

Mathematical

The screenshot shows a spreadsheet with the following content:

- Newton's Law of Cooling:**
 - Equation: $\int_{T_0}^T \frac{dT}{T - T_s} = -K \int_0^t dt$
 - Equation: $[\ln(T - T_s)]_{T_0}^T = -K(t - 0) = -Kt$
 - Equation: $\ln(T - T_s) - \ln(T_0 - T_s) = \ln\left(\frac{T - T_s}{T_0 - T_s}\right) = -Kt$
- Hooke's Law:**
 - Equation: $[\text{stress}] = \mu[\text{strain}]$
 - Equation: $\frac{F}{A} = \mu \frac{dl}{l}$
 - Equation: $F = m\ddot{x} = -kx$

Visual elements include a blue square with white snowflakes on the left and a vertical spring on the right.

Modeling in Science

Using spreadsheets to create mathematical models and address scientific inquiry

by **Robert M. Horton and William H. Leonard**

The national science and mathematics standards both emphasize process as well as content. These processes engage students in problem solving, communicating, reasoning logically, representing ideas in various ways,

and connecting ideas within and between disciplines. A fundamental part of these processes is inquiry, in which students address a question by gathering data, formulating and testing conjectures, and making inferences.

In science, inquiry is used as students explore important and interesting questions concerning the world around them. In mathematics, one contemporary inquiry approach is to create models that describe real phenomena. Creating mathematical models using spreadsheets can help students learn at deep levels in both science and mathematics, and give teachers an excellent opportunity to integrate these two disciplines.

In many schools, students in math classes create regression models using graphing calculators. The use of spreadsheets rather than calculators, however, has several advantages. First, users are not limited to the models built into the calculators; instead they can build a model of any type. Second, with spreadsheets, students can estimate and modify the parameters for a model. Consequently, students will better understand the model and be able to tie each parameter to its real-world meaning. Third, spreadsheet regression can facilitate a deeper understanding of statistical error in mathematical models. In fact, spreadsheet regression allows students to define total error in any way they choose (e.g., squared errors or absolute errors). Yet another advantage spreadsheets have over calculators is the quick availability of a hardcopy reference.

The project described in this article involves a melding of mathematical and scientific processes. The two scientific concepts explored involve physics, specifically Hooke's law and Newton's laws of heating and cooling, though the modeling described can be used in many places in the science curriculum. The mathematics, which is appropriate for students at the introductory algebra level or higher, includes regression analysis, the parameters in a model, and the errors that inevitably occur.

Hooke's law

Consider a spring and its total length as it is stretched by various masses. Ask students if they expect there to be a relationship between the mass and the length. If so, what type and how strong of a relationship do students anticipate? After a discussion that includes predictions and hypotheses, students are ready to collect data. The sample data we gathered are shown in Figure 1 (p. 42); however, teachers should encourage students to gather their own data. We suggest that students work in groups of two. All they need is a spring, a ring stand and ring clamp, a meterstick, and various weights.

The regression discussed here is done in Microsoft Excel but other spreadsheet programs can provide similar results. (For teachers new to Excel, see the sidebar "Excel tips.")

This first demonstration is done together as a class, with the teacher encouraging discussion and help from students before entering information into Excel. Demonstrating for the class, enter mass (g) in column A and the total spring length (cm) in column B. Mea-

suring the distance the spring is stretched is an option, but for the math classes using the total length develops a meaningful y -intercept. After all data are entered, ask students to comment on any observed patterns. In some contexts, students may need to sort the data first. Here, however, assuming mass is continually added,

Excel tips.

Cell identification: Cells are identified by their column and row labels. For example, G2 refers to the value in column G, row 2.

Formulas: Formulas begin with an equal sign. This tells Excel to compute a numerical value depending on the formula. For example, the formula $=A3+B3$ computes the sum of the values stored in cells A3 and B3. If the entered values in cells A3 and B3 are changed, then the sum will be changed automatically. This holds true for any formulas, many of which are built into Excel. By selecting INSERT and then FUNCTION, you will find many formulas that you can use and which will be updated automatically as the selected cells are changed.

