Venus Express ASPERA-4 ELS Background Data Examples

Written by: Dr. Rudy Frahm (rfrahm@swri.org)

This document attempts to answer the following questions:

1. What is background and how should I treat background?

If you are measuring an object and your measurement device is perfect, then you will recover the environmental signal. However, our measurement devices are not perfect, and as a result, what we measure is a mixture of the environmental signal modified by the device in some manner by mixing the environmental signal with some type of coherent or incoherent noise. In this archive we deal with a source of noise which is produced when we make electron measurements of the environmental energy spectrum. This noise is called instrument background. The type of instrumental background which will be dealt with is caused by any energy independent source of noise which adds to the signal from the environment.

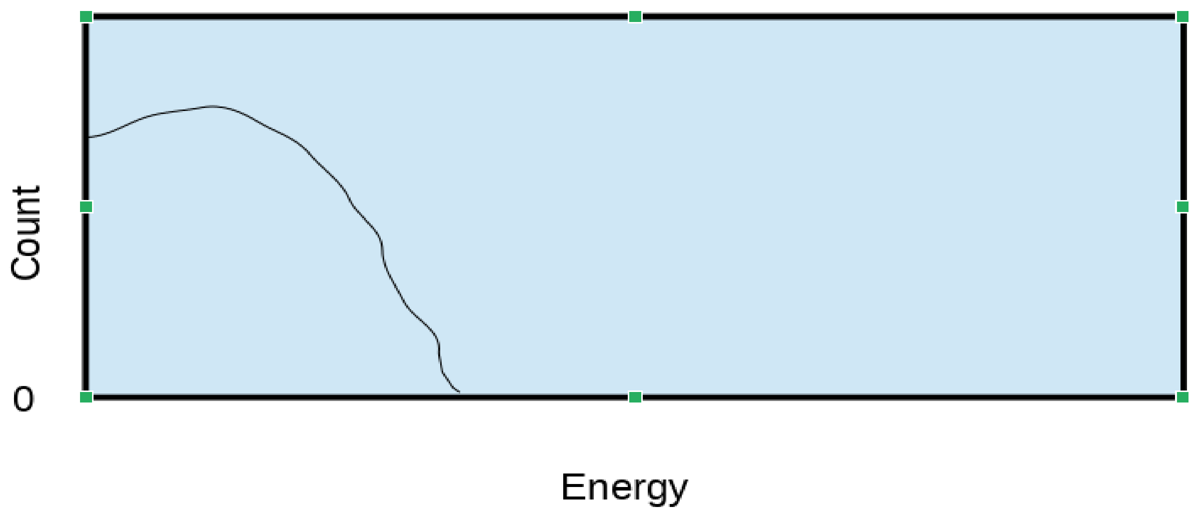

Figure 1. Measurement of an energy spectrum with a perfect device. For a perfect detector measuring the energy spectrum in the environment, the measurement corresponds to the signal obtained directly from the environment.

If the instrument which we use to measure the energy spectrum were perfect, then we would recover the environmental signal. Say this environmental signal looks like the energy spectrum shown in Figure 1. Since the detector is really not perfect and adds a source of an energy independent noise to the signal, the measured spectrum will be modified as shown in Figure 2. Here we address one class of background noise, one which adds a noise signal to the environmental signal and this noise is energy independent. This means that the source of noise has an equal probability of adding to any energy measurement. These types of noise sources can be sensor related (e.g. the temperature may randomly produce extra counts within the detector), electronic related (e.g. electronics may produce leakage currents which get amplified and interpreted as counts), or geophysical (e.g., penetrating radiation through the sides of the instrument can cause a signal within the detector), as these are examples of noise sources which are energy independent.

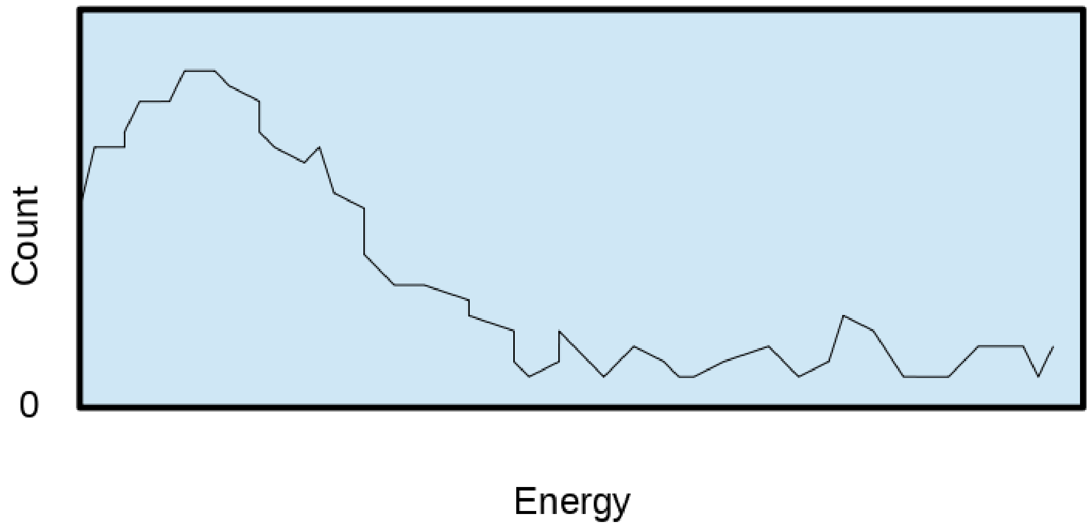

Figure 2. Environmental signal modified by an energy independent source of noise. For a non-perfect detector measuring the energy spectrum in the environment, the measurement corresponds to a combination of the signal obtained directly from the environment and the background noise created by the detector.

