ULF Waves and Diffusive Radial Transport of Charged Particles

by

Ashar Fawad Ali

B.S., California State University, Northridge, 2007

M.S., California State University, Northridge, 2009

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Applied Mathematics

2016

This thesis entitled:
ULF Waves and Diffusive Radial Transport of Charged Particles
written by Ashar Fawad Ali
has been approved for the Department of Applied Mathematics.

Dr. Scot R. Elkington

Dr. Jem Corcoran

Dr. Howard Singer

Dr. Juan Restrepo

Dr. William Kleiber

Date: _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Ali, Ashar Fawad (Ph.D., Applied Mathematics)

ULF Waves and Diffusive Radial Transport of Charged Particles

Thesis directed by Dr. Scot R. Elkington

The Van Allen radiation belts contain highly energetic particles which interact with a variety of plasma and magnetohydrodynamic (MHD) waves. Waves in the ultra low-frequency (ULF) range play an important role in the loss and acceleration of energetic particles. Considering the geometry of the geomagnetic field, charged particles trapped in the inner magnetosphere undergo three distinct types of periodic motions; an adiabatic invariant is associated with each type of motion. The evolution of the phase space density of charged particles in the magnetosphere in the coordinate space of the three adiabatic invariants is modeled by the Fokker-Planck equation. If we assume that the first two adiabatic invariants are conserved while the third invariant is violated, then the general Fokker-Planck equation reduces to a radial diffusion equation with the radial diffusion coefficient quantifying the rate of the radial diffusion of charged particles, including contributions from perturbations in both the magnetic and the electric fields.

This thesis investigates two unanswered questions about ULF wave-driven radial transport of charged particles. First, **how important are the ULF fluctuations in the magnetic field compared with the ULF fluctuations in the electric field in driving the radial diffusion of charged particles in the Earth's inner magnetosphere?** It has generally been accepted that magnetic field perturbations dominate over electric field perturbations, but several recently published studies suggest otherwise. Second, **what is the distribution of ULF wave power in azimuth, and how does ULF wave power depend upon radial distance and the level of geomagnetic activity?** Analytic treatments of the diffusion coefficients generally assume uniform distribution of power in azimuth, but *in situ* measurements suggest that this may not be the case.

We used the magnetic field data from the Combined Release and Radiation Effects Satellite (CRRES) and the electric and the magnetic field data from the Radiation Belt Storm Probes

(RBSP) to compute the electric and the magnetic component of the radial diffusion coefficient using the *Fei et al.* [2006] formulation. We conclude that contrary to prior notions, the electric component is dominant in driving radial diffusion of charged particles in the Earth's inner magnetosphere instead of the magnetic component. The electric component can be up to two orders of magnitude larger than the magnetic component. In addition, we see that ULF wave power in both the electric and the magnetic fields has a clear dependence on Kp with wave power decreasing as radial distance decreases. For both fields, the noon sectors generally contain more ULF wave power than the dawn, dusk, and the midnight magnetic local time (MLT) sectors. There is no significant difference between ULF wave power in the dawn, dusk, and the midnight sectors.

Dedicated to everyone who has taught me to learn.

Acknowledgements

First of all I would like to thank my research adviser, Dr. Scot Elkington, for continuously supporting my graduate studies and research, for being extremely patient and generous, and advising me on all aspects of academia in preparing me for a lifelong career as a scientist. I would like to acknowledge my good friend Amrik Sen, with whom I have spent many pleasant hours and who has always given me the right advice at the right time. Without his friendship, completion of this degree might not have been possible. I also want to express my gratitude to Mark Joseph with whom discussions are always lively and thought provoking. His insights have always been necessary and invaluable to my growth as an individual. I would like to thank my thesis committee members for reviewing my work and suggesting improvements. I would like to acknowledge Dr. Harvey Segur, who has always taken an interest in my success and encouraged me when needed. I want to thank Allison Jaynes, David Malaspina, Frederick Wilder, Quintin Schiller, Joshua Murphy, Thiago Brito, Alexander Drozdov, and Ksenia Orlova. These are my friends and colleagues, who have always helped me no matter how much of a burden it may have been. I want to thank my beautiful wife and my adorable daughter for providing the momentum that I needed to finish my thesis successfully. Most importantly, I want to thank my parents, who have sacrificed tremendously to make sure that I could become who I want to be, to provide me with as many opportunities as possible, and to ensure that I succeed and make them proud.

The research presented in this thesis was funded by NASA grants NNX15AF59G, NNX14AC04G, NNX13AE39G, NNX09AI05G, the NASA Earth and Space Science Fellowship (NESSF) grant NNX13AO43H, and a Los Alamos National Laboratory sponsored Vela Fellowship.

Contents

Chapter

Figures

Figure

Chapter 1

Introduction and Background

1.1 Introduction

The radiation belts contain energetic particles, ranging from hundreds of keV to hundreds of MeV, trapped in the geomagnetic field in the form of two concentric tori around the Earth. The inner radiation belt is below 2 R_E and contains protons and electrons while the outer radiation belt is typically between 3-7 R_E and consists mostly of electrons [*Tascione*, 1988; *Kivelson and Russell*, 1995; *Baumjohann and Treumann*, 1997; *Treumann and Baumjohann*, 1997; *Cravens*, 2004]. Between these zones is a region called the slot region, kept relatively free of particles because of high-frequency wave-particle interactions [*Lyons et al.*, 1972]. These high-energy particles are hazardous to both technology and astronauts [*Baker*, 2001]. With our increasing dependence on space-based technology for communication, navigation, and weather prediction, a clear, predictive understanding of the radiation belts and how they are populated and depleted is increasingly urgent. In the magnetosphere, there are several types of plasma and magnetohydrodynamic (MHD) waves which can result in acceleration and loss of energetic particles. Very low-frequency (VLF) waves include whistler-mode chorus, whistler-mode hiss, and magnetosonic waves [*Temerin et al.*, 1994; *Li et al.*, 1997; *Summers et al.*, 1998; *Horne and Thorne*, 1998] while ultra low-frequency (ULF) waves include electromagnetic ion cyclotron (EMIC) waves [*Hudson*, 2013].

Geomagnetic disturbances produce MHD waves with power in the ULF range which then interact with radiation belt particles via wave-particle interactions and change their energy [*Schulz and Lanzerotti*, 1974]. Several researchers have shown that solar wind conditions are highly cor-

related with ULF wave power observed in the magnetosphere [*Kivelson and Southwood*, 1988; *Hartinger et al.*, 2013]. This correlation is good evidence that the strong velocity shear present near the flanks drives Kelvin-Helmholtz instability there, which in turn is the source of ULF wave power at the flanks [*Cahill and Winckler*, 1992; *Mann et al.*, 1999; *Claudepierre et al.*, 2008; *Claudepierre*, 2008]. On the dayside, solar wind dynamic pressure variations contribute to ULF wave power [*Kivelson and Southwood*, 1988]. We will investigate how these ULF waves interact with the radiation belts and drive particle diffusion.

1.2 Definitions

In this section, we will define and describe some of the terms and acronyms used throughout this thesis.

$\mathbf{B}_{\parallel}$ refers to the compressional component of the magnetic field which denotes the oscillations in the magnetic field in the direction of the magnetic field, i.e., waves or changes in the field parallel to the field. This component is used to calculate D_{LL}^{B}.

$\mathbf{E}_{\phi}$ refers to the azimuthal component of the electric field which denotes the oscillations in the tangential direction in the magnetic equatorial plane of the Earth when the electric field is represented in cylindrical coordinates. This component is used to calculate D_{LL}^{E}.

CRRES refers to the Combined Release and Radiation Effects Satellite, a combined effort between NASA and the United States Air Force. The mission was launched on July 25, 1990 and it lasted until October 12, 1991. In this thesis, we use the magnetic field data from CRRES to estimate the magnetic component of the radial diffusion coefficient, D_{LL}^{B}[CRRES], presented in Chapter 3. The electric field data has been used by *Brautigam et al.* [2005] to estimate the electric component of the radial diffusion coefficient, D_{LL}^{E}[CRRES].

$\mathbf{D_{LL}}$ refers to the radial diffusion coefficient, which quantifies the mean radial displacement of an ensemble of particles. The total radial diffusion coefficient is a measure of the net radial

diffusion of charged particles in the radial direction over timescales much longer than the characteristic timescales associated with the third adiabatic invariant. The radial diffusion coefficient has units of time^{-1}.

D_{LL}^{B} refers to the magnetic component of the radial diffusion coefficient. This component quantifies contributions only from the magnetic field, with no electric field components [*Fei et al.*, 2006].

D_{LL}^{E} refers to the electric component of the radial diffusion coefficient. This component quantifies contributions from the total electric field: the convective electric field and the inductive electric field, with no magnetic field component [*Fei et al.*, 2006].

D_{LL}^{EM} refers to the electromagnetic component of the radial diffusion coefficient. This component quantifies contributions from the magnetic field as well as the inductive electric field [*Fälthammar*, 1965, 1968].

D_{LL}^{ES} refers to the electrostatic component of the radial diffusion coefficient. This component quantifies contributions only from the convective (electrostatic) component of the electric field [*Fälthammar*, 1965, 1968].

GSE refers to the geocentric solar ecliptic coordinate system which is a Cartesian coordinate system. The x-axis is the Sun-Earth line pointing to the Sun. The y-axis is in the ecliptic plane pointing towards dusk (MLT = 1800 hours, in a direction opposing planetary motion). The z-axis is parallel to the ecliptic pole and completes the right-hand system. The GSM coordinate system can be obtained by rotating the GSE coordinate system about the x-axis.

GSM refers to the geocenteric solar magnetospheric coordinate system which is a Cartesian coordinate system centered at the Earth. The x-direction is the Sun-Earth line pointing to the Sun. The y-axis is perpendicular to the Earth's magnetic dipole axis, so that the x-z plane contains the dipole axis. The z-axis is parallel to the dipole (positive North) axis. The

GSE coordinate system can be obtained by rotating the GSM coordinate system about the x-axis.

Kp refers to the planetary K-index. The K-index is a measure of geomagnetic activity and the resulting disturbance in the geomagnetic field. It is derived using maximum fluctuations in the geomagnetic field over a three-hour period, and can take values in

$$\{0, 0+, 1-, 1, 1+, 2-, 2, ..., 8-, 8, 8+, 9-, 9\}.$$

Kp is derived using weighted averages of measurements provided by several geomagnetic observatories all over the world. $Kp = 0$ indicates quiet times with not much geomagnetic activity, while $Kp = 9$ means severe activity.

L refers to the set of magnetic field lines which would exactly be L Earth radii away in the equatorial plane if the geomagnetic field were a pure, constant, symmetric, dipole field without any perturbations. There are various other definitions of the L parameter which return slightly different values of the parameter. For the CRRES study in Chapter 3, we use the McIlwain L [*McIlwain*, 1961], as that was the only parameter computed and provided in the CRRES dataset. For the RBSP study in Chapter 4, we use the Roederer L^* parameter [*Roederer and Zhang*, 2014] which is inversely proportional to the third adiabatic invariant Φ. Because of the asymmetry of the geomagnetic field, the L^* parameter is much more convenient to work with than the third adiabatic invariant Φ.

mGSE refers to the modified geocentric solar ecliptic coordinate system. This is commonly used to provide data from spacecraft because this coordinate system is local to the spacecraft. The mGSE system is similar to the GSE system except that the x-axis is the spin axis of the spacecraft instead of the Sun-Earth line. The spin axis is deliberately kept at a small angle from the Sun-Earth line. For CRRES, the spin axis was usually kept within 12° of the Sun-Earth line. For RBSP, the spin axis is kept within 37° of the Sun-Earth line.

MLT refers to the magnetic local time which is a component of a geomagnetic spherical coordinate

system. Ranging from 0000 hours to 2400 hours, the magnetic local time is analogous to local time, except that noon MLT and midnight MLT occur when the sun is in the meridian plane, defined using the geomagnetic poles instead of the geographic poles. A point at noon MLT is directly on the Sun-Earth line while a point at midnight MLT faces the other direction.

PSD refers to the power spectral density of a given signal. Another common name for the power spectral density is the periodogram, which shows the amount of wave power present in a signal at a given frequency. PSD commonly refers to the phase space density in space physics, but in this thesis PSD is not used with this meaning.

RBSP refers to the NASA mission Radiation Belt Storm Probes, a set of two identical probes with nearly identical orbits. RBSP is the first mission since CRRES dedicated to the study of the Van Allen radiation belts. After a successful launch on August 30, 2012, the mission was renamed the Van Allen Probes. In Chapter 4 we use the magnetic and the electric field data from RBSP to estimate the magnetic and the electric components of the radial diffusion coefficient, D_{LL}^{B}[RBSP] and D_{LL}^{E}[RBSP].

1.3 Earth's Magnetosphere

It has long been known that the Earth has an intrinsic magnetic field, which originates from the interior of the planet and extends far out into space where it interacts with the solar wind, a stream of charged particles and a magnetic field with solar origins [*Tascione*, 1988; *Kivelson and Russell*, 1995; *Baumjohann and Treumann*, 1997; *Treumann and Baumjohann*, 1997; *Cravens*, 2004]. It is believed that the geomagnetic field is the result of a geodynamo produced by the motion of molten iron in the Earth's outer core. The geomagnetic field is approximated reasonably well below and near the surface by a dipole magnetic field which is known to change over time. Currently, the dipole axis is tilted at an angle of about 11° with respect to Earth's rotational axis and shifted up from the center of the Earth. As we move away from the surface into space, plasma

flows present in the magnetosphere form currents which contribute to the geomagnetic field, and the dipole magnetic field becomes very distorted.

The Earth's magnetosphere is a region above the ionosphere which extends tens of thousands of kilometers into space and is distorted by interaction with the solar wind. On the dayside, the side directly in sunlight, the solar wind exerts a pressure on the magnetosphere and compresses it. The region of space where the solar wind pressure is balanced by the geomagnetic field pressure is considered to be the boundary of the magnetosphere, called the magnetopause. On the nightside the magnetosphere is stretched out in a magnetotail and the magnetosphere is therefore not symmetric. On the dayside the magnetosphere extends to a distance of about 10 R_E (Earth radii), while on the nightside the magnetosphere extends beyond 200 R_E. On the dayside, sunward of the magnetopause, there is the bow shock, a region where the solar wind velocity vector changes and the solar wind speed slows down abruptly. The region between the bow shock and magnetopause is called the magnetosheath. The inner magnetosphere contains a region called the plasmasphere which contains low-energy particles (a few eV). The plasmasphere is located above the ionosphere; its outer boundary, called the plasmapause, is marked by a sudden drop in plasma density. Figure 1.1 shows a simple diagram of the Earth's magnetosphere.

1.4 The Van Allen Radiation Belts

The Van Allen radiation belts are structures in the inner magnetosphere, in the shape of two concentric tori around the Earth [*Tascione*, 1988; *Kivelson and Russell*, 1995; *Baumjohann and Treumann*, 1997; *Treumann and Baumjohann*, 1997; *Cravens*, 2004]. These layers of charged particles are formed by particles trapped by the geomagnetic field. Between these two radiation belts, there is a region of space called the slot zone which is kept relatively free of charged particles [*Lyons et al.*, 1972]. The radiation belts are named after Dr.James Van Allen who discovered them using the data collected from scientific instruments on board the Explorer I, Explorer III, and Pioneer III satellites in 1958. A point to note here is that although other planetary bodies with intrinsic magnetic fields may have similar radiation belt structures, the name Van Allen radiation

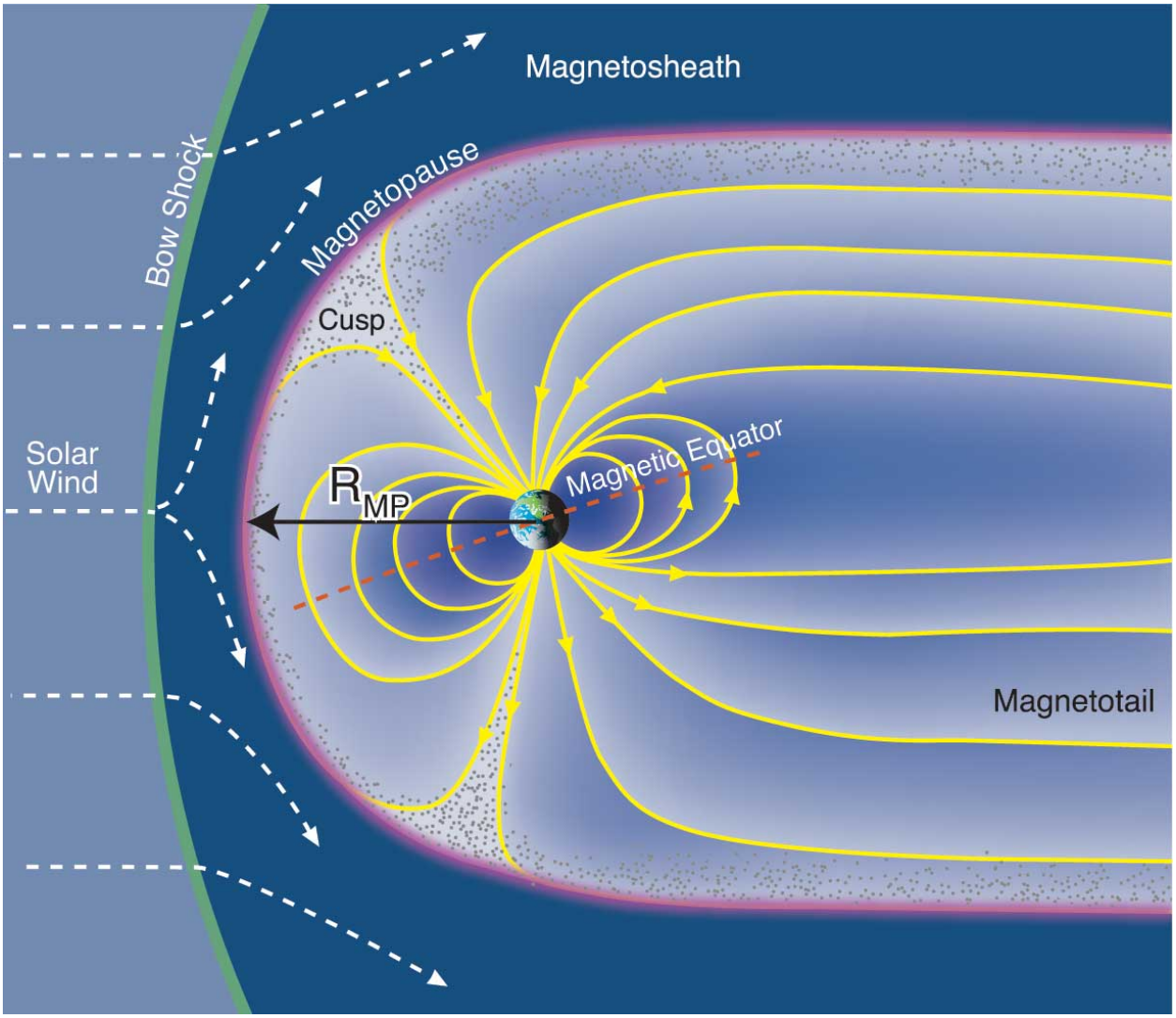


Figure 1.1: Diagram of the Earth's magnetosphere. [Courtesy: Fran Bagenal & Steve Bartlett, LASP]

belts specifically refers to the radiation belts surrounding the Earth.

The inner Van Allen radiation belt extends typically from 1.01 R_E to about 2.5 R_E. The particle population consists mostly of high-energy protons along with some lower-energy electrons. Typical energies for protons are hundreds of MeV while electrons are in the range of hundreds of keV. Because the magnetic dipole axis is tilted and off-center, the inner radiation belt is closest to the surface of the Earth in the South Atlantic region, forming the so-called South Atlantic Anamoly. The outer Van Allen radiation belt occupies a region of space typically between 3-7 R_E and is mostly populated by high-energy electrons with energies between 0.1 and tens of MeV. Inward radial diffusion [*Fälthammar*, 1965; *Schulz and Lanzerotti*, 1974; *Elkington et al.*, 2003] and local acceleration [*Horne et al.*, 2005] of radiation belt electrons due to the energy transfer from whistler-mode plasma waves are two of the primary causes for the creation of the outer radiation belt [*Friedel et al.*, 2002; *Shprits et al.*, 2008a,b]. Electrons are also removed from the outer radiation belt by outward radial diffusion, losses to the magnetopause, and collisions with the atmosphere.

The outer radiation belt occupies a region of space much larger than the inner radiation belt, and its population and size are highly variable relative to the population and the size of the inner belt. The inner radiation belt is quite stable, while the outer zone fluxes can greatly increase and decrease due to geomagnetic storms. It should be noted here that recently *Baker et al.* [2013] reported a third storage ring which was observed to exist temporarily, distinct from the outer radiation belt, before being destroyed by an interplanetary shockwave. Furthermore, *Baker et al.* [2014] reports a "barrier" at the inner boundary of the outer radiation belt which cannot be penetrated by ultra-relativistic electrons with kinetic energies higher than 5 MeV. Figure 1.2 shows the Van Allen radiation belts along with the location of a few spacecraft.

1.5 Charged Particles in the Inner Magnetosphere

It is well known that the motion of a charged particle in magnetic and electric fields is described by the Lorentz equation,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{1.1}$$

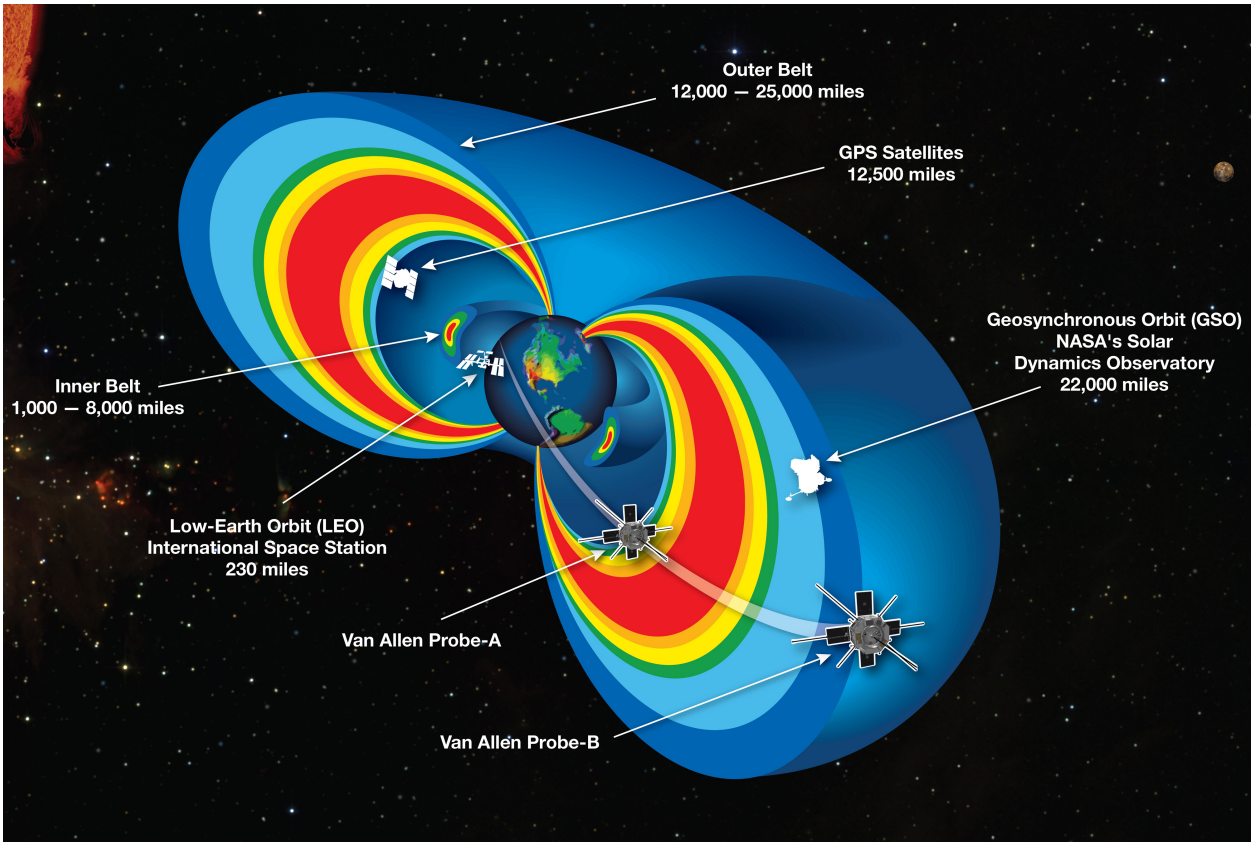


Figure 1.2: Location of the two Van Allen radiation belts along with the Van Allen Probes and other satellites in the Earth's inner magnetosphere. [Courtesy: NASA]

where q and $\mathbf{v}$ represent the charge and the velocity of the particle and $\mathbf{B}$ and $\mathbf{E}$ are the magnetic and the electric vector fields respectively [*Schulz and Lanzerotti*, 1974; *Tascione*, 1988; *Kivelson and Russell*, 1995; *Baumjohann and Treumann*, 1997; *Treumann and Baumjohann*, 1997; *Cravens*, 2004; *Roederer and Zhang*, 2014]. In the case of a uniform magnetic field and $\mathbf{E} = 0$, the particle will move in a circular path in a plane perpendicular to $\mathbf{B}$ if the velocity parallel to the field is zero. This motion is called the gyro motion. An important quantity is the gyro radius or the cyclotron radius,

$$r = \frac{mv_\perp}{Bq}, \tag{1.2}$$

where $v_\perp$ is the speed of the particle perpendicular to the field, and B is the strength of the magnetic field. If the particle has some velocity parallel to $\mathbf{B}$ then the particle will trace out a helical trajectory with the parallel component of the velocity $\mathbf{v}_\parallel$ being a constant for a constant $\mathbf{B}$. The center of the circular motion, called the guiding center, will simply drift with velocity $\mathbf{v}_\parallel$. The pitch angle of the helix is the angle between the particle velocity and the magnetic field and is given by $\alpha = \arctan(v_\perp/v_\parallel)$. If the pitch angle for a given particle is close to 90°, then the particle will move nearly in a circular fashion with $v_\parallel$ being extremely small. If the pitch angle is close to 0° then the helical trajectory will be stretched.

Since the gyro motion is a periodic motion, there is an adiabatic invariant associated with it. The canonical momentum $\mathbf{P}$ of a charged particle in a magnetic field is given by

$$\mathbf{P} = m\mathbf{v} + q\mathbf{A} = \mathbf{p} + q\mathbf{A}, \tag{1.3}$$

where m, q, $\mathbf{v}$, and $\mathbf{p}$ represent the mass, charge, velocity, and (kinetic) momentum of the particle and $\mathbf{A}$ is the magnetic vector potential of the magnetic field $\mathbf{B}$ such that $\mathbf{B} = \nabla \times \mathbf{A}$. Integrating the canonical momentum $\mathbf{P}$ over the periodic orbit of any periodic motion gives us the adiabatic invariant associated with that motion. The invariant is adiabatic because if the field is not changing at all then the invariant will be a constant. If however the field is changing very slowly compared to the frequency of the periodic motion then the invariant will almost remain a constant. For the gyro motion, if $\mathbf{P}$ is integrated over a gyro orbit we obtain what we will call the (relativistically

corrected) first adiabatic invariant,

$$M = \frac{p_\perp^2}{2m_0 B}, \tag{1.4}$$

where $p_\perp$ is the momentum of the particle in the direction perpendicular to the magnetic field, m_0 is the rest mass of the particle, and B is the strength of the magnetic field. In the Earth's inner magnetosphere, a charged particle will gyrate about a magnetic field line with the characteristic timescale being on the order of a few milliseconds for an electron.

If the background magnetic field $\mathbf{B}$ is not uniform in space and has a gradient $\nabla\mathbf{B}$ then a magnetic mirror effect may occur for a charged particle, depending on the velocity of the particle. If a charged particle moves along a field line from a region of low magnetic field strength to a region of higher magnetic field strength, then the particle will decelerate in the direction of the gradient due to the conservation of energy and the first invariant M. If the gradient is large enough, then the parallel velocity can be reduced to zero, at which point the particle reverses its direction. This can be seen as follows. If M is conserved, then increasing the strength of the magnetic field B must necessarily increase $p_\perp$. Since the total kinetic momentum $p = p_\parallel + p_\perp$ must also remain constant, the parallel momentum $p_\parallel$ must decrease in magnitude. This results in a decrease in parallel velocity. If the decrease is large enough then the particle may come to rest in the parallel direction and reverse its direction.

If two magnetic mirrors are placed together this results in a magnetic bottle in which a charged particle can be temporarily trapped. This is precisely the case in the Earth's inner magnetosphere. If a particle has a nonzero velocity parallel to the Earth's magnetic field, then in addition to the gyro motion, the particle will also move latitudinally along a magnetic field line between the North and the South magnetic poles. As the particle moves from the equator towards one of the poles, it moves into regions of higher magnetic field strength. If we assume that the first invariant is conserved, then the particle decelerates in the parallel direction. It eventually reaches a mirror point where the parallel velocity is zero and it is then reflected back towards the equator. The same process occurs near the other pole and the particle is trapped, mirroring back and forth

between the two mirror points. This second type of periodic motion is called the bounce motion. For an electron the typical timescales are between a tenth of a second and a few seconds. The adiabatic invariant associated with this periodic motion is

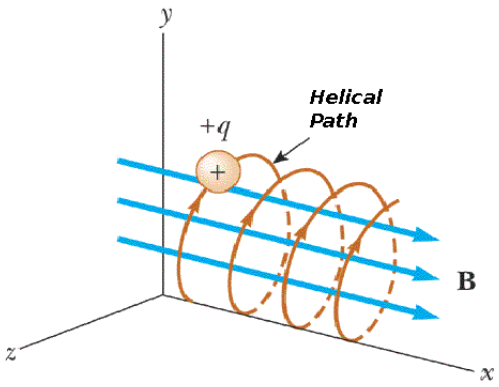
$$J = \int_{m_1}^{m_2} p_{\parallel} ds, \tag{1.5}$$

where m_1 and m_2 are the two mirror points, $p_{\parallel}$ is the momentum of the particle parallel to the magnetic field, and ds being the element of length along the field line.

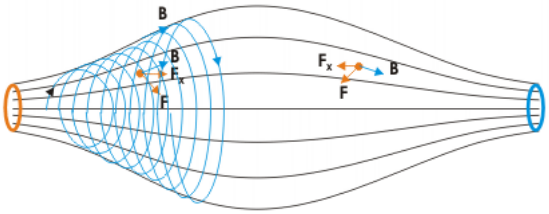
The third type of periodic motion for a charged particle in the Earth's inner magnetosphere is the longitudinal drift of a geomagnetically trapped particle around the Earth. This drift is caused by the both the curvature drift and the gradient drift, with the drift being eastward for electrons and westward for positively charged ions. The curvature drift is due to the fact that while following a curved field line, a charged particle experiences a centripetal force, resulting in a drift velocity perpendicular to the curvature. The gradient drift exists because as a particle moves inward into a region of higher field strength, the gyro radius decreases. As the particle moves outward toward lower field strength, the gyro radius increases resulting in a net drift of the guiding center. Figure 1.3 illustrates the gyro motion, the magnetic mirroring, the gradient drift, and the curvature drift of a charged particle.

One thing to note here is that curved parallel magnetic field lines are not possible in reality. Curved magnetic field lines must necessarily be either converging or diverging, and therefore the curvature drift will always be accompanied by the gradient drift. Inclusion of a nonzero electric field results in an "E cross B" drift as the electric field results in the motion of the guiding center in the plane perpendicular to both $\mathbf{E}$ and $\mathbf{B}$. There are other drifts such as the gravitational drift of charged particles in the inner magnetosphere which we ignore because of their effects are negligible. For a general force $\mathbf{F}$, the drift velocity is given by

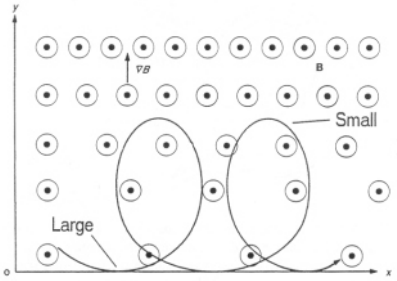
$$\mathbf{v}_{\text{Drift}} = \frac{1}{q} \cdot \frac{\mathbf{F} \times \mathbf{B}}{B^2}. \tag{1.6}$$



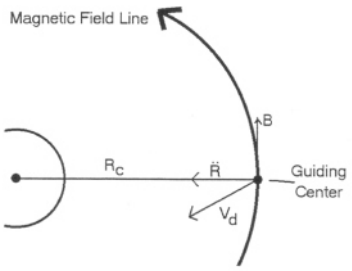
Gyro Motion



Magnetic Mirror



Gradient Drift



Curvature Drift

Figure 1.3: Illustrating gyro motion, magnetic mirroring, gradient drift, and curvature drift of a charged particle in various magnetic field configurations.

The third adiabatic invariant associated with this periodic motion is

$$\Phi = \oint \vec{B} \cdot ds, \tag{1.7}$$

the magnetic flux enclosed by the drift path, the characteristic timescale for which is on the order of a few minutes to tens of minutes. This third invariant is most useful in describing drift paths during slow changes in the geomagnetic field. For example, slow compressions or expansions of the geomagnetic field result in the trapped particles moving inward or outward in order to conserve the third invariant. If the changes are slow enough then these effects are reversible, and upon restoration of the geomagnetic field the trapped particles will return to their original positions. If however these changes occur faster than the characteristic timescale of the third invariant, then the change in Φ will be permanent. Many physical processes in the magnetosphere happen on such a timescale that the first two invariants are conserved while the third invariant is violated causing a net change in the radial distance of a charged particle, having a diffusion-like effect on the particle in the radial direction.

Because of the irregularity and the asymmetry of the geomagnetic field, it is much more convenient to work with the related quantity

$$L^* = -\frac{2\pi k_0}{\Phi R_E}, \tag{1.8}$$

as defined in *Roederer and Zhang* [2014], with R_E being the Earth's radius, and k_0 being the Earth's dipole moment, assumed to be located at the center of the Earth and pointing southward. Physically speaking, $L^* = L$ describes the set of magnetic field lines which would be exactly L Earth radii away in the equatorial plane if the geomagnetic field were a pure, constant, symmetric, dipole field without any perturbations. It should also be noted here that perturbations which cause violation of the third invariant while preserving the first and the second invariant also cause a net change in the perpendicular energy of a particle. A particle will gain (lose) energy if it diffuses inwards (outwards) into regions of higher (lower) magnetic field strengths. Since the first invariant is conserved, energy increases if and only if B increases.

In summary, the geometry of the geomagnetic field results in three distinct types of periodic motions for a charged particle in the inner magnetosphere: the gyro motion, the bounce motion, and the drift motion. These motions have characteristic timescales on the order of a few milliseconds, between a tenth of a second and a few seconds, and a few minutes, respectively, for an electron. These periodic motions result in particles being trapped in the inner magnetosphere. The net motion of a particle is the sum of all three of these periodic motions, as shown in Figure 1.4.

1.6 The Fokker-Planck Equation

The derivation presented in this section is described in detail in *Schulz* [1996] and *Roederer and Zhang* [2014]. Consider an ensemble of charged particles in an electromagnetic field. If we denote the number of particles per unit $\mathbf{x}$ at time t by $f(\mathbf{x}, t)$, then f is called the phase space density function for the given particle population. In order to fully describe the dynamics of the particle population, the vector $\mathbf{x}$ consists of six components, three of which are spatial coordinates and the other three velocity (or momentum) coordinates. This is the traditional six-dimensional phase space common in statistical mechanics. The phase space density function can be integrated over the appropriate variables to give the number of particles in a given region of space, or the number of particles within a given range of velocities (or momenta).

Now let us introduce a probability function $F(\mathbf{x}-\Delta\mathbf{x}, \Delta\mathbf{x}, t, \Delta t)$ which is the probability that a particle at coordinates $\mathbf{x} - \Delta\mathbf{x}$ at time t will have its coordinates increased by $\Delta\mathbf{x}$ in the time interval Δt. This probability function represents the random physical perturbations affecting the particle population and causing small variations in their coordinates $\mathbf{x}$. The physical perturbations can be from electric field accelerations, particle collisions, perturbations due to the asymmetry of the magnetic field, or just random perturbations. Since F is a probability function, we have that $\int F d(\mathbf{x}) = 1$ for all time t where the integral is taken over all possible values of $\mathbf{x}$.

Using the probability function F, we can obtain the time evolution of the phase space density

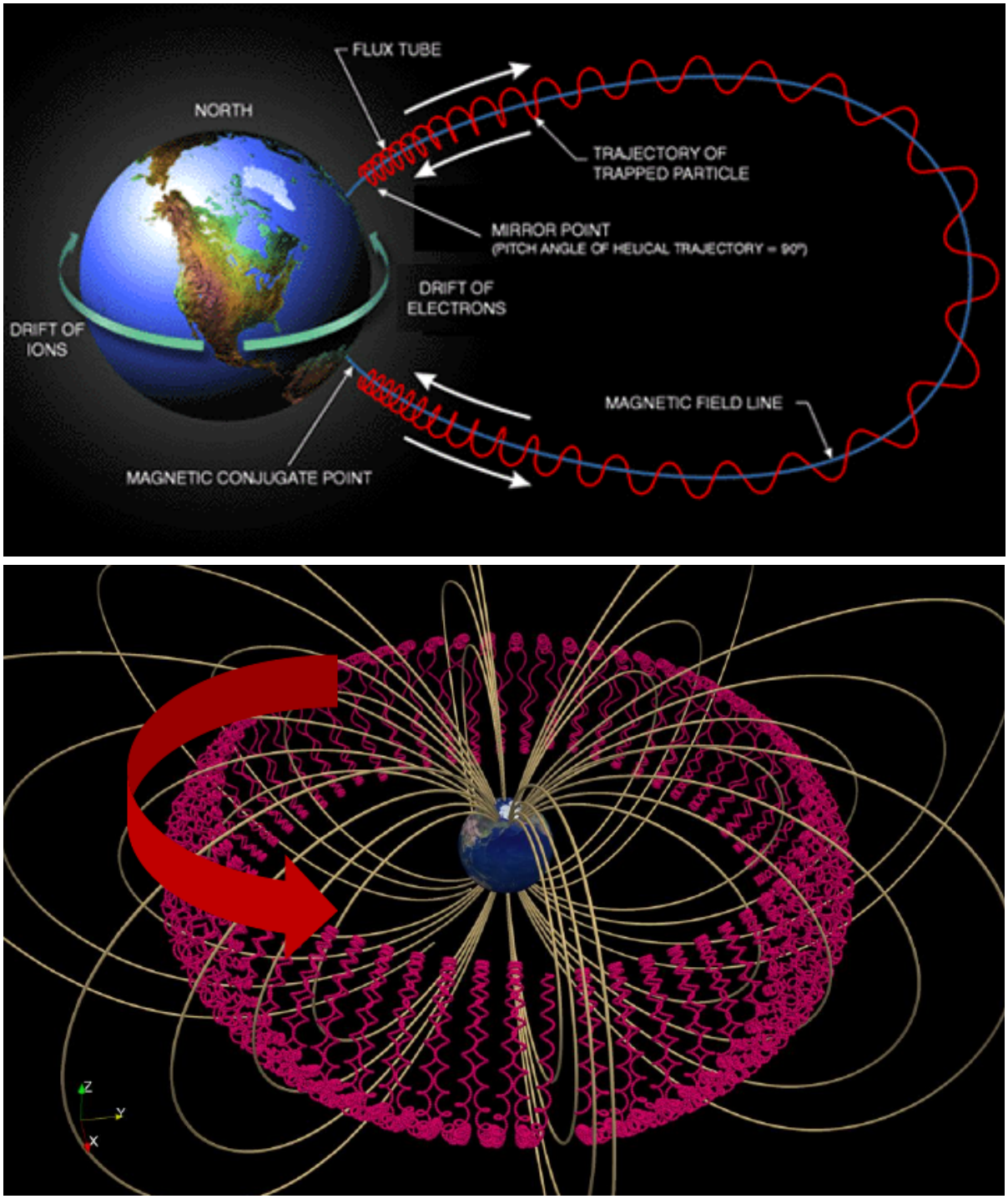


Figure 1.4: The three periodic motions of a charged particle in the Earth's inner magnetosphere.

f from time t to time $t + \Delta t$ as

$$f(\mathbf{x}, t + \Delta t) = \int f(\mathbf{x} - \Delta\mathbf{x}, t) F(\mathbf{x} - \Delta\mathbf{x}, \Delta\mathbf{x}, t, \Delta t) d(\Delta\mathbf{x}). \tag{1.9}$$

Let us now expand Equation (1.9) using Taylor series. Let us expand both f and F in the spatial coordinates $\mathbf{x}$ up to the second order terms on the right-hand side, in addition to expanding f in t to the first order term on the left. Grouping and rearranging the terms then results in the Fokker-Planck equation,

$$\frac{\partial f}{\partial t} = -\sum_i \frac{\partial (D_{x_i} f)}{\partial x_i} + \sum_i \sum_j \frac{\partial^2 (D_{x_i x_j} f)}{\partial x_i \partial x_j}, \tag{1.10}$$

where x_i are the components of the coordinate vector $\mathbf{x}$. The D_{x_i} are called the first order (Fokker-Planck) diffusion coefficients and are defined as

$$D_{x_i} = \frac{\langle \Delta x_i \rangle}{\Delta t} = \frac{1}{\Delta t} \int x_i F dx_i. \tag{1.11}$$

They represent the average change in coordinate x_i per unit time for one particle. The $D_{x_i x_j}$ are called the second order diffusion coefficients, defined as

$$D_{x_i x_j} = \frac{\langle \Delta x_i \Delta x_j \rangle}{2\Delta t} = \frac{1}{2\Delta t} \int \int x_i x_j F dx_i dx_j. \tag{1.12}$$

The second order diffusion coefficients represent the "strength" or the "rate" of the diffusion processes. We would like to point out here that in general, even if the first order diffusion coefficients are zero, diffusion mechanisms may still be present. In the case of the second order diffusion coefficients, however, diffusion occurs if and only if the coefficients are nonzero.

