A Simplified Ab Initio Cosmic-ray Modulation Model with Simulated Time Dependence and Predictive Capability

K. D. Moloto1, N. E. Engelbrecht1,2, and R. A. Burger1

1 Center for Space Research, North-West University, Potchefstroom, 2522, South Africa
2 National Institute for Theoretical Physics (NITheP), Gauteng, South Africa; n.eugene.engelbrecht@gmail.com

Received 2018 March 6; revised 2018 April 27; accepted 2018 April 27; published 2018 May 30

Abstract

A simplified ab initio approach is followed to model cosmic-ray proton modulation, using a steady-state three-dimensional stochastic solver of the Parker transport equation that simulates some effects of time dependence. Standard diffusion coefficients based on Quasilinear Theory and Nonlinear Guiding Center Theory are employed. The spatial and temporal dependences of the various turbulence quantities required as inputs for the diffusion, as well as the turbulence-reduced drift coefficients, follow from parametric fits to results from a turbulence transport model as well as from spacecraft observations of these turbulence quantities. Effective values are used for the solar wind speed, magnetic field magnitude, and tilt angle in the modulation model to simulate temporal effects due to changes in the large-scale heliospheric plasma. The unusually high cosmic-ray intensities observed during the 2009 solar minimum follow naturally from the current model for most of the energies considered. This demonstrates that changes in turbulence contribute significantly to the high intensities during that solar minimum. We also discuss and illustrate how this model can be used to predict future cosmic-ray intensities, and comment on the reliability of such predictions.

Key words: cosmic rays – diffusion – solar wind – Sun: heliosphere – turbulence

1. Introduction

Astronauts in space encounter severe risks to their health due to their greater exposure to ionizing high-energy cosmic radiation during flights (see, e.g., Badhwar et al. 2001a), in the International Space Station itself (e.g., Cucinotta 2014), and especially when engaged in spacewalks (e.g., Zapp et al. 1998). These risks include an increased probability of getting cancer (Cucinotta & Durante 2006), developing issues with their central nervous systems (Cucinotta et al. 2014), and suffering damage to their eyes (Cucinotta et al. 2001a). This radiation also poses a risk to the integrity of electronic systems (Adams 1985; Holmes-Siedle & Adams 2009), which in turn could lead to catastrophic mission failures. These risks become even greater when possible manned missions to Mars are considered (e.g., Cucinotta et al. 2001b; Hellweg & Baumstark-Khan 2007; Zeitlin et al. 2013). Long transit times equal greater exposure to cosmic rays (CRs; Badhwar et al. 2001b; Wilson et al. 2001), and upon arrival, astronauts would receive almost no shielding from the Martian atmosphere, further increasing their exposure to radiation (Zeitlin et al. 2004). Given these risks, it is essential in the near future to have some indication of expected CR intensities so as to attempt to minimize the radiation exposure of future manned missions (see, e.g., Schwadron et al. 2014).

The development of a model to predict galactic CR intensity is the subject of this study. Such a project would require, due to the complexity of the various processes involved in CR transport and modulation, a fully three-dimensional, energy- and time-dependent study of CR modulation (e.g., Qin & Shen 2017; Strauss & Effenberger 2017, and references therein). This technique has the advantage over finite difference techniques previously used to study CR modulation in that it does not suffer from stability issues, which necessitated a reduction in dimensions or the assumption of a steady state in these prior models (see, e.g., Burger et al. 2008). As to the second condition, most existing models employ ad hoc expressions for diffusion coefficients, which are varied to achieve model agreement with spacecraft observations of CR intensities. Although this approach can lead to excellent agreement with data and some information on the rigidity dependence of diffusion coefficients (see, e.g., Potgieter 1996; Zhang et al. 2007), it is exceedingly difficult to extrapolate the possible future behavior of these diffusion coefficients except in very broad terms. Therefore, the second condition requires an ab initio approach to modulation. In this approach, diffusion and drift coefficients are derived from first principles using various scattering theories (e.g., Teufel & Schlickeiser 2003; Shalchi 2009; Ruffolo et al. 2012; Qin & Zhang 2014), using as inputs for these coefficients results from turbulence transport models (such as those proposed by, e.g., Breech et al. 2008; Oughton et al. 2011; Wiengarten et al. 2016; Weygand et al. 2016; Zank et al. 2017), with outputs in agreement with turbulence observations throughout the heliosphere (for a review of these, see Bruno & Carbone 2013). This has been done with some success by Engelbrecht & Burger (2013a, 2013b), who computed intensity spectra for galactic protons, antiprotons, electrons, and positrons. They found reasonable agreement with existing observations of the same in various parts of the heliosphere, using the same model and diffusion coefficients (taking into account the effects of dissipation range turbulence for low-mass leptons), albeit using a steady-state 3D Alternating Direction Implicit solver for the Parker TPE. The development of a CR modulation model with the particular purpose of providing an estimate of the CR exposure in
spaceflight has been the subject of previous studies (see, e.g., Badhwar & O’Neill 1994; O’Neill 2006; Golge et al. 2015; Miyake et al. 2017), but these modulation codes have generally solved the Parker TPE in one spatial dimension or employed the highly simplified force-field approximation of Gleeson & Axford (1968). This is a severe limitation, as all of the relevant processes involved in CR modulation, such as drifts (see, e.g., Jokipii & Thomas 1981; Kóta 2013; Moraal 2013; Potgieter 2013), simply cannot be taken into account using such approaches.

