LA-UR-15-26289

Title: DNDO Report

Author(s): Liegey, Lauren Rene
Wilcox, Trevor
Mckinney, Gregg Walter

Intended for: Report

Issued: 2015-08-07

DNDO REPORT

Lauren R. Liegey, Trevor A. Wilcox, & Gregg W. McKinney

Project Dates: June 1-August 7, 2015

EXECUTIVE SUMMARY

My internship program was the Domestic Nuclear Detection Office Summer Internship Program. I worked at Los Alamos National Laboratory with Trevor A. Wilcox and Gregg W. McKinney in the NEN-5 group. My project title was "MCNP Physical Model Interoperability & Validation". The goal of my project was to write a program to predict the solar modulation parameter for dates in the future and then implement it into MCNP6. This update to MCNP6 can be used to calculate the background more precisely, which is an important factor in being able to detect Special Nuclear Material. We will share our work in a published American Nuclear Society (ANS) paper, an ANS presentation, and a LANL student poster session. Through this project, I gained skills in programming, computing, and using MCNP. I also gained experience that will help me decide on a career or perhaps obtain employment in the future.

THE PROJECT

As a first exercise, I wrote a program outside of the MCNP6 code that would fit a sine curve to a set of data. My mentor also wrote a very similar program based on my work, and since his was easier to alter, we took his as our baseline and added on from there. To complete part of the algorithm, I had to write a subroutine that found the inverse of a matrix through lower-upper decomposition. I also wrote several programs to help optimize the curve fitting algorithm, including a program that found the correct number of data points to use in creating a fit and a program that determined the error in our fit compared to actual data. I did many tests and tweaks to the program before trying to compile it with MCNP6 to make sure it was giving reasonable results. Then I took the subroutines my mentor and I had created and added them into a module in MCNP6 that deals with cosmic background radiation. I also used more current solar modulation data in my version of the code. After some initial debugging of problems caused by re-formatting the subroutines, we ran several input decks of test problems both with the regular version of MCNP6 and with a version which had our added updates for the solar modulation parameter. The results of my project and more details about my work are included in my paper, which I have also included in the next sections of this report, and will be soon be submitting to be published by ANS.

Introduction

The annual values of the solar modulation parameter, φ, have been determined since 1936, and their fluctuation is seen to be roughly sinusoidal over time. Knowledge of this fluctuation could be used in several applications. For example, if one wished to detect a small amount of radioactive substance, knowledge of exactly how much radiation was coming from cosmic or terrestrial background would be essential for forming a threshold for detection. A measurement above this threshold would indicate a detection of the substance and a measurement at or below the threshold would indicate that no substance was present. This example experiment would be easier to carry out if the solar modulation parameter's value was known. If one wished to predict or simulate the outcome of the experiment in the future, however, a prediction for the solar modulation would be necessary. The current data available only reports solar modulation until the year 2014. The current version 6.1.1 of MCNP only contained the data up through 2005. We replaced this data with newer data, from I. G. Usoskin et al. [1], and also added the data for years 2006-2014 to MCNP (see figure 1).

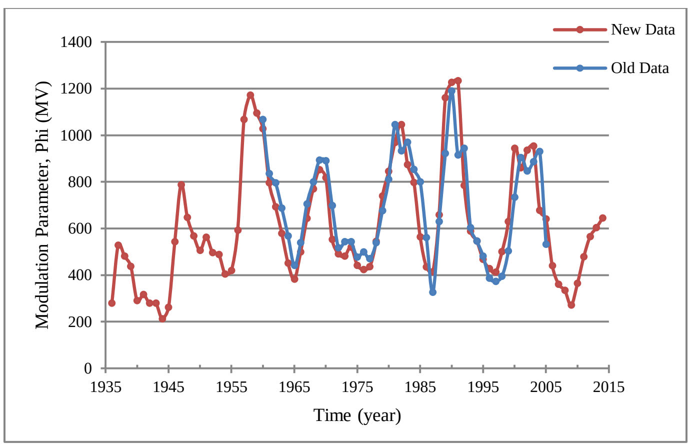

Fig. 1. The old data (in blue) was already implemented in MCNP [2]. We replaced it with the new data (in red) in preparation for implementing our program to predict solar modulation.

