The Astrophysical Journal, 795:11 (13pp), 2014 November 1 doi:10.1088/0004-637X/795/1/11

LATITUDE SURVEY INVESTIGATION OF GALACTIC COSMIC RAY SOLAR MODULATION DURING 1994–2007

W. Nuntiyakul[1,2,3], P. Evenson[4], D. Ruffolo[1,2], A. Sáiz[1,2], J. W. Bieber[4], J. Clem[4], R. Pyle[4], M. L. Duldig[5], and J. E. Humble[5]
[1] Department of Physics, Faculty of Science, Mahidol University, Bangkok 10400, Thailand; w.nuntiyakul@gmail.com, david.ruf@mahidol.ac.th, alejandro.sai@mahidol.ac.th
[2] Thailand Center of Excellence in Physics, CHE, Ministry of Education, Bangkok 10400, Thailand
[3] Faculty of Science, Chandrakasem Rajabhat University, Bangkok 10900, Thailand
[4] Bartol Research Institute and Department of Physics and Astronomy, University of Delaware, Newark, DE 19716, USA; evenson@udel.edu, jwbieber@bartol.udel.edu, clem@bartol.udel.edu, pyle@bartol.udel.edu
[5] School of Physical Sciences, University of Tasmania, Hobart, Tasmania 7001, Australia; John.Humble@utas.edu.au, Marc.Duldig@utas.edu.au
Received 2014 March 19; accepted 2014 September 1; published 2014 October 8

ABSTRACT

The Galactic cosmic ray spectrum exhibits subtle variations over the 22 yr solar magnetic cycle in addition to the more dramatic variations over the 11 yr sunspot cycle. Neutron monitors are large ground-based detectors that provide accurate measurements of variations in the cosmic ray flux at the top of the atmosphere above the detector. At any given location the magnetic field of the Earth excludes particles below a well-defined rigidity (momentum per unit charge) known as the cutoff rigidity, which can be accurately calculated using detailed models of the geomagnetic field. By carrying a neutron monitor to different locations, e.g., on a ship, the Earth itself serves as a magnet spectrometer. By repeating such latitude surveys with identical equipment, a sensitive measurement of changes in the spectrum can be made. In this work, we analyze data from the 1994 through 2007 series of latitude surveys conducted by the Bartol Research Institute, the University of Tasmania, and the Australian Antarctic Division. We confirm the curious "crossover" in spectra measured near solar minima during epochs of opposite solar magnetic polarity, and show that it is directly related to a sudden change in the spectral behavior of solar modulation at the time of the polarity reversal, as revealed from contemporaneous variations in the survey data and a fixed station. We suggest that the spectral change and crossover result from the interaction of effects due to gradient/curvature drifts with a systematic change in the interplanetary diffusion coefficient caused by turbulent magnetic helicity.

Key words: cosmic rays – solar–terrestrial relations

1. INTRODUCTION

The Galactic cosmic ray (GCR) flux in the solar system is strongly influenced by solar variations (Forbush 1954), a process known as solar modulation. Solar modulation itself is dominated by the roughly 11 yr sunspot cycle in which the GCR flux decreases during sunspot maximum, the time period of maximum solar activity, and increases during sunspot minimum (Figure 1). This can be approximately described by a spherically symmetric model with a single "modulation" parameter related to the solar wind speed and the GCR diffusion coefficient, such as the force-field model (Gleeson & Axford 1968).

However, there is also a roughly 22 yr GCR variation corresponding to the solar magnetic cycle, wherein the dominant solar magnetic polarity reverses at each sunspot maximum (Thambyahpillai & Elliot 1953). In other words, 11 yr periods with opposite magnetic polarity exhibit distinct GCR variations. These effects are associated with a variety of interesting phenomena, such as guiding center drifts, latitudinal GCR gradients, particle charge sign dependence, and changing diffusion coefficients (e.g., Jokipii et al. 1977; Garcia-Munoz et al. 1986; Bieber & Chen 1991). These phenomena depend on the sign of qA, where q is the particle charge and A is the solar magnetic polarity.

Effects of solar magnetic polarity are clearly seen in the GCR flux as a function of time, which is more "pointy" or "flat-topped" in alternating sunspot cycles (Jokipii & Thomas 1981; Webber & Lockwood 1988; see Figure 1), and by comparing fluxes of particles of the same charge to mass ratio but opposite charge sign, such as electrons to positrons or protons to antiprotons (Bieber et al. 1999). There are many more subtle manifestations as well (Popielawska & Simpson 1990, 1991). One of the most puzzling aspects of modulation phenomenology is the so-called crossover in spectral form during opposite magnetic polarity epochs. The crossover is one manifestation of a dependence of spectral shape on solar magnetic polarity. This has been observed by means of ship-borne neutron monitor surveys that study the GCR spectrum by traveling across a wide range of geomagnetic cutoff rigidities. With significant reliance on the work of Webber & Lockwood (1988), Moraal et al. (1989) extensively discussed and characterized the effect by comparing neutron monitor response functions obtained at solar minima separated by 11 yr intervals (opposite magnetic polarity) and 22 yr intervals (same magnetic polarity). They reported a crossover at a rigidity (momentum per charge) of approximately 6 GV in spectra taken during opposite polarity epochs. Intensities at different energies are not uniformly higher or lower. In the negative polarity state the low energy fluxes are suppressed while the high-energy fluxes are enhanced, with just the opposite being true in the positive polarity state. This crossover phenomenon was confirmed by Lockwood & Webber (1996) and Bieber et al. (1997), albeit with some question as to the energy of the crossover. Reinecke et al.

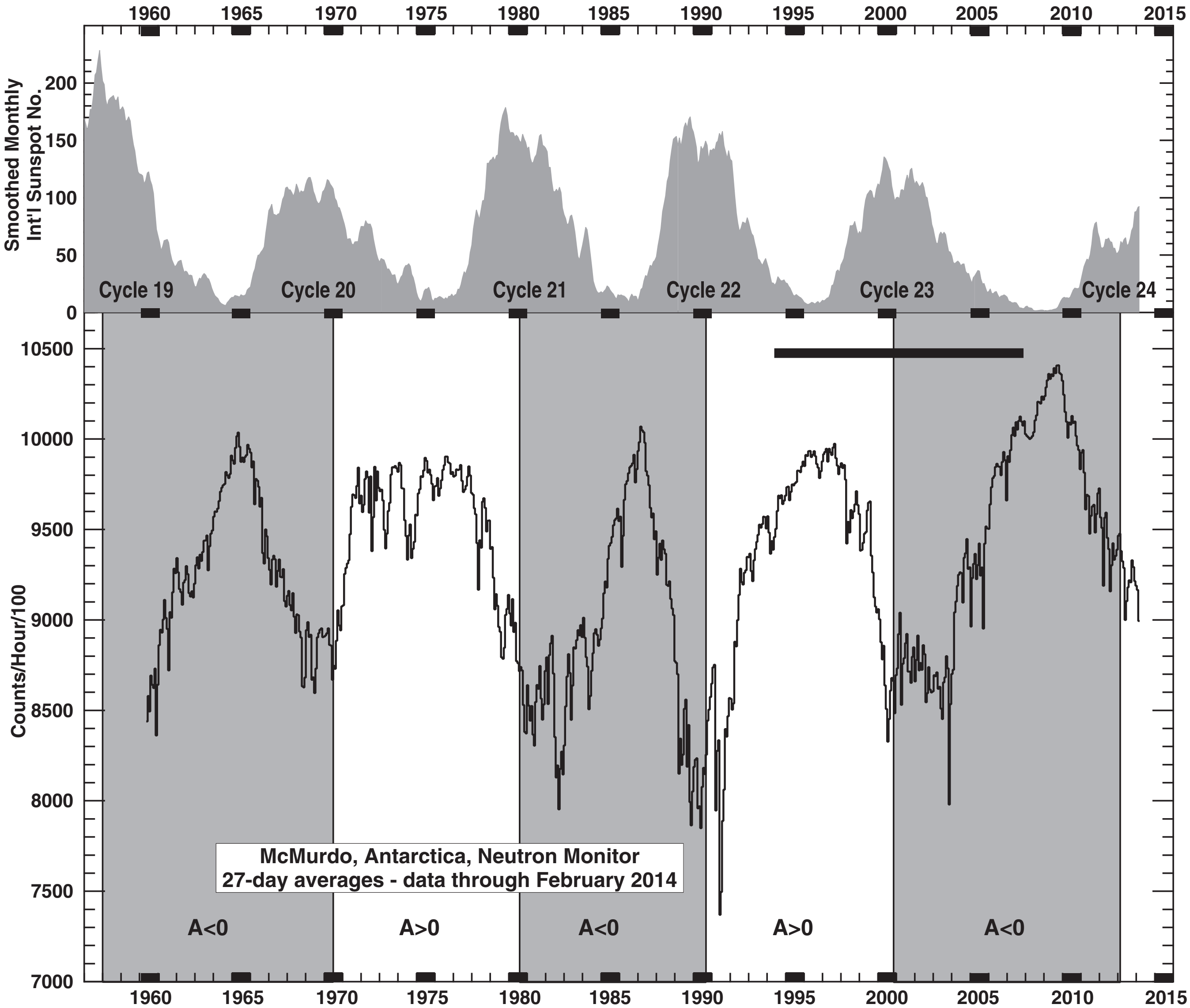

Figure 1. Smoothed monthly international sunspot number (using five-month boxcar smoothing) and McMurdo neutron monitor count rate as a function of time. The long-term drift at McMurdo has been corrected following Oh et al. (2013). A neutron monitor count rate indicates the Galactic cosmic ray flux, which undergoes "solar modulation" in association with solar activity. Solar modulation includes dramatic 11 yr variations with the sunspot cycle, and a 22 yr variation with the solar magnetic cycle, seen here in changes in the solar modulation pattern between positive ($A > 0$) and negative ($A < 0$) magnetic polarity. In this work we present observations of spectral changes in Galactic cosmic rays in association with solar modulation and changing solar magnetic polarity for the time period 1994–2007, indicated by a horizontal bar.

