CLUSTER ACTIVE ARCHIVE

USER GUIDE

Prepared by	H. Laakso
reference	CAA-EST-UG-0001
issue	1
revision	0
date of issue	15 May 2014
status	Draft
Document type	UG
Distribution	public

TABLE OF CONTENTS

1. REFERENCE DOCUMENTS

[RD-1] Instrument User Guides, Calibration Reports, and Interface Control Documents. http://caa.estec.esa.int/caa/ug_cr_icd.xml.

[RD-2] Laakso, H. et al. (Eds.), *The Cluster Active Archive - Studying the Earth's Space Plasma Environment*, 489 p, Springer Netherlands, 2010. http://www.springer.com/content/book/978-90-481-3498-4.

[RD-3] Cluster Exchange Format - Data File Syntax, DS-QMW-TN-0010, v2.4. URL: http://caa.estec.esa.int/caa/documentation.xml.

[RD-4] Cluster Metadata Dictionary, CAA-MDD-0001, v3.0, http://caa.estec.esa.int/documents/CAA-MDD-0001_v30.pdf.

[RD-5] Cluster Data Disposition System: Data Delivery Interface Document (CL_ESC_ID_2001-CDDID) http://caa.estec.esa.int/documents/CDDID3-0.pdf.

[RD-6] User Guide to the CSDS. Doc: DS–MPA–TN–0015-CSDS. http://caa.estec.esa.int/caa/ documentation.xml.

[RD-7] Raw Data Network Delivery Interface Document, Doc CAA-EST-ID-0001, v1.2, 28 November 2005. URL: http://caa.estec.esa.int/documents/CAA_EST_ID_0001_V01_2.pdf.

.

2. ACRONYMS

ASPOC	Active Spacecraft Potential Control (Cluster Experiment)
BM	Burst Mode (Cluster Telemetry Mode)
CAA	Cluster Active Archive
CEF	Cluster Exchange Format
CSA	Cluster Science Archive
CIS	Cluster Ion Spectrometry (Cluster Experiment)
CSDS	Cluster Science Data System
DDID	Data Delivery Interface Document
DSN	Deep-Space Network
DWP	Digital Wave Processing Experiment (Cluster Experiment)
EDI	Electron Drift Instrument (Cluster Experiment)
EFW	Electric Field and Waves (Cluster Experiment)
ESA	European Space Agency
ESOC	European Space Operations Centre (in Darmstadt, Germany)
ESTEC	European Space Research and Technology Centre (in Noordwijk, Netherlands)
FGM	Fluxgate Magnetometer (Cluster Experiment)
GSE	Geocentric Solar Ecliptic (coordinate system)
GSM	Geocentric Solar Magnetic (coordinate system)
HK	Housekeeping
ICD	Interface Control Document
ISR	Inverted Spacecraft Reference of frame
JSOC	Joint Science Operations Centre
NFS	Network File System
NIS	Network Information Service
NM	Normal Mode (Cluster Telemetry Mode)
PEACE	Plasma Electron and Current Experiment (Cluster Experiment)
PI	Principal Investigator
QSAS	Queen Mary Science Analysis System
Qtran	Queen Mary File Translator
RAL	Rutherford Appleton Laboratory (in Didcot, UK)
RAPID	Research with Adaptive Particle Imaging Detectors (Cluster Experiment)
RDM	Raw Data Media
STAFF	Spatio-Temporal Analysis of Field Fluctuations (Cluster Experiment)
SWT	Cluster Science Working Team
WBD	Wideband Data (experiment)
WHISPER	Waves of High Frequency and Sounder for Probing of the Electron Density by Relaxation (Cluster Experiment)

3. INTRODUCTION

The purpose of this document is to help the Cluster data users to take the full advantage of the auxiliary datasets of the Cluster mission as well as some science datasets produced by the CAA (Cluster Active Archive). The science datasets produced by the Cluster instruments are mostly explained in great detail in the individual instrument user guides [see RD-1]. The overview of the CAA system is presented in [RD-2], including the introduction to the CAA data format and metadata. Their details are explained in [RD-3] and [RD-4]. The previous user interfaces developed by the CAA have now been adopted by the CSA that provides similar services with a java client. This interface will ultimately become the final Cluster data archive portal once all the developments and science operations are over.

The CAA is a new type of data archiving activity for the Sun-Earth science community which, for the first time, provides this community with access to all calibrated full-resolution measurements from a space mission. The required activities and tasks have been very demanding and it has always been obvious that both the quality of the data and the services of the CAA will significantly improve over time. A good example is that in year 2014 as more knowledge has been gained in the instrument performances, some of the files and datasets will be soon regenerated, including even the files from the first year of Cluster operations, 2001.

3.1 CAA vs CSA

In summer 2010 it was requested that the long-term Cluster Science Archive (CSA) will be placed at ESAC, Madrid, similarly to all other science archived of Science Directorate, and one of the key requirements was that all services developed by the CAA during the past years will be implemented in the CSA. However, in long term it is the most cost effective that this archive is based on the architecture of the existing archives at ESAC which will ease the maintenance and changes of the system after the CAA team is not available.

Therefore a transfer plan and a set of requirements was generated which guided the implementation of the CSA. The architecture of the CSA is different from the CAA but nevertheless the database and services replicate those of the CAA. In 2014 the implementation was completed as most of the services were implemented at the CSA, so the user services to the CAA can be closed to public and the CSA acts as the sole access point to the Cluster data.

It is important to note that the CAA remains active in any other respect, only the user services are provided by the CSA. For instance, the CAA is responsible for the data calibration and generation as well as the maintenance of the metadata dictionary. Note that all data files are first validated and ingested into the CAA database, after which the files will be transferred to the CSA database. Also the CAA keeps developing new tools and services as before. Since the CAA is responsible to most technical documents, including the present one, some parts of the CAA website will remain open.