Once the environmental signal has been modified by the noise signal, statistical methods must be used in order to recover the environmental portion of the signal from the measured signal (Figure 2). We must realize that it is not possible to recover the exact environmental signal because we do not know the exact energy distribution of the noise included within each measurement. However, if we can determine a portion of the energy spectrum where there is no environmental signal, then this portion is all noise signal and we can use statistical methods to remove the noise signal from the measured signal in order to recover an approximation to the environmental signal.

To determine the noise background, a portion of the energy spectrum in which we are sure there is no environmental signal is used so that we are characterizing the noise portion of the spectrum. Since the noise portion is energy independent, it should not matter where in energy we select, only that the energy measurements that we select are free of the environmental portion of the measured signal.

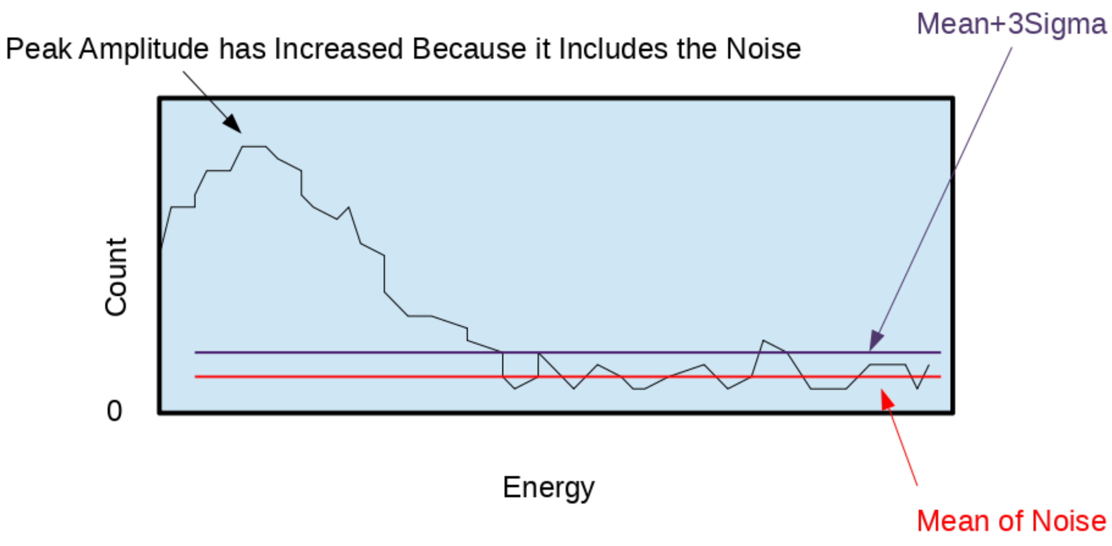

Figure 3. Measured spectrum where the environmental signal is increased due to adding noise to all energies. The background noise mean and mean+3*sigma levels are shown.

Expressed in terms of counts, the mean of the noise distribution, figured over the energy range where the measured signal does not include the environmental signal, is shown in Figure 3. The mean of the noise distribution is given by

$$Mean = \frac{1}{N} \sum_{i=Energy}^{N} Count(i) \tag{1}$$

where there are N samples of the noise signal located at Energy values (i). We recognize that the noise does contain a distribution of values, so that the standard deviation of the noise (σ) is the square root of the distribution variance (σ^2).

$$Variance = \frac{1}{N} \sum_{i=Energy}^{N} (Count(i) - Mean)^2 \tag{2}$$

The standard deviation describes how much variability there is in the mean

$$Standard\ Deviation = \sqrt{Variance} \tag{3}$$

Corrected spectra should remove the mean count rate background noise value from the measured distribution.

Researchers often take three times the mean to be the sensitivity level below which the signal is not useful. An example of this level is also shown in Figure 3. However, one must be cautious when using this method. If after using a cut level and setting spectral values less then this level to zero, a problem exists when averaging these corrected spectra. In the example shown in Figure 3, there is some part of the noise signal above the cut line. If one accumulates and averages these corrected spectra, then these small number of counts accumulate and you end up with additional level of noise which is not (and can never be) removed. If however, you maintain the negative values resulting from subtraction of the measured spectrum minus the mean of the noise background, then when you average