An ab initio approach, then, can more fully model potential time dependences in diffusion and drift coefficients. This is particularly important for light nuclei, where we expect strong time dependences for turbulence quantities. The present study attempts to construct a first version of just such a model, attempting to simulate large-scale (like the heliospheric magnetic field, HMF) and small-scale (such as the turbulence) heliospheric conditions observed during the last three solar minima and using diffusion and drift coefficients that are theoretically well motivated; in doing so, we try to reproduce simultaneously and self-consistently the galactic CR proton transport model solutions as was done by, e.g., Engelbrecht & Burger (2013a), is motivated by the inherent simplicity of such an approach and the ease of application in concert with the effective approach taken in modeling large-scale quantities in this study. The diffusion tensor used here will be introduced and motivated, and the effects of changes in the large- and small-scale heliospheric quantities on the parallel and perpendicular mean free paths (MFPs) as well as on the turbulence-reduced drift coefficients, will be demonstrated. The final section of this paper will present galactic CR transport calculated using this model for the years 1987, 1997, and 2009, with comparisons to spacecraft observations. Furthermore, the model will be used to make tentative predictions of CR intensities that may be observed during the next solar minimum, using a range of various heliospheric parameters predicted in several studies as motivation.

2. The Transport Model

The Parker (1965) CR transport equation (TPE) is solved here using the stochastic approach outlined by Engelbrecht & Burger (2015b). This technique is discussed in great detail by, e.g., Zhang (1999), Pei et al. (2010), and Strauss & Effenberger (2017). Ignoring sources of energetic particles (like, for instance, the Jovian source of low-energy electrons (Simpson et al. 1974; Eraker 1982)), the Parker TPE is given by

$$\frac{df_0}{dt} = \nabla \cdot \left( K \nabla f_0 \right) - V_{sw} \cdot \nabla f_0 + \frac{1}{3} \left( \nabla \cdot V_{sw} \right) \frac{\partial f_0}{\partial \ln p},$$ (1)

with $f_0(r, p, t)$ the omnidirectional CR phase-space density as a function of particle position $r$, momentum $p$, and time $t$. This quantity is related to the observed CR differential intensity through the relation $j_m^r = T_{oo}^2 f_0$ (see, e.g., Moraal 2013). In this equation, various processes act to modulate an incoming local interstellar spectrum (LIS). These are diffusion, drifts due to gradients and curvatures in the HMF as well as along the heliospheric current sheet (HCS), convection due to the solar wind (with speed $V_{sw}$), and adiabatic energy changes.

The quantity $K$ denotes the diffusion tensor, given in HMF-aligned coordinates as (see, e.g., Burger et al. 2008)

$$K = \begin{bmatrix} \kappa_1 & 0 & 0 \\ 0 & \kappa_{1,2} & \kappa_A \\ 0 & -\kappa_A & \kappa_{2,3} \end{bmatrix}.$$ (2)

Off-diagonal elements denote drift coefficients, while diagonal elements denote diffusion coefficients parallel and perpendicular to the background HMF. Note that these diffusion and drift coefficients can be related to a length scale such that $\kappa = v\lambda/3$, with $v$ the particle speed. In the stochastic approach to solving Equation (1), the equation can be written in terms of a set of equivalent Itô stochastic differential equations (see, e.g., Zhang 1999; Gardiner 2004; Strauss & Effenberger 2017),

$$dx_i = A_i(x_i)dt + \sum_j B_i(x_i) \cdot dW_j,$$ (3)

with $i \in \{r, \theta, \phi, E \}$, $x_i(t)$ describing Itô processes and $dW_j$ describing Wiener processes. The quantities $A_i$ and $B_i$ are
treated in exactly the same manner as described by Engelbrecht & Burger (2015b). In the present study, Equation (3) is solved as described by Engelbrecht & Burger (2015b) in a time-backward manner. In this approach, the evolution of $N$ pseudoparticles in phase space is iteratively traced from an initially specified point until they exit at a boundary, where-upon the average CR intensity at the initial point is calculated using (see, e.g., Strauss et al. 2011b; Engelbrecht & Burger 2015b; Strauss & Effenerger 2017)

$$j(x^o_i, t^o) = \frac{1}{N} \sum_{k=1}^{N} j_B(x^o_{i,k}, t^o_k),$$

(4)

where $(x^o_i, t^o)$ and $(x^o_{i,k}, t^o_k)$ denote the initial and final phase-space points and times, respectively, and $j_B$ is the boundary intensity. In the present study, we choose $N = 10,000$ pseudoparticles per energy bin, following, e.g., Strauss & Effenerger (2017) and motivated by the fact that the statistical error is proportional to $1/\sqrt{N}$ (see, e.g., Strauss et al. 2011a and references therein).