Description of the Actual Work

Our general algorithm consisted of the following steps: To start, we want to take in as input a date for which our program will predict the solar modulation. That input date will be processed in one of three ways. If the input year is before 1936, the modulation data from 1936 is used. If the input year is between 1936 and 2014, the program uses a linear interpolation to find the solar modulation. If the input year is after 2014 (or is after the most recent year for which solar modulation data has already been collected), several of the most recent modulation data points available (we'll call them the "predictor points") are used. A sinusoidal curve is fit to those predictor points, and from that curve, the desired year's solar modulation can be predicted. Note that our use of "sinusoidal" fit refers to a fit of the following form:

$$y = A\sin(wt) + B\cos(wt) + C \tag{1}$$

Where y is our estimate of the solar modulation, t is a time index related to the year, w is the angular frequency or 2π times the inverse of the solar modulation period, and A, B, and C are unknown real-valued constants.

Producing an Initial Fit

Now we will focus on how we obtained the fits for years beyond which we have data, which is currently 2014. We employed the common method of minimizing the mean-square errors (MSE) as seen in Tilden et al., 2001[3]. To avoid confusion, we called the solar modulation data from the predictor points φ_i, corresponding to the years t_i where i is an integer from 1 to the number of predictor points, which we'll call n for now. Now, to minimize the MSE, we want to minimize the difference between our fit values, and the actual values of the solar modulation parameter (the φ_i's). Mathematically, we want to minimize the following expression:

$$\sum_{i=1}^{n}(\varphi_i - A\sin(wt_i) + B\cos(wt_i) + C)^2 \tag{2}$$

To do this, we must know (or have an estimate for) w, and then determine the correct coefficients A, B, and C. Using Equation 2, we can create a matrix expression in which our calculated values for A, B, and C (A_0, B_0, and C_0 respectively) are collected into one vector of unknowns. If we let x_0 be that vector of unknowns such that

$$x_0 = \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} \tag{3}$$

and φ be the vector whose i^{th} component is φ_i, and D_0 be such that

$$D_0 = \begin{bmatrix} \sin(wt_1) & \cos(wt_1) & 1 \\ \sin(wt_2) & \cos(wt_2) & 1 \\ \vdots & \vdots & \vdots \\ \sin(wt_n) & \cos(wt_n) & 1 \end{bmatrix} \tag{4}$$

Then we have a much simpler looking matrix expression to replace Equation 2:

$$(\varphi - D_0x_0)^T(\varphi - D_0x_0) \tag{5}$$

If we make an educated guess at a value for w, call it w_0, then all that remains to do is minimize the expression in Equation 5 and solve for x_0 using the following equation from Tilden et al., 2001 [3].

$$x_0 = (D_0^TD_0)^{-1}(D_0^T\varphi) \tag{6}$$

Once we have w_0 and x_0, we have an initial guess at a fit! If we were confident in the accuracy of our guess for w_0, we would be done and our fit would be a good one. However, in this case, that confidence would be unfounded.

Since we know that the solar cycle's period is approximately 11 years [4], the angular frequency is ~ 0.57 /year. So we chose w_0 = 0.57 /yr. But a closer inspection reveals that if our estimate of the period is wrong by just ±0.5 yr (a reasonably probable occurrence) then we can have a change of ±0.02 in w, which changes our results significantly. In fact, when doing our fit, we found that the actual modulation data can have a local "period" of up to 12.3 years (see figure 1 around 1984-1993) or down to 8.4 years (see figure 1 around 1995-2006), though the overall average may be about 11 years.

