Modeling Data: Teacher Lesson Plan

By Matt Laird and Heather Murphy

Last Modified July 2019

Subject(s):
STEM
Physics
Pre-Calculus Analysis
Statistics
Computer Science
Earth and Space

Grade(s):
9th to 12th grade students.

Time:
Two or three 55 minute class periods, depending on technical proficiency.

Standard(s): *These performance expectations have been directly taken from the Next Generation Science Standards website at* https://www.nextgenscience.org/*.*

NGSS Standards

HS-PS4-3: Evaluate the claims, evidence, and the reasoning behind the idea that electromagnetic radiation can be described either by a wave model or a particle model, and that for some situations one model is more useful than the other.

HS-PS4-5: Communicate technical information about how some technological devices use the principles of wave behavior and wave interactions with matter to transmit the capture information and energy.

Michigan Math Standards

S-ID-1-4: Summarize, represent, and interpret data on a single count or measurement variable.

S-ID-5-6: Summarize, represent, and interpret data on two categorical and quantitative variables.

S-IC-1-2: Understand and evaluate random processes underlying statistical experiments.

Connections to Prior Learning:
● Atomic Structure
  ● Observing Radiation (See Multiwavelength Astronomy: Lesson Plan)
  ● Cosmic Gamma Rays (See Cosmic Gamma Rays: Lesson Plan)
● Basic Astronomy
  ● Star life cycle (See Star Evolution and Gamma Ray Sources: Lesson Plan)
  ● Pulsar, Blazar, ect. (See Gamma Ray Sources: Lesson Plan)
● Basic Electromagnetic spectrum (See Multiwavelength Astronomy: Lesson Plan)
● Graphs/Graphing
● HAWC Detector (See Cosmic Gamma Rays: Lesson Plan)

●   Spread-Sheet software

Concept/Topic to Teach:
●   Modeling real world data.
●   Statistical data analysis and evaluation.
●   Sources of electromagnetic radiation.

Specific Objectives: Student will be able to…
●   Download real world Gamma-Ray Astronomy data and upload into a Spread-sheet program like Google Sheets.
●   Plot and analyze data, then critique the findings.
●   Model and Interpret data to evaluate a specific question related to Gamma-Ray Astronomy.
●   Illustrate and present data in a professional manner.

Lesson Relevance:
●   Analyze, Model, Interpret, and Present the data.
●   Connecting of abstract concepts (Gamma-rays from celestial objects) to the individual student

Cross Curriculum Connections:
●   Physics: Modeling real world Electromagnetic particle Observation.
●   Math: Statistical analysis and graphing.
●   Astronomy: HAWC Data analysis of celestial objects.
●   Engineering: Statistical analysis and graphing. Problem solving real world issues.
●   Statistics: Statistical analysis and graphing.
●   Computer Science: Data analysis software.

Required Materials:
●   A computer with an internet connection
●   Text Editor software or Cloud Sharing access (Google Drive, ect.)
●   Spread-sheet software (Google Sheets, MS Excel, ect.)
●   Scoring Rubric
●   Websites:
■   https://www.hawc-observatory.org/
■   http://www.csgnetwork.com/julianmodifdateconv.html

Technology:
●   Graphing software
●   Implementing and manipulating Spread-sheet software

Modification/Accommodation*(ELL/IEP students)***:**

●   Printing out the list of abbreviations associated with this lesson

- Scaffolding of the process as the students follow along.
- Provide the data downloaded and modified.
- Modify the data so that the data includes a column that calculates the current date from the Modified Julian Date (MJD).
- Reducing the focus or the required work to only one dataset and graphs.
- Additional time for the assignment.
- Guided notes/Printed Google Slides Presentation.
- Reduction of Assignment (choose one EM spectra).
- Provide list of useful Vocabulary.
- Pairing with strong anchor student for project work.

Reteach/Extensions*(struggling/advanced students)***:**

- Reteach: Basic cell manipulation in Google Sheets.
- Reteach: Assisting students with function writing via the insert > function toolbar.
- Extension: Additional styles of graphs with appropriate interpretations within the conclusion section.
- Extension: Report that includes in depth research into the gamma-ray source celestial objects included within this lesson.
- Extension: Identify isolated events where one object was more active (higher flux) than the others.

Instructional Procedure:
Heading: Modeling the Data

Goal of the Day:
- Student will be able to model and interpret HAWC light curves from three celestial sources (Crab Nebula, Blazar Mrk 421, Blazar Mrk 501).
- Students will be able to compare Gamma-ray flux emissions from three sources on their birthday.
- Students will be able to present data in a professional manner.

Background Information *(Lesson Introduction)***:**
There are several celestial objects that produce Gamma-rays. Neutron stars, white dwarfs, nebulas, and black holes. Many of these sources can be further delineated into more descriptive names, describing their orientation or their interactions with neighboring celestial objects.