This study, like that of Engelbrecht & Burger (2015b) and Qin & Shen (2017), employs a boundary spectrum set inside of the nominal location of the heliospheric termination shock, in this case at 85 au. A similar approach is taken by Guo & Florinski (2016). This is due to the fact that a considerable amount of modulation has been observed to occur in the heliosheath (see, e.g., McDonald et al. 2000; Caballero-Lopez et al. 2010; Stone et al. 2013; Zhang et al. 2015). The input spectrum used in this study has been constructed to agree with Voyager observations reported by Webber et al. (2008) at 85 au and is given by

$$j_B(85 \text{ au}) = \frac{17.0(P/P_o)^{-2.4}}{2.2 + 2.1(P/P_o)^{1.7}},$$

(5)

in units of particles m$^{-2}$ sr$^{-1}$ MeV$^{-1}$, $P_o = 1$ GV, and $P$ in GV. As a first approach, we ignore the effects of charge-sign-dependent modulation on this input spectrum. Future studies will take this into account.

The present study seeks to model heliospheric conditions over consecutive solar minima. As such, during these periods, the solar wind speed has been observed to show a latitudinal dependence, assuming values of ~800 km s$^{-1}$ over the poles and ~400 km s$^{-1}$ in the ecliptic plane (see, e.g., McComas et al. 2000). This is modeled as a function of colatitude $\theta$ using a hyperbolic tangent function,

$$V_{sw}(\theta) = 400\left\{\begin{array}{ll}
\frac{3}{2} & -\frac{1}{2} \tanh[8(\theta - \pi/2 + \alpha + \delta_l)], \quad \theta \leq \pi/2; \\
\frac{3}{2} & + \frac{1}{2} \tanh[8(\theta - \pi/2 - \alpha - \delta_l)], \quad \theta > \pi/2
\end{array}\right.,$$

(6)

in units of km s$^{-1}$, with $\delta_l = \pi/9$ radians and $\alpha$ denoting the heliospheric tilt angle. The HMF is here described by the Parker (1958) field. The temporal variations of these large-scale plasma quantities, including those of the HCS tilt angle, are modeled using an effective-value approach (see, e.g., Nagashima & Morishita 1980a, 1980b). This approach takes into account the fact that it takes the solar wind and the HMF embedded in it over a year to reach the outer limits of the heliosphere (Nagashima & Morishita 1980b). Thus, even if a CR particle could traverse the whole heliosphere instantaneously, it would still experience approximately at least the last year’s worth of the solar wind, magnetic fields, and tilt angles. In this study, the intensities measured at a certain time are associated with the average of at least the preceding year’s worth of tilt angle values as measured close to the Sun. Thus, any intensity measured on a particular day must be associated with a running average of at least the previous year’s worth of tilt angle, magnetic field, and solar wind data, a fact supported by the heliospheric residence times of CRs reported by Strauss et al. (2011b). Figure 1 shows radial model values for the tilt angle as found on the Wilcox Solar Observatory Web site. The spot value, indicated by the blue line, shows the tilt angle as it is measured, and the 16 and 20 month effective tilt angles show the running average calculated over the 16 and 20 month window preceding the date shown, respectively, with these interval lengths being informed, as noted above, by CR residence times as well as by the time it takes the plasma to propagate from Earth out to 100 au. It is these 16 month-averaged tilt angles that we use as inputs for this quantity pertaining to particular points in time in the solar cycle. The short-term variations in the tilt angle are smoothed out, which is to be expected since we are applying an average. A consequence of applying the effective tilt angle is that it leads to a time shift with respect to the original data. The time shift is one-half of the period over which the effective tilt angle is calculated, thus taking a 16 month effective tilt angle is similar to taking an eight-month time shift. These time-shifted tilt angles are then used as inputs for the heliospheric current sheet, the angular extent of which is modeled here as in Burger (2012) and Engelbrecht & Burger (2015b) as

$$\theta_{nt} = \frac{\pi}{2} \tan^{-1}(\tan \alpha \sin \phi^*),$$

(7)

where $\phi^* = \phi + r \Omega / V_{sw}$. Current sheet drifts and drifts due to gradients and curvatures in the HMF are dealt with as proposed by Burger (2012). An identical effective-value approach is used for the solar wind speed and HMF magnitude at Earth, taken from OMNI spacecraft observations, and shown in the middle and bottom panels of Figure 1. More specifically, effective values for the HMF magnitude, solar wind speed, and tilt angle during the solar minimum periods of interest to this study are listed in Table 1.

As to the diffusion tensor, we assume as a point of departure that the composite slab/2D model for turbulence is valid (e.g., Bieber et al. 1994). The parallel MFP used here is constructed from the Quasilinear Theory (QLT; Jokipii 1966) results derived by Teufel & Schlickeiser (2003) and by Burger et al. (2008), and subsequently employed in several numerical modulation studies (e.g., Engelbrecht & Burger 2013a, 2015b). This parallel MFP expression, derived assuming a slab turbulence spectrum with a wavenumber-independent energy-containing range and a Kolmogorov inertial range with spectral index $-s = -5/3$, is given by

$$\lambda_J = \frac{3s}{(s-1) k_m} \left(\frac{R l_{pp}}{6 4 \pi^2 k_m} \left[ \frac{1}{4 \pi} + \frac{2 R^{-s}}{(2 - s)(4 - s)} \right] \right),$$

(8)

where $R = R_{l_{pp}}$, with $k_m$ the wavenumber at which the inertial range on the assumed slab spectrum commences, and $R_{l_{pp}}$
denotes the maximal proton Larmor radius. Furthermore, $\delta B^2$ denotes the total slab variance, while $B_0$ denotes the uniform background field.

