



Math Objectives

- Students will determine and analyze an exponential model for the data involving carbon dioxide in the atmosphere.
- Students will develop an improved model for this data by using the residuals associated with the exponential regression function.
- Students will apply mathematics to solve problems arising in everyday life, society, and the workplace (CCSS Mathematical Practice).
- Students will interpret mathematical results in the context of the situation and reflect on whether the results make sense, improving the model as necessary (CCSS Mathematical Practice).
- Students will use technological tools to explore and deepen understanding of concepts (CCSS Mathematical Practice).

Vocabulary

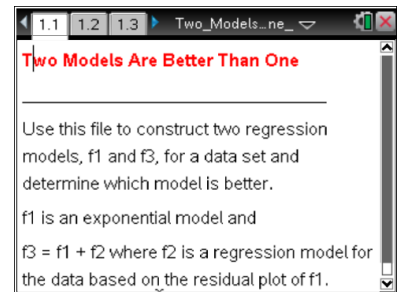
- regression equation
- residual
- residual plot
- SSE = sum of the squares of the residuals (errors)

About the Lesson

- This lesson involves modeling the amount of carbon dioxide in the air over a 12-month period
- As a result, students will:
 - Analyze the shortcomings of an exponential regression function fitting the given data.
 - Find an improved model by adding the original exponential function to one fitting the data in the residual plot of the original function.

TI-Nspire™ Navigator™ System

- Transfer a File.
- Use Quick Poll to assess students' understanding and compare equations.



TI-Nspire™ Technology Skills:

- Download a TI-Nspire document
- Open a document
- Move between pages
- Grab and drag a point

Tech Tips:

- Make sure the font size on your TI-Nspire handhelds is set to Medium.
- You can hide the function entry line by pressing **ctrl** **G**.

Lesson Files:

Student Activity

Two_Models_Are_Better_Than_One_Student.pdf

Two_Models_Are_Better_Than_One_Student.doc

TI-Nspire document

Two_Models_Are_Better_Than_One.tns

Visit www.mathnspired.com for lesson updates and tech tip videos.



Discussion Points and Possible Answers

Move to page 1.2.

The concentrations of carbon dioxide were measured monthly over a one-year period. The results are listed in the spreadsheet on Page 1.2.

- The first column contains the number of the month **[mon]** starting with January as 1.

The second column contains the concentration of carbon dioxide data collected during each month [in parts per million] **[conc]**.

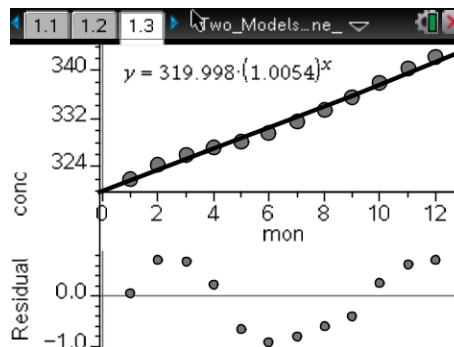
	mon	conc	resid1	resid2
1	1	321.8	0.07225	
2	2	324.2	0.733366	
3	3	325.9	0.685084	
4	4	327.2	0.227352	
5	5	328.1	-0.63988	

Move to page 1.3.

A scatterplot of the data from the spreadsheet is shown on Page 1.3.

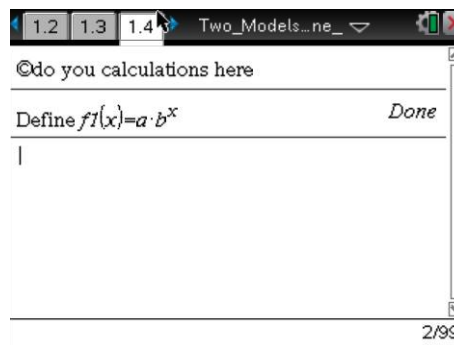
Using an exponential function to model the data, the regression equation is $y = 319.998 \cdot (1.0054)^x$. The residual plot is shown in the screen at the bottom of the page (the independent variable, month, is plotted along the horizontal axis, and the corresponding residual is plotted along the vertical axis.).