(1997) suggested that there could be two crossovers, at least at some times.

Reports of crossovers in the literature have usually compared observations at two different time periods "near solar minimum," but it has recently become clear that different solar minima can have different levels of solar modulation (Oh et al. 2013). The question of whether or at what rigidity the spectra intersect at successive solar minima depends on the relative level of modulation at the two time periods. Actually, the key physical issue is whether solar modulation of GCR spectra is independent of solar magnetic polarity for similar modulation conditions. Such a polarity dependence can be demonstrated by the existence of a spectral crossover at any similar modulation conditions for different magnetic polarities, or equivalently, if a model of solar modulation that matches the GCR spectral evolution during one solar magnetic polarity systematically deviates from that for the opposite polarity.

Additional questions that immediately come to mind involve the transition in solar modulation from one polarity state to the other. Is this transition smooth or abrupt? Alternatively, the crossover might be a phenomenon that appears as a result of some special conditions that appear and disappear only very near solar minimum. To investigate the nature of this transition, a series of neutron monitor latitude surveys was conducted between 1994 and 2007 by the Bartol Research Institute of the University of Delaware, the University of Tasmania, and the Australian Antarctic Division. We present two types of analysis of the survey data: the classic approach of determining response functions, which confirms a crossover when comparing similar modulation conditions near successive solar minima, and an analysis of contemporaneous variations in the survey data and data from a fixed NM station, which reveals a spectral change associated with a major transition in some aspect of modulation that occurs nearly simultaneously with the solar polarity reversal.

2. OBSERVATIONS

In a neutron monitor latitude survey the magnetic field of the earth is used as a spectrometer to explore the spectrum of cosmic rays striking the atmosphere. The count rate N of a surface mounted detector resulting from the impact of cosmic rays at the top of the atmosphere is described formally by

$$N(\Theta, \Phi, h, t) = \int_0^\infty \left[\sum_i G_i(P) M_i(P, t) Y_i(P, h) \right] \times T(P, \Theta, \Phi, t) dP. \quad (1)$$

Here the index i refers to the particle species (e.g., protons or alpha particles), and G_i is the GCR spectrum, also called the local interstellar spectrum (LIS), which is assumed to be time invariant and a function only of particle rigidity $P = pc/q$, describing the particle momentum per unit charge. The function M_i (for modulation) describes the time evolution of the spectrum due to interaction with magnetic fields in interplanetary space. T describes transmission through the magnetosphere of the Earth. It is a function of rigidity, latitude (Θ), longitude (Φ), and time (t). Seasonal and diurnal time dependence arises from the changing relationship of the offset and tilted dipole axis of the Earth to the flow direction of the solar wind. There is also explicit time dependence due to changes in the magnetosphere in response to fluctuations in the solar wind velocity and embedded magnetic field. Y_i is called the yield function, which expresses the inherent cascade processes in the atmosphere and the detector that produce a count rate in the presence of the flux at the top of the atmosphere. It is a function of rigidity (P) and depth in the atmosphere (h), which in many cases is equivalent to barometric pressure.

We follow the usual approach of representing the geomagnetic transmission T as a step function at an effective "cutoff" rigidity P_c, which depends on latitude, longitude, zenith angle, azimuth, and time. In this case the relation becomes

$$N(P_c, h, t) = \int_{P_c}^{P_L} \sum_i G_i(P) M_i(P, t) Y_i(P, h) dP. \quad (2)$$

We also introduce a limiting rigidity P_L as a numerical convenience because Y_i (as a function of P) increases without bound but G_i falls steeply at high rigidity. Taking T to be independent of arrival direction implicitly assumes that the incoming flux at the top of the atmosphere is isotropic, which it is not. We deal with this by using what we have defined as the "apparent cutoff," which takes obliquely incident particles into account by weighting effective cutoffs calculated for nine arrival directions. We calculated this individually at one hour intervals at the actual position of the ship with a time dependent model of the magnetic field. The apparent cutoff organizes mobile monitor data better than the commonly used vertical effective cutoff, and can be substantially higher. For example, our mobile neutron monitor surveys reached a maximum apparent cutoff of 18.08 GV, when the vertical effective cutoff was 17.42 GV. The two measures of the cutoff rigidity were compared in detail by Clem et al. (1997) and Bieber et al. (1997).

In our analysis, data from the mobile (ship-borne) neutron monitor are pressure-corrected to standard sea level (760 mm Hg), so we remove the dependence on h. The mobile monitor rapidly moved in P_c, while the solar modulation, M_i, varied slowly over a scale of years. (Note that temporary Forbush decreases due to solar activity are excluded from our analysis.) Therefore, $N(P_c, t)$ can be measured as a function of P_c at nearly constant t, and in this sense is commonly referred to as the (integral) response function for a neutron monitor at epoch t.

The integrand of this simplified equation can be referred to as the differential response function (DRF). Measuring the signal in the detector at two different geomagnetic cutoffs P_c and subtracting one measurement from the other determines the response of the detector to that little part of the spectrum lying between the two cutoffs. When divided by the (small) difference in cutoff this gives the DRF directly. For physical interpretation of the DRF, a common assumption is that G_i, M_i, and Y_i are independent of particle species i up to a constant, which is absorbed in the factors so that the sum in Equation (2) is replaced by the product $G(P)M(P, t)Y(P)$. (Note that measurement of the DRF does *not* require this assumption.) Then when an identical detector is used in successive surveys, a fractional change in the DRF is interpreted as a fractional change in the spectrum $G(P)$, which can be determined without knowledge of the yield function.

In our series of surveys, atmospheric neutrons were detected by a ship-borne mobile neutron monitor technically known as a 3NM64. NM64 refers to the basic design (Carmichael 1964) and "3" refers to our use of three independent neutron counters. These were installed in an insulated shipping container (called the "TasVan") as shown in Figure 2. The NM64 is constructed mainly from lead and polyethylene, with the neutron detectors inserted from the side. The detectors are Chalk River BP-28 proportional counter cylinders filled with boron trifluoride gas (enriched in the isotope ^{10}B). The boron nuclei react with neutrons and undergo nuclear fission. The reaction products (^{4}He and ^{7}Li) ionize the gas and eventually produce electrical pulses on a central anode wire maintained at ground, while a cylindrical outer cathode is at a potential of about −2800 V. These electrical pulses were counted by two independent and redundant data acquisition systems: a simple logger and electronics that were custom-designed and assembled by the Bartol Research Institute.

The 3NM64 mobile monitor was operated during 1994–2007 aboard one of two U.S. Coast Guard icebreakers, the *Polar Sea* or the *Polar Star*, which traversed the Pacific Ocean from Seattle, USA to McMurdo, Antarctica and back during a ∼6 month voyage. We refer to a "survey year" by the year in which the voyage began; each voyage then extended throughout the Austral Summer into the next calendar year. For example, data for "survey year 1997" refer to data of the voyage from 1997 October to 1998 April. We obtained data for 13 consecutive survey years (from 1994 to 2006). The geographic routes of all surveys are plotted in Figure 3; also shown for reference are contours of constant 1980 vertical effective geomagnetic cutoff rigidity.

During survey year 1994, the ship traveled from Hobart, Australia to McMurdo, maneuvered three quarters of the way around Antarctica, and then passed through the Straits of Magellan. Data taking ended near the Equator. In survey year 1995, the data were recorded starting in San Diego, USA, and from there to McMurdo, where the data taking stopped. In survey years 1996 and 1998, the ship started from Seattle and crossed the Pacific Ocean with stopovers in Honolulu, Hawaii and Hobart before going to McMurdo, then returning to Seattle with a stopover in Adelaide, Australia and passing near Thailand, the region of the world's highest cutoff rigidity. The routes of survey years 1997 and 1999 started from Seattle and crossed the Pacific Ocean

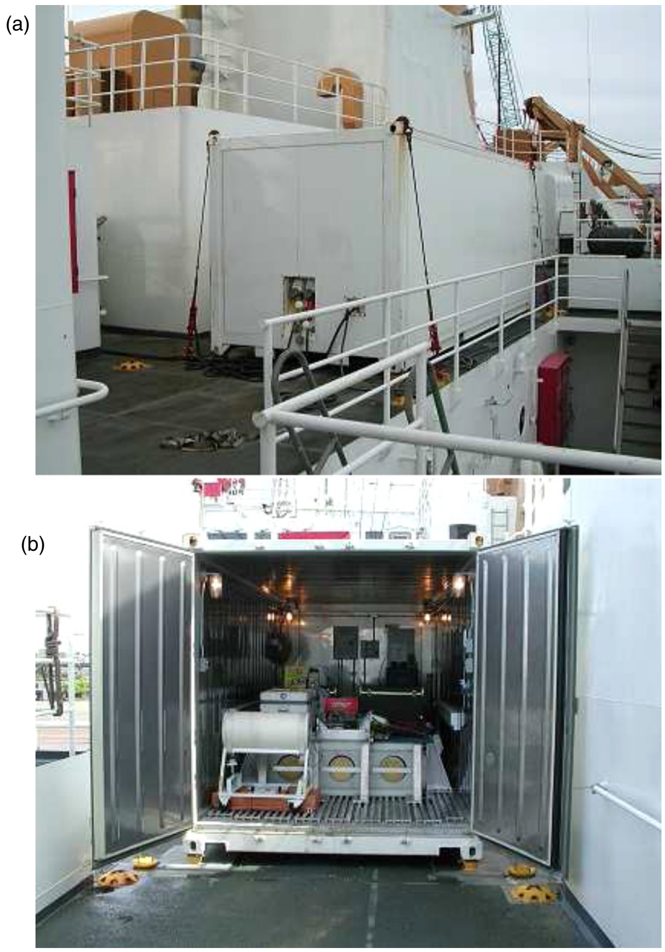

Figure 2. (a) Insulated container ("TasVan") used for 3NM64 neutron monitor latitude surveys on the *Polar Star* and *Polar Sea*. (b) 3NM64 installed inside the container. For this and several other voyages a small calibration neutron monitor (Krüger et al. 2008) was installed near the door.

with a stopover in Honolulu before going to McMurdo and back to Seattle passing through Mazatlán, Mexico. In survey years 2000, 2002–2004, and 2006, the ship traveled from Seattle via Honolulu to McMurdo and returned the same way, while in survey year 2001, the ship returned via Valparaíso, Chile. In survey year 2005, the data were recorded starting in McMurdo, and from there via Honolulu to Seattle.