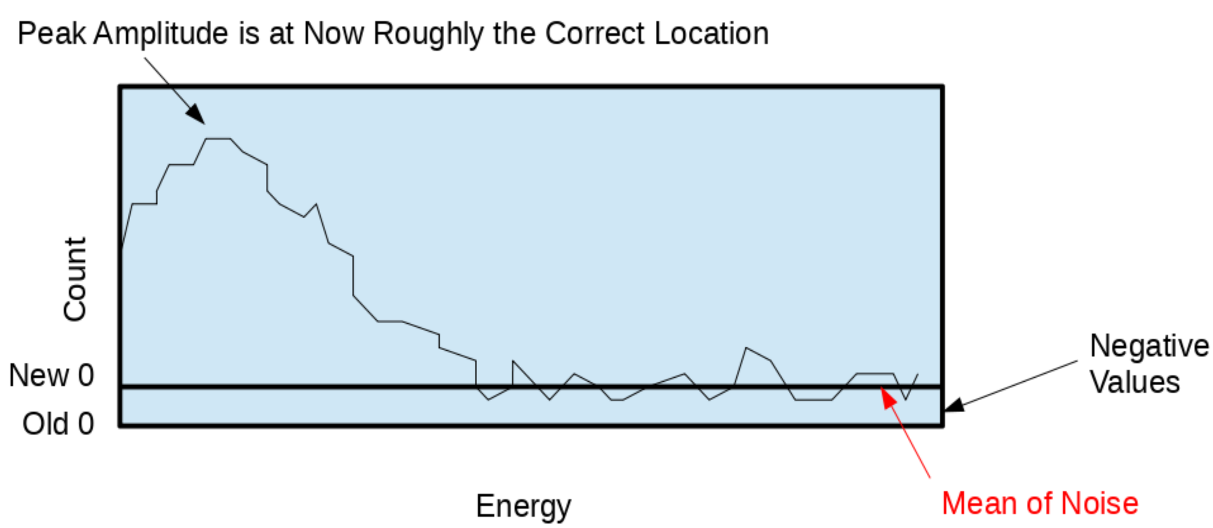

Figure 4. Negative values are obtained from measured values which are less than the mean. These values are statistically generated and result from numbers less than the mean. They are unphysical, but need to be kept for use in further statistical analysis.

over a large number of spectra, the noise values reduce. Figure 4 shows that the negative values arise from counts which are less than the mean. Averaging a large number of correct spectra causes the noise to approach the mean, which in the further corrected spectra, have values closer to zero. Thus, if you are going to average corrected spectra, then keep track of negative values and use them when determining the average spectra. This will give you the lowest possible amount of background noise within your corrected spectra. Since the negative values are purely statistical, be sure to remove any remaining negative values when you achieve a final product as the negative values are unphysical.

2. Can you show me an example of a background data file?

In this archive there are background data files which represent the energy independent noise averaged over 5 minute intervals. In most cases, the 5 min average time is adequate; however, there may be times when the researcher determines that the 5 min average background values are too high or too low for their research. For these cases, the researcher may want to generate new background values and remove them from the data. This procedure is described in the Venus Express ASPERA-4 ELS Description Document (VEX_ASPERA-4_ELS_Description.pdf) for this data set.

The background files are in table format (TAB). There are 36 columns in the table. The first column represents the beginning time the background data is valid. The second column represents the ending time when the background data is valid. Time is given in YYYY-DDDTHH:MM:SS.mmm format where YYYY is the year, DDD is the day of year (January 1 is day 1), HH is the hour of the day, MM is the minute of the hour, SS is the second of the minute, and mmm is the millisecond of the second. Accumulation times (the difference between the start and stop times) are not exactly 5 minutes due to data gaps and sampling rates. Always check the beginning and ending times for the exact time interval over which the background count rates apply.