If there are any issues with the data access, one should contact the CSA team. However, if one has problems with the data content and quality, one can contact the CAA by sending a message to caateam@rssd.esa.int. To facilitate this feedback, the CAA has a user feedback area at http://caa.estec.esa.int/caa/ feedback_intro.xml where users can report about problems with data or send suggestions. All problem reports on the user interfaces should be sent to the CSA team.

4. CAA STANDARDS

The CAA standards form the basis to the formatting and description of the digital data products being handled by the CAA. The CAA datasets are stored in CEF format (Cluster Exchange Format) [RD-3] and uses the CAA Metadata Dictionary [RD-4]. The users are strongly advised to become familiar with the CEF metadata dictionary in order to get the full benefit of the metadata information provided with the numerical data. The CAA data dictionary is available for download in the CAA documentation area.

4.1 CEF format

All calibrated data and most other data as well are stored in the CEF format on the CAA database. There are fast reading routines to read this format. In addition users can also download all data in CDF format if requested. This requires a translation on the fly and this request lasts a bit longer. However, many applications can read CDF format faster than CEF format which may be important particularly for large data files.

The Cluster Exchange Format (CEF) [RD-3] is a self-describing ASCII format consisting of a metadata header followed by data represented as a comma-separated list of values, which provides storage of science products in a robust and easily accessible form. Adoption of such a single common format allows to deliver the data products to users without the need for specific access software. The files in the CEF syntax must have names with the extension .cef to assist identification.

In CEF an exclamation mark is used as a comment marker, and all input to the right of this marker up to the new line character is ignored on input. A header must be included, preceding the data records, and provide sufficient metadata to fully describe the data and their formatting. The file syntax and header content of a CEF file is specified in [RD-4].

Entries in a record are comma separated and white space surrounding delimiters is ignored. Missing parameter values must be padded by a fill value that must be specified in the metadata. All data files are tabular and homogeneous - they have a sequence of records each with the same variables in the same order. Each record is ended by END_OF_RECORD_MARKER, which by default is a new line character (\n) but can be set to another character in order to allow multiline records.

The END_OF_RECORD_MARKER, tab characters, carriage return (\r), new line and white space characters at the start and end of records and surrounding delimiters should be ignored when read. Thus, data may safely be formatted with white space and end of line markers for human (text editor) readability, allowing for easier exchange between platforms where the end of line marker is variously \r\n under DOS, \n under Unix and \r under Mac OS.

For time series the records are ordered on the monotonically increasing time variable, and the time tag must be the first entry in a record. The other entries in the record are associated with that time via the metadata. Time tags are usually centered within the sampling interval, and the spacing between time tags is often the same as that sampling interval. Otherwise, Delta PLUS and Delta MINUS metadata describe the sampling or integration interval corresponding to the data and the location of the time tag within that interval.

4.1.1 CEF time format

Time is represented as a text string in the strict ISO 8601 ASCII calendar date time format: yyyy-mm-ddTHH:MM:ss.wwwZ where the trailing 'Z' is UTC designator. For example, 2001-02-01T01:23:00.000Z corresponds to 01 February 2001 at 01:23:00 UT. This format permits any number of digits after the decimal point in the seconds field, e.g., the example above is a valid time string to millisecond accuracy. This format is robust for arbitrary timing accuracy, but software developers must take care to ensure that their software retains the accuracy.

4.1.2 CEF to CDF translation

Nearly all CAA data files are stored in the CEF format and these datasets can be delivered to the user either in CEF or CDF format. The translation into CDF is done on the fly using Qtran software that is available for download in the CAA software area, and therefore the delivery of files in CDF is slower than that in CEF. However, many applications (e.g. IDL and matlab) can read CDF files faster than CEF files.

4.2 CAA metadata

Apart from quoted text data and variable names, metadata information is case insensitive. The open format also allows for the inclusion of any extra metadata deemed desirable by the generating team. Lines of metadata may be continued using '\' as a continuation marker, following one of the commas separating a list of values. A metadatum consists of a text string of the form "KEYWORD = value". It is recommended that the Cluster data user, whether he/she uses CEF or CDF format, becomes familiar with metadata dictionary as various keywords provide highly useful information.

The top-level metadata provides some general level information about the mission and the experiment:

- MISSION: all mission metadata are the same for all datasets
- OBSERVATORY: there are four OBSERVATORY header files, one for each Cluster spacecraft. These metadata are the same for all datasets from the same spacecraft. There is also an OBSERVATORY header for multi-spacecraft products.
- EXPERIMENT and INSTRUMENT: these metadata provide general information about the experiment and are identical to all the datasets of one experiment on one spacecraft. The instrument level general caveats are given here.
- DATASET: The most important part of metadata is given at the dataset level and is global to the all files of the given dataset. Some of the useful keywords are:

KEYWORD	EXPLANATION
DATASET_ID	unique identifier of the dataset, which is also used in the beginning of the filename. This ID is needed when requesting data with command line tool.
DATASET_TITLE	concise description of the dataset. This name is used to list the dataset on the web GUI.
DATASET_DESCRIPTION	detailed description of the dataset.
TIME_RESOLUTION	time interval between two data points, expressed in seconds. If not constant, the keywords MIN_TIME_RESOLUTION and MAX_TIME_ RESOLUTION must be used.
PROCESSING_LEVEL	level to which the data have been processed. Possible values are: Raw, Uncalibrated, Calibrated, Derived, and Auxiliary
DATASET_CAVEATS	miscellaneous information about the dataset. This may include references to external files.
FILE_CAVEATS	specific caveats of individual files. This may include references to external files.