To model the perpendicular MFP, the results for $\lambda_\parallel$ are used as inputs for the expression derived from the Nonlinear Guiding Center (NLGC) theory first proposed by Matthaeus et al. (2003) and employed in modulation studies by Burger et al. (2008), who modify the result presented by Shalchi et al. (2004) to take into account an arbitrary ratio of slab to 2D energy. This result is derived for a 2D turbulence power

Figure 1. Effective tilt angles (top panel), HMF magnitudes (middle panel), and solar wind speeds (bottom panel) at Earth employed in this study. The hatched bars separate the numbered solar cycles, and the gray bars denote periods of full solar maximum. Red and green lines indicate lagged values. Spot values (blue lines) indicate unlagged quantities as observed at Earth. See the text for details.
spectrum assumed to consist of a flat energy-containing range and a Kolmogorov inertial range only. This spectral form is not entirely realistic (see Matthaeus et al. 2007), and expressions for $\lambda_\perp$ assuming more realistic input power spectra have been previously derived (see, e.g., Shalchi et al. 2010; Engelbrecht & Burger 2013a, 2015b) by employing more recent scattering theories (see, e.g., Shalchi 2009, 2010; Qin & Zhang 2014). However, the NLGC expression still provides a tractable analytical expression that does not differ too greatly from the result derived by Engelbrecht & Burger (2013a) for a similar, yet more physically motivated, spectrum. The NLGC perpendicular MFP is given by

$$\lambda_\perp = \left[ \alpha^2 \sqrt{3\pi} \frac{2\nu - 1}{\nu} \frac{\Gamma(\nu)}{\Gamma(\nu - 1/2)} \lambda_{2D} \frac{\delta B^2}{B_0^2} \right]^{2/3},$$

(9)

where $\nu = 5/6$ denotes half of the assumed inertial range spectral index, $\delta B^2$ the total 2D variance, and $\lambda_{2D}$ the length scale corresponding to the wavenumber at which the inertial range on the assumed 2D turbulence power spectrum commences. We follow Matthaeus et al. (2003) in assuming that $\alpha^2 = 1/3$, based on the results of their numerical test-particle simulations of the perpendicular diffusion coefficient.

Numerical test-particle simulations (see, e.g., Minnie et al. 2007b; Tautz & Shalchi 2012) and theory (e.g., Burger 1990; Jokipii 1993; Fisk & Schwadron 1995; le Roux & Webb 2007) show that CR drift coefficients are reduced from the weak-scattering value $\kappa_A = \nu R_L / 3$ (Forman et al. 1974) in the presence of magnetic turbulence. Modeling this self-consistently, however, has proven to be difficult (see Engelbrecht & Burger 2015a, and references therein). In this study, an expression for the turbulence-reduced drift coefficient derived from first principles by Engelbrecht et al. (2017) is employed, providing results in reasonable agreement with numerical test-particle simulations for the range of turbulence conditions expected in the supersonic solar wind. Here, the length scale corresponding to the drift coefficient is given by

$$\lambda_D = R_L \left[ 1 + \frac{\lambda^2_\perp \delta B^2}{R_L^2 B_0^2} \right]^{-1},$$

(10)

where $\delta B^2$ denotes the total (slab and 2D) transverse variance. In conditions where turbulence levels are very low, this expression reduces to the Larmor radius, which is the weak-scattering drift length scale. Note that it is a function of the perpendicular MFP, and thus, from Equation (9), also a function of the parallel MFP.

The above diffusion tensor requires as inputs values for various turbulence quantities, such as the slab and 2D magnetic variances and turnover scales, throughout the heliosphere. The approach of this study entails using parameterized fits to the radial and colatitudinal profiles of the turbulence quantities yielded by the two-component turbulence transport model proposed by Oughton et al. (2011), as solved by, e.g., Engelbrecht & Burger (2013a) for generic solar minimum conditions. These fits are adjusted to be in agreement with observations of various turbulence quantities in different parts of the heliosphere as well as during different solar minima.

The magnetic variances reported by Zank et al. (1996) can be fitted with a simple power law as a function of radial distance. In doing this, however, the contribution to the slab variance from waves generated due to the formation of pickup ions (see, e.g., Zank 1999; Isenberg 2005) is omitted. This contribution is expected to theory to predominate at high wavenumbers (e.g., Williams & Zank 1994). This is borne out by some observations (e.g., Cannon et al. 2014; Aggarwal et al. 2016; Cannon et al. 2017) and thus would not greatly affect the transport of the highly energetic galactic CR protons considered in this study as they should not affect the level of the slab fluctuation spectrum in the inertial and energy-containing ranges. However, the effects of such waves on lower-energy particles may perhaps be significant (Engelbrecht 2017). The power law used in this study to model the total variance is given by

$$\delta B^2_T = \delta B^2 \left( \frac{r}{r_0} \right)^{\epsilon_1},$$

(11)

