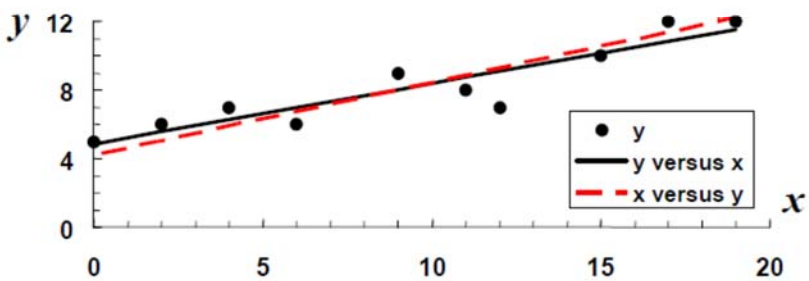
Problem Set #5 Solutions

17.3 The results can be summarized as

	y versus x	x versus y
Best fit equation	$y = 4.851535 + 0.35247x$	$x = -9.96763 + 2.374101y$
Standard error	1.06501	2.764026
Correlation coefficient	0.914767	0.914767

We can also plot both lines on the same graph



Thus, the "best" fit lines and the standard errors differ. This makes sense because different errors are being minimized depending on our choice of the dependent (ordinate) and independent (abscissa) variables. In contrast, the correlation coefficients are identical since the same amount of uncertainty is explained regardless of how the points are plotted.

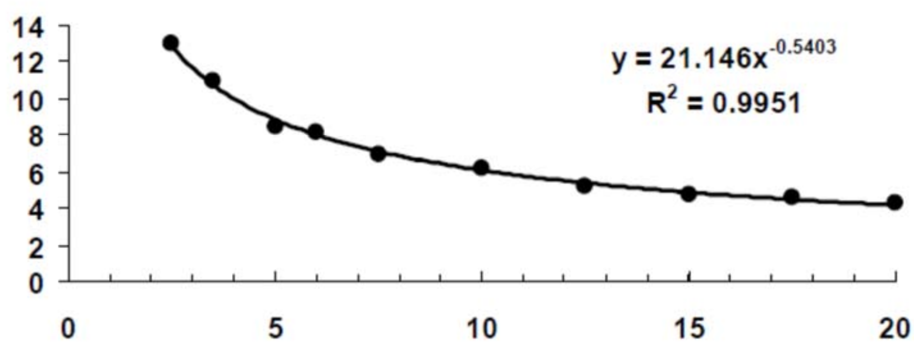
17.8 We regress $\log_{10}(y)$ versus $\log_{10}(x)$ to give

$$\log_{10} y = 1.325225 - 0.54029 \log_{10} x$$

Therefore, $\alpha_2 = 10^{1.325225} = 21.14583$ and $\beta_2 = -0.54029$, and the power model is

$$y = 21.14583x^{-0.54029}$$

The model and the data can be plotted as



The model can be used to predict a value of $21.14583(9)^{-0.54029} = 6.451453$.