Recall that the residual associated with the data point (x_i, y_i) is $y_i - f(x_i)$ where y_i is the actual value of the data and $f(x_i)$ is the value predicted by the regression function.



Move to page 1.4.

The function y from Page 1.3 has been saved as $f1(x) = a \cdot b^x$ where $a = 319.998$ and $b = 1.0054$.

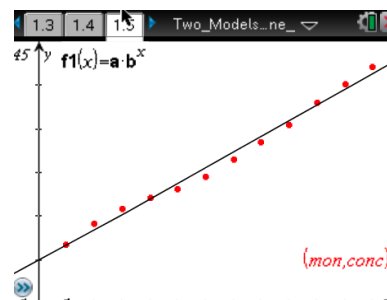




Move to page 1.5.

Page 1.4 shows a scatter plot of the data and a graph of $y = f1(x)$.

- Looking at the graph of $y = f1(x)$ on the scatterplot, is the function $f1$ a good fit for this data? Give reasons for your answer.



Sample Answer: Although the fit looks reasonably good, there are, alternately, data points above and below the graph of the exponential function indicating a possible pattern in the residual plot.

Move back to page 1.2.

To save the residuals in the spreadsheet, “ $= conc - f1(mon)$ ” has been entered into the cell right below the title of the third column, *resid1*. The residuals are displayed. In the residual, $y_i - f(x_i)$, y_i is the actual value of the concentration in month i , whereas $f(x_i)$ is the value of the concentration in month i predicted by the regression function, $f1$.

Move to page 1.4.

- What is the largest residual in absolute value (or magnitude)?

Answer: By scanning the list of twelve residuals or using $\max(abs(resid1))$, we find the largest residual in absolute value to be 0.91666303225.

Another measure of how well a function, f , fits a set of data is the sum of the squares of the residuals denoted by $SSE(f)$. Usually, the smaller the value of the $SSE(f)$, the better f fits the data.

- Compute $SSE(f1)$, the sum of the squares of the residuals for $f1$, by typing $sum(resid1^2)$, and then record its value here:

Answer: $SSE(f1) = 4.505$

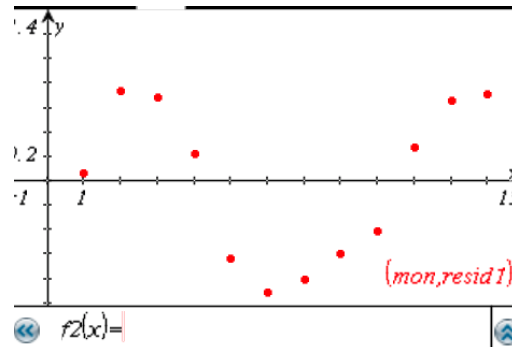


Teacher Tip: Students could use Scratchpad or the Calculator page to compute their answers for Questions 2 and 3.

Move to page 1.6.

Another copy of the residual plot is included on Page 1.6, $x = mon$ and $y = resid1$.

- Because the residual plot associated with the regression function $f1(x)$ suggests a pattern, there might be a better fit to the original data. As a first step to finding such a function, what type of function would be a good fit for the data in the residual plot? Give a reason for your answer.



Sample Answers: Either a sinusoidal function or quartic polynomial function is a possible fit since both functions can have two relative maxima for $0 \leq x \leq 13$.

Move back to page 1.4.

- Find the regression function, $f2$, of the type you chose in Question 4 by selecting **MENU > Statistics > Stat Calculations > the regression function, $f2$** . Use $x = mon$ and $y = resid1$. Save **RegEqn** to $f2$, and record this function here:

Sample Answers: $f2(x) = 0.838387 \sin(0.702634x - 0.13214) - 0.171035$ or
 $f2(x) = -0.003784x^4 + 0.104132x^3 - 0.913467x^2 + 2.71362x - 1.80806$.

TI-Nspire Navigator Opportunity: Quick Poll
 See Note 1 at the end of this lesson.

Move to page 1.6.

Check the fit of $f2(x)$ to the data in the residual plot. Open the entry line, move back up to $f2(x)$, and press **enter**. If $f2(x)$ does not appear to be a very good fit, make another choice for $f2$, and repeat the steps on Page 1.4 again.



