Chapter 11

The Meteosat/SEM-2 data base

This chapter contains the new analysis of the archived data of the Space Environment Monitor SEM-2 onboard Meteosat. This data set provides a record of many aspects of the geosynchronous orbit environment. It is almost continuous from Nov 1988 to Nov 1995. A description of the Meteosat mission and of the SEM-2 instrument can be found in Technical Notes 6 and 7 of the TREND-2 study (the final report of TREND-2 contains a summary of this description).

Since the completion of the TREND-2 study, a number of problems have been found concerning the Meteosat-3/SEM-2 onboard data and postprocessing (see Technical Note 9 for more details). These errors became apparent in the unnatural clustering of the polar-azimuthal data to certain preferred flux values, and in regular spikes seen in some elements of the polar arrays.

During the TREND-3 study the data base has been re-analysed to correct the identified errors. The paper by Rodgers (1991) on Meteosat-3 anomalies has been updated with the corrected data base. The updated text is reproduced in Technical Note 9 of this study.

11.1 The new data base

The data base covers the period November 1988 to November 1995, with missing months March 1991, April 1991, and July 1995. The data format is the same as that of the data base used in TREND-2, as described in Table 11.1. The data base files have file names of the types `mmmyy_archive_hr.dat` and `mmmyy_archive_lr.dat`, e.g. `nov88_archive_hr.dat`.

11.2 Effect of Data Correction on TREND-2 Results

11.2.1 Overview of the Error Correction

The effect of the numerous errors in the Meteosat SEM-2 data is evaluated here. Overall, the corrected data base does not differ substantially from the old one for most of the time.

Table 11.1. Record structure of the Meteosat/SEM-2 archived data set record

Record	Variable
1	Start time of bin in hours UT
2	End time of bin in hours UT
3	Total flux of electrons in energy range 42.9–300 keV, summed over all polar and azimuthal bins
4–8	Flux in each energy bin, summed over all polar and azimuthal bins: 4: energy range 201.8–300 keV 5: energy range 134.9–201.8 keV 6: energy range 90.7–134.9 keV 7: energy range 59.4–90.7 keV 8: energy range 42.9–59.4 keV
9	Spectral index: the slope of the logarithm of the energy spectrum, calculated using a least squares fit
10	Delta spectral index: the error on the above calculation
11–15	Polar flow: flux in one of five polar angle sectors of the analyser (bins are approximately $\pm 5°$): 11: $150°$ to spin axis 12: $120°$ to spin axis 13: $90°$ to spin axis 14: $60°$ to spin axis 15: $30°$ to spin axis
16–21	Azimuthal flow: flux in one of six azimuthal angle sectors of the analyser. The angles are (in spacecraft coordinates, where at 0° the spacecraft looks towards the Sun): 16: 300°–360° 17: 240°–300° 18: 180°–240° 19: 120°–180° 20: 60°–120° 21: 0°–60°
22–51	Polar-azimuthal flux: in 30 bins, for each polar angle sector across the azimuthal angles sector
52–56	Counts in each of the five energy ranges (see record 4–8)

Table 11.1. (Continued)

Record	Variable
57–86	Polar-azimuthal counts: again, in 30 bins
87	Anisotropy: the anisotropy index describes the angular shape of the plasma distribution relative to its axis of symmetry
88–89	θ And ϕ: angles describing the angular shape of the plasma distribution relative to its axis of symmetry
90	K_{p} Index
91	$K_{\mathrm{p}}(\tau)$: Weighted average of successive K_{p} (Wrenn 1987)
92	Latch: occurrence of latch-ups in the test RAM (random access memory)
93–96	MUM: Memory upset monitors give number of SEUs in the four memory zones of the test RAM.

- The onboard compression problem is the most common. Although capable of producing errors of 50%, it usually produced errors much lower than that. The effect is always a decrease in the observed flux. Because the output flux arrays are averages of many count values, the effects of this error are generally diluted. No significant changes in the average flux or the dependence on LT are expected.

- The `ISCALE` error, occurs only for those rare events when an overflow is flagged. The effect is usually a small depression of flux and because of averaging, it is again significantly diluted.