Each dataset contains a time variable and a number of additional variables. Each variable is described with a number of useful keywords where the most important ones are (for the complete list, see RD-4):

KEYWORD	EXPLANATION
COORDINATE_SYSTEM	acronym of a coordinate system (for vectors, tensors or their components). For allowed acronyms, see [RD-4]
DELTA_PLUS (DELTA_MINUS)	used in particular with time tags to provide the time interval where the data value is valid. It provides a value to be added to (subtracted from) the nominal value to obtain the upper (lower) limit of the parameter interval within which the data was acquired.
FILLVAL	used to replace bad or missing data
FRAME_VELOCITY	This is important to the physical parameters whose values depend upon the motion of the coordinate system in which they are measured, such as the DC electric fields. Three possible values are accepted: Observatory (default); Inertial (i.e. GEI); Earth-Corotating.
LABLAXIS	can be used to label the y-axis of a plot or to provide a column heading for a data listing. For more than one-dimensional parameters, the labels can be given in LABEL_i
QUALITY	provides a quality of the parameter, see RD-4 for the definition
SI_CONVERSION	conversion factor to change the archived value to the value in SI unit, e.g., "1.0E-9> T" is used to transform the magnetic field data stored in nanotesla (nT) into tesla (T).
SIGNIFICANT_DIGITS	provides the number of decimal digits required to preserve the precision of the parameter and can have a value of any positive integer
SIZES	provides the dimensions of the array, e.g. SIZES=3 is for a vector of 3 components; required for any parameter with more than one component (vectors, tensors, etc.)
TENSOR_ORDER	0 for scalars (default value), 1 for vectors (e.g., magnetic field or velocity), 2 for tensors (e.g., plasma pressure)
UNITS	this provides the unit of the parameter that can be indicated on the axes of a plot. If parameter has no unit, this must be specified by UNITS="unitless"
VALUE_TYPE	provides identification of the value type (essential for ASCII conversion) with possible values: CHAR ; DOUBLE ; FLOAT ; INT ; ISO_TIME ; ISO_TIME_RANGE

4.2.1 COORDINATE SYSTEMS

Although the CAA metadata accepts several different coordinates systems (for the complete list, see RD-4), most of the science datasets are given either in ISR2 or GSE. There are two important notes:

- for most of the time these two coordinate systems are very close to each other

- ISR2 is preferred coordinate system for many measurements as the the ISR2 xy-plane is the spin plane of the spacecraft

The details of the key coordinate systems are:

- SR2 (Spacecraft Reference of Frame):
  - This coordinate system is the despun Spin Reference coordinate system where its z-axis is the maximum principal inertia axis (spin axis) of the spacecraft, and its x-axis in the meridian containing the direction of the Sun.
- ISR2 (Inverted Spacecraft Reference of Frame)
  - Since the Cluster spin axes are near the direction of the southern ecliptic pole (within ~6º), rotation of the SR2 coordinate system through 180º about its x-axis brings it close to the GSE system (within ~6º).
  - It should be noted that as the satellite spin axes are not exactly parallel to each other and so the SR2 and ISR2 systems are slightly different for each spacecraft. The value of the angle varies slowly, for its value, see section 5.2
- GSE (Geocentric Solar Ecliptic)
  - x is directed towards the Sun; y lies in the ecliptic plane in the direction opposite to the Earth's velocity around the sun and is perpendicular to x; z completes the right-hand system ($z = x \times y$) which is essentially along the ecliptic north pole.
  - If one wishes to perform coordinate transformation between ISR2 and GSE, one needs to know the orientation of the spin axis which is given in "Auxiliary Data" (listed under Auxiliary datasets, see section 5.2).
- MFA and FAC
  - CAA also supports these two magnetic coordinates, which are used by a few datasets. For their definitions, see RD-4.

In addition the FRAME_VELOCITY is a very important related keyword. Particularly for DC electric fields and drift velocities it can be very important if the measurements is made in the moving spacecraft velocity (value=OBSERVATORY) or in the geophysical Earth frame (value=INERTIAL).

5. CLUSTER SPACECRAFT AND PAYLOAD

The layout of the Cluster spacecraft and its instruments are shown in Figure 5.1. Each of the four spacecraft carries an identical set of 11 instruments to investigate charged particles, electrical and magnetic fields.

Acronym	Experiment name	Key measurements/datasets
ASPOC	Active Spacecraft Potential Control experiment	○ Ion current emitted from the spacecraft into plasma
CIS	Cluster Ion Spectrometry experiment	○ Consists of two instruments: HIA (ions), CODIF (mass spectrometer for H+, He+, O+), each having two different sensitivities ○ Thermal ion distribution functions (0-30 keV) in different units ○ Ion pitch-angle distributions in different units ○ Ion moments
DWP	Digital Wave Processing experiment	○ Particle correlator (using individual PEACE counts)
EDI	Electron Drift Instrument	○ Electron drift velocity and DC E-field ○ Total magnetic field ○ Ambient electron flux (0.5 or 1 keV) at high time resolution
EFW	Electric Field and Wave experiment	○ **ExB** drift velocity and DC E-field at 25 (NM) or 450 (BM) Hz sampling ○ Spacecraft potential
FGM	Fluxgate Magnetometer	○ DC Magnetic field vector
PEACE	Plasma Electron And Current Experiment	○ Thermal electron distribution functions (0-30 keV) in different units and with two different sensitivities ○ Electron pitch-angle distributions in different units ○ Electron moments
RAPID	Research with Adaptive Particle Imaging Detectors	○ Consists of Energetic electron (30-450 keV) and ion (30keV – 4MeV) detectors ○ Electron and ion distribution functions ○ Electron and ion pitch-angle distributions ○ Electron and ion moments
STAFF	Spatio-Temporal Analysis of Field Fluctuation experiment	○ Search coil: calibrated waveform at 25 (NM) or 450 (BM) Hz sampling ○ Spectrum analyser: wave characteristics between 8 Hz – 4 kHz
WBD	Wide Band Data instrument	○ Electric field waveform up to 500 kHz
WHISPER	Waves of High frequency and Sounder for Probing of Electron density by Relaxation experiment	○ Natural and active electric field spectra between 2-80 kHz ○ Electron density from both natural and active spectra

5.1 Sun Reference Pulse

The spacecraft sun sensor is located at 26.2° counterclockwise from the Y_B axis, see Figure 5.1. The spin of the spacecraft is defined as the period between two consecutive pulses. These are stored in “Spin Time” dataset [C*_CP_AUX_SPIN_TIME, *=1-4] which is available under Auxiliary datasets.