Absolute reference: An absolute cell reference does not change as a formula is copied into other cells. To tell Excel that a cell reference should not change, use the \$ sign to fix the column and the row. For example, in our Hooke's law problem shown in Figure 2 (p. 42), we have our estimate for slope in cell G2, and we want the slope always to come from this cell—to establish this cell, in our formulas we have in $\$G\2 .

Relative reference: A relative reference changes as a formula is copied into other cells. For our Hooke's law example, we want similar formulas in our PREDICTED column, but we want the x values (the times) to change as we move down from row to row. Using the cell identification alone tells Excel that we want a relative reference. Note in Figure 2 that in column C the references to the time values change from cell A2 to cell A3 and so on (though we want absolute references to cells G2 and G4).

Fill down: Rather than copy and paste, we can quickly fill formulas down (or across) a set of cells. To do so, we click in the cell that has the formula to be filled into other cells. In Figure 2, after creating the initial formula in cell C2, we click in that cell. We can then drag the mouse down through the cells into which we want to copy the formula. In this case, we drag through cell C7. Then, select "fill down" from the menu (or, use the keyboard shortcut: on a PC, we simply hold the Ctrl key and press D). An even simpler way is to click in the cell with the formula, grab the "fill handle," which is the lower right corner of the cell, and drag the mouse down into the cells in which we want the formula to appear.

FIGURE 1 Sample data for stretching a spring.

Mass (g)	Length (cm)
0	4
15	10
30	14
45	21
60	27
75	34

the data are already sorted, so students should recognize the positive association.

Next, create a scatterplot, asking students again to discuss the type of relationship they see and compare it with what they anticipated. As you create the graph, students should take notes of the procedures so they will be able to create their own scatterplots and models later. The relationship appears linear; ask students whether the correlation is positive or negative, strong or weak, and why.

Now begin to build a linear model in the form $y = mx + b$, with x representing mass and y representing the length of the spring. If you ask students to estimate the slope m , they may initially struggle. Some may decide to pick two representative points to determine the slope, a technique from introductory algebra. More importantly, through teacher questioning, students should come to understand that the slope on average represents the increase in spring length for each additional gram.

Students should know that the y -intercept b represents the length of the spring before any weights are attached, however if they have studied math out of context, this may not be readily apparent. For these

students, working toward this understanding may provide an “aha” moment that will help them finally grasp an important concept.

On the spreadsheet, put the label “ m ” in cell F2 and the label “ b ” in cell F4. Then, put a student’s estimates (or guesses, if necessary) for these values in the adjacent cells, G2 and G4, respectively. Type “PREDICTED” in cell C1 to label the column and ask students how to compute the value for cell C2. Class discussion should lead all students to determine that the slope (cell G2) must be multiplied by the mass (cell A2) and then the y -intercept (cell G4) added.

When we tried this in our classroom, we found that distinguishing between absolute and relative cell references is not difficult for students, so we were able to arrive at the formula $=G2*A2+G4$ with relative ease. This computes the predicted value for the spring length in cell A2. Some students may not know how to fill the formula down the column, so we demonstrate this. Next we show students how the formula appears in each cell, noting which values in the formula stay the same (the parameters) and which values change (the variables). Again, we found that this presents minimal difficulty.