- The overflow error can have a substantial effect when flux is very high, i.e. during substorm injection events. This error causes a substantial decrease in flux and, because real overflows may occur one after another, may lead to sections of the flux time sequence being depressed. By affecting the peak height of events, this error may cause a change in variance of 43–60 keV electrons, because the flux of the most significant events will be higher. But whilst the heights of these events will change, there will not be changes in timing so that the dispersion seen by the superposed epoch analysis and the time correlation analysis will be unaffected.

  Although the overflow error was potentially very serious, there are believed to be only about 150 genuine overflows out of the 24,000 which were flagged over the seven-year data set. Out of a total data set of over 400,000 data points, this is a very small number of bad points, although statistically it is obviously more important than the number of points implies because the erroneous data points are not randomly distributed throughout the data set, but rather occur during the most energetic injection events.

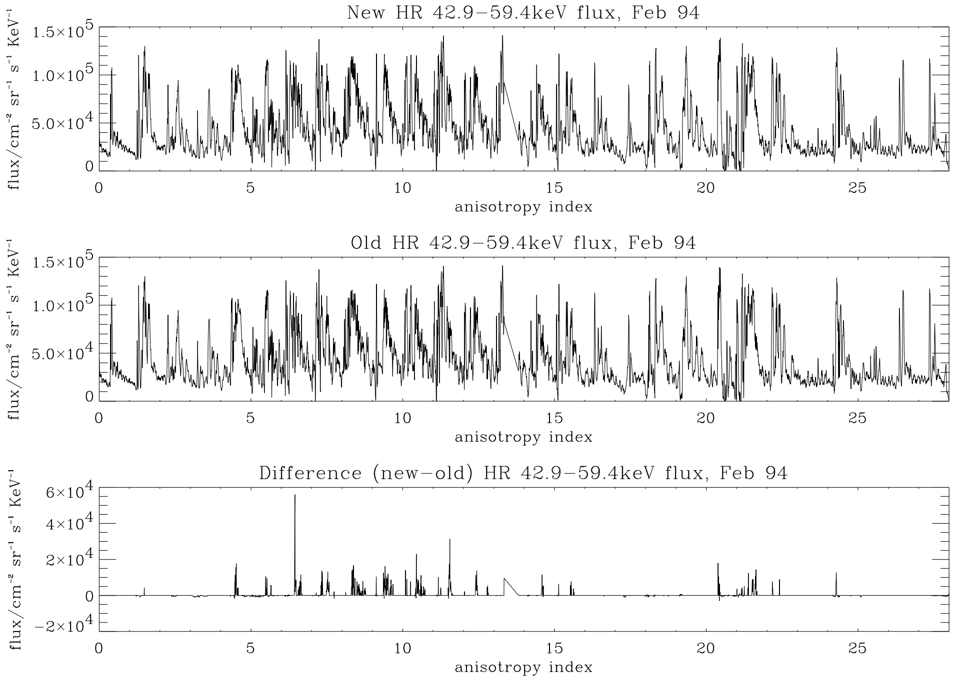

Figure 11.1. E5-E4 flux, before and after data set corrections

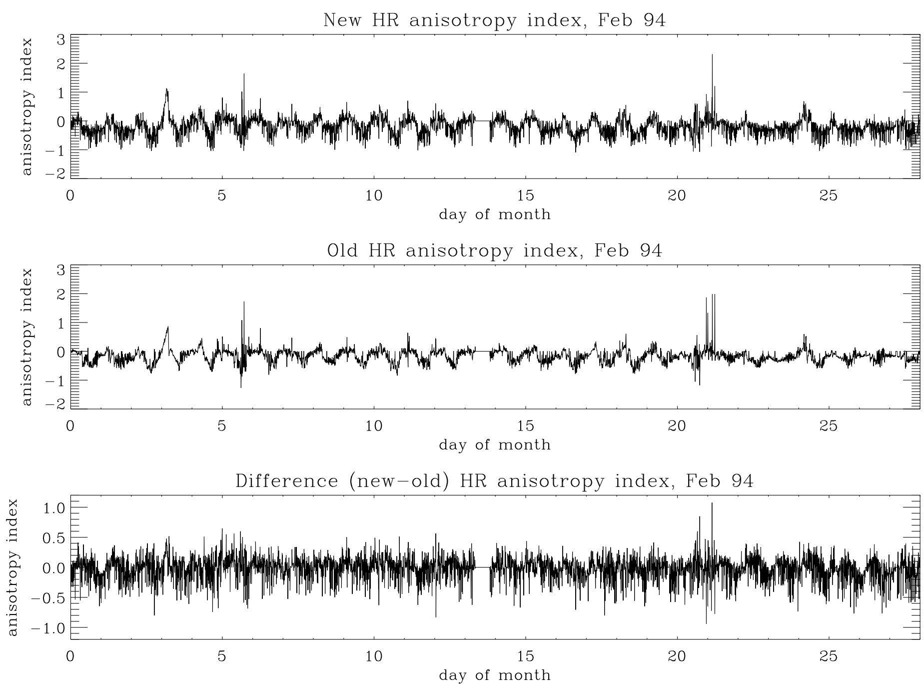

Figure 11.2. Anisotropy index, before and after data set corrections

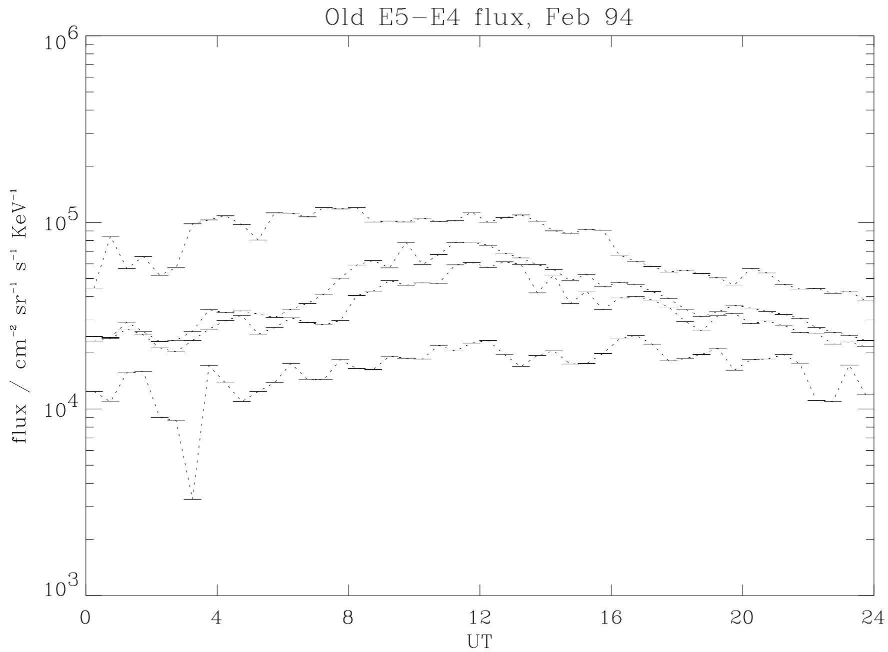

Figure 11.3. E5-E4 flux statistical plot, before data set corrections

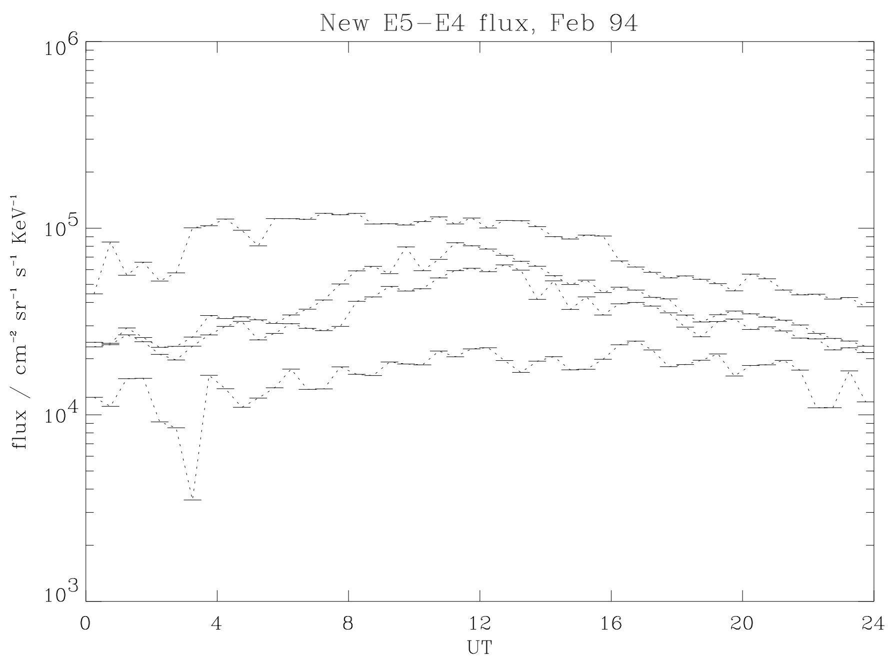

Figure 11.4. E5-E4 flux statistical plot, after data set corrections

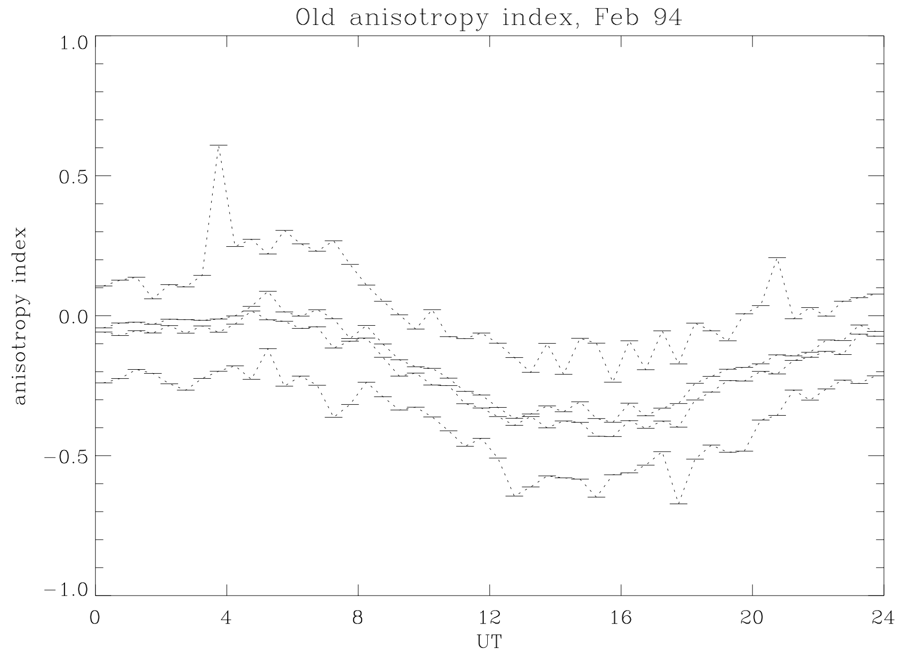

Figure 11.5. Anisotropy index statistical plot, before data set corrections

