Periodic and Quasi-periodic Responses of the Earth's Outer Electron Belt to the Solar Wind with Implications to Space Weather Modeling.

by

Manny Presicci

B.S., University of Rochester, 1978

M.S., Rensselaer Polytechnic Institute, 1983

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 Physics

2010

This thesis entitled:

Periodic and Quasi-periodic Responses of the Earth's Outer Electron Belt to the Solar Wind with
Implications to Space Weather Modeling
written by Manny Presicci
has been approved for the Department of Physics

_____
Daniel N. Baker

_____
David Newman

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.

Presicci, Manny (Ph. D., Physics)

Periodic and Quasi-periodic Responses of the Earth's Outer Electron Belt to the Solar Wind with

Implications to Space Weather Modeling

Thesis directed by Professor Daniel N. Baker

It has long been noted the distinct periodicities in the electron flux response in the outer belts are associated with solar wind speed periodicities. The well-organized high-speed streams in the solar wind drive the flux, most commonly resulting in 27, 13.5 and 9-day peaks in the periodogram during the declining phase of the solar cycle. The distribution of peak power across L-shells is calculated and a connection is made with the P_1 peak in the impulse response function.

Next, the periodogram estimates the power spectrum relating weak, diffuse periodicities in the number of coronal mass ejection occurrences to relativistic electron flux periodicities. Because coronal mass ejections occur preferentially during ascending phases of the solar cycle, this supplements the distinct periodicities from high-speed streams evident during declining phases of the solar cycle.

Finally, the successful estimation of exogenous coefficients in an autoregressive exogenous (ARX) model is shown to require excitation at all pole frequencies in the transfer function between the output and the input. Specifically, the solar wind speed should excite all pole frequencies in the electron flux response for complete identification. Additionally, because only the first two autoregressive coefficients are dominant, system identification of eight coefficients in an ARX time series model via the Kalman filter suggests a lower order ARX model, such as the stochastic linear oscillator may be more appropriate.

Acknowledgments

Many people contribute to the completion of a doctoral dissertation. First and foremost my research adviser Dr. Daniel N Baker has provided me guidance throughout seven years that it has taken to complete this dissertation. Dr. Baker has demonstrated the value of having a laser-like focus on a difficult problem, while giving me the freedom to explore unconventional trial solutions, some of which were fruitless. I must have been a temperamental student for him.

Second, I will certainly mention Dr. Shri Kanekal. Dr. Kanekal assuaged my self-doubt in getting started with the dissertation. Through him, I was able to overcome the "paralysis by analysis", which plagues researchers of my temperament.

There were many instances, where Dr. Xinlin Li was available to clear up technical misconceptions which I naively held. Dr. Scot Elkington was also valuable, when I often asked him what he thought of ideas I had, which were misdirected.

Edward Burin Des Roziers and Nathan Farr provided camaraderie as well as some guidance. I remember with fondness the time Edward and I left work aside one Saturday morning and explored some of the hiking trails in Chautauqua Park in Boulder. Nathan completed the Physics program like me, and I could ask for his perspective to see if it matched mine. This relieved some of the isolation from long hours.

Dr. David Newman, Dr. Mihaly Horanyi, Dr. Tobin Munsat, and Dr. Michael Wiltberger served on my graduate committee. I thank them for their service. I also thank Dr. Scott Robertson for serving on my COMPS II proposal. Dr. Wiltberger served as the required outside reader of the dissertation.

Most likely the person with the greatest professional impact on me was the mentor assigned to me at the Naval Surface Weapons Center in Dahlgren, Virginia. For over twenty-six years Don Boyer has been a good friend, and mentor, as well as an instructor of mathematics and systems theory to me. This has been a valuable relationship that I hold to this very day.

The staff of the Lester Library in the Duane Physics Building and fellow employees at LASP are too numerous to single out, but these "unsung" heroes are usually overlooked when reviewing the efforts they make towards a successful thesis completion. I appreciate their efficiency and effort.

On a less personal level, Dr. Michael Wiltberger reminds me to regard the role of data suppliers to the advancement of this thesis. Because I use onset times from the SOHO LASCO CME catalog, the data supplier requests that I include the following statement conditioned on its use.

"This CME catalog is generated and maintained at the CDAW Data Center by NASA and The Catholic University of America in cooperation with the Naval Research Laboratory. SOHO is a project of international cooperation between ESA and NASA."

Contents

Chapter

Tables

Table

Figures

Figure

1. Introduction

1.1 Roadmap:

The classical Greeks considered the two "ideal" plane shapes, the line and the circle. Extending these spatial shapes along a time axis, dynamic prediction is most natural when physical models place emphasis on either linear or cyclic behavior. This thesis treats cyclic and quasi-cyclic processes in the solar wind and electron radiation belt as well as linear predictive models for electron flux from records of solar wind speed. Because this effort has been preceded by similar investigations noting recurrence in solar wind and electron flux characteristics in the radiation belts, it is important to begin these investigations with a brief history. Next, the stage is set, presenting a background concerning features of the solar corona and characteristics of the fast and slow solar wind. Because different structures are responsible for elevating relativistic electron flux depending on the solar cycle phase, high-speed streams (HSS) effective during the declining phase are introduced. Second, Coronal Mass Ejections (CME) effective during ascending phases of the solar cycle, are discussed. Third, the literature is reviewed to establish comprehensively the role storms driven by co-rotating interaction regions (CIRs) and storms driven by CME's in enhancing the electron flux in the radiation belt. Three superposed epoch analyses (SEA's), from Miyoshi and Kataoka [1], Turner et al. [2], and Kanekal [3], illustrating the behavior of the solar wind speed (driver) during CIR and CME-driven storms are discussed. A follow-up of this comparison for solar wind is made for electron flux (response), using results from Borovsky and Denton [4], and Kanekal. These three sections conclude the Introduction.

The first section in Chapter 2, Dynamic comparison of the periodicities, describes data sources used to determine the power spectrum for the solar wind speed and relativistic electron flux. Because graphically illustrating periodicity with the periodogram requires estimating the power spectrum, the second section presents useful concepts necessary for power spectral estimation and "ideal" examples of power spectra from Gelb[5], although such ideal examples are seldom realized in practice.

It would be fortunate if power spectral estimation with actual data containing noise were straightforward, however there are many pitfalls for the uninitiated. Realistic power spectral estimation of noisy processes from actual measurements necessarily requires a tradeoff between consistent estimation and frequency resolution. This tradeoff is made by windowing, the subject of the third section.

It is only after making the tradeoff that the first original results showing periodograms for the relativistic electron flux and solar wind speed for a wide variety of L-shells and over the full length of SC-23 can be presented in the fourth section. In the fifth and final section, the Welch periodogram is used to determine the distribution of L-shell locations with the maximum power throughout 1994-2006 containing solar cycle 23 (SC-23). The peak in the power distribution is later associated with the P_1 peak in the impulse response function for electron flux with the solar wind speed obtained through singular value decomposition by Vassiliadis et al [6,7]. This observation reconfirms the role of HSS in the solar wind to enhancing electron flux noted by Vassiliadis et al.

Strong distinct peaks in the power spectrum with definite periodicity from HSS are not the only features of the electron flux and solar wind power spectra. The electron flux and solar wind power spectra from ascending phases of the solar cycle may also contain weak diffuse

peaks. The third chapter, Spectral estimation of CME number, provides evidence relating the weak 36-day peak in the electron flux power spectrum during the ascending phase of SC-23 to the periodicity in CME number.

A recent investigation by Lara et al. [8] describes short-period periodicity in CME number during SC-23, obtained by estimating the power spectrum with the maximum entropy method. Because this method often produces noisy peaks, Lara et al. reinforce their estimation with wavelet methods. This thesis further reinforces this estimation using the Lomb-Scargle [9, 10], and Welch's [11] periodogram. Because the SOHO LASCO catalog of CME onset times is incomplete, the Lomb-Scargle algorithm for estimating the power spectrum with unevenly spaced data is a natural choice. Alternatively, I generate synthetic data with the same statistical properties to fill in missing data and estimate the power spectra with the Welch periodogram. Specifically, missing entries in CME number are filled in drawn from a Poisson distribution after establishing the time between CME onsets follows an exponential distribution. These two methods further reinforce 94, 45, 36, 28, and 23-day periodicities found by Lara et al. Together these two results along with the maximum entropy periodogram from Lara .et al. establish a consensus. The periodicities established in the consensus are related to the electron flux periodogram during ascent phases in the solar cycle.

Chapter 4 moves the focus from establishing the periodicities to their potential significance for a space weather model using linear estimation. First, the autoregressive exogenous time series (ARX(4,4)) is chosen to model the electron flux response to the solar wind speed. The collection of eight system coefficients in the ARX(4, 4) model will defines the time series which best "fits" the input/output stream (i.e. solar wind speed /electron flux records). Subsequently the recursive nature of the Kalman filter is presented, and the criterion used to

optimally estimate the state (i.e. collection of eight system coefficients) under consideration. The next two sections 4.5 and 4.6 outline the nature of solutions to the Kalman filter equations, along with the ARX (4, 4) time series modeled by the signal model. Both, the series of state estimates from the Kalman filter, and the state of the system modeled by the time series will be Normal stochastic processes. Section 4.7 describes specifically how the optimal state estimate and error covariance from the Kalman filter can be obtained in a cyclic and recursive manner. This manner is most amenable to solving with a computer algorithm.

With the substantial mathematical formalism behind us, I use the computer algorithm to determine the system coefficients in a trivial case with zero-noise, and zero forcing. The signal model with known (i.e. prescribed) system coefficients generates a set of measurements to be input into the Kalman filter. Filter validation consists of identifying (in this case recovering known) system coefficients. In the zero-noise case, recovery of the autoregressive coefficients is sudden, and recovery of the exogenous coefficients will be impossible. The Kalman filter validates these results. Section 4.9 continues by examining the performance of the filter with white noise input (the autoregressive moving average (i.e. ARMA) process), and coefficient identification using periodic measurements.

The filter results will reveal the following.

1. The simulated ARX(4, 4) time series with prescribed autoregressive coefficients will have two particular frequencies corresponding to the poles in the transfer function. When the forcing function has a frequency remote from the two pole frequencies (case a), identification of the exogenous coefficients is unsuccessful.

2. When the forcing function has frequencies near one (either) of the pole frequencies (cases b and c), identification of the exogenous coefficients is unsuccessful.

3. When the forcing function is contains frequencies near both of the pole frequencies (case d), identification of the exogenous coefficients is successful.

4. Because white noise contains a wide spectrum of (ideally all) frequencies, white noise input allows for successful exogenous coefficients given sufficient time.

Collectively, results 1-3 along with our study of periodicities imply the success of a linear analysis to predicting coefficients in the ARX model will depend on frequency characteristics of the solar wind speed (forcing). If the frequency characteristics excite all of the system "modes", the exogenous coefficients can be identified. The dependence of power spectrum for the solar wind on solar cycle phase suggests system identification of exogenous (i.e. forcing) coefficients in a linear predictive model will also be solar cycle phase dependent.

As a final note, actual solar wind speed measurements and logarithm electron flux measurements from 1994-2004 are input into the Kalman filter, and the eight system coefficients in an ARX(4,4) model identified. The dominance in the first two autoregressive coefficient suggests a simpler model may be more appropriate.

This roadmap summarizes the organization used in the forthcoming thesis. I now proceed with a history of earlier investigations into solar wind and electron flux periodicities.

1.2   History of Periodicities in the Solar Wind and Radiation Belts.

Kristian Birkeland's contention in 1896, that the Sun emitted beams of electrons, which were responsible for geomagnetic disturbances on Earth was revolutionary for its time [12-14]. A. Harvey noticed that magnetic storms tended to occur at 27-day intervals and attributed this recurrence to electron beams, an idea first proposed by Birkeland [13, 15]. Four years earlier,

the eminent Lord Kelvin considered a terrestrial magnetic storm should require the Sun to radiate

160 trillion horsepower, which he reckoned was 346 times the total power in solar radiation [16].

(Today we know he vastly underestimated the Sun's luminosity).   He dismissed the association

between sunspots and magnetic storms as compelling but coincidental [14,16,17]. Putting aside

Kelvin's estimate for a stronger argument, Kivelson and Russell cite A. Schuster's remark that

an electron beam from the Sun should break up due to electrostatic repulsion [17]. The Sun-

Earth connection was far from established.

Full acceptance of particles from the Sun and the 27-day association between storms and

solar wind speed had to await the opening days of the space age. The launching of Mariner 2 in

1962 provided indisputable evidence for an omnipresent solar wind, and the association of a 27-

day period for solar wind speed with M-region solar storms. The M- regions (the term used by

Snyder, Neugebauer and Rao [15]) emit high-speed plasma from the Sun.

Our retrospective fifty years later reveals a few weaknesses in this characterization of the

solar wind. For example, Snyder et. al concluded that there was no strong correlation between

solar activity and solar wind speed, and no obvious relationship between cosmic ray intensity

and plasma speed; (these are both at odds with our current understanding).   Despite these false

contentions, today we appreciate the in-situ capability of spacecraft developed after 1957, which

has greatly improved space physics studies.

An early investigation by Baker, et al. [18] found both proton and electron fluxes at

synchronous orbit L = 6.6 Re. were elevated in response to the solar wind speed, (see

 Figure 1.1). The daily average proton fluxes from the lowest Hi-P energy channel 400-500 KeV

and (channels Hi-P 5,6,and 7) 1 MeV to 2.1 MeV from the charged particle analyzer (CPA)

instrument onboard the geosynchronous satellite 1976-059A are shown in the top panel. The

middle panel shows the daily average electron flux from the high- energy electron unit of the energetic particle sensor (EPS), which measures the flux above seven thresholds from 0.2 to 2.0 MeV [19]. The bottom panel shows 3-hour averages of the solar wind speed from the LASL-experiment onboard NASA satellites IMP-7 and IMP-8.
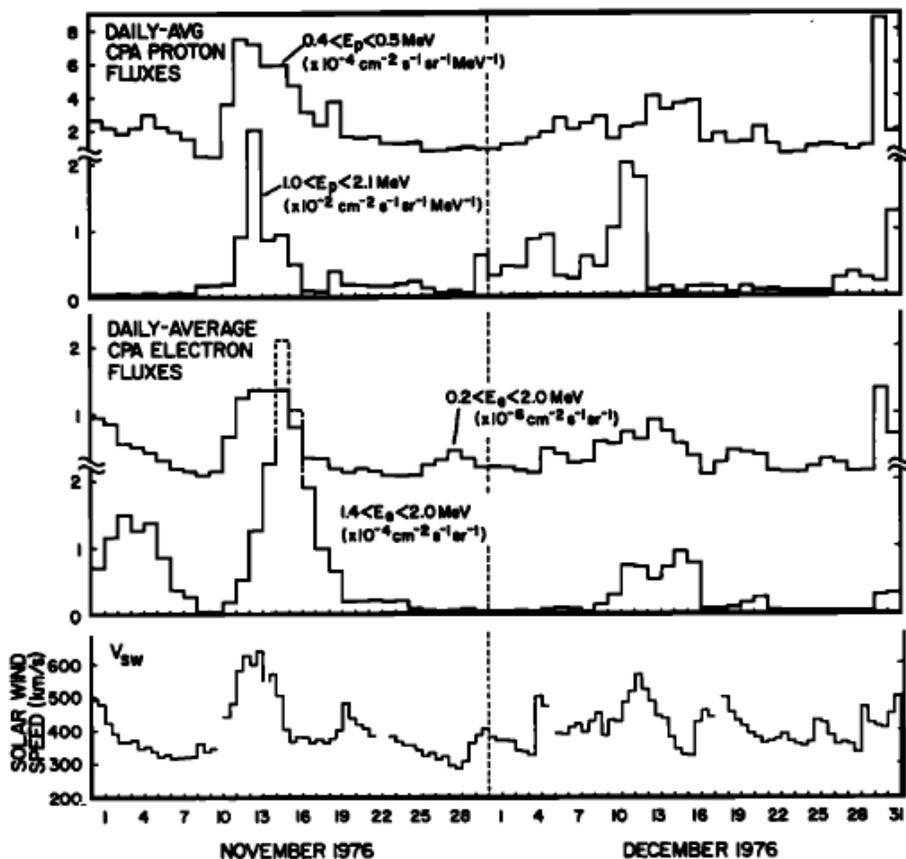
Figure 1.1: Solar Wind Speed Proton Fluxes Electron Fluxes, Baker et al. [18].

It is evident from Figure 1.1, that the electron and proton fluxes are elevated after a two to three day lag time to increases in solar wind speed.

Paulikas and Blake [20] also associated relativistic electron flux in the outer belt with a 27-day periodicity in solar wind speed from HSS. They investigated daily averaged electron flux measurements at low energy 140-600 KeV, and higher energy > 700 KeV, from the Aerospace Corporation Experiment onboard ATS-6 during July to December 1974 and 1976-7, during the

late declining phase of SC-20. Jack Gosling at Los Alamos provided them a tape of hourly solar wind velocity measurements, indicating unusually fast high-speed streams over 700 km/sec during solar rotation 1935, along with slower high-speed streams barely 600 km/sec during two solar rotations 1963 and 1964. They observed a strong correlation between the increase in electron flux and the passage of the high-speed stream. Between one and two days after the peak solar wind velocity from the stream, the electron flux builds to a maximum. It is also interesting that they noted the recovery back to the earlier electron flux level did not follow the fall in solar wind speed.

The correlation from Paulikas and Blake also indicated a 12-hr delayed response of the lower energy electron flux to the solar wind speed. At the higher energy, there was a 36-hr delayed response. These correlations demonstrated the electron flux response follows solar wind driving with an energy-dependent delay.

Baker, et al., [21] also observed recurrent 27-day flux increases using 3-10 MeV electron flux measurements from the spectrometer for energetic electrons (SEE) from the LANL spacecraft 1979-053 and 1982-019. After solar maximum, during the descent phase from late 1981-1984, they noted frequent flux increases. Their superposed epoch analysis (SEA) demonstrated the electron flux builds up for 2-3 days after the peak in solar wind velocity, and decays 3-4 days afterward. They noted the full width at half maximum for the peak was 2.5 days. They also advanced the possibility the high-energy electrons are energized externally and transported by channeling Jovian electrons between two successive co-rotating interaction regions (See Figure 1.2). A digression into the follow-up to this idea illustrates the point that observing periodicities can either suggest or discard relevant acceleration mechanisms, adding motivation for studying these periodicities.

For example, indications of external transport of high-energy electrons from Jupiter would require electron flux peaks at 13-month intervals, i.e. the synodic period of Jupiter as viewed from the Earth. Christon, et al. [22] failed to see any significant correlation in electron flux at geosynchronous orbit from the LANL charged particle
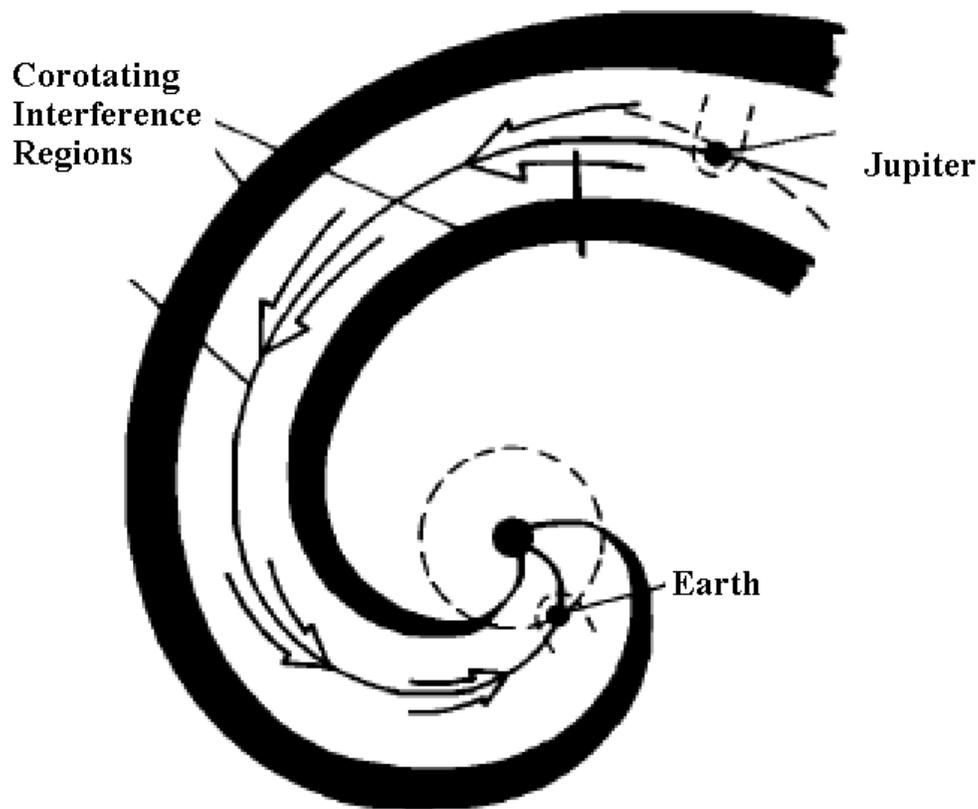
Figure 1.2: Co-rotating Streams Channeling Jovian Electrons to the Earth, Baker et al. [21]

analyzer onboard S/C 1976-059A (0.2-2 MeV electrons) or S/C 1979-053 and S/C1982-019 (3-7 MeV electrons) from 1976-1984 for 7 possible 13-month periods.

However, the absence of the 13- month periodicity is not the only evidence weakening the theory of Jovian origin. More recently, Li, et al. suggest external transport cannot be the sole

cause of energetic electrons. These electrons require internal acceleration mechanisms within the magnetosphere because there is insufficient phase space density in 20-200 KeV (i.e. seed) electrons in the solar wind to supply the outer radiation belt [23]. In any case, attributing the electron acceleration mechanism to the Jovian magnetosphere is only a change of venue and unsatisfying. It raises the question, how are electrons accelerated to relativistic energies at Jupiter?

A. Nishida presented a recirculation model for non-adiabatic acceleration of electrons in the Jovian magnetosphere [24]. One can consider it as similar to a "four-stroke" engine operating in a cycle. The four phases of the cycle are inward radial diffusion, pitch angle scattering, cross-L diffusion at low altitude, and isotropization (see Figure 1.3).

The key is that the all energy gained from betatron acceleration during inward diffusion would be lost during transport outward which would be required for the process to work in a cycle. However, pitch angle scattering changes the transverse kinetic energy of the electrons so that the transverse kinetic energy given back in the outward process; isotropization (the same amount gained in inward diffusion) does not match the current transverse kinetic energy of electron, potentially allowing the electron to keep some energy at the end of the process. Monte-Carlo simulations by Nishida have shown how this process could work in the Jovian magnetosphere [25]. However, the failure to find a significant 13-month interval peak in the electron flux weakens the support for a Jovian origin.

Recirculation continues to be a viable mechanism to explain electron energization. Fujimoto and Nishida [26], and Baker et al.[27] present evidence that a similar mechanism may be responsible for energizing electrons in the Earth's outer radiation belt, preserving recirculation as a possibility without requiring Jovian electrons.
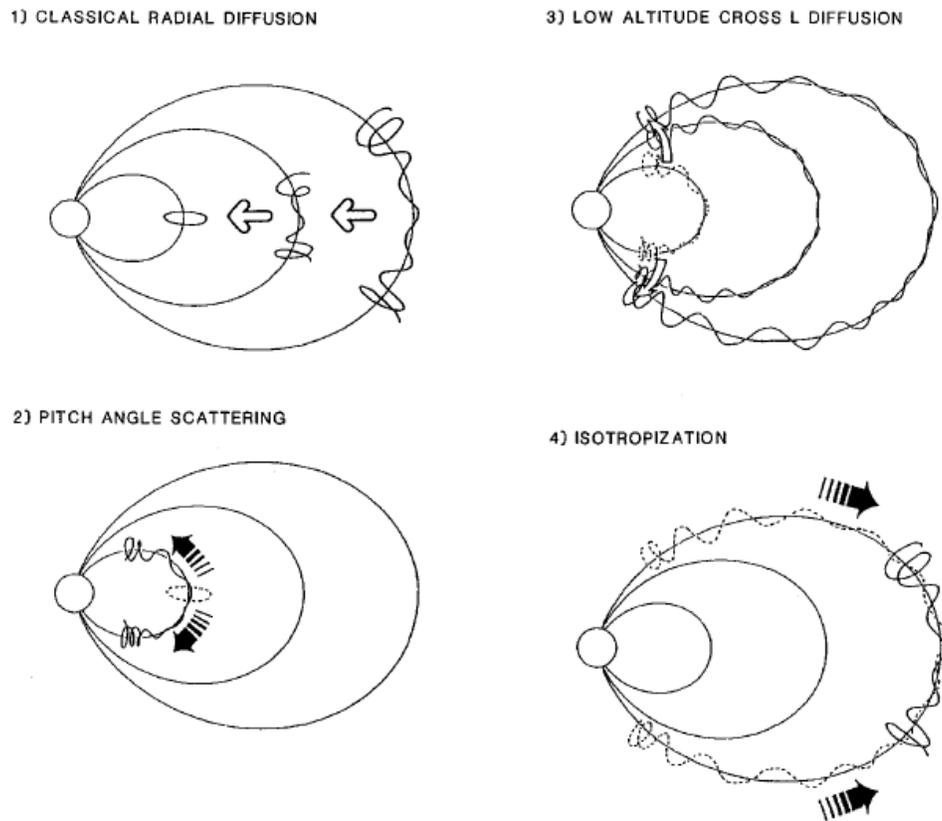
Figure1.3: The Four Elements of the Recirculation Process Energizing Jovian Electrons, adapted from Nishida [24]

This history for earlier investigations of periodicity in the solar wind and outer radiation belt continues with the introduction of linear predictive modeling to study periodicity in relativistic electron flux. Using a linear continuous-time (aka Weiner) filter Baker, et al. [28] noted 27, 54, and 81 day periodicities in the relativistic electron flux, peaks of 27 and 13 days in solar wind speed, and peaks of 27 and 54 days in geomagnetic indices from auto-correlations and cross-correlations for these quantities. They determined similar impulse response matrices using the planetary index, Kp, and the AE index as input time series. One motivation behind this extension was their possible use as a proxy if solar wind speed measurements were unavailable. The electron flux measurements investigated were from the Spectrometer for Energetic Electron

Sensor (SEE) onboard two geosynchronous LANL satellites S/C 1979-053 and 1982-019 from 1981 to 1984. The fluxes were small during 1982-1983, the early declining phase of SC-21, and were larger in the late declining phase during 1984-5. They determined the impulse response matrix from a matrix equation relating the auto-covariance of the input time series (e. g. the solar wind speed) to the cross-covariance from the input-output time series (from the solar wind speed with the relativistic electron flux together).

The impulse response in figure 1.4 from Baker et al. below shows the electron flux peaks 2 days after passage of the high-speed stream; the full width at half-maximum (FWHM) is ~ 2-1/2 days .
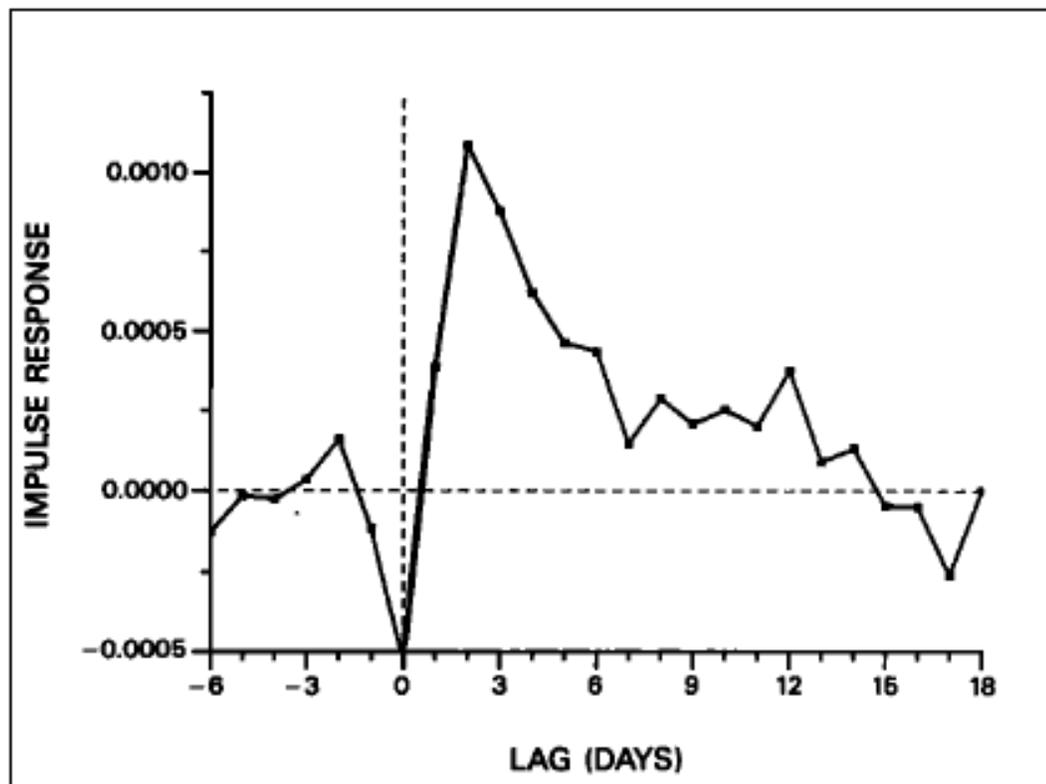
Figure 1.4: The Electron Flux Impulse Response to Solar Wind Speed, Baker et al. [28]

Finally, Baker et al. note the impulse response for electron flux from solar wind

driving strongly resembles similar impulse responses using geomagnetic indices as an input justifying their use as a proxy if needed. However, they remark the solar wind speed falls more rapidly to zero than auroral electrojet (AE) or planetary index (Kp). This indicates the solar wind speed is closer to the central cause for enhanced electron flux, than indices also driven by the flux.

This summary of earlier investigations into periodicities associated with solar rotation is representative, and not exhaustive. It could be supplemented using many references cited from these investigations. The last study that I will mention by Gibson et al. uses the declining phase from two different solar cycles to demonstrate differences in electron belt response [29].

