Variability of 40-300keV Electrons at Geosynchronous Orbit

Sarah Szita

Mullard Space Science Laboratory
Department of Space and Climate Physics
University College London

A thesis submitted to the University of London for the degree of Doctor of Philosophy

1998
Corrigenda

---

p15 line 1: 'Kivelson, 1995' should read 'Kivelson and Russell, 1995'

p19 line 6: Young citation missing from References section:

p46 line 27: 'geosyncronous' should read 'geosynchronous'

p48 line 18: 'will' should read 'with'

p72 line 7: 'desribed' should read described'

p73 line 17: 'ranges' should read 'range'

p150 line 24: 'Szita et al.' should read 'Szita et al. (1993)'

p157 line 5: 'sepctrum' should read 'spectrum'

p164 line 25: 'dervied' should read 'derived'

p171 line 27: 'highy' should read 'highly'

p175 line 22: 'develpoment' should read 'development'

p179 line 6: 'electons' should read 'electrons'

p186 figure 6.2: colour scale should be labelled 'linear'

p188 line 4: 'its makes' should read 'it makes'

p217 line 9: 'bein' should read 'being'

p221 line 15: 'trappedd' should read 'trapped'

p224 figure 7.8: 'tranforms' should read 'transforms' in caption

p226 figure 7.10: plot lines should be labelled:
upper line as 'E5-E4 flux', lower line as 'E1 flux'

p229 line 5: 'If is' should read 'If it'

p231 figure 7.12: x-axis should read labelled 'days'

p239 line 10: 'affect' should read 'effect'

p247 line 14: 'affect' should read 'effect'
for my father
Abstract

The energetic electron population at geosynchronous orbit is highly variable and affected by many different time-dependent processes. Substorm injections recur on time scales of hours, local time variations result from the geomagnetic field asymmetry, magnetic storms create periods of enhanced activity and quiet periods result in continual loss, seasonal variations are driven by changes in the Earth's magnetic field geometry around its orbit, and solar cycle variations occur on time scales of years. From 1988–1995, the SEM-2 (Space Environment Monitor) onboard the geostationary satellite Meteosat P2 detected 42.9–300keV electrons in five differential energy ranges. This energy range is ideal to investigate two important components of the geosynchronous environment: the trapped radiation belt population and the lower energy substorm injected electrons. The SEM-2's 30 look directions allow the determination of the symmetry axis of the particle distribution, and comparison with the Tsyganenko 89 magnetic field model shows that this provides a good indication of the magnetic field direction.

The dependence of radiation belt intensity and substorm injection signatures on local time and geomagnetic activity is quantified in models based on flux probability distributions, and the effects of including a time lag behind $K_p$ for the high energy data are explored. A longer term model is constructed to give the probable range of observed flux in terms of mission duration.

Injection frequency is investigated using wavelet analysis and found to depend on solar wind speed. Wavelets are also used to investigate a 16-day periodicity which may relate to previous observations interpreted as planetary wave signatures.

The solar cycle dependence of both populations is examined. The lower energy population peaks at solar maximum, but is also enhanced in the declining phase which has the dominant effect on the higher energy electrons.
Contents

1 The Earth's Magnetosphere .................................................. 13
  1.1 The Solar Wind ............................................................. 13
  1.2 The Magnetosphere ....................................................... 17
  1.3 The Radiation Belts ....................................................... 21
    1.3.1 Trapped Particle Motions ....................................... 22
    1.3.2 Drift Shell splitting ............................................... 31
    1.3.3 Quasi-trapping ..................................................... 34
  1.4 Magnetospheric Substorms ............................................ 39
    1.4.1 Magnetic Storms .................................................. 39
    1.4.2 Substorms .......................................................... 40

2 The Geostationary Environment ........................................... 50
  2.1 Observations at Geostationary Orbit .............................. 51
  2.2 Energy Spectra .......................................................... 65
  2.3 Applications of the SEM-2 Data .................................... 73

3 The Meteosat SEM-2 Data Set ............................................. 75
  3.1 Meteosat P2 ............................................................... 75
  3.2 SEM-2 .................................................................... 76
    3.2.1 Calibration ......................................................... 78
    3.2.2 Telemetry ........................................................ 81
  3.3 The SEM-2 Archived Data Set ....................................... 82
    3.3.1 Archived data .................................................... 82
    3.3.2 Calculation of the Anisotropy Index ....................... 87
7.2 Solar Rotation ................................................................. 223
7.3 16-day Periodicity ....................................................... 227
7.4 Mission duration .......................................................... 235
7.5 Conclusion ................................................................. 239

8 Concluding Remarks ...................................................... 242

A SEM-2 Errors ................................................................. 248
  A.1 Archiving Errors ............................................................ 248
     A.1.1 Omissions ............................................................... 248
     A.1.2 Data Clipping .......................................................... 248
     A.1.3 Polar Flux Asymmetry ............................................. 249
  A.2 Onboard Data Compression Error .................................. 250
     A.2.1 SEM-2 Data Compression Report ............................ 261
  A.3 Decompression Errors .................................................. 265
     A.3.1 Scaling Factor Error ............................................... 265
     A.3.2 Overflow Errors ..................................................... 266

B Coordinate Systems ....................................................... 268
  B.1 GEO Coordinates ........................................................ 268
  B.2 GSM Coordinates ....................................................... 269
List of Tables

2.1 LANL detector electron energy levels .......................... 52
2.2 LANL satellites and detectors ................................. 52
3.1 SEM-2's five differential energy levels ....................... 77
3.2 Elements of the archived SEM-2 data set record ........... 89
5.1 Total flux points in \( K_p \) ranges ......................... 147
7.1 SEM-2 yearly temperatures and densities ..................... 214
7.2 SEM-2 yearly temperatures and densities for high \( K_p \) ... 215
7.3 SEM-2 yearly temperatures and densities for low \( K_p \) ... 215
7.4 Yearly average diurnal variations ............................ 217
A.1 Percentage of counts affected by the compression error ... 260
A.2 Errors on ISCALE ............................................. 265
List of Figures

1.1 The Parker spiral .............................................................. 16
1.2 The Earth's magnetosphere .............................................. 18
1.3 Superposition of dawn-dusk and corotational electric fields .... 20
1.4 Radiation belts: electron and proton distributions ............... 23
1.5 The three trapped particle motions .................................... 24
1.6 The loss cone .................................................................... 27
1.7 Contours of $B$, $I$ and $L$ ............................................... 32
1.8 Split shells ........................................................................ 33
1.9 Shell splitting for particles starting on common field lines at noon 35
1.10 Shell splitting for particles starting on common field lines at midnight 35
1.11 Quasi-trapping regions in the magnetosphere .................... 36
1.12 The drift loss cone .......................................................... 37
1.13 Magnetopause shadowing ............................................... 38
1.14 Reconnection ................................................................. 44
1.15 Substorm variations of $H$, $AE$ and particle flux ............... 47

2.1 MPA data showing local time distribution of geosynchronous orbit regions ......................................................... 55
2.2 Cayton et al.'s energy spectra ............................................. 66
2.3 SEM-2 energy spectrum showing Cayton et al.'s data .......... 67
2.4 Energy spectrum of SEM-2 data ........................................ 69
2.5 Energy spectrum dependence on local time ....................... 71
2.6 Energy spectrum of AE-8 data ......................................... 72

3.1 Meteosat P2's longitude, 1988-1995 ................................. 76
5.9 Statistical plot of E1 flux versus local time ........................................ 130
5.10 Average and median flux versus local time .......................................... 132
5.11 Magnetic field contours in the Earth’s equatorial plane .......................... 133
5.12 Outer magnetosphere drift paths .......................................................... 134
5.13 LANL statistical plots ........................................................................ 137
5.14 Low energy flux in the absence of injections ......................................... 139
5.15 Flux dropouts ....................................................................................... 141
5.16 Local time dependence of flux decreases .............................................. 142
5.17 Adiabatic modelling of a flux decrease .................................................. 144
5.18 Statistical plots of E5-E4 flux in ranges of $K_p$ ..................................... 148
5.19 Statistical plots of E4-E3 flux in ranges of $K_p$ ..................................... 149
5.20 Statistical plots of E3-E2 flux in ranges of $K_p$ ..................................... 149
5.21 Statistical plots of E2-E1 flux in ranges of $K_p$ ..................................... 151
5.22 Statistical plots of E1 flux in ranges of $K_p$ .......................................... 151
5.23 Correlation of flux with $K_p$ ................................................................. 152
5.24 Statistical plots of E1 flux in ranges of $K_p$, time lag applied ............... 153
5.25 Statistical plots of E1 flux in ranges of $K_p(\tau)$ .................................. 154
5.26 Identification of average local time behaviour ...................................... 157
5.27 Anisotropy index versus local time ...................................................... 159
5.28 Anisotropy index versus local time in ranges of $K_p$ ............................ 160
5.29 Spectral index versus local time ............................................................ 162
5.30 Spectral index versus local time in ranges of $K_p$ ................................ 162
5.31 Total flux versus anisotropy index, in ranges of $K_p$ ............................. 165
5.32 Total flux versus spectral index, in ranges of $K_p$ ................................. 166
5.33 Anisotropy index versus spectral index, in ranges of $K_p$ ................. 167
5.34 Example pitch angle distributions ........................................................ 168
5.35 Calculation of pitch angle ................................................................. 170
5.36 Calculated pitch angle distribution dependence on local time ............. 172
5.37 Pitch angle distribution dependence on local time, low $K_p$ .............. 173
5.38 Pitch angle distribution dependence on local time, high $K_p$ ............ 174
5.39 Energetic electron pitch angle distributions in the near equatorial magnetosphere .............................................................. 176
5.40 Geometry of the injection boundary ........................................... 180

6.1 Morlet wavelet ..................................................................................... 184
6.2 Wavelet transforms of common signals .......................................... 186
6.3 Wavelet transform showing substorm and diurnal frequencies .......... 190
6.4 Wavelet determination of injection frequency ................................. 191
6.5 Injection frequency throughout the mission ..................................... 194
6.6 Monthly averaged flux for each SEM-2 energy range for the whole mission ............................................................................. 195
6.7 Monthly averaged solar wind speed for the duration of the mission 196
6.8 Monthly averaged $K_p$ for the duration of the mission .................... 196
6.9 Substorm frequency and solar wind speed ....................................... 197
6.10 Superposed epoch analysis of peaks, LT 2300-2400 ....................... 200
6.11 Superposed epoch analysis of peaks, LT 0600-0700 ....................... 201
6.12 Superposed epoch analysis of peaks, LT 1200-1300 ....................... 202
6.13 Superposed epoch analysis of peaks, LT 1800-1900 ....................... 203
6.14 Pitch angle distribution of injection peaks ...................................... 205
6.15 Peak flux versus $K_p$ across the nightside ...................................... 207

7.1 Monthly averaged sunspot number for the duration of the mission 211
7.2 Solar flare occurrence, 1991-1996 ...................................................... 213
7.3 Yearly temperature and density ............................................................. 216
7.4 Yearly E1 statistical plots ..................................................................... 218
7.5 Yearly E5-E4 statistical plots ................................................................. 219
7.6 Flux of $>1.4$MeV electrons versus solar wind speed ....................... 222
7.7 E5-E4 and E1 flux versus solar wind speed ....................................... 222
7.8 Wavelet transforms of daily averaged data .................................... 224
7.9 Average wavelet modulus of daily data ........................................... 226
7.10 The 27-day periodicity throughout the mission ............................... 226
7.11 Wavelet transforms over 10-20 day periods ................................... 230
7.12 16-day wavelet transforms ................................................................. 231
7.13 Histogram of total flux ....................................................................... 235
7.14 Average total flux versus accumulation time ..................................... 237
7.15 Year-long flux averages ..................................................................... 237

A.1 The polar flux error ............................................................................ 249
A.2 Polar-azimuthal flux distribution ........................................................ 250
A.3 Input ranges for SEM-2 data compression .......................................... 252
A.4 Decompression output numbers and errors ........................................ 253
A.5 Corrected input ranges and percentage errors for compression .......... 254
A.6 Compression errors for exponent 4 ...................................................... 256
A.7 Compression errors for exponent 5 ...................................................... 257
A.8 Compression errors for exponent 6 ...................................................... 258
A.9 Compression errors for exponent 7 ...................................................... 259
A.10 Effect of compression error on counts ............................................... 260

B.1 The GEO coordinate system ................................................................. 268
B.2 The GSM coordinate system ............................................................... 269
Chapter 1

The Earth’s Magnetosphere

This chapter discusses the near-Earth environment: the solar wind which streams out through the solar system, the magnetosphere which is the domain of the geomagnetic field, and the radiation belts, where energetic charged particles are trapped in the Earth’s magnetic field.

1.1 The Solar Wind

The solar wind comprises particles from the outermost region of the solar atmosphere, the corona, which have high thermal energies enabling them to escape the Sun’s gravitational field (Parks, 1991). The result is a near-radial outflow of plasma. According to Chen (1974), a plasma is a quasineutral gas of charged and neutral particles which exhibits collective behaviour.

Quasineutrality requires that the scale length of the plasma is large compared with the Debye length of the plasma. An isolated charge has an electrostatic potential $\phi$ given by:

$$\phi = \frac{q}{4\pi\varepsilon_0 r}$$  \hspace{1cm} (1.1)

where:

$q$ is charge
$\varepsilon_0$ is the permittivity of free space
$r$ is the distance from the charge.

Introducing an isolated charge into a plasma would result in a charge cloud of
the opposite sign collecting around it, which then shields the affect of the charge from the rest of the plasma. The Debye length $\lambda_D$ is a measure of shielding in the plasma. In this case $\phi$ is called the Debye screening potential:

$$\phi = \frac{q}{4\pi\varepsilon_0 r} e^{-r\lambda_D}$$  \hspace{1cm} (1.2)$$

where the Debye length $\lambda_D$ for an electron-proton plasma is given by:

$$\lambda_D^2 = \frac{kT_e\varepsilon_0}{n_0q_e^2}$$  \hspace{1cm} (1.3)$$

where:

- $k$ is Boltzmann's constant
- $T_e$ is the electron temperature
- $q_e$ is the electron charge
- $n_0$ is the plasma density (Parks, 1991).

If $r \ll \lambda_D$, then equation 1.2 shows that the potential becomes the Coulomb potential. If $r > \lambda_D$, then the potential decreases exponentially. Therefore outside the Debye sphere (a sphere of radius $\lambda_D$ centred on the charge) this potential decreases rapidly, so the Debye sphere can be regarded as the sphere of influence of the charge. Thus, if the plasma scale length is much larger than $\lambda_D$, it can shield out an applied potential in a distance very short compared to its size, i.e. most of the plasma is unaffected.

The density of the plasma is required to be low enough that short-range collisions between particles are negligible, so that long-range electromagnetic forces are dominant. The plasma exhibits 'collective' behaviour, where many of the charged particles react together to electromagnetic forces. The number of particles in the Debye sphere, $N_D$, is given by:

$$N_D = \frac{4\pi n_0 \lambda_D^3}{3}$$  \hspace{1cm} (1.4)$$

and the plasma parameter $g$ is defined by:

$$g = \frac{1}{N_D}$$  \hspace{1cm} (1.5)$$

The condition $g \ll 1$ is called the plasma approximation: this condition must be met for there to be enough charged particles present for them to be considered
a plasma, because there must be enough particles for Debye shielding (Kivelson, 1995). Equations 1.3, 1.4, and 1.5 combine to give:

\[ g \sim \frac{n_0^{\frac{1}{3}}}{T^{\frac{2}{3}}} \]  

(1.6)

So for \( g \ll 1 \), some combination of high temperature and low density is required. If \( g \) is small there are few collisions, and as \( g \) tends to zero the plasma becomes collisionless.

The solar wind plasma has the magnetic field from its point of origin, at the Sun, frozen into it: this is because in an electrical conductor magnetic fields diffuse slowly. As the Sun rotates 'away' from the outflowing particles and magnetic field, it produces the Parker spiral (figure 1.1). The particles stream out radially but the magnetic field 'winds up' since the field lines are rooted on the Sun. The latitude band \( \pm 30^\circ \) on the Sun is divided into sectors of opposite polarity, which creates sectors in the solar wind. The magnetic field carried by the solar wind is referred to as the Interplanetary Magnetic Field (IMF).

The solar wind streams out from the Sun through the entire solar system and envelopes the Earth's magnetosphere. At the Earth, solar wind electron and proton number densities are typically 5 particles cm\(^{-3}\), and the solar wind velocity is on average 400 km\(\text{s}^{-1}\). These parameters are dependent on solar activity and can vary widely. Where the Sun's magnetic field is weak, high speed solar wind streams can result. Coronal holes, coronal mass ejections and solar flares are all sources of high speed solar wind. Large coronal holes have been linked with geomagnetic disturbance (Neupert and Pizzo, 1974). The link between solar and geomagnetic activity was shown to have two components by Venkatesan et al. (1991). The \( K_p \) index, a measure of geomagnetic activity, has two maxima for each maximum in the sunspot number, which provides a measure of solar activity and peaks every 11 years. \( K_p \) sums for sudden storm commencement (SSC) days correlate well with the number of SSC days and the correlation peaks at solar maximum. \( K_p \) sums for non-SSC days also correlate with the number of non-SSC days but the peak correlation occurs 5 years after solar maximum. These two discrete components of geomagnetic activity result from flare activity, with the shock front from the flare causing a SSC, and corotating solar wind streams.
Figure 1.1: The Parker spiral (Smith and Jacobs, 1973). The thin arrows show the magnetic field direction, and the thick arrows show the direction of the motion of solar wind particles.
1.2 The Magnetosphere

The Earth’s magnetic field is confined by the magnetized solar wind, to a region around the Earth known as the magnetosphere. Figure 1.2 shows the main features of the Earth’s magnetosphere. The boundary between the domain of the geomagnetic field and the solar wind is called the magnetopause. The magnetosphere is compressed by the solar wind on the sunward side, with the magnetopause typically at about $10R_E$, and drawn out into a tail several hundred $R_E$ long in the anti-sunward direction. For this reason, although a dipole model for the Earth’s magnetic field, with the geomagnetic axis inclined to the Earth’s spin axis at 11.5°, is a reasonable approximation close to the Earth, it proves less accurate as altitude increases and the magnetosphere becomes increasingly asymmetric.

Meng (1970) showed that the position of the magnetopause correlates with substorm activity. From observations of magnetopause crossings by the elliptical satellite IMP-2 he deduced that during substorms the magnetosphere is compressed and so the magnetopause lies closer to Earth. Later Sibeck et al. (1991) studied 1821 magnetopause crossings from different satellite data sets and showed that the location of the magnetopause depended on the direction of the IMF and on solar wind dynamic pressure. In very disturbed conditions, the magnetopause has been observed inside geostationary orbit.

Because the solar wind is supersonic and superalfvénic when it encounters the Earth a bow shock forms, typically about $3R_E$ upstream of the magnetopause. This is a collisionless shock because the low density of the solar wind near the Earth means that the solar wind particles have a very long mean free path, of the order of the Sun–Earth distance, and so collisions are rare. Between the magnetopause and the bow shock lies the magnetosheath. A small fraction of the magnetosheath plasma ($\leq 1\%$) crosses the magnetopause into the magnetosphere and forms the boundary layer. This boundary layer plasma at low latitudes is called the low-latitude boundary layer whereas at higher latitudes it is swept into the tail and forms the plasma mantle (Lui, 1992).

The plasmasphere, bounded by the plasmapause, includes the ionosphere and
Figure 1.2: A schematic diagram of the Earth's magnetosphere in the noon-midnight plane (Parks, 1991).
extends to 3 or 4 $R_E$ from the Earth’s surface. The plasma within the plasmapause is the only part of the magnetosphere that is corotating with the Earth (Parks, 1991). The position of the plasmapause shifts according to geomagnetic activity: it generally lies inside $5R_E$ but is found further out when $K_p$ is very low. (Lyons and Williams, 1975a, and references therein). Particle energy increases from $<1$eV deep inside the plasmasphere to a few eV at the plasmapause (Young, 1983).

The radiation belts are populated by particles which are trapped in the Earth’s magnetic field, and are discussed further in the next section. The drift motion of trapped particles around the Earth produces the ring current. The peak current density lies between 4 and 5 $R_E$ or closer during times of increased activity. The geomagnetic tail is produced by the transfer of some of the solar wind’s momentum and energy to the geomagnetic field and has been observed to extend more than 200$R_E$. The plasma sheet is a current carrying region of the tail which lies between the low particle density tail lobes. The plasma sheet and magnetotail are important sources of the trapped radiation further inside the magnetosphere. The plasma sheet itself comprises hot (>1keV) tenuous ($\sim 0.5\text{cm}^{-3}$) plasma which is a combination of plasma originating in the ionosphere and in the solar wind (Young, 1983, and references therein). The earthward edge of the plasma sheet lies at a geocentric distance of about 6-8$R_E$, and closer to the Earth when geomagnetic activity is high. The electron energy density decreases exponentially inward from the inner edge of the plasma sheet, due to the average electron energy decreasing while the density remains fairly constant (Frank, 1971).

Figure 1.3 shows how the convection of plasma sheet particles earthward due to the dawn-dusk electric field is affected by the superposition of the corotational electric field. This produces effectively ‘forbidden’ paths for particles. Electrons of 0.1-1.0 keV are excluded from the dusk region of geosynchronous orbit, and for higher energies, this exclusion region is wider. In geomagnetically quiet times, when the convection electric field is smaller, this forbidden region for plasma sheet particles may be so large that plasma sheet particles are excluded from the geosynchronous region altogether (Thomsen et al., 1996a).
Figure 1.3: (a) Equipotentials of superposed corotational and dawn-dusk electric fields; (b) 3keV ions under the combined influence of the net electric field in (a) and the gradient-curvature drift in a dipole field; (c) same as (b) for 0.1keV electrons; (d) same as (b) but for 1keV electrons (Thomsen et al., 1996a).
The geosynchronous orbit at 6.6\(R_E\) lies in a dynamic region which may, at times of different geomagnetic activity, find itself in the plasmasphere or plasma sheet, or occasionally even outside the magnetopause, as well as seeing ring current particles (Garrett et al., 1981).

1.3 The Radiation Belts

The Earth’s radiation belts were discovered in 1958 by James Van Allen and his co-workers, who included Carl McIlwain and Ernie Ray, and hence are also known as the Van Allen belts. Explorer I data showed that particle intensities could be organized in terms of the magnetic field intensity, and therefore that the particles are controlled by the Earth’s magnetic field. The concept of inner and outer radiation belts originated with the interpretation of data supplied by Explorer IV, which detected high-energy protons (\(\geq 30\)MeV) and electrons (\(\geq 1.6\)MeV). The proton counts peaked in a region around 2\(R_E\), and the electron counts between 3 and 4\(R_E\), apparently defining two radiation zones. Subsequent data have shown that there are belts for each species and energy range, and that the radiation in fact has a continuous distribution stretching from the ionosphere to the magnetopause (Parks, 1991). There is however an electron slot separating the electron inner and outer zones which, at quiet times, has very low fluxes of electrons. The slot lies between 2–4\(R_E\), depending on electron energy, and has been observed to fill with particles during substorm injections. The slot exists because wave-particle interactions result in the slot region electrons being lost to the atmosphere at a high rate (Lyons et al., 1972 and Spjeldvik and Rothwell, 1985).

The radiation belts exist because the interaction of plasma with the geomagnetic field results in a population of trapped charged particles. The belts are toroidal in shape, having a rough symmetry about the Earth’s magnetic dipolar axis. The energetic particle distribution consists of high-energy protons in the inner zone and high energy electrons in the outer zone (figure 1.4). For particles to be permanently trapped in the radiation belts they must be able to drift
completely around the Earth. Anderson (1966) defined a *stable trapping zone* for particles within $8R_E$, which are completely trapped, and a *distant radiation zone* beyond $8R_E$ where, due to the non-dipolar nature of the geomagnetic field, particle drift paths are not complete: particles encounter the magnetopause on the dayside or the tail on the nightside. Rothwell and Lynam (1969) observed that the boundary between the stable and distant zones moved earthward as far as $6R_E$ during geomagnetic activity. Frank (1971) observed the trapping boundary close to geosynchronous altitudes during a magnetic storm. Particle trapping and drift motions are discussed later in this section.

The inner belt extends to latitudes $\pm 40^\circ$, and is closer to the Earth on the western hemisphere (600km) than the eastern hemisphere (1600km). Its inner and outer boundaries lie further away from the Earth nearer the equator. The outer belt stretches beyond latitudes $\pm 55^\circ$ and is also much thicker than the inner belt (Delobeau, 1971). The inner belt fluxes are quite stable and only affected by the most intense magnetic storms, whereas the outer belt is much more variable and flux can change by a few orders of magnitude quite suddenly (e.g. Lezniak et al., 1968) in the course of substorms. Northrop and Teller (1961) suggested that, since the two radiation belt maxima have different particle energy spectra, it is likely that their origins are also different. It is now believed that cosmic ray interactions with atmospheric particles produce neutrons which then decay, depositing protons and electrons in the inner radiation belt. Of the outer belt particles, the lower energy component is believed to be of atmospheric origin, whilst the higher energy particles are believed to come from the solar wind and from plasma sheet particles accelerated inward during substorm injections (Parks, 1991).

### 1.3.1 Trapped Particle Motions

A simplified description of radiation belt particle motions is provided by adiabatic theory (Spjeldvik and Rothwell, 1985). If there is little energy and momentum transferred between particles, or between particles and the electric and magnetic fields that surround them, then it is possible to find parameters which are
Figure 1.4: The approximate spatial distribution of trapped particles in the radiation belts: (a) $>30\text{MeV}$ protons (b) $>1.6\text{MeV}$ electrons (c) 1-5 MeV protons (d) $>40\text{keV}$ electrons \((\text{Hess, 1968})\).
Figure 1.5: A descriptive drawing of the three types of motion of particles trapped in the Earth’s magnetic field (Spjeldvik and Rothwell, 1985).

constants of the particle motion. These parameters are called adiabatic invariants. Trapped radiation belt particles execute three types of motion: they gyrate about the magnetic field lines, bounce between conjugate mirror points and undergo drift motion around the Earth. This is illustrated in figure 1.5. There is an adiabatic invariant associated with each of these three types of motion. This invariant will be conserved for as long as there are no changes on the time scale of that particular motion.

Cyclotron motion

The first type of particle motion, cyclotron motion, is a gyration about the magnetic field direction. A particle with charge $q$ and velocity $v$ in a constant magnetic field $B$ experiences a force $F$ given by:

$$ F = qv \times B $$  \hspace{1cm} (1.7)

The vector cross-product shows that the force $F$ experienced by the particle is perpendicular to both $v$ and $B$. $F$ is the Lorentz force and provides the centripetal force, in the plane perpendicular to $B$, for the gyration with radius $r_L$, the Larmor
radius:

\[
\frac{mv_{\perp}^2}{r_L} = qv_{\perp}B \tag{1.8}
\]

\[
r_L = \frac{mv_{\perp}}{qB} \tag{1.9}
\]

\[
r_L = \frac{v_{\perp}}{\omega} \tag{1.10}
\]

where $\omega$ is the cyclotron frequency.

The Larmor radius is therefore smaller where the magnetic field strength is greater. The effect of the Lorentz force combined with the particle's velocity along the field direction results in the particle following a helical trajectory. The angle between the particle's velocity vector $v$ and the magnetic field $B$ is known as its pitch angle.

The first adiabatic invariant is the magnetic moment $\mu$. If the cyclotron motion of a charged particle is considered equivalent to a current loop, then the magnetic moment is equal to the current multiplied by the area of the loop. The magnetic moment is defined by:

\[
\mu = \frac{p_{\perp}^2}{2m_0B} \tag{1.11}
\]

where $p_{\perp}$ is the particle's momentum perpendicular to the magnetic field, $m_0$ is its rest mass, and $B$ is the magnetic field intensity.

If $\mu$ is conserved then the total magnetic flux enclosed by the cyclotron motion is conserved. Conservation of $\mu$ holds as long as there are no significant changes in the magnetic field on the time scale of the cyclotron period. In considering a particle's motion on a scale larger than the cyclotron motion, it is usual to work in terms of the guiding centre motion rather than that of the actual spiralling particle. The guiding centre system is based on the moving frame of reference in which the motion of the particle is periodic and perpendicular to the magnetic field: the guiding centre for a particle executing cyclotron motion will be the centre of the circular motion.
Bounce motion

The second type of particle motion is a bounce motion, in which the particles are reflected by 'magnetic mirrors' formed by the stronger magnetic fields which exist at higher latitudes. As the particle spirals around a geomagnetic field line, it moves to regions of stronger magnetic field closer to the poles. Its motion perpendicular to the field is accelerated (equation 1.8), and since its total energy must remain constant in a static field, its parallel motion diminishes. At the mirror point, the parallel component becomes zero and the particle reverses and spirals back along the field line. This bounce motion continues between conjugate mirror points in either hemisphere. It is the bounce motion that leads to the trapping of particles in the geomagnetic field to form the radiation belts. Since, where \( \alpha \) is the particle's pitch angle:

\[
p_{\perp} = p \sin \alpha
\]  

conservation of \( \mu \) implies, from equation 1.11 that:

\[
\frac{\sin^2 \alpha_0}{B_0} = \frac{\sin^2 \alpha}{B} = \frac{1}{B_m} = constant
\]  

where \( \alpha_0 \) and \( B_0 \) are the equatorial pitch angle and field, and \( B_m \) is the mirror field, when \( \alpha \) is 90°. So a given particle always mirrors at a point with the same field strength. Note that if a particle's equatorial pitch angle is 90° it 'mirrors' at the magnetic equator.

There is a limit on the pitch angle a particle has at any point which will keep it mirroring far enough above the Earth's surface to avoid being lost to the atmosphere. This limiting pitch angle defines the loss cone, as illustrated in figure 1.6. Particles whose pitch angles are inside the loss cone will be lost before they can bounce back because they will collide with atmospheric particles before reaching their mirror points (e.g. Roederer, 1970).

The second adiabatic invariant, \( J \), is associated with the component of the particle's motion parallel to the magnetic field:

\[
J = \oint m v_{\parallel} ds
\]  

(1.14)
where $m v_\parallel$ is the particle's momentum in the direction of the field and $ds$ is an elemental path length along one complete bounce path. Conservation of $J$ holds as long as there are no significant changes in the magnetic field on the time scale of the bounce period.

**Drift motion**

The third type of particle motion is an azimuthal drift around the Earth. There are five types of drift motion:

- Gradient drift

Gradient drift occurs because the geomagnetic field is stronger on the earthward side of the particles’ gyration about the field lines. The radius of curvature of the gyration is therefore smaller on the earthward side (equation 1.9). Therefore per gyration, the particle moves a greater distance on the far side of its cycle than on the earthward side, which means that the particle experiences a net motion around the Earth. Because protons and electrons gyrate in opposite senses, this drift is charge dependent. Electrons drift eastward and protons westward.
• Curvature drift
Curvature drift arises because the geomagnetic field lines are curved between the poles, and this results in a centrifugal force on the particles in the direction radially outward from the centre of curvature. The resulting drift motion must be perpendicular to both the magnetic field and centrifugal force, i.e. azimuthal.

• Electric field drift
A convection electric field exists across the magnetosphere in the dawn dusk direction, induced by the solar wind flow past the Earth’s magnetic field. This is combined with a radial electric field produced by the Earth’s rotating dipole. The electric field produces a force which causes a drift. The drift motion is independent of a particle’s charge (protons and electrons drift together) and is independent of a particle’s mass and energy.

• Polarization drift
If the electric field is varying slowly with time, then the drift velocity also varies in time. A frame of reference moving with the drift velocity experiences acceleration and the associated force causes an additional drift. The polarization drift is in the direction of the electric field, and depends on both particle mass and charge.

• Gravitational drift
Gravitational drift results from the component of gravitational force perpendicular to the magnetic field. This drift is mass-dependent and heavier particles drift faster. It also depends on charge. The effect of gravitational drift is small compared to that of the other drifts.

For electrons with energies >10keV, as are detected by the SEM-2, the drift motion is largely a combination of gradient and curvature drifts, (Spjeldvoik and Rothwell, 1985). Electric field drift only becomes important for low energy particles (≤1keV) and the other drifts are negligible. The direction of both the
gradient and curvature drift components depends on the charge of the particle, electrons drift eastwards and protons westwards, forming the ring current.

The third adiabatic invariant, $\Phi$, is associated with drift motion. $\Phi$ is the magnetic flux enclosed by the drift shell of a particle. $\Phi$ is conserved as long as changes in the magnetic field occur on time scales longer than the drift period. Drift velocities are longitude dependent: beyond $5R_E$, since the gradient of the field is larger on the nightside, the gradient drift velocity will be greater, so the trapped particle will spend less time on the nightside than the dayside (Roederer, 1967).

**Violation of the adiabatic invariants**

The adiabatic invariants are violated when electric or magnetic field variations occur which have frequencies close to one of the three motions. Since the frequency of the drift motion is much lower than that of the bounce motion, which in turn is much lower than that of the cyclotron motion, it is possible for one of the invariants to be violated while the others are conserved. Northrop and Teller (1961) state that, in the absence of scattering, it is the violation of these adiabatic laws that is responsible for the loss of particles from the radiation belts. Although in general the first invariant will be conserved due to the short time scale of the spiralling motion, magnetic fluctuations on time scales comparable with the particle bounce motion and longitudinal drift motion will result in the second and third invariants not being conserved, and particles may, as a result, diffuse away or be lost to the atmosphere.

The transport of particles in the radiation belts can occur through diffusion processes, which result from the violation of one or more of the adiabatic invariants. The invariants may be violated through wave-particle interactions, through magnetic or electric field fluctuations, or through atmospheric collisions. Pitch angle diffusion can occur when field fluctuations violate the first adiabatic invariant, or the second, or both together. During pitch angle diffusion energy exchange takes place between the wave and particle and the particle’s direction changes. The result of pitch angle diffusion is to move the particle’s mirror point
along the field line. In doing so particles may enter the loss cone.

Radial diffusion allows particles to move across field lines in the radial direction: the particles move across drift shells. Pure radial diffusion occurs when field fluctuations on the time scale of the drift period violate the third adiabatic invariant but leave the first two intact. Because particles which diffuse radially inward will find themselves in regions of stronger magnetic field, then if the first adiabatic invariant is conserved, their kinetic energy will increase (equation 1.11).

Both radial and pitch angle diffusion can lead to the loss of particles from the radiation belts. In changing the mirror points of particles, pitch angle diffusion is the dominant mechanism by which particles are lost to the atmosphere. Radial diffusion may move the particle onto a lower L-shell where it may be lost to the atmosphere. In the Earth’s field, pitch angle diffusion is always accompanied by radial diffusion because if the particle’s mirror point changes, it finds itself on a different drift shell (Roederer, 1970).

Energy diffusion of particles can occur when particles are energized by or lose energy during interactions with waves. Energy changes can also result from collisions with particles in the exosphere (Spjeldvik and Rothwell, 1985).

Observations of electrons generally give a better idea of the structure of the outer magnetosphere because a proton of the same energy will be much more likely to suffer from violations of the adiabatic invariants (West, 1979).

The L Parameter

McIlwain (1961) introduced the L parameter which is a function of the second adiabatic invariant for static fields I (see next section) and the magnetic field B. The L parameter was found to organize measurements adequately along lines of force and was conceptually straightforward: L is the equatorial radius of a magnetic shell in R_E in a dipole field. Stone (1963) estimated that the L value is equal to the equatorial radius of a shell in Earth radii to an accuracy of a few percent. Points where B and L are constant form a ring in each hemisphere: the rings are connected by a shell such that a particle mirroring at this particular B and L drifts on this shell. Particles with different mirror points will not drift on
the same shell, and this is termed shell splitting. Frank (1965a) found that the \((B,L)\) coordinates failed to organize electron data adequately beyond \(6R_E\), and explained this as due to the degree of distortion of the geomagnetic field by the solar wind.

1.3.2 Drift Shell splitting

A particle's bounce motion combined with its drift motion traces out a drift shell. Because the geomagnetic field is not dipolar in the outer zone, there are magnetic field gradients in the azimuthal direction which cause particles to drift radially as they move round the Earth. The radial drift depends on pitch angle. If the first and second adiabatic invariants are conserved, particles which originate at a certain point will return to that point, but will drift on different shells according to their equatorial pitch angle (Northrup and Teller, 1960): this is called shell splitting. In the case of perfect azimuthal symmetry, two particles with different mirror points starting from the same origin will drift on shells which coincide, i.e. they share all the same lines of force: this is termed shell degeneracy. In an azimuthally asymmetric field, the shells do not coincide. Drift shell splitting becomes important beyond about \(5R_E\) since it is at larger altitudes that the dipole approximation becomes increasingly inaccurate. At low altitudes, a dipole model of the Earth's magnetic field with the geomagnetic axis inclined to the Earth's spin axis at 11.5° is a reasonable approximation. As altitude increases and the deformation of the magnetosphere by the solar wind becomes more pronounced, this approximation becomes invalid because there is extensive asymmetry between dayside and nightside field structures. Roederer (1967) predicted that in the outer zone equatorial pitch angles would be field-aligned on the nightside but perpendicular on the dayside as a result of shell splitting.

Drift shell splitting can be explained using the integral invariant \(I\), which is also referred to as the second adiabatic invariant for static fields. From equation 1.14, if the mirror points are \(l_1\) and \(l_2\), then:

\[
J = 2 \int_{l_1}^{l_2} m v_\parallel dl
\]
The integral invariant is defined by:

\[ I = \frac{J}{2mv} \]  

(1.16)

where \( I \) is given by:

\[ I = \int_{l_1}^{l_2} \sqrt{1 - B/B_m} \, dl \]  

(1.17)

\( I \) gives the length of field line between mirror points and is usually expressed in \( R_E \). Equation 1.17 implies that the total length of the particle trajectory is conserved. This holds in a static field in the absence of external forces. Figure 1.7 shows contours of constant \( B, I \) and \( L \). In order to satisfy both the first and second adiabatic invariants, a particle which has mirror points \( P \) and \( P^* \) with \( I = I_0 \) on one side of the Earth can only mirror at \( A_2 \) and \( A_2^* \) after drifting to the other side of the Earth. Since \( I \) must be conserved, mirror points \( A_1 \) and \( A_1^* \) (with \( I_1 < I_0 \)) and \( A_3 \) and \( A_3^* \) (with \( I_3 > I_0 \)) are excluded. Recall that by conservation of the first adiabatic invariant the mirror point field is conserved (equation 1.13), which means mirror points \( Q \) and \( Q^* \) are also excluded (O’Brien, 1963).

![Diagram showing contours of constant B, I, and L](image)

Figure 1.7: Contours of \( B, I \) and \( L \) (O’Brien, 1963).

Figure 1.8 shows shell splitting as two particles drift from longitude \( \phi \) to \( \phi + 180^\circ \). A particle mirroring at a field value \( B_m \) will also conserve \( I \) as it drifts;
a second particle starting on the same field line but mirroring at a lower field value $B'_m < B_m$ will also have a lower value of the integral invariant $I' < I$. If there is perfect azimuthal symmetry, these two shells will in fact coincide.

Figure 1.8: Split shells (Roederer, 1970).

Figure 1.9 shows particles drifting from the same initial field line at noon round to midnight. This figure was produced by Roederer (1970) using the Mead-Williams quiet time magnetic field model. In drifting from noon to midnight, shell splitting is directed radially inward. The figure shows how in drifting from noon to midnight the equatorial pitch angle decreases. Recall from equation 1.13 how conservation of the first adiabatic invariant means that the magnetic field intensity at the mirror point of a particle remains constant. Since the equatorial field strength decreases going from noon to midnight, then the equatorial pitch angle also decreases. Figure 1.10 shows a similar figure for particles drifting from the same initial field line at midnight round to noon. In this case shell splitting is directed radially outward and the equatorial pitch angle increases.

Roederer (1967) stated that one manifestation of shell splitting would be that equatorial pitch angles would be field-aligned on the nightside but perpendicular
on the dayside. Konradi (1965) had observed from Explorer 12 data that outer zone dayside trapped particle fluxes are peaked perpendicular to the magnetic field. Serlemitsos (1966) first identified field-aligned trapped electrons on the nightside from Explorer 14 data, and also saw dayside electron pitch angle distributions peaked perpendicular to the field direction. Haskell (1969) noted from Explorer 33 data that at the far edge of the outer belt, particle fluxes peaked along the field on the nightside and perpendicular to the field on the dayside. Pfitzer et al. (1969) used data from the geosynchronous satellite ATS 1 and the elliptical satellite OGO 3 to confirm the existence of shell splitting. Their observations showed the pitch angle distribution to be near-isotropic at noon but cigar-shaped at midnight. Using data from the geostationary satellite ATS 6, Kaye et al. (1978) observed 40° pitch angle flux to dominate 90° pitch angle flux on the nightside, and 90° flux to dominate 40° flux on the dayside.

Roederer suggested that pancaked dayside and cigar-shaped nightside distributions would result from the different drift paths followed by particles with different equatorial pitch angles. In the outer magnetosphere, only particles with certain pitch angles can survive drifting around the Earth: others are lost to the tail or magnetopause.

1.3.3 Quasi-trapping

Further modification of the pitch angle distribution of trapped outer zone particles arises from quasi- or pseudo-trapping. Anderson (1966) defined a stable trapping zone for particles within 8R_E, which are completely trapped, and a distant radiation zone beyond 8R_E where, due to the non-dipolar nature of the geomagnetic field, particle drift paths are not complete. Quasi-trapping was described by Roederer (1967, 1970): figure 1.11 shows the location of quasi-trapping regions in the magnetosphere. Since contours of constant magnetic field intensity from beyond 7R_E on the nightside magnetic equator do not close on the dayside, particle drift paths are not complete: the outermost closed contour is termed the limit of stable trapping. Similarly particles drifting from the dayside may be lost to the tail. Other authors report different distances for the trapping
Figure 1.9: Computed shell splitting for particles starting on common field lines in the noon meridian. Dots represent the particles’ mirror points, curves show the position of mirror points for constant equatorial pitch angle $\alpha_0$ (Roederer, 1967).

Figure 1.10: Computed shell splitting for particles starting on common field lines in the midnight meridian. Dots represent the particles’ mirror points, curves show the position of mirror points for constant equatorial pitch angle $\alpha_0$ (Roederer, 1967).
boundaries. Konradi (1965) placed the outer boundary of the trapping region at typically 10\(R_E\). Haskell (1969) questioned from his observations of pitch-angle distributions whether trapping was really impossible beyond 8-10\(R_E\), and Sermitsos (1966) reported trapped 100keV electrons near magnetic midnight and at low magnetic latitudes extending beyond 12\(R_E\). Observations of the trapping boundary will depend on geomagnetic activity and the extent to which the magnetosphere is compressed.

![Image of magnetic field lines and quasi-trapping regions](image)

Figure 1.11: Location of the quasi-trapping regions in the magnetosphere (Roederer, 1967).

The limit of stable trapping can be described by the *drift loss cone*: there is a limiting pitch angle for which particles mirror inside the quasi-trapping regions. Figure 1.12 illustrates how on the dayside, the drift loss cone is directed along the magnetic field direction (and includes the bounce loss cone) whereas on the nightside the drift loss cone is directed perpendicular to the magnetic field direction (Roederer, 1970).

