Automated Multi-Dataset Analysis (AMDA): An on-line database and analysis tool for heliospheric and planetary plasma data

Vincent Génot, E. Budnik, C. Jacquey, M. Bouchemit, B. Renard, N. Dufourg, N. André, Baptiste Cecconi, H. Si Hadj Mohand, C. Tao, et al.

Planetary and Space Science 201 (2021) 105214

Contents lists available at ScienceDirect

Planetary and Space Science

journal homepage: www.elsevier.com/locate/pss

Automated Multi-Dataset Analysis (AMDA): An on-line database and analysis tool for heliospheric and planetary plasma data

V. Génot[a,*], E. Budnik[a,b], C. Jacquey[a], M. Bouchemit[a], B. Renard[c], N. Dufourg[d], N. André[a], B. Cecconi[e], F. Pitout[a], B. Lavraud[a], A. Fedorov[a], M. Ganfloff[a], I. Plotnikov[a], R. Modolo[f], N. Lormant[c], H. Si Hadj Mohand[c], C. Tao[g], B. Besson[d], D. Heulet[d], D. Boucon[d], J. Durand[d], N. Bourrel[a], Q. Brzustowski[a], N. Jourdane[a], R. Hitier[a], P. Garnier[a], B. Grison[h], N. Aunai[i], A. Jeandet[i], F. Cabrolie[j]

[a] IRAP, Université de Toulouse, CNRS, CNES, UPS, Toulouse, France
[b] Noveltis, Toulouse, France
[c] AKKA, Toulouse, France
[d] CNES, Toulouse, France
[e] LESIA, Observatoire de Paris, CNRS, Université PSL, Meudon, France
[f] LATMOS, France
[g] NICT, Japan
[h] Department of Space Physics, Institute of Atmospheric Physics of the Czech Academy of Sciences, Prague, Czech Republic
[i] LPP, France
[j] GFI, France

ARTICLE INFO

Keywords:
Space plasmas
Analysis tool
Database
Interoperability

ABSTRACT

Accessing, visualizing and analyzing heterogeneous plasma datasets has always been a tedious task that hindered students and senior researchers as well. Offering user friendly and versatile tools to perform basic research tasks is therefore pivotal for data centres including the Centre de Données de la Physique des Plasmas (CDPP http://www.cdpp.eu/) which holds a large variety of plasma data from various Earth, planetary and heliophysics missions and observatories in plasma physics. This clearly helps gaining increased attention, relevant feedback, and enhanced science return on data. These are the key ideas that crystallized at CDPP more than 15 years ago and resulted in the lay-out of the concepts, and then development, of AMDA, the Automated Multi-Dataset Analysis software (http://amda.cdpp.eu/). This paper gives a description of the architecture of AMDA, describes its functionalities, presents some use cases taken from the literature or fruitful collaborations and shows how it offers unique capabilities for educational purposes.

1. Introduction

The genesis of AMDA[1] dates back to 2005 when the CDPP[2] team realized that no tool existed in the space physics domain to search for and visualize data from various origins and often with heterogeneous formats. This variety of formats (because of varied instruments, IT systems, habits, persons, …) was indeed a key obstacle to efficient instrumental inter-calibration campaigns, cross-mission and cross-instrument analysis and probably of enhanced scientific results. In brief, it appeared that a certain level of uniformity in data products, physically (format) and conceptually (structure), is needed for a database to be 1/more attractive and 2/able to build value-added services on its data. If those ideas prevailed at CDPP, they were in fact shared by others. In magnetospheric physics, this is the general vision that pre-dates to the births of the Cluster Active Archive (now Cluster Science Archive, Laakso et al. (2010)) and the CDAWeb (NASA CDAWeb Development Team, 2019), and in planetary sciences in a lesser measure (because services are more difficult to build from the diversity of data), the NASA/Planetary Data System

* Corresponding author.
E-mail address: vincent.genot@irap.omp.eu (V. Génot).

[1] http://amda.cdpp.eu.
[2] http://www.cdpp.eu.

https://doi.org/10.1016/j.pss.2021.105214
Received 25 August 2020; Received in revised form 15 March 2021; Accepted 19 March 2021
Available online 10 April 2021

(McMahon, 1996; Walker et al., 1996) and the ESA/Planetary Science Archive (Besse et al., 2018). In Japan the Data ARchives and Transmission System (DARTS) followed a similar path, but at yet another level as it also aggregates astrophysical data in the same system (Miura et al., 2000; Tamura et al., 2004). On the software side, the concepts behind SPEDAS (Angelopoulos et al., 2019) share many similarities with those that led to AMDA. The ability to integrate space physics scientific functions and models, a key element in AMDA, took its inspiration from the development of the QSAS software.[3] Finally the versatility for displaying different kinds of data permitted by Autoplot[4] is also an example (Faden et al., 2010).

At CDPP, a prototype was therefore rapidly put in shape from existing bricks: a data server, a rudimentary web interface, and some IDL codes for plasma data analysis. The initial guidelines that are still in use today are: 1/using a limited set of objects that are represented and manipulated by the user in the interface (parameters, TimeTables/catalogues, requests), 2/using standard formats for import/export (CDF, netCDF, VOTable, json, tabular ASCII), and 3/favoring interoperability with other similar systems (web-services, protocols). Being accessible via a browser was also a novelty at the time, and it proved to be quite efficient in term of maintenance over the years as javascript evolved but maintained backward compatibility. The CDPP approach to multi-source data analysis (applying the above principles) is illustrated in Fig. 1. It resulted in a flexible and attractive online tool, AMDA, which allows the user to combine and plot data from heterogeneous sources. As of today, AMDA serves more than 35 space missions and ground observatories (past, active or in preparation), various planetary and solar indices, and supporting models.

While displaying data is the main feature of the service, it also comes with other scientifically valuable features; among them is the “data mining” capability, enabling to search data with combinations of predefined mathematical functions, and initially called “conditional search”. This function produces “time tables” or lists of events (that comply with the conditional criteria applied on datasets and time intervals) which are a common and very much used object in space physics for which data mostly take the form of time series (data records indexed in time order). In the following TimeTable(s) will be used when referring to lists of events handled in AMDA; further description is given in Section 2.1.2, and the distinction with catalogues is provided in Section 2.1.3. The ability to produce, manage and share TimeTables among AMDA users became very popular such that a common, simple format was designed and implemented in different space physics tools (CL/CLWeb,[5] 3DView[6]). It was later extended to provide standardized catalogues (further discussions led to the HPEvent, see Section 2.1.3). Today with the advent of machine learning techniques which make a great use of catalogues of events, the role of TimeTables and catalogues in AMDA is more and more prominent.

Over the years, the number of AMDA users increased rapidly (it is close to a thousand to date, from many different countries) and the outcomes of this ‘success’ are at least two-fold. First, and most importantly as it was the initial goal, AMDA helps produce science results. This paper aims at showing that this goal is achieved by facilitating the life of scientists, from undergraduate students to senior researchers, which is very probably illustrated by 1/the increasing usage of the tool and the partly correlated number of publications (Section 6.1), and 2/the variety of use cases (Section 6.4) testifying of eased mass treatments, data mining, classification and vizualisation. Second, and it was not planned so, the achievement of AMDA helps the CDPP as a whole being recognized for its expertise in data distribution and exploitation, such that in return AMDA was, and is, involved in various data-centered projects (e.g. Helio, Europlanet, IMPEx, some of which are shortly described below).

The purpose of this paper is to describe the basics of AMDA, how it is built (Architecture, Section 2), what it is concerned with (Data, Section 3), what it performs (Functionalities, Section 4) and how it lives and communicates in the ever-growing ecosystem of tools and databases (Interoperability, Section 5). Finally an important part (How useful ? Section 6) is devoted to applications which show the variety of ways AMDA has been used by the scientific community. The description of the tool makes necessary the mentions of a number of concepts related to space science, IT technologies, and organizations which heavily rely on acronyms. Therefore a list of the most cited acronyms is provided in appendices to help the reader. This paper offers a significant extension of the description of AMDA which is already available in the literature but based on an out-dated version of the tool (Jacquey et al., 2010; Génot et al., 2010a,b). Finally, the conclusion offers some thoughts on running a long term development for the scientific community in changing times. Following the FAIR principles (Findable, Accessible, Interoperable and Reusable) the code is open[7] and under the GPLv3 licence; in essence AMDA is FAIR as well as the data it makes accessible.

2. Architecture

The size, abstraction and complexity of AMDA’s code base has greatly evolved over the years. The architecture presented here corresponds to the latest version of the tool, although the general concepts were established long ago. Ergonomics are a key characteristics of AMDA’s ambition. First, as an online tool, no installation is required. Data are readily accessed and manipulated by users through a combination of objects and graphical elements. These objects are viewed and handled through the Workspace Explorer, the central part of the Graphical User Interface (GUI) which is presented as laptop desktop. These workflows are made possible via a modular architecture organized in “layers” glued together by web-services and standardized protocols. These concepts are described in the following sections.

2.1. General concepts of AMDA

2.1.1. Parameter

A parameter is the central object in AMDA and represents a time dependent quantity. It may be a scalar, a vector or higher order array (2D, 3D), for instance a proton density, a magnetic field vector or a particle flux depending on energy and incidence angles. A parameter has a time resolution and may have a unit. Parameters have a common name (e.g., |*b*| or ‘density’) and an ID (‘*c1_btot*’ for the magnetic amplitude of Cluster 1 at 4s resolution). The user can build user-defined parameters from basic parameters available in the Workspace Explorer. For a user-defined parameter called ‘*name*’, the ID will simply be ‘*ws_name*’. This ID is visible in the plot layout (see Section 4.1) and is used to derive new parameters in the editing window (see Section 4.2).

For a given instrument on a particular observatory (in a broad sense: spacecraft or ground-based facility) the classical hierarchy is such that a parameter is the leaf of the data tree “mission > instrument > dataset > parameter”.

2.1.2. TimeTable

A TimeTable is a collection of couples (start time and end time) representing a list of events. It can be used to represent any list of intervals of user interest, from physical events like Interplanetary Coronal Mass Ejections or plasma boundary crossings to instrumental mode of operations, etc. A TimeTable has a name, a number of intervals, and optionally (but very recommended) a description composed of a free text.

A TimeTable in AMDA may come from different sources

[3] https://sourceforge.net/projects/qsas/.

[4] http://autoplot.org/.

[5] http://clweb.irap.omp.eu/.

[6] http://3dview.cdpp.eu.

[7] https://gitlab.irap.omp.eu/groups/CDPP.

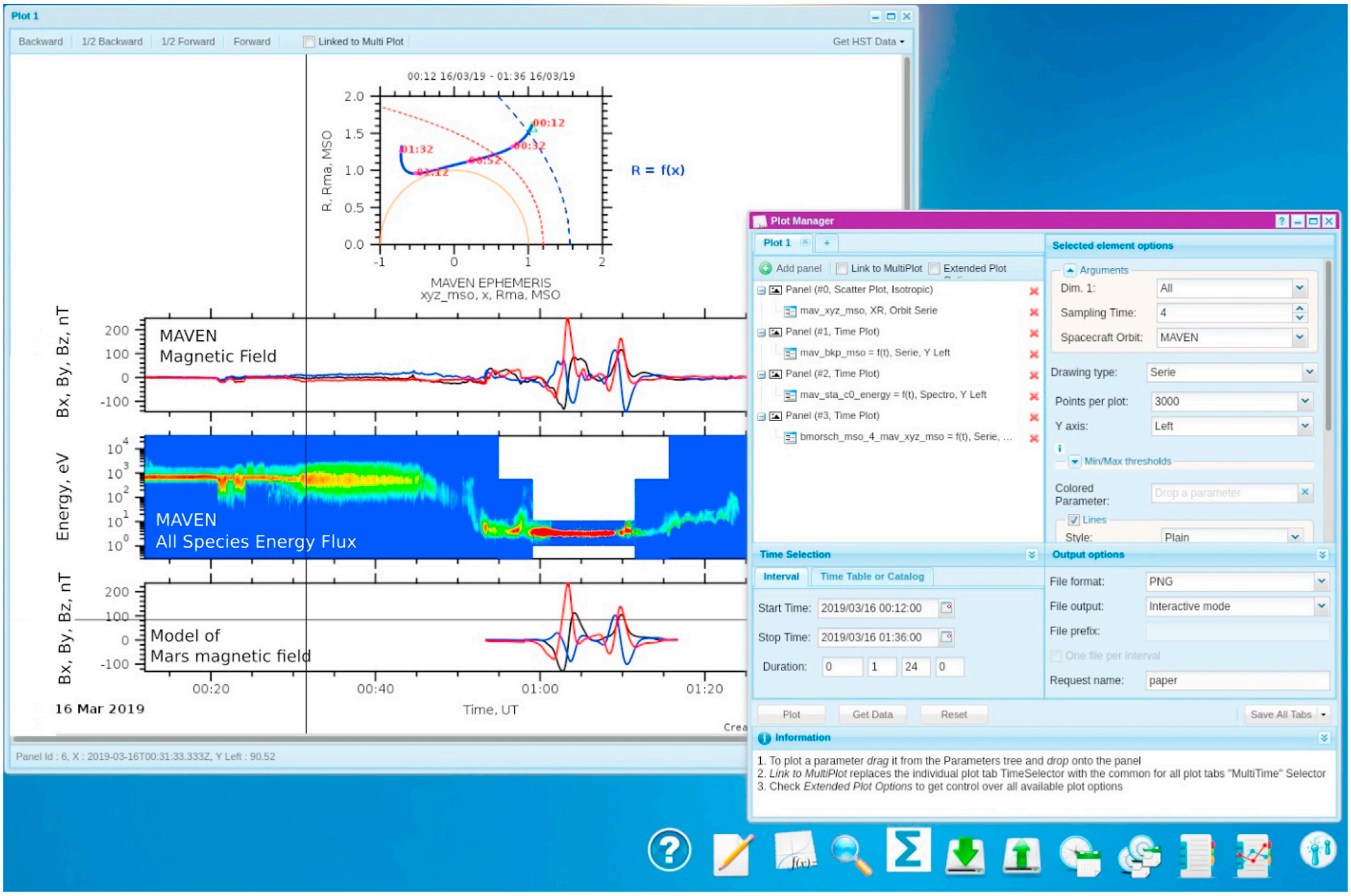

Fig. 1. A generic AMDA screen. The plot window shows the MAVEN orbit about Mars, MAVEN magnetic field data, MAVEN Energy-Time spectrogram of the ion energy flux, and the Morschhauser et al. (2014) model of the martian crustal magnetic field calculated along the MAVEN orbit. The right hand side window shows the "plot request" panel set up for this plot. Icons at the bottom right are short cuts to the different analysis activities described in the paper.

- conditional search (data mining), as discussed above, see Section 4.3
- upload, in various possible formats, see Section 4.5
- mouse selection for start and stop times directly on the plot window, see Section 4.1

TimeTables may be manipulated (merge, intersection) in a special window described below in Section 4.4. Statistics on the intervals of a given TimeTable are also available.

2.1.3. Catalogue

A catalogue is an extension of the TimeTable object. Both types of objects will eventually be merged into one single concept but, at present, for historical and technical reasons, there is still a distinction in AMDA. Catalogues have two first columns similar to those of TimeTables (i.e. a start time and a stop time) followed by an unlimited number of additional columns. This additional information of various types (float, integer, string, …) corresponds, for instance, to the mean value of a given parameter on the corresponding interval (or the min, max, median, or any other statistical quantity), a flag (instrument mode, plasma region), a comment, a URL, a DOI, etc.

Catalogues in AMDA may come from the statistics module (see Section 4.6) or from upload (see Section 4.5). Discussion on a common catalogue format in space physics took place under the auspices of the SPASE consortium (in which the CDPP participates) and led to the HPEvents format[8] which now adopted in AMDA.

2.1.4. Request

The interface layouts setup by a user for either plot, download, or data mining (see the following section for details) can be saved as corresponding "requests" (with a name chosen by the user). All parameters, settings, options are saved for future use and can be retrieved in the "operations" tab of the Workspace Explorer (see Fig. 2).

2.2. Navigating in the GUI: the Workspace Explorer

The Workspace Explorer is the main window used to navigate in AMDA resources and to monitor actions. It is unique to each user and is divided into three tabs as shown on Fig. 2: resources, operations, and jobs, all of which gather AMDA and user data together with the results of data treatments. Resources is the main location for data either coming from the AMDA database, or remote databases, or more generally all data uploaded by the user (My DataBase). It is also where new objects defined by the user are stored: derived parameters, TimeTables and catalogues. Objects shared between all AMDA users (TimeTables and catalogues) are also accessible there in "Shared TimeTables"/"Shared catalogues" directories. Each user is allocated a dedicated storage capacity of 200 Mb with no time limit; each session is logged and saved for future use. Some users regularly use and update their workspace since the beginning of AMDA 15 years ago. The second tab is "operations" where requests for plotting and data mining are saved (download and statistics requests are currently being implemented). Clicking on a saved request automatically

[8] http://www.spase-group.org/docs/conventions/HDMC-Event-List-Specification-v1.0.3.pdf.

