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
- Data analysis software

AP Course Topics
Exploring Data:
- Exploring Bivariate Data

NCTM Standards
Data Analysis and Probability

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.

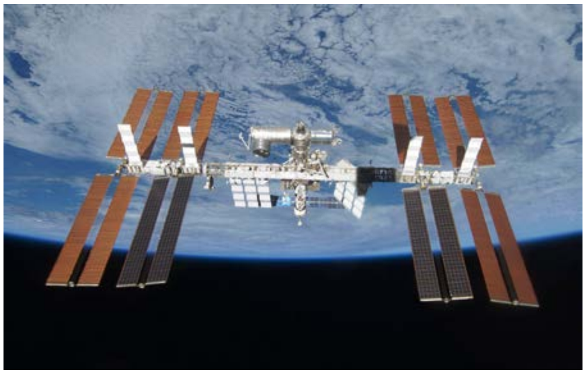


Figure 1: The International Space Station

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.

Table 1 shows abbreviated 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). A complete table is found in Appendix A and is also provided in an Excel file to use for data analysis.

Table 1: Abbreviated ISS Radar Tracking Data

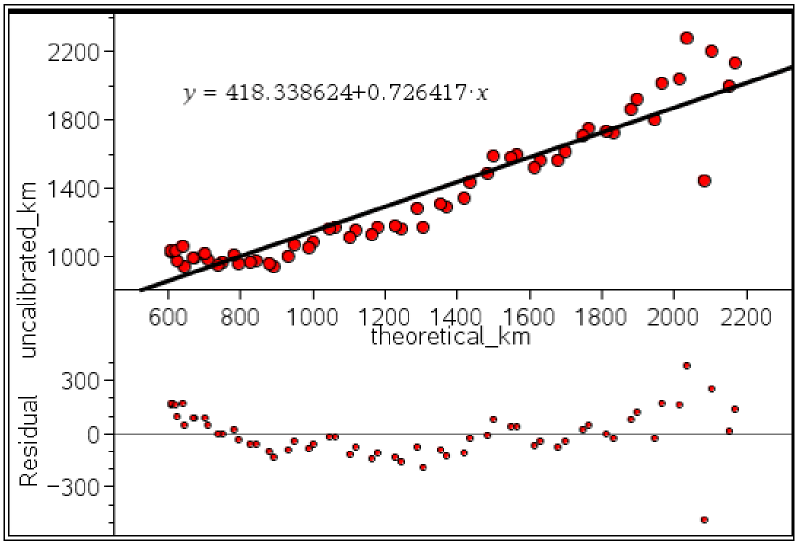
Time (sec)	Theoretical (km)	Uncalibrated (km)	Calibrated (km)
0	2170.111598	2129.940333	2170.09689
10	2102.040628	2200.362598	2102.040253
20	2034.114088	2281.359857	2033.983616
30	1966.355309	2013.767679	1966.452606
…	…	…	…

A. 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*

   The screenshot below shows an example of the plot, least squares regression line, and residual plot using data analysis software.



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.

$$\hat{y} = 418.338624 + 0.726417(1500)$$
$$\hat{y} = 1507.964124$$
$$\text{Uncalibrated Radar Range} \approx 1507.964\text{ km}$$

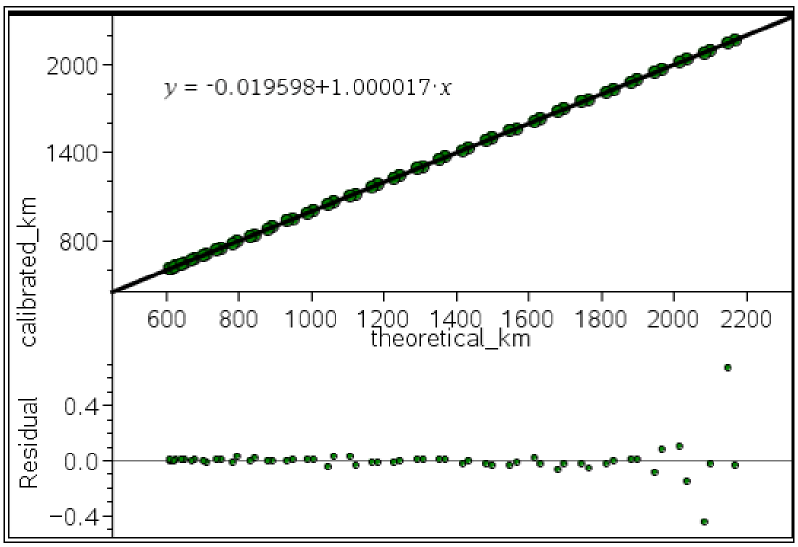
II. If the theoretical range data is 2000 km, estimate the range that the uncalibrated radar will predict.

