February 19, 2020

HELIOS 1 & 2 IN-SITU DATA ARCHIVE
Project Action Item List

Helios 1 & 2 rank amoung the most important missions in Heliophysics, and the more-than 11 years of data returned by its spacecraft remain of paramount interests to researchers. Their unique trajectories which brought them closer to the Sun than any spaceccraft before or since, enabled their diverse suite of in-situ instruments to return measurements of unprecedented scientific richness. Analyses of these measurements produced groundbreaking insights into the large-scale spacial and temporal variations in the inner heliosphere (*Marsch*, 1991a,b, 2006), solar wind turbulence across both MHD and kinetic scales (*Marsch*, 1991b; *Bruno and Carbone*, 2005, 2013), the effects of kinetic microinstabilities (*Marsch et al.*, 1982a; *Marsch and Livi*, 1987), the process of collisional thermalization (*Marsch et al.*, 1982b), and ongoing heating processes (*Schwartz and Marsch*, 1983).

Chadi Salem (SSL/UC Berkeley) has been funded by NASA, under a two-year HDEE grant (NASA grant NNX14AQ89G, 10/1/14 - 9/30/16) to work on an important project of aggregating, analyzing, evaluating, documenting and archiving the available Helios 1 and 2 in-situ data. He has been working closely on this project with colleagues at the University of Koln, at the University of Kiel, at the Imperial College in London and at the Paris Observatory.

Below is a list of action items for the Helios data archive project as discussed and defined at the June 2016 Helios Mini-Worshop in Cologne Germany. The focus of the workshop was largely on the magnetic field measurements from the E3 and E3 fluxgate magnetometers and the E4 search coil magnetometer.

The team present at this workshop is made of the following colleagues:
- Chadi Salem (PI, SSL/UC Berkeley)
- Fritz Neubauer (University of Köln)
- Lex Wennmacher (University of Köln)
- Joachim Saur (University of Köln)
- Jan Steinhagen (University of Kiel)
- Lars Berger (University of Kiel)
- Milan Maksimovic (LESIA, Observatoire de Paris)
- Olga Alexandrova (LESIA, Observatoire de Paris)
- Thierry Dudok de Wit (LPC2E, Orléans)
- Matthieu Kretzschmar ((LPC2E, Orléans)
- Lorenzo Matteini (Imperial College London).

People in the team not present at the workshop are:
- Bob Wimmer-Schweingruber (University of Kiel)
- Eckart Marsch (University of Kiel)
- Rainer Schwenn (Max-Planck-Institut für Sonnensystemforschung)
- Roberto Bruno (University of Rome)
- Tim Horbury (Imperial College London).

The comprehensive Helios data Archive, hosted at the Space Sciences Laboratory at UC Berkeley, is now up and running. The url is: `http://helios-data.ssl.berkeley.edu`. It is still a work in progress, as we gradually improve the archive, expand its content with more documentation, and more and newer datasets added as we proceed further with the project.

1 General

1.1 Scan the relevant PhD thesis and documents gathered at University of Koln by Fritz N. This includes the *orange book* (Lex) and the *yellow book* by *Musmann & Maier* (Fritz). The orange book that Lex has is very useful for Helios 1 only as it has all the information about the commisionning phases in English. **Fritz N & Students**. **DONE**.

1.2 Scan the 3 PhD thesis and the short papers in German that Fritz has. **Fritz N & Students**. **DONE**.

1.3 Scan all the instrument blue books found in Kiel. **Lars B**.

1.4 Gather the information about the attitude of the s/c (spin axis angle with respect to Ecliptic). **Lex W**.

1.5 Produce prediction Kernel for Helios 1 for the time missing ($>$ 1980-1985): orbital position. **Lex W**. **DONE**.

1.6 Produce tables (for Heliso 1 an 2) with YY MM DD HH, ephe time, cartesian coordinates, ecliptical coordinates, J2000, orbital Vx, Vy, Vz (s/c), vector from Sun to Earth, velocity of Earth. **Lex W**. **DONE**.

1.7 Start the construction of a website at SSL that will host the Helios data archive. First make it password protected and available to users in the team. **Chadi S**. **DONE**.