A Pulsar is a type of Neutron star. Pulsars emit radiation that is not oriented parallel with the Neutron stars magnetic field. The result is that the emitted radiation is observed periodically, similar to the light cast from a lighthouse. Similarly, a Blazar is a type of black hole found in the center of a galaxy that is oriented so that the radiation emitted from its accretion disc is observable from Earth.

Two of the celestial objects included in this lesson are Markarian 421 (Mrk 421) and Markarian 501 (Mrk 501), both are blazars. Mrk 421 is located in the constellation Ursa Major, and Mrk 501 is the brightest gamma ray object in the night sky. (Protheroe. et.al, 1997).

A supernova is a nebula that is essentially a cloud of gas that is the remnants of a dying star. The Crab Nebula located in the constellation Taurus, is a supernova that has a spinning neutron star at its center. According to NASA the neutron star in the Crab Nebula rotates at 30 revolutions/pulses per second.

Electromagnetic radiation can be described either by a wave model or a particle model. The wave model was accepted after the Young's Interference Experiment of 1801. The experiment consisted of passing light through a series of slits so that the light demonstrates phase relationships of constructive or destructive interference. Constructive and destructive interference are products of wave properties, notably wavelength and amplitude. Wave properties, such as frequency and wavelength, are currently applied to define regions of the electromagnetic spectrum (Radio, microwaves, infrared, visible, ultraviolet, x-ray, and gamma-ray).

Electromagnetic radiation is also modeled as a particle called a photon, a mass-less package of energy that has momentum. Gamma-ray astrophysicists uses the particle model of gamma-rays because the properties of gamma-rays allow them to be observed at high altitudes within the Earth's atmosphere.

Gamma-rays are electromagnetic waves that have a high energy and frequency but also a short wavelength. Gamma-rays are do not have a charge so they are not affected by electric or magnetic field. Gamma-rays are also small enough particles that they can be considered massless. These properties of gamma-rays allow them to travel in straight lines from the source like a particle. The particle model of the gamma-ray interact with particles within Earth's atmosphere producing a shower of particles with momentum and energy.

A group of scientists have developed a High Altitude Water Cherenkov Gamma-Ray Observatory (HAWC) in Parque Nacional Pico de Orizaba, Mexico. The HAWC observatory is designed to capture the shower effect of the gamma-rays entering the atmosphere as flux. The detectors measure electron volts (eV), counting the number of particles per time and area. This information can be analyzed to indicate the intensity, but also the direction of the gamma-ray emission.

The data collected has been analyzed by scientists in the HAWC group to remove background interference due to cosmic microwave background. Much like the sun would create light interference if someone was trying to observe the moon during the day time vs at night. Depending on a number of factors, there can be varying amounts of interference detected by the observatory that are not related to the source of the gamma-rays. The flux (Flux+) is then converted to Crab Units (CU). The CU is a unit of measure based on the gamma-ray flux density of the Crab Nebula. Hence, the Crab Nebula will have an average flux of one CU.

HAWC has made three sets of light curve data publicly available on their website. These are data sets that list the flux or light intensity vs time. Within this data set time is measured in Modified Julian Days (MJD), a common astronomical unit of time.

Error (-err) has been calculated for each independent flux measurement using a forward-fold likelihood curve, similar to one standard deviation analysis of a gaussian curve.

Transits are the percentage of available time in which the detector was acquiring data. A transit of 1 would indicate that the detector was acquiring data 100% of the available time. An analogy, would be the sun during the daylight hours of one day. If the sun was visible from dawn to dusk, during the daylight hours it would have a transit value of 1 (100%). If during the daylight hours there was a storm that prevented the sun from being visible then the transit value would

be less than 1 (less than 100%). It is suggested that a transit of 0.80 or greater, indicating that the detector was acquiring data for 80% or more time, is acceptable.

TS is an indicator of the likelihood the detected flux is not due to background interference, the larger the TS the less likely the flux is due to background interference.

A Histogram is used to identify outlier information and trends within the data. One should anticipate a bell shaped curve. Values that do not correlate with the bell shaped trend are considered Outliers. Outliers are possible flux values of interest.

Standard Deviation is calculated to indicate the extent of deviation for a group, a lower standard deviation indicates a constant state of gamma ray flux.

Procedure/Strategies:
The following procedure was written for Google Sheets program because it is free or inexpensive to students and schools. This lesson could be easily accommodated for other spread-sheet software such as Excel or Calc.

Guided Inquiry Modeling Activity: Elaborate
Guided Question: On your birthdate was there a significant gamma ray emission?

Step 1: Downloading
- Go to the HAWC Observatory website (https://www.hawc-observatory.org/)
- Go to 'Light Curves" in Public Data Sets and download data set with link on bottom of screen
- Save file in Google Drive. Open extract file. There will be four .dat files. A readme, and file for the Crab Nebula, a file for Makarian 421 and a file for Makarian 501.
- Rename file to .txt, this can be done by adding a .txt to the pre-existing file name. Google sheets requires imported files to be in the .txt format.
- Open each file with text editor (notepad). Select the find and replace function and replace all double spaces with single spaces. (Note: Double space is hitting the spacebar twice, and single space is hitting the spacebar once.). Save changes.
- Repeat for each of the data files.