Variable	Description
time_tags	UT Time, time of centre of the spine
spin_interval	Spin time range
spin_period	Spin period
algorithm	Spin determination data source/algorithm
flag	Spin determination flag

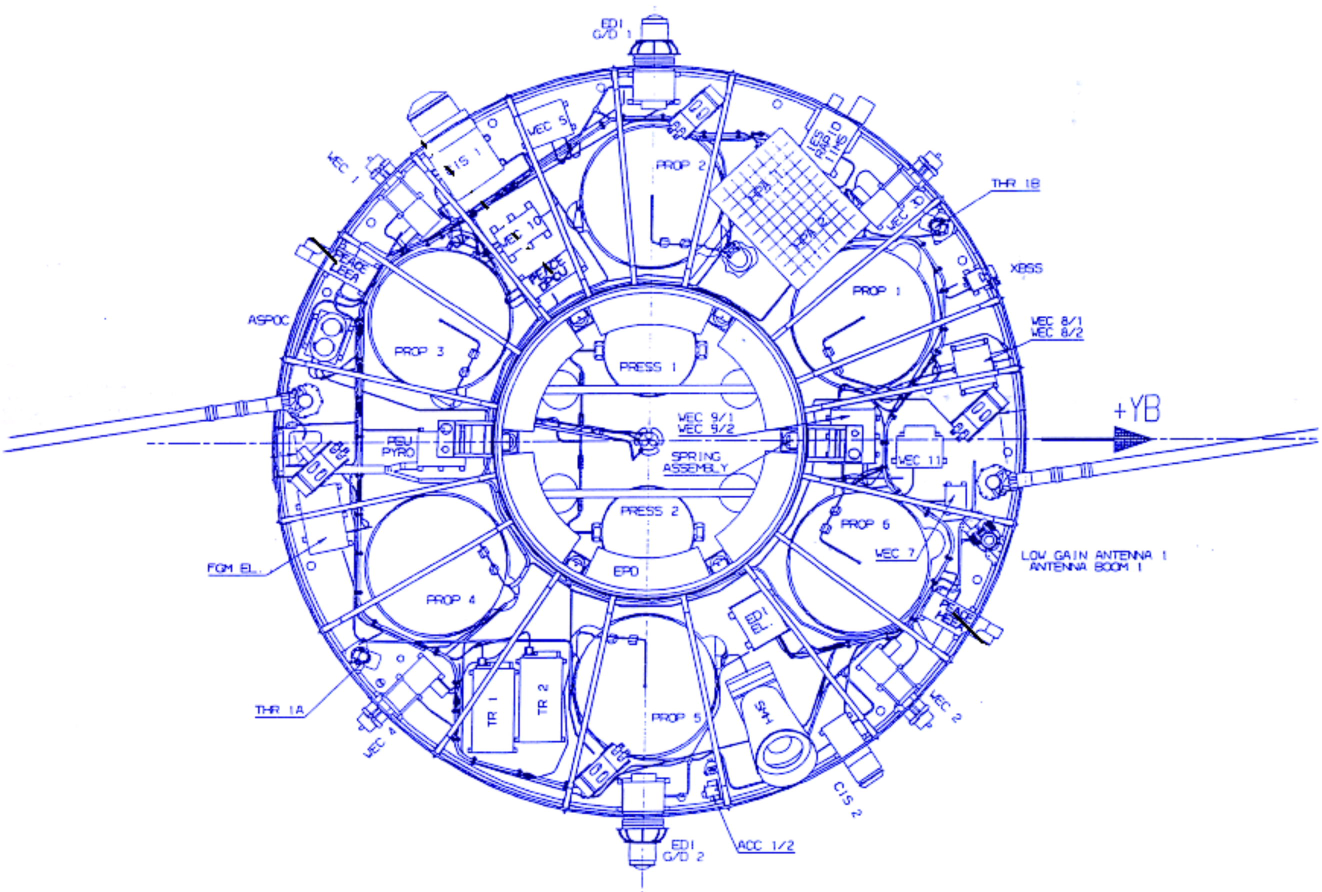

Figure 5.1. The sketch of the Cluster spacecraft with its instruments.

5.2 ESOC datasets on spacecraft performance

ESOC is responsible for the operations of the Cluster spacecraft and implements and monitors various subsystems of the spacecraft. As a result it produces a large set of datasets in various formats that have been converted into CEF by the CAA. These datasets are available under Auxiliary datasets with the following dataset titles

- ESOC - Eclipse Intervals
- ESOC - Battery Conditioning History
- ESOC - Ground Station Utilization Log
- ESOC - Experiment Power Sharing
- ESOC - Solar Array Performance
- ESOC - Spacecraft Event Log
- ESOC - Spacecraft CTU/RTU switch-over
- ESOC - SSR Bit Errors
- ESOC - Maneuvre Intervals

For many users, particularly interesting datasets are Eclipse Intervals, Solar Array Performance and Maneuvre Intervals which are explained in the following sections.