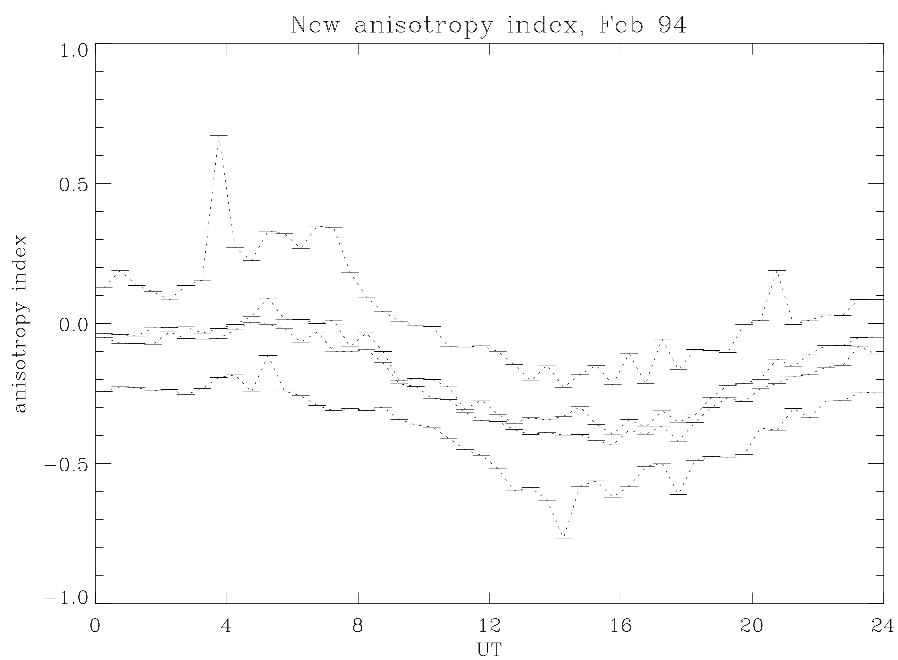

Figure 11.6. Anisotropy index statistical plot, after data set corrections

11.2.2 Comparison of 'Before' and 'After' data

Some example plots of data before and after all the errors were corrected are shown in Figs. 11.1 and 11.2. Figure 11.1 shows the high time resolution E5-E4 flux (42.9-59.4keV) for February 1994. This month was chosen because it contained the largest number of genuine overflows: a total of 16. The E5-E4 flux is shown since it is likely to contain the most genuine overflows and therefore the largest errors. Note that all energies include some errors, even the highest energy which does not overflow. Therefore, the differences between the new and old data for this month are likely to be the most severe. The upper two panels of Fig. 11.1 show the new and old data respectively. The diurnal variation can be clearly seen in the flux data, with peak fluxes of the order of 1.0–1.4×10^5cm^{-2}sr^{-1}s^{-1}keV^{-1} occuring daily. The bottom panel of Fig. 11.1 shows the difference between the new and old flux data, as *new minus old*. The maximum difference is less than 6.0×10^4cm^{-2}sr^{-1}s^{-1}keV^{-1}. The plot shows that in the old data set, where data points were wrong, their values were generally underestimated. The errors look relatively small, but because of the overflow error the largest errors occur where the flux was severely underestimated: the maximum difference point here corresponds to a flux value which, in the old data set, was 66% too low. However, since most of the erroneous points were not affected by the overflow error but by the less significant errors, most of the errors were much lower. In fact, less than 5% of the 4408 data points that make up February 1994, were out by more than 5%, and only 1.75% of points were out by more than 10%.

