GEM IM/S EUV Plasmapause Data: Directory Structure

The GEM IM/S Challenge events are 22-23 April 2001 and 21-25 October 2001. You will find two directories, one for each of these two events. In each event directory, you will find a "movies" directory and a "ppa" directory. (For example, for the 22-23 April event, the "ppa" directory is actually named "ppa_2001_Apr_22-23".)

The "movie" directory contains some animations and movies of EUV plasmasphere and plasmapause data.

The "ppa" directory contains the actual plasmapause data, as described below.

MOVIE DIRECTORIES INFORMATION

In each event folder there is a "movie" directory. (For example, for the 22-23 April event, the "movie" directory is actually named "movie_2001_Apr_22-23".) The movie directory contains four sub-folders: "mapped movies", "raw movies", "indices", and "subjective error".

EUV_raw_movies: These are the movies of "raw" images downloaded from the EUV instrument, with minimal processing. (These movies are obtainable via anonymous FTP from http://euv.lpl.arizona.edu/euv/.)

EUV_mapped_movies:
These movies contain "mapped images." A mapped image is created by taking the pixels of the raw EUV image, and assigning L and phi values to the vertices of each pixel. The L-value is assigned using the "minimum L" routine of *Roelof and Skinner [2000]* (see reference below). Thus, the pixels of the raw EUV image are mapped to the magnetic equator (in SM coordinates), presumably removing much or most of the perspective effects from the raw images.
There are two different movie files. One is just the mapped image movie, and the other has the extracted plasmapause points overplotted onto the mapped images. (The one with the extracted plasmapause has the string "PPA" in the name of the file.)

indices: This directory simply contains a .gif picture file of Final Dst obtained for the event by WDC Kyoto.

subjective_error: There are two movies in this directory, used to attempt to quantify the subjective error in the plasmapause extraction, as described in detail below. Briefly, the file labeled "MANUAL" shows comparisons between three different manual extractions, and the file labeled "AUTO" shows comparisons between the primary manual extraction and an automated extraction routine under development by Dennis Gallagher.

PPA FILES INFORMATION: Information about EUV plasmapause files

The following is a very brief description of the primary plasmapause location files, the "ppa" files.

They contain multiple points that represent the plasmapause as manually extracted from global snapshots of the plasmasphere obtained by IMAGE EUV.

Not Always Just "Plasmapause" Data.
It's worth noting that these points are not always purely "plasmapause" data. In most cases, the points do indicate a best guess for the outer edge of the plasmasphere. But there are two exceptions.
(1) No Steep Outer Gradient. If there is no clear outer boundary (e.g., the plasmasphere density has a gradual roll-off rather than a single steep well-defined density gradient), an attempt was made to represent the plasmasphere shape by clicking. In cases where there was no sharp density gradient, the click points do not define a single curve, but rather make up a scattered "band" (intended to represent the thickness of the poorly-defined outer edge), or in some cases, the click points indicate some degree of structure to the outer boundary. For a good example of this, look at the 22 Apr 2001 03:51 UT plasmapause data (look in the "PPA" movie for the 22-23 Apr 2001 event that has the plasmapause overplotted onto the mapped image, file name = "euv_22

-23Apr2001_PPA.mov"). In the post-midnight sector there is a nice, sharp plasmapause, and the click points therefore define a single curve. However, east of about 0530 MLT, the plasmasphere has no sharp outer gradient, but rather a diffuse, hazy outer boundary, and the click points therefore compose a scattered "band" of points that attempt to represent this hazy outer boundary.

(2) Multiple Edges or Interior Structure. In cases where there are multiple sharp edges, or some kind of interior structure, the click points do not define a single curve. For example, look at the movie snapshot of 21 Oct 2001 16:15 UT (In file name "euv_21-25Oct2001_PPA.mov"). While there is an outer boundary at L=3-4 discernable in the post-midnight sector, there is also a great deal of interior structure (e.g., a "V"-shaped plasma depletion interior to the plasmapause, that extends inward to about L=2).

File Naming Convention

Each file is named according to the convention YYYYDDDHHMM.ppa, where

YYYY = 4-digit year
DDD = 3-digit day of year (e.g., day 1=Jan 1)
HH = 2-digit UT hour
MM = 2-digit UT minute
".ppa" = 3-letter extension ("ppa" = shorthand for plasmapause).