For the velocity of a charged particle in the inner magnetosphere, the Fokker-Planck equation is identical to the drift-diffusion equation. The drift-diffusion equation is also called the advection-diffusion equation, convection-diffusion equation, or a scalar transport equation in various disciplines. The drift-diffusion equation in general describes a physical system in which particles are transported via the mechanisms of drift (or advection) and diffusion. Hence, it is common to refer to the first order diffusion coefficients as drift coefficients while the second order diffusion coefficients can simply be referred to as the diffusion coefficients. Since we derived the Fokker-Planck

equation using a six-dimensional coordinate space, we see that there are six drift coefficients from the drift vector and thirty-six diffusion coefficients counting all of the diagonal terms ($D_{x_i x_i}$) and all of the off-diagonal mixed terms ($D_{x_i x_j}$) in the diffusion tensor. This amounts to a total of forty-two coefficients in the Fokker-Planck equation.

First let us transform our coordinate system from position and velocity to the coordinate space of the three adiabatic invariants and the corresponding phases as the charged particles go through the three periodic motions in the inner magnetosphere. This transformation requires the Jacobian of the transformation to transform the Fokker-Planck equation as well as to rewrite the drift and diffusion coefficients in terms of the new variables. Essentially we have that

$$f(x, y, z, v_x, v_y, v_z, t) \rightarrow f(M, J, \Phi, \phi_M, \phi_J, \phi_\Phi),$$

where M, J, and Φ are the three adiabatic invariants as defined above. The phases for each of the three periodic motions are given by ϕ_M, ϕ_J, and ϕ_Φ.

Now let us assume that for an ensemble of particles, the particles do not "bunch up" anywhere (as seen in Figure 1.5) and that the particles are distributed uniformly over all phase angles. For any given value of M, the particles are evenly distributed over a gyro trajectory. For any given value of J, the particles are not concentrated at a particular latitude. For any given value of Φ, the particles are uniformly distributed for all longitudes. Then we do not need to consider the instantaneous phase of the particles and we can reduce our coordinate space from six variables to only three,

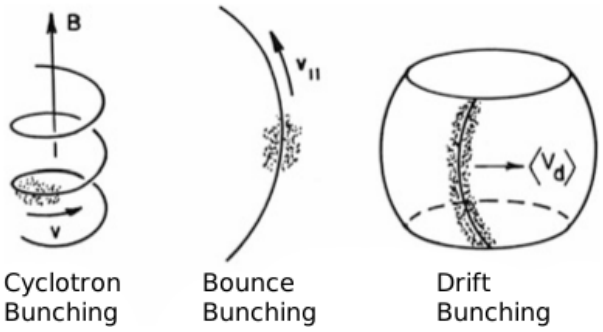


Figure 1.5: Phase "bunching" for each of the three periodic motions [*Roederer and Zhang*, 2014].

$$f(M, J, \Phi, \phi_M, \phi_J, \phi_\Phi) \rightarrow f(M, J, \Phi).$$

This lowers the number of drift and diffusion coefficients from forty-two to only twelve. Note that this assumption is not unreasonable even if an ensemble of particles initially all have the same

phase angles. Tiny random variations will result in a uniform distribution in phase after a time interval much larger than the characteristic timescale for an invariant. We also have that the diffusion coefficients are symmetric because $D_{x_i x_j} = D_{x_j x_i}$, further reducing the number of distinct coefficients from twelve to nine. Lastly, we consider Fick's first law. Particles diffuse from regions of higher density to regions of lower density. Diffusion is a smoothing process and it tends to eliminate spatial gradients in phase space density profiles. In other words, if particles move out of a volume of invariant coordinate space, then an equal number of particles must enter that volume assuming the steady state of the system. This allows us to write the drift coefficients in terms of the diffusion coefficients,

$$D_{x_i} = \frac{\partial D_{x_i x_i}}{\partial x_i}. \tag{1.13}$$

We have now eliminated another three independent coefficients, leaving a total of six, all of which are second order diffusion coefficients. The Fokker-Planck equation can now be written as

$$\frac{\partial f}{\partial t} = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial}{\partial x_i} \left(D_{x_i x_j} \frac{\partial f}{\partial x_j} \right) + Q - S, \tag{1.14}$$

where x_1, x_2, and x_3 represent the adiabatic invariants M, J, and Φ respectively and Q and S denote sources and sinks for particles.

1.7 ULF Waves in the Magnetosphere

The magnetosphere supports several varieties of plasma and MHD waves that dictate the dynamics of particles and their energization in the magnetosphere. Whistler waves [*Temerin et al.*, 1994; *Li et al.*, 1997; *Summers et al.*, 1998], named for the sounds they produce when detected by an audio amplifier, are a subset of very low-frequency (VLF) waves. Whistler-mode chorus waves are implicated in both heating and loss of charged particles while whistler-mode hiss waves, confined to the co-rotating region, are usually invoked as a loss mechanism [*Horne and Thorne*, 1998]. Figure 1.6 shows the location of several types of waves.

The ultra low-frequency (ULF) range is defined as 1.67 mHz - 5 Hz, with Pc-5 range being the subset 1.67 mHz - 6.67 mHz [*Jacobs et al.*, 1964]. ULF waves are one type of waves present in the

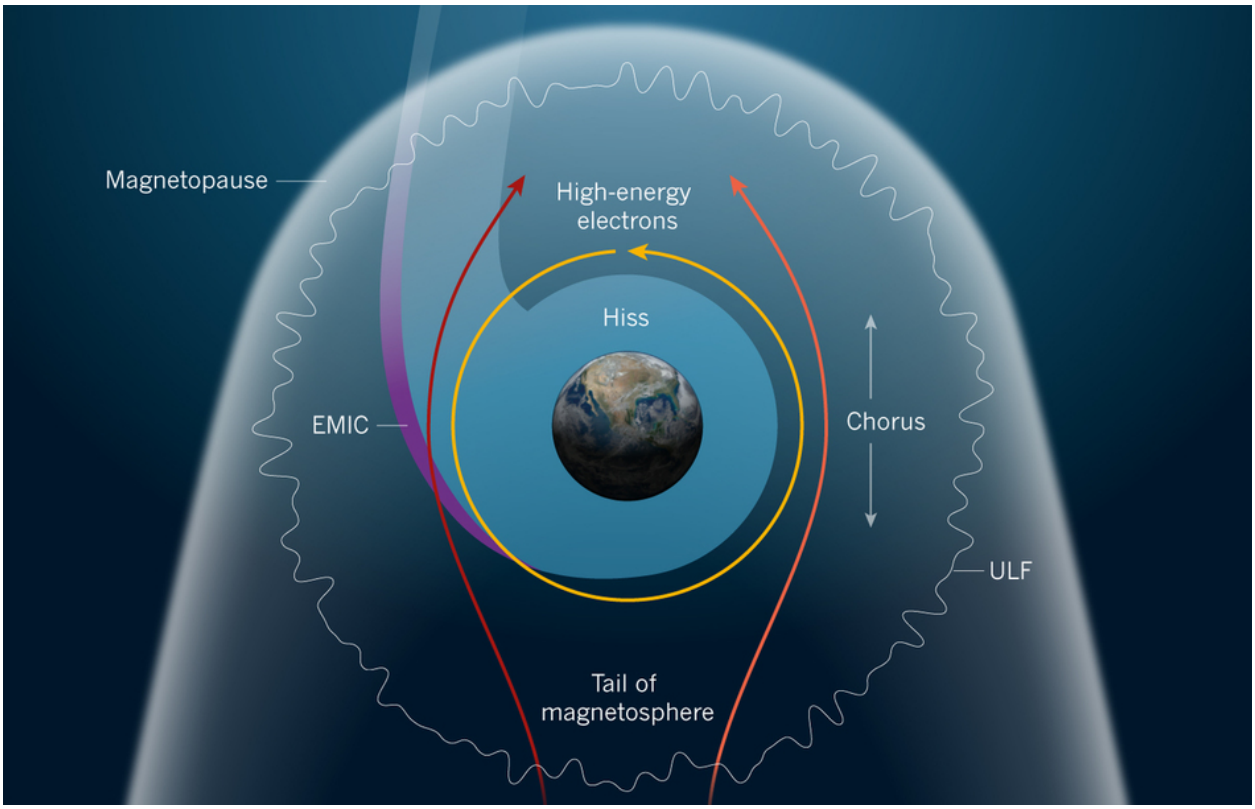


Figure 1.6: Location of various plasma and MHD waves in the magnetosphere [*Hudson*, 2013].

magnetosphere, and have several sources of origin. Mirror [*Hasegawa*, 1969] and drift [*Southwood et al.*, 1969] instabilities in the plasma drifting into the inner magnetosphere are among the internal processes thought to generate power at ultra low frequencies. External processes contributing to ULF wave power include shear flow instabilities along the magnetopause predominant at the flanks [*Cahill and Winckler*, 1992; *Mann et al.*, 1999; *Claudepierre et al.*, 2008], and solar wind pressure variations predominant on the noon side [*Kivelson and Southwood*, 1988]. More recently, *Hartinger et al.* [2013] showed transient ion foreshocks to be another possible source of Pc-5 ULF wave activity.

Pc-5 perturbations in the Earth's magnetosphere happen on timescales such that for a typical MeV electron in the outer radiation belt, the first two invariants would be conserved while the third invariant is violated. These circumstances will result in a net change in the radial distance at which the particle drifts. Since the first invariant is conserved, the perpendicular momentum of the particle increases if the particle drifts radially inward into regions of higher magnetic field strength, and decreases if it moves radially outward into regions of lower field strength. This radial motion of particles is most effective when Pc-5 ULF waves have frequencies matching the drift frequency of the particle in question. The resonance condition in an azimuthally-symmetric field configuration for this interaction is given by $\omega = m\omega_d$, where ω is the wave frequency, m is the azimuthal wave mode number, and ω_d is the drift frequency of the particle [*Fälthammar*, 1965; *Schulz and Lanzerotti*, 1974]. Under circumstances where the particles are acted on by a spectrum of waves at different frequencies, the particles' radial motion becomes stochastic and can be quantitatively described in a diffusion formalism.

1.8 Radial Diffusion Formalism

A large number of small, random perturbations in the electric and the magnetic field lead to a net change in a particle's radial position, which can be modeled as a diffusive process. If we assume that the first and the second adiabatic invariants of a particle are conserved, while the third adiabatic invariant is violated, then in Equation (1.14) the only nonzero term in the diffusion tensor would be $D_{\Phi\Phi}$, leaving us with only the term D_Φ in the drift vector. The three-dimensional Fokker-

Planck equation reduces to a simple one-dimensional diffusion equation, which after transforming the coordinate $\Phi \rightarrow L$ and utilizing the Jacobian, yields

$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2} \cdot \frac{\partial f}{\partial L}\right), \tag{1.15}$$

where f is the phase space density of the electrons averaged over all phase angles, and L denotes the Roederer L^* parameter [*Roederer and Zhang*, 2014]. Recall that the radial diffusion coefficient is defined as

$$D_{LL} = \frac{\langle (\Delta L)^2 \rangle}{2\Delta t}, \tag{1.16}$$

where Δt is a time period much longer than the characteristic timescale of the third invariant, and ΔL is the change in L at a time Δt units later. Hence D_{LL} is the mean square radial displacement of an ensemble of particles and quantifies the rate of diffusion of charged particles in the radial direction.

Theoretical analysis by *Fälthammar* [1965] used stochastic perturbations in the electric and the magnetic field to derive the appropriate transport coefficients, but the assumptions included only the fluctuations in the $m = 1$ mode, and in the case of magnetic field fluctuations, were valid only for non-relativistic particles. *Fei et al.* [2006] generalized the expressions to include contributions from all azimuthal wave mode numbers as well as relativistic effects. *Elkington et al.* [1999] and *Elkington et al.* [2003] have extended the theoretical framework for radial diffusion by including drift resonant interactions between electrons as well as different polarization (toroidal and poloidal) modes of ULF waves in an asymmetric field.

Since variations in both the magnetic and the electric field can cause radial diffusion, the diffusion coefficient D_{LL} can be written as the sum of both the electric D_{LL}^E and the magnetic D_{LL}^B components which respectively take into account the perturbations in the potential electric and the global magnetic field. In a dipole magnetic field both components are given by *Fei et al.* [2006],

$$D_{LL}^E = \frac{L^6}{8B_0^2 R_E^2} \sum_{m=1}^{\infty} P_m^E(m\omega_d), \tag{1.17}$$

$$D_{LL}^B = \frac{M^2 L^4}{8q^2 R_E^4 B_0^2 \gamma^2} \sum_{m=1}^{\infty} m^2 P_m^B(m\omega_d). \tag{1.18}$$

Here R_E is Earth's radius, B_0 is the strength of the equatorial geomagnetic field at the Earth's surface, M is the first adiabatic invariant, γ is the relativistic factor, q is the charge of the particle, ω_d is the drift frequency of the particle, and m is the azimuthal wave mode number. P_m^E and P_m^B are the power spectral densities (denoted PSDs for the rest of this thesis) of the azimuthal component of the electric field and the compressional component of the magnetic field, respectively. D_{LL}^E includes contributions from both the convective and the inductive electric field while D_{LL}^B considers only the magnetic field perturbations.

We would like to point out here that *Fälthammar* [1965] assumed a phase relation between the magnetic and the induced electric fields while *Fei et al.* [2006] made no such assumption. *Perry et al.* [2005, Figure 1] showed that the magnetic field phase and the induced electric field phase may not be independent as assumed by *Fei et al.* [2006]. *Perry et al.* [2005] developed a three-dimensional model to study the guiding-center trajectories of relativistic electrons. They used ground magnetometer data to specify the ULF wave power to drive the energetic particles, and reported that the electric and the magnetic field phases are directly correlated. But since it is very difficult to separate the total measured electric field from single point measurements in space [*Brautigam et al.*, 2005] into its convective and inductive components, we consider the total electric field to compute D_{LL}^E and only the magnetic field to compute D_{LL}^B for the rest of this thesis.

It should be noted here that Equations (1.17) and (1.18) require the summation of wave power from all contributing spatial modes m. Global numerical simulations such as *Elkington et al.* [2002, 2004]; *Claudepierre et al.* [2008]; *Claudepierre* [2008]; *Elkington et al.* [2012] enabled studies of modes $m > 1$ so that the distribution of wave power in various modes can be quantified. However, for *in situ* measurements, in order to study the wave power from the first m spatial modes we require a constellation of at least $2m$ satellites. Since that is rarely possible, we make the assumption that all of the wave power is in the first, $m = 1$ mode, and that the power in all other modes was zero. This assumption was made explicitly by *Brautigam et al.* [2005] when analyzing the CRRES electric field measurements and is further justified by the recent results presented by *Elkington et al.* [2012] and *Tu et al.* [2012]. *Elkington et al.* [2012] concluded that during the main

phase of a storm, the magnetic field power mostly resided in the first mode, while the electric field power had a significant fraction of its total power in the second and the third mode. However during the recovery phase, the first mode predominated in both the electric and the magnetic fields. *Tu et al.* [2012] showed that the power at $m = 1$ is always dominant, but less so during enhanced geomagnetic activity with high AE index. Equations (1.17) and (1.18) also implicitly assume a uniform distribution of wave power in azimuth. In reality, the azimuthal distribution of the ULF waves in the magnetosphere depends on the physical origin of the waves and their propagation characteristics. For example, ULF activity due to the Kelvin-Helmholtz instability is expected to be largely confined to regions near dawn and dusk flanks [*Claudepierre et al.*, 2008], while ULF waves resulting from pressure perturbations in the solar wind will be largely distributed across the dayside with less power appearing in the midnight sector [*Huang et al.*, 2010a].

1.9 Previous Results

The analytical treatment by *Fälthammar* [1965] and *Fälthammar* [1968] showed that the perturbations in the induced electric field ($\frac{8}{15}$ of the total magnetic flux variations) are roughly as important as the perturbations in the magnetic field ($\frac{7}{15}$ of the total magnetic flux variations) in driving particle diffusion. The electromagnetic radial diffusion coefficient, denoted D_{LL}^{M}, contains contributions from quasi-transverse fluctuations in both the magnetic field and the induced electric field. The electrostatic radial diffusion coefficient, denoted $D_{LL}^{E_{\text{static}}}$, contains contributions only from the quasi-electrostatic fluctuations in the convective electric field. The total time-dependent radial diffusion coefficient can then be computed as

$$D_{LL} = D_{LL}^{M} + D_{LL}^{E_{\text{static}}}. \tag{1.19}$$

Since the *Fälthammar* [1965] and *Fälthammar* [1968] expressions for the radial diffusion coefficients are functions of the power spectral density of compressional magnetic and azimuthal electric field components in the ULF range in the magnetic equatorial plane, there have been various attempts to measure the PSDs experimentally and to estimate D_{LL}. *Frank* [1965]; *Newkirk and*

Walt [1968]; *Lanzerotti et al.* [1970]; *Lyons and Williams* [1975]; *West et al.* [1981]; *Selesnick et al.* [1997] conducted studies using particle observation over long periods of time assuming a single D_{LL} fixed for all times independent of geomagnetic activity. *Mozer* [1971] and *Holzworth and Mozer* [1979] used data from balloon campaigns to determine the Kp-dependence of the electric field PSDs. Magnetic field power spectra have been analyzed and D_{LL}^{M} have been computed at $L = 4$ [*Lanzerotti and Robbins*, 1973; *Lanzerotti and Morgan*, 1973] and at $L = 6.6$ [*Lanzerotti et al.*, 1978; *Arthur et al.*, 1978; *Huang et al.*, 2010b; *Tu et al.*, 2012]. *Brautigam and Albert* [2000] showed that in order to model magnetic storm time behavior of relativistic electrons, an activity-dependent D_{LL} is necessary. Here they derived an analytic expression for the electromagnetic component of the diffusion coefficient as a function of Kp and L-shell value, using ground-based magnetic field measurements at $L = 4$ [*Lanzerotti and Morgan*, 1973] and $L = 6.6$ [*Lanzerotti et al.*, 1978]. They concluded that the electromagnetic diffusion coefficients are much larger than the electrostatic diffusion coefficients for particles with high energy in the outer radiation belt. For lower-energy particles with the first invariant $M < 100$ MeV/G and $L \leq 3$, D_{LL}^{E} can dominate D_{LL}^{M}. *Brautigam et al.* [2005] used *in situ* measurements of the electric field from the Combined Release and Radiation Effect Satellite (CRRES) to estimate the electrostatic component of the radial diffusion coefficient as a function of Kp and L-shell value, and concluded that

$$D_{LL}^{E}[\text{CRRES}] \leq D_{LL}^{E}[\text{BA}] \ll D_{LL}^{M}[\text{BA}] \tag{1.20}$$

where $D_{LL}^{E}[\text{BA}]$ and $D_{LL}^{M}[\text{BA}]$ represent the *Brautigam and Albert* [2000] estimates.

We would like to clarify here that despite *Brautigam et al.* [2005] using the *Fälthammar* [1965] formulation to compute $D_{LL}^{E}[\text{CRRES}]$, they made no effort to separate the convective and the inductive electric fields, and used the total electric field to estimate $D_{LL}^{E}[\text{CRRES}]$ in their study. This is simply because the CRRES electric field instrument made a single point measurement of the total electric field [*Wygant et al.*, 1992] and from a single point measurement it is not possible to separate the measured total electric field into its inductive and convective components. This assumption is explicitly stated in *Brautigam et al.* [2005, see paragraph 13].

Fei et al. [2006] assumed no phase relation between the electric and the magnetic field perturbations and derived expressions for the magnetic diffusion coefficient D_{LL}^{B}, containing only the contribution from the compressional magnetic field perturbations, and the total electric diffusion coefficient $D_{LL}^{E_{\text{total}}}$, containing contributions from both the convective and the inductive azimuthal electric field perturbations. The total time-dependent radial diffusion coefficient is then computed as

$$D_{LL} = D_{LL}^{B} + D_{LL}^{E_{\text{total}}}. \tag{1.21}$$

It is this formulation which has been used to compute radial diffusion coefficients in more recent efforts [*Huang et al.*, 2010b; *Ozeke et al.*, 2012; *Tu et al.*, 2012; *Ozeke et al.*, 2014; *Ali et al.*, 2015]. *Tu et al.* [2012] computed radial diffusion coefficients using global MHD studies for the March 2008 storm and reported that the electric diffusion coefficients in this event were dominant, contrary to the conclusions of *Brautigam and Albert* [2000]. *Tu et al.* [2012] used *in situ* measurements from THEMIS and GOES to validate their MHD results. The Lyon-Fedder-Mobarry (LFM) [*Lyon et al.*, 2004] MHD fields contained roughly the same amount of wave power in the magnetic field as the observations but the LFM MHD wave power in the electric field was an underestimate of the *in situ* measurements by about a factor of ten. *Ozeke et al.* [2012] and *Ozeke et al.* [2014] used *in situ* magnetic field measurements from AMPTE at $L = 3, 4$, and 5 [*Takahashi and Anderson*, 1992] and from GOES at $L = 6.6$, to compute D_{LL}^{B}[Ozeke] along with ground-based magnetic field PSDs mapped to electric field PSDs in space to compute D_{LL}^{E}[Ozeke]. They arrived at the same conclusion, that the magnetic diffusion coefficients are not as effective in driving radial diffusion as previously thought. The electric diffusion coefficients determined in this study were dominant, sometimes by as much as two orders of magnitude [*Ozeke et al.*, 2012, 2014].

The goal of this thesis is two-fold. First, using the measurements from the Combined Release and Radiation Effects Satellite (CRRES) and the Van Allen Probes (RBSP), we estimate the power spectral densities (PSD) of the electromagnetic fields in the Pc-5 range and study the distribution of electromagnetic wave power in azimuth, parameterizing the wave power with magnetic local time

(MLT), L, and level of geomagnetic activity Kp. Second, we use the *Fei et al.* [2006] formulation to compute the magnetic and electric radial diffusion coefficients using L and Kp as parameters. We then compare our results with the previous results of *Brautigam and Albert* [2000] and *Brautigam et al.* [2005], as well as some more recent estimates. Specifically, we compare D_{LL}^{B} and D_{LL}^{E} to determine the relative importance of each of the two components in driving radial diffusion.

Chapter 2

Perils of Using Least Squares

In order to change a nonlinear regression problem into an easier, linear regression problem, it is common to transform the data as well as the model. For example, if the model is a simple power law, an exponential, or a Gaussian function, the data can be mapped into log-log space and a nonlinear least squares problem is converted into a linear least squares problem. However, such transformations skew the error distributions and may violate the assumptions necessary for the least squares method to work. In this chapter, we present qualitative and quantitative arguments that the least squares method provides skewed and biased results if nonlinear transformations (such as the log transformation) are performed unreservedly before the fitting procedure. We then present several space physics examples which illustrate how varied the parameter estimates can be if the data is transformed. We conclude that least squares should be performed on the original data, even if it is a nonlinear problem assuming that all of the least squares assumptions are satisfied. The data may be transformed if the transformed data will satisfy the least squares assumptions. The data should not be arbitrarily transformed just to make a problem easier. Such transformations may lead to highly misleading results. We also present a flowchart for the reader to use in order to pick the appropriate least squares method, depending on which assumptions are satisfied by the data and which assumptions are violated.

2.1 Introduction

In this day and age of vast amounts of data collection and storage, there is no shortage of information. Data collection methods are ubiquitous and are used in all fields. However, this has only exacerbated the problem of data processing. After collecting data, how can it be processed so that meaningful information can be extracted from it? There are many methods available, depending on the nature of the data collected and on what kind of information is to be extracted. For scientific purposes, one of the most popular methods is that of least squares (LS) [*Gauss*, 1823; *Legendre*, 1805]. LS is a generalized regression model; its most common applications are curve fitting, which can then be use for interpolation [*Green and Kivelson*, 2001; *Lejosne et al.*, 2013; *Olson and Pfitzer*, 1974], detrending the data by removing slow variations [*Hudson et al.*, 2004], estimating the linear trend in a dataset [*Holzworth and Mozer*, 1979; *Huang et al.*, 2010b; *Ruohoniemi et al.*, 1991; *Elkington et al.*, 2003], and understanding parameter dependence [*Li et al.*, 2014; *Mead*, 1964; *Mozer*, 1971; *Perry et al.*, 2005; *Rae et al.*, 2012].

Consider the case of one independent variable x and one dependent variable y. Let (x_i, y_i) be a set of given/measured data points with the fitting function being $y = f(\mathbf{a}; x)$ where the components of the vector $\mathbf{a}$ are the unknown parameters we wish to estimate using LS. There are several assumptions needed to derive LS [*Gauss*, 1823; *Legendre*, 1805; *Draper and Smith*, 1966; *Hastie et al.*, 2009]. It is first assumed that the measured data involves some sort of error, which exists only in the dependent variable (y_i). The error in the independent variable (x_i) is either absent or negligible. Therefore the model that we assume is

$$y_i = f(\mathbf{a}; x_i) + \epsilon_i, \tag{2.1}$$

the additive error model in which the errors have been added to the observations. It is further assumed that ϵ is a random variable with mean zero (exogeneity) and no correlation between any of the observed errors ϵ_i. Note here that the distribution of errors is not specified. Errors being normally distributed is not at all needed for the validity of the LS method. However, in most real life applications measurements tend to have normally distributed errors, due to the central

limit theorem which states that under certain conditions the sum of many random variables will be approximately normally distributed. In particular when the error term ϵ_i is the sum of errors from several sources, we can assume that the ϵ_i are normally distributed [*Cramér*, 1946].

While taking measurements, we can also assume that only random error is present and that there is no systematic error of any kind unless there is strong evidence of its presence, in which case it must be removed. Therefore, the assumption that the error is normally distributed with mean zero and variance σ^2, $\epsilon \sim N(0, \sigma^2)$, is very reasonable and it enables LS to possess additional desirable properties. If normality of the errors is assumed, then LS is equivalent to the maximum likelihood estimator (which selects parameter values to maximize the agreement between the observed data and the model selected) [*Margenau and Murphy*, 1956]; it also achieves the Cramér-Rao lower bound [*Rao*, 1945, 1973; *Cramér*, 1946] and is therefore efficient in the statistical sense. LS being efficient implies that LS has the lowest possible mean squared error among the set of all unbiased estimators. Furthermore, constant variance across all measurements (homoscedasticity) is also not necessary in order to obtain unbiased estimates from the LS method. However, both the error distribution being normal and homoscedasticity assumptions are needed to construct confidence intervals, interval estimates for a statistic, such as those provided by all of the common statistical packages, such as MATLAB, IDL, R, and SciPy. If normality cannot be assumed, then the sample size must be large enough that we can appeal to the law of large numbers and the central limit theorem to obtain the confidence intervals [*Draper and Smith*, 1966; *Hastie et al.*, 2009].

The least squares method seeks to minimize the sum of the squared errors (SSE),

$$\text{SSE} = \sum_i (y_i - f(\mathbf{a}; x_i))^2, \tag{2.2}$$

as a function of the parameters a_j. The SSE is minimized when the gradient vector with respect to the parameters a_j is zero, namely

$$\frac{\partial(\text{SSE})}{\partial a_j} = 0. \tag{2.3}$$

If $f(\mathbf{a}; x)$ is a linear function in each of the components of $\mathbf{a}$ then the SSE will be a quadratic surface, and we know for a fact that a minimum value exists, it is unique, and it can be easily found by

solving the linear system given in Equation (2.3). This method is called linear least squares and has a closed-form solution [*Draper and Smith*, 1966; *Rao and Toutenburg*, 1999; *Hastie et al.*, 2009].

In contrast, if $f(\mathbf{a}; x)$ is a nonlinear function of any of the components of $\mathbf{a}$, then Equation (2.3) is a nonlinear system which may be algebraic, trigonometric, etc., depending on the form of $f(\mathbf{a}; x)$. The problem then becomes a generic optimization/root-finding problem. This method is called nonlinear least squares [*Draper and Smith*, 1966; *Rao and Toutenburg*, 1999; *Hastie et al.*, 2009] and in general it has no closed-form solution, though it is possible that a particular problem may have a closed-form solution. Usually, however, the problem requires an initial value for $\mathbf{a}$ and iterative methods with stopping criteria so that the solution can be refined iteratively and stopped when the stopping criteria are met. The stopping criterion may be a maximum number of iterations, a lower bound on the difference between two successive approximations, or a combination of both. Here the SSE surface can take on any form, depending on the data and on the form of the fitting function. There can be multiple minima, both global and local. In fact, there can be infinitely many global minima (such as for a trigonometric system); which minimum is found depends on the initial values and the algorithm used. There can be convergence issues depending on the gradient of the SSE in the neighborhood of the minimum. There are both gradient-based and gradient-free methods available for such nonlinear problems [*Press et al.*, 2007; *Kelley*, 1999]. Note here that the linearity of $f(\mathbf{a}; x)$ in x is not used in the classification at all. We are only concerned with the linearity of $f(\mathbf{a}; x)$ in $\mathbf{a}$.

2.2 Data Transformation

Often in the context of regression, the model we wish to fit is a nonlinear model. It is a common practice to transform the model nonlinearly so that a nonlinear regression problem becomes a much easier linear regression problem. Consider, for example, the simple power law model $y = a_1 x^{a_2}$, which is nonlinear in a_2 and necessitates the use of nonlinear LS. The quantity

we seek to minimize is

$$\mathrm{SSE}_{\mathrm{linear}} = \sum_i \left(y_i - a_1 x_i^{a_2}\right)^2 . \tag{2.4}$$

However, if we transform the model to $\ln(y) = \ln(a_1) + a_2 \ln(x)$ by taking the logarithm of both sides, we now have a linear problem in $\ln(a_1)$ and a_2 which can be solved easily with linear LS by minimizing

$$\mathrm{SSE}_{\log} = \sum_i \left(\ln(y_i) - (\ln(a_1) + a_2 \ln(x_i))\right)^2 . \tag{2.5}$$

Although we have transformed a nonlinear problem into a linear problem, we have also transformed the observed data as well as the measured errors included in the measurements. Instead of performing LS on y_i vs. x_i, we would be performing LS on $\ln(y_i)$ vs. $\ln(x_i)$ at this point. Clearly, whatever the distribution of errors in y_i, the distribution of errors in $\ln(y_i)$ will be different.

Specifically, if we assume the error distribution to be normal in y_i then the errors in $\ln(y_i)$ will not be normal. Therefore LS is no longer the most efficient method [*Rao*, 1945, 1973; *Cramér*, 1946]. But more importantly, if we assume the errors in y_i to have mean zero then the errors in $\ln(y_i)$ will certainly not have mean zero, and hence the LS method is rendered invalid. LS will provide biased estimates [*Draper and Smith*, 1966; *Rao*, 1973; *Rao and Toutenburg*, 1999]. This can be understood intuitively by seeing that the logarithm is not a linear transformation. The expressions for $\mathrm{SSE}_{\mathrm{linear}}$ and $\mathrm{SSE}_{\log}$, given by Equations (2.4) and (2.5), are different and the values where the minimum is achieved can be quite different. This is simply because the logarithm function is a strictly concave function and

$$\begin{aligned} &\text{for } x, y \in (0,1), \quad && |\ln(x) - \ln(y)| > |x - y|, \\ &\text{for } x, y \in (1,\infty), \quad && |\ln(x) - \ln(y)| < |x - y|, \end{aligned}$$

so a relatively small error in a small y_i ($y_i < 1$) will be magnified and a large error in a large y_i ($y_i > 1$) will be diminished in log-log space. Hence, LS will provide skewed results in log-log space.

To quantify the effects of the log transformation, we present some analysis. Let $X \sim N(\mu, \sigma^2)$ be a normal random variable with mean μ and variance σ^2. Recall that the probability distribution

function (PDF) for X is

$$f(\mu, \sigma; x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{2.6}$$

with the expected value given by the first moment

$$E(X) = \int_{-\infty}^{\infty} x f(\mu, \sigma; x) dx = \mu. \tag{2.7}$$

Simply taking the logarithm of X doesn't make sense because X can take negative values. So we consider $\ln|X|$. If μ is strictly positive and the variance σ^2 is small enough then taking the absolute value has no practical effect because the data will be concentrated near the positive mean μ. In order to compute the expected value of $\ln|X|$, we don't have to compute the probability density function of $\ln|X|$. We can just use the law of the unconscious statistician which allows us to compute the expected value of a function $g(X)$ of a random variable X without explicitly knowing the probability distribution function of $g(X)$ [*Ross*, 1981]

$$E(\ln|X|) = \int_{-\infty}^{\infty} \ln|x| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx, \tag{2.8}$$

giving us

$$E(\ln|X|) = \ln\left(\frac{\sigma}{\sqrt{2}}\right) - \frac{\gamma}{2} + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2^{n+1}} \cdot \frac{\sqrt{\pi}}{n\Gamma(n+1/2)} \cdot \left(\frac{\mu}{\sigma}\right)^{2n}, \tag{2.9}$$

where $\gamma = 0.5772156649...$ is the Euler-Mascheroni constant [*Conway and Guy*, 1996] and $\Gamma(x)$ is the analytic continuation of the factorial function [*Abramowitz and Stegun*, 1965].

Next, let us quantify what happens to the error distribution if the data is mapped to log-log space. Imagine a simple case in which we are trying to measure some constant quantity c by taking repeated measurements and collecting c_i. Each measurement has some random error ϵ_i which we assume to be normally distributed with mean zero and some constant unknown variance σ^2. Each measurement can therefore be expressed as $c_i = c + \epsilon_i$, after which we take its logarithm and subtract $\ln(c)$ to obtain the transformed errors. In order to study the error distribution in log space, we see that,

$$\ln(|c + \epsilon_i|) - \ln(|c|) = \ln\left(\left|1 + \frac{\epsilon_i}{c}\right|\right), \tag{2.10}$$

and the error ϵ being normally distributed with mean zero and variance σ^2 implies that the quantity $1 + \epsilon/c$ will be normally distributed with mean 1 and variance σ^2/c^2, i.e.,

$$\epsilon \sim N(0, \sigma^2) \Rightarrow \left(1 + \frac{\epsilon}{c}\right) \sim N\left(1, \frac{\sigma^2}{c^2}\right). \tag{2.11}$$

Substituting $X = 1 + \epsilon/c$ in Equation (2.9) gives us the expected value of our transformed errors,

$$E(\ln\left(\left|1 + \frac{\epsilon}{c}\right|\right)) = \ln\left(\frac{\sigma}{c\sqrt{2}}\right) - \frac{\gamma}{2} + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2^{n+1}} \cdot \frac{\sqrt{\pi}}{n\Gamma(n + 1/2)} \cdot \left(\frac{c}{\sigma}\right)^{2n}, \tag{2.12}$$

which is certainly not zero. The expected value depends on both the true value being measured, c, and on the original error variance σ^2, as can reasonably be expected. Since the transformed errors no longer have mean zero, LS is rendered invalid and should not be used on the transformed data. It will return biased results.

Figure 2.1 compares the error distribution before and after a log transformation. In this example, the true value which we are trying to estimate is $c = 1$ and the errors are normally distributed with mean zero and variance $\sigma^2 = 0.09$ ($\epsilon \sim N(0, 0.09)$). The panel on the top left shows the true value as a solid black horizontal line; the blue points are repeated measurements taken over time with random error ϵ_i added. The panel on the top right shows the logarithm of the measured data, with the black horizontal line representing $\ln(c)$. The nonlinearity of the log transformation is evident here, where the errors are no longer symmetrically distributed about zero. Measured values $c_i < 1$ have been magnified while measured values $c_i > 1$ have decreased in magnitude. The panel at the bottom compares the error distributions from both spaces. The blue curve is the probability density function (PDF) of the errors in linear space; we can see it is indeed normal about mean zero. The orange curve is the probability density function of the transformed errors in log-log space. The orange curve is clearly neither normal nor symmetric about its mean. The transformed error distribution is negative skewed; the left tail is heavier than the right tail.

2.3 An Illustrative Example

In this section we present a specific example of a regression problem in order to show how a least squares estimate can change if a nonlinear transformation is first applied to the data. We

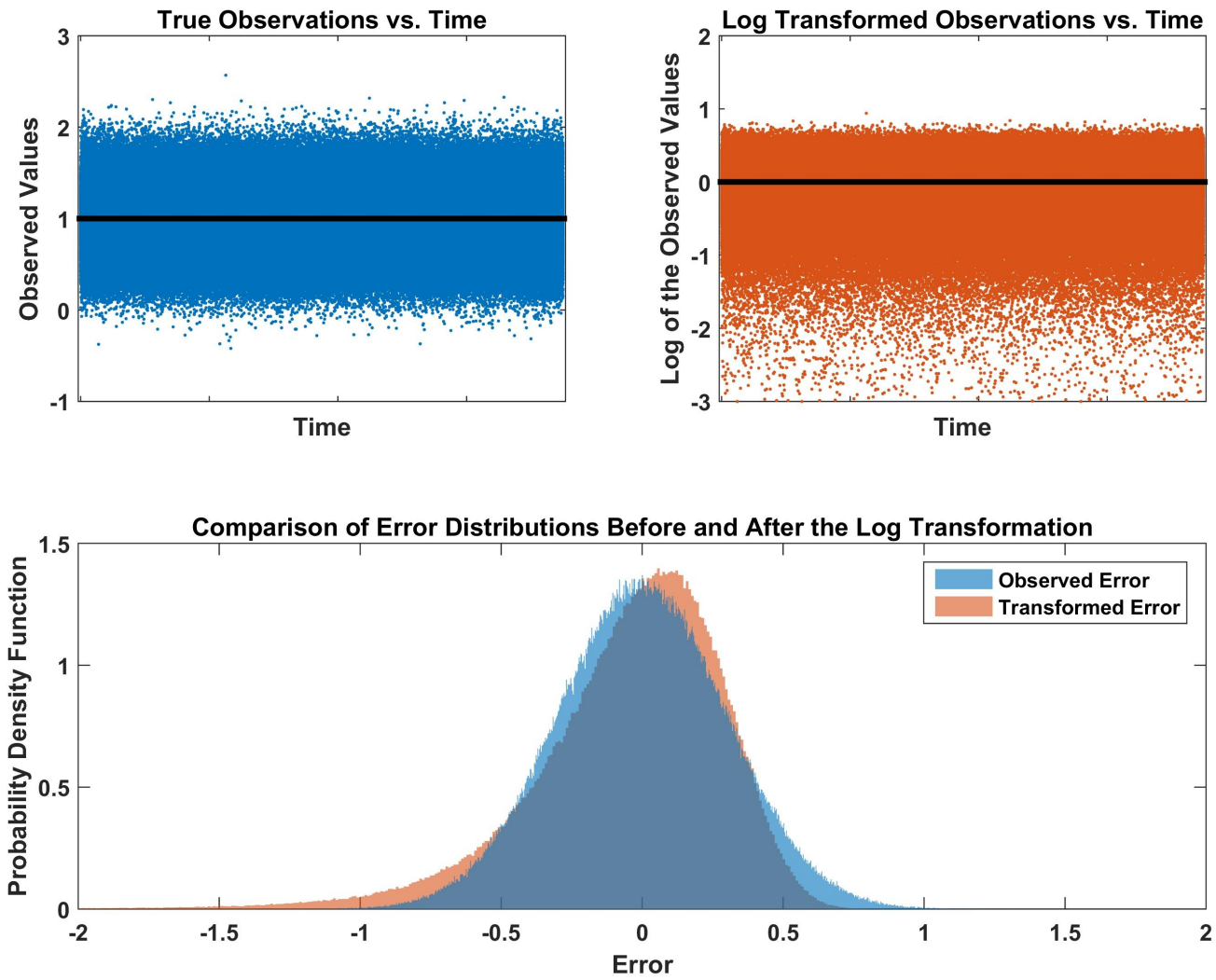


Figure 2.1: Distortion of errors and their distribution under the log transformation. In the top two panels, the solid black line represents the actual quantity that is being measured. The top left shows the actual measurements in blue while the top right panel shows the log-transformed measurements in orange. The bottom panel shows the probability density functions (PDF) of both the measured and the transformed errors.

present a simple example so that a closed-form solution can be obtained in all cases. This example is not meant to represent a model used in any real application. It is meant only to show how nonlinear transformations can affect the LS estimates. Let (x_i, y_i) be a set of given data points and let the form of the fitting function be $y = \frac{1}{ax}$. We wish to estimate a single parameter a. Since our fitting function depends nonlinearly on a, we have a nonlinear regression problem. If we were to apply ordinary least squares, our goal would be to minimize

$$\text{SSE} = \sum_i \left(y_i - \frac{1}{ax_i} \right)^2 . \tag{2.13}$$

If we take its derivative with respect to a, set it equal to zero, and solve for a, we obtain,

$$\begin{aligned} \frac{\partial(\text{SSE})}{\partial a} = \sum_i 2 \left(y_i - \frac{1}{ax_i} \right) \left(\frac{x_i}{(ax_i)^2} \right) &= 0 \\ \sum_i \left(y_i - \frac{1}{ax_i} \right) \left(\frac{1}{x_i} \right) &= 0 \\ \sum_i \left(\frac{y_i}{x_i} - \frac{1}{ax_i^2} \right) &= 0 \\ \sum_i \left(\frac{y_i}{x_i} \right) - \frac{1}{a} \sum_i \left(\frac{1}{x_i^2} \right) &= 0 \\ \frac{\sum_i (1/x_i^2)}{\sum_i (y_i/x_i)} &= a. \end{aligned} \tag{2.14}$$

Equation 2.14 presents the correct exact solution for the least squares problem. On the other hand, if we attempt to linearize the regression problem by taking the logarithm of both sides of the fitting function, we obtain,

$$\text{SSE}_{\log} = \sum_i \left(\ln(y_i) + \ln(a) + \ln(x_i) \right)^2 , \tag{2.15}$$

which can be minimized as,

$$\begin{aligned}
\frac{\partial(\mathrm{SSE}_{\log})}{\partial a} = \sum_i 2\left(\ln(y_i) + \ln(a) + \ln(x_i)\right)\left(\frac{1}{a}\right) &= 0 \\
\sum_i \left(\ln(y_i) + \ln(a) + \ln(x_i)\right) &= 0 \\
\sum_i \ln(y_i) + \sum_i \ln(a) + \sum_i \ln(x_i) &= 0 \\
\sum_i \ln(y_i) + \sum_i \ln(x_i) &= -n\ln(a) \\
\prod_i \left(\frac{1}{x_i y_i}\right)^{1/n} &= a, \qquad (2.16)
\end{aligned}$$

where $\prod$ is compact notation for multiplication analogous to $\sum$ denoting summation. Another transformation which we can perform on the fitting function is to take the reciprocal of both sides, giving us,

$$\mathrm{SSE}_{\mathrm{reciprocal}} = \sum_i \left(\frac{1}{y_i} - a x_i\right)^2, \qquad (2.17)$$

which can be minimized as,

$$\begin{aligned}
\frac{\partial(\mathrm{SSE}_{\mathrm{reciprocal}})}{\partial a} = \sum_i 2\left(\frac{1}{y_i} - a x_i\right)(-x_i) &= 0 \\
\sum_i \left(\frac{x_i}{y_i} - a x_i^2\right) &= 0 \\
\frac{\sum_i (x_i/y_i)}{\sum_i x_i^2} &= a. \qquad (2.18)
\end{aligned}$$

As we can clearly see, the correct ordinary least squares estimate given by Equation 2.14 is completely different from the estimates given by Equation 2.16 and Equation 2.18.

