

AMDA

Automated Multi Dataset Analysis

un outil du CDPP
cdpp-amda.cesr.fr

Vincent Génot



Contexte national

- Mission du CDPP : valoriser des données de missions satellites françaises ou à participation française dans le domaine de la physique des plasmas
 - Pour dynamiser l'utilisation du centre, il y eu nécessité de dépasser le cadre de la "simple" archive
- développement d'un outil au service de la communauté

Contexte international

- Les observatoires virtuels en physique magnétosphériques sont balbutiants
 - Implication du CDPP dès l'origine dans la mise au point de standard (modèle de données SPASE)
- "On peut se mettre dans la course"

Spécificités / contraintes de la discipline

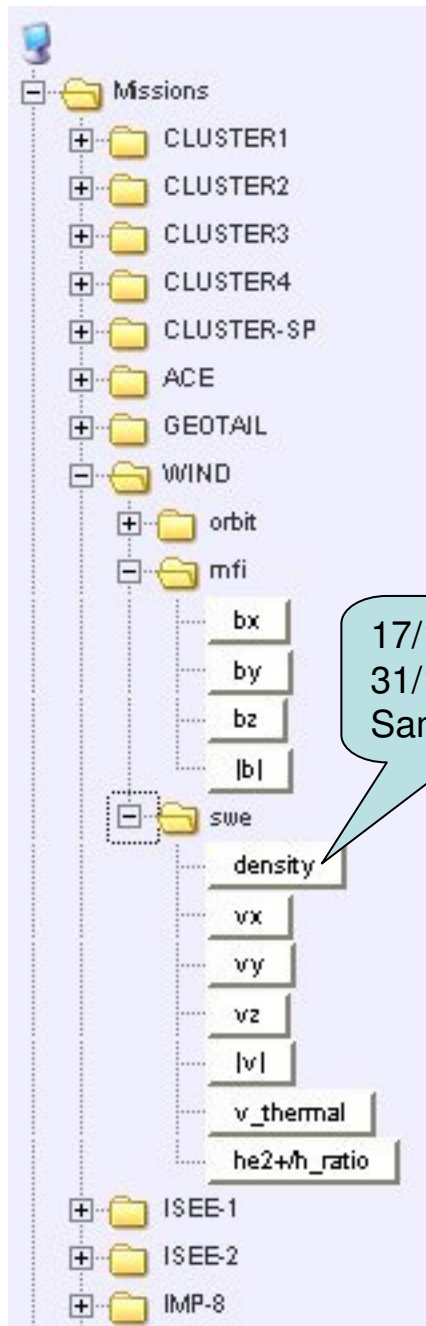
- Les données archivées sont des scalaires, des vecteurs, des tenseurs, des spectres, ... ordonnés par le temps
- Il n'existe pas un format de données standard adopté par tous
- Il existe une multitude de bases de données réparties
- De plus en plus le travail scientifique requiert une confrontation des résultats de plusieurs missions à la fois : études multi-échelles, en météorologie spatiale, héliophysique

Motivation / g n se du projet

- Initialement pens  pour valoriser les donn es du CDPP
- Une premi re phase : r utilisation de l'existant, test d'int gration
- Id e de l'espace de travail qui centralise toute la production de l'utilisateur
- Construit autour d'une base + outil de visualisation d velopp  pour la mission magn tosph rique CLUSTER (interface web)

D veloppement :

- Principalement 2 d veloppeurs "dans les murs" (IDL, C, Fortran, javascript)
- 1 re version en quelques mois
- Am lioration constante du prototype en relation avec les avis du CU du CDPP
 - Int gration de standards
 - Ouverture vers les VO



Le paramètre :

Un paramètre est une série temporelle :

- champ magnétique ou électrique
- moment particule
- indice
- ...
- *une combinaison*

→ valide sur un intervalle de temps

→ à une résolution temporelle donnée

→ utilisé pour :

- la représentation visuelle
- la recherche conditionnelle
- le téléchargement, ...