Move back to page 1.4.

6. Define $f_3(x) = f_1(x) + f_2(x)$, and record it here:

Sample Answers:

$$f_3(x) = 319.998 \cdot (1.0054)^x + 0.838387 \sin(0.702634x - 0.13214) - 0.171035 \text{ or}$$

$$f_3(x) = 319.998 \cdot (1.0054)^x - 0.003784x^4 + 0.104132x^3 - 0.913467x^2 + 2.71362x - 1.80806.$$

In the next few steps, you will verify that $f_3(x)$ is a better fit than $f_1(x)$ to the original data.

Move to page 1.5.

To view the graph of $f_3(x)$ on the scatterplot, open the entry line, move back up to $f_3(x)$, and press enter.

7. Looking at the graph of $f_3(x)$ on the scatterplot, give a reason why it is a better fit to the data than $f_1(x)$.

Sample Answer: This graph oscillates a little bit and fits the data much better.

Move back to page 1.2.

Record the residuals for $f_3(x)$ in the spreadsheet by selecting the cell right below the title of the fourth column, *resid2*, and entering = **conc** - **f3(mon)**. The values of the residuals will be displayed.

8. What is the largest residual in absolute value (or magnitude)?

Sample Answers: By scanning the list of twelve residuals or using $\max(\text{abs}(\text{resid1}))$, we find the largest residual in absolute value to be 0.27001580138 if

$$f_3(x) = 319.998 \cdot (1.0054)^x + 0.838387 \sin(0.702634x - 0.13214) - 0.171035;$$

or 0.21456876088 if

$$f_3(x) = 319.998 \cdot (1.0054)^x - 0.003784x^4 + 0.104132x^3 - 0.913467x^2 + 2.71362x - 1.80806.$$



Move to page 1.6.

Hide the graph of $f2$. Then change the function type to **scatterplot**. Display the residual plot of $f3$ by setting $x = mon$ and $y = resid2$.

9. By comparing the two residual plots, give a reason why $f3(x)$ is a better fit to the data than $f1(x)$ is.

Sample Answer: The residuals associated with $f3$ do not suggest a pattern, and their values are between -0.2 and 0.2 while the values of the original residuals are between -0.8 and 0.8.

Move to page 1.4.

10. Finally, find $SSE(f3)$ by typing $sum(resid2^2)$, and record the value here:

Sample Answers: $SSE(f3) = 0.254038$ if

$$f3(x) = 319.998 \cdot (1.0054)^x + 0.838387 \sin(0.702634x - 0.13214) - 0.171035;$$

$SSE(f3) = 0.127212$ if

$$f3(x) = 319.998 \cdot (1.0054)^x - 0.003784x^4 + 0.104132x^3 - 0.913467x^2 + 2.71362x - 1.80806.$$

11. Compare the values of and $SSE(f1)$ by completing the statement:

Answer: 5.6% if

$$f3(x) = 319.998 \cdot (1.0054)^x + 0.838387 \sin(0.702634x - 0.13214) - 0.171035;$$

2.8% if

$$f3(x) = 319.998 \cdot (1.0054)^x - 0.003784x^4 + 0.104132x^3 - 0.913467x^2 + 2.71362x - 1.80806.$$

12. Review your answers to the previous questions, and summarize the reasons why $f3(x)$ is a better fit to the data than $f1(x)$ is.

Answer: The graph of $y = f3(x)$ fits the data better; the residual plot associated with $f3(x)$ does not have a pattern and has a smaller range of values than the residuals of $f1(x)$; $SSE(f3)$ is less than 6% of $SSE(f1)$.



Wrap Up

Upon completion of the lesson, the teacher should ensure that students are able to understand:

- How to find a regression equation to fit real-world data.
- How to find an improved model to an original one by adding a function fitting the data in the residual plot of the original model.
- How to analyze how well a function fits a set of data using the graph of the function, the residual plot of the function, and the SSE value of the function.

TI-Nspire Navigator

Note 1

Question 5, Name of Feature: Quick Poll

You might want to send a Quick Poll to have students send in their regression equations. You can then compare the similarities and differences in their equations.