2.4 Comparison of the Estimates

In this section, we will present some specific examples of space physics data and show how the parameter estimates can change if we use linear least squares incorrectly. One of the most common nonlinear functional forms used in the physical sciences is the simple power law $y = a_1 x^{a_2}$. Instead of using nonlinear LS (with weights if necessary), the data is usually transformed into log-log space and then linear LS is used with $\ln(y) = \ln(a_1) + a_2 \cdot \ln(x)$. The scaling exponent a_2 is estimated as

the slope of the straight line in log-log space, while the coefficient a_1 is estimated as the exponential of the intercept of the straight line. Taking the log of the data, however, introduces a bias. If the original data satisfies the LS assumptions then taking the log of the data will most certainly violate those assumptions. Nonlinear and linear LS will provide different results, sometimes noticeably so, depending on how much bias is introduced by the log transformation.

Our first example is taken from Chapter 3 of this thesis, in which the fluxgate magnetometer data from the Combined Release and Radiation Effects Satellite (CRRES) is used to estimate the magnetic component of the radial diffusion coefficient D_{LL}^{B}[CRRES]. Kp is the activity-dependent parameter used and for various fixed values of the first invariant M, D_{LL}^{B} was computed as a function of L [*Roederer and Zhang*, 2014]. In order to investigate the dependence on L of the diffusion coefficients, we assumed the functional form to be a simple power law $D_{LL} = D_0 L^n$ and estimated the scaling exponent n, starting with $M = 500$ MeV/G and increasing to $M = 5000$ MeV/G with $\Delta M = 100$ MeV/G separated by Kp, as shown in Figure 3.10.

For Figure 2.2, each set of diffusion coefficients, with the first invariant M held constant, was fit first as a function of L using weighted nonlinear least squares in linear space without any kind of data transformation. The weights used were proportional to $1/y_i^2$. Next we transformed the diffusion coefficients to log-log space and used linear least squares to estimate the slope of the linear trend. The exponents from both methods are plotted in Figure 2.2 for comparison. The solid diagonal line in Figure 2.2 is the line $y = x$, used as a visual aid. The points are color-coded by Kp. For quiet Kp ($Kp = 0, 1$), low Kp ($Kp = 2, 3$), and moderate Kp ($Kp = 4, 5$) we see that the scaling exponents from both of the least squares methods are the similar because they fall on the identity line. For Kp high ($Kp = 6, 7$), however, the difference between the two estimates is quite obvious. For high Kp, nonlinear LS returns a scaling exponent of 11.5 while linear LS will return 10.25 for the same set of diffusion coefficients.

Our second example is taken from Chapter 4 of this thesis, in which we use the magnetic field data from the Electric and Magnetic Field Instrument Suite and Integrated Science (EMFISIS) [*Kletzing et al.*, 2013] and the electric field data from the Electric Field and Waves (EFW)

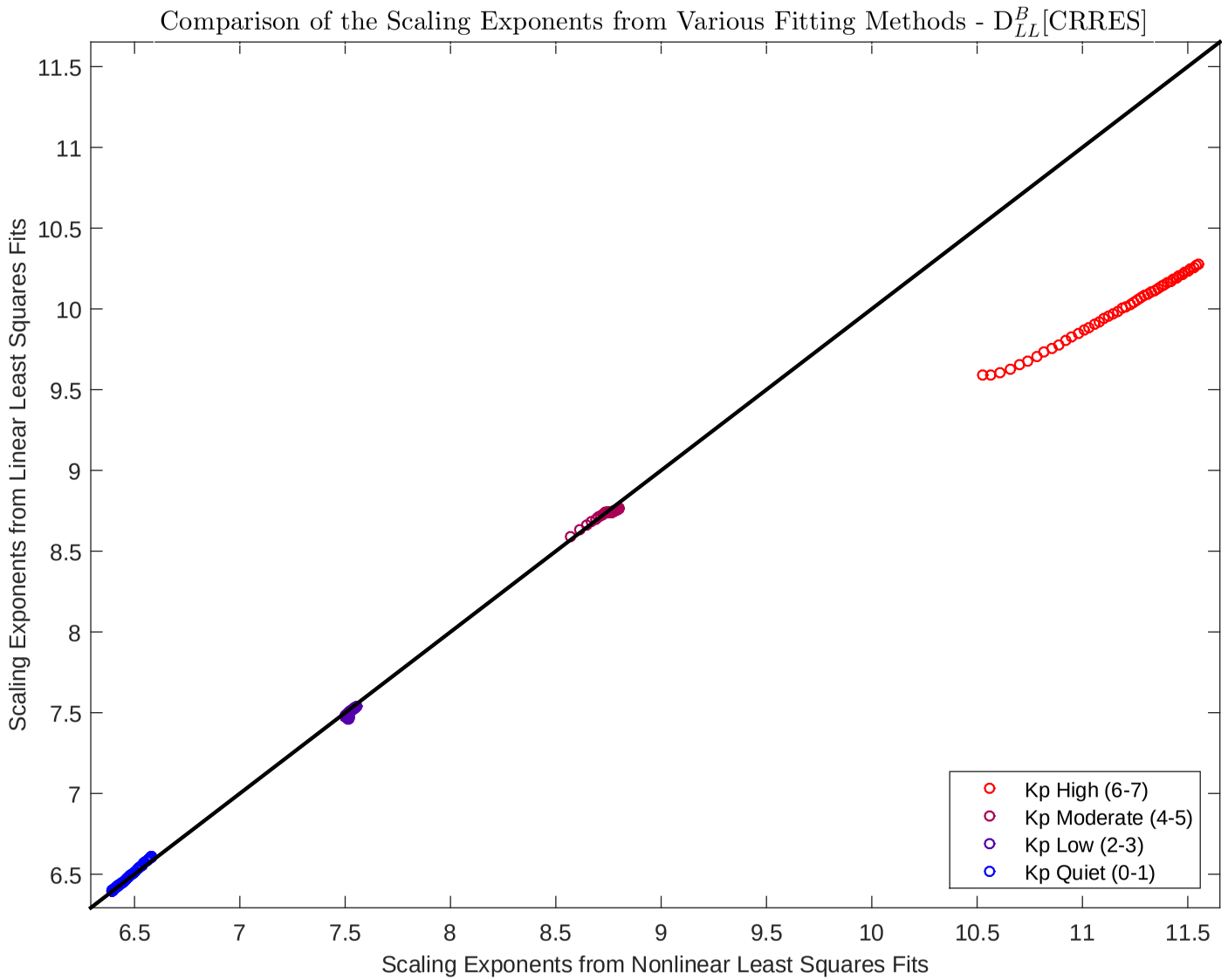


Figure 2.2: Comparison of scaling exponents n obtained using weighted nonlinear least squares without any transformations and linear least squares in log-log space, assuming $D_{LL}^{B}[\text{CRRES}] = D_0 L^n$. The weights used were proportional to $1/y_i^2$.

instrument suite [*Wygant et al.*, 2013] on board the Van Allen Probes. The magnetic and the electric field data were used to estimate the magnetic and the electric components of the radial diffusion coefficients, D_{LL}^{B}[RBSP] and D_{LL}^{E}[RBSP], using the *Fei et al.* [2006] formulation. For the first adiabatic invariant M between $M = 500$ MeV/G and $M = 5000$ MeV/G in steps of $\Delta M = 100$ MeV/G, M is held constant and the diffusion coefficients are fit to a simple power law as a function of L. Figures 2.3 and 2.4, using D_{LL}^{B}[RBSP] and D_{LL}^{E}[RBSP] respectively, show the comparison of the scaling exponents estimated using weighted nonlinear least squares and linear least squares in log-log space. In Figure 2.3, the scaling exponents differ significantly, even for quiet times ($Kp = 0$) while in Figure 2.4, the scaling exponents don't show much variation.

A final comparison between linear least squares and nonlinear least squares is performed on particle count observations taken from the Colorado Student Space Weather Experiment [*Li et al.*, 2013; *Schiller et al.*, 2014], a 3U CubeSat in a low-altitude, high-inclination orbit. The on-board particle telescope measures relativistic electrons in three nominal energy channels: 0.64, 1.84, and 3.41 MeV. In Figure 2.5, the top two panels show three samples taken from the full energy spectrum, which often obey a simple power law ($y = a_1 x^{a_2}$). The top panels include fits using both linear and nonlinear least squares methods. The top left panel compares both fits on a linear scale while the top right panel shows the observed samples with both of the fits plotted on a log-log scale to emphasize the differences. In this dataset, the 3.41 MeV channel has a higher noise floor than the other two observations. We can see that for this specific example, nonlinear LS returns the scaling exponent $a_2 = -5.96$ while the incorrect linear LS performed in log-log space returns $a_2 = -5.04$. The reason for this is that the data point at 3.41 MeV is weighted more heavily than it should be by the linear LS algorithm; this is due to the log transformation. The slope of the line is skewed to be more highly positive.

We take all such data points from the entire two-year mission and consider only the data points that fall within the outer radiation belt with $3 \leq L \leq 6$, where L is the radial distance of the particle during its equatorial crossing in a dipole magnetic field. The reason for considering only particles in the outer radiation belt is to make sure that there are as many counts in the particle

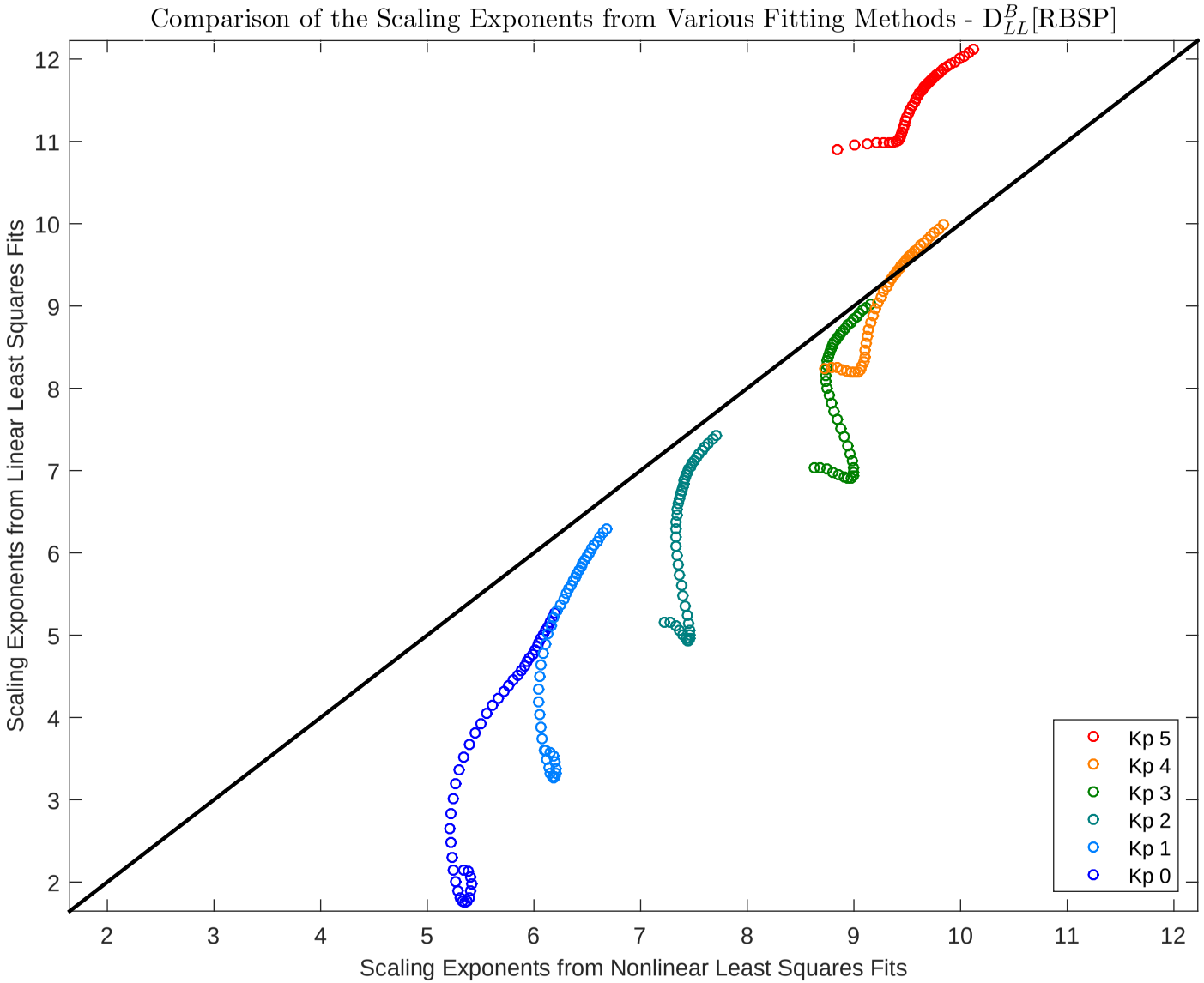


Figure 2.3: Comparison of scaling exponents n obtained using weighted nonlinear least squares without any transformations and linear least squares in log-log space, assuming $D_{LL}^{B}[\text{RBSP}] = D_0 L^n$. The weights used were proportional to $1/y_i^2$.

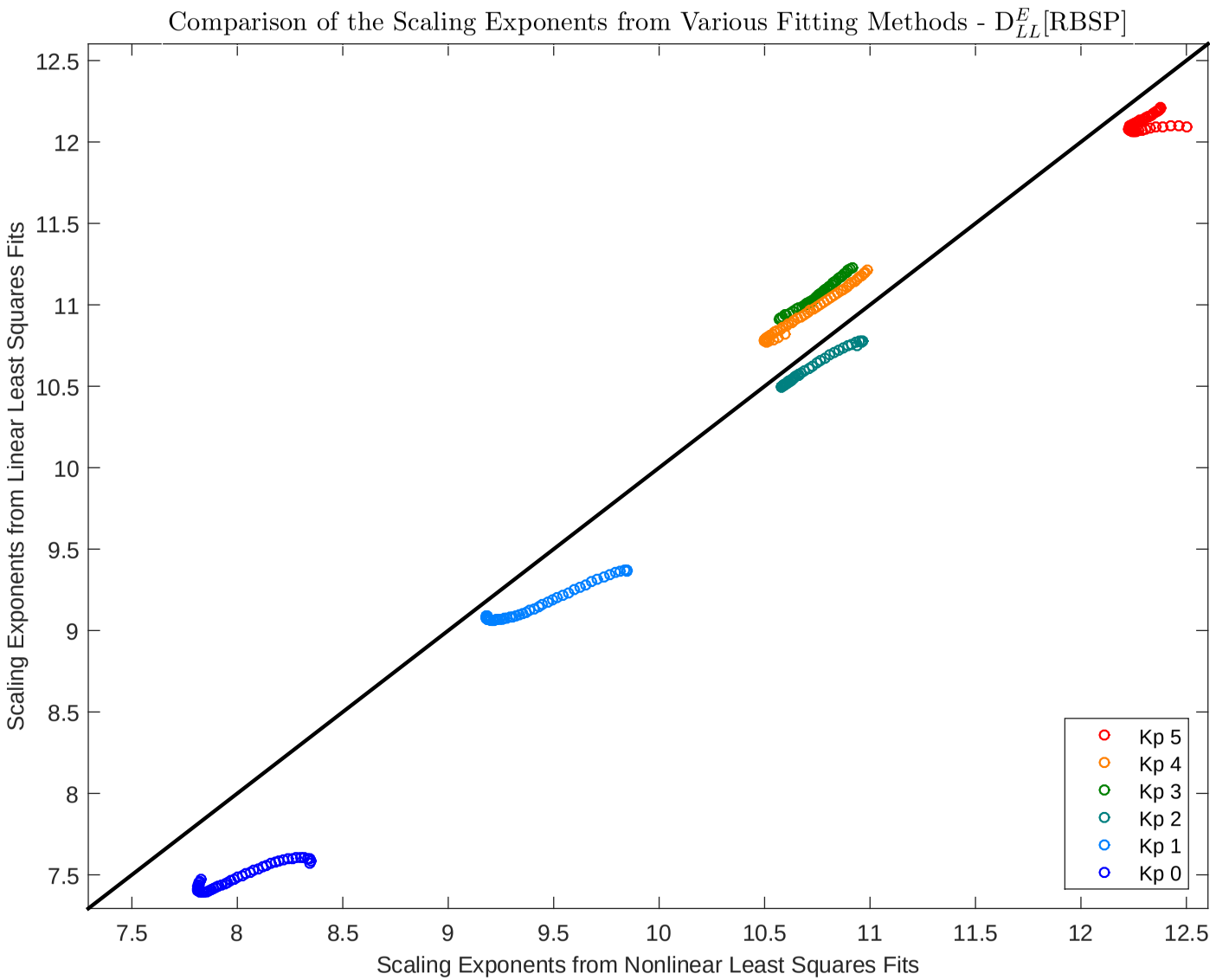


Figure 2.4: Comparison of scaling exponents n obtained using weighted nonlinear least squares without any transformations and linear least squares in log-log space, assuming $D^E_{LL}[\text{RBSP}] = D_0 L^n$. The weights used were proportional to $1/y_i^2$.

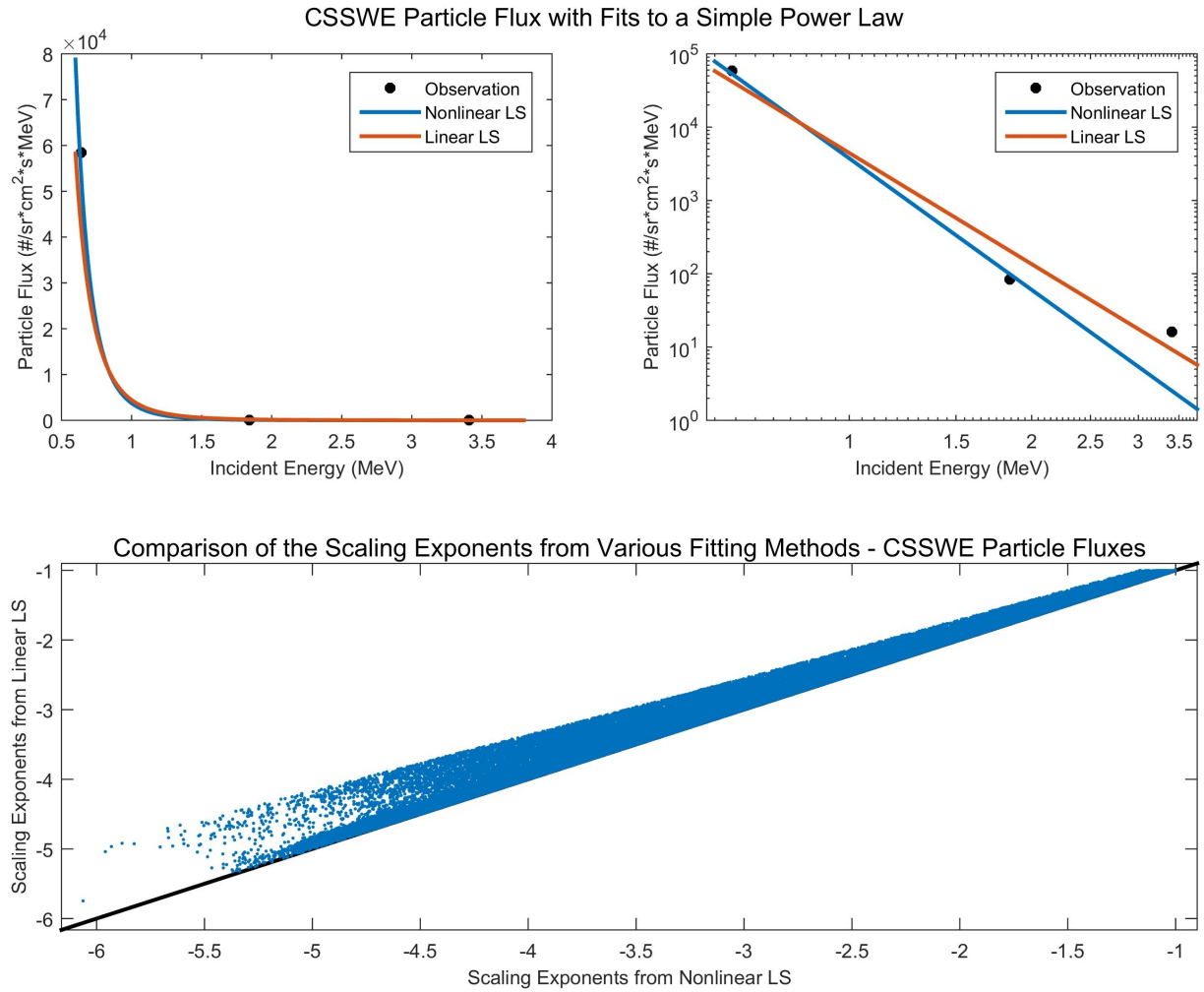


Figure 2.5: An arbitrary example spectrum, on linear and logarithmic scales, is shown in the top panels along with linear and nonlinear LS fits assuming the model $y = a_1 x^{a_2}$. The bottom plot compares the scaling exponents a_2 for about 144K such spectra using both linear and nonlinear LS methods. The use of nonlinear LS here is critical, as the vast majority of points do not lie on the identity line.

telescope as possible to avoid fitting to the noise. We then perform both linear and nonlinear LS to estimate the energy spectrum for each of the data points. We discard spectra with very poor fits ($r^2 < 0.5$) as well as non-physical fits of exponents larger than -1. The results are plotted in the bottom panel of Figure 2.5. The difference between the linear and the nonlinear least squares estimates can be as great as one. The nonlinear LS estimate is always equal to or greater than the linear LS estimate, due to a combination of instrumental noise and the additional weight on the 3.41 MeV observation. This deviation is exaggerated on the log-log scale, as we can see in the top right panel and, as a result, the linear LS algorithm returns a greater positive slope than does the nonlinear LS algorithm.

2.5 Cause of the Transformation Bias

The underlying cause of this bias can be explained as follows. Recall that the expectation operator returns the mean value of a random variable, and is a linear operator,

$$E(s_1X_1 + s_2X_2 + s_3) = s_1E(X_1) + s_2E(X_2) + s_3, \tag{2.19}$$

where s_1, s_2, and s_3 are arbitrary scalars and X_1 and X_2 are arbitrary random variables which are not necessarily independent or identically distributed. The expectation operator is, however, not multiplicative and it certainly does not possess functional equivariance, $E(g(X)) \neq g(E(X))$, in general. In fact, for a strictly concave function such as $g(x) = \ln(x)$, Jensen's inequality [*Jensen*, 1906] gives us,

$$E(g(X)) < g(E(X)), \tag{2.20}$$

so qualitatively we already see that the log transformation introduces negative bias.

Specifically, in the case of LS, if we assume that the mean error is zero, then the transformed errors will not have mean zero. It is this failure of the expectation operator to commute with nonlinear functions which is the cause of the transformation bias. The log transformation changes the error magnitudes and hence the objective function (sum of the squared errors) being minimized is also transformed. In fact, any regression technique where the objective function depends directly

on the errors is susceptible to transformation bias. For example, besides LS, both least absolute residuals (LAR) with the objective function $\sum_i |y_i - f(\mathbf{a}; x_i)|$ and minimax approximation with the objective function $\max_i\{y_i - f(\mathbf{a}; x_i)\}$ are vulnerable to transformation bias if the data is nonlinearly transformed before regression is applied.

2.6 Remedies

Before any statistical technique can be applied on a dataset, it must be verified that all of the assumptions necessary for the technique in question are satisfied. Otherwise, the technique is unreliable and may provide misleading results. The method of least squares is no exception. LS assumes that the error exists only in the dependent variable, with the error in the independent variables being either absent or negligible. LS additionally requires the additive error model with the mean error being zero, and no correlation between the errors. Error variance being constant and the error distribution being normal is not necessary for LS to provide unbiased estimates. But if the error distribution is normal, then LS is statistically efficient and provides the lowest mean squared error among the set of all unbiased estimators. Furthermore, the error distribution being normal with constant variance allows us to construct confidence intervals for the estimates. If the form of the fitting function $y = f(\mathbf{a}; x)$ is linear in all a_j then linear least squares is the method of choice. If the form of the fitting function is nonlinear in any of the a_j, then nonlinear least squares is more appropriate. In general, the data should not be nonlinearly transformed without a good reason as that can render the least squares method invalid by violating its assumptions.

In order to describe the remedies available if an assumption for ordinary least squares is violated, we first present the matrix formulation of a general linear least squares problem. Any linear least squares problem can be written as,

$$\mathbf{X}\beta = \mathbf{y}, \tag{2.21}$$

where the matrix $\mathbf{X}$ contains the independent values, the vector $\mathbf{y}$ contains the dependent values, and the vector β contains the unknown parameters we are trying to estimate. In least squares

problems, the matrix $\mathbf{X}$ is not a square matrix because we have more points than the number of parameters that we are trying to estimate. Therefore, the linear least squares solution is given by,

$$\beta = (\mathbf{X^T X})^{-1}\mathbf{X^T y}. \tag{2.22}$$

This is the solution to the linear least squares problem which always exists, is unique, and can be written down explicitly. We also note here that in practice the Gramian ($\mathbf{X^T X}$) is rarely inverted as it can be very computationally inefficient and expensive. Standard solvers are employed especially for large least squares problems. We now describe what remedies are available if an assumption for ordinary least squares is violated.

- If the error is present in both the independent and the dependent variables then total least squares is recommended. The total least squares method considers orthogonal errors instead of vertical errors in order to estimate the best fit curve.

- If the mean error in the data is nonzero, a systematic error is usually the cause. In this case the source should be investigated and the systematic error removed.

- If the errors are correlated, then the errors should be decorrelated. This is done by first constructing the conditional variance matrix $\mathbf{V}$. Then the least squares system is multiplied by $\mathbf{V^{-1}}$ after which we proceed as above and obtain the solution,

  $$\beta = (\mathbf{X^T V^{-1} X})^{-1}\mathbf{X^T V^{-1} y}. \tag{2.23}$$

  This method is called generalized least squares and has the disadvantage that the error variances and covariances must already be known.

- If, however, correlation is suspected but unknown then feasible generalized least squares is recommended. This proceeds in two steps. First, we perform ordinary least squares and estimate the error variances and covariances. Then we construct $\mathbf{V}$ and proceed as above.

- If the errors are uncorrelated but the variance is not a constant, then weighted least squares is the method of choice. The weight matrix $\mathbf{W}$ should be constructed as a diagonal matrix

where the diagonal entries w_{ii} are the weights to be given to the i-th observation, and all off-diagonal entries are zero. The least squares system can then be multiplied by $\mathbf{W}$ and we can solve for β as above to obtain,

$$\beta = (\mathbf{X^T W X})^{-1}\mathbf{X^T W y}. \tag{2.24}$$

If the variances can be estimated, then $w_{ii} = 1/\sigma_i^2$ is the best choice of weights, as stated by the Gauss-Markov theorem. If the variances are unknown, then perhaps a knowledge of the underlying processes or instrument design can lead to an appropriate weighting scheme. In addition to normalizing by the variances, weighted least squares can also be used if different observations are to be given different weights, for example, in the case of varying accuracy. Accuracy can degrade over time, or depend on the magnitude of the observations. Observations that are more accurate than the rest should be given more weight. The weights specified in weighted least squares are relative, so the absolute magnitudes of the weights are unimportant. It is only the ratio of the weights among different observations that matters. LS estimates are also rather insensitive to small changes in weights. The LS estimates will not change noticeably unless the weights change by orders of magnitude. Therefore a ballpark estimate of the weights suffices.

If nothing can be assumed about the variances, then usually for data spanning orders of magnitude we assume that the variance is proportional to the magnitude of the observation simply because smaller measurements tend to have smaller variances and larger measurements tend to have larger variances. Therefore, if a set of observations has a range of several orders of magnitude and we know nothing about the change in accuracy of the observations, then weighted least squares should used with $w_{ii} \propto 1/y_i^2$. Something as simple as $w_{ii} = 1/y_i^2$ is entirely appropriate in view of the Gauss-Markov theorem. Note that this scheme is equivalent to assuming $\sigma_i = k \cdot y_i$ where k is some small fraction such as $k = 0.1$. This gives us $w_{ii} = 1/\sigma_i^2 = 1/k^2 y_i^2$, but since the weights are only relative, the constant $1/k^2$ makes no difference. We should also warn the reader here that before the above prescribed weighting scheme is applied to the observed data, the data

should be inspected for measurements close to the noise floor of the instrument in question. If some measurements are just pure noise with a very small magnitude then the corresponding weights will be very large in magnitude, and will severely impact the LS estimates.

Note that all of the points mentioned above apply to both linear and nonlinear least squares problems. *Draper and Smith* [1966]; *Strutz* [2011]; *Rao* [1973]; *Rao and Toutenburg* [1999]; *Hastie et al.* [2009] are excellent references and discuss all of these points in detail. We would also like to point out here that nonlinearly transforming the data and then applying LS is not always wrong. If the regression model and the error distribution are such that the log-transformed data satisfies all of the assumptions of the least squares method, then LS is a perfectly valid option in log-log space. For example, if the errors are multiplicative instead of additive and if the errors have a distribution such that the transformed errors have mean zero, then LS will work on the transformed data. Depending on the distribution of the transformed errors LS may not be the most efficient method but it will provide unbiased estimates. Figure 2.6 provides a flowchart which can help the reader determine the correct least squares approach depending on which assumptions are satisfied by the data and which are violated.

2.7 Theil-Sen Method

Least squares is just one of many regression methods available. Its prevalence is due to its historical precedence and its ease of understanding and implementation in the pre-computer age. It was one of the first regression methods discovered [*Gauss*, 1823; *Legendre*, 1805] and applied systematically to physical problems. If, for a given dataset, LS is inappropriate or too complicated to be applied correctly, then there are other methods available today which can be used [*Rao and Toutenburg*, 1999; *Rousseeuw and Leroy*, 2003; *Hastie et al.*, 2009]. We present one such method, based on the median instead of the mean, which avoids all of the problems with LS, is easy to implement, and can be fully generalized to any number of variables in all linear and a few nonlinear cases.

The Theil-Sen (TS) method of estimating the linear trend for a given dataset is a very robust

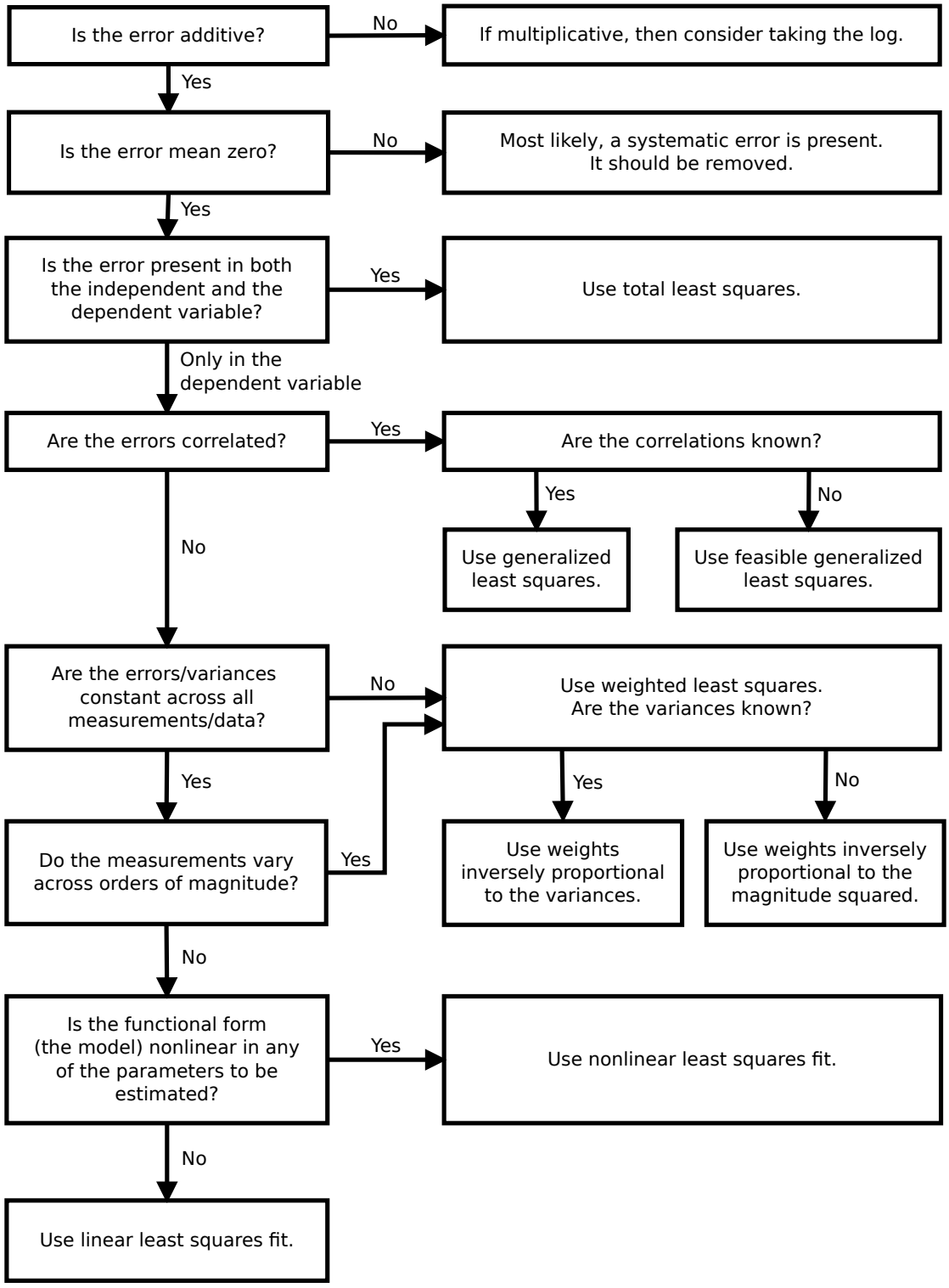


Figure 2.6: This flowchart describes the procedure that can be used to decide the correct least squares algorithm to apply to a given dataset.

and nonparametric method which assumes nothing about the distribution of errors besides the fact that the error probability distribution function be continuous [*Theil*, 1950; *Sen*, 1968]. The original method presented by *Theil* [1950] estimates only the slope of the linear trend in a two-dimensional dataset y_i vs. x_i. *Sen* [1968] later extended the method to include repeated observations. It can be shown that this simple TS method has its breakdown point at $1 - \frac{1}{\sqrt{2}} \approx 29\%$, meaning that up to 29% of the data can be arbitrarily corrupted before the TS estimator is affected [*Dietz*, 1987, 1989]. Compare this to the LS method, where a single outlier can skew the results [*Wilcox*, 2004]. To estimate the slope of the linear trend in a given dataset using TS, one simply computes all of the slopes between all possible pairs of N points and then takes their median as the estimator of the slope, i.e.,

$$\hat{m} = \text{Med}\left(\left\{\frac{y_i - y_j}{x_i - x_j} : 1 \leq i, j \leq N\right\}\right). \tag{2.25}$$

After $\hat{m}$ has been calculated, *Sen* [1968] proposed to take the median of all of the intercepts obtained by drawing a straight line through each of the N sample points with slope $\hat{m}$ to estimate the y-intercept of the best fit line, i.e.,

$$\hat{b} = \text{Med}\left(\{y_i - \hat{m}x_i : 1 \leq i \leq N\}\right). \tag{2.26}$$

Since this method is based on the median, any one-to-one monotonic transformation on the data, such as the log transformation, has no effect on these estimates [*Hald*, 1952]. So, for example, if the form of the fitting function is a power law ($y = a_1 x^{a_2}$), the data can be transformed into log-log space, and the TS estimates can be obtained and transformed back into linear space without any transformation bias,

$$\hat{a_2} = \hat{m}, \tag{2.27}$$

$$\hat{a_1} = \text{Med}(\{e^{b_i}\}) = e^{\text{Med}(\{b_i\})} = e^{\hat{b}}, \tag{2.28}$$

because of the functional equivariance of the median. In cases of skewed or heteroscedastic data, this method can be much more accurate than the LS method, and in cases where both LS and TS work, the TS method works just as well as the LS method [*Wilcox*, 1998a,b, 2001].

2.8 The Multivariate Median

Before the Theil-Sen method can be generalized to more than two parameters, we must generalize the notion of the univariate median to the multivariate median. Given a set of real numbers x_i, the median can simply be found by ordering the data and then selecting the value at the 50th percentile. The median is one of many measures of central tendency, along with the arithmetic mean and the mode. The mean is always unique while the median may not be. The mean is also very sensitive to outliers and has its breakdown point at 0%. The median in contrast is a very robust statistic with its breakdown point being at 50%. This means that up to half of our univariate data can be outlying before the median will be affected. In fact, the 50% breakdown is the maximum breakdown possible because if more than half of our observed data is considered outlying or corrupted then our underlying model, the assumptions we make, or the data collection methods themselves should be reconsidered [*Rousseeuw and Leroy*, 2003]. The univariate median achieves this maximum possible breakdown point.

On the other hand, it is very easy to generalize the notion of arithmetic mean from univariate to multivariate while it has proven difficult to generalize the median to higher number of dimensions. This is because we cannot order order the data in any meaningful way in dimensions higher than one. Given a set of data points with each point $\mathbf{x_i} = (x_i^1, x_i^2, ..., x_i^n)$ having n coordinates, the arithmetic mean can be found by simply taking the mean of each coordinate $\overline{\mathbf{x}} = (\overline{x^1}, \overline{x^2}, ..., \overline{x^n})$. The coordinate-wise multivariate mean preserves all of the properties of the univariate mean and in fact, it has been proven that the arithmetic mean is the only measure of central tendency with this property [*Donoho*, 1982]. However, the coordinate-wise median for $n > 2$, may not even be contained in the convex hull of the dataset in question. For a simple example, consider the three data points $\{(1,0,0),(0,1,0),(0,0,1)\}$ in $\mathbb{R}^3$. The coordinate-wise median of these three points is $(0,0,0)$ which is not even in the plane containing the three points.

