - \bar{X})^2 = (X_{ij} - \bar{X}_i)^2 + (\bar{X}_i - \bar{X})^2 + 2(X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X})$$

and then substitute the right side of this equation back into the summation term for SS_t

$$SS_t = \sum_{i=1}^h \sum_{j=1}^{n_i} \{(X_{ij} - \bar{X}_i)^2 + (\bar{X}_i - \bar{X})^2 + 2(X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X})\}$$

and expand the summation across the terms in the curly parentheses

$$\begin{aligned} SS_t &= \sum_{i=1}^h \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \\ &\quad + \sum_{i=1}^h \sum_{j=1}^{n_i} (\bar{X}_i - \bar{X})^2 \\ &\quad + 2 \sum_{i=1}^h \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) \end{aligned}$$

Note that this is not the case for the first two terms in this expanded equation for SS_t because these terms sum up the squares of the differences, which always are positive, not the differences themselves, which are both positive and negative.

The last of these terms is equal to zero because this always is the result when you sum up the difference between a mean and the values that give the mean; thus, we now have this simpler equation

$$SS_t = \sum_{i=1}^h \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 + \sum_{i=1}^h \sum_{j=1}^{n_i} (\bar{X}_i - \bar{X})^2$$

Finally, we note that

$$\sum_{i=1}^h \sum_{j=1}^{n_i} (\bar{X}_i - \bar{X})^2 = \sum_{i=1}^h n_i (\bar{X}_i - \bar{X})^2$$

because, for each of the h samples, the inner summation term simply adds together the term $(\bar{X}_i - \bar{X})^2$ a total of n_i times. Substituting this back into our equation for SS_t gives

$$SS_t = \sum_{i=1}^h \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 + \sum_{i=1}^h n_i (\bar{X}_i - \bar{X})^2$$

which is equivalent to $SS_t = SS_w + SS_b$.

18. (a) Using equation 14.28, our estimate for the relative standard deviation is

$$R = 2^{(1-0.5 \log C)} = 2^{(1-0.5 \log(0.0026))} = 4.9\%$$

- (b) The mean and the standard deviation for the data set are 0.257%w/w and 0.0164%w/w respectively. The experimental percent relative standard deviation, therefore, is

$$s_r = \frac{0.0164\%w/w}{0.257\%w/w} \times 100 = 6.4\%$$

Because this value is within the range of $0.5 \times$ to $2.0 \times$ of R , the variability in the individual results is reasonable.