During the solar minimum phase of SC-22, an international campaign, the Whole Sun Month (WSM), was launched to study high-speed streams at 1 AU from August 10, to September 8, 1996. The Whole Heliospheric Interval (WHI), March 20 to April 16, 2008, marked a similar second campaign during the late minimum phase of SC-23. Although the two campaigns studied solar minimum for succeeding solar cycles, the high-speed streams emanating from solar coronal holes were profoundly dissimilar and had dramatic effects on the > 2 MeV electron flux response, and auroral electron power (see figure 1.5).

Gibson et al. also presented coronagraphs in extreme ultraviolet (EUV) from the Solar Orbiting Heliospheric Observatory (SOHO) showing coronal holes as dark regions within a light background, and noticed differences in their morphology. The solar disk during WHI was marked with large low latitude coronal holes, where during WSM the coronal holes at solar minimum were smaller and polar. In general, during solar minimum, the Sun's magnetic field is strongly dipolar and aligned with the rotation axis. Fast solar wind is directed mainly along polar field lines. Consequently, the Earth in the ecliptic sees little of this fast solar wind. On the other

hand, the Earth, in the Sun's ecliptic can intercept high-speed streams emanating from low-latitude coronal holes so the average solar wind speed should be higher during WHI than WSM (because low latitude coronal holes were more frequent). This explains the higher fractional durations of sustained solar wind speed above 450 km /sec, during (WHI) 55%, than during (WSM) 31%.

In Figure 1.5, Gibson et al. show the power spectrum from a Lomb-Scargle periodogram by Emery et al. [29,30] for frequencies present in auroral power, solar wind velocity and radiation electron belt flux above 2 MeV, during WSM and WHI. During WSM distinct periodicities in the range from 3 to 100 days were not apparent, with only small peaks in power at 27 and 13 days. During WHI, the power spectrum showed sharp 27, 13.5 and 9-day periodicities in auroral power, solar wind velocity and radiation belt flux.
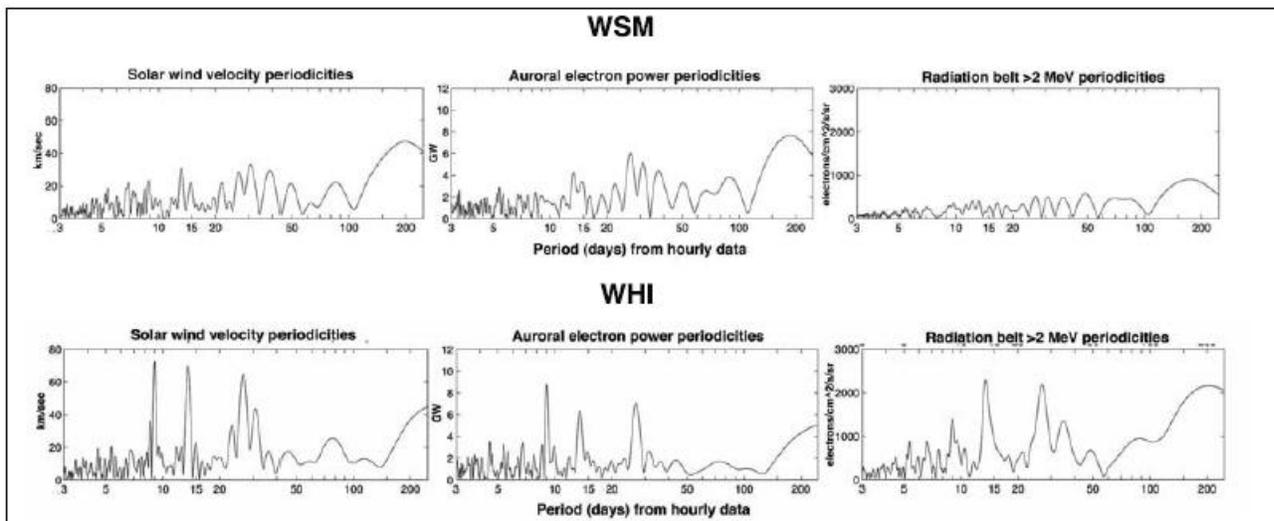
Figure 1.5: Lomb-Scargle Periodograms for Solar Wind Speed, Auroral Power and Electron during WSM and WHI, taken from Gibson et al. [29]

The 27-day periodicity in solar wind speed has a natural association with the rotation of the Sun's corona. The 13.5-day periodicity can be explained readily by a two-sector pattern in

the magnetic field from a tilted helio-sheet [31]. However, until recently the 9-day periodicity was somewhat mysterious. The study of areas from coronographs by Vrsnak et al. has found an association with a triangular distribution of coronal holes on the Sun [32, 33].

One may also note auroral electron power is not the only terrestrial effect exhibiting a 9-day periodicity of solar origin. Lei et al. has observed the 9-day periodicity in the neutral density in the thermosphere [34] during the descent phase of SC-23, including the WHI.

The earliest reference I could find noting the nine-day periodicity in the solar wind speed was from Verma and Joshi [35] during solar cycles 20 and 21 (1972-1984) from instrumentation (unspecified in the reference) onboard Pioneer 10. This early observation from SC-20, 21 along with the descending phase of SC-23 leading to solar minimum SC-24 suggests the 9-day period for solar wind speed is not an infrequent event, unlikely to recur.

Using daily averaged solar wind speed measurements from PIONEER obtained from the OMNIWEB database, Figure 1.6 shows a Welch periodogram estimating the power spectrum in solar wind speed for one year starting June 14, 1975 to Mid - June 1976, concluding SC-20. Features distinguishing the Welch periodogram from other power spectral estimations are provided in section 2.3.2. I used linear interpolation to fill in missing solar wind speed measurements. Because there are only 366 data points, resolving long periods is difficult. The period range shown is limited to [0, 40] days. The periodogram for solar wind speed shown in Figure 1.6 shows a strong 9-day peak, a weaker 13-day peak, and strong 27-day peaks, with increasing width.

1.3   Solar Cycle Phase Influences

Radiation and the solar wind are responsible for the energy, momentum and mass entry into the magnetosphere. Because this thesis focuses on the role of solar wind speed to energizing electrons in the Earth's outer radiation belt, I will only briefly discuss radiation; although the
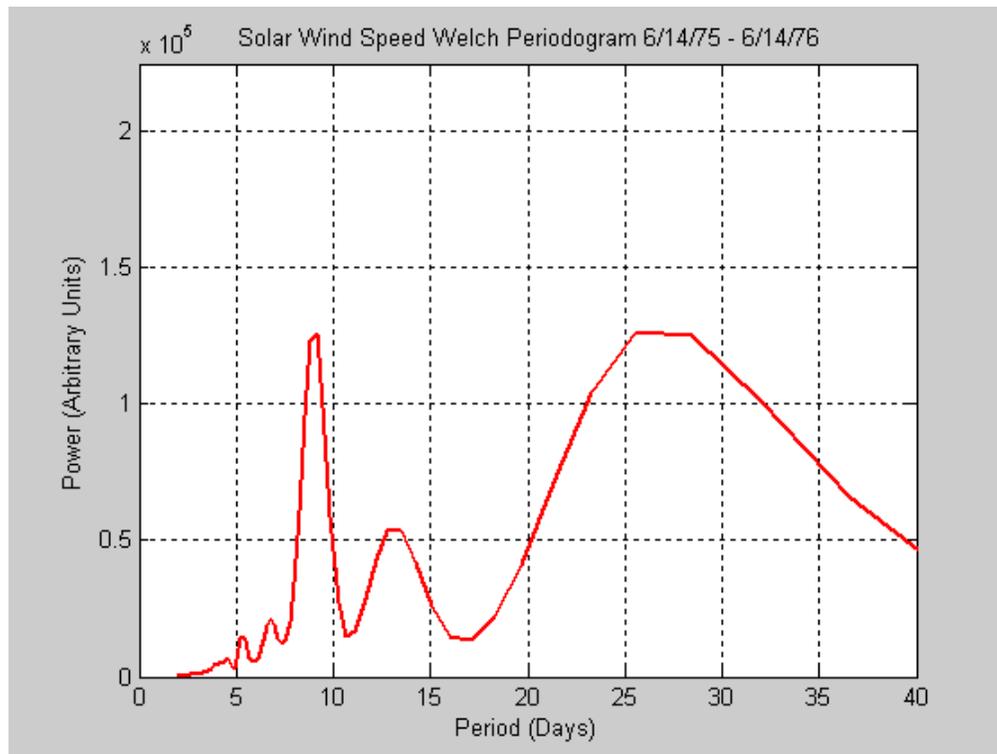
Figure 1.6: Welch Periodogram for the Solar Wind Speed Measured by Pioneer 10 June 14, 1975 - June 14, 1976

photographs below using visible light will illustrate crucial differences in the Sun's corona throughout the solar cycle.

The high temperature solar corona emits ultraviolet and x-rays; invisible radiation which is potentially hazardous, while the photosphere (surface) is visible to our eyes. The left photograph in figure 1.7, from the High Altitude Observatory (HAO), shows the corona during a solar eclipse on February 16, 1980 when the Sun was near the solar maximum of SC-20.
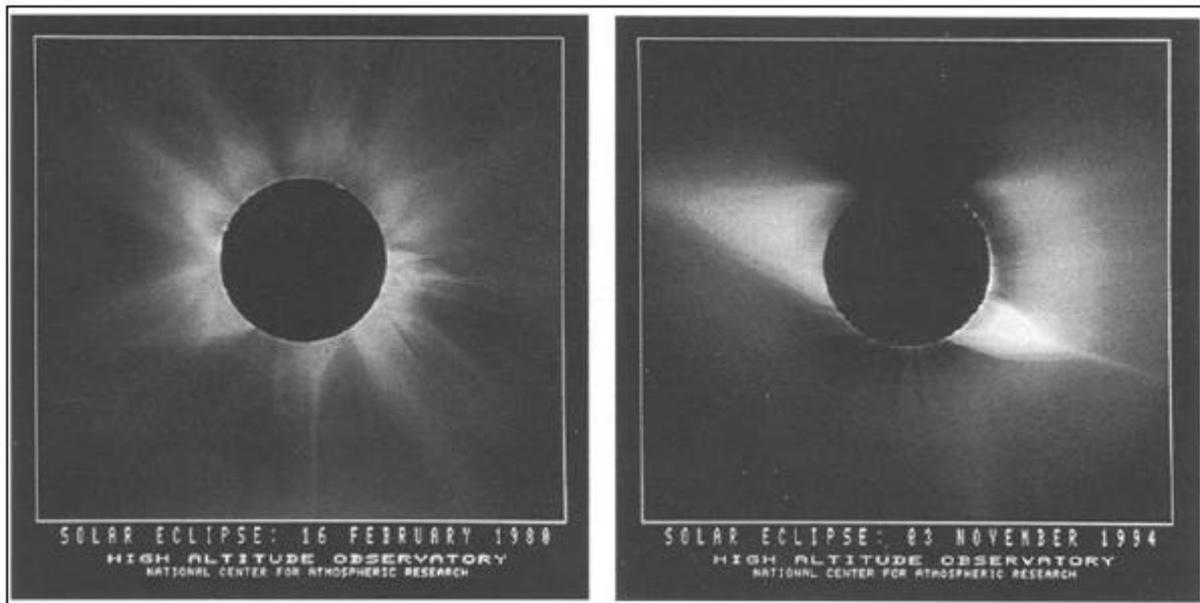
Figure 1.7: White Light Photographs of the Occulted Solar Corona. (left) Solar Maximum; (right), Solar Minimum from Low [36]

The pointed "church steeples" radiating from the corona are helmet streamers, which form a boundary between open field lines and closed ones within the streamer. Closed loops within the streamers are cooler than the surroundings and appear dark in contrast. The small bright spot at the base of many streamers in these diagrams indicates a solar prominence. At times, the solar prominence may erupt and blow out the streamer from the Sun, which is one-way to form a CME (Kivelson and Russell [17]). The association of CME's with solar flares suggests there may be other ways [17].

By contrast, the right photograph from HAO shows the corona during a period of solar minimum on November 3, 1994. The right of the photograph shows one helmet streamer. The extended dark regions at the poles of the Sun are coronal holes. These dark regions define open field lines to interplanetary space. Closed field lines cover the solar equatorial region.

Leaving aside for the moment transient phenomena such as CME's, the solar wind consists of fast and slow winds. A clock diagram for solar wind speed based on measurements from the SWOOPS (Solar Wind Observations Over the Poles of the Sun) experiment onboard Ulysses, is shown if figure 1.8. The left hand diagram illustrates the latitude dependence of the solar wind speed for solar minimum during the first orbit (1994-6). The right hand diagram illustrates the same dependence for solar maximum during the second orbit (2000-2002). The sunspot number below shows the solar activity for each Ulysses orbit.
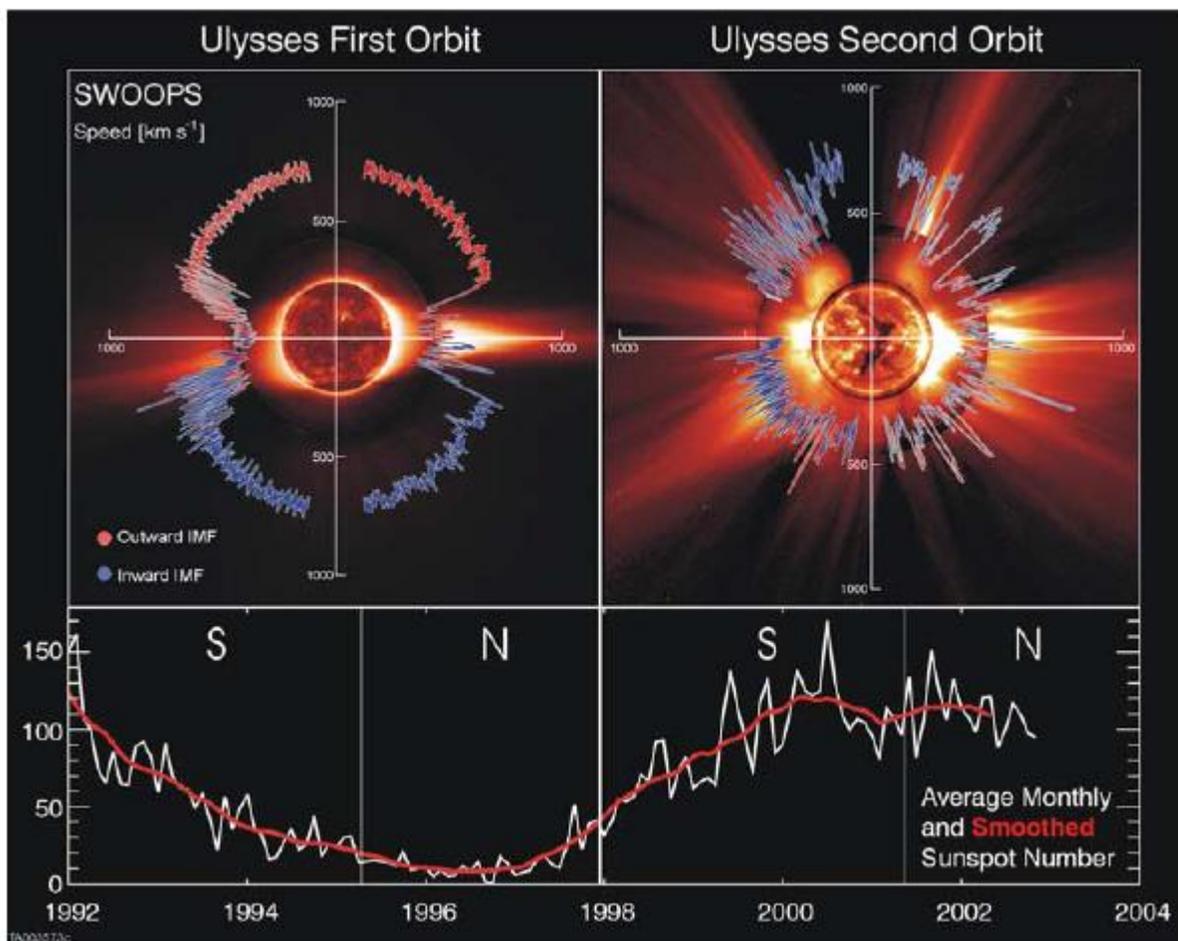
Figure 1. 8: Solar wind speed latitude dependence from Ulysses [37].

The Ulysses mission during solar minimum show the slow wind (~ 400 km/sec) with high variability fills the equatorial region. The fast wind (~ 800 km/sec) with low variability fills the remaining region. During solar maximum, the fast wind appears in many high speed-streams which are not limited in elevation and the slow wind appears everywhere outside of these streams. Because the Earth is in the ecliptic plane, closely aligned with the solar equatorial plane, it sees mainly a slow wind during solar minimum. However, HSS, which emanate from trans-equatorial coronal holes, may also occur during solar minimum. When the coronal holes persist for several solar rotations, the HSS in the solar wind periodically drive the Earth's magnetosphere including the radiation belt, the focus of this study.

The fixity of the 27-day periodicity in HSS may seem surprising given that the equatorial regions of the Sun rotate faster than Polar Regions. The rotation rate so described regards the Sun's surface (i.e. the photosphere). The corona and coronal holes rotate more rigidly than the photosphere and the period is only weakly latitude dependent [38]. Although the latitude dependence in rotation period varies with solar cycle phase, this dependence is small and ignorable for our purpose [39].

1.3.1.   High-Speed Streams (HSS) during the declining phase

Investigations of the 27-day periodicity cited earlier in the Introduction shows a definite trend towards the declining phase and solar minimum. In a critical assessment of these investigations, it is important to note negative results as well as positive ones to best illustrate this trend. As mentioned in the Introduction, Paulikas and Blake in their 1979 paper found periodicities in the electron flux from ATS-6 during the declining part of the SC-20. In an earlier

1976 paper, they associated the observed enhanced electron flux with the passage of magnetic sector boundaries during this declining phase. Because the magnetic sector boundaries in the solar wind rotate with the Sun, this association with enhanced electron fluxes was a natural explanation for the observed 27-day (or a sub-multiple) periodicity.

However, they also mention in that 1976 paper, an earlier effort to determine a relation using ATS-1 data from late 1966 to 1968 during an ascending phase of the solar cycle rather than descending phases. They found they could not correlate the sector boundaries with enhanced flux using ATS-1 measurements [40].

Later, in this 1976 paper they contend the solar activity level determines the extent of the modulation of the enhanced flux from the solar wind.   They concluded that during high solar activity, the presence of magnetic storms masks the periodic modulation of the enhanced electron flux. For another negative finding, Baker, Blake, Klebesadel, and Higbie, also note only infrequent and sporadic flux increases during the ascending phase of SC-21, from 1979-1981, (see reference 21). In view of both positive and negative findings, today we can say that a 27-day periodicity in the solar wind speed due to high-speed streams, most prevalent in the declining phase of the solar cycle, drives a similar periodicity in the relativistic electron flux.

1.3.2.   Coronal Mass Ejections (CME) during the Ascending Phase

Fast solar wind from HSS can efficiently accelerate electrons to relativistic energies. During 2003 to 2005 the time series for solar wind speed superposed on the time series of SAMPEX 2-6 MeV electron flux measurements strongly suggests the electron flux in enhanced

when the solar wind speed exceeds 500 km./sec [41]. However, CME's can also drive fast solar wind and enhance electron flux.

Miyoshi and Kataoka conducted a SEA; henceforth referred to SEA-MK of CME and CIR driven storms during SC-23. They distributed the events into CIR driven storms, CME driven storms and great CME driven storms. They defined their epoch time at minimum storm-time disturbance index (Dst). The maximum solar wind speed in CIR driven storms occurred with a twelve-hour lag from the epoch time and was approximately 570 km/sec. Alternatively, the maximum speed for CME driven storms occurred at the epoch time and was barely 500 km/sec. Miyoshi and Kataoka define "great" storms, if the minimum Dst was below -130 nT. Their SEA-MK showed the great storms had a maximum speed (also at epoch time) of 600 km/sec.

In general, the plasma speed in CME driven storms is somewhat lower than CIR driven storms; (note that the SEA-MK shows it still is at or exceeds 500 km/sec on average, and remains capable of enhancing electron flux). One should note, however that fast solar wind exceeding 800 km/ sec. is always the result of powerful CME's. For example, the solar wind from two CME's driven shocks during the Halloween storm October 29-30 2003, reached 1850 km /sec [42].

Because CME's can enhance relativistic flux, it is appropriate to examine their prevalence over the solar cycle. In Figure 1.9, the blue plot shows the total 7-day total CME number, obtained by distributing CME onset times from the (SOHO/LASCO) CME catalog into weeks throughout SC-23. The red plot shows the 7-day total sunspot number obtained from the OMNIWEB database, divided by 4 for ease of comparison. This scaling does not alter the conclusion.
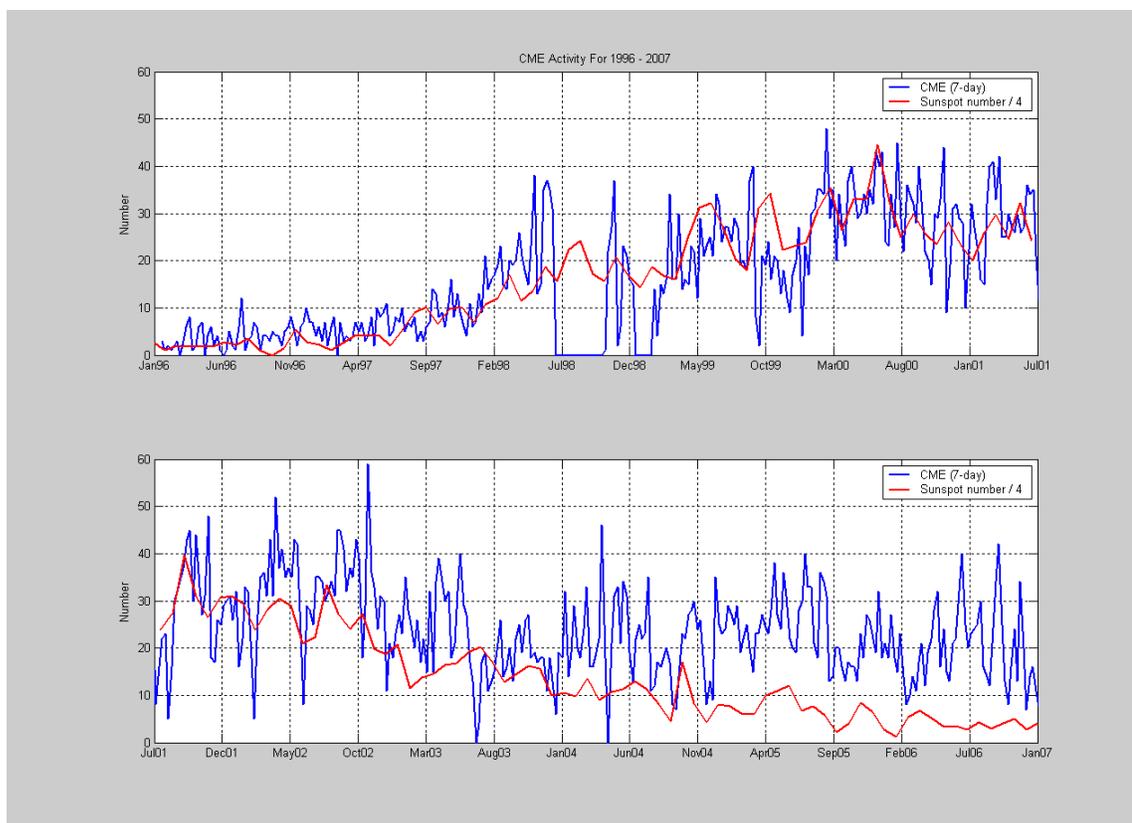
Figure 1.9 Comparing CME Number to Sunspot Number during SC-23.

I should warn, Figure 1.9 shows zero CME number during times when the LASCO coronagraph is not operating. The longest downtime interval was 114 days from late June to mid-October 1998. The second longest downtime was 44 days throughout late December 1998 to early February 1999. Removing these two intervals leaves 18 downtimes each from four to ten days, and the remaining 700 downtimes are shorter than 4 days. The 7-day averaging results in non-zero CME number in all but the two longest downtimes.

Figure 1.9 shows CME's occur most frequently during solar maximum. For SC-23 this was approximately May to August 2000. They may average more than 4 per day during solar maximum up from 2-3 per week during solar minimum. In addition, the CME number follows the solar activity cycle convincingly during the ascending phase of SC-23.   However, the

number of CME's after 2004 shown in Figure 1.9 is greater than anticipated from the diminishing sunspot number leading into SC-24.

Gibson et al. compared the minimum leading into SC-24 during WHI to the preceding solar minimum during WSM noting among other differences, more numerous and stronger high-speed streams, a weaker solar polar field, lower solar wind density and lower speed both measured at the Sun's poles for this most recent minimum SC-24. It would be interesting to see if the CME number tracks the sunspot number for a more typical solar minimum. Unfortunately, the SOHO launch in December 1995 was too recent to provide the SOHO/LASCO catalog of CME onset times sufficiently long before earlier solar minimums.

1.4 Superposed Epoch Analysis (SEA)

The superposed epoch analysis (SEA) is a graphical technique where several time series of a common class of data (e.g. solar wind speed) is plotted relative to an epoch time set at zero after defining an epoch time (e.g. the time of minimum Dst), common throughout the class of data. Several SEA's will be presented throughout the remainder of the introduction.

1.4.1. Relevant Solar Wind Driver Differences between CIR and CME driven storms.

Before introducing differences between CIR and CME drivers, I should qualify that storm occurrences are not equivalent to electron flux enhancement. Geomagnetic storms may elevate electron flux in the outer belt; however, only about ½ of the storms enhance radiation belt flux; ¼ lower the electron flux, and ¼ leave the flux unchanged [see Reeves et al., 43].

Nevertheless, Reeves et al. found high solar wind speed increases the probability of large flux enhancement in the outer belt.

Turner et al. conducts a SEA (referred to as SEA-T) similar to Miyoshi and Kataoka for solar wind conditions from 91 CIR-driven storms and 118 CME –driven storms from 1995-2005. The solar wind speed from CIR drive storms peaked approximately 12-15 hours after the minimum Dst epoch, just as in the (SEA-MK). The solar wind speed for CME driven storms experienced no delay. The peak solar wind speed from SEA-T for CIR driven storms was 550 km/sec, the peak for CME driven storms was 525 km/sec, close to the speeds from SEA-MK.

I supplement the SEA-MK and SEA-T with Kanekal's SEA for solar wind speed for the two years, 1994 and 1997. In order to avoid confusion I note, unlike the earlier SEA's, Kanekal defines the epoch time as the day where the solar wind velocity attains a maximum. The solar wind velocity is normalized to one at the epoch time. This SEA for 1994-5 for high-speed stream events (top) shown in figure 1.10 demonstrates the solar wind speed declines more slowly for HSS (i.e. CIR) events than for CME events.

The average solar wind speed was 685 km/sec and 556 km/ sec. for high-speed stream and CME events respectively. Minimum Dst values were -43 nT and -69 nT for high-speed streams and CME events respectively.

The SEA-MK already introduced the solar wind speed differences between CIR and CME driven storms. It has long been established, that the southern component of the magnetic field also plays an important role to electron flux enhancement. Paulikas and Blake (1976), use the half-wave rectifier as an analogy to energy entry into the magnetosphere, cite earlier references the role of southern Bz, and describe this importance as the commonly accepted view in their 1979 paper. The minimum southern Bz component from SEA-T averages -12 nT for CME-driven and

-7 nT for CIR driven storms. The southern Bz component for both CIR and CME storms from

SEA-MK is -10 nT, along with -15 nT experienced during great storms.
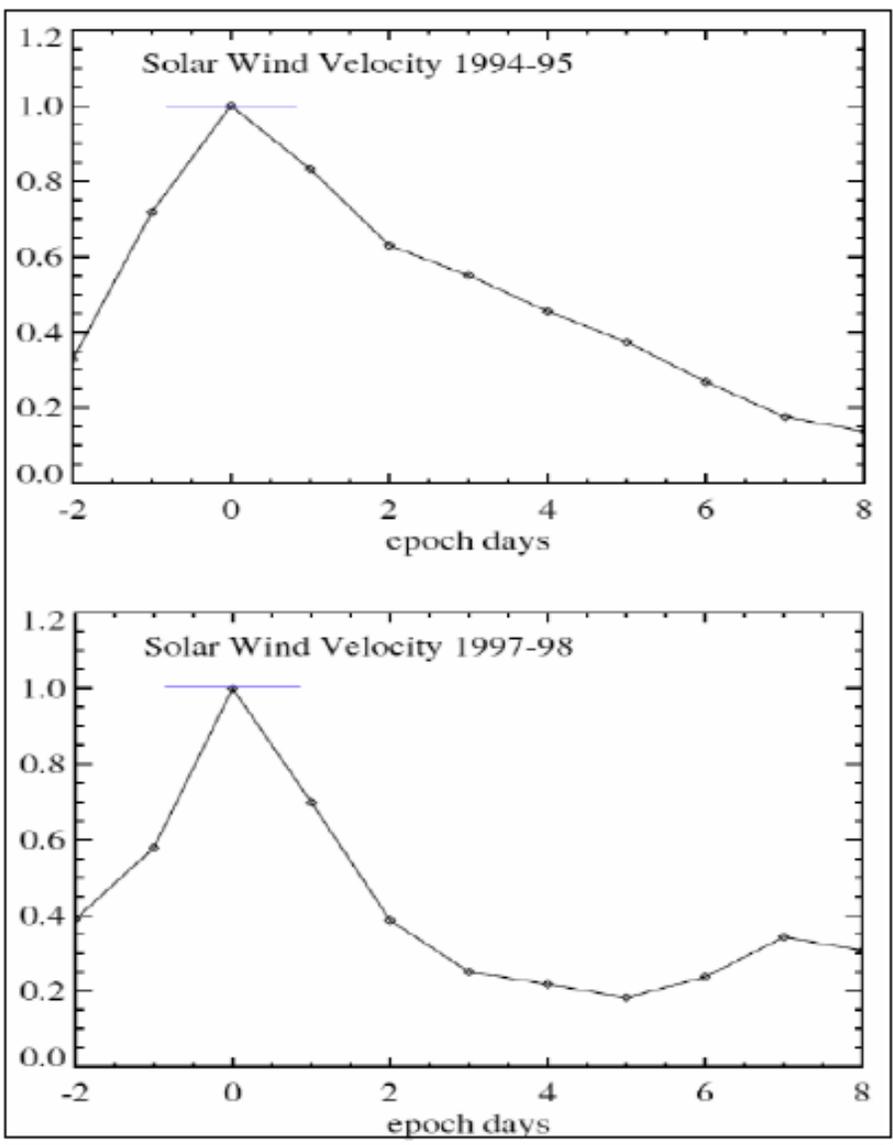
Figure 1.10: Superposed Epoch Analysis from Kanekal [3]. Top: SEA from HSS Events; Bottom: SEA from CME Events

The SEA-T shows the Bz becomes most negative (i.e. southern Bz peaks) approximately

1-hour before the minimum Dst for CIR driven storms, recovers to the pre-storm levels within

about 4-hours, and holds tightly to these levels. Alternatively, the CME driven storms show a prolonged recovery time of 1-day, with greater variance from the pre-storm levels. The SEA-MK finds the same prolonged recovery time to a lesser extent.