5.3 Solar Array Performance

ESOC determines the performance of the Solar Array normally once per orbit around the apogee. This information is available in dataset "Solar Array Performance" [C*_CP_AUX_ESOC_SAP, *=1-4]. The key parameters include Solar Array Power and Solar Aspect Angle. The dataset also contain useful orbital information. The parameters are:

Variable	Description
Perigee_UT	Perigee pass time (UT)
Rev_Num	Orbit revolution number
Perigee_Dist	Perigee distance
Apogee_UT	Apogee pass time (UT)
Apogee_Dist	Apogee distance
Long	Orbital elements at apogee: Longitude
SMA	Orbital elements at apogee: Semi-major axis
ECC	Orbital elements at apogee: Eccentry
Incl	Orbital elements at apogee: Inclination
Node	Orbital elements at apogee: Node
ArgPer	Orbital elements at apogee: Argument of Perigee
Eclipse	Orbital elements at apogee: Eclipse Interval
SAP_Time	SAP measurement interval (UT)
SAP_MainBusVoltage	SAP - Main Bus Voltage
SAP_Power	Solar Array Power
SAA	Solar Aspect Angle
SA	Solar Array Temperature

6. CLUSTER SPACECRAFT ORBIT AND ATTITUDE

6.1 Cluster orbit

The Cluster spacecraft were originally placed in a 4 × 19.6 R_E polar orbit (orbital period ~ 56 hours), with the single spacecraft orbits designed to enable a perfect tetrahedron to be formed at particular parts of the orbit, aiming at science investigations in such as the bow shock, polar cusp, and plasma sheet. After 2006, due to the limited fuel reserves, it was decided not to maintain the original orbits and so over the following years the line of apsides has rotated further below the ecliptic (see Figure 1). This has allowed the mission to cross the ecliptic plane much closer to the Earth and so to enter regions that were not in the original mission plan (e.g. the low-latitude magnetopause, near-Earth plasma sheet, auroral acceleration regions and radiation belts), thus enhancing the originally anticipated scientific output.

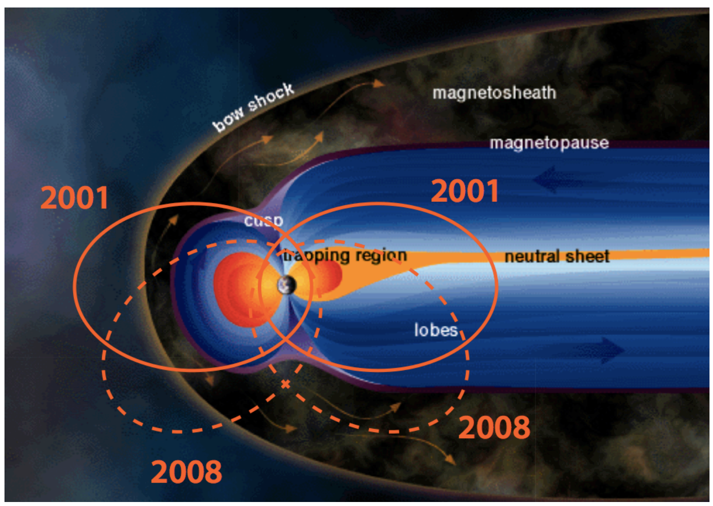

Figure 1. Cluster orbit in the x-z GSE plane during the period 2001-2009.

6.2 Cluster spacecraft separation strategy

The separation strategy during the entire mission to date is detailed in Table 1 and Figure 2. In 2001-2, the spacecraft separation was altered approximately every six months (purple line in Figure 2), whereas for the following three years (January 2003-December 2005) the manoeuvres were limited to once per year to reduce operational costs and instrument down time and to help ensure uniform spatial coverage of scientifically interesting regions at the same magnetic local time over the calendar year.

The Cluster spacecraft separation has gradually increased over the years, allowing the same regions of the magnetosphere to be sampled at different scales. This is a one-way strategy, as the spacecraft cannot return to a small tetrahedron configuration due to diminishing fuel reserves that no longer allow substantial separation distance changes, particularly perpendicular to the direction of the orbit. However, capability does exist to move spacecraft ‘along track’ in the direction of the orbit, which has facilitated the implementation of the multi-scale operations since January 2006. In this configuration, three of the spacecraft (Cluster 1, 2 and 3) form a 10,000 km sided triangle, while Cluster 3 and 4, whose orbits are very similar, can be drifted with respect to one another and their separation can be varied from a few 10s of km to 10,000 km. In other words, with such minor manoeuvres only, Cluster will have fuel to continue operations for a number of years.

6.3 Spacecraft manoeuvres

The spacecraft constellation changes require several long manoeuvres to be performed during which most of the payload are turned off. The manoeuvre intervals can be downloaded as separate datasets C*_CP_AUX_ESOC_THRUSTER, where *=1-4, that are listed under "Auxiliary and support data". The datasets describe in great detail what kind of manoeuvre has taken place, what was its duration, which tanks were used etc.

6.4 Eclipses

Until the end of year 2005, the spacecraft were kept strictly in a polar orbit (90º inclination) and so the mission had two clear eclipse seasons. Around March, when the spacecraft apogee was in the dayside, the spacecraft had short eclipses, each lasting about half an hour or less. Around September, when the spacecraft apogee was in the night sector, the spacecraft had long eclipses, each lasting up to about 3 hours.

After 2006 the spacecraft are not kept in their original orbits (in order to save fuel) but were allowed to drift, as a result of which the inclination of the spacecraft orbit has reduced causing more frequent eclipses so that in 2011-12 they tend to occur for most of the orbits.