where $\delta B^2_T$ is the value this quantity assumes at Earth ($r_0 = 1$ au), and $\epsilon_1$ is a constant. Effective values for $\delta B^2_T$ corresponding to the solar minimum years considered here are listed in Table 1, as reported by Burger et al. (2014). Note that stream-shear effects due to latitudinal increase of the solar wind speed (see, e.g., Breech et al. 2008) are not taken into account in this treatment, in contrast to what is done by Engelbrecht & Burger (2013a, 2015b) when full solutions to the Oughton et al. (2011) TTM are taken into account. The slab/2D anisotropy (see, e.g., Bieber et al. 1994) assumed in the ecliptic plane is that reported by Bieber et al. (1996; although these values vary considerably; see, e.g., Oughton et al. 2015). This ratio is held to a different value over the poles, as turbulence in the fast solar wind has been found to be different from that observed in the slow solar wind (Bavassono et al. 2000a, 2000b). Motivated by the findings of Dasso et al. (2005), who report a preponderance of fluctuations with wavenumbers quasi-parallel to the background magnetic field, we assume a 90/10 slab/2D ratio at high latitudes. Values for $\epsilon_1$ are chosen differently in the ecliptic region as opposed to those in the poles, since they are motivated by outputs yielded by the Oughton et al. (2011) TTM. These, and the values assumed for the slab/2D ratios, are listed in Table 2. Variances at 1 au over the poles are scaled up by a factor of 2 from the corresponding ecliptic values in Table 1 using a hyperbolic tangent function of the form of Equation (6), following the Ulysses observations of increased variances at high latitudes reported by, e.g., Forsyth.
et al. (1996) and Erdős & Balogh (2005). This approach to the latitudinal dependence of magnetic variances is markedly different from that taken by Qin & Shen (2017), who argue, based on observations reported by Perri & Balogh (2010), that this quantity would decrease as one moves toward the polar regions. We choose instead to follow the observations of, e.g., Erdős & Balogh (2005), as such a scaling has been found, when used in the numerical CR modulation model of Engelbrecht & Burger (2013a), to yield galactic CR proton latitude gradients in reasonable agreement with Ulysses observations of the same, as reported by Heber et al. (1996). The variances thus modeled are shown as a function of heliocentric radial distance in the top panels of Figure 2, in the ecliptic plane (right panel) and over the poles (left panel). In the inner heliosphere in the ecliptic plane, the modeled variances fall well within the range of Voyager observations as reported by Zank et al. (1996), whereas larger values are assumed over the poles.

The spatial dependence of the slab correlation scale yielded by the Oughton et al. (2011) TTM is the most complicated of all the turbulence quantities considered, as solutions where the effects of pickup-ion fluctuations are not ignored are parameterized here. This is under the assumption that, although fluctuations driven by pickup-ion formation may not affect the slab spectral level at the lower wavenumbers where galactic CR protons resonate, they will still affect the correlation function and hence the correlation scale. This is modeled as a combination of power laws that are a function of radial distance in three stages, each with its own power-law index:

$$\lambda_s = \lambda_s^E \frac{r}{r_0} \left[ 1 + \left( \frac{r}{r_1} \right)^{s_1} \frac{s_1}{\nu_1} \right] \left[ 1 + \left( \frac{r}{r_2} \right)^{s_2} \frac{s_2}{\nu_2} \right],$$

(12)

where $\lambda_s^E$ is the value of the 2D correlation scale at Earth, at a radial distance $r = r_0$, under the assumption that the slab correlation scale at Earth is $\sim 2.3$ times the 2D scale (within the error bars of the observations reported by Weygand et al. 2011); $s_1$, $s_2$, and $s_3$ are the exponents of the three different radial dependences; $r_1$ and $r_2$ are the radial distance where the radial dependence changes from $r^{s_1}$ to $r^{s_2}$ to $r^{s_3}$ respectively; and $f_1(\geq 0)$ and $f_2(\geq 0)$ determine how sharp these transitions are. Large values for $f$ result in abrupt transitions, while smaller values ensure a smoother transition between the different stages.
Values result in smoother transitions. Values for the parameters used are given in Table 3. In contrast to the slab correlation scale, the 2D correlation scale is modeled as a single power law with indices following the radial dependence of the Oughton et al. (2011) TTM and remaining within the range of the observations of this quantity reported by Smith et al. (2001) such that

$$\lambda_{2D} = \lambda_{2D}^E \left( \frac{r}{r_0} \right)^{f_1}$$

with its value at Earth, as well as fitting parameters, given in Table 4. Observed correlation scales for components of the magnetic field at Earth (Wicks et al. 2013) show virtually no change from one solar minimum to the next. Therefore, the 1 au values employed here are kept the same for each solar minimum period considered here. Note that for both slab and 2D correlation scales, different fitting parameters are used in the polar regions, with values at 1 au set to have the ratio of the 2D to slab correlation scales agree with that reported by Dasso et al. (2005) and Weygand et al. (2011) for fast solar wind speed data intervals at Earth. This is motivated by observations indicating that the behavior of turbulence in these conditions is similar to that in the polar regions of the inner heliosphere, where the fast solar wind dominates (Bavassono et al. 2000a, 2000b). Latitudinal changes for both correlation scales are modeled using a hyperbolic tangent function of the form of Equation (6), as with the variances, to change values for $\lambda_{2D}^E$ and $\lambda_{2D}$.