The southern Bz allows magnetic field lines to reconnect more readily, increasing energy transfer so difference in the peak and recovery of Bz, demonstrates differences in the effectiveness of CIR's and CME's to enhanced electron flux. Great storms with Bz below -15 nT, should enhance the relativistic electron flux even further. Approximately 1/3 of CME's have magnetic clouds within them [44]. This cloud contains frozen-in magnetic field lines, which may rotate through large angles. In figure 1.11, taken from Burlaga [45], the second panel shows the elevation angle of the field lines relative to the ecliptic as a function of time for June 19-22, 1980. In figure 1.11, the elevation angle δ changes from +90 degrees to –90 degrees in a 16-hour interval on June 20, establishing a 4-hour period with southern Bz component, suggesting this cloud in particular should be geo-effective. One can almost visualize the opening of the Earth's dayside magnetopause with the southern components in the cloud as a "key turning in a lock".

In addition to the characteristics associated with the solar wind, it is appropriate to examine the electron belt response to the storm-time disturbance index Dst, a measure of ring current energy and a definitive indicator of geomagnetic storms. The SEA-T determines the minimum Dst for as -70 nT, and -130 nT for CIR and CME driven storms respectively. The SEA-MK gives -130 nT for both values, with -200 nT for great storms. Both analyses show the Dst for CIR driven storms recovers more slowly to pre-storm levels. Although the Dst given in SEA-T is limited to two days after epoch, the Dst given in SEA-MK show Dst recovery for CIR-storms remains incomplete after 4 days.
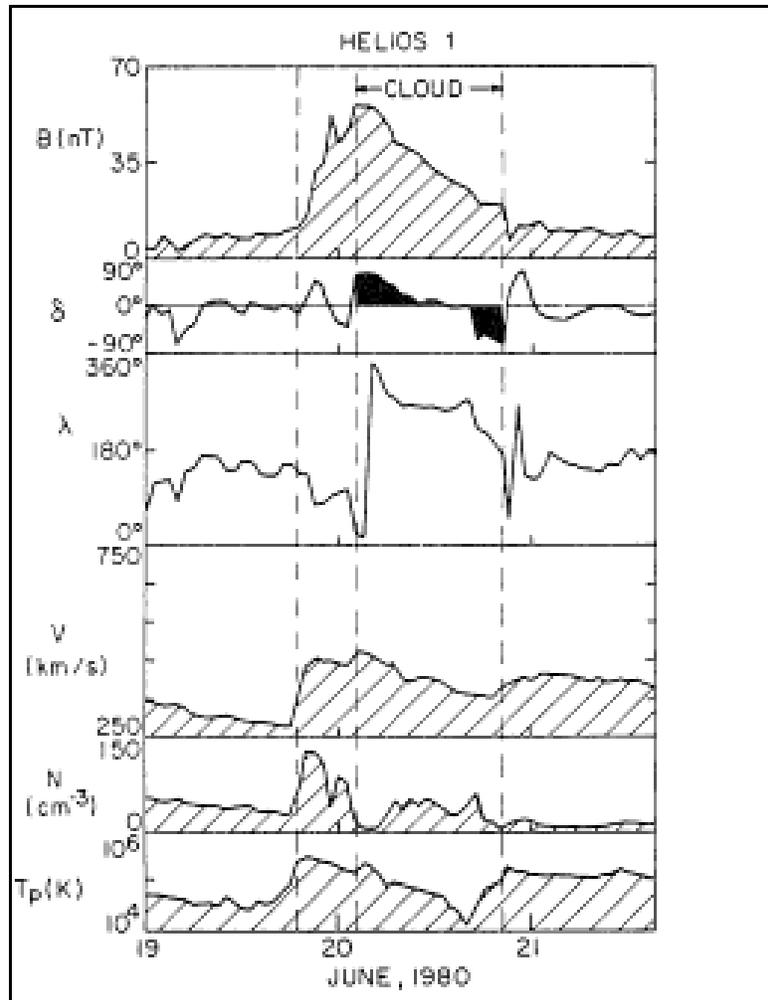
Figure 1.11: Magnetic Cloud Characteristics from Helios 1 Adapted from Burlaga [45]. Panels 1-6 From Top: 1. Magnetic Field; 2. Elevation Angle; 3. Azimuth Angle; 4. Speed; 5. Density; 6. Temperature

1.4.2. Different Electron Flux Responses

The picture beginning to emerge from SEA-T and SEA-MK shows CIR driven storms as

weaker (a Dst which is less negative) with a longer recovery period than CME driven storms.

Differences in the pattern of minimum Dst and recovery time affect the electron flux in the

radiation belts. Borovsky and Denton summarize twenty-one differences between CIR and CME

driven storms [4]. This section will discuss only differences relevant to electron flux enhancement.

Borovsky and Denton reconfirm the difference in prevalence of both CIR and CME driven storms with solar cycle phase, and confirm the CIR driven storms associated with HSS have a 27-day period, where the CME driven storms have irregular recurrence (quasi-periodicity). The most relevant difference from Borovsky and Denton relates directly to the relativistic electron flux.

Figure 1.12 shows the period from 1993 to 2002 sorted in 27.27-day long rows, associated with the Carrington rotation. The black squares in the grey background illustrate periods of enhanced geomagnetic activity evidenced by a planetary index Kp exceeding 4^+. The red lines indicate the lines where the geosynchronous electron flux in the 1.1 - 1.5 MeV channel of the Synchronous Orbit Particle Analyzer (SOPA) energetic particle instrument, exceeds 30 counts/cm^2/ sec/ster /KeV. The yellow dots indicate CIR's for years 1994-6 from the McPherron catalog of stream interfaces but the catalog does not include a "block" of black squares from mid 1999 to early 2000, from day 19 to 24, which shows a 27-day recurrence not associated with solar minimum.

Another notable feature is the two to three day delay between CIR streams and electron flux enhancement shown in the figure as the interval from the yellow dot to the red bar is a commonly observed feature and is consistent with the impulse response delay obtained from the linear predictive analysis by Baker et al. Figure 1.12 demonstrates CME driven storms are less effective in enhancing electron flux.
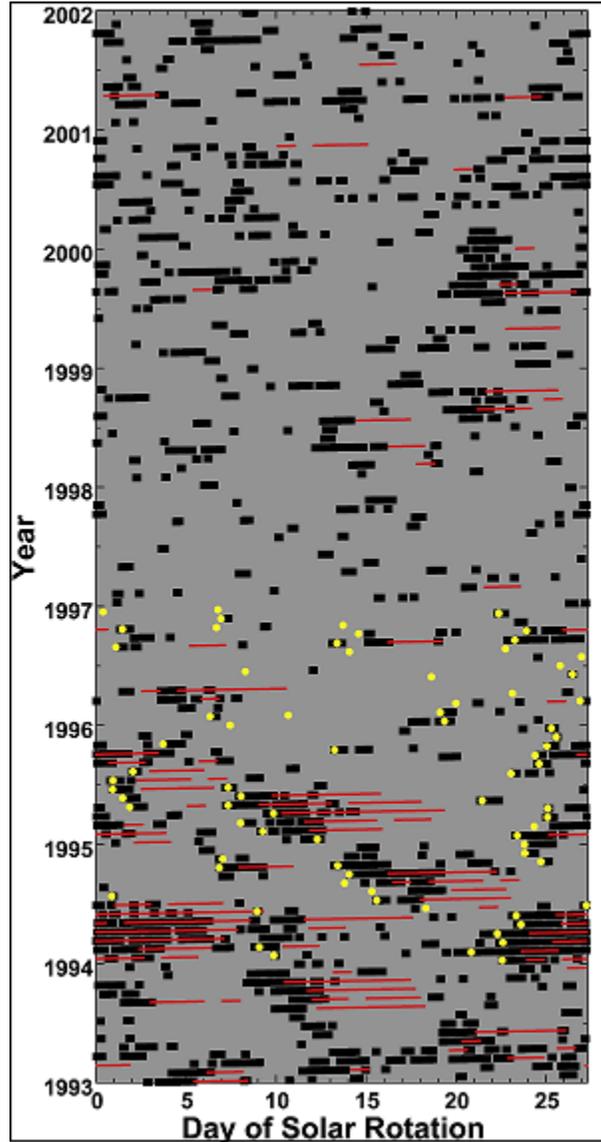
Figure 1.12: High Electron Flux Associated with CIR and CME-driven storms, from Borovsky and Denton [4]

Relativistic Electron flux storm-time behavior shown in Figure 1.12 is limited to geosynchronous altitude L = 6.6. The SAMPEX 2-6 MeV electron flux measurements available at L-shells from 1.1 to 10, has allowed Kanekal to complete a SEA for relativistic electron flux.

Figure 1.13 shows the L-shells between 2 and 7 during 1994 when high-speed streams were prevalent (top), and during 1997 when CME,s were prevalent (bottom).

Note the epoch time is defined as the time where the solar wind achieves its maximum speed. The rise time and decay time for relativistic electron flux from CIR-driven events with associated high-speed streams is 2 and 2 ½ days respectively. For CME events, the rise time is approximately 3 days and the decay time is four days. This decay time is longer for CME events than for CIR events.

As pointed out by Kanekal, the longer rise and decay times for electron flux response noted for CME events relative to CIR events is quite puzzling in light of the SEA from Kanekal for the solar wind speed shown in figure 1.10. This figure indicates the solar wind driver has a lengthier rise and decay time for CIR events than CME events. It is reasonable to expect the relativistic electron flux response (rise and decay times) to mimic the solar wind driver.

Moreover, Kanekal's result is not an isolated case. In their SEA for electron flux > 2 MeV, Miyoshi and Kataoka also find the recovery and enhancement of electron belt flux is faster and stronger in CIR-driven storms than in CME-driven storms, in concert with Kanekal's finding.

Summarizing, this subsection 1-2 has discussed solar cycle influences to the solar wind driver and the relativistic electron belt response. High-speed streams are prevalent in the declining phase of the solar cycle. Coronal mass ejections are more frequent during the ascending phase of the solar cycle.

This closes a lengthy introduction presenting the history of investigations of periodicities in solar wind speed and relativistic electron flux. In addition, the introduction presented

differences in prevalence during solar cycle phase and in effectiveness of HSS and CME's

storms to energizing electrons in the radiation belt.
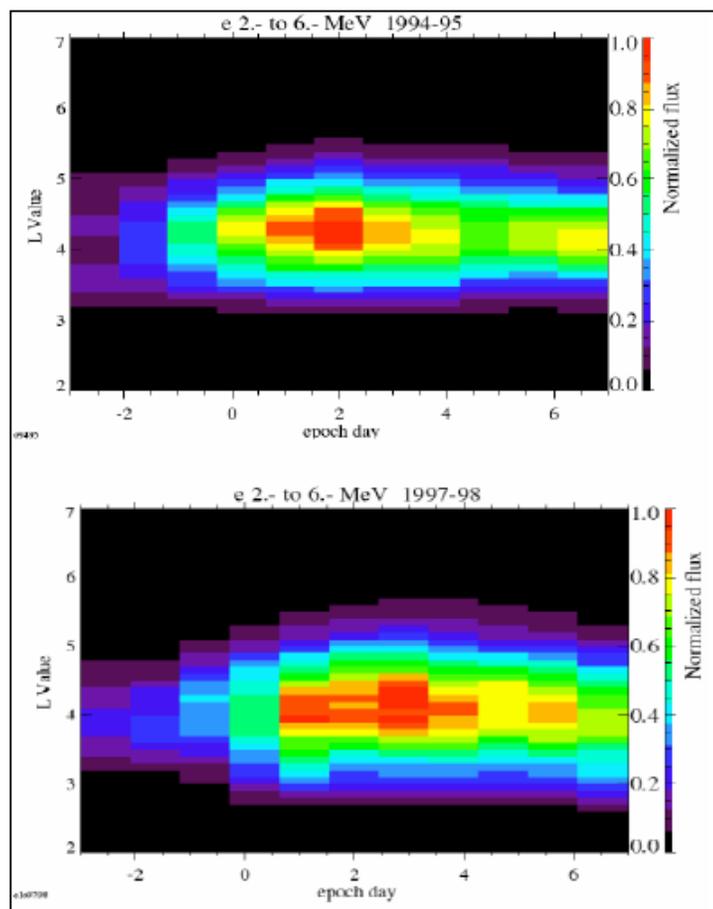
Figure 1.13: Superposed Epoch Analysis of Electron Flux. ILWS Workshop 2006 GOA, February 19-24, 2006 [3]

2. Dynamic Comparison of Periodicities

Investigating periodicities in the solar wind speed and relativistic electron flux in the outer belt requires that we estimate the power spectrum for these measurements. This first section will detail the data sources for these measurements.

2.1. Data Sources

In brief, solar wind speed measurements for the 1994-2010 are available on the OMNIWEB database. The Proton-Electron Telescope (PET) onboard the Solar Anomalous, the Magnetospheric Particle Explorer (SAMPEX) provides logarithmic electron flux measurements from $2 - 6$ MeV at ninety L-shells between 1 and 10. The SAMPEX measurements are incomplete after 2004, although the two year period from 2004-2006 has only a 30 day data gap. In general after 2004, I supplement (i.e. replace for 2004-2006) the SAMPEX measurements with measurement from two GOES satellites, which provide geosynchronous electron flux measurements greater than 2 MeV, from 2004-2010. Unfortunately, GOES relativistic electron flux measurements are limited to geosynchronous altitude $L = 6.6$. The detailed description for the three data sources for this section is provided below.

2.1.1 The Proton Electron Telescope (PET) onboard SAMPEX.

The SAMPEX satellite was launched July 3, 1992 into a high inclination (81.7-degree) orbit with a 98-minute period. The altitude above the Earth surface varies from 520 to 670 km. [46]. This nearly polar orbit cuts the Earth's dipole magnetic field to allow the instruments to record electron flux measurements for a large range of L-values.

The PET instrument is a stack of solid-state detectors to measure the flux for electrons in three energy ranges: The ELO channel measures 2-6 MeV, the EHO channel measures 4-15 MeV, and the EWG (Electron Wide Geometry) channel that measures energies 4-30 MeV [47]. The only SAMPEX measurements presented in this thesis are from the ELO channel.

The electron flux measurements are highly variable. This thesis treats the SAMPEX electron flux using a procedure outlined in Kanekal et al. [48]. A summary of this procedure follows. First, the electron flux measurements are distributed into 90 L-shell bins 0.1 L wide to span the range from L = 1.1 to 10. This sorting is based on satellite location using the International Geomagnetic Reference Field (IGRF). Daily averaging for ~ 16 orbits smoothes the relativistic electron flux. Kanekal et al. take the base 10 logarithm before averaging to decrease the sensitivity of the average to unreliable measurements, i.e. outliers. In addition, Kanekal et al. remove outliers, which are defined as measurements over an order of magnitude higher than the measurements in the neighboring bins. Most often, these outliers are due to a faulty dead-time correction. Finally, a 3-bin running average smoothes these binned values.

2.1.2 The Energetic Particle Sensor Onboard GOES

The GOES (Geostationary Operational Environmental Satellite) are weather satellites, located over the Earth's equator at geostationary altitude (35790 km). This thesis uses electron flux > 2 MeV measured from the Energetic Particle Sensor (EPS) onboard GOES-11 for years 2004-6, and 2009; and uses the EPS onboard GOES-12 for years 2006-2008. GOES-11 is located at 135 degrees west longitude, over the mid-pacific ocean, and GOES 12 is located 75

degrees west longitude over the Amazon river. Because of its restricted altitude and zero inclination orbit, measurements of particle flux are limited to L = 6.6.

The EPS on GOES satellites consist of a telescope detector and three dome detectors (numbered 3, 4, and 5 in Baker et al. [46]. The telescope detector measures lower energy protons, and alpha particles. Dome detector three onboard GOES separately measures the relativistic electron flux > 0.6 MeV, and > 2.0 MeV. This consists of three sets of silicon barrier detectors 1500 μm. thick, surrounded by an aluminum moderator to provide the energy threshold for electrons [49]. The specific daily averaged electron flux measurements from 2004 to 2010 for GOES, for energy channels > 0.6 MeV (unused) and > 2.0 MeV, were obtained from the National Weather Service, Space Weather Prediction Center website [50].

2.1.3 OMNIWEB

The daily averaged solar wind speed was obtained from the OMNI WEB Data Explorer on website, maintained by the National Space Science Data Center (NSSDC) [51]. In general, the solar wind variables are compiled from 15 satellites, and three spacecraft (ISEE-3, WIND, and ACE). A listing of every spacecraft along with the time span for inclusion into the database is given in Table 2 of the OMNI 2 documentation [52]. Because I consider the period from 1994 to 2010, that only WIND, ACE, IMP-8, and Geotail are operational and can potentially supply the solar wind speed over any part of that period. After 2005, only ACE and Wind were operational. The WIND orbit is difficult to characterize, but it orbits mostly between L1 and L2 Lagrange points. The ACE spacecraft loiters around the L1 Lagrange point, approximately 1.5 million km sunward from the Earth; IMP-8 is an Earth satellite. Measurements from these WIND, ACE and Geotail are time shifted to the Earth and

hourly averaged before they are included in the OMNI 2 database. Measurements from the IMP-8 satellite are local and do not require time-shifts.

In addition to solar wind speed, the OMNI 2 Database has supplied me with the 27-day averaged sunspot number, and the interplanetary magnetic field. The sunspot number characterizes the phase in the solar cycle shown in figure 1.7.

2.2 Estimating the Power Spectrum

The standard practice in treating periodic processes is to use an interpretation in terms of frequency, instead of period. The two interpretations although inversely related are equivalent. Because it is likely to be cumbersome discussing periodicity without adhering to standard practice, I will discuss the concepts in this section, in terms of frequency, not period.

The introduction has already provided examples of several tools used for the associating relativistic electron flux measurement to solar wind speed measurement time-series. Plotting the time series usually gives the first indications of a recurring pattern of measurements. This technique is especially effective when a physical basis suggests specific frequencies. For example, observing that the solar wind speed from high-speed streams recurs every 27 days, might suggest an association with the rotation of the Sun. It might also suggest looking for a corresponding frequency in the relativistic electron flux. This correspondence can be motivated by considering the electron flux to be the response to the solar wind speed (driving force or driver).

As an example, we may examine this idea by suggesting at least tentatively, that the relativistic electron flux in the outer belt responds linearly to the solar wind. It is well known the

response of a linear system to an impressed frequency, matches the impressed frequency in steady state. Therefore, it is reasonable to look for frequencies in the electron flux, which match frequencies in the solar wind (i.e. examine the time series for the same 27-day recurrence period).

Plotting the time series and overlaying sinusoids with judiciously chosen frequencies and comparing is quite effective. However, there are better mathematical algorithms and concepts to determine frequencies present in time series, which are less subject to human discriminating abilities. One such concept, can be interpreted in continuous time (Fourier transform) or discrete time (Fourier analysis). Because the discrete time concepts are easier to interpret, a brief discussion of Fourier analysis follows. After defining the periodogram and its characteristics, I give simple and useful examples for the power spectral density (PSD) and auto-correlation from Gelb [5].

2.2.1 The periodogram, power spectral density and autocorrelation

One can usually assume a time series of data (usually physical measurements) x(n), consists of a sum of (circular) trigonometric functions with discrete frequencies.

$$X(e^{j\omega}) = \sum_{n=0}^{N-1} x(n)\ e^{-j\omega n} \Leftrightarrow \int_0^\infty x(t)\ e^{-j\omega n}\ dt \qquad 2.1$$

The (discrete) Fourier series, $X(j\omega)$ in equation 2.1 of the data is given as a function of complex frequency. The Fourier analysis becomes the Fourier transform in the continuum limit shown after the arrows on the far right side of the equation.

For statistical applications, a least-squares analysis is often desired. More broadly, for some physical applications, squaring the time series is justified. For example, if the time series were current measurements, squaring the series (for constant resistance) would give us power. This motivates referring to the periodogram as the PSD. The preference to treat real rather than complex variables provides yet another reason to square the time series.

The average square of Fourier transform of the time-series $X(j\omega)$ defines the periodogram, $I_N(\omega)$ below:

$$I_N(\omega) = \frac{1}{N}\left|X(e^{j\omega})\right|^2. \qquad\qquad 2.2$$

Redefining a double summation index m, leads to equations 2.3 and 2.4. The equivalent form for the periodogram is given by:

$$I_N(\omega) = \sum_{m=-(N-1)}^{(N-1)} c_{xx}(m)\ e^{-j\omega m}, \qquad\qquad 2.3$$

where the coefficients $c_{xx}(m)$, are the autocorrelation estimate defined as:

$$c_{xx}(m) = \frac{1}{N} \sum_{0}^{N-|m|-1} x(n)\,x(n+m). \qquad\qquad 2.4$$

Equation 2.3 demonstrates the periodogram is the Fourier transform of the auto-correlation estimate (equation 2.4) of the square of the time series. The autocorrelation and PSD are dual to each other. That is, the PSD represents the time series in the frequency domain, the autocorrelation represents the time series in the time-domain. The transformation between representations is the Fourier and Inverse Fourier transform respectively.

2.2.2 Common Examples of Power Spectral Density and Autocorrelation

Figure 2.1 from Gelb shows a few useful examples for the autocorrelation function and the power spectral density for four common stationary stochastic processes. The abscissa for the autocorrelation is the time lag, the abscissa for the power spectrum is the frequency. The ordinate is amplitude for autocorrelation and power for the PSD for each example.

White noise shown in the top panel, is perfectly correlated with itself for time $t = 0$, otherwise there is identically zero (i.e. no) correlation. Such a series is entirely unpredictable, and contains no predictive information. The Fourier transform of the delta-function autocorrelation results in the power spectrum for white noise, which is constant i.e. all frequencies contribute equally to the power spectrum. This mathematical artifice is useful as an approximation but white noise across all frequencies would contain infinite total power, which would be unphysical.

The term "white" noise is borrowed from optics, where the term "white" light is often used. Sir Isaac Newton recognized that a glass prism separates "white" sunlight into its constituent frequencies (i.e. colors). The solar radiation spectrum is approximately flat for frequencies receptive to the human eye (4000-7000 A). Glass in the prism filters the light selectively by presenting a frequency dependent index of refraction to the light resulting in separating the observed colors. Hence, white light is an approximately equal mixture of all frequencies.

The delta function autocorrelation from white noise has an infinitesimal correlation time τ (lag) approaching zero. Earlier values of the random variable are entirely independent of later values. Alternatively, one can conceive of processes where the value of a random variable

Process	Autocorrelation Function (ϕ_{XX})	Power Spectral Density (Φ_{XX})
White Noise	$\phi_{XX}(\tau) = \phi_0\delta(\tau)$	$\Phi_{XX} = \phi_0$
Markov Process	$\phi_{XX}(\tau) = \sigma^2 e^{-\beta|\tau|}$	$\Phi_{XX} = 2\beta\sigma^2/(\omega+\beta^2)$
Sinusoid	$\phi_{XX}(\tau) = (A^2/2)\cos\omega_1\tau$	$\Phi_{XX}=(\pi/2)A^2[\delta(\omega-\omega_1) + \delta(\omega+\omega_1)]$
Random Bias	$\phi_{XX}(\tau) = m^2$	$\Phi_{XX} = 2\pi m^2\delta(\omega)$
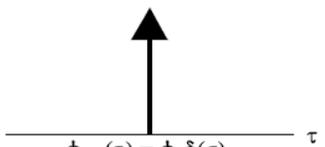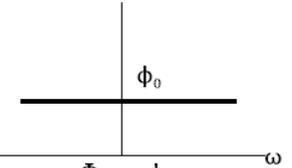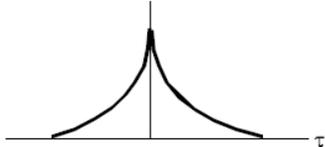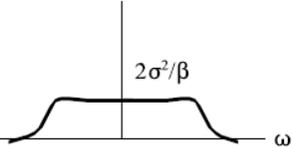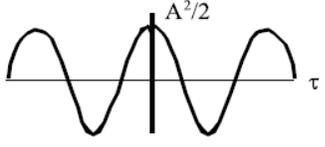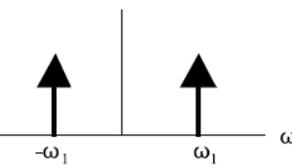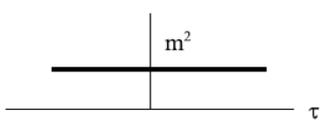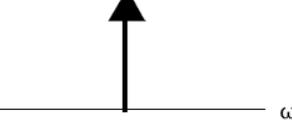
Figure 2.1: Power Spectrum and Autocorrelation for some common stochastic Properties Adapted from Gelb [5].

relates more strongly to earlier values   For example, it may be natural to describe the process by introducing a finite autocorrelation time τ so that value of the random variable at some later time relates in a steadily decreasing manner to the earlier value.

The second panel down on the left shows the autocorrelation for a first-order Markov process with finite autocorrelation time τ. The autocorrelation decreases as an exponential function with increasing time lag τ (i.e. autocorrelation time). The white-noise process investigated earlier is equivalent to this process in the infinite beta limit, where beta shown in the

autocorrelation equation in the second panel is the inverse of the e-folding time parameter. The panel on the right shows (a rather poor graphic representation of the proper equation shown, the representation appears to dip at zero) the corresponding power spectral density with the shape of a Cauchy distribution. This PSD is often called "red" noise since it favors low frequencies over high ones. This particular stochastic process is often obtained from the output from passing white noise (input) through a low-pass filter. The process can be simulated with a RC filter model

The third panel, down on the left shows the autocorrelation for a sinusoidal process, which is the focus of this section. The PSD is a pair of frequencies located at discrete points on the abscissa. It is appropriate to note that because our focus is real-valued data, the PSD is even with respect to the y-axis, so that in the remaining sections, I consider only positive values. Reflection around the ordinate will provide negative frequencies if needed. Characteristics of mechanical and electrical components in systems may drift and be treated by random bias shown in the fourth panel. I will not use random bias in this thesis.

2.3 The Tradeoff between Consistent Estimation and Frequency Resolution

The autocorrelation and PSD for the few examples from preceding sections along with definition and characteristics of the periodogram should facilitate the discussion of spectral estimation. Spectral estimation with the periodogram consists of determining frequencies in the mean square of the time series (see Equation 2.2). It would seem finding Fourier coefficients should be a simple matter. In practice the mathematical procedure, which may be arduous involving integrals calculated analytically often done as textbook exercises, is beset with

difficulties when applied to actual time series data even after evaluating the numerical integrals. The following section addresses the two difficulties encountered in spectral estimation. The first difficulty encountered is obtaining a consistent estimate of the power spectrum can be circumvented by averaging periodograms. However, averaging periodograms leads to lower frequency resolution. This tradeoff is treated with the examples below.

2.3.1 Difficulties obtaining a consistent estimate

Several textbooks demonstrate the difficulty in obtaining a consistent estimate by calculating the variance and covariance assuming a white noise signal [11, 53, 54]. Oppenheim and Schaefer provides the variance and covariance of the periodogram below [11]:

$$Var[\,I_N(\omega)\,] \quad = \sigma^4 \left[1 + \left(\frac{\sin(\omega N)}{N \sin(\omega)} \right)^2 \right]$$

and

$$Cov[\,I_N(\omega_1),\ I_N(\omega_2)\,] = \sigma^4 \left[\left(\frac{\sin(\pi(k-l))}{N\sin(\pi(k-l)\,/\,N)} \right)^2 + \left(\frac{\sin(\pi(k+l))}{N\sin(\pi(k+l)\,/\,N)} \right)^2 \right],$$

where $\omega_1 = 2\pi\,k\,/\ N$, and $\omega_2 = 2\pi\,l\,/\ N$.

The variance σ^2, and similarly the square of the variance is equal to 1 for a un variance white-noise signal. The equation 2.6 demonstrates the covariance is zero unless $\omega_1 = \omega_2$. For $\omega_1 = \omega_2$, equation 2.5 demonstrates the variance is approximately one, independent of N. Because the variance maintains itself at approximately one as N increases, but the covariance approaches zero

at intervals 2 pi / N, as N increases, long record lengths with large N exhibit large fluctuations. This observation is contrary to the expectation that the estimate of the power spectrum should improve with larger sets of measurements.

Oppenheim and Schaefer gives a white-noise example for N = 14, 51, 135, and 452, showing increasing fluctuations with increasing record length N. Figure 2.2 below shows results from Oppenheim and Schaefer. Although, the expected power spectrum for the white-noise power spectrum should be flat, the simulation demonstrates the variance around (any) chosen constant value increases as N increases.

This difficulty is not limited to the special case of white noise. Oppenheim and Schaefer argue for similar difficulties in obtaining an estimate of the power spectrum are encountered when "colored" noise processes are considered by passing white-noise through a linear system. I also offer the relevant statement from Brown and Hwang [53], "Recall that it is the average periodogram that is a measure of the spectral density function. Averaging may not be essential in the analysis of deterministic signals, but it is for random signals." … "In either event, analog or digital, some form of averaging is essential when analyzing noise".