The eclipse intervals can be found in datasets C*_CP_AUX_ESOC_ECLIPSE, where *=1-4 indicates the spacecraft number. These datasets are available under "Auxiliary and support data". The key parameters are the timings of the entry and exit of the penumbra and umbra.

CAVEAT: One should notice that the spacecraft spin period cannot be determined in eclipse, and so the vector directions cannot be resolved even if the instrument was kept in operation during eclipse. For some instruments, the data files were produced in eclipse assuming the spin period before eclipse (or interpolated using the spin period before and after eclipse). Both cases produce highly incorrect spin phase and the measurements are not correct and should not be used. Many datasets provide a caveat file on this matter.

6.5 Spacecraft Anomalies

To be written once the ESOC anomaly reports are included in the CAA database.

7. SCIENCE OPERATIONS OF THE CLUSTER MISSION

The Joint Science Operations Centre (JSOC), located at the Rutherford Appleton Laboratory, supports the Cluster Project Scientist in coordinating the complex multi-spacecraft and multi-instrument science operations of the Cluster mission.

The data production rate on Cluster depends on the spacecraft operation mode. In the normal mode (NM) the data production rate is 17 kbit/s while in the burst mode (BM) it increases to 105 kbit/s. All these data are recorded onboard on the SSMM and are transmitted to the ground once a day. In addition, the WBD experiment can produce data at 220 kbit/s during short intervals (for a few hours every few days); these data are not recorded onboard but are transmitted directly to the NASA DSN. There are a varying number of DSN stations available. The telemetry modes are given in dataset "Telemetry Mode" (dataset id: Cx_CT_AUX_TMMODE) given in Auxiliary datasets.

7.1 Normal mode (NM)

In the normal mode one needs to reduce the measurements significantly onboard in order to fit to the TM allocation and the SSMM allocation. Most importantly (1) the particle experiments cannot transmit large distribution functions, (2) DC electric and magnetic field experiments can return vectors at 25 and 22 Hz, respectively, and (3) the wave experiments can return wave spectra at lower rates. The main difference between NM1-3 is a different share of telemetry between CIS and PEACE. NM3 has a high share for PEACE, and on C2 it was used after it was found that CIS will not function while NM1 was used on the other spacecraft. Later the concept of NM3 was modified to give more telemetry to RAPID as well.

7.2 Burst Mode (BM1)

In the burst mode, the experiments can return more complete measurements, such as (1) particle experiments can return their large particle distribution functions, (2) DC electric and magnetic fields are sampled at 450 and 66 Hz, respectively, (3) the wave experiments have high cadence in the wave spectra. For more details, see the user guides of the experiments, [RD-1].

7.3 BM3 modes

The most important additional operation mode is BM3. They are short (6-minute) periods to dump the instruments' internal burst memory and they occur twice per orbit. In particular the EFW experiment uses this opportunity regularly to return very high frequency sampling of electric fields.

There are important consequences to the CAA datasets in BM3 such as the FGM experiment does not return magnetic field measurements during this period, and therefore any other product that requires the knowledge of magnetic field vector is not available such as pitch-angle distributions.

8. CAA DATASETS

The CAA creates over 400 different datasets from each Cluster spacecraft. Most of the datasets have been calibrated and validated by the CAA teams to the level suitable for optimum scientific return. However, all data cannot be validated as well as desired because of limited resources available, and thus the data files can still contain features that represent instrumental effects rather than physical processes. It is therefore important that the users read all documentation and caveat information related to the data and contact the CAA for specific questions or problems. Many products also include quality flags to help the user to apply the datasets in their research.

Currently the data files are delivered roughly a bit less than 12cmonths delay to the CAA. The datasets that require more demanding cross-calibration may take a longer time whereas the datasets that are straightforward to produce are delivered with 3-4 months delay. It is aimed that in summer 2014 nearly all observations up to the end of year 2013 have been calibrated and ingested. In addition, it is foreseen that a massive reprocessing of many/most data products is still under planning for some instruments in order to improve quality and consistency of data, which highlights the benefit of the cross-calibration working group, which has helped to identify and solve calibration issues.

The availability of the data on any given day can be seen in inventory plots. These plots are based on inventory files that contains data intervals and gaps produced at the time of ingestion reading the time array of each file. A gap is defined to be an interval of no data for more than 5 minutes.

The total number of files that are currently active (i.e. latest version) is in excess of 3.5 million with a total uncompressed volume equivalent to about 100 TB. The pre-generated graphical panels amount to more than 7 million files that require about 40 TB of disk space.

8.1 Cluster raw data media (RDM)

For the second extension of the Cluster mission, starting in January 2006, the distribution system of the raw data was changed. Earlier, the raw data collected by ESOC were transferred to CD-ROM and mailed to about 70 recipients. Now the data are being transferred over the network, approximately ten days after the measurements, from ESOC to the CAA from where the RDM can be accessed by the users. The CAA also provides a full online archive of the RDM from the entire mission by copying the existing CDs. Notice that these data are non-calibrated in TM units and therefore primarily useful to those who have the PI software available for the decommutatation and analysis of the raw data. The CAA does not hold and run the PI software.

For automated access of the RDM data, one can use HTTP GET method either within a web browser address bar or using a command line based http browser (e.g. wget) tool:

http://caa.estec.esa.int/cgi-bin/caa_rdm?key1=value1&key2=value2&...

Both the keyword and the value are treated as case insensitive. The files delivered are direct copies of the contents of the original RDM volume. The details of the utility including the keywords and their values are explained in [RD-7].

8.2 Timing accuracy

The standard timing accuracy of the Cluster data packets is ±2ms, as stated in [RD-5]. Better knowledge of timing is needed by several wave experiments. Improved timing information is provided in the DWP_TCOR dataset which are available among the DWP datasets. This correction is used to have the time tag of full-resolution data to an accuracy of ~20 microsecond. The instruments that require this kind of high time resolution has already taken an advantage of this dataset, and have corrected the timing in order to ease the usage of such datasets.

