EUROPEAN SPACE AGENCY

Extension of the SOHO Mission

1. Introduction

Since its launch on 2 December 1995, SOHO has provided an unparalleled breadth and depth of information about the Sun, from its deep core to the outer corona, the solar wind, and the interaction with the interstellar medium. Here we propose a fourth mission extension, of 3 years, covering the period from January 2010 through December 2012. There is a unique opportunity to measure critical parameters of the Sun's interior, its atmosphere, and the impacts of solar activity in the heliosphere over an entire 22-year magnetic cycle, but only if SOHO is able to continue observing in conjunction with the Solar Dynamics Observatory (SDO), the lineal descendent of SOHO and the first flagship mission of NASA's Living With a Star (LWS) program. SDO will replace some of SOHO's measurements; this will allow a descoping of the original SOHO mission and significant reductions in operations cost. SOHO's continued measurements of total solar irradiance, integrated EUV solar output, low-frequency global solar oscillations, energetic particles and solar wind plasma parameters coupled with SOHO's unique capabilities for UV and white light coronagraphy, and UV and EUV spectroscopy will be critically important if we want to understand how and why the Sun varies – one of the key objectives of the International Living With a Star (ILWS) program.

SOHO is operated from the SOHO Experimenters' Operations Facility at NASA Goddard Space Flight Center (GSFC), and NASA covers the major part of the operating costs. In the 2008 NASA Heliophysics Senior Review[1] SOHO was ranked in the "excellent" category for both science merit and for relevance to NASA's overall Heliophysics goals. The panel recommended continued funding through FY 2012. The spacecraft and payload are in good condition and there are no technical limitations which would prevent SOHO from observing for several more years.

2. Science Case for Extension

2.1 Introduction

SOHO has provided a nearly continuous record of solar and heliospheric phenomena for an entire activity cycle, i.e. half a magnetic cycle. These observations have revolutionized our understanding of the Sun and space weather research. Since late 2006, SOHO observations have been complemented and enhanced by those from NASA's STEREO and JAXA's Hinode missions (and vice versa), and the synergies between these missions will be even greater during the onset and rise of solar cycle 24. After the launch of SDO (possibly in June 2009, or as late as January 2010), SOHO will fill a critical lacuna in its instrument complement. After the descoping of the coronagraph on SDO there is no other mission (even in planning) that can provide observations of the Sun's corona and coronal mass ejections (CMEs) from the Sun-Earth line. NASA's stated main interest in continued SOHO operations is to maintain these critically important ILWS and space weather measurements.

2.2 Impact of mission to date

Publications: SOHO enjoys a remarkable "market share" in the worldwide solar physics community: Over 3,150 papers in refereed journals since launch (not counting refereed conference proceedings), representing the work of over 3,000 individual scientists. With 352 publications, the last year (2007) has been the most productive. A searchable SOHO publications database[2] is available on the SOHO home page. It is not too much of an exaggeration to say that every living solar physicist has had access to SOHO data.

[1] http://umbra.nascom.nasa.gov/soho/sr08/soho_sr_2008_nobudget.pdf
[2] http://soho.esac.esa.int/publications/publications.html

Data accessibility: All the SOHO experiments make all their data available online through the SOHO archive, at PI sites, and via the Virtual Solar Observatory (VSO). A typical PI site, the EIT Web catalog, has served over 1.4 Tbyte of data in response to over 1,000 requests since early 2006 — and the EIT database was only 700 Gbyte at the end of 2007. The larger MDI database, which includes several levels of computationally expensive, higher-level data products, has served over 17 Tbyte in response to over 10,000 online data requests in the last two years. (This total does not include an equal volume of larger data exports shipped to Co-Investigator and Guest Investigator sites on tape.) In addition to professional access, amateurs routinely download LASCO data to search for new comets. As a result, over half of all comets for which orbital elements have been determined (since 1761) were discovered by SOHO, over two thirds of those by amateurs accessing LASCO data via the Web. 128 of the 177 comet discoveries in the last year confirmed by the IAU were made with SOHO observations.

Research access: All SOHO instruments' scientific data are accessible through a single interface[2] to the SOHO master archive, which has been developed at GSFC by the ESA SOHO Project Scientist Team. This interface searches both the general SOHO archive at GSFC and the MDI high-rate helioseismology archive at Stanford. The holdings of the SOHO archive are identical to those used by the PI teams, and are current (i.e. to within a month or two before present, to allow time for "Level-Zero" data delivery.) Mirrors of the SOHO archive are maintained at the Institut d'Astrophysique Spatiale (France), the University of Torino (Italy), and ESAC for faster access by European researchers. SOHO data were among the first whose metadata, including browsable images for EIT and MDI, became searchable via the VSO. The VSO is designed to deliver data via the original servers, so the download traffic still occurs at those sites.

Public interest: Public interest in SOHO (as measured by Internet traffic) has constantly increased over time. The SOHO webservers operated by the ESA Project Scientist Team at GSFC now serve an average of 25 million requests and over 14 Tbyte of data each month. Since launch, the SOHO servers have served over 315 Tbyte of images and data in response to over 1.4 billion web page requests.

2.3 Expected science in extension

In the following paragraphs we will highlight some of the unique science opportunities afforded by continued SOHO operations during the onset and rise of the next solar cycle, the maximum of which is expected to occur in 2012. Given the limited space available, we can address only a very small subset of the numerous science projects currently ongoing or planned for the future.

TSI variations and solar "forcing" of terrestrial climate. A long-term, accurate, absolutely calibrated record of total solar irradiance (TSI) is essential for answering questions about the degree to which solar forcing affects climate on earth. The TSI record from the SOHO VIRGO radiometers now extends over 12 ½ years. Significant differences in their designs have made it possible to determine and eliminate instrumental aging from the record of VIRGO, and the methods could also be applied to earlier TSI records used for a composite that now spans nearly 30 years. Even though solar minimum had not occurred by mid of 2008, TSI dropped significantly below the solar minimum value of 1996. Relative to the cycle amplitude the present minimum is almost 20% lower than the previous one when SOHO started. Continued operation of the VIRGO instruments will allow redundant, overlapping measurements with new radiometers (e.g. on Picard), which is critically important to maintain the nearly 30-year record of TSI measurements, and it will allow scientists to see whether the rise of the new cycle is similar to the previous one after 1996.

[2] http://soho.esac.esa.int/data/archive/

Global oscillations and the dynamics of the Sun's core. From the analysis of 10 years of SOHO GOLF data, a Spanish-French team of scientists concluded that the power in a range of periods around 24 minutes was consistent with a model of g-modes. Their results also suggest a solar core rotating significantly faster than the rest of the radiative zone. The degree of significance, however, is far from conclusive and there is a divergence of opinion in the g-mode community regarding the significance of this result as a detection of the long-sought modes believed to exist deep in the solar interior. Clearly, more research is needed in this area with improved and extended data. It is important to note that, after the announced closure of the BISON network, GOLF and VIRGO will remain the only instruments providing measurements of the Sun's global p- and g-mode oscillations.

Transition region dynamics. SOHO SUMER and CDS are the only UV and EUV spectrometers that can be used to measure transition region lines ($10{,}000 \leq T \leq 600{,}000$ K) and thus study this important interface between the chromosphere and corona. They are an indispensable complement to Hinode EIS, which observes only coronal temperatures. These capabilities will remain unique. No other UV or EUV spectrometer covering this important temperature range is in development or even planned before the Solar Orbiter mission.

Acceleration of the solar wind. SUMER can measure ion flow speeds at the coronal base and the initial acceleration, and only UVCS can measure supersonic outflow speeds in coronal holes and streamers. UVCS Doppler Dimming observations combined with STEREO and Hinode data will give a comprehensive picture of the fast wind origin, in plumes and interplume regions.

White light corona. We want to understand the initiation of CMEs, their propagation into the heliosphere, and the large-scale structure of the corona through a full magnetic cycle. In order to understand the initiation of CMEs we will need the higher cadence EUV and vector magnetic field information from SDO and the white light coronagraphy from SOHO LASCO. Similarly, we can only hope to reconstruct the propagation of CMEs into the heliosphere accurately using as many viewpoints as possible, for example the two STEREO spacecraft and SOHO. And only LASCO can give a continuous record of the large-scale corona over a full cycle from one viewpoint.

Coulomb drag and solar wind abundances. Using SOHO CELIAS and ACE/SWICS data, the CELIAS team determined the absolute abundances in the slow solar wind during solar maximum and solar minimum. Surprisingly, they found a solar cycle effect, with a depletion of heavier elements at solar minimum that appears to be well organized by the Coulomb drag factor (but not by e.g., mass, or mass-per-charge). Confirmation of this trend through cycle 24 would establish the importance of Coulomb drag in the low-speed solar wind.

