Background analysis for IMA sensor on ASPERA instrument

Fernando J. Capalbo

IRF Technical Report 054
February 2010

ISSN 0284-1738

INSTITUTET FÖR RYMDFYSIK

Swedish Institute of Space Physics

Kiruna, Sweden

Background Analysis for IMA sensor on ASPERA instrument

Fernando J. Capalbo

Lulea University of Technology, SpaceMaster program
and Swedish Institute of Space Physics, Kiruna

fernando.capalbo@gmail.com

February 2010
Version 3.1

Fernando J. Capalbo
IRF Technical Report 054
ISSN: 0284-1738
Swedish Institute of Space Physics
P.O. Box 812
SE-981 28 Kiruna
SWEDEN
fernando.capalbo@gmail.com

Printed in Sweden by Swedish Institute of Space Physics, Kiruna 02/2010.

Contents

Abstract

Data from the IMA-ASPERA (Ion Mass Analyser - Analyser of Space Plasmas and Energetic Atoms) sensor on board Mars Express and Venus Express is a very important contribution to interplanetary, Venusian and Martian atmosphere studies. In addition to hot plasma measurements in the Martian and Venusian environments, a background channel of this sensor could be used to monitor high energetic particles events. This would serve as a proxy of Solar Energetic Particles events and therefore contribute to space weather surveillance. In this work, data obtained by Mars Express and Venus Express IMA sensor is analysed and an appropriate background channel is determined for both instruments. The channel is a subset of an energy-mass matrix obtained from the data. A software for background calculation is explained and was implemented to serve as part of the ESA (European Space Agency) ASPERA data processing chain. Time series of background values calculated with data in the Mars Express and Venus Express database showed the feasibility of using this product as a proxy for space weather monitoring.

1 Purpose of this document

This document describes the procedure to determine the background caused by penetrating radiation for the measurements from the Ion Mass Analizer (IMA) detector of the Analyser of Space Plasmas and Energetic Atoms (ASPERA) instrument, on board satellites Mars Express and Venus Express. It includes description of the preliminary data analysis and results, determination of the background channel and calculations of the background values time series. The software developed and its implementation are described. This document can also be used as a user manual for the software, to be implemented at the European Space Agency (ESA) pipeline for the ASPERA background data processing.

2 Introduction

Satellites Mars Express (launched in 2003) and Venus Express (launched in 2005) have elliptic orbits around Mars and Venus. The instruments ASPERA-3 and ASPERA-4 are part of the payload of Mars Express and Venus Express respectively. Their scientific objectives are to study the interaction between the solar wind and the planets' atmosphere (Futaana et al., 2008). Both instruments were built by the Swedish Institute of Space Physics (IRF). The ASPERA instruments are identical and consist of two units (Barabash et al. (2007); Futaana et al. (2008)): the IMA sensor and the Main Unit (MU), composed in turn of three sensors: Neutral Particle Detector (NPD), Neutral Particle Imager (NPI) and the Electron Spectrometer (ELS). The IMA is the object of the present study.

Analysis of the background data caused by penetrating radiation for the IMA sensor is very useful, not only to evaluate fluxes of energetic particles but also to have an idea of the sensor health and degradation. Moreover, the increase of the count rate in the background channel could be applicable in the detection of high energy particles, possibly related to Solar Energetic Particle (SEP) events (Futaana et al., 2008). Although some background values were used in the previous paper (Futaana et al., 2008), before this work there were no systematic studies of background available.

2.1 The IMA sensor

The IMA sensor measures ion properties with energy, mass and angular resolution. As an example, IMA on ASPERA-3 provides ion measurements in the energy range 0.01 - 36 keV/q for the main ion components H^+, H_2^+ , He^+, O^+, and for the group of molecular ions $20 < M/q < \sim 80$. The IMA field of view is 90°(polar) x 360°(azimuth).

The IMA sensor consists of four main components (Barabash et al., 2007) (see Figure 1): electrostatic deflection system to provide polar (or elevation) sweep, top-hat electrostatic energy analyser, permanent magnet based velocity analyser, and a Micro Channel Plate (MCP) detector with a position sensitive anode. The energy range is swept over E=96 steps. The mass of the particles is resolved in M=32 rings in the MCP detector. The angles of incidence of the particles are determined by Z=16 azimuthal sectors in the detector and P=16 polar angle steps (refer to (Barabash et al., 2007) and Figure 1 for detector and geometry definitions). The sampling time of each energy step is 125 ms. In T_m=192 seconds the instrument performs a complete measurement cycle and provides a data set consisting of 32 mass rings x 16 azimuthal sectors x 96 energy steps x 16 polar angles. By on board processing this set can be re-binned according to the mode of operation. For example, one mode can result in data with E=96, M=32, Z=16, P=8, where two polar angle steps are summed up. A measurement cycle results in a 4-D array, consisting in particle counts per 125 ms sampling time. In this report the elements of this array are called bins. E.g. the bin (energy,mass)=(96,32) with value "5", means that there are 5 particle counts for that mass and energy. In this example azimuth and polar dimensions were not specified and usually this will denote that all azimuth and polar angles were taken into account, summing over them. This information is stored with 16-bit resolution.

As a clarifying example, Table 1 shows these values for a typical operation mode of IMA on ASPERA-3, on board Mars Express.

It is worth noting that the indices for the measurement steps start with 0, however, they start with 1 in the software developed (and its documentation) due to the 1-base matrix indexing of the programming language used.

In this work, by analysing and filtering ASPERA-IMA data from the beginning of the Mars Express and Venus Express missions, a background channel is determined. After this a background extraction software is implemented. The software is tested on the existing ASPERA-IMA data and prepared to run in the ESA ASPERA production chain. Section 3 briefly describes the

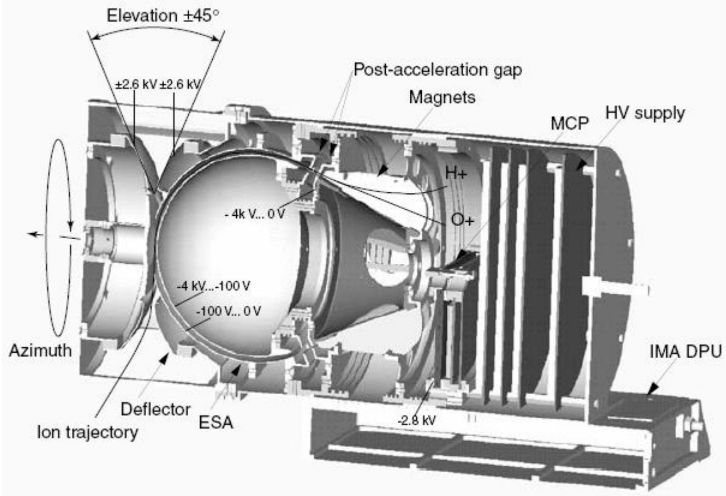

Figure 1: The IMA instrument cross-section. The elevation angle in the figure is referred to as polar angle in the text. From Barabash et al. (2007).

Parameter	Steps	Physical Range
E	0-95	36,000 eV - 10 eV
M	0-31	1, 2, 4, 8 amu/q
P	0-15	-45°- +45°
Z	0-15	0-360 °
T_m		192 s

Table 1: Parameter values for a typical operation mode of Mars Express Aspera-3 IMA sensor.

sensor operations and the data format. Section 4 describes the data set used for analysis and how this data were handled. Section 5 explains the background channel determination process. In Section 6 an outline of the software developed is presented. Its implementation is explained in Section 7 and the full documentation and source code are referenced in appendix B. Conclusions and further work are in Section 8. This work was performed during the 6-week period as part of a summer job at IRF.

3 IMA data

3.1 Data interpretation

There are several ways to treat IMA data. Energy-mass and energy-time plots are common in studies using this type of data. Examples of energy-mass plots can be seen below. Figures 2 and 3 show two measurements each, for two extreme positions in the orbit of Mars Express and Venus Express respectively: outside the magnetosheath (left), where detected particles are mainly from the solar wind, and close to the surface (right), where detected particles are mainly from the planet's ionosphere. The data are in energy-mass-matrix form, summing over all azimuth and polar angles. The black curves are energy-mass contour lines, from the calibration measurements, for the same post acceleration voltage, shown as a reference.

