

LA-UR-16-26115

Approved for public release; distribution is unlimited.

Title: DNDO Report: Predicting Solar Modulation Potentials for Modeling
Cosmic Background Radiation

Author(s): Behne, Patrick Alan

Intended for: Report

Issued: 2016-08-08

Disclaimer:

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

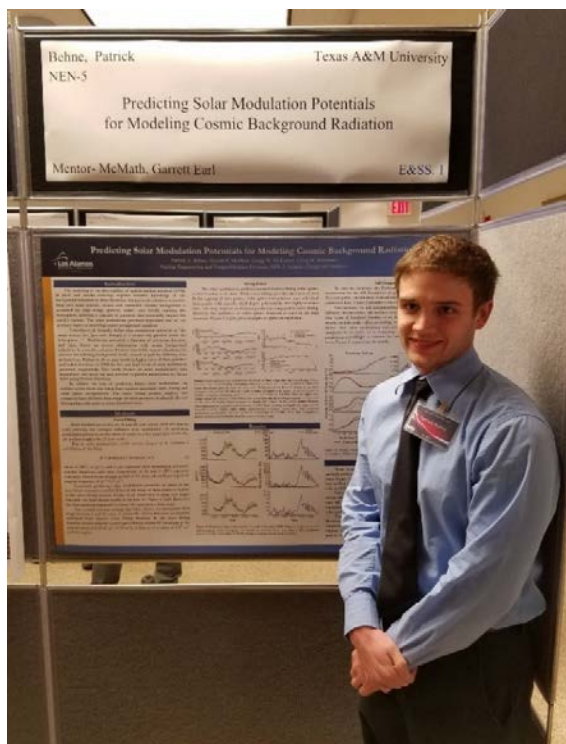
DND0 Report: Predicting Solar Modulation Potentials for Modeling Cosmic Background Radiation

Patrick Behne

Hosting Site: Los Alamos National Laboratory

Mentors: Garrett McMath and Gregg McKinney

Abstract. The modeling of the detectability of special nuclear material (SNM) at ports and border crossings requires accurate knowledge of the background radiation at those locations. Background radiation originates from two main sources, cosmic and terrestrial. Cosmic background is produced by high-energy galactic cosmic rays (GCR) entering the atmosphere inducing a cascade of particles that eventually impact the earth's surface. The solar modulation potential represents one of the primary inputs to modeling cosmic background radiation. Usosokin et al. formally define solar modulation potential as “the mean energy loss [per unit charge] of a cosmic ray particle inside the heliosphere...” Modulation potential, a function of elevation, location, and time, shares an inverse relationship with cosmic background radiation. As a result, radiation detector thresholds require adjustment to account for differing background levels, caused partly by differing solar modulations. Failure to do so can result in higher rates of false positives and failed detection of SNM for low and high levels of solar modulation potential, respectively. This study focuses on solar modulation's time dependence and seeks the best method to predict modulation for future dates using Python. To address the task of predicting future solar modulation, we utilize both non-linear least squares sinusoidal curve fitting and cubic spline interpolation. This material will be published in transactions of the ANS winter meeting of November, 2016 [1].



1. Internship Project

INTRODUCTION

The modeling of the detectability of special nuclear material (SNM) at ports and border crossings requires accurate knowledge of the background radiation at those locations. Background radiation originates from two main sources, cosmic and terrestrial. Cosmic background is produced by high-energy galactic cosmic rays (GCR) entering the atmosphere inducing a cascade of particles that eventually impact the earth's surface. The solar modulation potential represents one of the primary inputs to modeling cosmic background radiation.

Usosokin et al. formally define solar modulation potential as “the mean energy loss [per unit charge] of a cosmic ray particle inside the heliosphere...” [2]. Modulation potential, a function of elevation, location, and time, shares an inverse relationship with cosmic background radiation. As a result, radiation detector background thresholds require adjustment to account for differing background levels, caused in part by differing solar modulations. Failure to do so can result in higher rates of false positives and failed detection of SNM for low and high levels of solar modulation potential, respectively. This study focuses on solar modulation's time dependence and seeks the best method to predict modulation for future dates. The best method will be incorporated into NINESIM, a discrete event simulator. An in-depth description of NINESIM is beyond the scope and purpose of this paper.