1.8 Ask SSL for disk space to host the Helios data archive. **Chadi S**. **DONE**.

1.9 Make things available to the team on this website. **Chadi S**. **DONE**.

1.10 ISSI Team on the Helios Archive Project. Bob Wimmer-Schweingruber will lead this initiative. As of Nov 21, 2016, Bob and Chadi have a draft proposal that they are working on. The project will include all helios instruments, beyond the in-situ E1 to E5 instruments. **Bob W and Chadi S**. **NO**.

1.11 Tim Horbury mentioned also the Horizon 2020 proposals. Are we thinking of putting a Helios Archive proposal forward? **Tim H**. **NO**.

2 E2 Fluxgate mag data

2.1 Determine what is exactly the noise level (0.1 or 0.4 nT/$\sqrt{Hz}$?). Reference day: Jan. 6, 1978 (Olga's birthday). **Thierry DdW**.

2.2 Send to the group the E2 and E3 data for the following 2h interval of data: Jan 6, 1978, starting at 5:37:06.201 from Helios 2. **Chadi S**.

2.3 Estimate the digitization noise for E2. **Lars B**.

2.4 Estimate the noise floor from the data. **Thierry DdW**.

2.5 Look into the 50nT saturation level in the high resolution data, not seen in the 40.5 sec NSSDC data. **Lex W and Chadi S**. **DONE – See Below**.

2.6 Try to correct the source code and also try to recover all useful data. **Lex W and Chadi S**. **Obsolete**.

2.7 A good example of an interval illustrating the 50nT cut is the shock day on Helios 2: 1977, Days 110-112. One can look at the IGM data and compare to lower resolution data. Also, check if there is a correlation in time between range switchings and cuts at 50nT. **Lex W and Chadi S**. **Obsolete**.

2.8 Compile files with the previously fabricated lower resolution (40.5 sec and 8 sec) averages of the E2 mag data made in Köln. Lex and Chadi found out that they were unaffected by the 50nT cut. => Produce ascii files of these 8 sec and 40.5 sec averages (v2 & v3). These will be very useful. **Lex W**. **DONE**.

2.9 Make sure that the Cologne 4Hz data is complete and not a CD or so is missing from what was given to Kiel. **Lex W**. **DONE**.

2.10 Produce ASCII files for all the averages available in Cologne (1 sec, 8 sec, 30 sec, 40.5 sec, 60 sec, etc.). **Lex W**. **DONE**.

2.11 Produce and add (at all resolutions) also the mag data for when Spice Kernels were not available. Lex produced the spice kernels missing and he will add the corresponding mag data to the rest of the data set. **Lex W**.

2.12 Try to correct the zero-level on E2 Bz. From discussions with Fritz N., it seems like it is difficult to do and it would take a lot of time. This action item depends on a few items listed in Section 5, namely items # 5.1 – 5.4.

2.13 Flipping Problem discussed by Fritz: Is it only in the early phase of the mission? During the commissioning phase, where the instruments were being tested and calibrated? If so, IDENTIFY a safe start date for the use of the data after which we can decide that the s/c was spinning nominally and the data is good! Or identify the exact dates of commissioning phase. **Lex W**.

2.14 Identify calibration times (where magnetic field is applied and they tried to measure it) and these intervals are still in the data. There is no documentation on when these were happening. Fritz says at the beginning of the mission and further in the orbit. We can check the plots and identify big jumps in all three components (are they at the same time?). **Lex W**.

2.15 Check also for internal sensitivity calibration in E2. They are different most certainly than for E3. See below in E3 section. **Lex W**.

2.16 **50nT cut:** Lex and Chadi worked on this issue of the 50nT clipping of the data and eventually found the source of the problem. Document what we found out: process of peak removal and threshold used was 25nT then changed to 50nT. **Lex W and Chadi S**. **DONE**.

2.17 **50nT cut:** See if we can locate the missing data on CDs or hard drives. Lex has already found some of the removed peaks, but only a fraction. **Lex W**. **DONE**.

2.18 Contact the person Lex and Chadi talked about for more information. **Lex W and Chadi S**. **DONE**.

2.19 Try to reconstruct the data with removed peaks. If no trace of removed peaks, then perhaps use 8 sec average. **Lex W**. **NO**.