$$\hat{y} = 418.338624 + 0.726417(2000)$$
$$\hat{y} = 1871.172624$$
$$\text{Uncalibrated Radar Range} \approx 1871.173\text{ km}$$

I. 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.

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

$$\hat{y} = 1500.005902$$

$$\text{Calibrated Radar Range} \approx 1500.006\ km$$

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 ≈ 2000.014 *km*

B. 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.

C. The error of the uncalibrated radar is determined by finding the difference in the uncalibrated range values and the theoretical range values. Figure 3 is the error plot of the uncalibrated radar 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.

Uncalibrated Radar Error vs. Time

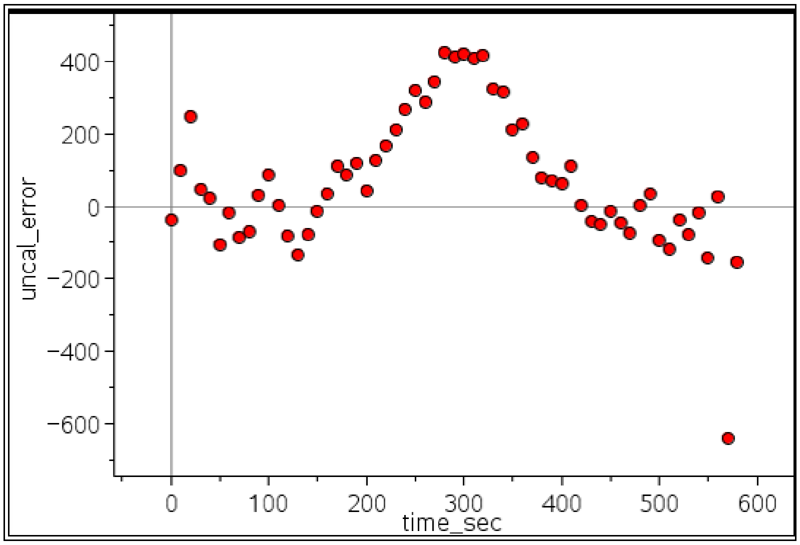


Figure 3: The Error plot of uncalibrated radar error versus time

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 found in Table 2 on Appendix A 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	*Essentially Correct (E)*	All graphs are labeled and interpreted correctly, all questions are answered correctly, and all statistical information is complete and accurate.
	Partially Correct (P)	Graphs are labeled and interpreted correctly, all questions are answered correctly, but some statistical information is incomplete or inaccurate.
	Incorrect (I)	Graphs are not labeled correctly, some questions are answered correctly, but most statistical information is incomplete or inaccurate.
B	*Essentially Correct (E)*	All graphs are labeled and interpreted correctly, all questions are answered correctly, and all statistical information is complete and accurate.
	Partially Correct (P)	Graphs are labeled and interpreted correctly, all questions are answered correctly, but some statistical information is incomplete or inaccurate.
	Incorrect (I)	Graphs are not labeled correctly, some questions are answered correctly, but most statistical information is incomplete or inaccurate.
C	*Essentially Correct (E)*	All statistical information is presented and interpreted correctly, the question is answered without error, and all statistical information is accurate and complete.
	Partially Correct (P)	Some statistical information is presented and interpreted correctly, the question is answered with inconsistencies, and some statistical information is incomplete or inaccurate.
	Incorrect (I)	The statistical information that is presented is incorrect, with the question answered incorrectly or with incorrect interpretation of the statistics calculated.
D	*Essentially Correct (E)*	All graphs are labeled and interpreted correctly, all questions are answered correctly, and all statistical information is complete and accurate.
	Partially Correct (P)	Graphs are labeled and interpreted correctly, all questions are answered correctly, but some statistical information is incomplete or inaccurate.
	Incorrect (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. *OR* 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. *OR* One part is essentially correct and two or three parts are partially correct. *OR* Four parts are partially correct.
1 pt	**Minimal Response**	One part is essentially correct and one or no parts are partially correct. *OR* 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 T^3 (Teachers Teaching with Technology™) National Instructor, Hamburg High School, Hamburg, NY