DESCRIPTION OF THE ACTUAL WORK

We utilize two approaches to address the task of predicting future solar modulation: curve fitting and interpolation/extrapolation.

Curve Fitting

This approach follows the methodology of Liegey et al [3]. Solar modulation occurs on 11 and 22 year cycles, with the shorter cycle exerting the stronger influence over modulation. In predicting modulation potential on the order of weeks to a few years into the future, we neglect the 22 year cycle.

Due to solar modulation's cyclic nature, Liegey et al. assume a correlation of the form

$$\phi = A \sin(\omega t) + B \cos(\omega t) + C \quad (1)$$

where ϕ (MV), ω (yr^{-1}), and t (yr) represent solar modulation potential, angular frequency, and time,

respectively. A, B, and C (MV) represent constants. Since this equation models the 11 year cycle, an expected value of $\omega = 0.57 \text{ yr}^{-1}$ may be estimated.

The effect of solar modulation on background neutron count rate may be approximated by differentiating McKinney et al.'s linearity equation [4], giving

$$\frac{dC}{d\phi} \cong -230 \frac{\text{counts}}{\text{hr} \cdot \text{MV}} \quad (2)$$

Since Usoskin et al.'s modulation data ranges from 170 MV to 2016 MV, this suggests a background neutron count rate difference of up to 4.2×10^5 counts/hr when comparing background on different dates.

Accurately predicting solar modulation potential on dates in the near future requires a careful choice of the range of data points to include in the curve fitting process. Data over too large a time span results in a poor fit in the region of interest because the local period of measured solar modulation varies from 8.4 to 12.3 years. On the other hand, a curve fitted from data over too small a duration does not give the curve fitting process enough information for an accurate fit. Liegey et al. concluded that using data over the range of a few years longer than the span of the most recent local maximum and minimum (half the local period) results in the best fit.

Instead of following the Fortran algorithm that Liegey et al. developed, this study takes advantage of Python's built-in non-linear least squares curve fitting optimization function to fit a curve to known solar modulation data.

Choosing Data for the Best Fit

At the time of writing, Usoskin et al. provide averaged monthly solar modulation data for the time period 1936 – 2015. Due to the data's averaged nature, we assume that the given solar modulation occurs during the middle of each month.

The developed Python function receives the date of desired solar modulation potential and returns the corresponding modulation. In accordance with Liegey et al., an algorithm seeks information about the local period to determine the range of data to pass into the curve fitting process. If the input date occurs after the latest value in the data set (i.e., in the future), the algorithm seeks the two most recent local extrema in the set. If the input date occurs within the set (i.e., in the past, but between known data points), the algorithm seeks the two closest local extrema on either side of the input date.

If the local extrema seeking algorithm attempts to use the original Usoskin et al. data, it returns meaningless extrema due to noise and data fluctuations caused by the 28 day solar rotation cycle [4]. Methods that coarsen and smooth the data remedy this issue.

Coarsening the data removes some of its fluctuations while retaining its general sinusoidal trend. This study employs two coarsening methods. The first retains every n th data point while discarding all other points. The second method averages, before discarding, all data points into every n th point. The value of n determines the coarseness of the data. For example, if $n = 6$, a coarse point occurs every 6 months. The local extrema seeking algorithm then receives the coarse data and returns coarse local extrema. These extrema represent true extrema more accurately than those given by the original noisy data.

Data smoothing, an alternative to data coarsening, attempts to reduce fluctuations via Python's built-in Savitzky-Golay filter. The filter smooths the data by fitting a third order polynomial to a window size of every m data points. Figure 1 shows examples of the two coarsening methods and the smoothing method.