Step 2: Import the Data: (How to communicate to the computer the fields for importing into excel/sheets)
- Open new google sheet. Name appropriately.
- Go to file then import, then upload. Select the HAWC_fluxlc_Crab_2014-11-26_2016-06-02_1days_si2.63_c010000000_int1.00TeV_mintransits0.50.dat.txt data file.
  - Import location set to Insert new sheet(s)
  - Separator type set to custom (put in a single space)
  - Convert text to number, dates, and formulas set to Yes
  - Select Import data
- Fix row 5 so that the headings are listed correctly.
  - A should read MJD Start (*Modified Julian Date*)
  - B should read MJD Stop (*Modified Julian Date*)

- C should read Flux + [SI] (*HAWC Reading in photons/s*cm^2*)
- D should read -err [SI] (*error in photons/s*cm^2*)
- E should read Flux + [CU] (*HAWC Reading in Crab Units (Normalized to Crab Nebula)*)
- F should read -err [CU] (*error in Crab Units*)
- G should read Transits (*% of observation time (quality of data set)*)
- H should read Index (*Unnecessary*)
- I should read Cut-off (*Unnecessary*)
- J should read TS (bkg) (*Higher value means higher significance over background noise*)

At this point your spreadsheet should look something like this:



Step 3: What does the data mean: Read me file

The Readme file contains information on the dates that the data was acquired. Notable information also includes : the length of time for an average transit time, location that the transit was taken, and the units within the data file.

Conduct a group discussion concerning the units within the spreadsheet: the Modified Julian Date (MJD), Crab units (CU), and transits specifically. Additional time should be taken to discuss the wave-particle duality of gamma-ray emissions and how the HAWC observatory measures the particle properties of gamma-rays in the unit of Teraelectrovolts (TeV).

Step 4: Create a filter to only include relevant data. If then Statement (remove low Transits (below 80%)

- ● The Transits (column G): The transit value is the percentage of the transit that data was available. So values below 80% (0.80) are not desirable, because the detector was not collecting information for over 20% of the time.
- ● In Column K write an If then statement to remove data that has a transit value below 80%.
  - ■ =IF(G6>0.8,E6," ").
  - ■ Click bold box in the lower right corner of cell to drag the formula down to the end of the data (K477 in the Crab nebula data sheet).
  - ■ Label K5: Flux [CU] (Transits>0.8)
- ● At the bottom of column K (K478) Calculate the average by typing =AVERAGE(K6:K477) into cell K478. In J478 type the label AVERAGE.
  - ■ The average for the Crab Nebula should be 1.025 CU. It is anticipated to be around 1 because the Crab Units (CU) are standardized off the Crab Nebula.
- ● Below the Average calculate the Standard Deviation (K479). Calculate the Standard Deviation by typing =STDEV(K6:K477) into cell K479. In J479 type the label StadDev.
  - ■ The StadDev for the Crab Nebula should be 0.295. This is relatively small due to the Crab Nebula's constant state of emission.
- ● Highlight the data in column K. This is done by clicking the column K at the top of the spreadsheet. Go to Insert then select chart.
  - ■ Multiple chart types will be interesting. Select a chart of your preference. It is recommended to select a column, line, or scatter plot chart.
  - ■ Right click on the chart and select Chart & axis title. Make sure the appropriate axises are labeled.