Ensuring Fit Quality

Since the modulation period of the sun is not exactly 11 years, we need a more accurate estimate for w. To find it, we changed our matrices so that after iteration, they would yield to us a Δw. That is, the amount by which the estimate of w changes due to the iteration process. Said another way, we will use our initial values of w_0, A_0, B_0, and C_0 as a starting point and find new values, w_1, A_1, B_1, and C_1 using Δw. Once we have these new estimates, we can iterate the process of finding them, each time using the newest, improved set of values to find an even better set. We took as an assumption that through the process of this iteration, our coefficients' values would converge to a certain specific set of values which, if found, would give the very best possible fit. Since that convergence could take a long time, however, we decided to stop the iteration when the coefficients no longer changed by an amount greater than a certain specified amount, which we chose based on our desired level of precision. We also put in a limit on the amount that our final estimate for w could deviate from the average of 0.57 /yr. This way, we could prevent our fit from doing anything we would consider unreasonable. Since the highest frequency we found in the recent past was ~0.75 /yr and the lowest was ~0.51 /yr, we set the bounds for the frequency to be 0.45 /yr and 0.79 /yr. (Note that this corresponds to a change in the period of up to 3 years from the average of 11 years.) If, after the iterative process, the value of w was estimated to be outside the above bounds, our program would automatically reset the estimate for w back to 0.57 /yr since we deemed an estimate outside the bounds to be unreasonable. After all iterations or resets are complete, the final coefficients determine our fit and thus our estimate for the solar modulation for the input year.

Accuracy of the Prediction

Once we had our basic algorithm and were confident we could make good predictions, we looked for ways to quantitatively measure the accuracy of our prediction and then to make the prediction even more accurate.

Measuring Accuracy

In order to test our prediction, we needed to compare it to actual data. That means we needed to compare to data from the past - we needed to “predict” values that were actually in the past. We modified our program so that it would temporarily do a sinusoidal fit to “predict” years between 1936 and 2014. Then, by running our program with a year from the past as the input, we compared our “predicted” solar modulation value to the actual one for that year. We also looked at how far into the future the fit generated from that particular input year would be valid. We compared our predictions for the year we input, as well as the next three years after the input year to the actual data to see how our prediction matched reality if it was carried further into the future. We made a small program to calculate the absolute and relative differences for each of the four predicted years for a given input year, the results of which will be discussed later.

Making Improvements

When doing our initial tests of our fit and making plots, we used a fixed number of predictor points, namely eleven of them. We noticed that using any fixed number of predictor points to make our sinusoidal fit did not give the best results when used over multiple years. We determined through many tests that to get the best fit, we need to have the most recent local maximum and minimum from the measured data included in our predictor points, plus a couple of points on the far side of the second extrema. Since, as we mentioned before, the local frequency of the solar modulation can vary, just using eleven points every time to make our prediction did not always ensure that both the most recent maximum and minimum (an approximate period) would be included in those eleven points. Furthermore, for some years, eleven points seemed to be too many, perhaps including three of the most recent extrema, which also gave a bad prediction. So, to ensure that we used only the most recent local maximum and the most recent local minimum when finding our fit, we wrote a program to find the two most recent extrema. It then counted the number of predictor points that should be used, so that both extrema were included in the predictor points when calculating the fit. See figures 2 and 3 for a comparison and how only using a few less points drastically improved the fit for input year 2001. Note that depending on the input year, more or fewer predictor points may have needed to be used.

Additionally, since our method of finding the coefficients assumed that the value of the coefficient was converging to a single value, we noted that the program might not work well for a certain input year if one or more of the coefficients were not converging to a single value fast enough, or at all. Thus we put in a hard iteration limit, so that if after iterating many times the difference between subsequently calculated coefficients was not decreasing, the iterating process would stop anyway at a fixed number of iterations (namely 25) and issue a warning message to the user. Note that in some cases this could lead to higher errors but would force the program to always produce a result in a reasonable amount of time.