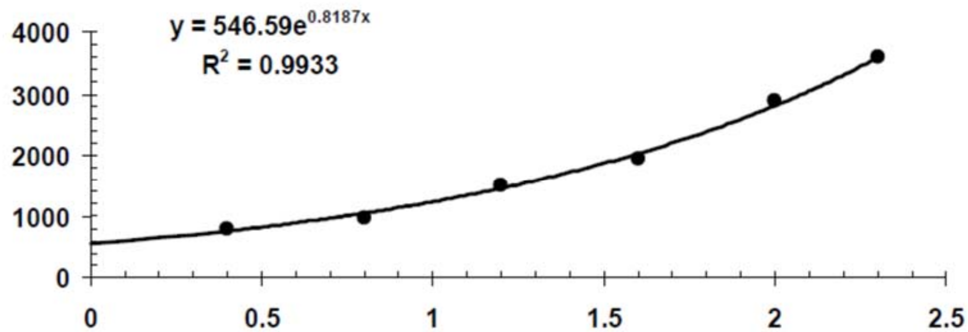
17.9 We regress ln(y) versus x to give

$$\ln y = 6.303701 + 0.818651x$$

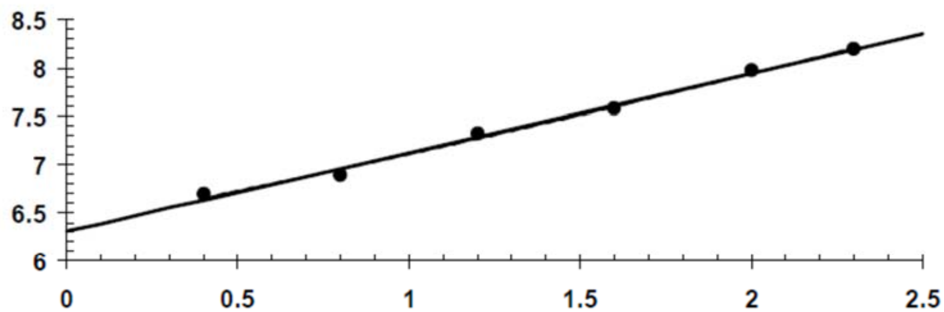
Therefore, $\alpha_1 = e^{6.303701} = 546.5909$ and $\beta_1 = 0.818651$, and the exponential model is

$$y = 546.5909e^{0.818651x}$$

The model and the data can be plotted as



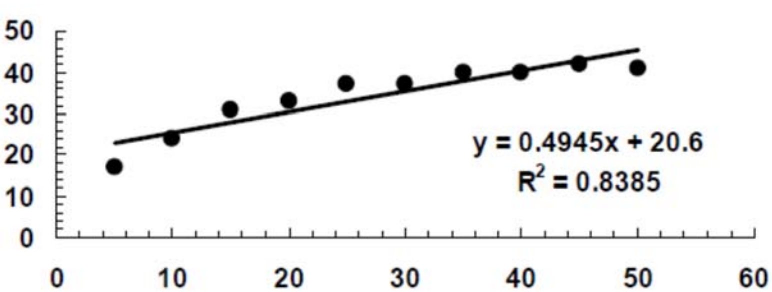
A semi-log plot can be developed by plotting the natural log versus x. As expected, both the data and the best-fit line are linear when plotted in this way.



17.16 (a) We regress y versus x to give

$$y = 20.6 + 0.494545x$$

The model and the data can be plotted as



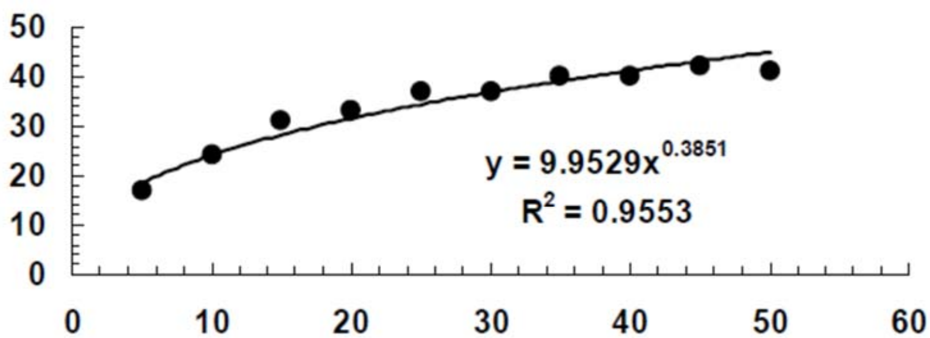
(b) We regress $\log_{10}y$ versus $\log_{10}x$ to give

$$\log_{10} y = 0.99795 + 0.385077 \log_{10} x$$

Therefore, $\alpha_2 = 10^{0.99795} = 9.952915$ and $\beta_2 = 0.385077$, and the power model is