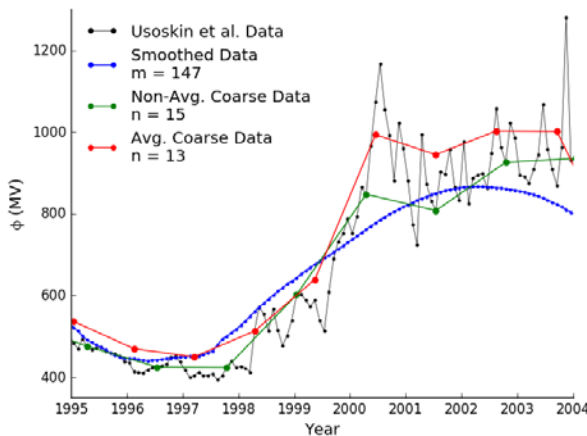


Figure 1 - Comparison of local extrema seeking methods for data taken 1995-2004. The original noisy data gives many false local extrema. Coarsening the data lessens this effect.

The extrema seeking algorithm uses one of these three methods to locate the local maximum and minimum relevant to the input date. To ensure the validity of the located extrema, the function requires a time range between the two extrema, half the local period, of 4 - 10 years. The larger upper bound allows the few additional years to the span of half the local period that results in the best fit. Once the algorithm identifies the relevant local extrema, the function passes the original Usoskin et al. data between these extrema into the curve fitter.

To obtain the best possible fit, the first version of the curve fitting function computes curves for numerous values of n (coarseness) and m (window size). In coarsening computations, n ranges between

6 and 84. The smoothing method sets m values from 71 to 349. The function returns the fit with the highest coefficient of determination (R^2 value).

Figure 2 illustrates how each method works. The legend may be interpreted as follows:

- Blue points:** Original Usoskin et al. data. Points are monthly averages of solar modulation potential.
- Red points:** Result of the coarsening or smoothing process. Passed into the extrema seeking algorithm.
- Green points:** Original Usoskin et al. data between the chosen local extrema. Passed into the curve fitting function.
- Cyan line:** The best curve fitted to the green points.
- Dark violet star:** Point on the fitted curve at the input date, the date of desired modulation potential. The function returns the predicted modulation potential at this point.

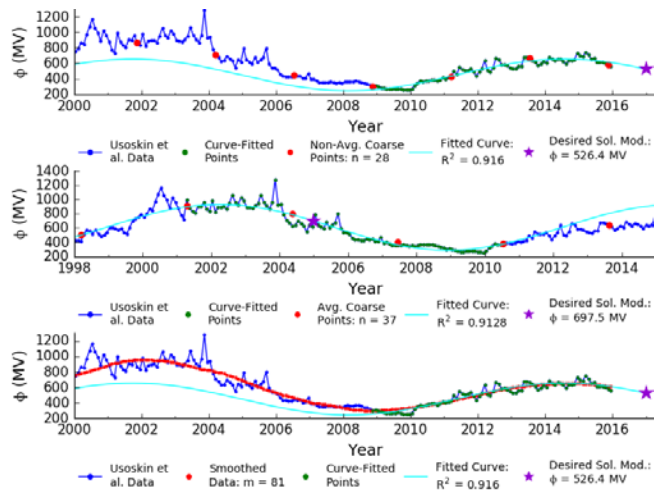


Figure 2 - Top: non-averaging coarsening method to predict modulation on 1/1/2017. Middle: averaging coarsening method to predict modulation on 1/1/2005. The algorithm finds local extrema with respect to the coarse points for these two methods. Bottom: smoothing method to predict modulation on 1/1/2017. The algorithm finds local extrema with respect to the smoothed points.

Note the relatively large values of n in Figure 2. As the value of n increases, the coarse points spread out, reducing the probability that a coarse point occurs precisely at the time of a true local extrema. Recalling that half the local period plus a few years results in the best curve fit, higher n values provide a better fit: although they do not locate extrema as accurately and precisely as smaller n (assuming the smaller n values are large enough to ignore noise and fluctuations), the extrema they do find add those few extra years to half the true local period.

Interpolation/Extrapolation

The other modulation predictor method involves fitting cubic splines to the Usoskin et al. data. While curve fitting provides the curve of best fit for a group of data points, cubic spline interpolation joins individual data points with smooth, third degree polynomials. For highly accurate data, this may improve modulation prediction compared to curve fitting. However, the usefulness of cubic splines decreases as noise in the data increases. Figure 3 provides examples of interpolation/extrapolation.