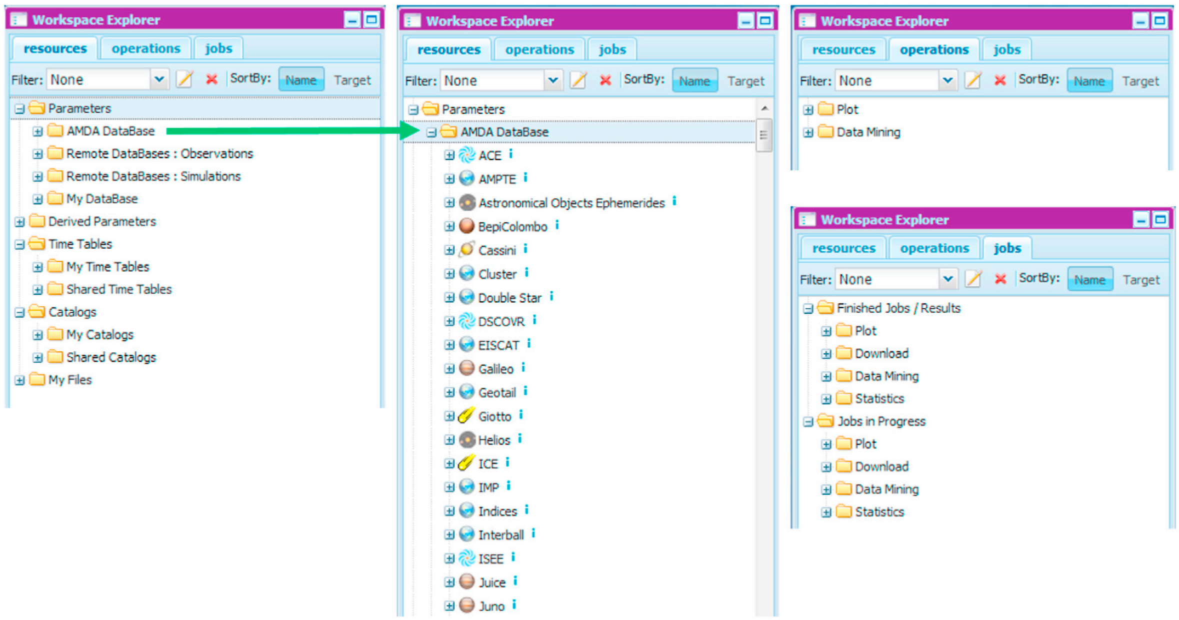

Fig. 2. The Workspace Explorer with the three tabs “resources” (left), “operations” (top right) and “jobs” (bottom right). The middle image is a view of AMDA database resources (by alphabetical order, from A to J only on this view) and shows directories for spacecraft missions, radar, indices, and ephemeris.

fills the plot/data mining interface with the corresponding settings. Finally the “jobs” tab enables users to monitor actions which take too long to be run interactively and have therefore been transferred to the batch mode. Active jobs are in “Jobs in Progress” and, when terminated (either correctly or with error), go into “Finished Jobs/Results”. Information on a given job is given in the tooltip.

The middle panel of Fig. 2 shows different types of AMDA resources. The icon before each mission name refers to the type of environment the resource is about: planet, the solar wind, comet. Most of these resources are space missions but radar data (EISCAT), celestial body ephemeris, geomagnetic indices are also available. More resources can be seen by scrolling down this menu.

2.3. AMDA software structure

Fig. 3 shows the AMDA software structure. Starting from the top of the figure, AMDA can equally be accessed by either a human through the web javascript interface or through its web-service; there are internal mechanisms that transparently handle both types of requests. The Web Server provides services which, in most cases, generate a request (see Section 2.1.4) for the Kernel. When a request is prepared by the Kernel Access Layer, the Web Server calls a corresponding script from the Kernel package. The Kernel keeps responsibility for some services (like TimeTable management and custom parameters rules) but the main task of the Kernel scripts is the Parameters preparation and manipulation. The Kernel Application Layer decodes the request and, if it has to prepare parameters, passes the corresponding commands to the Kernel Parameter Layer. The Kernel Parameter Layer can get the original parameters from several sources as follows:

1. AMDA local Data Base via conventional internet access
2. Remote data centres (e.g. NASA/CDAWeb)
3. IMPEx remote data centres via web services (see Section 3.2)
4. Users custom data

Note that the internal AMDA Data Base is able to manage a Local Base as well as data located in remote data centres such as NASA/CDAWeb. Also the user can provide his/her own files with custom data together with the custom parameters extraction and calculation rules. When the Kernel gets all necessary data, it calculates the parameter using the parameter creation rules, re-sampling and interpolation utilities. Then, if needed, the Kernel plots the parameter into the corresponding file (PNG, for instance) or prepares data for download. When the image or data are available, the Web Server sends them back for display or download accordingly. Expressions for user-defined parameters or data mining requests generate internal codes that are compiled as libraries which are loaded “on the fly” by the Kernel.

2.4. Web-services and communication

As it can be seen on Fig. 3 there are two ways of accessing AMDA data, the first one being through the GUI as it is explained in previous Sections. The second access is via webservices and is mostly used for machine to machine communication, for instance when a user wants to automate some data retrieval and treatment outside of AMDA. These webservices are developed with the two protocols REST and SOAP. For the simpler REST access, via a URL in a browser, dedicated documentation is available[9]; alternatively the SOAP implementation is also described[10]. The list of available webservices is the following:

- *auth*: returns a token to use as an API security parameter for getDataset, getOrbits and getParameter; this token has a 10min validity limit
- *getDataset*: provides a full dataset (one or a series of parameters) for a given period of time
- *getObsDataTree*: provides the hierarchy of public access data
- *getOrbits*: provides the trajectory of a spacecraft for a given period of time
- *getParameter*: provides a single parameter for a given period of time
- *getParameterList*: provides the list of AMDA parameters (including those from the “Derived parameters” directory if information on the user is provided)
- *getStatus*: gets the current status of the request according to a process ID
- *getTimeTable*: provides the contents of a TimeTable or Catalogue

9 http://amda.irap.omp.eu/help/apidoc.
10 http://amda.irap.omp.eu/help/Methods_AMDA.xml.

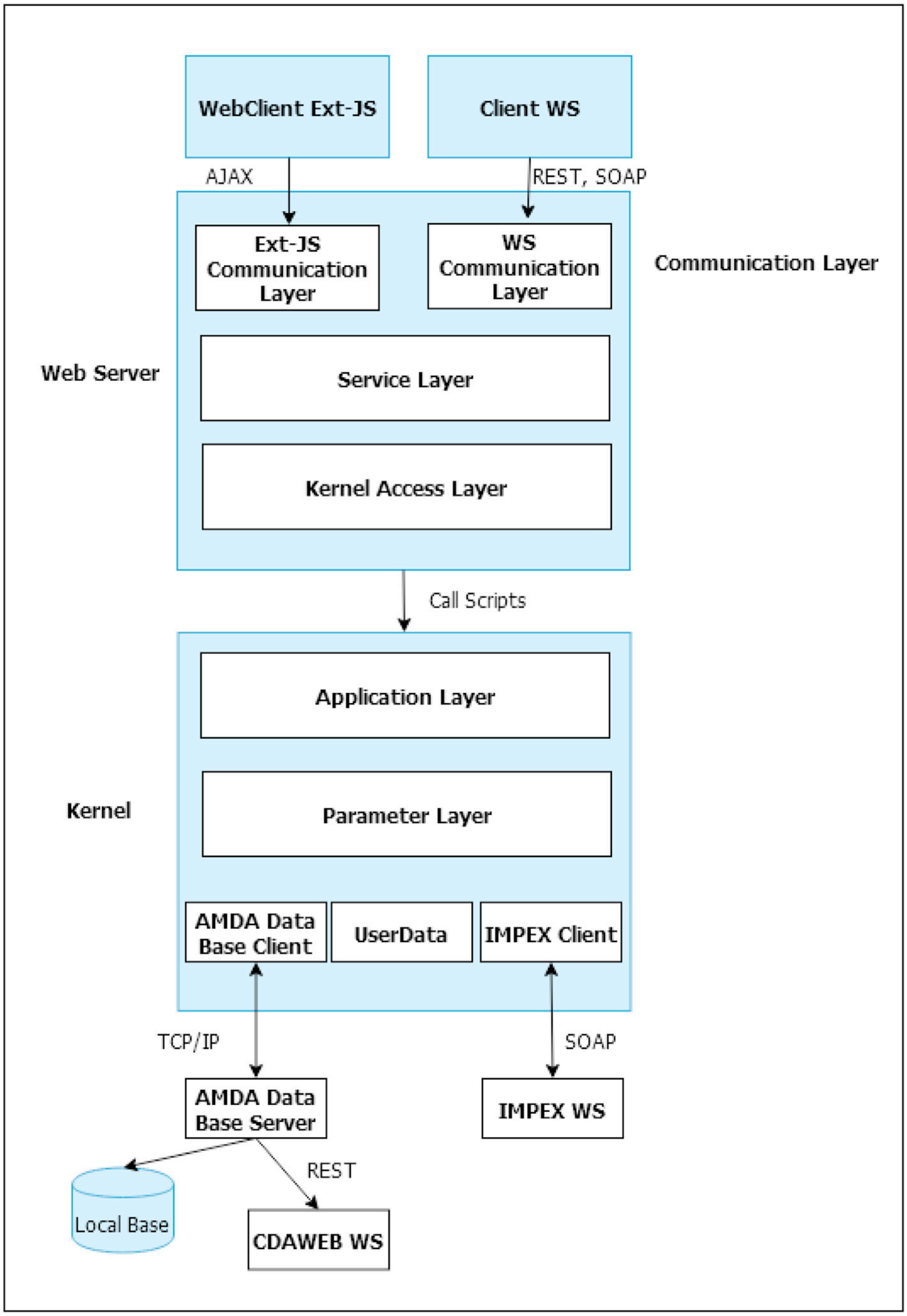

Fig. 3. General architecture of AMDA (see text for details).

•*getTimeTableList*: provides the private list of TimeTables owned by a user. When called without userID, this web-service returns the list of shared TimeTables.
•*isAlive*: checks whether AMDA services are available or not

These webservices were initially developed during the IMPEx project (see Section 3.2) and have been in use since then. Recently the SciQLop project (SCIentific Qt application for Learning from Observations of Plasmas https://github.com/SciQLop/SciQLop) has become one of the intensive users (see Section 6.1.3). The Propagation Tool (see Section 5.2) proposes a functionality to plot in-situ data from AMDA, and in this aim, an additional webservice *getPlot* has been specifically developed for predefined parameters. Finally these webservices have been used to implement the HAPI protocol (Heliophysics Application Programmer's Interface, see Vandegriff et al. (2019)). This data access specification is a RESTful API and streaming format specification for delivering time series data (see Section 5.1).

2.5. Technical implementation and operational service

On the technology front, AMDA has always been a mix of different approaches, not always the most trendy and up-to-date, but solid enough to ensure long term operations and very few failures (the service is operational more than 99,8% of the time). The technology for the GUI is extJS (a Javascript extension from the Sencha company), and the web services use PHP. The core system is coded in C++ including the mathematical functions for data treatment. The server runs on a CentOS 7 Linux machine. This technology mix ensures AMDA is working on all available OS platforms and with all browsers.

The system described above, with the functionalities, capabilities and results detailed below, has been in place for about 15 years with several new versions coming over the years. The development principle is a classic continuous integration with additions and bug fixes first tested on separated instances before being pushed to the operational instance. Changes are therefore done on a weekly or bi-weekly basis; the users are warned of the most important changes in the announcements on the login page.

AMDA accesses, and associated workspaces, have been granted to almost a thousand persons worldwide. This includes normal logins (when a user asks it at amda@irap.omp.eu), the logins put in place for ESA/SSA (see Section 6.1.4) and those done for summer schools, workshops (see Section 6.5). Public access (no login, no mail required) is also possible for limited session (5 h) and the workspace is not saved. Over recent years, over 400 AMDA sessions per month are run (this does not include the access via web services). The number of simultaneous sessions is limited by the number of jobs running at AMDA server at the same time which must be less than 100, and less than 10 for an individual user.

3. Data

The architecture detailed above is tailored to one main task: serving data. The next section is devoted to the explanation of how these data are exploited. This section shows the variety of data AMDA can deal with which is in line with one of the missions of CDPP: ease data exploitation of missions exploring natural plasmas. This encompasses both heliophysics and planetary plasmas.

3.1. Data from large archives

As explained above, the objectives of AMDA are to present a wide variety of plasma data in an homogeneous way such that mathematical treatments and science analysis can be easily performed with standardized methods. This is the reason why AMDA restricts itself to a limited number of parameters types: time series of arrays of one or several dimensions. Large public archives generally contain more types: distribution functions, altitude profiles, spectra, maps and images but cannot offer the ability to plot and analyze all of them. The science committee of AMDA therefore makes a choice of which public data are duplicated in AMDA from the various large, agency-led databases that mainly comprise of the Planetary Data System (NASA/PDS[11]), the Coordinated Data Analysis Web (NASA/CDAWeb[12]), the Cluster and Double Star Science Archive (ESA/CSA[13]), and the Planetary Science Archive (ESA/PSA[14]). A data conversion is also applied as the internal AMDA format is netCDF (CDF to netCDF and PD3/4 to netCDF converters are used). Smaller, mission-focused databases are also accessed for data duplication (THEMIS at Berkeley/SSL, SWARM at ESA/EO). Original data are retrieved, converted and copied but specific uses may require some processing, for instance producing a new, averaged data sets or omni directional flux from multiple anode datasets (information on these treatments are shown to the user).

Finally, an important aspect of the databases is to maintain and regularly update orbits, trajectories and attitudes of spacecraft and celestial bodies. This is done thanks to the SPICE kernels from NAIF and ESAC (meta kernels are used for recent ESA missions[15]).

3.2. Data from simulations and models

Numerical simulation results are an increasingly important part of the means used to understand space physics processes. Modeled data, from which time series can be derived, are produced for case studies as well as for long term analysis.

AMDA was one of the central tools in the EU/FP7 IMPEx[16] project which promoted comparison between observations and simulations of planetary plasma environments (Romanelli et al., 2018; Modolo et al., 2018). Several types of simulation codes and analytical models were involved. The IMPEx environment offered the ability to fly virtual spacecraft into 3D simulations and models and to perform one to one comparison between simulated and observed time series of plasma parameters.[17] Similar features have been simultaneously developed in the 3DView tool (see Section 5.2) and some use cases are detailed in Génot et al. (2018a,b,c) including one on Mercury plasma environment which also involves AMDA data mining functionality (see Section 4.3). The numerical values of the time series are not stored within the AMDA database but accessed remotely via web-services. They can be accessed from the Workspace Explorer in the directory “Parameters/Remote Database: Simulations”. The articles cited above give full information on the IMPEx standards and protocols developed and used during the project and they will not be repeated here. An important point is that the list of data providers can be easily extended and it was done so at several occasions after the project itself ended. It should be noted that a prototype access to simulations at CCMC[18] was operational in AMDA before this facility started reworking a data model inherited from SPASE (see Section 5.1); this shift should now facilitate a larger scale deployment of this access. More recently the French ANR-funded project “TEMPETE” extended the IMPEx concepts to other codes and for time-dependent simulations (Romanelli et al., 2019). The new simulations will be equivalently visible in AMDA by re-employing IMPEx concepts.

At CDPP, along with regular updates of data from external databases, models are also run routinely and the results are accessible via AMDA. This is the case of the 1D MHD solar wind propagation model (Tao et al., 2005) which provides plasma and magnetic parameters at all planets. The 1D nature of the model permits rapid execution and light volumes of data such that datasets are produced from Tao et al. (2005)1990 onward for all planets. Solar wind conditions at spacecraft in cruise phase (Juno, BepiColombo) and around comet (Rosetta) are also presented. In AMDA these data are available in the “Solar Wind Propagation Models” directory. The same datasets are also accessible in the Heliopropa service detailed in Section 6.1.4 where more information on the execution architecture is given.

3.3. Community data

In some occasions data are produced by partnering laboratories or individuals who deliver data to AMDA, occasionally in combination with other data repositories (see Table 1). These data are transformed in netCDF by the AMDA team if needed. Table 2 summarizes the missions, instruments, producers and types of physical parameters which have been delivered over the years.