One way to circumvent the difficulty in obtaining a consistent estimate of the PSD is to average K periodograms each with length M = N / K, to keep M small and the variance bounded. This averaging is part of Bartlett's procedure. The variance of the average of K periodograms is calculated in detail in Oppenheim and Schaefer. The resulting variance is inversely proportional to K so that as K gets large, the variance goes to zero. Therefore, the average of K periodograms is a consistent estimate of the power spectrum, as desired. However, for fixed record length N, increasing K, means decreasing M. As a result, the bias in the average of K periodograms is

larger than the bias for the periodogram with length N. This bias shows up as a decrease in

spectrum resolution when averaging the K periodograms.
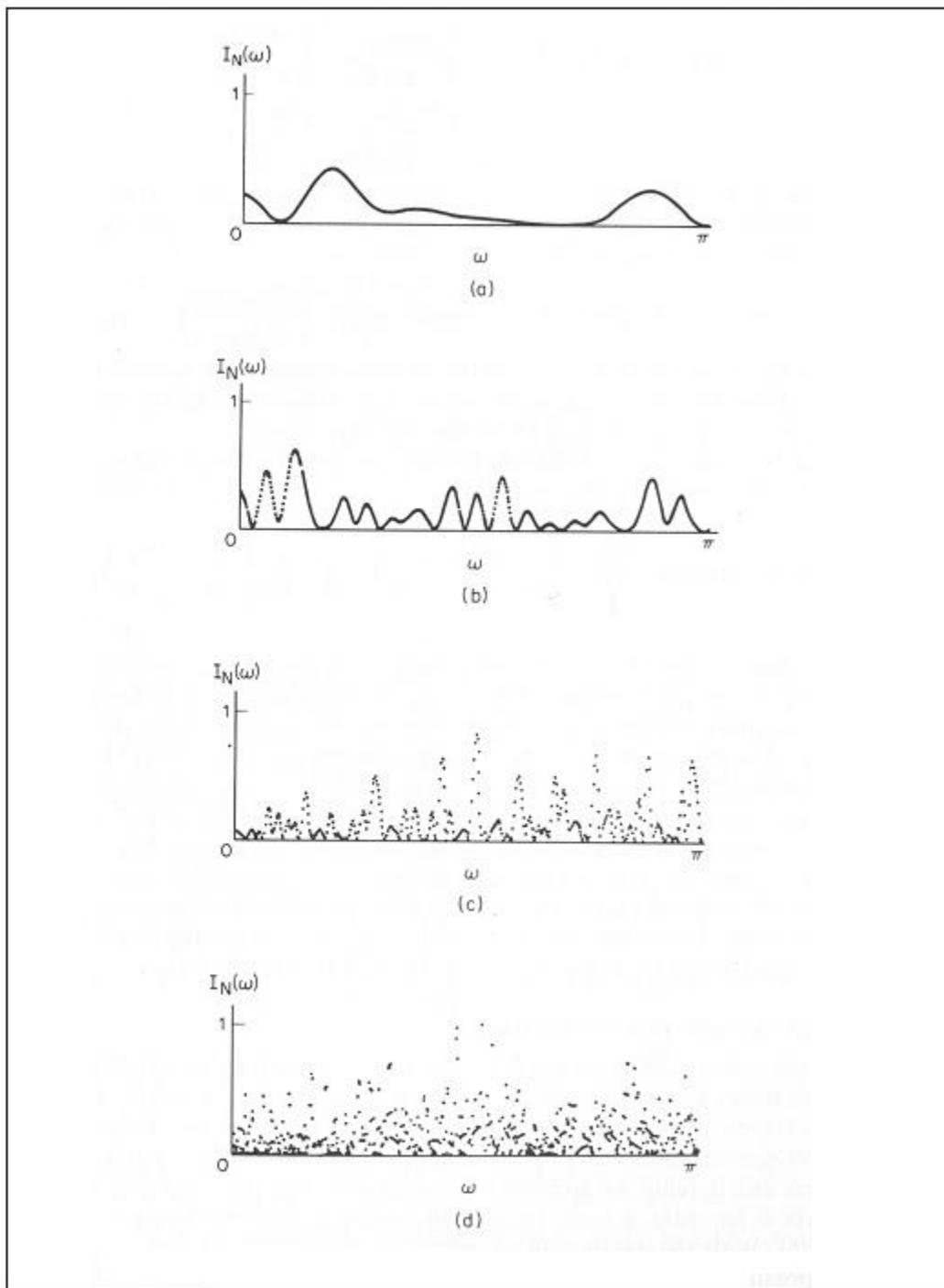
Figure 2.2: Consistent power spectrum estimation becomes more difficult as the record length
N, increases

Figure 2.3, shows the Fourier spectrum for daily averaged electron flux from SAMPEX for L-shell 4.5 during the two-year period 1994-1996. The cyan curve in the right hand plot results from averaging 731 daily averaged values / 20 (the length of each periodogram) or approximately 36 periodograms. The black curve averages over fewer periodograms etc. The figure demonstrates the best frequency resolution comes from averaging the fewest number of periodograms, or conversely, the as the number of periodograms averaged increases (length L decreases), the frequency resolution becomes poorer. The poor resolution asserts itself when many periodograms are averaged, although the most consistent estimate of the frequency occurs when the number averaged is high.
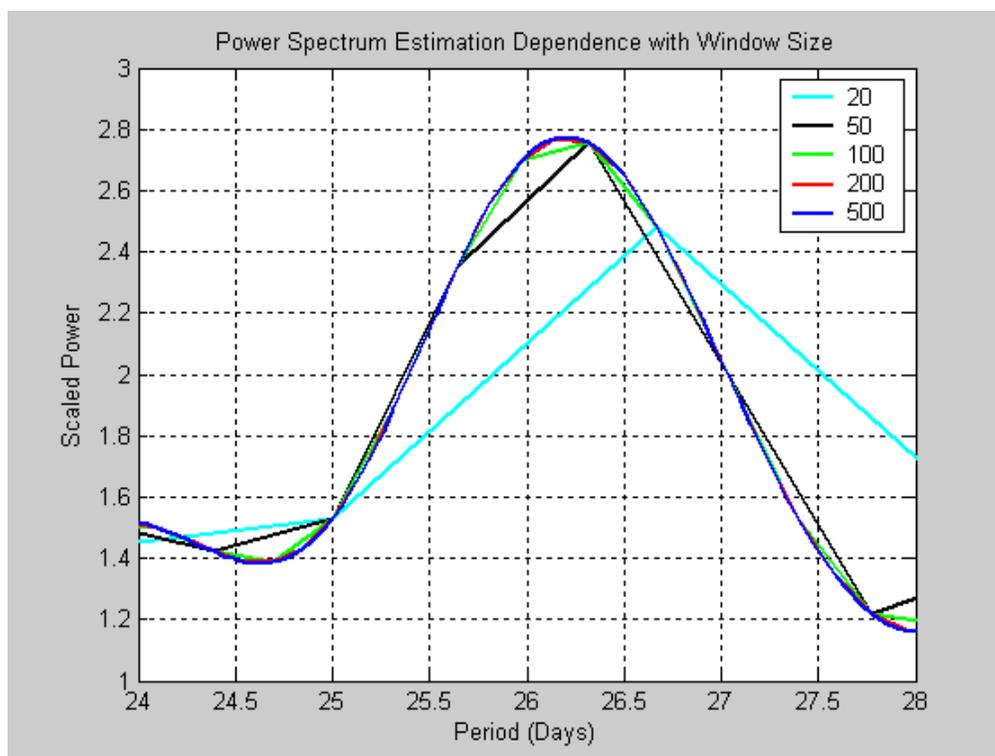
Figure 2.3: Poor Frequency Resolution Resulting From Averaging Over Many Periodograms

Summarizing, estimating the power spectrum from data necessarily involves a tradeoff. When the number of data points (N) becomes large, the variance in the periodogram (i.e. spectral

estimate) does not approach zero and the periodogram is not a consistent estimate of the power spectrum. To circumvent this problem, averaging or "windowing" is a common technique. The data with length N is divided into K segments of segment length M, and the K periodograms are averaged. This leads to consistent estimation. However, dividing the data into K segments and averaging, results in poorer frequency resolution as the segment length M decreases.

Power spectrum estimation requires choosing the length M and the window type appropriately. In practice, the size of the window influences the periodogram more strongly than the window type. The type is important to eliminating sidelobes, but I will not consider this issue. Some specific types are: 1. Modified Bartlett, 2. Daniell, 3. Tukey-Hamming, 4. Tukey-Hanning, 5. Parzen, and 6. Bartlett-Priestly. Only small differences in the power spectrum were apparent between types. Figure 2.4 shows the scaled power spectrum for the solar wind speed near the main peak at 27 days along with a periodogram for the electron flux. Differences in type for other period intervals (regions) which are not shown in the figure supported the conclusion reached in this region. Because type differences did not result in significantly different power spectrum estimates, the periodograms for solar wind speed from OMNIWEB and relativistic electron flux from SAMPEX and GOES, use the Hamming window.

2.3.2 The Welch Periodogram

Welch's algorithm estimated periodograms used in this chapter. This algorithm averages the data segments first before taking Fourier transforms. In addition to averaging (windowing), Welch's algorithm uses overlapping data segments so the periodograms for each segment are not independent. Together, both features lead to smoothing the power spectrum and variance reduction; both are desirable [11]. The degree of overlap was controllable for our

implementation, but I chose to accept the 50% overlap between data segments, the default for the routine PWELCH in the MATLAB signal processing toolbox.
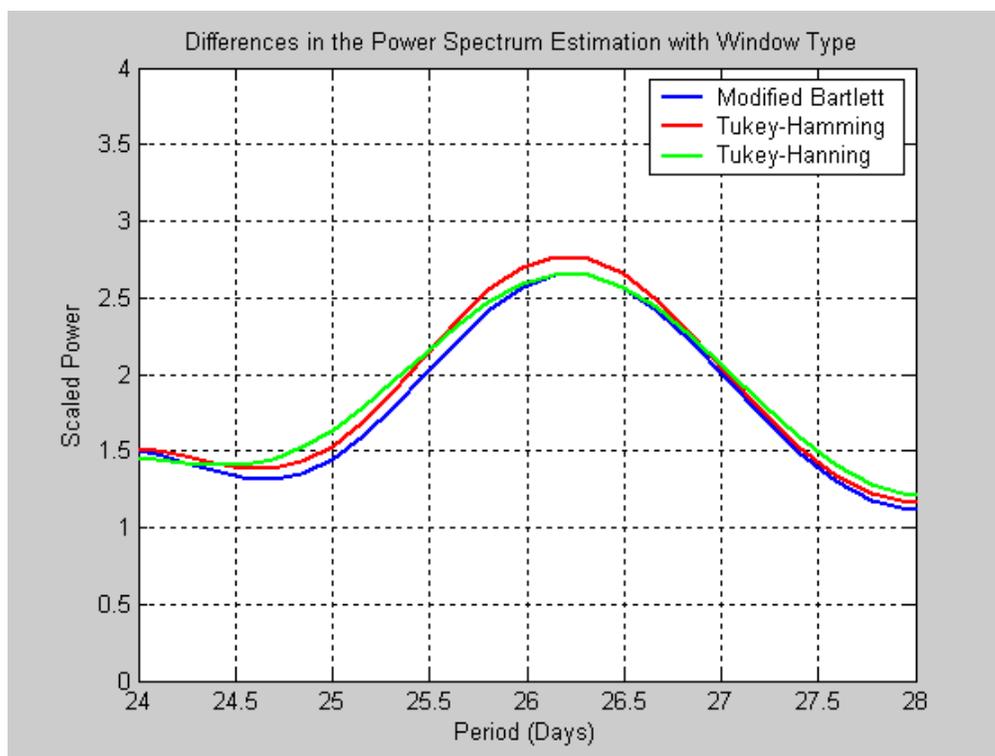
Figure 2.4: The Power Spectrum for the Solar Wind Showed Only Small Differences in Window Type

Figure 2.5 shows the power spectrums for relativistic electron flux in the ELO range (blue) from SAMPEX at $L = 6.5$, and daily averaged solar wind speed (red) during 1994-6 when the solar cycle was at solar minimum. Along with the main 27-day peaks, narrow peaks at sub-harmonic peaks at 13.5 and 9 days also show up in the figure for both solar wind speed and relativistic electron flux. Our choices for overlap and window size effectively reproduce the expected peaks in the solar wind speed power spectrum, and show remarkable agreement between solar wind speed and electron flux. Because we expected a 27-28 day periodicity in the solar wind speed, we used the zero-mean time series in daily averaged solar wind speed to

choose a good window size for the Fourier transform algorithm, PWELCH. The size M = 200 (approximately 28% of the total 730 day data set) was sufficiently large to demonstrate the 27-28 day peak with some clarity (see the red trace in figure 2.5 below). Unless otherwise stated, the criteria for window size will be set at approximately 30% of the total data length N.
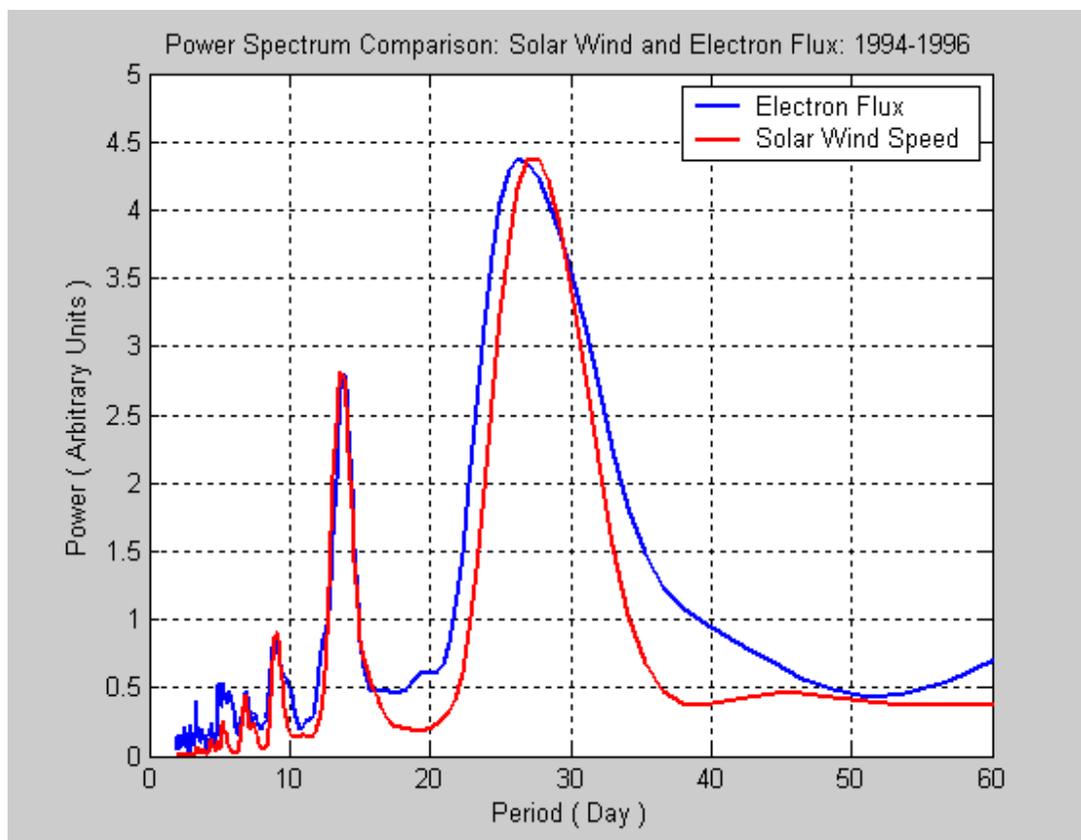
Figure 2.5 Periodogram for Solar Wind Speed (Red) and Electron Flux (blue) at L = 6.5, during the descending solar cycle phase leading to minimum SC-23.

With the choice of algorithm; PWELCH, the window size; (30%), and the window type; Hamming, all characteristics needed to define the periodogram to estimate the power spectrum are set. The next section will show periodograms demonstrating relations between power spectra of solar wind speed and electron flux which vary with L-shell and throughout the solar cycle.

The record of relativistic electron flux from SAMPEX is logarithmic (to base 10)

imposes a degree of artificiality to the process of implementing the measurements and estimating

the power spectrum. The base 10 choice is accidental; another choice of basis is related to that

choice by a particular multiplicative factor. This introduces a multiplicative factor to the power

spectrum, which allows rescaling. The rescaling for the electron flux periodogram shown in

figure 2.5 is set so the power (i.e. amplitude) of the scaling in the 27-day main peak matches the

main peak in the solar wind during the time interval considered. This method of rescaling will

be maintained throughout this thesis, unless otherwise stated.

2.4   First Results: Solar Wind Speed / Electron Flux Periodicities

2.4.1 HSS during the descent to minimum phase of SC-23.

Although HSS were seen at practically all heliospheric latitudes during solar maximum

(see figures 1.7 and 1.8), these streams are disorganized and do not lead to distinct periodicities.

During the descent to minimum phase of the solar cycle, the Earth in the ecliptic (near the solar

equator), generally encounters a slow wind. When HSS do emanate from a trans-equatorial

coronal hole, and they persist over several solar rotations, they are not masked by disorganized

streams as they might be during solar maximum. Instead, evidence of these streams show up in a

distinct series of peaks in the power spectra. These peaks are displayed in Welch periodogram

estimates of the power spectra for solar wind speed (driving force), and in the electron flux

(response).

Figure 2.6 a-f show periodograms for the solar wind speed plotted along with

periodograms for the electron flux at six L-shells during the descent to solar minimum
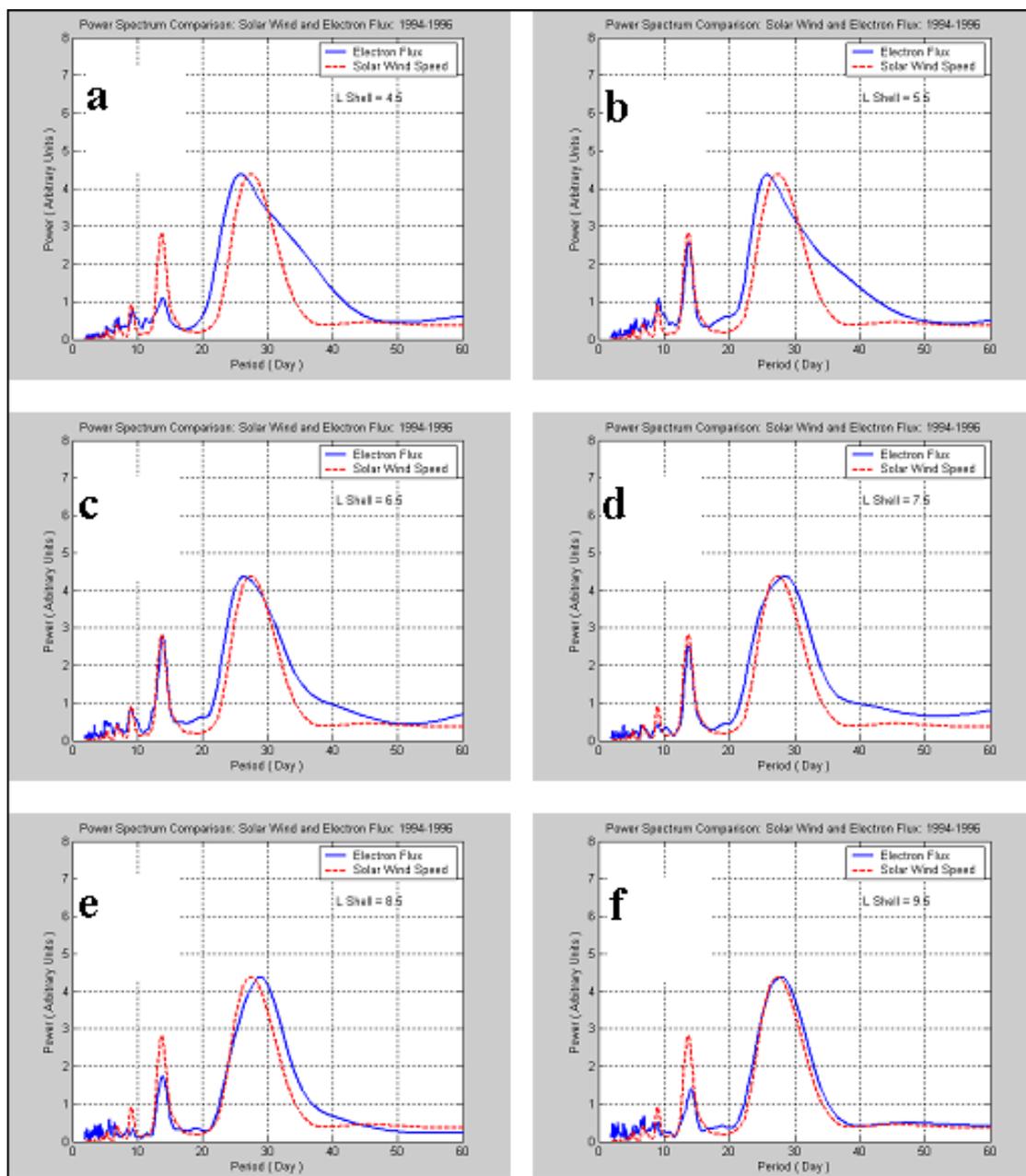
Figure 2.6: Solar Wind Periodograms and Electron Flux Dependence on L-shell (1994-1996).

phase leading into SC-23. The periodograms shown in figure 2.5 are repeated in panel c of

figure 2.6 for the L-shell corresponding to the field line near geosynchronous orbit. Every panel

shows the same solar wind speed power spectrum in order to compare with the 2-6 MeV electron flux from SAMPEX/PET. Both electron flux at all L-shells and solar wind speed periodograms estimate a spectrum with 27-day and 13.5-day peaks.

During 1994-1996, the shape of the broad peak in the electron flux periodogram agrees with the periodogram of solar wind speed at high L-shells, indicating a similar electron flux response for frequencies near the main peak, as well as directly at the main peak where rescaling enforces the periodograms to have equal power. It is interesting to note that the periodograms at lower L-shells display "tails" at slightly longer times than the main period, which are not reflected in the solar wind speed. The tail in the periodogram for electron flux at higher L-shells disappears as L increases to agree with the shape of the solar wind speed spectrum at the highest L-shell shown (L = 9.5). I have no explanation for this tail.

The power in the secondary (13.5-day) peak in the electron flux agrees with the secondary peak in the solar wind speed for intermediate L-shells with best agreement at approximately L = 6.5. Tertiary and higher order peaks at sub-harmonic frequencies (sub-multiples of the 27-day period) for electron flux and solar wind speed are weak.

Comparing the electron flux responses can suggest physical content to examining frequency domain analyses to the solar wind and electron flux. For any particular periodicity, the power transfer from solar wind to electron radiation belts across a variety of L-shells can be examined by comparing the differing electron flux responses from different L-shells noting they respond to the (same) solar wind speed.

For example at L = 6.5 shown in panel c of figure 2.6, the periodogram for electron flux approaches the periodogram for solar wind speed near the 13-day peak. This suggests the power with this frequency (inverse period) in the solar wind drives the electron flux at the same

frequency efficiently. Panels b and d of figure 2.6 show the electron flux periodogram for at L = 5.5 and 7.5 respectively, has slightly less power near the 13-day peak than the electron flux periodogram at L = 6.5, which nearly approaches the power in the solar wind speed periodogram (the same for all L-shells). This suggests the power transferred to the radiation belts from the solar wind near the 13-day peak at these L-shells is slightly less than the power transferred at L = 6.5. Similarly, panels a and e indicate the power transferred near the 13-day peak at L = 4.5 and L = 8.5 respectively is much less than the power transferred to the L-shell at 6.5 or L-shell at 5.5.

2.4.2 HSS during the descent to minimum phase of SC-24

Dissimilarities in the solar minimum leading into SC-24 from SC-23 cited in the introduction motivate us to examine the periodograms for the descent to solar minimum phase during SC-24, shown in Figure 2.7 for the one-year period 2005-2006. Because the interval has been reduced to one-year, the range of periods in the spectral estimation has been reduced from 0 to 60 days to 0 to 30 days. Focusing on the solar wind speed periodograms, the most striking observation is the appearance of a strong narrow tertiary peak at a 9-day period. Broader main and secondary peaks in the power spectrum for solar wind speed have lower strength, although the main peak averaged over its broad frequency range may contain more power than the narrow 9-day peak.

The 9-day peak before Solar Minimum during this time interval has already been noted (see refs [29-35]), and is associated with high-speed streams from trans-equatorial coronal holes (see ref [33]). As before, the electron flux periodogram is rescaled so that the power in the main

peak matches the power in the main peak in the solar wind speed. No special significance should

be attributed to power in the main peak of the electron flux periodogram, which is shown in
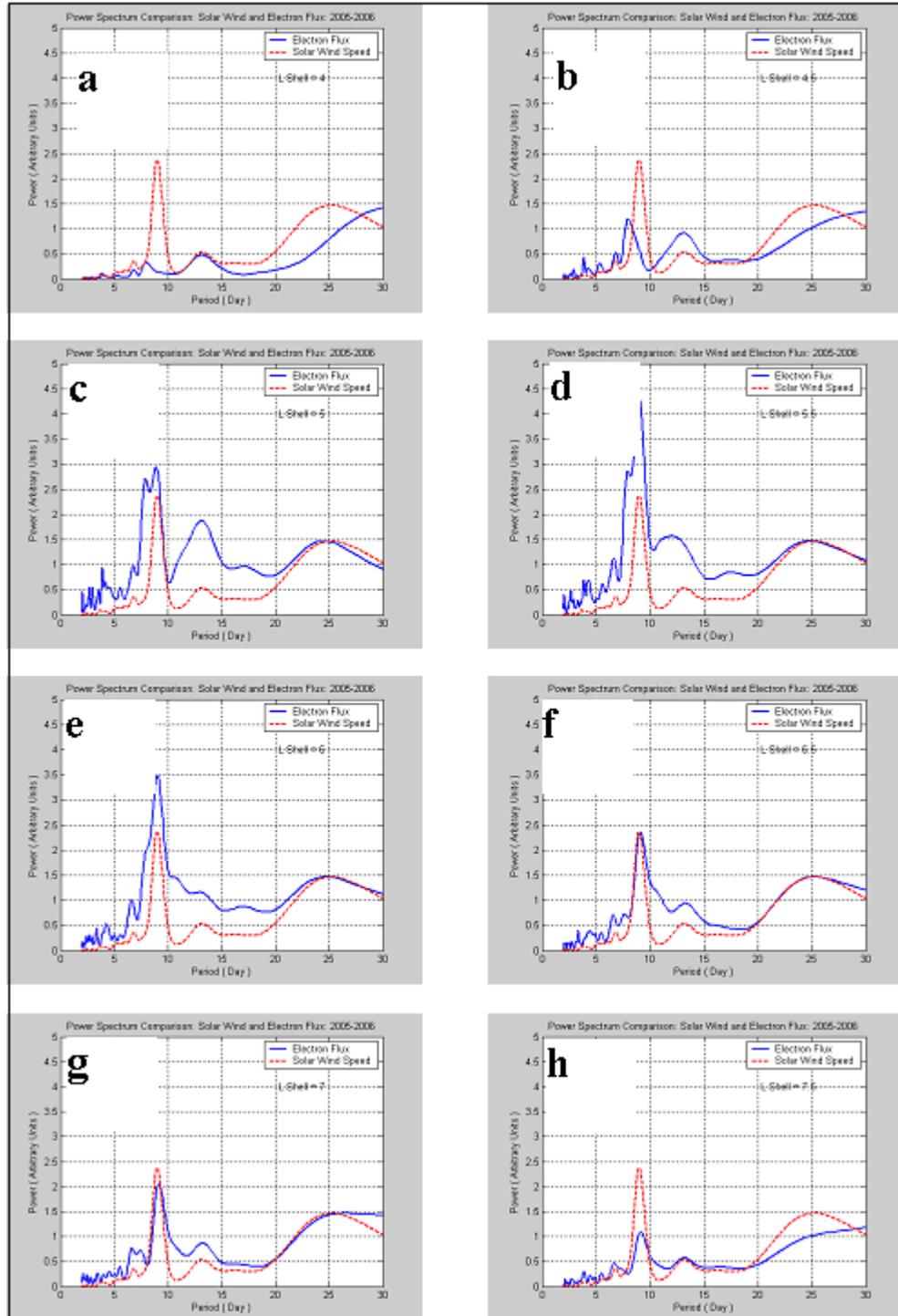
Figure 2.7: Solar Wind Speed and Electron Flux Periodograms during 2005 (Solar Minimum)

arbitrary units. Only ratios of the power in the periodograms at the various frequencies are relevant. However, there is one caveat which should be brought to attention.

In Panels a and b the main peak in electron flux is scaled to the main peak in the solar wind speed periodogram, but the main peak in the electron flux occurs at a significantly different frequency at L-shells below than the main peak for solar wind speed. Artificial rescaling the electron flux periodogram to the main peak in the solar wind speed is misleading in panels a and b.

Unlike the earlier interval approaching solar minimum for SC-23, this interval leading to SC-24 shown in the figure, demonstrates a broad main peak in electron flux which best matches the solar wind speed power spectrum at intermediate, not the highest, L-shells. Diverting away from the main peak and focusing on the entire periodogram shows the solar wind speed and electron flux power spectrum best agree for L-shells at 6.5. The agreement in (relative) strength in the power spectra at L = 6.5 for the tertiary peak is notable and not artificial, because the power is rescaled to the strength of the main peak, rather than the strongest (i.e. 9-day) peak.

The periodograms for electron flux and solar wind speed at L = 6.5 (panel f of figure 2.7) best agree in the periods of the peaks (placement along the abscissa), and relative strength of the tertiary to main peak. The periodogram for electron flux at L = 5.5 (panel d of figure 2.7) displays the greatest peak power at the 9-day period in response to the solar wind speed at a similar period. It is inconsequential that the power at the tertiary peak shown in this periodogram (as well as panels c and e) exceeds the power at the tertiary peak for the solar wind speed. Only ratios of power relative to a fixed reference, in this case the power at the period of the main (~27-day) peak in solar wind speed is important. Therefore, the power (~ 4.3 units) at

the tertiary peak in the electron flux from the solar wind speed is greatest at L = 5.5, somewhat less (~3.5 units) at L = 6.0, etc. The power transfer to the radiation belts due to the solar wind is greatest at L = 5.5, and is reduced to approximately 2.35 at L = 6.5, and reduced to 1.2 at L = 4.5.