Greenspan et al. (1985) observed from ISEE 1 data that 90° electrons predominated in the morning whilst 90° protons predominated in the afternoon.
To explain the morning-afternoon pitch angle asymmetry they introduced the concept of the magnetopause as an absorber. Figure 1.13 (Sibeck et al., 1987) illustrates how drift shell splitting can lead to the loss of 90° particles from the outer zone at the dayside magnetopause. Because of the directions in which the particles drift, electrons eastward and protons westward, 90° electrons will be lost at the pre-noon magnetopause and 90° protons will be lost at the post-noon magnetopause. This process is called *magnetopause shadowing*. As a result of this there should be a depletion of 90° electrons in the post-noon region in the outer zone. In times of high geomagnetic activity the magnetopause can move earthward and so this depletion would be observed closer to the Earth. Magnetopause shadowing cannot be responsible for a morning-afternoon pitch angle asymmetry on drift paths which do not intercept the magnetopause.

*Sibeck et al.* (1987) used a drift shell splitting model to explain unusual pitch angle distributions observed during storms and substorms. Shell splitting leads to particles injected on the nightside drifting to different dayside radial positions depending on their pitch angle: 90° particles will end up furthest away from the...
Figure 1.13: Qualitative sketch of magnetospheric drift paths of energetic particles with 30° and 90° pitch angles in the equatorial plane. Particles with 90° pitch angles follow contours of constant equatorial magnetic field strength, but those with lower pitch angles follow more circular paths. Particles with 90° pitch angles drift radially outward as they move to the dayside magnetosphere, leaving lower pitch angle particles on inner magnetospheric drift paths. (Sibeck et al., 1987)
Earth. Thus the outer edge of the injection signature will show a surplus and the inner edge a deficit of $90^\circ$ particles once drifted to the dayside.

1.4 Magnetospheric Substorms

The magnetospheric substorm is so called because the main (second) phase of the magnetic storm consists of many of them. Magnetic storms are discussed in section 1.4.1, and substorms, which are smaller and more common than storms, are discussed in section 1.4.2

1.4.1 Magnetic Storms

The solar wind has varying speed and when fast solar wind catches up with slow solar wind, the plasma and the magnetic field carried with it are compressed. This, combined with a strong, southward IMF, is a major factor in storm occurrence (Fairfield, 1992). A magnetic storm is a prolonged disturbance of the magnetosphere caused by solar wind variations. Until recently, the primary cause of storms was believed to be solar flares. Gosling (1993) showed that during high activity, fast coronal mass ejections (CMEs) were in fact responsible and that flares were not a factor in producing large magnetic storms. A CME is the sporadic ejection of a large mass of coronal particles by the Sun into interplanetary space, which can produce an interplanetary shock depending on the speed of the ejection. Whilst the CME events produce particles comparable to the solar wind in terms of elemental abundances and ionization states, flare-associated events are rich in heavy elements which are highly ionized (Reames, 1996). The trigger for the CME is not understood. The shock wave disturbance carries the conditions conducive to magnetic storms: high flow speeds and strong magnetic fields, often with a strong southward component.

The resultant storm is identified in three phases, two of which energize the magnetosphere and the third of which is a recovery phase. Following compression of the magnetosphere by a solar wind shock front, the first phase begins with a sudden commencement which is observed at mid-latitudes as a sharp rise (taking
a few minutes) in the horizontal component of the magnetic field. This phase lasts a few hours during which time the magnetosphere remains compressed. The second phase correlates with the southward turning of the IMF, allowing magnetic reconnection at the magnetopause and the transference of energy from the solar wind to the magnetosphere. The ring current is enhanced, causing large decreases in the horizontal component of the magnetic field over most of the Earth, and the magnetosphere inflates. This phase typically last several hours. During the last phase, the recovery, the ring current diminishes through loss processes and is not replenished. Over a day or so the magnetic field returns to the pre-storm value.

1.4.2 Substorms

The substorm is the process responsible for dissipating the energy which the magnetosphere has extracted from the solar wind. Substorms usually occur at least daily, around local midnight. From the Earth they are identified by their auroral signatures and negative bays in the magnetic field. At geostationary orbit the substorm is observed as a sharp increase in particle flux called an injection. By this process each substorm contributes particles to the ring current. In fact, substorm injections are a major source for populating the outer radiation belt and outer ring current (Young, 1983). Greenspan et al. (1985) showed that the injection signature could not be a stationary spatial structure that the satellite passed through because the elliptical satellite ISEE 1 saw high energy particles arriving first on both inbound and outbound passages.

The Near-Earth Neutral Line (NENL) model of substorms is probably the most widely accepted model (Russell and McPherron, 1973). When the IMF turns southward dayside reconnection occurs and magnetic flux is swept over the poles and added to the tail lobes. The nightside plasma sheet and tail current thin, and move earthward. Reconnection occurs in the tail and as the field lines are cut a plasmoid is ejected tailward. Although the convection electric field formed from the action of the solar wind on the magnetopause moves particles earthward from the tail, this is not sufficient to cause injections which are
impulsive processes driven by inductive electric fields due to magnetic field dipolarizations (e.g. Mauk and Meng, 1987). Another popular type of model is based on the disruption of the tail current in the thin current sheet resulting in the diversion of currents into the ionosphere, a phenomenon called the substorm current wedge. The substorm injection boundary model has also been suggested, in which particles tailward of a boundary are accelerated but those earthward of the boundary are not disturbed. Mauk and McIlwain (1974) found from ATS-5 observations of low energy electrons that the injection boundary had a spiral structure and a scale dependent on $K_p$, such that injections were observed earlier at high $K_p$. Lopez et al. (1990) explain the spiral shape and $K_p$ dependence by means of the magnetic perturbation due to the cross-tail current: the boundary is at a critical ratio of the perturbation magnetic field to the dipole field. Moore et al. (1981) suggested the existence of an injection front which propagates earthward, the energized plasma behind it. Since spectrograms show how some pre-existing parts of the particle distribution are left undisturbed by adjacent (in energy), newly injected plasma, this points to a sharp spatial boundary between the new and pre-existing plasma (e.g. Mauk and Meng, 1987). The injection front was interpreted as being due to a travelling compression wave from the tail. The original injection boundary model made no statement about how the particles are energized. The injection boundary and propagating front models of the substorm have been combined and put in the context of the NENL model, by assuming the compression wave is launched by reconnection in the tail (Baker and Pulkkinen, 1991). The results of Friedel et al. (1996) confirm that the injection region has sharp boundaries in magnetic local time, but may extend over several $R_E$ radially.

Two theories exist concerning the mechanism of substorm energy release. The loading and unloading model has solar wind energy stored in the magnetosphere up to a point when it is all released, then the build-up starts again. In the directly driven model the magnetosphere responds to solar wind changes: there is some evidence to support this model in the strong relation of the IMF $B_z$ to the $AE$ index, although there is nothing which could account for the very sudden
substorm onset as a directly driven feature (Fairfield, 1992).

Wing and Sibeck (1997) correlated $B_z$ and solar wind dynamic pressure with the geosynchronous magnetic field. Increases in dynamic pressure enhance the field on the dayside but decrease it on the nightside: this is due to enhanced magnetopause and cross-tail currents respectively. The effect of $B_z$ was strongest when southward and reduced when northward. Thus the solar wind dynamic pressure and $B_z$ strongly control the geosynchronous field magnitude and direction.

An energy coupling function $\epsilon$ describing the degree of coupling between the solar wind and magnetosphere has been defined in terms of solar wind speed $v$, the IMF intensity $B$, and its orientation $\theta$ (polar angle in GSM coordinates):

$$\epsilon = vB^2\sin^4\left(\frac{\theta}{2}\right)l_0^2$$

where $l_0$ is equal to 7RE (Akasofu, 1979 and Akasofu, 1980). A good correlation exists between $\epsilon$ and the total energy generated in the magnetosphere during magnetic storms, and with the $AE$ index. When $\epsilon$ reaches a critical value substorm onset occurs. If $\epsilon$ is much larger, a huge enhancement of the ring current occurs. Subsequent consideration of $\epsilon$ shows that at a sudden storm commencement, $\epsilon$ and $AE$ increase together, as does the energy stored in the tail. Akasofu suggests that with real-time solar wind parameters available, $\epsilon$ may provide a means for predicting substorms and ring current enhancement.

A double peak in the response time of the $AL$ index to the solar wind input parameter $vB_z$ ($B_z$ equal to $-B_z$ if $B_z$ is negative, or zero if $B_z$ is positive) showed shorter response times (~20 minutes) occurred at high magnetospheric activity and longer response times (~60 minutes) at moderate activity (Bargatze et al., 1985). The interpretation was that 20 minutes may be the typical response time for a solar wind coupling mechanism, and 60 minutes the typical time for the release of stored magnetotail energy. This would then imply that both driven and unloading aspects are important to substorm occurrence.

The 'trigger' for substorms has long been sought. Bruening and Tanskanen (1987) observed a substorm apparently triggered by a short northward excursion of the IMF $B_z$ after a period of southward $B_z$ during which a taillike field.
developed and geosynchronous particle flux decreased.

Substorms are commonly described in three phases: the growth phase, the expansion phase and the recovery phase.

**Growth phase**

The IMF is typically northward during magnetically quiet magnetospheric conditions. During the growth phase, the southward turning of the IMF and the resultant reconnection at the dayside magnetopause lead to the storage of energy in the geomagnetic tail, which is later released suddenly in the substorm expansion phase (Fairfield, 1992, and references therein). Energy storage continues as long as the IMF remains southward or until a substorm occurs, and if the IMF turns northward then energy storage levels off but resumes when the IMF is southward again (Baker et al., 1982a).

Magnetic reconnection provides a mechanism for the transfer of mass, energy and momentum from the solar wind to the magnetosphere. Dungey (1961) first proposed reconnection at the dayside magnetopause and in the tail. Through reconnection, energy stored in the geotail is converted to the kinetic energy of particles. Support for the theory of reconnection lies in the increased magnetospheric activity observed when the IMF turns southward. Reconnection should serve as a mechanism for accelerating plasma away from the site of reconnection, termed the X-line or reconnection line, and this has been observed on the dayside by Paschmann et al. (1979). A simple schematic of reconnection is shown in figure 1.14.

In the ionosphere, the growth phase is marked by enlargement of the polar cap and enhancement of ionospheric electrojets. During the growth phase, 1–1.5 hours before the expansion phase onset, the geomagnetic field becomes more taillike. The cross-tail current moves earthward, and trapped particles follow contours of constant magnetic field closer to Earth, which causes flux *dropouts* to be observed at geostationary orbit. These flux decreases are often seen prior to injection events and were first observed by Erikson and Winckler (1973). They typically last 40-60 minutes, before the sudden increase in flux at the expansion
Figure 1.14: Schematic representation of the reconnection model. A southward directed interplanetary magnetic field is shown connected to magnetospheric field lines (Dungey, 1968.)
phase onset (Erikson et al., 1979). Sauvaud and Winckler (1980) noted that these particle dropouts observed out of the magnetic equatorial plane by ATS 6 (having a magnetic latitude of 12°) appeared greater than when observed in the equatorial plane by ATS 1. They also observed that the dropouts are greater for higher energy particles.

Expansion phase

At expansion phase onset, negative magnetic bays are seen at the auroral zone and positive magnetic bays at lower latitudes (Nagai, 1983). The expansion phase lasts 0.5 to 1 hour. The expansion phase onset is characterised by a surge of tail plasma earthward: this injection consists of plasma sheet particles which have been accelerated by an induced electric field produced by the collapse of the magnetic field (Winckler, 1970). The trigger mechanism for this release is not fully understood. At the same time trapped particles, which were drawn earthward during the growth phase, move tailward on lines of constant magnetic field. The magnetosphere becomes more dipole-like. These two mechanisms lead to an increase at geostationary orbit of both the high energy trapped particles and lower energy particles injected from the tail. Sauvaud and Winckler (1980) saw in ATS 6 data that lower energy particles (32-51keV) are observed at greater flux after the dropout than before it, and therefore must have been accelerated, whereas higher energy particles simply returned to the levels of flux they had before the dropout.

The injection site is believed to be near local midnight and close to geostationary orbit, since all particle energies are observed to arrive simultaneously when injections are observed near local midnight i.e. before much dispersion can have taken place. At later local times, the injections are observed to have undergone energy dispersion, with the lower energy particles, which have lower drift speeds, arriving later (McIlwain and Whipple, 1986). McPherron et al. (1973) and later Nagai (1987) proposed models in which the cross-tail currents convert to field-aligned currents which connect to the polar ionosphere. This can happen because the field lines occupied by geostationary orbit map down to the auroral
zone. McLwain (1974) used ATS-5 data to show that the injections originate at an injection boundary which maps down to the equatorward edge of the auroral oval.

Recovery phase

The recovery phase lasts 1 to 2 hours. In the near tail, drift echoes of injected particles are observed. In the mid-tail, the plasma sheet thickens and in the far tail, electrons are seen to stream tailward. Precipitated particles have been observed in the morning sector during the recovery phase, which may be due to the earthward part of the tail flipping back after the ejection of a plasmoid.

The recovery phase has often been regarded as a passive phase, when the magnetosphere slowly returns to its quiet state following an active period. Some authors, e.g. Kopányi and Korth (1995), have argued that the recovery phase itself contains structural changes in the magnetic field configuration. Kopányi and Korth observed electron and ion dropouts lasting 10-30 minutes in the local morning sector during the recovery phase, in the course of multiple substorm events, at geosynchronous orbit. These dropouts were distinct from magnetopause crossings (the satellite GEOS-2 remained well inside the magnetosphere during the observation) and from the midnight sector growth phase dropouts caused by the earthward motion of the radiation belt. During these dropouts, the magnetic field intensity increased and the tailward component turned sunward and then back again. Their explanation of the dropout is that during multiple substorms the plasma sheet moves earthward and thins to such an extent that a geostationary spacecraft finds itself in the lobes. This dynamic behaviour is seen in the recovery phase. Kopányi and Korth also noted that the recovery phase of one substorm can overlap the growth phase of the next substorm during a multiple event.

In summary, figure 1.15 is a schematic representation of the behaviour of geosynchronous flux, the $AE$ index and the geosynchronous $H$ component of magnetic induction during a substorm (Sauvaud and Winckler, 1980). The flux decrease occurs in the growth phase accompanying the stretching of the tail, and
Figure 1.15: Schematic representation of the simultaneous variations of the H component of the magnetic induction at 6.6R_E (antiparallel to the dipole axis), the AE index and the energetic particle flux at 6.6R_E during a substorm preceded by a quiet period (Sauvaud and Winckler, 1980)
the injection and field dipolarization are seen at onset.

The abrupt increases in particle flux observed at geosynchronous orbit at substorm onset have been shown to have a one to one correlation with substorms (Deforest and McIlwain, 1971). Injections are the geosynchronous manifestation of the substorm. Of the injected particles, those with small pitch angles (in the loss cone) will be precipitated into the atmosphere, whilst those with large pitch angles become trapped, intensifying the ring current. Multi-satellite observations of the injection region (e.g. Reeves et al., 1990) serve to show how particles are injected inward and subsequently undergo drift motion around the Earth.

Arnoldy and Chan (1969) observed the injection of 50–150keV electrons with ATS-1 at geostationary orbit (the next channel, 150–500keV, did not always see the injection) near midnight and their subsequent drift on closed field lines around the Earth. They noted that at these energies the group of electrons injected can rarely be observed for more than one complete drift, the electrons either dispersing so that the group is no longer identifiable, or else lost from either the energy range of the detector or from synchronous altitudes. High energy (0.4–2 MeV) electrons were observed by Lanzerotti et al. (1967) to drift round the Earth a few times before fading, and their observations agreed well will calculated drift velocities.

Since there is no disturbance in the midtail in the growth phase, and none in the midtail or distant tail in the growth or expansion phases, this suggests that the substorm must originate in the near tail and effects propagate down the tail as the storm progresses (Lui (1991) and references therein).

The Aurora

A large amount of the energy released in a substorm is dissipated during aurorae, when energetic particles from the plasma sheet precipitate into the high latitude atmosphere. The incident particles ionize and excite atmospheric particles which, on recombining or returning to their ground states, emit visible light in the characteristic auroral colours, most commonly red and green. Some of the incident particles' energy is also converted to bremsstrahlung X-radiation. The
precipitation of particles into the ionosphere increases ionospheric conductivity and eastward and westward ionospheric currents – electrojets – flow to the order of millions of Amperes. These currents can significantly affect the magnetic field measured at the Earth's surface.

The auroral substorm begins with quiet arcs drifting equatorward. The most equatorward arc brightens suddenly in the pre-midnight sector, which marks the onset of the auroral substorm. During the expansion phase the brightening expands poleward and westward into a dynamic auroral display spanning the midnight sector, and at the westward edge a westward-travelling surge is produced and an auroral bulge forms at the boundary between dipole and tail field lines. When the expansion poleward ends the expansion phase is over. The recovery phase of the auroral substorm lasts about 90 minutes, during which the disturbed region dims and the westward-travelling surge diminishes, and quiet arcs are re-established.
Chapter 2

The Geostationary Environment

Meteosat P2 is in a geostationary orbit, which is a special case of geosynchronous orbit. A satellite in geosynchronous orbit has an orbital period the same as the Earth's rotation period. In geostationary orbit, the satellite's inclination to the Earth's geographic equator is zero, so the spacecraft remains over the same point on the Earth's surface throughout its orbit. Geostationary orbit is therefore widely used for communications, weather prediction and surveillance applications. Geostationary orbit lies at 6.6\,R_E from the Earth's centre, at an altitude of \( \sim 3.6 \times 10^4 \, \text{km} \), on the far side of the outer radiation belt and beyond its maximum particle intensity.

Geostationary orbit is positioned at the outer edge of the trapped radiation zone and at the inner edge of the magnetotail plasma sheet. In quiet times geostationary orbit is generally earthward of the plasma sheet, but in active times it is in the plasma sheet and observes fresh injections of particles during substorms. This therefore presents a highly variable environment since the plasma sheet is highly dependent on geomagnetic activity. In addition the response of the magnetosphere to activity means that at active times when the magnetosphere is strongly compressed a geostationary satellite may encounter the dayside magnetopause \((e.g. \ Korth \ et \ al., \ 1982)\) and at quiet times it may cross the plasmapause \((e.g. \ McComas \ et \ al., \ 1993)\).
2.1 Observations at Geostationary Orbit

In this section previous observations of the geostationary environment are discussed, including recent observations from the Los Alamos National Laboratory (LANL). The SEM-2 has the same detector unit as the low energy electron detector, the Lo-E, on the LANL Charged Particle Analyser (CPA) which was flown on various geosynchronous satellites as part the Defense Support Program of satellites (Aiello et al., 1975). The CPA measures electrons of energies 30keV to 2MeV in 12 energy channels and protons 75keV-200MeV in 26 energy channels. LANL have therefore obtained a large amount of data on electrons in the SEM-2 energy range at geosynchronous orbit since the launch of the first CPAs in 1976 (e.g. Baker et al., 1981a).

Other LANL detectors on synchronous satellites, namely the MPA (Magnetospheric Plasma Analyser) which measures ions and electrons from 1eV to 40 keV per unit charge, and the SOPA (Synchronous Orbit Plasma Analyser) which measures 50keV-1.5MeV protons, electrons and helium ions, have added greatly to the knowledge of this region. The MPA has six detectors covering polar angles 25° to 255°, and the spin of the satellite is divided into 24 azimuthal sectors of 15° each. The spin axes of the satellites point to the centre of the Earth, and a complete distribution is obtained every 10s spin, ions and electrons alternately. An energy sweep through 40 logarithmically spaced energy channels from ~40 keV/charge down to ~1eV/charge is performed within each azimuthal sector. The SOPA has three detectors at polar angles 30°, 60° and 120° to the satellite spin axis, each with a field of view of about 11°, and measurements are made in 64 azimuthal sectors every 10s spin. There is on board magnetic field measurement but the field orientation has been derived from the shape of the MPA particle distributions (Thomsen et al., 1996). Table 2.1 shows the electron energy levels of three LANL detectors and table 2.2 shows which instruments are on which satellites. Several of these geosynchronous satellites have operated at the same time over a range of longitudes, providing continuous data from 1976 to the present. Results from these and other instruments are reviewed here.

Numerous other satellites have contributed to the geosynchronous data set.
### Table 2.1: LANL detector electron energy levels

<table>
<thead>
<tr>
<th>CPA LoE (keV)</th>
<th>CPA HiE (MeV)</th>
<th>SOPA (keV)</th>
<th>SOPA (MeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-300</td>
<td>0.2-2</td>
<td>50-75</td>
<td>0.75-1.1</td>
</tr>
<tr>
<td>45-300</td>
<td>0.3-2</td>
<td>75-105</td>
<td>1.1-1.5</td>
</tr>
<tr>
<td>65-300</td>
<td>0.4-2</td>
<td>105-150</td>
<td>&gt;1.5</td>
</tr>
<tr>
<td>95-300</td>
<td>0.6-2</td>
<td>150-225</td>
<td>0.7-1.8</td>
</tr>
<tr>
<td>140-300</td>
<td>0.9-2</td>
<td>225-315</td>
<td>1.8-3.5</td>
</tr>
<tr>
<td>200-300</td>
<td>1.4-2</td>
<td>315-500</td>
<td>3.5-6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500-750</td>
<td>6.0-7.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.8-10.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.8-26</td>
</tr>
</tbody>
</table>

### Table 2.2: LANL satellites and detectors

<table>
<thead>
<tr>
<th>CPA</th>
<th>MPA and SOPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976-059</td>
<td>1989-046</td>
</tr>
<tr>
<td>1977-007</td>
<td>1990-095</td>
</tr>
<tr>
<td>1979-053</td>
<td>1991-080</td>
</tr>
<tr>
<td>1981-025</td>
<td>1994-084</td>
</tr>
<tr>
<td>1982-019</td>
<td>LANL-97A</td>
</tr>
<tr>
<td>1984-037</td>
<td></td>
</tr>
<tr>
<td>1984-129</td>
<td></td>
</tr>
<tr>
<td>1987-097</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: LANL satellites and detectors
The SMS (Synchronous Meteorological Satellite) and GOES (Geostationary Operational Environmental Satellite) series of satellites carry SEM systems which include the EPS (Energetic Particle Sensor) which has one high energy integral electron channel and several proton channels, and magnetometers. ATS (Applications Technology Satellite) 1, 5 and 6 also provided both measurements of energetic protons and electrons, and magnetic field data.

Other valuable observations of the outer radiation belt have come from elliptical satellites, such as the recent CRRES (Combined Release and Radiation Effects Satellite) mission, which occupied a geostationary transfer orbit.

The following review of results is divided into sections, for electron energies below, within and above the SEM-2's range.

Low energies

Figure 2.1 shows summarized results from MPAs on three geosynchronous satellites for six weeks of data; even though the geosynchronous environment is highly variable, there are clear local time preferences for observing each of the various regions (McComas et al., 1993). McComas et al. found that at quiet times ($K_p < 2$) the satellites' orbits could lie completely inside the plasmasphere, whereas during more active times the plasmasphere could only be seen in the afternoon to evening. In addition, the dusk side plasmaspheric bulge was seen to vary significantly with activity. Seven distinct regions are observed: Firstly, the cool, dense plasmasphere (Psp on figure 2.1, and 13.1% of the data) which is mostly observed at the dusk bulge, but could be seen at all LTs when the magnetosphere was very quiet. The hot plasma sheet (PSh, 40.3%) is observed mainly between 2000 and 0800 LT, but also reaches round to the dayside. The warmer, less dense plasma trough (PT, 22.5%) is seen over a similar LT range to the plasmasphere, and a combination of the plasma trough and plasma sheet (PT/PSh, 18.6%) is seen mostly pre-noon. An empty trough region, devoid of plasma sheet, plasmasphere or plasma trough populations (Empty, 4.3%) is seen between the plasma sheet and either plasmasphere or plasma trough, on the dawn or dusk side. These trough regions were explained by changes in the convection
electric field, causing different particle orbits to open or close, trapping or untrapping different populations (McComas, 1996). The magnetosheath and/or low latitude boundary layer (Msh/BL, 0.7%) is observed mostly on the dawn side. The magnetospheric lobe is seen only rarely, between 0000 and 0800 LT (Lobe, 0.3%). Lobe observations, characterized by no ions of energy above a few eV and no electrons of energy above a few keV, tended to correlate with magnetopause crossings, implying they occur when the magnetosphere is compressed. Magnetopause encounters were found to be more common pre-noon than post-noon. The low energy end of the ring current was also a constant feature of the MPA observations at the top part of its ion energy range.

Moldwin et al. (1996) used multiple MPA and SOPA data to observe ‘plasmaspheric intervals’, regions where plasmaspheric plasma is observed at geosynchronous orbit. The intervals, identified by cold (~1eV), dense (10-100cm⁻³) plasma, are generally seen around dusk, although at low $K_p$ they are larger and are seen post-dusk, whereas at high $K_p$ they are smaller and seen close to noon. They result from the spacecraft traversing the plasmasphere’s dusk-side bulge. From a survey of 15 plasmaspheric intervals and their relation to substorm injection events, the authors conclude that plasmaspheric regions are brought out to geosynchronous altitudes by the reconfiguration of the duskside magnetosphere during magnetospheric substorms. Since the thermal ion motion of the plasmaspheric population is governed by $E \times B$ drift, the outward motion of the plasmasphere requires an eastward electric field, which is generated in the midnight region by the earthward motion of the plasma sheet boundary during the substorm growth phase.

Thomsen et al. (1996a) studied plasma sheet particles at geosynchronous orbit using data from three MPAs. The warm electrons of the plasma sheet would appear, lower energies first, after the spacecraft observed the cold ions of the plasmasphere in the dusk sector. Sometimes this energy dispersion was not seen and this was interpreted as either sunward motion of the plasma sheet boundary, or as a sudden particle injection. Both ion and electron plasma sheet densities were found to peak in the evening and night sectors, the minimum occurring in the
Figure 2.1: Local time distributions of the various magnetospheric regions observed by the MPA on three geosynchronous satellites. From McComas et al. (1993).
noon/dusk region. The distribution was found to be generally anisotropic \((T_\perp > T_\parallel)\) but most anisotropic in the morning sector, and least anisotropic at dusk. Superposed epoch analysis of the plasma sheet at geosynchronous orbit in the course of substorms was carried out by Birn et al. (1997) using MPA and SOPA data. The zero epoch time for the analysis was established from dispersionless injections of 50keV to a few hundred keV ions, then the plasma sheet flux at 30eV to 40keV was superposed. Generally, the plasma sheet electron population increases in temperature and density at substorm onset. The anisotropy shows the injection is associated with a brief parallel electron distribution, and the axis of symmetry obtained from the anisotropy calculation shows the stretched field geometry undergoing dipolarization at onset.

The tail lobes are occasionally observed at geosynchronous orbit when geomagnetic activity is high. Lobe encounters at geosynchronous orbit, identified by no ion flux above 1eV and no electron flux above several hundred eV, are a manifestation of the distortion of the tail current sheet away from the magnetic equator during high geomagnetic disturbance (Thomsen et al., 1994). Lobe encounters were further investigated by Moldwin et al. (1995). Two types of encounter, flank and midnight events, were both found to occur at high \(K_p\). Flank encounters were often associated with magnetopause crossings on the same day and high IMF \(B_z\), indicating a compressed magnetosphere, whereas midnight events were believed to be associated with the substorm growth phase: simultaneous observations from other geosynchronous satellites showed that the tail field was extremely taillike at the time of the lobe events, and that injections were frequently associated with their observation.

**SEM-2 energies**

Many observations of the substorm growth phase and particle injections from geostationary orbit have been made at the energy range of the SEM-2. At geosynchronous orbit \(\leq 300\text{keV}\) electrons track the solar wind conditions in the respect that they are controlled by substorm activity which is in turn controlled by the solar wind velocity and IMF (Baker et al., 1997).
Garrett et al. (1981) used 30eV - 80keV electron data from ATS-5 and ATS-6 to examine different geosynchronous populations and their behaviour for different local times and geomagnetic activity levels. Using a double Maxwellian (i.e. two population) representation, they identified a hot component related to trapped electrons, a cooler but denser population seen primarily at midnight which they relate to the plasma sheet, and a second cooler component seen at all local times which they refer to as either high energy plasmasphere electrons or residual electrons from previous injections. Cayton et al. (1989) divided the 30-2000keV electron energy spectrum into two components, one below 300keV dominated by substorm injections, and the other showing little influence of them.

The geosynchronous environment experiences extreme variability as the result of magnetospheric substorms. During the growth phase energetic particle flux is seen to decrease, then at onset increase suddenly, often by orders of magnitude. It had been suggested that the substorm growth phase might be a manifestation of a previous substorm expansion or recovery phase, but Baker et al. (1981b) argued against this by showing that an isolated substorm exhibited growth phase characteristics which could not be associated with any other substorm. The injection was observed in geosynchronous >30keV electrons and coincided with large negative bays in auroral zone magnetograms, indicating onset. Multi-point (geosynchronous and ground-based) measurements by Baker et al. (1982a) also showed that growth phase features were peculiar to one substorm and not 'left over' from a previous one.

The onset of the isolated substorm studied by Baker et al. (1981b) was preceded for an hour and a half by the stretching of the field to a taillike configuration and the development of the cigar phase at geosynchronous orbit. Baker et al. (1982a) also found the early growth phase is weakly cigar-like, the late growth phase strongly cigar-like and after onset the distribution is strongly pancaked. Shell splitting leads to a nightside depletion of 90° particles in the growth phase, seen prior to pre- and post-midnight injections.

Baker et al. (1978) defined an anisotropy sequence for 30-300keV electrons in the course of substorms with the Los Alamos LoE data. There was no in situ
magnetic field data so the symmetry axis of the distribution was calculated using measurements from the LoE's five sensors at 30° intervals to the spin axis, combined with the spacecraft spin, by fitting the particle data to spherical harmonics (Higbie et al., 1979), and the symmetry axis was then assumed to be the field direction. Baker et al. found that 1-2 hours before onset the electron distribution is field aligned accompanying a tail-like field, then at onset changes abruptly to a pancaked distribution and the field is dipolarized, and this distribution persists after the injection. The electron flux is seen to decrease prior to onset, then after the injection return to a level higher than before the substorm, although the size of the enhancement decreases with increasing energy. During the dropout the electron distribution evolves from isotropic to cigar-shaped. The injection corresponds to local auroral zone ground observations of magnetic bays indicating substorm onset. The median time between the cigar-shaped distribution setting in and the injection was found to be 95 minutes. The interpretation of the electron anisotropy sequence is that as the tail becomes stressed, the electrons become field-aligned through their drift paths in the distorted field. Magnetopause shadowing may also contribute. It was noted that this sequence was only observed prior to a substorm: if there was no substorm, then the cigar phase did not occur and near isotropy was maintained. Therefore the cigar phase can be seen as a manifestation of the growth phase and may provide a predictive tool for substorm occurrence. It was also noted that 70% of the cigar phases were associated with southward turnings of the IMF. The cigar-like distribution was observed about 25% of the time, on the nightside. At high $K_p$, the cigar phase was more pronounced, and began as early as noon, and the pre-noon distribution was more pancaked. During quiet times a local time dependence was still evident, with the distribution being less pancaked at midnight than at noon. The 30-300 keV electron flux was observed to peak at noon in quiet times but in the morning sector at moderate $K_p$, which is a combination of the effects of shell-splitting, magnetopause shadowing and substorm injections. Drift path tracing showed that for higher energy particles, because their radial gradient is higher, the probability of developing a cigar phase is higher. The cigar phase is more
pronounced at dusk than dawn because the distribution is masked by the more isotropic, freshly injected particles.

Baker et al. (1982a) found from IMP-8 solar wind data that the growth phase corresponds to a southward turning of the IMF. Energy storage is enhanced by the southward IMF, and this energy storage results in a taillike field and cigar-shaped distribution. If the IMF turns northward, the energy storage ‘levels off’, as does the development of the geosynchronous substorm features. In one case observed, the northward turning of the IMF after a period of being southward may have triggered the substorm. In the case of multiple substorms observed, each one was preceded by a period of southward IMF and geosynchronous cigar phase. These observations support the energy storage and release (unloading) model of substorms: energy is imparted to the magnetosphere from the solar wind when the IMF is southward, and stored as tail currents. Energy input stops if the IMF turns northward, and the magnetosphere stays in its stressed form until the IMF turns southward again and energy storage resumes, or until a substorm occurs and the energy dissipates in accelerating particles, heating plasma and diverting large currents through the ionosphere. The switch to a pancaked distribution at onset sometimes seemed to coincide with IMF reversals but not always, so this could not be identified as a substorm ‘trigger’.

Solar wind-magnetosphere coupling and tail energy storage was studied by Baker et al. (1985) using geosynchronous satellites in conjunction with solar wind and tail observations. It was found that the magnetosphere responds to the southward turning of the IMF with the first signs of the growth phase within 10-15 minutes; dissipation then usually occurs with the expansion phase onset after about an hour of loading. These observations point neither to a directly driven model, since onset does not occur until energy loading has been going on for about an hour, nor specifically to an energy loading and release model, since low-level energy dissipation occurs throughout the growth phase. Instead, both types of model have factors to contribute.

Arnoldy and Chan (1969) observed geostationary electrons for substorms occurring near midnight during moderate activity with ATS-1. They found the
injection spectrum to be soft, since the injection was observed in 50-150keV electrons but not always seen in 150-500keV electrons. They observed the substorm electrons to be produced near midnight and drift, and that the drifting electron bunch was associated with precipitation: from this they concluded that the injection may cause the stable trapping limit to be exceeded.

Following a substorm injection, magnetospheric waves can resonantly interact with particles, scattering them into the atmospheric loss cone. Since the waves which do the scattering are most effective at the geomagnetic equator, Baker et al. (1981c) attempted to correlate auroral zone riometer data with geosynchronous electron (>30keV) data. The study showed that substantial precipitation is always accompanied by the substorm injection of newly energized particles. These particles are not just a displaced, pre-existing population which moves across geosynchronous orbit as the magnetosphere expands or contracts, but are accelerated at substorm onset. The field is observed to be stretched to a tail-like configuration on the nightside 1-2 hours before onset which causes shell splitting. The particle distribution is field-aligned leading up to onset. During weak to moderate precipitation events, the pitch angle distribution at the outer zone geomagnetic equator is peaked at 90°, consistent with the predictions of the weak diffusion case. Strong precipitation events occur when the flux at geosynchronous orbit exceeds the Kennel-Petschek weak diffusion stable trapping limit. Although the flux at geosynchronous orbit rarely exceeds the strong diffusion limit, when it gets close very large precipitation events are seen at the conjugate riometer station. These large precipitation events are accompanied by isotropic distributions at geosynchronous orbit, primarily at low energies (30-60keV) consistent with whistler mode cyclotron resonance scattering. During strong diffusion, strong, fast (~10 second) fluctuations in the loss cone flux (not at other pitch angles) are seen at the geomagnetic equator. Strong diffusion was seen only at local times 2300-0900, and where most likely (0000-0800LT) was seen only 10-15% of the time, most often when the solar wind speed was high. Baker et al. (1982a) observed a substorm for which the precipitation of auroral electrons was delayed, and this was interpreted as being the time for the stable trapping limit to be
A large substorm, the last in a sequence of four following a global SSC, was studied by Baker et al. (1982b) using multiple satellites at three different local times at geosynchronous orbit. A flux dropout was observed 25 minutes prior to the injection in 30-630 keV electrons at midnight, due to the spacecraft moving into the plasma sheet prior to onset where the particle flux is much lower. The same injection observed later at LT 0700 was preceded by a gradual decline rather than a dropout, and the injected particles were energy-dispersed. Still later observations (LT 1300) showed further energy dispersion, with the leading edge of the injection being sharp only for energies >80keV. Similar signatures were observed for protons and at >0.4MeV drift echoes were observed; tracing back each drift echo gave similar times for the injection of close to midnight. For this same injection, gradient anisotropy information was used to find the origin of the injected particles. For a local injection, using 100-200keV protons which have gyroradii ~0.1R_E it is possible to tell, from which side of the spacecraft sees the flux and density enhancement first, which direction the injection comes from. The injection was found to originate from beyond geosynchronous orbit.

Reeves et al. (1990) observed a substorm injection using data from the CPAs on three synchronous satellites simultaneously. The event was isolated and occurred in a fairly quiet time. When first observed at LT 0300, the electron signature showed no energy dispersion, showing that the observation took place very close to the injection site. The second satellite, at LT 1330, observed the electrons to be energy dispersed and the leading edge of the signature to be less sharp. The third satellite, at LT 1745, observed the energy dispersion to be more advanced. The observations are consistent with the injection occurring some time near midnight and the electrons then drifting eastward. Since ions drift in the opposite sense to electrons, the ions reached the satellite at LT 1745 first, which observed a fairly sharp peak; some dispersion could be seen by the time the ions reached the satellite at LT 0300. Observations of the westward-drifted ion injection signature were traced back using a field model, and their projected injection time coincided with the electron injection and the ground magnetograms. By
using the ions to locate the western edge of the injection region and the electrons
to locate the eastern edge, the injection region was found to span 90° around
midnight.

Drift shell tracing of injected particles in a model field was carried out by
Reeves et al. (1991) for three simultaneous observations of an injection by geosyn-
chronous satellites, two of which were on the same drift shell at the time. The
drift paths were found to extend further on the nightside and were closer to Earth
on the dayside when $K_p$ was high. The tracing showed that the electrons were
injected further east than the ions were. Tracing back from all three satellites
pinpointed the time of observation, and pointed to an injection region 50° wide.

Pc pulsations have been observed in the >30keV electron flux at geosyn-
chronous orbit by Baker et al. (1980). There are two principal kinds of geomag-
netic ULF pulsations: Pi (irregular) and Pc (continuous). They have periods of
less than one second to several hundreds of seconds, and amplitudes of tenths to
hundreds of nanotesla, and have been shown to correlate with particle precipi-
tation and auroral intensity. Saka et al. (1992) showed that the flux increase
of 30-200keV electrons at geosynchronous orbit is almost coincident with Pc5
pulsations and precipitation measured at the ground, and suggested that the pu-
slations are in fact excited by the injection. The pulsations observed by Baker et
al. were Pc4 and Pc5 pulsations with periods between 1 and 10 minutes. The
observations showed strong dependences on magnetic latitude and on season.

Multiple onset substorms at geosynchronous orbit were studied by Nagai et
al. (1983). Activity at synchronous orbit was observed in a discrete pre-midnight
longitude sector, with no signs observed elsewhere on the orbit, then propagated
east and west with successive onsets. Substorms often have multiple expansion
phase onsets, characterized by injections, each with its own associated auroral
signatures and auroral zone magnetic field negative bays. The onset on the
ground correlates with the injection, which follows a dropout. Their observations
also suggest that small substorms may occur without complete dipolarizations,

i.e. with some energy still left in the tail.

The frequency of substorm injections was investigated using multiple geosta-
tionary satellites by Borovsky et al. (1993). Using observations close to midnight, a sample of 1001 inter-substorm times produced an average time between onsets of 5.74 hours, and a most probable time of 2.75 hours. The conclusion was that substorms may become ‘periodic’ if conditions are favorable. These higher frequency substorms are interpreted to be at the fastest energy loading and unloading time of the magnetosphere. The lower frequency substorms were interpreted to be ‘random’ events, where perhaps a solar wind trigger causes the onset. Half of all substorms were found to fall in each category. An alternative interpretation is that two random processes, one solar wind and one magnetospheric, act to trigger substorms, and when both probabilities are high, substorms can be quasi-periodic.

**High energies**

Observations of geosynchronous MeV electrons, *e.g.* from the CPA HiE, show different behaviour from electrons of energies seen in substorm injections, and these may have some features in common with the high energy end of the SEM-2 range.

The relativistic electrons at geosynchronous orbit have been shown to exhibit solar cycle dependence, with minimum flux occurring at solar maximum (Baker et al., 1986, and Belian et al., 1996). Belian et al. showed that although the relativistic electron flux correlates with solar wind velocity in the short term, the long term correlation is not good. Also, although high solar wind speeds are required to produce a high flux of relativistic electrons, they will not necessarily do so. Baker et al. (1990) established that relativistic electron enhancements at geosynchronous orbit lagged 2-3 days behind solar wind velocity and the geomagnetic indices $K_p$ and $AE$. Correlating the electron flux and geomagnetic indices produced peaks at multiples of 13 and 27 days. Blake et al. (1992) argued that large flux enhancements require a substantial solar wind velocity increase *and a southward turning IMF*: if the IMF is northward then the high speed pulse has no effect on the relativistic electrons.

Baker et al. (1997) studied the response of relativistic electrons at geosyn-
chronous orbit to weak geomagnetic storms, which are caused by high speed solar wind streams (as opposed to coronal mass ejections which are responsible for strong geomagnetic storms). Because of the Sun’s rotation the weak storms may recur at 27 day intervals and this is reflected in the relativistic electrons. In a sequence of three ‘recurrent’ storms observed, the third and largest was seen to produce a lasting enhancement in the >2MeV electrons. It was thus shown that even weak storms can increase the relativistic electron flux by a factor 10 or more on a time scale of the order of a day.

Large, persistent enhancements in highly relativistic electron flux (3-10MeV) at geosynchronous orbit were found to be rare at solar maximum but frequent on the approach to solar minimum, where a 27-day periodicity associated with solar wind streams could be detected (Baker et al., 1986). Superposed epoch analysis showed that these relativistic enhancements had rise times of 2-3 days and decay times of 3-4 days. The suggested mechanism behind this is that high speed solar wind streams cause geomagnetic activity, and the resultant substorms produce ≤ 1MeV electrons which are then internally accelerated by the magnetosphere. An alternative theory, that Jovian or solar wind electrons enter the distant tail and are accelerated earthward during substorms is supported by the fact that these electrons have a 27-day periodicity due to their transport by the solar wind. However, sometimes the geosynchronous relativistic electron flux obviously does not correlate with the Jovian or solar wind sources, which does point to an internal acceleration mechanism (Baker et al., 1989). The production of relativistic electrons from the internal acceleration of substorm injected electrons was questioned by Blake et al. (1992) since the ‘seed’ population is always present, even in the absence of southward IMF turnings which, combined with fast solar wind, produces relativistic enhancements.

A possible mechanism for generating the relativistic electrons observed at geosynchronous orbit begins with substorm injected electrons undergoing radial diffusion inward which violates the third adiabatic invariant, resulting in a large energy gain transverse to the field. Then, pitch angle scattering at low L, which violates all three adiabatic invariants, puts the electrons on higher shells which
puts them back in the outer zone. This may recur to 'pump' the electron energy up (Baker et al., 1989 and Fujimoto and Nishida, 1990). If this recirculation process is occurring, then the pitch angle distributions should be butterfly at relativistic energies, but not at lower energies: Baker et al. report butterfly distributions of relativistic electrons and pancaked distributions of lower energy electrons simultaneously.

2.2 Energy Spectra

In this section energy spectra are produced and compared with those of Cayton et al. (1989), which were obtained from CPA HiE and LoE data. The geosynchronous plasma is not strictly Maxwellian, but the fitting of the particle distribution in this energy range to two Maxwellians has been shown to be adequate. Cayton et al. divided the 30-2000keV electron energy spectrum into two components; a 'soft' component (30-300keV) characterized by substorm injections and a 'hard' component (300-2000keV) which shows little variability on the time scale of substorms. The differences in the flux in the SEM-2 energy ranges reveals these two different electron populations: the lower energies show large substorm injections, but the E1 range shows little evidence of them. SEM-2's four lowest energy bands (42.9-201.8keV) identify with the 'soft' electron component, whereas the highest energy band (201.8-300keV) identifies more with the 'hard' component.