The bottom panels of Figure 2 show the slab and 2D correlation scales as a function of radial distance in the ecliptic plane and over the poles. Note the decrease in the slab correlation scale beyond ~4 au for both the colatitudes shown in the figure, which, as noted above, models the effects of pickup-ion formation and occurs at different radial distances depending on latitude, reflecting the outputs yielded by the Oughton et al. (2011) TTM. This fit also takes into account that the slab correlation scale modeled by Engelbrecht & Burger (2013a) relaxes in the outer heliosphere at the resonant scale corresponding to the wavenumber at which the energy due to the formation of pickup ions is injected into the slab fluctuation spectrum (see, e.g., Oughton et al. 2011). The same panels show the monotonically increasing 2D correlation scale, which is consistent with the consistently decreasing 2D variance shown in the top panels of Figure 2. We also show in Figure 2 the corresponding solutions to the full Oughton et al. (2011) TTM as solved by Engelbrecht & Burger (2013a, 2015b) for generic solar minimum conditions usually assumed in CR modulation studies and not specific to a particular solar minimum, as required in this study. The variance scalings employed in this model do not greatly differ in magnitude from those yielded by the TTM, but reflect the different solar-cycle-specific values at Earth employed for this quantity. In the ecliptic plane (top-left panel of Figure 2), the radial dependences are somewhat different, but nevertheless yield results within the spread of the Zank et al. (1996) observations. As to the correlation scales in the ecliptic plane (bottom-left panel of Figure 2), the approximations and full solutions show very similar radial dependences, with deviations in magnitude well within the error bars of the Weygand et al. (2011) observations relevant to the slow solar wind. Over the poles (bottom-right panel of Figure 2), the parameterized solutions differ in magnitude from the TTM solutions, due to the fact that the parameterized solutions were set so as to agree with the Weygand et al. (2011) observations relevant to the fast solar wind. The radial dependences of both the TTM outputs and the parameterized solutions remain, however, very similar.

The effects of the turbulence quantities described above on the spatial and rigidity dependences of the parallel and perpendicular MFPs as well as the corresponding drift length scales described by Equations (8)–(10) are illustrated in Figure 3 for the parameters corresponding to the solar minima of 1987, 1997, and 2009, respectively. The quantities corresponding to these years are denoted in this figure by solid, dotted–dashed, and dashed lines, respectively, while the parallel MFP, perpendicular MFP, and drift scale curves are blue, orange, and green, respectively. Overall, these results are very similar to those reported by Engelbrecht & Burger (2013a), confirming that the current approach, regardless of its simplicity, is a reasonable alternative to the full TTM. The top panel of Figure 3 shows the radial dependences of the 1 GV values of these quantities in the ecliptic plane. Below ~10 au, $\lambda_{1}$ remains relatively constant as a function of radial distance, as the monotonic decrease in the slab variance is matched by a corresponding increase in the slab correlation scale. Beyond this distance, however, the effects of the pickup-ion-induced decrease in the slab correlation scale modeled in Equation (12), combined with the continued decrease in the slab variance, cause the parallel MFP to increase steeply until ~40 au, where the increase in the slab correlation scale seen in Figure 2 causes $\lambda_{1}$ to flatten out somewhat. The increase of $\lambda_{\perp}$ with radial distance is less prominent, as the 2D correlation scale is here modeled to increase monotonically with radial distance. The slight kink in the perpendicular MFP beyond ~10 au is, however, due to the $\lambda_{3/5}^{1/3}$ dependence seen in Equation (9). The drift scale at this rigidity assumes weak-scattering values beyond ~2 au, differing between the different solar minima due to the different values assumed for the HMF magnitude, as listed in Table 1. This is simply due to the fact that, from Equation (10), turbulence levels are too low to significantly reduce the drift coefficient, only becoming large enough to do so in the very inner heliosphere. Note that, for the parameters used here, the drift scale becomes larger than the perpendicular MFP at ~2 au. In terms of the temporal differences in these quantities, the 1987 parameters yield the smallest drift scales.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>TTM Results Pertaining to the Slab Correlation Scales in the Ecliptic Plane as well as at High Heliographic Latitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{E}^{1}$ (au)</td>
<td>$\epsilon_1$</td>
</tr>
<tr>
<td>Ecliptic</td>
<td>$1.54 \times 10^{-3}$</td>
</tr>
<tr>
<td>Polar</td>
<td>$6.7 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>TTM Results Pertaining to the 2D Correlation Scales in the Ecliptic Plane as well as at High Heliographic Latitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{2D}$</td>
<td>$\epsilon_{2D}$</td>
</tr>
<tr>
<td>Ecliptic</td>
<td>$6.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>Polar</td>
<td>$9.4 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
and parallel MFPs, while the 2009 parameters yield the largest values for these quantities, effectively bounding the curves for the 1997 parameters. The perpendicular MFPs for these years remain very similar, with only slight changes discernible.

Considering the rigidity dependences of these quantities at Earth, as shown in the middle panel of Figure 3, the proton parallel MFP shows the expected $P^{1/3}$ dependence throughout the rigidity range considered and remains slightly above the Palmer (1982) consensus range (black box) for the rigidity range considered in this study. The Palmer consensus range, however, does not take into account possible solar-cycle dependences of the MFPs (see, e.g., Bieber et al. 1994 for more details). Chen & Bieber (1993), from an analysis of CR intensities, and Burger et al. (2014) and Zhao et al. (2018), from direct analyses of solar wind turbulence, do, however, report larger MFPs during solar minimum, in qualitative agreement with what is reported here. The perpendicular MFP displays a relatively flat $P^{1/9}$ dependence, in accordance with the Palmer consensus and as expected from Equations (8) and (9). Turning to the drift scale, this quantity displays the $P^4$ dependence expected of the weak-scattering length scale (which is equal to the Larmor radius of the particle in question) beyond $\sim 2$ GV, deviating significantly from that dependence below this rigidity. This implies that, for the parameters of this model, drift-reduction effects due to turbulence will only play a significant role in the transport of lower-energy CRs. Temporally, the picture here is the same as when the radial dependences are considered, with the use of the 1997 parameters yielding results intermediate between the larger 2009 and the smaller 1987 parallel MFPs and drift scales, with relatively little effect on the perpendicular MFPs.