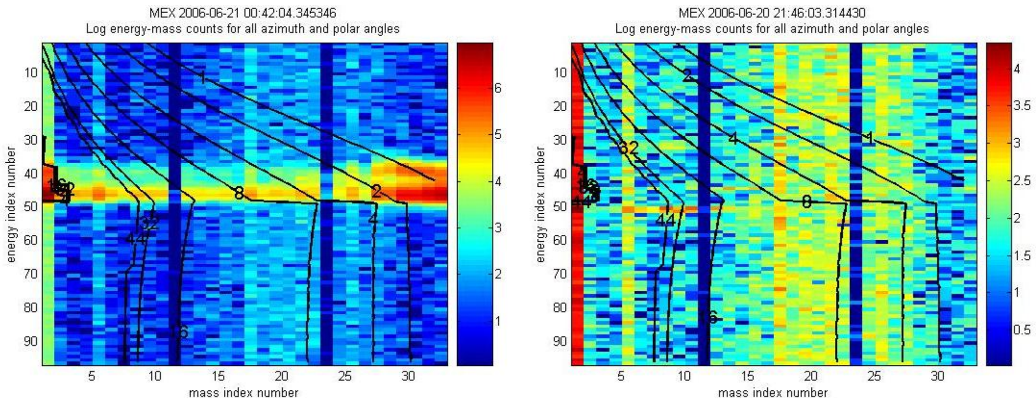

Figure 2: Mars Express data in energy-mass form, for all azimuth and polar angles. Data from solar wind measurements (left) and close to the surface (right). Black numbered lines are mass lines for different m/q = 1:H^+, 2:He^{++} or H_2^+, 4:He^+, 8:O^{++}, 16:O^+, 32:O_2^+, 44:CO_2^+. Orbit 3140 in 2006-06-20/21.

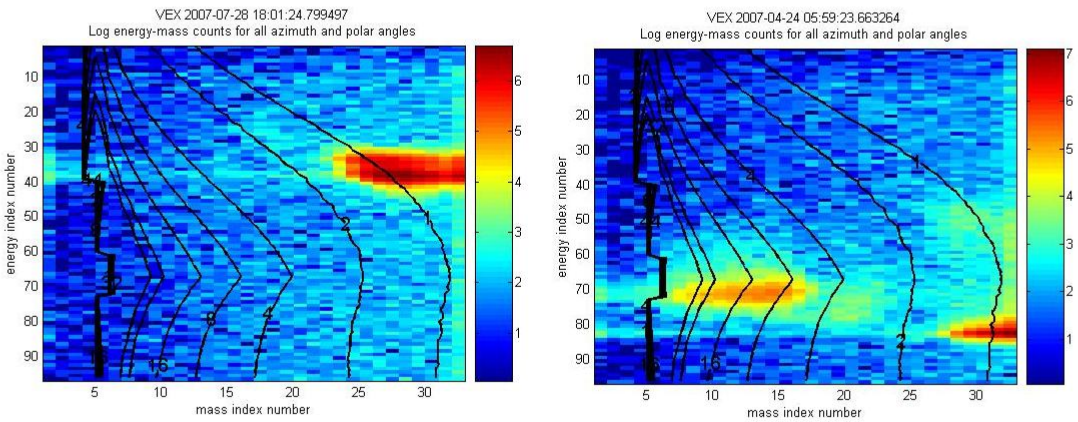

Figure 3: Venus Express data in energy-mass form, for all azimuth and polar angles. Data from solar wind measurements (left) and close to the surface (right). Black numbered lines are mass lines for different m/q = 1:H^+, 2:He^{++} or H_2^+, 4:He^+, 8:O^{++}, 16:O^+, 32:O_2^+, 44:CO_2^+. Orbit 463, year 2007-07-27/28.

Figures 4 and 5 show the data of the same orbit as Figures 2 and 3 in energy-time form, summing over all masses, azimuthal and polar angles. The energy index number was converted to energy units and measurements cycles for different consecutive times are shown. Data are from the same extreme orbital positions as in Figures 2 and 3.

In this work the data were analysed in the form of energy-mass matrices, where the values of the matrix represent the particle counts for a certain energy-mass bin.

There are also energy tables and post acceleration factors to be considered (Barabash et al., 2007), and that would add extra dimensions. For example, in the Mars Express case, the energy

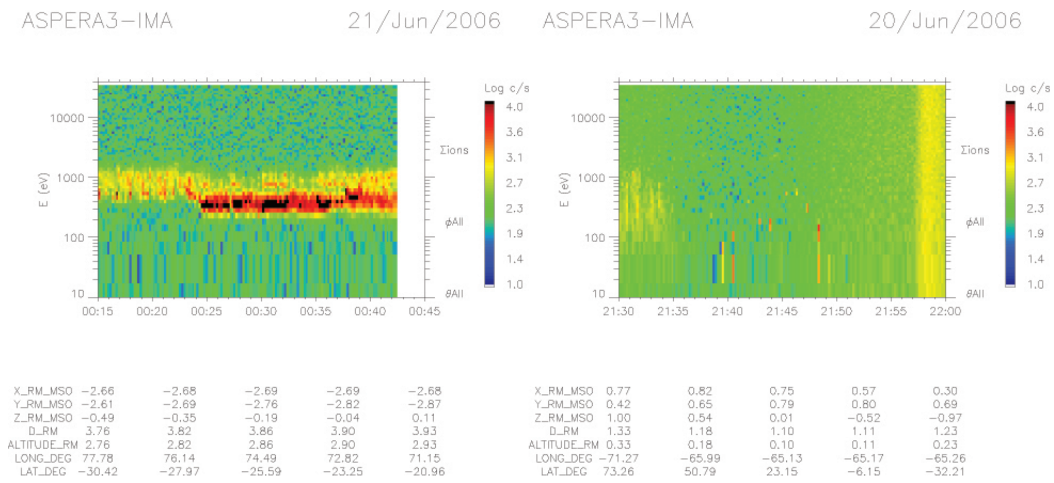

Figure 4: Mars Express data in energy-time form, for all masses, azimuth and polar angles. Data from solar wind measurements (left) and close to the surface (right). Orbit 3140 in 2006-06-20/21.

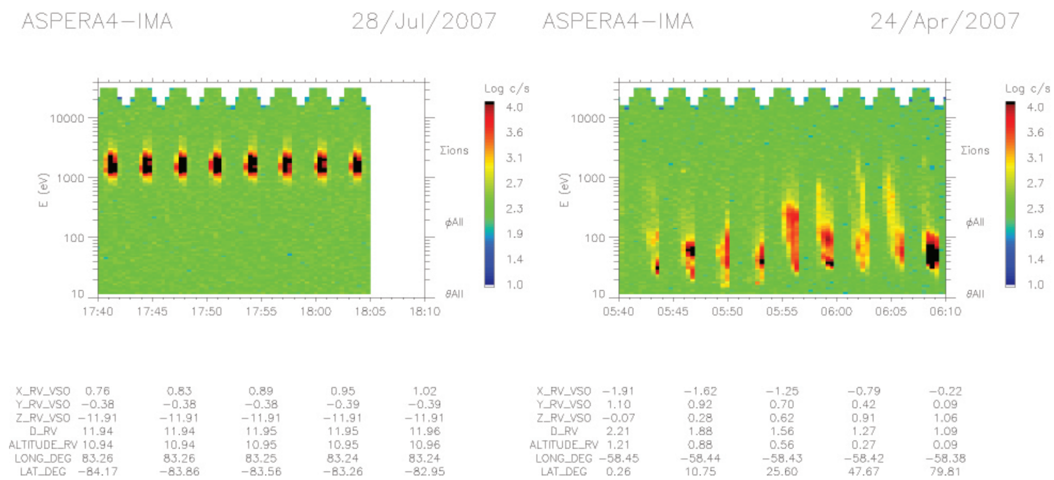

Figure 5: Venus Express data in energy-time form, for all masses, azimuth and polar angles. Data from solar wind measurements (left) and close to the surface (right). Orbit 463, year 2007-07-27/28.

table used to measure the ions changed in April 2007. In this study these dimensions were not considered. By using unit-less indexes for energy and mass and not physical units such as amu or eV, all post acceleration and energy re-scaling (tables) are included in the analysis without discrimination. The determination of the background channel is therefore independent on these dimensions and applicable for any existing energy table and post acceleration setting in June 2009.

The raw data from the satellite are processed into intermediate text files, which are the data files handled in this work. Brief comments on these intermediate data files follow to allow comprehension of the next sections and the software.

3.2 Data format

The data files are named "E_IMyyyymmddnnnnnn.txt' where yyyymmdd stands for the date in year-month-day format and nnnnnn is the start time of operation. Each ASCII file contains a set of N measurement cycles of 192 seconds each corresponding to a certain time period. This time period depends on instrument operation and therefore a file can contain measurement cycles from different days. In this report data from one file will be referred to as data from the day in the files name.

There are two format versions for these text files: v4 and v5. This is due to a time tag conversion problem with the v4 data file format. On the other hand, v5 has a disadvantage to v4 on its performance of conversion from raw data. It is decided to use v5 for the released software because the accuracy of time-tag is mandatory for the future scientific analysis. All analysis and comments in subsequent sections refer to the v5 data files unless otherwise stated. However, there is a possibility to fix the time tag problem in v4 in the future, thus, the software was made compatible with both versions of the file format, but was optimized for v5, which is default. Both formats are briefly described below.