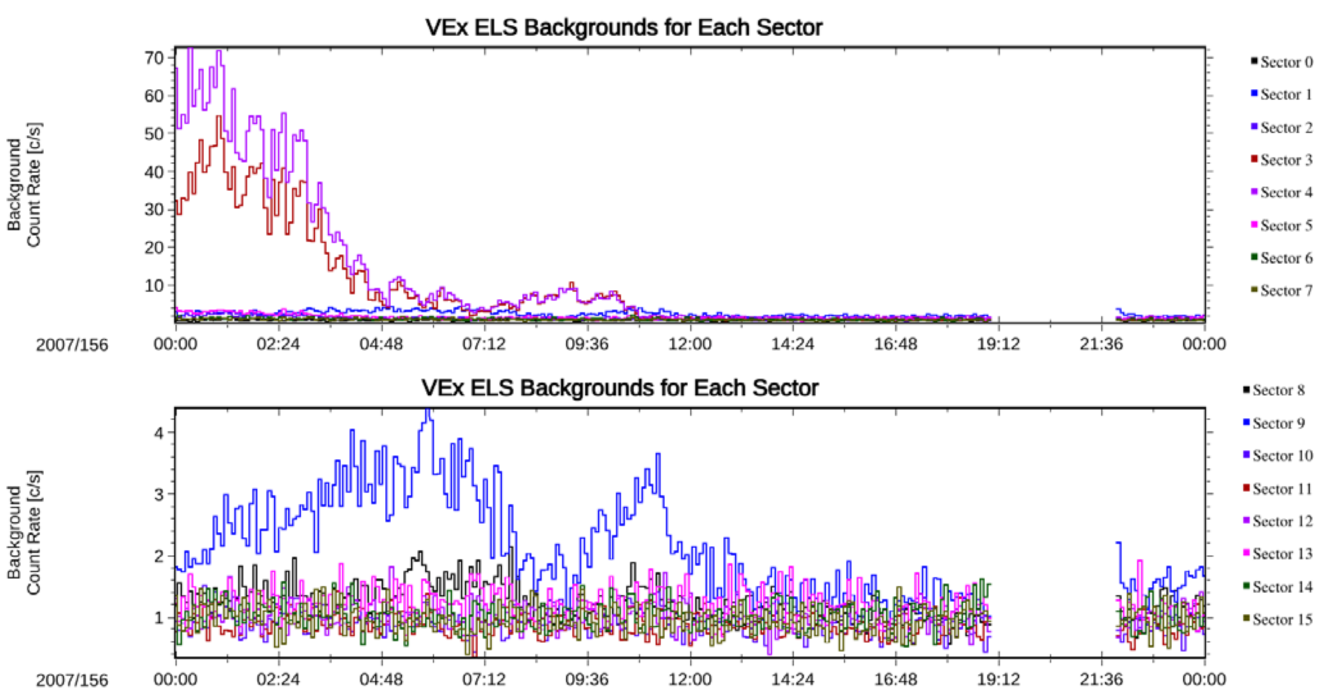

Figure 5. Background count rate for day 2007156. Count rates are shown for the first 8 ELS sectors in the top panel and the second 8 in the bottom panel. Count scales differ between the two panels so that all data is included. Some sectors shown in the upper panel have much higher count rates then other sectors. Count rates in the lower panel and upper panel show variability both in time and shape.

Columns 3 through 19 represent the background count rate in counts/second for the total MCP and then each of the 16 anodes from the Electron instrument (ELS) on the Venus Express (VEx) Analyzer of Space Plasmas and Energetic Atoms (ASPERA-4) experiment (Note: ELS sectors are numbered 00-15). The background values are kept separately because some ELS sectors contribute more to background count rates than other sectors. A total background rate for the MCP is also given so that it can be easily seen how much total noise is generated by the MCP. An example of the variability in the background count rate for each sector is shown in Figure 5.

Columns 20 through 36 represent the standard deviation of the background count rate for the total MCP and then each of the ELS sectors (Note: ELS sectors are numbered 00-15). The standard deviation of the count rate describes the width of the count rate distribution for each of the 5 minute averaged times and for the ELS sector. Standard deviations are larger if there is a large spread in the count rate distribution and smaller for a narrower distribution. An example of the standard deviation of the count rate distribution is shown in Figure 6.

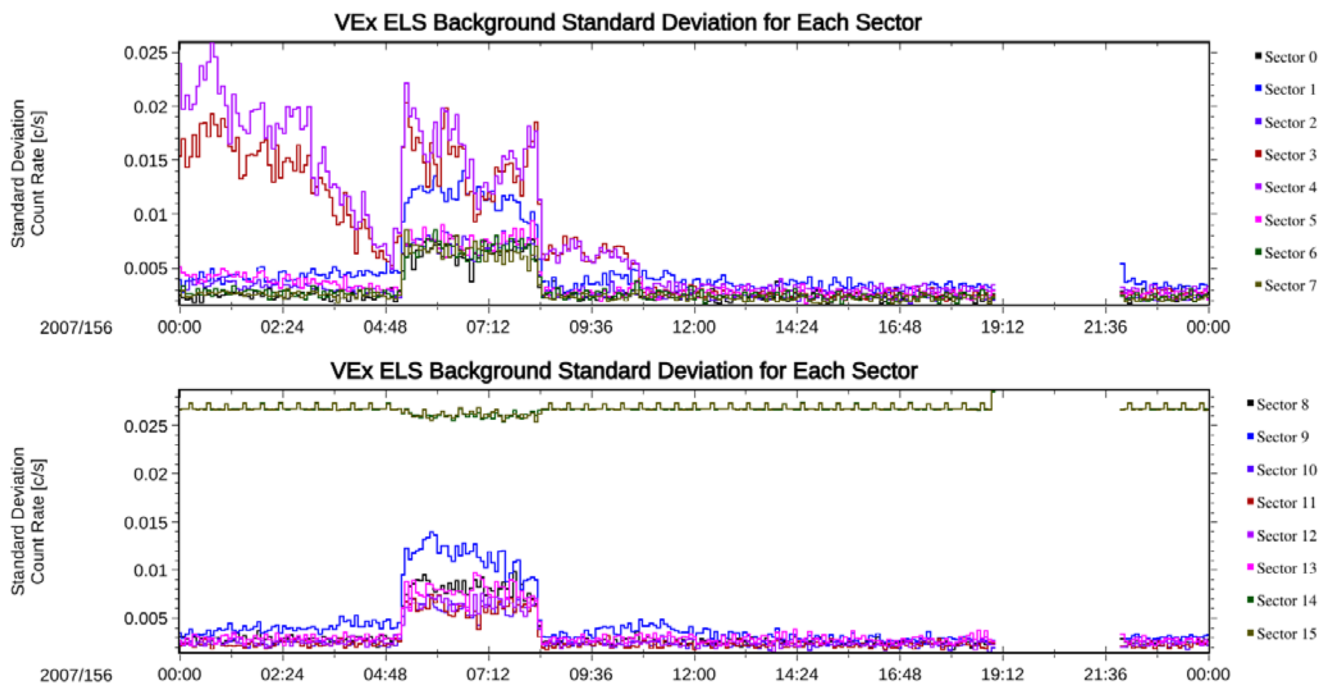

Figure 6. The standard deviation of the background count rate distributions from day 2007156. Although the format is the same as Figure 5, the standard deviations show a difference in the distribution shapes of the background count rates for each sector. Time differences in the distribution shapes are also observed and are most prevalent in the lower panel.