The high time resolution anisotropy index is shown in Fig. 11.2. The anisotropy index is affected by the errors because it is calculated from the polar-azimuthal count arrays, which were found to contain strong effects of the decompression and `ISCALE` errors. These effects are much less obvious in the flux arrays because of the averaging that has been done to produce them. The maximum difference between the new and old values is about 1.0 compared with peak anisotropy index values of about 2.0. Many more anisotropy points are affected than flux points.

In the course of the TREND studies, many statistical models were produced using long-term averaging of the SEM-2 data. Some sample plots have been produced to show how these models may have been affected by the errors.

Figures 11.3 and 11.4 show statistical plots of the E5-E4 flux data for February 1994, before and after error corrections respectively. These plots use the low time resolution data (30 minute averages). Note that the x-axis is UT rather than LT (Meteosat was at longitude 74° west at this time, so local midnight is at about 0500 UT on these plots). These plots show four lines: the upper and lower lines are the levels below and above which 95% of the data are observed, and the middle two lines show the median $\pm$5% of observations. Although the plots use only one month of data and, therefore, can be expected to show a fair amount of scatter, the two plots show hardly any differences.

Similarly, Figs. 11.5 and 11.6 show statistical plots of the anisotropy index before and after error corrections respectively. Again, the plots show very few differences.

11.2.3 Conclusions

We conclude that the errors did not affect the results of the previous TREND studies. Although the various errors could have significant effects on individual data points, most of the study results relied on statistical studies where the effect of including a few erroneous points in a large sample of data would have been small.

The example plots of ‘before’ and ‘after’ data presented here have shown that the statistical models produced in the TREND reports are unlikely to change significantly with the correction of the errors. The important results are the local time dependence and overall flux range, neither of which should change significantly. This conclusion should hold for all the local time dependence models, for the Fourier and wavelet analysis, for the correlation analyses showing the lag between different energy flux, and the flux probability versus mission duration models. It is our belief that reproducing these models and analyses with the corrected data set would yield very similar results and the same conclusions.

The work most likely to have been affected by the errors would have been the studies of the flux peaks due to substorm injections which used superposed epoch analysis. This work used a sample of 200 flux peaks from the year 1989. This year was found to have 42 genuine overflows, some of which will have been included in the sample. However, generally only the lowest energy bin is affected, and although the flux level will rise, its other attributes such as duration should not change unless, of course, there is a sustained period of high flux.