In addition to the mobile neutron monitor and data acquisition systems, a variety of additional equipment was carried in the TasVan. A high-precision Digiquartz barometer (from Paroscientific, Inc.) was used to provide precise barometric data to enable correction of the neutron monitor count rate. An AIR (brand name) barometer provided a redundant (though less precise) measurement of the barometric pressure, as a precaution in case of failure or drift in the Digiquartz measurement. A temperature sensor was used to measure the temperature inside the TasVan. A special feature of the electronics, started in the 1995 survey year, was the measurement of histograms of neutron time delay, i.e., the time interval from the detection of one neutron to the next (Bieber et al. 2004). The time delay spectrum is related to the energy distribution of cosmic rays impinging on the atmosphere. It depends strongly on the cutoff rigidity, and in principle can be analyzed to provide some additional information about the cosmic ray spectrum during the surveys (P.-S. Mangeard et al., in preparation).

During survey year 1994, the ship's "sea state" was recorded as the measure of pitch and roll. In order to directly monitor orientation effects on the mobile monitor count rate we installed clinometers, i.e., sensors of inclination, for the remaining surveys (starting in survey year 1995) to measure the pitch and roll each second. The clinometers were composed of redundant sensors located on a circuit board in the card cage in the electronics rack. The output of each sensor provided three analog signals giving the inclination along two axes and the unit temperature. Counts from the 3NM64 were recorded once per second, together with data from the clinometers. Once per minute, pressure data and the GPS-derived latitude, longitude and time were recorded. In our subsequent data analysis, the pitch and roll were found to have only a minor effect so the data in the present analysis were not corrected for the pitch and roll.

During certain survey years, additional neutron detection equipment was included. Bare neutron counters were mounted in the 1995–1996 survey year, which permitted a calibration of the bare counter to neutron monitor count ratio as a function of the cutoff rigidity. A compact calibration neutron monitor was mounted near the access door during several surveys to test its operation and calibrate it against the standard monitor (Krüger et al. 2008).

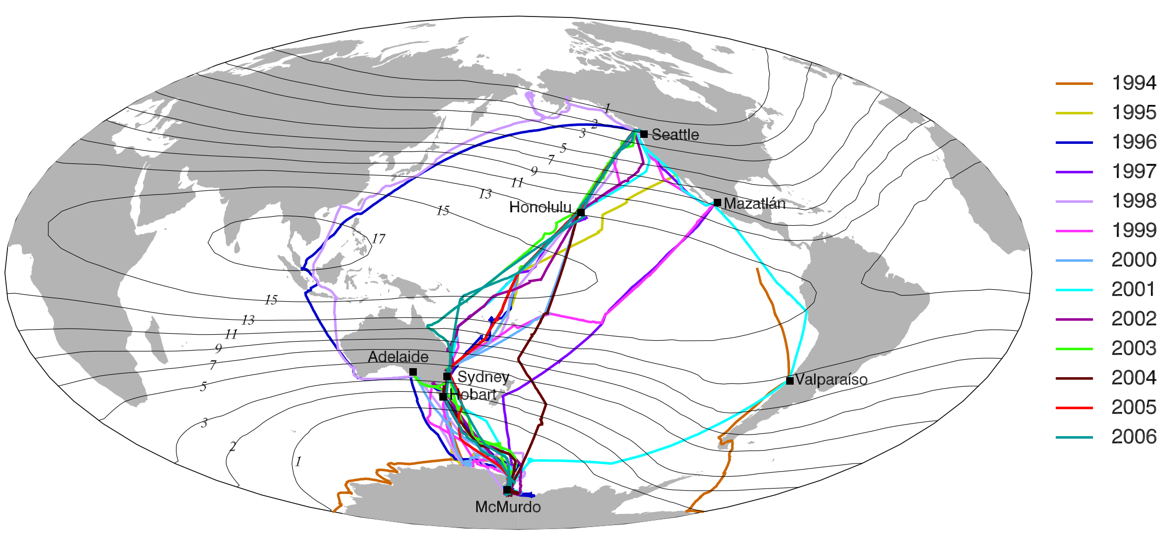

Figure 3. Tracks of the ship-borne neutron monitor latitude surveys for 1994–2007, superimposed on contours of the vertical cutoff rigidity in GV.

3. DATA REDUCTION

The count rates of the three detector tubes (which we refer to as T1, T2, and T3) were corrected in several steps. In the first step, we considered the three count ratios T1/T2, T2/T3, and T3/T1 for each data interval. The data interval duration was 1 hr for survey year 1994 and 30 minutes for subsequent survey years. The rough and changing conditions on board the ships caused the response of individual tubes to occasionally change, become noisy, or even stop completely. In order to correct for these effects during the surveys we used the inherent redundancy of the three detectors. We identified time intervals with anomalous tube ratios by comparing the tube ratios with the average tube ratios for the whole survey, which we call S1/S2, S2/S3, and S3/S1. Then we calculated a corrected count rate from the actual count rate of the properly operating detectors. If only one tube was ignored, e.g., if T1 was removed, we calculated the corrected count rate from [(S1+S2+S3)/(S2+S3)]×(T2+T3). If two tubes were removed we similarly used the remaining tube to determine the corrected count rate. If none of the tubes were operating properly the result was a data gap. For the second step of correction, we calculated the daily average tube ratios for all surveys and plotted them on one long plot as a function of time. With their greater statistical accuracy, the daily average tube ratios provide a much more sensitive indication of tube drifts or noise. We identified anomalous time periods by eye and applied corrections when possible; otherwise, the data were excluded.

Although the basic monitor assembly was quite similar each year, the TasVan was not always mounted on the ship in the same location or orientation. Additionally, in December of 2001 we moved the detector tubes to a nominally identical lead/polyethylene assembly in a new shipping container. (The old one had become leaky and impossible to maintain properly.) To develop a correction for such systematic changes, the count rate (corrected for pressure) of the McMurdo neutron monitor was used in a regression analysis with the mobile monitor count rate (corrected independently for pressure) when the ship was near McMurdo, i.e., in a roughly square 220 × 220 km^2 area centered on the McMurdo neutron monitor station, defined as $(77.85 \pm 2)^\circ$S, $(166.67 \pm 9.5)^\circ$E. We used the mobile/McMurdo count rate ratio for the 2006 survey year, 0.194 ± 0.001, as a reference. During most survey years, the ratio when the ship was near McMurdo was close to this value, but there were five survey years for which the data needed to be normalized. The normalization factors for survey years 1995, 1997, 1998, 1999, and 2000 were 1.015, 1.040, 1.039, 1.047, and 1.026, respectively.

Proper correction for variations in barometric pressure is vital to the success of our project for two reasons. First, and most obvious, is the need to remove short-term fluctuations from the data. Less obvious is the correlation of average barometric pressure with geographic latitude, causing a correlation with cutoff. We corrected to 760 mm Hg using an empirical pressure coefficient β in units of percent per mm Hg varying with cutoff rigidity as determined from our survey data:

$$\beta = 1.006 - 0.0153 P_c, \tag{3}$$

$$C_{\rm TP} = C_T e^{\beta(p-760)}, \tag{4}$$

where C_T is the mobile monitor count rate corrected for tube anomalies and normalized to McMurdo, $C_{\rm TP}$ is the mobile monitor count rate corrected for tube anomalies, normalized to McMurdo, and corrected for pressure, P_c is the apparent cutoff rigidity in units of GV, and p is the barometric pressure in units of mm Hg.

To generate a summary set of response functions we also applied a correction for short-term variations in modulation level over the course of the survey based on variations in the McMurdo count rate. First we determined regression coefficients between the normalized mobile neutron monitor count rate corrected for tube ratios and pressure and the McMurdo count rate corrected for pressure for each P_c bin, i.e., 0–1 GV, 1–2 GV,..., 17–19 GV (where 17–18 GV and 18–19 GV were grouped together). We found that the regression coefficient S as a function of apparent cutoff rigidity P_c was well fit by

$$S = 0.211 e^{-0.141 P_c}. \tag{5}$$

We then used this value of S to correct each survey year's mobile monitor data to the average McMurdo value for that survey year according to

$$C_{\rm TPM} = C_{\rm TP} - S(m - \overline{m}), \tag{6}$$

where $C_{\rm TPM}$ is the normalized mobile monitor count rate corrected for tube anomalies, pressure, and short-term GCR variations, m is the McMurdo count rate, and $\overline{m}$ is the average McMurdo count rate for that survey year. Then $C_{\rm TPM}$ was used to determine the response functions. Figures 4(b) and (c) show examples of the mobile monitor count rate for the 2002 survey year after applying all corrections.