$$y = 9.952915x^{0.385077}$$

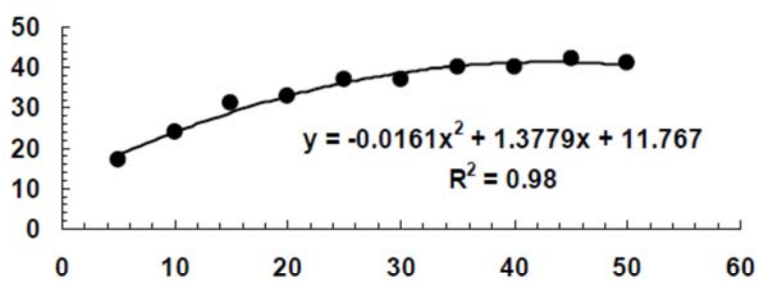
The model and the data can be plotted as



(d) We employ polynomial regression to fit a parabola

$$y = -0.01606x^2 + 1.377879x + 11.76667$$

The model and the data can be plotted as



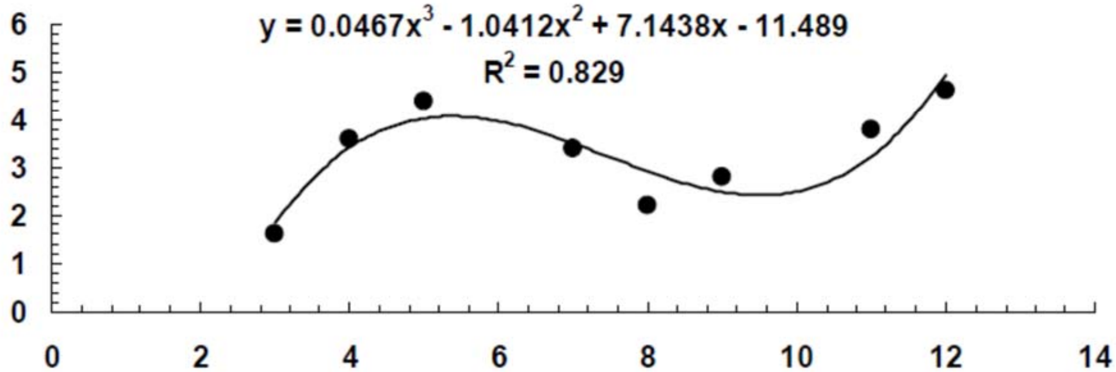
<u>Comparison of fits:</u> The linear fit is obviously inadequate. Although the power fit follows the general trend of the data, it is also inadequate because (1) the residuals do not appear to be randomly distributed around the best fit line and (2) it has a lower r^2 than the saturation and parabolic models.

The best fits are for the saturation-growth-rate and the parabolic models. They both have randomly distributed residuals and they have similar high coefficients of determination. The saturation model has a slightly higher r^2. Although the difference is probably not statistically significant, in the absence of additional information, we can conclude that the saturation model represents the best fit.

17.17 We employ polynomial regression to fit a cubic equation to the data

$$y = 0.046676x^3 - 1.04121x^2 + 7.143817x - 11.4887 \qquad (r^2 = 0.828981; s_{y/x} = 0.570031)$$

The model and the data can be plotted as



18.1 (a)

$$f_1(10) = 0.90309 + \frac{1.0791812 - 0.90309}{12 - 8}(10 - 8) = 0.991136$$

$$\varepsilon_t = \frac{1 - 0.991136}{1} \times 100\% = 0.886\%$$

(b)

$$f_1(10) = 0.9542425 + \frac{1.0413927 - 0.9542425}{11 - 9}(10 - 9) = 0.997818$$

$$\varepsilon_t = \frac{1 - 0.997818}{1} \times 100\% = 0.218\%$$

18.2 First, order the points

$x_0 = 9 \quad f(x_0) = 0.9542425$
$x_1 = 11 \quad f(x_1) = 1.0413927$
$x_2 = 8 \quad f(x_2) = 0.9030900$

Applying Eq. (18.4)