3.4. User data

The users can upload their own files (see Section 4.5 for specific formats) and hence creates their own database (this is even facilitated via a system of file masks and wildcards). The parameters which are thus defined can be plotted and combined like any other (internal) AMDA parameters. They are visible in “Parameters/My DataBase”.

[11] https://pds.nasa.gov/.
[12] https://cdaweb.sci.gsfc.nasa.gov/.
[13] https://www.cosmos.esa.int/web/csa/.
[14] https://archives.esac.esa.int/psa/.
[15] https://www.cosmos.esa.int/web/spice/.

[16] *http://impex-fp7.oeaw.ac.at/is no longer online*, see also http://impex.latmos.ipsl.fr/.
[17] This has been illustrated in an “ESA story” available at https://sci.esa.int/web/solar-system/-/60416-exploring-planetary-plasma-environments-from-your-laptop.
[18] https://ccmc.gsfc.nasa.gov/.

Table 1
Model contributions to AMDA database.

Model	Type	Producer	dataset/parameter(s)
Tao et al. (2005)	MHD 1D	C. Tao (NICT, IRAP)	magnetic field and plasma
Morschhauser et al. (2014)	analytic	A. Beth (IRAP)	Martian crustal magnetic field
Khurana's	analytic	A. Sicard (ONERA)	Jovian magnetic field
Connerney et al. (2018)	analytic	AMDA Team	Jovian magnetic field
Tsyganenko and Stern (1996)	analytic + data	AMDA Team	Earth's magnetic field
Ballistic propagation	analytic	AMDA Team	time delay

3.5. Private instances

In specific cases, an instrument team arranges for its data to be visible in AMDA only to a restricted list of individuals. This can happen during the proprietary phase of the data or during the generation of higher level (L3) data products or during the review of some specific datasets (following cross-calibration for instance). AMDA administration can grant access to a restricted list of users based on the logins. This enables a team with limited funding to access all AMDA functionalities without developing its own dedicated, costly system. Such an example, on a large scale, was the case of the Rosetta Plasma Consortium (see Section 6.2 for details).

4. Functionalities

4.1. Plotting data

This is the main functionality both in terms of usage frequency and in terms of design and coding. There is a large variety of plot types that can help visualizing space physics data, and AMDA had to specialize in a selection of them. The main type is "time series" where quantities are displayed as a function of time on the abscissa. Here time series include scalar, vector or spectrogram type of representation. For a spectrogram the ordinate can typically be 'energy', 'frequency', or 'altitude' and the third dimension is color-coded; a color bar is therefore mandatory and is usually displayed on the right of the figure. On time series plots one can also overplot constant values, filled area, or bounded intervals. AMDA allows all these representations. Apart from time series, scatter plots are also common and enable the representation of one quantity as a function of another one, as time evolves. A particular type of scatter plot is orbit plots where typically a reduced radius (i.e. $\rho = \sqrt{y^2 + z^2}$) is plotted as function of the planet-Sun distance (x). Over the life of AMDA, the plot interface has greatly evolved until it reached the current level of complexity.

On the interface, parameters may be dragged and dropped in the upper left part of the Plot Manager window which displays the plot layout (see Fig. 4); the bottom left part is for the time selection (a single interval, one or several TimeTables or catalogues). On the upper right, customized options are visible and pertain to the level of the plot layout which is highlighted (see also below). The bottom right part is for the output options (interactive or in file with format PNG, PDF, PS or SVG) and saving the layout for future uses; the saved requests will be displayed in the Operations tab of the Workspace Explorer (see Section 2.2). Finally the plot is produced by clicking on the "Plot" button. The data selected in the layout can also be downloaded directly from this layout with the "Get data" button. A pre-filled Download interface will then appear with the same parameters and on the same interval as on the Plot request. From the window presenting the resulting plot, it is possible to zoom in/out (with the mouse) or move the interval backward/forward in time (see the buttons visible on the upper left corner of Fig. 1).

Table 2
Dataset contributions to AMDA database by institute/individual.

Mission	Instrument	Producer	dataset/parameter(s)
Cassini	RPWS	LESIA (L. Lamy, B. Cecconi)	SKR/total emitted power, flux, polarisation, phase
Cassini	RPWS	IRAP (M. Holmberg)	Langmuir Probe/ion density proxy
Cassini	RPWS	LESIA (P. Schippers)	QTN/thermal electron
Rosetta	IES	SwRI (J. Burch, B. Trantham)	ion & electron spectra
Rosetta	MIP	LPCE/Imperial College	signal properties, density proxy
Geotail	EPIC	JHU/APL	ion & electron
Phobos 2	ASPERA, MAG	IRF (R. Ramstad)	ion & electron spectra, magnetic field
Galileo	EPD	IRAP (Q. Nenon)	ion & electron
STEREO	SWEA	IRAP	electron spectra
Venus Express	ASPERA-3	IRAP (A. Fedorov, E. Budnik)	ion & electron
Mars Express	ASPERA-4	IRAP (A. Fedorov, E. Budnik)	ion & electron
AMPTE-CCE	MEPA/MFE	RAL (C. Perry)	particles and magnetic field
AMPTE-IRM	MAG/PLA/PWS/SUL	RAL (C. Perry)	particles, plasma, waves and magnetic field
AMPTE-UKS	ELX/FGM/ION/PWS	RAL (C. Perry)	plasma, waves and magnetic field
ISEE	FVA	IRAP/CDAWeb/IAS	energetic particle fluxes
FERMI	LAT	IRAP (I. Plotnikov)	solar gamma-ray

Many more advanced customizations may be reached by clicking the "Extended plot options" above the plot layout. The detailed description of all options is however out of the scope of the present paper.

Finally, several Plot tabs can be defined, each of them resulting in a different plot window.

4.2. Constructing new parameters

One of the originalities within AMDA is its calculator, or parameter editor, which enables parameters to be combined from heterogeneous instruments and produce a new, user-defined parameter. The temporal resolution of this new parameter is also decided by the user. Once saved, the parameter can then be used in plots, conditional search (data mining, see next section) or be inserted in the construction of another derived parameter. Technically the derived parameter is not calculated and inserted in the database; it is rather internally defined as an XML file and computed on-the-fly each time it is used for one of the purposes listed above. Examples of common derived parameters are plasma β, Mach number, Alfvén velocity, solar wind - magnetosphere coupling functions … which require the combination of magnetic field and plasma moments.

Fig. 5 shows the general interface for this functionality with fields where users enter the name, the resolution, the unit and label which will appear on plots (Y title). Boolean operators (see below), classical constants (like μ_0, ε_0, planetary radii, speed of light, particle masses, …) and mathematical, vector and statistical functions are available in the menu in the top right of the interface.

"Conditional parameters" can also be defined. For instance $param2$, defined by $param2 = (param1 > 0)$, is equal to 0 for negative and null values of $param1$, and to 1 otherwise. One can then construct $param2 = param1 * (param1 > 0)$ which is equal to $param1$ when $param1$ is strictly positive, and to 0 otherwise. For instance, this can facilitate easy construction of variously classified plasma regimes of the solar wind from in-situ parameters (Xu and Borovsky, 2015). More generally this helps labelling data in a convenient manner, which is very useful in the statistical learning domain where feeding algorithms, like neural networks, with catalogues is a necessary step (see Section 6.1.3).

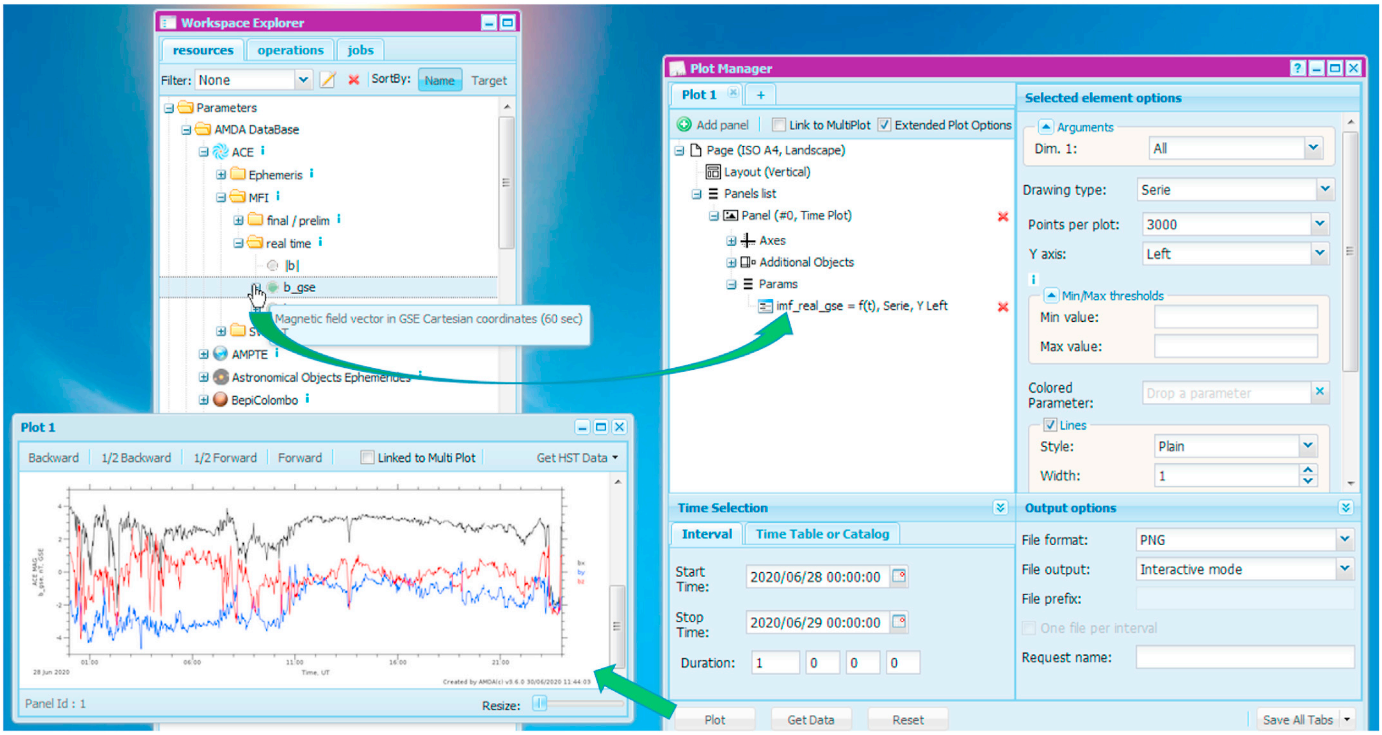

Fig. 4. Plotting a parameter in AMDA is realized by drag and dropping a parameter from the Workspace Explorer to the Plot Manager (upper green arrow). The plot interface shows the panel composer (upper left), the time selection (bottom left), the plot customization options (upper right) and output selection (bottom right). Activating the "Plot" button produces a figure (second green arrow). The example shows "Real Time" ACE magnetic field vector in GSE (with internal parameter ID = 'imf_real_gse') selected for 1 day. A tooltip is shown when the mouse is over a parameter with the full parameter name, resolution and unit.

4.3. Data mining

This functionality also came with the earliest version of AMDA in 2005. Together with the easy "drag&drop" plotting capability it was the most salient feature of AMDA at the time. Digging into large amount of data with minimal coding from a web interface was a novelty. The goal of this functionality is to produce TimeTables, or lists of events.

A data mining condition is a mathematical expression which is assessed (true/false) over a given time interval, or over a number of TimeTables. When the condition is held 'true' over a series of consecutive records, the interval constituted by these records is returned as an 'event'; usually multiple such events form the TimeTable (one single event is also possible). In the editor, the expression is syntactically built from the name of parameters as they are stored in AMDA (see Section 2.1.1), pre-defined mathematical functions and boolean operators ($<$, $>$, &, |, = , ! =) as shown in Fig. 6. In the interface the user is asked to enter the name of the data mining request together with the time step over which all AMDA parameters used in the expression will be re-sampled (averaged/interpolated) before calculations. The expression may be constructed in a similar fashion to the derived parameters as explained in the previous section. Once the request is complete the "Do Search" button must be hit. Depending on complexity of the request and the size of the time intervals to be scanned, the results are quickly displayed or the request goes into batch mode. In any case the resulting TimeTable may be edited and/or saved for future use.

4.4. Operations on TimeTables and catalogues

In the TimeTable GUI some statistics are provided; for each TimeTable the following quantities may be displayed: the minimum, maximum, mean, and median durations together with the "density", i.e. the proportion of all intervals over the global interval (defined with the minimum start time and the maximum stop time). The intervals of a TimeTable can be extended or shifted by a given amount, and overlapping intervals can be merged. These actions are logged into the 'Operation log' windows.

There is also a dedicated window for applying "merge" (union) and "intersect" operations on multiple TimeTables.

Catalogues can also be edited in a dedicated window or manipulated for visualization in another window. At this stage only simple plotting functionalities are available for catalogues, and are restricted to scatter plot between parameters (basically one column of the catalogue is plotted against another one).

4.5. Upload and download

All AMDA objects described above may be uploaded and/or downloaded. Of course uploading is possible for data in specific formats, and similarly downloading is also constrained, but AMDA is able to manage most of both syntactic and file formats which are of common use in space physics. This means for upload,

- for data files: ASCII, netCDF, CDF, VOTable
- for TimeTables and catalogues: ASCII, VOTable

for download,

- for data files: ASCII, Json, CDF, VOTable
- for TimeTables and catalogues: ASCII, VOTable

4.6. Statistics

The last functionality exploits all concepts presented before and enables generation of catalogues from the combined use of TimeTables, parameters and statistical functions. Basically, for a given TimeTable, a column of the catalogue will be formed from the result of applying a

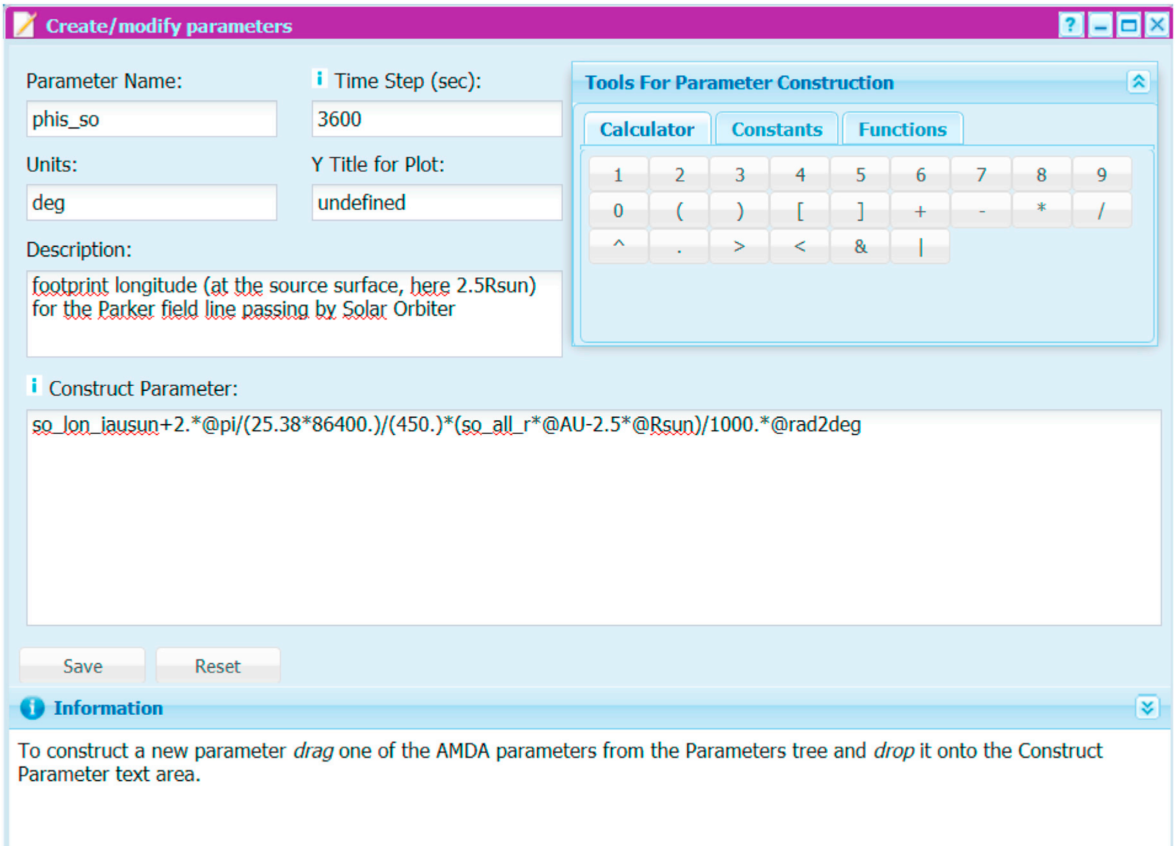

Fig. 5. The parameter editor interface. The example shows how to construct the magnetic footprint longitude φ_s of the Solar Orbiter spacecraft as given by Equation (1) (see Section 6.1.5); the time resolution is 1 h and the unit 'deg' for degree. The parameter IDs which appear in the "Construct Parameter" edit window have been obtained by drag&dropping of the corresponding (explicit) parameters from the Workspace Explorer (similar to the drag&drop action for the Plot, see Fig. 4). The variables starting with "@" are built-in constants found in the "Constants" tab (common physical constants, planetary radii, …).

statistical function (mean, max, min, median, standard deviation, skewness, kurtosis) on a chosen parameter on all the intervals of the TimeTable. The catalogue can then be saved in "My Catalogue" directory.