Cayton et al.'s spectra are shown in figure 2.2. The undisturbed period (a) was at 0530 UT and the disturbed period (b) was at 1100 UT: the longitudes of spacecrafts 1982-019, 1984-129 and 1984-037 put them at local times 0300, 1030 and 1930 respectively for the undisturbed event, and at 0830, 0100, 1600 respectively for the disturbed event. These spectra therefore show that the characterization of the plasma by two Maxwellians is valid for different local times and different disturbance levels.

For a plasma in thermal equilibrium, each species may be described by a
Figure 2.2: Synchronous orbit 30-2000keV electron energy spectra: (a) undisturbed ($K_p \simeq 0$) period on 13 January 1986, and (b) disturbed ($K_p \simeq 5$) period on 27 January 1986. Circles are from 1982-019, squares 1984-129 and triangles 1984-037. From Cayton et al. (1989).
Figure 2.3: SEM-2 energy spectrum showing Cayton et al. 's data. The dashed and dotted lines show data from spacecraft 1984-129.
Maxwellian distribution $f_j(E)$ (Cayton et al., 1989, and references therein):

$$f_j(E) = \frac{n_j(2m_jc^2)^{1/2}}{4\pi(kT_j)^2} \frac{\exp(-E/kT_j)}{\alpha \exp(\alpha)K_2(\alpha)}$$

where:

- $E$ is energy
- $n_j$ is the density of species $j$
- $m_j$ is the mass of species $j$
- $c$ is the speed of light
- $k$ is Boltzmann's constant
- $T_j$ is the temperature of species $j$
- $K_2(\alpha)$ is a modified Bessel function of the second kind of argument $\alpha$

where $\alpha \equiv (m_jc^2/kT_j)$.

The differential flux $J(E)$, divided by $p^2/2m_0$, gives the distribution function $f_j(E)$ multiplied by $\sqrt{2/m_0}$. On a graph of the natural logarithm of flux versus linear energy, the Maxwellian is a straight line with gradient $a$:

$$a = -(1/kT_j)$$

and intercept $b$:

$$b = \frac{2c}{4\pi(kT_j)^2\alpha \exp(\alpha)K_2(\alpha)} n_j$$

Therefore, from the gradient and intercept, the temperature and density of the plasma can be found.

Figure 2.3 shows the SEM-2 data, with Cayton et al.'s data from spacecraft 1984-129 to aid comparison with figure 2.2. The SEM-2 points are averages for all local times and for quiet ($K_p \leq 1+$) and active ($K_p \geq 6-$) conditions. The placement of data points agrees well. The difference between the high and low $K_p$ lines in the SEM-2 data is less than for Cayton et al.'s two intervals, but since the SEM-2 points are long-term averages this is not surprising. Although the E1 point falls in the region of 'crossover' between the two Maxwellians on figure 2.2, it is regarded as the part of the SEM-2 energy range representing the trapped population because it contains little evidence of substorm injections.
Figure 2.4: Energy spectrum of SEM-2 data, showing high and low $K_p$ averages

The average energy spectrum for SEM-2 data is shown in figure 2.4. The SEM-2 has only five energy ranges, which span the energies around where the two Maxwellians cross in figure 2.2. In Cayton et al.'s spectra the transition between the two populations was at 200-500 keV depending on activity. The fact that there are few SEM-2 points and that they lie close to this crossover position is likely to limit the accuracy in determining temperatures and densities from the SEM-2 data for the aforementioned hard and soft components. The lower energy component may be adequately defined by the lowest three energy ranges of the SEM-2 (42.9-134.9keV), although since lower energy points are not present, the line which is fitted to these three points may be slightly too shallow. This would mean that the temperature determined from the line may be too high and the density too low. The higher energy component is expected to be less well defined: as can be seen on figure 2.2 there may be points close to the crossover of the two Maxwellians which actually lie on neither line. If this is the case with the SEM-2 data then fitting to the two highest energy points can only give the limiting case, because the line fit is likely to be too steep: the lower limit on temperature and the upper limit on density will be found.
Also shown in figure 2.4 are spectra for different $K_p$ conditions. The gradient of the whole spectrum is higher for high $K_p$. At low energies this is due to increased activity producing more substorm injected particles at geosynchronous orbit. The high $K_p$ line also has the highest energy point at a significantly lower flux value than at lower $K_p$. This is due to the magnetosphere being compressed during strong activity, so that the radiation belts are moved earthward and a lower flux of radiation belt particles will be measured at geosynchronous orbit on the far side of the outer belt. This point will also include magnetopause encounters which cause the flux to drop out. The low $K_p$ spectrum shows flux at all energies is lower than the average case: few substorms are occurring so energetic electrons are not being delivered to the geosynchronous region.

Cayton et al.'s results for the soft (low energy) component suggest a density of $5 \times 10^{-3} \text{cm}^{-3}$ and a temperature of 25 keV. The hard (higher energy component) was found to have a density of $10^{-4} \text{cm}^{-3}$ and a temperature of 200 keV. To calculate parameters from the SEM-2 spectra the transition between the hard and soft components in figures 2.4 has to be placed between the third and fourth points ($E_3\rightarrow E_2$, $E_2\rightarrow E_1$). As noted previously, in Cayton et al.'s spectra the transition between the two curves is gradual: there are points in figure 2.2 (a), close to where the curves cross, which lie on neither curve. Since it is known that the $E_1$ energy range is not generally influenced by substorm injections, then the $E_1$ point on the energy spectra is expected to give the best approximation to Cayton et al.'s hard component, but two points are used to obtain a fit. The line formed by the three lowest energy SEM-2 points is used to define the soft component.

The results obtained for the soft component are a temperature of about 30 keV and density of $3 \times 10^{-3} \text{cm}^{-3}$, and for the hard component a temperature of 105 keV and density $1 \times 10^{-3} \text{cm}^{-3}$. Comparison with Cayton et al.'s results shows that for the soft component the agreement is fairly good, but for the hard component the results are less similar. So the lines fitted here are shallower for the soft component (temperature higher, density lower) and steeper for the hard component (temperature lower, density higher).

In figure 2.5 the dependence of the SEM-2 energy spectrum on local time is
The shape of the energy spectrum changes with local time because the two electron populations have different diurnal variations. The high energy population has peak flux at midday and low flux at midnight, whereas the low energies peak in the morning sector. The steepest spectrum is therefore seen in the morning sector where substorm injections are occurring and increasing the flux of low energy particles, and the shallowest spectrum is observed a few hours before midnight. This local time dependence explains some of the features on figure 2.2. The disturbed plot shows a vast difference in the energy spectra observed by the three spacecraft for the low energy population. But the steepest line was observed at 0100LT, the next steepest at 0830 LT and the shallowest line at 1600 LT, which when compared with in figure 2.5 is explained by the influence of injected particles at different local times. The higher energy population is observed with lower flux closer to midnight.

Figure 2.6 shows an energy spectrum for AE-8 model output. This was produced using the RADMODLS package written by A. L Vampola. The spectra follows a similar shape to the SEM-2 spectra, but shows that higher energy flux is overestimated. In comparison with both the SEM-2 and Cayton et al.’s spectra,
the low energy flux is underestimated. The spectra of course depend on activity, and whereas Cayton et al.'s represent very short time intervals, the SEM-2 spectra shown here are long-term averages. AE-8 was based on a moderate solar cycle so quiet or active cycles may show differences (Vampola, 1989). Although AE-8 has versions for solar maximum and solar minimum, at geosynchronous altitudes there is no difference between them (Pierrard and Lemaire, 1996).

A double Maxwellian fit to the electron population described by the AE-8 model was carried out by Pierrard and Lemaire (1996) and proved a good fit over a range of L values including geostationary orbit, for energies up to 4MeV. Garrett et al. (1981) used a two Maxwellian fit to the electron distribution function for geosynchronous electrons of 30eV-80keV data from ATS-5 and ATS-6. They identified a more energetic component, corresponding to trapped electrons, and two cooler components, one identified as plasma sheet electrons (seen near local midnight) and the other as either high energy plasmasphere electrons or as previously injected electrons which have drifted in local time.
2.3 Applications of the SEM-2 Data

Although the SEM-2 data have relatively low time resolution, the long data set provides a vast resource for the study of time dependent features with time scales ranging from tens of minutes to years.

From the review of previous observations of geosynchronous electron features in section 2.1 it is clear that the SEM-2 energy range is ideal for observing new plasma as it is injected from the magnetotail during substorm reconfigurations of the geomagnetic field, as well as the resident radiation belt electrons. The last section produced energy spectra which showed the presence of these two populations. Substorm injections are an obvious candidate for study, and although the SEM-2 data may miss fine detail, much can be found out about the size and frequency of injections and their distribution in local time. Strong local time behaviour exists in both the substorm injected particles and the trapped particles at geostationary orbit. Analysis of this local time behaviour with such a large data set can provide much information about substorms and the geomagnetic field structure. The E1 energy range only will be used to represent trapped (non-substorm influenced) electrons, and the E5-E4 ranges used for the study of substorm-related phenomena. Sometimes it may be possible to infer things about the electron environment by the presence of dropouts of the SEM-2 energy ranges: if the high electron flux at radiation belt energies disappears, then the SEM-2 is observing outside the radiation belt, perhaps in the tail lobes or beyond the magnetopause.

The SEM-2 data set is seven years long which provides observations over a large part of a solar cycle, from just before solar maximum to just before solar minimum. Solar activity has a strong effect on geomagnetic activity and this will be seen in the radiation belts. Long periodicities such as that associated with the 27-day solar rotation may be studied. Also, because the Earth’s geographic and magnetic axes do not coincide, geostationary orbit lies between $+11^\circ$ and $-11^\circ$ magnetic latitude. During periods of high activity, when the inner edge of the plasma sheet moves inside geostationary orbit and the field is stretched into a tail-like configuration, a satellite at $\pm11^\circ$ magnetic latitude may find itself on
a field line stretching far into the tail (Kivelson and Russell, 1995). This brings another component of variability, on a seasonal time scale, into the expected environment at geostationary orbit.
Chapter 3

The Meteosat SEM-2 Data Set

3.1 Meteosat P2

The geostationary satellite Meteosat P2 was launched in June 1988, one of a series of European Space Agency weather satellites. It began as a prototype of Meteosat 2 and was later made flight-worthy. It was spin stabilised at 100 rpm and equipped with a visible-infrared radiometer and a meteorological data collection system.

The first Meteosat satellite experienced operational anomalies which showed a correlation with geomagnetic activity and local time. The first Spacecraft Environment Monitor (SEM-1) was included in the payload of the second Meteosat spacecraft, Meteosat F2 (Johnstone et al., 1985), to investigate differential charging which was at that time believed to be the cause of the anomalies. Although the SEM-1 confirmed that spacecraft charging was occurring, the anomalies showed no correlation with the charging events. The third Meteosat spacecraft, Meteosat P2, was equipped with the second Spacecraft Environment Monitor, SEM-2, to investigate higher energy electrons. The SEM-2 data were used to confirm a correlation between the Meteosat P2 anomalies and energetic electron flux through a process of deep dielectric charging (Coates et al., 1990 and Rodgers, 1991).

Although in a geostationary orbit, Meteosat P2's location was shifted several times: figure 3.1 shows how its longitude varied between 1988 and 1995. Since the SEM-2 archived data is recorded at universal time, a time adjustment equivalent
to the satellite’s departure from 0° was made to convert to local time. The conversion from universal time to local time required a positive time adjustment for a positive longitude value (i.e. eastward of 0°) and a negative time adjustment for a negative longitude value (westward of 0°). The size of the adjustment is one hour of time for each 15° of longitude. During the mission, the inclination of the orbit did not exceed 1.7°.

In November 1995 Meteosat P2 was switched off and taken out of geostationary orbit. The SEM-2 data set therefore provides an almost complete seven-year record of the electron environment at geostationary orbit. From November 1988 through to November 1995, only three months (March 1991, April 1991 and July 1995) are missing from the archived data set.

### 3.2 SEM-2

The detectors used in the SEM-2 were provided by the Los Alamos National Laboratory (LANL). They are the same detectors that were used in the Low Energy Electron unit flown on the Defense Support Program series of satellites.
Table 3.1: SEM-2’s five differential energy levels

(Aiello et al., 1975). The Mullard Space Science Laboratory (MSSL) supplied the data processing unit which incorporated an energy-level discriminator, the power circuitry and the ground support equipment for the SEM-2.

The SEM-2 consists of five telescopes positioned at 30°, 60°, 90°, 120° and 150° to the spacecraft spin axis. Each telescope comprised a collimator which defines a half angle of approximately 5°, an aluminized mylar window (which does not admit light or protons with energy below 300keV) and a surface-barrier detector which has an active depth of 700µm, corresponding to the range of a 300 keV electron. Internal baffles were designed to reduce scattering from the walls of the collimator. Whilst the telescope positions provide polar coverage, azimuthal coverage was provided by the 600ms spin of the spacecraft.

The total energy range of the detector, 42.9-300keV, is divided into five differential energy ranges by the discriminators E1-E5 (table 3.1). Incident electrons produce a current pulse the size of which depends on their energy. Note that the uppermost 300keV threshold is nominal in that it corresponds to the highest energy particle that is stopped within the detector. Electrons of energy greater than 300keV produce a 300keV pulse but pass through the detector without depositing any more energy.
3.2.1 Calibration

The calibration of the SEM-2 and its integration on to the spacecraft were carried out by MSSL. Calibration was performed at the Goddard Space Flight Center (GSFC) by A. J. Coates and B. K. Hancock of MSSL, supported by S. Brown and C. Smith (GSFC) and T. A. Fritz (LANL). The discriminator thresholds were measured in a vacuum chamber using two electron beam systems with energies 30-120keV and 120keV+. Each telescope was aligned with the electron beam tube in turn and counts were sampled for chosen electron beam energies. All five thresholds were only determined for one of the telescopes due to lack of time; three were determined for each of the other telescopes. The thresholds were found to provide energy bandwidths which were correctly spaced logarithmically, except for the lowest energy one which is slightly smaller. The thresholds were found to be consistent between the telescopes to 10% or better, showing that the system gains were well matched. An azimuthal scan was done at two energies to determine the angular passband of the sensor, which is about 10° in each case.

A sequence was run to study scattering: results showed that a 106keV beam scattered no more than 0.03% in the telescope tested. All angular and scattering tests were therefore deemed to be satisfactory (Coates et al., 1990).

Some of the previous Low Energy Electron units had been found not to behave as expected. Rodgers et al. (1993) compared 45-65keV electron data from the LANL instruments on five geostationary satellites: 1976-059, 1977-007, 1979-053, 1981-025 and 1984-037. They found significant differences existed between the LANL data sets themselves: local time averages of electron flux differed by a factor of 10 or more, from which Rodgers et al. concluded there were problems with the LANL calibration. Rodgers et al. also made a comparison of the LANL data sets with the SEM-2 data. Although this comparison was done prior to the correction of the SEM-2 software errors (see section 3.3.3) in terms of long-term averaged local time profiles of the data, the errors had little effect on the SEM-2 data (Rodgers and Szita, 1993). To begin with, the SEM-2 calibration performed by MSSL produced calibration factors which did not agree well with factors supplied by LANL. Rodgers et al.'s comparison showed that the
LANL flux data were generally much larger (by orders of magnitude) than the SEM-2 data. In the same study Rodgers et al. also compared the high energy SEM-2 data with data from MEA (Medium Energy Analyser) onboard CRRES (Combined Release and Radiation Effects Satellite), and found that these data sets compared favourably.

Baker et al. (1981a) published a compilation of CPA data from 1976 to 1978. This featured data from satellites 1976-059 and 1977-007. A comparison of their local time averages carried out in Chapter 5 shows fairly good agreement with the SEM-2 data, certainly within the same order of magnitude.

Data from other more recent LANL geosynchronous satellites, namely 1984-129, 1987-097, 1989-046, 1990-095 and 1991-080 is now available and this compares much more favourably with the SEM-2 data. An example is shown in figure 3.2. The energy levels shown are at the right of each plot, and the vertical bars on these plots indicate local midnight for each satellite. A substorm injection occurred at 1100 UT which was observed as a sharp peak by 1989-046 which was at midnight, but as an energy dispersed signal by 1990-095 which was at 0930 LT.

The SEM-2 flux data for same day is shown in figure 3.3. The individual LANL energy ranges are not identical to the SEM-2 levels, but the whole energy range is about the same (except for the first two LANL plots which have an extra, lower energy range). Local midnight for Meteosat P2 is shown by the bar at about 0440 UT. Similar features to the LANL plots can be seen including the 1100 UT injection which shows some energy dispersion. The highest energy flux level compares well with the LANL plots. The lowest energy flux is higher on the LANL plots, but this difference is less than an order of magnitude and is not as severe as for the earlier LANL data sets discussed by Rodgers et al. The small discrepancy there is may be partly explained by a problem in the SEM-2 data handling which is discussed in Chapter 8. The middle energy levels also compare reasonably well.
Figure 3.2: Example LANL plot (see text)
3.2.2 Telemetry

Although full 3-dimensional distributions were obtained onboard, these were not telemetered down because of restrictions on the telemetry rates. Instead one 1-dimensional distribution of flux as a function of energy (summed over both angles) and one 2-dimensional distribution of flux as a function of both polar and azimuthal angle (summed over energy) were transmitted. These distributions were transmitted at successive 500, 500 and 600s intervals.

A basic measurement is made by taking counts from each of the five polar angle sensors P1-P5 operating at each energy range for 4ms. So the first five readings are from the first (30°) sensor P1, with the discriminator set at energies E1 to E5, the next five are from the second (60°) sensor, P2, and so on for all five sensors in turn. This data collation takes $25 \times 4\text{ms} = 100\text{ms}$.

The experiment cycle works as follows: first, the time $t$ to an internally generated 600ms pulse, used to mark the start of the sequence for collecting a set of data, is measured to 4ms accuracy. Next, an azimuthal measurement is made over one spin. The measured start-time is later referenced to the spacecraft time so that an absolute time for the start of the azimuthal measurement can be found. The azimuthal measurement is a basic measurement as described above, with the data binned to produce six azimuthal samples, each a sum over 100ms. If the summation overflows, an overflow word is set to indicate which sample overflowed.
Polar and energy measurements are then taken, using the same basic measurement, but with the sequence repeated 6 times to cover a whole 600ms spin. Polar measurements are stored as the sum over energy of each polar bin, and energy measurements are stored as the sum over polar angle of each energy range. The totals are stored after each spin for 10 spins. Should an overflow occur before the end of the 10 spins, an overflow word is set which indicates in which bin and spin the overflow occurred. When an overflow occurs, measurements stop at the end of that spin and the overflow and number of spins accumulated are taken into account later when the data are processed. The overflow capacity effectively increases the dynamic range of the instrument.

The data are compressed at the end of 11 spins (6.6 seconds). There are 16 data words to transmit (6 azimuthal, 5 energy and 5 polar). Each of these is a 16-bit word which is compressed to an 8-bit word using a quasi-logarithmic compression scheme. The compression results in a small uncertainty in the counts, since a range of input numbers will correspond to each output number. A complete experiment cycle including data compression and building up a table of data to transmit is 100.664 seconds long.

### 3.3 The SEM-2 Archived Data Set

Since SEM-2 was turned on in orbit on 22 June 1988, data were collected and archived in monthly units. With the exception of three months (March and April 1991 when the instrument was switched off, and July 1995) complete monthly archive files exist from November 1988 to November 1995 (prior to November 1988 they exist in a different format which is not compatible with current processing software).

#### 3.3.1 Archived data

Daily and monthly summary plots were produced showing grey-scale flux, spectral index, polar and azimuthal flow, anisotropy index, $K_p$ index and memory upset monitoring. These data were archived each month. An example daily plot is
shown in figure 3.4. Shown on the summary plot are:

**Total flux**

This is the electron flux in the total energy range of the detector, 42.9-300keV. The transmitted data are in the form of counts above the 5 energy thresholds (table 3.1), which are then subtracted to give counts in each range. Flux is then obtained from counts using the following equation (Coates et al., 1990):

\[
j = \frac{\text{counts}}{tG\epsilon\delta E}
\]  

(3.1)

where:

- \(j\) = flux in \(\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\)
- \(t\) = accumulation time in s
- \(G\) = geometric factor = \(3.6 \times 10^{-3}\text{cm}^2\text{sr}\)
- \(\epsilon\) = efficiency
- \(\delta E\) = energy width of bin in keV

The accumulation time \(t\) for an energy measurement is 0.12s (4ms for each of the 5 polar sensors, times 6 azimuthal bins per spin) multiplied by the number of spacecraft spins which will be ten, unless there is an overflow.

**Greyscale flux**

This is the flux in the five differential energy ranges in table 3.1, shown as a greyscale plot.

**Spectral index**

The spectral index describes the shape of the energy spectrum. It is the slope of the log of the energy spectrum, and error bars are shown which give the quality of the least squares fit. The spectral index is generally negative, with small negative values indicating a hard spectrum with many particles above the analyser's energy range, and large negative values indicating a soft spectrum where the population reduces as energy increases.
Polar flow

This is a greyscale plot of the flux in each of the five polar angle sectors of the analyser. The polar look directions are at 30°, 60°, 90°, 120° and 150° to the spacecraft spin axis, and each bin covers approximately ±5°.

Azimutal flow

This is a greyscale plot of the flux in each of the six azimuthal angle sectors of the analyser, which cover 0-60°, 60-120°, 120-180°, 180-240°, 240-300° and 300-360°, where at 0° the spacecraft looks towards the Sun.

Anisotropy

This is the second order anisotropy index, which is calculated from the flux in the 30 polar-azimuthal bins of the SEM-2 analyser. The second order anisotropy index indicates how field-aligned or perpendicular the distribution is. A positive value of the second order anisotropy index denotes a field aligned distribution, whereas a negative value corresponds to a distribution which is enhanced perpendicular to the magnetic field direction. Trapped or loss-cone distributions therefore have a negative anisotropy index. Isotropic distributions will have a second order anisotropy index of zero. The calculation of this index is described in the next section.

Axis

The spherical polar angles \( \theta \) and \( \phi \) of the axis of symmetry of the particle distribution in spacecraft coordinates are found from the anisotropy index calculation. This is useful because there is no magnetometer on Meteosat P2 so in situ magnetic field data is not available. Since the radiation belt electrons are organized by the magnetic field, then the symmetry axis of the distribution should give the magnetic field direction.
**$K_p$ index**

$K_p$ is a 3-hourly index of geomagnetic disturbance, having values ranging from 0 to 9 for low to high disturbance respectively. $K_p$ is a planetary index and is based on measurements from many stations, each referenced to their own quiet conditions.

**MUM**

The Memory Upset Monitor shows the occurrence of latch-ups ('Latch') where a test RAM (random access memory) attempted to draw excessive current, and the number of single event upsets ('Upset 1-4') in 4 memory zones of the test RAM.

On figure 3.4 local midnight falls shortly before 0500UT. An injection is seen at approximately 0530UT which is dispersionless and preceded by a dropout. A second injection observed at 0800-0900UT shows energy dispersion. The spectral index shows that the energy spectrum is generally harder on the dayside than the nightside and injections cause the spectrum to soften. The anisotropy index shows that the electron distribution is generally pancaked on the dayside and more field-aligned on the nightside. The substorm signature is investigated further in Chapter 6. The distribution axis is difficult to interpret in spacecraft coordinates but it can be seen that it is much more stable on the dayside than the nightside. A thorough evaluation of the axis determination is given in Chapter 4. No memory upsets are recorded for this day.

Monthly summary plots are also produced. An example monthly plot is shown in figure 3.5. Memory upsets are recorded for day 12. Diurnal variations can be discerned particularly in the anisotropy index and total flux, and in the axis orientation. These plots are therefore of the same format as the daily plots except that $K_p(\tau)$ is plotted instead of $K_p$. This is a weighted average of $K_p$ devised by Wrenn, (1987) which recognizes that the level of activity in the recent past may be important rather than just the present value. $K_p(\tau)$ gives more weight to the
most recent $K_p$ and less to earlier $K_p$, where $\tau$ is a 3-hour attenuation multiplier. Because $K_p$ has a logarithmic scale its linear equivalent $ap$ is used then $ap(\tau)$ converted back to $K_p(\tau)$:

$$ap(\tau) = (1 - \tau)[ap + \tau ap_{-1} + \tau^2 ap_{-2} + \tau^3 ap_{-3} + ...]$$ (3.2)

Thus the contribution of each $ap$ value to $ap(\tau)$ decays as the time increases. The value of $\tau$ used in the SEM-2 processing is 0.95 since this was found by Wrenn to give the best correlation between $K_p(\tau)$ and the Meteosat-F2 anomalies. The contribution of each $ap$ value to $ap(\tau)$ is therefore reduced to $1/e$ of its value after approximately 2.4 days.

Two archive files are produced for each month, one high time resolution and one low time resolution. In the high time resolution files, the data are stored with the maximum time resolution that the raw data were telemetred down at, in records of 500s, 500s, and 600s successively. Lower time resolution files were created with 30 minutes time resolution. These files are particularly convenient for the study of local time behaviour. Both types of archive file contain records comprising 96 elements of 4 bytes each and are identical in format. The 96 variables contained in each record are shown in table 3.2. Where data are missing, a flag of -1 is used. The archiving software was written at MSSL by M. A. Birdseye, based on the original plotting package by H. E. Huckle, and subsequently modified by M. A. Birdseye, S. Szita and N. J. Flowers.

### 3.3.2 Calculation of the Anisotropy Index

A pitch angle distribution can be described by an infinite series of spherical harmonics. The method for doing this has been described by Sanderson and Page (1974), Sanderson and Hyde (1977) and Higbie and Moomey (1977). Sanderson and Page showed that, for a single detector on a spinning spacecraft, the three-dimensional distribution could not be measured unambiguously and therefore neither could the anisotropy, at least not without making assumptions about the particle distribution, such as at what angle relative to the magnetic field the peak flux occurs. Sanderson and Hyde subsequently considered the use of
METEOSAT-P2 Space Environment Monitor
MONTHLY SUMMARY FOR MAR 1993

Total Flux

Flux

Spectral Index

Polar Flow

Azimuth

Anisotropy

Axis

Kp(\tau)

MUM

Figure 3.5: SEM-2 monthly summary plot

* $\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{KeV}^{-1}$
<table>
<thead>
<tr>
<th>element variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 start time of bin in hours UT</td>
</tr>
<tr>
<td>2 end time of bin in hours UT</td>
</tr>
<tr>
<td>3 total flux of electrons in energy range 42.9-300 keV, summed over all polar and azimuthal bins</td>
</tr>
<tr>
<td>4-8 flux in each energy bin, summed over all polar and azimuthal bins</td>
</tr>
<tr>
<td>4 : 201.8-300 keV</td>
</tr>
<tr>
<td>5 : 134.9-201.8 keV</td>
</tr>
<tr>
<td>6 : 90.7-134.9 keV</td>
</tr>
<tr>
<td>7 : 59.4-90.7 keV</td>
</tr>
<tr>
<td>8 : 42.9-59.4 keV</td>
</tr>
<tr>
<td>9 spectral index</td>
</tr>
<tr>
<td>10 the error on the spectral index calculation</td>
</tr>
<tr>
<td>11-15 polar flow</td>
</tr>
<tr>
<td>16-21 azimuthal flow</td>
</tr>
<tr>
<td>22-51 polar-azimuthal flux: in 30 bins, for each polar angle sector across the azimuthal angles sector</td>
</tr>
<tr>
<td>52-56 counts in each of the five energy ranges (see 4-8)</td>
</tr>
<tr>
<td>57-86 polar-azimuthal counts: again, in 30 bins</td>
</tr>
<tr>
<td>87 the second order anisotropy index</td>
</tr>
<tr>
<td>88-89 axis θ and φ: angles defining the axis of symmetry of the plasma distribution (obtained from the anisotropy index calculation)</td>
</tr>
<tr>
<td>90 $K_p$ index</td>
</tr>
<tr>
<td>91 $K_p(\tau)$ (Wrenn, 1987)</td>
</tr>
<tr>
<td>92 latch</td>
</tr>
<tr>
<td>93-96 MUM</td>
</tr>
</tbody>
</table>

Table 3.2: Elements of the archived SEM-2 data set record
multiple detectors to measure the three-dimensional particle distribution. They concluded that three detectors would be best arranged at 45°, 90° and 135° to the spacecraft spin axis to obtain the best description of the distribution. However, Higbie and Moomey pointed out that a more suitable configuration for measuring the distribution of trapped electrons would be an asymmetric arrangement of detectors with respect to the spin axis; in this case the magnetic field direction can vary widely but the particle distribution should be symmetrical.

The SEM-2 detectors provide three-dimensional flux distributions and so allow the determination of the anisotropy of the electron distribution. To begin with the axis of symmetry is found. A transformation to this axis is performed in order to remove uncertainties in the sign of the second order anisotropy index introduced by the fitting process. The index is then recalculated. This calculation produces three orders of terms. Each order of the spherical harmonic expansion requires more terms and presents more information about the distribution: the zeroth order index gives the average intensity of the distribution, the first order index describes the flow of the distribution, and the second order anisotropy index indicates how field-aligned or perpendicular the distribution is. In order to solve for the second order, nine unknowns are needed, so nine equations are required. The SEM-2 data provides 30 such equations via its polar-azimuthal flux bins, therefore the equations are over-determined. The following calculation is taken from Sanderson and Page, and Higbie and Moomey.

Given that a pitch angle distribution can be described by a set of spherical harmonics, the particle intensity $\psi$ in directions $(\theta, \phi)$ is given by:

$$\psi(\theta, \phi) = \sum_{m,n} [A_{mn} Y^m_n(\theta, \phi) + B_{mn} Y^0_{mn}(\theta, \phi)]$$

(3.3)

where $Y^m_n$ and $Y^0_{mn}$ are spherical harmonics. Note that these $\theta, \phi$ give the flux direction, not to be confused with the axis direction in the last section. This can also be expressed by:

$$\psi(\theta, \phi) = \sum_{n=0}^{\infty} C_n P_n(\cos \omega_n)$$

(3.4)
where \( P_n(\cos \omega_n) \) comes from the addition theorem for spherical harmonics:

\[
P_n(\cos \omega_n) = \sum_{m=0}^{n} \epsilon_m \frac{(n-m)!}{(n+m)!} [Y^{e}_{mn}(\theta, \phi) Y^{e*}_{mn}(\theta_n, \phi_n) + Y^{0}_{mn}(\theta, \phi) Y^{0*}_{mn}(\theta_n, \phi_n)]
\]

(3.5)

and where \( \omega_n \) is the angle between \((\theta, \phi)\) and \((\theta_n, \phi_n)\).

Here \( \epsilon_0 = 1 \) \((m = 0)\) and \( \epsilon_m = 2 \) for \( m \geq 1 \).

From equations 3.4 and 3.5, the coefficients \( A_{mn} \) and \( B_{mn} \) from equation 3.3 can be expressed as:

\[
A_{mn} = C_n \epsilon_m \frac{(n-m)!}{(n+m)!} Y^{e}_{mn}(\theta_n, \phi_n)
\]

(3.6)

\[
B_{mn} = C_n \epsilon_m \frac{(n-m)!}{(n+m)!} Y^{0}_{mn}(\theta_n, \phi_n)
\]

(3.7)

The zero order coefficient, \( C_0 \), is obtained with \( n = 0 \). Here \( C_0 = A_{00} \), which therefore is not a function of direction, so \( C_0 \) gives the isotropic component of the anisotropy, or the average of the distribution. The first order coefficient \( C_1 \), obtained with \( n = 1 \), gives the 'unidirectional' component of the anisotropy which describes the net flow of the plasma. The sign of the second order coefficient, \( C_2 \), is physically significant in describing the shape of the distribution relative to its symmetry axis; it is positive for a cigar-shaped particle distribution but negative for a pancaked distribution.

Each of the spherical harmonics can be expressed in Cartesian coordinates as functions of \((x, y, z)\), where the \( z \) axis is the spacecraft spin axis. These coordinates come from \( z = \cos \theta \) where \( \theta \) is the colatitude of the detector look angle and \( y/x = \tan \phi \), where \( \phi \) is the rotation angle of the detector look angle. The first nine harmonics (which is as far as is needed for the calculation of the second order anisotropy index) are expressed thus in Cartesian coordinates:

\[
Y^{e}_{00} = 1 \\
Y^{e}_{01} = z \\
Y^{e}_{11} = x \\
Y^{e}_{02} = \frac{1}{2}(2z^2 - x^2 - y^2) \\
Y^{e}_{12} = 3xz \\
Y^{e}_{22} = 3(x^2 - y^2) \\
Y^{0}_{11} = y \\
Y^{0}_{12} = 3yz \\
Y^{0}_{22} = 6xy
\]

(3.8)
The calculation puts the flux values into five polar bins each 36° wide starting at 18° (centre), and six azimuthal bins each 60° wide starting at 30°. These bins define the angles \( \theta \) and \( \phi \) respectively. These angles deliver \((x, y, z)\) and the spherical harmonics can be calculated. Once the symmetry axis has been found, the flux distribution is transformed to a frame of reference with the symmetry axis as the z-axis, and the process of fitting the harmonics is repeated. The fit is performed twice because the sign of the second order index is difficult to determine. The second order anisotropy index \( C_2 \) is found from the coefficients \( A_{mn} \) and \( B_{mn} \):

\[
|C_2| = [A_{02}^2 + 3(A_{12}^2 + B_{12}^2) + 12(A_{22}^2 + B_{22}^2)]^{0.5}
\] (3.9)

The sign of \( C_2 \) comes from examining the signs of the coefficients (Table 1 in Higbie and Moomey). The spherical polar angles of the axis of symmetry of the particle distribution are also defined by the coefficients from the anisotropy index calculation:

\[
\theta_{axis} = \tan^{-1} 4[/(A_{22}^2 + B_{22}^2)/(A_{12}^2 + B_{12}^2)]^{0.5}
\] (3.10)

\[
\phi_{axis} = \tan^{-1}(B_{12}/A_{12})
\] (3.11)

This is the closest it is possible to get to finding the magnetic field direction since there was no magnetometer on Meteosat P2. This inferred magnetic field direction is compared with the Tsyganenko 89 magnetic field model in Chapter 4.

### 3.3.3 Corrections Made to the Archived Data Set

Several errors were found in the SEM-2 data set in the course of this thesis. The error analysis and corrections were carried out with B. Hancock, M. Birdseye and D. J. Rodgers (MSSL). Programming errors were discovered in three stages of the SEM-2 data processing: in the archiving software, in the onboard data compression software and in the decompression software. A brief summary of the errors discovered is given here and a full description is included in Appendix A.
Archiving errors

The archiving program was found to omit certain variables, recording them as zeros. Also, flux values were 'clipped' because the boundaries from the daily and monthly summary plots were imposed on them in the archiving process.

A third archiving error occurred in the polar flux: Rodgers (1991) had reported an asymmetry in polar flux sectors which he ascribed to an obscured or inefficient detector. Closer examination of the polar flux revealed a normalization error in the treatment of one polar bin which left every third point far too high. There was a further error in the calculation of the polar flux from the polar counts, such that the polar flux was a factor 25-30 too high in all five bins.

Onboard Data Compression Error

A serious error was discovered in the SEM-2 onboard data compression routine, meaning that the data was already corrupted before it was telemetred down. The problem originated from an erroneous bit-shifting operation. Prior to compression the number of counts exists as a 2-byte (i.e. 16-bit) input word, which is compressed down to an 8-bit number to be transmitted. An 8-bit processor is used so each of the 2 bytes in the input word must be handled separately. The compressed 8-bit number consists of a 4-bit mantissa which gives the most significant bits of the number, and a 4-bit exponent which records how many zeros must be added to the end of the number. In some cases bits from both bytes of the input word are used to create the mantissa, and in these cases it is necessary to perform bit-shifts on both bytes. In these cases the bit-shift on the second byte of the input word was one bit short, leaving an ‘empty’ bit. This meant that one bit was wrongly fixed at zero, and the following bits were misplaced.

This meant that there was a set of numbers for which wrong values were always assigned in the SEM-2 data, and the assigned value was always too low. Also, certain numbers could never occur. The E1 (201.8keV threshold) counts were worst affected by the error and were significantly modified by it. The E5 counts (42.9keV threshold) were least affected since most E5 observations occurred above the range of count values affected.
This error could be corrected but with a loss of precision: full details are given in Appendix A. Briefly, the position of the shift error could be obtained from the value of the exponent, and any numbers in the mantissa after the zero shifted forward into their correct places. The precision is decreased because the last bit of the mantissa is irrevocably lost, but the maximum error is still less than 6%.

Decompression Errors

Two decompression errors were found. Normally ten spacecraft spins are used to accumulate each value of counts, but when flux is high the counts can overflow and an overflow flag is set so that the overflow can be taken into account when the counts are decompressed. When an overflow occurs, counting stops at the end of the spin, then the counts are scaled up to give a value for 10 spins (so if an overflow occurs in the second spin, counting stops at the end of the two spins and the total count is multiplied by five). The first decompression error concerned the factor used to scale up the counts in the event of an overflow, which was defined as an integer instead of a floating point number. This could cause an error as great as 67% in any one data point when the number of spins is 6.

A second and more serious error was that in the decompression software all the overflow flags were ignored. This had been done because initial inspection of the data showed some obviously spurious overflow flags in the azimuthal and polar counts. Although the polar and azimuthal overflow flags are still believed to be spurious, as are most of the energy overflow flags, some of energy overflows appear genuine. When the overflow flags were examined it was determined that when the energy overflow flag was set on its own, the overflow was genuine: counts were high with the overflow generally occurring in spin 9 or 10 of the accumulation. If the overflow flag was set in the polar or azimuthal array at the same time, the energy overflow was not genuine: the counts were not high enough to have overflowed, and the false overflow was generally flagged early on in the 10-spin accumulation. On this basis an algorithm was designed to identify where the counts had really overflowed.

Since more counts occur at lower energies making overflows more likely, the
lower energy ranges of the detector were worst affected, and the resulting counts could be orders of magnitude too low.

### 3.3.4 Advantages of the SEM-2 Data Set

The errors described have been corrected, and the analyses presented in this thesis use the corrected data. These data compare favourably with the current LANL data sets and with MEA. The main strengths of the SEM-2 data set are:

- A long continuous time series exists: the data set is 7 years 1 month long, with 3 months missing.
- The energy range of the SEM-2 allows observations of both substorm injected electrons and the trapped population.
- Directional information is available using the five detector look directions and the spin of the spacecraft. This allows the calculation of an anisotropy index and limited pitch angle information.
- The low time resolution archive is in the form of 30-minute averages which is particularly advantageous for averaging or binning data in local time.

The time resolution of the data is however fairly low: SEM-2's highest time resolution data is 8-10 minutes, which means that detail is lost when sudden changes occur. Although three-dimensional distributions are available, they are for the total energy range of the detector, not for the differential energy ranges. Therefore, although the flux of two different electron populations is observed, only their collective anisotropy is available.

Two regions of interest present themselves as pertinent to the two electron populations in the SEM-2's energy range. The first is substorm injections, seen with the low energy end of the SEM-2 range. The second is the geosynchronous magnetic field configuration which organizes the higher energy particles.

The magnetic field structure may be probed using the axis direction angles derived in calculating the anisotropy index. Doing this also enables the method for calculating the anisotropy index to be tested, by comparing the resulting field
directions with those of an established magnetic field model. This analysis is presented in Chapter 4.

Local time analysis of the high energy population also describes the magnetic field structure at geosynchronous orbit. The same analysis applied to the lower energy electrons describes the local time dependence of injection events. Local time models are derived in Chapter 5. The sequence of electron flux and anisotropy in the course of a substorm injection, and the frequency of injections, is addressed in Chapter 6. Longer term dependences for both populations are investigated in Chapter 7.
Chapter 4

Evaluation of the Calculated Magnetic Field Direction

No magnetometer was flown on Meteosat P2, so the magnetic field direction has to be inferred from the particle distribution. If the particle distribution is gyrotropic then it will have an axis of symmetry and finding this axis of symmetry should give the magnetic field direction. In order to test the values obtained for the field direction, they are compared in this chapter with directions obtained using the combined Tsyganenko 89 and IGRF field models.

4.1 Modelling the Magnetic Field

The Tsyganenko 89 models compute the components of the external contributions to the geomagnetic field in the geocentric solar magnetospheric (GSM) coordinate system. The GSM coordinate system is shown in Appendix B. The model covers geocentric distances of up to $70R_E$, and includes contributions from the tail currents, ring current, Chapman-Ferraro and Birkeland currents. T89a was based on data from IMP-A,C,D,E,F,G,H,I,J (1966-1974), HEOS-1 and -2 (1969-1974).

The Tsyganenko 82 model (Tsyganenko and Usmanov, 1982) was based on the merged IMP-HEOS experimental data set and used separate contributions from the ring current, magnetopause and magnetotail current sheet to produce the total field at different levels of geomagnetic activity and different solar wind
conditions. The Tsyganenko 87 model (Tsyganenko, 1987) followed the 82 model and incorporated more magnetotail field data, making it valid up to 70\( R_E \) down-tail as opposed to about 20\( R_E \) for the earlier model.

The Tsyganenko 89 model improved on previous Tsyganenko models by taking into account the warping of the tail current sheet due to the geodipole tilt and by allowing the current sheet to vary in thickness across and along the tail (Tsyganenko, 1989). For non-zero values of the dipole tilt angle (the angle the dipole axis makes with the GSM z-axis) the tail current sheet warps: near midnight the plane of the current sheet gradually departs from the plane of the dipole equator towards the solar wind stream direction, and the current sheet bends in the cross-tail section, the current plane moving above (below) the GSM equatorial plane at the centre of the tail and below (above) it at the tail flanks when the dipole tilt angle is greater than (less than) zero. Previous Tsyganenko field models displaced the whole current sheet rigidly as the tilt angle varied. This seasonal dependence, such that the current sheet is raised above the equatorial plane in Northern hemisphere summer, was also noted by Voight (1984).

The external field model used for this analysis was the Tsyganenko T89c model, which differs slightly from the T89a version in two ways: ISEE-1 and 2 data were added to the existing IMP-HEOS dataset, and the tail field was modified to respond to changes in the dipole tilt angle.

The T89c code used for field calculations was written by N. A. Tsyganenko. The input parameters for this model are the GSM cartesian coordinates of the point, the geodipole tilt angle, and a parameter which defines the \( K_p \) level at the time. The model has seven \( K_p \)-dependent states ranging from 0o to 0+ at its quietest to 6- and above at highest \( K_p \). The output parameters from the model are the GSM cartesian components of the magnetic field at the point specified.

The Tsyganenko model does not include contributions from the Earth's main (internal) field. The internal geomagnetic field model used was the IGRF (International Geomagnetic Reference Field). The inputs for this model are the year and spherical geographic coordinates of the point, and the outputs are spherical geographic (GEO) components of the Earth’s internal field. The GEO coordinate
The GEOPACK suite of subroutines was used for accessing internal and external field models and for performing the necessary coordinate transforms. It was written by N. A. Tsyganenko and adapted by M. Peredo (both at Hughes STX and NASA/GSFC). The internal field code was also written by N. A. Tsyganenko and modified by M. Peredo.

Unlike the Earth’s internal field, the external component is highly variable. Variability can arise from changes in solar wind pressure and in the IMF, and magnetic storms and substorms can reconfigure the field. The T89c model addresses the problem of variability solely by using the $K_p$ index. Since the $K_p$ index provides a measure of magnetic disturbance at the Earth’s surface, it may often not be indicative of what is happening at large distances from the Earth. However the $K_p$ index is probably the most widely used and readily available indicator of geomagnetic disturbance.

4.2 Previous Testing of the Tsyganenko 89 Field Model

The Tsyganenko magnetic field model has been found to represent the geomagnetic field well, although a few weaknesses have been identified.