In v4 format, each measurement record consists of a header and corresponding data. The header includes (but is not limited to) date, time of the beginning of the scan and the number of steps (binning) corresponding to each of the dimensions commented above (e.g. E=96, M=32, A=16, P=16). The data are stored in the matrix form, with Z columns corresponding to the azimuth angles and the other dimensions folded in M x E x P rows. The first dimension to vary in the rows is the mass, followed by the energy and then the polar angle. Thus, the mass index increases in every row, the energy index increases after M rows and the polar index after M x E rows. The indexes for the rows in one measurement could be schematized as shown in Table 2.

rows 1-3072	rows 3073-46080	rows 46081-49153
M, E, P	M, E, P	M, E, P
0, 0, 0	0, 0, 1	0, 0, 15
...	...	...
31, 0, 0	31, 0, 1	31, 0, 15
0, 1, 0	0, 1, 1	0, 1, 15
...	...	...
31, 1, 0	31, 1, 1	31, 1, 15
...	...	...
...	...	...
0, 95, 0	0, 95, 1	0, 95, 15
...	...	...
31, 95, 0	31, 95, 1	31, 95, 15

Table 2: M, E, P indexes for the lines in the version 4 format ASCII files.

As an example, the first 10 lines of the v4 file 'E_IM20050212070650.txt' can be read as follows:

```
-----------------------------------------------------------------------
is:173548 ius:075000 s:56272780 fs:35061 2005-02-12 07:19:46 bhv:1 sh:1 r:0 pa:1 ml:0 de:0 m:24
UT:173548 TIME:0173548375 MODE:Exm-0 INDEX:24 M:32 E:96 A:16 P:16
-----------------------------------------------------------------------
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

The v5 format is based on Planetary Data System (PDS) format of the ASPERA-3 archive. v5 is a comma-separated-values format, which includes (but is not limited to) date, time and azimuth, mass and polar angle indexes. The data that follows this information correspond to E energy steps (e.g. E=96 in the nominal case). The other dimensions are folded in the following Z x M x P lines. In this case, the first dimension to vary in the rows is the azimuth, followed by the mass and then the polar angle. Thus, the azimuth index increases in every line, the mass index increases after Z lines and the polar index after Z x M lines. The indexes for the lines in one measurement could be schematized as shown in Table 3.

rows 1-512	rows 513-7680	rows 7681-8192
Z, M, P	Z, M, P	Z, M, P
0, 0, 0	0, 0, 1	0, 0, 15
...	...	...
15, 0, 0	15, 0, 1	15, 0, 15
0, 1, 0	0, 1, 1	0, 1, 15
...	...	...
15, 1, 0	15, 1, 1	15, 1, 15
...	...	...
...	...	...
0, 31, 0	0, 31, 1	0, 31, 15
...	...	...
15, 31, 0	15, 31, 1	15, 31, 15

Table 3: Z, M, P indexes for the lines in the version 5 format ASCII files.

As an example, the first 10 lines of the v5 file 'E_IM20050212070650.txt' can be read as follows:

2005-02-12T07:12:49.691761,1,1,0,4,0,00,00,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,01,01,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,02,02,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,03,03,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,04,04,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,05,05,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,06,06,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,07,07,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,08,08,00,00,00,00,00000,00000,00000,...,00000
2005-02-12T07:12:49.691761,1,1,0,4,0,09,09,00,00,00,00,00000,00000,00000,...,00000

In this sample '...' was used to replace real data for reasons of limited space, but 96 groups of 5 counting data must be understood in every line.

4 Sample data and data handling

4.1 Data subsets

The background channel analysis was performed using a subset of data, considered to be representative of typical data from the IMA sensor. This sample data consists of 19 files, covering the period from January 2004 till December 2007 for Mars Express and 18 files, covering May 2006 till April 2009 for Venus Express. Figures 6 and 7 show the sum of counts for all the measurements in the sample files for Mars Express and Venus Express respectively. The data is in energy-mass-matrix form, summing over all azimuth and polar angles.

Appendix A describes the files used.

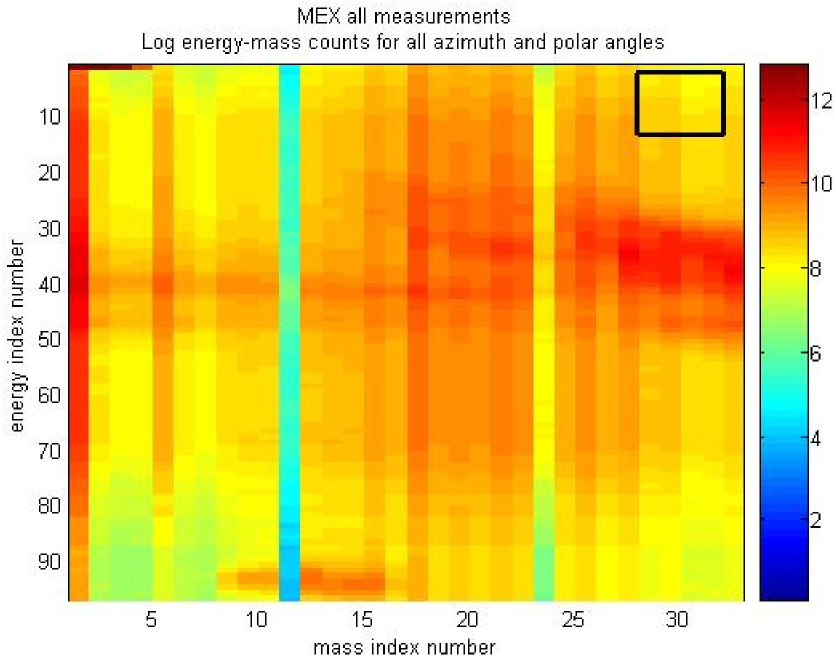

Figure 6: Mars Express data in energy-mass form, for all azimuth and polar angles. Data from all non-saturated measurements are included. The rectangular in the upper right corner represents the final background channel.

4.2 Data handling and usage

The texts files are processed in different ways depending on the use of the data. For the background analysis part of this work, data in the text files was first saved in MATLAB ".mat" format for fast access. For the background calculation the text files ".txt" are used. Figure 8 shows the data flow.

Some of the measurements were saturated for certain mass and energy values (see Figure 9 for an example of the saturated values). If these saturated measurements were in the mass and energy ranges from where statistical calculations are done they would bias the result, making the whole set of data meaningless for the interpretation. These measurements were identified and the corresponding whole set of 192 s was eliminated from the background analysis. The criterion used for the filter was to exclude measurements with more than a certain number of saturated bins, i.e., elements of the data array with values 65536. Due to the data characteristics (see section 3), if more than E x M values were saturated this means that there could be a saturated value for all the mass-energy combinations for at least one polar angle and one azimuth angle, spoiling the values of the whole energy-mass matrix. This would be a very unlikely case, as usually the saturation values are contiguous. However, even in the case of contiguous values, E x M saturated bins would imply saturation for all energy values for some masses, azimuths and one polar angle, which would also spoil the analysis for those mass columns in the energy-mass