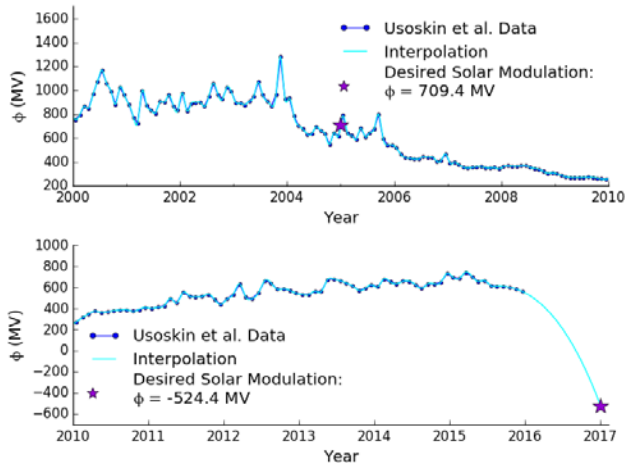


Figure 3 – Cubic spline interpolation to predict modulation. Top: predicted modulation on 1/1/2005. Bottom: predicted modulation on 1/1/2017. Extrapolation poorly predicts modulation for dates far outside the data range.

Function Development

To test the quality of the curves that version 1 generates, a script ran the function for 243 dates in the period 1931 – 2020. Figure 4 summarizes the results.

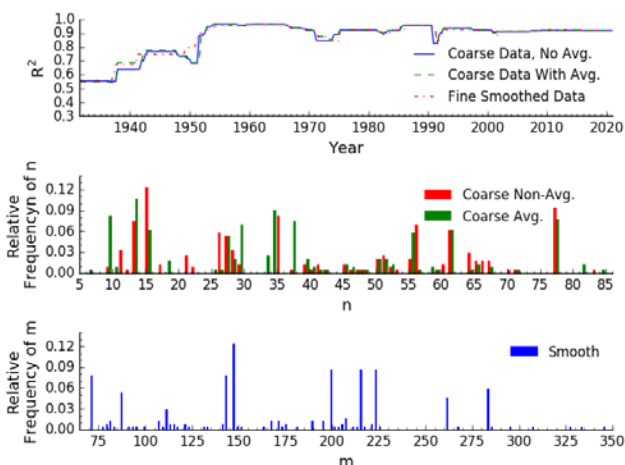


Figure 4 - Testing of version 1 of the curve fitting function. From top to bottom: curve R^2 values, relative frequency of n values (coarse data), and relative frequency of m values (smoothed data).

According to the R^2 values after 1951, the coarsening and smoothing methods seem sound. Usoskin et al. explain that because pre-1951 data is measured differently, it “...is characterized by large

uncertainties and should be taken with caveats.” This explains the higher noise in the earlier data, resulting in lower R^2 values. In addition, the use of another data acquisition method during the period 1951 – 1964 results in uncertain data. From this point, this study excludes data occurring prior to 1965.

Version 2 sacrifices accuracy to decrease computation time. Instead of using $n \in [6, 84]$ and $m \in [71, 349]$, we retained the top 10% most frequently occurring n and m values from the version 1 test, which, when combined, occurred 61.7% (coarse no avg.), 62.6% (coarse avg.), and 89.7% (smooth) of the time.

The same test on version 2 yields an R^2 plot similar in shape to, but slightly less in magnitude than that in version 1. Excluding the 88 pre-1965 dates, versions 1 and 2 return average R^2 values of 0.9186 and 0.9100, respectively. In addition, version 2 runs 6.0 times faster than version 1.

Version 3 seeks to further reduce computation time. If the curve fitting function cannot compute a converged solution within 10^4 iterations (10^6 in versions 1 & 2) or the solution gives $\omega \notin [0.45 \text{ yr}^{-1}, 0.79 \text{ yr}^{-1}]$, it fixes ω to a value of 0.57 yr^{-1} and tries again. On average, version 3 runs 1.3 times faster than version 2.