Tsyganenko (1989) himself tested the model against observations at geosynchronous orbit. A year of GOES-2 magnetic field data was used to examine the local time dependence of the magnetic field inclination, in three ranges of $K_p$. It was shown that the Tsyganenko 89 model was closer to the data than either the Tsyganenko 87 or Mead-Fairfield model. However, the Tsyganenko 89 model failed to show the observed dawn-dusk asymmetry of field line stretching which is strongest at high $K_p$. The Tsyganenko 89 model was also closer than the Tsyganenko 87 model at finding geosynchronous midnight values of the GSM $B_z$ component at various $K_p$.

Tsyganenko (1995) identified three main weaknesses of his 1989 model. Firstly, the magnetopause, although defined fairly realistically on the dayside, became less
accurate and less stable further down the tail due to sparse data coverage. The magnetopause however is simply a fit to the data: its shape and standoff distance are not specified as they are in, for example, the Olson-Pfitzer models (Jordan et al., 1992). Secondly, the use of $K_p$ to define the level of geomagnetic activity was crude. Thirdly, there were inaccuracies in the equatorial magnetotail magnetic field values, which caused the model to underestimate the extent of field line stretching in the inner tail.

Jordan et al. (1992) compared various magnetic field models with CRRES observations during both quiet and storm conditions on August 26, 1990. They found the Tsyganenko 89 model to be in good agreement with the CRRES data on average, but incapable of modelling dynamic storm-time changes. For such dynamic events, models based on $K_p$, which is only available every three hours, are limited to averages at intervals throughout the storm. In addition, the model was found to be fast to run and therefore suitable for large data sets. Since the SEM-2 data set is both large and of relatively low time resolution, the Tsyganenko 89 model should be well suited to this kind of analysis. A problem with the model addressed by Jordan et al. is that the Tsyganenko 89 model is not suitable for low altitudes, since the tail current sheet has no inner edge (it gradually decreases earthward) and also there is no eastward ring current component. Significantly, no data were available on which to base the model within $4R_E$. At geosynchronous orbit these concerns should not apply.

Thomsen et al. (1996) used observations from several geosynchronous satellites to test the Tsyganenko T89a magnetospheric field model. The data came from the Los Alamos Magnetospheric Plasma Analyser (MPA) on the three satellites 1989-046, 1990-095 and 1991-080. The MPA has better spatial resolution than the SEM-2: six detectors provide six polar bins between $25^\circ$ and $155^\circ$, and the satellite spin divides the full azimuth angle into 24 sectors of $15^\circ$ each. The time resolution of the MPA data is also better than that of the SEM-2, with a full three-dimensional distribution obtained in one 10s spin, covering an energy sweep through 40 bins between 40keV and 1eV. Using the parallel and perpendicular temperatures of the distribution the symmetry axis was identified, and
this was then assumed to be the magnetic field direction. This method was previously proven using magnetometer data for solar wind measurements by Phillips et al. (1989). Thomsen et al.'s data suggest that the derived field angles are most accurate when the anisotropy is significant, but more scattered where the anisotropy is low, as would be expected. Generally, their calculated field directions agreed well with the Tsyganenko model, both on the magnetic equator and several degrees away. The only significant difference was that on the dayside and away from the magnetic equator, the Tsyganenko model predicted that the angle between the field direction and the orbit plane would be less than the calculated values showed: this suggests that the T89a model may not sufficiently take into account the dayside compression of the magnetosphere by the solar wind. The magnetic field angle modelled on the dayside hardly varies between all stretching levels of T89a. The conclusions of their study were that away from the magnetic equator, the T89a model tends to be more stretched than observations suggest, whilst on the equator the model was overstretched and understretched with fairly equal occurrence. The findings were generally true for all seasons. However, the analysis was limited in the sense that observations were compared as a whole with the range of outputs from the T89a model for between lowest and highest \( K_p \): no comparison was made for periods of high \( K_p \) or low \( K_p \) individually.

Hones et al. (1996) tested the T89a model using data from three polar DMSP (Defense Meteorological Satellite Program) satellites and two geosynchronous satellites. The method they used was to project the DMSP orbit onto the equatorial plane. If at 6.6\( R_E \) its orbit came within 10° of one of the geosynchronous satellites, it became a candidate for a magnetic conjunction. Particle spectra were then compared for the two spacecraft to find the best match. At the time of the best match, the path of DMSP spacecraft was projected onto the ionosphere along with the drift path of the geosynchronous particles calculated with T89a (at the observed \( K_p \)). If the paths intersected then the model and observations were said to agree. Of 47 calculated conjunctions looked for, 20 were found to give good agreement between both satellites in that they showed good spectral similarity for a few seconds. Of these 20, the calculated ionospheric footprint
was within 1° of the observed location in 8 cases, but more than 6° out in 5 cases. All of the bad matches occurred between 0700-0900 LT, perhaps identifying a weakness in the Tsyganenko code.

Weiss et al. (1997) also compared T89a with observational data. Using a set of magnetic conjunctions between low-altitude DMSP satellites and geosynchronous satellites, investigations were made of the ionosphere to geosynchronous altitude field line mapping and of the degree of stretching in the model which is $K_p$ dependent. The conjunctions were identified as intervals having closely matched distributions observed at both satellites. The comparison of the data with the T89a model suggests that T89a underestimates the degree of stretching in the dusk and dawn local time sectors. To fit the observed stretch in the field required a higher value of $K_p$ than occurred at the time of the conjunction. In many cases, the most extreme T89a conditions were not sufficient to fit the observations: some low $K_p$ observations suggested a field configuration less stretched than the most relaxed T89a version, and some high $K_p$ observations suggested a field configuration more stretched than the most stretched T89a version.

4.3 Calculation of the Magnetic Field Direction

The magnetic field direction is assumed to be the symmetry axis of the electron distribution, since the radiation belt electrons are organized by the magnetic field.

Knowing the 30 polar-azimuthal look directions of the SEM-2 allows calculation of the symmetry axis. The SEM-2 timing for azimuthal data collection uses an internally generated spin pulse every 600ms. There is no Sun sensor on the SEM-2 but the internal timer pulse can be referenced to the spacecraft time in the ground processing. The spacecraft spin period can vary slightly: Meteosat orbit and attitude reports produced by ESA confirm that the spin period may be 599ms or 601ms sometimes. Even a 1ms difference in the spin period would change the Sun direction significantly from one data accumulation to the next. This problem is addressed in the processing software, which references the spacecraft time and takes into account the azimuthal offset arising from the slightly
faster or slower spin. The spacecraft spin rate remained at 601ms throughout the year of data used for this analysis.

The $K_p$ index input to the model is the real index at the time of each observed data point. For the calculation of Meteosat P2’s position, it is assumed that the inclination of the satellite orbit to the geographic equator is 0° and that the satellite altitude is fixed at $6.6R_E$.

![Graph](image)

Figure 4.1: GSM theta and phi for day 110-118 of 1994. The values calculated from the SEM-2 data are shown by the solid line. The model values (see text) are shown by the dotted line.

### 4.4 Results

For this analysis one full year of data, 1994, was compared with the Tsyganenko field model.

Figure 4.1 shows GSM theta and phi spherical polar coordinates for the T89 model field (dotted line) and the SEM-2 calculated field (solid line) directions for an eight day sequence. The data and the model agree fairly well. Theta is the angle from the $z$-axis and phi is the angle from the $x$-axis in the $x-y$ plane (see Appendix B). The average $K_p$ for this time was between 2- and 20
and peaked at 4+. Note that local midnight on this plot falls at about 0500UT. Phi switches from 360° to 0° suddenly at local midnight as the observed tail field direction in the equatorial plane crosses the x axis. Theta is close to 0° on the dayside where the field is more dipolar and inclined at between 45° and 90° on the nightside where the field is more stretched in the direction of the Sun-Earth line. This quality of fit to the model parameters is fairly typical of the calculated parameters: about 90% of the year's data shows this degree of similarity with the model.

Local time dependence

The differences between the modelled values for the field direction and those calculated using the SEM-2 data are addressed further in figures 4.2 and 4.3. As before the parameters theta and phi are measured in the GSM spherical polar coordinate system. The whole year of low resolution data are shown divided into local time quadrants. The difference in the two values is the SEM-2 calculated value minus the Tsyganenko model value, therefore, where theta or phi is positive, the model has underestimated the true value, and where theta or phi is negative, the model has overestimated the true value. For theta, which is in the range 0-90°, the difference values are shown in 10° bins. For phi, which is in the range 0-360°, the difference values are shown in 20° bins.

It is clear that both theta and phi are more accurately predicted by the model on the dayside, with a much greater spread of values on the tail side. On the dayside, the calculated theta is within 10° of the model value in 80% of cases, whereas on the nightside only 40% of values are within a 10° difference. Similarly for phi, 76% of values are within a 20° difference on the dayside compared with 37% on the nightside. This reflects the fact that the geometry of the tail is more variable than that of the dayside magnetosphere, changing from a stretched tail configuration to a more dipolar shape in the course of substorms, so that theta can have a much wider range of values on the nightside. As was seen in section 4.2, the degree of tail stretching in the model has raised concern with most users.

Some peculiar features arise on the phi plots, namely small peaks in the
Figure 4.2: The difference between the model value of theta and the value calculated using the SEM-2 data, divided into local time quadrants, for one year of data.

Figure 4.3: The difference between the model value of phi and the value calculated using the SEM-2 data, divided into local time quadrants, for one year of data.
difference values at approximately $\pm 180^\circ$. These 'sidebands' occur more often on the nightside and at high $K_p$. At midnight where the model predicts a switch in phi from $360^\circ$ to $0^\circ$ as the field direction crosses the $x$ axis, the change in the data is less sharp, rarely reaching the $0^\circ$ and $360^\circ$ marks. The geometry of the tail field appears to be much more dynamic than a simple, stretched dipole model suggests. Often phi seems to drop to the $180^\circ$ level a few hours before midnight, then further to zero a few hours after midnight (see figure 4.1). Thus phi is greatly overestimated by the model shortly before midnight and underestimated shortly after. This is the cause of the two sidebands on figure 4.3. These points can also be seen on the dayside, since whenever theta tends to zero, as it may do close to noon, phi may again switch between very different values.

It is also noted that the investigation of these points showed that when phi is found in these sidebands, the anisotropy index is strongly peaked at low values (i.e. near isotropy). In a near isotropic distribution, the algorithm is forced to choose an axis and has a much greater chance of error. Thus the symmetry axis of the distribution is best established when there is considerable anisotropy, which was also seen by Thomsen et al. (1996), The anisotropy index calculated from the SEM-2 data generally lies between -0.6 and 0.2 during 1994, with no values below -1.3 or above 2.7. At low values of the second order anisotropy index (defined as between -0.1 and 0.1, giving a sample of 4421 data points) theta was within $10^\circ$ of the model value in 30% of cases, and phi was within $20^\circ$ of the model value in 28% of cases. At high values of the second order anisotropy index (defined as below -0.4 or above 0.4, giving a sample of 3564 data points) theta was within $10^\circ$ of the model value in 74% of cases, and phi was within $20^\circ$ of the model value in 67% of cases. This indicates that the axis is better determined from the shape of the particle distribution at larger anisotropy.

There may therefore be a combination of effects resulting in the model and calculated fields having vastly different values of phi. Certainly a large factor is the departure of the true field from the idealized geometry of the modelled tail field. The dynamic tail behaviour is difficult to model with merely the $K_p$ index to dictate the degree of disturbance. An additional effect is that the symmetry
axis of the particle distribution is difficult to establish during low anisotropy conditions.

**$K_p$ dependence**

It might be expected that the greatest differences between the model field direction and that inferred from the SEM-2 data would occur at high $K_p$. Both the highest and lowest $K_p$ versions of the model have caused concern in the past (section 4.2). Figures 4.4 and 4.5 show the distribution of the differences between the modelled and observed GSM theta and phi in four increasing $K_p$ groups. At lowest $K_p$ theta is more likely to be overestimated by the model (52% of cases) whereas at high $K_p$ theta is much more likely to be underestimated (65%). The chance of phi being 180° out increases with $K_p$.

To see how the dayside and nightside field angles are affected by $K_p$, figures 4.6 and 4.7 were produced showing the average local time profiles of the field angles for the model and data.

For figure 4.6 GEO theta was used. It is conceptually simpler to look at theta in GEO coordinates for this form of graph rather than GSM, since in GEO theta is the angle of the magnetic field to the normal to the orbit plane (i.e. geographic North). By considering theta in GEO coordinates it is found that the model does underestimate the stretching of the field at high $K_p$. Figure 4.6 shows GEO theta versus local time for increasing $K_p$. The match between the model and the SEM-2 value is best at moderate $K_p$ (second panel), and less close at lowest and high $K_p$. Note that at lowest $K_p$ around midnight, the model appears to overestimate theta, a feature also seen by Thomsen et al. (1996). The dip centred on midnight will be discussed later. Another feature which can be seen in these plots is an asymmetry about noon: the Tsyganenko model has GEO theta at a minimum at noon whereas the data has the minimum value at 1000, and possibly earlier at higher $K_p$. In the light of this theta can be slightly overestimated in the morning sector for most $K_p$ values.

Figure 4.7 is a similar plot for phi in GSM coordinates. It is more intuitive to consider phi in GSM coordinates since in GEO the $x$-axis rotates with the
Figure 4.4: The difference between the model value of theta and the value calculated using the SEM-2 data, divided into ranges of $K_p$, for one year of data.

Figure 4.5: The difference between the model value of phi and the value calculated using the SEM-2 data, divided into ranges of $K_p$, for one year of data.
Figure 4.6: GEO theta (in degrees) versus local time (hours) for increasing $K_p$.
The solid line shows the values calculated from the SEM-2 data. The dotted line shows the Tsyganenko model values.

Figure 4.7: GSM phi (in degrees) versus local time (hours) for increasing $K_p$.
The solid line shows the values calculated from the SEM-2 data. The dotted line shows the Tsyganenko model values.
0° meridian, whereas in GSM the x-axis is the Sun-Earth line. At low to moderate $K_p$ the model and data fit well except in the region ±5 hours of midnight. The deterioration in fit is due to nightside tail field departure from the idealized model situation which was seen in figure 4.1, which causes the averaged nightside profiles to tend towards 180°. The numbers of hours of bad fit in the tail increases with $K_p$ since the degree of stretching is increased and the field more taillike. Again there is also a slight asymmetry about noon for most $K_p$ values which is not seen in the model.

Seasonal dependence

A magnetic field model will show dependence on the time of year because the orientation of the Earth's spin axis to the Sun-Earth line changes as the Earth moves around the Sun. This means the orientation of the dipole axis with respect to the Sun-Earth line also changes. When deformed by the solar wind, compressed on the sunward side and drawn into a tail on the other side, the geomagnetic field adopts different geometries at the solstices and equinoxes.

Figures 4.8 and 4.9 show the distribution of theta and phi differences for each season. The solid lines shows data for ±45 days of the equinox or solstice, the dotted lines ±15 days. These plots show that the difference in behaviour between the calculated magnetic field axis and that predicted by the model shows seasonal dependence. The reasonably equal and opposite effects seen in summer and winter, a bias which is not as extensive in spring or autumn, suggest a shortcoming of the Tsyganenko model. Also the ±180° sidebands in the phi plot, which are connected with differences between the model and real tail configurations, are larger at the equinoxes than at the solstices.

The next two figures 4.10 and 4.11 show the seasonal dependence of the field angles. Figure 4.10 shows GEO theta: SEM-2 calculated values are shown by the solid line, and Tsyganenko model values by the dotted line. Summer and winter local time dependence is better predicted by the model than are spring or autumn. This is particularly evident on the nightside where, at midnight, the data shows theta dipping to a much lower value. In autumn especially the SEM-2
Figure 4.8: The difference between the model value of phi and the value calculated using the SEM-2 data, divided into seasons. The solid line shows values for ±45 days of the equinox or solstice, the dotted line ±15 days.

Figure 4.9: The difference between the model value of phi and the value calculated using the SEM-2 data, divided into seasons. The solid line shows values for ±45 days of the equinox or solstice, the dotted line ±15 days.
calculated theta is frequently observed to drop to zero at midnight, a feature not present in the model (this feature is not real and is discussed below). The same noon asymmetry is seen as before, with minimum theta occurring before noon rather than at noon as the model predicts. In spring and summer theta may be underestimated in the morning sector.

Figure 4.11 shows GSM phi. As was seen before, the SEM-2 nightside average differs substantially from the model, due to fluctuations in the non-ideal tail geometry. A small noon asymmetry is apparent in the SEM-2 values which is not seen in the model, where phi crosses the 180° point slightly before noon in winter but slightly after noon in summer. Winter provides the closest match of model and data values.

The seasonal dependence of the geosynchronous field can be explained by considering the position of the geographic and magnetic equators throughout the year. As the inclination of the Earth’s axis to the Sun-Earth line changes, so does that of the geomagnetic axis. The seasonal effect in the synchronous orbit magnetic field was addressed further by McPherron and Barfield (1980). As shown in figure 4.12 the magnetic equatorial plane can be distorted during active times. This figure was constructed for a synchronous satellite at 70° west longitude which is close to the location of Meteosat P2 when this data was collected. When geomagnetic activity is high, the tail current moves earthward, where it ‘hinges’ to the magnetic equator; in summer or winter the earthward motion therefore causes the tail sheet to move below or above its normal position. McPherron and Barfield found that a superposed epoch analysis of the synchronous orbit magnetic field for winter lacked the large diurnal variation seen in summer and autumn, as might be expected if the satellite was in a plane of magnetic symmetry.

A further seasonal discrepancy between the Tsygankenko model field and the calculated field which has already been mentioned concerns the value of theta at midnight close to the equinoxes. This anomaly revealed an oversight in the ground data handling. At the points in question, theta was calculated from the data to be 0° where the model predicted it would be 40-50°. An example of this
Figure 4.10: GEO theta (in degrees) versus local time (hours) for each season. The solid line shows the values calculated from the SEM-2 data. The dotted line shows the Tsyganenko model values.

Figure 4.11: GSM phi (in degrees) versus local time (hours) for each season. The solid line shows the values calculated from the SEM-2 data. The dotted line shows the Tsyganenko model values.
is shown in figure 4.13. This drop is a sudden event (one or two data points) on a curve which otherwise fits the model well. It occurs repeatedly at these times of year and so is unlikely to be a calculation anomaly. At the time of these theta points, phi is experiencing its sudden flip from 360° to 0° at midnight. Although there is no particular value of flux associated with these points, a closer look at the polar-azimuthal count arrays gives a clue to the cause. The azimuthal array only has zero at these points, whereas the polar and energy arrays are unaffected. The azimuthal measurement is the first thing done in the experiment cycle. At equinox at these times, the satellite is in eclipse. Although the SEM-2 timing is from an internally generated pulse and not a sun sensor, the timing of the pulse is referenced to the spacecraft time at the start of each experiment cycle, so that the absolute spin time can be implemented when the azimuthal data are processed. When the spacecraft time is referenced on the ground (in the archiving software) then the zeros are created. Up until this point, the azimuthal raw counts do not contain the zeros points. The interruption of the spacecraft's sun-sensor interferes with the spacecraft housekeeping data words, so that the timing later used as a reference for the SEM-2 spin causes the error. The rest of the SEM-2's experiment cycle, i.e. polar and energy measurements, are unaffected. Because the spin timing of the polar and energy measurements (over 10 spins) is not critical as for the azimuthal measurement (the first spin only), they do not need to reference the spacecraft time.

4.5 Summary

- The difference between the Tsyganenko model field direction and the direction calculated from the observed particle distribution shows dependence on local time, $K_p$ index and time of year.

- The GSM spherical polar coordinates theta for the SEM-2 calculated field direction was found to lie within 10° of the Tsyganenko model equivalents in 64% of cases overall for one whole year of data. Phi was found to lie within 20° in 55% of cases.
Figure 4.12: The effect of season and magnetic activity on the magnetic equator at synchronous orbit (McPherron and Barfield, 1980). Magnetic disturbance distorts the equator by moving the tail current earthward, so in summer the satellite's effective magnetic latitude is increased and in winter it is decreased. Earth's spin axis is shown by $\Omega$, dipole axis by $\mu$.

- The data gave a better fit to the Tsyganenko model on the dayside than on the nightside.
- The degree of field line stretching in the tail is underestimated at high $K_p$ and may be overestimated at low $K_p$ by the Tsyganenko model.
- The calculated symmetry axis is closer to the Tsyganenko model field direction when the degree of anisotropy is high.
- Asymmetry about noon was observed in the SEM-2 data which is not present in the Tsyganenko model.
- GSM phi showed significant differences between the data and the model on the nightside. Two factors are thought to be important here: the less predictable geometry of the tail compared to the dayside field, and the contribution from low anisotropy points.
- The differences between the model and data showed seasonal dependence.
Figure 4.13: A consistent difference between the Tsyganenko model and SEM-2 calculated field shows up a SEM-2 software problem: for several days around the equinoxes, the midnight value of theta drops suddenly to zero. These are eclipse points (see text).
Approximately equal but opposite differences for summer and winter suggest a weakness in the model.

- At the equinoxes a data handling problem, due to the satellite being in eclipse, caused the measured value of theta to drop briefly to zero from a value of $\sim 50^\circ$.

### 4.6 Conclusion

The method used to calculate the symmetry axis of the particle distribution, assumed to be the magnetic field direction, has been validated by the generally good agreement of the SEM-2 data and the Tsyganenko T89c magnetic field model. The advantage of using a large data set (a full year of 30-minute averages) is that persistent differences between the model and the data can be identified. The Tsyganenko magnetic field model has also been shown to provide a good approximation of the Earth's magnetic field, although a few weaknesses of the model have been identified, as discussed in section 4.2. The one year of SEM-2 data used in this analysis, 1994, was a particularly active year. Since the Tsyganenko model has been shown to be less accurate at high $K_p$, this was a good test for the model.

The fact that the difference between the Tsyganenko model field direction and the field direction calculated from the SEM-2 data appears to vary in specific ways with local time, $K_p$ index and season, indicates that there are consistent differences between the modelled field and real field which are dependent on these factors. The dynamic nature of the true field compared with a relatively static model inevitably means a degree of scatter on the predicted field would be expected, and this is observed, but in addition the differences show persistent local time, $K_p$ and seasonal dependence.

The largest discrepancy between the observed field and the model is the tail field. Instead of varying smoothly the field direction was seen to flip between opposite field directions repeatedly. As noted by Jordan et al. (1992), one weakness of the Tsyganenko T89 model is its inability to react swiftly to sudden changes,
since, with the exception of dependence on a three-hourly index \((K_p)\) it is not truly dynamic. The degree of stretching of the tail is seen to be underestimated above moderate \(K_p\), and at lowest \(K_p\) there is a suggestion that the tail is over-stretched. The same conclusion was reached by Weiss et al. (1997).

The differences between the data and the Tsyganenko model showed seasonal dependence, by which roughly equal but opposite differences arose for summer and winter which did not appear at the equinoxes. The Tsyganenko 89 model was compared with data from each season by Thomsen et al. (1996) but they reported only modest seasonal differences. The seasonal differences arise because of the orientation of the Earth’s spin axis, and therefore magnetic axis, to the Sun-Earth line.

A seasonal anomaly in the SEM-2 data occurs at the equinoxes when the satellite is in eclipse. The polar angle theta was observed to go briefly close to zero from a value of \(~50^\circ\). The reason for this is that at the start of each experiment cycle, a one-spin azimuthal measurement is made and the SEM-2’s internal timing is referenced to the spacecraft timing. The eclipse is therefore responsible for interrupting the spacecraft’s Sun sensor and causing the timing problem, which is only imposed on the SEM-2 data on the ground in the unpacking software. Similar points where theta approaches zero at midnight close to the equinoxes have been found in the GOES-6 magnetometer data.

Reversals of the tail field (seen in GSM phi flipping by \(180^\circ\)) were seen in the data more around the equinoxes than the solstices. Some of these points will undoubtedly be due to the equinox values addressed in the last paragraph. But they may also describe a dynamic feature of the tail near the equatorial plane, since at the equinoxes the spacecraft is closer to the central tail (due to the streamlining effect of the solar wind on the magnetosphere).

Overall, the solstices were fitted better by the model than the equinoxes were. In addition, winter was found to be better fitted than summer, especially in terms of phi. This can be explained by the fact that the latter part of 1994 was less active. Several large magnetic storms in the first half of 1994 led to high \(K_p\) values: there were hardly any \(K_p\) values in the \(> 6\) category for winter.
It was shown in the analysis that high \( Kp \) values were less well fitted by the model so this explains why summer was a worse fit. The outstanding question is why the equinoxes are not as well modelled as the solstices. The erroneous zero points seen in eclipse, although having large errors (40-50°) only account for one or two data points per day, and local time analysis of the seasonal dependence shows that the differences between the data and the model are not restricted to midnight but cover a range of local times. The difference is undoubtedly to do with some aspect of seasonal distortion of the magnetosphere. It is possible that the streamlining of the tail is handled better for the extreme cases of the solstices than elsewhere.

The calculated symmetry axis direction was found to be closer to the Tsyganenko model field direction when the anisotropy was high. This points to a weakness in the method for calculating the symmetry axis of the SEM-2 electron distribution. The algorithm has difficulty choosing an axis when the distribution is near-isotropic. Thomsen et al. (1996) found the same result when deriving the symmetry axis from parallel and perpendicular electron temperatures.

Overall several deficiencies in the Tsyganenko model are suggested by this analysis. There are seasonal dependences which the model does not adequately present, which may be due to insufficient influence of the dipole tilt angle, and also to the bending of the magnetic equatorial plane by the passage of the solar wind. The tail field is dynamic in nature and therefore hard to model, but the degree of stretching in the model generally does not change sufficiently with \( Kp \). On the dayside, the field was found to be asymmetrical about noon, whereas the Tsyganenko model has near symmetry. This feature can be seen in GOES-2 data in Tsyganenko’s (1989) paper but was not commented on, since that analysis concentrated on the tail field. McComas et al. (1993) were able to demonstrate this asymmetry using the first simultaneous pre-noon and post-noon observations from the MPA on board three geosynchronous satellites.

The overall conclusion of this analysis is that the method used for finding the field direction from the SEM-2 data, namely, finding the symmetry axis of the particle distribution, is generally successful. The Tsyganenko model is also
validated by the results. The analysis has also helped to identify some weaknesses in the axis calculation method, and in the Tsyganenko model itself. Since the data used were from a particularly active year the discrepancies between the model and the data may well be worst case.
Chapter 5

Local Time Dependences:
Energy Dependence and the Effect of Geomagnetic Activity

This chapter focuses on the local time behaviour of the geosynchronous electron flux. Diurnal variations are to be expected at geosynchronous orbit due to the asymmetry of the geomagnetic field and the nightside occurrence of substorm particle injections, and have been observed with many satellites and for many particle energies. For the SEM-2 energy range, two distinct local time variations arise because of the different behaviours of the two populations described in Chapter 2.

A sample of total (42.9-300keV) flux for approximately 12 days in 1989 is shown in figure 5.1(A). The flux in the five differential energy ranges is shown in order of increasing energy in (B) to (F). Plot (G) shows the $K_p$ index for the same time period. The satellite was at 0° longitude at this time so the x-axis tickmarks correspond to local midnight. The total flux shows greatest similarity to the lower energy flux and has marked differences from the highest energy range. This is because most counts occur at the lower energies so the total energy range is dominated by these low energy counts.

A diurnal variation is clearly visible in all energy ranges. In the lower energy ranges (plots (B) and (C)) a sharp increase in flux is often seen close to midnight.
This is sometimes preceded by a flux dropout, as is seen before the injection at day 144, or a less severe flux decrease, as on day 151. The large flux peaks are often seen to recur on timescales of a few hours but their amplitude, which peaks within a few hours of midnight, generally decreases the later in local time they are observed. The highest energy range (plot (F)) shows a different diurnal variation. The daily flux profile from day 147 onward is more symmetrical about noon and shows little influence of the large flux peaks seen in the low energy flux.

The low energy flux is dominated by the occurrence of substorm injections. The flux level is enhanced by a factor 5-10 above its level before the substorm on day 143. Electrons of energy 30-80keV observed by ATS-5 and ATS-6 showed maximum flux at LT 0200-0400 and minimum at 1500-1800 [Garrett et al., 1981], with a difference of up to an order of magnitude. The high energy flux is less affected by the injections and is dominated by the magnetic field: the diurnal variation here is due to trapped particle motions in the distorted field geometry. Explorer 14 observations of >1.6MeV electrons in the outer zone showed a local time dependence with maximum electron intensity close to midday and minimum close to midnight (Frank, 1965b). Low altitude observations from Injun I showed that the flux of energetic trapped electrons was two orders of magnitude higher on the dayside than the nightside (O’Brien, 1963). ATS-1 observations of 0.4-2 MeV electrons showed the diurnal variation to be dependent both on energy, with higher energy electrons undergoing a more severe nightside decline, and on $K_p$, with the diurnal variation being larger at high $K_p$ (Lanzerotti et al., 1967).

Before day 147 the profile is more complex. A large enhancement of the $K_p$ index occurred in the afternoon and evening sectors of day 143. At the time of the enhancement the flux in all energy ranges drops out. A few hours later a sharp flux peak occurs in the lower energies. The highest energy flux does not show the same increase following the dropout, but about two days after the $K_p$ peak, the high energy flux is greatly enhanced and remains so for several days.

Of the three intermediate energy ranges, all show some sign of the injections, the magnitude of which decreases with increasing energy. This feature is also manifested in the way the local time profiles become smoother with increasing
Figure 5.1: (A): A sample of total flux (42.9-300keV) data from 1989 illustrating the diurnal variation. (B)-(F) the same time series in the five differential energy ranges, in order of increasing energy, E5-E4 first through to lowest E1. (G) The $K_p$ index for the same time period.
energy. Only the E2-E1 flux (panel E) shows some resemblance to the local time behaviour seen in the highest energy range.

The $K_p$ index quantifies the level of geomagnetic disturbance. The large rise in $K_p$ on day 143 is associated with the large injection which follows at midnight. The highest energy range generally shows only the largest injections, and following the dropout, whereas the low energy flux is enhanced well above its pre-dropout level, the highest energy flux is not. The high energy flux exhibits a delayed response to geomagnetic activity which will be discussed further in the next section.

The low time resolution SEM-2 archive data is ideal for the study of local time behaviour because it consists entirely of successive 30-minute averages. In this chapter, statistical profiles of the local time behaviour of electron flux, spectral index and anisotropy are produced, and the effect of geomagnetic activity is examined using the $K_p$ index.

5.1 Energy Dependence

Figure 5.2 is a scatter plot showing low time resolution E5-E4 flux data versus local time. Each data point is a 30-minute average and the whole data set (1988-1995) is plotted.

The range of flux values observed across local night through to morning is much greater than is seen on the dayside, particularly post-noon. The highest flux occurs from around 2200 LT to dawn. These local times are where the geosynchronous population is most variable, when substorm injection events can cause order of magnitude increases in the low energy flux. In addition, the greatest concentration of low flux values occurs either side of midnight. Therefore it is important to know, in addition to the average situation at geosynchronous orbit, what range of flux might be expected and also what the probability is of observing a certain value of flux. It is clear from figure 5.2 that all of these aspects depend on the local time of the observation.

The large variability seen around midnight is produced by injections of parti-
Figure 5.2: Scatter plot of low resolution E5-E4 flux (42.9-59.4keV) versus local time

Figure 5.3: Statistical plot of low resolution E5-E4 flux (42.9-59.4keV) versus local time (see text)
cles during magnetospheric substorms, preceded by flux decreases in the stretched tail field. As a result of reconfigurations of the geomagnetic field, particles are accelerated earthward from the magnetotail. These injection events occur near midnight and the injected electrons then drift eastward, toward dawn (Pfitzer and Winckler, 1969), the drift being energy dependent such that higher energy particles drift faster. The amplitude of the event decreases the later it is observed in local time (Lezniak et al., 1968 and Parks et al., 1968). This is because the particles are observed from an energy range, and will drift at different speeds according to their energy within the range. Thus a fairly sharp peak seen close to the injection site will ‘spread out’ at later local times.

Statistical plots are produced to show the variability in the flux data. The same data as in figure 5.2 are presented in a different form in figure 5.3, which consists of four curves. The upper curve marks the level below which 95% of observations occurred, the lower curve marks the level above which 95% of observations occurred and the middle two curves show the median ± 5% of observations. The upper and lower curves are henceforth referred to as the 95% and 5% curves respectively. These plots are similar to those produced by Baker et al. (1981a) from early CPA data except that their plots show the average, 10% and 90% flux. Plotting the data in terms of the flux level at which a specific proportion of the observations occur demonstrates the variability in the data and how it depends on local time. It also presents the probability of observing a particular flux value at any local time: e.g. at dusk the probability of observing a flux of electrons in the E5-E4 energy range above 50000cm⁻²sr⁻¹s⁻¹keV⁻¹ is only a few percent.

The E5-E4 flux therefore exhibits strong local time behaviour with greatest variability occurring either side of midnight and peak flux occurring in the post-midnight quadrant. Figure 5.2 shows these high flux observations start as early as 1900 LT, though figure 5.3 shows that in general they are observed from 2200 LT, so although large substorm injections may be observed as early as ~1900 LT at geosynchronous orbit they are generally seen from 2200 LT onwards.

The fact that the 95% line in the statistical plot of E5-E4 flux peaks at ~0400 LT does not mean that the largest injections occur here. It shows that
most injections are seen here. This is illustrated by figure 5.4 which shows that the largest peaks are seen earlier in local time. As was seen in figure 3.2 when peaks are observed later in local time, they are energy dispersed. Figure 5.5 shows the differences that exist between local and dispersed injections. This plot is a superposed epoch analysis of peaks in the high time resolution E5-E4 data. The solid line shows peaks observed between 00-01 LT: at each energy, the peak occurs at about the same time (actually, this example may show slight energy dispersion, but since the time resolution of the data is about 10 minutes, it is not conclusive). The dotted line shows peaks observed between 05-06 LT: energy dispersion is clearly seen. The lowest energy peak is at zero on the $x$-scale whereas the E2-E1 peak is at -0.5 hours. This occurs because higher energy particles drift faster. The nature of the curves either side of the peak is related to the average diurnal variation.

Figure 5.4: Distribution of peaks in E5-E4 flux in local time (units of flux are $\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$)

The fact that most peaks are observed a few hours after local midnight, combined with the energy dispersion signature, shows that the injections are generally produced beyond geosynchronous orbit. By the time they reach geosynchronous
orbit, the electrons have already begun drifting eastward. Since larger, non-dispersed peaks are seen closer to midnight, this points to the injection site itself being close to local midnight. Another interpretation is in terms of the injection boundary model (e.g. Mauk and Meng, 1983), which would mean that there is a most common crossing point of geostationary orbit by the injection boundary. However, Moore et al., (1981) state that injections are often observed well after crossing the boundary. Therefore the point where the injections are most commonly observed does not give the boundary location.

Figure 5.3 contains a lot of information about substorm injections. It shows where the injections occur and characterizes the effect of subsequent drift and energy dispersion: if an injection is observed at 0400 LT with a flux peak height a little over \(10^5\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\), then the same event when it drifts round to dusk will be observed with a flux peak of around \(4 \times 10^4\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\).

The lowest flux points shown by the 5\% curve, which occur from a few hours before midnight until just after, indicate the large flux decreases seen prior to injections. In figure 5.3 it can be seen that the minimum flux is reached earlier
Figure 5.6: Statistical plot of low resolution E4-E3 flux (59.4-90.7keV) versus local time (see text)

Figure 5.7: Statistical plot of low resolution E3-E2 flux (90.7-134.9keV) versus local time (see text)
Figure 5.8: Statistical plot of low resolution E2-E1 flux (134.9-201.8keV) versus local time (see text)

Figure 5.9: Statistical plot of low resolution E1 flux (201.8-300keV) versus local time (see text)
than the peak flux, at 2200-2300 LT. These low points are due to dropouts, which are commonly seen prior to injections (dropouts are discussed at the end of this section).

Figures 5.6 to 5.9 show similar statistical plots for the E4-E3 to E1 energy ranges. The shape of the 95% and median lines shows the decreasing influence of the injections as energy increases, changing from the morning sector peak to roughly symmetrical about noon. The position of the peak flux appears to shift to later local time with increasing energy as the trapped particle influence becomes comparable with that of the injected electrons. The morning peak is seen up to the E2-E1 energy range but not in E1: this shows that the electrons >200keV are generally not injected. This is also seen in the decreasing fraction of the total range of the observations occupied by the injected flux peak. In the low energy ranges E5-E4 and E4-E3, the injected flux is large and may cause rapid order of magnitude increases, whereas for the higher E2-E1 energy range the injections have a smaller effect on the flux local time profile determined by the magnetic field. The shape of the local time profiles represents a superposition of the substorm injections and the trapped population in each case.

The average and median flux of each energy range are plotted in figure 5.10. This plot contains several pieces of information. At each energy and for all local times, the average flux is higher than the median flux. This shows that the higher half of flux observations have a greater range than the lower half. At low energies the high flux events, i.e. the injections, have the dominant effect on the average. Also, for lower energies the difference between the average and median flux is dependent on local time, being larger from shortly before midnight until mid-morning. This is due to the fact that substorm injection flux is dependent on the local time of observation. This is not the case at the highest energy, where the difference between the average and median flux shows little dependence on local time.

In figure 5.9, the statistical plot for the E1 flux, there is little evidence of the injections signatures which dominate the E5-E4 flux. The distribution is roughly symmetrical about a point close to noon (1000-1100 LT at all flux levels).
Figure 5.10: Average (solid lines) and median (dotted line) flux versus local time, for total flux and flux in each energy range.
with the greatest variability being on the nightside, where significantly lower flux may occur. A similar diurnal variation was observed for 0.4-2 MeV electrons by ATS-1 (Lanzerotti et al., 1967). At this energy the magnetic field configuration has the dominant effect on the electrons. In the outer radiation belts the field departs significantly from a dipolar configuration, due to the actions of magnetospheric currents: on the dayside the Chapman-Ferraro current which flows eastward along the magnetopause tends to increase the magnetic field strength inside the magnetopause, while on the nightside the tail current is westward which reduces the field. The magnetosphere is asymmetric, compressed on the dayside and drawn out into a tail on the nightside. The asymmetric field has different effects on particles depending on their pitch angles. A particle with 90° pitch angle (i.e. mirroring at the equator) drifts on constant magnetic field contours. This means it will be closer to Earth on the nightside than the dayside, as can be seen from figure 5.11 which shows contours of constant magnetic field strength in the Earth’s equatorial plane (Fairfield, 1968).

Figure 5.11: Experimentally determined average contours of constant magnetic field intensity in the equatorial surface (Fairfield, 1968).
Figure 5.12: Drift paths in the outer magnetosphere

But for particles with small $\mu$ (equation 1.11) and large $J$ (equation 1.14), the magnetic field experienced on its bounce path is smaller than its mirror field. If $J$ is large then $I$ (equation 1.17) tends to $\sqrt{B_{m}S}$ (where $S$ is the bounce path length). If $\mu$ is conserved then $B_{m}$ must be conserved, which means that for constant $I$, $S$ is also conserved. Due to the stretching of the field, lines of equal length will take the particle further away from the Earth on the nightside than the dayside (Kivelson and Russell, 1995).

Figure 5.12 shows a set of drift paths calculated for 0000UT at the 1989 winter solstice in the plane of the geographic equator. The drift paths were calculated by tracing large sets of field lines and then conserving $B_{m}$ and $I$ for a 45° particle. The model field used was the Tsyganenko 89c external magnetic field model combined with the IGRF internal field. To access the field models the GEOPACK suite of programs, written by N. A. Tsyganenko and M. Peredo, was used. It contains subroutines to trace field lines and to make the necessary coordinate transforms (since the Tsyganenko model works in GSM coordinates and the IGRF model in GEO).
The inputs for the Tsyganenko field model are a parameter describing the level of the $K_p$ index, the dipole tilt angle, and the cartesian GSM coordinates of the point. The outputs are then the GSM cartesian components of the external field. For the IGRF model the inputs are the year and GEO cartesian coordinates of the point, and the outputs are the GEO cartesian components of the internal field. The coordinate systems are described in Appendix B.

Field lines were traced at 16 longitudes and at radial distances between 2.0 and 10.0$R_e$ in increments of 0.01$R_e$. The chosen range of $K_p$ was 2- to 2+. The starting point for each line was the magnetic equator at each specified radial distance. The magnetic field was found at this point from the Tsyganenko and IGRF models, allowing the mirror point magnetic field to be calculated, using the conservation of the first adiabatic invariant:

$$\mu = \frac{mv^2}{2B} = \frac{mv^2 \sin^2 \alpha}{2B}$$

(5.1)

where at any point on the field line:

- $\mu$ is the first adiabatic invariant
- $m$ is particle mass
- $v$ is velocity and $v_\perp$ its component perpendicular to the field
- $B$ is the magnetic field intensity
- $\alpha$ is the pitch angle.

So, if $\mu$ is conserved:

$$\sin^2 \alpha = \frac{\sin^2 \alpha_M}{B_M}$$

(5.2)

$$\sin^2 \alpha = \frac{B}{B_M}$$

(5.3)

Here, $\alpha_M$ and $B_M$ are the pitch angle and magnetic field intensity at the mirror point (so $\alpha_M$ is 90°).

Starting with a 45° particle at the magnetic equator, the mirror field was calculated for each longitude and radial distance. To follow an electron on its drift around the Earth, its mirror field value must be conserved and so must the value of $I$ (the second adiabatic invariant for static fields), given by:

$$I = \int_{m1}^{m2} \left[1 - \frac{B(s)}{B_M}\right]^{\frac{1}{2}} ds$$

(5.4)
which is computed along the field line between the mirror points $m_1, m_2$.

The method for finding drift shells was therefore to produce a vast number of field lines at different longitudes and different radial positions in small increments. Then it is a case of finding matches between lines at each longitude for $B_M$ and $I$. The resulting set of field lines describes a drift shell. Once the set of field lines was identified, the point at which the lines cross the geographic equator was identified and this therefore produced the drift paths for geostationary orbit. The errors in producing this plot are small since a fine radial grid was used. Perhaps the largest error comes from the integration of $I$ along the field line, but the number of steps along the field line is such that this error is negligible on the scale of the plot.

This figure illustrates how the paths of particles are brought closer together on the dayside but drawn apart on the nightside, leading to more intense flux being observed on the dayside. The drift shell geometry changes with season, because the spin axis orientation changes with respect to the Sun-Earth line (and the magnetic axis is at $\sim 11^\circ$ to the spin axis), and with UT since the magnetic axis rotates with the Earth’s spin.

Baker et al. (1981a) produced sets of plots of the early Los Alamos data showing the average flux, 10% and 90% flux versus local time, one set each for satellites 1976-059 and 1977-077. Some of these plots are shown in figure 5.13. There are differences between the two sets of plots but they are not large and since they apply to different epochs and different magnetic latitudes some variation would be expected. The lowest energy plots have a resemblance to the SEM-2 low energy plots: although the LANL plots are integral flux, the low energy counts dominate. The highest LANL energy range shown resembles the E1 plot.

The comparison between the and SEM-2 plots is not straightforward since the data from the LANL plots is not available. Also since those plots were integral flux ($IF$) in $cm^{-2}sr^{-1}s^{-1}$ then in order to compare them quantitatively with the
Figure 5.13: LANL results from satellites 1976-059 (left) and 1977-077 (right), for three integral energy ranges
SEM-2 differential flux \((DF)\) plots in \(\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\) a conversion is necessary:

\[
DF(E1 : E2) = \frac{IF(E1) - IF(E2)}{E2 - E1}
\]

where \(E2, E1\) are the higher and lower energy levels respectively.