There have been many attempts to generalize the univariate median to a multivariate median by selecting what may be considered by some the most important or desirable property of the

univariate median and then extending that notion to higher dimensions. One of the properties of the univariate median is that it minimizes the sum of the Euclidean distances to each data point x_i. Minimizing the quantity $\sum_i |y - x_i|$ results in $y = \text{Med}(\{x_i\})$. Extending this to higher dimensions, $\mathbf{y}$ which minimizes the sum of all straight line distances from $\mathbf{y}$ to each $\mathbf{x_i}$, i.e. the quantity $\sum_i \|\mathbf{y} - \mathbf{x_i}\|_2$, is called the geometric median (GM) [*Weber*, 1909]. In the case of three points in $\mathbb{R}^2$, the geometric median is also called the Fermat point or the Torricelli point. Another way to interpret the straight line distance in $\mathbb{R}$ is to consider the straight line between any two points in $\mathbb{R}$ to be the one-dimensional simplex with the length of the line being its one-dimensional volume. The two-dimensional simplex would be a triangle with its two-dimensional volume being its area while a three-dimensional simplex is a tetrahedron. An n-dimensional simplex in $\mathbb{R}^n$ requires $n+1$ points. Given $\{\mathbf{x_i}\}$ in $\mathbb{R}^n$, the point $\mathbf{y}$ which minimizes the sum of the n-dimensional volume of all of the n-dimensional simplices with $\mathbf{y}$ being one of the vertices is called Oja's simplex median [*Oja*, 1983]. If we consider the univariate median to be the number that minimizes the maximum number of data points on either side, it leads us to the Tukey or the halfspace median [*Tukey*, 1975; *Donoho and Gasko*, 1992]. If we instead consider the univariate median as the quantity remaining after we repeatedly discard two outermost points until we are left with one or two centermost points, we obtain the multivariate analog by computing and deleting the convex hull of our data until we have sufficiently few points left "in the middle" and then computing their average. This is called the convex hull peeled median [*Shamos*, 1976; *Barnett*, 1976]. Another property of the univariate median is that it is contained in the greatest number of intervals formed by $\{x_i\}$. In higher dimensions we can find the point $\mathbf{y}$, which is contained in the greatest number of simplices constructed from $\mathbf{x_i}$. This is called the simplical median [*Liu*, 1990]. There are many more generalizations; for the interested reader *Small* [1990] and *Aloupis* [2006] are good starting points.

Different notions of multivariate medians have various advantages, disadvantages, and variations in computational complexity, breakdown point, uniqueness, difficulty of implementation in practice, and equivariance under rotations, scaling, and affine transformations of the data. None

of the multivariate medians mentioned above has a closed-form solution and most require intensive computations for even a moderately sized dataset. Furthermore, each of these definitions usually gives a slightly different estimate of central tendency from all of the others. The two simplest and easiest to compute multivariate medians are the coordinate-wise median and the geometric median. The coordinate-wise median is trivial to implement; we simply take the median of each coordinate. The geometric median also has a very simple iterative algorithm called Weiszfeld's algorithm [*Weiszfeld*, 1937], which can be easily implemented and converges quite rapidly. For regression purposes, both the coordinate-wise median and the geometric median appear to do quite well. For space physics applications, however, generalized Theil-Sen (GTS) combined with the geometric median performs moderately better then does the coordinate-wise median. We leave the decision up to the reader to choose between the two multivariate medians. In some cases the geometric median provides better estimates at a reasonable cost of implementation and some extra computation.

2.9 Generalized Theil-Sen Method

There are several ways in which the Theil-Sen method can be generalized to more than two parameters [*Wilcox*, 2004]. *Pegoraro* [1992] presented a straightforward generalization to multiple parameters in the linear case, which we summarize here. Suppose that there are N data points given and our model $y = f(\mathbf{a}; x)$ has n parameters which need to be estimated. There are $K = \binom{N}{n}$ subsets possible with n data points in each subset. In the case of linear regression, for every one of the K subsets, we form an $n \times n$ linear system with n equations in n unknowns $(a_1, a_2, ..., a_n)$ and then solve for each of the a_js. After solving K linear systems, we obtain K values for each of the a_js and we simply take the multivariate median of all K values of a_j as an estimate of the true value of a_j.

As an example, let us be given ten data points $\{(x_1, y_1), ..., (x_{10}, y_{10})\}$ with the fitting function being a quadratic function $y = a_1x^2 + a_2x + a_3$ in which case $N = 10$, $n = 3$, and $K = \binom{10}{3} = 120$.

For each of the 120 combinations possible, we form and solve a linear system,

$$y_i = a_1 x_i^2 + a_2 x_i + a_3, \tag{2.29}$$

$$y_j = a_1 x_j^2 + a_2 x_j + a_3, \tag{2.30}$$

$$y_k = a_1 x_k^2 + a_2 x_k + a_3, \tag{2.31}$$

for $1 \leq i, j, k \leq N$, with three equations in three unknowns. After repeating this process for each of the $K = 120$ combinations, we have 120 estimates for each of the a_1, a_2, and a_3. Our final estimate will be

$$\hat{\mathbf{a}} = (\text{Med}(\{\hat{a_1}\}), \text{Med}(\{\hat{a_2}\}), \text{Med}(\{\hat{a_3}\})), \tag{2.32}$$

if we wish to use the coordinate-wise median, or

$$\hat{\mathbf{a}} = \text{Geometric Median}(\{\hat{a_1}, \hat{a_2}, \hat{a_3}\}), \tag{2.33}$$

if we wish to use the geometric median.

Note that the method described works only for linear regression problems. This is because given any n unique points, we can always find values for $\mathbf{a}$ such that the curve $y = f(\mathbf{a}; x)$ goes through all n of the points. In our example above, given any three distinct points, it is always possible to find a parabola that passes through all three of those points. On the other hand, if we were trying to fit our data to the functional form $y = a_1 x^{a_2} + a_3$, there is absolutely no guarantee that given any three arbitrary points we can find a_1, a_2, a_3 such that a curve of the form $y = a_1 x^{a_2} + a_3$ goes through all three points. However, since this generalization of Theil-Sen is not susceptible to transformation bias, we can use it for some specific forms of nonlinear regression problems. Any nonlinear regression problem which can be linearized after a suitable transformation can be solved by this method. The idea is to transform the model into a linear form and to transform the data appropriately, and then apply this method. Reverse transform the parameter estimates, and then compute their multivariate median. The order of the last two steps is crucial. The estimates must be reverse transformed first and then their median computed. Since the coordinate-wise median is equivariant only under one-to-one monotonic transformations and the geometric median is not

equivariant at all, for the sake of simplicity and implementation it is much easier to remember always to reverse transform and then compute the median; this will work in all cases.

For example, for a given a dataset we wish to find the best fit Lorentzian curve $y = \frac{a_1}{a_2+(x-a_3)^2}$. We can change this nonlinear regression problem into a linear regression problem by considering the multiplicative inverse $\frac{1}{y} = \frac{a_2+(x-a_3)^2}{a1}$, which expands into a parabola, and so we use generalized Theil-Sen to fit $(x_i, \frac{1}{y_i})$ to the parabola $\frac{1}{y} = \alpha_1 x^2 + \alpha_2 x + \alpha_3$. After all of the estimates of α_1, α_2, and α_3 are obtained, they are reverse transformed using

$$a_1 = \frac{1}{\alpha_1}, \tag{2.34}$$

$$a_2 = \frac{\alpha_3}{\alpha_1} - \left(\frac{\alpha_2}{2\alpha_1}\right)^2, \tag{2.35}$$

$$a_3 = -\frac{\alpha_2}{2\alpha_1}. \tag{2.36}$$

Then we find the estimates $\hat{a_1}$, $\hat{a_2}$, and $\hat{a_3}$ using the coordinate-wise median or the geometric median as desired.

It is well known that the number of all possible sets of a certain size increases very rapidly as the total number of data points increases. Given 500 data points, if we wish to obtain the best fit cubic polynomial with four unknown parameters we would have to solve $K = \binom{500}{4} = \frac{500!}{4!\cdot 496!} = 2,573,031,125$ linear 4×4 systems. Such computational intractability has been another reason why Theil-Sen type methods have not been widely used. While it is computationally feasible today to solve millions of small systems, the general recommendation is to somehow make the problem size smaller [*Wilcox*, 1998a,b, 2001]. *Wilcox* [2004] recommends that out of K subsets, TS should be applied only to a group of randomly chosen subsets. *Pegoraro* [1992] presents another method where instead of considering a problem of size N, it might be preferable to consider p problems of size N/p each where p is a positive integer. This is because $p \cdot \binom{N/p}{n}$ decreases very rapidly as p increases. The parameter p is restricted to $1 \leq p \leq N/n$ because N/p is restricted to integral values with $n \leq N/p \leq N$. We point out that as attractive as it may seem, we should not pick the largest value of p that is possible but rather the smallest value of p possible. The number of subsets reduces rapidly with increasing p and our estimates of a_j may suffer from small sample

size effects if p is too large. After an appropriate value of p has been picked, we should divide our dataset into p roughly equal subsets and then apply generalized Theil-Sen to each of the p subsets individually. Then all of the estimates for each a_j are combined into a single list and the median of the combined list is selected as an estimate of a_j.

Suppose that we are given a dataset D with 261 data points and we wish to find the quadratic curve which best represents D. We decide that we are willing to solve at maximum only $100,000$ systems trading off between a large sample size and computational effort. $K = \binom{261}{3} = 2,929,290$ is too large and we note that $p = 6$ is the smallest p such that $p \cdot \binom{261/p}{3} < 100,000$. We divide D into six subsets as follows,

$$\begin{aligned}
D_1 &= \{(x_1, y_1), (x_7, y_7), ..., (x_{253}, y_{253}), (x_{259}, y_{259})\}, \\
D_2 &= \{(x_2, y_2), (x_8, y_8), ..., (x_{254}, y_{254}), (x_{260}, y_{260})\}, \\
D_3 &= \{(x_3, y_3), (x_9, y_9), ..., (x_{255}, y_{255}), (x_{261}, y_{261})\}, \\
D_4 &= \{(x_4, y_4), (x_{10}, y_{10}), ..., (x_{256}, y_{256})\}, \\
D_5 &= \{(x_5, y_5), (x_{11}, y_{11}), ..., (x_{257}, y_{257})\}, \\
D_6 &= \{(x_6, y_6), (x_{12}, y_{12}), ..., (x_{258}, y_{258})\},
\end{aligned}$$

and indeed we see that $3\binom{44}{3} + 3\binom{43}{3} = 76,755 < 100,000$. Now we perform generalized Theil-Sen on each of the $D_1, ...D_6$ individually. Then for each a_j, we merge all of the estimates of a_j into a single list and then compute the median of the combined list as a point estimate for a_j. Note that despite reducing the problem to 2.6% of its original size, this method still utilizes each and every data point, and is deterministic instead of being randomized. Before dividing D into its subsets, D should be ordered with respect to the independent variable because the interlacing scheme is used as a mixing mechanism. Otherwise D can be divided in any manner.

Generalized Theil-Sen is immune to transformation bias because the method does not fit to errors. The objective function being minimized is not a function of the measurement errors. In the Lorentzian example above, it doesn't matter what the measurement error itself is or how the errors are transformed. We simply wish to solve many small problems of finding a Lorentzian

curve which intersects any three given points, and that problem can be easily solved by considering the multiplicative inverse of y_i. In addition, this method has the additional desirable property of robustness. The Generalized Theil-Sen method has a positive breakdown point, while LS has a 0% breakdown point so that a single outlier will influence the results. As mentioned above, for the original Theil-Sen method the breakdown point is $1 - \frac{1}{\sqrt{2}} \approx 29\%$. The general formula for the breakdown point for Theil-Sen type methods can easily be shown to be $1 - b^{1/n}$, where b is the breakdown point of the median being used and n is the number of parameters that are being estimated. The original Theil-Sen estimates the linear trend ($y = a_1 x + a_2$) in $\mathbb{R}^2$ with two parameters to be estimated ($n = 2$) and utilizes the univariate median with $b = 1/2$ giving us $1 - \frac{1}{\sqrt{2}} \approx 29\%$. Both the coordinate-wise median and the geometric median have their breakdown points at 50% and hence the generalized Theil-Sen breaks down at $1 - \frac{1}{2^{1/n}}$, which decreases with increasing n but is nonetheless positive and advantageous compared with least squares and its 0% breakdown point. This allows TS type methods to be immune to a certain number of outliers in the data. However, if the main goal is robust regression, there are other, better methods. Recall that least squares minimizes the sum of squared residuals $\sum r_i^2$ which is equivalent to minimizing the mean squared error $\frac{1}{N}\sum r_i^2$. Other regression methods which are more robust to outliers include the least absolute residuals (LAR) method minimizing $\sum |r_i|$ and least median squared (LMS) minimizing $\text{Med}(\{r_i^2\})$. Bisquare weights is a weighted least squares method in which ordinary least squares is first performed and each point is assigned a weight inversely proportional to its distance from the LS curve. Then weighted least squares is performed and its estimates are returned. Least trimmed squares is a regression method where instead of summing all squared residuals, the residuals are squared and ordered, and then the lower half of the squared residuals from the zeroth percentile to the 50th percentile are summed. This method by definition has a 50% breakdown point because the points with large residuals are ignored and have no effect on the estimates. However, all of these methods suffer from transformation bias because, unlike the Theil-Sen methods, the objective functions being minimized all depend on the observation error. The interested reader is referred to *Rousseeuw and Leroy* [2003] and *Hastie et al.* [2009] for further reading if interested in

robust methods of regression.

We caution the reader here that when using generalized Theil-Sen to solve even small linear systems (i.e., $n = 2, 3$), stable numerical methods such as Gaussian elimination (LU factorization) should be used [*Press et al.*, 2007]. Often in regression problems, as for example in the case of a power law model, it is possible for two measured values to be extremely close to one another or to be orders of magnitude apart. Numerical instabilities can cause too many estimates (higher than the breakdown point) to be corrupt, and hence this method can fail. All of the scientific computing environments used today have excellent linear solvers which are both fast and numerically stable. Environments such as IDL provide many functions from which the user can select an appropriate solver if something is known in advance about the structure of the coefficient matrix. Other environments such as MATLAB provide functions such as `mldivide` which can automatically select an appropriate solver depending on the structure of the linear system.

2.10 Conclusion

This practice of log-transforming the data and fitting a straight line using linear least squares dates from the days before the computer age, when fitting a straight line by hand was preferable to solving a nonlinear system iteratively. In this age of modern computers with GHz processors, there is no need for anyone to avoid nonlinear least squares fitting. All of the common data processing and statistical packages, including MATLAB, IDL, R, and SciPy provide routines for nonlinear least squares which are fast, efficient, and easy to use. Some of the more commonly used functional forms such as the power law and the Gaussian function have their own dedicated routines which include algorithms to optimally select the starting point, the iterative algorithm to use, and the convergence criteria. The generic algorithms allow for custom equations so that any functional form can be used. The gradient-based methods can even numerically estimate the gradient and the Hessian so that the user doesn't have to provide those if the function is complicated. There are also various gradient-free methods such as the simplex methods, which avoid all gradient-related problems. *Press et al.* [2007] and *Kelley* [1999] provide an excellent compendium of such methods

with detailed discussions. If one wishes to perform a least squares fit on a dataset, then it should be performed on the original data only, with some linear transformations allowed if necessary [*Draper and Smith*, 1966; *Hastie et al.*, 2009; *Strutz*, 2011]. Nonlinear transformations (such as the log transformation) can render the least squares method invalid and return biased results unless the error distribution is such that the transformed data satisfies all of the assumptions necessary for the least squares method.

If the proper application of the least squares method is difficult or entirely inappropriate, there are other regression techniques which can be used [*Rao*, 1973; *Rao and Toutenburg*, 1999; *Rousseeuw and Leroy*, 2003; *Hastie et al.*, 2009]. We presented one such technique here Theil-Sen [*Theil*, 1950; *Sen*, 1968], which is nonparametric and based on the median instead of on the mean. It is a robust technique [*Dietz*, 1987, 1989] which can be applied in full generality to all linear and some nonlinear regression problems [*Pegoraro*, 1992]. It addresses all of the shortcomings of the least squares method [*Wilcox*, 1998a,b, 2001] and is straightforward to implement. Depending on the size of the dataset, it may require more computational effort than least squares, but *Pegoraro* [1992] presents a method, which we expand to some nonlinear cases, to reduce the computational effort significantly. Compared with the mean, the median possesses equivariance under a much broader class of functions[*Hald*, 1952]. This provides Theil-Sen type methods additional advantages which least squares methods lack, and presents Theil-Sen type methods as an attractive alternative to least squares.

Chapter 3

Combined Release and Radiation Effects Satellite

The purpose of this chapter is to use the magnetic field data from the Combined Release and Radiation Effects Satellite (CRRES) to estimate the magnetic component of the radial diffusion coefficient (D_{LL}^{B}) in a manner analogous to the study conducted by *Brautigam et al.* [2005], which used the electric field data from CRRES to compute the electric component of the radial diffusion coefficient (D_{LL}^{E}). Before the Van Allen Probes, CRRES had been the last mission dedicated to the study of the Van Allen radiation belts, but despite a gap of two decades between the two missions the CRRES magnetometer data had not been utilized by the scientific community in a study of D_{LL}^{B}. This research project was initiated before the launch of the Van Allen Probes and consisted of studying the distribution of ULF wave power in azimuth in the inner magnetosphere as well as computing D_{LL}^{B} so that we can quantify the relative importance of the magnetic and the electric components in driving radial diffusion of charged particles. Results from this chapter have been published in *Ali et al.* [2015].

3.1 Mission Details and Instruments

The Combined Release and Radiation Effects Satellite (CRRES) was launched into a geosynchronous transfer orbit on July 25, 1990. The orbital period was about 9.4 hours, with a perigee of 350 km, an apogee of 36,000 km, and an inclination of 18° in geographic latitude. The satellite orbit was designed so that the local time at apogee decreased by 2.5 minutes per day starting at 0800 MLT just after launch, and would then return to this position in nineteen months. This was

done so that the magnetosphere could be studied at different local times throughout the mission. However, after only about fourteen months contact with CRRES was lost, on October 12, 1991, presumably due to on-board battery failure [*Giles et al.*, 1995]. For the duration of the mission, the apogee precessed from 0800 MLT through midnight to around 1330 MLT [*Giles et al.*, 1995].

After the initial spin-up period, the satellite spin period stabilized at 2.2 rpm with the spacecraft spin axis aligned within 12° of the Earth-Sun line. CRRES had instruments to measure both electric and magnetic fields. The electric field instrument (EFI) [*Wygant et al.*, 1992] used both a pair of spherical and a pair of cylindrical probes to take electric field measurements and was orthogonal to the CRRES spin axis, sampling at a rate of 32 Hz. The signal was filtered at 10 Hz, providing measurements of the two-dimensional electric field in the CRRES spin plane. This chapter focuses on the magnetic field readings provided by the fluxgate magnetometer [*Singer et al.*, 1992], which contained three orthogonal sensors located at the end of a 6.1 m boom and sampled the magnetic field vector eight times per second. The magnetometer operated in low-gain mode for $L < 3.5$ with a dynamic range of $\pm 45,000$ nT and a resolution of 22.0 nT. The high-gain mode was used for $L > 3.5$ with a dynamic range of ± 850 nT and a resolution of 0.4 nT, with the switching between the two modes being fully automated, triggered at 850 nT. The full-resolution signal was then spin fit to remove the variations resulting from the spin of the spacecraft. Since the CRRES spin axis was roughly parallel to the x-axis of the GSE coordinate system, the CRRES magnetometer data used in this study was converted from local spacecraft coordinates to a modified GSE (mGSE) coordinate system provided by the Virtual Magnetospheric Observatory (VMO) hosted by IGPP/UCLA with open access rights (http://vmo.igpp.ucla.edu/). A copy of the CRRES data utilized for this study has also been made available at the Virtual Radiation Belt Observatory (http://virbo.org/∼ali). In mGSE, the x-axis is the CRRES spin axis pointing towards the Sun, the y-axis intersects the ecliptic and the spin planes and points towards dusk, and the z-axis completes the coordinate system. These coordinates are very similar to GSE because the CRRES spin axis was always within 12° of the Earth-Sun line.

3.2 Data Preparation and Processing

We started with the spin fit 30-second resolution magnetometer data in the mGSE coordinate system separated into individual orbits. Due to routine spacecraft maneuvers, and uncertainties and errors introduced at various levels of data processing done on the CRRES magnetometer data, not all of the magnetometer data from the entire mission was used for this study. For example, an orbit in which attitude adjustment took place had to be discarded because of artifacts induced in the measurements due to boom oscillation. The apogee immediately after launch was at 0800 MLT and precessed through midnight to 1330 MLT. However, because of the initial spin-up period and the October 9, 1990 magnetic storm [*Brautigam and Albert*, 2000], the first orbit to be included in this study was orbit 190, with its apogee at around 0500 MLT, then continuing until the end of the mission as shown in Figure 3.1. The rest of the dataset was inspected visually and random errors and spikes in the measurements over small intervals (less than five minutes with the cadence of the data being 30 seconds) were interpolated so that we would have the maximum possible number of temporally continuous data segments for subsequent Fourier analysis. Orbits containing errors and spikes over a longer time interval as well as systematic errors of any kind were discarded. Furthermore, orbits with the apogee on the dayside with $B_z < 50$ nT were eliminated because of possible magnetopause crossings.

Since we are concerned with ULF wave power in the Pc-5 range, with frequencies on the order of a few mHz, the large background magnetic field combined with the large gradient of the magnetic field measurements as CRRES moved inbound and outbound from its perigee made it quite difficult to estimate and subtract the static background fields. PSD estimation is a smoothing process, so any attempt to estimate the PSD in the Pc-5 range resulted in severe leakage from the lower frequency channels as a result of these static background fields. Furthermore, high velocities at low L-shells as well as high quantization error in the fluxgate magnetometer's low-gain mode introduced further noise. For all of these reasons, data with $L < 3.5$ was completely ignored. We would like to point out here that for the rest of this chapter we use the McIlwain parameter L

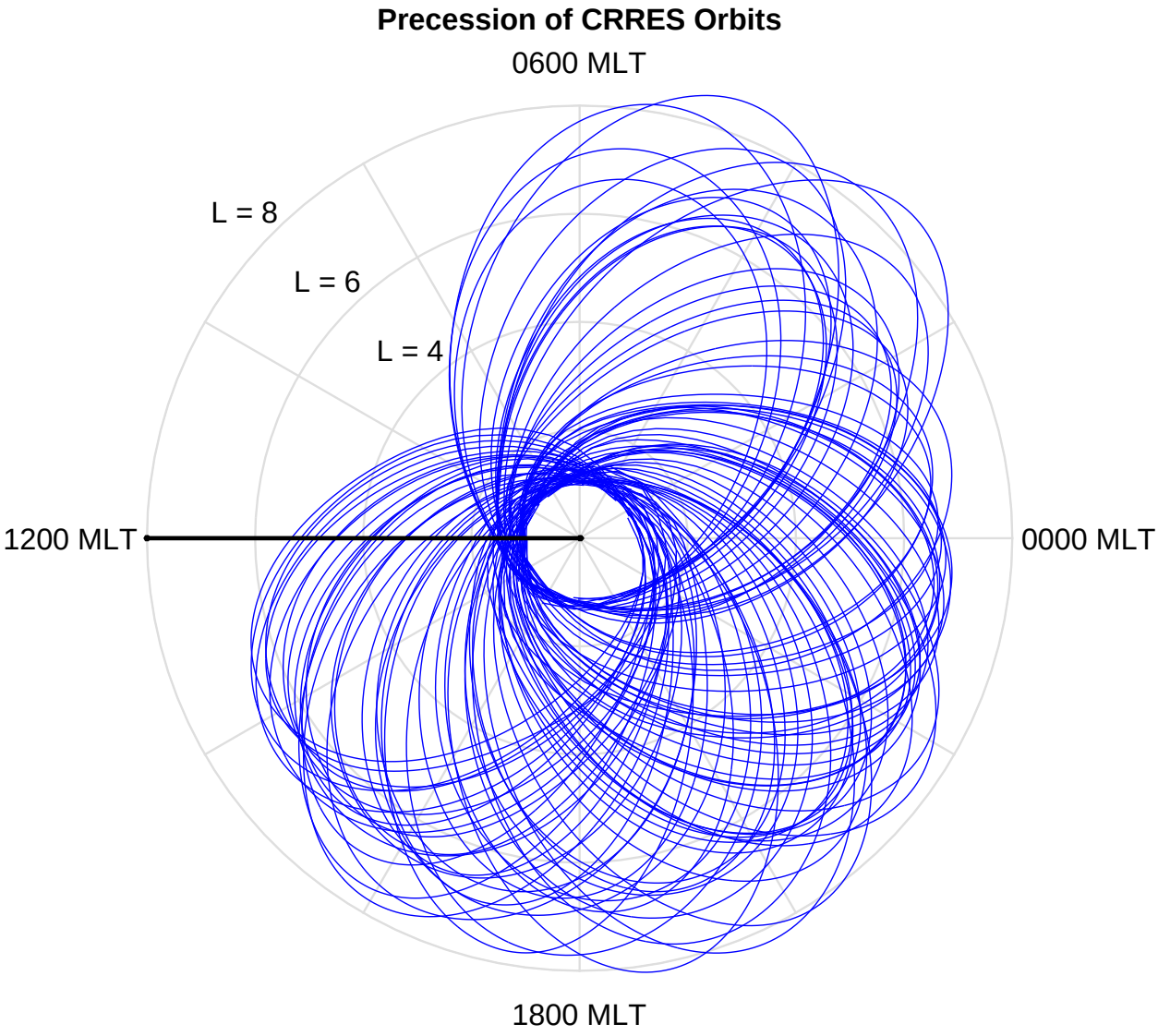


Figure 3.1: Precession of CRRES orbits. Only the orbits used for this study are shown, with the apogee starting at around 0500 MLT (orbit 190) and precessing through midnight until 1330 MLT (orbit 1062) when the mission ended.

[*McIlwain*, 1961], which may differ slightly from the Roederer L^* parameter [*Roederer and Zhang*, 2014]. In the end, we utilized approximately two-thirds of the entire CRRES magnetometer dataset for this study. Also note here that because CRRES apogee did not precess through all magnetic local times, as depicted in Figure 3.1, there were fewer measurements in the noon sector than at other magnetic local times, especially at higher L-shells.

Since the magnetic diffusion coefficients require power spectral density of the compressional component of the magnetic field vector $\mathbf{B}$ at frequency $m\omega_d$, we needed to compute the variations in $\mathbf{B}$ in the direction of $\mathbf{B}$ i.e.,

$$\mathbf{B}_{\text{Compressional}} = \Delta\mathbf{B} \cdot \frac{\mathbf{B}}{\|\mathbf{B}\|} = (\mathbf{B} - \langle\mathbf{B}\rangle) \cdot \frac{\mathbf{B}}{\|\mathbf{B}\|}, \tag{3.1}$$

where $\langle\mathbf{B}\rangle$ is the average background field which must be removed from $\mathbf{B}$ to obtain $\Delta\mathbf{B}$. To estimate $\langle\mathbf{B}\rangle$, we used a low-pass digital filter utilizing the Hann window with a cutoff frequency of 0.8 mHz [*Press et al.*, 2007]. This cutoff frequency was chosen because it is the highest cutoff frequency for which the power spectral density in the ULF range was not noticeably affected by the background fields.

Given a digital discrete signal $x(t)$ in the time domain, our goal was to estimate its power spectral density. It is well known that if the signal is of finite length N with the sampling period Δt in seconds, then using the discrete Fourier transform (usually computed using the FFT algorithm) to estimate the periodogram is not a good estimate of the spectrum. This is because the periodogram is a biased estimator, due to the sharp truncation of the signal, and the variance does not decrease even if the length of the signal is increased. However, if the signal is first multiplied by a windowing

function w_j (tapered),

$$X_k = \sum_{j=0}^{N-1} x_j w_j e^{2\pi i j k/N}, \quad k = 0, 1, ..., N-1 \tag{3.2}$$

$$P(f_0) = \frac{|X_0|^2}{W_{ss}}, \tag{3.3}$$

$$P(f_k) = \frac{|X_k|^2 + |X_{N-k}|^2}{W_{ss}}, \quad k = 1, 2, ..., \left(\frac{N}{2} - 1\right) \tag{3.4}$$

$$P(f_{N/2}) = \frac{|X_{N/2}|^2}{W_{ss}}, \tag{3.5}$$

$$W_{ss} = N \sum_{j=0}^{N-1} w_j^2, \tag{3.6}$$

$$f_k = \frac{k}{N\Delta t}, \quad k = 0, 1, ..., \frac{N}{2} \tag{3.7}$$

so that the resulting signal gradually goes to zero at the end points, then the spectral bias can be reduced. Furthermore, in order to reduce the variance of the spectrum, one method is to obtain several statistically independent estimates of the spectrum from the same signal and then average them [*Press et al.*, 2007].

For this study we used the multitaper method which eliminates the use of subsequences of $x(t)$ (usually overlapping) to reduce the variance of the periodogram [*Thomson*, 1982]. Instead of using a single taper, we used multiple tapers (window functions) which are all mutually orthogonal vectors, and which when applied to the entire signal provide statistically independent estimates of the spectrum. The spectra obtained were then averaged to give us a final spectral estimate. This allowed us to obtain a reduction in bias and variance without losing frequency resolution. The tapers used here were the Discrete Prolate Spheroidal Sequences (DPSS) also known as the Slepian sequences [*Slepian*, 1978]. The DPSS are given by the eigenvectors of a positive self-adjoint semi-definite symmetric tridiagonal matrix. Therefore all eigenvalues are real, non-negative, and in this case bounded above by one. In addition, the eigenvectors corresponding to distinct eigenvalues are mutually orthogonal. An interesting feature of the DPSS is that they can take negative values [*Slepian*, 1978; *Thomson*, 1982; *Press et al.*, 2007]. Since computing DPSS and then estimating the PSD can become quite intensive computationally, a reasonable number of the first few DPSS

must be chosen to provide a useful tradeoff between variance reduction and computation time. Furthermore, with higher order DPSS, since the tapers are not zero at the endpoints, leakage can occur. We used the first seven DPSS as data tapers. Additional tapers showed only an insignificant change in the PSD estimates.

We took 20-minute data segments and estimated their PSDs, which were then stored with L-shell value, Kp, and MLT associated with the data point at the center of each data segment. Because we used 30-second resolution data with data segment lengths of 20 minutes, the resolved frequency range was $\sim$0.8 mHz to $\sim$16.2 mHz with a frequency resolution of $\sim$0.8 mHz. Since we are mainly interested in the Pc-5 frequency range, the domain of the periodograms considered was truncated at $\sim$8.1 mHz for the duration of this chapter. With the sole exception of the results presented in Figure 3.3 we ignore the power at higher frequencies.

3.3 Power in the Magnetic Field - MLT Dependence

In order to empirically study the magnetic field power spectral density we must decide how to parametrize the magnetic field power. Following the approach taken by *Brautigam et al.* [2005] we decided to study the dependence of magnetic field power on the McIlwain parameter L and the level of geomagnetic activity Kp. In addition, we examined the power distribution in magnetic local time. Ideally, we would have considered other time-dependent parameters as well, such as the solar wind velocity, but for the duration of the CRRES mission, Interplanetary Monitoring Platform (IMP-8) was the only source of solar wind velocity data and its time resolution was too low to be of any use in this study.

To study the ULF wave power dependence on L, Kp, and MLT, we organized the computed PSDs into various bins. The L bins were centered at $L = 4, 4.5, 5, 5.5, 6, 6.5$. The width of each bin in L was $0.5\,R_E$. In Kp, we had four bins,

$$\begin{aligned} \text{Q} &= \text{Quiet Activity} &&= \{0, 0+, 1-, 1, 1+\}, \\ \text{L} &= \text{Low Activity} &&= \{2-, 2, 2+, 3-, 3, 3+\}, \\ \text{M} &= \text{Moderate Activity} &&= \{4-, 4, 4+, 5-, 5, 5+\}, \\ \text{H} &= \text{High Activity} &&= \{6-, 6, 6+, 7-, 7, 7+\}. \end{aligned}$$

In MLT, we divided up the local time into four equal bins of six hours each, centered at midnight, dawn, noon, and dusk. For example, the noon bin was centered at 1200 MLT with the bin spanning 0900 MLT - 1500 MLT. Figure 3.2 indicates the number of data segments in each radial, local time, and activity bin. We should point out here that out of all the bins being used for this study, the dawn sector with high Kp contains the smallest number of data segments. This is simply because after the October 1990 storm there wasn't much high Kp activity as the CRRES apogee precessed through the dawn sector.

Figure 3.3 shows all of the data segments in the bin $L = 6.5$, with high Kp activity in the noon sector. All 1,082 PSDs are plotted in the background to emphasize the variability that can exist within a single bin. The PSDs in this bin span three orders of magnitude. The arithmetic mean, geometric mean, and the median PSD are plotted along with the inter-quartile range. Here we clearly see that because of the outlying PSDs containing much more power than the rest of the dataset, the arithmetic mean is a severely biased measure of central tendency. For most of the frequencies the arithmetic mean is larger than 75% of the data. For data with such a large spread, the geometric mean or the median are much more appropriate measures of central tendency. We chose to use the median PSD from each bin as the representative PSD of that bin because of the ease of computation as well as the robustness of the median against outlying data.

Binning the data in MLT allows us to study the distribution of magnetic field power in azimuth. For this, we selected the median PSD from each bin and computed the integrated power of the median PSD over the truncated frequency range from 0 mHz to 8.1 mHz and then compared how the total magnetic field power depended on L, Kp, and magnetic local time. Figure 3.4 shows the total integrated power of the median PSD across all of the bins. The increase in power across all L-shells and MLT as the level of geomagnetic activity (Kp) increases is clearly shown, with

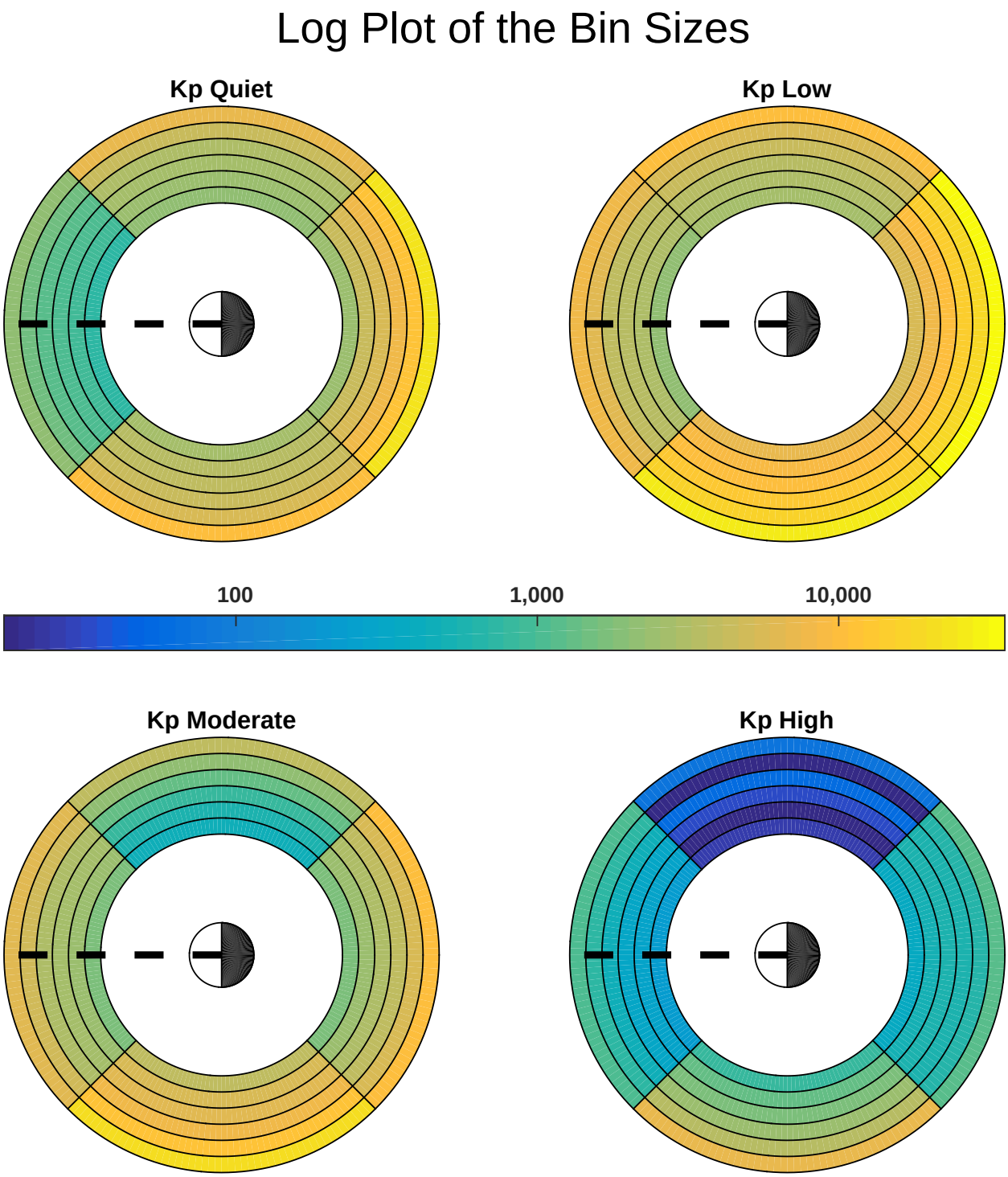


Figure 3.2: Log plots of the bin sizes separated by Kp with the innermost ring and the outermost ring corresponding to $L = 4.0$ and $L = 6.5$ respectively.

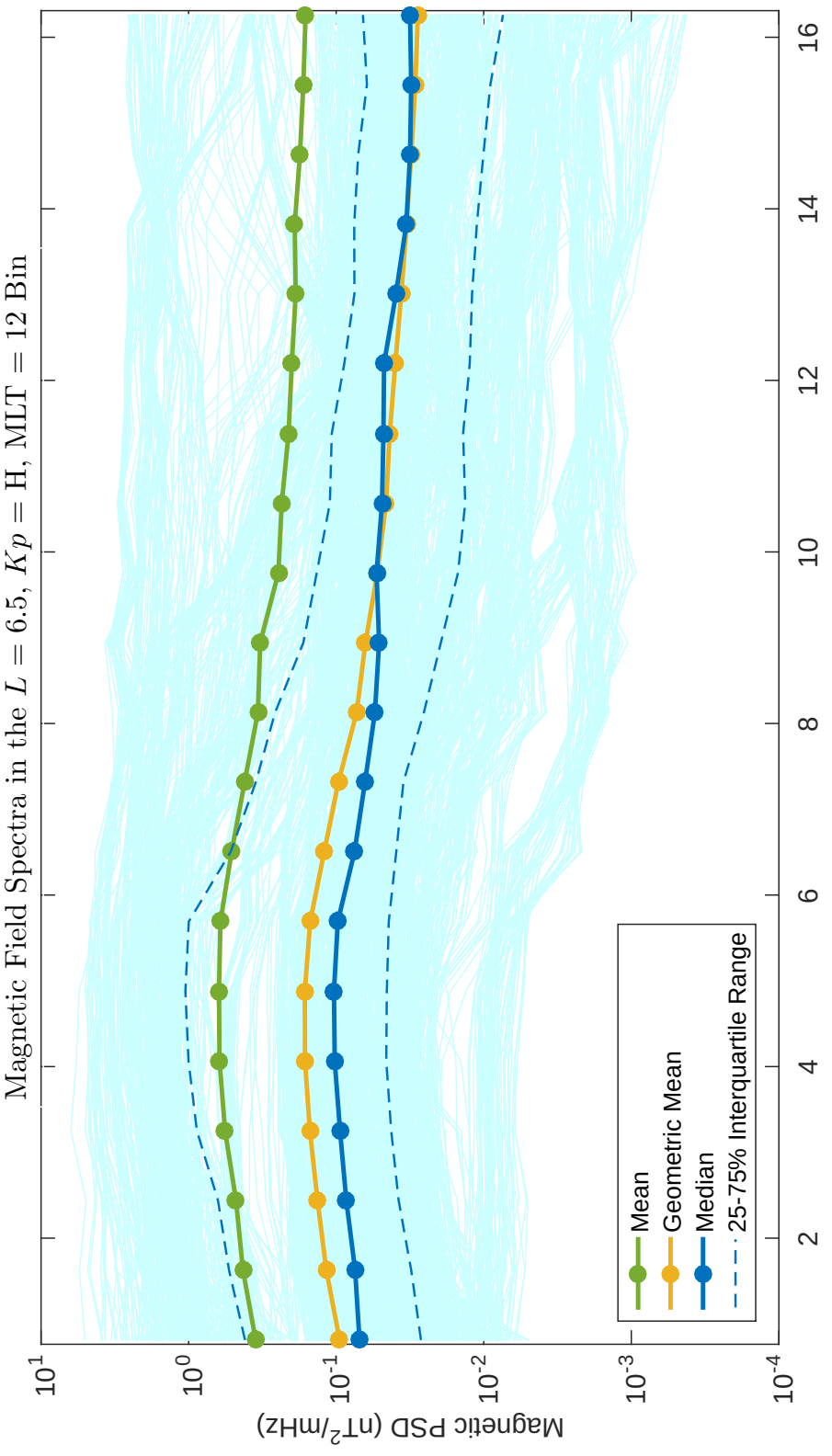


Figure 3.3: The bin $L = 6.5$, $Kp = $ H, MLT $= 1200$ with all of the magnetic wave power spectral densities in that bin on a log-log scale along with their mean, geometric mean, and the median with its interquartile range over the entire resolved frequency range.

quiet times having very low power and high activity exhibiting the greatest power. In azimuth, we generally see more power on the noon side than at other MLT, with no significant difference among the other MLT sectors. For quiet and low Kp we consistently see more power in the noon sector at all radial distances. However for moderate and high Kp, the noon sector is not necessarily dominant. In L, we generally see a decrease in power as L decreases. The only peculiarity to be addressed here is the unusually low power in the dawn sector at low L for high Kp. This may be an artifact of the small sample size in this bin, as indicated in Figure 3.2.