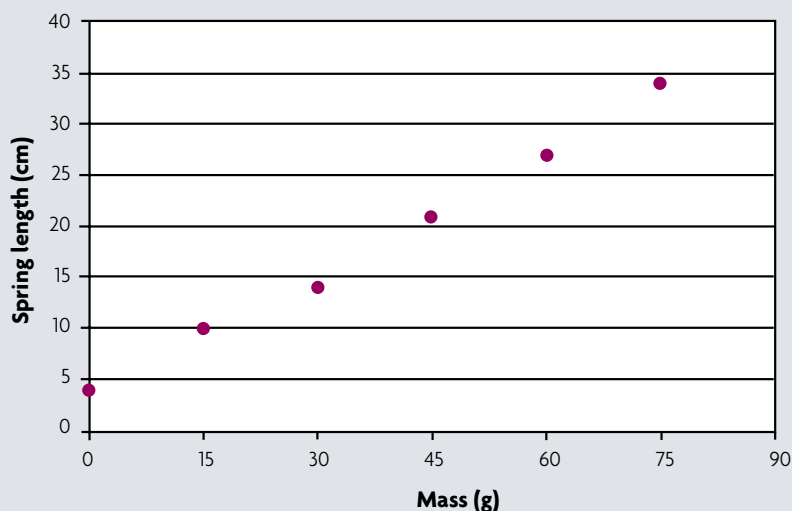
Now ask students to compare the differences between the predicted and actual spring lengths. Assuming each group has access to a computer, students should experiment with modifying their estimates for the slope and y -intercept, trying to derive a better fitting model. As they do so, students should appreciate the need to quantify the difference (the error) between the actual and predicted values. Ask students, “How else can we tell which mathematical model best fits our physical data? What do we really mean by best fit?”

Type “ERROR” in cell D1 as a column label. After a brief discussion of what positive and negative errors

FIGURE 2 Spreadsheet with equations.

	A	B	C	D	E	F	G
1	MASS (g)	LENGTH (cm)	PREDICTED	ERROR	SQ ERROR	MODEL:	$y=mx+b$
2			$=G2*A2+G4$	$=B2-C2$	$=D2^2$	m	(est for m)
3			$=G2*A3+G4$	$=B3-C3$	$=D3^2$		
4			$=G2*A4+G4$	$=B4-C4$	$=D4^2$	b	(est for b)
5			$=G2*A5+G4$	$=B5-C5$	$=D5^2$		
6			$=G2*A6+G4$	$=B6-C6$	$=D6^2$		
7			$=G2*A7+G4$	$=B7-C7$	$=D7^2$		
8							
9			SUM	$=SUM(D2:D7)$	$=SUM(E2:E7)$		

FIGURE 3
Spreadsheet models.



Estimated model						
MASS	LENGTH	PREDICTED	ERROR	SQ ERROR	MODEL: $y=mx+b$	
(g)	(cm)					
0.0	4.0	4.0	0.0	0.0	m	0.40
15.0	10.0	10.0	0.0	0.0		
30.0	14.0	16.0	-2.0	4.0	b	4.0
45.0	21.0	22.0	-1.0	1.0		
60.0	27.0	28.0	-1.0	1.0		
75.0	34.0	34.0	0.0	0.0		
		SUM	-4.0	6.0		

Best-fit model						
MASS	LENGTH	PREDICTED	ERROR	SQ ERROR	MODEL: $y=mx+b$	
(g)	(cm)					
0.0	4.0	3.47	0.523	0.274	m	0.396
15.0	10.0	9.41	0.580	0.337		
30.0	14.0	15.3	-1.36	1.85	b	3.47
45.0	21.0	21.3	-0.304	0.0928		
60.0	27.0	27.2	-0.247	0.0613		
75.0	34.0	33.1	0.809	0.655		
		SUM	0.0000124	3.27		

represent, students should understand that to compute the error for row 2, they need the formula =B2-C2, since the error is just the difference between the predicted (C2) and actual (B2) length. After filling this formula down, compute a sum of the errors by labeling cell C9 "SUM" and typing the formula =SUM(D2:D7) into cell D9. By changing values for m and b , students can note how the sum of error changes. However, they cannot readily see that this sum can be misleading because of the cancellation of positive and negative errors. To help students see this, compute the mean spring length and note how the sum of the deviations from the mean is necessarily 0, because the positive and negative deviations cancel each other out (not shown). In other words, if the mean spring length is used to predict the length regardless of the value of the attached mass, the total error is 0, but nothing is learned about the relationship between the mass and the spring length.

One way to eliminate the cancellation of positive and negative errors is by taking their absolute values. A second technique is to square the errors, which is what is typically done and what is shown here. Label column E "SQ ERROR" and then, in cell E2, type the formula =D2^2. Fill the formula down and sum these squared errors in cell E9 with the formula =SUM(E2:E7). As before, the goal is to minimize the total error. Again adjust the estimates for m and b , this time trying to minimize the sum of the squared errors. With each adjustment in m and b estimates, the spreadsheet will instantly recalculate the predicted lengths and the absolute and squared errors.

