UNIVERSITY OF MINNESOTA
MORRIS

Data Processing and Curve Fitting of Counts Data

ST. CATHERINE UNIVERSITY

Liam Taylor, Gordon McIntosh, and Alaina Swanson
Division of Science and Mathematics, University of Minnesota, Morris, MN, 56267

Erick Agrimson and Kaye Smith
Department of Physics, St. Catherine University, St. Paul, MN 55105

Abstract

Various data analysis methods were explored to more accurately and consistently determine the Regener-Pfotzer (RP)[1] maxima for high altitude cosmic ray intensity. The radiation has been measured during 15 balloon flights using Geiger-Mueller (GM) counters with five second accumulation times. Of the 15 flights, 10 of them included omnidirectional counts data, and 8 of them included vertical coincidence counts data. Count data from altitudes greater than 10 km were analyzed to determine the maxima. The data analysis methods used were moving average filtering and summation of GM counts into one minute intervals. Moving average filtering did not give reliable results, so the summation method was chosen. Once the data were summed, several different curves were fit to determine the RP maximum. The curves tested include second and third order polynomials as well as cubic spline interpolation of the data averaged over 1 km intervals. Second order polynomial fitting did not fit the data well. Third order polynomials and cubic splines gave better results. Third order polynomial fitting was chosen due to its ease of use and the similarity of the results given by the cubic spline interpolation (within 1%). The omnidirectional RP maxima occurred at an average altitude of 21.8 km ± 1.7 km, while the vertical coincidence RP maxima occurred at an average altitude of 18.5 km ± 1.1 km.

Background

Balloon payloads launched by the University of Minnesota, Morris and St. Catherine University are instrumented with GM counters, temperature, and pressure sensors to determine the RP maxima. Once the data are collected, it is often noisy and difficult to analyze, so data smoothing is necessary. Additionally, once the data are processed, the RP maximum is not obvious. Several methodologies of curve fitting and data smoothing were developed to mitigate these issues. These methods were developed in anticipation for the Total Solar Eclipse on August 21st, 2017.

Data Processing and Curve Fitting

The first method of data processing used was a moving average filter. This methodology was used by Harrison et al. [2] for their high altitude coincidence counts data. However, they use a miniature GM tubes as a part of their payload, while the University of Minnesota, Morris and St. Catherine University both use AWARE Electronics Model RM-60 and RM-80 GM counters The RM-60 and RM-80 GM counters have much larger collection areas than a miniature GM tube. As such, a moving average filter is not an applicable method of data smoothing in this case.

The next method of data processing was reducing the counts into counts per minute in Excel using the equation

SUM(OFFSET(**input cell**,ROWS(**reference cell**:reference cell)*n-n,,n)), (Eqn 1)

where n is the number of cells added together, the input cell is the first cell containing counts data, and the reference cell is the cell containing the equation. All of the bolded cell are static cells. To find the corresponding altitude, we modify equation 1 to be

AVERAGE(OFFSET(**input cell**,ROWS(**reference cell**:reference cell)*n-n,,n)) (Eqn 2).

With this, we were able to generate altitude vs. counts per minute plots.
Once the data were processed, several curves were fitted. Second and Third order polynomial fitting was done using Psi-Plot, and cubic spline interpolation[3] fitting was done using MatLab. Second order polynomials were ineffective at determining the RP maxima, while third order polynomials and cubic spline interpolation better fit the data. Third order polynomial fitting was chosen over cubic spline interpolation due to its ease of use. Cubic spline interpolation required additional data wrangling to be effective, and the RP maxima determined by the curves were within a 1% difference. Once a curve is fit to the data, the RP maximum is determined by finding where the slope of the fitted curve is zero.

		Aug 19th Omnidirectional (1)	Aug 19th Vertical (2)	Aug 21st Omnidirectional (3)	Aug 21st Vertical (4)
2nd order	Chi-Squared	53818.525	10547.329	17312.923	7150.247
	R^2	0.86618	0.66528	0.94183	0.57972
3rd order	Chi-Squared	27399.699	6088.939	11303.304	5207.213
	R^2	0.93187	0.80769	0.962027	0.6939

Table 1: Statistical parameters for the fitted curves in Figure 2.

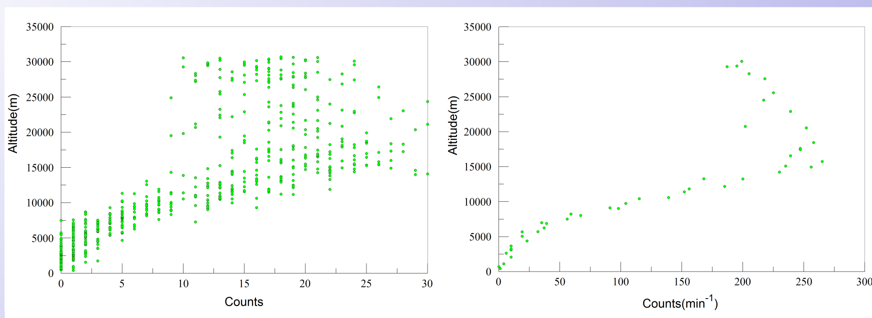

Figure 1: Comparison of raw vertical counts data (left) with the reduced counts per minute data (right). These data were taken from flight UMM 25 (June 11th, 2017).