This particular day (2007156) was chosen to show an example of the effect of low count rates on the standard deviation. Observed in sectors 14 and 15 from Figure 6 is an oscillating signal in the standard deviation count rate above 0.025 counts/sec. This is caused by the statistics of small integer numbers coupled with a variable accumulation rate and the results are valid. As a more detailed explanation of this effect, the standard deviation count rate is determined by normalizing the standard deviation count (equation 3) with the accumulation time for the valid values obtained during the sweep over roughly 5 minutes. This accumulation period can differ from sweep-to-sweep due to the requirement that the accumulation period is set by an integer multiple of sweeps coupled with the instrument duty cycle (missing 1 out of every 8 spectra), causing a difference in the valid accumulation period – 1 out of every 5 background values in the Figure 6 example. Because of low count rates, the standard deviation count is about the same from measurement to measurement. Therefore, dividing by a smaller time base causes a slightly larger value in the standard deviation count rate – in this case every 5th sample. Physically, this means that the width of the background count spectra has a slightly larger error every 5th sample because its determination is not as accurate as at other times. Thus, this translates into a few percent larger error in the background count rate every 5th spectra because it is determined from a smaller accumulated sample. This is a valid result.

3. Can you show me an example of how to visualize background corrected data?

The Venus Express ASPERA-4 ELS background corrected data is organized into comma separated variable (CSV) files, where fill values are not supplied but are implied by adjacent commas without a value between. In the case of the end of a line, the fill value is implied when the last comma is followed by the end-of-line character (carriage return/line feed). The CSV file is a NASA Planetary System Standard and is meant to be read by a spreadsheet program. A CSV file is an ASCII text file which can be viewed or read by a program written by the user; however, the fill values must be handled properly. When processing the data read from CSV data files, locations where a fill value occurs should be ignored as if no measurement was made at that location. Any value of fill can be used which would not normally be contained in the data. For these examples, the value of -1.0e+31 was used for the fill value as it could easily be identified and would not conflict with any negative statistical (geophysical) values contained within the data files.

There are 8 types of CSV files in this archive which fall into two groups. The two groups are referred to as high range data and the low range data. The high range data is prefaced by the letters "ELSHR" and the low range data by "ELSLR". For this instrument, the range refers to the power supply used internally within ELS to generate the stepped sweep (energy levels). There are 4096 possible values for high range power steps and 4096 possible values for low range power steps; however, only a small subset of values were used. The ELS stepping sweep decays from its highest value of energy to its lowest value of energy. Within a stepping sweep, there are 127 values with each step acquiring data for 28.125 msec and a latency between steps of 3.125 msec (power supply transition), for a maximum stepping rate of 32 steps per sec. There is one flyback step at the end of the sweep where no valid science data is collected and it is excluded from the PDS-4 data files. This gives a sweep having 127 steps (numbered 0 to 126). The 127 sweep steps may be made up of a combination of high and low range values in various proportions. For each high and low range data, the steps are translated into the energy of the diagnosed electron and stored in the CSV files which include the label "STPS" within their file name.

The file formats for the CSV files are similar. The first column contains the beginning time of the data presented on the rest of that same line. The second column represents the ending time of the data shown on the rest of the same line. Time is given in YYYY-DDDTHH:MM:SS.mmm format where YYYY is the year, DDD is the day of year (January 1 is day 1), HH is the hour of the day, MM is the minute of the hour, SS is the second of the minute, and mmm is the millisecond of the second. The third column is either the word "SCAN" or the word "SENSOR". "SCAN" refers to the step sweep data which are the energy levels (in eV) for each step of the energy. The word "SENSOR" refers to the data measurement shown in the file name. Besides the "STPS" keyword, the labels "CNTS", "THR", and "BCNF" are also available: "CNTS" refers to the counts which were measured, "THR" refers to the 1-count threshold differential number flux (cnts/[cm^2 s sr eV]), and "BCNF" refers to the background corrected differential number flux (cnts/[cm^2 s sr eV]). The fourth column is the spacecraft-sensor-sector-power range for the data (spacecraft is VEx, sensor is ELS, sector is 00 through 15, power range is either HR or LR). The fifth column is the units of the data. From the sixth column on, the data values are given. The length of the line is dependent on the operational mode of the instrument at the time the data was acquired and may not be a fixed value from file-to-file in each energy range; however, the total number of steps in the combination of high range and low range is always 127 values.

To use the data, the “STPS” files determine the energy level of the spectrum and the “BCNF” files determine the background corrected differential number flux. To reconstruct the energy sweep, high range data comes before low range data. Energy steps are given in sequential order and data values correspond 1-to-1 with the energy step values. The background corrected spectral data is shown in Figure 7. Also shown on Figure 7 is the uncorrected differential number flux to show an example of how much influence the background has on the uncorrected spectral data. In order to obtain the uncorrected differential number flux, multiply the value obtained from the “CNTS” file with those obtained from the “THR” file (e.g., uncorrected differential number flux = counts per accumulation * 1 count threshold). The values in the “CNTS” file represent the number of counts in an accumulation period (28.125 sec) and could be a fractional value if the instrument was in a telemetry compression mode. The values in the “THR” file represent the differential number flux obtained if there was 1 count obtained during the accumulation period. If generating average data as in Figure 7, be sure to generate the value to be averaged (in this case, the uncorrected flux) at each time step before accumulation, and then average the result. Be sure to exclude any times represented by filled values and include any negative flux values (as in the corrected differential number flux). The numbers obtained and shown in Figure 7 are presented in Table 1.