For instance, for the intervals of Table 3, the distance between Solar Orbiter and BepiColombo can be defined as a new parameter (Section 4.2) and a catalogue can be generated by computing the min/max/mean of this distance.

Whether a catalogue is produced from the Statistics functionality or imported by other means, AMDA offers the ability to represent it in various ways. The visualisations are not shown but consist of histograms (or the distribution of all values of one column) and scatter plots (one column of the catalogue is plotted as a function of another one). This enables a quick way to analyze/visualize a catalogue. Those functions are however not extended as they are not the core of expertise of AMDA; companion tools like Topcat[19] should be favored for refined tabular analysis (Taylor, 2005), for instance through the use of SAMP (see next Section).

5. Interoperability

5.1. Insertion in the IHDEA

The ability for a tool to broadcast/communicate its data is important, but integrating an eco-system of diverse tools where data and functionalities can be "mixed" is a further crucial step of development. It enables the tool to benefit from external developments (standards, protocols), and this also helps the tool getting known and used by a broader community. This is the general philosophy behind the establishment of the International Heliophysics Data Environment Alliance (IHDEA[20]) which recently (end 2019) edited its charter and bylaws with CNES (through the intermediary of CDPP) as a member. IHDEA, for space physics, joins the two other alliances for astronomy (IVOA) and planetology (IPDA).

Concerning standards, the early adoption use of SPASE for the proper description of AMDA data is certainly the most emblematic move toward interoperability. SPASE ("Space Physics Archive Search and Extract" http://www.spase-group.org/) started in 1997 and CDPP joined very quickly after if was created (1998). Roberts et al. (2018) has an extended discussion of the basic concepts and utilizations of SPASE, including those of AMDA. At the beginning, for historical reasons related to how the main server was built, the internal AMDA Data Model was proprietary. However, during a major rebuild of AMDA, the SPASE Data Model was adopted for the benefit of users and developers. As discussed previously, the missions in AMDA may be space missions as well as ground-based observatories or simulations. These concepts are mapped to SPASE resources: Observatory, Instrument, and Numerical Data. The SPASE XML files are stored in a registry and used to display information at each level in the Workspace Explorer (see Section 2.2); SPASE metadata are also displayed as titles for axes on the plot. The SPASE group has also promoted the development of the Heliophysics API (HAPI[21]) with the intention of providing a single route to all heliophysics time series data with easily adopted methods for servers and clients. Very recently

[19] http://www.star.bris.ac.uk/mbt/topcat/.

[20] https://ihdea.net/.

[21] https://github.com/hapi-server/.

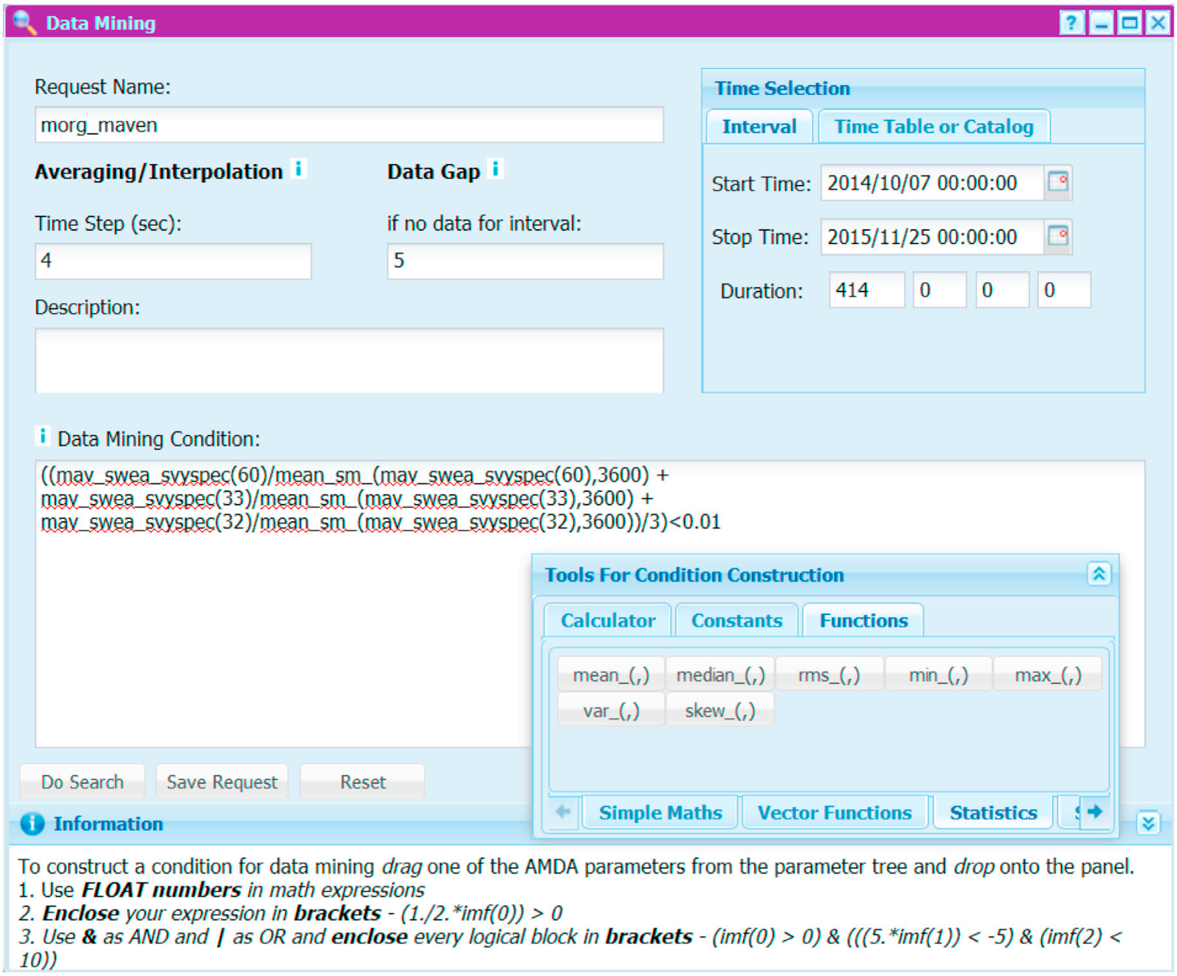

Fig. 6. The conditional search interface. The windows are filled with parameters extracted from the MAVEN use case detailed in Section 6.1.1.

Table 3
Time intervals of the magnetic alignments between BepiColombo and Solar Orbiter for Feb. 2020–Dec. 2025 with $u_{SW} = 450$ km/s obtained with AMDA data mining (see text for the condition).

Start time	Stop time
2021-07-18T14:30:00	2021-08-19T19:30:00
2021-10-08T20:30:00	2021-10-10T12:30:00
2022-07-16T04:30:00	2022-07-17T09:30:00
2024-02-15T03:30:00	2024-02-16T12:30:00
2024-04-05T18:30:00	2024-04-06T18:30:00

this protocol has been integrated in AMDA such that data are now distributed via HAPI thanks to the use of the official node-js HAPI server and the implementation of a binding to the AMDA REST web services.[22] This will enhance the visibility of AMDA and its datasets via a facilitated access; for instance plotting AMDA datasets in Autoplot (Faden et al., 2010), another popular plotting tool, is now straightforward.

From the earliest version of AMDA, the integration of the Simple Access Messaging Protocol (SAMP) developed by the astronomy community (IVOA) has been ensured. This is continued today and AMDA can now broadcast data (parameters), TimeTables and catalogues to companions tools connected via SAMP. More generally applications of this protocol within the CDPP tools have been discussed in Génot et al. (2014) and in Génot et al. (2018a,b,c). Recently however technical incompatibilities between SAMP and 'https' have been revealed[23] which may compromise the sustainability of its use (in AMDA and elsewhere).

Another protocol implemented in AMDA to open the tool to wider uses is EPN-TAP (Erard et al., 2018) developed in the frame of Europlanet projects via the VESPA consortium. VESPA is providing an infrastructure to search for data products based on science content and coverage query parameters (Erard et al., 2020). This is enabled through a data model called EPNcore (Europlanet core metadata), coupled with the Table Access Protocol (Dowler et al., 2018) developed and maintained by the IVOA. Many heliophysics related data collections are available through the VESPA query interfaces (either the main VESPA query portal, http://vespa.obspm.fr, or from interfaces embedded in tools). The mapping between the relevant EPNcore and SPASE dictionary elements has been done by the AMDA team for the internal database.

The same registry that is used in AMDA is also used for VESPA. The EPN-TAP protocol allows access to granules, which are typically files. Each file is described in a row of the relational database, and the SPASE XML registry is used to populate the rows, after a translation from SPASE to EPNCore. Some metadata not defined in EPNCore are added as optional for EPN-TAP. For example "observed region" (SPASE) is translated to "spase region" (EPNCore). Data from AMDA may be searched from VESPA, using the common registry. Users searching for data of interest through, for example, the VESPA portal may choose several search criteria, most of them compatible with SPASE (instrument, timespan …)

[22] http://amda.irap.omp.eu/service/hapi.

[23] See the IVOA report https://wiki.ivoa.net/internal/IVOA/InterOpMay2016-GWS/tlsamp.pdf.

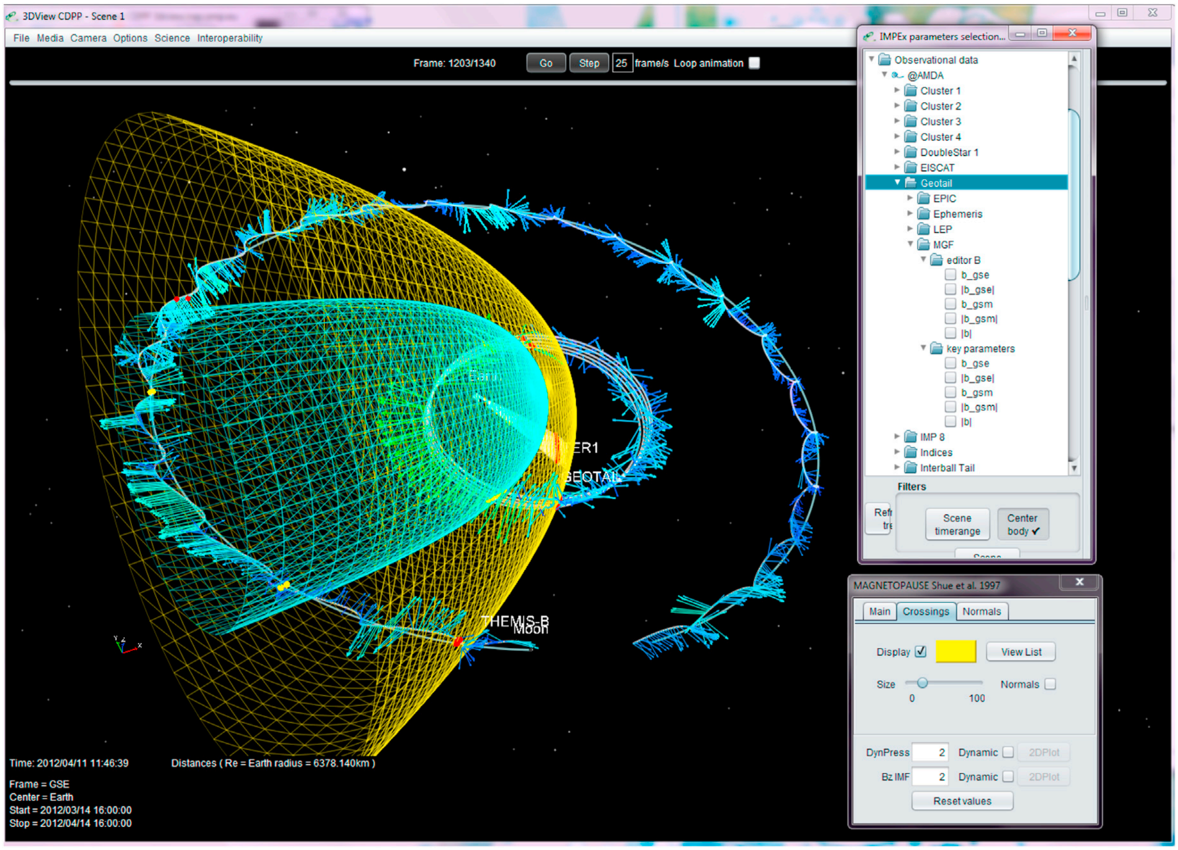

Fig. 7. Themis-B and Geotail magnetic field vectors obtained from AMDA and plotted in a magnetospheric scene in 3DView (the amplitude of the field is coded in both length and color). The main window shows about a month of data (mid-March to mid-April 2012): the trajectory of Themis-B orbiting around the moon, and several orbits of Cluster 1 and Geotail; also shown are the bowshock (Fairfield, 1971) and magnetopause (Shue et al., 1997) modeled surfaces in yellow and blue respectively. Red and yellow bullets are the crossing locations of theses two boundaries by orbiting objects, spacecraft and the moon. The inset windows on the right shows the 3DView access to AMDA (top) and the control panel for the magnetopause model bottom (bottom). Other control panels are not shown.

and get a list of files (one file per row) with a link to access the actual data. The same operations are possible with other VESPA clients like 3DView (see below).

5.2. Communication with other tools

The interoperability of CDPP tools at large is also discussed in the respective reference papers for 3DView (Génot et al., 2018a,b,c; Génot et al., 2018a,b,c) and the Propagation Tool (Rouillard et al., 2017). 3DView enables the display of 3D planetary and heliospheric scenes with spacecraft trajectories along which observed plasma data can be plotted as illustrated on Fig. 7. An access to AMDA data via its web-services has therefore been implemented into 3DView. For a given scene including spacecraft, the AMDA database is filtered such that the user is presented with the most relevant data. The Propagation Tool allows users to connect solar perturbations seen at the Sun and their in-situ observations at planets and probes by computing time shifts with several models (radial or SEP propagation, co-rotation). A built-in link to AMDA, based also on the web services, is presented to the Propagation Tool user to display in-situ data in a dedicated window as illustrated on the use case shown on Fig. 8. AMDA and Propagation Tool additional functionalities are complementary for following ICMEs in the heliosphere: the spacecraft hit by a given CME are predicted with the Propagation Tool, then the data from these spacecraft can directly be plotted in AMDA to analyze the in-situ plasma properties (as illustrated, for instance, in Witasse et al. (2017) and Grison et al. (2018)). Finally, in order to simplify the way users access CDPP data, the long-term CDPP archive[24] at CNES (called SIPAD, to be replaced by REGARDS mid 2021) has implemented the web services to give full exposure and access to AMDA datasets in a dedicated branch "Missions@AMDA"; this is illustrated on Fig. 9.

The implementation of EPN-TAP and SAMP (see previous Section) enables the direct visualization, within AMDA, of data from other compliant archives like, for instance, APIS[25] for auroral imagery at giant planets; such use cases have already been described in André et al. (2011); Génot et al. (2014). Similarly the CLIMSO[26] (Pitout et al., 2020) solar imagery data could be used in relation with in-situ measurements from Solar Orbiter or PSP.

To investigate micro-physics of plasma processes AMDA can also be used for large scale browsing and the production of TimeTables of localized events before more focused tools like CL (or CLweb[27]) are employed to display and analyze detailed distribution functions; indeed CL also complies with the same TimeTable format as AMDA does.

6. How useful?

In this last section, we present a number of use cases encompassing the variety of ways AMDA has been exploited for scientific purposes, data distribution and education in the space science and planetology

[24] https://cdpp-archive.cnes.fr/.

[25] http://apis.obspm.fr/.

[26] http://climso.irap.omp.eu/.

[27] http://clweb.irap.omp.eu/.