Figure 3.5 shows plots of the median magnetic field PSDs in the noon sector plotted on a log-log scale as a function of frequency. These plots indicate a very weak sublinear trend in frequency, along with a peak at around 5 mHz indicative of a field-line resonance. A field-line resonance occurs when the eigenfrequency of a geomagnetic field line is the same as the frequency of the exciting ULF wave, causing a resonant increase in wave amplitude [*Chen and Hasegawa*, 1974; *Southwood*, 1974; *Glassmeier et al.*, 1999]. For any fixed L, we see a clear Kp-dependence. In the Pc-5 ULF range, power is directly correlated to the level of geomagnetic activity with wave power being three orders of magnitude higher at high Kp activity than at quiet times. In addition, in the Pc-5 ULF range we see a very slight frequency dependence with power increasing slightly as a function of frequency. For higher frequencies, however, there is a very sharp drop off of wave power.

3.4 Drift-Averaged Magnetic PSDs

Although binning Pc-5 activity by local time provides information on the global distribution of wave power, in order to compute the magnetic radial diffusion coefficients we need to determine a drift-averaged magnetic power spectral density for each Kp and L bin over all MLT. Since the CRRES apogee did not precess all magnetic local times, there was a bias in our data as coverage in the noon sector was not as extensive as in other sectors. In order to compute a drift-averaged PSD, we took a weighted arithmetic mean of the median PSDs normalized by CRRES dwell time in each MLT sector. For example, for Kp quiet, for any fixed L, the noon sector has the fewest number of PSDs and was therefore the most heavily weighted sector. As shown in Figure 3.6,

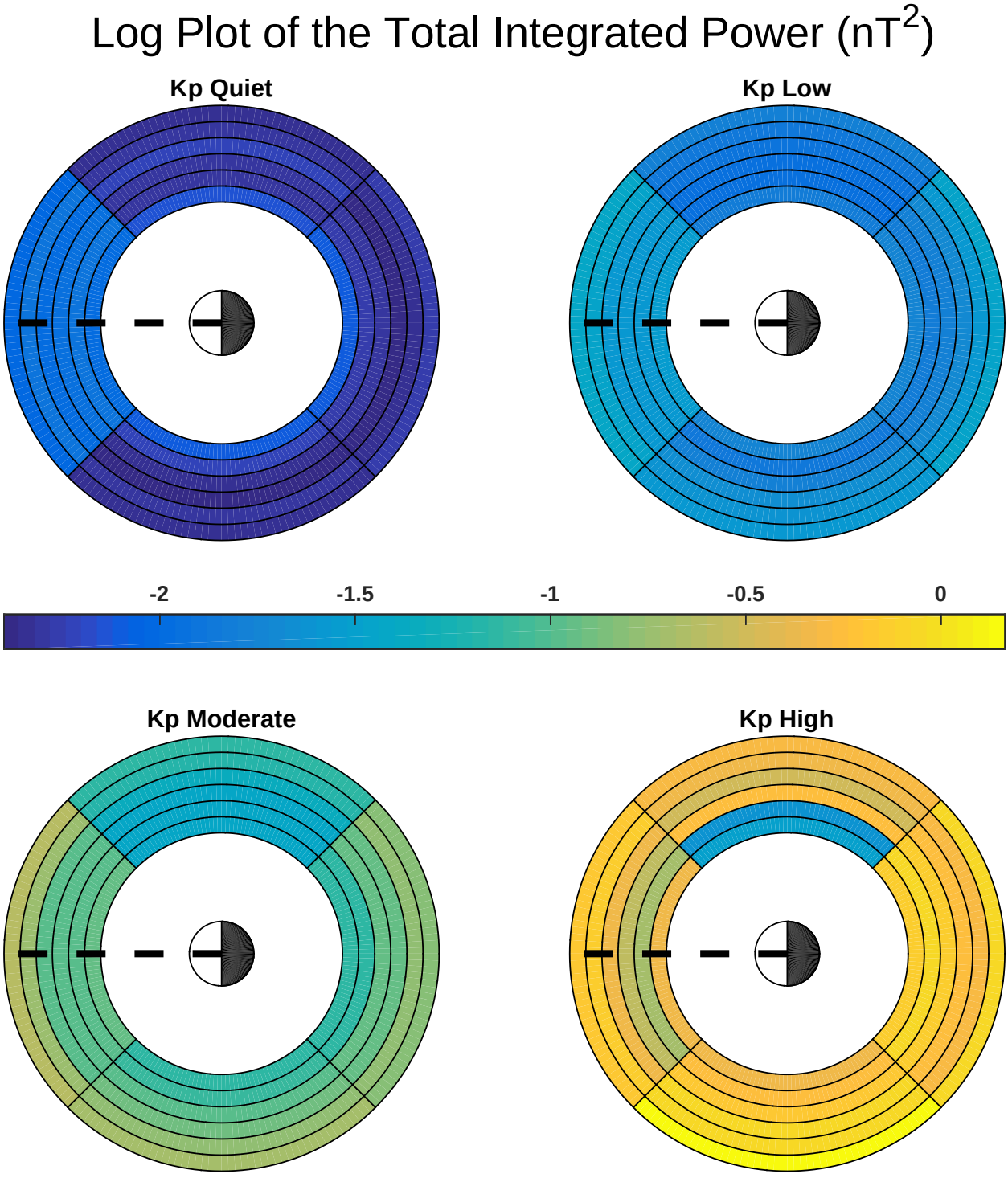


Figure 3.4: Log of the total integrated power of the median PSDs from each bin separated by Kp. The power spectral densities are integrated over the range 0.8 mHz to 8.1 mHz.

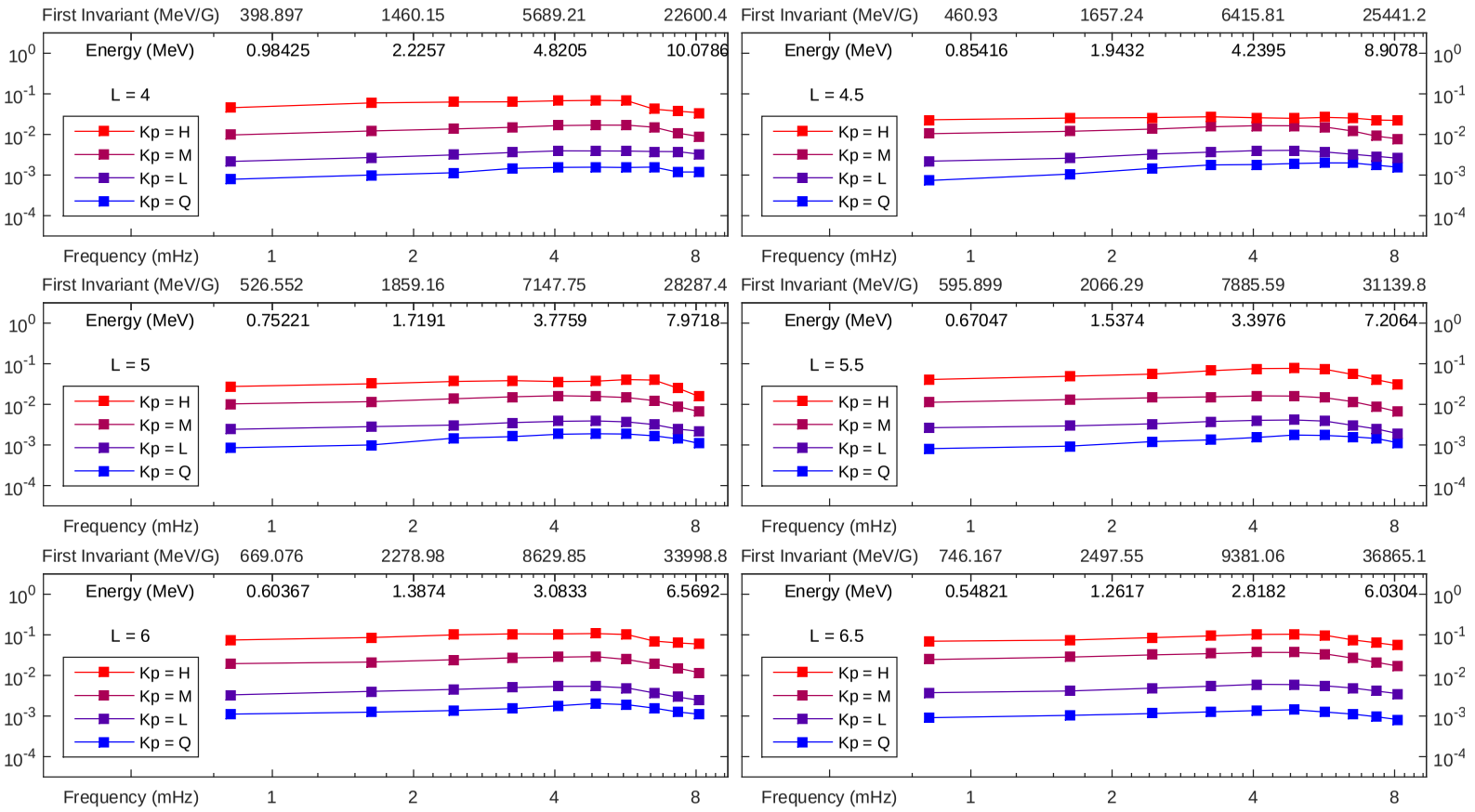


Figure 3.5: Median magnetic field PSDs in the noon sector for various Kp and L. The energies and first invariant values presented are computed assuming that $\omega = \omega_d$ for an electron.

the drift-averaged PSDs show many of the same characteristics as MLT-dependent PSDs. For ultra low-frequencies we see a relatively flat PSD with a weak maximum, possibly due to field-line resonance, at around 5 mHz. Then the power drops off noticeably as frequency increases. There is a clear dependence on the level of geomagnetic activity Kp, with higher power being directly correlated to higher Kp. Furthermore, the PSDs show a very weak dependence on L, with the general trend being increasing PSDs with increasing L.

3.5 Magnetic Radial Diffusion Coefficients

Using the *Fei et al.* [2006] expression given by Equation (1.18), we used the drift-averaged magnetic field PSDs to compute the magnetic component of the radial diffusion coefficients. In Figure 3.7 we see the computed values of D_{LL}^{B}[CRRES] for various L-shell values and Kp as a function of frequency. The diffusion coefficients show a clear frequency dependence for frequencies in the Pc-5 range. They scale as a power law for lower frequencies, leveling off at higher frequencies. Figure 3.8 shows D_{LL}^{B}[CRRES] for several fixed values of the first invariant as a function of L; we note that D_{LL}^{B}[CRRES] increase as M increases, with the effect being more pronounced as Kp increases.

Direct comparison with other previously published results is difficult because of differing field models and observations used for the diffusion coefficient calculations. Various studies have used different methodologies as well as different parameters to compute and parametrize the diffusion coefficients. *Brautigam and Albert* [2000] assumed that the root mean square of the electric field amplitude is a linear function of Kp and used the *Cornwall* [1968] expression to compute D_{LL}^{E}[BA] as a function of L, M, and Kp. *Elkington et al.* [2003] used a compressed dipole analytical field to drive a test particle simulation at a single point $L = 6.6$, while *Fei et al.* [2006] used a compressed dipole analytical field to solve the diffusion equation for $2 \leq L \leq 10$ and presented D_{LL}^{B} assuming an $L^{8.5}$ static power law. *Brautigam et al.* [2005] assumed a dipole field and used the CRRES electric field measurements to estimate D_{LL}^{E}[CRRES] as a function of both L and Kp. *Huang et al.* [2010b] used LFM runs to drive test particle simulations and compute D_{LL}^{B}. More recent studies

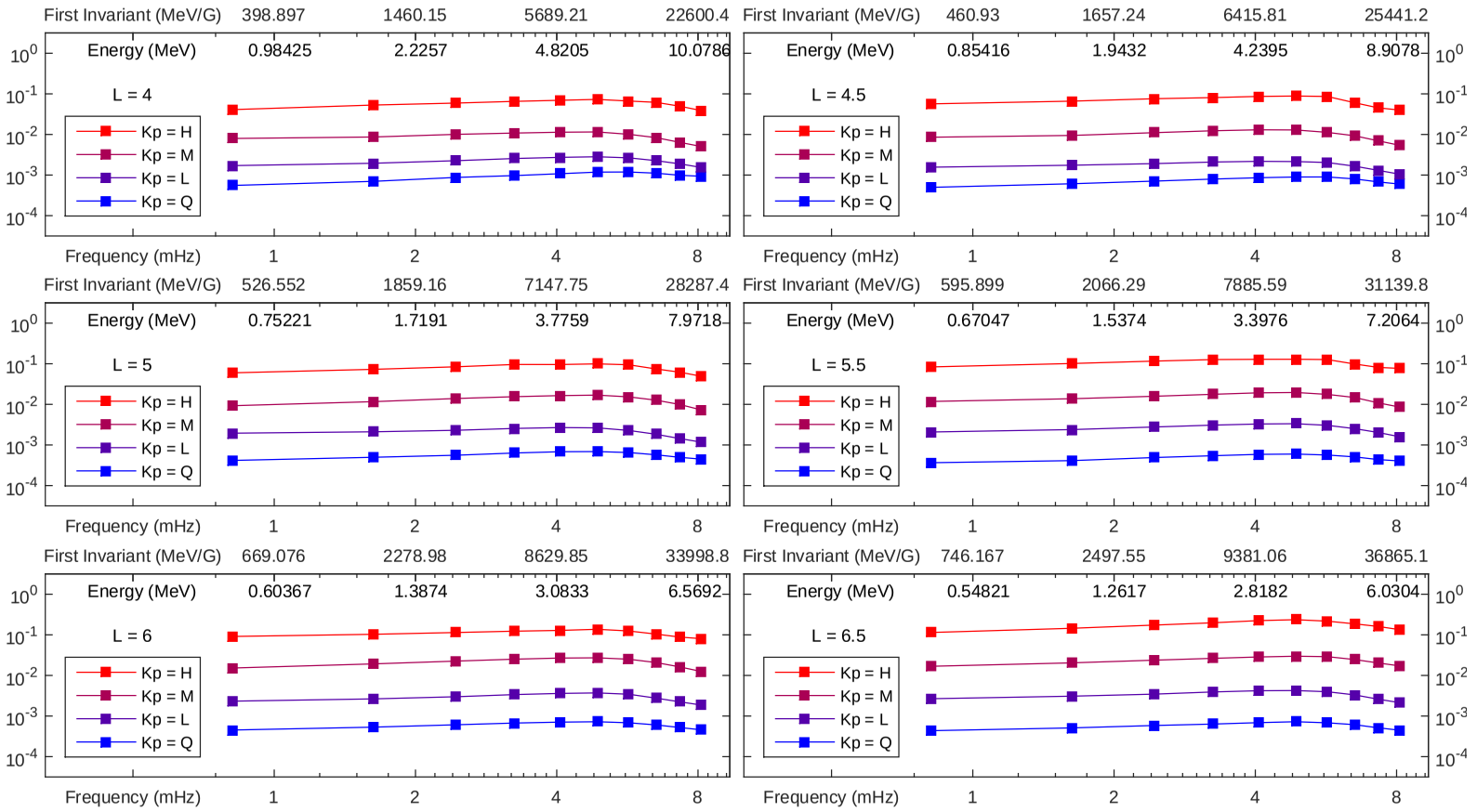


Figure 3.6: Drift-averaged power spectral density of the magnetic field at various L and Kp as a function of frequency. The energies and first invariant values presented are computed assuming that $\omega = \omega_d$ for an electron.

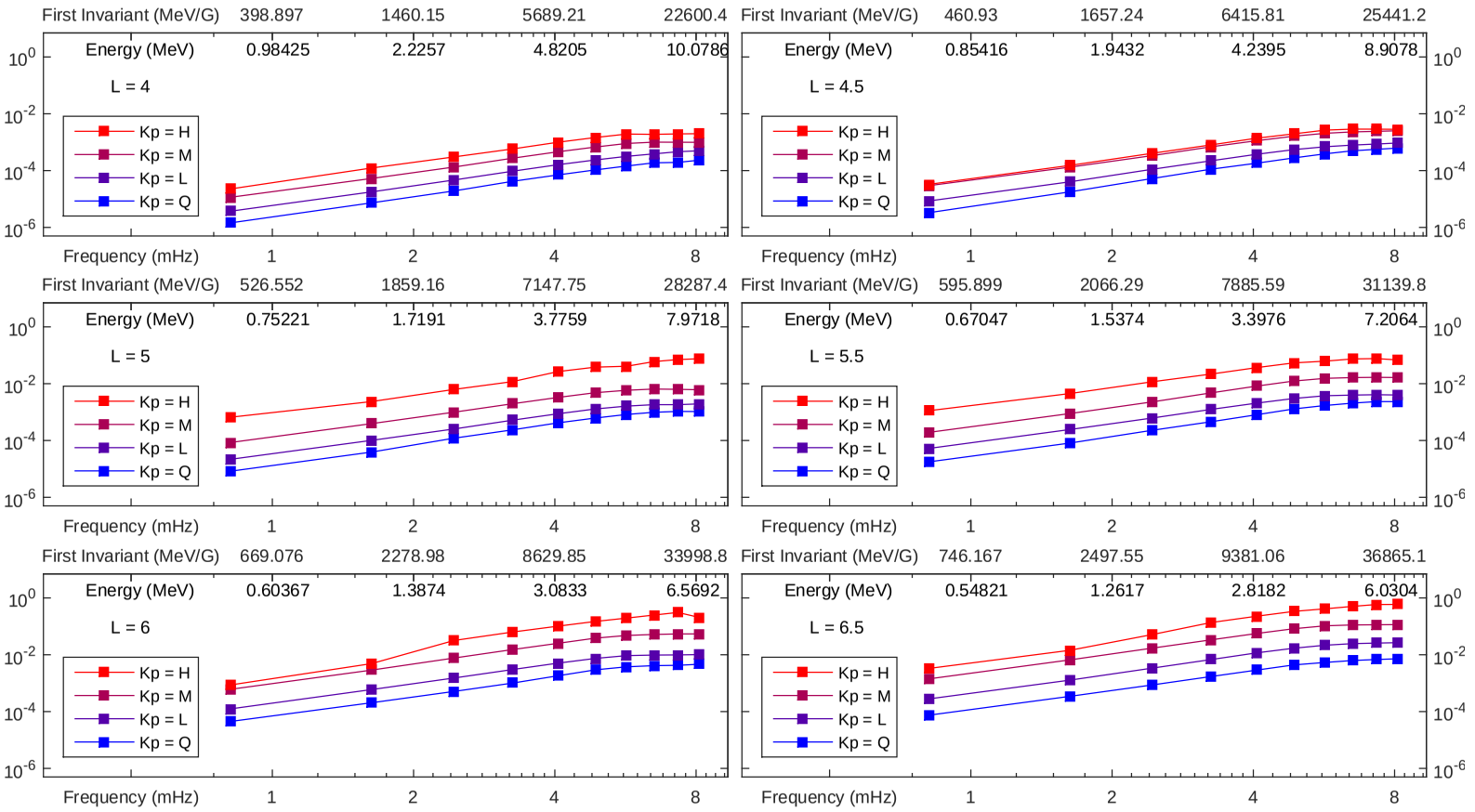


Figure 3.7: D_{LL}^{B}[CRRES] as a function of frequency on a log-log scale for various values of L and Kp. Corresponding values of energy and the first invariant are also presented for ease of comparison, computed assuming $\omega = \omega_d$ for an electron.

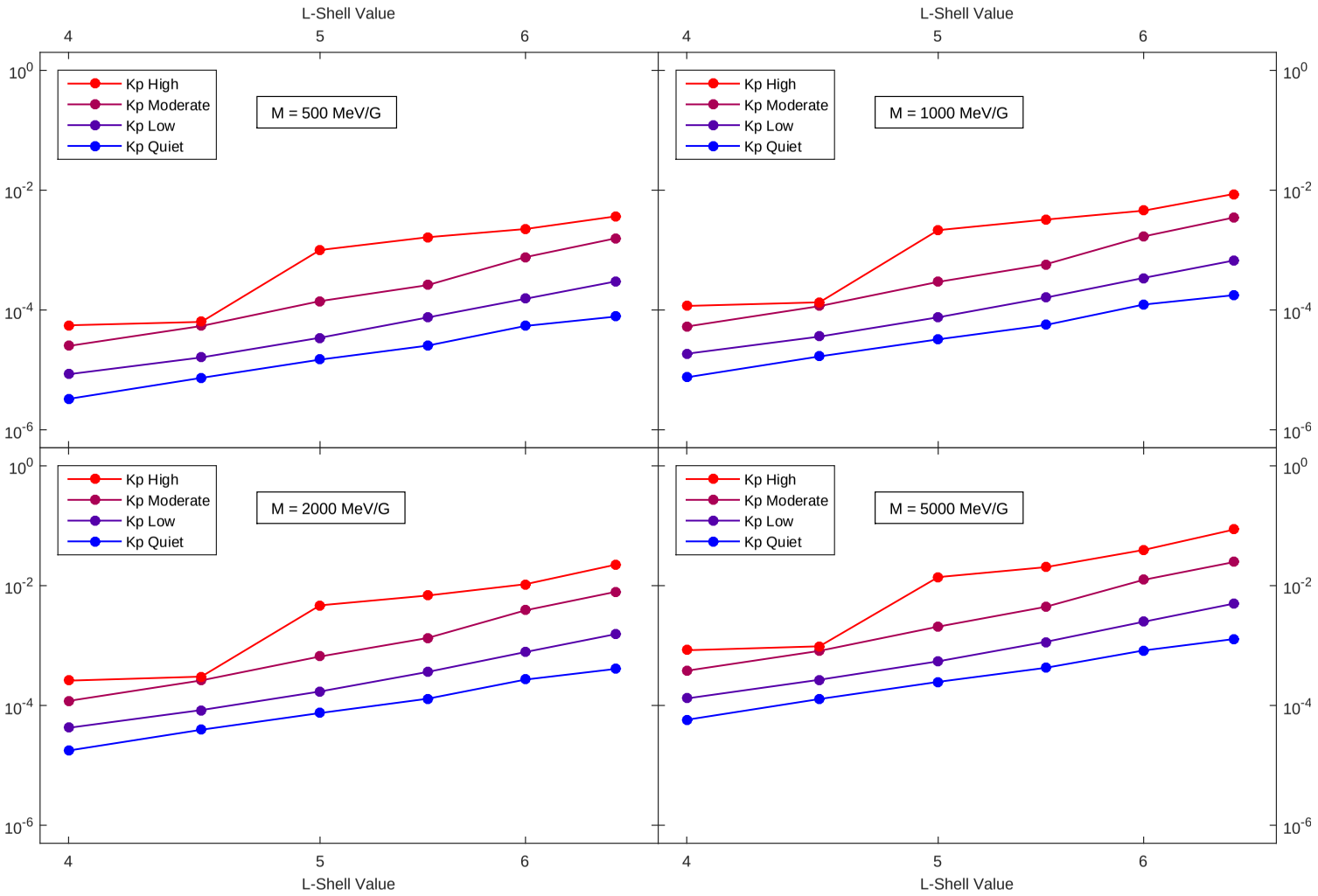


Figure 3.8: D_{LL}^{B}[CRRES] as a function of L on a log-log scale for various fixed values of the first invariant M and Kp.

include *Ozeke et al.* [2012] and *Ozeke et al.* [2014], which used *in situ* magnetic field measurements from AMPTE and GOES to compute D_{LL}^{B}[Ozeke], and ground magnetic field PSDs mapped to electric field PSDs in space to compute D_{LL}^{E}[Ozeke]. *Tu et al.* [2012] used LFM MHD simulations along with observations from GOES and THEMIS satellites to estimate both the electric and the magnetic components of the diffusion coefficients.

For our study, during the lifetime of CRRES, Interplanetary Monitoring Platform (IMP-8) was the only source of solar wind parameters. This data was too sparse to be of significant use in a systematic study such as ours. Therefore we chose the level of geomagnetic activity Kp as our time-dependent parameter to study the radial diffusion coefficients. Furthermore, the magnetic radial diffusion coefficients are presented as a function of frequency, where the frequency f varies directly with M and inversely with L^2,

$$f = \frac{3M}{2\pi\gamma q L^2 R_E^2}, \tag{3.8}$$

$$\gamma = \left[1 + \frac{2BM}{W_{\text{rest}}}\right]^{1/2}, \tag{3.9}$$

$$= 1 + \frac{W}{W_{\text{rest}}}, \tag{3.10}$$

$$W = W_{\text{rest}}(\gamma - 1), \tag{3.11}$$

$$= W_{\text{rest}}\left[\left(1 + \frac{2BM}{W_{\text{rest}}}\right)^{1/2} - 1\right], \tag{3.12}$$

where we assume SI units and a symmetric dipole field resulting in the factor of 2π in the expression for f. For any other magnetic field model, the factor 2π must be replaced by integrating the period function over one complete drift orbit. Here M is the relativistically corrected first invariant, γ is the relativistic correction factor, q is the charge of the particle, R_E is the Earth's radius, $W_{\text{rest}} = m_0 c^2$ is the rest energy of the particle, and B is the strength of the local magnetic field which, assuming a symmetric dipole field, can be approximated by B_0/L^3 where B_0 is the strength of the magnetic field on the surface of the Earth. Since we have two algebraic equations relating four variables L, the first invariant M, frequency f, and energy W, this system has two degrees of freedom. This

means that given the values of any two of the variables it is a straightforward matter to compute the other two variables.

As Figures 3.7 and 3.8 show, D_{LL}^{B}[CRRES] as a function of L spans several orders of magnitude, even for a fixed M or Kp. Obtaining a global model in which all of the magnetic diffusion coefficients can be represented as a function of L, M, and Kp is challenging. Obtaining a relatively simple model which would capture most of the variation in the magnetic diffusion coefficients is considerably more challenging. Therefore, in order to first obtain a simple model, we used a genetic algorithm with the fitness function being the sum of the squared residuals in the least squares sense. Numerous runs of the algorithm resulted in numerous models which were then sorted and ranked according to the complexity of the model. The complexity of a model can be gauged by counting the number of terms presented in the model, how many operations are required to evaluate the model, and the complexity of the individual terms in the model. For example, a monomial term is simpler than a term involving a hyperbolic function. After selecting a reasonably simple model, we performed weighted least squares on the entire dataset and obtained

$$D_{LL}^{B}[\text{CRRES}] = \exp(L + 0.00060104 \cdot M + 0.10003 \cdot Kp \cdot L - 16.618), \tag{3.13}$$

where the units of the second coefficient are G/MeV. Here M is the first adiabatic invariant (in units of MeV/G), L is the unitless McIlwain parameter and Kp is the unitless planetary K-index, which is a global measure of a geomagnetic disturbance. Note here that this model has no physical basis and is presented solely as a compact presentation of CRRES magnetic diffusion coefficients in order to aid the community with its modeling efforts. This model is much more succinct and easier to use than large tables containing numerical values of D_{LL}^{B} or dozens of fitted curves separated by Kp and M. The domain used to fit this model in L was $4.0 \leq L \leq 6.5$. In Kp, the quiet, low, moderate, and high levels of activity were represented by $Kp = 1$, 3, 5, and 7 respectively. The domain for the first invariant M ranged from 500 MeV/G to 5000 MeV/G. The diffusion coefficient D_{LL}^{B}[CRRES] is returned in the units of days^{-1}. Since the model is continuous in all three parameters it provides an excellent approximation, not only for interpolation inside the

domain, but for extrapolation as well. Kp can take any value in $\{1, 1.3, 1.7, 2, 2.3, ..., 6.3, 6.7, 7\}$ while L and M can vary continuously in their respective domains. Furthermore, the trends inside the domain for each parameter continue in a reasonable fashion outside the domain. For Kp, it is possible to plug in any value in $\{0, 0.3, 0.7, 1, 1.3, ..., 8.3, 8.7, 9\}$ while the limits on L and M for extrapolation are left to the individual reader to decide. Obviously, if we move far enough away from the boundary of the domain, the diffusion coefficients just simply can no longer remain physical. The reader has the option of just using the nearest neighbor approximation for points outside of the domain. For example, value of $Kp = 7$ can be used for all $Kp > 7$, value of $L = 4.0$ can be used for lower L-shell values, and so on.

In Figure 3.9 we present a comparison of our results to some of the previously published results. Here D_{LL}^{E}[CRRES] is presented from *Brautigam et al.* [2005], computed using the CRRES electric field data; D_{LL}^{B}[CRRES] has been computed using the CRRES magnetometer data and is the main subject of this chapter; D_{LL}^{E}[BA] and D_{LL}^{M}[BA] are presented from *Brautigam and Albert* [2000], computed using ground-based measurements mapped onto the equatorial plane; and D_{LL}^{E}[Ozeke] and D_{LL}^{B}[Ozeke] are presented from *Ozeke et al.* [2012], computed using ground-based measurements as well as *in situ* measurements from AMPTE and GOES.

Comparing our magnetic diffusion coefficients with previous estimates, we see in Figure 3.9 that D_{LL}^{B}[CRRES] is smaller than some of the previous estimates. Specifically, in comparing with D_{LL}^{E}[CRRES] from *Brautigam et al.* [2005], we present evidence against previous notions that the magnetic component is dominant over the electric component in driving particle diffusion [*Brautigam and Albert*, 2000; *Brautigam et al.*, 2005]. Simulations conducted using various numerical solvers [*Shprits et al.*, 2005; *Varotsou et al.*, 2008; *Albert et al.*, 2009; *Tu et al.*, 2009; *Chu et al.*, 2010; *Su et al.*, 2010, 2011a,b; *Subbotin et al.*, 2011a,b; *Kim et al.*, 2012; *Tu et al.*, 2013; *Li et al.*, 2014] used the *Brautigam and Albert* [2000] and *Brautigam et al.* [2005] analytic expressions to drive radial diffusion in their simulations and hence implicitly assumed that the magnetic diffusion coefficients were much larger than the electric diffusion coefficients.

Our estimates of D_{LL}^{B}[CRRES] are much closer to D_{LL}^{B}[Ozeke]. *Ozeke et al.* [2012] and *Ozeke*

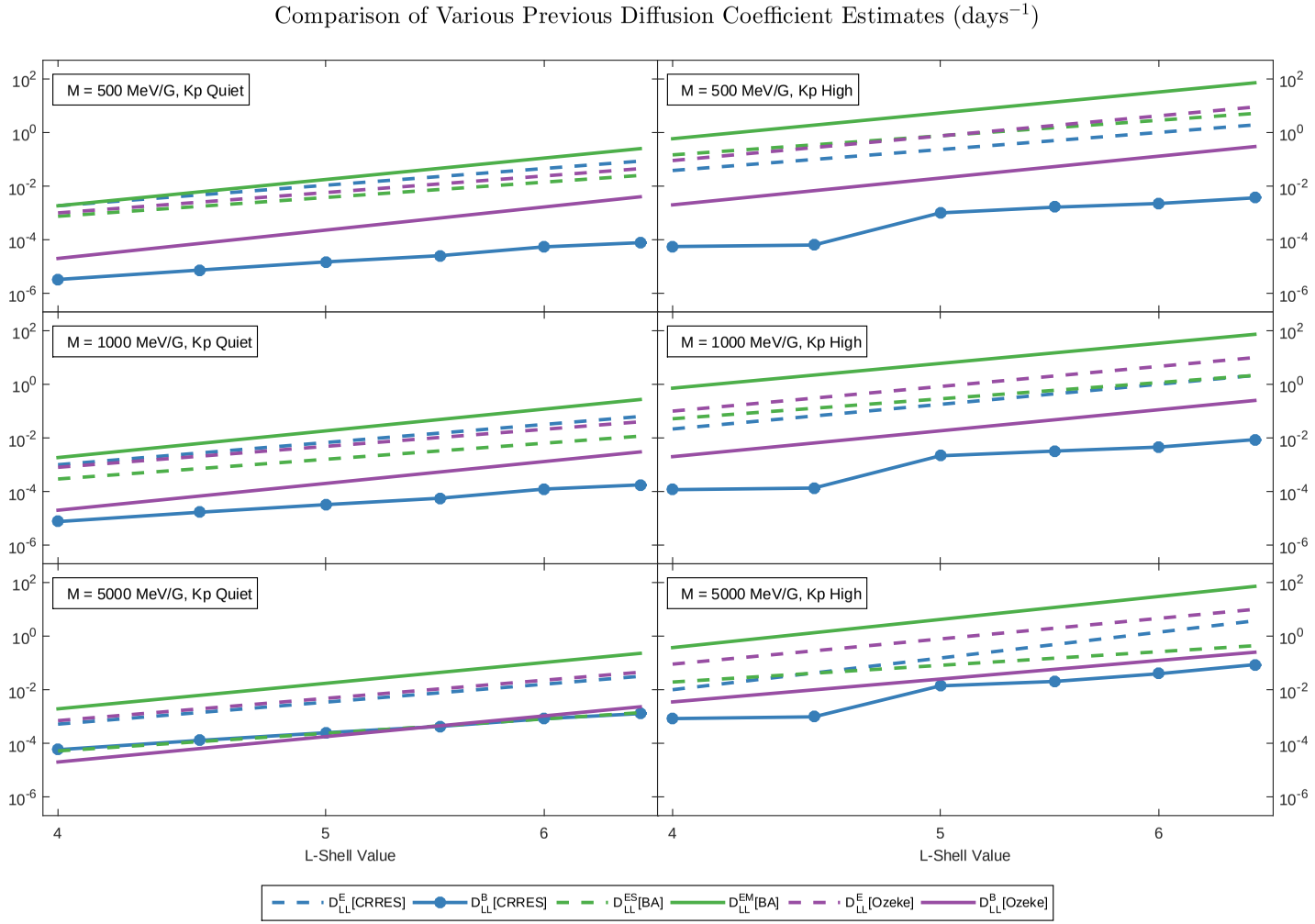


Figure 3.9: A comparison of D^{B}_{LL} for various values of the first invariant M, with some previous estimates of the radial diffusion coefficients. D^{E}_{LL}[CRRES] estimates from *Brautigam et al.* [2005] are shown, along with our estimates of D^{B}_{LL}[CRRES]. *Brautigam and Albert* [2000] and *Ozeke et al.* [2012, 2014] estimates are also presented here.

et al. [2014] corroborate our conclusions about the relative magnitudes of D_{LL}^{B} and D_{LL}^{E}. They computed the electrical diffusion coefficients using more than fifteen years of ground magnetometer measurements, and then mapped these magnetic field measurements to the electric field in the equatorial plane. The magnetic diffusion coefficients in *Ozeke et al.* [2012] were derived using *in situ* measurements from AMPTE (CCE) and GOES (East and West) spacecraft. *Ozeke et al.* [2014] included some additional data from THEMIS, but utilized only a limited L-shell range from $L = 5$ to $L = 7$. Figure 3.9 shows that D_{LL}^{B}[CRRES] increases as M increases and D_{LL}^{B}[CRRES] has excellent agreement with D_{LL}^{B}[Ozeke] for high M.

In Figure 3.10 we show D_{LL}^{B}[CRRES] as a function of L for various values of Kp and the first invariant M. In each of the four panels the bottommost curve shows D_{LL}^{B} vs. L, with M held fixed at 500 MeV/G. M is then incremented by $\Delta M = 100$ MeV/G until it reaches a maximum value of 5000 MeV/G represented by the topmost curve. In the fourth panel, with high Kp activity, we see that the diffusion coefficients scale much faster than they do for quiet Kp. This suggests that the L-dependence may not be a constant but rather a function of Kp. The diffusion rates increase as M increases, with the effect becoming more and more pronounced as the level of geomagnetic activity increases. This is in contrast with the results of *Ozeke et al.* [2012] and *Ozeke et al.* [2014], which reported a slight decrease of magnetic diffusion coefficients as M increases.

Figure 3.11 shows the scaling exponent n as a function of the first invariant M and Kp. For quiet and low Kp the scaling exponents are almost constants at $n = 6.5$ and $n = 7.75$, respectively. For moderate and high Kp the scaling exponents stabilize at around $n = 8.75$ and $n = 9.5$, respectively, as M increases. We assumed the functional form of a simple power law,

$$D_{LL}^{B} = D_0 L^n, \tag{3.14}$$

where the parameters D_0 and n are functions of both Kp and the first invariant M. Since the diffusion coefficients do not exhibit a constant variance (homoscedasticity) as a function of the independent variable L, we cannot use the ordinary least squares method [*Draper and Smith*, 1966]. Since it is difficult to estimate the true variance here, weighted least squares is also not

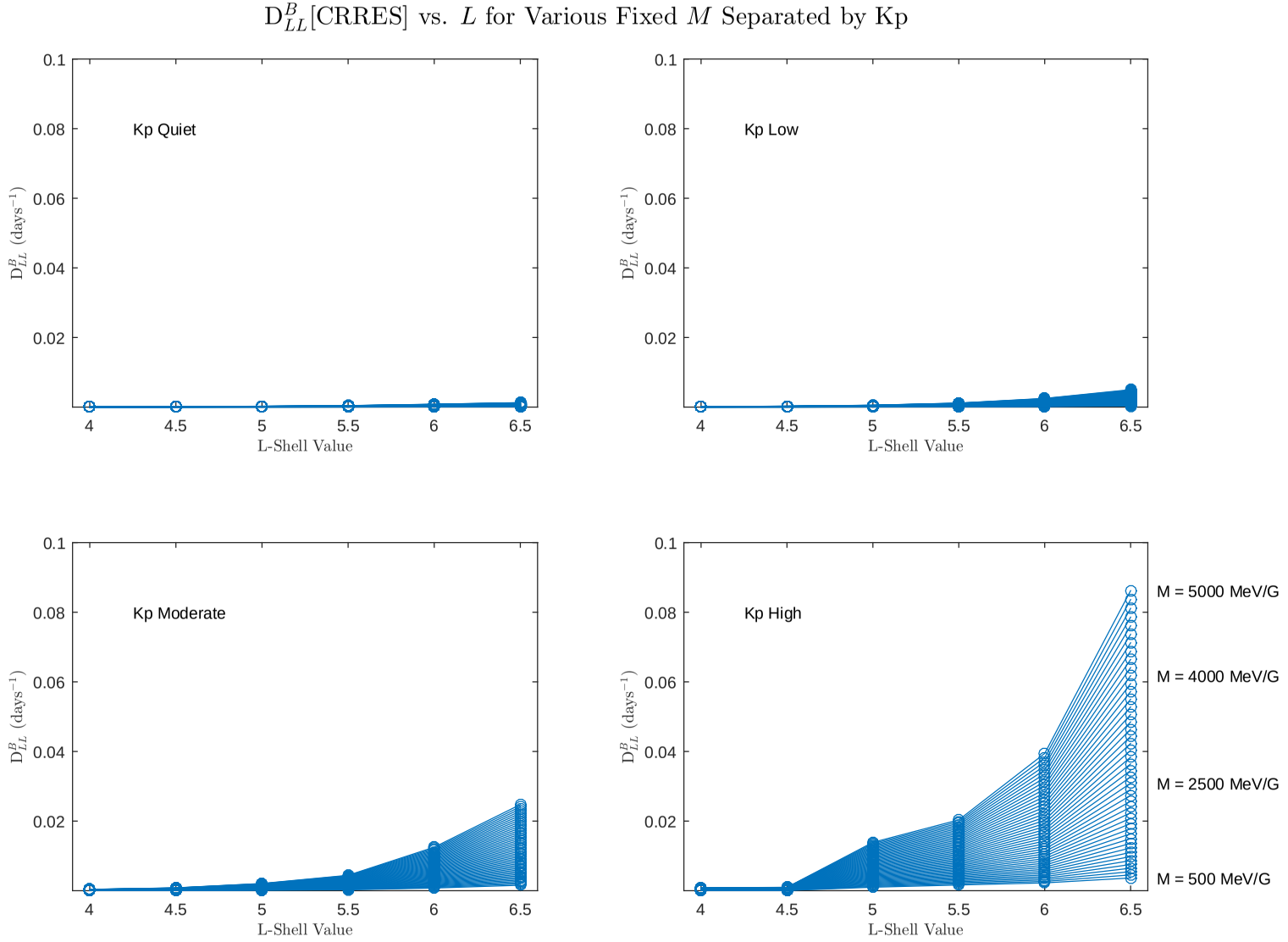


Figure 3.10: Plots of D^B_{LL}[CRRES] vs. L as the first invariant M increases from $M = 500$ MeV/G to $M = 5000$ MeV/G separated by Kp.

recommended. Therefore we use the Theil-Sen method [*Theil*, 1950; *Sen*, 1968] to estimate the slope of the best fit line after transforming the data into log-log space as described in Section 2.7. The method is quite easy to implement. In order to estimate the linear trend in a given set of data, we compute the slopes between all possible pairs of points and then pick the median slope. This method is a very robust and nonparametric (distribution-free) method, which is why log-transforming the data has no effect on the Theil-Sen estimate. This method can be significantly more accurate in the case of skewed or heteroscedastic data [*Wilcox*, 1998a,b, 2001].