The energies of the Los Alamos plots shown are \(>30\text{keV}, >140\text{keV}\) and \(>200\text{keV}\). Therefore it is possible to make a comparison between the SEM-2 \(E2-E1\) range \((134.9-201.8\text{keV})\) and the differential flux calculated from the LANL \(>140\text{keV}\) and \(>200\text{keV}\) ranges using equation 5.5.

As a high energy example, the average flux at noon for satellite 1976-059 is \(\sim 4\times10^5\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\) at \(>200\text{keV}\), and \(\sim 6\times10^6\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\) at \(>140\text{keV}\). This gives a rough figure of \(3 \times 10^3\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\) for the differential flux. From figure 5.10 the noon average of the SEM-2 \(E2-E1\) range is \(5 \times 10^3\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\). For a low energy example, the noon average at \(>30\text{keV}\) is \(3 \times 10^6\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\) which combined with the value at \(>140\text{keV}\) gives a differential flux of approximately \(2 \times 10^4\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\). This energy range corresponds roughly to the SEM-2’s three lowest energy ranges, the noon average for which is between \(1 \times 10^4\) and \(3 \times 10^4\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\). A similar degree of agreement is obtained for the limits on the two sets of plots. Although the comparison is coarse it is clear that the greater than order of magnitude differences seen by Rodgers et al. (1993) do not exist in this version of the LANL data set.

In a recent paper, Ivanova et al. (1996) argued that the SEM-2 low energy population pattern of low flux seen pre-midnight, followed by the higher flux in the morning sector, is due to spatial structure rather than repeated injections. The reason for their argument is that density profiles of 0.1-13keV electrons measured by GORIZONT-91/2 show repeated increases which correspond with the timing of the statistical curves for the \(E5-E4\) range shown in figure 5.3. However, although an average flux versus local time profile exhibits this local time behaviour, even for quiet time averages, this pattern is not seen during a time sequence devoid of injections. This was shown in Szita et al. (1996). An example of this behaviour is shown in figure 5.14. The flux decreases monotonically as if there is no local time dependence at all. If there is a local time dependence its effect cannot be seen above the decrease in flux in the absence of injections. In
Figure 5.14: E5-E4 flux in the absence of injections: $K_p$ was $\leq 1$ for more than 2 days in this example.

The total absence of injections, the familiar local time profile of the low energy electrons is not seen. This in turn supports the conjecture that it is the injections which are predominantly responsible for the local time structure at these energies, rather than a spatial structure with enhanced nightside density. If this was the case then in quiet times a local time variation would still exist, but when injections do not occur, the E5-E4 flux simply decreases.

The data from the statistical plots provide a model of the observed range of flux values in this energy range at all local times: 90% of all flux observations can be expected to lie within the two outer curves. The statistics here are good since the models are based on a large data set. However, since the radiation belts, and geomagnetic activity itself, show solar cycle dependence, these models may change between solar maximum and minimum.

The high flux observed at low energies in substorm injections causes the low energy 95% curves to differ significantly from the high energy 95% curves. The same effect can be seen to a lesser extent in the median curves. This means that
more than half the time, the low energy flux is influenced by injection events. The 5% curves in each of the five statistical plots have a similar shape and this is because the low flux observations in each case define the magnetic field structure: the dropouts and growth phase decreases depend on the field, not on the injections, so all energies are affected similarly.

**Dropouts**

Flux decreases or dropouts are commonly observed at geosynchronous orbit during the substorm growth phase, as a substantial decrease in the flux of particles with energies above several tens of keV. They are due to a combination of effects associated with increased geomagnetic activity. The displacement of the radiation belts earthward during the substorm growth phase (Bogott and Mozer, 1973, and Sauvaud and Winkler, 1980) has been observed in particle flux gradients which indicate the inward motion of the radiation belt particles (Walker et al., 1976). Because 6.6\(R_E\) is past the peak intensity of the outer belt, this means that a geosynchronous spacecraft sees reduced flux. Another contribution to the growth phase dropouts comes from the thinning of the near-earth plasma sheet (Reeves et al., 1993, and references therein). During the growth phase the cross-tail current intensifies and the field lines become more taillike. Therefore, a satellite which is not on the magnetic equator finds itself on field lines which extend further down the tail: it is connected to regions of much weaker particle flux. In the course of a substorm, the field will dipolarize and the injection of particles means that the flux observed after the dropout will be higher than before.

Figure 5.15 shows two sections of flux data and their corresponding anisotropy index values. The anisotropy index has been smoothed with a 5-point boxcar. These plots are taken from 1989 when in both cases Meteosat P2 was within 1° of midnight so local midnight is at the day tickmarks. The anisotropy index shows a pancaked dayside distribution and a nightside distribution which is isotropic to field-aligned. In plot (A) the flux at all energies drops to very low levels. The \(K_p\) index at the time was \(\geq 8\) throughout the event. This was therefore a time
Figure 5.15: Flux ‘dropouts’ and corresponding anisotropy index; x-axis tickmarks are at local midnight. (A) shows a magnetopause encounter, (B) a nightside dropout.
of greatly enhanced geomagnetic activity and the likely explanation is that the plot shows a magnetopause encounter spanning several hours of local time. Also on this plot are smaller flux decreases preceding injections near midnight on days 71 and 74. Plot (B) shows a deep nightside dropout on day 116. Again there are smaller flux decreases preceding injections near the three other midnights.

During multiple substorms each injection may be preceded by a dropout. In figure 5.15 there is an example of this on day 74. Multiple-onset substorms were observed by Nagai et al. (1983); a series of flux peaks occurring within a dropout were interpreted as small substorms which do not cause major reconfigurations of the geomagnetic field. This explains a feature seen in the 5% line on the statistical plots, namely that the low flux curve experiences an extended minimum around midnight.

Figure 5.16: Local time dependence of flux decreases for SEM-2’s highest and lowest energy ranges.

The anisotropy index shows different behaviour for different examples of flux decreases. Smaller decreases preceding injections show positive anisotropy index values. This corresponds to a taillike field producing a field-aligned distribution in the growth phase. The injections which follow cause the anisotropy index to
reach zero or lower. In contrast, during the magnetopause encounter in (A) and the deep dropout in (B), the anisotropy index remained fairly isotropic.

The growth phase decreases in flux are often not complete dropouts, as may be observed when encountering the magnetopause on the dayside or the low flux tail lobes. Flux in these cases drops to low values but real counts are still being recorded. In the case of a magnetopause or lobe encounter the flux drops much lower, effectively to the background levels of the instrument.

In figure 5.16 the distribution of low flux points have been plotted as a function of local time. The thresholds have been chosen in an attempt to classify flux decreases as are seen prior to injections (E5-E4, E1 flux below 5000, 100 cm$^{-2}$sr$^{-1}$s$^{-1}$keV$^{-1}$ respectively), and to classify dropouts as might be seen on the dayside during a magnetopause encounter (E5-E4, E1 flux below 100, 10 cm$^{-2}$sr$^{-1}$s$^{-1}$keV$^{-1}$ respectively). Of course the flux seen in dropouts varies widely but it is hoped that only complete dropouts fall into the more restrictive category.

In the upper panel of figure 5.16 the E1 low flux points show a small asymmetry about midnight, with more low flux points before midnight than after. The E5-E4 flux shows greater asymmetry about midnight, with less low flux points seen from midnight to 0800 than between midnight and dusk. These nightside low flux points are seen when the spacecraft is beyond the higher flux seen in the radiation belt, or passes into the plasma sheet. This happens when geomagnetic activity is high and the radiation belts move closer to Earth. The E5-E4 flux is less affected post-midnight since when substorms occur, injections raise the particle flux.

The lower panel of figure 5.16 is scaled to show the dayside low flux points at lower flux thresholds, chosen to pick out possible magnetopause encounters. When the magnetosphere is strongly compressed, geosynchronous orbit may pass outside the magnetopause on the dayside. There are few magnetopause encounters in the SEM-2 data, but due to the low time resolution of the data, any brief encounters (less than several minutes) will be missed. Figure 5.16 shows that the magnetopause encounters occurred generally between 0800-1500 LT, so there is a slight asymmetry about noon. The high energy flux is the first to decrease and
Figure 5.17: Adiabatic modelling of a flux decrease (see text)
the last to recover during complete dropouts such as magnetopause encounters. The reason for this is that a higher energy particle has a bigger gyroradius so it will encounter a boundary before a lower energy particle on the same field line with the same pitch angle.

A way of testing to see if the dropouts are due simply to the stretching of the magnetotail is to make use of Liouville's theorem, which states that a trapped particle's phase space density is conserved along its trajectory, and the conservation of the adiabatic invariants. Baker et al. (1982b) used this method for a severe dropout preceding an injection to show that non-adiabatic effects were important. Even though their example was not a complete dropout (flux stayed above background levels) it could not be explained simply by the effect of the magnetic field changes on trapped electrons. This was interpreted as the spacecraft passing into the plasma sheet.

Attempts to use the same method for the SEM-2 data to show a conservation of phase space density have proved inconclusive. An example is shown in figure 5.17. The first panel shows magnetic field intensity: the problem here of course is that there is no in situ magnetic field data, and the closest readily available was from GOES-7 a few hours away. This obviously is not ideal but it is the best option available. The next panel shows E5-E4 flux. The method is first to fit the energy spectrum for the low energy population with a Maxwellian as in Chapter 2. Next, the value of $\mu$ is chosen: if the process is adiabatic, $\mu$ is conserved. A range of suitable $\mu$ values is (Baker et al., 1982b):

$$\frac{E_{\text{min}}}{B_{\text{max}}} \leq \mu \leq \frac{E_{\text{max}}}{B_{\text{min}}}$$ (5.6)

The problem is simplified for 90° particles since $p = p_{\perp}$, so it is assumed that the omnidirectional flux available is representative of the 90° flux. For each point in time then, a value of relativistic momentum can be found for each magnetic field value, using the relation for the first adiabatic invariant:

$$p^2 = 2m_0\mu B$$ (5.7)

From this the kinetic energy can be found and therefore, using the fit of the
energy spectrum, the flux \( j \) is found to be used in:

\[
f = \frac{m}{v^2} j
\]  

This gives phase space density \( f \), shown in the third panel. The two traces are for the two limiting values of \( \mu \) chosen. The result suggests that the injection is not adiabatic but that the dropout may be, since the value of phase space density just before the injection is close to that before the dropout started, but what happens prior to the flux dropout is not clear. Several examples were tried but no 'typical' signature emerged.

The problem is that the same particles are not seen by a geosynchronous spacecraft throughout the dropout. The more the tail is stretched, the wider the range of field lines that are sampled on the orbit. The approximation of assuming that the same particles are seen along the orbit does not hold. What is needed to solve this problem properly is measurements at different radial distances used in conjunction with a model magnetic field. Such a method was employed by Kim and Chan (1997) in a study of storm time relativistic electron flux using CRRES observations. Relativistic electron flux decreases by orders of magnitude in the main phase of magnetic storms, then increases to well above the pre-storm value. Their results showed that whilst adiabatic effects could account for a significant portion of the flux decrease, non-adiabatic effects were clearly contributing.

5.2 Dependence on Geomagnetic Disturbance

In this section the local time dependence of the flux data is examined in terms of the level of geomagnetic disturbance as defined by the \( K_p \) index. \( K_p \) is a three-hourly index which provides a measure of the level of planetary magnetic disturbance. It has values ranging from 0 at low disturbance to 9 at high disturbance on a logarithmic scale, using + and − for between-integer values, i.e. 0o, 0+, 1−, 1o, 1+, 2−, etc..

The value of \( K_p \) is calculated from observations by 12 stations between geomagnetic latitudes 48° and 63°, which also provide good longitudinal coverage. Each station records the deviation of a magnetic trace from the quiet situation
for that particular observatory. Allowing for seasonal variations at each station, the results are then averaged for all the observatories to produce the planetary index (Parks, 1991 and Knecht and Schuman, 1985). The $K_p$ index is widely used as an indicator of geomagnetic activity and is readily available.

The flux data were divided up into four groups according to the value of $K_p$ at the local time of each flux measurement, ranging from low to high magnetospheric activity. The four $K_p$ ranges chosen were 0o to 1+, 2- to 3+, 4- to 5+ and 6- to 9o. The number of good data points that fell into each of the four $K_p$ ranges is given in table 5.1. Of the four chosen $K_p$ groups, most observations occur in the 2- to 3+ group and least in the 6- to 9o group. Although the highest $K_p$ group only contains 4% of the data, this is still almost 100 points per local time bin.

<table>
<thead>
<tr>
<th>$K_p$ range</th>
<th>no. of data points</th>
<th>% of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>0o to 1+</td>
<td>31015</td>
<td>26</td>
</tr>
<tr>
<td>2- to 3+</td>
<td>57478</td>
<td>49</td>
</tr>
<tr>
<td>4- to 5+</td>
<td>25349</td>
<td>21</td>
</tr>
<tr>
<td>6- to 9o</td>
<td>4471</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 5.1: Total flux points in $K_p$ ranges

Results

Statistical plots were produced as described in section 5.1 for each energy range divided into the four $K_p$ ranges defined above. These are shown in figures 5.18 to 5.22.

The E5-E4 energy range (figure 5.18) shows that there is little variability in the low energy component at low $K_p$ and substantial variability at high $K_p$, and that the diurnal variation is more pronounced at high $K_p$. There is a small morning increase at lowest $K_p$ showing that few injections are occurring. This increases with $K_p$ and the change in shape of the median line also shows the influence of more injections as $K_p$ increases. The 95% and median curves also show that the time that the injections are observed gets earlier as $K_p$ increases: at
Figure 5.18: Statistical plots of E5-E4 flux (42.9-59.4keV) versus local time, in ranges of $K_p$.

lowest $K_p$ the 95% curve peaks at around 0600 LT whereas at highest $K_p$ the peak is at 0300 LT. The increasing number of injections towards higher $K_p$ also has the effect of raising the median up towards the 95% curve. The 5% curve progresses to lower nightside values and higher dayside values as $K_p$ increases (except for the highest $K_p$ range; see next paragraph). As was seen in the previous section, these low flux observations depend on the geomagnetic field structure, and here they show a more compressed dayside magnetosphere and more tail-like nightside at high $K_p$. At lowest $K_p$ the nightside decreases are not occurring, which fits with Baker et al.’s (1978) interpretation of them as a growth phase feature.

The highest $K_p$ plot shows a feature not seen at lower $K_p$ in the 5% line. The flux drops to very low values around noon. This is due to magnetopause encounters which occur occasionally at geosynchronous orbit when the magnetosphere is highly compressed. There is an asymmetry about noon with more low flux values pre-noon than post-noon. This is consistent with the electrons drifting eastward: ions would be expected to ‘drop out’ post noon since they would meet the magnetopause from the other direction (Sibeck et al., 1987). The same fea-
Figure 5.19: Statistical plots of E4-E3 flux (59.4-90.7keV) versus local time, in ranges of $K_p$.

Figure 5.20: Statistical plots of E3-E2 flux (90.7-134.9keV) versus local time, in ranges of $K_p$. 
ture was seen in figure 5.16. These encounters only happen at highest $K_p$ when the magnetosphere is strongly compressed.

Figures 5.18 to 5.22 show the same progressions seen in figures 5.3 to 5.9, namely the decreasing flux due to injections with increasing energy. The effect of the injections seems to come in at higher $K_p$ for increasing energy. This suggests that bigger substorms result in the injection of a greater range of particle energies. The features which depend on the magnetic field structure, such as the drop in the 5% line around noon at high $K_p$ due to magnetopause encounters or the nightside dropouts associated with the substorm growth phase, are seen at all five energies in figures 5.18 to 5.22.

For all energies, as $K_p$ increases, the range of flux observations increases, which reflects both the increased occurrence rate of substorm injections and the growing day/night asymmetry of the geomagnetic field. The local time variations therefore becomes more pronounced at high $K_p$.

The E1 plots in figure 5.22 show the range of trapped flux observed increases with $K_p$. This is again obvious in the larger decreases on the nightside and is due to the development of a more taillike field. At highest $K_p$ there is a suggestion in the 95% line and median that some particles of this energy are being injected, since there appears to be more variability in the morning than later. The lowest $K_p$ plot does not show the large nightside decreases because here the field is at its most dipolar.

Geomagnetic activity correlates with $K_p$ so the low energy E5-E4 data might be expected to be well organized with respect to $K_p$. The high energy E1 population however has a delayed response to geomagnetic activity, e.g. Szita et al. found that the SEM-2 E1 flux lagged behind $K_p$ by about 2 days, whereas the E5-E4 flux lagged by between 0 and 1 day. The same report found that the E1 flux lagged behind the E5-E4 flux by about a day. In figure 5.1 the E1 flux was seen to become enhanced about two days after a large peak in the $K_p$ index. The correlation is repeated in figure 5.23. This figure shows the value of the correlation coefficient calculated for different time lags between two times series of data.
Figure 5.21: Statistical plots of E2-E1 flux (134.9-201.8keV) versus local time, in ranges of $K_p$.

Figure 5.22: Statistical plots of E1 flux (201.8-300keV) versus local time, in ranges of $K_p$. 
The longest continuous section of the whole data set was used, i.e. May 1991 to May 1995, with small gaps filled. The correlation has a local time dependence due to the diurnal variations of the flux so the peak value of the correlation coefficient has been plotted per day (therefore the E1 to \(K_p\) curve is effectively \(\pm 12\) hours on the x-axis and the E5-E4 to \(K_p\) curve about \(-4\) to \(+20\) hours). It can be seen that whereas the E5-E4 flux changes promptly with \(K_p\), the peak correlation between the E1 flux and the \(K_p\) index is between 1.5-2.0 days lag. Therefore, if the E1 flux is to be presented in a form which shows how its variability depends on geomagnetic activity, then incorporating a time lag may organize the data better in terms of the \(K_p\) index. The plots obtained by imposing a lag on the flux of 1.75 days behind the \(K_p\) index are shown in figure 5.24.

![Figure 5.23: Correlation of E1 and E5-E4 flux with the \(K_p\) index and with each other.](image)

A second approach along the same lines was also tried: instead of using instantaneous \(K_p\) to categorize the E1 flux data, \(K_p(\tau)\) was used (Wrenn, 1987). This is a weighted average of \(K_p\) which was designed because the recent history of geomagnetic activity is often important rather than just the present value. The derivation of \(K_p(\tau)\) is given in Chapter 3. \(K_p(\tau)\) gives most weight to the recent past and less to earlier times. In Chapter 3 it was seen that the contribution
of each \( ap \) value to \( ap(\tau) \) is reduced to \( 1/e \) of its value after approximately 2.4 days, which is a time scale comparable to the correlation lag seen here between \( K_p \) and the El flux. The plots obtained using \( K_p(\tau) \) are shown in figure 5.25.

A consequence of producing \( K_p(\tau) \) is that less data points will now fall in the extreme (highest and lowest) divisions: only 8% of the data has \( K_p(\tau) \) in the 0o to 1+ range, and only 1% of the data has \( K_p(\tau) \) equal to 6- or above. Therefore a model based on \( K_p(\tau) \) may well require different limits to be set than for a model based on \( K_p \).

Figure 5.24: Statistical plots of El flux (201.8-300keV) versus local time, in ranges of \( K_p \) with a 1.75 day lag applied.

It is at first perhaps surprising that the El flux does not appear to be better organized by the use of \( K_p(\tau) \) or by introducing a time lag. Therefore in order to establish which is the best El model, it is necessary to consider in detail what the models show. The problem with the El flux is that there are two different physical processes at work which affect the same population in different ways and on different time scales. Firstly, the high energy flux is enhanced after a lag of 1.5-2 days after the \( K_p \) index and low energy flux is enhanced. The popular theory is that the low energy electrons are 'pumped' in energy to form the high
energy population and that the time taken for the energization is the lag time. Secondly, the geomagnetic field configuration is related to $K_p$ in the sense that, at high $K_p$, the field is distorted and at low $K_p$ the field is more dipolar: the position of a geosynchronous spacecraft in the radiation belt is therefore dependent on the current degree of activity. This effect obviously does not involve the 2 day time lag. Therefore, it might be expected that the high flux component associated with large $K_p$ enhancements requires a time lag in the model, but the low flux component, which depends on the magnetic field configuration, does not.

There are two reasons why the original plot which used instantaneous $K_p$ to organize the E1 data gave a good result. One is that often sustained periods of high and low activity occur. The second, and probably most important one, is that the spacecraft position in the magnetic field depends on the current $K_p$ value. Whether the E1 flux is greatly enhanced or not, if the tail is highly stretched then a geosynchronous spacecraft will still see decreased nightside flux.

A good model for the response of the field to the instantaneous $K_p$ value should therefore be detectable in the 5% line, since the taillike nature of the field
is picked out here. The degree of stretching in the tail is expected to increase with $K_p$, therefore the model should show the range in the 5% line increasing. For the high flux component, a good model should show the peak flux increasing with geomagnetic activity, allowing for any delayed response. These two effects working together serve to complicate attempts to model this population. If the E1 flux is greatly enhanced 2 days after a period of high $K_p$, the extent to which the enhancement may be observed at geosynchronous orbit is heavily dependent on the geomagnetic field configuration at the time.

Not surprisingly, it appears that low flux observations, as seen in the 5% line, are best organized in figure 5.22 (instantaneous $K_p$) since in this figure there is progression from low to high $K_p$ in which the nightside flux gets lower, reflecting the degree of field distortion. This sequence is less well defined for the $K_p(\tau)$ and is particularly bad when $K_p$ is used with a time lag. Also, not using instantaneous $K_p$ means that the magnetopause encounters no longer show up.

The 95% line however shows that the model of the E1 flux does benefit from imposing a lag on its response to $K_p$ or by using $K_p(\tau)$. This is manifested by the progressively higher flux attained by the 95% line with increasing $K_p$. In this case, the time-lagged data shows the best progressive increase in flux with $K_p$, in that it reaches the highest flux of the 95% lines and maintains a diurnal variation.

The fact that the high E1 flux can be well organized by the $K_p$ index with a lag of 1.75 days, or by $K_p(\tau)$, means that statistical plots of this type may serve as a predictive tool for the radiation belt flux of 200-300keV. In terms of peak flux, this is indeed best achieved with the time-lagged data.

Similar statistical plots for the intermediate energy ranges were shown in figures 5.19 to 5.21, in terms of instantaneous $K_p$. The energy levels between E5-E4 and E1 show an increasing lag in the best correlation between flux and the $K_p$ index. This shows the increasing time lag required to 'pump' the energy of the injected particles. In these cases, implementing specific time delays between the $K_p$ index and flux for each energy should better organize the high flux component as for the E1 flux.
5.3 The Diurnal Flux Variation

The SEM-2 observations show strong local time behaviour which is dependent on particle energy. Figure 5.1 and subsequent statistical plots showed how the lower of SEM-2's energy ranges is dominated by substorm injections and is strongly peaked post-midnight. The highest energy range is almost symmetrical about the noon-midnight line, with higher flux on the dayside than the nightside.

This section is concerned with identifying the shape of the average diurnal variation, which is shown in figure 5.10. A suitable tool for this is Fourier analysis.

Fourier Analysis

Fourier analysis is a technique widely used for finding periodicities in a time series of data. A periodic signal can be decomposed into a series of sine and cosine waves of different amplitudes and phases. The highest frequency which can be determined from Fourier analysis is called the Nyquist frequency \( f_c \) and is given by:

\[
f_c = \frac{1}{2\Delta}
\]

where \( \Delta \) is the time spacing of the sampling. Therefore the minimum period which can be determined is twice the sampling period i.e. a minimum of two points per wavelength are needed to identify a waveform. This means that from the low resolution SEM-2 data, the minimum period which could be found would be 1 hour.

There may be a problem with carrying out a Fourier transform on a time series if there were higher frequencies in the data than the sampling frequency, i.e. shorter periods than the sampling period. This produces an error known as aliasing: the power in these higher frequencies will be 'folded back' into the output of the transform, i.e. the power in frequencies higher than the Nyquist frequency will be moved into the frequencies lower than the Nyquist frequency. This should not be a problem with this analysis since a year of data is being used, and there are unlikely to be any such short period, long term waveforms.

The software package used for the Fourier analysis was ARK, written by Alan
Identifying the diurnal variations

The results of the Fourier analysis showed that the frequency spectrum for the E1 flux featured a large peak at a period of 24 hours. After filtering out this sine wave at the relevant amplitude and phase, the average local time profile was fairly flat. The average diurnal variation of the E1 flux could therefore be adequately described by a single sine wave of period 24 hours. This fit is shown in the left hand plot of figure 5.26 for the 1989 data.

For the E5-E4 flux, initial investigation showed that although the average diurnal variation is sinusoidal in appearance, a single sine wave does not provide an adequate fit: a sine wave of the correct frequency lags the real data for half a cycle and leads for the other half. In this case two peaks in the Fourier spectrum showed that a combination of two sine waves with periods 24 hours and 12 hours were required. The right hand plot of figure 5.26 shows how well the composite wave fits the 1989 data. Again, filtering out this composite wave left the average local time profile flat.
Any periodic waveform can be represented by the summation of its fundamental frequency plus its harmonics: the \( n \)th harmonic of a wave has a frequency \( n \) times the fundamental frequency (Newland, 1993). The fundamental frequency suffices to describe the average local time behaviour of the E1 flux but in the case of the E5-E4 flux a combination of the fundamental frequency plus the second harmonic is required. The difference in the local time variations is caused by the fact that the magnetic field dictates the E1 flux behaviour whereas the injected electrons dominate the E5-E4 flux.

### 5.4 Anisotropy Index and Spectral Index

In this section the anisotropy index, and spectral index, are presented in the same statistical plot format as has been used for flux. Finally, scatter plots are used to relate both indices to flux and show the dependence of all three parameters on \( K_p \).

#### 5.4.1 Anisotropy Dependence on Local Time

The same statistical plotting technique is applied to the second order anisotropy index in figure 5.27. Recall that the anisotropy index is positive for a cigar-like distribution and negative for a pancaked distribution.

The positive values of the anisotropy index across the nightside reflect the taillike field: if a field line becomes more taillike, its equatorial value of magnetic field is reduced, and if \( \mu \) is conserved then its equatorial pitch angle is also reduced. So the nightside electrons at geosynchronous orbit are more field-aligned than on the dayside. Baker et al. (1978) suggest that magnetopause shadowing may also contribute to the field-aligned nightside distribution. Figure 5.27 shows that on the dayside the distribution is pancaked and has a rough symmetry about \( \sim 1100 \) LT rather than noon as was also seen in the high energy electron flux. The injections have an effect on the 95\% curve which will be discussed further in Chapter 6. Where a high flux of injected electrons was observed, from just before midnight until dawn, the anisotropy index becomes less positive. There may be
a combination of effects at work here: the pitch angle distribution of the injected electrons may affect the overall distribution, and in addition, substorms result in field dipolarizations. When this happens, stretched field lines revert to roughly dipole positions, so as a result particles on the lines will have their equatorial pitch angles increased.

![Anisotropy index versus local time](image)

Figure 5.27: Anisotropy index versus local time

The anisotropy data are divided into ranges of $K_p$ in figure 5.28. As noted by Higbie et al. (1979) in CPA observations of 30-300keV electrons, even at low $K_p$ local time behaviour exists showing the distribution to be less pancaked at midnight than at noon. The cigar phase starts earlier at high $K_p$: at lowest $K_p$ the sharp upturn in the 95% curve occurs at 2000 LT whereas at highest $K_p$ it occurs at 1600 LT. This is in common with the injections of electrons starting earlier at high $K_p$. At higher $K_p$ a greater degree of anisotropy is reached than at low $K_p$, because the tail becomes more stretched. In addition the end of the steep fall in the 95% curve which is at 0200-0300 LT at low $K_p$ shifts to midnight at high $K_p$. The local time behaviour of the anisotropy index therefore shows that a distribution indicative of a taillike field appears earlier in local time as $K_p$ increases, and covers a larger local time sector. Higbie et al. also noted that
Figure 5.28: Anisotropy index versus local time in ranges of $K_p$

The cigar phase was stronger towards dusk than towards dawn. It can also be seen, particularly in the median curves, that the dayside pancaked phase becomes narrower in local time with increasing $K_p$. The afternoon shift is greater than pre-noon, which may show that magnetopause shadowing is occurring. Overall, at high $K_p$, the magnetosphere as a whole is more asymmetric.

The anisotropy features described are reminiscent of Baker et al.’s (1978) observations of the anisotropy sequence for 30-300keV electrons in the passage of a substorm. This is studied in further detail in Chapter 6. An hour or two before onset Baker et al. saw a cigar phase develop, accompanying a taillike field, which generally did not happen if there was no substorm. It was suggested therefore that the cigar phase may serve as a predictive tool for substorm occurrence.

Following the injection the distribution became pancaked. The median time delay from the onset of the cigar phase and the injection was $\sim 95$ minutes. By comparing figures 5.28 and 5.18, the development of the field-aligned distribution precedes the subsequent substorm injections by a decreasing amount of time for increasing $K_p$, as seen in the 95% line. It is also seen that the earlier the cigar phase is observed, the greater the range of flux subsequently observed and the
95% line reaches higher flux.

5.4.2 Spectral Index

Energy spectra were presented earlier in Chapter 2. The spectral index stored in the archive files is the slope of the logarithm of the energy spectrum calculated using a least squares fit. The archiving program calculates the spectral index at each point using the flux in SEM-2's five energy ranges. The spectral index is usually negative because there are generally more lower energy than higher energy particles: in the SEM-2 low resolution archive data, the value of spectral index generally lies between 0 and -6. A strongly negative spectral index corresponds to a soft energy spectrum, where the particle population dies away with increasing energy, whereas a weakly negative spectral index corresponds to hard spectrum, with many particles above the energy range of the analyser. In the light of the two electron populations present, the spectral index will describe the relative influence of the high and low energy electrons on the total energy range. The most negative spectral index will be expected when injections occur and flood the geosynchronous orbit region with low energy electrons. The least negative spectral index will be expected post noon when the effect of the injection is lowest.

Figure 5.29 shows the statistical plot of spectral index versus local time. The most negative values of spectral index, indicating when the energy spectrum is softest, occur from shortly before midnight through to the morning where injection events were seen in the low energy flux (figure 5.3): the abundance of the lower energy injected electrons softens the energy spectrum. The least negative values are seen on the dayside, post-noon. By this local time the injected electrons are dispersed in energy and the flux of the high energy, trapped electrons is higher than on the nightside.

When divided into ranges of $K_p$ (figure 5.30), the spectral index shows an increased range of values with increasing $K_p$ and the diurnal variation has increased amplitude. Thus the greatest ratios of high to low and low to high energy electrons are found at high $K_p$. The most negative spectral index is observed closer to midnight with increasing $K_p$: this fits with the fact that peak flux of injected
Figure 5.29: Spectral index versus local time

Figure 5.30: Spectral index versus local time in ranges of $K_p$
electrons were observed closer to midnight at high $K_p$.

### 5.4.3 Flux Related to Spectral and Anisotropy Indices

The final type of plot presented shows the mutual dependences of electron flux, the anisotropy index and the spectral index, using colour to provide the third dimension of the $K_p$ index. Flux from the total energy range (42.9-300keV) is used because both indices apply to the whole energy range.

Figure 5.31 contains four scatter plots of the total flux versus the second order anisotropy index, each for an hour of local time from midnight, dawn, midday and dusk. Each dot is a high time resolution data point and the whole data set was used. The colour scale is used to show the $K_p$ index: purple for 0 to 1+, bright blue for 2- to 3+, yellow for 4- to 5+, red for 6- and above. The high $K_p$ points are plotted last over the low $K_p$ points. On the dayside, as seen in the dawn and noon plots, the anisotropy index is predominantly negative so the electron distribution is pancaked. At noon there are a few low flux points at a range of anisotropy values, all at high $K_p$, which must correspond to magnetopause encounters. By dusk a significant portion of the data has positive anisotropy corresponding to the field becoming taillike on the nightside; these points have low to moderate flux and are associated with electrons becoming more field-aligned in the stretched tail. These points are also seen at midnight. The high $K_p$ points at midnight occupy two regions on the graph: some, at high flux, correspond to injections with the anisotropy index tending to be negative, and others at very low flux. So at moderate $K_p$ at midnight the cigar phase is seen but at high $K_p$ dropouts occur, and these two phenomena occupy different regions on the graph. Flux in the total energy range generally drops to $\sim 10^3\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$ or so prior to the midnight sector large injections. Flux $< 10^3\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$ is less common and is observed occasionally on the nightside or dayside: these points are associated with magnetopause or lobe encounters.

Figure 5.32 contains four local time scatter plots of total flux versus the spectral index, also showing $K_p$ dependence by colour. These plots show that at higher $K_p$, and higher flux, the range of values taken by the spectral index be-
comes both larger and more negative. This is most obvious at midnight and at dawn. This shows the effect of injected electrons on the energy spectrum.

Figure 5.33 shows the data from the last two plots plotted as the anisotropy index versus the spectral index. At midnight, the more positive the anisotropy index points get, the narrower the range of spectral index values they occupy. These points have moderate $K_p$, whereas at dusk, high $K_p$ points are seen here. The high $K_p$ points at midnight take the most negative spectral index values, and are fairly isotropic. At all local times, the spectral index range of values becomes more negative with increasing $K_p$. The anisotropy index is more negative on the dayside, and more isotropic/field-aligned on the nightside.

On all three sets of scatter plots, the high $K_p$ points show different behaviour at midnight and dawn: although injections are seen at both midnight and dawn, the dropouts are seen only at midnight. Therefore at high $K_p$ there is a much greater range of flux observed at midnight than dawn. From the midnight and dawn plots therefore the range of anisotropy and spectral index associated with injection events is clear: the injected electrons are fairly isotropic and the spectral index is most negative for the high $K_p$ events.

In summary then, at midnight, higher flux is seen with higher $K_p$ from midnight to dawn where injections occur, and these points occupy the most negative spectral indices because of the high injected flux at low energies. Also at midnight, low flux, field-aligned points are seen corresponding to the taillike growth phase, which have a fairly narrow range of spectral indices. Very low flux points, seen at higher $K_p$ than the growth phase decreases, correspond to severe flux dropouts, and are more isotropic with more negative spectral indices than the growth phase flux decreases. The anisotropy derived in these cases may have more to do with the lack of counts than any real distribution, but this feature helps to distinguish between two different types of flux decrease.

At dawn the field-aligned, growth phase flux decreases are absent and the flux is very well organized by the $K_p$ index. At high $K_p$, high flux injections are occurring. There are a few very low flux, high $K_p$ points which show that the
spacecraft is outside the radiation belts, perhaps in the lobe.

The noon ranges of flux, spectral index and anisotropy index are all more restricted than at midnight. The flux points reach lower values than at dawn because of energy dispersion of the injected electrons. A few high $K_p$ points have very low flux and are therefore likely to indicate magnetopause encounters. The anisotropy index is mostly negative showing the trapped distribution. The spectral index tends to be more negative at high $K_p$, which must be due to injected electrons persisting through extended active periods, and also reduced high energy flux seen when the dayside magnetosphere is compressed.

At dusk a few high flux points are seen at highest $K_p$ showing that a few injections occur this early, with spectral and anisotropy index values the same as the midnight injections. These events are clearly distinct from injections which occurred nearer midnight then drifted round this far. The very low flux dropouts are seen, as for midnight at high $K_p$, but the growth phase flux decreases which were seen at moderate $K_p$ at midnight occur at higher $K_p$ here, which shows that higher geomagnetic disturbance is required to create a taillike field this early in local time.

Figure 5.31: Total flux versus anisotropy index, in ranges of $K_p$. Purple, blue, yellow, red represent the four ranges of low to high $K_p$. 
Figure 5.32: Total flux versus spectral index, in ranges of $K_p$. Purple, blue, yellow, red represent the four ranges of low to high $K_p$.

5.5 Pitch Angle Distributions

Since the SEM-2 has 30 polar-azimuthal look directions, and since the magnetic field direction can be inferred from the particle distribution, it is possible to obtain coarse pitch angle data. Although the anisotropy index can show where the distribution is pancaked or cigar-like, it should be possible to go further to see how the distribution develops in local time using calculated pitch angles. Because the polar-azimuthal data is only available for the whole energy range of the SEM-2, pitch angles cannot be calculated for each population separately.

The pitch angle distribution (PAD) of a plasma may be one of several characteristic types. A distribution which is peaked at the $90^\circ$ pitch angle is called a normal or pancaked distribution. A field-aligned distribution which peaks at low pitch angles is referred to as cigar-shaped. An isotropic distribution features all pitch angles with equal significance. A head and shoulders distribution is similar to a normal distribution but has an excess of $90^\circ$ pitch angle particles. A butterfly distribution has peak fluxes of particles between $0^\circ$ and $90^\circ$ and between $90^\circ$
Figure 5.33: Anisotropy index versus spectral index, in ranges of $K_p$. Purple, blue, yellow, red represent the four ranges of low to high $K_p$.

and 180°. Figure 5.34 illustrates some of these pitch angle distributions (from Sibeck et al., 1987).

The PAD can be modified by many different processes. Particles with small pitch angles may be lost from the radiation belts to the atmosphere because their mirror points occur at low altitudes: this limit on the pitch angles defines the loss cone. Particles with large pitch angles may be lost to the magnetopause on disturbed days. Drift shell splitting and quasi-trapping also influence the observed PAD. The typical pitch angle distribution at geosynchronous orbit is discussed later in this section.

**Pitch angle calculation**

To calculate pitch angles, the directions of the electrons and the magnetic field direction are needed. There are 30 look directions which define the direction the electrons are travelling in: five polar bins at 30°, 60°, 90°, 120° and 150° to the spacecraft spin axis, by six azimuthal bins which cover six 60° sectors between 0° and 360°.
Figure 5.34: Example pitch angle distributions: (a) normal, (b) butterfly, (c) isotropic, and (d) head and shoulders (from Sibeck et al., 1987).

Figure 5.35 illustrates the geometry of the pitch angle calculation. The spacecraft spin axis is the z-axis of the Cartesian coordinate system. The two vectors shown represent the magnetic field ($\mathbf{B}$) and the electron flux ($\mathbf{j}$), which is defined by the look-direction of the detector telescope. The orientations of these vectors in the spherical polar coordinate system are described by the polar and azimuthal angles $\theta_B$ and $\phi_B$ for $\mathbf{B}$, and $\theta_j$ and $\phi_j$ for $\mathbf{j}$. The angle between the directions of the vectors $\mathbf{j}$ and $\mathbf{B}$, $\alpha$, is the electron’s pitch angle.

The angle between two vectors can be found from their scalar product. The angle $\alpha$ between $\mathbf{B}$ and $\mathbf{j}$ is found from the following calculation. The scalar product of $\mathbf{B}$ and $\mathbf{j}$ is:

$$\mathbf{B} \cdot \mathbf{j} = |\mathbf{B}||\mathbf{j}| \cos \alpha$$  \hspace{1cm} (5.10)

The scalar product can also be expressed as:

$$\mathbf{B} \cdot \mathbf{j} = B_x j_x + B_y j_y + B_z j_z$$  \hspace{1cm} (5.11)
where the components of the vectors $\mathbf{B}$ and $\mathbf{j}$ are given by:

\[
B_x = |\mathbf{B}| \sin \theta_B \cos \phi_B \\
B_y = |\mathbf{B}| \sin \theta_B \sin \phi_B \\
B_z = |\mathbf{B}| \cos \theta_B \\

\dot{j}_x = |\mathbf{j}| \sin \theta_j \cos \phi_j \\
\dot{j}_y = |\mathbf{j}| \sin \theta_j \sin \phi_j \\
\dot{j}_z = |\mathbf{j}| \cos \theta_j
\]  

Substituting for the components of $\mathbf{B}$ and $\mathbf{j}$ from equations 5.12 into equation 5.11, and by equating 5.10 and 5.11, an expression for $\alpha$ is obtained:

\[
\alpha = \cos^{-1} \left( \sin \theta_B \cos \phi_B \sin \theta_j \cos \phi_j + \sin \theta_B \sin \phi_B \sin \theta_j \sin \phi_j + \cos \theta_B \cos \theta_j \right)
\]  

Each pitch angle was calculated using the centre of the polar-azimuthal bin to define $\theta_j$ and $\phi_j$. The viewing angle of each telescope is $10^\circ$ meaning that the assumed value for $\theta_j$ is $\pm 5^\circ$. However, $\phi_j$ is $\pm 30^\circ$ because each azimuthal bin is one sixth of a spacecraft spin i.e. $60^\circ$. This means that each pitch angle calculated has a maximum error of just over $30^\circ$. The calculated pitch angles were averaged in $30^\circ$ bins between $0^\circ$ and $90^\circ$ for plotting: because the sense of the magnetic field direction is not available from the axis calculation, it is not possible to tell if the electrons are moving in the direction of the field or against it. Although the maximum error on any one point is a little over $\pm 30^\circ$, the actual error may be anywhere between this and zero: the average error is nearer $\sim \pm 15^\circ$. Given this fact and the fact that there is a vast data set justifies binning down to $30^\circ$.

**PAD dependence on local time and $K_p$**

Figure 5.36 shows pitch angle distributions for each hour of local time over the whole data set. The distribution from midnight to 0200 LT is fairly isotropic, with
Figure 5.35: Calculation of pitch angle: knowing the local magnetic field (B) direction and sensor look direction, the particle pitch angle $\alpha$ can be found (see text).
flux increasing. The flux increase is due to the injections. From 0200 or 0300 LT the distribution starts to become increasingly pancaked until 1100, during which time the small pitch angle flux decreases and the large pitch angle flux increases. The situation then reverses and the distribution is isotropic again by 1900 LT. After this, a field aligned or butterfly distribution develops. Given the coarseness of the pitch angle data it is impossible to see what is happening in the loss cone which is only a few degrees wide.

Therefore on the dayside the pitch angle distribution is pancaked and on the nightside it is isotropic to field aligned. This would be expected from the effects of shell splitting in an asymmetric field. More information can be gained by looking at times of low and high geomagnetic activity. To define these, the $K_p$ index is again used.

The same plot was produced for low $K_p$, defined as $K_p \leq 1+$, shown in figure 5.37. Flux is lower at all local times with the greatest difference between the flux here and at all local times (figure 5.36) is in the post-midnight sector, due to less injections occurring at low $K_p$. The same general pattern of the distribution being pancaked on the dayside and isotropic to field aligned on the nightside is seen, but the development of a pancaked distribution starts later at 0300 or 0400 LT, and the return to an isotropic distribution is reached later at about 2100 LT. The few hours preceding midnight are more isotropic at low $K_p$ than for all $K_p$. This all points to the field being more dipolar at low $K_p$.

Figure 5.38 shows the plots for high $K_p$, defined by $K_p \geq 6+$. Peak flux at 0200LT shows where the injections are occurring. The dayside pancaked distribution appears to develop later than at other $K_p$, the first sign being at 0500 LT. The decline of the pancake after noon is swifter and by 1700 LT the distribution is cigar-shaped, and remains so until 2300 when it becomes fairly isotropic. Thus pre-midnight the field is highly stretched, and post-midnight the effect of more injections occurring at high $K_p$ makes the distribution more isotropic. At most times the flux is higher than on figure 5.36, but between 0900 and 1500 UT their levels are comparable, and from 1500 until dusk they are significantly lower. A flux drop is seen at 0900 LT, which, accompanied by the reduced 90° flux at 1000
Figure 5.36: Calculated electron pitch angle distributions for all local times
Figure 5.37: Calculated average electron pitch angle distributions for all local times, $K_p \leq 1+$
Figure 5.38: Calculated average electron pitch angle distributions for all local times, $K_p \geq 6$—
LT, suggests that magnetopause shadowing has reduced the high $K_p$ flux.