During the previous considered solar minimum leading to SC-23 (years 1994-1996), there was agreement at the 9 day peak for L-shells between 4.4 and 7.0, although this agreement is less noticeable since the 9-day peak power throughout this interval did not dominate the other peaks to a comparable degree. Although generalizing from one two-year time interval (1994-1996) and one one-year interval (2005) can only be tentative, the solar wind speed has the greatest power transfer to electron flux at L-shells ~ 5.0 – 7.0 for frequencies near the tertiary peak at solar cycle phases leading to solar minimum.

In both the descent phases leading to SC-23 and SC-24 the relative power in the secondary and tertiary peaks from L = 5 - 7 for solar wind speed and electron flux agrees quite closely. Outside of this range in L, the peak locations for solar wind speed and electron flux along the abscissa agree, although there is sometimes considerable disagreement in the relative power and power transfer. For the most part, there is similar frequency response of the electron flux to the solar wind speed for a broad range of L-shells, i.e. the outer radiation belt exhibits a large degree of spatial coherence to the solar wind forcing [55-57].

Up to this point, spectral estimation has been restricted to looking at the descending phase of the solar cycle leading to solar minimum. It is time to examine the frequency response of the electron flux to the solar wind speed for ascending phases of the solar cycle. It will be evident in the following analysis that the electron flux response during the ascending phases of

the solar cycle does not maintain the tight connection and may exhibit frequencies, not present in the solar wind speed. This is likely to be the result of quasi-periodicities related to CME's

2.4.3 Electron Flux Response Dependence on Solar Cycle Phase

The panels a-f in Figure 2.8 show periodograms comparing the power spectrum for solar wind speed with the power spectrum for 2-6 MeV logarithmic electron flux from SAMPEX at L = 6.5 for six two–year intervals from 1994 – 2006. Panel g in Figure 2.8 plots the 216-day running average of the 27-day sunspot number during this total 12-year period to roughly illustrate the phase of the solar cycle for each of the panels a-f. The six two-year periods demarcate SC-23 into the following separate phases, a) descent SC-22 leading into solar minimum SC-23, b) solar minimum SC-23, c) Ascending SC-23, d) solar maximum SC-23, e) early descent after solar maximum SC-23, and f) descent SC- 23 leading to solar minimum SC-24.

Panel a shows remarkable agreement regarding both the power and peak placement between the solar wind speed periodograms (red dotted), and the log electron flux periodogram (blue solid) during the descending phase leading into SC-23. The discussion of figure 2.7 has already noted the small tail in the electron flux periods longer than the main peak at 27 days becomes smaller as higher L-shells are considered.

The power (y-axis) in the next three panels b-d, for ascending and maximum phases of SC-23 is rescaled down by a factor of 3 in the same units as panel a. If the same scale were used, these periodograms would appear flat when compared to panel a, obscuring distinct frequencies,

which are present. Note that panels e and f rescale upward by a factor of two compared to panels

b-d.
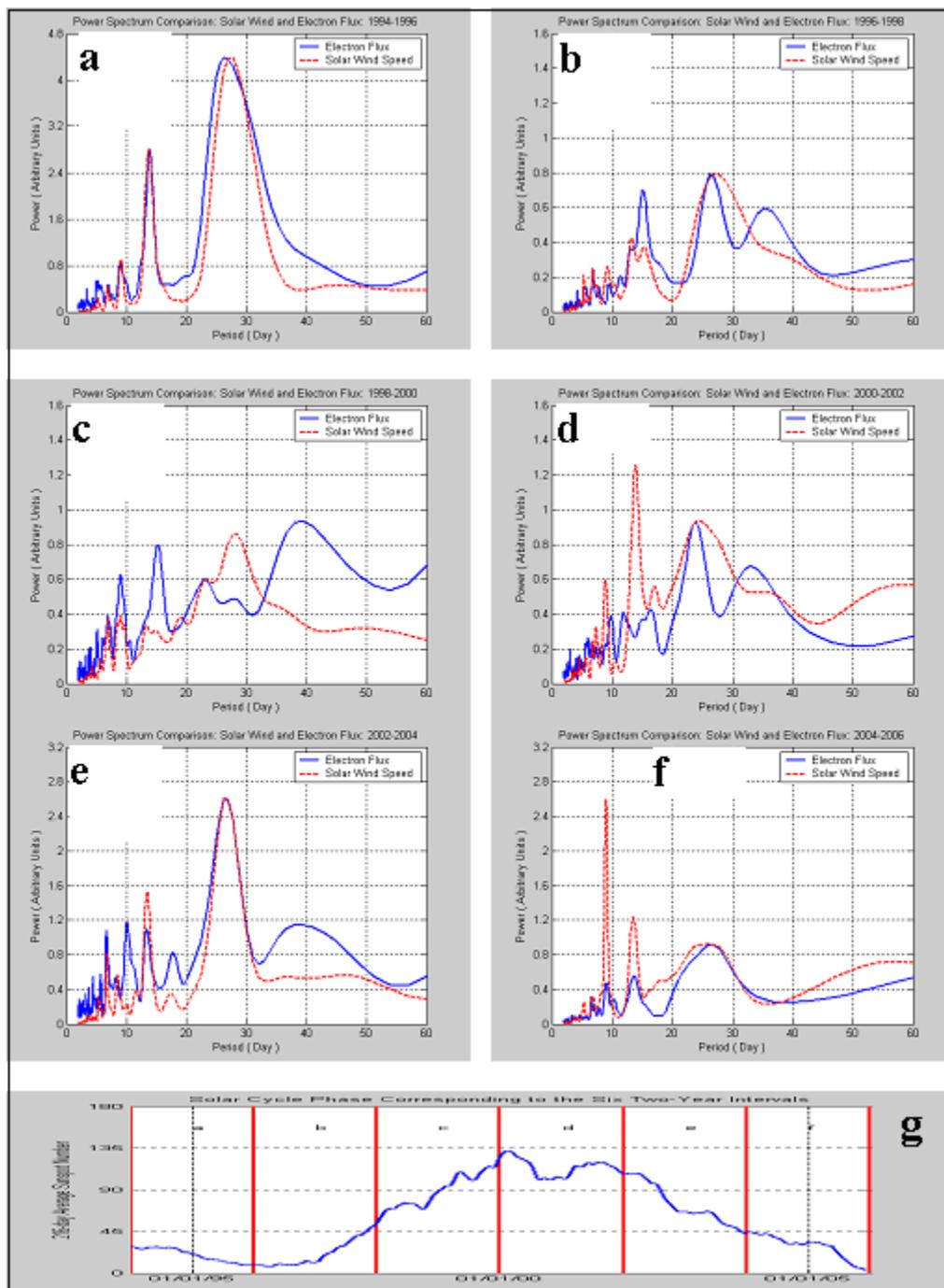
Figure 2.8:   Welch periodograms estimating power spectral density at L = 6.5 in the range [0,60] days over the two-year intervals a) 1994-1996; b) 1996-1998; c) 1998-2000; d) 2000-2002; e) 2002-2004; f) 2004-2006, and g) Corresponding Solar Cycle Phases.

During solar minimum SC-23 (panel b) the electron flux periodogram shows a peak with a 36-day period which is not present in the periodogram for the solar wind speed, although it is interesting to note the solar wind speed periodogram has developed an asymmetry (i.e. a "tail"). In the next section of this thesis, quasi-periodicity will suggest this rather weak (recall the power rescaling exaggerates the peak size relative to panel a) peak may be the result of a CME occurrences rather than high-speed streams.

The electron flux periodogram during the ascending phase of the solar cycle from panel c displays peaks at 16 and 39 days that are unmatched in the solar wind speed. Conversely, the solar wind periodogram displays a 27-day periodicity, which is unmatched in the electron flux. I can present no satisfactory explanation for these features, although the rescaling of power by a factor of three relative to panel may be exaggerating the significance of the disagreement in the periodograms. In addition, I can note the caveat mentioned earlier. The 39-day peak in the electron flux periodogram in panel c was rescaled to match in amplitude with the 27-day peak in the solar wind periodogram. Rescaling by matching amplitudes in main peaks when they are located far apart may be misleading.

The solar wind periodogram during solar maximum from panel d illustrates a narrow 13-day periodicity along with a broader 23-day periodicity. The strong 13-day period in the solar wind is unmatched in the electron flux. This disagreement in electron flux with solar wind is difficult to explain. The weak 23-day periodicity in the electron flux may be associated with CME occurrences (see next section).

The peak locations along the abscissa for periodograms for solar wind speed and electron flux in descending phases after solar maximum (panels e and f), agree for periods between twelve and thirty days. The solar wind periodogram in panel f shows a strong narrow 9-day

periodicity, a somewhat broader and weaker 13-day periodicity, and a broad and still weaker 27-day periodicity. The power from the electron flux periodogram agrees at the main (27-day) peak, but is weaker for secondary and tertiary peaks, opposite the trend exhibited by peaks in the solar wind speed periodogram. One hopes for periodograms leading into the minimum SC-24 (panel f) which are as similar in organization as the periodograms leading into minimum SC-23 (panel a). The atypical nature of the SC-24 alluded to in the introduction may be partly responsible for the discrepancies shown in panel f.

2.5 The Electron Flux Power Distribution across L-Shells

Rescaling the logarithm of the electron flux measurements so that the power in the main peak of the power spectrum matches the power in the main peak in the solar wind speed power spectrum facilitates the comparison between force and response. However, because each two-year time interval rescaled differently, the rescaling makes it difficult to compare the electron flux response in one interval to the response in another interval. In this section, the electron flux power is compared without rescaling. All two-year intervals considered use a base-10 logarithm of the SAMPEX electron flux, allowing the direct comparison of the power spectra for electron flux for different intervals.

The ninety L-shells and six two-year data intervals suggest a forbidding set of potential comparisons for electron flux power. The comparisons should be limited to investigating a particular relevant feature. Here, the maximum power for the range of L-shells from 4 to 8 for each of the two-year intervals is the feature. Figure 2.9 shows the maximum power for any peak in the power spectrum with periods between 10 and 60 days, for each of 41 equally spaced L-shells in the range from 4.0 to 8.0. The three panels in the figure shows the maximum power

curves obtained for the six two-year intervals grouped in pairs, selected because of the similarity in the shape of the plotted curves. The periodograms estimating the peaks in the power spectrum use the same type and window size as earlier periodograms. The power in arbitrary units displays a different scale than the earlier periodograms because the estimated power spectral density is not rescaled and the base-10 logarithm of the electron flux measurements is used instead.

The top panel in Figure 2.9 shows the maximum power in any peak in the range from 10 to 60 days for solar minimum (red) and solar maximum (black). The L-shell with the maximum power overall is in the range from 4.7 to 5.1 for either case. Similarly, the middle panel shows the maximum power during the two-year interval leading to solar minimum SC-23 (blue), and the two-year interval following the solar maximum SC-23 (magenta). The L-shell with maximum power is approximately 5.5 to 5.8 in both of these cases. The remaining two-year intervals (cyan and green) are shown in the bottom panel of Figure 2.9. The power is distributed nearly equally, across the L-shells from 5-8, with greater power distributed to increase roughly linearly moving towards lower L-shell from 5 to 4. The maximum power occurs at L = 4.

The top (middle, bottom) panel shows the greatest (least, intermediate) agreement between curves. However, this assessment must be weighed along with the observation that the two-year intervals are separated by four years in the top panel, eight years in the middle panel, and six years in the bottom panel.

Summarizing, the two-year intervals corresponding to the solar minimum and maximum phases and the two-year intervals preceding and following these phases have a maximum power within the [0,60] day interval at approximately L = 5.0 and L = 5.6 respectively. The two remaining two-year intervals suggest a transition interval where the power distribution is mostly
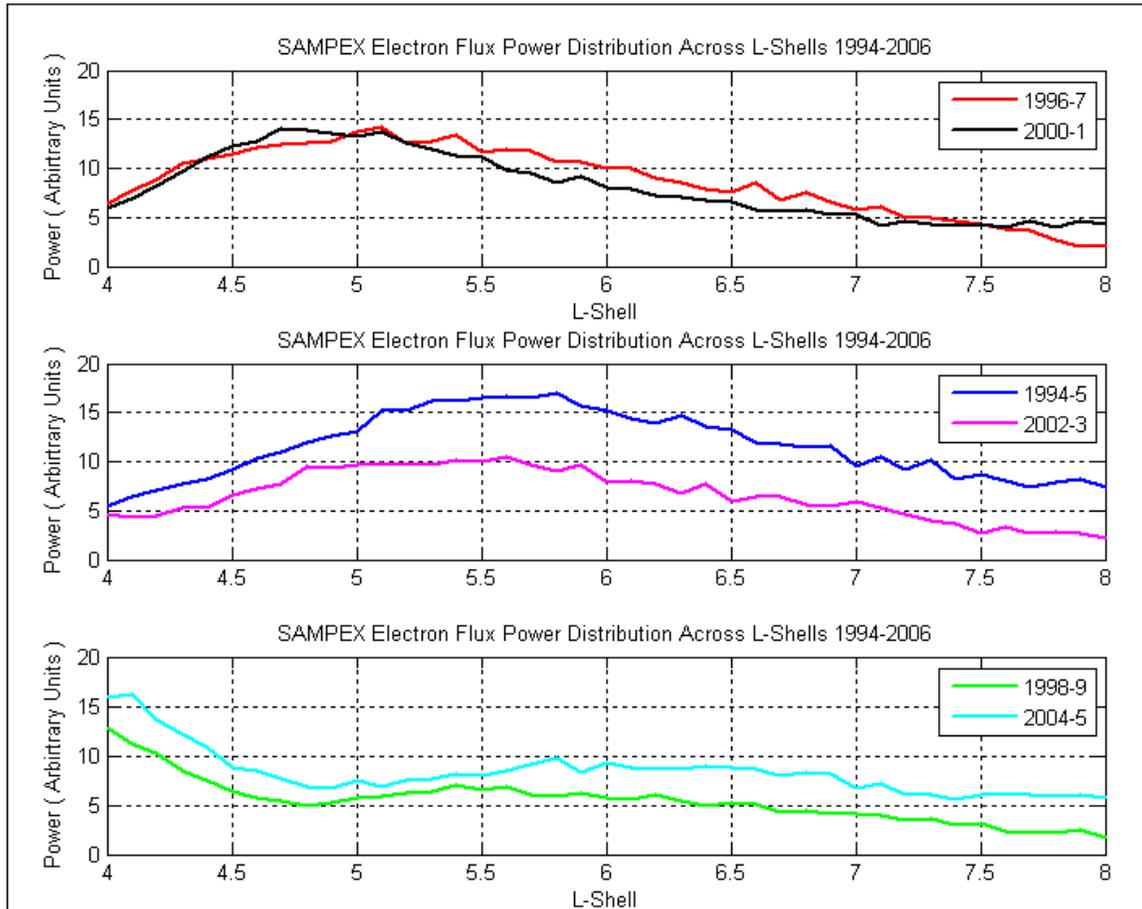
Figure 2.9: Power Distribution Across L-shell Throughout SC-23.

flat across L-shells from five to eight, but peaks strongly at lower L from 4 to 5. I will not

discuss the transition interval further, and will concentrate on locations (from L = 5.0 to L = 5.6)

of maximum power in the four two-year intervals, 1994-1996, 1996-1998, 2000-2002, 2002-

2004. In particular, I offer these L-shell locations are associated with the peak in the impulse

response function from Vassiliadis et al. [6,7].

Vassiliadis et al. give the impulse response, H (τ;L) as a function of the L-shell and a

time lag τ, defined through the relation below:

$$j_e(t;L) = \int_{-T_s}^{T} H(\tau;L) \; V_{sw}(t-\tau) \; d\tau,$$

where, $T_s = 5$ days, $T = 20$ days, V_{sw} are solar wind measurements and j_e are electron flux measurements. The (output) SAMPEX measurements from $1993 - 2001$ and (input) solar wind speed measurements determine the impulse response function, H through singular value decomposition. A plot of the impulse response in two variables is given in Figure 2.10 below from Vassiliadis et al. We may focus on a constant time lag of ~2.5 days and examine the location of the P_1 peak at approximately $L = 5.4$. This is reasonably close to the L-shell containing the most power from the top two panels in Figure 2.8. Although this agreement could be coincidental, it is difficult to find another connection (e.g. plasma pause location, locations of maximum relativistic flux, etc.) for this particular L-value. Moreover, it is not difficult to posit a connection between the L-shell of the maximum electron flux response to a unit disturbance in the solar wind speed (i.e the impulse response), with the L-shell of the maximum power spectral density for the electron flux.

However, I am obliged to offer contrary evidence to this connection. Vassiliadis et al. (2003) present a plot showing the dependence of the radial extent of the P_1 region with solar cycle phase. The P_1 peak moves inward (towards smaller L-shell) as time progresses from 1996-2001, from $L = $ ~6.0 to $L = 4.2$, where our maximum power L-shell varies from $L = $ ~5.7 to $L = 5.1$, and does not vary significantly from solar minimum to solar maximum. Despite this difference in variation, the connection between the P_1 peak and the maximum power location is likely.
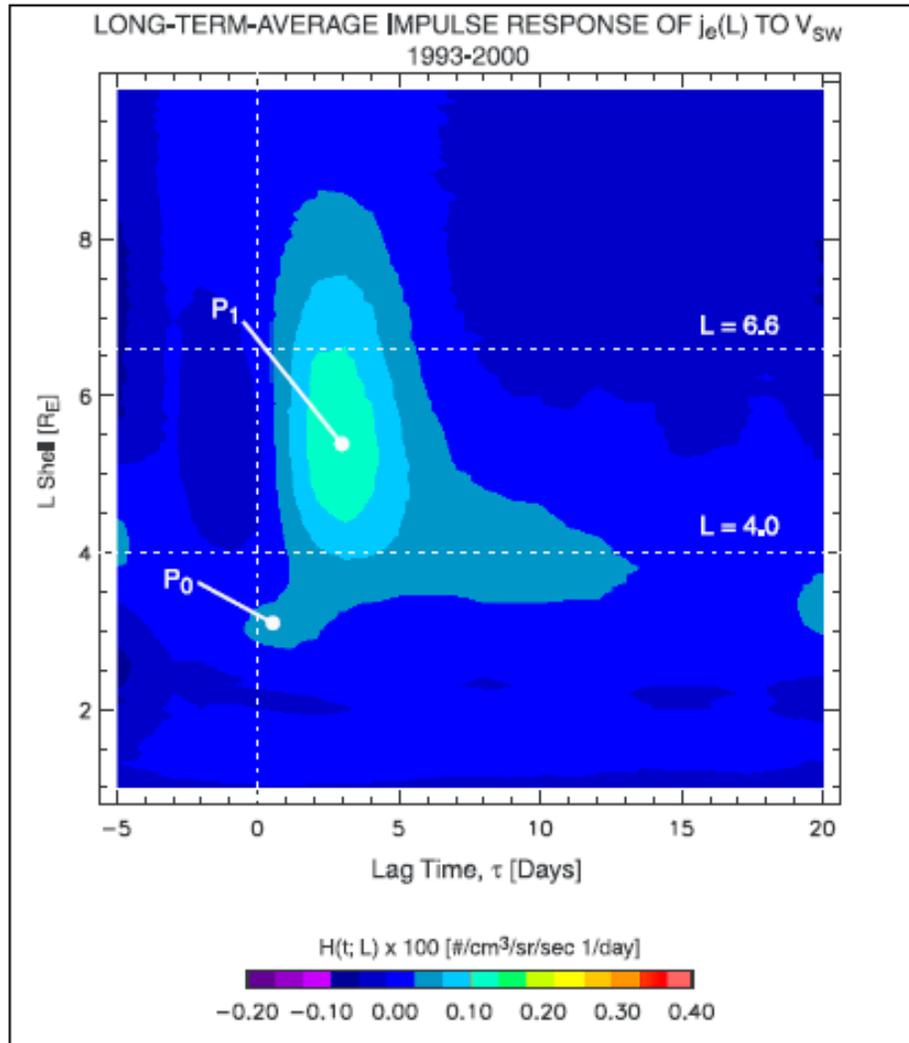
Figure 2.10 Impulse Response Function from Solar Wind to Electron Flux from Vassiliadis et al. [7]

In light of a possible connection between the P_1 peak in figure 2.10 and the maximum power distribution with L-shell given in figure 2.9 begs the question: What causes the peak in the P_1 region? Vassiliadis et al. suggests ULF waves generated by hydrodynamic shear may energize seed electrons from magnetic substorms by resonant absorption. More broadly, Vassiliadis (2003) attributes the P_1 peak in the impulse response to high-speed streams in the solar wind, shocks, and products of interplanetary coronal mass ejections (ICME's) which may replenish seed electrons necessary for energization.

Although many issues are unresolved, it is appropriate to turn our attention away from distinct strong periodicities from high-speed streams, to consider what power spectral estimation can determine regarding weak "quasi-periodicities". This is the subject of the following section.

3. The Spectral Estimation of CME number

The previous section demonstrated strong periodicities in the solar wind velocity drive strong periodicities in the electron flux in the electron radiation belt. High-speed streams from coronal holes during the declining phase of the solar cycle cause the distinct peaks with 27-day, 13-day, and 9-day periods seen in the periodograms for solar wind speed and electron flux. However, as mentioned in the introduction, CME's can also drive solar wind and enhance electron flux in the radiation belt. In addition, an earlier comparison between solar wind speed and electron flux periodograms during 1996-1998 noted an electron flux 36-day peak, which was not present in the solar wind speed periodogram (figure 2.8, panels b, d). Therefore, we may conclude that there may be periodicities in the electron flux, which may not be present in the solar wind. In addition, high-speed streams are not the only driver of fast solar wind and elevated electron flux.

However, the 36-day peak in the electron flux periodogram in panels b and d, are weak periodicities occurring during or shortly after the solar minimum phase of SC-23. The appearance of the peak may be misleading because the power is scaled to highlight differences in the solar wind speed and electron flux power spectrum. Plotted along the same scale as panel a, the periodograms in panel band d would both appear flat, with no strong dominant peaks (i.e. quasi-periodicity). Quasi-periodicity and the possible connection with CME's provide motivation for applying spectral estimation techniques to CME characteristics.

At the outset, I must grant a concession and restatement, which I addressed in the introduction. Not every CME will enhance high-energy electron flux, however occurrences of CME's increase the probability for electron flux enhancement (see Reeves). Therefore, I chose

to calculate periodograms to estimate the power spectrum for CME number in order to compare them to weak periodicities evident in the electron flux in the outer belt.

Investigating periodicities in the CME number is not unprecedented. Lara et al. observed 193, 94, 45, 36, 33, 28, 25, and 23 day periodicities in CME number estimating their power spectrum using the maximum entropy method [8]. Presenting the power spectral estimation without a brief explanation for this method, would be unsatisfying. The procedure for this method follows.

3.1 The Maximum Entropy Method [58]:

First, the frequencies f, are mapped into the z-plane using the equation:

$$z = \exp(2\pi i f \Delta).$$

The Δ shown in the equation is the sampling time. Given this mapping as a function of frequency f, the transfer function as a function of frequency f, is given by:

$$H(z) = \frac{1}{\sum_{j=0}^{P} a_j z^{-j}}.$$

The upper index on the summation, P, is the order of the maximum entropy method model. In order to determine the transfer function H(z) as a function of frequency, the coefficients a_j are required. These coefficients are equal to the autocorrelation function evaluated at k = 0, 1, 2, …P.

$$R_{xx}(j) = a_j.$$

The square of the absolute value of the transfer function $H^2(z)$, as a function of f is the power spectral density. A more comprehensive treatment of the algorithm and concepts supporting this method appears in Numerical Recipes [59].

Figure 3.1 from Lara et al. shows their result for the frequency spectrum for the number of CME's. The numbers above the peaks in the figure correspond to the period in days. The level dashed line provides the three-sigma level of confidence. The blue dash-dotted line shows the periodogram obtained using their wavelet method. The broadness in the peak in the frequency range corresponding to 23-50-day periods offers little support for the frequencies they obtain with the MEM method. The Lomb-Scargle and Welch periodograms affirm their frequencies more convincingly.

Before discussing the data procedure used by Lara et al. in detail, we may examine why weak discrete periodicities are evident in CME number. CME's transport magnetic field flux out from the solar interior to the solar wind and ultimately into the outer heliosphere. Because the interior, photosphere and corona are rotation, (all at different rates), the magnet flux is being wound and twisted. CME's are one mechanism for removing contorted magnetic flux from the Sun [60].

Lara et al. treats small data gaps by averaging the number of CME's over a period determined by the Carrington rotation period divided by 9, or approximately 3 days. Major gaps were padded with zeros. This generated a 1219-point series. Because large numbers of observations generate noisy peaks via the maximum entropy method, they generate the power spectrum with 300 coefficients in their autocorrelation function. This is 25 % of the entire data set. This smoothes the resulting power spectrum somewhat. The technique of selecting the number of coefficients is reminiscent of choosing window size for the Welch periodogram.

Nevertheless, their power spectrum contains still several noisy peaks, a limitation of this method. With regard to the maximum entropy method, Numerical Recipes in FORTRAN contains the warning, "Some experts recommend the use of this algorithm in conjunction with more conservative methods …and to avoid being fooled by spurious spectral features." [59]. Although the Lomb-Scargle periodogram I will present will be at least as noisy as this one, the warning provides justification for reexamining these spectral peaks with two more methods to lend support to their results.

The first method I will use will supply missing measurement data from the CME onset time record with synthetic data, in order to use the Welch periodogram with windowing as in the earlier section. Next, I will use the Lomb-Scargle periodogram, developed specifically for
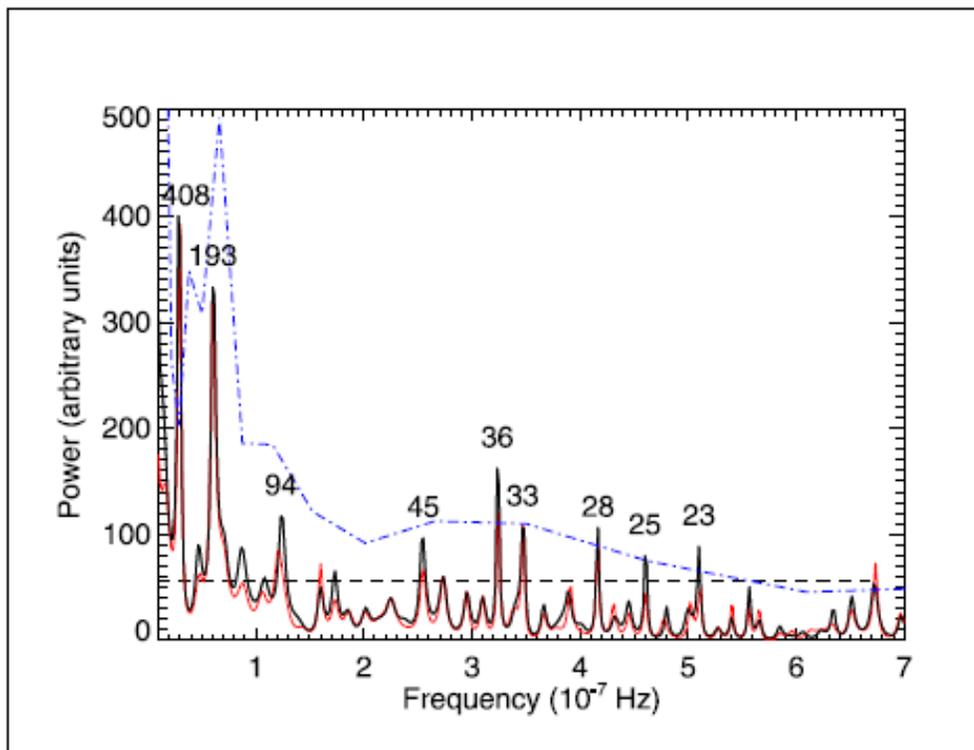
Figure 3.1: Maximum Entropy Method Periodogram for CME Number from Lara et al. [8]

unevenly spaced data, to estimate the power spectrum for CME number. I will then compile a list of periodicities estimated from all three methods Lara's maximum entropy, Lomb-Scargle, and Welch to reach a consensus. My spectral estimation begins by enumerating and sorting CME onset times from the Solar Heliospheric Observatory / The Large Angle and Spectrometic Coronagraph (SOHO/LASCO) CME catalog.

3.2 The SOHO / LASCO CME Catalog

The Virtual Solar Observatory (VSO) tool on website provides CME onset times from the CME catalog [61]. The SOHO spacecraft operates in a halo orbit around the L1 Lagrangian point. LASCO onboard SOHO are a set of three occulting coronographs, which examine the sun and detect CME's. A spectrometer is also included in LASCO for composition studies [62].

In order to generate sufficient statistics, all CME's in the catalog were enumerated in this study with either, C2 or C3 visibility. The SOHO / LASCO CME catalog also supplies a file of downtime intervals in which the LASCO was not operating. It is important to keep downtime in mind when searching for periodicities in the CME number obtained from the onset times. Performing Fourier transforms on data sets with missing data can introduce spurious frequencies in the periodogram.

There were 11406 onset times for C2 / C3 CME occurrences during 1996 to 2007 listed in the SOHO/LASCO Catalog. However, there were 720 data gaps listed with durations varying between 3 hours to 110 days. Sorting these gaps by decreasing duration provided the most serious downtimes highest in the list. Continuous operating times or "gaps" in the downtimes

were established, this list sorted so the longest operating interval would appear first. Table 1 lists

the 20 longest operating intervals and their respective durations.