RESULTS

Short Term Prediction

To test its accuracy, the function predicted solar modulation for Usoskin et al. points using both a past and future mode. Past mode provides the function with all available Usoskin et al. data, minus the point occurring on the prediction date, which occurs within the date range of the data set. Future mode passes only the data occurring prior to the prediction date into the function, with prediction dates occurring a month after the most recent data point. We set the future mode prediction date period to 1975 – 2015 (492 prediction dates) to ensure that future mode did not consider pre-1965 data in the curve fitting process. The past mode prediction date period was set to 1970 – 2010 (480

Stat.		Mode	Coarse Non- Avg.	Coarse Avg.	Smooth	Interp.
Relative Runtime		Past	37.8599	42.4576	101.7844	1.0000
		Future	13.8017	16.344	125.8868	1.0000
R ²	Mean	Past	0.8143	0.8168	0.8400	N/A
		Future	0.7839	0.7811	0.8167	
	Max.	Past	0.9508	0.9508	0.9489	N/A
		Future	0.9492	0.9478	0.9196	
	Min.	Past	0.2035	0.2035	0.6716	N/A
		Future	0.0541	0.1315	0.5794	
% Error (%)	Mean 4.52	Past	9.83	9.81	9.69	5.37
		Future	16.06	17.16	17.57	7.3E4
	Max. 10.20	Past	52.44	45.03	58.05	37.24
		Future	66.35	60.68	59.4	3.2E7
	Min. 1.29	Past	0.04	0.00	0.00	0.00
		Future	0.01	0.01	0.15	0.08

prediction dates) for the same reason, and to ensure that later prediction dates rested between two local extrema. Table I summarizes the resulting R^2 values and percent errors between predictions and Usoskin et al. data.

Long Term Prediction

Figure 5 illustrates the curve of best fit when passing data prior to December 1998 into version 3 of the function.

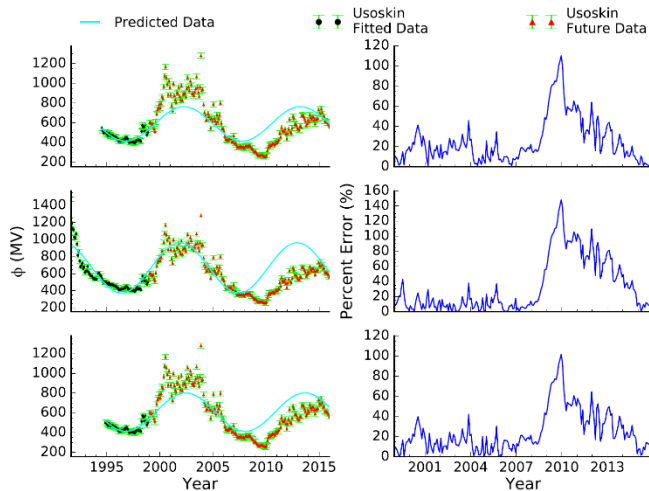


Figure 5 - Predicting solar modulation for dates after November 1998. From top to bottom: coarse non-averaging, coarse averaging, and smoothing methods.

Furthermore, to aid understanding of prediction behavior under various circumstances, we averaged prediction percent error in four types of localized Usoskin et al. data regions. The localized data regions include peak, valley, ascending, and descending data. Thus, if the user knows that solar modulation currently resides at a maximum or minimum in its cycle, or is increasing or decreasing, they may scale predictions accordingly to account for the Python function's systematic error. Figure 6 summarizes the results.

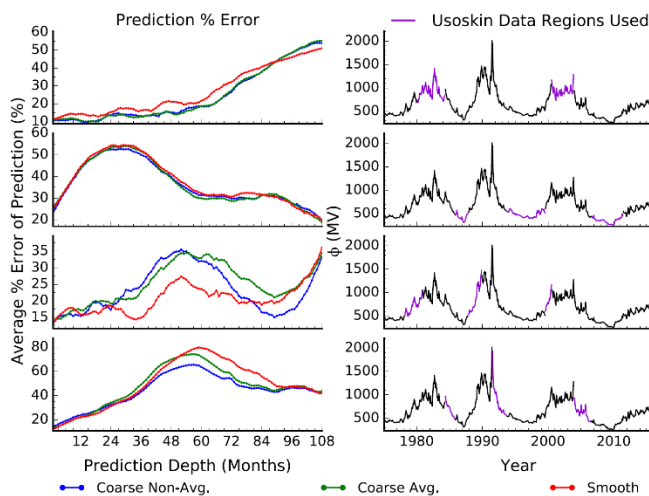


Figure 6 - Average expected percent errors and the Usoskin et al. data regions used in their computation. From top to bottom: peaks, valleys, ascensions, and descensions.