At this point your spreadsheet should look something like this:
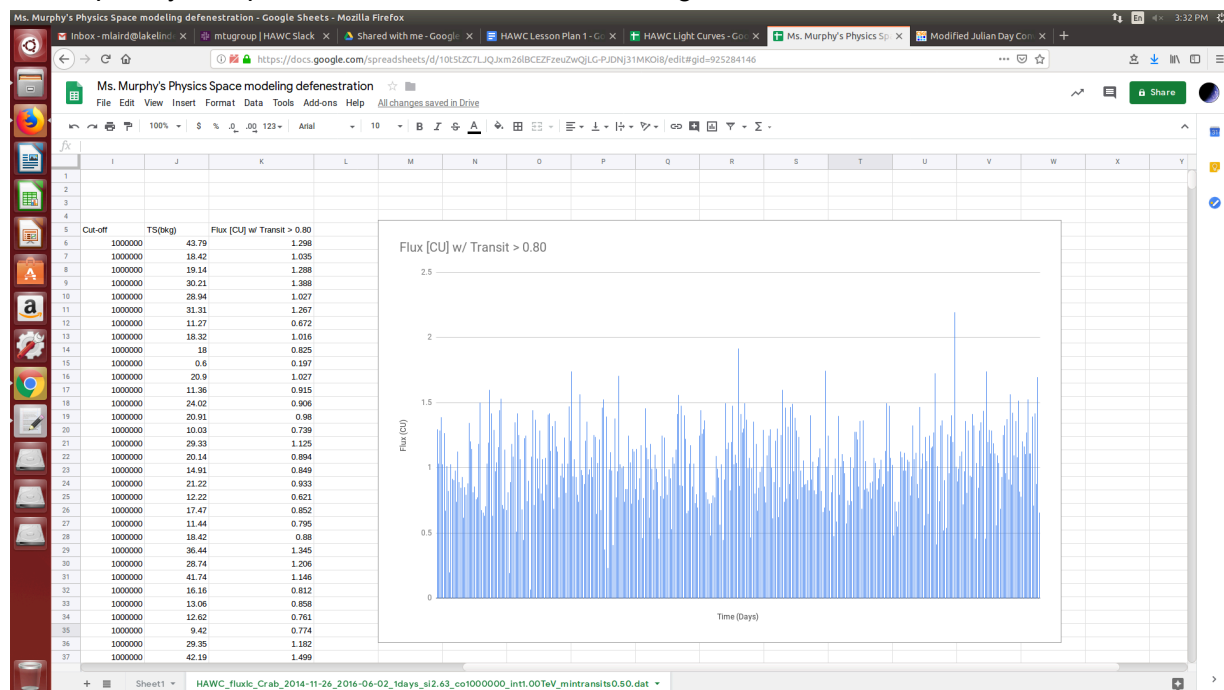
At this point students should record their observations of the graph.

After recording their observations students should compile as many questions as possible related to the data. The more questions generated the better. This should take no less than 5 minutes.

Step 5: Create a Chart (Histogram) To analyze the relevance of the data.
- Highlight the data in column K. This is done by clicking the column K at the top of the spreadsheet. Go to Insert then select chart.
  - Select Histogram chart.
  - Right click on the chart and select Chart & axis title. Make sure the appropriate axises are labeled.

At this point your spreadsheet should look something like this:
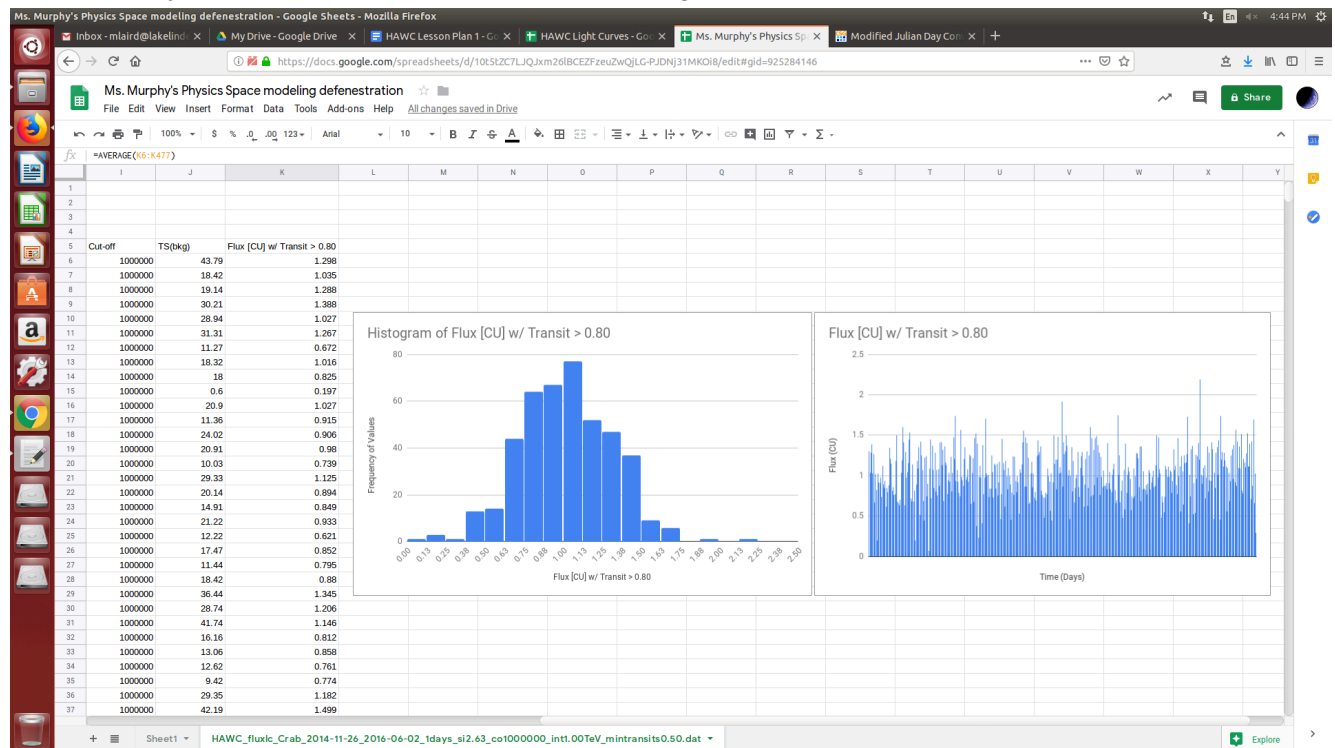
A Histogram is used to identify outlier information and trends within the data. One should anticipate a bell shaped curve. Values that do not correlate with the bell shaped trend are considered Outliers. Outliers are possible flux values of interest.
The Crab Nebula is considered to be an example of constant state data.
- Repeat steps 2 to 5 for Blazar Mrk 421, Blazar Mrk 501