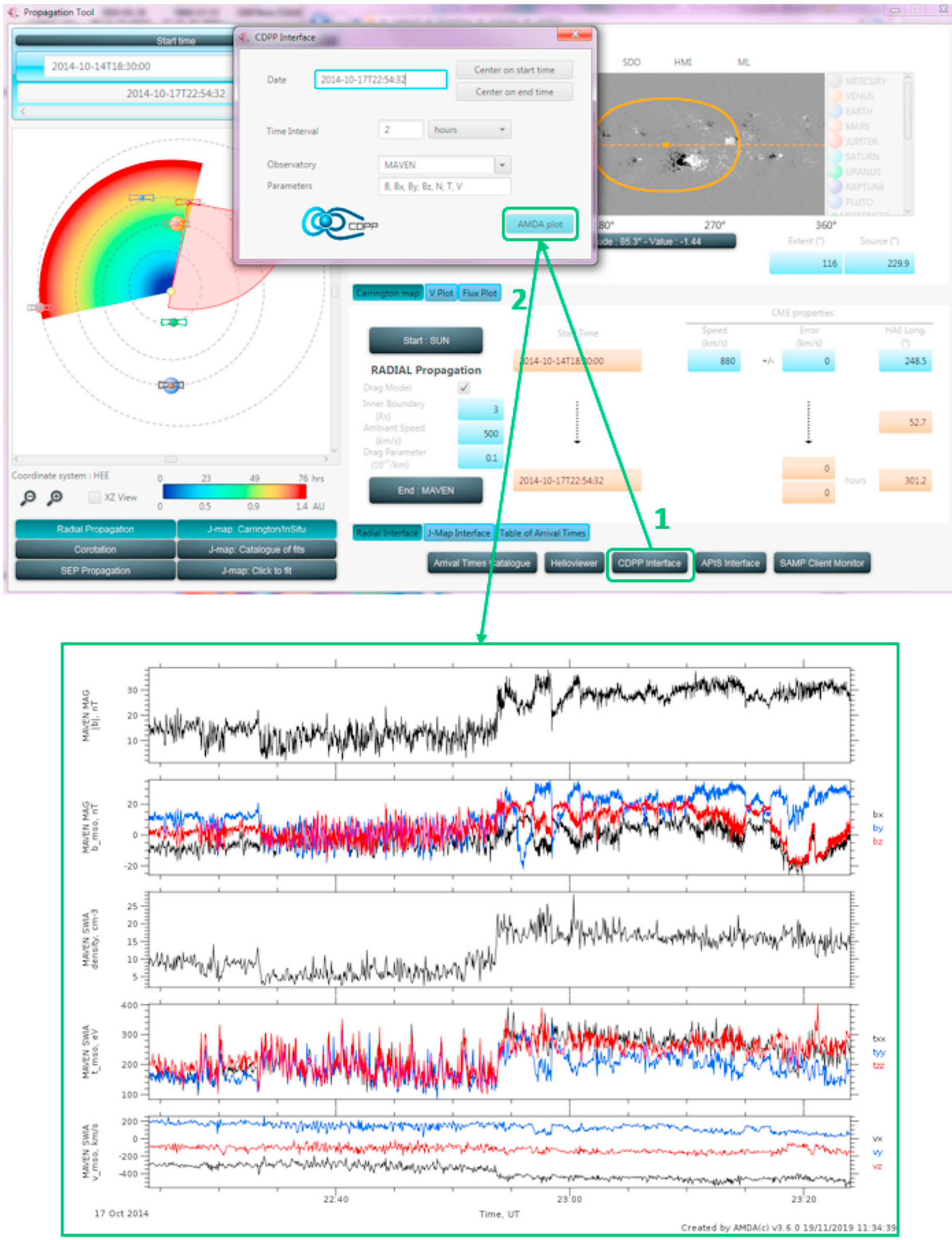

Fig. 8. Study of the 14 October 2014 (Sun time) Interplanetary Coronal Mass Ejection (ICME) with the Propagation Tool and AMDA; a full report is available in Witasse et al. (2017). The above plot shows the Propagation Tool interface in which the ICME parameters have been filled to compute the propagation from the Sun to MAVEN with a predicted arrival on 2014-10-17T22:54:32. By clicking on "CDPP interface" (1) the small inset window enables users to choose which in-situ parameters are plotted in AMDA. Here MAVEN magnetic field and plasma moments are chosen for 2 h centered on the arrival time. Clicking on "AMDA plot" generates the bottom plot (2) in the browser: within a few tens of second from the centre time (predicted by the Propagation Tool) the jumps in all parameters are typical of the ICME passage.

Fig. 9. The SIPAD interface with a similar “look & feel” as the main CDPP server. On the left part the data tree shows the data directly available at the SIPAD on the top level whereas the bottom level (starting at the red insert) shows a replication of the AMDA tree. When AMDA data are ordered from the SIPAD the AMDA web-service is called to fulfill the command.

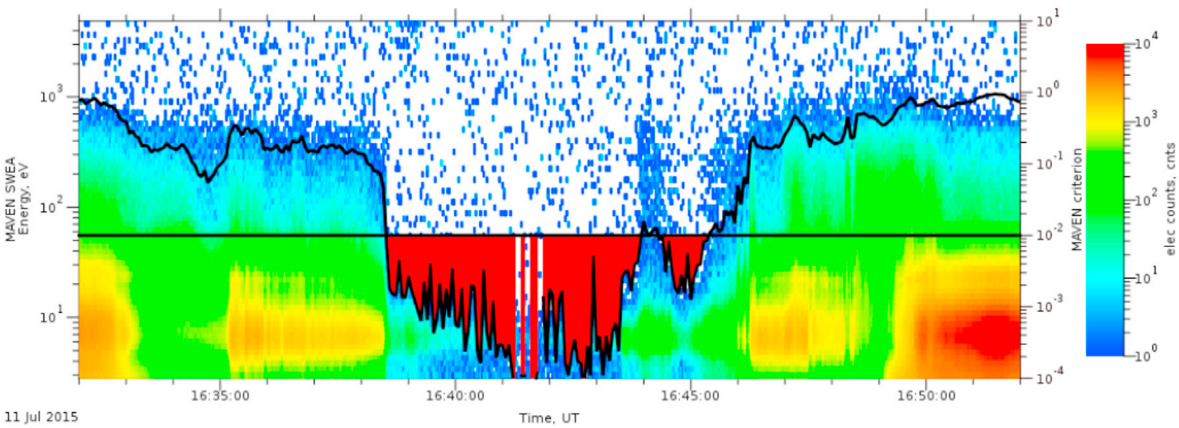

Fig. 10. Electron depletions (red regions between black lines) seen by MAVEN and detected in AMDA from the criterion proposed in Steckiewicz et al. (2017) (see the text for details). The spectrogram is from the SWEA instrument (electron counts).

communities.

6.1. A variety of usages: use cases and community work

6.1.1. Cataloguing, selection of events: joint analysis of MAVEN, MEX, and MGS data

Steckiewicz et al. (2017) analyzed observations of electron depletions from three different Martian missions: Mars Global Surveyor (MGS), Mars Express (MEX) and Mars Volatile EvolutioN (MAVEN), in order to better characterize their altitude and geographical distributions and understand their formation processes. These structures, first observed with MGS data (Mitchell et al., 2001) that shows abrupt drops of the electron count rates in the nightside ionosphere, are among the features that reveal how the nightside ionosphere is irregular, spotty, faint and complex. Steckiewicz et al. (2015) then showed with MAVEN data that these structures are not plasma voids but suprathermal electron depletions, and these depletions are favored by the presence of magnetic crustal field sources. Steckiewicz et al. (2015) and Steckiewicz et al. (2017) tuned a detection algorithm directly in AMDA to automatically detect suprathermal depletions, based on the ratio between raw electron count rates of the spacecraft and their temporal sliding means (over 1 h). The criterion used for MAVEN (Equation (1) of Steckiewicz et al. (2017)) is coded in AMDA. It is seen on Fig. 6 in the "Data Mining Condition" window. In this window the values 60, 33, 32 correspond to the energy channels used for the detection (21.2 eV, 95.04 eV, and 103.25 eV respectively as mentioned in Steckiewicz et al. (2017)), mean_sm is the sliding mean function in AMDA (computed over 3600s) and mav_swea_svyspec is the AMDA parameter for the electron counts of MAVEN/SWEA.

The search through this criterion resulted in thousands of events and one of them is illustrated on Fig. 10. Steckiewicz et al. (2017) were thus able to provide the first multi-spacecraft analysis of electron depletions covering seventeen years of Martian exploration, offering a comprehensive view of the phenomenon based on the same detection method for the three missions.

The results showed that electron depletions are spread on the nightside of the Martian environment at altitudes between 110 and 900 km. The aerographic distributions of electron depletions for each mission produced results in agreement with each other: electron depletions are strongly linked with the horizontal crustal magnetic fields. Nevertheless, the crustal field influence disappears below a transition region near 160–170 km altitude regardless the hemisphere. The crustal fields act indeed as a barrier preventing the replenishing of the electron depleted areas by external incoming electrons, the depletion being due to the absence of solar extreme ultraviolet (EUV) photoionization that could counterbalance the loss by the interaction with the CO_2. However, below the transition region, MAVEN data revealed globally homogeneous distributions of electron depletions: at low altitudes, crustal magnetic fields are no longer predominant in the creation of electron depletions, the dense atmospheric CO2 population becomes responsible for creating the depletions at those altitudes by electron absorption processes.

Electron depletions can be observed to some extent until the terminator region. Since the absence of replenishing is essentially due to the absence of photoionization by EUV photons in the nightside, the limit of detection of these depletions can be used to detect the martian EUV terminator. Steckiewicz et al. (2019) thus used MAVEN depletions with the same AMDA detection algorithm to calculate the location of the EUV terminator, and were able to show it is on average at 120 km above the optical terminator. Moreover, it is higher on the dusk side than on the dawn side at equinox, thus providing insight into the atmospheric seasonal asymmetries.

6.1.2. Cataloguing and selection of events: characterization of the ion flux intensities around L2 in the perspective of the ATHENA mission

ATHENA is the next ESA L-class observatory for X-Ray observations with a planned launch in the early 2030s. The goal of the ATHENA mission is to provide X-ray imaging and spectroscopy capabilities with sensitivity and resolution much better than previous missions. For reaching this goal, the reduction of the instrumental background is a fundamental issue and comprises one of the main mission requirements. The background of the X-ray instruments is strongly affected by the radiative environment and in particular by the "soft protons and ions" in the range of tens to hundreds of keVs (typically 40–200 keV). As with many other ESA missions dedicated to astronomy (Planck, Herschel, Gaia, …) the orbit of ATHENA was initially foreseen around the L2 Sun-Earth Lagrangian point. However the Earth's magnetosphere is an important particle accelerator especially when the geomagnetic activity is enhanced, generating energetic ion populations which are expected to significantly contribute to the flux intensities at L2.

In this frame, a key task of the ESA AREMBES project (contract No 4000116655/16/NL/BW. ACE EPAM/LEMS120) was dedicated to the systematic analysis and the characterization of the particle fluxes in the 40–200 keV energy range. These results have then been used as input to simulations of ATHENA instrument performance, such as the X–IFU detector (Lotti et al., 2017). This original study has been performed by using AMDA for massively exploiting the GEOTAIL/EPIC (Williams et al., 1994) data in a few steps:

- The first step was to analyze and to treat Christon's catalog (Christon et al., 1998) which divides the 2 years "distant tail" GEOTAIL campaign with respect to its transit inside the various magnetospheric regions (solar wind, dawn & dusk magnetosheath,

magnetotail, boundary layer, plasma sheet, lobes). From this catalog, 14 time-tables have been generated, each one associated with a particular region and a level of geomagnetic activity.[28]
• These time-tables were then used to map the likelihood of encounter of the various magnetospheric regions around L2. The time-tables were used to extract the data from AMDA in respect to the observed region. This allowed to compute the occurrence frequency of GEOTAIL (or any other probe) in each particular region.
• The time-tables were also used to extract (from the AMDA database) subsets of detailed (including species, energy and angular information) GEOTAIL particle data with respect to the observed regions and geomagnetic activity. The data were then analyzed outside of AMDA for characterizing the statistical properties of the flux intensities, fitting the spectra and computing their cumulative fraction with respect to the regions, the particle species, their energy, their streaming direction and the geomagnetic activity.

A comparable analysis has been conducted (Budjáš et al., 2017). However the AMDA-based analysis above reached more advanced conclusions, notably on the direction anisotropy, the ion species (not only the protons), the link to the geomagnetic activity, and the magnetospheric regions (not only the magnetosheath).

The AREMBES project has later been extended to analyze the GEOTAIL/EPIC data obtained in conditions representative of those encountered around L1 through a full solar cycle. This has been done by building a first time-table by selecting the periods of transit of GEOTAIL inside the solar wind. Then this time-table was reduced by excluding the time-intervals when GEOTAIL was likely in the foreshock by testing the connectivity of the spacecraft to the shock from the magnetic field measurement and a simplified shock model (depending only on the ram pressure). This time-table was finally ingested into AMDA to extract the detailed particle GEOTAIL/EPIC data and perform a study similar to the one realized to characterize the particle environment at L2. The results obtained from GEOTAIL measurement were thereafter confirmed by extending the study with the ARTEMIS data. AMDA was then used for searching the periods during which the ARTEMIS probes provide simultaneous survey of the solar wind and of the distant tail.

The main conclusions of these studies, illustrated on Fig. 11, indicate that choosing an orbit around L2 does not bring advantage with respect to the 40–200 keV proton fluxes:

• the Earth's magnetosphere does not offer any shielding at the L2 location,
• ions accelerated inside the magnetosphere consist of an additional population which contributes significantly to the ion fluxes at these energies.

Another study was performed inside the AREMBES project where AMDA was used as data provider. This work was dedicated to the characterization of the higher energy ion fluxes with the help of multi-spacecraft measurements (Laurenza et al., 2019).

6.1.3. AMDA in the machine learning era: use of web-services

The large amount of data held in the AMDA database associated with the availability of (REST) web-services which enables easy access to parameters makes AMDA a very suitable repository for machine learning applications. The 'learning' phase of these techniques indeed requests the ingestion of numerous 'events' characterized by one or multiple parameters. From simple command lines one or multiple parameters can be retrieved on one or multiple intervals (TimeTables). The example below shows how to build the command in a Python code.

One strength of the AMDA architecture is that it offers an identical access method for parameters from the database as well as user defined parameters. This means 'paramID' in the command above can be the magnetic field magnitude measured by Cluster1 or whatever complex parameters that the user has defined in their workspace (for instance an integration over energy channels of a particle spectrogram). This provides a powerful means to embed pre-processing in a pipeline where this pre-processing is done by the AMDA server. This has the advantages of reducing the data transfer over the network while the coding is lighter (in the example above only the scalar integral will be transferred instead of the full spectrogram).

To date, the AMDA database and web-services have been used for several machine learning applications including plasma boundaries detection around Mars/Venus (from Mars & Venus Express, MAVEN data, see (Garnier, submitted)) and electron diffusion region associated with magnetic reconnection at the magnetopause (MMS data, see (Lenouvel et al., 2021)). More applications have also been investigated in the frame of a contract between CNES and AKKA (IT company). The ultimate goal is to provide AMDA users with catalogues of events which are automatically updated upon new data ingestion. This kind of development is currently in a prototyping phase. For instance on Fig. 12 an illustration of the global pipeline for bow shock detection at Mars is illustrated. On the right upper part of the Figure, time series of plasma parameters are plotted with a colour chart describing the detected regions: solar wind, bow shock or planetary environment. The general principle will be the same for each envisioned applications: a model is trained (beforehand by the AMDA team and collaborators) to recognize regions or events and produce a dedicated catalogue. The catalogue is then interpreted (like on Fig. 12 with the colour chart) to give the user an enhanced view of the raw data. Proposing value-added products like these predictions is one of the main objectives of applying machine learning techniques on the AMDA large database.

This novel type of application is at the core of the development of the SciQlop application. This new software, which shares some similarities with AMDA on the general objective of simplifying plasma physics data analysis, mainly focuses on data tagging and labelling with the aim of easily producing and sharing catalogues of events ready to be serialized in machine learning pipelines. To do so, SciQLop embeds AMDA web-services to access data and will present an ergonomic GUI to manually select or automatically detect events of interest with learning techniques mentioned above. The on-going partnership between AMDA and SciQlop could eventually lead to a new tool pushing space data analysis a step further towards massive statistical studies of plasma processes.