Strong Forbush decreases (FDs) clearly did not follow the same regression relation as other short-term rate fluctuations, so we excluded them from the data set, treating the time intervals as data gaps. The criterion for rejection was a maximum percentage decrease (%D) > 10 in the McMurdo neutron monitor, which occurred three times during our mobile monitor surveys. The intervals that were excluded from our analysis due to the three FDs were: 1) from 2004 November 7 to 2004 November 18 (%D = 11.6), 2) from 2004 January 17 to 2004 January 26 (%D = 14.3), and 3) from 2006 December 6 to 2006 December 25 (%D = 10.8).

We also organized the data into survey "segments," where a segment is defined as a transition (in either direction) between high geomagnetic latitude (with low cutoff) and the cosmic ray (geomagnetic) equator (CRE; with high cutoff). We define "segment A" from the US West Coast to the CRE, "segment B" from the CRE to McMurdo, "segment C" from McMurdo to the CRE and "segment D" from the CRE to Seattle. Each survey therefore could have up to four segments, which provide four distinct measurements of the response function. However, equipment failure or major Forbush decreases resulted in fewer than four available segments in the 1994, 1995, 2005, and 2006 survey years. In total, the 13 surveys provided 44 segments of data.

4. RESPONSE FUNCTIONS AND CONTEMPORANEOUS VARIATIONS

The corrected mobile monitor count rate as a function of apparent cutoff rigidity represents the integral response function of a 3NM64 neutron monitor. This is plotted for each survey year in Figures 5(a)–(d), 6(a)–(e), and 7(a)–(d) for three modulation periods approximating solar minimum conditions (survey years 1994–1997), the transition from solar minimum to solar maximum (1998–2002), and the return

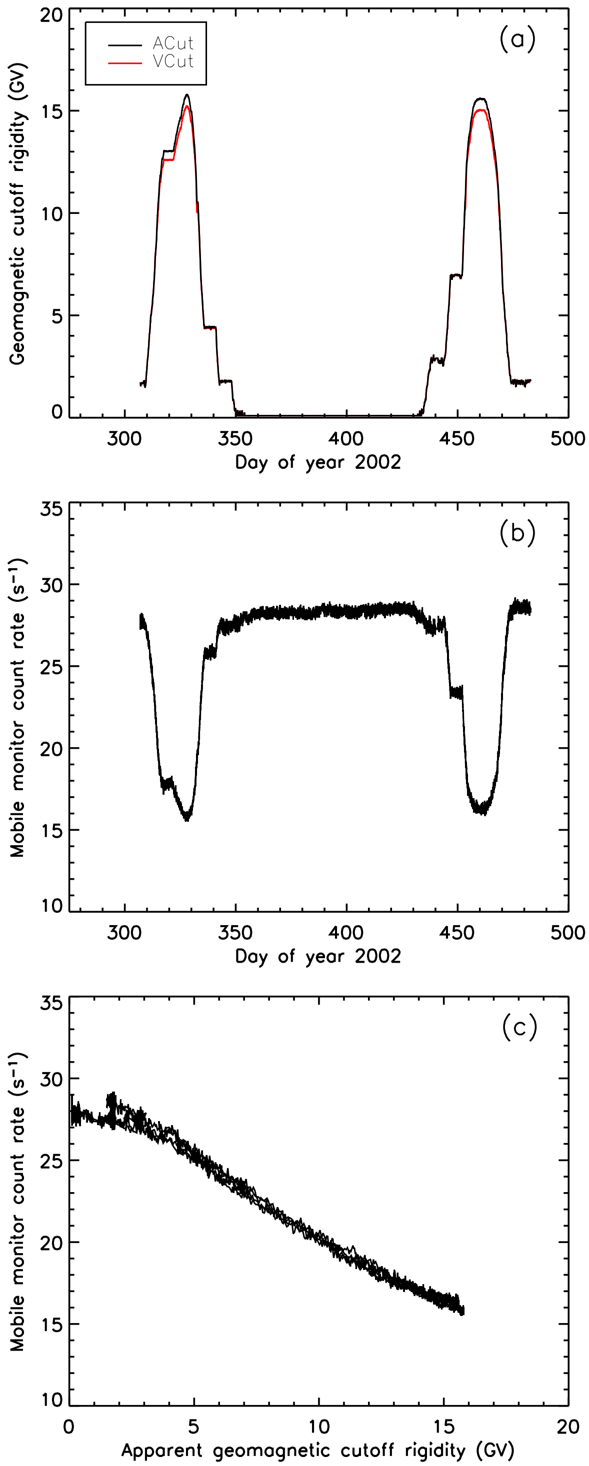

Figure 4. Example of data for the 2002 survey year. (a) Geomagnetic cutoff rigidity as a function of time. The black line traces the apparent geomagnetic cutoff rigidity while the red line shows the vertical effective cutoff rigidity. (b) Mobile (ship-borne) neutron monitor count rate after all corrections as discussed in the text. (c) Corrected count rate as a function of apparent geomagnetic cutoff rigidity for all four segments (i.e., the four transitions between low and high cutoff rigidity).

to near solar minimum (2003–2006), respectively (see also Figure 1). Averages over 0.5 GV cutoff rigidity bins, averaged over all segments in one survey year, are plotted against apparent geomagnetic cutoff rigidity. If there are multiple segments, the average is indicated by a solid symbol and there is an error bar to represent the standard error among the determinations for different segments, which indicates the systematic (reproducibility) uncertainty. Actual statistical errors are negligible. An open symbol indicates that data were available for only one segment, and the systematic uncertainty has not been determined. The results are also fitted to a commonly used parameterized function of apparent geomagnetic cutoff, the "Dorman function" (Dorman et al. 1970). Although the Dorman parameterization has no physical content it provides an excellent representation of the integral response function N and can be differentiated to determine the DRF:

Table 1
Derived Dorman Parameters

Survey Year	N_0	α	κ
1994–1995	3.08E+01	9.05E+00	8.99E-01
1995–1996	3.09E+01	9.05E+00	8.99E-01
1996–1997	3.18E+01	7.99E+00	8.66E-01
1997–1998	3.14E+01	7.95E+00	8.62E-01
1998–1999	3.06E+01	8.63E+00	8.80E-01
1999–2000	2.88E+01	1.01E+01	9.03E-01
2000–2001	2.80E+01	1.10E+01	9.26E-01
2001–2002	2.81E+01	1.02E+01	9.03E-01
2002–2003	2.79E+01	1.01E+01	8.98E-01
2003–2004	2.83E+01	9.74E+00	8.89E-01
2004–2005	2.97E+01	9.57E+00	9.03E-01
2005–2006	3.14E+01	8.81E+00	8.94E-01
2006–2007	3.17E+01	8.74E+00	8.94E-01

$$N = N_0(1 - e^{-\alpha P_c^{-\kappa}}), \tag{7}$$

$$N = \int_{P_c}^{\infty} (\mathrm{DRF})dP \tag{8}$$

$$\mathrm{DRF} = N_0 \alpha P^{-\kappa-1} \kappa (e^{-\alpha P^{-\kappa}}), \tag{9}$$

where N_0, α, and κ are free parameters. The values of these three "Dorman parameters" for each survey year from analysis of our fully corrected data for all surveys are shown in Table 1. Our fits to determine Dorman parameters excluded data from rigidities below 0.15 GV, and also excluded all hours of data where the cutoff had not changed from the previous time, i.e., when the ship was not moving. The Dorman function has an analytic derivative which immediately gives the differential response functions, denoted as "DRF" in Equations (8) and (9) and shown in Figures 5(e)–(h), 6(f)–(j), and 7(e)–(h) for different survey years.

As noted in the introduction, the objective of the present paper is focused on characterizing the evolution of GCR spectra from one solar magnetic polarity to the other. In Figure 8 we show the DRF for the two surveys closest to solar minima of opposite magnetic polarity, which clearly manifest the crossover phenomenon. While previous studies have examined the crossover at times near solar minimum, our data for 13 consecutive survey years allow us to examine in detail the transition in spectral shape with solar magnetic polarity. In principle we could address the questions of whether the transition is smooth or abrupt, or a phenomenon that appears as a result of some special conditions that occur only very near solar minimum, and therefore are not part of a continuum at all. However, while solar minimum is a relatively stable time with weak variability, the response functions from other survey years are generally not comparable because of different modulation levels, and cannot be expected to exhibit a literal crossover.