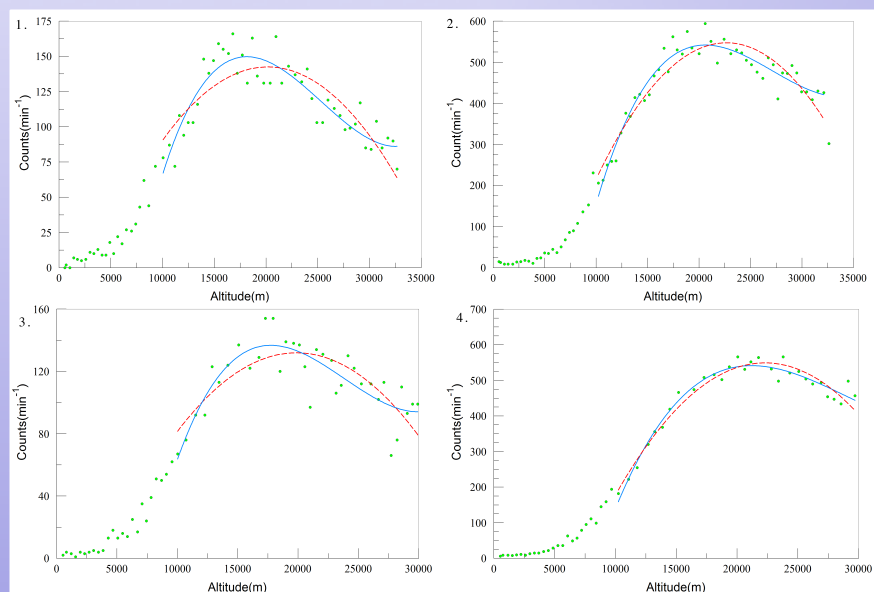

Figure 2: Comparison of fitted curves from several flights. All counts data were reduced into counts per minute. Graphs 1 and 3 contain vertical counts data, and graphs 2 and 4 contain omnidirectional counts data. The data from graphs 1 and 2 were taken on August 19th, 2017 and the data from graphs 3 and 4 were taken on August 21st, 2017. The dashed red curves are second order polynomials, and the solid blue curves are third order polynomials.

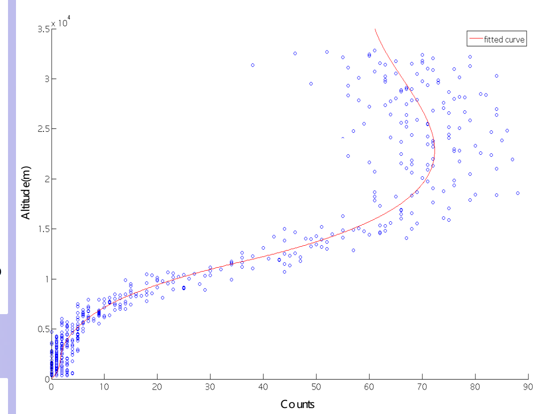

Figure 3: Cubic spline interpolation of data averaged over 1 km intervals. The spline was fit to data taken on April 19th, 2016. The RP maximum given by the spline is 21.6 km. Third order polynomial fitting gives a RP maximum of 21.9 km.

Future Steps

The methods of data reduction and curve fitting have reduced the uncertainty in determining the RP maximum. These methods are not without downsides, however. Moving to higher order polynomials would give better goodness of fit, but such curves make determining an underlying function difficult. Future work includes,

- Exploring different curve fitting to increase goodness of fit.
- Reduce noise
- Developing other methods of determining RP maxima using pressure, and temperature data.

As these methods continue to be developed, they will be applied to a wider range of data sets to analyze different aspects of RP maxima.

References

[1] Carlson, P., and A. A. Watson. "Erich Regener and the Ionisation Maximum of the Atmosphere." *History of Geo- and Space Sciences*, vol. 5, no. 2, 2014, pp. 175–182., doi:10.5194/hgss-5-175-2014.

[2]Harrison, R.G., Nicoll, K.A, Lomas, A.C., 2013. Geiger tube coincidence counter for lower atmosphere radiosonde measurements. *Rev. Sci Instrum*, 84, 076103.

[3]Harrison, R.G., Nicoll, K.A., Aplin, K.L., 2014. Vertical profile measurements of lower troposphere ionisation. *J. Atmos. Sol.-Terr. Phys.* 119, 203-210.