3D solar UV irradiance. SOHO SWAN full-sky images are used to derive the 3D distribution of the solar UV flux. From these maps and using the Sun's rotation, it is possible to predict the UV flux received by the Earth two weeks in advance. It is also possible to compute the UV flux emitted toward any planet or object in the solar system. These values are produced on a regular basis and distributed through the SWAN web page. Further, SWAN data are used in a project that aims at integrating the SWAN farside imaging technique and the MDI/GONG farside helioseismology results to recreate the solar activity pattern at any time and any point on the Sun. Such a tool will be used to predict the values of the various activity indices used in space weather. One application is the prediction of the thermospheric temperature, which is the main parameter used to compute the drag effect on satellites on low earth orbit. This project is funded until 2012.

Hydrogen ionization rates in the heliosphere. SOHO SWAN observations are used to derive the three-dimensional distribution of the interplanetary hydrogen ionization rate and the solar wind mass flux. The results will be used to support the analysis of the data from the Interstellar Boundary Explorer (IBEX; launch in late 2008), which will study the distribution function of energetic neutral atoms (ENAs) produced in the heliospheric interface region and coming back toward the Sun. For the success of this study, it is necessary to compute the transmission function of ENA's between the production region (80-100 AU) and the observer (1 AU). This transmission function is derived directly from the hydrogen ionization rate determined by SWAN.

Energetic neutral atoms. ENAs of hydrogen and helium are measured by SOHO CELIAS in the energy range ~60 - 90 keV and will thus provide an extension of the lower energy IBEX measurements (0.01 – 10 keV) to higher energies. This will provide additional constraints for theories of the dominant sources of ENAs and for the modeling parameters of heliospheric plasma simulations.

SEP seed particles. Any successful theory of solar energetic particle (SEP) production by CME shocks must account for the large observed variations in SEP spectral characteristics and elemental abundances. It has been proposed that this variability arises from an inherently variable population of suprathermal seed particles, which have never been measured, though. UVCS has been demonstrated to have the sensitivity to distinguish between a Maxwellian velocity distribution of the coronal protons and a suprathermal distribution. Between 2005 and 2007, though, no significant departures from a Gaussian distribution were seen by UVCS. This is not surprising at solar minimum, where there have been very few SEP events. The situation should be different during the rise of solar activity to the next solar maximum. The direct observation of the presence and distribution of a coronal suprathermal seed population would enable an entirely new SEP predictive capability.

Solar Storm Forecasting. Sudden increases in the fluence of >30 MeV protons in SEP events pose a hazard to human space activities and robotic space missions. A new method, based on SOHO COSTEP measurements of relativistic (150 keV - 10 MeV) electrons, permits for the first time up to an hour of warning for the later arriving protons in SEP events. The electrons act as test particles by probing the continuously changing heliospheric transport conditions in the same region of the heliosphere through which the slower-moving protons have to propagate. The new method was for the first time tested under operational conditions during the February 2008 Atlantis mission, which transported ESA's Columbus laboratory to the International Space Station. NASA plans to use this method routinely in the future. Currently, only COSTEP can provide the required measurements. In the space weather communities there is a recognition that SOHO's space weather capabilities are international resources that should continue uninterruptedly, even now with STEREO and after the launch of SDO.

2.4 Issues of complementarities and uniqueness

SOHO, with its unique and comprehensive complement of instruments, has played a central role in the "Heliophysics Great Observatory" and it will continue to do so, even after the launch of SDO. Currently, there are four other solar remote sensing missions in orbit, each complementary to SOHO:

i. TRACE, launched in April 1998 and often considerd the 13th instrument on SOHO because of its tight integration in SOHO science operations provides high resolution EUV images complementary to SOHO's UV and EUV spectrometers. It will be operated until the launch of SDO.
ii. RHESSI, launched in February 2002, is a high-energy mission to study solar flares and particle acceleration using X-ray and γ-ray spectroscopy. CDS and SUMER spectral diagnostics combined with RHESSI high-energy measurements provides unique diagnostic capabilities to study solar flare dynamics.
iii. Hinode, launched in September 2006, carries a 50 cm optical telescope for high-

resolution imaging and magnetometry, a full disk soft X-ray telescope complementing and extending EIT imaging to higher temperatures, and an EUV Imaging Spectrometer (EIS). While EIS supersedes CDS to some degree, CDS can access lines from several more species and at lower temperatures. Unlike EIS, CDS and SUMER can produce high cadence time series with excellent pointing stability over long time intervals, and its selection of targets on the disk and low corona is fully flexible while EIS must observe wherever Hinode is pointed. Only CDS and SUMER can currently provide measurements of Doppler velocities of the cool (log $T_e \sim 5.5 - 6.0$) filament material associated with nearly all CMEs.

iv. STEREO, launched in October 2006, carries particle and plasma instruments as well as an EUV imaging/coronagraph suite similar to the one of SOHO. Its orbit is very different, though, with one spacecraft drifting ahead of Earth, the other following it. The separation between the two STEREO spacecraft is increasing by 45° per year. In 2010 they will be a ±90° from the sun-earth line, from where SOHO will provide complementary information as STEREO's "3rd eye".

While SDO will replace some of SOHO's remote sensing measurements (see Section 4), it lacks both a coronagraph as well as a spectrometer. CDS and SUMER will be the only available source of spectroscopic characterization of the underlying, cool plasma involved in eruptive events observed with STEREO and SDO during the rise of the new solar cycle. Further, SDO does not carry any in-situ instruments. The combination of in-situ instruments on STEREO and SOHO ERNE, COSTEP and CELIAS provides unique science opportunities to measure the composition and energy spectra of energetic particle events from multiple locations, which should narrow the range of acceleration geometries.

GOLF and VIRGO global solar oscillation measurements will remain unique. There are no instruments in planning that would continue these invaluable long-time records of the Sun's interior structure. SWAN data will also remain unique. There is no other instrument in orbit or planned that would provide similar data. In the future, SWAN data will offer the unique opportunity to correct IBEX data for effects in the inner heliosphere.

Finally, there is no other ultraviolet coronagraph spectrometer around (or in planning). SOHO UVCS will therefore be the only instrument that can measure key plasma properties in the acceleration of the solar wind and in CMEs. Those measurements are essential to testing models of coronal heating, solar wind acceleration, SEP production in CMEs, and CME evolution. Combined observations from UVCS and LASCO on SOHO, RHESSI, Hinode, STEREO, and SDO should give a much more complete picture of plasma heating, ion acceleration, and CME shocks than was possible in the same phase of the previous solar cycle.

3. Health of Spacecraft and Instruments

All spacecraft equipment is running on the main branch, and the gyroless and intermittent recording software patches perform flawlessly. Solar array degradation so far amounts to only 19.7%, which corresponds to 1.6% per year (budget was 4% per year). This ensures sufficient power for at least another 6 years. The fuel reserves (117 kg) are also more than comfortable. The average use was less than 1 kg per year in recent years. Nothing indicates that the spacecraft cannot support an extension through December 2012.

Almost all of the 12 instruments on SOHO are operating at full scientific capability, and can be expected to do so for the next years (for details see Annex A). On the instrument side, the only significant change since the previous request for extension (ESA/SPC(2006)28) is a further degradation of the two UVCS detectors and the loss of one of the two redundant SUMER detectors. The progressive, expected fall-off in throughput of some of the other instruments is not worrying, and should not prevent SOHO from returning unique science data for many more years.

4. Changes to Ops/SciOps concept/management

Payload. SDO, scheduled for launch in mid 2009, carries a vastly improved version of SOHO's MDI and EIT instruments (plus an EUV irradiance monitor). After the completion of the cross-calibration of MDI and EIT on SOHO with their new counterparts on SDO (foreseen for mid 2010), MDI will be hibernated and EIT will be used for no more than a few synoptic images per day (mainly for comparing long-term response changes with that of the EUV telescopes on SDO). The other 10 instruments will continue to provide unique data and will be operated in their current mode.

Ground system and spacecraft. Starting in late 2006, in response to budget pressures at NASA, SOHO engineers began an in-house reengineering effort to automate SOHO mission operations. This required developments of new ground software (pass generator, anomaly detection and notification) as well as modifications of the Central On-Board Software. Following three months of engineering trials and testing, the SOHO Flight Operations Team (FOT) began the routine automation of all GSFC-local nighttime DSN contacts on October 8, 2007. The plan is to proceed to full automation of all contacts by the end of September 2008.