File Contents

In each file, you will find all the plasmapause location points corresponding to a single EUV image. That is, all the points in a given file are simultaneous global "measurements" of the plasmapause.

The .ppa file consists of ASCII text, 2 columns consisting of:

L = L-shell location (defined as the radial distance in the magnetic equator in Earth radii RE).
phi = angular location in radians, defined so that phi=0 at 1200 MLT (noon), phi=pi at 2400 MLT, etc.

At the end of the .ppa file, you will find a Fourier expansion of the (L,phi) points. This Fourier expansion is useful in cases where one wishes to determine electric fields from the plasmapause motion (as described in Goldstein et al., [2004]; see reference below.). It is not advised that you use this Fourier expansion without ***first consulting with Jerry Goldstein (jgoldstein@swri.edu)***. The reason for this warning is that the Fourier expansion is only a good representation of the plasmapause under certain restrictive circumstances. Probably the most important requirement for a "good" Fourier expansion is that the plasmapause (L vs phi) curve must be single-valued. If there is a drainage plume, for example, it is most likely that the plasmapause curve is *NOT* single-valued. A secondary consideration is that the Fourier expansion is only really good inside a range of MLT (i.e., phi) values for which there were actually extracted points. For example, if there are only plasmapause points extracted on the nightside, one should not expect, in general, the Fourier expansion to be accurate on the dayside. *Improper use of the Fourier expansion can provide a very innaccurate plasmapause!*

Notes on Manually-Extracted Plasmapause Location: "Click Points"

SUBJECTIVE ERROR

Something to keep in mind is that the (L,phi) points have been manually extracted. There is inescapably some subjectivity to the extraction, although comparison with in situ data has shown that the manually extracted plasmapause is reasonably accurate (within a few tenths of an RE) under most circumstances.

We have attempted to provide an estimate of the subjective error by having four different EUV investigators extract the plasmapause. By comparing these four different extractions, we can get a feel for how much error is introduced by individual subjectivity.

Manual extractions were performed by Jerry Goldstein ("JG"), W. Terry Forrester ("WTF"), and Bill R. Sanel ("BRS"). An automated extraction technique still *under development* was provided by Dennis L. Gallagher ("DLG"). Note that the "JG" extraction is the one that appears in the .ppa files.

In the "movies" directory, you will find a subdirectory called "subjective_error". This subdirectory contains two files, one comparing the manual extractions of three EUV investigators, and the other comparing the JG manual extraction to the automated extraction of DLG.

Comparing Different Manual Extractions

The "MANUAL" movie file shows snapshots for which three manual extractions were performed. Each frame of the movie plots L (of the plasmapause) versus MLT. The three manual extractions of JG, WTF and BRS are color- and symbol-coded as given in the legend. The horizontal dotted lines are separated by 0.2 L; the vertical dotted lines are separated by 1 MLT hour.

You can see from the MANUAL comparisons that in most cases where a plasmapause was clearly identifiable, the three manual extractions differed by not very much (0.1 L or less!). However, there were many instances where the three manual extractions differed by as much as 1 L-shell. In these cases, one of the following was true:

(a) There was no clear outer boundary gradient, but rather a gradual density drop-off or perhaps some significant amount of interior plasmasphere density structure;

(b) There was significant background noise or significant sunlight contamination in the image that partially obscured the plasmapause gradient, leading to increased subjectivity.

A nice example illustrating this is 22 Apr 2001 03:51 UT (in the file "euv_22-23Apr2001_MANUAL.mov"). Where the plasmapause was nice and sharp (between 0000 MLT and 0530 MLT), the three manual extractions differ by no more than 0.2 L. However, east of 0530 MLT, the outer density gradient was more gradual (as reflected in the scattered "band" of green points), and in this MLT sector of gradual density drop-off, the WTF and JG extractions differ by about 0.4 L, and the BRS and JG extractions differ by about 1 L-shell!

Comparing Manual to Automated Extraction

The "AUTO" movie shows snapshots similar to the "MANUAL" described above, but here we're comparing the JG manual extraction (green points) to the DLG automated extraction (blue points).

The automated extraction generally is close to the manual extraction when there's a well-defined plasmapause with good image quality (not a lot of background noise or sunlight contamination). However, the automated technique cannot distinguish some instrument artifacts very well, leading to "plasmapause" data that is really indicating camera edges, or image streaking due to sunlight contamination.