Curtis Brown – Mathematics Coordinator, National Math and Science Initiative

Appendix A

Table 2: Complete ISS Radar Tracking Data

Time (sec)	Theoretical (km)	Uncalibrated (km)	Calibrated (km)
0	2170.111598	2129.940333	2170.09689
10	2102.040628	2200.362598	2102.040253
20	2034.114088	2281.359857	2033.983616
30	1966.355309	2013.767679	1966.452606
40	1898.790593	1921.106824	1898.811696
50	1831.449734	1724.927367	1831.458216
60	1764.366652	1746.803512	1764.32018
70	1697.580162	1610.180969	1697.565336
80	1631.134911	1560.000619	1631.126878
90	1565.082533	1596.034702	1565.083081
100	1499.483065	1587.106636	1499.454591
110	1434.406705	1436.385902	1434.41691
120	1369.935989	1288.517779	1369.947788
130	1306.168507	1172.969065	1306.177142
140	1243.22028	1163.3699	1243.22306
150	1181.229982	1166.093591	1181.223399
160	1120.364188	1155.131971	1120.32941
170	1060.82387	1170.182927	1060.854167
180	1002.852341	1087.373574	1002.85862
190	946.7448002	1066.028503	946.749758
200	892.8594117	936.3822804	892.8543871
210	841.629467	968.3546735	841.642961
220	793.5753681	958.754003	793.5970401
230	749.3138548	960.7191578	749.3149203
240	709.5599223	978.3150454	709.5453366
250	675.1145587	993.6148357	675.119861
260	646.8298591	935.9674692	646.8281996
270	625.5446379	970.5768827	625.548281
280	611.9912695	1034.422285	611.9840092
290	606.6887521	1019.743088	606.6749012
300	609.8519051	1028.81792	609.8533493

Time (sec)	Theoretical (km)	Uncalibrated (km)	Calibrated (km)
310	621.3500195	1030.47825	621.3381634
320	640.7321732	1057.180291	640.7339037
330	667.3085399	991.2794741	667.303696
340	700.2565526	1014.299819	700.2487126
350	738.7195776	950.2074156	738.7286407
360	781.8790252	1008.907245	781.8646839
370	828.9961307	964.130932	828.9889734
380	879.4290518	956.4328668	879.4284295
390	932.6337487	1001.568104	932.628382
400	988.1560542	1050.021734	988.1670513
410	1045.620058	1156.405567	1045.58054
420	1104.715822	1107.951618	1104.743459
430	1165.187953	1122.845519	1165.182357
440	1226.825655	1175.308416	1226.817647
450	1289.454389	1275.617015	1289.471902
460	1352.929025	1305.717708	1352.941031
470	1417.128306	1342.639245	1417.111756
480	1481.950382	1484.466204	1481.936056
490	1547.309241	1582.025328	1547.28961
500	1613.131838	1520.325417	1613.157553
510	1679.355792	1559.017114	1679.296946
520	1745.927537	1709.641215	1745.914351
530	1812.800822	1734.836372	1812.788487
540	1879.935499	1861.90462	1879.96188
550	1947.296533	1802.97012	1947.229642
560	2014.853197	2042.265532	2014.974438
570	2082.578405	1443.108218	2082.147312
580	2150.448165	1994.737904	2151.146419