6.1.4. Model run at AMDA: the HelioPropa service

Following a collaboration with C. Tao (now at NICT, Japan), a version of her 1D MHD code (Tao et al., 2005) for solar wind propagation has been implemented in AMDA. The architecture is the following: solar wind observations at L1 (magnetic field and plasma moments) available in AMDA are taken as inputs to the model which is run automatically every day to produce estimates of solar wind conditions at distant locations, planets and spacecraft in their cruising phases (for such large scale estimates, conditions at the planet and at orbiting spacecraft are identical). Backward propagation is applied toward Mercury and Venus, and also toward the inner heliosphere spacecraft, Solar Orbiter and Parker Solar Probe. Cruising spacecraft include Juno, Rosetta and BepiColombo. The sources of (observed) inputs are from STEREO A/B, ACE Real Time, DSCOVR and OMNI. Therefore solar wind conditions are computed at a target (planet/probe) for all source-target pairs. This amounts to a rather unique and complex product which can serve as a heliospheric monitor of solar wind conditions over a broad range of time (parameters are available from 1990 onwards). It directly exploits L1 data held in AMDA to produce new datasets. To attract more exposure for this service, Europlanet RI funding was employed to develop a dedicated interface (in the

[28] These TimeTables are accessible in the AMDA "Shared TimeTables/Earth" directory.

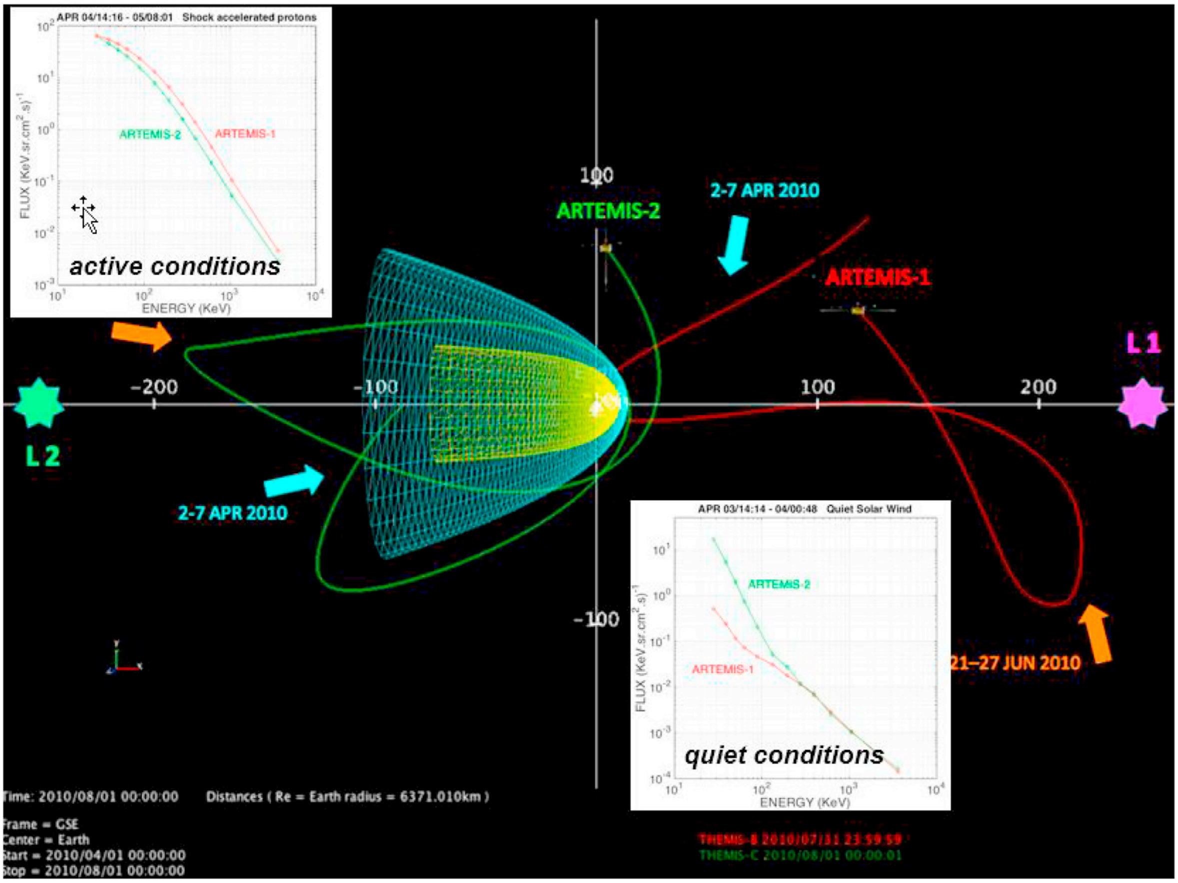

Fig. 11. Comparison of the spectra obtained by ARTEMIS-1 in the solar wind and ARTEMIS-2 in the distant tail or magnetosheath during active (upper left) and quiet (bottom right) solar conditions. Because the particle environment at L2 is likely to be worse in terms of ATHENA background, the AMDA-based analysis showed there is no benefit to placing ATHENA at L2 versus L1 (see text for details). The background image is a 3DView XZ GSE view of the magnetosphere between L1 and L2 (magnetopause and shock models are seen in yellow and blue respectively) with trajectories of both ARTEMIS spacecraft.

```
def buildParamURL(start_time, end_time, paramID, token):
 get_url = REQUEST_URL + '?' + 'startTime='+start_time + '&stopTime='+end_time+
 '&parameterID='+paramID+'&token='+token+'&sampling='+str(SAMPLING_TIME)+
 '&userID='+USERNAME+'&password='+PASSWORD
 return get_url
```

frame of the PSWS work-package[29]). This interface is available at http://heliopropa.irap.omp.eu/and shown on Fig. 13. To complement the solar wind magnetic field and plasma moments at planets and probes, an heliospheric view in HEE coordinates is available to help figure out what is the planetary configuration at a given time. This service, with its underlying daily computation, is an example of the diversity of products that AMDA architecture can offer, beyond those of a simple database. Note that HelioPropa is also available through the ESA/SSA Heliospheric Weather portal[30].

6.1.5. BepiCoordObs

A very recent application of AMDA took place in the frame of a working group set up by ESA (lead. L. Hadid, LPP, France) to find windows of opportunity for coordinated analysis between BepiColombo and other heliospheric spacecraft, mainly Solar Orbiter and Parker Solar Probe. AMDA was used to find relevant orbital configurations in order to study Coronal Mass Ejection (CME) or Solar Energetic Particles (SEP) propagation for instance and for which switching on BepiColombo instruments would make a difference. No science was indeed planned during the cruise phase and it was initially envisaged to have BepiColombo off in that respect. The purpose of the BepiCoordObs working group is therefore to propose interesting intervals where this stringent condition could be relaxed. Two main 'spacecraft and planet' alignment configurations were sought for

- radial alignments between 2 or more bodies, i.e. when the bodies were in a small latitude/longitude cone from the Sun
- magnetic alignments between 2 or more bodies, i.e. when the footprints of Parker field lines passing through the bodies were in a small latitude/longitude region at the source surface (R_s).

For the first alignment a typical condition was defined by a cone of 10° in longitude and 5° in latitude. For the second alignment the longitude of the Parker field line at the source surface has first to be computed

[29] http://planetaryspaceweather-europlanet.irap.omp.eu/.
[30] http://swe.ssa.esa.int/heliospheric-weather.

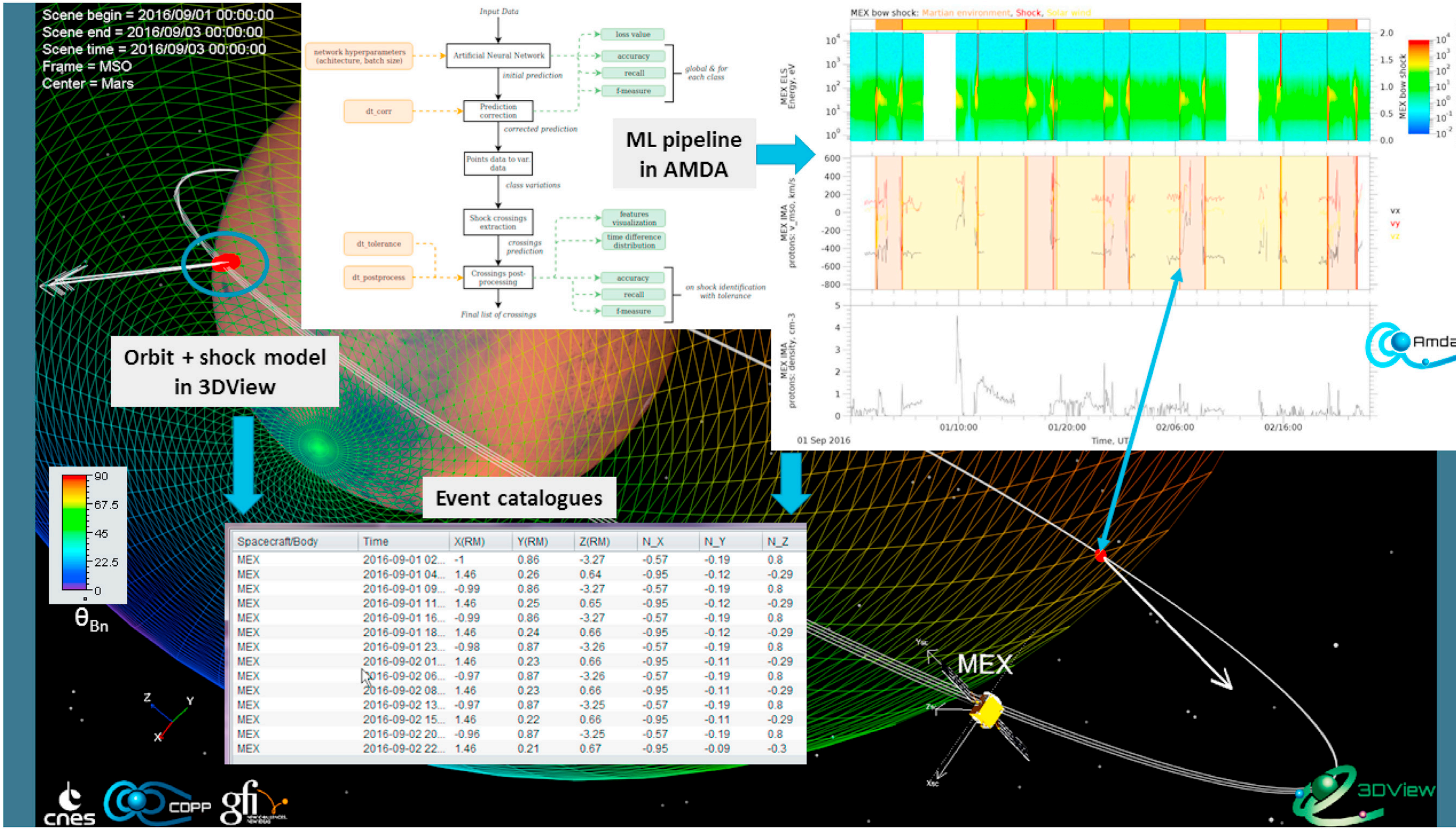

Fig. 12. Detection of plasma regions around Mars from Mars Express data: a supervised machine learning algorithm (top left) has been trained to detect solar wind/bow shock/planetary environment. The 'predicted' regions form a catalogue (bottom left) are shown to the AMDA user both as a colour chart above the data and with coloured intervals (top right) as an help to understand the observational context. The underlying illustration is a 3DView scene which makes use of the same catalogue for 3D display of the shock crossings and normals by the MEX spacecraft. The wired surface is a shock model coloured by the values of the shock angle (θ_{Bn}).

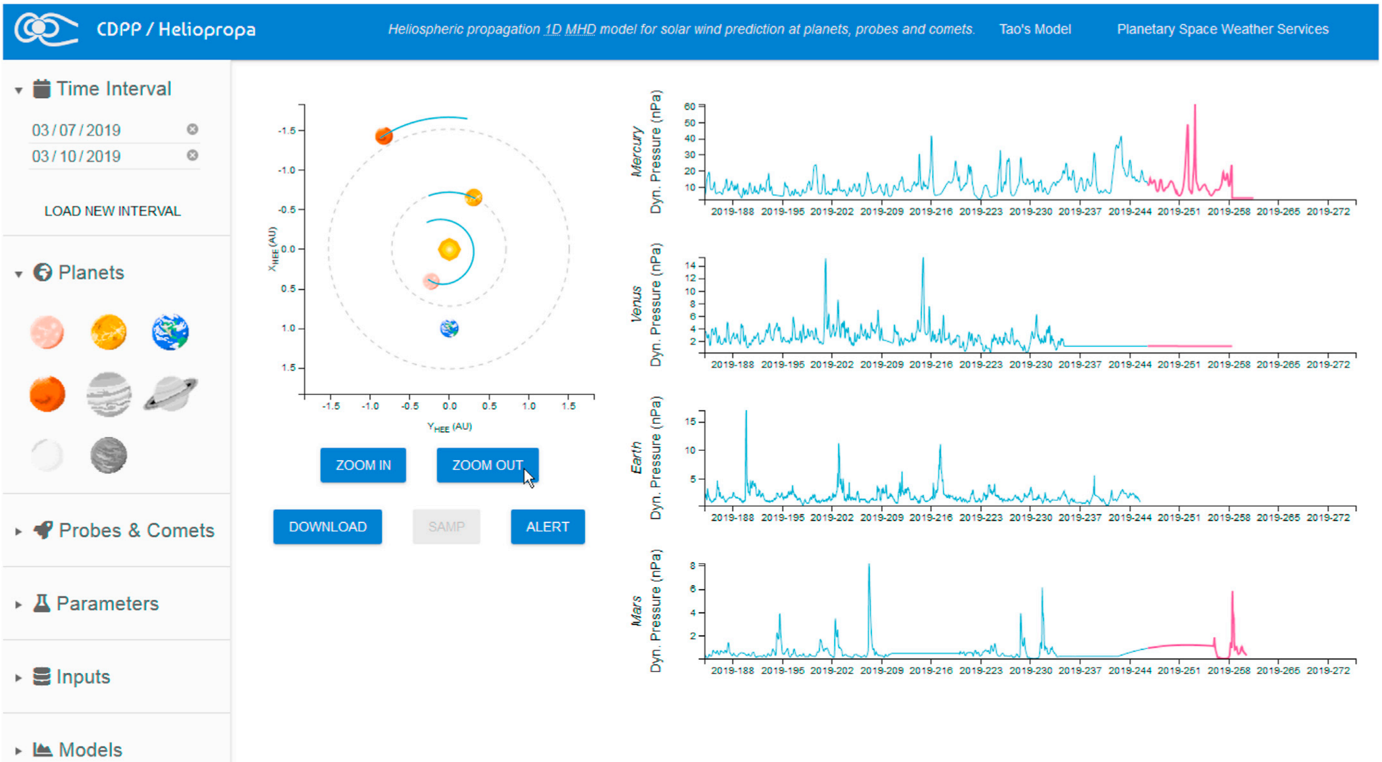

Fig. 13. A view of the HelioPropa service that displays solar wind conditions at planets and probes computed from observed conditions at L1 with the 1D MHD model by Tao et al. (2005). The heliospheric configuration is shown on the left for the same interval as the time series of magnetic field and plasma moments on the right.