Operating Interval	Duration (Days)
2000/07/17 19:54 - 2000/03/29 00:14	110.82
2003/01/31 19:25 - 2002/10/23 04:26	100.62
2000/10/25 19:40 - 2000/07/20 21:28	96.92
1999/09/19 12:30 - 1999/06/19 11:58	92.02
2004/03/20 07:40 - 2004/01/06 09:20	73.93
1999/05/17 21:26 - 1999/03/07 01:45	71.82
2003/05/06 00:50 - 2003/02/27 18:16	67.27
2007/05/24 01:24 - 2007/03/19 21:42	65.15
2005/02/28 16:54 - 2004/12/25 19:48	64.88
2001/11/06 20:53 - 2001/09/04 20:36	63.01
2001/06/27 21:54 - 2001/04/26 20:28	62.06
2002/01/14 21:43 - 2001/11/15 08:07	60.57
2001/04/20 19:31 - 2001/02/23 21:18	55.93
2009/01/12 17:54 - 2008/11/20 21:12	52.86
2005/08/08 17:46 - 2005/06/18 18:17	50.98
1997/05/31 12:28 - 1997/04/13 16:36	47.83
1997/10/24 10:18 - 1997/09/08 01:15	46.38
2002/05/31 23:26 - 2002/04/16 01:40	45.91
2006/06/02 07:36 - 2006/04/19 20:53	43.45
2001/01/14 20:30 - 2000/12/02 19:31	43.04

Table 1: Days of Uninterrupted LASCO Service Sorted by Duration

3.3 Approximating Missing Data

Table 1 shows the longest uninterrupted service from March to Mid-July 2000 is 110

days. This interval is separated by roughly three days from the third longest interval of 97 days.

Merging these two intervals would give us 207 days with only one 3- day interruption roughly

midway. However, on average one can expect about 4 CME's per day during this period of solar

maximum and about 1 per day during solar minimum. Estimation using the Welch periodogram is sensitive to inaccuracies caused by data gaps.

Generating synthetic data to fill in gaps, rather than ignoring them is another way to confront the missing data problem. It is best practice to use known statistical characteristics of the data present to generate and fill in the missing data. The first step is to plot CME onset times on a timeline, examine the distribution of inter-arrival times, and determine the statistical characteristics exhibited by CME onset times.

After following this procedure, the exponential distribution provided a good fit to waiting (inter-arrival) times between CME onsets in the catalog, during the longest uninterrupted period, March to Mid-July 2000. The mean waiting time was 4.47 hours, or roughly five CME's per day. Next, the exponential distribution was fit to the waiting time distribution during a period near solar minimum. The mean waiting time determined by the fit was 20.67 hours, or roughly one CME per day (see Figure 3.1).

The e-folding time (the mean waiting time between CME onsets) fully characterizes the exponential distribution. Reexamining figure 1.9 from the introduction leads us to expect differences in the waiting time with solar cycle phase, because clearly the 27-day number of CME's displayed has strong time dependence.

Given that the waiting time between CME's is an exponential distribution, one seeks a discrete probability distribution with the property that the inter-arrival time between events is an exponential distribution. The Poisson distribution (equation shown below) has this property [e.g. 63,64]:

$$P(\lambda) = \frac{\lambda^k}{k!} \exp(-\lambda), \qquad k = 0, 1, 2,... \quad . \tag{3.1}$$

The mean waiting time λ, characterizes the Poisson distribution. The index k can takes all non-negative integers. The routine DRNPOI from the IMSLLIB FORTRAN library provides non-negative random integers to simulate the discrete probability distribution and fill in missing data with specified statistical characteristics. Specifically, every interval of missing data is associated with phases of the solar cycle, and the Poisson random number generator uses the corresponding mean waiting times to generate the synthetic (simulated) data with the requisite characteristics. After filling in missing measurements, the power spectrum can be estimated as before with Welch periodograms.
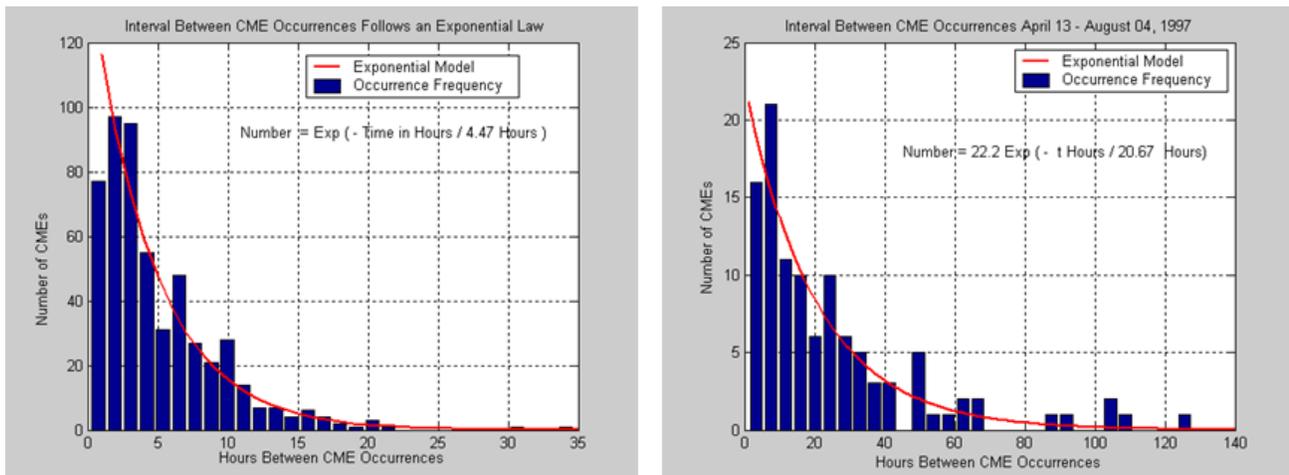
Figure 3.2: Exponential Waiting Time Distribution between CME Occurrences during Solar Maximum (left) and Solar Minimum (right)

Although this process is preferable to ignoring or linearly interpolating missing data, estimating the power spectrum with actual measurements is desirable. Including actual measurements would reinforce periodicities in the data. This procedure is similar to adding white noise to a signal, which will increase the zero-frequency component. This will degrade any periodicities present, but it will not introduce new frequencies (periods) in our periodograms.

3.4 The Lomb-Scargle Periodogram [9,10]

This periodogram was developed for astronomical applications where regularly spaced data was the exception, not the rule. Press and Rybicki provide a good informal description of its workings [65]. Briefly, given observation times t_i and data values h_i, the time-offset τ is defined as:

$$\tan(2\omega\tau) = \frac{\sum_j \sin(2\omega t_j)}{\sum_j \cos(2\omega t_j)}.$$

The Lomb-Scargle estimate of the power spectrum as a function of frequency is defined below:

$$P(\omega) = \frac{1}{2\sigma^2}\left[\frac{\left(\sum_j (h_j - \bar{h}) \cos \omega(t_j - \tau)\right)^2}{\sum_j \cos^2 \omega(t_j - \tau)} + \frac{\left(\sum_j (h_j - \bar{h}) \sin \omega(t_j - \tau)\right)^2}{\sum_j \sin^2 \omega(t_j - \tau)}\right],$$

where, the overbar on h indicates the sample mean, and square sigma is the unbiased sample variance.

The time offset defined above allows the power spectrum estimate $P(\omega)$ to be independent of shifting all times t_i by any particular constant (for all the observation times). This translation time invariance is a comforting feature, however more importantly the appendix supplied in Scargle, demonstrates the power spectrum estimate $P(\omega)$ has the same probability distribution as one obtained from evenly spaced data.

There are several shortcomings of the power spectrum estimate $P(\omega)$ given above. We shall see the Lomb-Scargle periodogram is very noisy when compared to the Welch periodograms. Moreover, the noisiness does not diminish as the number of observations increases. The frequencies of the peak of the spectral estimate also suffer from spectral leakage

and aliasing, so that one is not sure that all frequencies with significant amplitude in the spectrum are actual or aliased from higher frequencies. Scargle addresses many of the criticisms, which several earlier investigators have advanced. However, it is important to keep in mind the limitations of the Lomb-Scargle periodogram, and if possible, reinforce it with other spectral estimations and include windowing.

3.5 Spectral Estimation for CME Number

3.5.1 Lomb-Scargle

This section provides the spectral estimation from the Lomb periodogram calculated using the FORTRAN subroutine LSONE from Numerical Recipes. Because the Lomb-Scargle algorithm may be applied to unevenly spaced data, filling in the missing daily records of CME number throughout SC-23 with synthetic data will degrade the estimate and is unnecessary.

The Lomb-Scargle periodogram determined from sorting the LASCO CME onset times into numbers per day is shown in Figure 3.2, for two different ranges of periods. Figure 3.2 shows a close-up of periods less than 40 days, and 26 b. extends that range to 100 days. The estimated periodogram is very noisy, but we can interpret the peaks by comparing with the earlier results by Lara et al.

Peaks at 94, 45, 36, 33, 28 and 23 days found earlier with the maximum entropy method are evident. The Lomb-Scargle periodogram did not reveal a peak at 25 days as Lara et al. found. The Lomb-Scargle periodogram shows peaks at 39, 54, 60, and 66 days, and possibly others, which were not found using the maximum entropy method. In the range considered, their 94, 36, 33 and 28 day peaks are highest, in the Lomb-Scargle result the 36, 23, 60 and 45-

day peaks are the highest. Averaging by windowing (if it were available), both periodograms

might smooth them, and increase the agreement by down-selecting spurious noisy frequencies.
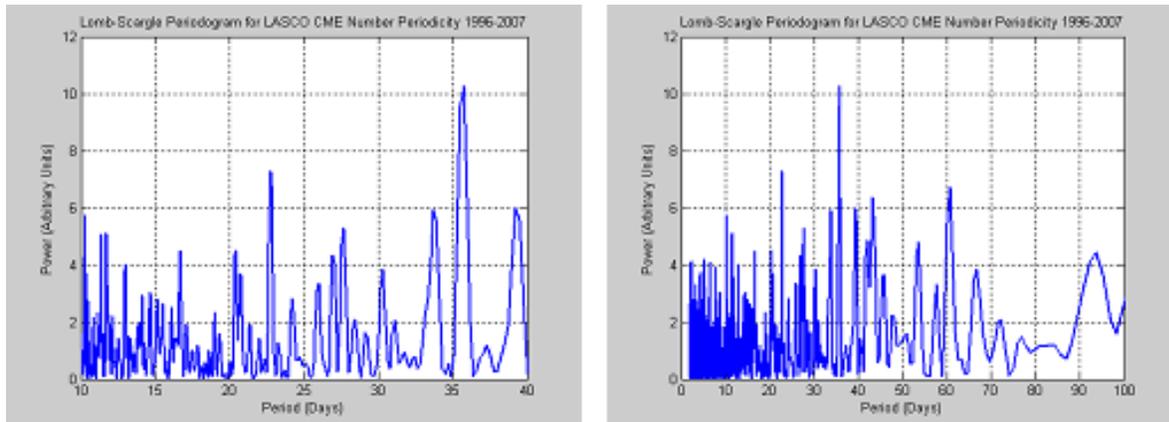
Figure 3.3: Lomb-Scargle Periodogram for CME Number For SC-23 (1996-2007).

Given that we are comparing results using different measurement input with two noisy

estimators of the power spectrum, there is rough agreement in the periods listed between the

Lomb-Scargle and the maximum entropy method.

3.5.2 Welch's method with Synthetic Data.

Finally, I present results from Welch's method to estimate the power spectrum from the

LASCO CME record. The missing data gaps are filled with synthetic data as described earlier.

First, in the top panels of Figure 3.3, I present the results without windowing to retain the

noisiness of the spectral estimation in order to compare on equal terms with the maximum

entropy and Lomb-Scargle methods. In the lower panels of figure 3.3, I use a window M = 500.

With daily CME numbers for 11 years, N is approximately 4000 points. A 50% window overlap

averages over approximately 15 periodograms (Eight periodograms from 4000/500, with 7

overlapping periodograms).

Comparing the lower figures to the upper ones, shows the effect of windowing reinforces

some periods, notably one at 45 days, and weakens others. For example, the 23-day peak has

dropped considerably, the 12-day peak is absent, and the 30, 36, and 94-day peaks are

broadened. The 27 and 30-day peak has merged into one broad peak at approximately 28 days.

In view of the requirement that consistent power spectrum estimation requires windowing, the

lower panels in figure 3.3 are better representations of the power spectrum.
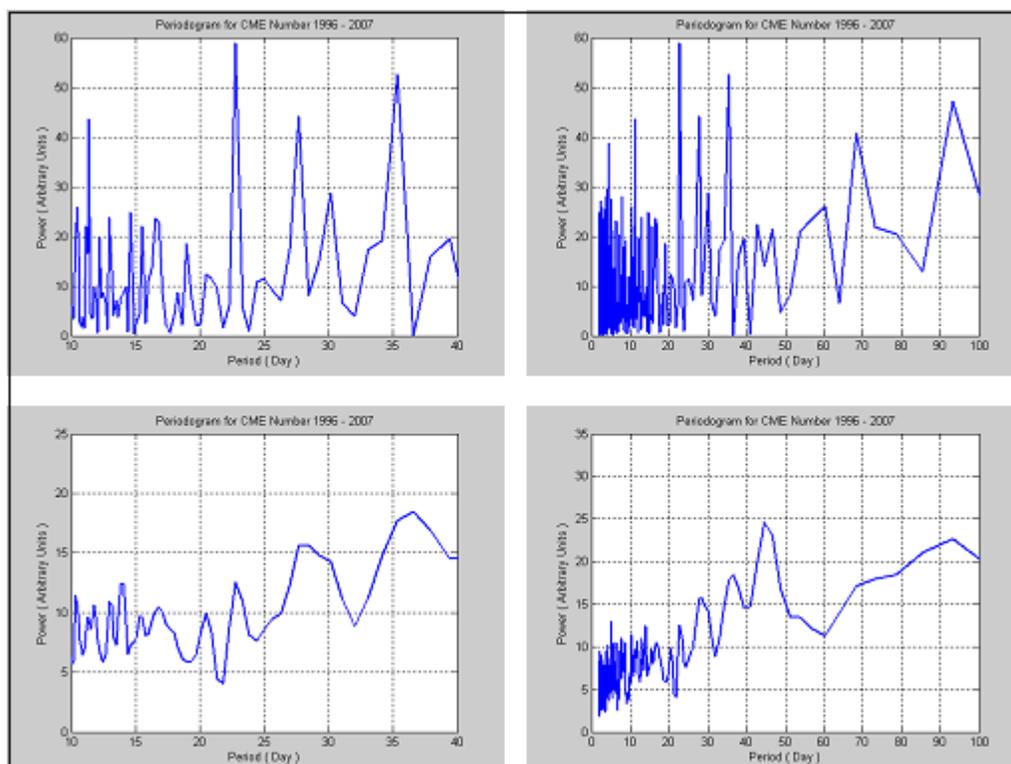
Figure 3.4: Welch Periodogram for CME Number Including Synthetic Data: Top Panels
show noisy periodograms without averaging, bottom panels show windowed periodograms

3.5.3 The Consensus

Table 2 below demonstrates the agreement between the maximum entropy, the Lomb Scargle, and Welch periodograms. The set intersection of the first three columns provides the consensus given in the last column.

The agreement in the location of peaks in the power spectrum between the maximum entropy method and Welch's method with synthetic data is striking. By a two to one vote, the Lomb periodogram seems to include many spurious frequencies. The maximum entropy method includes 25 and 33-day peaks, where Welch's method does not, but given the warning in Numerical Recipes, the 25 and 33-day peaks could be splittings from the 36 and 28-day peaks.

In summary, the consensus of all three methods includes peaks at 94, 45, 36, 28, and 23 days. Quite dissimilar methods for determining peaks in the CME number power spectrum were successful in reaching a common consensus.

Maximum Entropy	Lomb-Scargle	Welch	Consensus
94	94	94	94
	60		
	54		
45	45	45	45
	39		
36	36	36	36
33	33		
	30		
28	28	28	28
25			
23	23	23	23

Table 2: Comparison of CME Number Periodicity with PSD Estimation Method

3.6   The assessment of electron flux quasi-periodicities during ascending solar cycle phases.

    The consensus from the three power spectral estimation methods provided 5 peaks with periods 94, 45, 36, 28, and 23-days. Lomb-Scargle and maximum entropy methods suggest this list should contain a 33-day peak, since the power of that peak from these methods is quite strong. Armed with this list, it is appropriate to revisit the electron flux periodograms in figure 2.8, i.e. panels b-d for ascending and maximum phases of the solar cycle at $L = 6.5$.

    We may first note these panels show the electron flux periodogram scaled so that the maximum power along the y-axis is low relative to the power along the y-axis in the remaining panels for descending and minimum phases. Therefore, distinct peaks evident in panels b-d will be weak quasi-periodicities. If these quasi-periodicities in the electron flux persist in the declining phases of the solar cycle (panels a, e and f), they are hidden by the stronger distinct periodicities from electron flux responses to high-speed streams. Therefore, it is reasonable to confine our attention to panels b-d alone.

    Panel b (near solar minimum to early ascent SC-23) shows an electron flux periodogram with three peaks at 15, 27, and 36-days. Panel c (ascending to solar maximum phase SC-23) shows a electron flux periodogram with three peaks at 16, 23, and 39-days. Finally, panel d (solar maximum phase SC-23) shows an electron flux periodogram at 23 and 33-days.

    None of the three methods estimating the power spectrum address periods shorter than 20-days. The 28-day quasi-periodicity is close to the 27-day synodic rotation period from the Sun, so it could easily be confused with a much stronger periodicity, so this peak will not be considered in the assessment. This leaves the 36-day peak in panel b, the 23 and 39-day peaks in

panel c, and the 23 and 33-day peaks in panel d as candidates to be identified with periods from the consensus. We have already seen the consensus includes quasi-periodicities with 36-day, 23-day and (possibly) 33-day peaks. Apparently there is significant potential for CME quasi-periodicities to explain the weak periodicities at $L = 6.5$ during ascending and maximum phases of SC-23.

However, in closing, it is important to temper the success of this connection, by not only what periodicities are present in the electron flux, but also by what periodicities are not present in the electron flux. The list of quasi-periodicities from the consensus calls for a 45-day peak, but none of the panels in figure 2.8 b-d show evidence for this peak. The 45-day peak is somewhat weak using the maximum entropy method but it is strong for the Lomb-Scargle and Welch's method. Nevertheless, on balance these results suggest the quasi-periodicities in the electron flux are probably associated with CME number or a variable closely related to CME number.

This exposition of periodicities and quasi-periodicities in the solar wind and electron flux has been extensive. Determining existing recurrence of solar wind variables and electron flux has obvious importance in a predictive space weather model. Aside from this importance, the last section of this thesis "changes gears" from establishing the periodicities, to assessing to what extent these periodicities are helpful in identifying coefficients in a space weather model given a record of input and output (measurements). During the review of earlier effort given in the introduction, I mentioned the impulse response function obtained by singular value decomposition by Vassiliadis et al. The next section will treat system identification via the Kalman filter including (input) periodicities in the solar wind, and (output) electron flux measurements.

4. Periodicities and Successful System Identification

This chapter first presents a linear time-series model relating solar wind speed (input) to logarithm electron flux (output), through a series of system coefficients. The goal will be the successful identification (i.e. determination) of the system coefficients given the known input and output. Because noise is present in observations of the input and output, and because noise is present in the modeled process, the determined system coefficients will not be exact. A least squares solution for the system coefficients may be the best solution available.

The least squares solution for the coefficients given a complete record of solar wind speed and log electron flux is straightforward; however, the Kalman filter algorithm for system identification presented will be more ambitious. The system coefficients will evolve on time (one measurement at a time) from an initial, i.e. chosen set of coefficients.

The Kalman filter provides a recursive least squares solution after the final measurement is processed that is equivalent to the batch solution. The advantage in the recursive approach is the filter can be used in real-time. The system coefficients identified use measurements available at the time they are calculated, and do not use information, which may be available in the future. In addition, the identified system coefficients are optimal, i.e. best in the least squares sense, up to the time of the provided measurement.

One application demonstrating the advantage of a real-time predictive model might envision a satellite equipped to measure relativistic electron fluxes, and solar wind speed. Based on this information, the satellite might predict hazardous levels of electron flux, which might warrant a temporary suspension of activity or shut-off. It may be of little value to wait until the hazard is predicted from a lengthy record of fluxes and solar wind speed processed off-line. By that time, it may be too late. Recursive estimation and prediction provides an "up to the

moment" assessment of the hazard. Moreover, when the entire record of input and output is available, the solutions computed recursively are identical with the solutions computed off-line.

This section introduces a linear model for predicting electron flux from solar wind speed. The system coefficients in this model need to be determined. The system coefficients comprise the state estimated by the Kalman filter. The errors in estimating the coefficients from input and output measurements are minimized as described later in greater detail. That is to say the coefficients are optimal.

The least-squares approach to data analysis is sufficiently important so that if the Kalman filter were only a recursive version equivalent to the off-line (batch) version of this approach, it would not be lightly dismissed. I shall present evidence in this section demonstrating periodicities in the solar wind speed have a dramatic effect on the effectiveness of system identification with the Kalman filter. This conclusion is also true processing the measurements off-line, but it is not so apparent.

The first step is to introduce the ARX(4,4) time series model for the electron flux. I will note that this model has four autoregressive coefficients pertaining to previous sampled values of logarithmic relativistic electron flux and four exogenous coefficients pertaining to sampled values of solar wind speed forcing. If needed, elaborate models with a larger set of coefficients may be constructed at the cost of greater computing power. However, even this model is elaborate when compared to earlier ones such as an ARX(2,1) model, proposed by Klimas et. al for predicting storm-time disturbance index, Dst from the product of solar wind speed and the Southern component of the interplanetary magnetic field [66].

Klimas et al. shows this three-coefficient model is related to a stochastically driven linear damped harmonic oscillator. Similarly, this ARX(4,4) model can be related to a stochastically

driven coupled harmonic oscillator. The electrical analog to this system would be an inductively coupled transformer. Suffice it to say the physical underpinnings of these models are complex despite their simple forms. However, all that is important for our purpose is the ARX(4,4) model is sufficiently complex to demonstrate the periodicities in the input affect the effectiveness of coefficient identification in the model via the Kalman filter.

4.1 The Fourth-Order Autoregressive Exogenous Model

Equation 4.1 provides the defining system equation for the ARX(4,4) model:

$$
\begin{aligned}
z_k + a_1\, z_{k-1} + a_2\, z_{k-2} + a_3\, z_{k-3} + a_4\, z_{k-4} \\
= d_1\, u_{k-1} + d_2\, u_{k-2} + d_3\, u_{k-3} + d_4\, u_{k-4} + w_k
\end{aligned}
\tag{4.1}
$$

The solar wind speed measurements enter through the control variables u. The autoregressive coefficients are a1, a2, a3, a4, the exogenous coefficients are d1, d2, d3, d4, the (integer) k index subscript defines the time k under consideration, and the z variables are logarithm electron flux measurements. The w variable is white noise added to the model to reflect the (imperfect) measurement process. Noise may also enter the process itself through fluctuations in the system coefficients (autoregressive and/or exogenous) but this will not be considered for now. That is, let us tentatively assume the system coefficients are constant, although filter estimates of these coefficients will vary widely at early time (low k) when few measurements have been processed.

The goal is to best estimate the eight system coefficients, given measurements of the solar wind speed and log electron flux at time $k > 3$. Clearly linear regression techniques can be

used, but the Kalman filter can estimate the system coefficients recursively highlighting the advantages already discussed.

4.2 The Kalman Filter

Much of the mystery behind the Kalman filter can be illustrated using a simple example. Suppose we are given a series of 1000 numbers and we assume the numbers are best represented by their mean. That is to say, the knowing the mean is important to us. We find the mean of the 1000 numbers is (for example7.5). Then, we are given an additional number; say 7, and we want to compute the (new) mean of 1001 numbers.

The less statistically sophisticated among us would recompute the mean by (re)adding the 1000 numbers, adding the last number and dividing by 1001. A more reasonable approach would be to multiply the known mean by 1000, add the last number and divide by 1001. We note that the new mean is related to the old mean and the last number linearly by:

$$\bar{x}_N \;=\; \frac{N-1}{N}\; \bar{x}_{N-1} + \frac{1}{N}\; x_N, \qquad with \quad N = 1001. \qquad\qquad 4.2$$

The second approach is less computationally burdensome, and requires less storage of previous values. Analogously, but with greater complication, the Kalman filter estimates a state and state error covariance linearly from the earlier estimates of the state and state error covariance along with the newest measurement, recursively and provides the optimal estimate and optimal

covariance on that estimate. In order to discuss the role of the Kalman filter to optimal state estimation, background on noise and the signal model is presented first.

4.3   The Signal Model [67]:

The signal model generates a stochastic process in order to approximate the naturally occurring process being considered, in our case the ARX(4,4) time series for the solar wind speed and electron flux. The signal model consists of a measurement equation, which relates the state of the system to the measurement, and a state evolution equation, which describes the system dynamics.

The discrete time measurement equation is given below:

$$z_k = H_k\, x_k + \nu_k \qquad\qquad 4.3$$

The integer k subscript indicates the time considered, specifically the day. The current measurement is z_k. The symbol, ν_k is white noise at time k added from uncertainties in the measurement process. The state of the system at time k is x_k, usually an N x 1 vector, where N is the number of state components, (here N = 8). The observation z may have multiple output with M variables or single output M = 1, as in our case. The measurement matrix H will be an NxM matrix with M linear equations in N variables. The purpose of the Kalman filter is to optimally estimate the state at each time k, given observations z made up to and including time k.

Kalman filtering assumes white Normal distributed measurement noise with zero mean. Specifying the measurement noise variance as R_k, completely determines the additive noise

process "corrupting" the ideal measurement. The autocorrelation function for measurement

noise process is given below:

$$E[v(k)v(k')] = R_k \, \delta(k,k').$$
4.4

The E in equation (4.4) denotes expectation value. The Kronecker delta accords with the loose

interpretation that noise values observed are uncorrelated with prior or later noise values. The

expectation value for continuous functions is defined by the integral of the argument in brackets

weighted by the distribution function as follows:

$$E[x] \;=\; \int_{-\infty}^{\infty} x \, f(x) \, dx$$
4.5

Although the time is discrete, the values taken from the distribution are real so that the

expectation defined in equation above is appropriate. Having presented noise characteristics in

the measurement equation, we now regard the state evolution equation in the signal model.

The state evolution may also be a noisy process. In the literature, plant noise refers to

this term. The term originates from early uses in control theory to automate factories or "plants".

The term, process noise (distinct from the measurement process) is now in common use. The

state evolution equation in discrete time is provided below:

$$x_{k+1} = F_k \, x_k + G_k \, \omega_k.$$
4.6

The state at the successive time steps is related to the state at the prior step through the NxN matrix F_k. The symbol ω represents a vector of white noise components with matrix covariance Q_k. The covariance for the process noise is given below:

$$E[\omega(k)\omega(k')] = Q_k\,\delta(k,k').$$

4.7

The matrix G_k relates the process noise to the system state at time $k + 1$. Successively applying the state evolution equation propagates the state from the initial state $k = 0$, to any desired time k.

The state propagation equation and the measurement equation constitute a pair of equations, necessary to describe the system along with the measurement process. This pair comprises the signal model. Hereafter, the state vector x will have N components and the measurement z will have M components. The matrices F, G, and H, in the signal model will be NxN, Nx1, and NxM, respectively to conform to rules of matrix multiplication.

4.4 Objective of the Discrete Time Kalman Filter

There are several ways to address the discrete time Kalman filter. One approach is to treat the subject from a statistical point of view equivalent to recursive least squares estimation. Another approach is a systems analytic point of view as an optimal linear observer. All reasonable approaches use the Kalman filter to provide an optimal estimate of the state of the system in the presence of noise given information from a time series of measurements related to the state. The approach presented here is closer to systems analysis.

The term optimal estimate used in the criterion for the Kalman filter has the following meaning. The (column) state vector, x, at any time k can be estimated with (vector) values e from a probability distribution, determined as some linear function of the measurements z1, z2, … zk, up to time k. At each k, one such estimate e, will minimize the quadratic form below:

$$E ([x_k - e]^\dagger \ S \ [x_k - e]).$$

The dagger indicates matrix transpose so that the first factor in the quadratic form is a (single) row vector with N entries. The matrix S in the expression is a weighting function. This weighting is an arbitrary positive definite symmetric matrix, which may selectively weigh some components of the state vector in preference to others. As long as S is positive definite (i. e. has positive eigenvalues) and symmetric this expression is minimized for any S, by the same estimate e. The next paragraph, while not a mathematical proof, suggests why this should be the case.

Because S is symmetric, it has real eigenvalues. S is positive definite so it has positive eigenvalues. An orthogonal transformation can transform S into diagonal form. Orthogonal transformations are a subset of linear transformations, a fortiori, the expectation which is invariant to linear transformations is also invariant to orthogonal transformations. This allows us to replace the positive definite matrix S with the diagonal matrix D in the quadratic form. State vector x and estimate e will likewise have to be transformed (say, to y and f) by the orthogonal transformation in order to remain consistent. Next, a scale transformation can be applied to transform diagonal matrix D into the unit matrix. This scale transformation must also transform the earlier transformed state y and estimate f (say, to z and g resp.). The combination of an orthogonal transformation and a scale transformation rids us of matrix S in the quadratic form.

Handwaving further, the eigenvalues in S, (i.e. the diagonal elements in D, which must be positive) ensure that any linear transformation will not turn a global or local minimum to a global or local maximum). This paragraph suggests why the same estimate can minimize the quadratic form for arbitrary positive definite symmetric S. A Rigorous proofs are supplied in textbooks (e.g. Brown and Hwang, or Grewal), but are omitted here for brevity. This paragraph should be helpful to understand a proof uncovered during follow-up by the ambitious or interested reader.

Recapitulating, the signal model has been introduced. The purpose of this model is to approximate a naturally occurring stochastic process under study. This model consists of a measurement equation and a state propagation equation. The objective of the Kalman filter has been stated. Specifically, the (discrete) Kalman filter provides the optimal estimate of the state of the system at time k, given specific measurements up to and including time k, which relate to the state through the measurement equation. The estimate is optimal in the sense that the weighted square of the difference in the state from its estimate is minimized for each time k. This weighting must be symmetric and positive definite but is otherwise unspecified. (This freedom is remarkable)

4.5 The Nature of the Solution to the Signal Model.

It is reasonable to reexamine the signal model and consider the nature of the solutions. First, we need two important characteristics of Normal distributions, which will allow a generalization to Normal stochastic processes.