Have students compare and compile as many observations and then questions as related to the data and graphs. The questions should be collected and evaluated. Questions are

evaluated for the quantity of questions each student presented and the quality or level of the questions. The levels of the questions are classified as low order of thinking (LOT) or higher order of thinking (HOT) level questions based on the content of the question. LOT questions are questions that have one simple answer, that could be answered with minimum research. These are questions answered by reading the introduction for example. HOT questions are questions that could be answered through further experimentation. These questions required the student to put ideas together to construct a new experiment or research ideas.

Students should now discuss and evaluate their findings with their classmates and hypothesize reasons for discrepancy.

Step 6: Select a set of data of interest: Convert the modified Julian Calendar to Modern Gregorian Calendar so that the students can select their birthdate.

- In a separate window go to the web site http://www.csgnetwork.com/julianmodifdateconv.html This web site will allow you to convert the Modified Julian Calendar used in the data set to the Modern Gregorian Calendar familiar to your students. Students should select three days before their birth day in the year 2015, their birthday, and three days after their birthday. There should be a total of seven days selected.
- Select the appropriate days for each data set Crab, Mrk421, and Mrk501. Copy the flux+ [CU] and its -err [CU] into a new sheet. Column E and F respectively in the data set. Remember to label the data as it is pasted into a new sheet. The -err number represents the error of each measurement to one standard deviation. This are calculated individually based on Forward-Folded Likelihood Analysis curves for the measurement. This is important in that it indicates how accurate the single measurement is. In other words, the actual measurement can be found anywhere within the error bars created in this step technically.
- To create a chart with error bars the data needs to be rearranged. First establish a data table with the x-axis labeled 1 to 7, and the y-axis labeled 1 to 7. Then separate the first flux data into the 1:1 location the second flux data into 2:2 and continue for each of the seven data points.