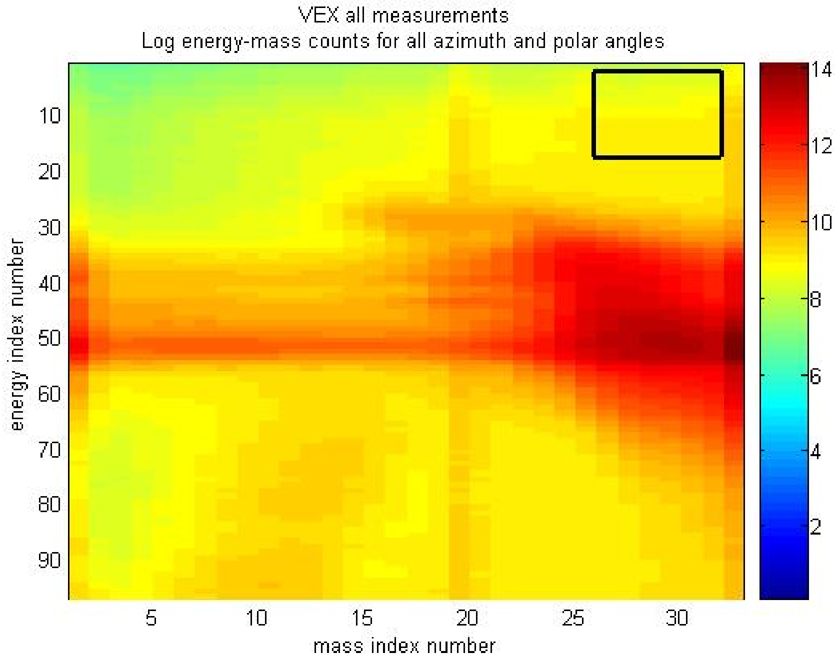

Figure 7: Venus Express data in energy-mass form, for all azimuth and polar angles. Data from all non-saturated measurements are included. The rectangular in the upper right corner represents the final background channel.

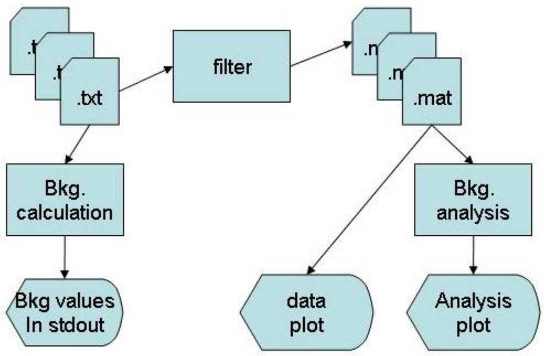

Figure 8: Data flow. Background analysis (right branch) and background calculation (left branch).

matrix. Other threshold values were considered for the filter. The value E x M gave acceptable visual results (see Figures 6 and 7) with almost no saturated bins and this is the value used in the final software version, i.e., if more than 96 x 32 values are saturated, the 192 s measurement is excluded from the analysis.

In case saturated values were observed in some filtered measurement, the whole file containing that measurement (and not only the corresponding 192 s set) was excluded from the background analysis. See section 6 for the detailed software implementation of the filter.

All the measurements will be used for the background values calculation in the final product.

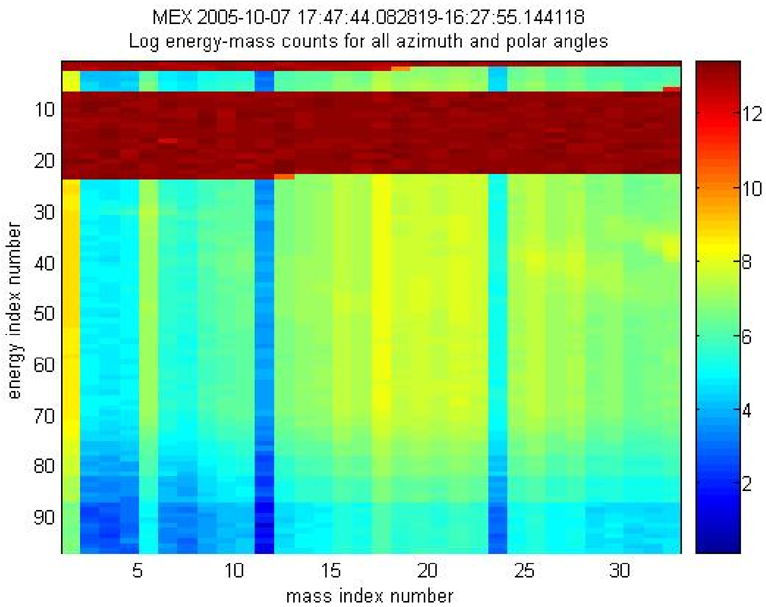

Figure 9: Mars Express data in energy-mass form, for all azimuth and polar angles. Data from all (including saturated) measurements for the date shown in the graph. Dark red regions are saturated bins.

5 Background channel

5.1 Analysis

The first step is to determine the background channel. Counts recorded by the sensor can be divided into two categories. Nominal counts, due to particles from the solar wind or the planetary ionosphere, are within the instrument ranges and are expected to be measured (Barabash et al., 2007). The other category are counts due to background. The background counts are caused by high energy particles that penetrate the sensor mechanical structure, reaching the detector, and by white noise (e.g. dark counts from the MCP). Nominal counts are restricted to be present at certain mass and energy indexes, depending on the mass, charge and energy of the incoming particle (around the M/q contour lines superimposed in Figures 2 and 3). Thus, there is an area in the energy-mass matrix were no nominal counts are expected. This can be confirmed by visual inspection of the sample data. Considering the energy-mass matrix, with the bin (0,0) in the upper left corner, the subset with no expected nominal counts is in the upper right quarter (as seen in Figures 6 and 7), and can be considered as a background channel.

The background channel should have certain characteristics. Regarding the particular case of the IMA sensor, and considering average measurement conditions and normal operation , the background counts should be much smaller than and independent on nominal counts. Little variation along time is desired.

A background channel was used previously for this instrument (Futaana et al., 2008), however, this was a preliminary approach that included some bins with possible nominal counts, therefore a different approach is followed here.

5.2 Determination and results

Several options were considered to determine the appropriate background channel. The first step being the same for all of them. Energy-mass bins in the region commented above for all the measurements in the sample data were considered and its quality as background channel was analysed. This was repeated for different region sizes and shapes. For simplicity, all the regions considered were rectangular, with N_E energy rows x N_M mass columns. N_E varied from 2 to 37 and N_M from 2 to 17, in steps of 5 bins. This values are arbitrary. All the regions had limits in the smallest energy index (corresponding to the highest energy) and biggest mass index (corresponding to the smallest mass). In practice, the border energy and mass values were excluded due to instrumental effects (see last column and first row on the Figures 6 and 7). No theoretical derivations were made for the background channel estimation, an experimental approach was used to determine the quality of the region under consideration. Three methods were implemented for this. In all the methods, the background value is calculated as the mean of all the values in the background channel area. Only 8 data files and up to 12 channel sizes were used for clarity of the plots shown here, but the analysis leading to the results commented in this and subsequent sections was performed with all the data in the sample files (measurements with saturated values were filtered out from the analysis as explained in section 4) with the channel size varied as commented above. All results for Venus Express and Mars Express are shown.

The first method compares the background counts with the total counts in the upper right quarter of the energy-mass matrix for each measurement. As the background is expected to be independent on the total counts (which include nominal counts), an increase in the background counts with the total counts means that the background channel is overlapping the data channels, and the shape and/or size of the channel is not appropriate. Results for this method are shown in Figure 10 for Mars Express and Figure 11 for Venus Express, where background counts are plotted against the total counts for different channel shapes and sizes. Total counts are the sum of the values of the elements in the upper right quarter of the energy-mass matrix.