Cost saving efforts. We appreciate the fact that there are enormous budget pressures on the ESA science programme and that SOHO is an older mission. At this stage of the mission we therefore accept an increased risk in order to reduce mission operations cost. Starting in 2010, engineering support for SOHO will be implemented remotely from Europe. The current resident ESA technical support team manager will not be replaced after his retirement in 2010. He will continue to provide his expertise on a part-time contract basis. The resident EADS Astrium engineer will also return to Europe at the end of 2009. He will support SOHO from Toulouse, with occasional visits to GSFC for maneuver support and other contingencies. Science operations will also be simplified, alleviating the requirements for an ESA Science Operations Coordinator (whose contract will expire at the end of January 2010).

5. Status and Prognosis of National Funding

For CDS, the currently agreed STFC support will continue until December 2009. A further extension will be applied for in 2009. All other teams expect continued funding at the current level, which is sufficient for safe operation of their instruments, data validation and archiving.

6. Finances

With above cost-saving efforts and additional savings until 2010 we will be able to extend SOHO operations for an additional 3 years from 1 January 2010 to 31 December 2012 with an increase of 1.986 M€ in the CAC, i.e. less than 700 k€ per year. This includes provisions for final archiving of the scientific data.

7. Conclusions

SOHO has revolutionized solar physics in the last 12 ½ years and will continue to make unique and critically important contributions to the "Heliophysics Great Observatory" during the onset and rise of the next solar cycle. The additional cost to ESA is very small and represents excellent value-for-money in return for a significant enhancement of the scientific harvest from the SOHO mission.

Annex A: SOHO Instrument Status

GOLF

- Operating nominally, with data continuity ~98% outside SOHO 1998-1999 "vacation" periods, including no losses during telemetry "keyholes"
- Overall throughput down by a factor of <7 since launch, but:
    - largest noise source is the Sun itself, so negligible adverse effect over most of the frequency range, including that in which the g-modes are expected
    - significant reduction in signal to total noise ratio in a region around 1 mHz
- No reason to doubt that GOLF can continue to function in its present mode for several years
- Complete redundant channel still available, though unused since initial, on-orbit commissioning

VIRGO

- All VIRGO instruments (the two types of radiometers: PMO6V and DIARAD, the filter radiometers SPM, and the luminosity oscillation imager LOI), are fully operational and performing properly. The degradation of sensitivity is still relatively small and all the instruments are still able to achieve the same accuracy and precision as at launch.

MDI

- ~90,000,000 images; after on-board computations, ~15,000,000 raw data images downlinked
- Expected degradation in total light throughput due to changes in the front window; compensated via increased exposure time.
    - mean annual degradation: 4%, appears constant
- Electronics anomaly in 2007 was corrected by instrument power cycle
- Exposure time uniformity: sudden drop in 2000 March, from a part in 12000 to a part in 4000
    - affects helioseismology only for $l < 4$
    - adds some noise to zero point of photospheric magnetic field measurements; correctable
    - No variations above the one part in 4000 level since 2002 February reduction in optics package temperature
    - No detected change in the CCD flat field except for variations with focus change
- The drift in central wavelength of the Michelson's has nearly stopped
- The drift in best focus position has moved the nominal focus setting back almost to the design point. Shortly after launch it was at the limit of the adjustment range.
    - This drift has also apparently slowed
- In summary, no known limit to MDI's useful life

SUMER

- Pointing mechanism has worked flawlessly during recent years and restrictions on pointing has been released, since detector lifetime is regarded as the most critical resource
- Detector A not used anymore because of MCP anode electronics anomaly
- (redundant) Detector B fully operational and is expected to remain radiometrically calibrated for another 2-3 years, based on extrapolation from past performance

CDS

- GIS nominal; no recalibration or changes to high voltages have been necessary in the past 3 years.
- NIS nominal; microchannel plate current anomaly in 2005 July appears to have self-healed after a series of tests and is being used for regular observations again; sensitivity in short wavelength channel 40 - 80% of pre-launch levels; expect drop to no worse than 20 - 60% over the next 5 years
- Electronics nominal; trending shows no aging of components
- Mechanisms: Some 'stickyness' when rastering the GIS slits necessitated a small restriction on the range of movements. This has now been compensated for by improved ground planning software that moves the allowed range of movements to outside of the restricted area. This issue no longer impacts science. All other mechanisms continue to operate nominally.
- Thermal: As with all other components of SOHO, the sunward side of CDS shows a secular increase in temperature, but analysis of the science data shows that the NIS wavelength calibration remains within tolerances.
- Onboard software: No issues

EIT

- EIT is nominal
- Instrument throughput decrease stopped and reversed since 2003 (see: http://umbra.nascom.nasa.gov/eit/eit_guide/euv_degradation.html)
    - Reversal due to long bakeouts occasioned by telemetry keyholes
    - CCE loss can be tracked accurately with calibration lamp images
    - Degradation now understood and modeled
    - Present exposure times range from 12 s (195Å) to 2 min (284Å): lots of latitude left
    - Current throughput at 195 Å is comparable to that in mid-1999

UVCS

- O VI detector: In late December 2005 the O VI detector began exhibiting an anomaly in its spatial imaging direction similar to what SUMER began to experience earlier: there are groups of adjacent dark rows, with the “missing” counts going into adjacent bright rows. The most likely cause of the anomaly is radiation damage of the analog to digital converter integrated circuit of the XDL detector. The spectral directions functions normally, i.e. the quality of the spectra is not affected. The first 87 detector rows provide a 10 arcminute field of view that functions normally and is used for high spatial resolution observations. Above detector row 87, the combined counts of adjacent 64 pixel rows provide 7.3 arcminute resolution bins, which are used for observations that can be done with that spatial resolution.
- Ly-α detector: This detector is experiencing a problem similar to that of the OVI detector. As of this writing, the Ly-alpha detector has four bright rows: numbers are 87, 151, 215, and 279. The dark regions extend from 88 through 150, 152 through 214, 216 through 258, and 280 through 293. It is being used primarily for supra-thermal seed particle population measurements. The rows below row 87 behave normally, are expected to continue to do so, and provide a 10 arcminute field of view for high resolution observations. The higher numbered rows that continue to function normally (i.e., rows outside the dark regions) can be used as well. The practical impact of the new problem is that the field of view of UVCS Ly-α observations with the Ly-α detector has been reduced to 10 arcminutes from 40 arcminutes.

- Visible light detector: This detector system experienced an anomaly in its housekeeping data telemetry systems and was turned off in April 2004. Since its principal function of verifying the LASCO electron density measurements and co-registration has been accomplished, the risk of further operation is not justified.
- Mechanisms: All mechanisms continue to behave nominally except for the Ly-alpha grating drive, which is slow to respond when commanded. This behavior has not prevented this channel from being used for high priority science.
- Radiometric Response: Changes in system radiometric responsivity are being accurately tracked using observations of stars. The changes depend on the observed heliographic height, which deternines the unvignetted aperture. For example, the current in flight calibration is within 50% of the laboratory calibration for observations at 2.5 solar radii. The changes are believed to be caused by telescope mirror reflectance changes.

LASCO

- Thernisien et al. (2005) have performed a detailed analysis of the intensity of a set of about 50 moderately bright stars that transited the C3 field of view
  - These 50 stars generated about 5000 observations during the lower cadence in the first three years of SOHO operations and about 15000 observations thereafter
  - All stars have spectra well known from 13-color photometry
  - Using these stellar spectra as standards and the observed LASCO count rates, the photometric calibration factors of the C3 coronagraph was derived for all five color filters with an absolute precision of ~7%
  - Decrease in the instrument sensitivity found to be only ~3.5% over the 8 years studied or < 0.5% per year
  - C2 response changes similar
- The Fabry Pérot interferometer in the C1 coronagraph did not survive the extreme cold the instrument experienced (-80°C) during the 1998 SOHO off-pointing

CELIAS

- MTOF/PM, STOF/HSTOF, SEM nominal
  - MTOF, PM efficiency degradation of 2 (Fe) to 5 (H); still extremely high signal-to-noise
  - STOF performance stable, MC degradation compensated for by increase in HV
- CTOF impaired since 1996 October (HV power supply hardware failure)

COSTEP

COSTEP consists of two sensors, the Low-Energy Ion and Electron Instrument (LION), and the Electron, Proton, and Helium Instrument (EPHIN). Both instruments have suffered some degradation but continue to generate valuable scientific data and fulfill their scientific goals.

- LION: Unexpectedly high noise level in the LION detectors since shortly after launch have resulted in the loss of the lowest energy channels (< 80 keV). In the course of the mission, three of the four LION sensor heads developed disturbances, some of which can be mitigated by careful data analysis. The disturbed periods are well documented in the level-2 data specification document. As of January 2008 one sensor head is still functioning nominally.
- EPHIN: Detector E of the EPHIN instrument showed steadily increasing noise levels throughout 1996, and had to be switched off (on October 31, 1996) to guarantee reliable measurements with the instrument. By changing the instrument

configuration, the EPHIN measurements can still be achieved with slightly degraded energy resolution in a limited energy range (3-10 MeV for electrons and 25-41 MeV/n for ions). This detector behavior causes no significant degradation of the scientific goals of EPHIN.