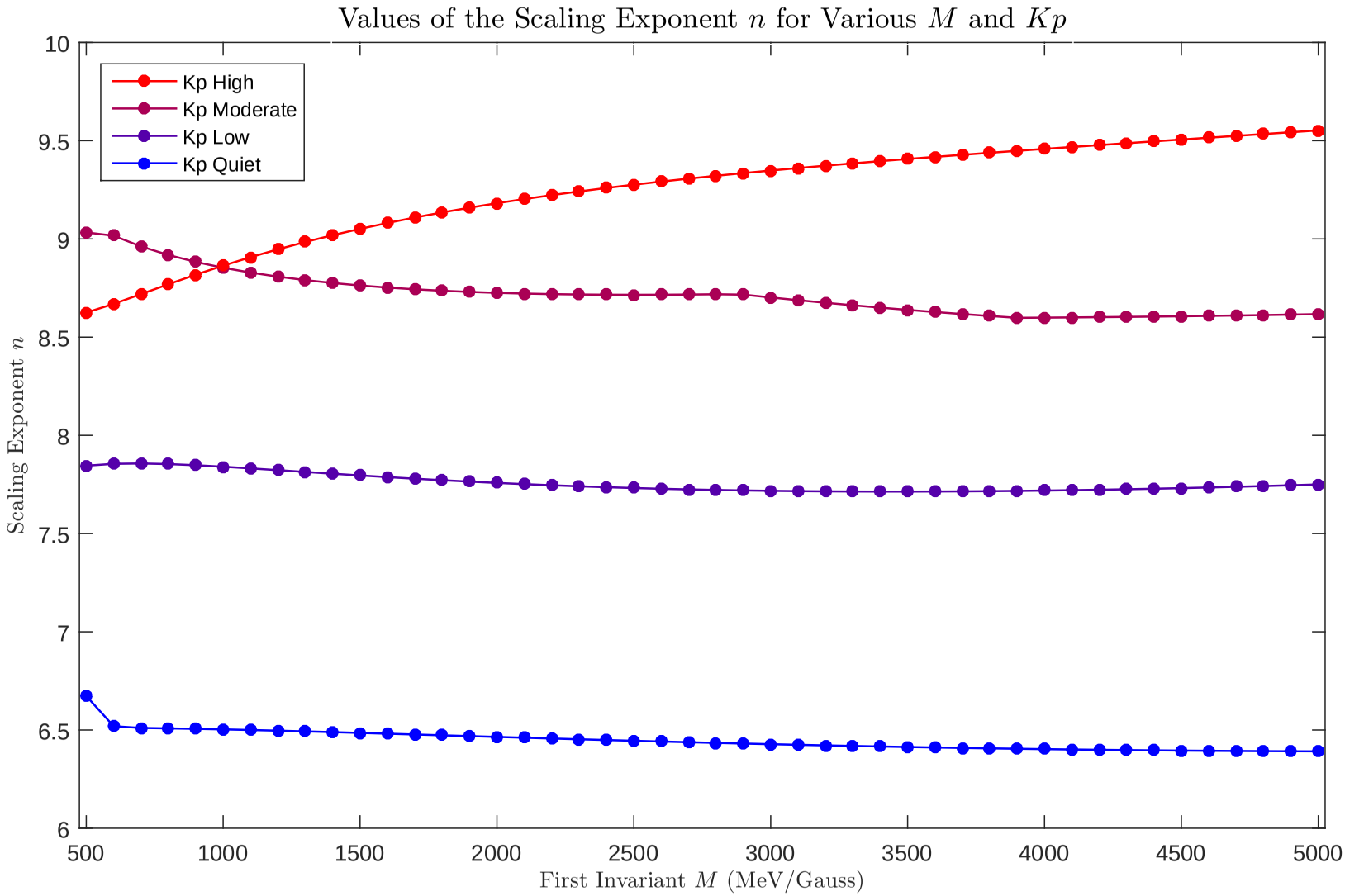


Figure 3.11: Plot of the radial scaling exponent n as a function of the first invariant M separated by Kp.

Chapter 4

Van Allen Probes

In this chapter, we analyze the magnetic and the electric field measurements returned by the instrument suites on board the Van Allen Probes, formerly known as the Radiation Belt Storm Probes (RBSP). This mission is the first NASA mission since CRRES which is dedicated to the study of the Van Allen radiation belts. It includes state-of-the-art instruments which have provided us with the most detailed and accurate field and particle measurements to date. We used the field measurements along with the *Fei et al.* [2006] formulation to estimate both the magnetic and the electric components of the radial diffusion coefficients. Since both components were processed and analyzed with identical techniques, we can compare and contrast both components without having to allow for differing techniques. The large amount of data available for this statistical study also allowed us to include the error estimates to gauge how much variability such estimates can display when real *in situ* measurements are used.

4.1 Mission Details and Instruments

The Van Allen Probes, formerly known as Radiation Belt Storm Probes (RBSP), is a NASA mission dedicated to the study of Van Allen radiation belts. The mission consists of two identical spacecraft, which were launched on August 30, 2012. The initial mission duration was two years, plus an extended mission. The probes have nearly identical orbits, with perigee at approximately 500 kilometers and apogee at approximately 30,000 kilometers (5.7 R_E), an inclination of almost ten degrees, and an orbital period of about 9 hours. Both probes are spin stabilized with the spin

rate being about 5 RPM. The Electric and Magnetic Field Instrument Suite and Integrated Sciences (EMFISIS) on board each of the Van Allen Probes includes a tri-axial fluxgate magnetometer with a dynamic range of 0.008 nT to 65,536 nT, providing the full magnetic field vector with a cadence of 64 measurements per second [*Kletzing et al.*, 2013]. The Electric Field and Waves (EFW) suite on board each of the Van Allen Probes consists of four spin plane booms and two spin axis booms to measure the ambient electric field. The spin plane component of the electric field is provided with the dynamic range being 0.05 mV/M to 1 V/m with a cadence of 32 measurements per second [*Wygant et al.*, 2013]. The spin fit level 2 electric field data contains only the E_y and E_z components in modified GSE (mGSE) coordinates. The modified GSE (mGSE) coordinate system is defined with the x-axis being the Van Allen Probes spin axis, which is always within 37° of the Sun-Earth line, the y-axis intersecting the ecliptic and the spin plane and pointing towards dusk, and the z-axis completing the coordinate system.

4.2 Data Preparation and Processing

The data used in this statistical study spans three years, from September 2012 to August 2015. The magnetic ephemeris parameters MLT and L^* were provided as a data product by Los Alamos National Lab (LANL) (www.rbsp-ect.lanl.gov/science/DataDirectories). MLT refers to the magnetic local time of the spacecraft in eccentric dipole coordinates given in units of hours. L^* is the Roederer L-shell value [*Roederer and Zhang*, 2014] corresponding to a 90° pitch angle computed using the Tsyganenko storm time magnetic field model (TS04D) [*Tsyganenko and Sitnov*, 2005]. The activity-dependent parameter used for this study is three-hour averaged Kp obtained from NASA's OMNI database (cdaweb.gsfc.nasa.gov/). The index Kp was chosen over other activity indices because it is a widely used planetary index indicative of magnetic activity on a global scale. Furthermore, the use of Kp facilitates comparison with other previous studies.

For the magnetic field data, we started with the level 3 (L3) EMFISIS fluxgate magnetometer data provided at four-second resolution in solar magnetic (SM) coordinates for both RBSP-A and RBSP-B. The data is hosted by the University of Iowa EMFISIS respository at

emfisis.physics.uiowa.edu. After deleting orbits which were incomplete (i.e., the provided magnetometer data spanned an interval much smaller than 9 hours), each magnetic field measurement was assigned the corresponding MLT, L^*, and Kp value using interpolation. At this point, all of the remaining orbits were inspected for spikes or invalid magnetic field measurements. Orbits with too many corrupted data points were removed while those with only a few corrupted points were fixed using interpolation. This was necessary because the subsequent Fourier analysis requires continuous data segments as well as to retain as much of the data as possible for a detailed statistical study. Since thruster firing events affect the fluxgate magnetometer readings, magnetic field measurements taken during spacecraft maneuvers should be discarded. During the course of our study it was noted that after a thruster firing event, all three components of the magnetic field recorded a "ringing" which introduced spurious broadband wave power between 1 mHz and 4 mHz and sometimes persisted for several days. Upon further examination, we saw that this ringing was not present in the compressional component of the magnetic field. Therefore, we deleted only the orbits in which maneuvering occurred without deleting any additional orbits. This resulted in about a 4% data loss for each spacecraft.

In order to make the comparison between the magnetic and the electric component of the diffusion coefficients easier, we decimated the magnetic field vector by a factor of three, in order to degrade the time resolution from 4 seconds to 12 seconds. For this purpose, we digitally filtered all three components of the magnetic field using a low-pass infinite impulse response (IIR) filter so that there wouldn't be any aliasing effects when we retain only every third data point [*Press et al.*, 2007]. Since we are only concerned with ULF wave power in the Pc-5 range (1.67 mHz - 6.67 mHz), the decrease in Nyquist frequency from $\sim$ 120 mHz to $\sim$ 40 mHz was inconsequential. In order to estimate D_{LL}^{B}, we need to estimate the power spectral density of the compressional component of the magnetic field vector $\mathbf{B}$, which in turn requires the variations in $\mathbf{B}$ in the direction of $\mathbf{B}$. In order to obtain $\Delta\mathbf{B}$, we employed a digital high-pass infinite impulse response filter after which we obtain $\langle\mathbf{B}\rangle$ and consequently the compressional component $\mathbf{B}_{\parallel}$. Because of our focus on the ultra low-frequency range, the large ambient magnetic field combined with a large gradient of the

magnetic field as the spacecraft moved inbound and outbound with high velocities at low L^*, it proved daunting to remove the static ambient field and to isolate the relatively small amplitude variations in which we were interested. We therefore deleted all data with $L^* < 2.5$. Figure 4.1 shows the B_z measurements from a typical orbit. The measurements are from perigee to perigee. We can see that the B_z component spans order of magnitude and also exhibits very large gradients as the spacecraft moves inbound and outbound close to the Earth at perigee. These large gradients made it quite difficult to estimate the background trend (shown in red in the top panel) which was to be removed in order to compute the compressional component of the magnetic field. The bottom panel shows B_z after it has been high-pass filtered by a digital filter. This is the only solution we found which was effective. As we can see in the bottom panel, the filtered signal exhibits the desired small amplitude oscillations but the amplitude does grow very large near perigee. For this reason, the data near both edges (at low L^*) was deleted.

We then employed the multitaper method to estimate the power spectral density in a manner similar to the one employed for the CRRES study. The tapers (window functions) chosen were the Discrete Prolate Spheroidal Sequences (DPSS), also known as the Slepian sequences [*Slepian*, 1978]. They are all mutually orthogonal vectors which provide statistically independent periodograms, which can then be averaged to obtain the final spectral estimate with reduced bias and variance. Data segments of the compressional component of the magnetic field of 20 minutes duration were utilizes to estimate the power spectral density using the multitaper method as described above. We chose the interval length of 20 minutes because that is how long the spacecraft typically take to cross an L^* bin of width 0.5 R_E. Since the resolution of the data segments is 12 seconds, the resolved frequencies lie between $\sim$ 0.825 mHz and $\sim$ 40 mHz with the frequency step size being $\sim$ 0.825 mHz. Since we only wish to consider the ULF Pc-5 range (1.67 mHz - 6.67 mHz), we truncate the power spectral density estimates at $\sim$ 8.25 mHz and ignore wave power at higher frequencies for the rest of this chapter. After estimating the power spectral density of each data segment, we assigned the MLT, L^*, and Kp values corresponding to the central data point of each data segment, and saves them to be binned later.

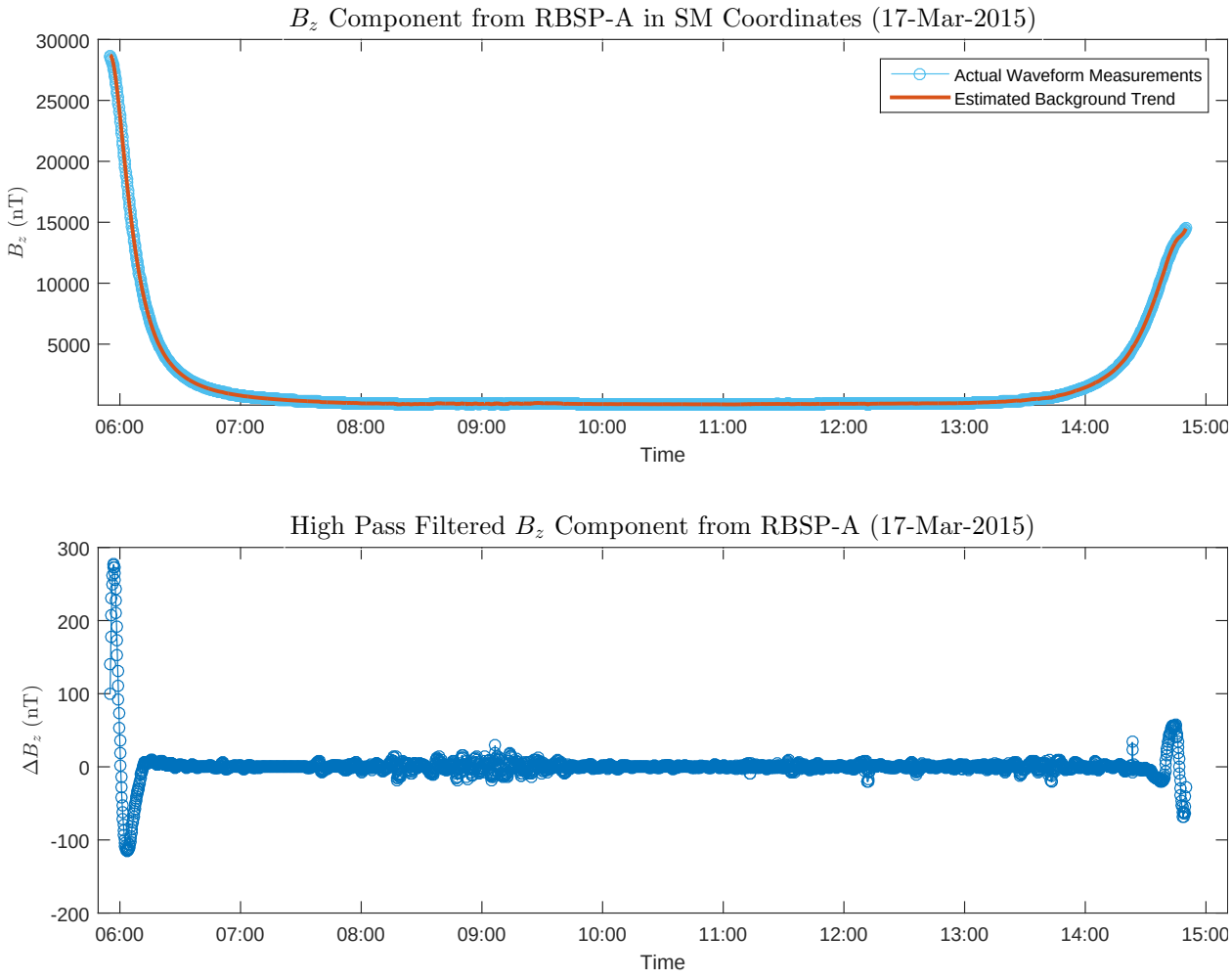


Figure 4.1: Plot of B_z measurements from a complete orbit. The top panel shows B_z along with the background trend which must be removed in order to compute the compressional component of **B**. The bottom panel shows the high-pass filtered B_z.

For the electric field data, we started with the level 2 (L2) EFW data provided at 12-second resolution in the mGSE coordinate system for both RBSP-A and RBSP-B. Note here that only the E_y and E_z components of the electric field were provided. This data was obtained from the University of Minnesota EFW data repository (www.space.umn.edu/rbspefw-data/). After deleting orbits that were corrupted or incomplete, each electric field measurement was assigned the corresponding MLT, L^*, and Kp values using interpolation, and the two components were inspected for invalid values or unphysical spikes. Orbits with too much corrupted data were eliminated and the rest were fixed using interpolation so that we could have as much data as possible for our study. The electric field data is affected by spacecraft charging events and eclipsing in addition to thruster firings. Similar to the magnetic field data preparation, some broadband spurious wave power was observed in the electric field data following spacecraft maneuvers but that wave power was not noticeable in the azimuthal component of the electric field. Therefore we inspected the electric field data for charging, eclipsing, and thruster firing events and deleted the orbits when these events occured. Because of the absence of spurious wave power following thruster fires, we only deleted the orbits in which a maneuver occurred without deleting any additional orbits. This resulted in a further 14% and 11% reduction, respectively, in the data available from RBSP-A and RBSP-B for this study.

Since the E_x component is missing in the L2 EFW data provided, we used the EMFISIS fluxgate magnetometer data to compute E_x. The magnetometer data was converted from the SM (solar magnetic) coordinate system to the mGSE (modified geocentric solar ecliptic) coordinate system; then using the assumption that $\mathbf{E} \cdot \mathbf{B} = 0$, we were able to compute the E_x component. Since small uncertainties in B_x result in large errors in E_x, not all of the values of E_x obtained in this fashion could be used. We started with the given accuracy of the B_x, B_y, B_z, E_y, and E_z measurements and used the standard error propagation formulas to estimate the error in E_x through the transformation of $\mathbf{B}$ from SM coordinates to mGSE coordinates and then through the relation $\mathbf{E} \cdot \mathbf{B} = 0$ [*Squires*, 2001]. After obtaining ΔE_x, we computed the elevation angle that the

magnetic field vector forms with the spin plane of the spacecraft in mGSE coordinates,

$$\theta = \left| \arctan\left(\frac{B_x}{\sqrt{B_y^2 + B_z^2}} \right) \right|. \tag{4.1}$$

Since we know that the smaller the elevation angle, the larger the uncertainty in E_x, we therefore expected ΔE_x to be a decreasing function of θ, because a large elevation angle implies a small error in E_x. This was indeed the case. In fact, ΔE_x decreased rapidly as a function of θ and $\Delta E_x = 0.5\text{mV/m}$ when $\theta \approx 5.75°$. Therefore this is the cutoff we used in our study. All data with an elevation angle $\theta < 5.75°$ was deleted. After this, the electric field data was converted from mGSE coordinates to SM coordinates, followed by the data's conversion from Cartesian coordinates to polar coordinates giving us the azimuthal component of the electric field E_ϕ which is necessary to compute the electric component of the radial diffusion coefficient D_{LL}^E. We would also like to warn the reader here that the EFW L2 data provided does not have a uniform cadence in time. The cadence is slowly changing as a function of time and oscillates around 11 seconds instead of the reported 12 seconds. This creates two challenges. First, classical Fourier analysis cannot be performed if Δt is not a constant. Second, even if it is erroneously applied to the EFW data assuming that $\Delta t = 12$ seconds, then the periodogram will be skewed, showing incorrect wave power for a given frequency. Because of this, we resampled E_ϕ to a constant cadence of 12 seconds before we estimated the power spectral density. E_ϕ was then digitally filtered in exactly the manner as the EMFISIS data so that wave power leakage into the ULF range could be avoided. The multitaper technique was then used with 20-minute data segments to estimate the power spectral density. The specifications for the digital filter and power spectral density estimation were identical for both the magnetic and the electric field measurements so that the two results may be meaningfully compared and combined.

In Figure 4.2, we see a typical time series for the level 2 electric field measurements for March 17, 2015. The data shown here is as it was obtained from the EFW data repository without any additional processing by us. The top panel shows the E_y component while the bottom panel shows the E_z component in mGSE coordinates, which are local spacecraft coordinates. The E_x data is too

unreliable and is not provided with the level 2 dataset. The time series exhibits a few effects which can invalidate the electric field measurements. The straight slanted lines exist because the data in that interval has been removed due to the eclipsing of the spacecraft and linearly interpolated. The several groups of large amplitude spikes are most likely due to spacecraft charging. Both of these types of events render the electric field measurements useless. Since this data is unreliable, we removed such data before proceeding. Figure 4.3 shows the azimuthal component computed for one of the orbits on March 17, 2015. In addition to the eclipsing and charging, we see a new effect (roughly at 07:00), where the azimuthal component seems to be diverging. This is because of the small magnitude of B_x. Since the E_x component was not provided, we had to use the approximation $\mathbf{E} \cdot \mathbf{B} = 0$ to estimate it. This requires division by B_x. If the magnitude of B_x is small, then small uncertainties in B_x will result in large uncertainties in E_x. Therefore we must discard the electric field data in this case as well. The bottom panel in Figure 4.3 shows the azimuthal component digitally filtered with the same IIR filter employed above for the magnetic field. The digital filter is very effective at removing all sorts of long term trends including the slope of the interpolating line. The spikes however still remain.

4.3 ULF Wave Power Distribution

In order to understand the dependence of ULF wave power on MLT, L^*, and Kp, it is necessary to parametrize the magnetic and electric field power. In MLT, we decided to use 4 bins of equal size. The MLT bins are centered on 0000 hours, 0600 hours, 1200 hours, and 1800 hours, with each bin spanning 6 hours. The noon bin, for example, spans 0900 MLT hours to 1500 MLT hours. In L^*, we use 6 bins which are centered on $L^* = 3, 3.5, 4, 4.5, 5, 5.5$. Each L^* bin is 0.5 R_E wide. The first bin, for example, spans $L^* = 2.75$ to $L^* = 3.25$. In Kp, we are only able to use 6 bins centered on $Kp = 0, 1, 2, 3, 4, 5$ because of a dearth of high geomagnetic activity during the

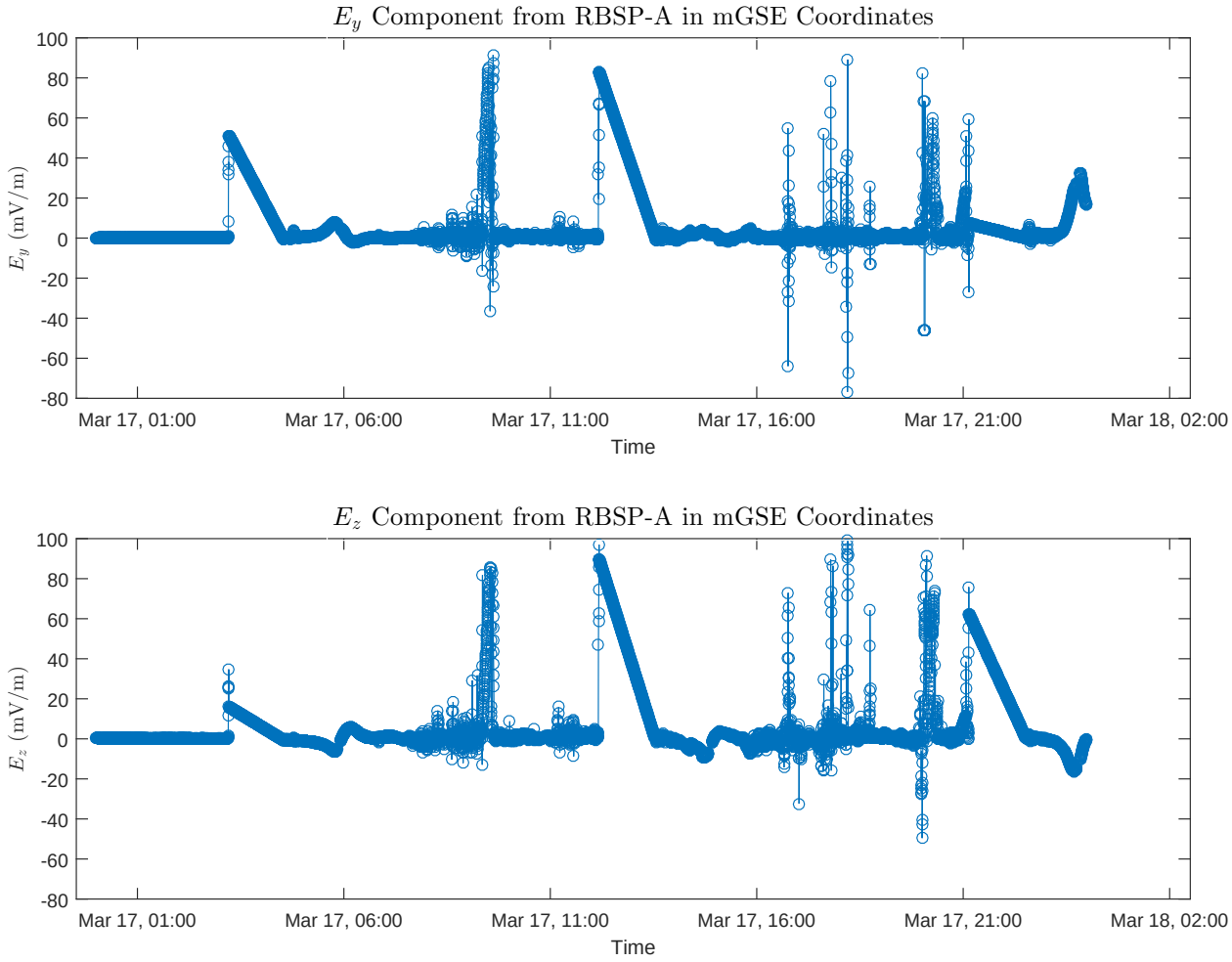


Figure 4.2: Plot of the level 2 electric field data for March 17, 2015. The top panel shows the E_y component while the bottom panel shows the E_z component.

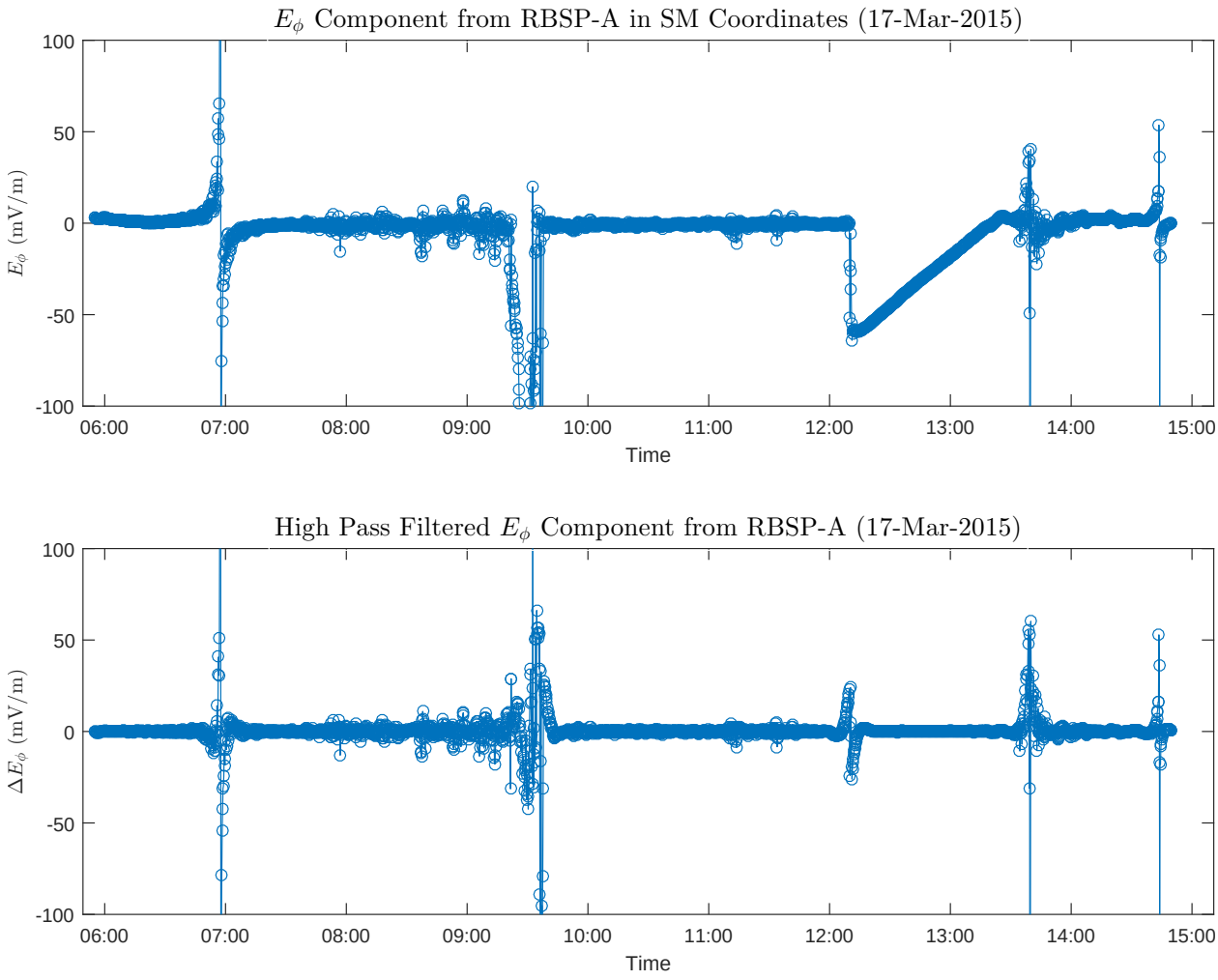


Figure 4.3: Plot of the azimuthal component of the electric field for an orbit. The top panel is the computed E_ϕ and the bottom panel is the high-pass filtered E_ϕ component.

Van Allen Probes era. The Kp bins are defined as follows,

$$\begin{aligned} Kp = 0 &\Rightarrow \{0, 0+\}, & Kp = 3 &\Rightarrow \{3-, 3, 3+\}, \\ Kp = 1 &\Rightarrow \{1-, 1, 1+\}, & Kp = 4 &\Rightarrow \{4-, 4, 4+\}, \\ Kp = 2 &\Rightarrow \{2-, 2, 2+\}, & Kp = 5 &\Rightarrow \{5-, 5, 5+, ...\}, \end{aligned}$$

where the last bin $Kp = 5$ includes all Kp values larger than 5. Figure 4.4 shows the number of spectra in the various bins after the magnetic field spectra and the electric field spectra have been binned. For a given spectrum, the MLT, L^*, and Kp values assigned to the center of the corresponding data segment determine the bin to which that spectrum is assigned. We can see that the statistics degrade as Kp increases. The higher the Kp, the smaller the amount of data available for both instruments.

Figure 4.5 shows the magnetic and electric field spectra in $L^* = 5.5$, $Kp = 2$, noon sector bin. All of the power spectral densities contained in that bin are shown in the background. The plots also show the mean, the geometric mean, and the median power spectral density with the interquartile range. The range of the spectra being several orders of magnitude shows that there can be significant variation even within a single bin. Since the extreme outlying spectra can take on very large values, the mean is a very inadequate measure of central tendency. In Figure 4.5 we can clearly see that the mean is at the 75th percentile. In such skewed distributions, the geometric mean or the median are much more appropriate as measures of central tendencies. For the ease of calculation and the robustness of the median, we chose the median power spectral density as the representative from each bin to proceed with this study.

Parameterizing the ULF wave power using magnetic local time allows us to study the ULF wave power distribution in azimuth. After selecting the median magnetic and electric spectra from each bin, the spectra are integrated over the truncated frequency range ~ 0.8 mHz to ~ 8 mHz giving us the total wave power which can then readily be compared across bins. Figure 4.6 shows the distribution of total ULF wave power estimates from the magnetic and electric field data. We clearly see that in general, the wave power increases as Kp increases which is to be expected as a higher level of geomagnetic activity results in larger perturbation amplitudes. The ULF wave power

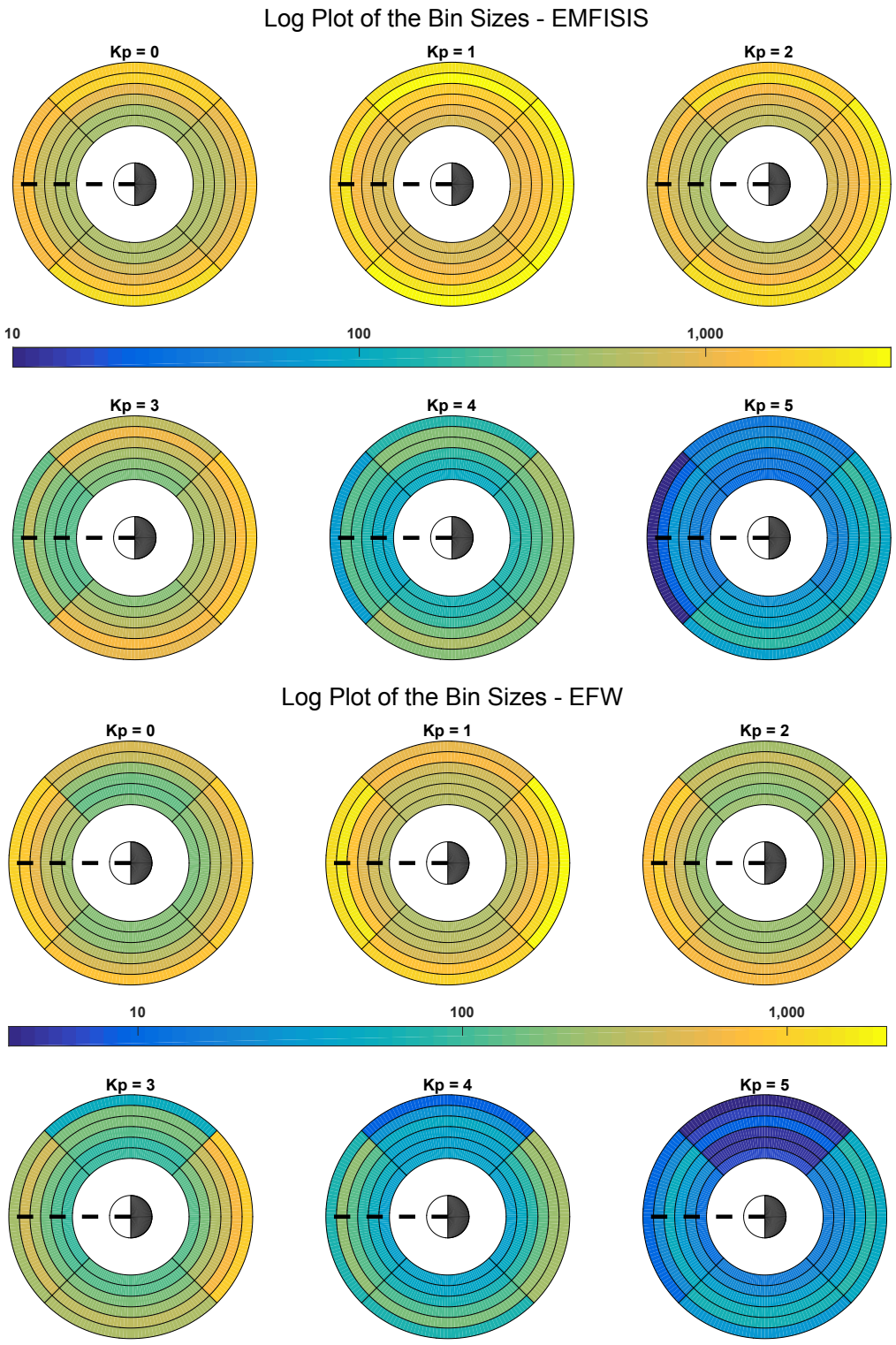


Figure 4.4: Log plots of the bin statistics for the EMFISIS magnetic field spectra and EFW electric field spectra, separated by Kp. The rings correspond to L^* bins between $L^* = 3$ and $L^* = 5.5$ with the dashed line being the Sun-Earth line.

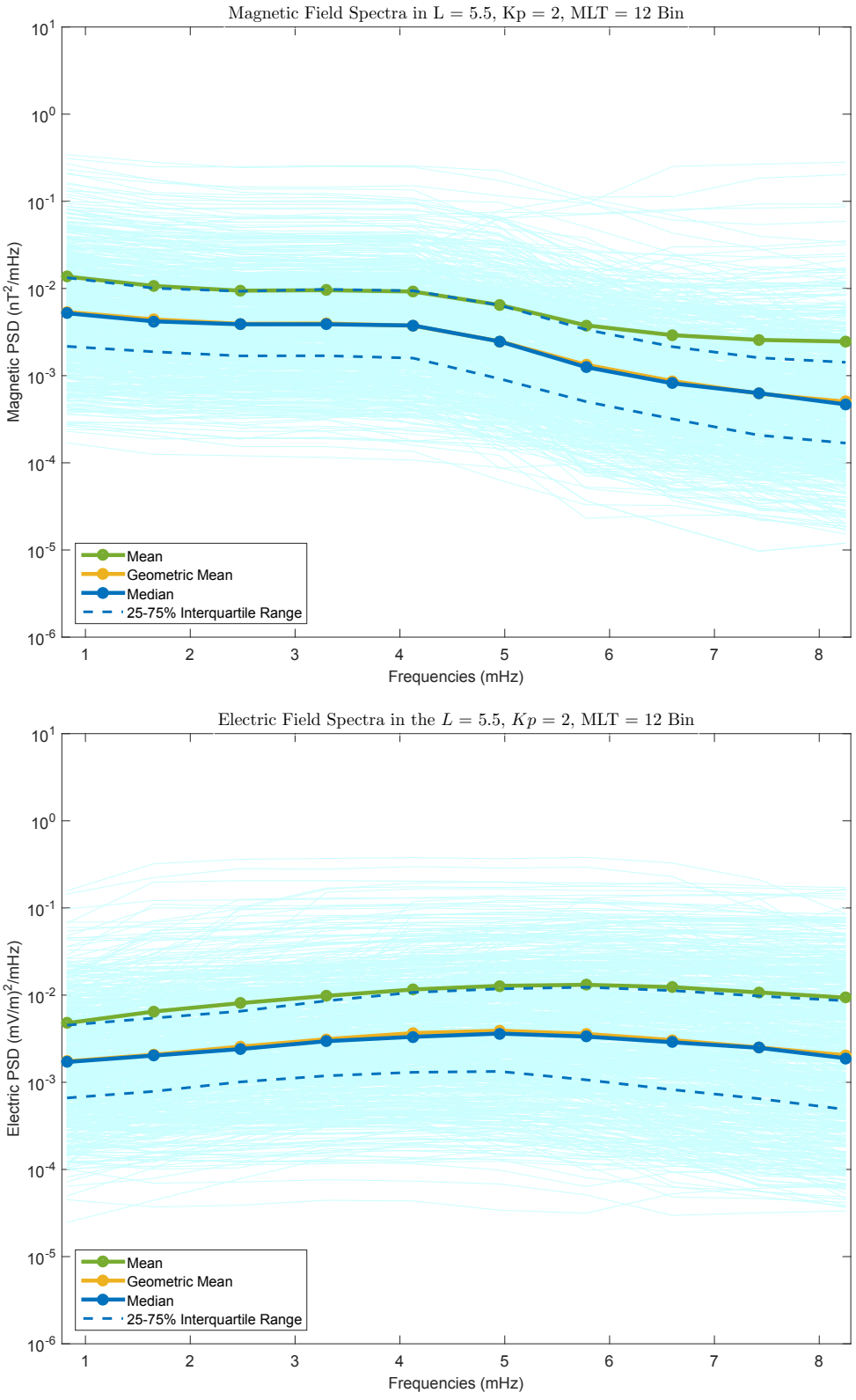


Figure 4.5: Magnetic and electric field spectra belonging to the $L^* = 5.5$, $Kp = 2$, noon sector bin with the mean, geometric mean, and median PSD.

varies directly with radial distance, as we see an increase in power as L^* increases. In magnetic local time, we see that the noon sector tends to have higher power than the other three sectors with no appreciable differences in ULF wave power between the dawn, dusk, and midnight sectors. These distributions are similar to the distributions presented in Figure 3.4 derived from CRRES magnetic field data.

4.4 Drift-Averaged Spectra and the Radial Diffusion Coefficients

Since the *Fei et al.* [2006] formulation requires the power spectral density of the compressional component of the magnetic field and the azimuthal component of the electric field to be a constant in azimuth, we need to estimate some sort of a drift-averaged spectrum. Since Figure 4.4 shows how the bin statistics vary in azimuth for any given L^* and Kp, taking a weighted average of the spectra seems an appropriate method to obtain a drift-averaged power profile. Since we do not want the measurements in a given local time sector to influence the drift-averaged profile unduly, we set the weights to be inversely proportional to the number of measurements in a given MLT sector, and for each L^* and Kp obtain a drift-averaged power spectral density for both the magnetic and the electric field measurements. Figure 4.7 shows the drift-averaged power spectral densities in frequency space, separated by L^* and Kp obtained from magnetic and electric fields measurements. In Figure 4.7, the corresponding energies as well as the relativistically corrected first invariant values are also calculated and shown, assuming the resonance condition in an azimuthally symmetric field configuration $\omega = m\omega_d$ and the azimuthal wave mode number being $m = 1$ for an electron. For the magnetic spectra, we see that for all L^* and Kp, the wave power is flat for low frequencies but above about 5 mHz the wave power decreases in a power law fashion. For the electric field, we see that the spectra are independent of energy under almost all conditions in addition to being independent of L^*. There seems to be a very weak L^*-dependence at high frequencies, but at lower frequencies the spectra have little or no L^*-dependence.