When using version 2 to compute localized data region percent errors, we found consistently higher percent errors than in version 3 – up to 400% higher. Because version 2 allows any value of ω , deeper predictions suffer in accuracy when compared to version 3 because unphysical ω values eventually shift out of phase with data with a different ω value.

These results suggest that, in general, the three curve fitting methods perform comparably, although the smoothing method takes much longer than the coarsening methods. Figure 6 indicates that the best method depends on the prediction depth and where the last recorded solar modulation potential resides in its sinusoidal cycle. For predictions in the very near future (i.e., a month), the user should consider version 2 over version 3.

Cubic spline interpolants give more accurate answers than curve fitting when predicting modulation within the data set. However, the non-sinusoidal nature of the splines render them useless in predicting future solar modulation potentials.

2. Impact of Internship on My Career

The majority of my internship involved developing and debugging Python code to predict future solar modulation. This taught me new Python skills, such as exception handling, and improved previous skills, such as using Matplotlib. In addition, verifying my code's output required me to expand my critical thinking skills. Finally, I attained basic knowledge of MCNP 6.1.1 and increased my previous knowledge of Monte Carlo methods through a week long intermediate course sponsored by the lab.

Although I had many positive experiences this summer, I made a few mistakes which taught me valuable lessons. First, I spent too much time testing my code in ways not as relevant to the code's overall purpose as other tests were. For example, I spent too much time working with data sets other than the Usoskin set for verification and validation purposes, when the Usoskin set accomplishes the same goal. Second, I kept adding features to my code because I was never satisfied with it. As a result of both mistakes, I ran out of time towards the end of my internship.

Prior to beginning my internship, I considered pursuing a career in computation. The computational experience I gained through this internship has convinced me to continue in the field of computational nuclear engineering. As a result, I will most likely attend Texas A&M University for graduate school, which hosts one of the best nuclear engineering computational groups in the country.

3. Acknowledgments

I'd like to thank my mentors, Mr. McMath and Dr. McKinney and my directed studies professor, Dr. Marianno, for aiding with my research. In addition, I'd also like to thank DNDO and LANL for having me in their internship program.

This work has been supported by the U.S. Department of Homeland Security, Domestic Nuclear Detection Office, under competitively awarded contract/IAA HSHQDC-15-X-B0005. This support does not constitute an express or implied endorsement on the part of the Government.

4. References

- [1] P. A. BEHNE, G. E. MCMATH, G. W. MCKINNEY, and C. M. MARIANNO, "Predicting Solar Modulation Potentials for Modeling Cosmic Background Radiation," Proc. ANS Annual Meeting, Las Vegas, NV, November 6-10, 2016 (accepted for publication).
- [2] I. G. USOSKIN, G. A. BAZILEVSKAYA, and G. A. KOVALTSOV, "Solar modulation parameter for cosmic rays since 1936 reconstructed from ground-based neutron monitors and ionization chambers," Journal of Geophysical Research, 116 (2011).
- [3] L. R. LIEGEY, J. R. TUTT, T. A. WILCOX, and G. W. MCKINNEY, "Predicting Future Solar Modulation and Implementation in MCNP6," Proc. ANS Annual Meeting, New Orleans, LA, June 12-16, 2016, Vol. 114, p. 660, American Nuclear Society.
- [4] G. W. MCKINNEY, D. J. LAWRENCE, T. H. PRETTYMAN, R. C. ELPHIC, W. C. FELDMAN, and J. J. HAGERTY, "MCNPX Benchmark for Cosmic Ray Interactions with the Moon," Journal of Geophysical Research, 111 (2006)
- [5] E. ROZANOV, T. EGOROVA, W. SCHMUTZ, AND T. PETER, "Simulation of the Stratospheric Ozone and Temperature Response to the Solar Irradiance Variability during Sun Rotation Cycle", Journal of Atmospheric and Solar-Terrestrial Physics, 68, 18, 2203 (2006).