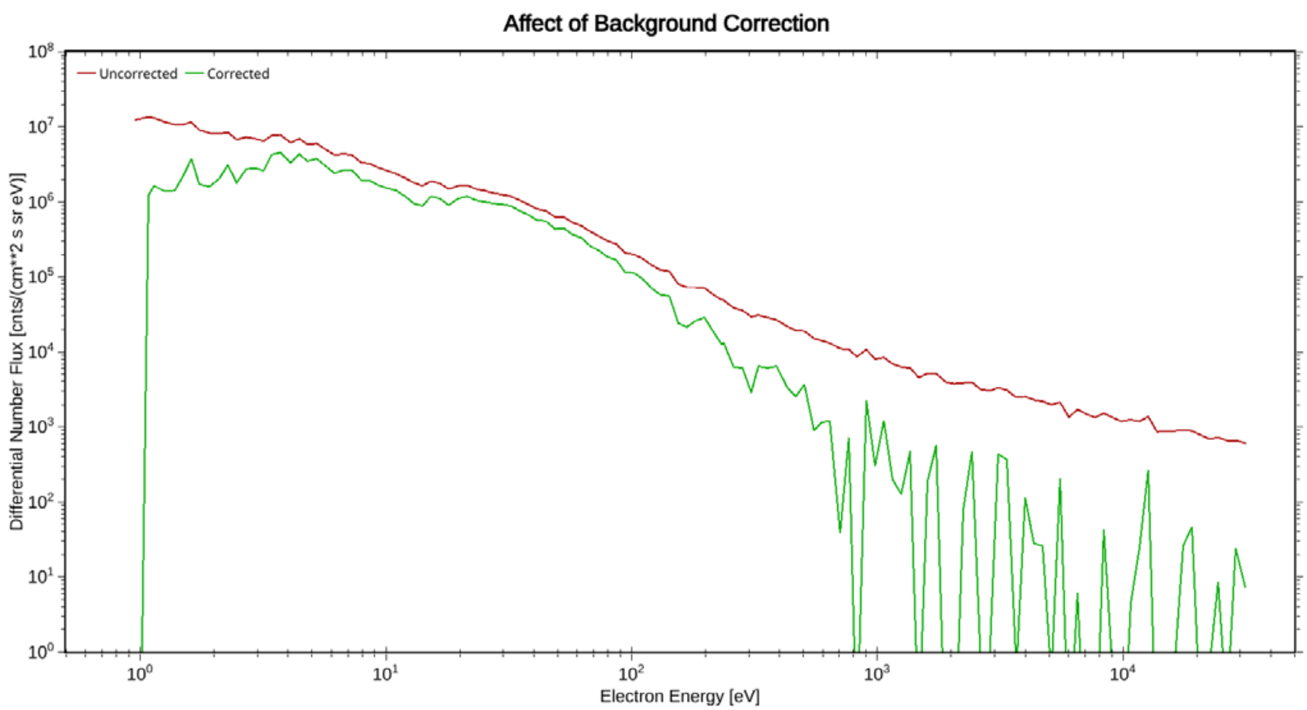

Figure 7. Influence of background correction. The uncorrected spectrum is shown in red for ELS sector 08 while the background corrected spectral data is shown in green. In order to display these data, negative values were set to 1. These data are the result of a straight average from the time range 2013-064 at 14:14:59.866 to 2013-064 at 14:20:19.302.

The corrected differential number flux is also obtained by subtracting the accumulation time adjusted background value from the same time frame as the count per accumulation, and then multiplying by the 1 count threshold (e.g., corrected differential number flux = [count per accumulation – background * accumulation period] * 1 count threshold); however, the corrected differential number flux has been provided in the "BCNF" file (recall that the accumulation period is 0.028125 sec and the background is given in counts/sec). The counts and threshold values are given in case (1) there is a need to determine the uncorrected differential number flux (as shown in Figure 7) or (2) the user determines a need to use their own computation of the background. Determination of background is given elsewhere in the documentation which is associated with this archive. Reasons for the user wanting to generate a different time cadence may involve either not enough background has been removed (e.g., the background determined by the 5 minute average is too small/large for the event, and a larger or smaller time base needs to be used to give a better estimate of the background) or too much background has been removed (e.g., the background determined by the 5 minute average is too small/large for the event, and a larger or smaller time base needs to be used to give a better estimate of the background). These are not the only conditions under which the user may require the use of a different background value; the choice has to be determined by the user.

Table 1. Corrected and uncorrected average result shown in Figure 7. These data are the result of a straight average from the time range 2013-064 at 14:14:59.866 to 2013-064 UT at 14:20:19.302 UT. The last column performs a check on the differential number flux average shown in column 3 by using the formula corrected differential number flux = [count per accumulation – background * accumulation period] * 1 count threshold. This computation was performed before averaging. Fill values have been excluded.