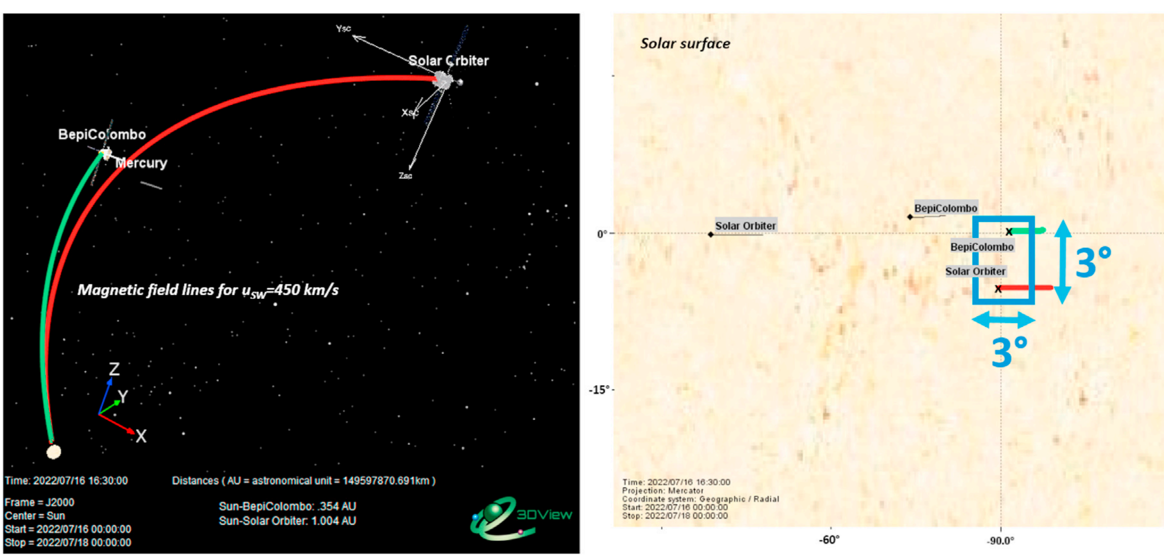

Fig. 14. 3DView illustration of the solar magnetic connectivity of BepiColombo (green) and Solar Orbiter (red) for the third interval of Table 3 when the 2 spacecraft have magnetic footprints (crosses on the right-hand side at the solar source surface) co-located in a 3° × 3° longitude/latitude box. The lines are displayed for u_{SW} = 450 km/s. The diamond symbols are the radial footprints (sub-spacecraft points).

for a given Solar Wind velocity u_{SW}; it is given by the following formula:

$$\varphi_s = \varphi(r) + \Omega(r - R_s)/u_{SW} \quad (1)$$

where $\varphi(r)$ is the spacecraft heliocentric longitude at distance r, $\Omega = 2\pi/25.38 days$, and $R_s = 2.5 R_{Sun}$. Several values of u_{SW} were used in the range 300–800 km/s. At the source surface a small region of 3° × 3° in longitude/latitude is considered for the footprints to be said co-located (see Fig. 14). The footprint of a given spacecraft at a given time is defined by its longitude (φ_s) and by its latitude (equal to the one of the spacecraft as a Parker field line is inscribed on a cone of constant latitude). Below the source surface the Parker approximation to model the magnetic field generally fails as the coronal magnetic field becomes more complex;

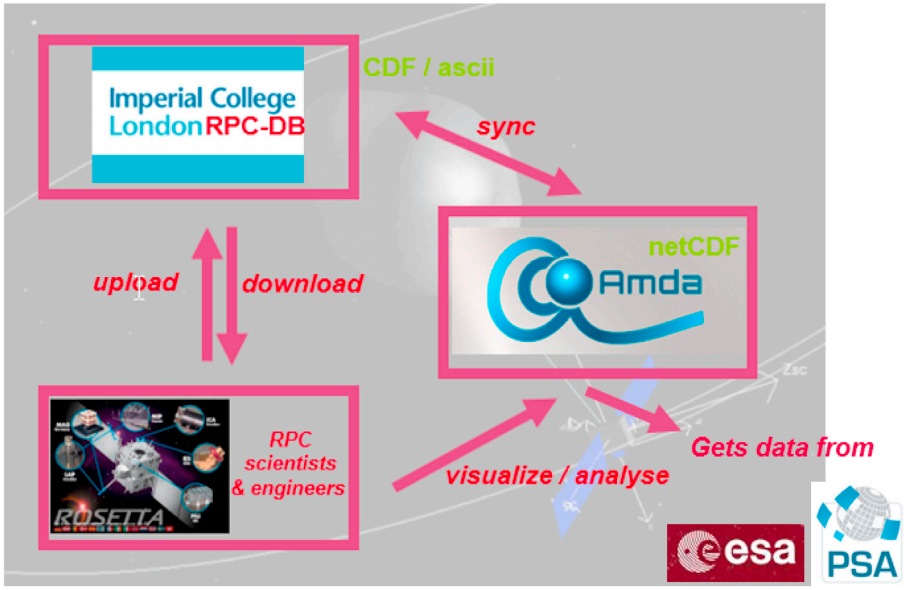

Fig. 15. Architecture of the Imperial College-AMDA collaboration for Rosetta Plasma distribution to the Consortium during the operational phase. In the post-operational phase the data were obtained from PSA when available (bottom right).

refined models that take into account observed/modeled magnetograms (e.g. PFSS) can then be used. This is what is done, for instance, in the "Connect Tool" developed at IRAP.[31] Nevertheless the first order approach explained above and used in AMDA by the working group is a quick way to derive times of interest for future coordinated analysis, and more specifically for studying Solar Energetic Particles (SEP) travelling on contiguous magnetic field lines from active regions at the Sun.

By considering inner heliosphere spacecraft (including JUICE in cruise phase, together with STEREO A, Solar Orbiter, PSP and Bepi Colombo) and planets, more than 1000 events were found for these different configurations. An example is given in Table 3 below for the magnetic connectivity of BepiColombo and Solar Orbiter; the third interval is illustrated on Fig. 14 with 3DView. The complete catalogue of events (time intervals plus columns related to statistics of orbital information on spacecraft and planets) is available to all AMDA users as a "Shared catalogue" (CATALOGS_BepiCoordObs_WindowsOpportunity_times).

6.2. Data distribution: Rosetta Plasma Consortium

In 2014 the CDPP was contacted by the Imperial College (IC) Rosetta Plasma Consortium (RPC) team (C. Carr, S. Schwartz, A. J. Allen) and the then-acting RPC PI (J.-P. Lebreton, LPC2E) with a request which radically changed the way AMDA was used up to then. The RPC was eager to access and visualize plasma data from all their five instruments in an unified way. IC had built the architecture of a temporary archive (the final one being now at ESAC) for the operation phase, centralizing files generated by the instrument teams and providing them in Common Data Format. However IC system was missing an online visualization tool. The requested functionalities fitted those proposed by AMDA and, following the formal request of the PIs, CDPP team provided a first prototype after a few months of work. Beyond visualization, the RPC imposed some constraints (i.e. the data should remain PI team restricted, download and parameter combination should be switched off, …) which, given the architecture of AMDA, were easily enforced. The list of users with "RPC access right" was kept by AMDA admin and extended upon PI request (when new PhD student, post-doc, engineer joined a given RPC team). Technically the synchronization between IC (CDF archived files) and AMDA was done on a daily basis. After some tests a web meeting was organized during which the AMDA functionalities were exposed to the RPC community. As the main need of RPC members was to access, plot and browse multiple datasets at once, AMDA was perfectly tailored for this task. Regular feedback between the PI team, the AMDA team and the RPC community on data content and quality ensured that AMDA was used extensively during the whole exploitation phase and after. The IC-RPC-AMDA collaboration (illustrated on Fig. 15) finally proved to be very successful as the number of publications making use of and acknowledging the facility testifies (see Section 6.4 and, for instance, the RPC papers published in the 2016 Rosetta issue of MNRAS Mandt et al. (2016); Beth et al. (2016); Nemeth et al. (2016); Yang et al. (2016); Goetz et al. (2016); Edberg et al. (2016)). As the mission ended and the final data were progressively delivered to the ESA/PSA, the datasets held in AMDA were also replaced and made public.

6.3. Community feedback

From the start AMDA has been conceived of as a community tool which means that additions of both the data and the analysis functionalities are mostly done upon users' suggestions. This worked particularly well during the infancy of the tool when everything has to be built from scratch and ideas were gushing. With time passing, and while the tool was becoming more complex, the need for prioritization and scheduling became urgent. The role of the CDPP User Committee became vital for the long term planning. On shorter time scale the community feedback through less formal means was also developed. It was made possible directly through the interface itself (the "feedback" button in the lower right corner), or by mailing at amda@irap.omp.eu, but also during all kinds of meetings, workshops, …during which AMDA was presented and promoted. This also includes feedback from students which were taught basics of space physics and data analysis thanks to AMDA during hands-on sessions at the Masters' level or during summer schools (as in the 2019 summer school of Les Houches[32]).

[31] http://connect-tool.irap.omp.eu/.

[32] https://plasmas2019.sciencesconf.org/resource/page/id/7.

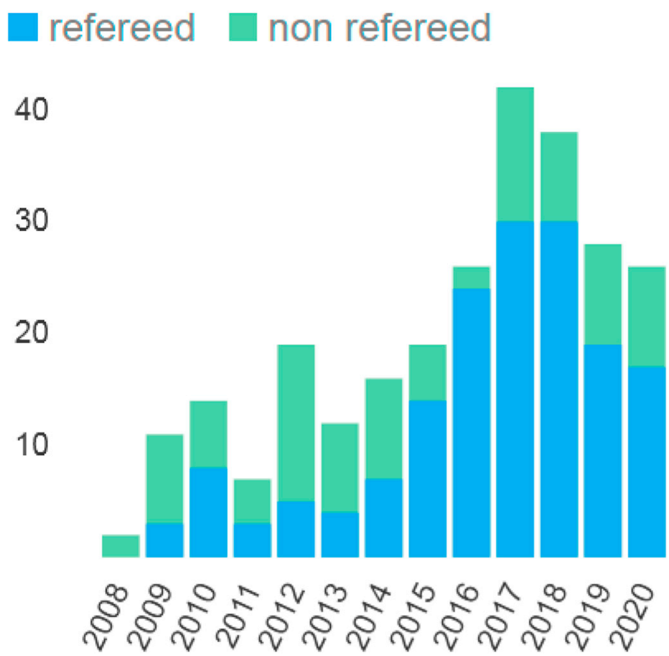

Fig. 16. Distribution of papers acknowledging AMDA over the period 2008–2020 obtained from the SAO/NASA Astrophysics Data System (as of 07/12/2020).

6.4. Promoting science

Fig. 16 shows the number of publications[33] acknowledging AMDA since 2005. The Astrophysical Data System (ADS) algorithm scans the full text of a paper so the occurrence of 'AMDA' may appear in the body of the paper as well as in the acknowledgement section. The surge in 2009 is linked to the involvement of CDPP in international projects (EU or ESA) which made AMDA known and used outside of France. The large increase of references after 2014 comes from the wide spread use of AMDA by the Rosetta Plasma Consortium members for data access and analysis (see Section 6.2). ADS may not reference some journals and therefore the numbers of publications are lower limits. Acknowledging AMDA in publications is crucial for many aspects (visibility, reporting, proposals, …) and a standard text to do so is proposed to authors on the AMDA web page (tab "Rules of the road").

The CDPP web site also centralises resources (references, publications, and conferences presentations) describing AMDA in the various contexts it has been used.

6.5. Space science education

Since AMDA gives easy access to many in-situ parameters of the heliosphere and planetary – including Earth – environments, it is very convenient to use it for teaching space science at various levels and on several circumstances.

The universities of Paris and Toulouse routinely use AMDA in their space physics courses at Master level, for instance to verify experimentally the Rankine-Hugoniot relations at Earth's bow shock with data from the Cluster spacecraft (a tutorial is available[34]). The students may then access real data to work on their hands-on exercise or research projects.

Also, AMDA is regularly used in space physics practical courses at summer schools for graduate students and young researchers. For instance, in space weather schools organized in Maghreb and West Africa since 2011 by the International Space Weather Initiative,[35] students follow an extreme solar event from the Sun to the Earth through interplanetary, magnetospheric and ground-based data. They may compare their findings to other CDPP services such as the Propagation Tool and 3DView.

Scientific unions, such as the International Astronomical Union, encourage research projects conducted in schools with the help of one or several scientists. The goal is to foster astronomy and space sciences education through research. In that perspective, high schools students also utilise AMDA for year-long projects on space science and space weather. AMDA gives them the opportunity of working as real scientists with real data, which is a well-known source of motivation.

To educate schoolteachers about space science, AMDA is also used during the yearly CLEA[36] teacher training sessions in France (Comité de Liaison Enseignants - Astronomes (Pitout and Ferrari, 2020)).

7. Conclusion

AMDA started with simple and sound concepts (use of a web browser, homogeneity of the parameters, multi-disciplinarity), and the initial implementation grew with extensive development into a mature tool that is useable in many science analysis, data mining, and educational situations. A few lessons learned about the full life cycle development costs are also worth describing. On average, over the last 15 years, about 3 engineers (FTE, full time equivalent) were employed to develop and maintain AMDA, with peaks around 5 FTE. The implication of scientists (for design, suggestion, validation) is harder to estimate as it is spread out over many individuals but is probably around 1 FTE. The maintenance and evolution have required sustained funding which was regularly provided by CNES and CNRS over the years and, on dedicated tasks, by specific projects. This continuing support allows management of issues pertaining to long term projects in term of software development, staff turn-over and knowledge/information integrity. In a similar context, the widespread diffusion of SPEDAS was helped by the fact that it is now officially supported by NASA Heliophysics as part of its data environment infrastructure (Angelopoulos et al., 2019).

As a concluding note, with many space exploration missions to come we can foresee a bright future for AMDA. Basic functionalities will of course continue to be improved but we think the real added value in coming years will come along a more seamless interoperability with companion services. In particular the 60+ years of accumulated space physics data are increasingly processed to yield catalogues of events which in turn feed learning algorithms or help in data assimilation techniques for simulations. Facilitating these approaches is certainly of one the future slots for AMDA. The recent partnership with Sciqlop is already along this line. Another route of development is the growing interest and need for tracking research data to ensure full reproducibility of published scientific results. The ability to replay a given plot or analysis from a DOI, or an equivalent, more complex, perennial, but still to be defined, identification process also looks promising and AMDA has all elementary bricks to achieve this. Indeed AMDA has been developed with FAIR principles from the beginning, making space physics data "FAIR" themselves in return, and this general philosophy will be pursued in the future.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[33] An up-to-date version of this figure is visible at https://tinyurl.com/ADS-AMDA-stat which points to the ADS statistics page for AMDA papers.
[34] http://impex.latmos.ipsl.fr/TP_MIRE.pdf.
[35] http://www.iswi-secretariat.org/.
[36] http://clea-astro.eu.

Acknowledgments

The Centre de Données de la Physique des Plasmas (CDPP) is supported by CNRS, CNES, Observatoire de Paris and Université Toulouse 3 Paul Sabatier, France. AMDA team thanks all active users who provided very useful feedback over the years and therefore contributed to enhance the quality of the system; a very special 'thank you' goes to the CDPP User Committee (UC) members who did not count their time during specific test phases and provided valuable suggestions; current and past members are:

- K. Issautier (former UC president, LESIA).
- M. Maksimovic (former UC president, LESIA).
- J.-L. Pinçon (former UC president, LPC2E).
- B. Grison (current UC president, IAP, Prague).
- R. Maggiolo (BIRA).
- C. Foullon (Exeter University).
- O. Lecontel (LPP).
- A. Retino (LPP).
- M. Barthélémy (IPAG).
- E. Astafyeva (IPGP).
- O. Santolik (IAP, Prague).
- P. Henri (LPC2E).
- N. Aunai (LPP).
- R. Modolo (LATMOS).
- C. Lathuillère (IPAG).
- P. Robert (LPP).
- A. Marchaudon (IRAP).
- P. Louarn (IRAP).

The AMDA team also thanks the following individuals for fruitful technical discussions and general support: E. Penou, A. Barthe, C. Hillembrand, L. Jahan, J.-M. Glorian and people at OV-GSO, and the support teams from NASA/SPDF, NASA/NAIF, NASA/PDS, VESPA, ESAC, 3DView at GFI.

A Acronyms

AMDA	Automated Multi-Data Analysis
API	Application Programming Interface
CDPP	Centre de Données de la Physique des Plasmas
CL	visualization software (also CLweb)
DOI	Digital Object Identifier
EPN-TAP	EuroPlanet Table Access Protocol
FAIR	Findable, Accessible, Interoperable, Reusable
GPLv3	GNU Public Licence version 3
GUI	Graphical User Interface
ICME	Interplanetary Coronal Mass Ejection
IHDEA	International Heliophysics Data Environment Alliance
IMPEx	Integrated Medium for Planetary Exploration
IPDA	International Planetary Data Alliance
IVOA	International Virtual Observatory Alliance
MHD	Magnetohydrodynamics
PSWS	Planetary Space Weather Services (EuroPlanet)
RPC	Rosetta Plasma Consortium
SAMP	Simple Application Messaging Protocol
SciQlop	SCIentific Qt application for Learning from Observations of Plasmas
SPASE	Space Physics Archive Search and Extract
SSA	Space Situational Awareness (ESA)
VESPA	Virtual European Solar and Planetary Access (EuroPlanet)
X–IFU	X-ray Integral Field Unit

References

André, N., Hitier, R., Cecconi, B., Gangloff, M., Jacquey, C., et al., 2011. Hst auroral campaign observations of jupiter and saturn enabled by the cdpp/amda and ivoa/aladin tools. In: Proceedings of PV-2011. Retrieved from. http://typhon.obspm.fr/idis/docs/Andre_PV2011.pdf.

Angelopoulos, V., Cruce, P., Drozdov, A., Grimes, E.W., Hatzigeorgiu, N., King, D.A., et al., 2019, Feb. The space physics environment data analysis system (spedas). Space Sci. Rev. 215 (1), 9. https://doi.org/10.1007/s11214-018-0576-4.