At this point your spreadsheet should look something like this:
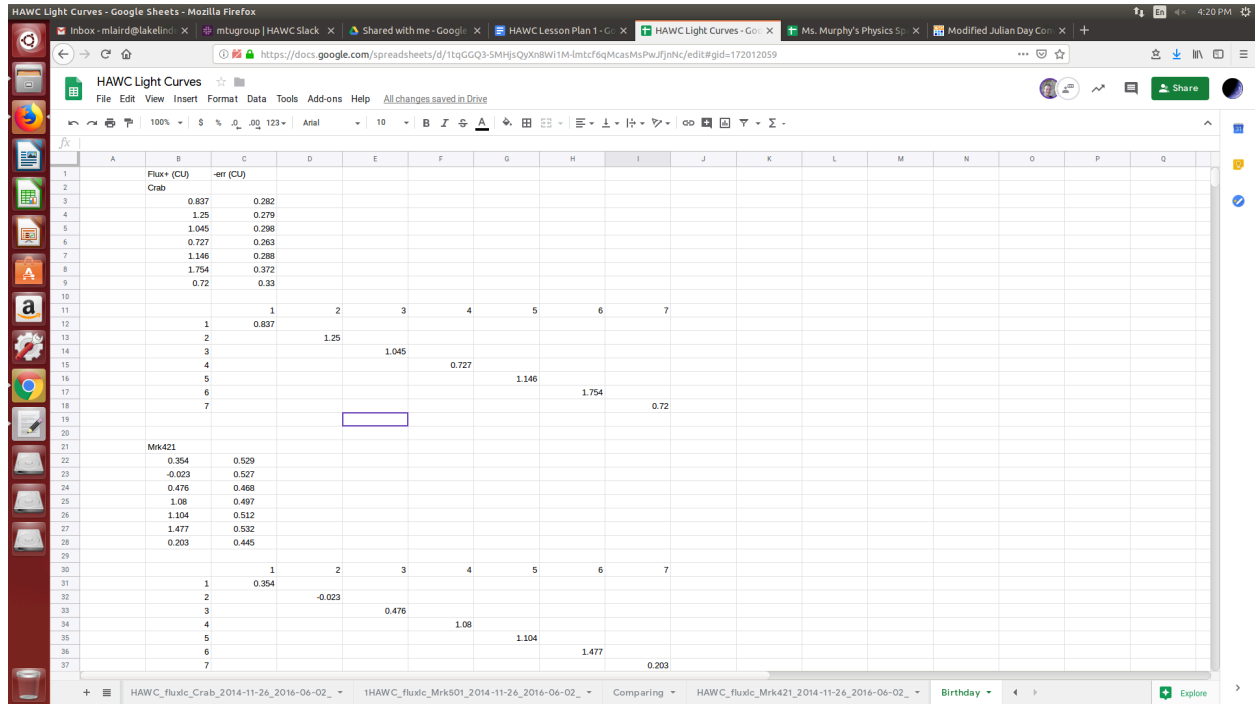
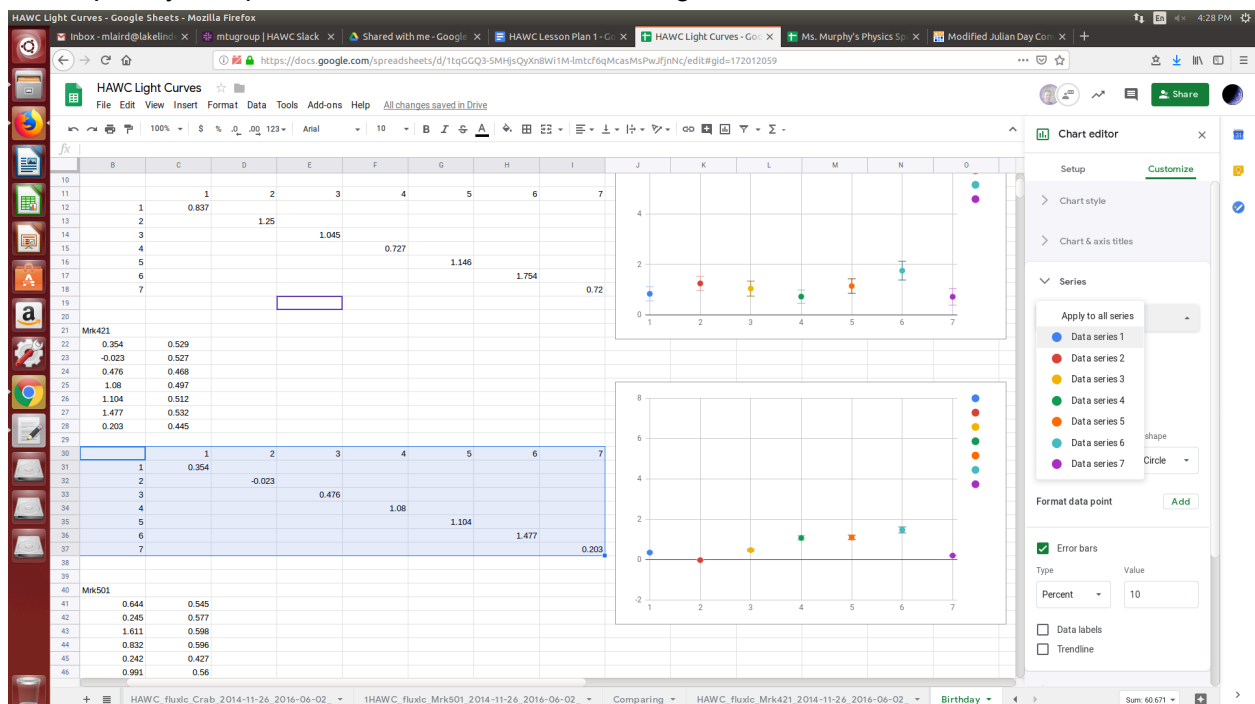- Highlight the the seven by seven data just created. Go to Insert then select chart.
    - Select Scatter Chart.
    - In Chart editor, select Customize then series, then check error bars.

