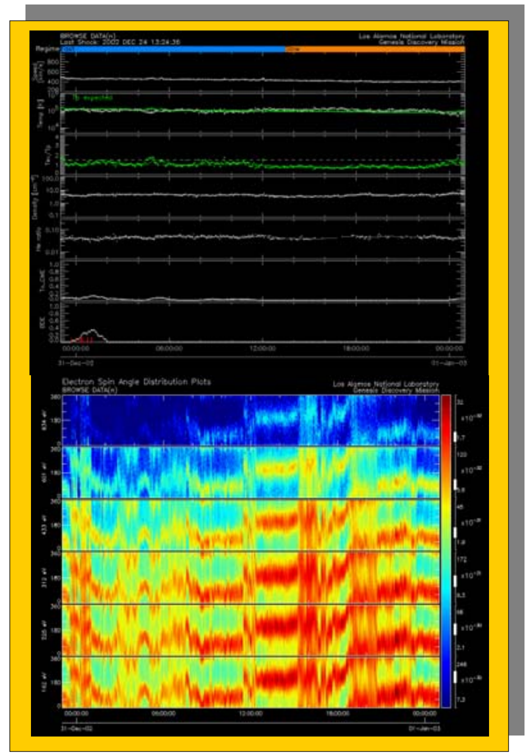
Data Analysis and Generalizations

Exploring Data

TEACHER GUIDE

BACKGROUND INFORMATION

The *National Science Education Standards* "Science as Inquiry" standard calls for students to learn how to analyze evidence and data. "The evidence they analyze may be from their investigations, other students' investigations, or databases. Data manipulation and analysis strategies need to be modeled by teachers of science and practiced by students. Determining the range of the data, the mean and mode values of the data, plotting the data, developing mathematical functions from the data, and looking for anomalous data are all examples of analyses students can perform. Teachers of science can ask questions such as, 'What patterns did you detect in the data? Were there any surprises or anomalies in the data? How confident do you feel about the accuracy of the data?'"



Los Alamos National Laboratory

Top: Solar Wind Data Plot
Bottom: Electron Spin Distribution Plot

In this activity, students will take a "first look" at actual solar wind data that has been collected and sent to Earth by the Genesis spacecraft. The data summary sheets contain seven panels: Speed, Temperature, Expected Temperature/Measured Temperature, Density, Helium Ratio, Coronal Mass Ejection (CME), and Bi-Directional Electron Indicator (BDE). In this activity, students will be studying the first five panels. All of these plots are determined autonomously by the Genesis onboard flight software. The averaging information is based on one-hour time frames. The label on the upper left-hand corner indicates the time of the last shock identification, referring to interplanetary shocks. Interplanetary shocks are easily identified as simultaneous, abrupt jumps in the speed, density, and temperature of solar wind. The term "planet" in interplanetary shocks can be misleading. Interplanetary in this regard really means "in the middle of nowhere." In other words, these shocks are not associated with planets at all, so the term "interplanetary" distinguishes these shocks from bow shocks associated with planets, asteroids, and any other objects. Interplanetary shocks result from fast wind plowing into slower wind. An interplanetary shock occurs rapidly, within a period of minutes. A vertical line extending through all of the panels indicates a shock identification by Genesis. This detection is retrospective, meaning the identification typically occurs about 30 minutes after the actual time of shock passage.

When viewing the solar wind summary data plots, note the color code for the solar wind regimes: yellow indicates slow inter-stream solar wind; blue indicates coronal hole fast solar wind; red indicates coronal mass ejection solar wind; and gray indicates that the regime is unknown (this occurs for short intervals when the algorithm is re-initialized. The white, dotted lines on all panels indicate the average of the measured data. These white, dotted lines are not always easily distinguishable on black and white printouts.

- The first panel is the Speed panel. Speed is measured here in kilometers per second. The white points indicate the speed of solar wind in kilometers per second, while the white, dotted line indicates the running average speed.
- The second panel is the Temperature panel, which measures the temperature in kelvins (K). There are three data sets on the Temperature panel. When viewing these data summaries on a computer or color printout, white points represent the measured temperature readings; the green points represent the temperatures expected for normal (non-CME) solar wind a given speed; and the white dots represent the average of the measured temperature readings.

- The third panel (Tex/Tp) shows the expected temperature to measured temperature readings as green points, while the green, dotted line indicates the average ratio.
- The Density panel is the fourth panel on the page. Density is measured in particles per centimeters cubed. The white points represent the proton (hydrogen nucleus) density. The white, dotted line represents the running average proton density.
- The Helium ratio panel is the fifth panel on the page. The white points show the ratio of alpha particle (helium nucleus) number density to the proton number density. The white, dotted line shows the running average of alpha/proton ratio.