1. If v is Normal distributed and T is a fixed matrix, T * v is again Normal distributed.

2. If v1 and v2 are two Normal distributed random variables, then their sum is a Normal distributed.

Together these conditions reveal the Normal distribution is linear.

These points allow us to consider the qualitative solutions, which follow. First, consider the measurement equation without noise. This is a matrix equation for output z. Point 1 establishes this product is Normal distributed. Next, measurement noise with variance R is added to this scalar. Using point 2, this sum, the solution to the measurement equation, is a Normal distribution.

Similar consideration is given to the state propagation equation. Before the process noise is added, the first term in the equation, $F * x_0$ will be an N dimensional vector. The process noise term is a Normal (vector) stochastic process added to this term. The state x_1 will have a Normal probability distribution surrounding vector $F * x_0$.

Applying the measurement equation to this new state x_{k+1} involves matrix multiplying this Normal distribution x_1 by H. Because matrix multiplication is linear, this will again lead to a Normal distribution. It is also evident that applying the state evolution equation, involving a matrix multiplication and adding two Normal distributions to the following time step will again result in a Normal distribution for the state. The upshot of this heuristic argument is that qualitative solutions to the state propagation and measurement equations are Normal distributed random variables at all times. Considering the process along the entire time 'axis', rather than individual times, show us the solutions to the state propagation and measurement equations are Normal stochastic processes.

Although, the process is always Normal; however, it is not prudent to assume that because the process and measurement noise are white that the state and measurement vectors are white, for all time. Matrices F, G, and H may introduce time-correlations in the states (alone), the measurements (alone), and between states and measurements.

It is now evident that solutions to the signal model presented will be Normal stochastic processes. With the parallel development of signal model and Kalman filter presented in this section the natural question arises, will optimal state estimates from the Kalman filter also be Normal distributed? Baye's rule allows us to show the answer is yes.

Optimal estimates given the specific history of realized (measured) observations are conditional probabilities. We may consider the probability distribution for the state x at any time k+1 given the condition that measurements z_0 z_1 z_2 ... z_{k+1} were observed from the measurement equation as a conditional probability. Baye's rule for conditional probability is given in most textbooks on probability theory (e.g. Papoulis [63]):

$$P(x_{k+1} | z_{k+1}) = \frac{P(z_{k+1} | x_{k+1})\ P(x_{k+1})}{P(z_{k+1})} \qquad 4.8$$

The left hand side of equation (4.8) is a probability for the estimate of the state at time k+1 given the latest measurement at time k+1. This represents the optimal state estimate from the filter. Applying Baye's is sufficient to derive the Kalman filter equations. Evaluating the right hand side of equation (4.8) will demonstrate the conditional density for the state given measurements up to current time is Normal distributed. The aggregation of time series (realizations) from the filter is a Normal stochastic process.

Summarizing, the "optimality" criterion for state estimates from the filter has been presented. Solutions to the signal model and Kalman filter are Normal stochastic processes. The filter assumes the signal model presented along with white measurement and process noise. We expect the optimal state estimate to belong to a Normal distribution and it will minimize the weighted state variance for a general class of symmetric positive definite weight functions. Because the mean and variance characterize the Normal distribution, these two parameters are sufficient to determine the distribution for the optimal estimate via the Kalman filter.

4.6 The Discrete Time Kalman Filter Equations

Just as process noise and measurement noise were assumed to be white noise processes with variances Q and R, we may assume a white state with variance can be set to P so that:

$$E[x(k)x(k')] = P_k\, \delta(k,k'). \qquad\qquad 4.9$$

We may recall from the definition (4.5) that the expectation operator E is linear. The estimator of state x_k may replace the mean of state x_k, leading to the first equation for the mean of the state in equation (4.10). The second equation for the measurement mean follows because the expectation operator is linear, and the measurement noise has zero mean. Therefore we have equation 4.10 below:

$$E\,[x_k] = \hat{x}_k, \quad and$$
$$E\,[z_k] = E\,[H_k\, x_k + \upsilon_k] = H_k\, \hat{x}_k + E\,[\upsilon_k] = H_k\, \hat{x}_k. \qquad 4.10$$

The covariances and cross-covariances of these the state and measurement may be computed similarly from definitions and the linear property of the expectation.

$$E[x_k\, x_k{}^\dagger] = P_k,$$
$$E[x_k\, z_k{}^\dagger] = H_k{}^\dagger P_k,$$
$$E[z_k\, x_k{}^\dagger] = H_k\, P_k, \quad and$$
$$E[z_k\, z_k{}^\dagger] = H_k\, P_k\, H_k{}^\dagger + R_k.$$

4.11

The conditional probability for the measurement z_k, given the state was at time k was x_k is equivalent to the probability that the measurement noise was v_k. The mean of the measurement under this condition is equal to $H_k\, x_k$. The distribution function for this event is given (with dummy index k+1 replacing k for later use in equation (4.8)):

$$f(z_{k+1} \mid x_{k+1}) = N_z(H_{k+1} x_{k+1},\ R_{k+1}).$$

4.12

The $N_z(a,b)$ denotes a Normal distribution in variable z with mean a and variance b.

The remaining distribution functions on the right hand side of Baye's Rule obtained by using equations (4.11) are:

$$f(x_{k+1}) = N_x(\hat{x}^-{}_{k+1},\ P^-{}_{k+1}) \quad and,$$
$$f(z_{k+1}) = N_z(H_{k+1}\hat{x}^-{}_{k+1},\ H_{k+1}\, P^-{}_{k+1}\, H_{k+1} + R_{k+1}).$$

4.13

Superscript minus signs on the state estimates x and covariance P denote a time before information from the measurement is used to improve the estimate. Baye's Rule allows us to solve the equations for the optimal estimate of the state as the conditional probability on the left hand side of equation (4.8). After some algebra including completing the square, the conditional probability is given by the Normal distribution below:

$$f(x_{k+1} \mid z_{k+1}) = N_x(\hat{x}_{k+1}, \ P_{k+1}), \qquad with \qquad \qquad 4.14$$

$$\hat{x}_{k+1} = \hat{x}^-_{k+1} + P^-_{k+1} H^\dagger R_{k+1}^{-1} \ (z_{k+1} - H \hat{x}^-_{k+1}), \quad and$$

$$P_{k+1} = P^-_{k+1} - P^-_{k+1} H^\dagger \ (H P^-_{k+1} H^\dagger + R_{k+1})^{-1} H P^-_{k+1}$$

$$4.15$$

Equations (4.14) and (4.15) provide the distribution for the optimal state estimate at each time-step k. This distribution function uses information from measurements from an initial time $k = 0$ to the time k+1. It is now apparent that as expected, the distribution for the optimal state estimate remains Normal for all k. The mean of the distribution is centered on the state estimate and is given in terms of the estimate before the measurement is used, (the estimate with the - superscript), by the first equation in (4.15). The second equation provides the variance for the optimal state measurement as a result of the information obtained from the observed measurement z_{k+1}.

Kalman filter equations describe how the mean and variance for the Normal distribution for the state estimate changes with every time-step in terms of the mean and variance at the preceding time-step. It is clear that this process can be used recursively.

It is interesting to note that the second of the Kalman filter equations describing the state covariance evolution does not involve the measurements z. Specific measurements are only necessary to calculate state estimates in time. This freedom suggests the remarkable possibility to determine the error in the state estimates without regard to the measurements themselves. This is the case for many Kalman filter applications. One may run the filter

"offline", if data is unavailable, in order to see what covariance history (i.e. performance) can be expected. Later, when measurements are available, re-running the filter with the measurements can determine the history of optimal state estimates.

However, for our specific application to system identification this will not be possible. In our application, the measurements will build up the measurement matrix H and system coefficients will comprise the state we are interested in estimating. The covariance will no longer evolve independent of the observed measurements for our implementation.

4.7 Natural Forms of the Kalman Filter Equations.

To some degree, there is a degree of arbitrariness in using the term natural. The Kalman filter equations describing how state and covariance evolve in time presented in equations (4.14-5) in terms of measurements and system coefficients are not easy to interpret conceptually. One equivalent interpretation regards the Kalman filter as a cyclic process (see Fig 4.1).

First, the filter requires an assumed initial state and state error covariance. The enter prior estimate and error covariance x_0^- and P_0^- indicate this input at the top of the figure. The minus superscript denotes the value before the measurement z is processed. For simplicity, we will omit references to the subscripts and superscripts and the figure will remind us of their presence. The Kalman gain is computed from the state error covariance P, the measurement noise variance R, and measurement matrix H, using the equation below for $k = 0$.

$$K_k = P_k^- H_k^\dagger \ (H_k P_k^- H_k^\dagger + R_k)^{-1} \ , \qquad\qquad 4.16$$

The dagger superscript denotes matrix transpose.

Kalman Filter Equations

Enter prior estimate $\hat{x}_0^-$
and error covariance P_0^-

Measurement

Best
Estimated **State**
and Covariance

Z_k

Compute Kalman Gain
$$K_k^- = P_k^- H_k^T (H_k P_k^- H_k^t + R_k)^{-1}$$

X k+1, P k+1

Update Estimate with
Measurement z_k
$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H_k \hat{x}_k^-)$$

Project Ahead:
$$\hat{x}_{k+1} = \phi_k \hat{x}_k$$
$$P_{k+1}^- = \phi_k P_k \phi_k + Q_k$$

Compute Error Covariance
for Updated Estimate
$$P_k = (I - K_k H_k) P_k^-$$
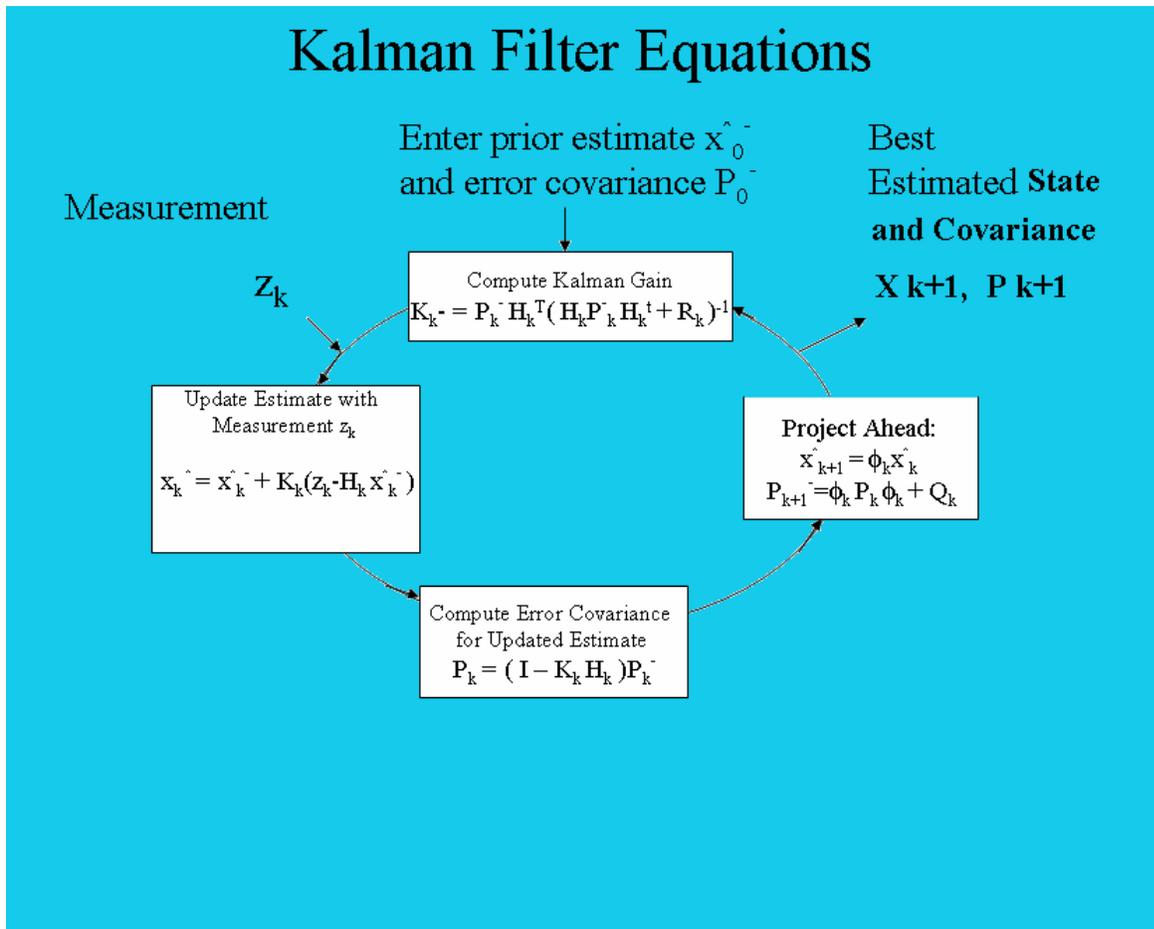
Figure 4.1: The Discrete-Time Kalman Filter Loop

Calculating the Kalman gain requires a matrix inverse. The dimension of the matrix inverse is equal to the dimension of sensor noise covariance R, which in turn is equal to the dimension of the measurement z. We are concerned with single output systems so that R is 1x1 matrix (i.e. a number), rather than a larger matrix. In this case, the reciprocal of the expression in parentheses instead of the matrix inverse can be used to calculate the Kalman gain K. This is a major advantage to the case where many measurements are needed at once such as (batch) off-line processing. Finding the inverse of large dimensioned matrices introduces the possibility the

matrices will become ill-conditioned and greatly increases memory storage and computational demands. Recursion illustrates major advantages for on-line "adaptive" processing over batch processing.

Proceeding counterclockwise in the figure, the measurement z is introduced to update the estimate of the state. Reordering terms in the equation for the state estimate, we see that:

$$\hat{x}_k = (I - K_k H_k)\, \hat{x}_k^- + K_k\, z_k. \qquad\qquad 4.17$$

This reordering clarifies that the updated state estimate is the result of combining the measurement with the previous state estimate. The Kalman gain determines to what extent information from the measurement changes the state from the earlier estimate before the measurement to the estimate after the measurement is available. If the Kalman gain is large, this is often the case early in the processing (i.e. small k) , there is more confidence in information provided by the measurement than there is in estimates from the existing model based on the few earlier measurements. When the Kalman gain is small, there is more confidence in the estimate from the model containing information from many earlier measurements, rather than information from the one measurement that still needs to be incorporated into the updated state estimate. Intuitively, this idea is sensible.

After computing the Kalman gain, the updated state error covariance is computed (bottom of figure 4.1). The updated state estimate is not needed in order to update the error covariance, see the paragraph before the beginning of this section for a discussion. The updated error covariance brings us around the box at the bottom of this figure. Continuing counterclockwise, we need to project the updated state estimate and the updated state error covariance from k = 0 to k =1, by using the transition matrix Φ. In our implementation, the state

transition matrix, Φ is taken to be the identity matrix. As indicated in the box on the right hand side, process noise Q increases the variance on the updated state error covariance P.

The optimally estimated state and error covariance are available after the box on the right. Specifically, the updated state estimate x(k=1), and state error covariance estimate P(k=1), will be output. These states will also reinitialize the cycle to commence with k = 1, at the top of the figure. The full cycle continues until all measurements are incorporated, and the filter computes the entire history of optimal state and (state) error estimates up to the present time.

These outputs will be optimal is the weighted least square discussed earlier. The state and the state error covariance were "optimized" in the selection of the Kalman gain. The Kalman filter equations after considering Baye's rule provided the method for calculating this gain.

The cyclic process presented is in the Kalman filter loop in figure 4.1 is conceptually superior to the Kalman filter equations (4.14-5). However, the loss of precision from computer round off errors and matrix stability issues militate towards reformulating this series of matrix inverses, multiplications and additions from the state error propagation equation at the bottom and right of figure 4.1 to the equivalent Joseph form. This symmetric form of the Kalman filter equation for the optimal state error covariance is given by:

$$P_{k+1} = (I - K_{k+1}H_{k+1})P^-_{k+1}(I - K_{k+1}H_{k+1})^\dagger + K_{k+1}R_{k+1}K_{k+1}^\dagger \qquad 4.18$$

Summarizing, this section has established the optimality criterion the Kalman filter uses to estimate the state based on noisy measurements related to the state by the measurement equation. This section has derived the discrete Kalman filter equations using Baye's rule and characteristics of noise and expectation. The Kalman filter cycle including the Joseph form

details the way the optimal state estimate and state covariance is computed as a series of matrix additions, multiplications, and inverse (when necessary).

In addition, this section has introduced the ARX(4,4) time-series model as the signal model we will use to calculate the optimal state estimates from the Kalman filter. The remaining step in our implementation is the ARX(4,4) time-series model must be recast in terms of the measurement equation and state evolution equation in the signal model, equations (4.3), and (4.6). This is the subject of the next subsection.

4.8 The state space form of the ARX(4,4) equation.

Matrix multiplication can verify the state space representation shown in equations 4.3 and 4.6 are equivalent with the general ARX (N,N) equation (not shown), when the following identifications are made for matrices F, G, and H in equations 4.3 and 4.6:

$$
x_{k+1} = F_k\, x_k + G_k\, u_k; \quad F_k =
\left(
\begin{array}{ccc|ccc|cc}
-a_1 & -a_2 & -a_3 & . & . & . & -a_{N-1} & -a_N \\
1 & 0 & 0 & . & . & . & 0 & 0 \\
0 & 1 & 0 & 0 & . & . & 0 & 0 \\
\hline
. & . & . & . & 0 & 0 & 0 & 0 \\
. & . & . & 1 & . & 0 & 0 & 0 \\
. & . & . & 0 & 1 & . & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{array}
\right)
; \quad G_k =
\left(
\begin{array}{c}
1 \\
0 \\
0 \\
\hline
0 \\
0 \\
0 \\
\hline
0 \\
0
\end{array}
\right)
;
$$

4.19

$$
z_k = H_k\, x_k + \upsilon_k; \quad H_k = \left(d_1 - a_1 \quad d_2 - a_2 \quad d_3 - a_3 \quad . \quad . \quad . \quad d_{N-1} - a_{N-1} \quad d_N - a_N \right)
$$

With N = 4, (four autoregressive and 4 exogenous coefficients) the specific signal model becomes:

$$x_{k+1} = F_k x_k + G_k u_k; \quad F_k = \begin{pmatrix} -a_1 & -a_2 & -a_3 & -a_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad G_k = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \quad \text{4.20}$$

$$z_k = H_k x_k + \upsilon_k u_k; \quad H_k = (d_1 - a_1 \quad d_2 - a_2 \quad d_3 - a_3 \quad d_4 - a_4)$$

Matrix multiplication can verify this form with these identifications of F, G, and H is equivalent with equation 4.1.

This signal model (state evolution and measurement equation) will allow us to generate a ARX(4,4) time-series from an (arbitrary) initial state x0, with preset autoregressive and exogenous coefficients (a_1, a_2, a_3, a_4, d_1, d_2, d_3, d_4), and a prescribed set of control variables u_k (such as solar wind speed) with preset periodicity (e.g. 27-days). This generated time-series with known characteristics will later be input as measurements into the Kalman filter for state estimation in order to assess the performance of the filter (i,e. validation).

4.9 The Kalman Filter Implementation to System Identification

The preceding subsection has identified F, G and H matrices in the signal model necessary to generate the time series with desired statistical characteristics. It is time to focus our attention on optimal state estimation rather than signal generation.

First, the Kalman filter is an optimal state estimator of state vector, x. The state vector, we wish to estimate consist of the collection of 8 system coefficients; 4 autoregressive (AR) and 4 exogenous (X) coefficients. The state vector is defined as the column in equation 4.21.

The measurement equation determines how the measurement relates to the state vector.

The ARX(4,4) process measurement equation for the filter is given below with $v_k = \omega_k$:

$$z_k = \begin{bmatrix} z_{k-1} & z_{k-2} & z_{k-3} & z_{k-4} & u_{k-1} & u_{k-2} & u_{k-3} & u_{k-4} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} + \omega_k = H_k x_k + v_k \qquad 4.21$$

The measurement matrix, H, at time k consists of the four previous measurements and the four previous values for the (periodic) controls generating the process. When identifying the coefficients in this implementation, we expect that the first four measurements and controls fill up the measurement matrix. Subsequent measurements and controls refill the measurement matrix left to right from most recent to most past. At every time step after completion, the measurement and value at the most distant past is lost; they fall out of the measurement vector, which contains only the 4 most recent measurements and controls.

Turning from the measurement equation to the state evolution equation for the Kalman filter, consider how the state is assumed to evolve. If the system coefficients are considered to be constant, the state transition matrix, Φ, in this implementation is the identity matrix I. The state at time k+1 is equal to the state at time k. In addition, if the state are truly constant, we should not add white noise in the state evolution equation below:

$$x_{k+1} = F_k \ x_k + G_k \ w_k = I \ x_k = \ x_k, \ \ with \ \ w_k \ = \ 0 \ or \ equivalently \ with \ Q = 0. \quad 4.22$$

Setting the process noise variance Q to zero is equivalent to zeroing the noise term in the state evolution equation.

In practice, even constant system coefficients are often modeled as a white noise process and are given a small but finite value for the process noise variance Q. This artifice enhances filter stability. (Note: The Kalman filter equations involve the reciprocal or matrix inverse. If both noise variances Q and R are too small, this can cause matrices in the filter to be ill-conditioned). The values for the states and observables are roughly ~ 1. The small but finite noise variances we choose in the validation are 10^-12, which is small in comparison to the states or observables. This is defined as the zero noise case.

4.9.1 The zero input case (no forcing)

Considering the case with zero input is instructive. The initial state prescribed in the Kalman filter is usually chosen arbitrarily, because as more measurement information becomes available, the optimal state estimates become insensitive to (independent of) the initial state, i.e. steady state is reached. Without any a priori information, the state is usually taken to be the zero vector, i.e. all system coefficients are taken to be zero initially

The Kalman filter also requires an initial state error covariance. In practice, this matrix is taken to have a large value. After several iterations, the optimal estimates are largely insensitive to this value. Results for the state error covariance $P_0 = 99.99999$ are discussed below. The results for 999999.99999 are similar. Because the state components (system coefficients) are of the order unity, the state error covariance should be sufficiently large when set to 100.

The controls taken to be zero (no forcing) and there zero noise results in the following matrix equation:

$$z_k = \begin{bmatrix} z_{k-1} & z_{k-2} & z_{k-3} & z_{k-4} & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}. \qquad 4.23$$

The Kalman filter estimates the state based on comparisons between true measurements (observed) and measurements expected based on the calculation in equation 4.23. In this special zero noise and zero input case, (equation 4.23), the measurement expected has zero contribution from the exogenous coefficients, d_{1-4}, independent of their value. Once these coefficients are set to their initial values, they never get a correction from the filter, so that they stay at their presumed initial values, in this case zeros. Without any forcing u_k or noise, the Kalman filter loses its ability to identify the exogenous system coefficients.

Considering only AR coefficients, we have at k = 5, one equation with 4 unknowns coefficients a_{1-4}. The filter can update the initial values (0), based on this limited information. At step k = 6, a second equation is available to the filter. This is only one additional equation, but the information from the first equation at k = 5 was already used to determine the first estimate. All told, after step k = 6, the filter has used two equations for the 4 unknown autoregressive coefficients.

By step 8, the filter has the information content of 4 equations for the 4 unknown autoregressive coefficients. In the zero-input, zero noise case, we anticipate a perfect

identification for the four AR coefficients. On the other hand, because there are no noise sources

and no input on the right hand side of the ARX(4,4) equation, there is no identification for the

exogenous coefficients, which remain at their presumed values, zero.
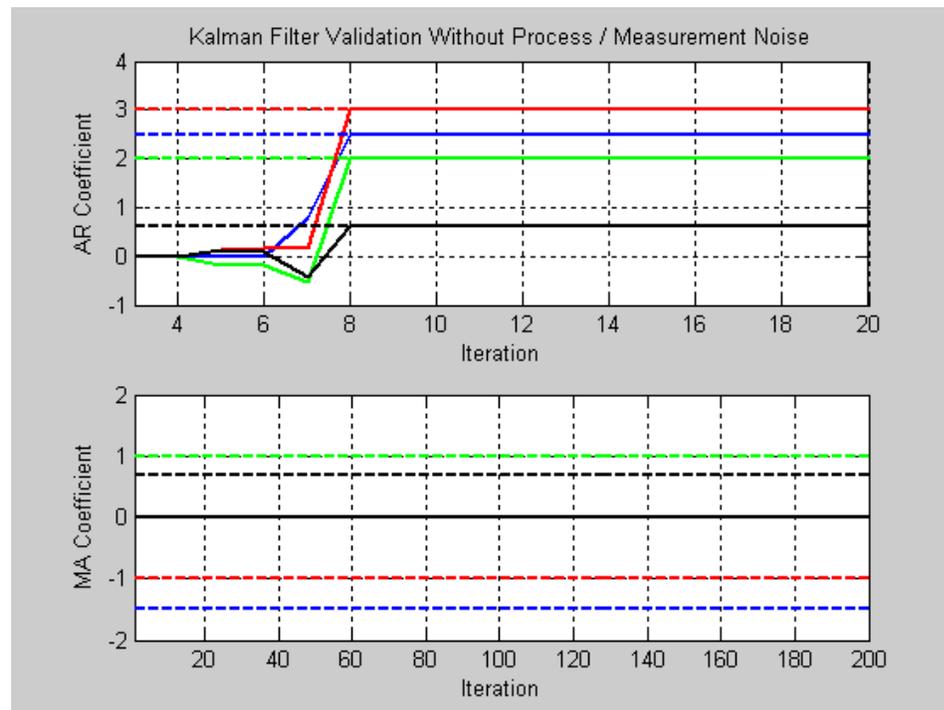
Figure 4.2: Perfect Identification of AR Coefficients; No Identification for Exogenous
Coefficients (labeled MA Coefficient).

Figure 4.2, shows the result of a zero-noise validation check on the Kalman filter code

and implementation. The signal model with zero noise and zero input generates the

measurements input into the Kalman filter. The four AR coefficients, a 1-4, and exogenous

coefficients 1-4, are (blue, red, green, and black) traces respectively. The dotted lines indicate

the values of the system coefficients used in the signal model. As expected, the traces begin at

initial values (zero). Until $k = 4$, the measurement vector is filling and the filter does no

correction. The iteration number $k = 5$ in the process, shows the first correction to the state

estimate. The iteration k = 8, results in a perfect fit for the AR coefficients. Because there is

zero noise and zero forcing, the exogenous coefficients are never adjusted from their initial zero

values. Only the (last) black trace at zero is visible in the plot of the exogenous coefficients,

because it overplots the other traces. This behavior of the optimal estimates is expected for the

zero noise, zero control case, and the Kalman filter implementation to system identification in

the ARX(4,4) process passes the first test.

The state error covariance history should also be examined. The error covariance for

iterations 4-8 are respectively (The Joseph form enforces the symmetry so that only the upper

triangular elements are shown): (The matrix at K = 8 is zero to at 8 decimal places)

99.999999	0	0	0
0	99.999999	0	0
0	0	99.999999	0
0	0	0	99.999999

99.985286	0.588148	-0.882223	0.588148
-	76.474040	35.288924	-23.525949
-	-	47.066603	35.288924
-	-	-	76.474040

78.058167	27.581703	-3.360347	-30.670770
-	43.243409	38.339639	14.955592
-	-	46.786534	31.756153
-	-	-	31.911868

15.891491	26.205305	23.612173	9.610242
-	43.212935	38.936823	15.847432
-	-	35.083851	14.279258
-	-	-	5.811711

0	0	0	0
-	0	0	0
-	-	0	0
-	-	-	0

It is easy to see from the covariance that the state estimate improves every time step. There is zero error after iteration 8, because there is a perfect fit between the estimates and the state itself. Without noise and forcing, determining the AR coefficients is simply a matter of fitting four points with four linear equations. Adding consistent equations after this point does not correct the estimate further.

Several other sets of system coefficients validated the Kalman identification of AR coefficients in the ARX time series models with similar results for the zero-noise zero input case.

4.9.2 The ARX Model driven by White Noise: (The ARMA(4,4) model)

It is somewhat more interesting to add noise to the signal model and / or filter and consider the effect of noise on the filter's ability to provide optimal estimates. In this case, the signal model generates measurements with a noise variance of $R = 10^{-3}$. The signal model also perturbs the coefficients by adding white noise with $Q = 10^{-6}$ and generating the measurements with these system coefficients. The measurement noise variance put into the filter was $R = 10^{-3}$, the same measurement noise variance as the model. The process noise variance for the filter $Q = 10^{-6}$ for the AR and MA coefficients.

The test indicates the AR coefficients converge to their set values more quickly than the MA coefficients. Both sets of coefficients require more iterations to converge to their set values

in the presence of noise. We note by adding a non-zero process noise in the signal model, we are perturbing the system coefficients. A normal random number generator simulates the perturbations to the ARMA(4,4) coefficients as white noise with zero mean and unit variance. The random variables are scaled to have variance

10^{-6}. Because the coefficients "jitter", the traces may approach the dotted lines indicating the means of the system coefficients but they do not converge except in a probabilistic sense.