ERNE

- Secular increase in temperatures at front of spacecraft has caused increased detector leakage currents. Including radiation effects, the increase during the last five years has been roughly 20 % per year.
- One of the detector channels of the topmost ERNE/HED detector layer malfunctioned on 2000 November 21. Updated onboard software accounts for this issue: the geometrical acceptance (view cone) of the detector is unaffected, as is the measurement of the heavy nuclei (carbon and heavier). Also the light nuclei are unaffected up to an energy of ~ 20 MeV/n. Between 20 MeV/n and 120 MeV/n (maximum energy measured by ERNE), both the coordinate and energy values of the affected detector become increasingly unreliable. This, however, has no effect on particle identification and produces only marginal statistical fluctuation on the total energy of these particles that deposit most of their energies in the lower detector layers.

SWAN

- Instrument status unchanged since 2001
    - All four motors nominal
    - +Z hydrogen absorption cell nominal; -Z cell empty: no absorption when activated (loss occurred in 2001)
    - Both sensors calibrated using HST STIS reference spectra: +Z sensor response constant since 1998 (outside of adjustments for high voltage [HV] setting), -Z sensor response shows decline of ~10% per year. HV setting changed to compensate as much as possible

Annex B: SOHO Mission Archiving Plan

Executive Summary

With some minor exceptions, *SOHO* active archiving is robust: level-0 or -0.5 data are publicly available within minutes to hours of their ground receipt, level-1 and higher products within a few months to a year, and documentation and calibration databases are also publicly available. The existence of multiple copies of the archive with public Internet access insures the survival of the current state of the archive, and facilities in both Europe and the US used for active archiving will serve as resident archives after the mission ends. Instrument resource Web pages direct potential users to publications, calibration information, software descriptions, and user's guides, as well as to a variety of methods for accessing the data. All *SOHO* data are accessible *via* the Virtual Solar Observatory.

This document describes in somewhat more detail the current status of the archiving efforts and documentation by instrument, as well as descriptions of the expected contents of the final archive and the effort needed to produce that archive. We also discuss, where appropriate, the tools commonly used for the analysis of the instruments' data. Finally, we describe the likely shape of the resident archives that will serve *SOHO* data, documentation, and analysis tools after the end of the mission.

I. Existing Data Products and Documentation

I.A Online Resources

I.A.1 Websites. The *SOHO* Website, http://soho.nascom.nasa.gov/, contains extensive information on the *SOHO* mission, instruments, spacecraft ephemeris and attitude, and access to both the *SOHO* mission archive (less MDI helioseismology observations) and the *SOHO* publications database. The MDI helioseismology are available at the Stanford Solar Center: http://soi.stanford.edu/data/data_request.html .

I.A.2 Instrument resources pages. Each *SOHO* instrument Principal Investigator (PI) team has a Webpage with links to standardized resources, including, but not necessarily limited to: published articles in the refereed literature describing the instrument; initial results; operational constraints; data file format description; metadata; reformatting levels; algorithms for reading the data files; recommended data access and analysis software; software and databases for calibrating lower-level data products; a user's guide to the instrument and data; and PI team contact information. Links to all of these instrument resources pages can be found at: http://soho.nascom.nasa.gov/data/archive/instruments.html .

I.A.3 Data access methods. The *SOHO* "data/archive" access page (http://soho.nascom.nasa.gov/data/data.html) contains links not only to a search facility for the mission archive at NASA Goddard, but also to: ancillary data such as attitude and ephemeris information, level-0 scientific telemetry since 2002, near realtime images and movies, and related information on space weather, comets, &c. The *SOHO* archive access page, http://sohodata.nascom.nasa.gov/cgi-bin/gui , uses fill-in forms and scrolling menus to establish metadata search parameters. Alternatively, all *SOHO* remote-sensing data, including those at Stanford, are also accessible *via* the Virtual Solar Observatory (VSO), at http://virtualsolar.org/, through any of the multiple access methods (physical observable, data source, data provider, &c.) offered by the VSO.

I.A.4. Mission data products. Attitude, orbit, and related ancillary files are available on the *SOHO* Website, with an explanation of contents at:
http://sohowww.nascom.nasa.gov/data/ancillary/

I.A.5 Existing data products. The following tables describe the data products for each *SOHO* experiment currently held in the *SOHO* archive at GSFC and the MDI archive at Stanford, as well as the mission data products.

Data product
Uncalibrated, high time resolution data per photomultiplier channel (level-1)
Three versions of level-2 calibrated data (whole Sun average Doppler velocity in m/s), calibrated using three different methods

*Table I-1. Current **GOLF** data products served by the SOHO archive.*

Data product
Level 1 high resolution data organized per detector
1 minute cadence time series of calibrated VIRGO total solar irradiance deduced from PMO6V data
1 hour cadence time series of calibrated total solar irradiance, PMO6V and DIARAD corrected
1 day cadence time series of calibrated total solar irradiance, PMO6V and DIARAD corrected, with DIARAD IRMB level 2 data
1 minute cadence time series of all three SPM channels, calibrated
1 hour cadence time series of all three SPM channels, calibrated
1 day cadence time series of all three SPM channels, calibrated
LOI 12 pixel images filtered with a 2 day triangular filter

*Table I-2. Current **VIRGO** data products served by the SOHO archive.*

Data product	Format
Original telemetry as received by the SSSC	SFDU
Level-0 MDI image data, organized by Data Product Code (DPC), the onboard-generated data tags describing unique instrument configuration and observing sequence (649 different possibilities)	FITS
Level-1.0 data which is calibrated data in units of m/s, gauss, etc., stored in datasets of (usually) hourly data organized by DPC. (332 distinct Level-1 data series)	FITS
Level 1.4/1.5 data which are collected into datasets named by observables, which may have contributions from multiple DPCs. The 1.4 data are in telemetry order of cropped images, while the 1.5 data are two-dimensional image arrays, reformatted from level 1.4 data if necessary. The 1.5 and higher data do not require special tables to use. There are 22 Level-1.4 series and 86 level-1.5.	FITS
Level 1.7/1.8 data which are "recalibrated," best available data created on demand from the Level-1.4/1.5 data. Level-1.8 data is the lowest processing level that is recommended for use for most science data analysis, and include the latest information on image scale, image center position, magnetic zero level, &c. Level 1.7/1.8 data are usually recalibrated on the fly at time of export. Final recalibrated values will be used when the data are ingested into a resident archive.	FITS
Level-2 data result from further standard processing for particular purposes. These data include products such as tracked data cubes of time-series of *e.g.* 15-degree regions in heliographic coordinates followed for 1664 minutes. These data are used for input to "time-distance" and "ring" local helioseismology analysis. There are 120 Level-2 products	FITS
SHT data, Spherical Harmonic Transform data: projections of Dopplergrams onto spherical harmonics. They are used for global helioseismology. There are 45 types, with varying ranges of mode degree-*l*.	FITS
LOI 12 pixel images filtered with a 2 day triangular filter	ASCII

*Table I-3a. **MDI** data products served by the SOI Science Support Center (SSSC) at Stanford University, as of mid-2007. The number may increase during the mission. A much smaller subset of the most frequently requested MDI data is also served by the SOHO archive at NASA Goddard (see following table).*

Data product	Format
Calibrated full solar disk 1024x1024 resolution 1 minute cadence magnetograms (Gauss)	FITS
Calibrated full solar disk 1024x1024 resolution 96 minute cadence average magnetograms (Gauss). (FITS)	FITS
Full solar disk 1024x1024 resolution 6 hour continuum images in arbitrary intensity units. (FITS)	FITS

*Table I-3b. Current **MDI** data products served by the SOHO archive; the Stanford Helioseismology Archive also makes Dopplergrams available online and higher-level products available nearline (online requestable).*

Data product	Format
Uncalibrated solar spectra organized by study	FITS
Calibration files, software, and documentation	*various*

*Table I-4. Current **SUMER** data products served by the SOHO archive.*

Data product	Format
Uncalibrated solar spectra organized by raster	FITS
Calibration files, software, and documentation	*various*

*Table I-5. Current **CDS** data products served by the SOHO archive.*

Data product	Format
Uncalibrated, full-disk (and a small number of subfield) images in 171, 195, 284, or 304 Å	FITS
Flat fielding, degridding, and calibration files Software, and documentation	FITS *various*