2.20 Compile a list of B>50nT gaps using the 8 sec averages. **Lex W and Chadi S**.

NOTES on E2:
We discussed the making of the merged mag-plasma data with Rainer Schwenn on the phone from Lex's office and we believe that the 40.5 sec mag data averages present in the plasma data in Kiel was made using the 8 sec averages provided by Fritz N. to Rainer S. back then. We found that the 40.5 averages in the files containing distribution functions have a factor 10 and have lots of dips that can be explained by an averaging done with a constant number of points in each average performed using 8 sec or 4Hz data, regardless of the presence of gaps.

3 E3 Fluxgate mag data

That dataset is available at the resolution of 6 sec only. The higher resolutiion data was stored in magnetic tapes at the University of Rome but they were discarded 1 or 2 decades ago because of tape deterioration. So they are gone forever. The 6 sec E3 averages available were stored at the NSSDC and were retrieved from there.
Warning! Load the data with a fixed number of columns and fixed format (fixed number of characters): so count the characters!

3.1 Look into the big drops of -150nT in Bz (and other components?) found and shown by Jan S. How often do they happen? Do they happen for each component simultaneously? Or is it just Bz? Is it due to some internal sensitivity calibration? **Who**?

3.2 Ask Len Burlaga if there are any E3 internal sensitivity calibration data. Perhaps Ask Roberto Bruno to ask/email him directly? **Roberto Bruno**?

3.3 Are these drops rather due to an error in file/column formatting? When/where would this have happened? **Who**?

3.4 Determine what is exactly the noise level for E3. Digitization noise and noise floor from data. **Roberto Bruno**?

4 E4 Search Coil data

4.1 Provide the best E4 data we have from Lex in ASCII format. It was decided that Lex's E4 data set will be the reference data set for E4. **Lex W**.

4.2 Cross-check Lex's E4 data with the same data set that Thierry and Matthieu read independently from binary files for sanity check. **Thierry DdW and Matthieu K**.

4.3 Redo the analysis by Neubauer et al. (Beinroth & Neubauer 1981, histogram of power spectral density at a given frequency) with the current data set and compare with the results from this 1981 paper. Try to reproduce Fritz's result. Hopefully we can obtain the same results. **Fritz N and Lex W and Matthieu K**.

5 Inter-Instrument Comparisons and Calibrations

5.1 E2 and E3 do not provide the same measurements of the magnetic field. For the sake of comparison, make the histogram of the 6/12/60 sec averages of 4Hz Bz (and Bx and By) and compare them to the histograms of the E3 Bz (and Bx and By). **Jan S**.

5.2 Do a cross-correlation between E2 and E3 data (Lex W. suggestion) to test whether the time agreement between E2 and E3 is good: for example, a shock time interval is the same between E2 and E3. At least get a grasp of how large the error/time difference is. Ideally it is Zero. **Jan S/Lex W/Chadi S**.

5.3 Re-do Jan's plot on the deviation between E2 and E3 in percentage at 12 sec resolution, and produce quality flags where the deviation is high. After doing this analysis, share with the team so we can decie what the right procedure to flag the data intervals is **Jan S**.

5.4 To understand the origin of the measurement differences or deviations between E2 and E3, Fritz and Chadi discussed post-workshop a few ideas on whiteboard and Chadi will try to perfom the corresponding data analysis. **Chadi S**?

Calculate Bz(E2) - Bz(E3):
– 1h averages of Bz(E2,4Hz) and Bz(E3) as a function of time: it should be quiet, no spikes
– decrease the time averaging to 60s then 6s, if 1h shows nothing
– look at these differences for different time intervals, like at perihelion etc. for H1 and H2 (H1 had a big temperature problem).

If deltaBz = Bz(E2,4Hz) - Bz(E3), then:
– Plot deltaBz as a function of Bz
– Plot deltaBz as a function of R
– Plot deltaBz/Bz as a function of R
– Compare from aphelion to aphelion? Compare early or late in mission?