The pancaked dayside distributions and field-aligned nightside distributions are consistent with shell-splitting theory and have been observed by many authors (e.g. Sibeck et al., 1987). The dayside normal distribution exists for a shorter time at high $K_p$ because during high activity the magnetosphere is highly asymmetric, and the nightside field is more taillike. This is why the PAD becomes field-aligned earlier at high $K_p$. The PAD then tends back to isotropic from field-aligned earlier than at lower $K_p$ because of the influence of substorm injected electrons.

Figure 5.39 shows the pitch angle distribution for energetic electrons (79-822 keV) in the near equatorial magnetosphere obtained from Ogo 5 data (West, 1979). Data from periods of major magnetic storms were excluded from the survey. The morning sector shows a normal distribution (peaked at 90°) out to the magnetopause. In the afternoon, butterfly distributions (depleted at 90°) appear from the magnetopause in to about 5$R_E$ and seen closer to Earth towards dusk. On the nightside the distribution is normal close to earth, tending to butterfly or isotropic distributions further away. The crossover from normal to butterfly distributions is shown by the semicircles of dashes: the inner one at 5-6$R_E$ for 822 keV electrons and the outer one at 7-8$R_E$ for 79 keV electrons. Outside these limits the distribution tends to butterfly in the pre-midnight sector, with more isotropic distributions in the post-midnight sector. This plot is in fair agreement with the SEM-2 results, although the hourly plots are better for showing the development of the distribution, e.g. from figure 5.36 the dayside normal starts a couple of hours earlier than implied by figure 5.39. Shell splitting is responsible for the change from normal dayside distributions to butterfly nightside distributions beyond 5$R_E$. Sibeck et al. (1987) observed a normal nightside distribution for particles $>25$ keV inside 6$R_E$ but butterfly distributions further out. Sibeck et al. remark that magnetopause shadowing can only therefore account for the butterfly distributions seen on the nightside where the drift paths have intercepted the magnetopause. In cases where the drift paths could not have encountered the magnetopause, these distributions are due to shell splitting in the presence of a strong negative radial flux gradient. In the inner magnetosphere the flux
gradient is positive radially outward, but in the outer magnetosphere (beyond $5R_E$) the flux gradient is negative. Starting from the dayside, $90^\circ$ particles will drift to nightside locations closer to Earth than lower pitch angle particles. This creates a deficit of $90^\circ$ particles on the nightside which cannot be compensated by other particles drifitng inward from the dayside because of the negative flux gradient. Therefore around $6R_E$ and beyond a quiet time butterfly distribution forms on the nightside. Since for higher energy particles the change from a positive to a negative flux gradient is closer to Earth, the inner edge of the butterfly distribution will lie closer to Earth than for less energetic particles.

Figure 5.39: Survey of energetic electron pitch angle distributions in the near equatorial magnetosphere as determined by the Lawrence Livermore Laboratory’s experiment on Ogo 5 (West, 1979).
5.6 Summary

Analysis of the local time behaviour of electron flux at geostationary orbit has yielded the following results:

- Statistical plots were produced showing the local time dependence of the data and how it depends on energy. Lower energies (< 200keV) show increased variability and peak flux in the several hours after midnight due to substorm injected electrons. Above 200keV the injections have a small effect and the local time behaviour reflects the geomagnetic field structure, having weaker flux and greatest variability on the nightside.

- The flux due to injections increased and the time of peak flux moved earlier in local time towards midnight as $K_p$ increased.

- The delayed response of the high energy component to the $K_p$ index was investigated. The overall shape of the local time distribution was well organized by instantaneous $K_p$, but using a time lag organized the peak flux better.

- Nightside flux decreases associated with growth phase dropouts and dayside decreases associated with magnetopause shadowing were seen at all energies.

- The anisotropy index showed a pancaked dayside distribution which extended further in local time at low $K_p$ than at high $K_p$. The night side distribution was more field aligned, with the most field aligned sector moving earlier in local time as $K_p$ increased.

- Coarse pitch angle information was calculated using the polar-azimuthal flux data and the assumption that the magnetic field direction is given by the symmetry axis of the particle distribution. The PAD progressed from pancaked on the dayside to isotropic/field-aligned on the nightside. At high $K_p$, the PADs showed the field to be more asymmetric.
• The pancaked dayside and field-aligned nightside populations agree with previous observations and with the predictions of Roederer's shell-splitting theory. In addition, very low flux indicative of nightside flux dropouts and dayside magnetopause encounters can be distinguished. Substorm injections are seen as high flux, isotropic to pancaked events.

5.7 Conclusion

Each of the statistical plots presented in this chapter forms a local time dependent model of the range of electron flux observed at geostationary orbit. A comprehensive picture of the geosynchronous orbit region at different levels of geomagnetic activity is provided by the local time models of flux, anisotropy index, spectral index and pitch angles. At low $K_p$ levels there are few injections and the geomagnetic field is at its most symmetrical whereas at high $K_p$ the day-night asymmetry is extensive. Pancaked dayside distributions and more field-aligned nightside distributions, in agreement with the predictions of shell-splitting theory, are seen even in quiet times, and in active times the extent of the pancaked region decreases and the magnetosphere becomes more taillike.

The low energy population is dominated by substorm injections and has peak flux in early morning. The E1 range generally does not see the injections, but they are occasionally observed in this range when $K_p$ is high. The models show that injections are seen only at higher $K_p$ for higher energies, so only when the magnetosphere is highly disturbed are electrons $>200$ keV injected by the substorm process. Higher $K_p$ also leads to an increased range of flux observations at all energies due to enhanced drift shell splitting and dropouts in the substorm growth phase. At low $K_p$ these growth phase decreases are absent so there is a reduced range of observations.

The flux of injected particles can be successfully modelled in ranges of $K_p$, since few injections occur at low $K_p$ whereas during high activity the E5-E4 flux can be greatly increased. A large enhancement in the E5-E4 flux may be followed a day or two later by an enhancement in the E1 range. However, the local time
dependence of the E1 flux has been shown to depend on the magnetic field geometry which varies directly with geomagnetic activity and the $K_p$ index. The flux of higher energy trapped particles therefore depends both on the instantaneous $K_p$ index, which changes with the geomagnetic field configuration, and on the level 1.75 days beforehand, which seems to be the time scale for pumping the energy of the injected electrons up to the E1 value. Consequently there are deficiencies in organizing the high energy population both in terms of instantaneous $K_p$ or with a time lag. It was shown that the E1 low flux observations are best organized by instantaneous $K_p$ but that high flux is best organized by imposing a lag on the observation behind the $K_p$ index.

Because the correlation of flux with $K_p$ contains the diurnal variation of the flux it is difficult to time exactly the lag between $K_p$ and the flux at each energy. The peak lag for successively higher energy flux does increase from ~0 to ~2 days between the E5-E4 and E1 energy ranges. Moreover, these correlation peaks are spaced farther apart at higher energy: it takes longer to increase the energy of an electron in the E4-E3 range to the E3-E2 range than it does to increase the energy of an electron in the E3-E2 range to the E2-E1 range. The energy ranges get bigger as energy increases so it seems it is taking longer to increase the energy by a larger amount. Thus if each recirculation step is associated with a characteristic magnetic field increase, causing a characteristic particle energy gain, this pattern of correlation lags might be expected. Even taking into account the effect of the shift of the diurnal maximum, the 40-300keV electrons do not behave as one population in this respect as has been reported for 3-40MeV electrons by Baker et al. (1990). Their analysis found for 3-40MeV electrons in four channels the time lag was 2-3 days behind $K_p$, the same for each channel. This time lag is longer than that found here for the E1 range so may still be consistent with a recirculation model.

The term 'dropouts' covers a range of events which cause reduced flux to be observed by a detector. The dropouts observed by the SEM-2 can be roughly divided into two categories. In the substorm growth phase flux decreases because of the stretching of the magnetotail. In addition the spacecraft may pass into
the plasma sheet which results in significantly reduced flux but not 'complete' dropouts (Baker et al., 1982b). The second category are effectively complete dropouts where the flux measured by the detector drops to background levels, as may occur when the SEM-2 enters the lobe or crosses the magnetopause. The distinction is shown clearly by scatter plots of flux versus the anisotropy index (figure 5.31). Dropouts that fit the picture of the substorm growth phase have a positive anisotropy index value, corresponding to field-aligned particles in a taillike field. The complete dropouts, which include magnetopause and lobe encounters, do not exhibit the field-aligned structure, although this may stem from the low counts measured rather than an accurate determination of a real distribution. This therefore can serve as a tool in distinguishing the types of dropout. Attempts at adiabatic modelling of the dropouts were inconclusive. The method is based on Liouville's theorem and therefore assumes that the same particles are seen along the orbit, which of course is not true. In a substorm growth phase when the magnetotail is being stretched, this approximation is at its furthest from reality. The conclusion is that the problem cannot be resolved solely by geosynchronous observations.

Figure 5.40: Geometry of the injection boundary proposed by Mauk and Meng (1983). The inflection point is situated an hour or two past midnight.
In terms of the injection boundary model of the substorm, Mauk and Meng (1983) have concluded that the inflection point of the injection boundary is displaced towards dawn. The morning peak of the low energy statistical plots supports this theory. The substorm injection boundary model defines a spatial threshold beyond which particles are accelerated at onset. This is shown schematically in figure 5.40. Mauk and Meng's data suggest the boundary's closest point to Earth is found between 0100 and 0200 LT at geostationary orbit, which is inferred from particle dispersion signatures.

It was seen that injections are observed by the SEM-2 earlier as \( K_p \) increases. Since the near-Earth part of the plasma sheet is important for substorm development (Baker and Pulkkinen, 1991) it would be expected that when geomagnetic activity is high, and the plasma sheet is drawn closer to Earth, then substorms would be observed earlier. In a standard model of the substorm, dayside reconnection in the presence of a southward IMF leads to the transport of magnetic flux to the tail lobes. The plasma sheet thins and moves earthward. Reconnection in the tail then leads to particles being accelerated earthward, which is termed a substorm injection. The evidence here is that when geomagnetic activity is high the dynamic processes of the substorm must occur closer to Earth.

ATS-5 observations interpreted in terms of the injection boundary model (Mauk and McIlwain, 1974) also show the boundary lies closer to Earth at high \( K_p \). Observations at a range of radial distances by Lopez et al. (1990) also pointed to a similar \( K_p \)-dependent injection boundary. The time the peaks are observed does not necessarily give the location of the boundary however, since generally injections will be observed some time after crossing the boundary (Moore et al., 1981).

It was also suggested by the statistical plots that the delay between the cigar phase onset and the subsequent injection was shorter for higher \( K_p \). This suggests that the energy loading time is shorter when \( K_p \) is high. This is also supported by the results of the wavelet analysis which show that in active times substorm frequency increases. Solar wind-magnetosphere coupling models (Akasofu, 1981) propose that substorms occur as soon as the stored energy exceeds a threshold.

181
value. In this model, energy being stored at a higher rate would lead to quicker releases. Indeed, Bargatze et al. (1985) found that the response time of the AL index to the solar wind input parameter $vB_s$ peaked earlier at high activity than for low activity, showing that the magnetospheric response time is shorter when activity is high.
Chapter 6

Substorm Injections

6.1 Substorm Injection Frequency

The frequency of substorm occurrence relates either to the loading and unloading time in the energy storage and release model of substorms, or to the disruption and recovery time in the directly driven model. Although substorms have been regarded as unpredictable and sporadic events, recent observations point to periodic behaviour when conditions are favorable (e.g. Borovsky et al., 1993).

Wavelet analysis is an ideal tool for identifying the frequency of substorms. Whereas Fourier analysis is suitable for finding long-term frequencies in a time series, and could easily miss a small train of periodic signals in a long data set, wavelet analysis can locate transient frequencies and their temporal location.

Wavelet Analysis

Fourier analysis is a technique best suited to the analysis of stationary signals, i.e. signals whose properties do not change with time, such as a combination of sine waves. Fourier transforming a non-stationary signal leads to any abrupt changes in the signal being spread out over the whole frequency spectrum.

Wavelet analysis is a relatively new technique which allows the frequency decomposition of a signal whilst keeping its temporal localization. It does this by using a wave packet or wavelet which is non-zero for only a finite time interval. Therefore, whereas Fourier analysis may be thought of as a correlation of the
whole signal with a sine wave, wavelet analysis may be thought of as correlating sections of the signal with a wave packet. This means that wavelet analysis can locate transient frequencies in a signal which might not be detected with Fourier analysis.

Wavelets are sets of functions of the form:

$$\psi_{a,b}(x) = |a|^{-1/2} \psi\left(\frac{x-b}{a}\right)$$  \hspace{1cm} (6.1)

where $a$ is a dilation parameter which dilates or contracts the wavelet $\psi$, and $b$ is a translation parameter which shifts the wavelet along the time axis. $\psi(x)$ having $a=1$ and $b=0$ is called the 'mother' wavelet, (Szu and Caulfield, 1992 and Ruskai et al., 1992).

![Figure 6.1: Morlet wavelet](image)

Dilating or contracting the wavelet, i.e. changing its frequency, is known as scaling. Whereas Fourier transforms work in the time-frequency domain, wavelet transforms use the time-scale domain. The scale parameter which is used instead of frequency is inversely proportional to frequency. Figure 6.1 shows the Morlet wavelet which was used for this analysis. The equation used for this Morlet
The continuous wavelet transform $S(a,b)$ of the real signal $s(t)$ is then given by:

$$S(a,b) = \frac{1}{\sqrt{a}} \int \overline{g}(\frac{t-b}{a}) s(t) dt$$  \hspace{1cm} (6.3)$$

where $\overline{g}$ is the complex conjugate of the wavelet function $g$. The wavelet transform can then be expressed in the following form:

$$S(a,b) = M(a,b)e^{i\phi(a,b)}$$  \hspace{1cm} (6.4)$$

where $M(a,b)$ and $\phi(a,b)$ are the modulus and phase of the transform respectively.

The modulus of the wavelet transform at each point in time shows the significance of the particular wavelet scale at each point along the time series. The wavelet package used for this analysis was Ondelet, supplied by the Observatoire Royal de Belgique and written by J-L. Modave and F. Collin. This program obtains the continuous wavelet transform at discrete values of $a$ and $b$ for a time series of data sampled at a constant rate.

**Scalograms**

Results from the wavelet analysis are presented as *scalograms*. Three scalograms are shown in figure 6.2 to illustrate the response of the wavelet transform to typical features of a time series. Each wavelet figure comprises two panels: the upper panel shows the time series used in the analysis as a white-on-black trace, and the lower colour scale plot shows the modulus of the wavelet transform. Here the colour scale illustrates how good the correlation of the wavelet is with each particular section of the time series: in these three cases the colour scale is arbitrary since it is the pattern of the response that is being illustrated. The colour plots show which periods are significant for each moment in time along the time series. The upper and lower panel share the horizontal axis which is linear in time. The vertical axis of the colour scale plot is linear in period, increasing
Figure 6.2: Wavelet transform of (A) a pulse, (B) a sine wave and (C) a decaying sine wave
downwards. In each colour scale plot, the colour scale is set to suit the range of the modulus in the plot.

Plot (A) shows a pulse: one data point in this time series has a non-zero value. The characteristic signature of the pulse in the wavelet transform spreads out in time as the period of the wavelet increases, but the response decreases as the wavelength of the wavelet becomes large compared to the pulse duration. Because of the $1/\sqrt{a}$ normalization factor in equation 6.3, the energy in the transform at each scale is the same. A pure sine wave is shown in plot (B): the modulus of the wavelet transform peaks in a horizontal line corresponding to the frequency of the wave, and frequencies close to the real frequency show a reduced response. Plot (C) shows a sine wave starting abruptly then decaying to zero. Since the sine wave has an abrupt start the characteristic signature of a pulse is seen where it begins. The horizontal line signature of a sine wave of constant frequency is then seen, with the modulus of the wavelet transform decreasing with the amplitude of the sine wave.

Wavelet transforms of E5-E4 flux data

Figure 6.3 (A) shows a sample of the E5-E4 flux data from 1989 (upper panel) with the scalogram showing the modulus of its wavelet transform (lower panel, colour scale plot). The x-axis covers 120 days and is the same for both panels. The y-axis of the colour scale plot shows the period of the wave being detected, from 1-35 hours. Since the low time resolution data (30 minute averages) is being used, the minimum periodicity that can be identified is 1 hour. This plot is dominated by the diurnal variation centred on 24 hours, which is itself quite variable: where there are fewer or lower flux peaks, e.g. around day 10, the wavelet transform shows the diurnal variation to be less significant. Some peaks are seen at periods of a few hours which correspond to particle injections.

The problem with attempting to find a frequency of substorm occurrence is that the size of the injection pulse varies with local time. Drifting from the injection site, energy dispersion causes the pulse to spread out spatially and so the height of the flux peak decreases. The response from the wavelet transform
will therefore also change with local time. It is therefore necessary to compensate for the diurnal variation in order to observe the injections with equal significance at all local times. The average diurnal variation was identified in Chapter 5, and it makes sense to use such a variation in customizing the data. However, as seen in plot (A) the diurnal variation is itself variable so it was decided that a ‘quiet’ time variation should be used to filter the data. This may mean that for active times the diurnal variation is underestimated, but using a stronger variation for the filter would impose an inverse local time variation on quiet parts of the time series. The quiet diurnal variation was determined, as before, from Fourier analysis of the local time averaged flux profile for data at times when $K_p$ was in the range 0o to 1+. This variation was found to be adequately described by a combination of two sine waves with periods 12 and 24 hours, as before, but with reduced amplitude. Figure 6.3 (B) shows the data shown in (A) after it has been filtered to remove a quiet diurnal variation. The wavelet transform shows that the dominance of the 24 hour periodicity has been only slightly diminished. The reason for this is that the diurnal variation itself comes predominantly from the injections, and removing a quiet time average does little to change their local time dependence.

Producing a statistical plot of the filtered data of the type presented in Chapter 5 showed little difference from figure 5.3, showing that the local time dependence was still in the data. A method was designed to compensate for the local time dependence of injection events more effectively, so that they could be located with equal significance irrespective of local time. This was achieved by ‘stretching’ the data. For each 30-minute local time bin two factors were calculated using the statistical plot data: the first was the factor needed to bring every point on the 95% curve up to the level of the highest point on the 95% curve, and the second factor was that needed to bring each point on the 5% curve down to the level of the lowest point on the 5% curve. Then for each time bin, the points above the median are stretched upward using the first factor, and the points below the median are stretched down using the second factor. Repeating the same statistical plot with this modified data set showed the 5% and 95% lim-
its of the data points were now approximately constant for all local times. This method for stretching the data cannot change the periodicities in the data: only the size of any oscillations present is changed, so what the stretching effectively does is remove the local time dependence of the injection signatures. Smaller events observed near midday will be brought up to equal significance with those observed close to midnight.

Figure 6.3 (C) shows the wavelet transform of the data from (B) after it has been stretched in this manner. Short period, transient waveforms are now much more obvious along the whole time axis. The remnants of the 24-hour periodicity have been removed.

In order to summarise the information in each scalogram, the modulus of the wavelet transform was averaged along the whole time axis for each period of oscillation: this was done for periods of integer hours between 1 and 35 hours.

The dashed line in figure 6.4 shows the average of the moduli for the scalogram of the raw flux data shown in figure 6.3 (A) for integer hour periods. A dominant peak is seen at 24 hours, the amplitude being comparable with that of the average local time profile. There is also a much smaller peak at 4 hours which may correspond to the most common injection periodicity. The dotted line shows the modulus average for the filtered data scalogram shown in figure 6.3 (B). The 24 hour period is still dominant but reduced. Other, shorter periods remain unchanged. The solid line in figure 6.4 shows the modulus average for the stretched data scalogram shown in figure 6.3 (C): the shorter periods are now enhanced and the local time dependence has been almost completely removed. There is a steep rise in modulus up to a period of 4 hours, then a more gradual decay. This implies that the separation time between injection events is most commonly around 4 hours.

There is another advantage to the method of averaging along the time axis to summarise the information in the scalogram. Recall the wavelet transform of the 'spike' signature in figure 6.2. It is possible that the scalograms contain many such signatures caused by the sudden flux increases associated with injections. This might 'contaminate' the scalogram, in that the colour is registering discontinuities
Figure 6.3: Wavelet transform showing substorm and diurnal frequencies: (A) raw data; (B) filtered data; (C) stretched data (see text). Units of flux are \( \text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1} \).
Figure 6.4: Wavelet transform determination of injection frequency. The average modulus of the wavelet transform for each integer hour period for the same sample of data as figure 6.3 in the three stages of the analysis (see text).

as well as true signatures of injections repeating at high frequencies. However, if a summation along the x-axis is carried out for a spike signature such as figure 6.2(A), the modulus sum for each period is the same. So, although a more intense response to a spike in the data is seen at low periods and a weaker but more extensive response at long periods, this effect is removed by summing the modulus along the x-axis. Therefore, plots such as figure 6.4 are not contaminated by discontinuities in the data.

The processing undertaken has increased the significance of the injections at later local times by effectively removing the diurnal variation. Hence it is found that for this section of data, the most common repeat time for injections is 4 hours.

Other authors have used different methods for estimating the time between substorm injections. Davis and Sugiura (1966) used observations of the $AE$ index to determine that polar disturbances tended to repeat with time intervals of about 4 hours, with the most active phase lasting about an hour. This periodicity
agrees well with the result obtained here. Borovsky et al. (1993) looked at the time between successive substorm onsets and found that under favourable conditions substorms could occur repeatedly and were therefore 'periodic', with a most probable time between onsets of 2.75 hours which was interpreted as the time for loading and unloading energy in a storage/release model of the substorm, or the time for disruption and recovery in the directly driven (by the solar wind) model of the substorm. Further 'random' (non-periodic) events were found to occur at lower frequencies: the average time between onsets was 5.74 hours. Prichard et al. (1996) used both geosynchronous electron data and AE data in a similar analysis and concluded that roughly periodic trains of substorms did occur, and therefore that the time to the next substorm could have some dependence on the time since the last one.

The response time of the AL index to the solar wind input parameter $vB_s$ ($B_s$ equal to $-B_z$ if $B_z$ is negative, or zero if $B_z$ is positive) was found to peak at 20 minutes and again at 60 minutes by Bargatze et al. (1985). The 20 minute peak was stronger for high activity whilst the 60 minute peak was stronger for moderate activity. So the magnetosphere responds quicker during high activity. Their interpretation is that 20 minutes may be the typical response time for a solar wind coupling mechanism, and 60 minutes the typical time for the release of stored magnetotail energy. This interpretation implies that both driven and unloading aspects apply to substorms. In contrast, Klimas et al. (1994) found that the substorm occurrence rate did not reflect the energy loading rate to the magnetosphere but remained constant with a 50 minute average period. With the SEM-2 data, 60 minutes is the shortest period that can be detected; however, here it is the most common injection repeat period which has been found rather than the shortest repeat time.

In modelling constant and time-varying solar wind energy input to the magnetosphere, Lee et al. (1985) found that tail reconnection occurred every 2–4 hours, and that the tail reconnection itself was a driven (i.e. forced) process.

From these examples it is apparent that the substorm frequency can vary by a few hours. Substorms and their associated injections may occur in quick
succession or there may be no injections for a day or two when the magnetosphere
is quiet (see figure 5.14). Here a value for the most common repeat period is
derived, rather than an average repeat time. The pre-processing (stretching) of
the data was useful because otherwise the wavelet transform would show local
time dependence in common with the injections. At periods of few hours, the
modulus of the wavelet transform would be highest midnight through to morning
and lowest in the afternoon.

6.2 Long Term Dependence of the Substorm In-
jection Frequency

This section looks at the long-term dependence of the substorm injection fre-
quency by using monthly-averaged wavelet transforms of the low energy SEM-2
flux at short periods. The data were processed as described in the previous sec-
tion, i.e. a quiet time average was filtered, then the data stretched to remove the
local time variation. Wavelet transforms were then performed on each month of
data.

The fact that the transform has been carried out on the processed data should
help prevent a particular ambiguity in the result. A high value of the wavelet
modulus could mean that the substorm frequency is higher, or that the frequency
is the same but the injections are larger. Since the data are stretched to the same
degree, the first case is likely to be true.

The top plot in figure 6.5 shows the monthly averaged wavelet modulus for
integer hour periods of 1 to 6 hours. The most common periods are 3-6 hours,
with 2 hours being less common and 1 hour less common still. In addition, each
line on the plot follows a similar pattern: when the 3-6 hour periods peak, the
shorter periods also peak.

Referring back to figure 6.2 (B), frequencies close to the 'best' frequency will
show a response but at a reduced amplitude. The lower plot in figure 6.5 shows the
ratio of the 1 to 6 hour curves from the top plot. This should remove the spread
of the response to nearby frequencies and show where the highest frequencies are
more important. On this plot the most noticeable feature is the large peak in the first half of 1994.

For comparison with these two plots, figure 6.6 shows the monthly averaged SEM-2 flux in each energy range, figure 6.7 shows the monthly averaged solar wind speed and figure 6.8 shows the monthly averaged $K_p$ index for the duration of the mission. The average $K_p$ is calculated from average $A_p$ then converted to $K_p$ using the standard conversion table (Knecht and Shuman, 1985).

![Graph of wavelet transform modulus for period of 1-6 integer hours for the duration of the mission (top) and the ratio of the 1:6 hour transforms.]

Figure 6.5: The monthly averaged wavelet transform modulus for period of 1-6 integer hours for the duration of the mission (top) and the ratio of the 1:6 hour transforms

Comparing figures 6.5 to 6.8 shows some agreement between peak substorm frequency, solar wind speed and $K_p$; for example, early 1989, mid-1991 and the first half of 1994 all have high wavelet amplitudes which coincide with times of high solar wind speed and $K_p$. At the end of 1990 the solar wind speed and $K_p$
Figure 6.6: Monthly averaged flux for each SEM-2 energy range, highest at the top, for the whole mission duration
Figure 6.7: Monthly averaged solar wind speed for the duration of the Meteosat P2 mission

Figure 6.8: Monthly averaged $K_p$ for the duration of the Meteosat P2 mission

are lower and so is the wavelet modulus. Similar features are seen in the E5-E4 flux. Figure 6.9 shows the monthly averaged solar wind speed and the substorm frequency amplitude ratio from figure 6.5, both smoothed with a 5-point boxcar and normalized. Parts of the graph show good agreement but the middle section does not: the overall correlation coefficient is only 0.4. However, for each of the six traces in the top plot of figure 6.5, the correlation coefficient with solar wind speed is between 0.6 and 0.8 for the smoothed traces (all with significance $>95\%$). Similar results are obtained for the $K_p$ index: for periods 1-6 hours the correlation coefficient with the wavelet transform is about 0.6-0.7 but worse for 1:6 hour ratio. Also not surprisingly the wavelet amplitude correlates better with the level of E5-E4 flux than E1 flux.

The results point overall to more frequent substorms occurring when solar wind speed is high, although there is obviously not a perfect correlation and other factors are involved. Energy coupling functions for the solar wind-magnetosphere interaction have been suggested which involve the southward component of the
Figure 6.9: Substorm frequency amplitude ratio from figure 6.5 (dashed line) and solar wind speed (solid line), smoothed and normalized.

IMF as well as solar wind speed (Akasofu, 1980, 1981). On the monthly time scale used here, no better correlation was found between the wavelet amplitude at substorm frequencies with empirical coupling functions \((vB_s)\) and \(v^2B_s\), where \(v\) is solar wind speed and \(B_s\) is equal to GSM \(B_z\) if the field is southward, or zero if the field is northward, see Bargatze et al. (1985) and references therein) than for solar wind speed. The use of long term data sets based on daily-averaged data is not ideal for this purpose. Another factor may be that the IMP-8 solar wind data is extremely patchy: about twelve days of data are missing per month in gaps of several days, and smaller gaps are common.

To investigate the coupling function on a shorter time scale, two months were studied with hourly-averaged solar wind data. The chosen months were August 1991 and April 1994 which both showed high moduli for 1-6 hour period wavelet transforms, and both experienced average solar wind speed over 500kms\(^{-1}\). The difference is that the 1:6 hour ratio is much higher for April 1994. For both months, maxima in \(vB_s\) and \(v^2B_s\) do correspond to peak values of the modulus of the wavelet transform for periods of a few hours. Daily averaged solar wind speeds for the two months are comparable. \(B_s\) was generally below 5nT for April 1994 but took several excursions above 10-15nT for August 1991. On average then, the coupling functions gave higher values for August 1991, so this does not explain why more frequent substorms seem to occur in April 1994, as is suggested by the ratio plot of figure 6.5.
6.3 Superposed Epoch Analysis of Peaks

The method of superposed epoch analysis (SPEA) uses ‘key times’ in a set of time series which have a common structure (e.g. Samson and Yeung, 1986). The key times $t_k$ in the times series $y(t)$ are used to find the estimator $\hat{s}(t)$ for the unknown common structure $s(t)$, by superposing the shifted time series $y_k(t - t_k)$:

$$\hat{s}_t = \frac{1}{n} \sum_{k=1}^{n} y_k(t - t_k)$$  (6.5)

where $n$ is the number of samples. The aim of the SPEA is to ‘average out’ differences between the samples whilst reinforcing the common structure.

The data used for this analysis were the lowest energy range (E5-E4) high time resolution data. Two criteria were used for identifying a peak in the data. Firstly, to be a peak the data point must be higher than the points before and after it. Secondly, the height of the peak was defined. Choosing a fixed flux threshold would lead to uneven samples with respect to local time: the flux is more likely to reach a given peak level in the post-midnight sector than at noon, due to the energy dispersion the injected electrons undergo as they drift. The threshold was therefore defined using the local time analysis in Chapter 5 as illustrated by figure 5.3. The level used was the 75-percentile line, i.e. the line below which 75% of observations occurred. This gave reasonably large and equal samples with $n$ between 1000-1200 peaks per hour of local time. The large sample ensured that, although the standard deviation on the sample was generally large, the standard error on any data point was small. It is important to carefully identify the key times in the data, since real effects may be ‘averaged out’ or spurious effects overemphasized (Reiff, 1993). For example, using a related criterion to identify when a substorm occurs, such as the development of the growth phase cigar-shaped distribution (Baker et al., 1978), would not yield the required results because an injection may or may not occur, the event time may be uncertain and the peak may be any size. Performing the SPEA based on the peaks themselves ensures that only true events are used and that the key times are correct.

Once the peaks had been found in the E5-E4 flux data, the key times $t_k$ were set at the times of the peaks. The same key times were then used to superpose
the flux in the other four energy bins and the spectral and anisotropy indices. Because the observation of substorm injection events has been shown to have a strong local time dependence the analysis was done separately for peaks occurring within one hour local time bins.

Some results are shown in figures 6.10 to 6.13. Figure 6.10 shows the SPEA results for the hour before midnight; figures 6.11 to 6.13 show the SPEAs for the hour after dawn, midday and dusk respectively. The first column of plots in each figure shows the SPEA for ±12 hours around the key times, for the flux in each energy bin and for the spectral and anisotropy indices. The dotted lines on these figures show the standard error which is small since the number in the sample was large. The next column of plots shows the overall local time average for each variable. For the lower energies the local time average peaks in the morning sector due to high flux of injected electrons and for the highest energy it peaks near midday where there is a high flux of trapped electrons. The average anisotropy index local time profile is positive on the nightside where the field is tail-like and the cigar phase develops prior to injections, and most negative around midday reflecting the trapped dayside distribution. The last column of plots shows the difference between the SPEA and the average profiles, as the SPEA minus the average. This subtraction was done because the 24 hour profiles will obviously depend on local time.

The first thing to notice about each of figures is that a diurnal variation of the same form as the average remains in the data after the average is subtracted. This means that when the high flux peaks occur, the local time variation in flux is more severe. So although flux is higher than usual at all local times, the difference in the day and nightside flux is greater than it would normally be. An experiment to subtract a high $K_p$ diurnal variation showed that, in order to get enough points for a smooth curve (by not making the $K_p$ threshold too exclusive) the curves became very close to the ones used here, so the results of the analysis would not change greatly. Secondly, at all local times, the subtraction of the average diurnal variation brings the E5-E4 and E4-E3 flux closer together. This effect is instrumental and is discussed further in Chapter 8. Large flux dropouts at all
Figure 6.10: Superposed epoch analysis of flux peaks over 40000 cm\(^{-2}\)sr\(^{-1}\)s\(^{-1}\)keV\(^{-1}\) observed between LT 2300-2400. The five lines on the flux plot are E5-E4, E4-E3, E3-E2, E2-E1, E1 from top to bottom.
Figure 6.11: Superposed epoch analysis of flux peaks over 40000 cm\(^{-2}\)sr\(^{-1}\)s\(^{-1}\)keV\(^{-1}\) observed between LT 0600-0700. The five lines on the flux plot are E5-E4, E4-E3, E3-E2, E2-E1, E1 from top to bottom.
Figure 6.12: Superposed epoch analysis of flux peaks over 40000 cm$^{-2}$sr$^{-1}$s$^{-1}$keV$^{-1}$ observed between LT 1200-1300. The five lines on the flux plot are E5-E4, E4-E3, E3-E2, E2-E1, E1 from top to bottom.
Figure 6.13: Superposed epoch analysis of flux peaks over 40000 cm$^{-2}$sr$^{-1}$s$^{-1}$keV$^{-1}$ observed between LT 1800-1900. The five lines on the flux plot are E5-E4, E4-E3, E3-E2, E2-E1, E1 from top to bottom.
energies before the injection are seen in the LT 0000 and LT 0600 plots. Energy dispersion can be observed in the LT 1200 with the peak observed latest at lowest energy. In the dusk curves the energy dispersion is partly masked because a few local, larger injections occur.

When an injection is observed near local midnight, the particle distribution switches suddenly from field-aligned to pancaked when the injection occurs. The anisotropy index at LT 2300 shows that this switch is more severe for large injections: the anisotropy index is more positive before the injection and more negative when the injection occurs. At later local times the effect on the anisotropy index is diminished as indicated by the reduced range covered by the fluctuation in the index at the injection time. Later in local time the fluctuations in anisotropy index also appear more sporadic. But what can be deduced is that the injected particles seem to have a characteristic anisotropy value: when the anisotropy index would normally be positive (the LT 2300 peak), the injection causes a dip in the anisotropy profile. When the anisotropy index would normally be strongly negative (the LT 1200 peak) the injection causes a peak in the anisotropy profile. The dawn and dusk peaks have less well defined effects because the anisotropy index is between the extremes at these times.

Figure 6.14 shows the low time resolution pitch angle distribution of the injected electrons, for all peaks $> 8 \times 10^4 \text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$ in 1990. There were 28 such peaks between 0000-0100LT, 50 between 0200-0300LT and 53 between LT 0400-0500. The figure confirms that following the injection the electron distribution has a slightly pancaked distribution. Baker et al. (1978, 1982a) observed the characteristic change from a cigar phase growth phase to a pancaked distribution at the expansion phase onset at geosynchronous orbit.

The average local time behaviour of the spectral index shows that the electron energy spectrum is hardest on the dayside, post-noon, and softest at around 0400 LT when peak low energy flux is seen, due to substorm injections. For an injection at 2300 LT (figure 6.10) the spectrum is harder on the dayside but softer on the nightside than in the average case. This represents the more compressed dayside and more taillike nightside field. The injection signature observed in the spectral
index, a steep drop, is well defined at all local times but does diminish in depth at later local times. This means that the injection causes a sudden softening of the spectrum locally, but that the local spectrum becomes less soft as the injected bunch of electrons drifts.

Overall, it is seen that when large injections occur, conditions in the magnetosphere as determined from electron flux and the spectral and anisotropy indices are more severe than in the average case. The fact that there is a greater difference between the electron flux on the dayside and on the nightside shows that the geomagnetic field is more compressed and shell splitting more pronounced than in the average case. The fact that the anisotropy index was more positive before the 2300 LT injection than it would be on average shows that, prior to large injections, the field is more taillike. The energy spectrum being harder on the dayside but softer on the nightside than in the average case reinforces this picture of more extreme dayside-nightside asymmetry, with more high energy particles observed on the dayside and less on the nightside than usual. These

Figure 6.14: Pitch angle distribution of injection peaks, for E5-E4 peaks $> 8 \times 10^5 \text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$ in 1990. Solid line 0000-0100 LT, dotted line 0200-0300 LT, dashed line 0400-0500 LT.
factors all lead to the conclusion that the largest injections accompany the most active conditions in the magnetosphere.

A further conclusion is that the injected electrons seem to have a characteristic anisotropy index (or range of anisotropy index) which although negative is between the extremes associated with the average diurnal variation. Calculated pitch angles illustrate the weakly pancaked distribution.

Finally, the data for the injection peaks can be used to show how the level of geomagnetic activity is manifested in the injection observations. The peak E5-E4 flux points which were used to find the key times for the superposed epoch analysis, i.e. the top 25% of flux observations, are plotted as a function of $K_p$ for each hour of local time across the nightside in figure 6.15. The lower cutoff of the sample changes from plot to plot because the position of the 75-percentile line changes with local time. It is clear from the distribution of these flux points that there are two components in this sample: a high flux component which is electrons being injected, and a lower flux component which is pre-existing. The two components are seen in the pre-midnight plots because less injections happen here, whereas at the time when peak flux occurs in the statistical plot in Chapter 5 (figure 5.3), i.e. between about LT 0300-0600, most of the top 25% of E5-E4 flux points are associated with new injections. The 1800-1900 LT bin is the earliest that the new injection component is seen. From dawn, the flux of the electron bunch decreases round to the dusk level. The interesting point here is that it is possible to see fresh injections of electrons with flux $\geq 1e5 \text{cm}^{-2} \text{sr}^{-1} \text{s}^{-1} \text{keV}^{-1}$ across the nightside, although they are scarcer pre-midnight, if $K_p$ is high enough. The high flux points occupy a smaller range of $K_p$, concentrated towards higher $K_p$, the earlier they are observed. This is why the peak of the 95% line in figure 5.3 moves earlier in local time at high $K_p$.

6.4 Summary

- The repeat period for substorms was found to be most commonly about 4 hours.
Figure 6.15: Peak flux versus $K_p$ across the nightside: the top 25% of E5-E4 flux observations are plotted versus $K_p \times 10$ for each hour of local time across the nightside.
• Evidence was presented that substorm frequency depends at least partly on solar wind velocity.

• Substorm injected particles are shown to occupy a specific range of the second order anisotropy index, which may be observed as a decrease or increase in the local value of the index depending on the local time of the observation, corresponding to a weakly pancaked distribution.

• Injections occur earlier at high $K_p$.

6.5 Conclusion

Wavelet analysis was applied to the question of substorm frequency. Recent publications propose that series of periodic substorms can occur when conditions in the magnetosphere are favorable (Borovsky et al., 1993, and Prichard et al., 1996). Wavelet analysis is a useful tool for this application since it can locate short durations of a particular frequency in a long data set. The most common period for substorm recurrence found by this method is 4 hours, which is comparable with the results of other authors (e.g. Borovsky et al. suggest the most probable and average inter-substorm times to be 2.75 and 5.74 hours respectively).

Better results may be obtained by using an asymmetric wavelet for the detection of substorms. The Morlet wavelet which was used here is a symmetrical wave packet which therefore makes it very versatile. However, a wavelet with a shape closer to a typical substorm signature, i.e. having a dip corresponding to a growth phase flux decrease followed by a large peak, might prove more effective.

Another possible application for wavelets lies in the filtering of data. In order to remove the diurnal variation from the flux time series, the method of filtering a quiet average followed by stretching the data was used. Although effective this method is laborious and there is also the problem of establishing a 'quiet' average. A wavelet transform would provide the amplitude of a 24 hour oscillation at every point along the time series, which could form the basis of a filtering algorithm tailored to the level of activity.
Superposed epoch analysis was used to demonstrate the typical signature of a flux peak in the anisotropy and spectral indices. The softening of the energy spectrum was observed to decrease as the injected electrons drifted and were energy dispersed. The injected particles were observed to occupy a specific range of anisotropy indices, indicating an isotropic to pancaked distribution. It was also seen that for large peaks, the injections were seen in the E1 range, although this is generally not the case. It was also seen that injections are observed as early as dusk when $K_p$ is high.
Chapter 7

Long Term Dependences

So far, the variability of the geosynchronous electron environment has been evaluated on the time scales of substorms and local time. In Chapter 4 seasonal behaviour was also seen which showed differences from the predictions of the Tsyganenko 89 magnetic field model. In this chapter the geosynchronous electron environment is considered on time scales longer than the diurnal variation. A pattern is clearly seen on an annual time scale in figure 6.6, most clearly in the E5-E4 flux. The explanation for the semi-annual variation in activity having maxima at both equinoxes and minima at the solstices is that at the equinoxes the Earth’s magnetic axis makes the largest possible angle with the normal to the plane of the Earth’s orbit ($\sim 35^\circ$). Thus the solar wind $B$ in the plane of the Earth’s orbit will have a component $B \sin 35^\circ$ which, if southward, leads to significant coupling between the magnetosphere and solar wind (Kivelson and Russell, 1995, and references therein). Solar related periodicities discussed here include the 11-year solar cycle and the 27-day solar rotation period.

7.1 Solar Cycle

The concept of a solar cycle was based on sunspot number: many sunspots are observed at solar maximum but few at solar minimum. We are now in solar cycle 23. Since cycle number 1 in 1755, the length of the solar cycle has varied in length from 9.0 to 13.7 years, but is on average about 11 years. The rise time of the
cycle is on average 4.3 years, shorter than the fall time which is on average 6.7 years, and the last solar maximum occurred in July 1989 (Priest, 1995). Since the last solar minimum occurred in 1996, the SEM-2 data set covers a period from solar maximum through the declining phase of the cycle. In fact the Sun's magnetic field reverses every 22 years, leading to the conclusion that the 11-year cycle is half the fundamental solar periodicity.

![Figure 7.1: Monthly averaged sunspot number for the duration of the Meteosat P2 mission](image)

Figure 7.1 illustrates part of the solar cycle in the diminishing sunspot number from the maximum in 1989 towards the minimum. Comparison with figures 6.6, 6.7 and 6.8 illustrates some of the solar cycle interdependences. Monthly averaged flux in all the SEM-2's energy ranges was enhanced at solar maximum and decreased over the next couple of years. However a second enhancement is seen during 1993-1994 particularly at higher energies, which is accompanied by increased solar wind speed and $K_p$, although sunspot number was low. The explanation for this is in two components of solar activity. Firstly, at solar maximum, the number of both sunspots and solar flares is high; the occurrence of flares for part of solar cycle 22 is illustrated by figure 7.2. Sims (1992) showed that the number of substorms occurring per month correlated with the number of solar flares for a five year period in the declining phase of solar cycle 21. By 1994 both sunspot number and flare occurrence are reduced, but solar wind speed and the SEM-2 geosynchronous flux have increased. The second component of solar activity is high speed solar wind streams (HSSWSs) associated with coronal holes. Coronal holes are regions of the corona which are cooler and less dense than their surroundings, and are on open magnetic field lines; because the field
lines are open, plasma can stream out at high speed, decreasing mass and energy at the hole. Geomagnetic disturbances result from the Earth’s path crossing the HSSWSs rooted in the coronal holes. Solar coronal holes were identified as a source of recurrent geomagnetic disturbances by Neupert and Pizzo (1974). Magnetospheric 27-day periodicities were originally believed to be due to the Earth’s crossing of the IMF sector boundaries, but are now believed to be caused by HSSWSs, and because the HSSWSs may persist for many solar rotations, 27-day recurrences in geomagnetic activity are observed (e.g. Lindblad et al., 1989). The HSSWSs are characterized by a steeply increased solar wind velocity which lasts for several days. Coronal holes have been observed to cause extended periods of geomagnetic disturbance in the declining phase of solar cycles. Large magnetic storms were reported for late 1993 and early 1994, and these were recurrent storms caused by HSSWSs (Watari, 1997). The geomagnetic field was observed to be continuously disturbed by a high-speed solar wind stream for several solar rotations and these disturbances were found to be associated with a large coronal hole expanding from the south pole of the Sun and causing high solar wind speed.