Shown as a function of colatitude at 1 au in the bottom panel of Figure 3, the 1 GV parallel MFP assumes values considerably larger in the ecliptic plane than over the poles, as expected from the slab variance dependence of Equation (8) and in qualitative agreement with the findings of Erdös & Balogh (2005). The perpendicular MFP behaves in a similar fashion, which, from the variance dependence in Equation (9) and the fact that the 2D variance in the polar regions is here modeled to be larger than that in the ecliptic plane, would appear to be counterintuitive, but is simply due to the $\lambda_{\perp}^{1/3}$ dependence of $\lambda_{\perp}$. Although the colatitudinal dependence of these MFPs is by construction very similar to that of these quantities as reported by Engelbrecht & Burger (2013a) as seen in Figure 5 of that paper, with larger values for $\lambda_{\parallel}$ in the ecliptic plane than over the poles, and with values for $\lambda_{\perp}$ in the ecliptic plane being smaller than those over the poles. Note that for the MFPs used in this study, there are no increases in these quantities at intermediate colatitudes as reported by Engelbrecht & Burger (2013a), as stream-shear effects due to the latitudinal increase of the solar wind speed are not taken into account here. Drift scales behave quite differently from the perpendicular MFP, assuming larger values over the poles than in the ecliptic plane, reflecting the change in HMF magnitude more than the increase in turbulence levels due to the larger Larmor radius of 1 GV CRs. For lower-energy particles, this changes, with smaller drift scales over the poles than in the ecliptic plane.

In the following section, results from the complete modulation model described here will be presented.
3. Modulation Results and Discussion

Galactic CR proton intensities at Earth computed for the three different solar minimum years considered in this study, using the model as described above, are shown in the top panel of Figure 4. Observations shown are from IMP-8 (McDonald et al. 1992) and PAMELA (Adriani et al. 2013) reported at Earth, and Voyager 2 at 85 au (Webber et al. 2008). The light blue line represents the boundary spectrum used in the modulation code, Equation (5), which passes through the Voyager data, as it is constructed to do. Model results for the solar minima of 2009 ($A < 0$) and 1997 ($A > 0$) are in excellent agreement with observations at all energies considered, while those for 1987 ($A < 0$) agree best with observations at higher energies. At higher energies (beyond $\sim$0.3 GeV), the 1987 results are slightly larger than those for 1997, in accordance with neutron monitor observations (see, e.g., Potgieter 2008 and references therein), with the opposite being true at lower energies, characteristic of the effects of drifts on CR modulation (e.g., Kota & Jokipii 1983). Given that the turbulence input to the present modulation model is based on reasonable assumptions and is guided by observations, the fact that the present model yields results in good agreement with the unusually high 2009 intensities leads us to conclude that the higher than expected CR intensity during the 2009 solar minimum can be quantitatively linked to turbulence parameters that differ from solar minimum to solar minimum, in agreement with the conclusions drawn by Zhao et al. (2014) and Moloto (2015).

Figure 4. Top panel: computed galactic cosmic-ray proton intensities at Earth for the years 1987 (dark blue), 1997 (red), and 2009 (green). The light blue line indicates the input spectrum used (Equation (5)) at 85 au. Also shown are spacecraft observations for the relevant periods, as reported by McDonald et al. (1992; IMP-8), Adriani et al. (2013; PAMELA) and Webber et al. (2008; Voyager 2). Bottom panel: the same, but with the gray range denoting possible predicted intensities during the next solar minimum when the magnetic variance and HMF magnitude is varied up (gray squares) or down (gray circles) by 20%. The orange line denotes intensities calculated for an $A > 0$ magnetic polarity cycle under the assumption of heliospheric conditions identical to those prevalent in 2009.
Given the ab initio nature of the present model and its ability to reproduce observed intensity spectra during the previous three solar minima, the question arises as to what predictions can be made for the next solar minimum. This process, however, requires making extrapolations regarding the behavior of the HMF magnitude, tilt angle, and turbulence quantities during that time. Hathaway & Upton (2016) and Cameron et al. (2016), using surface flux transport models to predict the Sun’s axial dipole strength during the next sunspot cycle minimum, argue that solar cycle 25 would be very similar to solar cycle 24. Therefore, a reasonable point of departure would be to assume that all the parameters relevant to modulation remain as they were modeled for 2009, except that now, a positive magnetic polarity cycle is assumed. The differential intensities calculated thus are shown as the orange dashed line on the bottom panel of Figure 4. Interestingly, the $A < 0$ 2009 spectrum remains larger than the equivalent $A > 0$ spectrum down to $\sim 0.07$ GeV. As this crossover occurs at larger energies when observations from previous solar minima are considered, from the simulations of Reinecke & Potgieter (1994), this implies that for this set of parameters, diffusion effects play a larger role than drift effects. As there is a large degree of uncertainty when predictions on future levels of solar activity are made (Cameron et al. 2016), and indeed it has even been predicted that cycle 25 could be even less active than cycle 24 (Ahluwalia 2016), the model was run with a change in HMF magnitude and variance of 20%. This yields the gray band in Figure 5.
Figure 4, where the gray squares denote the solution with a 20% larger HMF magnitude and variance and the gray dotted line the solution with a 20% smaller HMF magnitude and variance. The larger values for these quantities lead to computed intensities similar to the PAMELA observations from 2009, with this particular $A > 0$ solution crossing the $A < 0$ solution at an energy of $\sim 0.2$ GV, higher than the crossing energy for the $A > 0$ intensities calculated for 2009 conditions. Decreasing the variance and HMF magnitude by 20% leads to intensities significantly lower than the 2009 values, with a crossover occurring only at $\sim 0.01$ GeV. The reason for this behavior can be deduced from the drift and diffusion coefficients. The corresponding changes to the drift length scale and the MFPs affected by the different projected solar cycle 25 parameters, as well as the total variances, are shown in Figure 5 as a function of radial distance, rigidity, and colatitude. As both the HMF magnitude and the magnetic variances are changed simultaneously, the ratio of these quantities remains unchanged, leading to only relatively small changes in the MFPs. The largest changes are to be seen in the drift coefficients, with the 20% increase leading to a larger drift scale and the 20% decrease leading to a smaller drift scale. This would imply that in the former case, drift effects would play a larger role in the modulation of galactic CRs, leading, as shown by Reinecke & Potgieter (1994), to a shift of the crossover point in the spectra to a higher energy. The converse also holds, as the relatively decreased effects of drift implied by the smaller drift scale acquired when the HMF magnitude and variance are decreased by 20% lead to a shift toward a lower energy of the crossover point, again in agreement with the findings of Reinecke & Potgieter (1994).