Similarly: if deltaBx = Bx(E2,4Hz) - Bx(E3), and deltaBy = By(E2,4Hz) - By(E3)
– Plot deltaBx as a function of Bx and deltaBy as a function of By
– Look for effect of thermal issue
– Effect on B measurement close or far to the Sun (not serious close to Sun: dB/B smaller than at 1AU?)
– Long term trend: deterioration of the solar panels (cells)

5.5 Provide, as a caveat to the users, one curve (or several) showing the (dis-)agreement between E2 and E3. What time scale? at 6 sec? or 12 sec? **Jan S**.

5.6 Solve the question: is there a factor 10 between E3 PSD and E4? What about E2 PSD and E4? **Thierry DdW**?

5.7 Inter-connection between the E2 spectra and E4 spectra, and the E2 noise level. How can we reconcile the spectra between 0.01 and 10 Hz? **Chadi, Matthieu, Thierry, Olga and Lars**.

6 E1 plasma – ion & electron – data

The E1 experiment data existing today consists of ion and electron distribution functions (3D and/or 1D/reduced for the ions and 2D only for the electrons) and corresponding proton and alpha particle moments (density, velocity, temperature, etc.).

All of the E1 data available has been produced by Prof. Eckart Marsch and Prof. Rainer Schwenn sometime in the 90s (**????**) in the format of ascii files, each file corresponding to a single time tag and distribution function (ion and electron or sometimes one or the other). Each file, for a single time tag, a series of parameters, **describe those from the README file provided with the data**, including values for the magnetic field averaged at the resolution of the plasma data. The entire data sets as it is today has been provided to the Archive by the Team at the University of Kiel. Note that the original or raw E1 data are gone forever.

As far as the analysis of the E1 plasma data goes, there has been a preliminary work with Team at the University of Kiel to assess the uncertainty and reliability of the plasma data (proton/alpha moments) and associated magnetic field data in those files. It was not clear then how the available moments were calculated exactly (from the 3D or the reduced 1D distribution etc.). The original instrument blue book (in German) was provided by Eckart Marsch and another internal document (also in German) describing the processing of the plasma moments was provided by Rainer Schwenn.

From the 1D reduced ion distribution functions from the I1a instrument only, the protons and the alphas were separated by a cut in the distribution at a certain energy corresponding to a dip between the proton and the alpha populations. Then each distribution was integrated to produce its associated moments (basically density, speed, and total temperature). Then by integrating the 3D distribution functions from I1a, a velocity vector, is finaly provided. The original code that provided these moments in Fortran is included (in print format) in the document provided by R. Schwenn.

A careful analysis of the plasma data has revealed issues with the accuracy of these available plasma moments (ion density, velocity and temperature etc), as well as a few issues with the processing into the files of the various more developed products such as 1D (reduced) particle distributions and 3D distribution functions from the raw data. The plasma data, in particular, the ion moments, have been widely used in the past and continue to be used to date, without knowledge or understanding of their issues and limitations.

A few of these issues are: - Low energy bins contains what seems to be the instrument noise floor for the ion distributions (this is mentioned in the blue book). - the units of the 1D reduced distrutions (either for the I1a distributions or I1b current distributions) are not clear. For example, if we calculate the reduced distribution from the 3D I1a one, we get a factor that we could not figure out. - When the ascii data files were produced, the parts of the distribution functions with zero counts were omitted. So all zero counts are missing. This makes the reading of the data files a bit challenging because the files do not contain the same grid for all distributions (this info changes from file to file, i.e. from a measurement to the other). Of course this issue also affects the calculation of moments. - the electron distribution functions were corrected by a spacecraft potential of the form $\phi = const. \log N_p$ (as indicated in the original fortran programs), which is a very simplified form for the spacecraft potential. Therefore we can easily uncorrect from this s/c potential and use a more sophisticated spacecraft potential for any further processing of the electron distribution functions.

