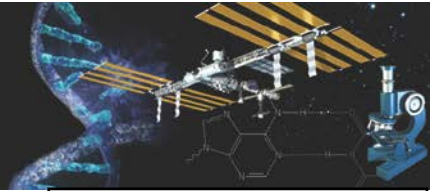




MATH AND SCIENCE @ WORK

AP* STATISTICS Educator Edition



SPACECRAFT RADAR TRACKING

Instructional Objectives

Students will

- calculate and interpret the meaning of the correlation coefficient, r , and the coefficient of determination, r^2 ;
- write a regression equation, and interpret the meaning of the slope and y -intercept in context to the problem;
- make predictions based on a least squares regression model; and
- plot residuals, and interpret their graphical form.

Degree of Difficulty

For the average student in AP Statistics, this problem is at a moderate difficulty level.

Class Time Required

This problem requires 50–65 minutes.

- Introduction: 5–10 minutes
 - Read and discuss the background section with the class before students work on the problem.
- Student Work Time: 35–45 minutes
- Post Discussion: 10 minutes

Background

This problem is part of a series of activities that applies Math and Science @ Work in NASA’s research facilities.

The Western Aeronautical Test Range (WATR), located at NASA Dryden Flight Research Center in Edwards, California, provides range engineering, technical expertise, and resources to support aerospace research, science, and low-Earth orbiting missions. High accuracy radar provides tracking and space positioning information on the International Space Station (ISS), as it has previously done with other research vehicles such as the space shuttle.

During previous space shuttle missions, the WATR’s Aeronautical Tracking Facility (ATF) provided telemetry (information collected via radio waves), radar, voice communication, and video support of ISS and space shuttle activities to NASA Johnson Space Center in Houston, Texas, using different telemetry tracking, space positioning, and audio communication systems.

Grade Level

11–12

Key Topic

Linear Regression

Degree of Difficulty

Moderate

Teacher Prep Time

5–10 minutes

Class Time Required

50–65 minutes

Technology

- TI-Nspire™ Learning Handhelds
- TI-Nspire document: *Radar_Tracking.tns*

AP Course Topics

- Exploring Data:
- Exploring Bivariate Data

NCTM Standards

Data Analysis and Probability

*AP is a trademark owned by the College Board, which was not involved in the production of, and does not endorse, this product.



The telemetry tracking system provided downlinked status information on the condition of the space shuttle and available video (from the pilot's point of view) to the NASA network via satellite. When required, this tracking system also provided uplinked command data to the space shuttle.



Figure 1: The International Space Station



Figure 2: WATR instrumentation radar

The space positioning system consists of two high-accuracy radars, differential global positioning system ground stations, and Federal Aviation Administration surveillance radar data. This system was used to track every space shuttle orbit, relaying time-space positioning information from launch to landing. The space positioning system also tracked the ISS—from the day prior to each space shuttle launch through the duration of each mission—providing critical information for the docking and undocking of the space shuttle.

Throughout each shuttle mission, voice communication was also enabled by the audio communications system. While a system of communication satellites used by NASA and other United States government agencies (known as the Tracking and Data Relay Satellite System, or TDRSS) provided the primary voice communication for the space shuttle, the WATR provided back-up support in case of a communications failure during a mission. The WATR also became the primary means of communication support in the event a space shuttle might be diverted locally to Edwards Air Force Base for a landing.

Mission data from these three systems were processed real-time (or near real-time), and were archived by NASA as a means of support for post-mission analyses.

AP Course Topics

Exploring Data

- Exploring bivariate data
 - Analyzing patterns in scatter plots
 - Least squares regression line
 - Residual plots

NCTM Standards

Data Analysis and Probability

- Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them
- Select and use appropriate statistical methods to analyze data



Problem and Solution Key (One Approach)

Great precision is used in tracking and recording sightings of the International Space Station (ISS) and other research vehicles. High accuracy instrumentation radars take measurements of the geometric range, or the distance to the vehicle. These radars track the moving vehicle from horizon to horizon—out to a distance of 4,828 kilometers (3,000 nautical miles), with amazing accuracies to 0.0006 degrees in angle and 9 meters (30 feet) in range. To achieve these levels, the radars must be carefully calibrated to remove atmospheric effects, location errors, and electronic equipment biases.

The following problem will investigate the reason for such great precision. How would inaccurate sightings affect the safety of the crew and the necessary communication and tracking between Earth and the ISS? To answer this question, you will investigate data that was collected from both calibrated and uncalibrated radars.