Step	**Energy [eV]**	**Differential Number Flux [cnts/(cm^2 s sr eV)]**	**Uncorrected Number Flux [cnts/(cm^2 s sr eV)]**	**Corrected Number Flux (Check)**
0	31422	7.5306E+00	6.1224E+02	7.3984E+00
1	28928	2.4365E+01	6.6242E+02	2.4365E+01
2	26641	-1.5071E+01	6.5851E+02	-1.5071E+01
3	24531	8.7052E+00	7.2037E+02	8.7050E+00
4	22590	-6.2387E+01	6.8984E+02	-6.2386E+01
5	20802	-1.0914E+01	7.8443E+02	-1.0914E+01
6	19152	4.6675E+01	8.8783E+02	4.6675E+01
7	17641	2.6267E+01	9.1562E+02	2.6267E+01
8	16244	-6.1722E+01	8.7869E+02	-6.1722E+01
9	14955	-1.1688E+02	8.7752E+02	-1.1688E+02
10	13773	-1.8719E+02	8.6382E+02	-1.8719E+02
11	12684	2.6342E+02	1.3741E+03	2.6342E+02
12	11679	2.4509E+01	1.1981E+03	2.4509E+01
13	10758	4.4379E+00	1.2439E+03	4.4378E+00
14	9906.1	-1.0857E+02	1.2006E+03	-1.0857E+02
15	9115.8	-5.4880E+01	1.3283E+03	-5.4880E+01

Step	Energy [eV]	Differential Number Flux [cnts/(cm^2 s sr eV)]	Uncorrected Number Flux [cnts/(cm^2 s sr eV)]	Corrected Number Flux (Check)
16	8394.5	4.3139E+01	1.5037E+03	4.3139E+01
17	7734.6	-1.9456E+02	1.3472E+03	-1.9456E+02
18	7120.7	-1.3507E+02	1.4935E+03	-1.3507E+02
19	6552.9	6.1628E+00	1.7272E+03	6.1618E+00
20	6038.8	-4.6522E+02	1.3523E+03	-4.6522E+02
21	5555.4	2.0646E+02	2.1288E+03	2.0646E+02
22	5118.0	-4.5471E+01	1.9868E+03	-4.5471E+01
23	4711.3	2.6311E+01	2.1772E+03	2.6310E+01
24	4343.0	2.7837E+01	2.3036E+03	2.7837E+01
25	3997.7	1.1298E+02	2.5252E+03	1.1298E+02
26	3683.1	-5.7220E+01	2.5002E+03	-5.7220E+01
27	3391.6	3.6205E+02	3.0770E+03	3.6205E+02
28	3123.0	4.3469E+02	3.3200E+03	4.3469E+02
29	2869.8	-6.8792E+01	3.0058E+03	-6.8792E+01
30	2647.3	-7.3181E+01	3.1976E+03	-7.3181E+01
31	2432.4	4.6600E+02	3.96E+03	4.6600E+02
32	2240.6	7.7917E+01	3.8091E+03	7.7919E+01
33	2064.1	-2.2729E+02	3.7617E+03	-2.2729E+02
34	1903.0	-2.8011E+02	3.9877E+03	-2.8011E+02
35	1749.5	5.7159E+02	5.1552E+03	5.7159E+02
36	1611.4	1.8797E+02	5.1103E+03	1.8797E+02
37	1480.9	-7.6110E+02	4.5429E+03	-7.6110E+02
38	1365.8	4.6471E+02	6.1705E+03	4.6471E+02
39	1258.4	1.2848E+02	6.2809E+03	1.2848E+02
40	1158.7	1.9634E+02	6.8440E+03	1.9634E+02
41	1066.6	1.2084E+03	8.4026E+03	1.2084E+03
42	982.17	2.9762E+02	8.0915E+03	2.9762E+02
43	905.44	2.2222E+03	1.0667E+04	2.2222E+03
44	828.71	-6.0572E+02	8.6233E+03	-6.0572E+02
45	767.32	7.2649E+02	1.0707E+04	7.2649E+02
46	705.93	3.8941E+01	1.0914E+04	3.8937E+01

Step	Energy [eV]	Differential Number Flux [cnts/(cm^2 s sr eV)]	Uncorrected Number Flux [cnts/(cm^2 s sr eV)]	Corrected Number Flux (Check)
47	644.55	1.1808E+03	1.3139E+04	1.1808E+03
48	598.51	1.1649E+03	1.4095E+04	1.1649E+03
49	552.47	9.0310E+02	1.4983E+04	9.0310E+02
50	506.43	3.6664E+03	1.9125E+04	3.6664E+03
51	468.07	2.5363E+03	1.9371E+04	2.5363E+03
52	429.7	3.4233E+03	2.1901E+04	3.4233E+03
53	391.33	6.4492E+03	2.6921E+04	6.4492E+03
54	360.64	6.0868E+03	2.8483E+04	6.0868E+03
55	329.95	6.5009E+03	3.1207E+04	6.5008E+03
56	306.93	2.8743E+03	2.9637E+04	2.8743E+03
57	283.91	6.1622E+03	3.5337E+04	6.1622E+03
58	260.89	6.2125E+03	3.8250E+04	6.2125E+03
59	237.87	1.3328E+04	4.8811E+04	1.3328E+04
60	233.78	1.2605E+04	4.8627E+04	1.2597E+04
61	215.28	1.8887E+04	5.8283E+04	1.8887E+04
62	198.24	2.8481E+04	7.1562E+04	2.8481E+04
63	182.58	2.6248E+04	7.3349E+04	2.6248E+04
64	168.12	2.1128E+04	7.2634E+04	2.1128E+04
65	154.81	2.4079E+04	8.0402E+04	2.4079E+04
66	142.52	5.6184E+04	1.1779E+05	5.6184E+04
67	131.25	5.7932E+04	1.2529E+05	5.7932E+04
68	120.89	7.0977E+04	1.4461E+05	7.0977E+04
69	111.32	9.5733E+04	1.7626E+05	9.5733E+04
70	102.52	1.1228E+05	2.0032E+05	1.1228E+05
71	94.390	1.1447E+05	2.1076E+05	1.1447E+05
72	86.921	1.6799E+05	2.7328E+05	1.6799E+05
73	80.054	1.8565E+05	3.0076E+05	1.8565E+05
74	73.669	2.2386E+05	3.4982E+05	2.2386E+05
75	67.886	2.6246E+05	4.0010E+05	2.6246E+05
76	62.465	3.3024E+05	4.8087E+05	3.3024E+05
77	57.526	3.6679E+05	5.3148E+05	3.6679E+05