The E1 experiment and associated data analysis and processing have not been discussed at the

Helios mini-workshop in Cologne in June 2016. Since Jan 2017, there is on-going effort to characterize and understand the issues and the processing of the data available as a close collaborative work between UC Berkeley and the Team at Imperial College London (David Stansby, Lorenzo Matteini, and Tim Horbury). The goal of this work is to reprocess the entire plasma data set, using more modern fit analysis techniques to calculate better and more accurate plasma parameters (density, velocity and temperature including anisotropies, and heat flux) for the various components forming the ion distribution function (proton core, proton beam and alpas) and the electron distribution function (core, halo and strahl). A proposal to the NASA HDEE 2017 Program has been submitted in July 2017 to this end, and funded. The work of reprocessing the ion data is still in progress. To date (Dec 2019), we have fitted the proton core portion of the ion distribution function and determined new, more accurate, densities, velocities (3 components), parallel, perpendicular as well as total temperatures of the proton core. In addition, we fitted the alpha particle population as well and determined similarly densities, velocities (3 components), parallel, perpendicular as well as total temperatures of the alpha particles. We are working on fitting the proton beam portion of the distribution. After that, we will analyze and fit the electron distribution functions and characterize core/halo and strahl.

Both the E1 experiment blue book and the document describing the original moment calculation have now been translated from German to English at Imperial College and are included in the Archive in the "E1 Experiment" directory. These documents contain valuable information on today's available dataset. A summary describing the various instruments and data sets has been compiled by David Stansby and can be located in the E1 experiment directory of the Helios Archive

7 Archiving Process

6.1 All data + documentation (papers, blue books, books, PhD thesis, reports, presentations, etc.) gathered to be organized and made available online in the Archive. **Chadi S**. **DONE**.

6.2 Write-up outcome/sumary of Mini-Workshop and data description of all known issues. See current document, sections above. **Chadi S**. **DONE**.

6.3 Archiving: make CDF (NASA Compliant) files with all the data we gathered and simple routines to read them (IDL?).

6.4 Keep original files as well.

6.5 Can all the binary files available online at the NSSDC/CDAweb be read (with the help of GSFC and the instrument blue books) and converted into readable/usable files?

8 References and Citations

References

Bruno, R., and V. Carbone (2005), The Solar Wind as a Turbulence Laboratory, *Living Reviews in Solar Physics*, *2*, 4–+.

Bruno, R., and V. Carbone (2013), The Solar Wind as a Turbulence Laboratory, *Living Reviews in Solar Physics*, *10*, 2, doi:10.12942/lrsp-2013-2.

Marsch, E. (1991a), *Kinetic Physics of the Solar Wind Plasma*, pp. 45–133, Physics of the Inner Heliosphere II. Particles, Waves and Turbulence, XI, 352 pp. 152 figs.. Springer-Verlag Berlin Heidelberg New York. Also Physics and Chemistry in Space, volume 21; 2, p. 45-133.

Marsch, E. (1991b), *MHD Turbulence in the Solar Wind*, pp. 159–241, Physics of the Inner Heliosphere II. Particles, Waves and Turbulence, XI, 352 pp. 152 figs.. Springer-Verlag Berlin Heidelberg New York. Also Physics and Chemistry in Space, volume 21; 2, p. 159-241.

Marsch, E. (2006), Kinetic Physics of the Solar Corona and Solar Wind, *Living Reviews in Solar Physics*, *3*, 1–+.

Marsch, E., and S. Livi (1987), Observational evidence for marginal stability of solar wind ion beams, J. Geophys. Res., *92*, 7263–7268, doi:10.1029/JA092iA07p07263.

Marsch, E., H. Rosenbauer, R. Schwenn, K.-H. Muehlhaeuser, and F. M. Neubauer (1982a), Solar wind helium ions - Observations of the HELIOS solar probes between 0.3 and 1 AU, J. Geophys. Res., *87*, 35–51, doi:10.1029/JA087iA01p00035.

Marsch, E., R. Schwenn, H. Rosenbauer, K.-H. Muehlhaeuser, W. Pilipp, and F. M. Neubauer (1982b), Solar wind protons - Three-dimensional velocity distributions and derived plasma parameters measured between 0.3 and 1 AU, J. Geophys. Res., *87*, 52–72, doi:10.1029/JA087iA01p00052.

Schwartz, S. J., and E. Marsch (1983), The radial evolution of a single solar wind plasma parcel, *Journal of Geophysical Research*, *88*, 9919–9932, doi:10.1029/JA088iA12p09919.