Finally, the class is ready to use the spreadsheet's Solver to create a best-fit model. (Note: On Mac versions of Excel, you must install the "Value Pack" to make the Solver available. On PCs, you may need to go to "Add-in" under the Tools menu to activate the Solver.) Highlight the sum of the squared errors in cell E9. Under "Tools," select "Solver." In the "Equal To" field, select "min", and under "By Changing Cells" select G2 and G4 by clicking cell G2, adding a comma, and then clicking cell G4. Click on "Solve." You can either replace the parameters or keep the estimates. Students should compare this new model with the one they came up with when they

tried to minimize the total by changing values for m and b on the spreadsheet, noting the smaller sum of squared errors. Figure 2 shows the shell of the spreadsheet, and Figure 3 (p. 43) shows our estimated and best-fit models. After students understand linear models, they should be ready for a more complex problem, this time completing the model themselves.

Newton's law of cooling

For this second part of the project, students are given a lot more responsibility, but you should be available to facilitate the process and answer questions. Ask the class what they think happens to the temperature of a heated liquid that is set out in a room. What kind of model might students expect when they explore the relationship between time and temperature? For our problem, we recorded the temperature of a warm liquid every 5 minutes for a period of 70 minutes (Figure 4).

A scatterplot reveals a curve, but the desired mathematical model is not something students have studied. An exponential model will not work because a decreasing exponential function approaches 0. Nevertheless, students can hypothesize that the temperature decreases a certain percentage each minute, so the time variable should be in the exponent. Our model takes the form $y = a \times b^x + c$. In the model, x represents the number of minutes the liquid has been cooling, and y represents the temperature, measured in Celsius for this example. The three parameters in the function are a , which represents the initial amount above room temperature; b , which represents the percent of temperature, relative to room temperature, the liquid retains each minute; and c represents the room temperature.

Do not tell students these parameters, but instead use guided inquiry to help them discover a , b , and c . (This model is based on Newton's law of cooling, which says that the difference between actual and room temperature is exponential. With our model, however, we can work directly with the temperatures, not the differences between the actual and room temperatures.) Understanding the model is challenging, but the exploration is rewarding. After we posed several questions and gave students an opportunity to explore, students seemed to grasp the reasoning behind the model:

- ◆ If the room temperature had been warmer, how would the model have changed? *The value for c would have been greater.*
- ◆ If room temperature had been the same, but the original temperature had been warmer, how would the model have been different? *The value for a would be greater.*

FIGURE 4 Measured Temperatures of our cooling liquid.

TIME (min)	TEMP (°C)
0	43.3
5	40.6
10	37.8
15	36.1
20	34.4
25	32.8
30	31.1
35	30.0
40	28.9
45	27.8
50	26.7
55	26.1
60	25.6
65	25.0
70	24.4

- ◆ What would a change in the b value indicate? *A greater value would indicate it retains more of its heat, implying that it would take longer to reach room temperature. Conversely, a lower value would indicate it retains less of its heat, meaning it would reach room temperature more quickly.*

As before, the spreadsheet requires estimates for the parameters. You might reason as follows. The room temperature might have been approximately 22°C, which serves as an estimate for c . (In reality, it was actually just a little above 20°C, which the final model predicts remarkably well. Note the advantage here over the direct use of Newton's law, in which we must first know the room temperature.) Because the temperature initially was 43.3°C, a should be close to 21°C, the difference between the estimated room temperature (the final temperature we expect) and the initial state.

Finally, you might guess that the temperature loses about 1% each minute, so you can estimate b to be 0.99 (the fraction of the temperature that it retains each minute). The model estimate thus becomes $y = 21 \times 0.99^x + 22$. Notice how much greater understanding in both science and mathematics students develop by estimating the parameters. In a computer lab, beginning with the values discussed in the class,

students can create spreadsheets that model the relationship between time and temperature. As in the Hooke's law demonstration, students should include a column labeled "PREDICTED" and another labeled "ERROR." They should also investigate the total error, using either the absolute values or the squared errors.