$b_0 = 0.9542425$

Equation (18.5) yields

$$b_1 = \frac{1.0413927 - 0.9542425}{11 - 9} = 0.0435751$$

Equation (18.6) gives

$$b_2 = \frac{\dfrac{0.9030900 - 1.0413927}{8 - 11} - 0.0435751}{8 - 9} = \frac{0.0461009 - 0.0435751}{8 - 9} = -0.0025258$$

Substituting these values into Eq. (18.3) yields the quadratic formula

$$f_2(x) = 0.9542425 + 0.0435751(x - 9) - 0.0025258(x - 9)(x - 11)$$

which can be evaluated at $x = 10$ for

$$f_2(10) = 0.9542425 + 0.0435751(10 - 9) - 0.0025258(10 - 9)(10 - 11) = 1.0003434$$

18.3 First, order the points

$x_0 = 9 \quad f(x_0) = 0.9542425$
$x_1 = 11 \quad f(x_1) = 1.0413927$
$x_2 = 8 \quad f(x_2) = 0.9030900$
$x_3 = 12 \quad f(x_3) = 1.0791812$

The first divided differences can be computed as

$$f[x_1, x_0] = \frac{1.0413927 - 0.9542425}{11 - 9} = 0.0435751$$
$$f[x_2, x_1] = \frac{0.9030900 - 1.0413927}{8 - 11} = 0.0461009$$
$$f[x_3, x_2] = \frac{1.0791812 - 0.9030900}{12 - 8} = 0.0440228$$

The second divided differences are

$$f[x_2, x_1, x_0] = \frac{0.0461009 - 0.0435751}{8 - 9} = -0.0025258$$
$$f[x_3, x_2, x_1] = \frac{0.0440228 - 0.0461009}{12 - 11} = -0.0020781$$

The third divided difference is

$$f[x_3, x_2, x_1, x_0] = \frac{-0.0020781 - (-0.0025258)}{12 - 9} = 0.00014924$$

Substituting the appropriate values into Eq. (18.7) gives

$$f_3(x) = 0.9542425 + 0.0435751(x-9) - 0.0025258(x-9)(x-11) \\ + 0.00014924(x-9)(x-11)(x-8)$$

which can be evaluated at $x = 10$ for

$$f_3(x) = 0.9542425 + 0.0435751(10-9) - 0.0025258(10-9)(10-11) \\ + 0.00014924(10-9)(10-11)(10-8) = 1.0000449$$

18.5 First, order the points so that they are as close to and as centered about the unknown as possible

$x_0 = 2.5$ $f(x_0) = 14$
$x_1 = 3.2$ $f(x_1) = 15$
$x_2 = 2$ $f(x_2) = 8$
$x_3 = 4$ $f(x_3) = 8$
$x_4 = 1.6$ $f(x_4) = 2$

Next, the divided differences can be computed and displayed in the format of Fig. 18.5,

i	x_i	$f(x_i)$	$f[x_{i+1},x_i]$	$f[x_{i+2},x_{i+1},x_i]$	$f[x_{i+3},x_{i+2},x_{i+1},x_i]$	$f[x_{i+4},x_{i+3},x_{i+2},x_{i+1},x_i]$
0	2.5	14	1.428571	-8.809524	1.011905	1.847718
1	3.2	15	5.833333	-7.291667	-0.651042	
2	2	8	0	-6.25		
3	4	8	2.5			
4	1.6	2				

The first through third-order interpolations can then be implemented as

$$f_1(2.8) = 14 + 1.428571(2.8-2.5) = 14.428571$$
$$f_2(2.8) = 14 + 1.428571(2.8-2.5) - 8.809524(2.8-2.5)(2.8-3.2) = 15.485714$$
$$f_3(2.8) = 14 + 1.428571(2.8-2.5) - 8.809524(2.8-2.5)(2.8-3.2) \\ + 1.011905(2.8-2.5)(2.8-3.2)(2.8-2.) = 15.388571$$