On the TI-Nspire™ handheld, open the document, *Radar_Tracking.tns*. Page 1.2 shows data collected during one pass of the ISS over the radar site. The radar started tracking the ISS as it came into target a couple of degrees above the horizon. It continued to track it as it passed over the site, and then lost the target a couple of degrees above the opposite horizon. The table shows three different geometric range values of the ISS measured in kilometers (km). The theoretical value represents a value seen from a perfect radar station, with no errors or biases (a scenario which exists only in simulations). The data in the last two columns are geometric range values measured by two different radars (uncalibrated and calibrated).

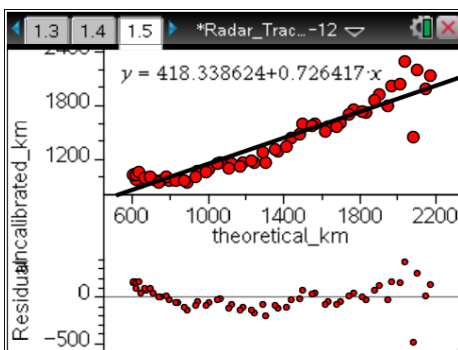
Read through the questions in the TI-Nspire document and perform calculations within the file. Record your answers either within the file or in the space provided on this handout.

Students are given the following problem information within the TI-Nspire document, Radar_Tracking.tns, which should be distributed to their TI-Nspire handhelds.

- A. Add a **Data & Statistics** page, and create a scatter plot using the theoretical range as the explanatory variable and the uncalibrated range data as the response variable. Determine the least squares regression line, and interpret its slope and y-intercept. What kind of accuracy does the linear model predict for the data? Give statistical justification for your answer.

As a means to substantiate their conclusions, students should:

- *determine the least squares regression line, and interpret its slope and y-intercept*
- *calculate values for r and r^2*
- *construct residual plots*



stat.results	
"Title"	"Linear Regression (a+bx)"
"RegEqn"	"a+b*x"
"a"	418.338624
"b"	0.726417
"r ² "	0.884896
"r"	0.940689
"Resid"	"{"...}"

Conclusions:

The correlation coefficient, r , shows a strong positive correlation of 0.940689. The coefficient of determination, r^2 , has a value of 0.884896. This means that about 88.4896% of the variation in the uncalibrated radar range is explained by the least square regression with the theoretical



radar range. This is not a very high value, so predictions using this data will be only somewhat reliable.

The slope value of the uncalibrated radar is 0.726417 (which is the change in range predicted by the uncalibrated radar over a change in the theoretical range). For every 1 km increase in the theoretical range, an average increase of 0.726417 km in the uncalibrated radar's range value is expected. The y-intercept is 418.338624 km, which is the distance the uncalibrated radar would measure when the theoretical distance to the target is 0 km.

The residual plot also shows a slightly curved appearance, which tells us the linear fit may not necessarily be the best model for this data. This supports our conclusion with our other findings that the data shows a somewhat linear trend with only somewhat good predictability.

- I. If the theoretical range data is 1500 km, use the linear model to estimate the range that the uncalibrated radar will predict.

$$\hat{y} = 418.338624 + 0.726417(1500)$$

$$\hat{y} = 1507.964124$$

Uncalibrated Radar Range \approx 1507.964 km

- II. If the theoretical range data is 2000 km, estimate the range that the uncalibrated radar will predict.

$$\hat{y} = 418.338624 + 0.726417(2000)$$

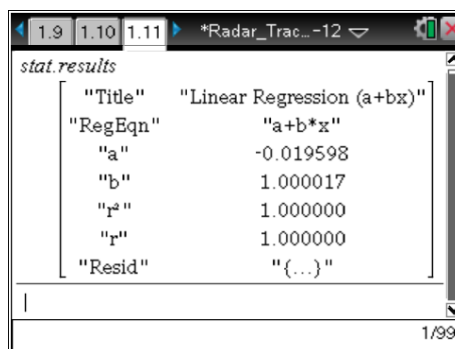
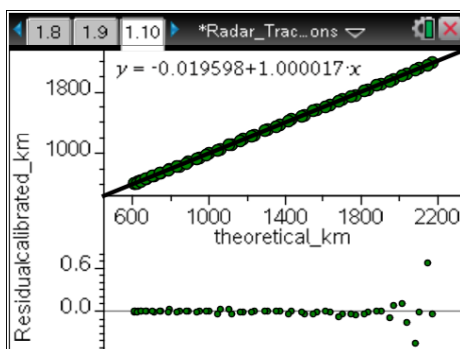
$$\hat{y} = 1871.172624$$

Uncalibrated Radar Range \approx 1871.173 km

- B. Add a **Data & Statistics** page, and create a scatter plot using the theoretical range as the explanatory variable and the calibrated range data as the response variable. Determine the least squares regression line, and interpret its slope and y-intercept. What kind of accuracy does the linear model predict for the data? Give statistical justification for your answer.