The solar wind speed and \( K_p \) increases in 1991 follow a large magnetic storm which occurred in March. Although there is no SEM-2 data for the time of the storm, figure 6.6 shows that the low energy electron component was boosted in the following months, although the high energy component is less affected. The \( K_p \) index was also high.

Substorm activity therefore has two solar cycle related components: one related to flare-associated sudden storm commencements (SSCs) which peaks at solar maximum, and one related to the high speed solar wind streams causing 27-day recurrent substorms, which peaks in the declining phase of the sunspot cycle (Venkatesan, 1991). It was seen in Chapter 6 that substorm frequency depends at least in part on solar wind velocity. The minimum in geomagnetic activity also correlates with solar wind velocity and occurs a year or so after sunspot minimum (Feynman, 1985).
Figure 7.2: Solar flare occurrence, 1991-1996, as measured by BATSE (Burst and Transient Source Experiment) on board the CGRO (Compton Gamma-Ray Observatory).
Table 7.1: SEM-2 yearly temperatures and densities

<table>
<thead>
<tr>
<th>year</th>
<th>soft component</th>
<th>hard component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>kT/keV</td>
<td>n/10^{-3}cm^{-3}</td>
</tr>
<tr>
<td>1989</td>
<td>30.8</td>
<td>4.31</td>
</tr>
<tr>
<td>1990</td>
<td>31.0</td>
<td>4.14</td>
</tr>
<tr>
<td>1991</td>
<td>29.3</td>
<td>3.99</td>
</tr>
<tr>
<td>1992</td>
<td>29.3</td>
<td>3.57</td>
</tr>
<tr>
<td>1993</td>
<td>30.5</td>
<td>3.52</td>
</tr>
<tr>
<td>1994</td>
<td>31.8</td>
<td>4.07</td>
</tr>
<tr>
<td>1995</td>
<td>30.7</td>
<td>3.47</td>
</tr>
</tbody>
</table>

Energy spectra

Energy spectra calculated for each year of SEM-2 data show trends in the behaviour of the hard and soft electron components. Temperatures and densities calculated from the fitting of two Maxwellsians as in Chapter 2 are shown in table 7.1 (1988 has been omitted since only two months of data are available). These spectra use averaged flux and therefore the temperature and density are those associated with this average flux, rather than the average temperature and density. Tables 7.2 and 7.3 show the results for high $K_p (> 6)$ and low $K_p (\leq 1+)$ evaluated separately.

Figure 7.3 summarises the results from the three tables. In the long term the temperature of the hard component $T_h$ is more variable than that of the soft component $T_s$. $T_h$ is lowest the year after solar maximum and highest in the declining phase of the solar cycle. $T_s$ is lowest mid-cycle, and higher at both solar maximum and in the declining phase. The density of the hard component $n_h$ is lowest mid-cycle, high at solar maximum and higher still in the declining phase. The density of the soft component $n_s$ is high in the declining phase but higher at solar maximum. Thus, it seems that the high geomagnetic activity at solar maximum has more effect on the low energy electron population whereas the activity in the declining phase has more effect on the high energy electron
### Table 7.2: SEM-2 yearly temperatures and densities for high $K_p$

<table>
<thead>
<tr>
<th>year</th>
<th>soft component</th>
<th>hard component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$kT/\text{keV}$</td>
<td>$n/10^{-3}\text{cm}^{-3}$</td>
</tr>
<tr>
<td>1989</td>
<td>30.4</td>
<td>5.57</td>
</tr>
<tr>
<td>1990</td>
<td>30.3</td>
<td>6.27</td>
</tr>
<tr>
<td>1991</td>
<td>28.8</td>
<td>4.99</td>
</tr>
<tr>
<td>1992</td>
<td>27.0</td>
<td>5.01</td>
</tr>
<tr>
<td>1993</td>
<td>30.3</td>
<td>5.66</td>
</tr>
<tr>
<td>1994</td>
<td>32.5</td>
<td>5.71</td>
</tr>
<tr>
<td>1995</td>
<td>30.6</td>
<td>5.26</td>
</tr>
</tbody>
</table>

### Table 7.3: SEM-2 yearly temperatures and densities for low $K_p$

<table>
<thead>
<tr>
<th>year</th>
<th>soft component</th>
<th>hard component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$kT/\text{keV}$</td>
<td>$n/10^{-3}\text{cm}^{-3}$</td>
</tr>
<tr>
<td>1989</td>
<td>32.4</td>
<td>2.79</td>
</tr>
<tr>
<td>1990</td>
<td>33.0</td>
<td>2.51</td>
</tr>
<tr>
<td>1991</td>
<td>31.2</td>
<td>2.36</td>
</tr>
<tr>
<td>1992</td>
<td>31.7</td>
<td>2.24</td>
</tr>
<tr>
<td>1993</td>
<td>31.4</td>
<td>2.49</td>
</tr>
<tr>
<td>1994</td>
<td>31.5</td>
<td>2.12</td>
</tr>
<tr>
<td>1995</td>
<td>31.2</td>
<td>2.05</td>
</tr>
</tbody>
</table>
Figure 7.3: Temperature and density for the two components for each year: top left $T_h$, top right $T_s$, bottom left $n_h$, bottom right $n_s$. Solid line for all $K_p$, dotted line for low $K_p$, dashed line for high $K_p$

population. This can inferred to some extent from figure 6.6 since the highest E1 monthly averaged flux occurs in 1994, whereas the highest E5-E4 monthly averaged flux occurs in 1989.

Cayton et al.'s (1989) geosynchronous electron data showed that on the time scale of a day, neither $T_h$ or $n_h$ varied significantly, although the hard component did exhibit a diurnal variation in $n_h$ (with higher density observed at noon than midnight), whereas $T_s$ and $n_s$ showed much more variability. On an intermediate time scale (~months) $T_h$ showed little variability whereas $n_h$ could be vastly reduced by large magnetic storms. The decreases in $n_h$ often coincided with large increases in $n_s$; both $T_s$ and $n_s$ showed large variations during times of high $K_p$ especially during substorm injections, when fluctuations of $n_s$ greater than two orders of magnitude were seen. A solar rotation (27-day) periodicity was also found in $n_h$. In the SEM-2 data, for time scales of years, the soft component appears to be more stable than the hard component but this may be partly due to a data handling problem which is discussed in chapter 8. However this problem
should not affect when the highest and lowest temperatures and densities occur.

In terms of the statistical plots which were presented in Chapter 5 both populations are seen to develop through the course of the solar cycle. The E1 plots are shown in figure 7.4 and the E5-E4 plots in figure 7.5. Each plot shows successive years as a solid line and the preceding year as dotted line. The 95% level of the E1 flux is lowest the year after solar maximum and highest in 1994, at the time of the recurrent storms. The same is true of the median, which drops again in 1995. The 95% level of the E5-E4 flux is reasonably constant, the only variable feature being that substorms seem to occur later at quarter-cycle and in 1995 than at maximum or in 1994. The median flux on the other hand shows a small morning enhancement at solar maximum and a larger rise in 1994. A curious feature of both the E1 and E5-E4 plots is that the effect of the nightside dropouts is small at solar maximum becoming more severe at quarter cycle. Closer examination of the data shows there are indeed far fewer nightside dropouts in the solar maximum year than at other times.

<table>
<thead>
<tr>
<th>year</th>
<th>E1 flux</th>
<th>E5-E4 flux</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>amplitude</td>
<td>T/LT</td>
</tr>
<tr>
<td>89</td>
<td>8.6e3</td>
<td>10.2</td>
</tr>
<tr>
<td>90</td>
<td>8.9e3</td>
<td>10.0</td>
</tr>
<tr>
<td>91</td>
<td>8.0e3</td>
<td>10.8</td>
</tr>
<tr>
<td>92</td>
<td>7.5e3</td>
<td>11.1</td>
</tr>
<tr>
<td>93</td>
<td>9.3e3</td>
<td>11.6</td>
</tr>
<tr>
<td>94</td>
<td>14.3e3</td>
<td>11.5</td>
</tr>
<tr>
<td>95</td>
<td>10.8e3</td>
<td>11.7</td>
</tr>
</tbody>
</table>

Table 7.4: Yearly average diurnal variations: the amplitude of the average diurnal variation in cm^{-2}sr^{-1}s^{-1}keV^{-1} and the time T of the peak flux in local time.

In Chapter 5 Fourier analysis was used to identify the form of the average diurnal variation. Applying this method to each year's E5-E4 flux data allows the comparison of the position of the peak amplitude of the diurnal variation in
Figure 7.4: Yearly El statistical plots. Each plot shows one year as a solid line and the preceding year as dotted line. The lines represent the 5%, 45%, 55% and 95% levels of observations.
Figure 7.5: Yearly E5-E4 statistical plots. Each plot shows one year as a solid line and the preceding year as dotted line. The lines represent the 5%, 45%, 55% and 95% levels of observations.
local time, which gives an idea of the influence of injections for each year. Table 7.4 gives the amplitude of the average diurnal variation in cm$^{-2}$sr$^{-1}$s$^{-1}$keV$^{-1}$ and the time $T$ of the peak flux in local time for the E1 and E5-E4 flux. For the E1 flux the average diurnal variation is approximated to a 24 hour period sine wave and for the E5-E4 flux it is a combination of two sinewaves with periods 24 and 12 hours. The amplitude of the E5-E4 flux local time variation is higher at solar maximum then declines, then peaks again in 1994. The time of peak amplitude is earliest at solar maximum, latest at quarter-cycle and early again in 1994. The E1 flux amplitude experiences a much stronger peak in 1994. The local time position of E1 peak flux is earlier around solar maximum than towards solar minimum.

In terms of substorms, the number of peaks above a fixed level of E5-E4 flux ($6 \times 10^4$cm$^{-2}$sr$^{-1}$s$^{-1}$keV$^{-1}$) between 2200 and 0400 LT is actually higher in 1994 (937 peaks) than 1989 (798 peaks). This figure is lower between these years and drops as low as 485 peaks for 1992. Of these peaks, a higher percentage occur before 0200 LT in 1994 (60% as opposed to 53% for 1989). The average $K_p$ was higher for 1994 than for 1989 (3+ as opposed to 3o) so this conforms with the conclusions of earlier chapters that injections are observed earlier at higher $K_p$.

In fact the $K_p$ index average for 1991 was the highest of the years shown here at 4-, but the activity in this year did not seem to have the sustained effect that is seen at either of the geomagnetic activity peaks: the number of similar peaks (normalised to twelve months of data) was 728, lower than for 1989 or 1994.

The high energy component of the geosynchronous electron population has been studied by several authors for extended time scales. High energy electron flux (>300keV) at geosynchronous orbit has an ~11-year periodicity which is out of phase with the sunspot cycle, with lowest flux observed at solar maximum and highest flux some time (1-2 years) before solar minimum (Baker et al., 1986, and Belian et al., 1996). For the E1 flux in figure 6.6 the maximum in the declining phase is indeed the highest flux feature, but high flux is also seen near solar maximum. It has however been shown that injections are sometimes seen
in this energy range during large substorms, which would explain the second enhancement in common with the E5-E4 flux at solar maximum.

Belian et al. looked for long term behaviour in relativistic electrons at geosynchronous orbit using CPA and SOFA data for a time sequence exceeding one solar cycle. Of three energies of electrons, the higher two (>300keV and >1.4MeV) showed convincing solar cycle behaviour, with minimum flux occurring at solar maximum, and maximum flux occurring between solar maxima, a couple of years before solar minimum. The lowest energy (>65keV) did not show a convincing solar cycle dependence. Given the differences between the E1 and E5-E4 flux in figure 6.6, this was probably due to a masking effect of the different solar cycle behaviour of the electrons at the low energy end of the range. Belian et al. found that high speed solar wind streams, although related to relativistic electron flux in the short term (flux lagging the solar wind speed by a few days) have a weak long-term correlation; this implies that there are other factors at work in establishing the solar cycle dependence of the energetic trapped electrons. Using 15 years of data it was seen that the solar wind velocity could stay high for extended periods during which the relativistic electrons declined. Thus high relativistic electron flux cannot occur at low solar wind speeds, but high solar wind speed does not automatically produce high flux, so although solar wind speed is obviously an important factor, it is not decisive.

Blake et al. (1997) presented representative cases of solar wind speed increases and their effect on the relativistic electron population of the outer zone during a solar minimum period. Satellite 1994-026 provided four electron channels (>1.5Mev to >8.5MeV) and WIND provided solar wind speed and density, and IMF measurements. It was found that a large increase in the solar wind speed, a pressure pulse and a southward turning of the IMF resulted in a strong decrease in the relativistic electron population for a few hours, followed by an increase over the next few days to a level higher than before the event. For a similar event where the IMF turned northward instead of southward, the relativistic electrons were unaffected. Thus the southward turning of the IMF seems to be an important factor for the relativistic electrons to be affected by changes
Figure 7.6: Flux of electrons >1.4MeV versus solar wind speed, 27-day averages, from Belian et al. (1996).

Figure 7.7: Flux versus solar wind speed, 27-day averages: stars E5-E4 flux, diamonds E1 flux.
in solar wind speed and density.

Belian et al.'s (1996) results for flux of electrons >1.4MeV from geostationary observations plotted versus IMP-8 solar wind speed are shown in figure 7.6. Figure 7.7 shows flux versus solar wind speed for the SEM-2 E5-E4 and E1 channels. The plot points are 27-day averages as used by Blake et al. but it is noted that the plot differs little from monthly averages. A similar conclusion is reached in that, although the distribution of points suggests a tendency for higher flux to occur at higher solar wind speed, the relationship is not straightforward and high speeds do not necessarily produce higher flux in either electron population.

The widely used AE-8 and AP-8 models have versions for solar maximum (AE-8 MAX, AP-8 MAX) and for solar minimum (AE-8 MIN, AP-8 MIN). However, for \( L > 5.5 \) the models are identical (Pierrard and Lemaire, 1993) so solar cycle variability in the outer zone is not described by the AE-8 and AP-8 models at all. Also, these models were based on data from a moderate solar cycle, so in the event of a much stronger cycle, these models could drastically underestimate the energetic particle populations occurring at solar maximum (Vampola, 1989).

7.2 Solar Rotation

The Sun undergoes differential rotation, taking from 24 days at the equator to 34 days at the poles for one rotation. Taking into account the Earth’s motion relative to Sun, there is an effective 27-day period for recurrent disturbances in the magnetosphere due to solar phenomena. The 27-day periodicity in geomagnetism has long been investigated; for example, Chree and Stagg (1927) analysed the distribution of quiet and disturbed days from a predetermined average situation and attributed the cause of the 27-day periodicity they identified to solar processes. A 27-day cycle in the intensities of trapped electrons with energies \( \geq 280\text{keV} \) throughout the outer zone \((L \geq 3.5)\) was found by Williams (1966), who attributed it to effects from the solar wind and IMF, and Baker et al. (1986) found a 27-day periodicity in 3-10 MeV electrons at geostationary orbit.

The 27-day periodicity in geomagnetic activity shows dependence on the
Figure 7.8: Wavelet transforms of daily averaged data: (A) E1 flux (B) E5-E4 flux (C) $K_p$ index. Units of flux are $\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$. Minimum of linear colour scale is zero, maximum for E1 flux is $2.2\times10^3$ $\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$, for E5-E4 flux is $1.2\times10^4$ $\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}$, for $K_p$ is 1.1
sunspot cycle. Using a 27-day recurrence index, Sargent (1985) showed that periods of 27-day recurrent behaviour tended to occur towards the end of the 11-year sunspot cycle. These periods had abrupt starts and ends, and lasted longer during even-numbered sunspot cycles, which identifies with the 22 year solar cycle comprising two non-identical 11-year cycles. Cliver et al.’s (1996) investigation of sunspot numbers and the aa index showed that geomagnetic activity is enhanced during the first half of odd-numbered solar (11 year) cycles and the second half of even-numbered cycles. Peak sunspot numbers were found to be higher in odd-numbered cycles whilst the 27-day recurrence in geomagnetic activity was found to be stronger in even-numbered cycles. This fits with SEM-2 observations of high energetic electron flux and more frequent substorms occurring in 1994 (Chapter 6), in the declining phase of cycle 22.

Baker et al. (1990) showed that cross-correlations of the geosynchronous 3-40MeV electron flux with the Kp and AE indices and with solar wind velocity had maxima at multiples of 27 days, and weaker correlations existed half-way between the 27-day multiples. They showed that whereas the indices Kp and AE respond directly to solar wind speed, the high energy flux peak 2 days after the solar wind streams maximum speed. The 200-300keV electron flux at geostationary orbit was shown to lag low energy (43-60keV) substorm-enhanced fluxes in Chapter 5. The high energy electron flux lags the low energy electron flux because whereas the low energy electrons are directly produced by substorms, the high energy electrons are not.

13 or 13.5 day periodicities have been attributed either to the existence of two solar wind streams at the same time, or else to the sectored structure of the IMF: the sectors have a low-velocity current sheet between them which separates the oppositely directed field lines along which the high speed streams flow. Verma and Joshi (1994) found a 9-day periodicity in the occurrence of high speed solar wind stream events, which is one third of the periodicity of solar rotation. Verma and Joshi suggest that the 9-day period is an energy build-up time for coronal holes to produce the high speed solar wind stream events.

Figure 7.8 shows wavelet transforms of daily averaged data for the whole of
Figure 7.9: Wavelet modulus average from figure 7.8.

Figure 7.10: The 27-day wavelet transform throughout the mission.
1990, for periods between 5 and 35 days. (A) shows E1 flux, (B) E5-E4 flux and (C) \( K_p \) index. All three plots show a periodicity between 25 and 30 days for the first hundred days or so of the plot; the E5-E4 flux and \( K_p \) index have a similar feature between 200-300 days which is much weaker in the E1 flux. Even though this is the year after solar maximum, there is clearly a 27-day periodicity although it is variable. Overall the E5-E4 flux and \( K_p \) index scalograms are very similar at all periods shown and have similar features, whereas the E1 flux has some similar features but also differences. A 9-day periodicity can be seen for about the last 100 days on the E5-E4 flux and \( K_p \) plots. Several features are also seen at 15-18 days in the middle section of the E5-E4 flux and \( K_p \) plots.

Figure 7.9 shows the transforms from the three scalograms in figure 7.8 averaged for each integer hour period. The three traces have been normalised. The E5-E4 flux has three peaks at 9, 16-17 and 27-28 days. The E1 flux has peaks at 10 and 23-24 days. The \( K_p \) index has the first two peaks in common with the E5-E4 flux but the third in common with the E1 flux. The 27-day period wavelet transforms for the whole E1 and E5-E4 flux data sets are shown in figure 7.10. The 27 day periodicity is variable and strongest during 1994 in the E5-E4 flux, and strongest both in early 1989 and 1994 for the E1 flux. On the E1 trace especially, it is clear that the 27-day periodicity gets stronger following solar maximum towards solar minimum.

### 7.3 16-day Periodicity

*Arquilla* (1993) noted from his observations of maximum daily voltages on IUE (International Ultraviolet Explorer), which he used to infer the maximum daily particle flux, the occurrence of a 15-17 day periodicity which he likened to the more commonly observed 13-day periodicity. No 15-17 day periodicity has been reported for the IMF (*e.g. Gonzalez and Gonzalez*, 1987). The wavelet transform in figure 7.8 showed enhancements at periods between 15 and 18 days and a peak at 16-17 days is seen in figure 7.9.

A 16-day periodicity has been observed in neutral wind data, in ground mag-
netic data and in ionosonde data, and is commonly interpreted as the effect of a planetary wave. Two important types of wave in the atmosphere are gravity waves and planetary or Rossby waves (Salby, 1992). Gravity waves are also called buoyancy waves, because the restoring force for the wave comes from buoyancy in atmospheric strata. These waves have short wavelengths (from metres to tens of kilometers) and periods of minutes to hours. Rossby waves occur on much larger scales (thousands of kilometers) and on time scales of days. At their largest dimensions they are often referred to as planetary waves, and are observed as undulations in global circulation patterns in the atmosphere or ocean. For these waves, the Coriolis force acts as the restoring force, due to the rotation of the Earth under the atmosphere.

Both these types of wave can propagate vertically and horizontally and thus can influence the atmosphere far from where they were originally excited. These waves are distinct from atmospheric tides, which are oscillations occurring at subharmonics of lunar or solar periods. Tides are an important part of the dynamics of the middle atmosphere, between 15-110km, and dominant in the upper mesosphere and lower thermosphere at 70-110km (Forbes, 1989).

Rossby Waves

Rossby waves were proposed by C. G. Rossby in the 1930s. They are a product of the Earth's rotation and have large wavelengths (thousands of kilometers) and slow movement. The cause of Rossby waves is the variation of the Coriolis parameter, $f$, with latitude. $f$ is given by:

$$f = 2\Omega \sin \phi$$

(7.1)

where $\Omega$ is the Earth's angular velocity and $\phi$ is latitude (Salby, 1992). $f$ is also called planetary vorticity. The vorticity of a fluid element in the atmosphere is defined as its spin about its own axis, and is a vector directed along the rotation axis, perpendicular to the surface of the fluid element. For horizontal fluid motion, what $f$ represents is the magnitude of the Earth's vorticity in the local vertical direction. The important factor in the creation of the Rossby wave is that, in the
absence of vorticity sources and sinks, the vorticity of an air parcel is conserved as it moves in the atmosphere. The absolute vorticity is the sum of the relative vorticity (i.e. that which is observed on the Earth’s surface) and the planetary vorticity (i.e. that due to the spin of the Earth) (Wells, 1986).

Consider an air parcel with zero vorticity on the equator. If it moves away from the equator it must gain vorticity because its latitude has changed (equation 7.1). But its absolute vorticity must be conserved (i.e. must remain zero), so the air parcel gains a component of relative vorticity equal to its planetary vorticity but of the opposite sign. In the Northern hemisphere $f$ is positive so the relative vorticity induced will be clockwise (positive vorticity is measured anticlockwise). So the air parcel moving North acquires a clockwise component of motion moving it back to the equator. This process will repeat about the equator causing a meridional oscillation about the equator. This is how a planetary wave forms: meridional displacement provides a restoring force, similar to that from buoyancy in gravity waves. More extreme weather can lead to increased amplitudes of Rossby waves (Wells, 1986). Global 'free' (i.e. unforced) oscillations formed in this manner are termed free Rossby waves or Rossby normal modes (Forbes, 1996). Most commonly observed are 5, 10 and 16 day period waves.

The mechanism whereby the planetary wave, which is excited in the lower atmosphere, can penetrate the ionosphere is not properly understood, but two theories have been put forward (Forbes, 1996). The first idea is that the wave might induce changes in the ionic composition and therefore the recombination processes in the ionosphere. The second, more favoured suggestion is that of the ionospheric wind dynamo. The dynamo region lies between 100-170km, and here the ion gyrofrequency is comparable to the ion-neutral collision frequency. In contrast the electron gyrofrequency is far in excess of the electron-neutral collision frequency. The neutral winds move ions across field lines, whereas electrons remain fixed to them, therefore a current is induced and this causes perturbations in the magnetic field measured at the ground.

Forbes et al. (1992) observed 16-day and 34-day period oscillations in $\Delta H$, the perturbation of the surface horizontal component of magnetic field during
Figure 7.11: Wavelet transforms over 10-20 day periods. From the top, E5-E4 flux, $K_p$, $H$ for four stations (from top BNG, NAQ, TUN, FAR), foF2 for two stations (from top TUN, FAR), and solar wind speed.
Figure 7.12: 16-day wavelet transforms: from the top, E5-E4 flux, $K_p$, $H$ for four stations, $foF2$ for two stations, and solar wind speed.
January/February 1979. The 34-day period was attributed to solar causes, since most of the spectral energy of 10.7 cm solar flux below the 60-day period frequency was found to lie between 20- and 40-day periods, i.e. connected with solar rotation. The 10.7 cm solar flux is a measure of ionizing solar radiation which changes the conductivity of the E region ionosphere. It is used as an index of solar activity and shows solar cycle dependence. No corresponding solar cause was found for the 16-day periodicity and this was interpreted by Forbes as being due to a free Rossby mode which had penetrated into the ionosphere from the stratosphere.

Parish et al. (1994) also observed several oscillations in $\Delta H$ with periods ranging from 2.5 days to 16 days, all of which could be due to the propagation of Rossby free modes into the ionosphere. These waves, observed throughout 1979, included 6, 7 and 9 day periods, with 16 day periodicity being one of the strongest planetary wave periodicities seen in the $\Delta H$ data. The 13.5 and 27 day periodicities associated with solar rotation were also seen.

Forbes et al. (1995) observed a 16-day oscillation in the mesosphere and lower thermosphere during January/February 1979. Data were 95 km altitude wind measurements obtained by radar from Obinsk, Russia (54°N, 38°E) and Saskatoon, Canada (52°N, 107°W). These show a ±10 m/s wind oscillation with a period of 16 days, and other oscillations with periods of 5 and 10 days; all these periodicities correspond to free Rossby modes. Although through modelling they show that the 16-day wave can penetrate up to these altitudes from the stratosphere and troposphere, their model does not account for significant penetration of the 16-day wave above 100 km. Their suggestion is that the 16-day wave observed in the ionosphere either results from a modulation of an upward propagating tidal wave or else is excited in situ. The propagation of the 16-day periodicity above 100 km has therefore yet to be convincingly explained.

Because of the presence of the 16-day wave in $H$ it was decided to compare some magnetic data with the SEM-2 flux. Figure 7.11 shows the average wavelet transform modulus for 1990, for periods of 10-20 days, for the E5-E4 flux, $K_p$, $H$ from four stations foF2 from two stations. The data used were daily averages from 1990. Peaks at 15-17 days are seen in the SEM-2 flux, $K_p$ index, and in $H$.  

232
although not in foF2.

H

Magnetic H data were obtained from Bangui (BNG, 4.3°N, 18.6°E) and Narsarsuaq (NAQ, 61.1°N, 314.8°E) observatories. These stations were chosen from a selection available through the World Data Centre for Geomagnetism: they were selected for being, respectively, approximately equatorial and auroral zone. The longitude was not considered as important since the waves are supposed to be global.

foF2

The foF2 parameter is obtained from soundings of the atmosphere. Radio wave propagation provides an important technique for remote sensing of the ionosphere (Hargreaves, 1979). For ionospheric sounding, pulses of radio waves are transmitted vertically and the echoes are timed. Unique relationships exist between sounding frequencies and the ionization densities which can reflect them: foF2 is the maximum radiowave frequency capable of reflection from the F2 region of the ionosphere. Similarly, critical frequencies of the E and F1 regions are called foE and foF1.

An ionosonde or ionospheric sounder has a transmitter and receiver slowly sweeping in frequency. The resulting ionogram, which gives the height of reflection versus frequency, can then be interpreted as a profile of electron density versus altitude, using the relation

\[ [f_N(kHz)]^2 = 80.5N(cm^{-3}) \]

and the height of the reflection can be obtained from the time of the echo.

Two stations were selected for which both magnetic data and foF2 were available. These were Tunguska (61.6°N, 90°E) and Faraday (65.3°S, 295.7°E).

Figure 7.12 shows the 16-day wavelet transforms for the length of 1990, for the same data used to construct figure 7.11. Similar features can be seen on
each graph, the largest peak being at about 100 days. In order to establish the
significance of the 16-day waveform in each variable it is necessary to compare the
amplitude of the wave to the data series itself. In each case the peak amplitude
of the 16-day wave is 10% or more of the maximum range of the data series used
for each variable. For the E5-E4 flux and $K_p$ index it is 17%. The 16 day wave is
strongest between about 80-220 days: for the E5-E4 flux, the amplitude exceeds
$5 \times 10^3 \text{cm}^{-2} \text{sr}^{-1} \text{s}^{-1} \text{keV}^{-1}$ throughout this time and this is comparable with the
amplitude of the average diurnal variation. The correlation coefficients for the $K_p$
and $H$ 16-day transforms with the E5-E4 flux transforms are 0.8-0.9. In general
these are better than are seen at the more common periodicities associated with
solar rotation at 27, 13.5 and 9 days. The foF2 correlations are weaker and less
similarity is seen.

Solar wind speed was also included in figures 7.11 and 7.12 to see if the
association with possible solar causes could be ruled out. In figure 7.12, although
the correlation with the other traces is lower, there are still a couple of features in
common with the other plots, namely the peaks near 100 and 160 days. The peak
amplitude here represents a quarter of the total range in the solar wind speed
data series used. A solar cause cannot therefore be ruled out from the results of
the wavelet transforms.

Previous investigations have used the comparison with solar ionizing radiation
to discount a solar connection for the 16-day wave. By using solar wind velocity
it has been shown that a solar connection may exist. Geomagnetic activity has
been observed to have effects on high latitude thermospheric winds and densities
(Forbes et al., 1993 and Forbes et al., 1996) which show changes with $K_p$. If it
is true that the 16-day wave originates in the solar wind then the planetary wave
connection is lost. However, without wind data, which were not available for this
study, it is not possible to tell if the 16-day periodicity is connected with that
which has previously been observed in winds, although several authors cite the
16-day wave in $H$ or foF2 as being due to a planetary wave without supporting
wind data. The work presented here presents some interesting questions regarding
the connections between the magnetosphere and atmosphere. A detailed study
encompassing the full set of wind, ground magnetic, geosynchronous electron and solar wind data is required to provide the answers.

7.4 Mission duration

This section presents an original model of the geosynchronous environment flux in terms of mission duration, which serves to quantify the observed variability in flux over the length of a satellite mission. The model was constructed as follows. The 10%, median and 90% observations were found for the 30 minute averaged flux data. This is illustrated in figure 7.13 which shows the distribution of the low resolution total flux observations. The data were then averaged over longer times. So for hourly averages, the first average is of the first and second data points, the second average is of the second and third data points, and so on. This forms a set of averages over a specific time duration, from which the 10% and 90% limits are found. As the length of the time bin is repeatedly doubled, the flux distribution histogram gets narrower and therefore the 10% and 90% limits move closer together.

![Figure 7.13: Histogram of low resolution total flux observations, showing the position of the median, 10% and 90% events.](image)

235
Figure 7.14 shows the results for this procedure, for time bin averages from 1 hour up to about 3.7 years. The implications of this model are that whereas the total dose accumulated over one year may be predicted with a small error margin, the dose accumulated over much shorter time periods is much harder to predict. In addition, the model implies that the chance of a flux of say, \(2 \times 10^4\text{cm}^{-2}\text{sr}^{-1}\text{s}^{-1}\text{keV}^{-1}\) occurring for several hours is only a few percent, and is unlikely to ever persist for periods of weeks.

Figure 7.15 shows the one year averages for 1989-1994 with the 90%, median and 10% limits from the model. The most active year on the whole, 1994, is on the 90% limit and the other years are within. Half fall above the median, half below.

Particle radiation can be hazardous to spacecraft through the effects of both penetrating radiation and spacecraft charging (Garrett, 1985), all of which depends strongly on the choice of the spacecraft orbit (Spjeldvik and Rothwell, 1985). The geosynchronous altitude is widely used for satellite applications: it is particularly useful for surveillance, communications and meteorological satellites. However, geosynchronous orbit lies within the outer radiation zone which is populated with energetic electrons and is highly variable. The particles may themselves cause damage or may produce damaging secondaries in the spacecraft material. In addition, when electrons are stopped in a material bremsstrahlung (braking) radiation is emitted; this is penetrating X-radiation which is difficult to shield against. Damage can result from the impact of single energetic particles (e.g. single event upsets) or from an accumulated dose of radiation (e.g. the degradation of space system components). Manned missions are also at danger from radiation damage to biological systems.

As an example, Baker et al. (1997) noted that the Telesat Canada Anik E1 communications satellite, in geostationary orbit, suffered severe operational failures coinciding with a large increase in the relativistic electron flux following a storm produced by a high speed solar wind stream. This storm was the latest in a sequence occurring twice per solar rotation and produced a long-lasting relativistic electron flux enhancement. Although the mechanism for the failure
Figure 7.14: Average total flux versus accumulation time

Figure 7.15: One year flux averages showing the 90%, median and 10% limits from the model.
in the power system was not established - it is not known whether the batteries, solar panels or the connecting circuits failed - they cite the relativistic electron enhancement as a possible cause for the failure.

Absorbed dose provides a measure of damage done to a material by penetrating radiation. It is equal to the energy deposited as ionization and excitation energy in a unit mass of a material, and depends on the incident fluence and the properties of the absorbing material. Although the absorbed dose gives a measure of the energy a material absorbs when exposed to radiation, the actual damage done to the material will also depend on the dose \( \text{rate} \) encountered, and on the type of particle (Dyer, 1991).

The object of this analysis was to work towards establishing limits on the radiation dose that would be encountered by a spacecraft in geostationary orbit during a given mission lifetime. This is of great practical importance to spacecraft design; for example, in establishing whether, for a short mission, the use of more costly radiation hard components is warranted, or how much shielding against radiation is required: too much shielding will increase the mass and therefore the launch cost of a spacecraft (e.g. Vampola, 1989).

**Summary**

- Solar cycle variations are seen in both electron populations. A double maxima in geomagnetic activity, one at solar maximum and one in the declining phase of the solar cycle, were observed in the electron flux.

- A 16-day wave is observed in electron flux which may identify with the 16-day wave previously observed in magnetic \( \mathbf{H} \) and identified as a planetary wave.

- The long SEM-2 flux time series has been used to model the average flux observed over increasing mission durations.
7.5 Conclusion

Overall the SEM-2 data gives a comprehensive view of the solar cycle influence on the energetic electron environment at geosynchronous orbit. In this case, strongest activity was seen in the declining phase. Obviously further study is needed to establish the variations between odd and even numbered solar cycles and cycles of different strength.

Solar cycle dependence was seen in both electron populations. Peak E5-E4 flux occurred at solar maximum and peak E1 flux in the declining phase in 1994. Thus solar maximum affects the low energy electrons more and solar minimum affects the high energy electrons more, although for the SEM-2’s narrow energy range, some affect of both peaks in activity was seen at all energies. The average diurnal variations increased in amplitude during these active periods and the E5-E4 variation peaked earlier, consistent with injections occurring earlier when $K_p$ is high.

An intriguing feature of this solar cycle is that close to solar maximum, 1989-1990, far fewer dropouts occurred than in other years. Of course further study is required to establish whether this is a common solar cycle feature. A possible explanation lies in the solar cycle behaviour of the magnetopause position due to changing solar wind pressure. The average magnetopause position is closer to Earth during the declining phase of the solar cycle when solar wind pressure is highest (Fairfield, 1991). Minimum solar wind pressure occurs at solar maximum which allows the magnetopause to expand sunward. Therefore the fact that the magnetosphere may be inflated at solar maximum may explain why a geosynchronous satellite is less prone to nightside flux dropouts at solar maximum. The increased solar wind pressure in the declining phase may also explain the shift in the peak E1 flux towards noon from earlier times at solar maximum.

A 16-day periodicity present in the geosynchronous electron flux shows similar time localisation to a 16-day periodicity in H. Previously authors have interpreted the 16-day periodicity in H as being due to an atmospheric planetary
wave which penetrates upward to the ionosphere, causing currents which then affect the ground magnetic field. There are several possible explanations for the 16-day wave in the SEM-2 data. One is that, like the 27-day periodicity, the wave actually has solar causes; this could not be ruled out by wavelet transforms of solar wind speed which show some features in common with the SEM-2 flux transforms.

A suggestion for further investigation is that this analysis should be expanded to cover the whole chain of events, i.e. to show whether the 16-day wave is present at the same time in neutral winds, in the ionosphere, in ground magnetograms and also in geosynchronous data. Neutral wind data were not made available at the time of this study. However, the amplitude of the 16-day wave is shown to be strong or weak with the same time localization in H, in foF2, and in the SEM-2 electron data. Further work should establish whether the 16-day periodicities in ground based data and in geosynchronous data point to a common cause. If so, then the cause may or may not be solar related. If the solar connection is not enforced, then the possibility that a planetary wave can influence the magnetosphere must be considered.

The fitting of energy spectra in the same manner as Cayton et al. (1989) is difficult for the SEM-2 data because there are only 5 data points which lie close to the point where the two Maxwellians fitted by Cayton et al. cross. Thus it is probable that the SEM-2 data cannot completely define either of those two populations. Of course this does not invalidate the fitting method for the SEM-2 data but the populations defined differ slightly to those of Cayton et al. The Maxwellian fit to the low energy points was more shallow ($T_{\text{high}}, n_{\text{low}}$) than Cayton et al. describe and the fit to the high energy points more steep ($T_{\text{low}}, n_{\text{high}}$). A better procedure for performing the fits for the two populations on such a limited spectrum may be to fit a curve through the five energy spectrum points. A tangent to the curve at the lowest and highest energy points may give a better indication of the gradient for high and low energy asymptotes.

A model of average flux for increasing mission duration was produced using
the long SEM-2 data set. The idea of the model is to quantify the range of flux expected over a mission lifetime: for short missions, there is a greater ‘margin of error’ in using the flux average to extrapolate for the whole mission. Providing 10% and 90% limits improves the predictability. Although using a large data set to produce this model ought to improve the statistics, the position of the five-year time series used in the solar cycle will affect the results. Either using data for a whole solar cycle or else treating solar maximum and minimum separately would produce a more generally useful model. As it is, the data used to construct it are from part of a solar cycle which includes both of the geomagnetic activity maxima, so it should cover the worst case scenario for this cycle. Subsequent cycles may of course prove to be more active.

The main problem with the model is that The SEM-2’s energy range is small: the model applies only to electron flux in the 40-300keV energy range. The model could therefore be improved by extending the energy range. This could be approximated by using the fit of two Maxwellians to the data. From these fits the flux at all energies can be calculated. This of course means estimating the total energy spectrum from the five SEM-2 points, which is likely to produce large errors. The alternative is to use a data set with a wider energy range from which the energy spectrum could be better determined. From this more complete model, the expected radiation dose for a mission could be calculated with more confidence.

Since this model is based on the variability in the SEM-2 data, such a model cannot be created from a ‘static’ representation of the electron environment as is provided by the AE-8 models.
Chapter 8

Concluding Remarks

It has been shown that for a geosynchronous environment model local time dependence is critical. Furthermore the variability in the data is an important factor which is not addressed by standard 'static' models. The radiation belt intensity peaks at noon whereas on the night side, injections cause low energy flux increases of orders of magnitude. The dayside field is dipolar or compressed and populated by pancaked radiation belt particles whereas the nightside field is often taillike and the particle distribution more field-aligned in the substorm growth phase. This situation changes suddenly at substorm onset with the injection of a slightly pancaked electron bunch and the dipolarization of the field.

The local time behaviour of both the substorm injected electrons and the trapped population depends on the level of geomagnetic activity. The day-night asymmetry in the radiation belt intensity is more pronounced at high $K_p$. At low $K_p$ there are few injections and few dropouts. Injections occur generally post-midnight but earlier when $K_p$ is high. The injected electrons generally have energies below 200keV, but at high $K_p$ higher energies are injected. The injection signature has a soft energy spectrum and slightly pancaked distribution which persist with the drifting electrons. When activity is high substorms occur more promptly: the injections happen sooner after the cigar phase onset which is characteristic of the growth phase.

Modelling the geosynchronous electron flux in terms of geomagnetic activity is complicated by the fact that the low energy E5-E4 flux responds promptly to
changes in $K_p$ whereas the progressively higher energy ranges exhibit delays of increasing lengths. However the flux of higher energy trapped particles at any point on the geosynchronous orbit also depends directly on the geomagnetic field configuration. Thus a successful model of geosynchronous high energy trapped electrons needs to incorporate both the instantaneous level of geomagnetic activity and the level from a day or two previously; for relativistic electrons the lag has been found to be 2-3 days behind $K_p$ (Baker et al., 1990). The low energy substorm injected component on the other hand is adequately organized by instantaneous $K_p$.

The flux probability distributions for different levels of geomagnetic activity underline the need for a dynamic model of the outer zone. The most widely used models of the outer zone are probably still the NASA AE-8 MIN and AE-8 MAX models for solar minimum and maximum respectively, which are static models and are identical for geosynchronous altitudes. A comparison of the SEM-2 data with the AE-8 models has shown that the high energy SEM-2 flux is lower than the model predicts but that the substorm injected flux is higher (Rodgers et al., 1993), which can also be seen in the energy spectra presented in Chapter 2. A recent ‘quasi-static’ model for outer zone relativistic electrons based on CRRES HEEF (High Energy Electron Fluxmeter) data uses a 15-day running average of the Ap index, $Ap_{15}$ (Brautigam et al., 1992). The model has 8 levels for different $Ap_{15}$ ranges. It was found that the model could broadly reproduce the CRRES observations. Comparison of the $Ap_{15}$ model with the NASA models found that the NASA models were particularly bad beyond $3R_E$ and worse for higher energies, with flux being significantly overestimated.

Brautigam et al. conclude that a single model can never adequately describe the outer radiation zone, and that separate models will always be required for quiet and active periods. The results for the E1 channel presented here suggest an alternative to Brautigam et al.’s running average approach, in that two components of activity may combine to provide a better description of the radiation belt configuration at a particular time. It has been shown that higher energy electrons react to enhancements in geomagnetic activity on a time scale related to their
energy (although Baker et al. (1990) note that relativistic electrons react ‘as a whole’), but it is also clear that the instantaneous state of activity defines the observed environment at a satellite. Therefore a model which uses both lagged (dependent on energy) and instantaneous activity levels may give more selective results than a model based on a broad average of recent activity levels.

The effect of the solar wind has been illustrated both through the injection frequency dependence on solar wind speed, and through the solar cycle behaviour of both electron populations. This emphasises that magnetospheric models cannot work in isolation, in that models of trapped radiation and of the geomagnetic field itself need to include the influence of solar wind conditions. The SEM-2 data suggests that the solar wind pressure may control the asymmetry of the trapped flux about noon and possibly also the frequency of dropouts. The level of activity in the T89 model is defined solely by $K_p$ which is inadequate both in the range of conditions that are modelled and in the 3 hour time scale for changes to take place. The latest Tsyganenko field model, T96-01, does use solar wind velocity and density as well as IMF parameters for inputs.

For solar cycle 22, strongest activity seemed to occur in the declining phase. The monthly-averaged low energy electron flux peaked at solar maximum, although more, and more frequent, injections occurred at the declining phase activity peak. The high energy electrons were affected most by the enhanced activity in the declining phase. It is noted that far fewer growth phase dropouts were observed close to solar maximum than in other years, although whether this is typical of the solar cycle cannot be established with the SEM-2 data. Since minimum solar wind pressure occurs at solar maximum (Petrinec et al., 1991), this may mean that the magnetosphere is inflated so that dropouts are seen infrequently by a geosynchronous spacecraft.