*Table I-6. Current **EIT** data products served by the SOHO archive.*

Data product	Format
Uncalibrated UV spectra organized per XDL detector (each file holds multiple individual exposures)	FITS
Uncalibrated visible light data (counts from the photomultiplier tube)	FITS
Calibration files Software and documentation	FITS; various

*Table I-7. Current **UVCS** data products served by the SOHO archive.*

Data product	Format
Uncalibrated Fe XIV and other forbidden line C1 inner coronal images (1996 - 1998); “level 0.5”	FITS
Uncalibrated white light, 1024x1024, coronal images grouped per coronagraph (C2 and C3); “level 0.5”	FITS
1024x1024 resolution white light coronal images calibrated to mean solar brightness, grouped per detector (C2 and C3); level 1	FITS

*Table I-8. Current **LASCO** data products served by the SOHO archive.*

Data product	Format
CTOF sensor rates: Start rate, Double Coincident rate, Triple Coincident rate, Solar wind speed, Proton rate, Helium rate (rates in Hz)	CDF
CTOF matrix elements	CDF
CTOF high resolution matrix rates, including velocity of solar wind in m/s	CDF
CTOF pulse height analyzer (raw energy spectra)	CDF
HSTOF matrix elements	CDF
HSTOF high resolution matrix rates	CDF
MTOF sensor rates (Front SEDA rate, neutral stop rate, ion stop rate, neutral double coincidence, ion double coincidence, ion start rate, multiple front SEDA rate, multiple double coincidence rate, neutral rates)	CDF

Data product	Format
MTOF pulse height analyzer: Neutral/Ion identification, amplitudes, time of flight	CDF
MTOF time of flight spectrum (far and near sides of MCP)	CDF
STOF high basic rate (including both STOF and HSTOF rates)	CDF
STOF low basic rate (including both STOF and HSTOF rates)	CDF
STOF matrix elements	CDF
STOF high resolution matrix rates	CDF
STOF pulse height analyzer (raw energy spectra)	CDF
PM radial spectra: solar wind speed and alpha (counts per 100 seconds)	CDF
PM theta array (counts per 100 seconds)	CDF
PM total rates (Hz for each step)	CDF
M radial-theta array (counts per 10 minutes)	CDF
SEM (photon counts)	CDF
Proton Monitor 5 minute averages (p speed, density, thermal speed, arrival direction, predicted He speed)	ASCII
Coronal mass ejection data (solar wind speed, kinetic velocity, density, position angle, ion densities)	CDF
Energetic neutral hydrogen atom fluxes (55 - 80 keV, with standard deviations)	CDF
Solar wind speed (for O, Si, Fe)	CDF
Solar wind densities (O, O^{+6}, O^{+7})	CDF
Fe freeze-in temperatures and mean charge states	CDF
Fe densities by charge state (from Fe^{+7} to Fe^{+16})	CDF
Si densities (for Si^{+7}, Si^{+8}, Si^{+9})	CDF
S^{+7} density	CDF

*Table I-9. Current **CELIAS** data products served by the SOHO archive.*

Data product	Format
Proton, deuterium, ^{3}He, ^{4}He: - 1 minute EPHIN counting rates given as intensities (in 1/cm^{2}/s/sr/MeV or 1/cm^{2}/s/sr/MeV/nucleon; 3 energy bands) - 1 minute EPHIN pulse height analysis (energy deposits in MeV) - 1 minute EPHIN rate correction (in counts). (ASCII)	ASCII
Proton, He: - 15 second LION counting rates given as intensities (units 1/cm^{2}/s/sr/MeV) in 3 energy bands for protons, 1 for He)	ASCII

*Table I-10. Current **COSTEP** data products served by the SOHO archive.*

Data product	Format
LED onboard analyzed counting rates (proton and He intensities 1/cm^{2}/s/sr/MeV, 10 energy bands)	ASCII
LED pulse height data (MeV)	ASCII
HED onboard analyzed counting rate (proton and He intensities 1/cm^{2}/s/sr/MeV, 10 energy bands)	ASCII
HED pulse height data (MeV)	ASCII

*Table I-11. Current **ERNE** data products served by the SOHO archive.*

Data product	Format
Uncalibrated full sky Ly a synoptic maps in ecliptic coordinates (one every three days: full time resolution)	FITS
Uncalibrated data organized per target of opportunity	FITS
Calibration files, software, and documentation	various

*Table I-12. Current **SWAN** data products served by the SOHO archive.*

Data product	Format
Spacecraft attitude files	CDF, FITS, ASCII, and SFDU
Spacecraft orbit files	CDF, FITS, ASCII, and SFDU
Time correlation log (onboard clock history)	ASCII
Daily report (from Flight Operations Team)	ASCII
Command history reports	ASCII
Spacecraft mode files	ASCII

*Table I-13. Current **Mission** data products served by the SOHO archive, available through the SOHO Website and, for FITS versions of attitude and orbit files, readable with SolarSoft routines.*

II. The Transition to Resident Archives

In the following pages, each PI team describes their archiving plan. Each should be achieved by no later than the end of FY10, and in many cases, well before. We assume that both the *SOHO* archive at the Solar Data Analysis Center (SDAC) at NASA Goddard and its European counterpart at the ESA Space Astronomy Centre (ESAC) in Spain will contain all of the final data products except for the bulk of MDI data which, as described below, will be ingested into the online data system being built for the SDO HMI database. All of these facilities have high-bandwith connections to the Internet and relatively stable funding with a horizon of at least a few years, corresponding well to the expectations for resident archives in the NASA Heliophysics Science Data Management Policy.

Data Access. Data will continue to be served *via* the "classic" *SOHO* archive interface, http://sohodata.nascom.nasa.gov/cgi-bin/gui and the Virtual Solar Observatory (VSO), http://virtualsolar.org . We will work with the Virtual Heliospheric Observatory to enable their service of as many heliospheric data sets directly as possible; in any case, all the *SOHO* data will be available *via* the VSO through translation of SPASE-based queries to VSO-native queries.

Metadata. Most of the *SOHO* remote sensing data are stored in FITS files which have extensive metadata available in their headers; it is in part upon those metadata that the VSO has based its data dictionary http://vso.stanford.edu/datamodel.html . The translation between SPASE and the VSO data model is an ongoing effort of the "Heliophysics Virtual Great Observatory," the consortium of virtual observatory efforts funded by the NASA Heliophysics MO&DA program.

GOLF

At present, we provide for public access daily files containing 24 hours of level-1 data. These are the cleaned-up versions of the series of intensities coming from the instrument. Conversion of these data into level-2 data is a research-based program, for which we have at present three versions. These produce three versions of the level-2 data. Level-2 is a continuous time series of the global solar surface velocity, and each is in the form of one file for the entire duration of GOLF observations to date. It is up-dated every year or so and all three versions are in the public archive. At the end of the mission observations a final update will be completed for at least one of the three versions, destined for the final mission archive. The level-2 data is in a simple form readily useable by the non-instrumentalist. The set of level-1 data will also be placed in the final archive. This allows the possibility for the more expert scientist to develop in the future possible improvements on the calibration method to arrive at level-2.

VIRGO

1. Description of the archived data with references to relevant articles

2. Level-1 data as they are on the VDC as daily files. The parameters are: PMO6V, DIARAD, SPM, LOI, HK in FITS ASCII files (these data are corrected for all known effects and are the basis for all further evaluations)

3. Level-2 data which include corrections for degradation, &.: *(i)* 1-minute data of VIRGO TSI (deduced from PMO6V and corrected as VIRGO TSI hourly and daily data). This will be a binary FITS file; *(ii)* hourly data of TSI (with records on VIRGO agreed, PMO6V and DIARAD corrected, PMO6V and DIARAD level 1.8 (what this means is in the header and will be explained in the general information). This file corresponds to virgo_tsi_h_v6_001_0801.dat and will be given as an ASACII file; *(iii)* daily data of TSI (with records on VIRGO agreed, PMO6V and DIARAD corrected, PMO6V and DIARAD level 1.8 (this will be explained in the general information) and the DIARAD IRMB level 2 data, which are their product and could be left out - I agree, they are wrong anyway). This file corresponds to virgo_tsi_d_v6_001_0801.dat and will be given as an ASCII file; *(iv)* 1-minute, hourly and daily values of the 3 SPM channels. The 1-minute file in FITS format, the daily and hourly as ascii files.; *(v)* LOI data high-pass filtered with a 2-day triangular filter as FITS file (this is an image, although with little spatial resolution).

4. The data will be made available at the VDC and at PMOD/WRC (radiometry and SPM) and at IAS, Orsay (LOI), as well as at the *SOHO* European (ESAC) and US (Goddard) archives.