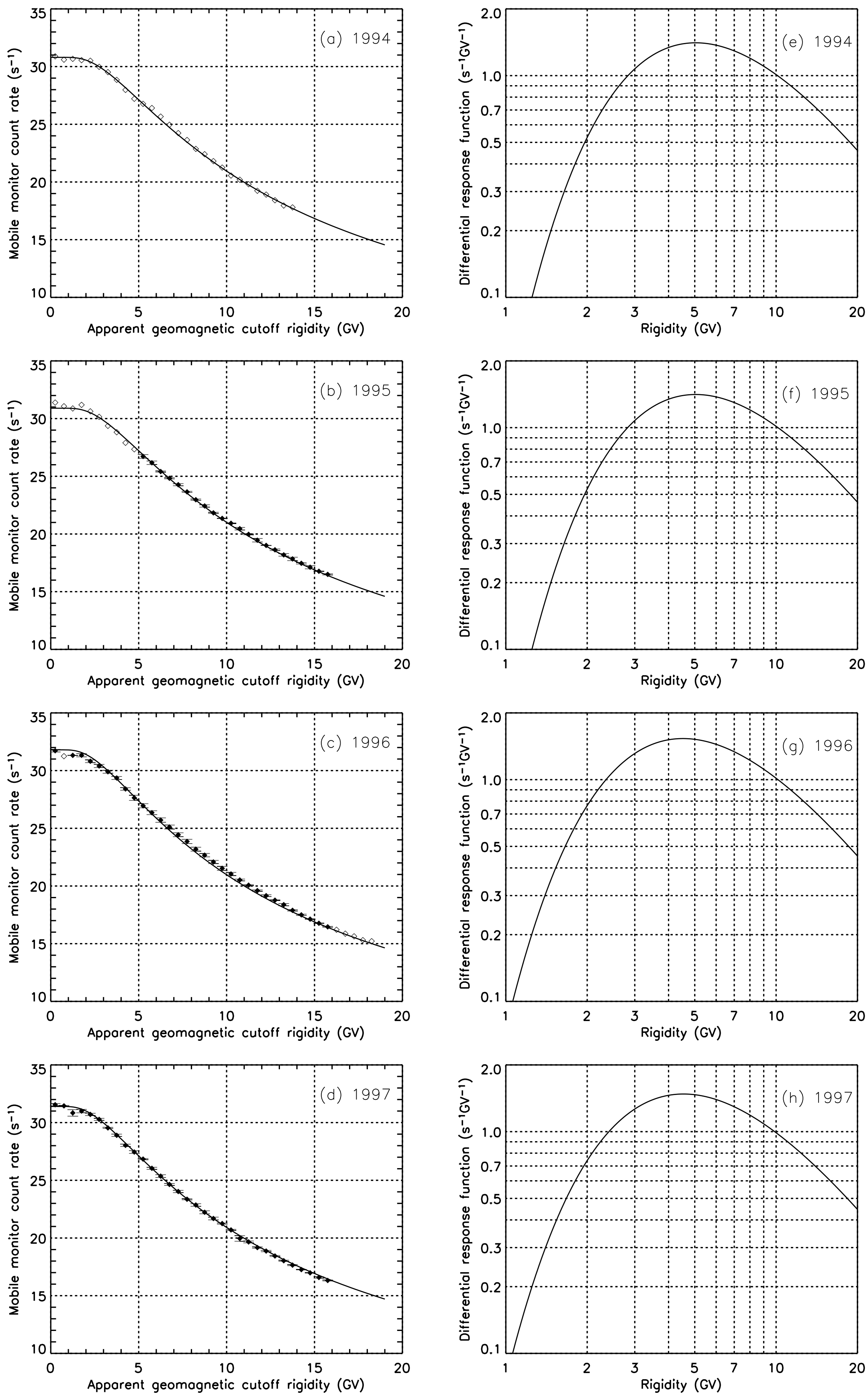

Figure 5. (a)–(d) Corrected mobile neutron monitor count rate for survey years 1994 to 1997 (solar minimum conditions) as a function of apparent geomagnetic cutoff (averaged over 0.5 GV rigidity bins). A solid symbol indicates that data were available for multiple voyage segments (as defined in the text). Vertical error bar represents the standard error between multiple segments; in many cases the error bar is smaller than the plot symbol. An open symbol with no error bar indicates that data were available for only one segment. Solid lines indicate Dorman function fits. (e)–(h) Inferred differential response functions, obtained as the derivatives of the Dorman function fits.

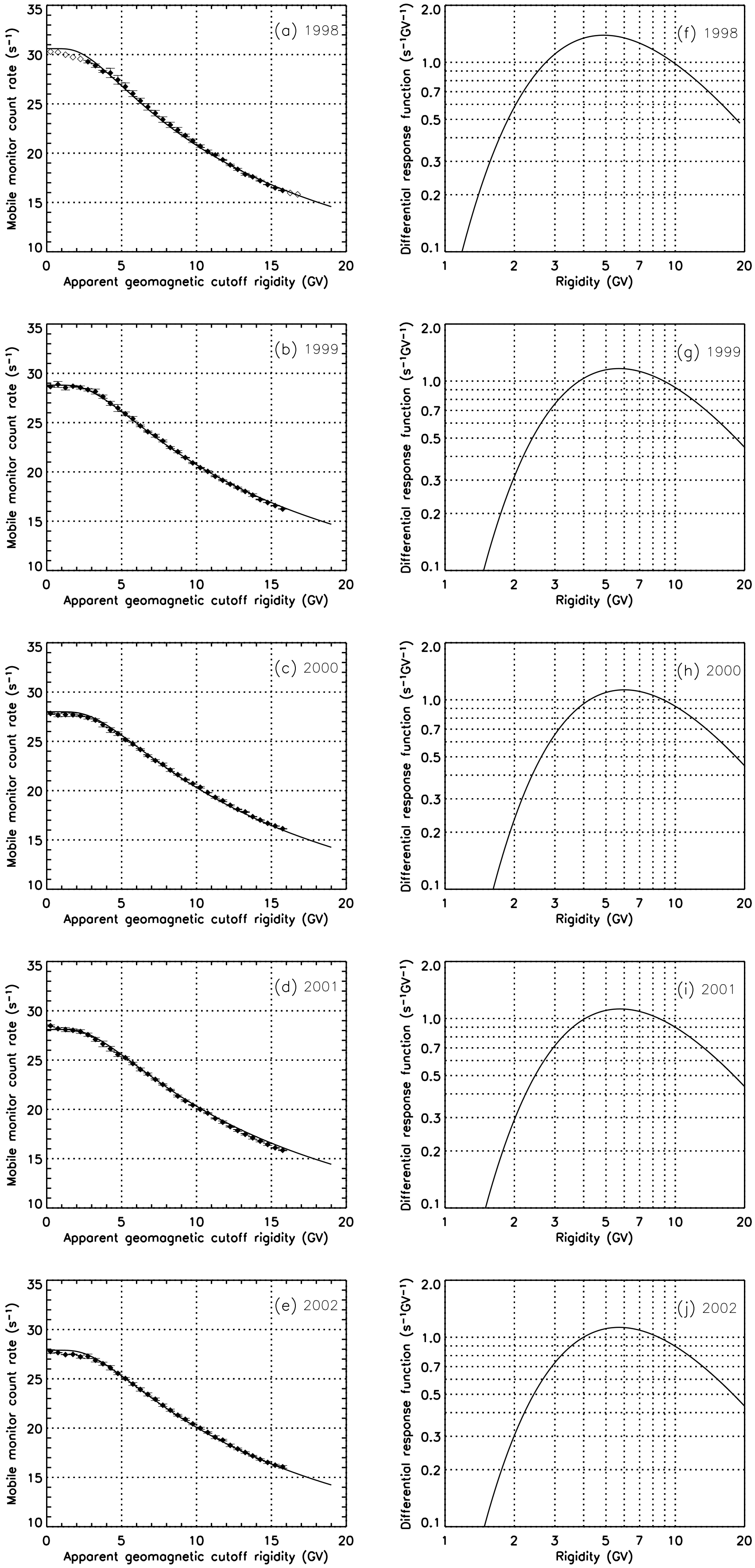

Figure 6. Same as Figure 5, for the 1998–2002 survey years, showing the transition from solar minimum to solar maximum conditions.