Step	Energy [eV]	Differential Number Flux [cnts/(cm^2 s sr eV)]	Uncorrected Number Flux [cnts/(cm^2 s sr eV)]	Corrected Number Flux (Check)
78	53.008	4.5061E+05	6.3057E+05	4.5061E+05
79	48.791	4.3846E+05	6.3535E+05	4.3846E+05
80	44.936	5.5387E+05	7.6912E+05	5.5387E+05
81	41.382	5.7303E+05	8.0838E+05	5.7302E+05
82	38.069	6.7631E+05	9.3394E+05	6.7631E+05
83	35.057	7.6144E+05	1.0432E+06	7.6144E+05
84	32.287	8.8315E+05	1.1912E+06	8.8315E+05
85	29.757	9.2115E+05	1.2576E+06	9.2114E+05
86	27.407	9.5292E+05	1.3208E+06	9.5292E+05
87	25.239	1.0002E+06	1.4025E+06	1.0002E+06
88	23.191	1.0505E+06	1.4914E+06	1.0505E+06
89	21.384	1.1846E+06	1.6660E+06	1.1846E+06
90	19.697	1.0946E+06	1.6208E+06	1.0946E+06
91	18.131	8.9857E+05	1.4743E+06	8.9857E+05
92	16.685	1.1141E+06	1.7440E+06	1.1141E+06
93	15.360	1.1828E+06	1.8719E+06	1.1828E+06
94	14.155	8.8199E+05	1.6348E+06	8.8199E+05
95	13.011	9.5924E+05	1.7841E+06	9.5924E+05
96	11.987	1.1810E+06	2.0825E+06	1.1810E+06
97	11.023	1.3958E+06	2.3830E+06	1.3958E+06
98	10.180	1.5122E+06	2.5884E+06	1.5122E+06
99	9.337	1.6813E+06	2.8632E+06	1.6813E+06
100	8.614	1.9351E+06	3.2250E+06	1.9352E+06
101	7.951	1.9280E+06	3.3347E+06	1.9280E+06
102	7.289	2.6670E+06	4.2129E+06	2.6670E+06
103	6.746	2.6674E+06	4.3484E+06	2.6674E+06
104	6.204	2.4113E+06	4.2520E+06	2.4113E+06
105	5.722	2.9625E+06	4.9718E+06	2.9625E+06
106	5.240	3.7752E+06	5.9855E+06	3.7752E+06
107	4.819	3.4643E+06	5.8850E+06	3.4643E+06
108	4.457	4.3167E+06	6.9510E+06	4.3167E+06

Step	Energy [eV]	Differential Number Flux [cnts/(cm^2 s sr eV)]	Uncorrected Number Flux [cnts/(cm^2 s sr eV)]	Corrected Number Flux (Check)
109	4.096	3.3074E+06	6.1944E+06	3.3074E+06
110	3.735	4.6219E+06	7.8130E+06	4.6219E+06
111	3.434	4.2162E+06	7.7118E+06	4.2162E+06
112	3.193	2.5987E+06	6.3811E+06	2.5987E+06
113	2.952	2.8294E+06	6.9476E+06	2.8294E+06
114	2.711	2.7123E+06	7.2286E+06	2.7123E+06
115	2.470	1.7900E+06	6.7858E+06	1.7900E+06
116	2.289	3.0700E+06	8.4947E+06	3.0700E+06
117	2.108	2.0222E+06	7.9526E+06	2.0222E+06
118	1.928	1.6065E+06	8.1418E+06	1.6065E+06
119	1.747	1.7245E+06	8.9955E+06	1.7245E+06
120	1.626	3.7665E+06	1.1623E+07	3.7665E+06
121	1.506	2.2471E+06	1.0787E+07	2.2471E+06
122	1.385	1.4083E+06	1.0756E+07	1.4083E+06
123	1.265	1.3757E+06	1.1692E+07	1.3757E+06
124	1.145	1.6327E+06	1.3130E+07	1.6327E+06
125	1.084	1.2039E+06	1.3395E+07	1.2039E+06
126	0.964	-1.3884E+06	1.2463E+07	-1.3884E+06