MDI

Unlike the other *SOHO* instruments where the primary archive is the *SOHO* archive at GSFC, the primary MDI archive during the mission has been the SSSC (*SOHO* / Stanford Science Support Center) at Stanford University. The bulk of the MDI (and therefore SOHO) data resides at the MDI data center. The current access is via: http://soi.stanford.edu/data/ .

The MDI data center is based on the DSDS (Data Storage and Distribution System) which is implemented in the SSSC. Data stored in the DSDS is in FITS file protocols with metadata also present in ".rdb" format which is tab delimited table data. The DSDS uses Ampex DD-2 19mm 330GByte tapes as its primary offline medium. (A copy of the Level-0 data are stored at the Joint Institute for Laboratory Astrophysics at the University of Colorado.) Metadata consisting primarily of the catalog of datasets is maintained in the DSDS Oracle database. MDI data is organized as "Dataseries" which are sequences of "Datasets". A Dataset is a set of "images" spanning usually a range of time such as 60 one-minute Dopplergrams. A Dataseries is the collection of all Datasets of a given type such as the set of all hours of one-minute Dopplergrams. The data is stored by "Dataset Name" which consists of three parts: Project, Level, and Series and indexed by Series number. The holdings consist of more than 150 TByte of MDI data with a total of more than 200 TByte including ancillary data. The contents of archive at Stanford includes the dataseries listed in Table I-3a, above.

In addition to data files, there are numerous data calibration tables, event lists, command sequences, command logs, &c., which are maintained on the MDI Webpages. While these ancillary products are not strictly needed for use of the Level-1.8 data they are very helpful in resolving issues related to non-normal observing conditions.

In addition to digital data, the MDI team maintains a collection of Technical Documents some of which have online versions. These describe the implementation of the MDI instrument and the DSDS data system as well as early calibration efforts.

The best information on MDI calibration issues resides in the published journals and conference proceedings and Ph.D. dissertations.

The software for operating the DSDS and analysis through the above data products is also available through DSDS web services.

Additionally, some MDI Co-Investigators have developed SolarSoft IDL software which while not part of and not maintained by the MDI team directly is nevertheless used by many investigators to aid in the use of some of the MDI data products.

Additional Holdings in the DSDS

In addition to the MDI data archive, the DSDS serves as a community active repository for related data in the Stanford Helioseismoogy Archive (SHA). Data in the SHA includes data such as partial set of groundbased Global Oscillations Network Group (GONG) data, data from both the 60-foot and 150-foot telescopes at Mt Wilson, and simulated solar data from the MSU program.

Plan for Disposition of the DSDS Archive

The primary plan for the MDI archive is to migrate it into the Solar Dynamics Observatory (SDO) Helioseismic and Magnetic Imager (HMI) and Atmospheric Imaging Array (AIA) Joint Science Operations Center (JSOC) Science Data Processing (SDP) activity. HMI and MDI share a Principal Investigator and many team members. They will ingest the DSDS holdings into the JSOC, which is designed to handle the bulk of the SDO data. HMI and AIA together generate about 1.4 TByte of compressed telemetry data per day. This will expand after decompression, calibration, recompression, &c., to about 2.0 PByte per year. Thus the DSDS holdings will comprise a few percent of the JSOC holdings.

The plan is to ingest DSDS Datasets each as a single record in the JSOC Data Record Management System (DRMS). The DRMS records will be accessed with the original DSDS dataset name. For example, the 6-hour sequence of low-rate Dopplergrams obtained on 1998 June 24 starting at 00:00 UT will move from the DSDS dataset name of:

prog:mdi,level:lev1.8,series:vw_V_06h[8000]

to

dsds.mdi_lev1_8_vw_V_06h[#8000]   also as
dsds.mdi_lev1_8_vw_V_06h[1998.06.24_00:00]

The DSDS software system will be modified to map the old DSDS names into the new ones such that existing MDI processing and analysis programs, web export, &c., will continue to function without change.

In addition to migrating the active MDI archive to the JSOC we will archive “snapshots” of the various web based support information and will maintain a library collection of the paper-only documents from early in the *SOHO* mission.

Enhanced Plan for Migration of MDI data into the JSOC

We have proposed to the current Virtual Observatories AO for support to fully migrate key MDI dataseries into the JSOC DRMS with per-image record access with full metadata support instead of the base per-dataset access. This plan will use the same data files but will take advantage of the DRMS ability to link to data segments from a record in one dataseries to a file in a data segment in a record in another dataseries. Thus only one copy of the data will be needed to allow per-record access to the DSDS holdings.

SUMER

Experience has shown that the continuous variation of the instrument performance and the incremental knowledge about this variation will not allow to achieve the ultimate quality for data products, which have been processed in an automated way. Therefore, SUMER data will be archived in three formats/levels (all stored in FITS files):

(a) As raw counts together with all procedures and calibration files needed to process the spectra and to convert them into physical units. All steps are fully documented.

(b) On an intermediate level (decompressed, reversed, geometrically corrected, flat-fielded using standard S/W tools).

(c) Data sets which allow radiometric calibration in an automated way will also be be provided in physical units. We estimate that this includes upward of 80% of the existing data.

The archive will provide browsing tools to search for data. It is organized in such a way that different search criteria and combinations of those can be applied: wavelength, date, solar region, campaign number, study number, SUMER user, and source code.

Most of the features described here are already implemented.

CDS

Current Status of the CDS Data Products and Software:

- **CDS Level-0 data products**:

CDS FITS files contain uncalibrated science data from the CDS NIS or GIS detector. Each FITS file may contain only one or many (up to several hundred) individual exposures. Level-0 data are stored in the standard FITS format, with metadata and selected housekeeping parameters inserted as the header of each file. The header defines the configuration of the instrument at the time that the data were taken. All up-to-date CDS FITS files are available through the SOHO archive at GSFC.

- **CDS data analysis software**

Large set of IDL-based routines used to read and manipulate Level-0 data. It is used to apply calibrations and corrections to Level-0 data in order to convert it into fully calibrated data in physical units and derive higher-level products such as spectral line intensities, line positions and widths. The CDS software is distributed through the SolarSoft package.

A description of the instrument, data, and software tools is available at:
http://solar.bnsc.rl.ac.uk/software/software.shtml
It includes the CDS User Guide and Software Notes.

- **Other CDS data products**

All CDS calibration files are included as part of the SolarSoft package. These include flat fields, wavelength calibrations, time dependent burn-in corrections, time dependent line-width correction factors for specific emission lines, and photometric calibration factors. All calibration factors, except the line width corrections, are automatically applied by the CDS calibration software.

Auxilliary data products include observing plans, data catalogs, atomic data for calculating emission measures, documentation, and housekeeping data.

CDS level-0 data together with the calibration files and CDS analysis software comprise a fully working package that has been used (with updates) for science analysis of CDS data by the solar community since the launch of SOHO.

Future archiving plan (subject to availability of funding from STFC):

- **CDS Level-1 data products**:

* The CDS level-0 FITS files will be converted to calibrated level-1 data in physical units by using the existing calibration routines, NIS_CALIB and GIS_CALIB. Level-1 filenames will be distinct from the current level-0 filenames to avoid confusion.
* The existing routine NIS_ROTATE will be used to remove small instrumental distortions (the spectral slant and the individual tilts of the spectral lines), so that the spectral and spatial axes are independent of each other.
* Where feasible, cosmic rays will be removed prior to calling NIS_ROTATE, to avoid problems with broadening of the cosmic ray tracks, making them harder to remove in subsequent processing.
* The data will be stored in the same FITS binary table format as is currently used. This format is a FITS standard, and is supported by many FITS software packages such as the FITSIO Fortran and C libraries.
* The current set of FITS keywords will be augmented by standard World Coordinate System keywords, based on Greisen and Calabretta (2002), to describe the axes of the data arrays.

CDS User Guide will be updated with instructions how to read Level-1 FITS files in IDL. The present version of the CDS analysis software will remain available in IDL. CDS analysis software in other programming languages or packages cannot be provided.

Level-1 FITS files will be made available to the ESA Project Scientist team who will be responsible for their inclusion, distribution and access through the SOHO archive. It is understood that Level-0 data will continue being accessible through the SOHO archive in parallel to Level-1 data.

Plans for completing the CDS calibrated data products.

Development and testing of software for converting Level-0 to Level-1 FITS files – by December 2008.

Data processing to convert all level-0 FITS files into calibrated Level-1 FITS files – by December 2009.