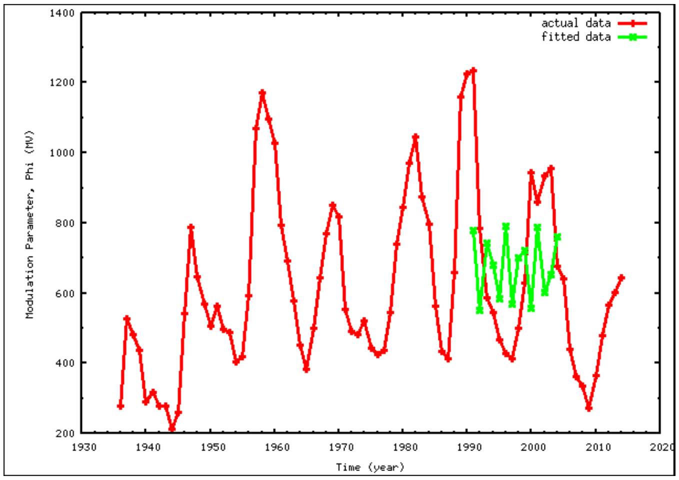

Fig. 2. Fit for years 2001-2004 using previous 11 points. The 11 predictor points and 4 predicted points are graphed in green.

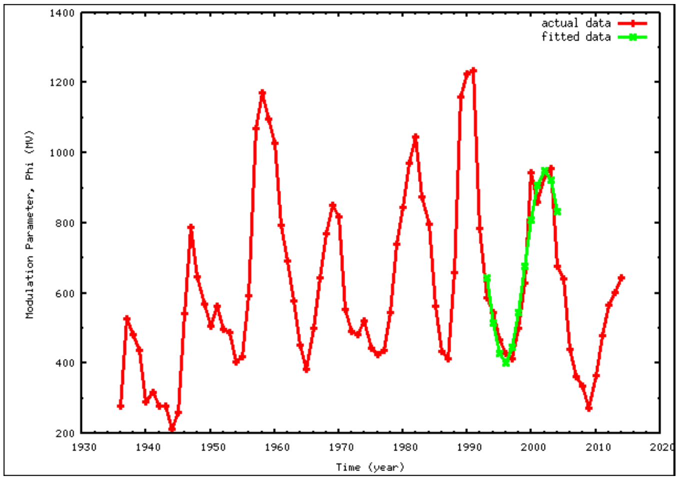

Fig. 3. Fit for years 2001-2004 using the number of points determined by a program that includes only the two most recent extrema. The 8 predictor points and 4 predicted points are graphed in green. The number of points will vary with input year.

Testing in MCNP

Once we had our programs written, we implemented them into MCNP to do more testing there. By running the code both with our updates and without, we could check to see that our additions did as we expected them to. We could also compare results from the updated and un-updated versions of the code to measured data.

First, we looked at just the value of the solar modulation parameter, φ, to see how much better our updated version of the code was able to predict φ than the version which had the old solar modulation data implemented. To do this, we ran MCNP for three cases, all with the input date January 1, 2014. In one case, we ran with all of the new data in place (our updated version of the code), so the value for φ was simply the last of the new measured data points (see figure 1). In another case, we ran the un-updated version of the code, which had the old data and algorithms and predicted the solar modulation for 2014 based on data from 2005 and earlier. In our final case, we ran our updated version of the code again, but this time we took out the data point from 2014 so that our code would predict the value for φ in 2014 using data from 2013 and earlier. We also did similar runs for the same date but at different latitudes. The results of the comparison of the values of solar modulation parameter we found are discussed later.

Next, we ran tests on MCNP to compare the results of our updated code, the un-updated code, and measured data. To do this, we used input decks which described a situation where actual data was taken and ran both versions of the code with these input decks. We again looked at the neutron flux energy spectrum for comparison. We used one input deck that was based off of an experiment done on the surface of the earth (see figure 5) and another input deck that was based off of an experiment done on a plane during flight (plot not shown here).