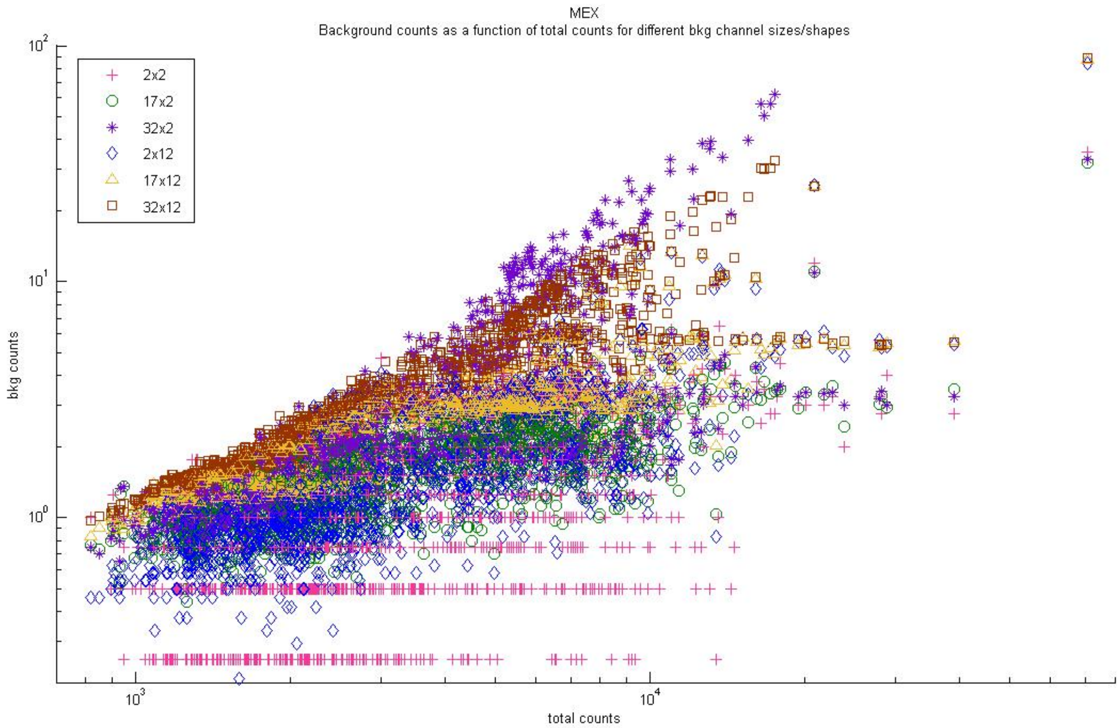

Figure 10: Background dependence on total counts. Background channel sizes are expressed as Energy bins x Mass bins. Measurements from 8 data files and 6 background sizes are shown. Mars Express data.

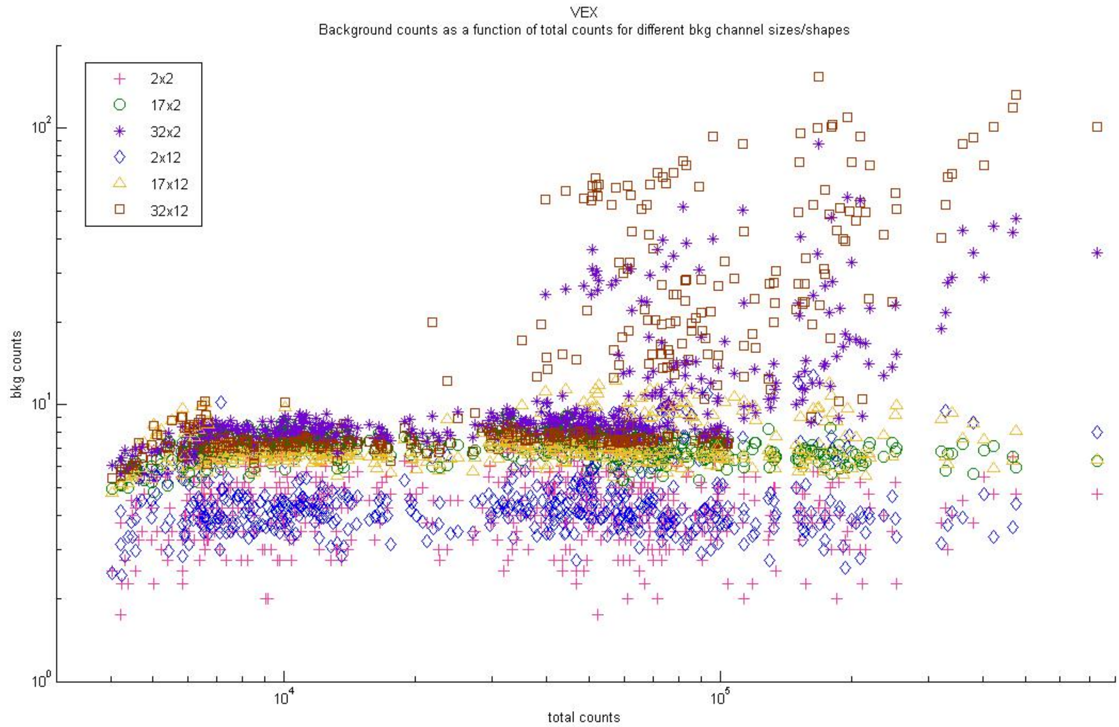

Figure 11: Background dependence on total counts. Background channel sizes are expressed as Energy bins x Mass bins. Measurements from 8 data files and 6 background sizes are shown. Venus Express data.

The criteria for the selection of appropriate background channels is as follows:

_ The channel sizes with a strong background dependence on total counts will not represent the background and are not background channels.
_ The sizes showing a strong dependence on time for constant total counts should also be excluded, as the background is expected to be independent on time over the time periods considered

for most of the measurements analysed.

Note that smaller channels are not necessarily better, as there is also a dependence on channel shape. The background increases towards the smaller masses (higher mass indexes) due to the fact that the physical sectors for detecting those particles are bigger, implying more counts for a uniform background flux. That makes the background counts sometimes higher than the mean corresponding to the upper right quarter in the energy-mass matrix (remember the background is calculated as the mean value of the bins contained in the background channel). This effect is also responsible for the fact that bigger background channels can result in lower background values than smaller channels , because the bigger ones include physically smaller mass channels, with less counts for the same flux.

Zooming in particular sections of this king of graphs, manually identifying and eliminating one by one the channel sizes that did not met the above criteria, it was determined that sizes of up to 22x12 bins would be appropriate for a background channel for Venus Express.

For Mars Express however, it was not possible to find an appropriate channel size. Even the 2x2 channel showed some dependence on the total counts.

The second method uses the same property of the background channel but a different analysis technique. The ratio between the background counts and the mean in the upper right quarter of the energy-mass matrix is considered for each measurement in the sample data. The background is expected to be much smaller than the quarter mean. An increase in the ratio with increasing background channel size means it is overlapping with data channels. However, this technique considers a uniform background, and this does not reflect the reality according to the above discussion regarding the increase in counts for bigger masses in a uniform background flux. Therefore, background/mean ratios bigger than one are possible.

This method was implemented anyway, assuming that channel sizes with ratios much bigger than one would be overlapping data channels and would not be appropriate.

Results for this method are shown in Figure 12 for Mars Express and Figure 13 for Venus Express, where ratios background/mean for a set of measurements in the sample data are plotted against background channel size. Ratios bigger than one are also shown in the bottom part of the figure, expressed in percentage of total ratios for each channel size.

The criteria for the selection of appropriate background channels is as follows:

_ Those channel sizes with a big ratio are not good background channels.

Those sizes that are left are candidates for the background channel. In this case the background shape dependence of the analysis is more evident.

Although for this analysis only the percentage of ratios smaller than one was used as selection criteria, all the data are presented for clarity of the procedure and future reference.

Looking at the percentage of ratios bigger than one in this kind of graphs, a size of 17x17 bins would be an appropriate background channel for Venus Express.

For Mars Express however, it was not possible to find an appropriate channel size, as most of the sizes had ratios between 1 and 1.5. Smaller and bigger sizes had the biggest ratios, suggesting that maybe more than the first energy channel should be left out for the calculations.

The third and last method is much simpler, it plots the value of the background channel as a function of the channel size and shape. In this case the background is computed over an energy-mass matrix resulting from the addition of all the values in the sample data files (In this case all the data are shown, not only the subset of 8 files as in the two previous methods).

Results for this method are shown in Figure 14 and Figure 15 for Mars Express and Venus Express respectively, where the background counts are plotted against background channel size.

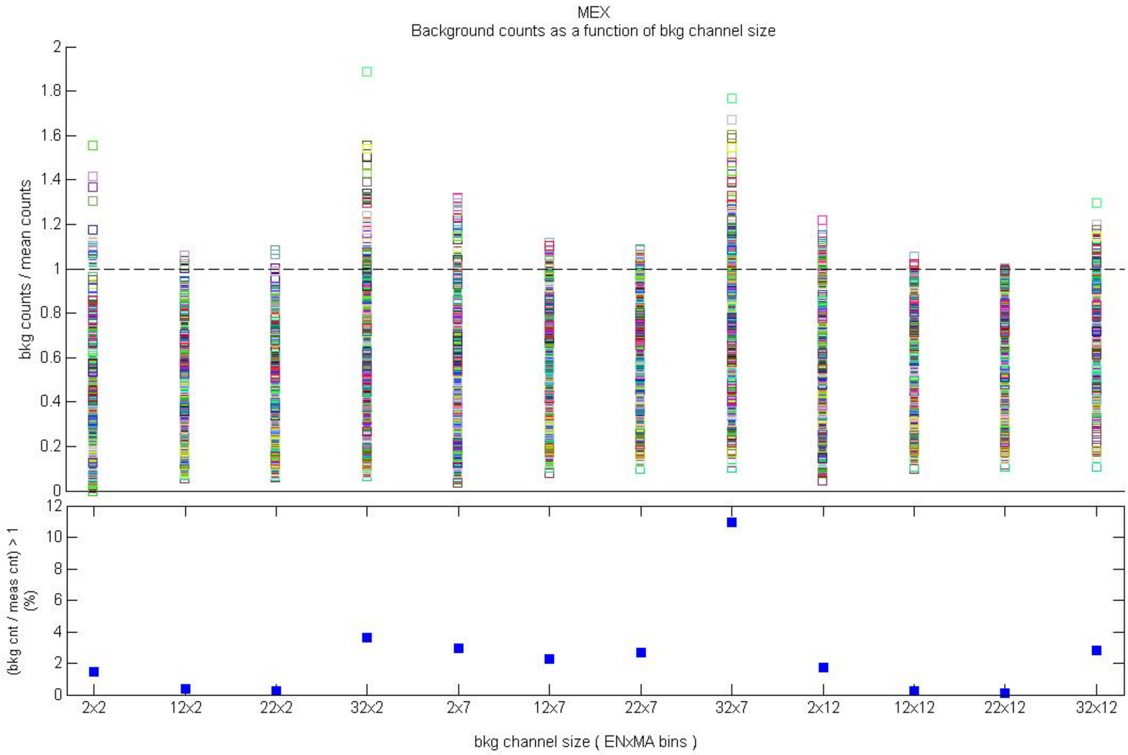

Figure 12: Background/total ratio dependence on the background channel size(top). Percentage of ratios bigger than 1 (bottom). Background channel sizes are expressed as Energy bins x Mass bins. Measurements from 8 data files and 12 background sizes are shown. Mars Express data.

The criteria for the selection of appropriate background channels is as follows:

- Those channel sizes with a big background value compared to the others are supposed to overlap data channels and are not good background channels.

Those sizes that are left are candidates for the background channel.