Update the CDS User Guide with references to Level-1 data – by December 2009

EIT

Currently, EIT produces "raw," level-0, 16-bit integer image data in FITS files and byte-scaled movies of images and image first differences in IDL save files. Those FITS files are available through the *SOHO* data archive Web interface, through a custom EIT Web interface (http://umbra.nascom.nasa.gov/eit/eit-catalog.html), and *via* the Virtual Solar Observatory. Documentation is available at the EIT instrument resource page, http://umbra.nascom.nasa.gov/eit/eit_resources.html , and both software and calibration database files are available as part of the standard SolarSoft distribution. The SolarSoft routine EIT_PREP reads the level-0 FITS files and outputs floating-point data numbers with flat field, degridding, throughput variation, and vignetting removed. All EIT documentation and top-level software are documented in the EIT User's Guide, http://umbra.nascom.nasa.gov/eit/eit_guide/ ; listings for all EIT SolarSoft routines are available in the User's Guide as well.

Within one year of the end of the *SOHO* mission (that is, at the beginnning of the Bogart mission), we will finalize the calibration database for the mission and produce level-1 FITS files with calibrated (floating-point) DN's. The movie database will be recast in one or more standard formats (*e.g.*FITS and Quicktime). All of the final archive will be served online at the *SOHO* archive at NASA Goddard; at the MEDOC in France; and at the *SOHO* European archive at ESAC (Spain). Additional, higher-level data products, such as synoptic (Carrington) maps may be added to the final archive if they can be validated with a reasonable level of effort. We also note that the original, level-0 FITS files will be part of the final archive as well, in case future users wish to refine the calibration.

UVCS

UVCS Level-1 data products. Spectral Data Files: Contain integrated count data from one XDL detector. Each Exposure Sequence produces one Spectral Data File per detector. Each file may contain only one or many (more than a hundred) individual exposures.

Visible Light Data Files: Contains integrated count data from the White Light Channel photomultiplier tube. Each Exposure Sequence produces one Visible Light Data File.

These Level 1 data products are produced from Level 0 data as the first stage of processing. They contain the (uncalibrated) science data and enough (uncalibrated) housekeeping data to define the configuration of the instrument at the time that the data were taken.

Level 1 data are stored in the standard FITS format, with the above mentioned housekeeping inserted as the "header" of each file.

***UVCS Level-2 data products*.** Fully Calibrated Ultraviolet Science Data Files and Fully Calibrated Visible Light Science Data Files: These Level 2 data products are produced from Level 1 data as the second stage of processing. They contain the best and final fully radiometrically and spectrometrically calibrated and corrected science data and all calibrated housekeeping data that are required to define the configuration of the instrument at the time the data were taken. Level 2 data are stored in the standard FITS format, with the above mentioned housekeeping inserted as the "header" of each file.

Other UVCS data products. Mission Log: Description of each observation. They are created at the time each observation is carried out.

Calibration Files: Contain information required to convert Level 1 data into Level 2 data. There is one file for each year of the mission in order to take care of time variations in the various calibrations.

Auxiliary Files: Set of files containing specialized housekeeping parameters used to perform small corrections to aspects of the instrument setup such as absolute pointing.

Detector background files contain detector background count rates for each pixel of the detector.

UVCS software. Data Analysis Software (DAS): Large IDL based code used to read and manipulate Level 1 and Level 2 data. It is used to apply calibrations and corrections to Level 1 data in order to convert it into Level 2 data.

Line Profile Curve Fitting Codes: Set of specialized IDL codes used to fit various functions to the spectral line data taking into account the instrument profile, stray light, and other effects.

Tutorial: Description of the instrument, data, tools, and how to use them.

Current status of data products and software tools. Visible Light Calibration Files are based on the laboratory measurements and observations of stars and Jupiter. They were finalized in 2002, and are available on the UVCS web site.

Auxiliary Files are available from the SOHO Archive from the beginning of the mission until 2 months prior to the current date.

Detector background files are considered part of the user observational data and are used by the user in analyzing the data. The DAS software provides the means to remove the detector background.

The latest version of the UVCS data analysis software (DAS-40) was released in April 2006. It is available on the UVCS web site. This version added the capability to do time-dependent calibrations as well as correcting the telescope mirror orientation value for crosstalk from the two digital impedance transducers of the grating drives. The mechanism calibrations are current as of May 2005 and are very stable. The wavelength calibration is not final and needs to be updated. The DAS is currently capable of removing detector background.

Line Profile Curve Fitting Codes are currently the responsibility of the user. SAO has line fitting codes that are available on request, but they are not very user friendly. The line fitting codes include a capability for removing the stray light contribution to the UVCS data.

The UVCS tutorial is available on the UVCS web site.

Plans for completing the UVCS data archive products. The first step is to improve the precision of the wavelength calibration as a function of time. This will be done first for beginning of mission through December 2005, and then for times up to June 2008, and finally up to June 2009.

Produce the fully calibrated Science Data Files for beginning of mission to December 2005. (by June 2008)

Update time variations of UVCS characterization and calibration through to the end of the mission (continue to observe stars and analyze the observations).

Produce the fully calibrated Science Data Files for times up to June 2008 by October 2008, and for times up to June 2009 (or the start of the Bogart mission) by October 2009.

Produce the user friendly line profile curve fitting code, which also removes stray light by June 2009.

Continue to provide Ultraviolet Spectral Data Files to the SOHO Archive.

Continue to produce and update the Mission Log.

Schedule for posting the final versions

By mid-June 2010:

Recommend/describe any changes required to the SOHO Archive in order to accommodate the complete UVCS data and tools packages.

Transfer the complete UVCS data and tools packages to the Archive.

Test the SOHO Archive Interface to be sure files are findable and retrievable.

Post references to publications describing the UVCS calibration and characterization.

LASCO

Complete processing of latest data to level-1 and level-2 (brightness, pB, colors). The LASCO images are converted from digital numbers (DN) to physical brightness units (mean solar brightness) using procedures and calibration data that are in SolarSoft. We will validate the calibration data for the entire mission, updating as required. We normally do this check 3-6 months after the final data are collected, but it would be worthwhile to perform a final validation check for the entire mission. This will ensure that the data collected at the end have the same quality and consistency as those collected at the start of the mission. The conversions that we normally perform are (1) the Level-1 conversion to Mean Solar Brightness (MSB) units for the standard images (2) conversion of the polarization triplet images taken at three angles of linear polarization into total brightness, degree of polarization and angle of polarization. We also compute the product of polarization and brightness, pB, but that is not a unique product. The calibration data are available in SolarSoft as well as the routines to perform the calibration.

Provide final documentation for hardware issues, calibration files, software, &c. There are a number of documentation tasks that will be undertaken to improve the understanding of the on-orbit instrument performance and to collect the information into a set of documents. These include:

1. LASCO Photometric Calibration (C2 and C3) over the SOHO mission. This would include (a) the photometric calibrations using stars, (b) the polarizer response (c) the failure of one of the C3 polarizers during the 1998 off-pointing, (d) the color filter transmittance relative to the standard color developed from stars, (e) the stray light, both the radially symmetric part and the non-radial part, (f) the polarized stray light component, (g) the objective vignetting (h) etc.

2. LASCO CCD performance over the SOHO mission for the C1, C2 and C3 CCDs. This would show (a) the evolution of the dark current, (b) the evolution of the electronic bias offset, (c) the evolution of dark and hot pixels, (d) the Charge Transfer Efficiency (CTE) variation, (d) Quantum Efficiency (QE), (e) etc.

3. C1 Calibration over the mission

4. LASCO on-orbit anomalies

5. LASCO Pointing and Roll History

Clean up analysis software and write analysis guide. The LASCO analysis software has evolved considerably over the mission and many of the routines in SolarSoft are either obsolete or not useful anymore. These routines should be updated or deleted. An analysis guide exists, but could be greatly improved upon.

CELIAS

A substantial fraction of CELIAS data is already publicly accessible through the SOHO archive (http://sohowww.nascom.nasa.gov/data/archive). This data set is continually updated from the University of Bern (UBe). The most recent CELIAS data is not currently available through the Archive interface (although the data files were sent by UBe) because of a recently uncovered technical problem, which we are told will be rectified soon. The data available include a coronal mass ejection file, diagnostic files, an energetic neutral hydrogen file, housekeeping information, matrix elements, matrix rates, pulse height information, sensor rates, and time-of-flight spectra. Although some of the above is in raw form, solar wind parameters from the Proton Monitor are in physical units, as are the SEM EUV fluxes and the CTOF data. The CTOF data include Oxygen freeze-in temperatures; speeds for solar wind Oxygen, Silicon, and Iron; O^{+6}, O^{+7}, and total Oxygen densities; Iron freeze-in temperatures and mean charge states; Iron densities by charge state from +7 to +16; and densities of S^{7+}, Si^{7+}, Si^{8+}, and Si^{9+}. The CTOF data cover the entire life of the CTOF instrument.