The zero-noise case argued earlier determined AR coefficients recursively with 4 or more linear equations and 4 unknowns. After adding noise, a good estimate of the AR equations requires just a few more equations (iterations). Fitting AR coefficients with a least squares
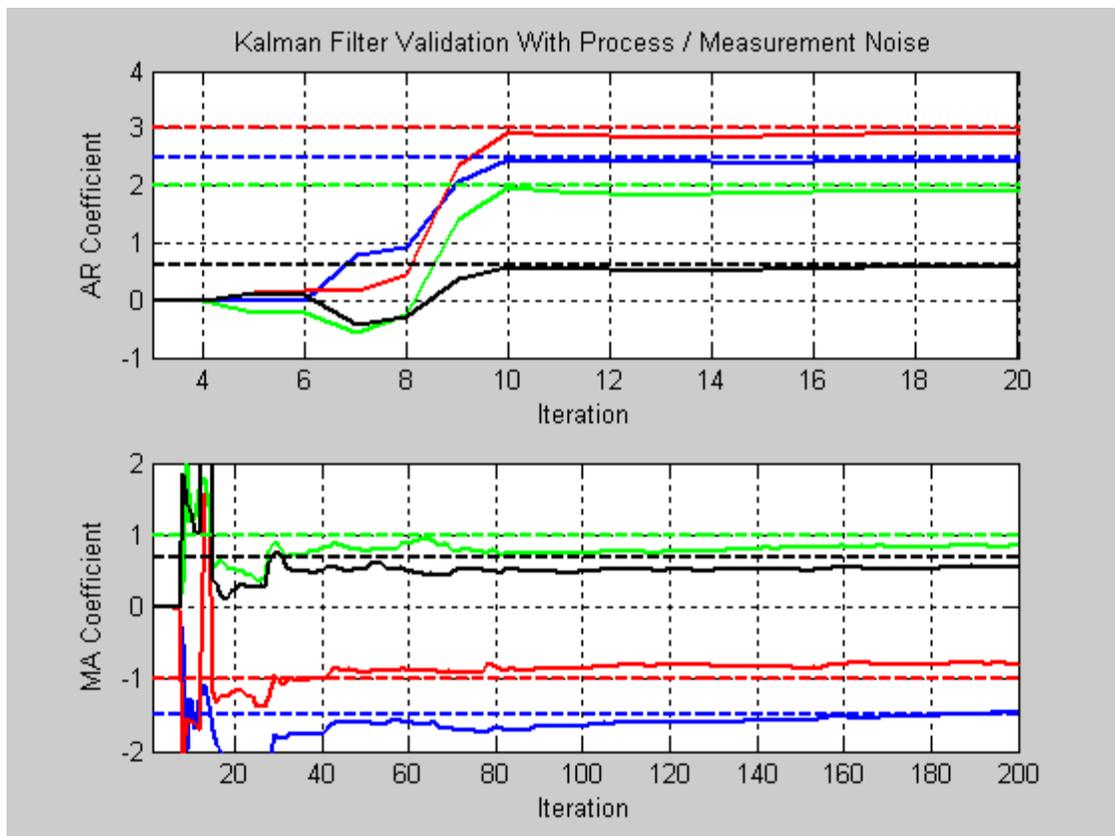
Figure 4.3: ARMA System Identification (white noise forcing)

criterion is a useful concept. We may recall any identification of the MA coefficients was impossible without some noise forcing the system. When noise is added, a good estimate of the MA coefficients may require many more iterations than the AR coefficients.

With these selected noise variances, the Kalman filter does eventually identify the MA coefficients after perhaps 10000 iterations, while only about 10 iterations are needed for the AR coefficients. In view of this result, we can also expect exogenous (X) coefficients with periodic forcing will require more iterations to converge than the AR coefficients. After discussing the ARX model results, we revisit this point.

The following table shows the optimal state estimate for the noisy ARMA coefficients after 200 iterations and after 10000 iterations. It is easy to see continual, but slow convergence to the actual values for the system coefficients.

Actual Coefficient	200 Iterations	10000 iterations
a1 = 2.5	2.367	2.48223889
a2 = 3.0	2.78697616	2.97138285
a3 = 2.0	1.80835931	1.96978281
a4 = 0.6	0.52174036	0.58854789
d1 = -1.5	-1.47203633	-1.50954257
d2 = -1.0	-0.78256965	-0.96604815
d3 = 1.0	0.84264216	0.98697945
d4 = 0.7	0.55112632	0.67543723

Table 3: System Identification of autoregressive and exogenous coefficients in an ARMA process

4.9.3 Kalman filter Identification in the ARX(4,4) Model

The ARX(4,4) model is formally similar to the ARMA(4,4) model except with prescribed controls forcing the system instead of white noise forcing. Here the white noise forces in the

state evolution equation are replaced with sinusoidal functions. Modeling the white noise

measurement error in the output equation still requires a white random number generator.
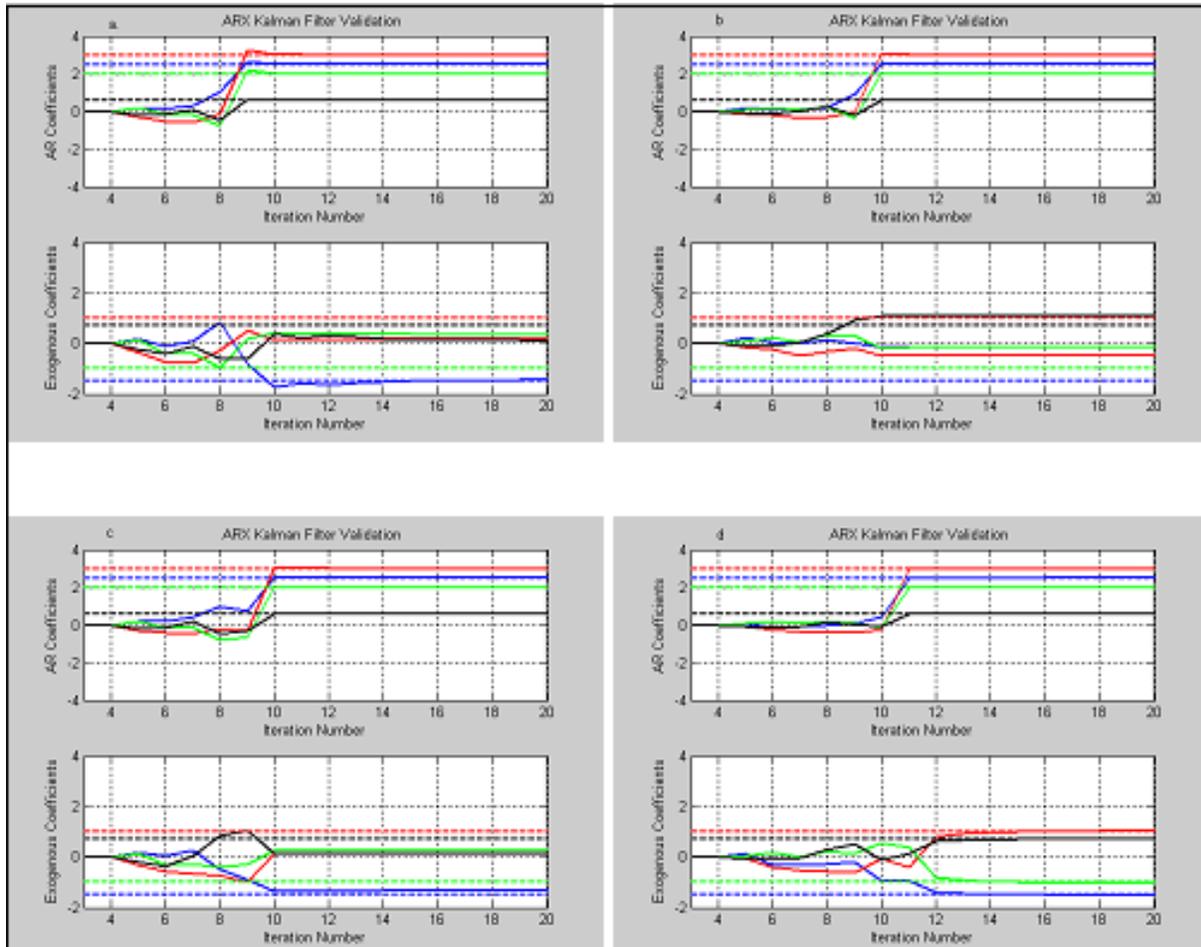
Figure 4.4: System Identification for an ARX(4,4) process with Periodic Forcing: a) frequency remote from system pole frequencies; b) frequency near the lower pole frequency; c) frequency near the higher pole frequency; d) frequencies near both lower and higher system pole frequencies

In this test, the Kalman filter estimated four AR states and four X states from the

ARX(4,4) model driven by unit amplitude sine wave(s) for four different sets of frequencies.

The frequencies chosen were a.) 0.04 / iteration; b.) 0.804 / iteration; c.) 0.254 / iteration; and the combination of frequencies d.) (0.6 * cos (0.804 * k) + 0.8 * sin (0.254 * k)).

The polynomial with autoregressive coefficients a1, a2, a3, and a4:

$$x^4 + a_1*x^3 + a_2*x^2 + a_3*x + a_4 = 0, \qquad\qquad 4.24$$

has roots are four complex numbers with complex conjugate imaginary parts, listed as follows, +/- 0.80735 and +/- 0.25333. The first case (a) is a frequency away from the complex root in the polynomial in the ARX(4,4) model. We refer to this as the pole frequency. The second case (b) is near one of pole frequencies, and the third case (c) is near the other pole frequency. Sinusoidal waves drive both pole frequencies in the fourth case (d).

The system coefficient estimates below show the AR coefficients are readily identified in all of these cases. However, good estimates of the exogenous coefficients result only in the fourth case where both pole frequencies are excited.

In general, this result suggests the Kalman filter can effectively identify the coefficients in a general ARX (N,N) model if there is input power at frequencies corresponding to all modes of the system. The roots of the polynomial formed by the N autoregressive coefficients determine these modes.

The moving average coefficients in the ARMA model are associated with white noise sources rather than sinusoidal forces. White noise sources have power across all frequencies. This power is not concentrated at the pole frequencies as in case d, so the estimation will not be as efficient. Nevertheless, enough power is present near the two pole frequencies to provide a good estimate of the MA coefficients eventually as shown in Figure 4.3, and Table 3.

We may emphasize the time-series from the signal model was prepared to have prescribed autoregressive and exogenous coefficients in the ARX case with sinusoidal forcing, or autoregressive and moving average coefficients in the ARMA case (ARX case with white noise forcing). These prepared cases are ideal in the respect that we know a priori the results the filter should estimate, if it is performing effectively. We saw results to suggest the filter cannot estimate the exogenous coefficients unless all modes of the system (determine by roots of the polynomial formed by the autoregressive coefficients) are excited. These results would not be weakened if we used (unfortunate) solar wind speed forcing and logarithmic electron flux output instead of the prepared input and output, which did not excite all modes. All that would change is it would be difficult to know the filter would be performing poorly, since we would not know a priori what to expect. The input / output must be taken as given, not carefully prepared to have known coefficients, to readily assess the filter performance.

In addition, these results suggest periodicities in the input at frequencies that do not excite all system frequencies may hinder successful identification. It is better to have white noise (or colored noise with a sufficient frequency spread) input, than a periodic signal input that does not excite all system frequencies.

4.9.4 Implications to Space Weather Modeling

The periodograms in Figure 2.8 show variation in the frequencies associated with solar wind speed forcing (input) with solar cycle phase. In view of these filter results; this variation would suggest the performance of Kalman filter system identification is also likely to have solar cycle phase dependence. In addition, because the Kalman filter implementation is equivalent

with recursive least squares, any linear analysis for system state identification will behave similarly.

As a final conclusion to this study, it is appropriate to provide the four autoregressive and four exogenous system coefficients estimated by the Kalman filter with actual electron flux (output) and with actual solar wind speed (input) measurements in the ARX (4,4) model. Figure 4.5 provides the system coefficients as they vary in time from 1996 to August 2003. The first autoregressive coefficient, a1 = 0.5 to 0.6. The second autoregressive coefficient, a2 is approximately 0.1. The remaining two autoregressive coefficients are nearly zero after the year 2000. Because the first two autoregressive coefficients are dominant over a3, and a4, the ARX(4,4) model should probably be simplified to a simpler model and reformulated. The ARX(2,2) or ARX(2,1) model, which Klimas et al. interprets as a stochastically driven damped linear oscillator may be a more appropriate model here. The first two exogenous coefficients are also the most significant, but to a lesser degree. It is surprising that the coefficients do not vary to a greater degree.
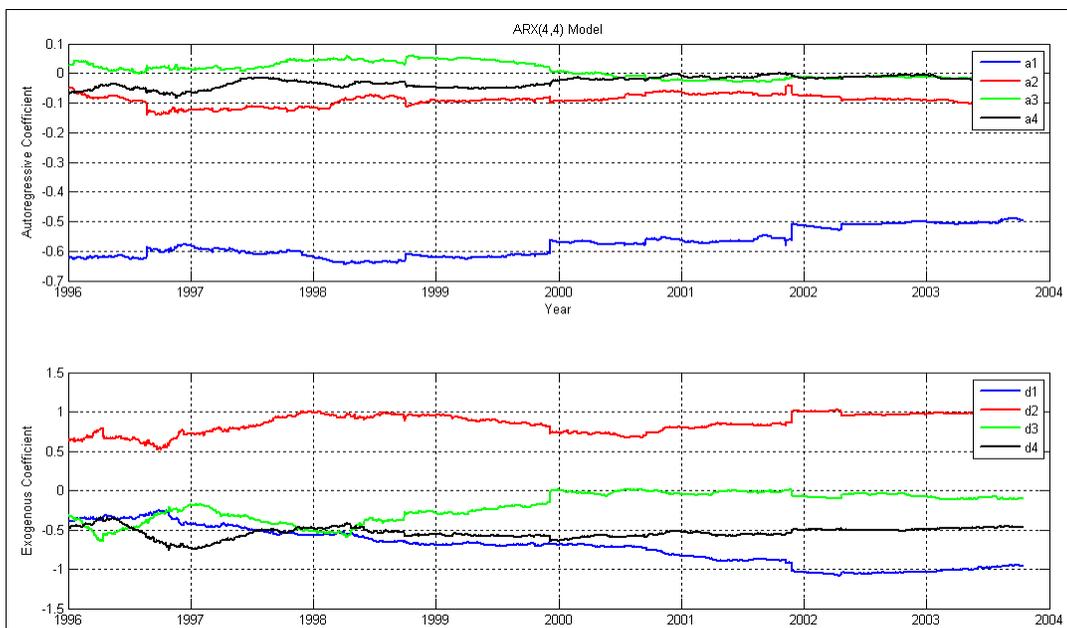
Figure 4.5 System Identification for Electron Flux with Solar Wind Speed Forcing Coefficients.

Summary:

The central theme in this thesis and effort has been the relationship between frequencies in the solar wind speed forcing to frequencies in the electron radiation belt response. During the declining phase of the solar cycle leading to solar minimum SC-23, the power spectrum in solar wind speed exhibits strong distinct peaks at periods of 27, and 13.5-days. In addition to the main (27-day) and secondary (13.5-day) peaks, a narrow (tertiary) peak with a 9-day periodicity was evident during 2003-5 during the declining phase of the solar cycle leading to solar minimum SC-24.

High-speed streams from trans-equatorial coronal holes are the cause of these peaks in the power spectrum with these periodicities. In each case, the relativistic electron flux response to the high-speed streams occurred at similar frequencies (i.e. similar periodicity). By examining different relativistic electron flux responses at various L-shells to the same solar wind forcing during the declining phase, it was often possible to determine qualitatively what at L-shell the solar wind speed transferred the power most effectively. For example at the 13.5-day peak in electron flux during the declining phase leading to SC-23 (1994-1996), the at $L = 6.5$ closely matched the corresponding peak in the solar wind power spectrum, implying efficient power transfer from solar wind speed to electron flux. Off that peak L at $L = 5.5$ or at $L = 7.5$, there is a small reduction in the peak in the electron flux periodogram indicating a slightly reduced power transfer. This reduction becomes dramatic at $L = 4.5$ or at $L = 8.5$. Similar results hold for the 9-day peak during 2005-6 at a peak $L = 5.5$. Thus qualitatively, L-shells with peak power transfer between solar wind speed and electron flux are likely to be in the range between 5 and 7.

More quantitatively, the peak power in the relativistic electron flux distributed across L-shells is shown in figure 2.9. During solar minimum and solar maximum (for SC-23), the peak power is at L = 5.0. During the two-year interval leading to solar minimum SC-23 and during the two-year interval after solar maximum, the L-shell with peak power has increased to L = 5.7. During the remaining two-year intervals, L-shell with peak power has moved inward. In each case, the associated pairs show a power distribution across L-shells with quite similar shape. In general the L-shell with peak power (during the intervals where a peak in figure 2.9 is evident) is distributed in the range from 5.0 - 5.7. This indicates a likely association between the L-shell with peak power with a P_1 peak in the impulse response function from Vassiliadis et al. I note Vassiliadis also used SAMPEX electron flux measurements and solar wind measurements to obtain his impulse response function so perhaps this is to be expected. The agreement in the association is still comforting. The association Vassiliadis et al. makes for the P_1 peak with high-speed streams is also reinforced with the power distribution across L-shells presented in figure 2.9.

During ascending phases in SC-23, the periodicities in the CME number obtained by Lara et al. are strongly supported by periodicities found using the Lomb-Scargle periodogram, and Welch periodogram with synthetic data (filled in measurements) having the same statistical characteristics. The weak 36-day periodicities in the relativistic electron flux shown in the panels b and d of figure 2.8 may be the result of the 36-day CME number periodicity reached by the consensus of spectral estimation algorithms.

A summary of results from linear estimation theory via the Kalman filter indicate successful identification of the exogenous (forcing) coefficients should require measurements generated, naturally or artificially, e.g. synthetic measurements, with frequencies present to

excite all (in our contrived case both) system modes. High-speed streams and CME's occurrences which are responsible for periodicity in the solar wind speed vary strongly with solar cycle phase. Thus the performance of a Kalman filter with solar wind speed forcing terms is most likely dependent on solar cycle phase, because frequencies exhibited by the solar wind speed depend strongly on these features.

System coefficients were identified using actual solar wind speed measurements and actual electron flux measurements. The system coefficients seem to be remarkable constant throughout the solar cycle during the years from 1996-2004. Given the dominance from the first two autoregressive coefficients, and the economy of a simpler model, the ARX(2,1) model or ARX(2,2) model is better justified.

REFERENCES:

[1] Miyoshi Y., and R. Kataoka, Ring Current Ions and Radiation Belt Electrons During Geomagnetic Storms Driven by Coronal Mass Ejections and Co-rotating Interference Regions, Geophysical Research Letters Vol 32, L21105 (2005)

[2] Turner, N. H., W. D. Cramer, S. K. Earles, B. A. Emery, Geoefficiency and energy partitioning in CIR and CME driven storms, Journal of Atmospheric and Solar-Terrestrial Physics 71 (2009) 1023-31.

[3] Kanekal S. G., A review of recent observations of relativistic electron energization in the Earth's outer Van Allen radiation belt, ILWS Workshop 2006 GOA, February 19-24, 2006

[4] Borovsky J. E., and M. H. Denton, Differences between CME driven Storms and CIR driven storms, JGR Vol 111, A07508 (2006).

[5] Gelb, Arthur, Applied Optimal Estimation, The MIT Press, 1974.

[6]   Vassiliadis D., A. J. Klimas, R. S. Weigel, D. N. Baker, E. J. Rigler, S. G. Kanekal, T. Nagai, S. F. Fung, R. W. H. Friedel, and T. E. Cayton, Structure of Earth's Outer Radiation Belt Inferred from Long-term Electron Flux Dynamics. Geophysical Research Letters Vol 30, No 19, 2015. (2003)

[7]   Vassiliadis D., A. J. Klimas, S. G. Kanekal, D. N. Baker, and R. S. Weigel, Long-term Average, Solar Cycle, and Seasonal response of magnetospheric energetic electrons to the solar wind speed, Journal of Geophysical Research, Vol 107, No 0 (2002)

[8] Lara A., Andrea Borgazzi, Odim Mendes Jr., Reinaldo R Rosa, Margarete Oliveira Domingues, Short Period Fluctuations in the Coronal mass Ejection Activity During Solar Cycle 23. Solar Phys(2008) 248 155-166 DOI 10.1007/s11207-008-9153-4

[9] N. R. Lomb, Least-squares frequency analysis of unequally spaced data, Astrophysics and Space Science, 39 #2 February 1976. Springer, Netherlands. Pp 447-62

[10] Jeffrey D Scargle, Studies in Astronomical Time Series II. Statistical Aspects of Spectral Analysis of Unevenly Spaced Data., The Astrophysical Journal 263 835-53 1982 December 15

[11]   Alan V Oppenheim and Ronald W Schafer, Digital Signal Processing, Prentice Hall Inc. Englewood Cliffs, New Jersey (1975)

[12]   Jago Lucy, The Northern Lights, Vintage Books, a division of Random House, New York 2002.

[13]   Greaves W M H., W. H. Newton, On the Recurrence of Magnetic Storms, Monthly Notices Royal Astronomical Society 89 641-65 1929.

[14] Nicole Meyer-Vernet, Basics of the Solar Wind, Cambridge University Press, (2007)

[15] Conway W. Snyder, Marcia Neugebauer, and U. R. Rao, JGR 68 Dec 15, 1963 6361

[16], Sir William Thomson (Baron Kelvin), Public Lectures and Addresses Vol II. New York; Macmillan and Co, from the New York Times May 24, 1894

[17] Kivelson and Russell, Introduction to Space Physics, Cambridge University Press (1995)

[18] Baker D. N., R. D. Belian, P. R. Higbie, and E. W. Hones Jr., High-Energy Magnetospheric Protons and Their Dependence on Geomagnetic and Interplanetary Conditions. Journal of Geophysical Research, Dec 1, 1979. pp. 7138.

[19] Higbie, P. R., D. N. Baker, E. W. Hones Jr., and R. D. Belian, Pitch Angle Distributions of > 30 keV. Electrons at Geostationary Altitudes, 1979 from Olsen W. P. Quantitative Modeling of Magnetospheric Processes, 1979

[20] Paulikas and Blake 1979 from Olsen W. P. Quantitative Modeling of Magnetospheric Processes, 1979

[21] Baker D. N., J. B. Blake, R. W. Klebesadel, and P. R. Higbie, Highly Relativistic Electrons in the Earth's Outer Magnetosphere. 1. Lifetimes and Temporal History 1979-1984. JGR Vol 91, No A4, pg. 4265-76 April 1, 1986.

[22]   Christon S. P., D. L. Chenette, D. N. Baker, and D Moses, Relativistic Electrons at Geosynchronous Orbit, Interplanetary Electron Flux and the 13 Month Jovian Synodic Year. GRL Vol 16 No 10, pages 1129-1132 October 1989.

[23] Li, Xinlin, D. N. Baker, M. Temerin, D. Larson, R. P. Lin, G. D. Reeves, M Looper, S. G. Kanekal, and R. A. Mewaldt, Are Energetic Electrons in the Solar Wind the Source of the Outer Radiation Belt. GRL Vol 24 No 8, 923-6 April 15, 1997.

[24] Nishida A., Outward Diffusion of Energetic Particles from the Jovian Radiation Belt, JGR Vol 81 No 16, April 1 1976.

[25] Nishida A., Non-adiabatic Acceleration of Magnetospheric Particles, Adv. Space Research Vol 12 No 8, pp (8)15-8(22) 1992

[26] Fujimoto M. and A Nishida, Energization and Anisotropization of Energetic Electrons in the Earth's Radiation belt by the Recirculation Process JGR Vol 95 No A4 April 1, 1990 pg 4265-70.

[27]   Baker D. N., J. B. Blake, L. B. Callis, R. D. Belian and T. E. Cayton, Relativistic Electrons Near Geostationary Orbit: Evidence of Internal Magnetospheric Acceleration, GRL Vol 16 No 6 pg 559-62 June 1989

[28] Baker D. N., R. L. McPherron, T. E. Cayton and R. W. Klebesadel, Linear Prediction Filter Analysis of Relativistic Electron Properties at 6.6 Re. JGR Vol 95, A9, pg 15133-15140. Sept 1, 1990.

[29] Gibson S. E., J. U. Kozyra, G. de Toma, B. A. Emery, T. Onsager, and B. J. Thompson, If the Sun is So Quiet, Why is the Earth Ringing? A Comparison of Two Solar Minimum Intervals.

[30] Emery B. A., I G. Richardson, D Evans, and F J Rich, Solar Wind Structure Sources and Periodicities of Auroral Electron Power Over Three Solar Cycles. Journal of Atmospheric and Terrestrial Physics 71 (2009) 1157-1175.

[31] Mursula K., and B. Zieger, The 13.5-day periodicity in the Sun, solar wind, and geomagnetic activity: The last three solar cycles. Journal of Geophysical Research, Vol 101, No A12. pp 27077-27090. Dec 1, 1996.

[32] Vrsnak, Temmer, and A. Veronig, Coronal Holes and Solar Wind High-Speed Streams, I. Forecasting the Solar Wind Parameters, Solar Physics (2007) 240: 315-330.

[33] Temmer M., B. Vrsnak, A. Veronig, Periodic Appearance of Coronal Holes and the Related Variation of Solar Wind Parameters, Solar Physics (2007) 241: 371-83.

[34] Lei J., J. P. Thayer, J. M. Forbes, E. K. Sutton, and S. Nerem, Rotating Coronal Holes and Periodic Modulation of the Upper Atmosphere GRL V35 (2008) .

[35] Verma V. K. and G. C. Joshi, On the Occurrence Rate of High-Speed Solar Wind Events, Solar Physics 155   pg, 401-4 (1994)

[36] Low B. C., Solar Activity in the Corona, Invited Review. Solar Physics Vol 167 No 1-2 (1996) pg 217-65.

[37]   McComas D. J., H. A. Elliot, N.A. Scwadron, J. T. Gosling, R. M. Skoug, and B. E. Goldstein, The three-dimensional solar wind around solar maximum, GRL Vol 30 No 10, 1517 (2003)

[38] J. E. Insley, and V. Moore and R. A. Harrison, "The differential rotation of the corona indicated by coronal holes", Solar Physics 160 1-18, 1995.

[39] Richard C Altrock, "A study of the rotation of the solar corona", Solar Physics 213, 22-37, (2003)

[40] G. A. Paulikas and J. B. Blake, "Modulation of Trapped Energetic Electrons at 6.6 Re. by the Direction of the Magnetic Field", GRL Vol 3, #5, May 1976

[41] Baker D. N., and S. G. Kanekal, Solar Cycle Changes, Geomagnetic Variations, and Energetic Particle Properties in the Inner Magnetosphere. Journal of Atmospheric and Solar-Terrestrial Physics 70, (2008) 195-206.

[42] Skoug R. M., J. T. Gosling, J. T. Steinberg, D. J. McComas, C. W. Smith, N. F. Ness, Q Hu, and L. F. Burlaga, Extremely High Speed Solar Wind 29-30 October 2003, JGR 109 A09102 (2004).

[43]   Reeves G. D., K. L. McAdams, R. H. W. Friedel, and T. P. O'Brien, Acceleration and Loss of Relativistic Electrons During Geomagnetic Storms.

[44] Kanekal S. G., D. N. Baker, J. B. Blake, B. Klecker, G. M. Mason, and R. A. Mewaldt, Magnetospheric Relativistic Electron Response to Magnetic Cloud Events of 1997. Adv. Space Res Vol 25, No 7/8 pp 1387-92 (2000)

[45] Burlaga L. F. Physics of the Inner Heliosphere, V. 2., Schwenn R. and E. Marsch ed. Springer Berlin (1991).

[46] Baker D. N., G. M. Mason, O. Figuroa, G. Colon, J. G. Watzin, and R. M. Aleman, An Overview of the Solar, Anomalous and Magnetospheric Particle Explorer (SAMPEX) Mission, IEEE Transactions on Geoscience and Remote Sensing 31, #3 May 1993 531.

[47] M. H. Lennard, J. E. Mazur, and G. M. Mason, Sampex Level 2 data products NSSDC submission description Ver 1.3.

[48] S. G. Kanekal, D. N. Baker, and R. L. McPherron, "On the seasonal dependence of relativistic electron fluxes", Ann. Geophys. 28, 1101-6 2010.

[49] Onsager, T. G., R. Grubb, J. Kunches, L. Matheson, D. Speich, R. Zwickl, and H. Sauer, Operational Uses of the Goes Energetic Particle Detectors. SPIE Vol 2812 pg 281-90.

[50] http://www.swpc/noaa.gov/

[51] http://omniweb.gsfc.nasa.gov/form/dx1.html.

[52] Joe King, and Natalia Papitashvili, OMNI 2 Documentation, SPDF Goddard Space Flight Center

[53] Brown, R. Grover, and Hwang Patrick Y. C., Introduction to Random Signals and Applied Kalman Filtering, 2nd ed. John Wiley and Sons Inc. New York 1992.

[54] Grewal Mohinder S. and Andrews Angus P. , Kalman Filtering: Theory and Practice, 1993.

[55] Baker, D.N., X Li, J. B. Blake, and S. Kanekal, Strong Electron Acceleration in the Earth's Magnetosphere, Advances in Space Research, vol 21.no 4. 609-13 (1998)

[56] Vassiliadis D. A. , R. S. Weigel, A. J. Klimas, S. G. Kanekal, R. A. Mewaldt, Modes of energy transfer from the solar wind to the inner magnetosphere, Physics of Plasmas, Vol 10 No 2, Feb 2003.

[57] Kanekal S., D. N. Baker, and J. B. Blake, Multi-satellite measurements of relativistic electrons: global coherence Journal of Geophysical Research 106 (A12), 29721-32, (2001)

[58] Cover, T, and Thomas J, Elements of Information Theory, John Wiley and Sons, Inc. (1991)

[59] Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T., and Flannery Brian T. Numerical Recipes: The Art of Scientific Computing, FORTRAN version, Cambridge University Press (1992)

[60] Benevolenskaya, E. E. J. T. Hoeksema, A. G. Kosovichev, and P. H. Scherrer, The interaction of new and old magnetic fluxes at the beginning of solar cycle 23, The Astrophysical Journal 517:L163-L166 June 1, 1999.

[61] http://vso.nso.edu/cgi/catalogui

[62] V. Domingo, B Fleck, and A. I. Poland, The SOHO Mission: An Overview, Solar Physics V 162 No. 1-2 December 1995.

[63] Papoulis, Athanasios, Probability, Random Variables, and Stochastic Processes, McGraw-Hill Book Company, (1965)

[64] Ross, Sheldon, Introduction to Probability models, Academic Press, (1997).

[65] William H Press, and George B Rybicki, Fast Algorithm for Spectral Analysis of Unevenly Spaced Data, The Astrophysical Journal 338 277-280 1989 March 1.

[66] Klimas A. J., D. Vassiliadis, and D. N. Baker, Data Derived Analogues of the Magnetospheric Dynamics, Journal of Geophysical Research Vol 102, No A 12. pp 26993-27009.

[67] Anderson B. D. O. and J. B. Moore, Optimal Filtering Prentice-Hall, 1979