From the results, sizes of up to 22x17 bins would be appropriate for a background channel for Venus Express. It is clear from Figure 15, and interesting to note, that the background counts increase with increasing energy bins in the background channel, but there is almost no variation with increasing mass bins for the same energy bins. For sizes over 32 energy bins, there is a dramatical increase in the background counts. For Mars Express (Figure 14) however, this dependence appears in both energy and mass, and the background counts increase every time the dimension increases. Therefore it was not possible to find an appropriate channel size in that case.

Besides these analyses, a visual analysis of the sample data helped to restrict the size and shape of the background channel, especially for the Mars Express case, where the methods above did not prove to be useful for this. Figures 6 and 7, as well as plots for individual dates and times were used for this purpose. The visual analysis allowed to exclude channels which presented anomalies in several measurements, like mass channel 23 for Mars Express and 25 for Venus Express.

One criterion used to determine the background channel was to consider regions in the upper right of the energy-mass matrix including all the measurements where the counts where about 3 orders of magnitude smaller than in regions with nominal counts.

Based on all the preceding analysis, and by the time of writing, a background channel size

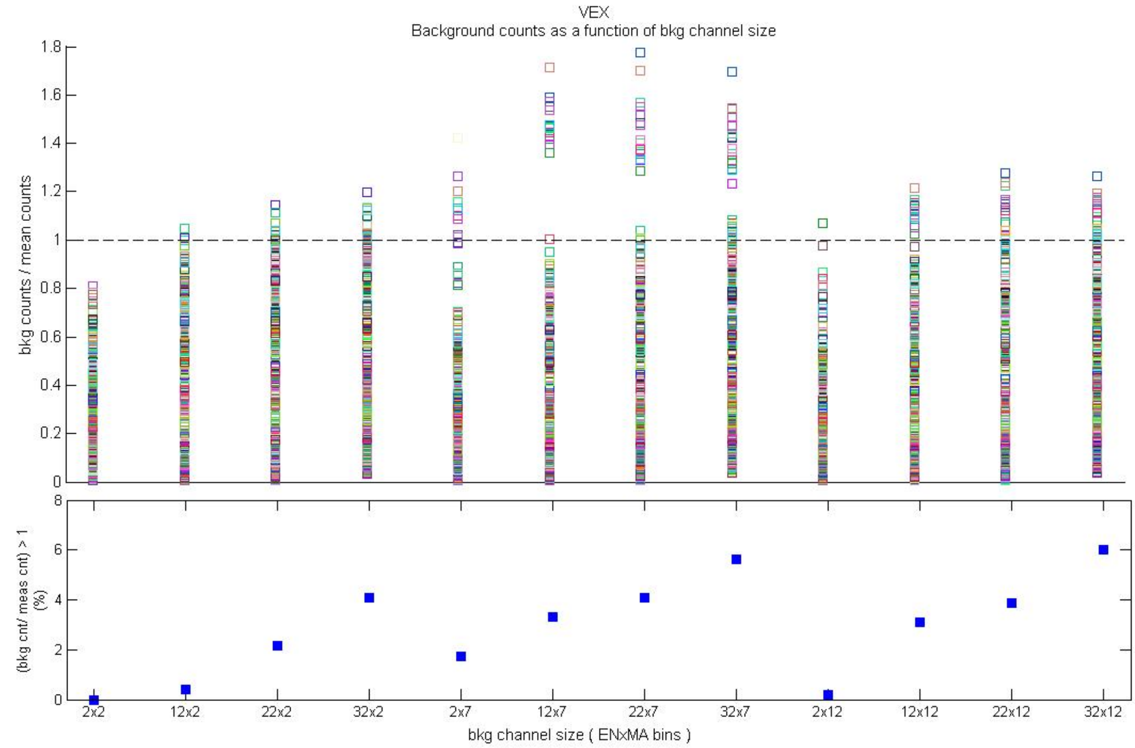

Figure 13: Background/total ratio dependence on the background channel size(top). Percentage of ratios bigger than 1 (bottom). Background channel sizes are expressed as Energy bins x Mass bins. Measurements from 8 data files and 12 background sizes are shown. Venus Express data.

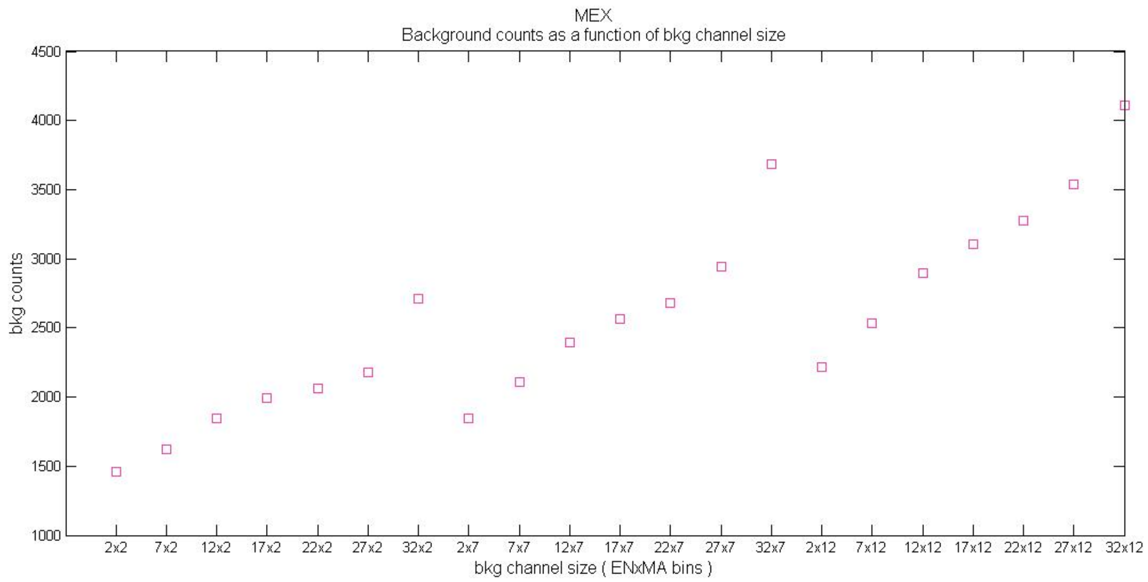

Figure 14: Background counts dependence on the background channel size. All data files and 32 background channel sizes are shown. Sizes increase in steps of 5 bins in each dimension. Some tick labels were excluded for clarity. The sizes are EnergyBins x MassBins. Mars Express data.

of 14x6 bins and 11x4 bins have been chosen for Venus Express and Mars Express respectively. It is worth mentioning that this is not the optimum channel regarding shape or size, which determination would take a more detailed analysis and time than the time available for this project. However, the results obtained are enough to provide an effective and safe background channel for the Mars Express and Venus Express missions.

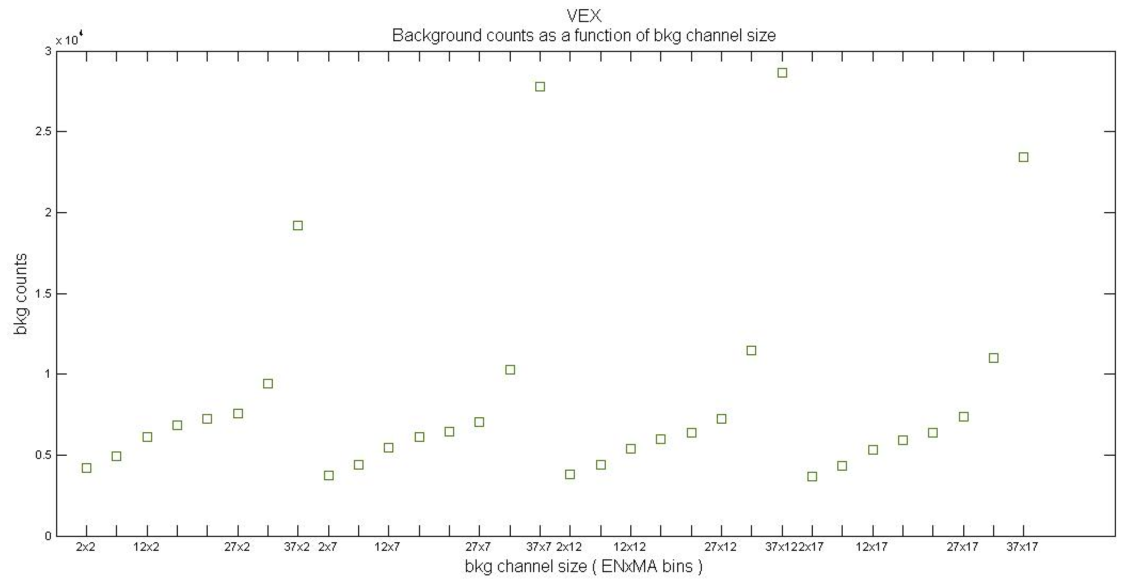

Figure 15: Background counts dependence on the background channel size. All data files and 32 background channel sizes are shown. Sizes increase in steps of 5 bins in each dimension. Some tick labels were excluded for clarity. The sizes are EnergyBins x MassBins. Venus Express data.

5.3 Time series

With the background channels specified, background values were calculated for all the sample data files and plotted as a time series. Figures 16 and 17 show background values and total counts from the whole energy-mass matrix as a function of time for Mars Express and Venus Express respectively. In the time axis, date labels show the beginning of the data collected that day, with different colours for different days and separated by vertical lines. Only some dates are shown as reference. The data plotted is chronologically ordered but not consecutive, because so were the sample files.

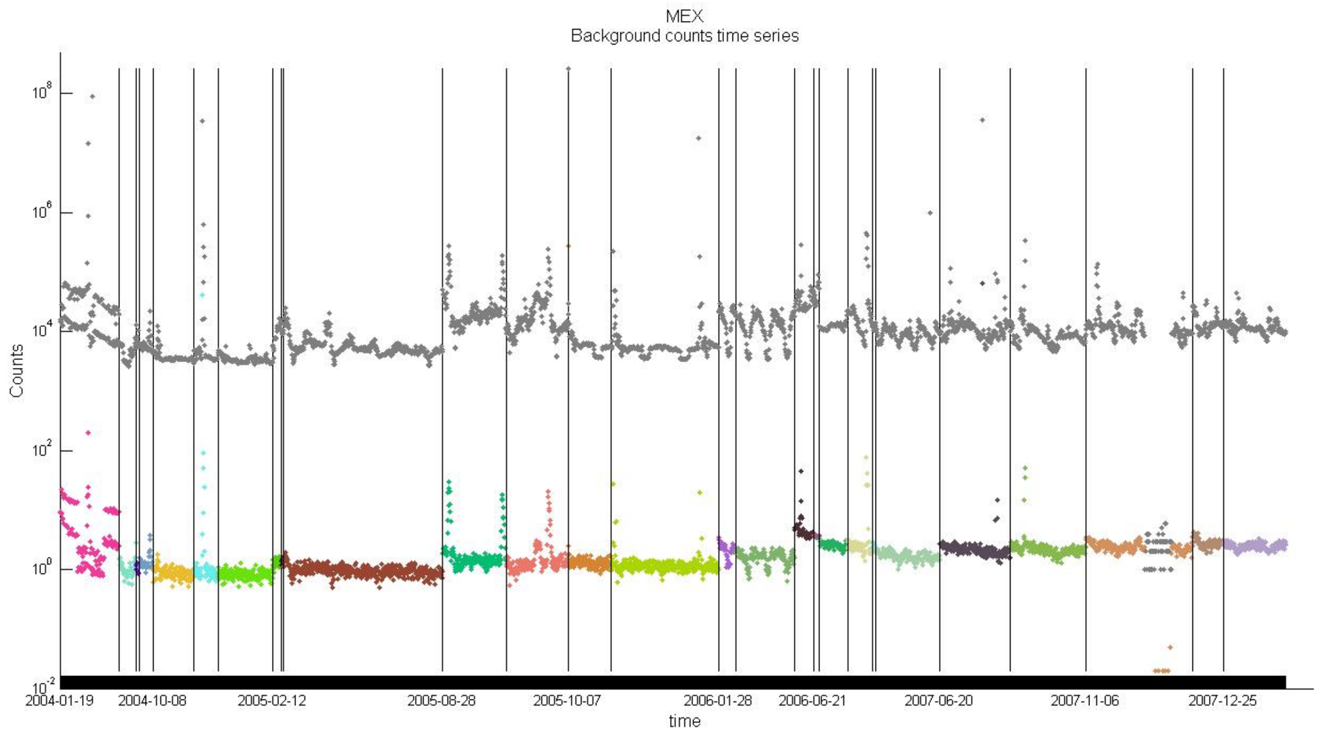

Figure 16: Background time series for Mars Express. All the data from the sample files are included. Days are displayed in different colours and separated by vertical lines. Only some dates labels are shown for clarity. The sample files dates are not consecutive.

From the preliminary analysis the data shows that the background is mainly constant in time, with a value of a few counts. Values on the order of 10^5 represent saturated measure-

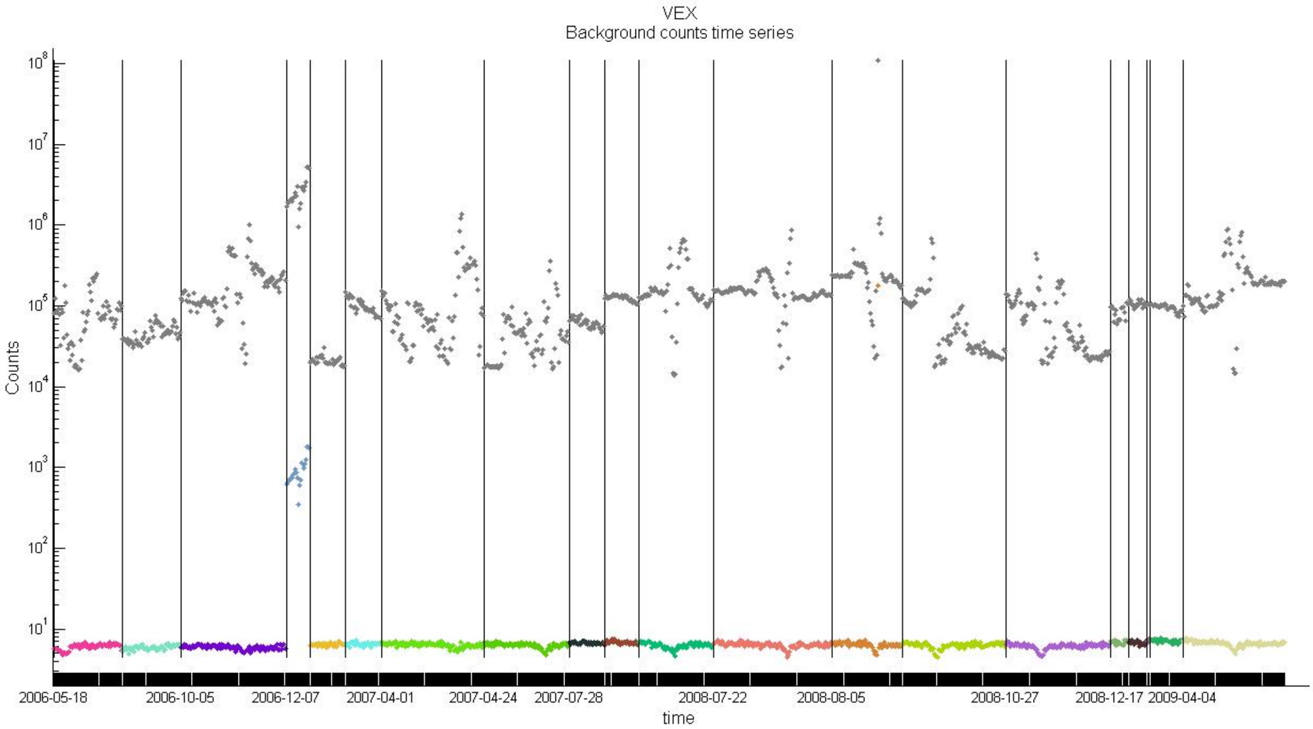

Figure 17: Background time series for Venus Express. All the data from the sample files are included. Days are displayed in different colours and separated by vertical lines. Only some dates are shown for clarity. The sample files dates are not consecutive.

ments. An increase of background in time is clearly seen in the Mars Express case, which could represent the ageing of the MCP detector. There are some other features that can be analysed in further studies. In particular, it is worth noting the increase in Venus Express background for measurements on 2006-12-07 (shown in Figure 17). Around this day a SEP event was reported (Futaana et al., 2008). Therefore, the developed software could be used as a proxy of SEP events and, eventually, of solar activity.

There are several bursty count increases in a short time scale (typically within 20 minutes) up to 100 cnt/s, both in the background and nominal counts. The bursty structures must not be related to high energy particles because they appear systematically at the beginning and at the end of the session. Moreover, such short time scale events of high energy particle illumination have not been found before. The lack of bursts in VEX background data also supports the interpretation that the bursts in MEX data are not caused by high energy particles. This may be caused by ultra-violet illumination of the sensor under a specific spacecraft attitude. Further detailed analysis may be needed to fully understand these structures.

6 Background software

6.1 Software organization