SEM fluxes are also available via the web in text format at http://www.usc.edu/dept/space_science/semdatafolder/semdownload.htm. Data are available at 15-second resolution by day, 5-minute data by month, 10-minute data by month, and daily averages by year for the full mission. In addition, plots of near-real time SEM are available at http://umtof.umd.edu/sem/ and http://umtof.umd.edu/semflux/.

MTOF will provide solar wind Oxygen, Magnesium, and Iron elemental densities at 5-minute resolution for the entire mission.

We currently routinely provide 5-minute Proton Monitor (PM) data to the Virtual Heliospheric Observatory project for use in their VHO web site http://vho.nasa.gov. The VHO project retrieves (via Rsync) text files containing the most recent 5-minute PM data from the University of Maryland (UMd) on a daily basis. We also provide maximum resolution 30-second data for a full year to VHO manually. The VHO site is currently being updated, so at present many links are not functioning and the PM data are unavailable using their Data Browser.

A Web page at UMd to provides access to the most recent level-1 data: the entire Proton Monitor data set is available at http://L1.umd.edu. This site allows a visitor to view or download 5-minute data for a range of days in any year, or to download full-year text files at either 5- minute or 30-second resolution. In addition, the Proton Monitor home page (http://umtof.umd.edu/pm/) provides access to plots of near-real time PM data, data organized by Carrington rotation, and a list of interplanetary shocks identified by the PM since launch.

Solar wind parameters are derived from the Proton Monitor (PM) rate data via a complex sequence of algorithms and were calibrated partially by comparison with the Wind/SWE Faraday Cup solar wind instrument as described in Ipavich et al (1998). The software modules (in Fortran) used to derive the parameters will be provided, as will a description of the algorithms.

We will also provide the software package used to derive the MTOF elemental densities mentioned above, along with a description of the algorithms, effects and efficiencies that are incorporated. Calibration files obtained pre-launch with MTOF and post-launch with the MTOF spare unit will also be provided.

The PM and MTOF data files will be provided in both ASCII and CDF format to whichever archive site is requested by the project.

References

Ipavich et al., Journal of Geophysical Research, 103, 17205, 1998

COSTEP

EPHIN

Data will be delivered as a set of standard ASCII files containing the following information and file structure:

- Intensity-time profile (naming convention: EPxyydoy.RL2) with X equal h for data from 1995 to 1999 and I for data from 2000 on, yy a two digit *e.g*. 99 and doy a three digit for the day of year *e.g*. 005. The contents of these files is described in detail in the file http://umbra.nascom.nasa.gov/soho/sr08/costep/l2_spec_ephin.doc .
- Correction files (with names such as EPxyydoy.KOR with X equal h for data from 1995 to 1999 and I for data from 2000 on, yy a two digit *e.g.* 99 and doy a three digit for the day of year *e.g.* 005) and contents also described in the same file l2_spec_ephin.
- Pulse Height Analysis: Energy losses in each detector, in files named PHA file: EPxyydoy.PL2 with X equal h for data from 1995 to 1999 and I for data from 2000 on, yy a two digit e.g. 99 and doy a three digit for the day of year e.g. 005 and as described in the same file l2_spec_ephin

We are currently negotiating with Arik Posner to include the Spectrogram-time-profile to the above data set.

LION

Data will be delivered as a set of standard ASCII files described in http://umbra.nascom.nasa.gov/soho/sr08/costep/l2_spec_lion.doc :

- Housekeeping and counting rates file: LIxyydoy.SL2 with X equal o for data from 1995 to 1999 and p for data from 2000 on, yy a two digit *e.g.* 99 and doy a three digit for the day of year *e.g.* 005

All these data will be stored in Kiel in a database, access via a Web interface or anonymous ftp will be given; the best method of replicating these data in the *SOHO* archive is still TBD.

ERNE

The ERNE archived data will be an enhanced version of the current level-2 export data.

Data will be stored in daily ASCII files. Time resolution will be 1 minute, except where indicated otherwise. In the file names 'yy' are the two last digits of the year and 'doy' is three digit running day count for the year (the first day of the year is '001')

The current level-2 export data consists of five types of data files:

1) ERNE status files (ESUyydoy.STS)
2) LED onboard analysed p and He spectra (LEDyydoy.SL2)
3) HED onboard analysed p and He spectra (HEDyydoy.SL2)
4) LED pulse height data (LEDyydoy.PL2)
5) HED pulse height data (HEDyydoy.PL2)

The final archive data will have the following modifications and additions:

a) The highest energy particles (HED/D3-stopped) will be included in the final HED pulse height data HEDyydoy.PL2.

b) New filetypes will be added containing LED and HED heavy particle spectra for selected periods and elements.

c) New file type will be added containing HED anisotropy observations in a form of a single anisotropy index, describing the strength of the observed anisotropy, supplemented with pitch angle distributions from selected periods.

SWAN

SWAN Level 1 data are available in the *SOHO* archive. These files consist of the raw data divided into observing modes. The instrument parameter values have been transformed from FITS digital values into engineering values using the ground calibration tables. These files can be used to compute maps of the Lyman a background using the SWAN IDL software availbale on the SWAN instrument resources page.

SWAN Level 2 data consist of calibrated maps of the Lyman a background. Each map corresponds to a 24-hour observation or less and combines the data of both sensor units. The maps are stored in FITS files. The header defines the start and end times of the observation, the sensor engineering settings (HV values, integration time, scan step values) . The name of the file containing the level 1 data used to compute the map is also given in the header. The map is defined by the coordinate system used (ecliptic heliocentric J2000), the bin size (both axes, usually 1°) and the limits of the image. The maps combine data from both sensors. The cross calibration factor used is given in the header. The background values are given in counts per second. The absolute calibration factor to be applied to the data in counts per second to obtain rayleigh values is also defined in the header, with a version number in case there is a change in the future. These processed data files already exist in IDL save format and are used within the SWAN team. They will be transferred into FITS format.

Data access will follow the standard distribution rules. We expect that the calibrated maps older than one year are available to everyone. For current observations, we will wait a one-year period before making them available outside the SWAN team. Images in JPEG or PNG format are posted on the SWAN web page a few days after the observations and we encourage people who wish to use the calibrated data to contact us directly. Data will be made available upon request even if they have not been delivered to the archive at the time.

III. Long-term archiving

Europe. A new ESA SOHO long term mission archive is being deployed at the European Science Astronomy Centre (ESAC), the site which is ESA's focal point for science operations and data archiving for the missions of ESA's Science Programme. The new archive is based on the technical infrastructure already developed by the ESAC Science Archive Team for astronomy and planetary missions, and has the capability of interfacing with the virtual observatories being deployed around the world. It will expand, but not replace, the services offered by the current archive at Goddard Space Flight Center. The science archives at ESAC are the permanent holding places for ESA's scientific data, including SOHO's, and they are, therefore, long term archives which hold their data in excess of 10 years. The first release of this new SOHO archive is estimated to become public in the fall of 2008.

US. The Heliophysics Science Data Management Policy calls for transitioning mission archives from Resident Archives to "a facility determined in collaboration with the NSSDC" when it is no longer cost effective to retain the data in the Resident Archives. The Policy also notes, however, that the SDAC has become a "center for excellence in providing multi-project, cross-disciplinary access to data and tools to support the broad range of science possible with the "Heliophysics Great Observatory." It is therefore likely that the *SOHO* archive will continue to be served by the SDAC for some years after the end of the mission, while a backup copy of all data, documentation, and software is deposited in a deep archive facility to be determined in consultation with the NSSDC. In this connection, it should be noted that Earth Science missions are already exploring long-term archiving with the National Archives and Records Administration. In addition, we have begun discussions with Google over the possible ingest of at least parts of the *SOHO* archive into the planned Google Research multi-Pbyte archive Palimpsest.

IV. Software

Some of the *in situ* instruments archive plans in Section II, above, as well as MDI, describe software that can be used with their data; most of the other remote sensing instruments have extensive SolarSoft code libraries for data reduction and analysis (*cf.* the SolarSoft "soho" tree). The contents of the *SOHO* branch of the SolarSoft tree is available online at: http://sohowww.nascom.nasa.gov/solarsoft/soho/ . It should be noted that considerable parts of the SolarSoft tree for each instrument are used to "prep" (*e.g.* flat-field, degrid, rectify, and/or calibrate) the raw data stored in most of the current FITS files. When the fully calibrated and corrected files are available along with the level-0 data in the final archive, the various <instrument>_PREP routines will be useful only for testing alternate calibration approaches. The entire point of widely adopted, standard formats such as FITS is to allow calibrated data to be used with anyone's software, so long as it incorporates a FITS reader. In principle, CDF should allow the same facility.

In practice, such generalizability is only possible with good metadata; the *SOHO* FITS files use a standard set of keywords agreed on before launch that still form the basis for similar standards for the VSO, the STEREO mission, &c.