As a means to substantiate their conclusions, students should:

- determine the least squares regression line, and interpret its slope and y-intercept
- calculate values for r and r^2
- construct residual plots





Conclusions:

The correlation coefficient, r , shows a very strong positive correlation of 1. The coefficient of determination, r^2 , has a value of 1. This means that about 100% of the variation in the calibrated radar range is explained by the least square regression with the theoretical radar range. This is a very high value, so predictions using this data will be much more reliable.

The slope value of the calibrated radar is 1.000017. For every 1 km increase in the theoretical range, an average increase of 1.000017 km in the calibrated radar's range value is expected. The y-intercept is -0.019598 km, which is the distance the calibrated radar would measure when the theoretical distance to the target is 0 km.

The residual plot shows no noticeable pattern; therefore, the trend of the data appears to be linear. The data shows a linear trend with good predictability.

- I. If the theoretical range data is 1500 km, use the linear model to estimate the range value the calibrated radar will predict.

$$\hat{y} = 0.019598 + 1.000017(1500)$$

$$\hat{y} = 1500.005902$$

Calibrated Radar Range \approx 1500.006 km

- II. If the theoretical range data is 2000 km, estimate the range value the calibrated radar will predict.

$$\hat{y} = 0.019598 + 1.000017(2000)$$

$$\hat{y} = 2000.014$$

Calibrated Radar Range \approx 2000.014 km

- C. Compare the uncalibrated and calibrated radars.

- I. Based on your findings from both calibrated and uncalibrated radars, which radar is more accurate in tracking range data? Explain your reasoning using statistical justification.

The calibrated radar appears to be superior to the uncalibrated radar. The tracking from the calibrated radar has less variation, as is evidenced by the points being closer to the regression line. The residuals plots also show less variation. The correlation coefficient and the value of r^2 are both 1, indicating a better fit and higher confidence in predictability.

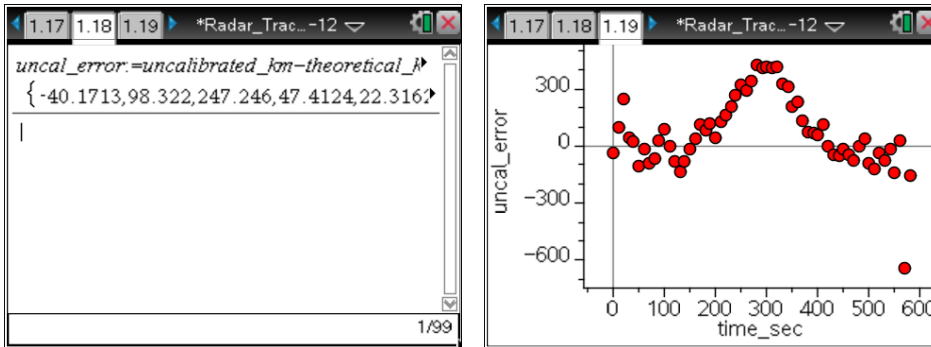
- II. What would you expect the model to be for a radar that was 100% accurate? Explain your reasoning.

If the radar was 100% accurate, then its range values would be identical to the theoretical values, which would make the regression line $y = x$, with a slope of 1, and a y-intercept of 0.

For vehicles like the ISS, NASA already has very good models that can predict where the vehicle will be from pass to pass. In order to calibrate radars, NASA flight dynamics specialists compare radar data to the existing model by creating an error plot. After analyzing this plot, they make adjustments to the calibrations. They continue to compare the data and make adjustments until they have removed all errors and biases (other than the small electronic noise that cannot be filtered out). The shape, spread, and offset of the residuals gives insight into exactly what might be causing calibration errors.



- D. Add a **Calculator** page and define a new variable that will represent the error of the uncalibrated data by typing: **uncal_error:=uncalibrated_km – theoretical_km** and press **Enter**. Then add a **Data & Statistics** page, and create a plot of the uncalibrated data error versus time. Remember that the radar tracks the vehicle from horizon (0 sec) to horizon (580 sec). The vehicle is furthest from the radar at the horizons, and is closest to the radar when it is directly over the radar.