After adjusting the parameters to see how small they can make the total error, students can, as before, use the Solver to minimize the error. In the completed model (Figure 5), we interpreted the parameters as

follows. The value for c is approximately 20.5°C , suggesting that the room temperature was approximately 20.5°C . The value for a is approximately 22.8°C , suggesting that initially the liquid was approximately 22.8°C above room temperature. (Simple addition suggests this to be 43.3°C , which, remarkably enough, is precisely what our initial temperature was.) Finally, the value for b according to our model is approximately 0.975 , or 97.5% . This tells us that, on average, we lose about 2.5% of the temperature difference between the actual and room temperature each minute. In Figure 5, SSE stands for "Sum of Squared Errors." Students

should be able to complete the model with minimal help and answer questions about what they did, why they did it, and what it means.

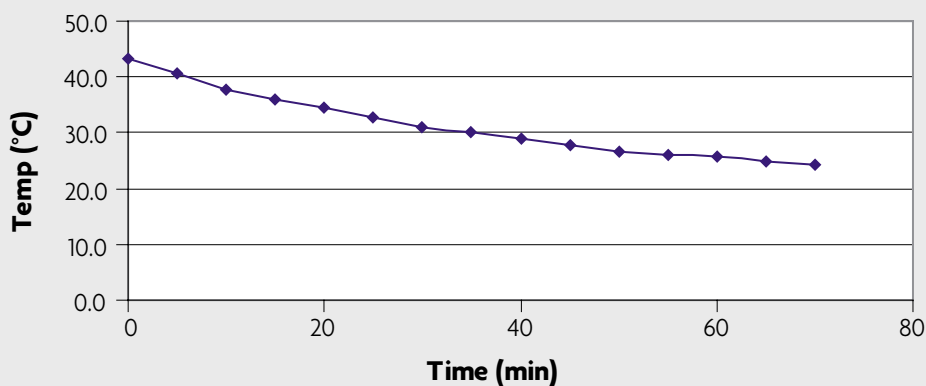
The use of spreadsheets can help students develop a solid understanding of important scientific concepts and a more sophisticated understanding of mathematical modeling. The contexts for which this strategy can develop meaning are almost limitless, from virtually every branch of science. Consider, for example, rates of reactions in chemistry, population growth in biology, or fossil deposition in the Earth sciences.

In our experience, students not only enjoyed what they were doing but also surprised us with how quickly they mastered the techniques, which allowed us to focus on the connections between the mathematical models and scientific ideas they were investigating. The process goals in the math and science standards have much in common. Teachers must take advantage of this and the power of modern technology to make learning as rich as possible. ■

Robert M. Horton (e-mail: bhorton@clemson.edu) is an assistant professor of education and William H. Leonard (e-mail: Leonard@Clemson.edu) is a professor of education and biology, both at Clemson University, 409B Tillman Hall, Clemson, SC 29634.

FIGURE 5 Newton's Law of Cooling actual experimental data.

A Cooling Liquid



TIME (min)	TEMP (°C)	PREDICTED TEMP (°C)	SQ ERROR	MODEL:	$y=a*b^x+c$
0	43.3	43.2	6.66E-03	a	22.8
5	40.6	40.5	4.07E-03	b	0.975
10	37.8	38.2	1.37E-01	c	20.5
15	36.1	36.1	2.84E-03		
20	34.4	34.2	2.50E-02		
25	32.8	32.6	3.32E-02		
30	31.1	31.2	7.27E-03		
35	30.0	29.9	6.16E-03		
40	28.9	28.8	8.71E-03		
45	27.8	27.8	5.41E-04		
50	26.7	27.0	6.54E-02		
55	26.1	26.2	8.20E-03		
60	25.6	25.5	7.14E-03		
65	25.0	24.9	6.39E-03		
70	24.4	24.4	2.75E-05		
		SSE	3.19E-01		