We used the drift-averaged magnetic and electric field power spectral densities along with the *Fei et al.* [2006] expressions given by Equations (1.18) and (1.17) to compute the magnetic and the

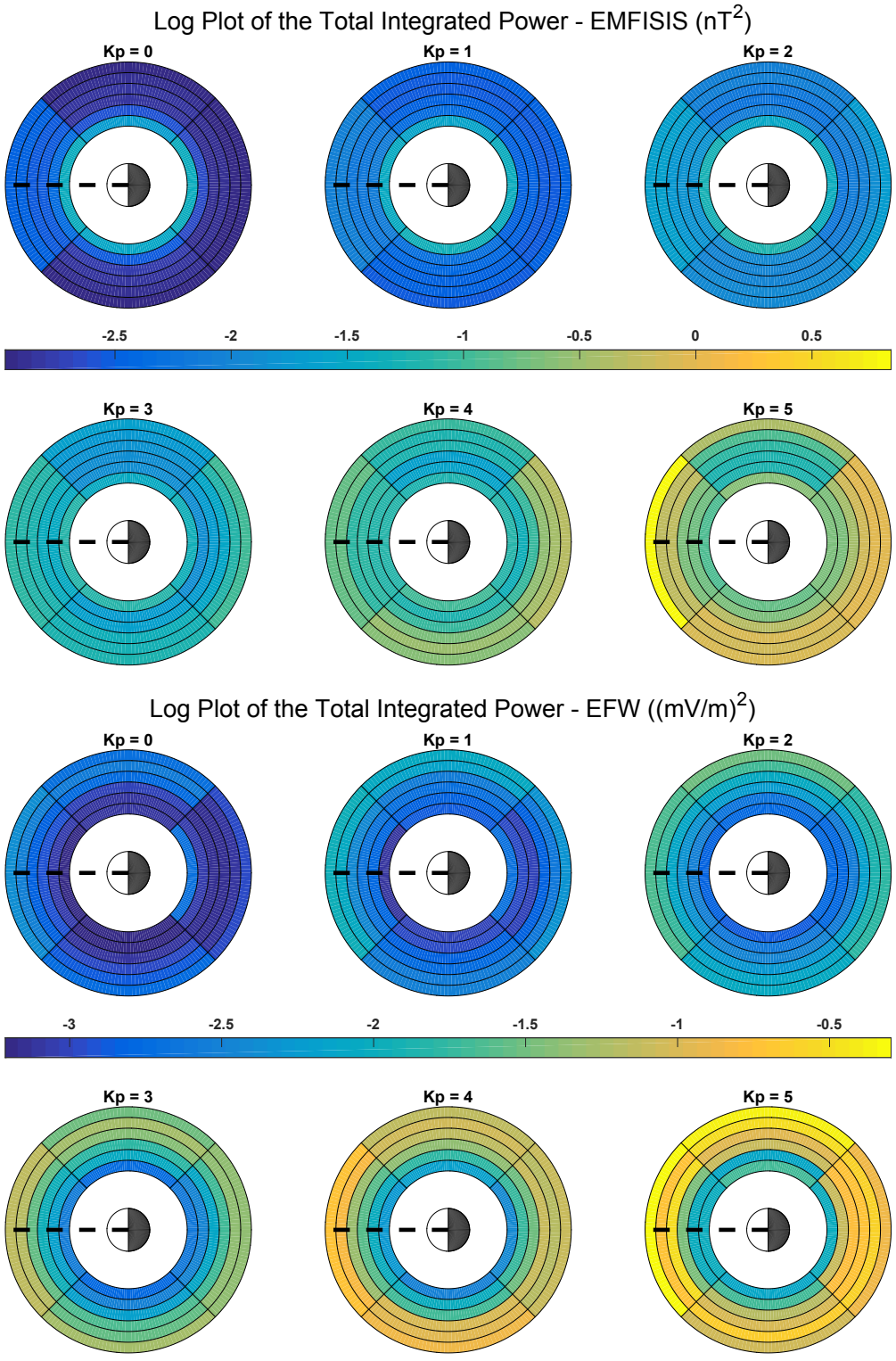


Figure 4.6: Distribution of the total ULF wave power from the magnetic and electric field measurements integrated from $\sim$ 0.8 mHz to $\sim$ 8 mHz. The dashed line is the Sun-Earth line.

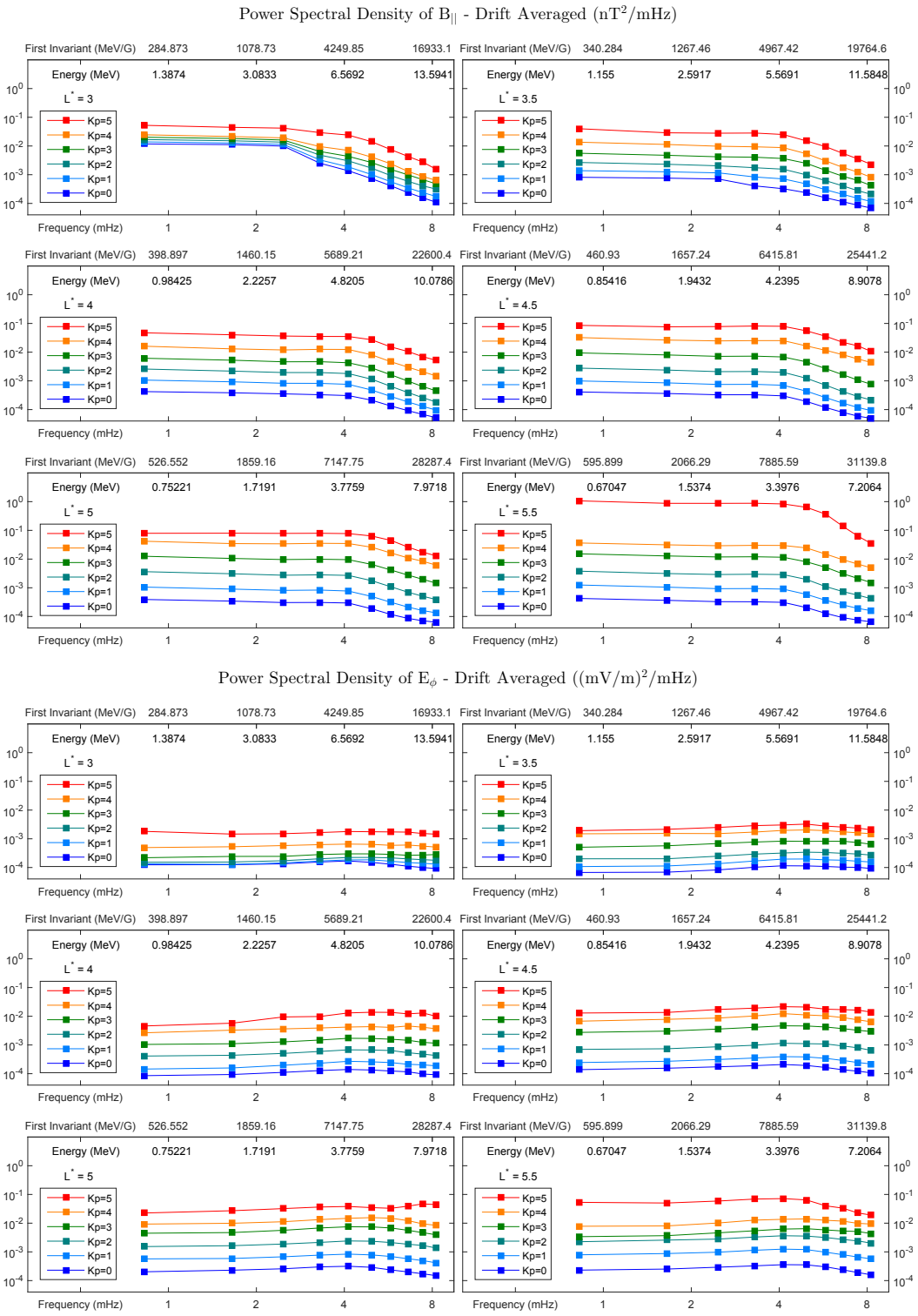


Figure 4.7: Drift-averaged spectra from the magnetic and electric field measurements as a function of electron drift frequency separated by L^* and Kp. The electron energies and relativistically corrected first invariant values are computed assuming $\omega = \omega_d$.

electric components of the radial diffusion coefficient, D_{LL}^{B}[RBSP] and D_{LL}^{E}[RBSP]. Figure 4.8 shows the magnetic and the electric components of the radial diffusion coefficient as a function of frequency, separated by L^* and Kp. For the magnetic component, there is a clear energy dependence, with the magnetic component, at low energies, increasing with increasing energy. After a very weak maximum at about 4 mHz, the magnetic component then decreases very slowly as energy continues to increase. One thing to note about the electric component is that it is independent of energy for all L^* and Kp. The electric diffusion rates are flat most of the time, with energy dependence being very minute and negligible when it does occur at high energies. A second thing to note here is that the electric component is much larger than the magnetic component, sometimes by two orders of magnitude.

Figure 4.9 quantifies the relative contribution of the electric component to driving radial diffusion of charged particles. Figure 4.9 shows the quantity $100\% \cdot (D_{LL}^{E}/D_{LL}^{Total})$ as a function of energy separated by L^* and Kp. We see that the electric component contributes more to radial diffusion than does the magnetic component. The contribution of the electric component can be as high as 99% which implies that the electric component is two orders of magnitude larger than the magnetic component. The only exception is the $L = 5.5$ case for $Kp = 5$. We must remind the reader here that for $Kp = 5$, the statistics are very poor, especially for the electric field measurements (as can be seen in Figure 4.4). Thus the implication that the magnetic field perturbations are dominant or just as important as the electric field perturbations for high Kp at high L^* may very well be false. Therefore we submit here that, in general, the electric component is dominant over the magnetic component in driving radial diffusion of charged particles. Both of the components should be considered when radial diffusion is used to model the Van Allen radiation belts. Considering Figure 4.9, if one of the components is to be ignored then it should be the magnetic component D_{LL}^{B}. The magnetic component contributes very little most of the time and the approximation,

$$D_{LL}^{Total}[\text{RBSP}] \approx D_{LL}^{E}[\text{RBSP}], \tag{4.2}$$

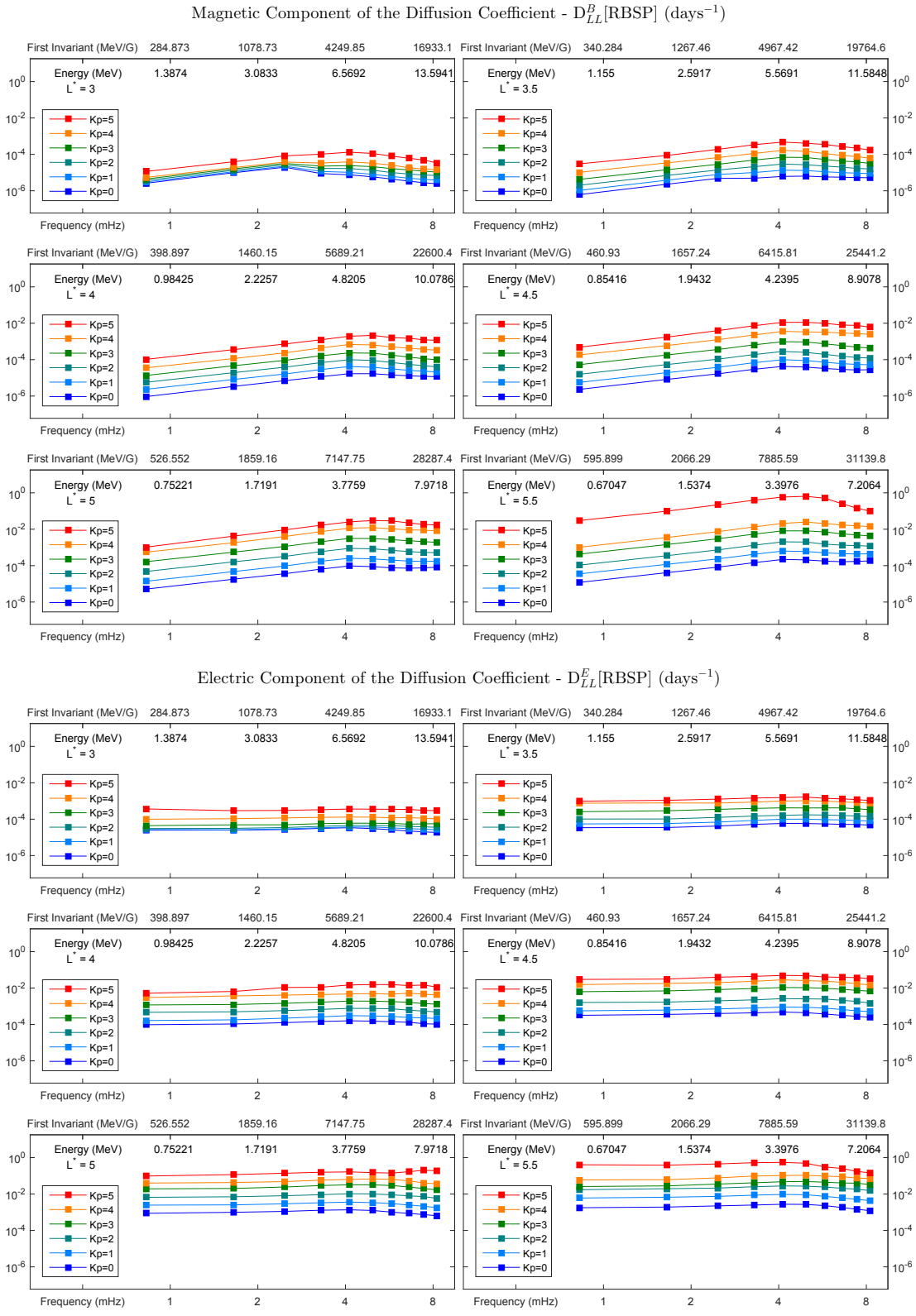


Figure 4.8: The magnetic and electric components of the radial diffusion coefficient in frequency space separated by L^* and Kp.

may be used if an approximation is desired by the modeling community. We would also like to clarify here that since the magnetic component is relatively inconsequential in driving radial diffusion, the total radial diffusion coefficient plots are almost identical to the electric radial diffusion plots. For this reason, with the exception of Figure 4.12, we present no figures showing the total radial diffusion coefficients. If, for example, the reader wishes to see D_{LL}^{Total}[RBSP] as a function of L^* for a fixed first invariant, then the reader is referred to the bottom panel of Figure 4.10. The curves for D_{LL}^{E}[RBSP] are identical to the curves for D_{LL}^{Total}[RBSP].

As Figure 4.10 shows, D_{LL}^{B}[RBSP] and D_{LL}^{E}[RBSP] span several orders of magnitude as a function of L^* even for a fixed M or Kp. Similar to the CRRES study, it is quite difficult to derive a model which will be a good description of both components as a function of L^*, M, and Kp. Therefore, we employed a genetic algorithm to obtain a reasonably simple model which captures all of the desired features from both components. We then performed weighted least squares for both components and derived,

$$D_{LL}^{B}[\text{RBSP}] = \exp(a_1 + b_1 \cdot Kp \cdot L^* + L^*), \tag{4.3}$$

$$D_{LL}^{E}[\text{RBSP}] = \exp(a_2 + b_2 \cdot Kp \cdot L^* + c_2 \cdot L^*), \tag{4.4}$$

where the constants are given by

$$a_1 = -16.253, \quad b_1 = 0.224,$$

$$a_2 = -16.951, \quad b_2 = 0.181, \quad c_2 = 1.982,$$

with L^* and Kp being unitless parameters. Both D_{LL}^{B}[RBSP] and D_{LL}^{B}[RBSP] components are returned in units of days^{-1}.

We would like to emphasize here that neither of these models has any physical basis, and they are presented solely as compact representations of the estimated diffusion coefficients in order to aid the community with its modeling efforts. These models are much more succinct and easier to use than large tables containing numerical values of D_{LL}^{B} and D_{LL}^{E} or dozens of fitted curves. The domain used while fitting both of these models was $3.0 \leq L^* \leq 5.5$ and $0 \leq Kp \leq 5$. The data used to fit both models corresponded to M between 500 MeV/G and 5000 MeV/G. Since both

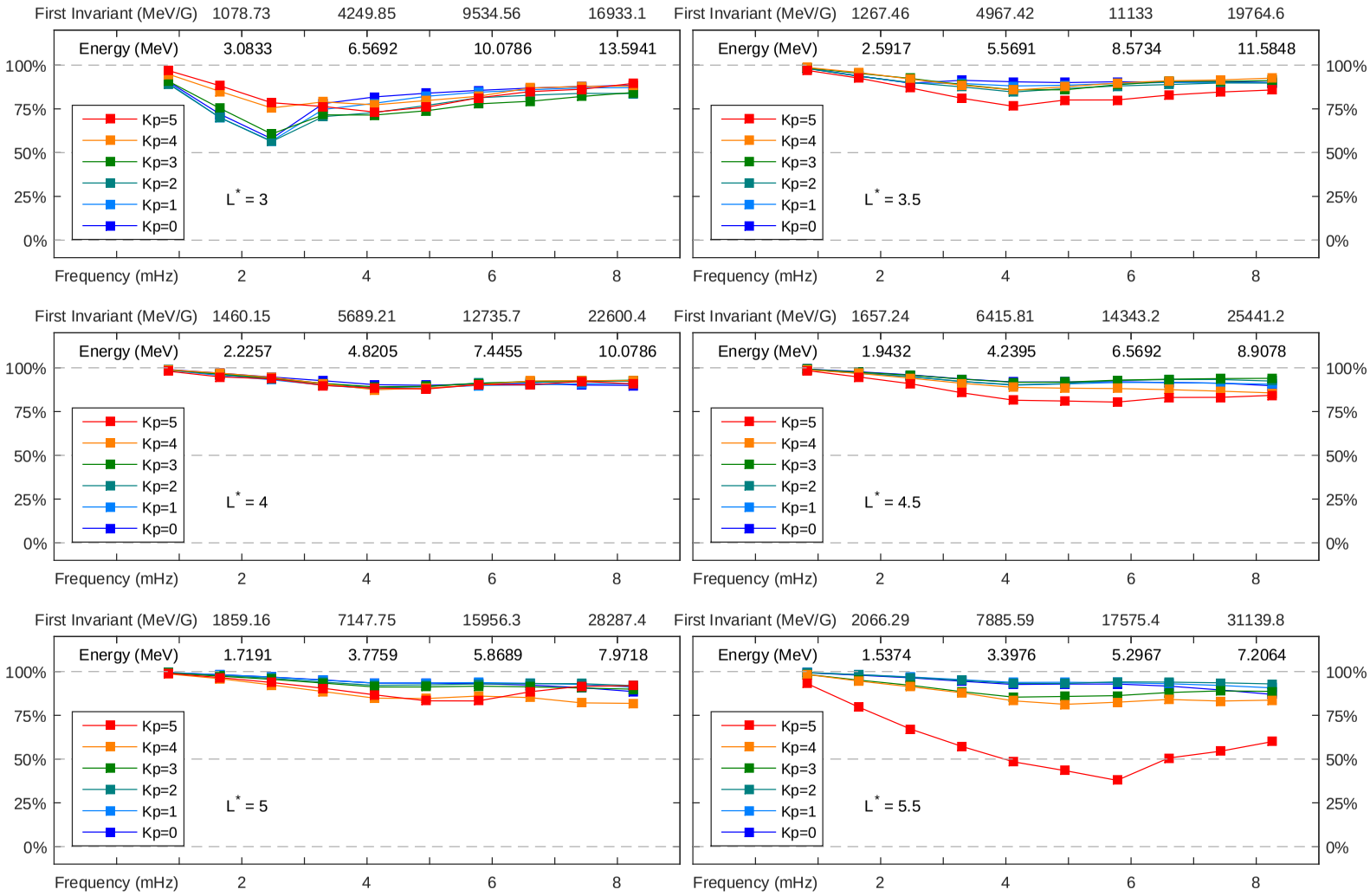


Figure 4.9: The relative contribution of the electric component of the diffusion coefficient to the total diffusion rate. It is clear that electric field perturbations are almost always dominant in driving radial diffusion.

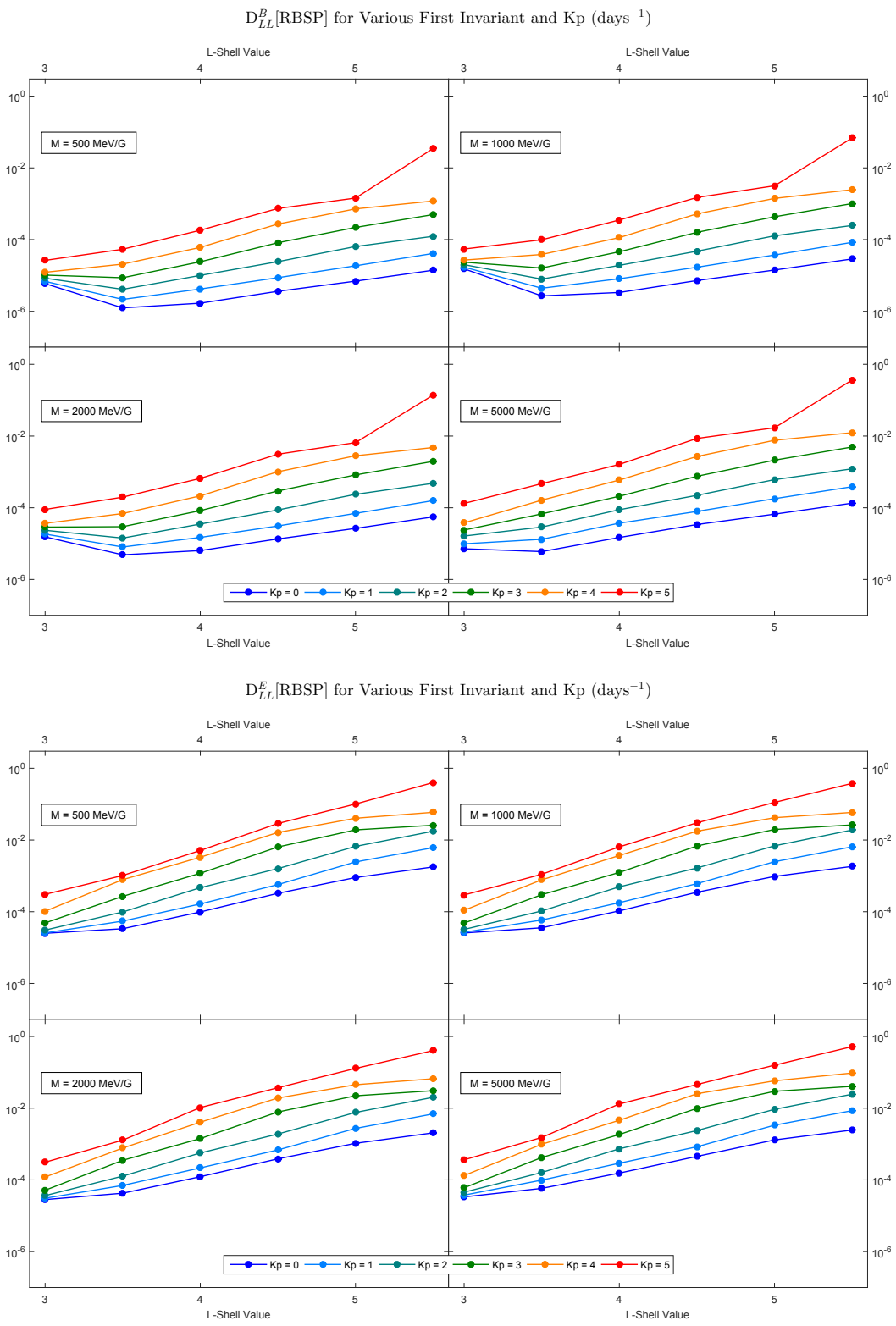


Figure 4.10: The magnetic and electric components of the radial diffusion coefficient for various constant values of the first invariant M as a function of L^* separated by Kp.

components exhibited a very weak energy dependence, models returned by the genetic algorithm reflected this fact. If a model contained an energy term, then the coefficient of that term would be small and the term would explain very little of the variance in the data. The increase in complexity of the model generated by including an energy term was certainly not worth the negligible reduction in error it provided. Therefore, for both components we present models without an M-dependence.

Both models provide excellent results not only for interpolation inside the domain but also for extrapolation outside the domain since the models are continuous in both L^* and Kp. Kp can take any value in $\{0, 0.3, 0.7, 1, ..., 4.3, 4.7, 5\}$ while L^* can vary continuously for $3.0 \leq L^* \leq 5.5$. Furthermore, the trends inside the domain for both parameters, continue in a reasonable manner outside of the domain. For Kp, it is possible to plug in any value in $\{0, 0.3, 0.7, 1, 1.3, ..., 8.3, 8.7, 9\}$ while the limits on L^* are left to the individual reader to decide. A similar warning holds for the parameter M. These models assume that for M between 500 MeV/G and 5000 MeV/G, the data exhibits a very weak M-dependence. It is left up to the individual reader to decide how far outside the M domain this assumption holds. Obviously, if we move far enough away from the boundary of the domain, the diffusion coefficients just simply can no longer remain physical. The reader has the option of just using the nearest neighbor approximation for points outside of the domain. For example, value of $Kp = 5$ can be used for all $Kp > 5$, value of $L^* = 3.0$ can be used for lower L^*-shell values, and so on.

Early research efforts [*Fälthammar*, 1965, 1966a,b, 1968; *Cornwall*, 1968; *Birmingham*, 1969] to estimate the radial diffusion coefficients using magnetic and electric field models resulted in an L^6 dependence of the radial diffusion coefficient. Assuming a symmetric dipole magnetic field and a convective electric field containing wave power at frequencies which are multiples of the drift frequency ($\omega = m\omega_d$), the electrostatic D_{LL}^E was derived to be proportional to L^6. *Fälthammar* [1968] also presented the electromagnetic diffusion coefficient D_{LL}^M proportional to L^{10} for particles with non-relativistic energies in a dipole magnetic field including the contributions from the induced electric field. Therefore, investigating the L-dependence of empirically derived diffusion coefficients is interesting in its own right. Considering Figure 4.10, both components of the diffusion coefficients

show a linear trend with respect to L^* in log-log space but the scaling exponent seems to depends on the level of geomagnetic activity. Therefore we fit both components of the diffusion coefficients as functions of L^* to a simple power law $D_{LL} = D_0 L^n$, separated by Kp. The scaling exponents n are given by,

	Level of Geomagnetic Activity (Kp)					
	0	1	2	3	4	5
D_{LL}^B[RBSP]	5.41	6.50	7.82	9.28	9.23	8.90
D_{LL}^E[RBSP]	7.67	9.13	10.66	10.91	10.87	11.58

where we see that the estimates of n are Kp-dependent and generally increase as Kp increases. Since we assumed the diffusion coefficients to be independent of M, the scaling exponents are also a constant as a function of M. This is in contrast with the conclusions of *Ozeke et al.* [2012] and *Ozeke et al.* [2014], that the magnetic coefficients decrease as M increases.

There are many previously published estimates of diffusion coefficients, but due to differences in models and observations used for diffusion coefficient calculations, a comprehensive comparison is difficult. Differences in methodologies as well as in parameters chosen to parametrize diffusion coefficients may make such comparisons of little value. *Brautigam and Albert* [2000] started with the assumption that the root mean square of the electric field amplitude is a linear function of Kp and then used *Cornwall* [1968] expressions to estimate D_{LL}^E using L, M, and Kp as parameters. *Elkington et al.* [2003] used a compressed dipole field with analytic expressions to drive a test particle simulation at geosynchronous distance. *Brautigam et al.* [2005] assumed a purely dipole field and used the electric field measurements from CRRES to numerically estimate D_{LL}^E using L and Kp as parameters. *Fei et al.* [2006] used a compressed dipole field with analytic expressions to solve the diffusion equation for $2 \leq L \leq 10$ and computed D_{LL}^B assuming a static L-dependence with the scaling exponent being 8.5. *Huang et al.* [2010b] used LFM model runs to conduct particle simulations and estimate D_{LL}^B. *Ozeke et al.* [2012] and *Ozeke et al.* [2014] used *in situ*

measurements from both AMPTE and GOES to compute D_{LL}^{B}, along with ground magnetic field measurements mapped to electric field PSDs in space to compute D_{LL}^{E}. *Tu et al.* [2012] used LFM MHD simulations and observations from GOES and THEMIS to estimate both D_{LL}^{B} and D_{LL}^{E}. *Ali et al.* [2015] used the CRRES magnetometer data to estimate D_{LL}^{B} as a function of L and Kp in direct comparison with the *Brautigam et al.* [2005] study.

Since Kp is a common index of activity used in previous diffusion rates studies, we chose Kp as our activity-dependent parameter. Unfortunately, there has been a dearth of high geomagnetic activity during the Van Allen Probes era to date and there are virtually no measurements for Kp higher than 5 which could be useful to us for this study. Figure 4.11 shows the comparison of our electric and magnetic components of radial diffusion coefficients with some of the previously published estimates. D_{LL}^{B}[RBSP] and D_{LL}^{E}[RBSP] are the diffusion coefficients computed using the Van Allen Probes magnetic and electric field measurements. D_{LL}^{B}[CRRES] were presented in *Ali et al.* [2015] and were computed using the CRRES fluxgate magnetometer data with techniques very similar to the current RBSP study. D_{LL}^{E}[CRRES] estimates are taken from *Brautigam et al.* [2005] which used the electric field measurements from CRRES. We would like to point out here that although *Brautigam et al.* [2005] used *Fälthammar* [1965] diffusion formulation, they made no effort to separate the inductive and the convective component of the electric field measurements, since it is quite difficult to do so with single point measurements from a spacecraft such as CRRES. Therefore, D_{LL}^{E}[CRRES] contains contributions from the total electric field. D_{LL}^{EM}[BA] and D_{LL}^{ES}[BA] are the electromagnetic and electrostatic components from *Brautigam and Albert* [2000]. D_{LL}^{B}[Ozeke] and D_{LL}^{E}[Ozeke] are taken from *Ozeke et al.* [2012] and *Ozeke et al.* [2014].

Figure 4.12 shows the total diffusion coefficients obtained by adding up the two individual components. The D_{LL}^{Total}[RBSP] and D_{LL}^{Total}[BA] diffusion components can just be added as they are, because both components have identical L domains; no extrapolation or truncation is needed. For D_{LL}^{Total}[CRRES], the electric component had to be truncated to $4 \leq L \leq 6.5$ before the magnetic component could be added to it. For D_{LL}^{Total}[Ozeke], the magnetic component had to be interpolated and truncated while the electric component had to be truncated so that the total diffusion coefficient

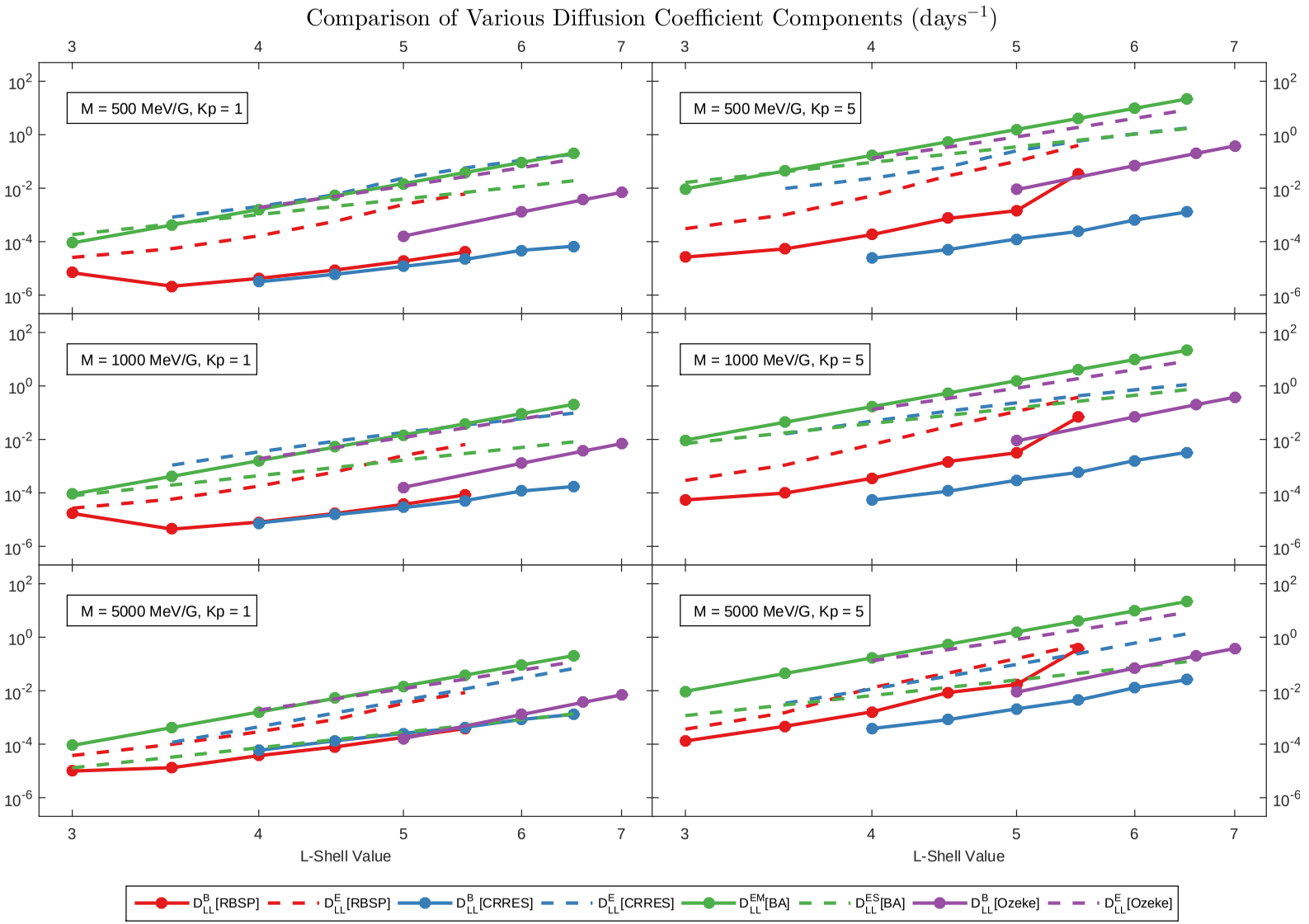


Figure 4.11: A comparison of various radial diffusion rates. D^B_{LL}[RBSP] and D^E_{LL}[RBSP] are presented here with D^B_{LL}[CRRES] from *Ali et al.* [2015] and D^E_{LL}[CRRES] from *Brautigam et al.* [2005] in addition to *Brautigam and Albert* [2000] and *Ozeke et al.* [2012, 2014] estimates.

can be obtained for $5 \leq L \leq 6.5$. The total diffusion coefficient is perhaps better for comparison between different estimates because of the different formulations used. The *Fälthammar* [1965, 1966a,b, 1968] formulation divides the total radial diffusion coefficient into its electrostatic and electromagnetic components, while the *Fei et al.* [2006] formulation partitions it into its magnetic and electric components. Since the electromagnetic component contains contributions from the magnetic field as well as the inductive electric field, the electromagnetic component will tend to be larger than the purely magnetic component and the electrostatic component will tend to be smaller than the total electric component. Adding up the components before comparison alleviates this problem. The total diffusion coefficient includes contributions from the magnetic field, convective electric field, and the inductive electric field once and only once. Hence, the comparison becomes more meaningful in a certain sense.

The shaded region in Figure 4.12 represents the variability in the D_{LL}^{Total}[RBSP] estimates. In Figure 4.5, where we presented all of the magnetic and electric field spectra in a given bin, we saw that the range of the spectra can be three to four orders of magnitude even if particular values of L^* and Kp are fixed. We picked the median power spectral density as an appropriate indicator of central tendency, but it is also important to note the large degree of variability present in each bin. Therefore, in addition to the median, we took the data at the 5th and the 95th percentile from each bin and estimated the upper bound and the lower bound on both the D_{LL}^{B}[RBSP] and D_{LL}^{E}[RBSP] estimates. These bounds on both of the components were then combined to estimate the uncertainty in D_{LL}^{Total}[RBSP]. The variability at each step of the calculation was estimated using standard error propagation formulas [*Squires*, 2001].

The distribution of the spectra within a bin appears to be a log-normal distribution although we cannot verify this in a statistically significant way. Attempts to use a statistical test to determine if the spectra are distributed log-normally resulted in rejection of the null hypothesis in about half of the cases, and we therefore could not make a quantitative statement in general about all of the bins. Some of the bins are clearly deficient because there simply aren't enough spectra in them to guess their probability distribution function. But for most bins, the distribution in log

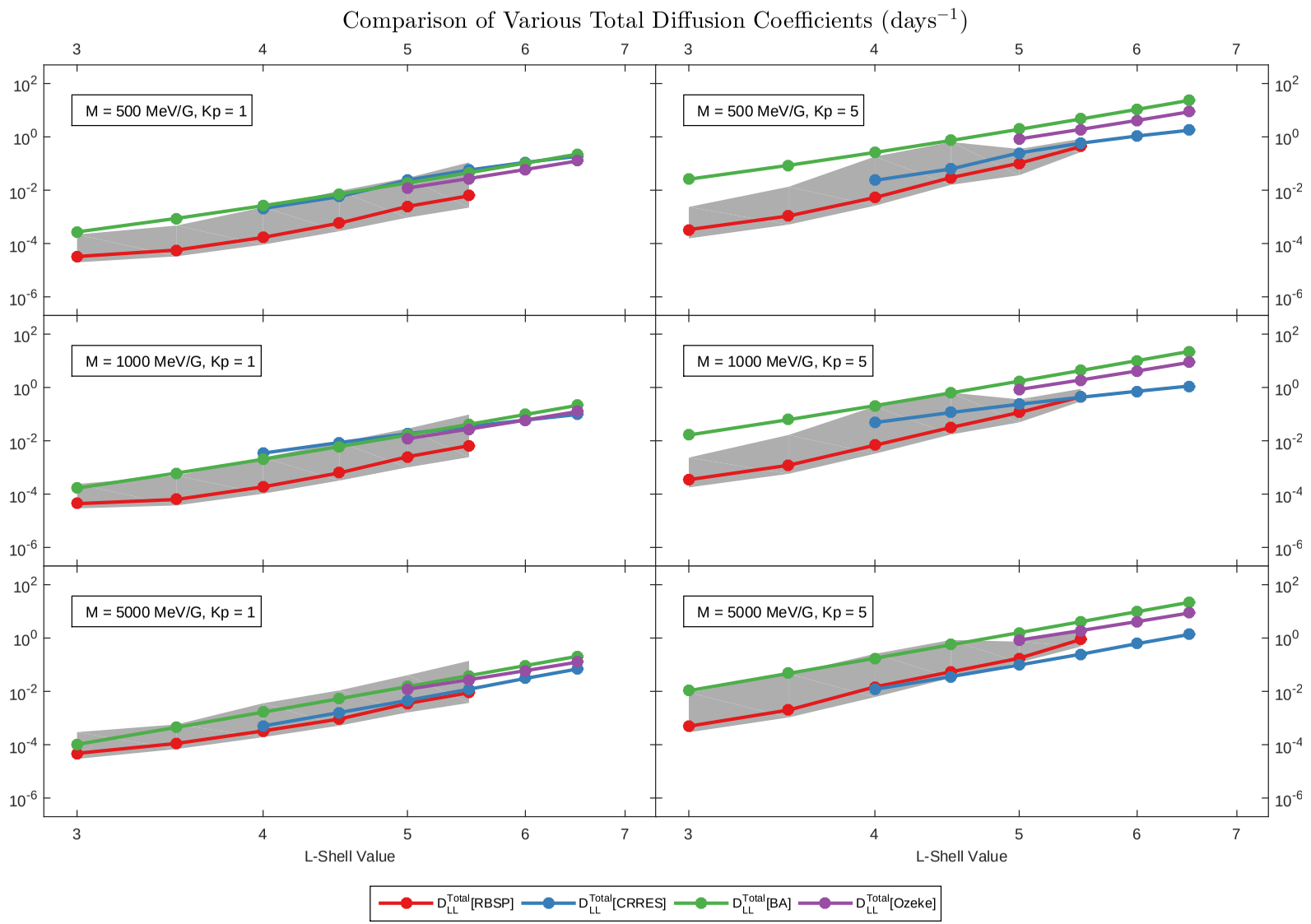


Figure 4.12: A comparison of various total radial diffusion rates. The shaded area is the region between the 5th and the 95th percentile for the D_{LL}^{Total}[RBSP] estimates.

space appears to be symmetric and bell-shaped. This implies that in linear space the mass of the distribution is concentrated on the left, near zero, with very small magnitudes which is indeed the case. The tail on the right is long, and contains some spectra with extremely large magnitudes. This results in the 5th percentile being very close to the median while the 95th percentile can be an order magnitude larger than the median. The effect clearly propagates all the way through to Figure 4.12, where we see the upper limit on D_{LL}^{Total}[RBSP] being rather large while the lower limit is barely visible. We would like to point out here that all of the data processing and the statistical analysis presented in this chapter was done in linear space. The data is only plotted in log space, when necessary, to make some of the relationships clear. Therefore in Figure 4.5, the mean, the geometric mean, and the median with the interquartile range were computed in linear space, which is why the mean is at the 75th percentile. This is precisely the reason why we chose the median as the measure of central tendency, rather than the arithmetic mean. The arithmetic mean is a very inadequate measure of the location for this distribution.