8.3 Cluster science datasets

The core of the CAA activity is to calibrate and produce calibrated full-resolution science datasets as well as various important auxiliary datasets. These datasets are produced by the eleven PI teams who have been financially supported by the CAA since year 2004. These datasets are explained in detail in the instrument user guides, calibration reports and ICD documents (see RD-1).

8.4 Cluster Auxiliary datasets

8.4.1 Orbital information

The orbital information of the Cluster satellites are available in two datasets, "Predicted Position and Velocity" (C[n]_CP_POSVEL, where [n]=1-4) and "Auxiliary Data" (CL_SP_AUX), both of which contain data at one minute resolution. The "Auxiliary Data", listed under Cluster 3, contains the position and velocity of the reference spacecraft (Cluster 3) in GSE coordinates. In addition, the separations of the other three spacecraft with respect to the reference spacecraft are given in GSE.

In addition the "Definite Position and Velocity" dataset contains the spacecraft position and velocity in GSE and the dataset is available for each spacecraft.

There are several datasets that are related to trajectories of the four Cluster spacecraft.

- Definitive Position and Velocity
- Predicted Geometric Positions
- Predicted Magnetic Position
- ECLAT - Cluster Magnetic Footprint (T96 model)
- ECLAT - Cluster Magnetic Footprint (TS-05 model)

The last two datasets are results of the EU/FP7 ECLAT project that provides the footprinf of each Cluster spacecraft using two different Tsyganenko models.

8.4.2 Magnetic coordinates of the spacecraft

The "Predicted Magnetic Position" (C*_JP_PMP, where *=1-4) dataset provided for each spacecraft contains the predicted invariant latitude, magnetic local time and L shell of each satellite at five minute resolution. This dataset has been produced by JSOC.

8.4.3 Sun reference pulse and spin period

The "Spin timing" (C*_CP_AUX_SPIN_TIME, where *=1-4) dataset contains the timings of the sun reference pulse, extracted from the spacecraft HK data. This dataset is provided for each spacecraft. Further information about the timing and sun reference pulse can be found in document "Spin Timings and Offsets for Cluster Prime Parameters" (DS-QMW-TN-0007) that can be downloaded from the CAA documentation area.

8.4.4 Spacecraft spin axis direction

The direction of the spacecraft spin axis is given in C*_CP_AUX_SPIN_AXIS, where *=1-4, that gives the two angles (latitude and longitude) of the spacecraft spin axis in GSE measured with respect to Ecliptic Plane for all four spacecraft.

Tilt campaign. In May 2008 a special operation was implemented to measure the 3rd component of the electric field along the spin axis. This was achieved by tilting Cluster 3 and Cluster 4 with respect to one another by 45 degrees, while they were separated by only 40 km. During the tilt campaign the ISR2 and GSE coordinate systems are not similar and so using ISR2 as a GSE proxy is not valid during this interval.

8.4.5 Tetrahedron parameters

The "Auxiliary Data" (CL_SP_AUX) dataset contains a number of parameters that describe the quality of tetrahedron, such as

Variable	Description
sc_config_QG	Tetrahedron Quality G, (vol/ideal) + (surf/ideal) + 1
sc_config_QR	Tetrahedron Quality R, norm*(vol/sphere vol)^1/3
sc_geom_size	Tetrahedron size L, characteristic size
sc_geom_elong	Tetrahedron Elongation E, elongation
sc_dr_min	Minimum Distance between Spacecraft
sc_dr_max	Maximum Distance between Spacecraft
sc_geom_planarity	Tetrahedron Planarity P
sc_geom_E_dir_gse	Direction of Elongation, direction cosines
sc_geom_P_nor_gse	Normal of Planarity, direction cosines

The definition of these variables can be found in "User Guide to the CSDS" that is available in the CAA documentation area.

8.4.6 Earth's magnetic field

The "Auxiliary Data" dataset contains a number of parameters that help user to analyze the measurements in magnetic coordinate systems, such as in GSM. The key variables are

Variable	Description
gse_gsm	Rotation angle GSE to GSM, positive from +z towards +y
dipole_tilt	Dipole Tilt in GSM z-x Plane, positive from +z towards +x

8.5 CSDS datasets

It was recognized in the 90'ies that the Cluster mission must have a fast and reliable exchange of datasets between the widely scattered experimenters. As a result, the Cluster Science Data System (CSDS) was created that consists of a set of nationally distributed data centres to generate and maintain selected datasets. These datasets are called PP and SP parameters that are called preliminary datasets in the CAA database, where PP values are of spin resolution and created for each spacecraft while SP parameters are one-minute averages and calculated for the reference spacecraft (Cluster 3). These datasets are explained in detail in CSDS guide [RD-6].

These preliminary values are listed as ancillary datasets because of the limited calibration routines done to create them. In all case better calibrated parameters can be found among the CAA datasets. Therefore the PP/SP datasets are not meant for scientific use but rather for searching events. The users are asked to use them with great care and rather use the CAA datasets that have been calibrated to much higher quality.

8.6 Quick-look Plots

The CAA produces various quick-look plots to support the usage of the Cluster observations.

8.6.1 Cluster Summary plots

As a part of the CSDS activity, the CSDS UK data center produces Cluster quick-look plots with a very short delay on the reference spacecraft (normally Cluster 3) based on default calibration values. These so-called CSDSweb plots can be accessed via CAA/CSA interface.

In addition the CAA team create the same plots for each spacecraft using the calibrated CAA datasets. Currently the following plot types are available:

- Overview
- Fields
- Particles1
- Particles2

These are available at 6-hr and orbital intervals.