Although the last two panels (CME panel and BDE panel) are not used in this activity, a brief description follows: the CME panel uses a fuzzy logic scheme to determine the likelihood that a measured solar wind is coronal mass ejection related. Higher values indicate greater likelihood of CME detection. The BDE panel (Bi-Directional Electron indicator) indicates whether or not the measured electron spectrum shows the characteristics of bi-directional electron streaming. A value of 1 indicates continuous bi-directional streaming over the past hour, while a value of 0 indicates no BDEs over the last hour. Intermediate values indicate a small amount of BDEs detected over the last hour. For a more complete description of the solar wind summary data plots, visit http://genesis.lanl.gov/plots/test/mec/description.html.

In "Exploring Data," students explore Genesis solar wind data by studying information on the Los Alamos National Laboratory (LANL) Genesis data Web site. Students record data patterns and anomalies from the predicted values, and think about questions that might arise from this study.

NATIONAL SCIENCE STANDARDS ADDRESSED

(Source – *National Science Education Standards*)

Grades 9-12

Science As Inquiry

Abilities necessary to do scientific inquiry
Understandings about scientific inquiry

(View a full text of the National Science Education Standards.)

PRINCIPLES AND STANDARDS FOR SCHOOL MATHEMATICS

Grades 9-12

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
Develop and evaluate inferences and predictions that are based on data
Understand and apply basic concepts of probability

(View a full text of the Principles and Standards for School Mathematics.)

MATERIALS

For each group of three to four students:

- Student Activity, "A First Look"
- Student Text, "Exploring Data"
- Data printout from one week of solar wind readings from the LANL Web site
- Data printouts for December 19-22, 2002, from the LANL Web site

Alternate Strategy Tip

When printing the data from the LANL Genesis Web site, it is best to make single-sided copies so that students can see the entire week's data by lining up the pages.

For materials management, provide a file folder that contains the printed sheets for each group. Have the pages numbered ahead of time in order to keep track of any pages that might need to be replaced.

PROCEDURE

1. Explain to students that they are going to be exploring actual solar wind readings from the Genesis spacecraft for both a one-day and a one-week time period. Tell them that the purpose of this activity is to become familiar with readouts, and to start looking for patterns and asking questions. Distribute the Student Activity, "A First Look," and the printouts of the solar wind readouts to each group of four students.

2. In part one, students observe a solar wind summary data plot information for one day. Allow student groups to observe the solar wind summary plots and fill in the answers to the questions on the student activity sheet. Circulate around the room and offer assistance when needed. In question 1, students need to indicate the regime that is present on that day. In order to understand what solar wind regimes are, they should have read the Student Text, "Solar Wind," from the "Briefing" section. For questions 3, 7, and 10, students should describe the trends. In some of the data sets, there may be more than one answer on a particular day. Have students do the best they can in verbally describing the trends they see. If your students are having difficulties with question 4, explain that quantitative observations involve numeric data, and qualitative observations involve verbal descriptions of data trends. For question 5, a possible benefit of quantitative observations is that it is precise information that more completely describes an object or an event. A benefit of qualitative observations is that trends can be more easily visualized with the use of word descriptions. A limitation of quantitative observations is that there may be an error in the accuracy of the numeric information. A limitation of qualitative observations is that it does not provide precise information about the object or event. For question 11, students should conclude that the higher ratio number should have the higher helium nuclei to hydrogen nuclei ratio.

3. In part two, students observe solar wind summary data plot information for at least one week. In the first chart, students identify the date and time of such changes. In the second chart, students note any patterns or anomalies that may exist during these changes. Using these observations, students develop questions and inferences that may be studied in more detail in the next phase of this module.

4. Once students have completed this student activity, distribute the Student Text, "Exploring Data," in which there is detailed information about how to develop a testable question, how to make interpolated and extrapolated predictions using data, and the need for careful data analysis during conclusion phase of a scientific investigation. Students will also need to have the solar wind summary data for December 19–22, 2002. They will need these data summaries for examples of data trends and making interpolated and extrapolated predictions.

5. The data trends in the second paragraph of the student text show examples of data relationships graphically. You may need to provide additional examples of trends in direct relationships, indirect relationships, and null relationships. A good example of a direct relationship is illustrated when you drive a car. As you increase the pressure on the gas pedal, the speed of a car increases. An example of an indirect relationship is illustrated when you eat a meal. As the amount of food eaten increases, the hunger that you experience decreases. An example of a null relationship might be found in studying for a test. At a certain point, the amount of time you study increases, but there is no effect on the score of the exam. Ask students to provide additional examples of these qualitative data trends.

6. For the paragraphs about extrapolated and interpolated predictions, ask students to write their readings and predictions in a lab notebook or on a piece of paper. You may want to complete this as a large group and have volunteers provide readings and predictions. Write down the responses on the board. For the last question in these paragraphs, students are asked why they think interpolated predictions are usually more accurate than extrapolated predictions. Students might suggest that extrapolated predictions are outside of what is known. Students might suggest that interpolated predictions are especially more accurate when there are clear trends in the data. For example, if there is a direct relationship shown graphically, the interpolated prediction is often on the regression line. However, this is not always the case. Sometimes, there are readings that do not fall where we would expect them. Use the graphics on the student text to make this point.

Alternate Strategy Tip

Ask students to find examples of each of these trends, and extend the experience to have them quantify and provide graphs of their examples.