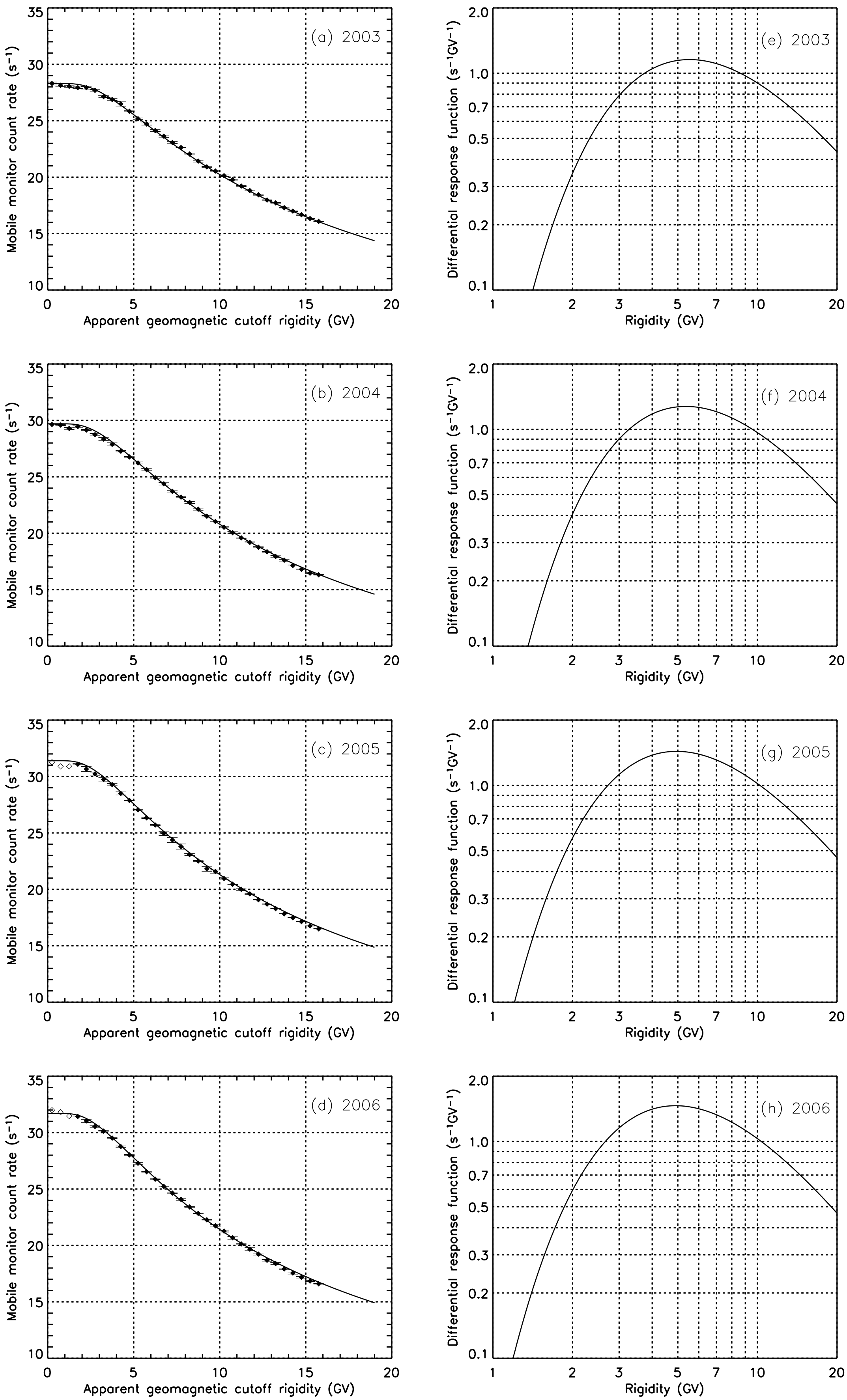

Figure 7. Same as Figure 5, for the 2003–2006 survey years, showing the transition from solar maximum to solar minimum conditions.

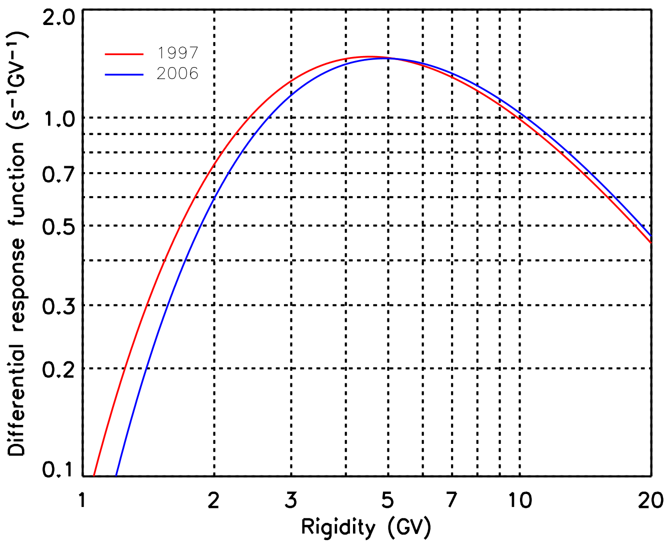

Figure 8. Differential response functions for two survey years, near solar minimum, of opposite polarity and similar modulation level. A crossover is apparent near 5 GV.

Instead, we believe that a startlingly simple analysis shows that the crossover is a natural consequence of a pronounced, persistent shift in the structure of modulation that occurred rather suddenly at the time of the polarity reversal in the year 2000.

Our evidence comes not from the details of the evolution of the DRF, which we leave to future work, but rather from a regression analysis of the mobile monitor count rate for various cutoff rigidity ranges against the contemporaneous pressure-corrected McMurdo count rate. For this analysis we use C_{TP}, the count rate uncorrected for short-term modulation variations, and we still exclude large Forbush decreases. Figure 9 summarizes the analysis, which consists of straight line fits to the data, divided into apparent cutoff rigidity bins of width 1 GV. To cut down on clutter, Figure 9 represents each rigidity bin in each survey by a single point at the average value for all data in all segments, whereas the actual fitting was performed for 1 hr or 30 minute averages.

There were two surprises in this analysis. The first is how well the regression for each individual rigidity bin against McMurdo for a given magnetic polarity is fitted by a straight line. This will be addressed in the following section. The second surprise is how cleanly a systematic change in the slopes of these lines is delineated by the solar polarity reversal in the year 2000. This effect can be seen in Figure 10, where for clarity the data are shown for every third rigidity bin, including all survey years for both magnetic polarity states.

Table 2 gives the coefficients of the regression analysis for each rigidity bin before and after the polarity reversal. Survey year 2000 itself is excluded from the regression analysis, as were all data with a McMurdo count rate below 147 s^{-1}, so as to exclude the transition period. In the table, C_1 is the mobile monitor count rate at a reference McMurdo count rate of 167 s^{-1}, and the years refer to survey years. Note that in the 0–1 GV bin there is, and should be, no noticeable change in slope with the solar polarity reversal. In this rigidity range, both the mobile neutron monitor and the McMurdo neutron monitor (with apparent geomagnetic cutoff ≈0.1 GV) had a yield function dominated by the atmospheric cutoffs, not the geomagnetic cutoff, and the count rates should be directly proportional to one another, regardless of cosmic ray spectral

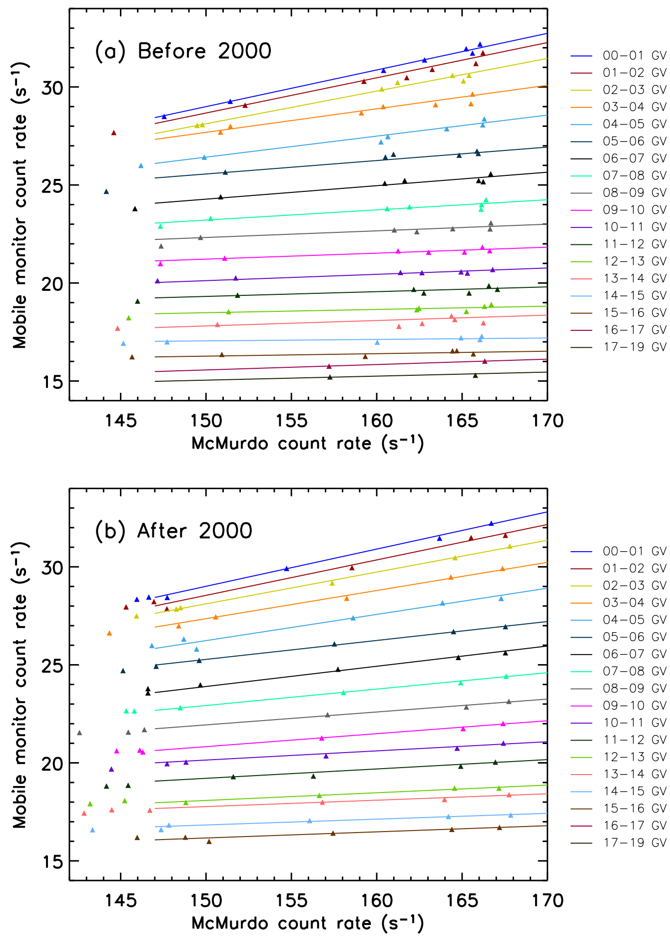

Figure 9. Regression of count rates for the mobile monitor in different cutoff rigidity bins against the count rate of the McMurdo neutron monitor during (a) $A > 0$ solar magnetic polarity (before 2000) and (b) $A < 0$ solar magnetic polarity (after 2000). Symbols indicate average values over each survey year.

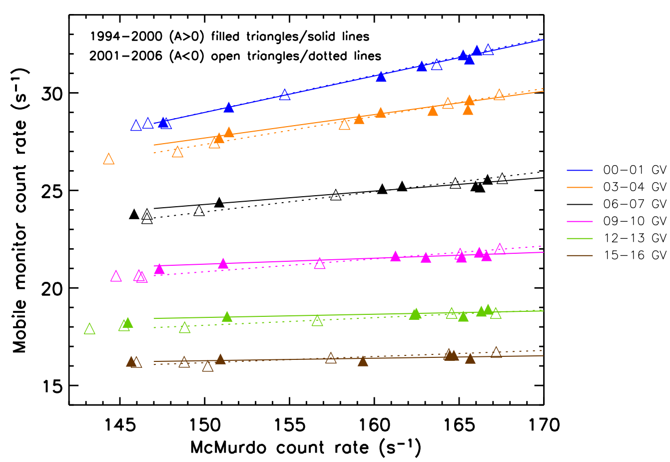

Figure 10. Alternative presentation of the data in Figure 9 using every third rigidity bin for clarity and superimposing the data for different solar magnetic polarities. Filled triangles are used to indicate positive ($A > 0$) solar magnetic polarity with solid lines showing the linear fits. Open triangles indicate data for negative ($A < 0$) solar magnetic polarity while the dotted lines are linear fits to these data. There are clear differences in cosmic ray modulation before and after the solar magnetic polarity reversal.

variations. We do find that the intercept of the linear fit is nearly zero (which is not the case for other rigidity bins). Indeed, the slope in this rigidity range reflects the smaller size of the mobile monitor (3 counter tubes) compared with the McMurdo

Table 2
Coefficients of a Linear Regression Analysis between Mobile Monitor Data in Multiple Cutoff Rigidity Intervals and McMurdo Neutron Monitor Data