At this point your spreadsheet should look something like this:
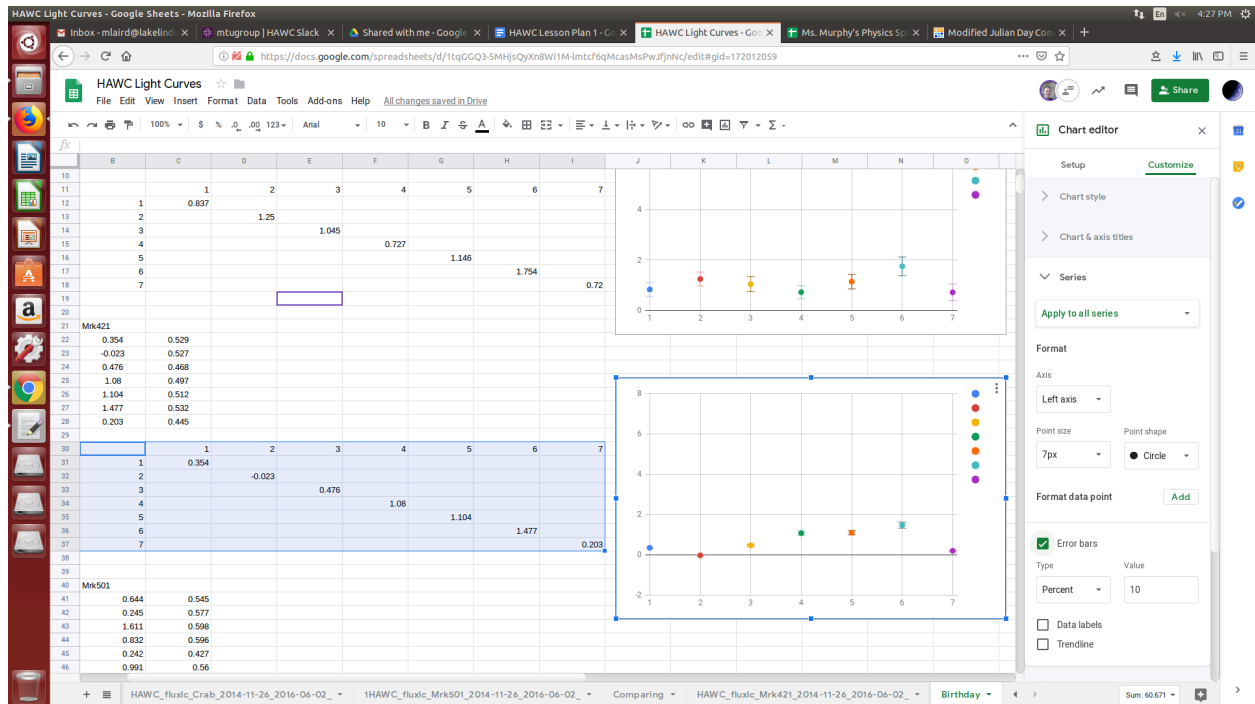
○   Under series selector drop down window change to data series 1.

At this point your spreadsheet should look something like this:

○   Under Error bars change dropdown window to Constant and type the error value for that flux value in the Value box.

At this point your spreadsheet should look something like this:
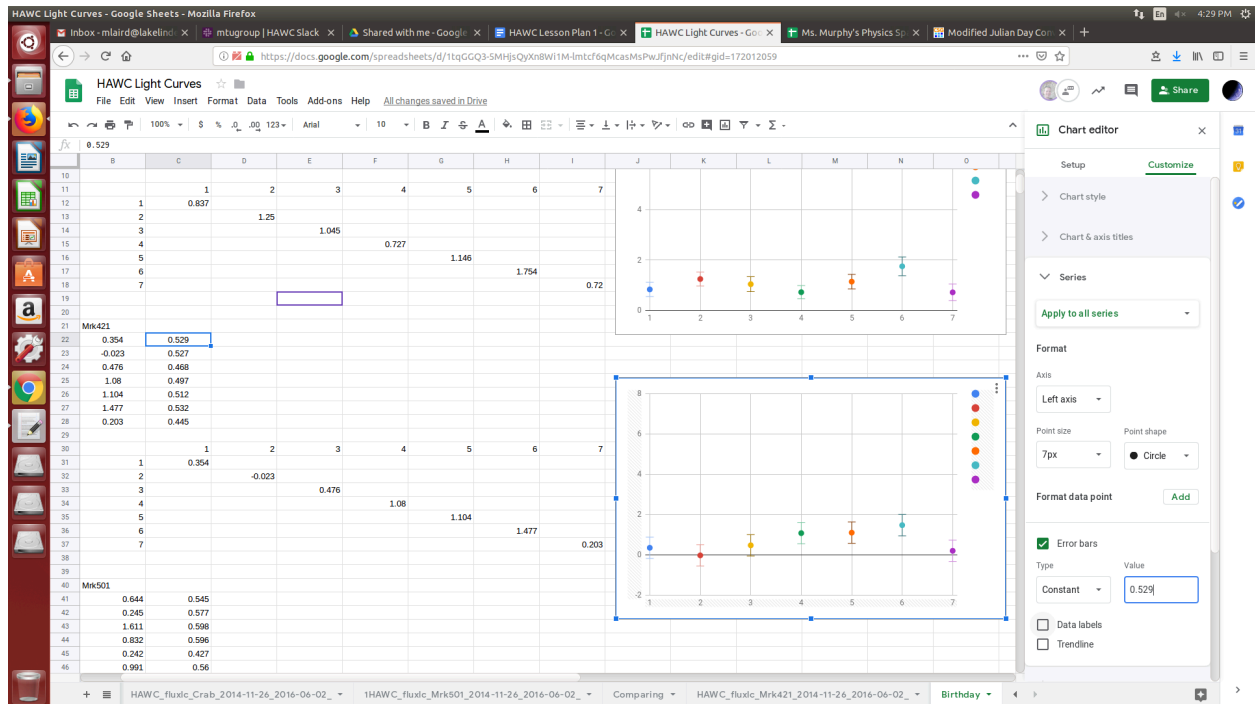
  ○   This process will need to be repeated for every data point for all three data sets.