18.6 First, order the points so that they are as close to and as centered about the unknown as possible

$x_0 = 3$ $f(x_0) = 19$
$x_1 = 5$ $f(x_1) = 99$
$x_2 = 2$ $f(x_2) = 6$
$x_3 = 7$ $f(x_3) = 291$
$x_4 = 1$ $f(x_4) = 3$

Next, the divided differences can be computed and displayed in the format of Fig. 18.5,

i	x_i	$f(x_i)$	$f[x_{i+1},x_i]$	$f[x_{i+2},x_{i+1},x_i]$	$f[x_{i+3},x_{i+2},x_{i+1},x_i]$	$f[x_{i+4},x_{i+3},x_{i+2},x_{i+1},x_i]$
0	3	19	40	9	1	0
1	5	99	31	13	1	
2	2	6	57	9		
3	7	291	48			
4	1	3				

The first through fourth-order interpolations can then be implemented as

$$f_1(4) = 19 + 40(4-3) = 59$$
$$f_2(4) = 59 + 9(4-3)(4-5) = 50$$
$$f_3(4) = 50 \ + 1(4-3)(4-5)(4-2) = 48$$
$$f_4(4) = 48 + 0(4-3)(4-5)(4-2)(4-7) = 48$$

Clearly this data was generated with a cubic polynomial since the difference between the 4th and the 3rd-order versions is zero.

18.13 For the present problem, we have five data points and $n = 4$ intervals. Therefore, $3(4) = 12$ unknowns must be determined. Equations 18.29 and 18.30 yield $2(4) - 2 = 6$ conditions

$$4a_1 + 2b_1 + c_1 = 8$$
$$4a_2 + 2b_2 + c_2 = 8$$
$$6.25a_2 + 2.5b_2 + c_2 = 14$$
$$6.25a_3 + 2.5b_3 + c_3 = 14$$
$$10.24a_3 + 3.2b_3 + c_3 = 15$$
$$10.24a_4 + 3.2b_4 + c_4 = 15$$

Passing the first and last functions through the initial and final values adds 2 more

$$2.56a_1 + 1.6b_1 + c_1 = 2$$
$$16a_4 + 4b_4 + c_4 = 8$$

Continuity of derivatives creates an additional $4 - 1 = 3$.

$$4a_1 + b_1 = 4a_2 + b_2$$
$$5a_2 + b_2 = 5a_3 + b_3$$
$$6.4a_3 + b_3 = 6.4a_4 + b_4$$

Finally, Eq. 18.34 specifies that $a_1 = 0$. Thus, the problem reduces to solving 11 simultaneous equations for 11 unknown coefficients,

$$\begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6.25 & 2.5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6.25 & 2.5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 10.24 & 3.2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10.24 & 3.2 & 1 \\ 1.6 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 16 & 4 & 1 \\ 1 & 0 & -4 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 1 & 0 & -5 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6.4 & 1 & 0 & -6.4 & -1 & 0 \end{bmatrix} \begin{Bmatrix} b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \\ a_3 \\ b_3 \\ c_3 \\ a_4 \\ b_4 \\ c_4 \end{Bmatrix} = \begin{Bmatrix} 8 \\ 8 \\ 14 \\ 14 \\ 15 \\ 15 \\ 2 \\ 8 \\ 0 \\ 0 \\ 0 \end{Bmatrix}$$

which can be solved for

$$
\begin{array}{lll}
b_1 = 15 & c_1 = -22 & \\
a_2 = -6 & b_2 = 39 & c_2 = -46 \\
a_3 = -10.816327 & b_3 = 63.081633 & c_3 = -76.102041 \\
a_4 = -3.258929 & b_4 = 14.714286 & c_4 = 1.285714
\end{array}
$$

The predictions can be made as

$$f(3.4) = -3.258929(3.4)^2 + 14.714286(3.4) + 1.285714 = 13.64107$$

$$f(2.2) = -6(2.2)^2 + 39(2.2) - 46 = 10.76$$

Finally, here is a plot of the data along with the quadratic spline,