The tasks required file input and output routines, data analysis as well as visualization routines. MATLAB® provides these capabilities in a flexible way and so was the chosen programming language for this work. The software consists of MATLAB files run within its environment. The software was developed with MATLAB 7.0(R14) under Windows Vista™ SP1 (32-bit) operating system with 3GB RAM (3.8GB page file), Intel® Core™2Duo T5750@2GHz processors. The software was tested in a Linux™ Ubuntu™ server and proved to work with no problems.

Due to the different tasks to be performed it was useful to divide the software into two main branches: the analysis branch and the implementation branch (see Figure 8).

The analysis branch includes the software devoted to prepare and handle the data for fast analysis. This software includes conversion from text files into MATLAB format for fast access and less storage size, reshaping and folding the data into 2D form, plotting the data and performing background channel analysis. As this branch is to be used for analysis, user intervention and proper input preparation is expected, avoiding some input or arguments validation for the sake of practicality.

The implementation branch of the software provides the background from the text files. It includes the necessary programs and scripts to perform the background extraction as part of the ESA pipeline for the ASPERA data. As this is the branch that is expected to run without supervision, parameter and values validations are performed along the process to avoid unexpected execution stops.

There are some common modules for both branches. Refer to Appendix B for details on the software organization.

6.2 Functionality description

A brief description of the functionality, intended usage and interrelation of the main functions implemented follows. For the detailed description of each function, including declaration, dependencies and variables see Appendix B.

Analysis branch

The function *alltxt2mat* reads all the text data files in a directory and saves the contents of the files in separate MATLAB files (extension ".mat") for further processing. It uses the *file2array_v5* function for this task. Measurements with too many saturation values (see section 4) are filtered out of the process. The function *showdata* shows the data in the MATLAB files contained in a directory. There is the option of displaying a list of dates and times (Every time being a different measurement) for a certain file, plot the data for a certain time in a file or the sum of data from all the times in a file. These options make use of the function *showdate*. It is also possible to plot the data for all the times in all the files in a directory. In this case the program makes use of the function *sumdates*. The data can be plotted in an energy-mass form or energy-time form (see section 3), according to the mode specified. Currently, only the energy-mass mode is implemented. The plots are performed with the functions *showenergymass* or *showenergytime*, which uses *plot_energy_mass* and *plot_energy_time* respectively. The conversion of the stream of data for one measurement into an energy-mass matrix is done with the function *sum_energy_mass*. The function *bg_chnl_size_anal* performs the analysis of the background channel described in section 5. There is the possibility to perform the analysis with the different methods commented in that section, and also it is possible to select to perform the analysis for one file or all the files in a directory.

Implementation branch

The implementation branch makes use of the function *bg_from_file* and *sum_energy_mass*. There can be also a script to run this files in batch mode, but this is not documented here.

Auxiliary routines

There are some routines developed for fast visualization, background analysis and calculation that are not documented but are included in the software package resulted from this work, that accompanies this report. This routines can be used as examples of usage and further reference.

7 Software user manual and implementation

This section provides a “quick start” to the main functions in the software, detailed description of each function can be found in Appendix B. All the software files must be in the directory where MATLAB can find them (working directory or directory in MATLAB path).

Analysis branch

The analysis branch of the software allows viewing the data in the files in the directory and performing the background analysis. The functions must be run within MATLAB environment. The first necessary step is to convert the text files to MATLAB files. This is done by executing:

alltxt2mat(< path >), where *< path >* is the relative or absolute path to the directory containing the files to be processed. The resulting files appear in the same directory as the original.

The data can be visualized with the function *showdata*. The call of this function can be:

showdata(<date>, <time>, <satellite>, <bins>, <plotmode>, <txt_ver>)).

See Appendix B for details.

Implementation branch

The implementation branch calculates the background for a specific file. The execution command for this is:

bkg_from_file(<file_name>, <satellite>, <txt_ver>)

This executes the program for the file specified. *<file_name>* can use wildcards. *<satellite>* can be ‘MEX’ or ‘VEX’. *<txt_ver>* is the version of the satellite data text file format (‘v5’ as default).

The result of the execution is a 2-line header followed by five columns containing a line header (fixed string ‘BKG’ to ease the further processing), date, time, background value and total counts for all the measurements in the file. For example:

```
BKG # Background for file ../Data/mars5/E_IM20040119100000.txt from satellite MEX
BKG # Date Time bkg_counts Tot_counts
BKG 2004-01-19 10:47:27.969737 8.86 16076
BKG 2004-01-19 10:50:56.032236 9.32 14636
BKG 2004-01-19 10:54:08.000984 22.25 28421
BKG 2004-01-19 10:57:20.994733 8.66 16414
BKG 2004-01-19 11:00:32.994732 8.89 14652
BKG 2004-01-19 11:03:44.988480 19.23 24064
BKG 2004-01-19 11:06:57.057229 7.45 12990
BKG 2004-01-19 11:10:09.057228 6.89 12513
...
```

The background channel has a fixed absolute size. This size is determined by the variables: CONSTANTS.BKG_MA_BINS, and CONSTANTS.BKG_EN_BINS, declared in the *set_const.m* file. There are 2 pairs of these constants, one for Mars Express and one for Venus Express.

8 Conclusions and comments

From the analysis of the ASPERA-IMA data the appropriate background channel size, shape and position were identified. The resulting background channel is a rectangular subset of the energy-mass matrix obtained from the data. Two subset sizes, 14x6 bins and 11x4 bins, were determined for Venus Express and Mars Express respectively. These values correspond to the index ranges in the energy-mass matrix.

Improvements in the analysis methods for background channel determination could include taking into account the increase in counts for higher mass indexes due to increase in the physical size of the detection area. Then a normalized count could be used. In this case, ratios smaller than one are expected in Figures 12 and 13.

It is interesting to note the dependence of background counts with energy seen in Figures 14 and 15, and its explanation also requires further study.

The implementation branch of the software is fully operative. Due to the lack of time, however, the analysis branch can still be extended. Modularization of the function *bkg_chnl_size_anal* could improve readability and ease further extension. Implementation of the energy-time mode for plotting could make the software more flexible for general purpose data visualization and analysis. It is worth mentioning that the code is actually prepared to include these additions with simple changes.

With the software developed, background values for ASPERA-IMA data on Mars Express and Venus Express can be obtained from the intermediate text data files in a straightforward way. Time series of background counts in the period of the sample data showed that this information can be used as a proxy for high energetic particles detection and thus, space weather monitoring.

9 Acknowledgments

The author thanks useful comments and supervision from Dr. Yoshifumi Futaana. I also thank the people from the Solar System Physics and Space Technology research program at the Swedish Institute of Space Physics (IRF), especially Prof. Stas Barabash for the opportunity of participating in this project.

Bibliography

Barabash, S., et al., 2007. The Analyser of Space Plasmas and Energetic Atoms (ASPERA-3) for the Mars Express mission. Space Science Reviews 126, 113–164.

Futaana, Y., et al., May 2008. Mars Express and Venus Express multi-point observations of geoeffective solar flare events in December 2006. Planetary and Space Science 56 (6), 873–880.

Appendices

A Sample Data Files

In the tables of this section a list of all the data text files used in this work can be found. Text format version 5 files were more in detail analysed and some of them were excluded from the background channel analysis, but not from the background values calculation and time series. Most of the files excluded presented saturated measurements even after being processed with the saturation filter commented in section 4. In these cases, although only some measurements were saturated, the whole file was excluded.

Mars Express data files (format version 5)	Excluded because of
E_IM20040119100000	saturation values
E_IM20040424140000	
E_IM20040615011515	
E_IM20041007203702	
E_IM20041027024154	
E_IM20041208104318	saturation values
E_IM20050212070650	
E_IM20050805233307	
E_IM20050828090108	
E_IM20051007162812	saturation values
E_IM20060128182224	
E_IM20060620205523	
E_IM20060801132257	
E_IM20061113093251	
E_IM20070119231525	
E_IM20070620033431	saturation values
E_IM20070711101730	
E_IM20071106004100	strange values
E_IM20071224125939	

Table A-1: Mars expres sample data files. Text format version 5.

Venus Express data files (format version 5)	Excluded because of
E_IM20060518004324	
E_IM20060519112924	
E_IM20061004213255	
E_IM20061207044816	saturation values
E_IM20070328144014	
E_IM20070331143753	
E_IM20070401020713	
E_IM20070424022954	
E_IM20070728154243	
E_IM20071213172825	
E_IM20080428011438	
E_IM20080722003406	
E_IM20080805023730	saturation values
E_IM20081006053122	
E_IM20081027061935	
E_IM20081217223401	
E_IM20090214231044	
E_IM20090404004145	

Table A-2: Venus Express sample data files. Text format version 5.

B Code Documentation and Source Code

The documentation of the developed software can be found at the IRF web site www.irf.se.

Mars Express data files (format version 4)	Venus Express data files (format version 4)
E_IM20040119100000	E_IM20060518004324
E_IM20040625232245	E_IM20060902144841
E_IM20040817161536	E_IM20060910203325
E_IM20050212070650	E_IM20070304145926
E_IM20060215162757	E_IM20070328144014
E_IM20060302011038	E_IM20070605214156
E_IM20061113093251	E_IM20071211172430
E_IM20061222030411	E_IM20080323130651
E_IM20070105163226	E_IM20080325130459
E_IM20070119231525	E_IM20080428011438
E_IM20071123225454	E_IM20080501132008
E_IM20080422200538	E_IM20080504152717
E_IM20080629015645	E_IM20080629133326
E_IM20080717224409	E_IM20080919052003
E_IM20080818212057	E_IM20080925212346
E_IM20090205010929	E_IM20081014170708
E_IM20090308153235	E_IM20081028210951
E_IM20090401114735	E_IM20090214231044

Table A-3: Mars Express and Venus Express sample data files. Text format version 4.

The source code of the developed software can be found at http://www.irf.se/software, as a part of the complete software. The latest version of the complete software is called 'aspera_bkg_v7.1.tar.gz'. This archive includes the analysis and the implementation branches described in this report, together with a batch file to execute a series of pre-processing software (to convert the raw file format to the intermediate v5 format) and a post-processing software to re-format the output from *bkg_from_file*.

Index

Institutet för rymdfysik

Swedish Institute of Space Physics

Swedish Institute of Space Physics
Box 812, SE- 981 28 Kiruna, SWEDEN
tel. +46-980-790 00, fax +46-980-790 50, e-post: irf@irf.se

www.irf.se