Apparent Geomagnetic Cutoff	Best Fit, 1994–1999		Best Fit, 2001–2006	
Rigidity (GV)	C_1	Slope	C_1	Slope
0–1	32.2	0.187	32.2	0.190
1–2	31.7	0.180	31.6	0.181
2–3	31.0	0.167	30.9	0.162
3–4	29.7	0.120	29.8	0.144
4–5	28.2	0.107	28.5	0.134
5–6	26.7	0.069	26.9	0.097
6–7	25.4	0.069	25.7	0.103
7–8	24.0	0.039	24.3	0.084
8–9	22.9	0.034	23.1	0.066
9–10	21.7	0.030	22.0	0.066
10–11	20.7	0.032	20.9	0.047
11–12	19.7	0.024	20.0	0.048
12–13	18.8	0.017	18.7	0.039
13–14	18.3	0.028	18.3	0.033
14–15	17.2	0.007	17.3	0.030
15–16	16.5	0.013	16.7	0.031
16–17	16.0	0.028	...	...
17–19	15.4	0.020	...	...

monitor (18 counter tubes), hence the slope close to $1/6$. For other low rigidity bins the mobile monitor response was still similar to that at McMurdo, with little change in slope upon polarity reversal. Solar modulation is generally weaker at higher rigidity, hence the decreasing slope when the mobile monitor was at higher rigidity. Because there is little polarity difference for low rigidity and a small slope that was difficult to measure for high rigidity, the difference in slope is clearest at intermediate rigidity (3–13 GV).

5. MODELING

To put our observations into the context of conventional modulation analysis, we have examined the predictions of a simple force field model. In particular, we examine whether the linear character of the plots in Figures 9 and 10 is consistent with such a model. The force field model of solar modulation of GCRs is a well known spherically symmetric model of solar wind convection, adiabatic deceleration, and diffusion. The combined effect of these processes is expressed by a single "modulation parameter" that depends only on position (radius) and time. The force field approximation results from a differential equation for the evolution of the phase space distribution function f of cosmic rays as a function of radius and momentum p, where the rigidity spectrum $j(P)$ is related to $f(p)$ by $j(P) \propto P^2 f(p)$. Our modeling approach is based on that of Caballero-Lopez & Moraal (2004).

The force field model itself (Gleeson & Axford 1968) comes from simplifications of the equation to a form that can be solved by the method of characteristics, so that the solution $f(r, p)$ is constant along contours of the characteristic equation. This relates the spectrum at a position within the heliosphere to the spectrum at a different momentum on the outer boundary; this is the LIS. With some further assumptions the characteristics reduce to the form $P_b - P = \phi$, where ϕ is the time-dependent "force field parameter" or "modulation parameter" that expresses a decrease in rigidity from the outer boundary to the location of interest, which is a reasonable approximation

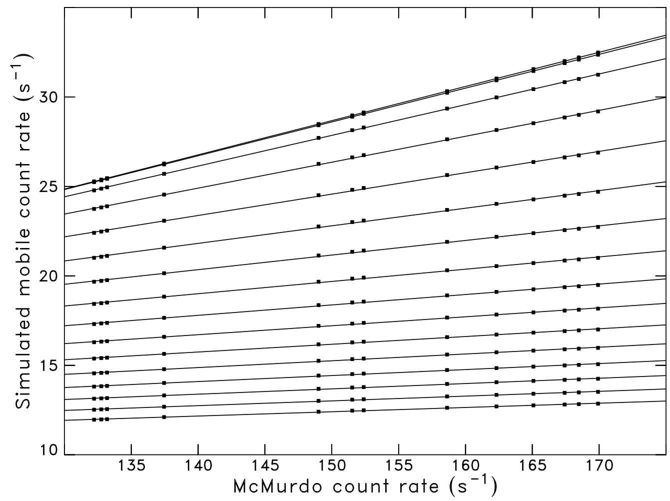

Figure 11. Like Figure 9, for calculations using a force field model as described in the text. Each line is a linear fit to model results for a 1 GV bin in apparent cutoff rigidity, ranging from the 0–1 GV bin at the top to the 15–16 GV bin at the bottom.

to the adiabatic deceleration in some circumstances (Caballero-Lopez & Moraal 2004). While the distribution function f is constant along characteristics, the expression for the rigidity spectrum j has the additional factor of P_b^2/P^2 reflecting the difference in the magnitude of the momentum at the two places, so that

$$j(P) = \frac{P_b^2}{P^2} j_b(P_b), \quad P_b(P, t) = P + \varphi(t). \tag{10}$$

We use yield functions for protons and alpha particles from Clem & Dorman (2000) that are parameterized, along with several other potentially interesting functions, in Table 1 of Caballero-Lopez & Moraal (2012), though we use a different definition of DRF. This paper also has a discussion of the "crossover" and useful formulae for manipulating spectra and yield functions. For the input spectra j_b, we adopted the LIS for protons and alpha particles from the Appendix of Burger et al. (2008); they reported spectra in terms of kinetic energy per nucleon T as a function of rigidity, which we have multiplied by dT/dP to obtain the rigidity spectrum j_b. For the force field parameter we used the tabulated values of Usoskin et al. (2011), which are based on an analysis of neutron monitors with cutoff of 6 GV or less during the neutron monitor era. Specifically, we used their value for December of the survey year (e.g., we used their 2006 December value for survey year 2006).

With these components, via numerical integration we generated a simulated series of surveys. There is an overall absolute normalization to estimate count rates of the mobile neutron monitor and McMurdo neutron monitor. From these simulations we prepared Figure 11, which should be compared to the actual data in Figure 9. As before, each point is plotted according to the count rates for each cutoff rigidity bin during each survey year, with a lower mobile monitor count rate at higher cutoff rigidity. Note that in the simulations, each survey takes place "instantly" and therefore at a constant McMurdo count rate. This is why the points for each survey year are vertically aligned. In the real surveys (Figure 9), the McMurdo count rate varies with time, so the points for a given survey at varying cutoff rigidity follow a more wandering pattern.

These simple simulations, and a comparison between Figures 9 and 11, show that the linear relationship in the plots

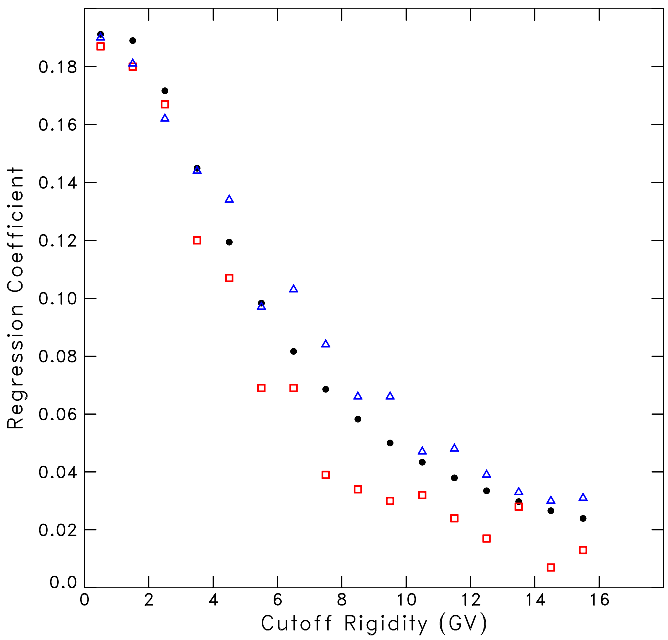

Figure 12. Regression coefficients for the lines in Figures 9 and 11 as a function of apparent cutoff rigidity. Red squares are from fits to data before 2000, blue triangles are from fits to data after 2000, and black circles are from fits to the force field calculation. There are clear differences in modulation before and after the solar magnetic polarity reversal.

is a natural consequence of the force field model over the range of modulation appropriate to the series of latitude surveys. For completeness, we note that the linear relationship fails over an extended (and possibly unphysical) range of modulation parameters. The regression line slopes over the cutoff rigidity range of interest are displayed as solid black circles in Figure 12, whereas our measured slopes from Table 2 are included as open symbols. The measured slopes before 2000 (red squares) and after 2000 (blue triangles) typically bracket the simulated values. Thus we see that the force field model can qualitatively explain solar modulation for a given solar magnetic polarity. However, there is nothing in the force field model that will produce a crossover or any differences in spectra before and after a magnetic polarity reversal. Thus we confirm that there must be a physical difference in the nature of solar modulation upon a change in solar magnetic polarity.

6. DISCUSSION

The crossover as typically defined has to do with spectral comparisons at "similar modulation levels." This is actually not very specific, and usually something like "solar minimum" or "minimum sunspot number" is used. The simple force field model has only one parameter. "Similar modulation levels" are defined as equal values of this parameter, which by construction produce identical DRFs. In practice, the DRF is determined for extended time periods at yearly intervals that often do not have similar modulation levels. Our analysis shows that there is a straightforward way of avoiding this ambiguity. By comparing survey data with contemporaneous data from a fixed neutron monitor (McMurdo), we find a consistent trend with a slope that changes with solar magnetic polarity. The "crossover" is a manifestation of this basic change in modulation that occurs suddenly at the time of the solar polarity reversal. Suddenly, in this context, is of course defined as a year or so—which is also approximately the time it takes for the solar wind to carry the new polarity state to the outer reaches of the solar system. The sudden change of state of the heliosphere is quite like that reported for the relative abundance of protons and antiprotons (Asaoka et al. 2002) and electrons and positrons (Clem & Evenson 2002, 2004). One immediate conclusion from the observation of a rapid change in solar modulation is that we are seeing a phenomenon related to the classic solar wind, not the very distant reaches of the heliosphere at or beyond the termination shock.