As an example, look at 23 Apr 2001 04:30 UT (in the movie file "euv_22-23Apr2001_AUTO.mov"). Where there are green points, there was indeed a nice plasmapause. However, there are gaps in the green points centered about midnight MLT (i.e., 2300 to 0100 MLT), and east of 0600 MLT. These gaps are due (respectively) to the presence of the Earth's shadow near midnight, and the presence of sunlight contamination that obscured the dayside plasmapause. The automated extraction (blue points) shows no gaps in these sectors because the automated algorithm mistakenly identifies an edge in the shadow region, and also mistakenly identifies the drop-off in dayside intensity (due to sunlight contamination) as the plasmapause. In fact, on the dayside, the plasmapause was located *just outside* the sunlight contamination edge, as later EUV images (e.g., 5:11 UT) show.

Why are we still using manual extractions?

A common question raised is, "Why hasn't anyone devised a routine for automated plasmapause extraction? Wouldn't an automated routine be better? One argument is that the automated extraction might be described as "not subjective" (since a computer isn't usually thought of as subjective). Another is that the automated routine would be quicker, not so labor intensive. The answer to this question is that the human eye does as well, or better, at recognizing the plasmapause location, when compared to any automated routine developed at the present time. It's a challenging problem to design an automated routine that can ignore shadows, noise, sunlight contamination, etc. So for now, the best we have is manual extraction, and we do our very best to minimize the subjectivity.

Error Bars on Plasmapause Data Points?

Ideally, we would provide not just plasmapause L,MLT locations, but also error bars (delta L). At this time, individual error bars for individual poijnts have not been determined. Reasons:

(a) We didn't do multiple extractions for all the available EUV images--there are many more .ppa files than there are WTF, BRS and DLG extractions.

(b) Even if we had multiple extractions to go with every .ppa file, we still would have a tough time assigning quantitative error bars--it would require a great deal of work. We may get to this eventually, but for now, we recommend the following.

If you have questions about the error bar of a particular plasmapause location, it is recommended that you go to the "MANUAL" or "AUTO" movie snapshot that is closest in time to the time you are interested in, and find the MLT sector that is closest to the point of interest, and estimate the error as the scatter between the JG, WTF and BRS extractions, or between the JG and DLG extractions. For example, suppose you were interested in the plasmapause of 22 Apr 2001 04:22 UT, at about 0400 MLT. Go to 22 Apr 2001 04:42 UT in the "MANUAL" movie and you'll see that the BRS and WTF extractions differ from the JG extraction (which is the

one in the .ppa file) by 0.2 L or so. Go the 04:42 UT in the "AUTO" movie and you'll see that the JG manual extraction and the DLG automated extraction agree to within 0.02 L or less. Therefore, for this particular MLT location at this particular time, the subjective error in the plasmapause L-location is probably no more than 0.2 L.

Admittedly, this is a very cumbersome way to get error bars! Ideally, you'd like individual error bars, but they are just not available at this time. Another way is to scroll through the MANUAL and AUTO movies for a given event and try to get a feel for the minimum and maximum error, noting times with particularly bad background noise or sunlight contamination, and taking this into account. As a general rule of thumb, the October 2001 event has worse background noise & signal contamination.

IMPORTANT NOTE ON USING PLASMAPAUSE DATA

In cases where a well-defined plasmapause edge existed, single plasmapause location points can be trusted. However, in cases where the click points were intended to represent a gradual density gradient, or interior features, care must be exercised in interpreting single points. In fact, the best way to use the EUV extracted "plasmapause" data is to do global comparisons between the *entire plasmapause shape* seen by EUV and by your simulation/model.

If you have any **questions** about how to use EUV plasmapause data, *please contact* ***Jerry Goldstein (jgoldstein@swri.edu).***

Linux Compatibility

For users of Linux or UNIX, some may have problems with the .ppa files that were created using IDL 5.3 for Mac OS 9.1. For these users, a separate directory of Linux/UNIX compatible .ppa files has been created for each event. The name of this directory will have the string "_Linux" at the end of it.

References

Goldstein, J., R. A. Wolf, B. R. Sandel, and P. H. Reiff, Electric fields deduced from plasmapause motion in IMAGE EUV images, *Geophys. Res. Lett., 31*(1), L01801, doi:10.1029/2003GL018,386, 2004.

Roelof, E. C., and A. J. Skinner, Extraction of ion distributions from magnetospheric ENA and EUV images, *Space Sci. Rev., 91*, 437, 2000.