For mission planning long term radiation belt models are required, and these need to incorporate solar cycle conditions. Different energy populations are enhanced at different times in the solar cycle. Radiation dose for a mission duration may be modelled from existing data sets using the approach in Chapter 7; for
longer time intervals the total dose may be predicted with more confidence. As it stands the model is limited by its narrow energy range but this could be extended with other data sets or extrapolated to other energies.

Comparison with the widely-used Tsyganenko 89 magnetic field model has shown that the calculated symmetry axis of the particle distribution gives a good indication of the magnetic field direction, although as would be expected, the determination of the axis direction is most accurate when there is considerable anisotropy. This method for finding the symmetry axis is thus generally validated which is useful for instruments such as the SEM-2 which are not accompanied by a magnetometer. Having a reliable indicator of the field direction allows pitch angle information to be extracted from a three-dimensional particle distribution. The analysis of the symmetry axis direction identified persistent differences between the calculated field direction and the Tsyganenko model field direction, suggesting shortcomings of the model. On the dayside, the data shows an asymmetry about noon which is not reproduced by the model. The stretching in the tail field was found to be underestimated at high $K_p$, and possibly also overestimated at low $K_p$, and the dynamic behaviour of the tail is inadequately modelled. Seasonal differences, of opposite effect for summer and winter, were also observed, but on the whole the solstices are better modelled by T89 than the equinoxes.

Wavelet analysis is a powerful and promising tool for time series analysis. Its strength lies in keeping the time localization of the frequency decomposition so that transient frequencies can be detected. Wavelet analysis was applied here to the problem of substorm frequency and found the most common inter-substorm time to be 4 hours, although the actual frequency varies and depends to some extent on solar wind conditions. It is suggested that wavelets of other shapes may provide better results, for example, a shape close to the peak signature revealed by superposed epoch analysis of injections.

A second application found for wavelets was the study of a 16-day wave observed in the geosynchronous electron flux. The time localization of the 16-day
periodicity episodes suggested a connection with similar frequency waves in H. A 16-day wave in H has been interpreted previously as being due to the penetration of a planetary wave into the ionosphere (e.g. Forbes and Leveroni, 1982). An interesting point is that solar influence was ruled out because 16-day periodicities were not present in solar ionizing radiation. Wavelet analysis of solar wind speed suggests that the solar influence cannot readily be discounted, although obviously further study is required to establish whether the solar wind is the driving force for a 16-day wave or not, and whether or not this is the same wave that is identified in neutral winds.

The mechanism by which a neutral wind periodicity can be transferred to H is through the flow of ionospheric currents causing perturbations in the magnetic field observed on the ground. It is difficult to formulate a mechanism whereby the same periodicity can spread to the radiation belts, although changing the conditions at the foot of a field line in the auroral zone would change the conditions along the whole field line. If the 16-day wave comes from the solar wind, then the radiation belt flux and ground magnetic field would be affected in the same way that the 27-day solar rotation period is observed in the outer zone flux and in $K_p$. Since it is known that neutral winds are affected by $K_p$ the changes in H may then reflect magnetospheric periodicities.

During the course of this work, many errors were found in the SEM-2 data and the whole data set had to be reprocessed twice. Errors existed in every level of the data processing: in the onboard software, in the ground decompression software and in the data archiving programs. All the errors (described in Appendix A) were corrected for the work presented here.

Two further errors were however discovered during completion of this thesis, both in the ground data handling software. The first affects relatively few data points: only the azimuthal counts near midnight at the equinoxes, when the satellite is in eclipse. The problem arises from the time in the spacecraft housekeeping data being interrupted, so that when the reference time is needed for the SEM-2 spin measurement, the error occurs. The erroneous azimuthal
points subsequently affect the axis calculation and anisotropy index, and the polar-azimuthal flux.

The second outstanding error concerns the efficiency of the SEM-2 in its five energy ranges. Firstly, although the calibration tests showed that the efficiency could not really be $\sim$100% (Coates et al., 1990) this was the value used in the processing software. From equation 3.1 it is clear that assuming too high an efficiency will give flux values that are too low. This is probably responsible for the fact that the SEM-2 flux seems slightly low when compared with the LANL instruments. A worse assumption made about the efficiency was that all five energy ranges were equally efficient. The calibration report shows that this is not true: the lowest energy range has a lower efficiency. With hindsight, the energy spectra in Chapter 2 do suggest that the E5-E4 points are at a lower flux than expected, which is a consequence of this error.

The affect of the eclipse problem is to record zero values of azimuthal and polar-azimuthal flux and anisotropy index, although few points are affected. The efficiency error has a larger effect in that the flux is underestimated and most importantly, the E5-E4 flux is too low relative to the other energy ranges. This will also affect the spectral index in the archived data set, which is calculated from the slope of the energy spectrum which will therefore be too shallow. However, such an error should not affect most of the work in this thesis qualitatively. Most of the analyses are carried out using individual energy ranges so only the absolute values of flux will be affected (they are out by a factor), although the total flux contains an error since it was calculated assuming all energy ranges to have equal efficiency.

In the next version of the SEM-2 data, it is recommended that the eclipse points are flagged with -1 for the affected variables, as for missing data, if the timing problem cannot be resolved. It is also desirable that more realistic efficiencies are implemented, or at least that the lower relative efficiency of the lowest energy range is taken into account.
Appendix A

SEM-2 Errors

Errors and omissions which were discovered in the original archived SEM-2 data, and corrected prior to the work in this thesis, are discussed here along with the effect that each error had on the data and its solution. It is important to understand how the data were affected by the software errors since the contaminated SEM-2 data have been available for analysis for many years.

A.1 Archiving Errors

A.1.1 Omissions

The archiving program failed to write certain variables, such as the polar-azimuthal counts, to the archive data files so that zero values were stored.

A.1.2 Data Clipping

Flux in the five differential energy ranges was found to be 'clipped', i.e. upper and lower limits were imposed. These limits corresponded to the plot boundaries used in the early SEM-2 software, which can be seen on the daily summary plot shown in figure 3.4. The archiving program was based on this earlier plotting program and the plot limits had been imposed on the archived data.
A.1.3 Polar Flux Asymmetry

Rodgers (1991) compared the flux observed in each of the five polar sectors. He found that whereas the 60° and 120° polar flux were of the same order of magnitude, the 150° polar flux was 1.38 times the 30° polar flux. From this it was concluded that either the 30° sensor had been wrongly calibrated, or that something on the spacecraft was obscuring part of its field of view. His recommendation was that the 30° flux should be multiplied by 1.38 to compensate.

![HR polar flux, day 1 of 1990](image)

Figure A.1: The polar flux error

Closer examination of the high resolution polar flux reveals an error in the 150° polar flux data: figure A.1 shows a sample of this data. The dotted line shows the flux in the 30° sensor, and the solid line is the 150° flux. Every third point in the 150° flux is raised well above that in the 30° flux, whereas the rest of the points seems to agree quite closely. This is the cause of the factor 1.38 difference seen by Rodgers. The error causing the polar flux asymmetry was traced to the archiving software where a normalization had been omitted. It was also found that the polar flux was wrongly calculated from the polar counts, such that it was a factor 25-30 too high in all five polar bins.
A.2 Onboard Data Compression Error

![Histogram of polar azimuthal fluxes](image)

Figure A.2: Histogram of polar azimuthal fluxes: all 30 bins, one year (1990) of data. Each of the 30 bins individually showed similar behaviour.

The SEM-2 onboard data compression routine contained a serious error, which meant the error was already present when the data were received at the ground. The error was seen in each of the 30 polar-azimuthal flux bins, but disguised in other parts of the data set by the averaging of the data over polar or azimuthal or both angles. This error caused certain values of counts and flux to occur more commonly than others in the polar-azimuthal arrays. This is illustrated by figure A.2, a histogram of the low resolution flux observations in all the polar-azimuthal bins. Note that the high resolution polar-azimuthal flux shows the same peculiarities, and each one of the 30 polar-azimuthal bins considered individually contains the same error.

Prior to compression the number of counts exists as a 2-byte input word. This number is then compressed down to an 8-bit number to be transmitted, and can then be decompressed to restore the input number of counts. Because an 8-bit processor is used, each byte (8 bits) in the 2-byte input word must be handled separately. The first byte is referred to as the most significant byte and
the second as the least significant byte. The compression algorithm reduces the 16 input bits to 8, a 4-bit mantissa and a 4-bit exponent. The mantissa gives the next four significant bits of the input number after the first significant (i.e. non-zero) bit, and the exponent tells how many zeros to add to the end of this number.

The compression is carried out as follows. The first significant bit of the 16 is shifted forward to set an overflow flag: the number of shifts required to do this provides the exponent. The next four bits are used to form the mantissa. In some cases it is necessary to use bits from both the most significant and least significant bytes of the input word to create the mantissa, and in these cases it is necessary to perform bit-shifts on both bytes. It is in these cases that the error occurs, whenever the exponent is between 4 and 7. The input ranges of numbers and the output decompressed numbers which are affected can be seen in the tables in figures A.3 and A.4. The error is due to the bit shifting operation being carried out incorrectly: the bit-shift in the least significant byte was one bit short, leaving an 'empty' bit. This meant that one bit, the first one contributed by the least significant byte, is wrongly fixed at zero, and the following bits, although correct, are misplaced.

This meant that there was a set of numbers for which wrong values were always assigned in the SEM-2 data. It also meant that certain compressed numbers could never occur (those which should have had a 1 where the forced zero was), which inevitably meant that certain decompressed numbers could never occur. These missing values, combined with others which were over-subscribed with wrongly shifted values, caused the peculiarities seen in figure A.2. Barry Hancock (MSSL), who coded the original compression software, has written a report describing the error and correction in detail which is included in this appendix.

Note that the compression process itself imposes an uncertainty on the data since a compressed number, and therefore a decompressed number, will correspond to a range of input numbers. Had the compression error not been made, the ranges of input numbers and corresponding output numbers would be given by the table in figure A.3, and the maximum percentage errors caused by the
Figure A.3: Input ranges for the SEM-2 compression algorithm: 'low' and 'high' are the limits for each code.

**LS nibble**  
(hez)  

<table>
<thead>
<tr>
<th>low</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>low</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>high</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>low</td>
<td>32</td>
<td>34</td>
<td>36</td>
<td>38</td>
<td>40</td>
<td>42</td>
<td>44</td>
<td>46</td>
<td>48</td>
<td>50</td>
<td>52</td>
<td>54</td>
<td>56</td>
<td>58</td>
<td>60</td>
<td>62</td>
</tr>
<tr>
<td>high</td>
<td>32</td>
<td>34</td>
<td>36</td>
<td>38</td>
<td>40</td>
<td>42</td>
<td>44</td>
<td>46</td>
<td>48</td>
<td>50</td>
<td>52</td>
<td>54</td>
<td>56</td>
<td>58</td>
<td>60</td>
<td>62</td>
</tr>
<tr>
<td>low</td>
<td>64</td>
<td>66</td>
<td>70</td>
<td>72</td>
<td>76</td>
<td>80</td>
<td>84</td>
<td>88</td>
<td>92</td>
<td>96</td>
<td>100</td>
<td>104</td>
<td>108</td>
<td>112</td>
<td>116</td>
<td>120</td>
</tr>
<tr>
<td>high</td>
<td>64</td>
<td>66</td>
<td>70</td>
<td>72</td>
<td>76</td>
<td>80</td>
<td>84</td>
<td>88</td>
<td>92</td>
<td>96</td>
<td>100</td>
<td>104</td>
<td>108</td>
<td>112</td>
<td>116</td>
<td>120</td>
</tr>
<tr>
<td>low</td>
<td>128</td>
<td>130</td>
<td>134</td>
<td>136</td>
<td>138</td>
<td>140</td>
<td>144</td>
<td>146</td>
<td>148</td>
<td>150</td>
<td>152</td>
<td>154</td>
<td>156</td>
<td>158</td>
<td>160</td>
<td>162</td>
</tr>
<tr>
<td>high</td>
<td>128</td>
<td>130</td>
<td>134</td>
<td>136</td>
<td>138</td>
<td>140</td>
<td>144</td>
<td>146</td>
<td>148</td>
<td>150</td>
<td>152</td>
<td>154</td>
<td>156</td>
<td>158</td>
<td>160</td>
<td>162</td>
</tr>
<tr>
<td>low</td>
<td>256</td>
<td>258</td>
<td>262</td>
<td>264</td>
<td>266</td>
<td>268</td>
<td>270</td>
<td>272</td>
<td>274</td>
<td>276</td>
<td>278</td>
<td>280</td>
<td>282</td>
<td>284</td>
<td>286</td>
<td>288</td>
</tr>
<tr>
<td>high</td>
<td>256</td>
<td>258</td>
<td>262</td>
<td>264</td>
<td>266</td>
<td>268</td>
<td>270</td>
<td>272</td>
<td>274</td>
<td>276</td>
<td>278</td>
<td>280</td>
<td>282</td>
<td>284</td>
<td>286</td>
<td>288</td>
</tr>
<tr>
<td>low</td>
<td>512</td>
<td>514</td>
<td>516</td>
<td>518</td>
<td>520</td>
<td>522</td>
<td>524</td>
<td>526</td>
<td>528</td>
<td>530</td>
<td>532</td>
<td>534</td>
<td>536</td>
<td>538</td>
<td>540</td>
<td>542</td>
</tr>
<tr>
<td>high</td>
<td>512</td>
<td>514</td>
<td>516</td>
<td>518</td>
<td>520</td>
<td>522</td>
<td>524</td>
<td>526</td>
<td>528</td>
<td>530</td>
<td>532</td>
<td>534</td>
<td>536</td>
<td>538</td>
<td>540</td>
<td>542</td>
</tr>
<tr>
<td>low</td>
<td>1024</td>
<td>1026</td>
<td>1028</td>
<td>1030</td>
<td>1032</td>
<td>1034</td>
<td>1036</td>
<td>1038</td>
<td>1040</td>
<td>1042</td>
<td>1044</td>
<td>1046</td>
<td>1048</td>
<td>1050</td>
<td>1052</td>
<td>1054</td>
</tr>
<tr>
<td>high</td>
<td>1024</td>
<td>1026</td>
<td>1028</td>
<td>1030</td>
<td>1032</td>
<td>1034</td>
<td>1036</td>
<td>1038</td>
<td>1040</td>
<td>1042</td>
<td>1044</td>
<td>1046</td>
<td>1048</td>
<td>1050</td>
<td>1052</td>
<td>1054</td>
</tr>
<tr>
<td>low</td>
<td>2048</td>
<td>2050</td>
<td>2052</td>
<td>2054</td>
<td>2056</td>
<td>2058</td>
<td>2060</td>
<td>2062</td>
<td>2064</td>
<td>2066</td>
<td>2068</td>
<td>2070</td>
<td>2072</td>
<td>2074</td>
<td>2076</td>
<td>2078</td>
</tr>
<tr>
<td>high</td>
<td>2048</td>
<td>2050</td>
<td>2052</td>
<td>2054</td>
<td>2056</td>
<td>2058</td>
<td>2060</td>
<td>2062</td>
<td>2064</td>
<td>2066</td>
<td>2068</td>
<td>2070</td>
<td>2072</td>
<td>2074</td>
<td>2076</td>
<td>2078</td>
</tr>
<tr>
<td>low</td>
<td>4096</td>
<td>4098</td>
<td>4100</td>
<td>4102</td>
<td>4104</td>
<td>4106</td>
<td>4108</td>
<td>4110</td>
<td>4112</td>
<td>4114</td>
<td>4116</td>
<td>4118</td>
<td>4120</td>
<td>4122</td>
<td>4124</td>
<td>4126</td>
</tr>
<tr>
<td>high</td>
<td>4096</td>
<td>4098</td>
<td>4100</td>
<td>4102</td>
<td>4104</td>
<td>4106</td>
<td>4108</td>
<td>4110</td>
<td>4112</td>
<td>4114</td>
<td>4116</td>
<td>4118</td>
<td>4120</td>
<td>4122</td>
<td>4124</td>
<td>4126</td>
</tr>
<tr>
<td>low</td>
<td>8192</td>
<td>8200</td>
<td>8208</td>
<td>8216</td>
<td>8224</td>
<td>8232</td>
<td>8240</td>
<td>8248</td>
<td>8256</td>
<td>8264</td>
<td>8272</td>
<td>8280</td>
<td>8288</td>
<td>8296</td>
<td>8304</td>
<td>8312</td>
</tr>
<tr>
<td>high</td>
<td>8192</td>
<td>8200</td>
<td>8208</td>
<td>8216</td>
<td>8224</td>
<td>8232</td>
<td>8240</td>
<td>8248</td>
<td>8256</td>
<td>8264</td>
<td>8272</td>
<td>8280</td>
<td>8288</td>
<td>8296</td>
<td>8304</td>
<td>8312</td>
</tr>
<tr>
<td>low</td>
<td>16384</td>
<td>16392</td>
<td>16400</td>
<td>16408</td>
<td>16416</td>
<td>16424</td>
<td>16432</td>
<td>16440</td>
<td>16448</td>
<td>16456</td>
<td>16464</td>
<td>16472</td>
<td>16480</td>
<td>16488</td>
<td>16496</td>
<td>16504</td>
</tr>
<tr>
<td>high</td>
<td>16384</td>
<td>16392</td>
<td>16400</td>
<td>16408</td>
<td>16416</td>
<td>16424</td>
<td>16432</td>
<td>16440</td>
<td>16448</td>
<td>16456</td>
<td>16464</td>
<td>16472</td>
<td>16480</td>
<td>16488</td>
<td>16496</td>
<td>16504</td>
</tr>
<tr>
<td>low</td>
<td>32768</td>
<td>32776</td>
<td>32784</td>
<td>32792</td>
<td>32800</td>
<td>32808</td>
<td>32816</td>
<td>32824</td>
<td>32832</td>
<td>32840</td>
<td>32848</td>
<td>32856</td>
<td>32864</td>
<td>32872</td>
<td>32880</td>
<td>32888</td>
</tr>
<tr>
<td>high</td>
<td>32768</td>
<td>32776</td>
<td>32784</td>
<td>32792</td>
<td>32800</td>
<td>32808</td>
<td>32816</td>
<td>32824</td>
<td>32832</td>
<td>32840</td>
<td>32848</td>
<td>32856</td>
<td>32864</td>
<td>32872</td>
<td>32880</td>
<td>32888</td>
</tr>
</tbody>
</table>

**MS nibble**  
(hez)  

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
</table>

**mantissa**
### Output Numbers (o/p) from Decompression Algorithm and the Maximum Percentage Error Associated with Each

<table>
<thead>
<tr>
<th>Nibble (Hex)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ERROR (%)</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>33</td>
<td>33</td>
<td>37</td>
<td>39</td>
<td>41</td>
<td>43</td>
<td>45</td>
<td>47</td>
<td>49</td>
<td>51</td>
<td>53</td>
<td>55</td>
<td>57</td>
<td>59</td>
<td>61</td>
<td>63</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>6.6</td>
<td>7.0</td>
<td>7.4</td>
<td>7.8</td>
<td>8.2</td>
<td>8.6</td>
<td>9.0</td>
<td>9.4</td>
<td>9.8</td>
<td>10.2</td>
<td>10.6</td>
<td>11.0</td>
<td>11.4</td>
<td>11.8</td>
<td>12.2</td>
<td>12.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>132</td>
<td>140</td>
<td>148</td>
<td>156</td>
<td>164</td>
<td>172</td>
<td>180</td>
<td>188</td>
<td>196</td>
<td>204</td>
<td>212</td>
<td>220</td>
<td>228</td>
<td>236</td>
<td>244</td>
<td>252</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>264</td>
<td>260</td>
<td>296</td>
<td>312</td>
<td>328</td>
<td>344</td>
<td>360</td>
<td>376</td>
<td>392</td>
<td>408</td>
<td>424</td>
<td>440</td>
<td>456</td>
<td>472</td>
<td>488</td>
<td>504</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>528</td>
<td>560</td>
<td>592</td>
<td>624</td>
<td>656</td>
<td>688</td>
<td>720</td>
<td>752</td>
<td>784</td>
<td>816</td>
<td>848</td>
<td>880</td>
<td>912</td>
<td>944</td>
<td>976</td>
<td>1008</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>1056</td>
<td>1120</td>
<td>1184</td>
<td>1248</td>
<td>1312</td>
<td>1376</td>
<td>1440</td>
<td>1504</td>
<td>1568</td>
<td>1632</td>
<td>1696</td>
<td>1760</td>
<td>1824</td>
<td>1888</td>
<td>1952</td>
<td>2016</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>2112</td>
<td>2240</td>
<td>2368</td>
<td>2496</td>
<td>2624</td>
<td>2752</td>
<td>2880</td>
<td>3008</td>
<td>3136</td>
<td>3264</td>
<td>3392</td>
<td>3520</td>
<td>3648</td>
<td>3776</td>
<td>3904</td>
<td>4032</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>4224</td>
<td>4480</td>
<td>4736</td>
<td>4992</td>
<td>5248</td>
<td>5504</td>
<td>5760</td>
<td>6016</td>
<td>6272</td>
<td>6528</td>
<td>6784</td>
<td>7040</td>
<td>7296</td>
<td>7552</td>
<td>7808</td>
<td>8064</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>8448</td>
<td>8960</td>
<td>9472</td>
<td>9984</td>
<td>10496</td>
<td>11008</td>
<td>11520</td>
<td>12032</td>
<td>12544</td>
<td>13056</td>
<td>13568</td>
<td>14080</td>
<td>14592</td>
<td>15104</td>
<td>15616</td>
<td>16128</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>o/p</strong></td>
<td>16896</td>
<td>17920</td>
<td>18944</td>
<td>19968</td>
<td>20992</td>
<td>22016</td>
<td>23040</td>
<td>24064</td>
<td>25088</td>
<td>26112</td>
<td>27136</td>
<td>28160</td>
<td>29184</td>
<td>30208</td>
<td>31232</td>
<td>32256</td>
</tr>
<tr>
<td><strong>error (%)</strong></td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
<td>2.6</td>
<td>2.5</td>
<td>2.4</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
</tr>
</tbody>
</table>

**MS nibble (hex):** 0 1 2 3 4 5 6 7 8 9 A B C D E F

**mantissa:**

**Exponent:**

*Figure A.4: Output numbers (o/p) from decompression algorithm and the maximum percentage error associated with each.*
Figure A.5: New input ranges and output numbers for compression. The maximum percentage errors are also shown.
compression by the table in figure A.4 (both from Coates et al., 1990). However, for the cases where the exponent is between 4 and 7, these tables have to be modified: the loss of precision incurred in correcting the compression error will result in a larger maximum percentage error in the output numbers. Table A.5 shows the new input ranges and output numbers, and the maximum percentage errors caused by the compression, for the corrected data compression routine.

Figures A.6 to A.9 show how numbers were wrongly recorded, and in each case the percentage error in the assigned value is stated. The upper row shows the true value that the decompressed counts should have been assigned to, and the arrows show to which number in the bottom row the count was assigned. In all cases where an error occurred, the assigned value was lower than the true value.

Figure A.10 shows how the counts for each threshold energy (as defined by discriminators E1 to E5) were affected. One month of counts data, January 1990, was used for this plot. The E5 counts (42.9keV threshold) were least affected since counts are higher at lower energies and so most E5 observations occurred above the range of count values affected. The E1 (201.8keV threshold) counts were mostly in the count range affected by the error so were significantly modified by it. The percentage of observations for which the decompressed number fell into the count bins affected by the error is shown in the second column of table A.1. Not all the observations in these count bins incorporated an error: figures A.6 to A.9 show that some counts in this range were correctly binned and had a zero error. To estimate what percentage of observations actually contained an error, it was assumed that where two true values were assigned to the same output value, the two true count values occurred the same number of times. So if one of the two paths leading to an output number was a zero-error path, half of the observations with the output count values were assumed to contain no error. The third column of table A.1 gives the resulting estimate of the percentage of observations containing an error.

The errors on the counts will be imposed on the calculated flux. Bands of flux concentrations are seen in the low resolution flux because each flux point
### Figure A.6: Errors in the output counts due to the compression error for exponent 4

<table>
<thead>
<tr>
<th>True Value</th>
<th>Assigned Value</th>
<th>0</th>
<th>6.06</th>
<th>0</th>
<th>4.88</th>
<th>0</th>
<th>4.08</th>
<th>0</th>
<th>3.51</th>
<th>0</th>
<th>3.28</th>
</tr>
</thead>
<tbody>
<tr>
<td>2112</td>
<td>2240</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2368</td>
<td>2496</td>
<td>0</td>
<td>5.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2624</td>
<td>2752</td>
<td>0</td>
<td>4.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2880</td>
<td>3008</td>
<td>0</td>
<td>3.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3136</td>
<td>3264</td>
<td>0</td>
<td>3.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3392</td>
<td>3520</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3648</td>
<td>3776</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3904</td>
<td>4032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

0 gives the percentage error in the assigned value

1184 shows a value which never occurs in the data due to the compression error
Figure A.7: Errors in the output counts due to the compression error for exponent 5

<table>
<thead>
<tr>
<th>true value</th>
<th>1056</th>
<th>1120</th>
<th>1184</th>
<th>1248</th>
<th>1312</th>
<th>1376</th>
<th>1440</th>
<th>1504</th>
<th>1568</th>
<th>1632</th>
<th>1696</th>
<th>1760</th>
<th>1824</th>
<th>1888</th>
<th>1952</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>assigned value</td>
<td>1056</td>
<td>1120</td>
<td>1184</td>
<td>1248</td>
<td>1312</td>
<td>1376</td>
<td>1440</td>
<td>1504</td>
<td>1568</td>
<td>1632</td>
<td>1696</td>
<td>1760</td>
<td>1824</td>
<td>1888</td>
<td>1952</td>
<td>2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 0 gives the percentage error in the assigned value
- 1184 shows a value which never occurs in the data due to the compression error
Figure A.8: Errors in the output counts due to the compression error for exponent = 6

0 gives the percentage error in the assigned value

1184 shows a value which never occurs in the data due to the compression error
**Exponent = 7**

<table>
<thead>
<tr>
<th>true value</th>
<th>264</th>
<th>280</th>
<th>296</th>
<th>312</th>
<th>328</th>
<th>344</th>
<th>360</th>
<th>376</th>
<th>392</th>
<th>408</th>
<th>424</th>
<th>440</th>
<th>456</th>
<th>472</th>
<th>488</th>
<th>504</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6.06</td>
<td>5.71</td>
<td>11.43</td>
<td>10.81</td>
<td>16.22</td>
<td>15.38</td>
<td>20.51</td>
<td>19.51</td>
<td>24.39</td>
<td>23.26</td>
<td>27.91</td>
<td>26.67</td>
<td>31.11</td>
<td>29.79</td>
<td>34.04</td>
<td></td>
</tr>
</tbody>
</table>

| assigned value | 264 | 280 | 296 | 312 | 328 | 344 | 360 | 376 | 392 | 408 | 424 | 440 | 456 | 472 | 488 | 504 |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|                | 0   | 6.06| 5.71| 11.43| 10.81| 16.22| 15.38| 20.51| 19.51| 24.39| 23.26| 27.91| 26.67| 31.11| 29.79| 34.04|     |

- **0** gives the percentage error in the assigned value.
- **1184** shows a value which never occurs in the data due to the compression error.

Figure A.9: Errors in the output counts due to the compression error for exponent 7.
Figure A.10: The effect of the compression error on the counts for each discriminator setting. The $x$ axis gives the output count and the $y$ axis gives the number of times the count occurred.

<table>
<thead>
<tr>
<th>discriminator</th>
<th>observations in affected bins (% of total observations)</th>
<th>estimate of observations containing a non-zero error (% of total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>86.2</td>
<td>64.0</td>
</tr>
<tr>
<td>E2</td>
<td>73.3</td>
<td>48.9</td>
</tr>
<tr>
<td>E3</td>
<td>47.5</td>
<td>29.3</td>
</tr>
<tr>
<td>E4</td>
<td>21.4</td>
<td>12.7</td>
</tr>
<tr>
<td>E5</td>
<td>11.6</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Table A.1: Percentage of observations which occurred in the count bin affected by the compression error
is calculated from more than one count value, so there is a range of preferential flux values that can result from different combinations of the preferential count values.

In each case where the compression algorithm error caused a count value to be wrongly recorded, the assigned value was too low. SEM-2's low energy ranges were least affected but in the highest energy range more than half the data was corrupted. The error on one point could be as much as 34%, although very few points were affected this badly. The error was such that the same numbers were affected and the same paths followed all the time, and in no case was the numerical order of the individual count values reversed. Therefore none of the data were 'scrambled', although sections of the data were reduced in magnitude. However, the error on the counts was propagated through to the flux and also to the anisotropy index which is calculated using the polar-azimuthal flux.

The compression error could be corrected for but with a loss of precision: full details of the correction are given in the official report on the error which follows.

A.2.1 SEM-2 Data Compression Report

The official report on the SEM-2 data compression error written by Barry Hancock of MSSL is included here.
METEOSAT SEM-2 DATA COMPRESSION PROBLEM

BKH 26/7/95
file: prob.doc

As a result of questions being raised about the SEM-2 data stream, I have done some analysis of the on-board software. I have found a problem with the coding of the data compression algorithm. The problem I believe is this:

The 1802 is an 8 bit microprocessor and has no 16 bit shift instructions, therefore, it is necessary to perform this shift separately on the most significant byte and the least significant byte of the input word. When the ms byte is shifted more than 4 places to the left some bits from the ls byte have to be used in the generation of the resultant mantissa. This involves shifting the ls byte right 8 - number of left shifts. Unfortunately the exponent is determined by the number of shifts minus one (or the number of leading zero's). The ls byte is shifted one place more than is required. This results in a corruption of the mantissa for four exponent values. These cases are detailed below.
I believe the data is largely recoverable with a small loss of precision.

From results of data from a number of Ampte orbits, it appears that its affected too. More research is required to confirm this.

Case 1 (exponent 4)
Shift ms byte 5 places left until overflow flag is set, then shift ls byte 4 places right (it should be 3 places).

<table>
<thead>
<tr>
<th>ms byte</th>
<th>ls byte</th>
</tr>
</thead>
<tbody>
<tr>
<td>b7  b6  b5  b4  b3  b2  b1  b0</td>
<td>0 0 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

add component of shifted ms byte to component of shifted ls byte.

<table>
<thead>
<tr>
<th>ms byte</th>
<th>ls byte</th>
</tr>
</thead>
<tbody>
<tr>
<td>m  m  m  0  l  l  l  l</td>
<td>0 0 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

add exponent (number of msb shifts -1).

<table>
<thead>
<tr>
<th>ms byte</th>
<th>ls byte</th>
</tr>
</thead>
<tbody>
<tr>
<td>m  m  m  0  e  e  e  e</td>
<td>0 0 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

The resultant value has a mantissa in which b4 is set to zero.

Solution
None.
Case 2 (exponent 5)
shift ms byte 6 places left until overflow flag is set, then shift ls byte 3 places right (it should be 2 places).

<table>
<thead>
<tr>
<th>o/f</th>
<th>b7</th>
<th>b6</th>
<th>b5</th>
<th>b4</th>
<th>b3</th>
<th>b2</th>
<th>b1</th>
<th>b0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>m</td>
<td>m</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

add component of shifted ms byte to component of shifted ls byte

<table>
<thead>
<tr>
<th>b7</th>
<th>b6</th>
<th>b5</th>
<th>b4</th>
<th>b3</th>
<th>b2</th>
<th>b1</th>
<th>b0</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>m</td>
<td>0</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
</tbody>
</table>

add exponent (number of msb shifts -1).

<table>
<thead>
<tr>
<th>b7</th>
<th>b6</th>
<th>b5</th>
<th>b4</th>
<th>b3</th>
<th>b2</th>
<th>b1</th>
<th>b0</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>m</td>
<td>0</td>
<td>l</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

The resultant value has a mantissa in which b5 is set to zero.

Solution
The value of b4 should be in b5, the intended value of b4 is lost.

Shift b4 one place to the left and zero b4.

<table>
<thead>
<tr>
<th>b7</th>
<th>b6</th>
<th>b5</th>
<th>b4</th>
<th>b3</th>
<th>b2</th>
<th>b1</th>
<th>b0</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>m</td>
<td>l</td>
<td>0</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

Case 3 (exponent 6)
shift ms byte 7 places until overflow flag is set, then shift ls byte 2 places (it should be 1 place).

<table>
<thead>
<tr>
<th>o/f</th>
<th>b7</th>
<th>b6</th>
<th>b5</th>
<th>b4</th>
<th>b3</th>
<th>b2</th>
<th>b1</th>
<th>b0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>m</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

add component of shifted ms byte to component of shifted ls byte

<table>
<thead>
<tr>
<th>b7</th>
<th>b6</th>
<th>b5</th>
<th>b4</th>
<th>b3</th>
<th>b2</th>
<th>b1</th>
<th>b0</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>0</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
</tbody>
</table>

add exponent (number of msb shifts -1).

<table>
<thead>
<tr>
<th>b7</th>
<th>b6</th>
<th>b5</th>
<th>b4</th>
<th>b3</th>
<th>b2</th>
<th>b1</th>
<th>b0</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>0</td>
<td>l</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>
The resultant value has a mantissa in which b6 is set to zero.

Solution.
The value in b4 and b5 should be in b5 and b6, the intended value of b4 is lost.

Shift b4 and b5 one place to the left and zero b4.

\[
\begin{array}{cccccccc}
\text{b7} & \text{b6} & \text{b5} & \text{b4} & \text{b3} & \text{b2} & \text{b1} & \text{b0} \\
\hline
\text{m} & \text{l} & \text{l} & \text{0} & \text{e} & \text{e} & \text{e} & \text{e}
\end{array}
\]

Case 4 (exponent 7)
shift ms byte 8 places until overflow flag is set, then shift ls byte 1 places (it should be zero shift).

\[
\begin{array}{cccccccc}
\text{off} & \text{b7} & \text{b6} & \text{b5} & \text{b4} & \text{b3} & \text{b2} & \text{b1} & \text{b0} & \text{b7} & \text{b6} & \text{b5} & \text{b4} & \text{b3} & \text{b2} & \text{b1} & \text{b0} \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
\]

add component of shifted ms byte to component of shifted ls byte

\[
\begin{array}{cccccccc}
\text{b7} & \text{b6} & \text{b5} & \text{b4} & \text{b3} & \text{b2} & \text{b1} & \text{b0} \\
\hline
0 & l & l & l & l & l & l & l
\end{array}
\]

add exponent (number of msb shifts -1).

\[
\begin{array}{cccccccc}
\text{b7} & \text{b6} & \text{b5} & \text{b4} & \text{b3} & \text{b2} & \text{b1} & \text{b0} \\
\hline
0 & l & l & l & e & e & e & e
\end{array}
\]

The resultant value has a mantissa in which b6 is set to zero.

Solution
The value of b4, b5 and b6 should be in b5, b6 and b7, the intended value of b4 is lost.

Shift b4, b5 and b6 one place to the left and zero b4.

\[
\begin{array}{cccccccc}
\text{b7} & \text{b6} & \text{b5} & \text{b4} & \text{b3} & \text{b2} & \text{b1} & \text{b0} \\
\hline
l & l & l & 0 & e & e & e & e
\end{array}
\]
Table A.2: Errors on ISCALE: the integer values used in place of the correct, real values

<table>
<thead>
<tr>
<th>ISCALE</th>
<th>integer ISCALE</th>
<th>real ISCALE</th>
<th>error / %</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/1</td>
<td>10</td>
<td>10.0000</td>
<td>0.0</td>
</tr>
<tr>
<td>10/2</td>
<td>5</td>
<td>5.0000</td>
<td>0.0</td>
</tr>
<tr>
<td>10/3</td>
<td>3</td>
<td>3.3333</td>
<td>11.1</td>
</tr>
<tr>
<td>10/4</td>
<td>2</td>
<td>2.5000</td>
<td>25.0</td>
</tr>
<tr>
<td>10/5</td>
<td>2</td>
<td>2.0000</td>
<td>0.0</td>
</tr>
<tr>
<td>10/6</td>
<td>1</td>
<td>1.6667</td>
<td>66.7</td>
</tr>
<tr>
<td>10/7</td>
<td>1</td>
<td>1.4286</td>
<td>42.9</td>
</tr>
<tr>
<td>10/8</td>
<td>1</td>
<td>1.2500</td>
<td>25.0</td>
</tr>
<tr>
<td>10/9</td>
<td>1</td>
<td>1.1111</td>
<td>11.1</td>
</tr>
<tr>
<td>10/10</td>
<td>1</td>
<td>1.0000</td>
<td>0.0</td>
</tr>
</tbody>
</table>

A.3 Decompression Errors

A.3.1 Scaling Factor Error

Normally, ten spacecraft spins are used to accumulate each value of counts for transmitting. However, when an overflow occurs, counting stops at the end of the spin, then the counts recorded so far can be scaled up to give a value for 10 spins. So if an overflow occurs in the second spin, counting stops at the end of the two spins and the total count is multiplied by five. The factor used to scale up the counts is called ISCALE and is equal to 10 divided by a number between 1 and 10. In the archiving software ISCALE was originally defined as an integer instead of a floating point number.

The ISCALE error therefore only occurred in overflowed data. In a sample year (1990), it was found that ISCALE was used 3577 times, which is about 6% of the data points. Of these points, when the value of ISCALE is 5 or 10, there is no error. It is also fair to assume that in half of the cases when ISCALE is 2, and in one fifth of the cases when ISCALE is 1, there is no error (table A.2). Therefore only ~3% of all records for 1990 included non-zero errors due
to ISCALE. However, because the error only affects the highest counts it is not randomly distributed in the data. The effect was therefore more significant than the percentages suggest since specific events and preferred local times, i.e. the large injection signatures seen after midnight, were most affected. Note that where ISCALE was wrong, it was always too low, so caused a decrease in counts.

### A.3.2 Overflow Errors

A further error was found in the decompression software concerning the handling of count overflows in the data. When flux is high, the counts can overflow and an overflow flag is set so that the overflow can be taken into account when the counts are decompressed. However, in the decompression software the overflow flags were ignored.

The reason why all the overflows had been ignored was that, when the data was first inspected, overflows were seen in the polar and azimuthal counts which were obviously spurious. The overflows in the azimuthal counts were obviously false because the count levels were generally too low to overflow, since the azimuthal measurement is made in only one spacecraft spin. Similarly for the polar array the counts were too low to have overflowed. It was also found that where an overflow was signalled in both the energy and polar counts, the energy overflow was false: the counts were too low to overflow, and the overflow was signalled early on in the 10-spin accumulation, before enough counts could have been observed to cause an overflow. However, some of the overflows in the energy arrays appear to be genuine. This is evident because the counts in the higher energy bins are also high, and come close to overflowing in the next lowest energy bin. In these cases the energy overflow was flagged on its own, and this provided the algorithm for distinguishing genuine overflows.

The effect of the overflow error was to reduce the levels of some of the highest peaks in the data for which the counts reach values above the overflow threshold. Since more counts occur at lower energies making overflows more likely, the lower energy ranges of the detector will be most affected. This error could result in counts being orders of magnitude too low.
As described in section A.2, each number is recorded as two 8-bit words, i.e. a 16-bit binary number. If all 16 bits are set to 1, this represents the number 65535. Adding one more means the number overflows and the overflow flag is set and counting starts again: therefore the number recorded should have 65536 added to it, which was not done when the overflow flag was ignored. The error can then be scaled up further since following an overflow, counting stops and the count is multiplied up to represent 10 spacecraft spins. So if the overflow occurred in spin 1, the deficiency of 65536 counts is multiplied by ten. In addition, a single flux value may be calculated from several count values which contain these large errors.

The effect of the overflow error was that the peak flux was clipped because the overflows were ignored. In order to correct this error, each point where an energy overflow was flagged was examined to see if it was genuine, using the algorithm described. When an energy overflow was flagged without a polar overflow, the energy overflow looked genuine: counts were high and the overflow generally occurred in spin 9 or 10 of the accumulation.

The cause of the spurious overflows has yet to be established.
Appendix B

Coordinate Systems

B.1 GEO Coordinates

The GEO (geographic) coordinate system is shown in figure B.1. The $x$-axis is in the Earth’s equatorial plane, fixed with the rotation of the Earth so that it passes through $0^\circ$ longitude. The $z$-axis is parallel to the Earth’s rotation axis, and the $y$-axis completes the right-handed system $y = z \times x$ (Kivelson and Russell, 1995).

Figure B.1: The GEO coordinate system
B.2 GSM Coordinates

The GSM (geocentric solar magnetospheric) coordinate system is shown in figure B.1. The $x$-axis points from the Earth to the Sun. The $y$-axis is perpendicular to the dipole axis, with the $x$-$z$ plane containing the dipole axis. The $z$-axis completes the orthogonal system, in the sense of the north magnetic pole (Kivelson and Russell, 1995).
Acknowledgements

I wish to start by thanking Alan Johnstone for giving me the opportunity to do this work and for being my supervisor. Thanks to the many other people at MSSL who provided help and advice, especially David Rodgers, Andrew Coates, David Walton, Nick Flowers and Geraint Jones. I am also indebted to Mark Birdseye, Howard Huckle and Rob Gowan who provided the SEM-2 archiving software, and Barry Hancock for his speed in locating and rectifying the SEM-2 data compression errors. Thanks also to Graziella Branduardi-Raymont for being on my PhD panel, and to my ESTEC and BIRA/IASB colleagues on the TREND-2 study: Eamonn Daly, Hugh Evans, Cecille Tranquille, Joseph Lemaire and Daniel Heynderickx.

Thanks are due to the many providers of the hardware, data and software: to T. Cayton (LANL) for providing the detectors used in the SEM-2 on Meteosat-3; to Ph. Guilleman (CNRS Marseille) for writing the wavelet software OndeletS and J-L. Modave and F. Collin (ORB Bruxelles) for adapting it; to A. L. Vampola for RADMODLS; to IUGG: Association of Geomagnetism and Aeronomy (International Service of Geomagnetic Indices) for supplying $K_p$ numbers; to the National Space Science Data Center’s (NSSDC) OMNIWeb facility, for $K_p$ numbers and solar wind parameters; to SPIDR (Space Physics Interactive Data Resource, a service provided by the National Geophysical Data Center, Solar-Terrestrial Physics Division) for GOES magnetometer data; to N.A. Tsyganenko and M. Peredo for GEOPACK which was obtained from NSSDC; to World Data Center C1 for geomagnetism (Copenhagen) for magnetic $H$ data, for Bangui supplied by Bureau Central de Magnetisme, Bangui, Central African Republic, for Narsarsuaq supplied by Solar-Terrestrial Physics Division, Danish Meteorological Institute, for Tunguska supplied by Academy of Sciences of Russia, and for Faraday Islands supplied by British Antarctic Survey; to World Data Centre C1 for Solar-Terrestrial Physics (Rutherford Appleton Laboratory) for foF2 data.

I am grateful to University College London for a grant in support of my studentship, and for providing the opportunity for me to do such interesting
work.

On a personal level, I wish thank my mother and my brother Steve for constant support and encouragement. Thanks to Eva for brightening up MSSL. Finally and most importantly, heartfelt thanks to the indispensible Ady James for much help and inspiration.

This thesis is dedicated to my father, with love. Thanks for everything, Dad.
References


[36] Collin, F., Manuel d'utilisation du programme 'Ondelet' écrit par Ph. Guilleman (CNRS Marseille) adapte et traduit en language C par J-L. Modave et
F. Collin (ORB Bruxelles), (User manual for the 'Ondelet' program written by Ph. Guilleman (CNRS Marseille) adapted and translated into the C language by J-L. Modave et F. Collin (Royal Belgian Observatory)), GPS Report No 14 (Observatoire Royal de Belgique), 1992.


285