  ○   Right click on the chart and select Chart & axis title. Make sure the appropriate axises are labeled.

Step 7: Interpret the data

   Guided Inquiry Modeling Activity
Guided Question: On your birthdate was there a significant gamma ray emission?

Remind the student the impact of error on their data when they are drawing their conclusion. Comparing the three flux graphs (Crab, Mrk421, Mrk501): Look for spikes that would indicate that a celestial object had more activity.

Step 8: Publish their results.
   Students should prepare a publication to communicate their results. Common publications could be power-point (Google Slide) presentations, laboratory reports, or posters. Attached are scoring rubrics and student handouts that may be modified for the requirement of the course.

Key Vocabulary and Academic Vocabulary:
● Crab Units (CU): "The standard specifications of flux density is a function of the spectrum of the source (relative to that of the Crab)." (NASA, 2019)
● Flux: "The flux is the amount of energy reaching each square centimeter of a detector (e.g., your eye, CCD, piece of the sphere) every second." (Properties of Stars, 2019)

● High Altitude Water Cherenkov Gamma-Ray Observatory (HAWC): An observatory in Parque Nacional Pico de Orizaba, Mexico, that utilizes Milagro detectors to observe uniformly-instrumented gamma-ray air showers. (HAWC, 2019)

● Histogram: "A graphical display where the data is grouped into ranges (such as "100 to 149", "150 to 199", etc), and then plotted as bars. The height of each bar shows how many are in each range". (Math is fun, 2019)

● Light Curve: "A light curve is a graph of light intensity of a celestial object or region, as a function of time." (Wikipedia, 2019)

● Modified Julian Date: "A modified version of the Julian date denoted MJD obtained by subtracting 2,400,000.5 days from the Julian date (JD). The MJD therefore gives the number of days since midnight on November 17, 1858." (Scienceworld.wolfram.com, 2019)

● Error bars: "graphical representations of the variability of data and used on graphs to indicate the error or uncertainty in a reported measurement." (Wikipedia, 2019)

Assessment: Summative:

Students will be assessed on the quantity and quality of questions generated, participation during the lesson, and publication of their results. Attached are scoring rubrics and student handouts that may be modified for the requirement of the course.

Acknowledgements

This lesson was created through the support of Michigan Technological University's Physics Department. The Physics Department sponsored a *Research Experience for Teachers*, which allowed the author to partner with scientists and post-secondary students working in the Michigan Tech group of the larger HAWC (High-Altitude Water Cherenkov Gamma-Ray Observatory) collaboration on gamma-ray astrophysics.

Resources:

Astronomy.com. Blazars explained. (2018). Accessed July 03, 2019.
http://www.astronomy.com/news/2018/07/what-is-a-blazar

HAWC. (2019). Accessed July 03, 2019. https://www.hawc-observatory.org/observatory/

Michigan Math Standards. (2010). Accessed July 03, 2019.
https://www.michigan.gov/documents/mde/K-12_MI_Math_Standards_REV_470033_7_550413_7.pdf

Math is fun. Histogram. (2019). Accessed July 03, 2019.
https://www.mathsisfun.com/definitions/histogram.html

Michigan/Next Generation Science. (2011). Accessed July 03, 2019.
https://www.nextgenscience.org/michigan

NASA. Messier 1 (The Crab Nebula). (2017). Accessed July 03, 2019.
https://www.nasa.gov/feature/goddard/2017/messier-1-the-crab-nebula

NASA. What is a Nebula?. (2019). Accessed July 03, 2019.
https://spaceplace.nasa.gov/nebula/en/

NASA. (2019).
https://heasarc.nasa.gov/docs/heasarc/ofwg/docs/general/ogip_93_001/ogip_93_001.html

OpenStaxCollege. College Physics. OpenStax College Physics and Rice University. Creative
Commons Attribution 4.0 International License. January 23, 2012.
https://opentextbc.ca/physicstestbook2/

Properties of Stars. (2019). Accessed July 03, 2019.
https://www.astronomynotes.com/starprop/s3.htm

Protheroe, Ray J.; C.L. Bhat; P. Fleury; E. Lorenz; M. Teshima; T.C. Weekes (12 October 1997).
"Very high energy gamma rays from Markarian 501" (PDF). arXiv:astro-pn/9710118v1.

Scienceworld.wolfram.com. (2019). Modified Julian Date -- from Eric Weisstein's World of
Astronomy. Accessed July 03, 2019
http://scienceworld.wolfram.com/astronomy/ModifiedJulianDate.html

Wikipedia. Light curve. (2019). Accessed July 03, 2019.
https://en.wikipedia.org/wiki/Light_curve

Wikipedia. Markarian 421. (2019. Accessed July 03, 2019.
https://en.wikipedia.org/wiki/Markarian_421