8.6.2 Comparison plots

There are several physical parameters that are measured by more than one instrument. The CAA creates hourly plots where two measurements are compared which helps the users to assess the similarities and differences of the two measurements. These plots can be downloaded or viewed via quick-look browsing tool.

So far the following comparison plots have been created:

Density Comparison	all three particle instruments vs Whisper • particles vs WHISPER instrument specific plots (active sounding and natural spectra are separated) • PEACE vs WHISPER • CIS-CODIF vs WHISPER • CIS-HIA vs WHISPER
Total Magnetic Field Comparison	total magnetic field • FGM vs EDI
Drift Velocity Comparison	comparisons made in ISR2. All vectors used are in inertial frame (i.e. in the Earth's frame, not in the moving spacecraft frame) • EDI vs EFW • EDI vs HIA • EDI vs PEACE • HIA vs EFW • HIA vs PEACE • PEACE vs EFW
DC Electric Field Comparison	As for drift velocity, the following comparisons are made in ISR2 and in inertial frame • EDI vs EFW • EDI vs HIA • EDI vs PEACE • HIA vs EFW • HIA vs PEACE • PEACE vs EFW

9. CROSS-CALIBRATION/CALIBRATION ACTIVITIES

The key area of CAA activity is related to the extensive instrument calibrations and cross-calibrations. Since autumn 2005, the CAA cross-calibration working group is actively supported by the instrument teams with two annual workshops; for the list of the past workshops, see

http://caa.estec.esa.int/caa/cross-cal.xml

where all the presentations and the minutes of the meetings are available to the users. This activity is invaluable for assessing and improving the quality of the datasets. Although the active participants in the working group come mainly from the instrument teams, anyone interested in the calibration issues of various measurements are welcome to participate in these workshops.

Among many areas of calibration activity, some of the most critical ones are:

- to improve the calibration of dc magnetic field measurements to a level of ~0.1 nT
- to determine the total plasma density using various measurements techniques, and to identify the limitations of each method
- to calibrate the dc electric field measurements using and comparing various data sources, and to identify the limitations of each technique
- to determine the **E**×**B**-drift velocity using and comparing various data sources, and to identify the limitations of each technique
- to compare the fluxes of electrons and ions measured by different experiments
- to compare power spectral densities of wave measurements, including both electric and magnetic field wave components

To help this activity, the CAA creates hourly comparison plots where two similar measurements are compared. This helps the user to assess the similarities and differences of the two measurements. These plots can be downloaded or viewed via quick-look browsing tool. For details of these plots, see section 8.6.2.

10. DOCUMENTATION OF THE CAA

10.1 Instrument documentation

The CAA documents can be found at

http://caa.estec.esa.int/caa/documentation.xml.

There are three key documents for each experiment that help the user to utilize the CAA datasets (available at http://caa.estec.esa.int/caa/ug_cr_icd.xml):

Acronym	Title	Description
ICD	Interface Control Document	this contains a brief description of the experiment and all detailed information about the syntax of the instrument datasets
UG	User Guide	this provides details of key science datasets, how to use them and how to apply the quality and caveat information of the datasets. This document improves the usability of the instrument datasets.
CR	Calibration Report	this provides the results from the CAA Cross-Calibration Working Group Activity and help the users to evaluate the quality of the datasets

10.2 ESOC documentation

The spacecraft operations are taken care of by the ESOC who provides various documents and reports. They are provided in the CAA documentation area.

10.3 Other documents

The spacecraft operations are taken care of by the ESOC who provides various documents and reports. They are provided in the CAA documentation area.

11. CAA REVIEWS

To ensure the high quality of the CAA data products and service to the science community, extensive reviews are regularly organized. Formal project reviews are held annually to verify that the key milestones have been met, that the technical aspects of the project are sound, and that the activities are progressing on schedule and within budget. The board reports and all presentations of all reviews can be found at http://caa.estec.esa.int/caa/reviews.xml. The following project reviews have been held:

Date	Review	Objectives
2003 Nov	System Specification Review	To approve the user requirements, to review conceptual design, to review proposed PI inputs, to agree the archive implementation plan, and to release instrument team contracts
2005 May	Implementation Review	to approve the archiving plans of the experiments and the technical implementation of the overall system
2006 May	1st Operations Review	status of data delivery and CAA services for years 2001-2
2007 May	2nd Operations Review	status of data delivery and CAA services for years 2003-4
2008 May	3rd Operations Review	status of data delivery and CAA services for years 2001-5
2009 March	CAA Peer Review	actvity lasted for two months in Jan-Mar. The objective was to evaluate the quality and usability of the datasets for years 2001-5.
2009 May	4th Operations Review	status of data delivery and CAA services for years up to 2007
2010 June	5th Operations Review	status of data delivery and CAA services for years up to 2008
2011 May	6th Operations Review	status of data delivery and CAA services for years up to 2010
2012 May	7th Operations Review	status of data delivery and CAA services for years up to 2011
2013 June	8th Operations Review	status of data delivery and CAA services for years up to 2012
2014 June	9th Operations Review	status of data delivery and CAA services for years up to 2013; first review of the Double Star datasets

In addition to help the implementation and transfer of the CAA database and services into the CSA, three specific reviews were held together with 6th-8th CAA Operations Reviews:

Date	Review	Objectives
2011 May	CSA Specification Review	to evaluate the requirements and transfer plan of the CSA
2012 June	1st CSA Implementation Review	to evaluate the progress of the CSA Implementation
2013 May	2nd CSA Implementation Review	to evaluate the progress of the CSA Implementation

From now on, as the CSA is operational, the CSA services are reviewed together the CAA Operations Review.