Owing to the large detector mass required to detect high-energy cosmic rays, ground-based instruments have remained the state-of-the-art method for studying time variations of these elusive particles (Simpson et al. 1953; McDonald 2000). There have been several direct observations of the full spectrum from balloon-borne and space-based instruments for specific epochs. Typically measurements have not been available for the same instrument and the same level of modulation in opposite polarities, so they have not confirmed a crossover per se. Clem et al. (2003) presented a compilation of proton and alpha particle spectra from several sources. The most direct comparison to the present work is provided by the series of proton spectra from the BESS payload (Mitchell et al. 2008) that span the year 2000 solar polarity reversal with observations from the same instrument. Mitchell et al. (2008) also fit their data to a force field model which, apart from normalization, has only one adjustable parameter. The fits to the BESS data may be good, but they are not perfect. Close examination reveals that the fits to spectra after the polarity reversal are not as good as those to spectra before the reversal. The fits in fact fail by overpredicting the data at low energy, just as we find in our results. We therefore consider that the BESS data are consistent with our results.

Gradient and curvature drifts are clearly established as an important factor in solar polarity dependence of modulation (Jokipii et al. 1977), but in many ways the full implication of this remains unclear (e.g., Potgieter 2013). In and of itself the drift picture does not predict a crossover. It is not intuitively obvious why reversal of the drift fields would increase modulation at one rigidity and decrease it at another. Our observation of a rapid switch from one state to another also indicates that the tilt angle of the heliospheric current sheet is not a major factor. As Clem et al. (2000) discuss, for observations at approximately 1 GV, the effect of current sheet tilt angle is clearly observable in precise measurements but it is tiny compared to the abrupt change with the polarity reversal itself.

There are other modulation effects that are sensitive to interplanetary magnetic field polarity. The "twenty-year wave" observed in the phase angle of the cosmic ray diurnal anisotropy (Forbush 1967; Bieber & Chen 1991) has also been attributed to particle drifts by Levy (1976). However Chen & Bieber (1993) showed that it is a consequence of their finding that the cosmic ray scattering mean-free path is systematically larger during epochs of negative solar polarity than during epochs of positive polarity.

Here we propose a specific mechanism to explain the different spectral evolution for different magnetic polarities. Diffusion coefficients can change radically with solar polarity because of helicity in the solar wind magnetic field and systematically organized magnetic helicity is actually observed in direct measurements of the interplanetary magnetic field. Bieber et al. (1987a) reported that the net helicity (integrated over wave number) of the magnetic field has a definite dominant sign, negative north of the heliospheric current sheet and positive south of it. The helicity dependence of the mean-free path is a function of the product of the sign of the helicity and the polarity of the large-scale field (Bieber et al. 1987b; Bieber & Burger 1990). This means that at any given time the effect upon the mean-free path is the same in both hemispheres (because the helicity and polarity *both* reverse sign across the sheet), but when the solar polarity reverses the magnetic polarity reverses sign, and the helicity does not. This will produce a larger mean-free path during negative solar polarity, consistent with the Chen & Bieber (1993) result.

An indication that magnetic helicity may become more important at neutron monitor energies was provided by Smith & Bieber (1993), who measured the magnetic helicity spectrum extended to very low frequency. They found that for frequencies above 5–10 Hz, the helicity fluctuates in sign and is statistically consistent with zero (Matthaeus & Goldstein 1982). At lower frequencies, however, the helicity assumes a definite dominant sign. It is this low-frequency turbulence that corresponds to wave numbers resonant with cosmic rays in the neutron monitor regime.

This produces a natural explanation for the crossover since the enhanced diffusion coefficients would work in the opposite direction of drifts. During the negative polarity state, when drifts operate to limit fluxes, a larger diffusion coefficient, particularly at the higher energies, permits enhanced access. With such competing effects, each having a different energy dependence, a crossover would be just an observational result that one effect dominates at low energy and the other dominates at high energy.

7. CONCLUSIONS

We have analyzed the results of 13 consecutive yearly latitude surveys during 1994–2007 and determined the yearly response function of the neutron monitor as a function of cutoff rigidity. We have shown that the curious "crossover" in spectra measured at solar minima during epochs of opposite magnetic polarity is actually a manifestation of a sudden change in the behavior of solar modulation at the time of the polarity reversal of the solar magnetic field. We suggest that this results from a systematic change in the interplanetary diffusion coefficient for cosmic rays.

Supported in part by the Thailand Research Fund via Royal Golden Jubilee fellowship PHD/0136/2552 and Basic Research Grant BRG5580001, by the United States National Science Foundation via awards OPP-0838839, PLR-1245939 and their predecessors, and by the Australian Antarctic Division. We thank the officers and crew of the USCG icebreakers Polar Star and Polar Sea for their assistance in conducting the surveys. We also thank Leonard Shulman, James Roth and Keith Bolton for their technical assistance in preparing and maintaining the TasVan. We are grateful to the anonymous referee for thoughtful comments that helped us to develop a clearer presentation of the results.

REFERENCES

Asaoka, Y., Shikaze, Y., Abe, K., et al. 2002, PhRvL, 88, 051101
Bieber, J. W., & Burger, R. A. 1990, ApJ, 348, 597
Bieber, J. W., Burger, R. A., Engel, R., et al. 1999, PhRvL, 83, 674
Bieber, J. W., & Chen, J. 1991, ApJ, 372, 301
Bieber, J. W., Clem, J., & Evenson, P. 1997, in Proc. 25th ICRC (Durban), Vol. 2, 389
Bieber, J. W., Clem, J. M., Duldig, M. L., et al. 2004, JGR, 109, 12106
Bieber, J. W., Evenson, P., & Matthaeus, W. H. 1987a, GeoRL, 14, 864
Bieber, J. W., Evenson, P. A., & Matthaeus, W. H. 1987b, in 20th ICRC (Moscow), Vol. 3, 175
Burger, R. A., Krüger, T. P. J., Hitge, M., & Engelbrecht, N. E. 2008, ApJ, 674, 511
Caballero-Lopez, R. A., & Moraal, H. 2004, JGR, 109, 1101
Caballero-Lopez, R. A., & Moraal, H. 2012, JGR, 117, 12103
Carmichael, H. 1964, IQSY Instruction Manual No. 7 (London: IQSY Secretariat)
Chen, J., & Bieber, J. W. 1993, ApJ, 405, 375
Clem, J., & Evenson, P. 2004, JGR, 109, 7107
Clem, J. M., Bieber, J. W., Evenson, P., et al. 1997, JGR, 102, 26919
Clem, J. M., De Angelis, G., Goldhagen, P., & Wilson, J. W. 2003, AdSpR, 32, 27
Clem, J. M., & Dorman, L. I. 2000, SSRv, 93, 335
Clem, J. M., & Evenson, P. A. 2002, ApJ, 568, 216
Clem, J. M., Evenson, P., Huber, D., et al. 2000, JGR, 105, 23099
Dorman, L. I., Fedchenko, S. G., Granitsky, L. V., & Rishe, G. A. 1970, in Proc. 11th ICRC (Budapest), Vol. 2, 233
Forbush, S. E. 1954, JGR, 59, 525
Forbush, S. E. 1967, JGR, 72, 4937
Garcia-Munoz, M., Meyer, P., Pyle, K. R., Simpson, J. A., & Evenson, P. 1986, JGR, 91, 2858
Gleeson, L. J., & Axford, W. I. 1968, ApJ, 154, 1011
Jokipii, J. R., Levy, E. H., & Hubbard, W. B. 1977, ApJ, 213, 861
Jokipii, J. R., & Thomas, B. 1981, ApJ, 243, 1115
Krüger, H., Moraal, H., Bieber, J. W., et al. 2008, JGR, 113, A08101
Levy, E. H. 1976, JGR, 81, 2082
Lockwood, J. A., & Webber, W. R. 1996, JGR, 101, 21573
Matthaeus, W. H., & Goldstein, M. L. 1982, JGR, 87, 6011
McDonald, F. B. 2000, SSRv, 93, 263
Mitchell, J. W., Abe, K., Fuke, H., et al. 2008, in 30th ICRC (Mérida), Vol. 1, 455
Moraal, H., Potgieter, M. S., Stoker, P. H., & van der Walt, A. J. 1989, JGR, 94, 1459
Oh, S., Bieber, J. W., Evenson, P., et al. 2013, JGR, 118, 5431
Popielawska, B., & Simpson, J. A. 1990, in Proc. 1st COSPAR Colloquium, Physics of the Outer Heliosphere, ed. S. Grzedzielski & D. E. Page (Oxford: Pergamon), 133
Popielawska, B., & Simpson, J. A. 1991, in Proc. 22nd ICRC (Dublin), Vol. 3, 660
Potgieter, M. S. 2013, SSRv, 176, 165
Reinecke, J. P. L., Moraal, H., Potgieter, M. S., McDonald, F. B., & Webber, W. R. 1997, in Proc. 25th ICRC (Durban), Vol. 2, 49
Simpson, J. A., Fonger, W., & Treiman, S. B. 1953, PhRv, 90, 934
Smith, C. W., & Bieber, J. W. 1993, in 23rd ICRC (Calgary), Vol. 3, 493
Thambyahpillai, T., & Elliot, H. 1953, Natur, 171, 918
Usoskin, I. G., Bazilevskaya, G. A., & Kovaltsov, G. A. 2011, JGR, 116, 2104
Webber, W. R., & Lockwood, J. A. 1988, JGR, 93, 8735