- I. Describe the error plot. What trends do you see?

The variability is greater at the beginning and end of the graph, and is lesser in the middle of the graph. The error tends to be smaller at the ends of the graph, and is much larger in the middle of the graph.

- II. Explore the theoretical range data in the table on the next page to see how the distance to the ISS changes during this time. What kind of relationship is shown between the error and the distance from the ISS? Hypothesize reasons that might explain this relationship.

There is some error at the beginning and the end of the graph when the ISS is farthest away from the radar. The error increases as the ISS moves closer to the radar—until the ISS moves farther from the radar and the error decreases. It appears there is more error in the data at closer proximity between the ISS and the radar.

Students' answers will vary on reasons that might explain the relationship found.

Teacher's note: Discuss with your students that there are a number of different errors that show themselves in different ways in this plot. Atmospheric effects usually show up as increased noise (i.e., spread) at low elevations and long ranges. The precise location of the tracking antenna on the Earth's surface must also be calibrated. This effect usually shows up as a bias (i.e., offset) at shorter ranges. When an object has a relatively small bias, it is not as noticeable from a far distance. As the object is approached, the more noticeable this discrepancy becomes.



Scoring Guide

Parts A, B, C, and D are scored as Essentially Correct (E), Partially Correct (P), or Incorrect (I).

Question	Score	Description
A	<i>Essentially Correct (E)</i>	All graphs are labeled and interpreted correctly, all questions are answered correctly, and all statistical information is complete and accurate.
	<i>Partially Correct (P)</i>	Graphs are labeled and interpreted correctly, all questions are answered correctly, but some statistical information is incomplete or inaccurate.
	<i>Incorrect (I)</i>	Graphs are not labeled correctly, some questions are answered correctly, but most statistical information is incomplete or inaccurate.
B	<i>Essentially Correct (E)</i>	All graphs are labeled and interpreted correctly, all questions are answered correctly, and all statistical information is complete and accurate.
	<i>Partially Correct (P)</i>	Graphs are labeled and interpreted correctly, all questions are answered correctly, but some statistical information is incomplete or inaccurate.
	<i>Incorrect (I)</i>	Graphs are not labeled correctly, some questions are answered correctly, but most statistical information is incomplete or inaccurate.
C	<i>Essentially Correct (E)</i>	All statistical information is presented and interpreted correctly, the question is answered without error, and all statistical information is accurate and complete.
	<i>Partially Correct (P)</i>	Some statistical information is presented and interpreted correctly, the question is answered with inconsistencies, and some statistical information is incomplete or inaccurate.
	<i>Incorrect (I)</i>	The statistical information that is presented is incorrect, with the question answered incorrectly or with incorrect interpretation of the statistics calculated.
D	<i>Essentially Correct (E)</i>	All graphs are labeled and interpreted correctly, all questions are answered correctly, and all statistical information is complete and accurate.
	<i>Partially Correct (P)</i>	Graphs are labeled and interpreted correctly, all questions are answered correctly, but some statistical information is incomplete or inaccurate.
	<i>Incorrect (I)</i>	Graphs are not labeled correctly, some questions are answered correctly, but most statistical information is incomplete or inaccurate.



Point Distribution

Suggested 4 points total to be given as follows:

- | | | |
|--------------|-----------------------------|---|
| 4 pts | Complete Response | All parts are essentially correct. |
| 3 pts | Substantial Response | Three parts are essentially correct and 1 or no parts are partially correct.
<i>OR</i>
Two parts are essentially correct and two parts are partially correct. |
| 2 pts | Developing Response | Two parts are essentially correct and no parts are partially correct.
<i>OR</i>
One part is essentially correct and two or three parts are partially correct.
<i>OR</i>
Four parts are partially correct. |
| 1 pt | Minimal Response | One part is essentially correct and one or no parts are partially correct.
<i>OR</i>
No parts are essentially correct and two to three parts are partially correct. |

Contributors

This problem was developed by the Human Research Program Education and Outreach (HRPEO) team with the help of NASA subject matter experts and high school AP Statistics instructors.

NASA Expert

Greg Holt, PhD – Flight Dynamics, Ascent/Entry Navigation Lead, NASA Johnson Space Center, Houston, TX

AP Statistics Instructors

Sharon Cichocki – Texas Instruments T³ (Teachers Teaching with Technology™) National Instructor, Hamburg High School, Hamburg, NY

Curtis Brown – AP Statistics Coordinator, Laying the Foundation, Inc., Dallas, TX