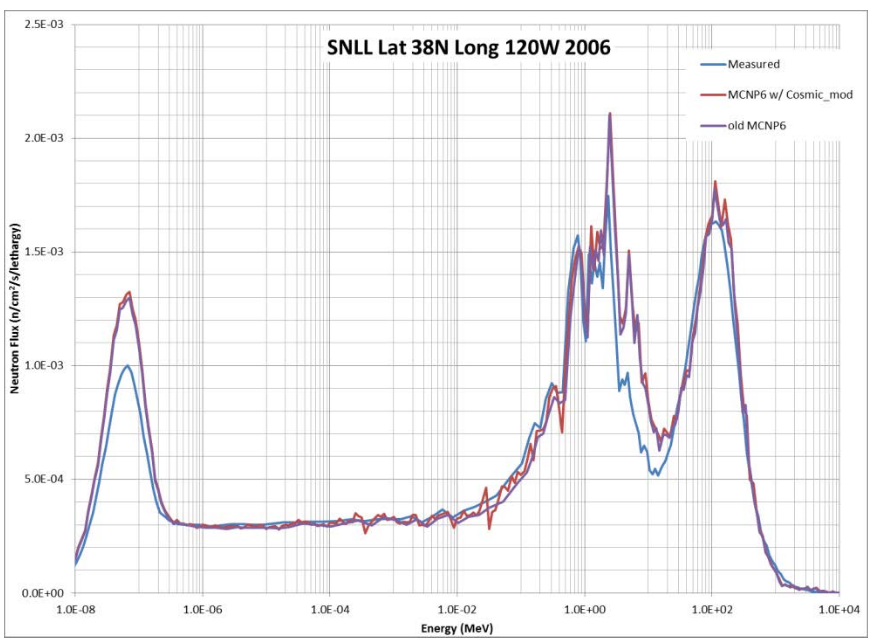

Fig. 5. Example of a comparison of the updated code and un-updated code to data measured in the U.S.

Results

Before discussing our results, it is important to keep in mind that the sun's cycles are not exactly sinusoidal and thus our aim to predict the modulation using a sinusoidal fit will always have shortcomings. One should certainly not expect to accurately predict what the solar modulation will be a decade in the future using this method. According our tests, however, after making a fit, one could expect an average difference from the actual modulation of around 8% for the first year in the future, a fairly good prediction. For the second, third, and fourth years in the future, however, one would expect an average difference of about 16%, 23%, and 29%, respectively. As is made evident by these values, the further in the future you try to predict, the worse you can expect your fit to be (see figure 6). We found these overall percent differences by using the method described earlier for each of the years between 1999 and 2008 (inclusive) and averaging them for each year step into the future.

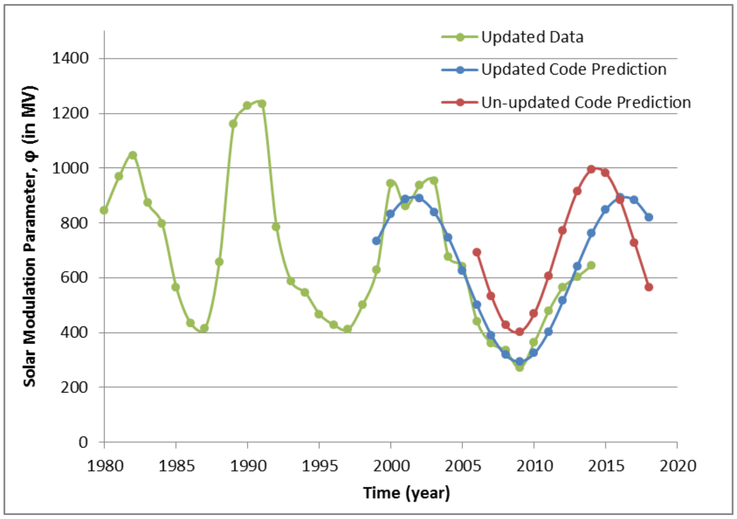

Fig. 6. Example of how well the updated code predicts the solar modulation parameter, compared to what a prediction by the un-updated code would look like for year 2018.

ACHIEVEMENTS

During the course of my project, the group I worked with, including all the students in the group and our mentors, met weekly to discuss our progress in our projects as well as any other concerns we had. Each week, every student gave a summary of the work they were currently doing, and discussed any problems and work plans for the coming week. My mentor also encouraged me to keep track of my progress by writing a paper detailing my work. Toward the end of my project I was able to fix this paper up and will be submitting it to ANS to be published, with me as the first author. I will be working on a presentation to go along with that paper so that either I or my mentor can present our work at the next available ANS conference. I also had the opportunity while at LANL to create a poster describing my project and I will be presenting that poster during the LANL Student Symposium before I leave. In this symposium I will be able to share my work from the summer and network with other students and mentors.