Comparing our estimates with other previously published results in Figure 4.11, we see that for quiet times the magnetic components from RBSP, CRRES, and *Ozeke et al.* [2012, 2014] are in good agreement with the agreement getting better as energy increases. For higher Kp, the agreement between RBSP and *Ozeke et al.* [2012, 2014] is better than with CRRES. One trend which we see in general is that the magnetic component is always smaller than the electric component lending credence to the fact that the perturbations in the electric field are more important than the perturbations in the magnetic field in driving radial diffusion of charged particles in the inner magnetosphere. This has important consequences for Van Allen radiation belt modelers [*Shprits et al.*, 2005; *Varotsou et al.*, 2008; *Albert et al.*, 2009; *Tu et al.*, 2009; *Chu et al.*, 2010; *Su et al.*, 2010, 2011a,b; *Subbotin et al.*, 2011a,b; *Kim et al.*, 2012; *Tu et al.*, 2013; *Li et al.*, 2014] who assume that the magnetic diffusion coefficients are much larger than the electric diffusion coefficients. Recent efforts by *Ozeke et al.* [2012, 2014] and *Ali et al.* [2015] along with the present study suggest that the reverse is true. The electric diffusion coefficient is dominant in driving radial diffusion sometimes by two orders of magnitude. Comparing the total diffusion coefficients from

Figure 4.12, we see that D_{LL}^{Total}[RBSP] is consistently smaller than the *Brautigam and Albert* [2000] estimates. For lower energies and for quiet times, the D_{LL}^{Total}[CRRES] and D_{LL}^{Total}[Ozeke] estimates agree quite well with the *Brautigam and Albert* [2000] estimates. But as energy and the level of geomagnetic activity increase, the D_{LL}^{Total}[CRRES] and D_{LL}^{Total}[Ozeke] estimates approach the D_{LL}^{Total}[RBSP] estimates. Note that the D_{LL}^{Total}[CRRES] and the D_{LL}^{Total}[Ozeke] estimates are always within the D_{LL}^{Total}[RBSP] error bounds albeit occasionally being at or near the upper boundary. On the other hand, the difference between D_{LL}^{Total}[RBSP] and D_{LL}^{Total}[BA] is simply too large at times, such that the *Brautigam and Albert* [2000] estimates are not within the error estimates of D_{LL}^{Total}[RBSP].

Chapter 5

Summary and Discussion

Understanding the relationship between the electric and magnetic diffusion coefficients is essential for understanding and designing ULF wave radial diffusion models and simulations. These will in turn help us see the larger picture of complex mechanisms in the radiation belts including the relative roles of transport and local heating. We used the CRRES and RBSP data to estimate the ULF wave power present in both the magnetic and the electric fields in the inner magnetosphere. The wave power was estimated in the compressional component of the magnetic field, which represents the oscillations in the geomagnetic field in the direction of the magnetic field, and the azimuthal component of the electric field which is the tangential direction of the electric field in the equatorial plane. We parametrized the ULF wave power as a function of L, Kp, and MLT to study the distribution of wave power. Note here that L represents the McIlwain parameter for the CRRES study while $L = L^*$ was the *Roederer and Zhang* [2014] parameter for the RBSP study. Figures 3.4 and 4.6 show clearly that ULF wave power is not uniform in azimuth, even though both the *Fälthammar* [1965] and the *Fei et al.* [2006] analytic treatments of radial diffusion assume a uniform distribution of power. The noon sector contains the highest amount of wave power while the other three sectors show no significant differences in all cases. Clearly, this warrants more attention and research from the space physics community.

We estimated the drift-averaged power spectral densities as a function L and Kp, since that is what the *Fei et al.* [2006] formulation requires for Equations (1.17) and (1.18). In order to do this, we took the weighted average of the PSDs, with the weights being inversely proportional to

the number of data segments in a given bin. In order to find the average value of a continuous function over a given interval, the function can be Riemann integrated and then divided by the length of the interval. Since the PSDs are a function of magnetic local time, computing the average of the PSDs is equivalent to numerical integration using four nodes, and is a reasonable estimate of the true drift-averaged PSD. We don't expect the violation of the uniform power in azimuth assumption to have a serious impact on our diffusion estimates.

Another assumption that we made is that the total measured electric and the magnetic field phases are independent of each other. *Perry et al.* [2005, Figure 1] shows that E_ϕ and $\partial B_\theta/\partial t$ have phases that are negatively correlated. *Perry et al.* [2005] concludes that this phase relation results in a reduced rate of radial diffusion. The main goal of these studies was to quantify the relative contribution of magnetic and electric field perturbations to radial diffusion of charged particles. Assuming no phase relation results in our magnetic estimates to be an overestimate of the true magnetic diffusion component. But since the electric component is so much larger than the magnetic component, assuming phase independence does not alter our main conclusion that the magnetic field perturbations contribute much less than the electric field perturbations in driving radial diffusion. We also assumed that all of the observed wave power is in the first, $m = 1$ mode. This assumption was made explicitly by *Brautigam et al.* [2005] and is entirely reasonable in light of the conclusions presented by *Elkington et al.* [2012] and *Tu et al.* [2012]. The effects of this assumption being violated are unclear because they depend on the true distribution of power in frequency and the mode number m. Another condition that we assumed is symmetric resonance $\omega = m\omega_d$. *Elkington et al.* [2003] studied the effects of asymmetric resonances $\omega = (m \pm 1)\omega_d$, induced by the noon-midnight asymmetry in the magnetosphere, on the radial diffusion coefficients. All three resonances considered simultaneously do increase the diffusion rates, but only by about a factor of two, and therefore has no effect on the main conclusions of this thesis.

Analysis of our results as well as comparison with other recently published studies shows consistently that the magnetic component is much smaller than the electric component, often by one to two orders of magnitude. This has major implications, for example, in diffusion simulation

codes which cite the conclusions of *Brautigam and Albert* [2000] and may underestimate the role of D_{LL}^{E} in driving radial diffusion [*Shprits et al.*, 2005; *Varotsou et al.*, 2008; *Albert et al.*, 2009; *Tu et al.*, 2009; *Chu et al.*, 2010; *Su et al.*, 2010, 2011a,b; *Subbotin et al.*, 2011a,b; *Kim et al.*, 2012; *Tu et al.*, 2013; *Li et al.*, 2014]. We observed very clear effects of Kp in driving diffusion. Higher levels of Kp activity drive faster diffusion of particles. This conclusion is in agreement with *Ozeke et al.* [2012, 2014] and *Tu et al.* [2012]. In contrast to these studies, we do observe the magnetic field PSDs to be weakly L-dependent, with magnetic field PSDs increasing slightly as L increases, while the electric field PSDs seem to be independent of L. Both of the magnetic diffusion coefficients have a very weak energy dependence, with the CRRES magnetic coefficients exhibiting slightly more energy dependence than the RBSP magnetic coefficients. The magnetic coefficients increase slightly as a function of energy for small energies, and then stay mostly flat after a small peak at $\sim$5 mHz. The RBSP electric coefficients as a function of energy are practically flat.

Figure 4.9 makes it clear that the electric radial diffusion component is always dominant over the magnetic radial diffusion component, sometimes by one to two orders of magnitude. This seems to hold independent of the radial distance (L^*) or the level of geomagnetic activity (Kp). Comparing the individual components in Figure 4.11, we see that for low M, our magnetic coefficients are consistently smaller than the estimates of *Ozeke et al.* [2012] and *Ozeke et al.* [2014], but as M increases our magnetic coefficients attain an excellent agreement with those of *Ozeke et al.* [2012] and *Ozeke et al.* [2014]. This also holds as Kp increases. The *Ozeke et al.* [2012] electric components are consistently higher than our RBSP electric components. We have also used the power spectral densities at the 5th and the 95th percentile in each bin to provide the error estimates in Figure 4.12. The RBSP total diffusion rates are smaller than the other three total diffusion rate estimates. The CRRES total diffusion rates do approach the RBSP total coefficients as energy and Kp increase. The total estimates of *Brautigam and Albert* [2000] are so much larger than the RBSP total estimates that they are sometimes larger than the upper error bounds, even if other estimates are within the error bounds.

The L-dependence of the radial diffusion coefficient components was investigated, as there

are conflicting answers in previously published literature. No doubt this is due to differences in conditions under which the data was collected, the data collection method, data processing methods, and numerical methods used. *Huang et al.* [2010b, see Table 1] provides a very concise summary of various estimates of the L-dependence for various L ranges and particle energies. We estimated the scaling exponent after the data was separated by Kp, as it was apparent that the scaling rates increase as Kp increases. This was indeed the case. The scaling exponents were well separated in Kp.

Many diffusion models and estimates have been published in trying to understand the role of electromagnetic waves in driving diffusion of charged particles trapped in the inner magnetosphere. CRRES and the Van Allen Probes are two of the missions dedicated to the study of the Van Allen radiation belts. The Van Allen Probes are still in operation and are equipped with state of the art instrument suites, providing an ideal source of high-quality measurements. Even though we seek to understand global oscillations in the electric and the magnetic fields, we used single point measurements from CRRES and the Van Allen Probes in the spirit of *Brautigam and Albert* [2000]; *Brautigam et al.* [2005]; *Ozeke et al.* [2012, 2014] to estimate the radial diffusion rates. We conclude that contrary to previously held notions, the magnetic component contributes very little to radial diffusion of charged particles in the Earth's inner magnetosphere. Both components should be considered, but if one component is to be ignored then it should be the magnetic component, rather than the electric component. This should be considered carefully when modelers design simulations to understand the dynamic evolution of the radiation belts.

Bibliography

Abramowitz, M., and I. A. Stegun (1965), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed., Dover, New York.

Albert, J. M., N. P. Meredith, and R. B. Horne (2009), Three-dimensional diffusion simulation of outer radiation belt electrons during the 9 October 1990 magnetic storm, *J. Geophys. Res.*, *114*(A09214), doi:10.1029/2009JA014336.

Ali, A. F., S. R. Elkington, W. Tu, L. G. Ozeke, A. A. Chan, and R. H. W. Friedel (2015), Magnetic field power spectra and magnetic radial diffusion coefficients using CRRES magnetometer data, *J. Geophys. Rev. Space Physics*, *120*, doi:10.1002/2014JA020419.

Aloupis, G. (2006), Geometric measures of data depth, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, *72*, 147–158.

Arthur, C. W., R. L. McPherron, L. J. Lanzerotti, and D. C. Webb (1978), Geomagnetic field fluctuations at synchronous orbit: 1. Power spectra, *J. Geophys. Res.*, *83*, 3859–3865, doi:10.1029/JA083iA08p03859.

Baker, D. N. (2001), Satellite anomalies due to space storms, in *Space Storms and Space Weather Hazards*, *NATO Sci. Ser.*, vol. 38, edited by I. A. Daglis, pp. 285–311, Kluwer Acad., Dordrecht, Netherlands, doi:10.1007/978-94-010-0983-6_11.

Baker, D. N., S. G. Kanekal, V. C. Hoxie, M. G. Henderson, X. Li, H. E. Spence, S. R. Elkington, R. H. W. Friedel, J. Goldstein, M. K. Hudson, G. D. Reeves, R. M. Thorne, C. A. Kletznig, and S. G. Claudepierre (2013), A long-lived relativistic electron storage ring embedded in Earth's outer Van Allen belt, *Science*, *340*, 186–190, doi:10.1126/science.1233518.

Baker, D. N., A. N. Jaynes, V. C. Hoxie, R. M. Thorne, J. C. Foster, X. Li, J. F. Fennell, J. R. Wygant, S. G. Kanekal, P. J. Erickson, W. Kurth, W. Li, Q. Ma, Q. Schiller, L. Blum, D. M. Malaspina, A. Gerrard, and L. J. Lanzerotti (2014), An impenetrable barrier to ultrarelativistic electrons in the Van Allen radiation belts, *Nature*, *515*, 531–534, doi:10.1038/nature13956.

Barnett, V. (1976), The ordering of multivariate data, *Journal of the Royal Statistical Society*, *A*(139), 318–355.

Baumjohann, W., and R. A. Treumann (1997), *Basic Space Plasma Physics*, Imperial College Press.

Birmingham, T. J. (1969), Convection electric fields and the diffusion of trapped magnetospheric radiation, *Journal of Geophysical Research*, *74*(9), 2169–2181, doi:10.1029/JA074i009p02169.

Brautigam, D. H., and J. M. Albert (2000), Radial diffusion analysis of outer radiation belt electrons during the October 9, 1990 magnetic storm, *J. Geophys. Res.*, *105*(A1), 291–309, doi: 10.1029/1999JA900344.

Brautigam, D. H., G. P. Ginet, J. M. Albert, J. R. Wygant, D. R. Rowl, A. Ling, and J. Bass (2005), CRRES electric field power spectra and radial diffusion coefficients, *J. Geophys. Res.*, *110*(A02214), doi:10.1029/2004JA010612.

Cahill, L. J., and J. R. Winckler (1992), Periodic magnetopause oscillations observed with the GOES satellites on March 24, 1991, *J. Geophys. Res.*, *97*(A6), 8239, doi:10.1029/92JA00433.

Chen, L., and A. Hasegawa (1974), A theory of long-period magnetic pulsations 1. Steady state excitation of field line resonance, *J. Geophys. Res.*, *79*(7), 1024, doi:10.1029/JA079i007p01024.

Chu, F., M. K. Hudson, P. Haines, and Y. Shprits (2010), Dynamic modeling of radiation belt electrons by radial diffusion simulation for a 2 month interval following the 24 March 1991 storm injection, *J. Geophys. Res.*, *115*(A03210), doi:10.1029/2009JA014409.

Claudepierre, S. (2008), Solar wind driving of magnetospheric ultra-low frequency pulsations, Ph.D. thesis, University of Colorado at Boulder, Boulder, Colorado, USA.

Claudepierre, S. G., S. R. Elkington, and M. Wiltberger (2008), Solar wind driving of magnetospheric ULF waves: Pulsations driven by velocity shear at the magnetopause, *J. Geophys. Res.*, *113*(A05218), doi:10.1029/2007JA012890.

Conway, J. H., and R. K. Guy (1996), *The Book of Numbers*, Springer-Verlag, New York.

Cornwall, J. M. (1968), Diffusion processes influenced by conjugate-point wave phenomena, *Radio Sci.*, *3*(7), 740–744.

Cramér, H. (1946), *Mathematical Methods of Statistics*, Princeton University Press, Princeton.

Cravens, T. E. (2004), *Physics of Solar System Plasmas*, Cambridge Atmospheric and Space Science Series, Cambridge University Press.

Dietz, E. J. (1987), A comparison of robust estimators in simple linear regression, *Communications in Statistics: Simulation and Computation*, *16*, 1209–1227.

Dietz, E. J. (1989), Teaching regression in a nonparametric statistics course, *American Statistician*, *43*, 35–40.

Donoho, D. (1982), Breakdown Properties of Multivariate Location Estimators, Ph.D. thesis, Harvard University, Boston, MA, Qualifying Paper.

Donoho, D., and M. Gasko (1992), Breakdown properties of location estimates based on half-space depth and projected outlyingness, *Annals of Statistics*, *20*, 1803–1827.

Draper, N. R., and H. Smith (1966), *Applied Regression Analysis*, John Wiley & Sons, New York, USA.

Elkington, S. R., M. K. Hudson, and A. Chan (1999), Acceleration of relativistic electrons via drift-resonant interaction with toroidal-mode Pc-5 ULF oscillations, *Geophys. Res. Lett.*, *26*(21), 3273–3276, doi:10.1029/1999GL003659.

Elkington, S. R., M. K. Hudson, M. J. Wiltberger, and J. G. Lyon (2002), MHD/particle simulations of radiation belt dynamics, *J. Atmos. and Sol. Terr. Phys.*, *64*, 607–615, doi:10.1016/S1364-6826(02)00018-4.

Elkington, S. R., M. K. Hudson, and A. A. Chan (2003), Resonant acceleration and diffusion of outer zone electrons in an asymmetric geomagnetic field, *J. Geophys. Res.*, *108*(A3), 1116, doi:10.1029/2001JA009202.

Elkington, S. R., M. Wiltberger, A. A. Chan, and D. N. Baker (2004), Physical models of the geospace radiation environment, *J. Atmos. and Sol. Terr. Phys.*, *66*, 1371–1387, doi:10.1016/j.jastp.2004.03.023.

Elkington, S. R., A. A. Chan, and M. Wiltberger (2012), Global structure of ULF waves during the 24-26 September 1998 geomagnetic storm, *Geophys. Monogr. Ser.*, pp. 127–138, doi:10.1029/2012GM001348.

Fälthammar, C.-G. (1965), Effects of time-dependent electric fields on geomagnetically trapped radiation, *J. Geophys. Res.*, *70*(11), 2503–2516, doi:10.1029/JZ070i011p02503.

Fälthammar, C.-G. (1966a), On the transport of trapped particles in the outer magnetosphere, *J. Geophys. Res.*, *71*(5), 1487–1491, doi:10.1029/JZ071i005p01487.

Fälthammar, C.-G. (1966b), Coefficients of diffusion in the outer radiation belts, in *Radiation Trapped in the Earth's Magnetic Field*, edited by B. M. McCormac, p. 398, NATO Advanced Study Institute.

Fälthammar, C.-G. (1968), Radial diffusion by violation of the third adiabatic invariant, in *Earth's Particles and Fields*, edited by B. M. McCormac, p. 157, NATO Advanced Study Institute, Van Nostrand Reinhold Publ. Co., New York.

Fei, Y., A. A. Chan, S. R. Elkington, and M. J. Wiltberger (2006), Radial diffusion and MHD particle simulations of relativistic electron transport by ULF waves in the September 1998 storm, *J. Geophys. Res.*, *111*(A12209), doi:10.1029/2005JA011211.

Frank, L. A. (1965), Inward radial diffusion of electrons of greater than 1.6 million electron volts in the outer radiation zone, *J. Geophys. Res.*, *70*, 3533–3539, doi:10.1029/JZ070i015p03533.

Friedel, R. H. W., G. D. Reeves, and T. Obara (2002), Relativistic electron dynamics in the inner magnetosphere - a review, *J. Atmos. and Solar-Terrestrial Phys.*, *64*, 265–282, doi:10.1016/S1364-6826(01)00088-8.

Gauss, C. F. (1823), *Theoria combinationis observationum erroribus minimis obnoxiae*, *Werke*, vol. 4, H. Dieterich, Göttingen.

Giles, B. L., M. A. McCook, M. W. McCook, and G. P. Miller (1995), CRRES Combined Radiation and Release Effects Satellite Program, *NASA Technical Memorandum 108494*, National Aeronautics and Space Administration, Marshall Space Flight Center, Alabama.

Glassmeier, K., C. Othmer, R. Cramm, M. Stellmacher, and M. Engebretson (1999), Magnetospheric field line resonances: A comparative planetology approach, *Surv. Geophys.*, *20*, 61, doi:10.1023/A:1006659717963.

Green, J. C., and M. G. Kivelson (2001), A tale of two theories: How the adiabatic response and ULF waves affect relativistic electrons, *J. Geophys. Res.*, *106*(A11), 25,777–25,791, doi: 10.1029/2001JA000054.

Hald, A. (1952), *Statistical Theory with Engineering Applications*, John Wiley & Sons Inc., New York.

Hartinger, M. D., D. L. Turner, F. Plaschke, V. Angelopoulos, and H. Singer (2013), The role of transient ion foreshock phenomena in driving Pc5 ULF wave activity, *J. Geophys. Res. Space Physics*, *118*, 299–312, doi:10.1029/2012JA018349.

Hasegawa, A. (1969), Drift mirror instability in the magnetosphere, *Phys. Fluids*, *12*, 2642.

Hastie, T., R. Tibshirani, and J. Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Series in Statistics, 2nd ed., Springer, New York, doi:10.1007/978-0-387-84858-7.

Holzworth, R. H., and F. S. Mozer (1979), Direct evaluation of the radial diffusion coefficient near L = 6 due to electric field fluctuations, *J. Geophys. Res.*, *84*(A6), 2559–2566, doi: 10.1029/JA084iA06p02559.

Horne, R. B., and R. M. Thorne (1998), Potential waves for relativistic electron scattering and stochastic acceleration during magnetic storms, *Geophys. Res. Lett.*, *25*(15), 1231, doi: 10.1029/98GL01002.

Horne, R. B., R. M. Thorne, Y. Y. Shprits, N. P. Meredith, S. A. Glauert, A. J. Smith, S. G. Kanekal, D. N. Baker, M. J. Engebretson, J. L. Posch, M. Spasojevic, U. S. Inan, J. S. Pickett, and P. M. E. Decreau (2005), Wave acceleration of electrons in the Van Allen radiation belts, *Nature*, *437*, 227–230, doi:10.1038/nature03939.

Huang, C.-L., H. E. Spence, H. J. Singer, and W. J. Hughes (2010a), Modeling radiation belt radial diffusion in ULF wave fields: 1. Quantifying ULF wave power at geosynchronous orbit in observations and in global MHD model, *J. Geophys. Res.*, *115*(A06215), doi:10.1029/2009JA014917.

Huang, C.-L., H. E. Spence, M. K. Hudson, and S. R. Elkington (2010b), Modeling radiation belt radial diffusion in ULF wave fields: 2. Estimating rates of radial diffusion using combined MHD and particle codes, *J. Geophys. Res.*, *115*(A06216), doi:10.1029/2009JA014918.

Hudson, M. K. (2013), Space physics: A fast lane in the magnetosphere, *Nature*, pp. 383–384, doi:10.1038/504383a.

Hudson, M. K., R. E. Denton, M. R. Lessard, E. G. Miftakhova, and R. R. Anderson (2004), A study of Pc-5 ULF oscillations, *Annales Geophysicae*, *22*, 289–302, doi:10.5194/angeo-22-289-2004.

Jacobs, J. A., Y. Kato, S. Matsushita, and V. A. Troitskaya (1964), Classification of geomagnetic micropulsations, *J. Geophys. Res.*, *69*, 180, doi:10.1029/JZ069i001p00180.

Jensen, J. L. W. V. (1906), Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Mathematica*, *30*(1), 175–193, doi:10.1007/BF02418571.

Kelley, C. T. (1999), *Iterative Methods for Optimization*, 18, SIAM Frontiers in Applied Mathematics, New York.

Kim, K., Y. Shprits, D. Subbotin, and B. Yi (2012), Relativistic radiation belt electron responses to GEM magnetic storms: Comparison of CRRES observations with 3-D VERB simulations, *J. Geophys. Res.*, *117*(A08221), doi:10.1029/2011JA017460.

Kivelson, M. G., and C. T. Russell (Eds.) (1995), *Introduction to Space Physics*, Cambridge University Press.

Kivelson, M. G., and D. J. Southwood (1988), Hydromagnetic waves in the ionosphere, *Geophys. Res. Lett.*, *15*, 1271, doi:10.1029/GL015i011p01271.

Kletzing, C., W. Kurth, M. Acuna, R. MacDowall, R. Torbert, T. Averkamp, D. Bodet, S. Bounds, M. Chutter, J. Connerney, D. Crawford, J. Dolan, R. Dvorsky, G. Hospodarsky, J. Howard, V. Jordanova, R. Johnson, D. Kirchner, B. Mokrzycki, G. Needell, J. Odom, D. Mark, J. Pfaff, R., J. Phillips, C. Piker, S. Remington, D. Rowland, O. Santolik, R. Schnurr, D. Sheppard, C. Smith, R. Thorne, and J. Tyler (2013), The Electric and Magnetic Field Instrument Suite and Integrated Science (EMFISIS) on RBSP, *Space Science Reviews*, *179*(1-4), 127–181, doi:10.1007/s11214-013-9993-6.

Lanzerotti, I. J., and C. G. Morgan (1973), ULF geomagnetic power near $L = 4$: 2. Temporal variations of the radial diffusion coefficients for relativistic electrons, *J. Geophys. Res.*, *78*, 4600–4610, doi:10.1029/JA078i022p04600.

Lanzerotti, I. J., D. C. Webb, and C. W. Arthur (1978), Geomagnetic field fluctuations at synchronous orbit: 2. Radial diffusion, *J. Geophys. Res.*, *83*, 3866–3870, doi:10.1029/JA083iA08p03866.

Lanzerotti, L. J., and M. F. Robbins (1973), ULF geomagnetic power near $L = 4$: 1. Quiet-day power spectra at conjugate points during December solstice, *J. Geophys. Res.*, *78*, 3816–3827, doi:10.1029/JA078i019p03816.

Lanzerotti, L. J., C. G. Maclennan, and M. Schulz (1970), Radial diffusion of outer-zone electrons: An empirical approach to third-invariant violation, *J. Geophys. Res.*, *75*, 5351–5371, doi:10.1029/JA075i028p05351.

Legendre, A. M. (1805), *New Methods for the Determination of the Orbits of Comets*, F. Didot, Paris, in French.

Lejosne, S., D. Boscher, V. Maget, and G. Rolland (2013), Deriving electromagnetic radial diffusion coefficients of radiation belt equatorial particles for different levels of magnetic activity based on magnetic field measurements at geostationary orbit, *J. Geophys. Res. Space Phys.*, *118*, 3147–3156, doi:10.1002/jgra.50361.

Li, X., D. N. Baker, M. Temerin, T. E. Cayton, E. G. D. Reeves, R. A. Christensen, J. B. Blake, M. D. Looper, R. Nakamura, and S. G. Kanekal (1997), Multisatellite observations of the outer zone electron variation during the November 3-4, 1993, magnetic storm, *Journal of Geophysical Research: Space Physics*, *102*(A7), 14,123–14,140, doi:10.1029/97JA01101.

Li, X., S. Palo, R. Kohnert, D. Gerhardt, L. Blum, Q. Schiller, D. Turner, W. Tu, N. Sheiko, and C. S. Cooper (2013), *Dynamics of the Earth's Radiation Belts and Inner Magnetosphere*, *Geophys. Monogr. Ser.*, vol. 199, chap. Colorado Student Space Weather Experiment: Differential flux measurements of energetic particles in a highly inclined low Earth orbit, pp. 385–404, AGU, Washington, D.C., doi:10.1029/2012GM001313.

Li, Z., M. Hudson, and Y. Chen (2014), Radial diffusion comparing a THEMIS statistical model with geosynchronous measurements as input, *J. Geophys. Res. Space Physics*, *119*, 1863–1873, doi:10.1002/2013JA019320.

Liu, R. (1990), On a notion of data depth based upon random simplices, *The Annals of Statistics*, *18*, 405–414.

Lyon, J. G., J. A. Fedder, and C. M. Mobarry (2004), The Lyon-Fedder-Mobarry (LFM) global MHD magnetospheric simulation code, *J. Atmos. Sol. Terr. Phys.*, *66*, 1333–1350, doi: 10.1016/j.jastp.2004.03.020.

Lyons, L. R., and D. J. Williams (1975), The quiet time structure of energetic (35-560 keV) radiation belt electrons, *J. Geophys. Res.*, *80*(7), 943–950, doi:10.1029/JA080i007p00943.

Lyons, L. R., R. M. Thorne, and C. F. Kennel (1972), Pitch-angle diffusion of radiation belt electrons within the plasmasphere, *J. Geophys. Res.*, *77*(19), 3455–3474, doi: 10.1029/JA077i019p03455.

Mann, I. R., A. N. Wright, K. J. Mills, and V. M. Nakariakov (1999), Excitation of magnetospheric waveguide modes by magnetosheath flows, *J. Geophys. Res.*, *104*(A1), 333, doi: 10.1029/1998JA900026.

Margenau, H., and G. M. Murphy (1956), *The Mathematics of Physics and Chemistry*, Princeton: Van Nostrand.

McIlwain, C. E. (1961), Coordinates for mapping the distribution of magnetically trapped particles, *Journal of Geophysical Research*, *66*(11), 3681–3691, doi:10.1029/JZ066i011p03681.

Mead, G. D. (1964), Deformation of the geomagnetic field by the solar wind, *J. Geophys. Res.*, *69*(7), 1181–1195, doi:10.1029/JZ069i007p01181.

Mozer, F. S. (1971), Power spectra of the magnetospheric electric field, *J. Geophys. Res.*, *76*, 3651–3667, doi:10.1029/JA076i016p03651.

Newkirk, L. L., and M. Walt (1968), Radial diffusion coefficient for electrons at $1.76 < L < 5$, *J. Geophys. Res.*, *73*, 7231–7236, doi:10.1029/JA073i023p07231.

Oja, H. (1983), Descriptive statistics for multivariate distributions, *Statistics and Probability Letters*, *1*, 327–332.

Olson, W. P., and K. A. Pfitzer (1974), A quantitative model of the magnetospheric magnetic field, *J. Geophys. Res.*, *79*(25), 3739–3748, doi:10.1029/JA079i025p03739.

Ozeke, L. G., I. R. Mann, K. R. Murphy, I. J. Rae, D. K. Milling, S. R. Elkington, A. A. Chan, and H. J. Singer (2012), ULF wave derived radiation belt radial diffusion coefficients, *J. Geophys. Res.*, *117*(A04222), doi:10.1029/2011JA017463.

Ozeke, L. G., I. R. Mann, K. R. Murphy, I. J. Rae, and D. K. Milling (2014), Analytic expressions for ULF wave radiation belt radial diffusion coefficients, *J. Geophys. Res. Space Physics*, *119*, doi:10.1002/2013JA019204.

Pegoraro, A. (1992), Revision of the Theil fitting method, *Astronomy and Astrophysics*, *259*, 386–393.

Perry, K. L., M. K. Hudson, and S. R. Elkington (2005), Incorporating spectral characteristics of Pc5 waves into three-dimensional radiation belt modeling and the diffusion of relativistic electrons, *J. Geophys. Res.*, *110*(A03215), doi:10.1029/2004JA010760.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2007), *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Cambridge University Press, New York.

Rae, I. J., I. R. Mann, K. R. Murphy, L. G. Ozeke, D. K. Milling, A. A. Chan, S. R. Elkington, and F. Honary (2012), Ground-based magnetometer determination of in situ Pc4-5 ULF electric field wave spectra as a function of solar wind speed, *J. Geophys. Res.*, *117*(A04221), doi:10.1029/2011JA017335.

Rao, C. R. (1945), Information and the accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.*, *37*, 81–89.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*, 2nd ed., John Wiley & Sons, New York.

Rao, C. R., and H. Toutenburg (1999), *Linear Models: Least Squares and Alternatives*, 2nd ed., Springer-Verlag, New York.

Roederer, J. G., and H. Zhang (2014), *Dynamics of Magnetically Trapped Particles*, Foundations of the Physics of Radiation Belts and Space Plasmas, 2nd ed., Springer-Verlag Berlin-Heidelberg, doi:10.1007/978-3-642-41530-2.

Ross, S. M. (1981), *Introduction to Probability Models*, 2nd ed., Academic Press.

Rousseeuw, P. J., and A. M. Leroy (2003), *Robust Regression and Outlier Detection*, 1st ed., Wiley-Interscience.

Ruohoniemi, J. M., R. A. Greenwald, and K. B. Baker (1991), HF Radar Observations of Pc5 Field Line Resonances in the Midnight/Early Morning MLT Sector, *J. Geophys. Res.*, *96*(A9), 15,697–15,710, doi:10.1029/91JA00795.

Schiller, Q., D. Gerhardt, L. Blum, X. Li, and S. Palo (2014), Design and scientific return of a miniaturized particle telescope onboard the Colorado Student Space Weather Experiment (CSSWE) CubeSat, in *35th IEEE Aerospace Conference*, 8.1102, doi:10.1109/AERO.2014.6836372.

Schulz, M. (1996), Canonical coordinates for radiation-belt modeling, in *Radiation Belts: Models and Standards*, *Geophysical Monographs*, vol. 97, edited by J. F. Lemaire, D. Heynderickx, and D. N. Baker, p. 153, American Geophysical Union, Washington, D.C.

Schulz, M., and L. Lanzerotti (1974), *Particle Diffusion in the Radiation Belts*, Springer, New York, doi:10.1007/978-3-642-65675-0.

Selesnick, R. S., J. B. Blake, W. A. Kolasinski, and T. A. Fritz (1997), A quiescent state of 3 to 8 MeV radiation belt electrons, *Geophys. Res. Lett.*, *24*, 1343–1346, doi:10.1029/97GL51407.

Sen, P. K. (1968), Estimates of the regression coefficient based on Kendall's Tau, *Journal of the American Statistical Association*, *63*(324), 1379–1389, doi:10.2307/2285891.

Shamos, M. (1976), *Geometry and statistics: Problems at the interface. Recent results and new directions in algorithms and complexity*, Academic Press, 251–280.

Shprits, Y. Y., R. M. Thorne, G. D. Reeves, and R. Friedel (2005), Radial diffusion modeling with empirical lifetimes: Comparison with CRRES observations, *Ann. Geophys.*, *23*(4), 1467–1471, doi:10.5194/angeo-23-1467-2005.

Shprits, Y. Y., S. R. Elkington, N. P. Meredith, and D. A. Subbotin (2008a), Review of modeling of losses and sources of relativistic electrons in the outer radiation belt I: Radial transport, *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*, 1679–1693, doi:10.1016/j.jastp.2008.06.008.

Shprits, Y. Y., D. A. Subbotin, N. P. Meredith, and S. R. Elkington (2008b), Review of modeling of losses and sources of relativistic electrons in the outer radiation belt II: Local acceleration and loss, *Journal of Atmospheric and Solat-Terrestrial Physics*, *70*, 1694–1713, doi: 10.1016/j.jastp.2008.06.014.

Singer, H. J., W. P. Sullivan, P. Anderson, F. Mozer, P. Harvey, and J. Wygant (1992), Fluxgate magnetometer instrument on CRRES, *J. Spacecr. Rockets*, *29*, 599–601, doi:10.2514/3.25506.

Slepian, D. (1978), Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case, *Bell Systems Tech. J.*, *57*, 1371–1430, doi:10.1002/j.1538-7305.1978.tb02104.x.

Small, C. G. (1990), A survey of multidimensional medians, *International Statistical Review*, *58*(3), 263–277.

Southwood, D. J. (1974), Some features of field line resonances in the magnetosphere, *Planet. Space Sci.*, *22*, 483, doi:10.1016/0032-0633(74)90078-6.

Southwood, D. J., J. W. Dungey, and R. J. Etherington (1969), Bounce resonant interaction between pulsations and trapped particles, *Planet. Space Sci.*, *17*, 349–361, doi:10.1016/0032-0633(69)90068-3.

Squires, G. L. (2001), *Practical Physics*, 4th ed., Cambridge University Press.

Strutz, T. (2011), *Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond*, Vieweg-Teubner Verlag.

Su, Z., F. Xiao, H. Zheng, and S. Wang (2010), STEERB: A three-dimensional code for storm-time evolution of electron radiation belt, *J. Geophys. Res.*, *115*(A09208), doi:10.1029/2009JA015210.

Su, Z., F. Xiao, H. Zheng, and S. Wang (2011a), CRRES observation and STEERB simulation of the 9 October 1990 electron radiation belt dropout event, *Geophys. Res. Lett.*, *38*(L06106), doi:10.1029/2011GL046873.

Su, Z., F. Xiao, H. Zheng, and S. Wang (2011b), Radiation belt electron dynamics driven by adiabatic transport, radial diffusion, and wave-particle interactions, *J. Geophys. Res.*, *116*(A04205), doi:10.1029/2010JA016228.

Subbotin, D. A., Y. Y. Shprits, M. Gkioulidou, L. R. Lyons, B. Ni, V. G. Merkin, F. R. Tofoletto, R. M. Thorne, R. B. Horne, and M. K. Hudson (2011a), Simulation of the acceleration of relativistic electrons in the inner magnetosphere using RCM-VERB coupled codes, *J. Geophys. Res.*, *116*(A08211), doi:10.1029/2010JA016350.

Subbotin, D. A., Y. Y. Shprits, and B. Ni (2011b), Long-term radiation belt simulation with the VERB 3-D code: Comparison with CRRES observations, *J. Geophys. Res.*, *116*(A12210), doi:10.1029/2011JA017019.

Summers, D., R. M. Thorne, and F. Xiao (1998), Relativistic theory of wave-particle resonant diffusion with application to electron acceleration in the magnetosphere, *J. Geophys. Res.*, *103*, 20,487, doi:10.1029/98JA01740.

Takahashi, K., and B. J. Anderson (1992), Distribution of ULF energy ($f < 80$ mHz) in the inner magnetosphere: A statistical analysis of AMPTE CCE magnetic field data, *J. Geophys. Res.*, *97*(A7), 10,751–10,773, doi:10.1029/92JA00328.

Tascione, T. F. (1988), *Introduction to the Space Environment*, Orbit Book Company, Inc.

Temerin, M., I. Roth, M. K. Hudson, and J. R. Wygant (1994), New paradigm for the transport and energization of radiation belt particles, in *Eos Transactions*, vol. 75.

Theil, H. (1950), A rank-invariant method of linear and polynomial regression analysis I-III, *Publications of the Statistical Department of the Mathematisch Centrum*, *53*, 386–392,521–525,1397–1412, Amsterdam.

Thomson, D. J. (1982), Spectrum estimation and harmonic analysis, in *Proc. IEEE*, vol. 70, pp. 1055–1096, doi:10.1109/PROC.1982.12433.

Treumann, R. A., and W. Baumjohann (1997), *Advanced Space Plasma Physics*, Imperial College Press.

Tsyganenko, N. A., and M. I. Sitnov (2005), Modeling the dynamics of the inner magnetosphere during strong geomagnetic storms, *Journal of Geophysical Research: Space Physics*, *110*(A3), doi:10.1029/2004JA010798, a03208.

Tu, W., X. Li, Y. Chen, G. D. Reeves, and M. Temerin (2009), Storm-dependent radiation belt electron dynamics, *J. Geophys. Res.*, *114*(A02217), doi:10.1029/2008JA013480.

Tu, W., S. R. Elkington, X. Li, W. Liu, and J. Bonnell (2012), Quantifying radial diffusion coefficients of radiation belt electrons based on global MHD simulation and spacecraft measurements, *J. Geophys. Res.*, *117*(A10210), doi:10.1029/2012JA017901.

Tu, W., G. S. Cunningham, Y. Chen, M. G. Henderson, E. Camporeale, and G. D. Reeves (2013), Modeling radiation belt electron dynamics during GEM challenge intervals with the DREAM3D diffusion model, *J. Geophys. Res. Space Physics*, *118*, 6197–6211, doi:10.1002/jgra.50560.

Tukey, J. (Ed.) (1975), *Mathematics and the picturing of data*, Proceedings of the International Congress of Mathematicians, Vancouver, 523–531.

Varotsou, A., D. Boscher, S. Bourdarie, R. Horne, N. P. Meredith, S. A. Glauert, and R. H. Friedel (2008), Three-dimensional test simulations of the outer radiation belt electron dynamics including electron-chorus resonant interactions, *J. Geophys. Res.*, *113*(A12212), doi:10.1029/2007JA012862.

Weber, A. (1909), *Theory of Location of Industries*, University Chicago Press, English translation by C. Friedrich (1929).

Weiszfeld, E. (1937), Sur le point pour lequel la somme des distances de n points donnés est minimum, *Tohoku Mathematical Journal*, *43*, 355–386.

West, I. H., R. M. Buck, and G. T. Davidson (1981), The dynamics of energetic electrons in the Earth's outer radiation belt during 1968 as observed by Lawrence Livermore National Laboratory's spectrometer on Ogo 5, *J. Geophys. Res.*, *86*, 2111–2142, doi:10.1029/JA086iA04p02111.

Wilcox, R. R. (1998a), A note on the Theil-Sen regression estimator when the regressor is random and the error term is heteroscedastic, *Biometrica J.*, *40*(3), 261–268.

Wilcox, R. R. (1998b), Simulation results on extensions of the Theil-Sen regression estimator, *Commun. Stat. Simul. Comput.*, *29*, 275–284.

Wilcox, R. R. (2001), *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*, Springer-Verlag, USA.

Wilcox, R. R. (2004), Some results on extensions and modifications of the Theil-Sen regression estimator, *British Journal of Mathematical and Statistical Psychology*, *57*, 265–280.

Wygant, J., J. Bonnell, K. Goetz, R. Ergun, F. Mozer, S. Bale, M. Ludlam, P. Turin, P. Harvey, R. Hochmann, K. Harps, G. Dalton, J. McCauley, W. Rachelson, D. Gordon, B. Donakowski, C. Shultz, C. Smith, M. Diaz-Aguado, J. Fischer, S. Heavner, P. Berg, D. Malsapina, M. Bolton, M. Hudson, R. Strangeway, D. Baker, X. Li, J. Albert, J. Foster, C. Chaston, I. Mann, E. Donovan, C. Cully, C. Cattell, V. Krasnoselskikh, K. Kersten, A. Brenneman, and J. Tao (2013), The Electric Field and Waves Instruments on the Radiation Belt Storm Probes Mission, *Space Science Reviews*, *179*(1-4), 183–220, doi:10.1007/s11214-013-0013-7.

Wygant, J. R., P. R. Harvey, D. Pankow, F. S. Mozer, N. Maynard, H. Singer, M. Smiddy, W. Sullivan, and P. Anderson (1992), CRRES electric field/Langmuir probe instrument, *J. Spacecr. Rockets*, *29*, 601–604, doi:10.2514/3.25507.