Besse, S., Vallat, C., Barthelemy, M., Coia, D., Costa, M., De Marchi, G., et al., 2018, Jan. ESA's Planetary Science Archive: preserve and present reliable scientific data sets. Planet. Space Sci. 150, 131–140. https://doi.org/10.1016/j.pss.2017.07.013.

Beth, A., Altwegg, K., Balsiger, H., Berthelier, J.J., Calmonte, U., Combi, M.R., et al., 2016, Nov. First in-situ detection of the cometary ammonium ion NH_4$^+$ (protonated ammonia NH_3) in the coma of 67P/C-G near perihelion. Mon. Not. Roy. Astron. Soc. 462, S562–S572. https://doi.org/10.1093/mnras/stw3370.

Budjáš, D., Nieminen, P., Jiggens, P., Santin, G., Daly, E., 2017. Soft proton fluxes in and around the earth's magnetotail. IEEE Trans. Plasma Sci. 45 (8), 1965–1971. https://doi.org/10.1109/TPS.2017.2672202.

Christon, S.P., Eastman, T.E., Doke, T., Frank, L.A., Gloeckler, G., Kojima, H., et al., 1998, Oct. Magnetospheric plasma regimes identified using Geotail measurements 2. Statistics, spatial distribution, and geomagnetic dependence. J. Geophys. Res.: Space Phys. 103 (A10), 23521–23542. https://doi.org/10.1029/98JA01914.

Connerney, J.E.P., Kotsiaros, S., Oliversen, R.J., Espley, J.R., Joergensen, J.L., Joergensen, P.S., et al., 2018, Mar. A new model of jupiter's magnetic field from juno's first nine orbits. Geophys. Res. Lett. 45 (6), 2590–2596. https://doi.org/10.1002/2018GL077312.

Dowler, P., Rixon, G., Tody, D., Demleitner, M., 2018. Table Access Protocol. version 1.1, IVOA Documentation. Retrieved from. http://www.ivoa.net/documents/TAP/.

Edberg, N.J.T., Alho, M., André, M., Andrews, D.J., Behar, E., Burch, J.L., et al., 2016, Nov. CME impact on comet 67P/Churyumov-Gerasimenko. Mon. Not. Roy. Astron. Soc. 462, S45–S56. https://doi.org/10.1093/mnras/stw2112.

Erard, S., Cecconi, B., Le Sidaner, P., Chauvin, C., Rossi, A.P., Minin, M., et al., 2020, March. Virtual European solar & planetary access (VESPA): a planetary science virtual observatory cornerstone. Data Sci. J. https://doi.org/10.5334/dsj-2020-016.

Erard, S., Cecconi, B., Le Sidaner, P., Rossi, A.P., Capria, M.T., Schmitt, B., et al., 2018, Jan. VESPA: a community-driven virtual observatory in planetary science. Planet. Space Sci. 150, 65–85. https://doi.org/10.1016/j.pss.2017.05.013.

Faden, J.B., Weigel, R.S., Merka, J., Friedel, R.H.W., 2010, Jun 01. Autoplot: a browser for scientific data on the web. Earth Sci. Inf. 3 (1), 41–49. https://doi.org/10.1007/s12145-010-0049-0 doi: 10.1007/s12145-010-0049-0. Retrieved from.

Fairfield, D.H., 1971, Oct. Average and unusual locations of the earth's magnetopause and bow shock. J. Geophys. Res. 76 (28), 6700–6716. https://doi.org/10.1029/JA076i028p06700.

Garnier, P. (submitted). AI techniques applied to planetary plasma boundaries. Planet. Space Sci..

Génot, V., André, N., Cecconi, B., Bouchemit, M., Budnik, E., Bourrel, N., et al., 2014, Nov. Joining the yellow hub: uses of the simple application messaging protocol in space physics analysis tools. Astron. Comput. 7, 62–70. https://doi.org/10.1016/j.ascom.2014.07.007.

Génot, V., Beigbeder, L., Popescu, D., Dufourg, N., Gangloff, M., Bouchemit, M., et al., 2018a. 3DView: Space Physics Data Visualizer.

Génot, V., Beigbeder, L., Popescu, D., Dufourg, N., Gangloff, M., Bouchemit, M., et al., 2018b. Science data visualization in planetary and heliospheric contexts with 3DView. Planet. Space Sci. 150, 111–130. https://doi.org/10.1016/j.pss.2017.07.007.

Génot, V., Renard, B., Dufourg, N., Bouchemit, M., Lormant, N., Beigbeder, L., et al., 2018c. TREPS, a tool for coordinate and time transformations in space physics. Planet. Space Sci. 150, 86–90. https://doi.org/10.1016/j.pss.2017.06.002.

Génot, V., Jacquey, C., Bouchemit, M., Gangloff, M., Fedorov, A., Lavraud, B., et al., 2010a. Space weather applications with CDPP/AMDA. Adv. Space Res. 45 (9), 1145–1155. https://doi.org/10.1016/j.asr.2009.11.010.

Génot, V., Jacquey, C., Budnik, E., Bouchemit, M., Gangloff, M., Fedorov, A., et al., 2010b. Science outputs of the CDPP on-line analysis tool AMDA. In: Boissier, S., Heydari-Malayeri, M., Samadi, R., Valls-Gabaud, D. (Eds.), Sf2a-2010: Proceedings of the Annual Meeting of the French Society of Astronomy and Astrophysics, p. 301.

Goetz, C., Koenders, C., Hansen, K.C., Burch, J., Carr, C., Eriksson, A., et al., 2016, Nov. Structure and evolution of the diamagnetic cavity at comet 67P/Churyumov-Gerasimenko. Mon. Not. Roy. Astron. Soc. 462, S459–S467. https://doi.org/10.1093/mnras/stw3148.

Grison, B., Soucek, Jan, Krupar, Vratislav, Píša, David, Santolík, Ondrej, Taubenschuss, Ulrich, Nemec, Frantisek, 2018. Shock deceleration in interplanetary coronal mass ejections (icmes) beyond mercury's orbit until one au. J. Space Weather Space Clim. 8, A54. https://doi.org/10.1051/swsc/2018043 doi: 10.1051/swsc/2018043. Retrieved from.

Jacquey, C., Génot, V., Budnik, E., Hitier, R., Bouchemit, M., Gangloff, M., et al., 2010, Jan. AMDA, automated multi-dataset analysis: a web-based service provided by the CDPP. Astrophys. Space Sci. Proc. 11, 239–247. https://doi.org/10.1007/978-90-481-3499-1_16.

Laakso, H., Perry, C., McCaffrey, S., Herment, D., Allen, A.J., Harvey, C.C., et al., 2010, Jan. Cluster Active archive: overview. Astrophys. Space Sci. Proc. 11, 3–37. https://doi.org/10.1007/978-90-481-3499-1_1.

Laurenza, M., Alberti, T., Marcucci, M.F., Consolini, G., Jacquey, C., Molendi, S., et al., 2019, Mar. Estimation of the particle radiation environment at the L1 point and in near-earth space, 873 (2), 112. https://doi.org/10.3847/1538-4357/ab0410.

Lenouvel, Q., Génot, V., Garnier, P., Toledo-Redondo1, S., Lavraud, B., Aunai, N., Nguyen, G., Gershman, D.J., Ergun, R.E., Lindqvist, P.-A., Giles, B., Burch, J.L., 2021. Identification of Electron Diffusion Regions with a Machine Learning approach on MMS data at the Earth's magnetopause. Earth Space Sci.

Lotti, S., Mineo, T., Jacquey, C., Molendi, S., D'Andrea, M., Macculi, C., Piro, L., 2017, Dec. The particle background of the X-IFU instrument. Exp. Astron. 44 (3), 371–385. https://doi.org/10.1007/s10686-017-9538-1.

Mandt, K.E., Eriksson, A., Edberg, N.J.T., Koenders, C., Broiles, T., Fuselier, S.A., et al., 2016, Nov. RPC observation of the development and evolution of plasma interaction boundaries at 67P/Churyumov-Gerasimenko. Mon. Not. Roy. Astron. Soc. 462, S9–S22. https://doi.org/10.1093/mnras/stw1736.

McMahon, S.K., 1996, Jan. Overview of the planetary data system. Planet. Space Sci. 44 (1), 3–12. https://doi.org/10.1016/0032-0633(95)00101-8.

Mitchell, D.L., Lin, R.P., Mazelle, C., Rème, H., Cloutier, P.A., Connerney, J.E.P., et al., 2001. Probing Mars' crustal magnetic field and ionosphere with the. MGS Electron Reflectometer 106 (E10), 23419–23428. https://doi.org/10.1029/2000JE001435.

Miura, A., Shinohara, I., Matsuzaki, K., Nagase, F., Negoro, H., Uno, S., et al., 2000. ISAS data archive and transmission system (DARTS). In: Manset, N., Veillet, C., Crabtree, D. (Eds.), Astronomical Data Analysis Software and Systems Ix, Asp Conference Proceedings. In: nadine, manset, christian, veillet, dennis, crabtree (Eds.), vol. 216. astronomical society of the pacific isbn 1-58381-047-1, 2000., p.180 (Vol. 216, p. 180).

Modolo, R., Hess, S., Génot, V., Leclercq, L., Leblanc, F., Chaufray, J.-Y., et al., 2018, Jan. The lathys database for planetary plasma environment investigations: overview and a case study of data/model comparisons. Planet. Space Sci. 150, 13–21. https://doi.org/10.1016/j.pss.2017.02.015.

Morschhauser, A., Lesur, V., Grott, M., 2014, Jun. A spherical harmonic model of the lithospheric magnetic field of mars. J. Geophys. Res.: Planets 119 (6), 1162–1188. https://doi.org/10.1002/2013JE004555.

NASA CDAWeb Development Team, 2019, Apr. CDAWeb: Coordinated Data Analysis Web.

Nemeth, Z., Burch, J., Goetz, C., Goldstein, R., Henri, P., Koenders, C., et al., 2016, Nov. Charged particle signatures of the diamagnetic cavity of comet 67P/Churyumov-Gerasimenko. Mon. Not. Roy. Astron. Soc. 462, S415–S421. https://doi.org/10.1093/mnras/stw3028.

Pitout, F., Ferrari, C., 2020. CLEA, a French pro-am initiative to foster astronomy education. In: Proceedings of the 1st IAU Astronomy Education Conference.

Pitout, F., Koechlin, L., López Ariste, A., Dettwiller, L., Glorian, J.-M., 2020, July. Solar surveillance with CLIMSO: instrumentation, database and on-going developments. J. Space Weather Space Clim. 10, 47. https://doi.org/10.1051/swsc/2020039.

Roberts, D.A., Thieman, J., Génot, V., King, T., Gangloff, M., Perry, C., et al., 2018, Dec. The spase data model: a metadata standard for registering, finding, accessing, and using heliophysics data obtained from observations and modeling. Space Weather 16 (12), 1899–1911. https://doi.org/10.1029/2018SW002038.

Romanelli, N., DiBraccio, G., Modolo, R., Leblanc, F., Espley, J., Gruesbeck, J., et al., 2019, Oct. Recovery timescales of the dayside martian magnetosphere to imf variability. Geophys. Res. Lett. 46 (20), 10977–10986. https://doi.org/10.1029/2019GL084151.

Romanelli, N., Modolo, R., Leblanc, F., Chaufray, J.-Y., Hess, S., Brain, D., et al., 2018, Jul. Effects of the crustal magnetic fields and changes in the imf orientation on the magnetosphere of mars: maven observations and lathys results. J. Geophys. Res.: Space Phys. 123 (7), 5315–5333. https://doi.org/10.1029/2017JA025155.

Rouillard, A.P., Lavraud, B., Génot, V., Bouchemit, M., Dufourg, N., Plotnikov, I., et al., 2017, Nov. A propagation tool to connect remote-sensing observations with in-situ measurements of heliospheric structures. Planet. Space Sci. 147, 61–77. https://doi.org/10.1016/j.pss.2017.07.001.

Shue, J.-H., Chao, J.K., Fu, H.C., Russell, C.T., Song, P., Khurana, K.K., Singer, H.J., 1997. A new functional form to study the solar wind control of the magnetopause size and shape. J. Geophys. Res.: Space Phys. 102 (A5), 9497–9511. https://doi.org/10.1029/97JA00196.

Steckiewicz, M., Garnier, P., André, N., Mitchell, D.L., Andersson, L., Penou, E., et al., 2017. Comparative study of the martian suprathermal electron depletions based on mars global surveyor, mars express, and mars atmosphere and volatile EvolutioN mission observations. J. Geophys. Res. (Space Phys.) 122 (1), 857–873. https://doi.org/10.1002/2016JA023205.

Steckiewicz, M., Garnier, P., Lillis, R., Toublanc, D., Leblanc, F., Mitchell, D.L., et al., 2019. Dawn/dusk asymmetry of the martian UltraViolet terminator observed through suprathermal electron depletions. J. Geophys. Res. (Space Phys.) 124 (8), 7283–7300. https://doi.org/10.1029/2018JA026336.

Steckiewicz, M., Mazelle, C., Garnier, P., André, N., Penou, E., Beth, A., et al., 2015. Altitude dependence of nightside Martian suprathermal electron depletions as revealed by MAVEN observations. Geophys. Res. Lett. 42 (21), 8877–8884. https://doi.org/10.1002/2015GL065257.

Tamura, T., Baba, H., Matsuzaki, K., Miura, A., Shinohara, I., Nagase, F., et al., 2004. Data archive and transfer system (DARTS) of ISAS. In: Ochsenbein, F., Allen, M.G., Egret, D. (Eds.), Astronomical Data Analysis Software and Systems (Adass) Xiii, Proceedings of the Conference Held 12-15 October, 2003 in Strasbourg, france. Edited by Francois Ochsenbein, Mark G. Allen and Daniel Egret. Asp Conference Proceedings, vol. 314. Astronomical society of the pacific, san francisco, p. 22.

Tao, C., Kataoka, R., Fukunishi, H., Takahashi, Y., Yokoyama, T., 2005, Nov. Magnetic field variations in the Jovian magnetotail induced by solar wind dynamic pressure enhancements. J. Geophys. Res. (Space Phys.) 110 (A11), A11208. https://doi.org/10.1029/2004JA010959.

Taylor, M.B., 2005. TOPCAT & STIL: starlink table/VOTable processing software. In: Shopbell, P., Britton, M., Ebert, R. (Eds.), Astronomical Data Analysis Software and Systems Xiv, vol. 347, p. 29.

Tsyganenko, N.A., Stern, D.P., 1996, Dec. Modeling the global magnetic field of the large-scale birkeland current systems. J. Geophys. Res.: Space Phys. 101 (A12), 27187–27198. https://doi.org/10.1029/96JA02735.

Vandegriff, J., Weigel, R., Faden, J., Grimes, E., Roberts, D.A., King, T., et al., 2019. Interoperability for heliophysics and planetary time series data via hapi. Earth Space Sci. Open Arch. 1 https://doi.org/10.1002/essoar.10501378.1 doi: 10.1002/essoar.10501378.1. Retrieved from.

Walker, R.J., Joy, S.P., King, T.A., Russell, C.T., McPherron, R.L., Kurth, W.S., 1996, Jan. The planetary plasma interactions node of the planetary data system. Planet. Space Sci. 44 (1), 55–64. https://doi.org/10.1016/0032-0633(95)00106-9.

Williams, D.J., McEntire, R.W., Schlemm, I.C., Liu, A.T.Y., Gloeckler, G., Christon, S.P., Gliem, F., 1994, Jan. GEOTAIL energetic particles and ion composition instrument. J. Geomagn. Geoelectr. 46 (1), 39–57. https://doi.org/10.5636/jgg.46.39.

Witasse, O., Sánchez-Cano, B., Mays, M.L., Kajdič, P., Opgenoorth, H., Elliott, H.A., et al., 2017, Aug. Interplanetary coronal mass ejection observed at STEREO-A, Mars, comet 67P/Churyumov-Gerasimenko, Saturn, and New Horizons en route to Pluto: comparison of its Forbush decreases at 1.4, 3.1, and 9.9 AU. J. Geophys. Res. (Space Phys.) 122 (8), 7865–7890. https://doi.org/10.1002/2017JA023884.

Xu, F., Borovsky, J.E., 2015, Jan. A new four-plasma categorization scheme for the solar wind. J. Geophys. Res.: Space Phys. 120 (1), 70–100. https://doi.org/10.1002/2014JA020412.

Yang, L., Paulsson, J.J.P., Wedlund, C.S., Odelstad, E., Edberg, N.J.T., Koenders, C., et al., 2016, Nov. Observations of high-plasma density region in the inner coma of 67P/Churyumov-Gerasimenko during early activity. Mon. Not. Roy. Astron. Soc. 462, S33–S44. https://doi.org/10.1093/mnras/stw2046.