SKILL AND KNOWLEDGE GAINED

Prior to this project I had only taken a one semester class in computer science, and this project was my first experience programming in FORTRAN. This alone was a big learning experience for me, since I didn't consider even the small amount of knowledge I gained from the computer science class I took to be all that useful. I think my new knowledge of programming will be very useful to me, as next year I will be starting graduate school to get my Ph.D. in Physics. In order to do this project, I was also exposed to Unix and High Performance Clusters, which were both new to me. I learned to use Unix in order to access FORTRAN and MCNP and I ran test problems on the clusters. While working on this program, I also had to brush up on my knowledge of linear algebra and some general math methods, which I'm sure will also be useful to me in the future. During the second week that I was at LANL, my mentor had signed me up for a week-long MCNP workshop, so I had the opportunity to learn quite a bit about making input decks and running the code. I even got good enough at creating geometries that I created a 3-D cartoon woman ("Fission Girl") that the instructor asked to keep and use as an example to show future students in the workshop. The knowledge I gained in this workshop and in the testing that I did with MCNP6 in the later stages of my project will be invaluable to me, as in general there are a limited number of people with the skills to work with MCNP. I believe that the skills I have collected during this internship will make me a very attractive applicant to potential employers and that I will be seen as a good addition to research groups I may want to join in the future.

THE EXPERIENCE

This internship experience has had a big impact on my academic and career planning. I think it really gave me a good idea of what working at a national lab could be like, which is something I have been interested in since I declared my intent to pursue a B.S. degree in Physics. It also gave me a chance to make connections here at LANL so that I may actually have the opportunity to work here in the future if that does end up being the career path that I choose. While here I got to tour both the TRINITY and DARHT facilities and hear about the work being done there. Our group of students (from NEN-5) was also taken to Albuquerque to tour the HERMES III, SPHYNX, and the Z-Machine at Sandia National Lab. I also got to attend many lectures at LANL. Most were given by people in my group who have been working here a while or who worked somewhere else but came to talk to our group's students about topics that may interest us specifically. Some others were given at other places in LANL and included topics such as Cyber Security and the search for Dark Matter. I also attended a picnic put on by the Student Association at LANL that was for all the students here during the summer and their mentors. It was a good time to socialize and have fun with my coworkers. I will be able to attend a picnic for my group, NEN-5, during the last week of my internship, which should also be a good time to network and have fun with other people doing work similar to me, including other students, our mentors, and other senior members of the group.

RESEARCH IDEAS

In general, it would be useful to DHS to invest in the modernization, improvement, and upkeep of simulation codes like MCNP, and not just funding projects to create new features. One thing I thought of that may be a possible area of development for MCNP is implementing more variables into the code that modulate the background, such as barometric pressure. This would serve the same purpose that updating the solar modulation did, and would hopefully make simulations of the background even more accurate. Another area that DHS might consider investing in is the research and development of Muon Tomography, which has been developed at LANL, and seems to be a promising technology.

REFERENCES

1. I. G. USOSKIN et al., “Solar modulation parameter for cosmic rays since 1936 reconstructed from ground-based neutron monitors and ionization chambers,” *Journal of Geophysical Research*, **116**, (2011). [points after year 2009 taken from the following site: http://cosmicrays.oulu.fi/phi/Phi_mon.txt
2. I. USOSKIN et al., “Solar modulation of cosmic rays since 1936: Neutron monitors and balloon-borne data,” *32nd International Cosmic Ray Conference*, Beijing, 2011, Vol. 11, p. 39.
3. S. TILDEN et al., “IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters,” IEEE Std 1241-2000, The Institute of Electrical and Electronics Engineers, Inc., June 13, 2001.
4. K. G. McCRACKEN and F. B. McDONALD, “A phenomenological study of the long-term cosmic ray modulation, 850-1958 AD,” *Journal of Geophysical Research*, **109**, (2004).