4. Summary and Conclusions

Taken as a whole, the agreement of model results with data for all three solar minima is good. Furthermore, the fact that careful, observationally motivated modeling of the turbulence quantities as they differ from one solar cycle to the next naturally leads to larger intensities relative to previous solar minima observed during the unusual solar minimum of 2009 leads us to conclude that the higher than expected CR intensity observed during this solar minimum can be qualitatively linked to turbulence parameters that differ from solar minimum to solar minimum. Furthermore, Figure 3 shows that, due to the lower turbulence levels and HMF magnitude in 2009, the drift length scale for this period was larger than that during previous solar minima, leading us to conclude that drifts still play a role during this period, contrary to what was argued by, e.g., Potgieter et al. (2015) and references therein, and more in line with the findings of Zhao et al. (2014).

Overall, the intensities predicted by the current model for solar cycle 25 remain, at larger energies, at or below the intensities observed in 2009. This differs from what has been reported by Miyake et al. (2017), who expect intensities 19% higher than those observed during the 2009 solar minimum. The differences in the results presented here may be due to the fact that the effects of basic turbulence quantities on the diffusion and drift coefficients of CRs are taken into account in this study. At lower energies, $A > 0$ cycle 25 intensities are expected to be larger than in 2009, but only moderately so. This is a consequence of the fact that most modulation of galactic CRs occurs in the heliosheath (see, e.g., Stone et al. 2013), and as such, this result is only expected to change if there were large, solar cycle and magnetic polarity related changes in the boundary spectrum at 85 au. Webber et al. (2008) do indeed report that Voyager observed intensities for protons above $\sim 150$ MeV during $A < 0$ that were a factor of 1.5–1.7 higher than the corresponding $A > 0$ intensities. As this is not taken into account in the present study, the present results may even be an upper bound to what can potentially be observed in 2009.

To conclude, then, the present model can, in the near future, at the very least give some indication of the expected CR intensities based on a realistic ab initio approach to the modulation of CRs, and predict that the contribution of galactic CR protons to the space radiation environment in solar cycle 25 will be very similar to, or slightly less than, that during cycle 24.

Future work will involve extending the present model to be fully time dependent, incorporating time-dependent current sheet, tilt angle, solar wind profile, HMF, and turbulence quantities to model several full solar cycles and to refine the predictions made in the present study. The incorporation of the effects of the heliosheath on CR modulation in an ab initio way is also a priority for future studies. Furthermore, the current model will be used to study the time-dependent, ab initio modulation of other species of CRs, including galactic electrons, positrons, and in particular Jovian electrons, which make up the majority of CR electrons observed at Earth (see, e.g., Ferreira et al. 2001a, 2001b) and which have not yet been studied in an ab initio manner.

This work is based on the research supported in part by the National Research Foundation of South Africa (grant number 111731). Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to the NRF.

The authors would like to thank the Centre for High Performance Computing (CHPC) in South Africa for providing computational resources for this study.

ORCID iDs

N. E. Engelbrecht @ https://orcid.org/0000-0003-3659-7956

References

Adams, L. 1985, Microelectronics Journal, 16, 17
Ahlulwalia, H. S. 2016, AdSpR, 57, 710
Badhwar, G. D., Nachtwey, D. S., & Yang, T. C. 2001b, AdSpR, 12, 195
Bavassono, B., Pietropaolo, E., & Bruno, R. 2000a, JGR, 105, 15959
Bavassono, B., Pietropaolo, E., & Bruno, R. 2000b, JGR, 105, 12697
Bruno, R., & Carbone, V. 2013, LRSIP, 10, 2