L'organisation en fichiers dans l'archive est cachée à l'utilisateur

Construction d'un paramètre 1/2

CDPP Home Page Welcome to AMDA

My Workspace Plot Data Download Data Conditional Search Web Services Help Data Inventory Graph F

Select parameters to construct new Workspace parameter

open all | close all

- Missions
 - CLUSTER1
 - orbit
 - fgm
 - bx
 - by
 - bz
 - lbl
 - cis-hia
 - cis-codif
 - efw
 - whisper
 - staff
 - cis-hia+fgm
 - CLUSTER2
 - orbit
 - fgm
 - bx
 - by
 - bz
 - lbl
 - efw

clics

Construct Your Parameter

Expression
b_c1(3)/b_c2(3)

Sampling time step
60 secs

Parameter name
bc1sbc2

Description
C1/C2 magnetic field ratio

Save to WS Reset

Construct Your Time Table

StartTime - StopTime
yyy-mm-ddThh:mm:ss yyy-mm-ddThh:mm:ss

Table Name

Date of Creation

Résolution temporelle

Nom du paramètre

Description

Syntax of Expression

arithmetic operators: + - * / ^

brackets: () []

functions: sin() cos() sqrt() atan() abs()

$\sin(\text{param}_1)^2 + \sqrt{\text{abs}(\text{param}_2)}^5$

Construction d'un paramètre 2/2

Select parameters to construct new Workspace parameter

open all | close all

- Missions
- Ground-based Indices
- Model Parameters along Orbit / Time Series
- Models along Orbit / Space

My WorkSpace

- MY PARAMETERS
 - bc1sbc2
 - AngleByBxDegres
 - valfRatio
 - TemperatureAnisotropy
- MY TIME TABLES

Construct Your Parameter

Expression
mom_c1(6)/mom_c1(5)

Sampling time step
60 secs

Parameter name
TemperatureAnisotropy

Description

nouveau paramètre dans l'espace de travail

Save to WS Reset

Des paramètres de

- différentes missions
- différents instruments
- avec différentes résolution

peuvent être combinés

Le nouveau paramètre est maintenant un élément virtuel de la base de données. Il sera calculé à la volée seulement quand il sera demandé pour un tracé, une recherche conditionnelle, le téléchargement

Time Shifting of SW Data

SWMonitor ACE Time Delay (secs) 3676 Save and Apply to Data

Target Cluster1 Year / Mon / Day 2001 / 03 / 15 Hour : Min 03 : 30 Calculate Delay

Quelques outils :

- 'time shifting' pour prendre en compte le décalage temporel entre instruments,
- 'resampling' pour prendre en compte les différence de résolution entre instruments.

Les tables d'événements 1/2

-Les tables d'événements (time tables, event lists) sont une collection de temps quand quelque chose se passe ou d'intervalles quand une condition est vérifiée

-Habituellement, elles sont produites et gérées "manuellement" par les scientifiques pour leur usage propre ou celui d'une communauté restreinte

-AMDA augmente les potentialités de ce vecteur d'information en offrant un outil générique pour créer, gérer et communiquer ces objets

Les tables d'événements peuvent être utilisées pour :

- Extraire une sous base de données,
- Exécuter des traitement massifs ou interactifs,
- Créer des catalogues,
- Servir de référence à la communauté

Créer une table d'événements : 1- Recherche conditionnelle

Select parameters to compose the condition

open all | close all

- Missions
 - CLUSTER1
 - orbit
 - fgm
 - cis-hia
 - dens
 - vx_gse
 - vy_gse
 - vz_gse
 - lwl
 - t_para
 - t_perp
 - cis-codif
 - efw
 - whisper
 - staff
 - cis-hia+fgm
- CLUSTER2
- CLUSTER3
- CLUSTER4
- CLUSTER-SP
- ACE
- GEOTAIL
- WIND

Construct Your Search Condition:

$\text{mom_c1}(6)/\text{mom_c1}(5) > 1.$

Syntax of Condition expression

arithmetic operators: + - * / ^
brackets: (), []
functions: **sin()** **cos()** **sqrt()** **atan()** **abs()**
relational operators: >, <
logical operators: &, |

Example
 $\text{sin}(\text{param1}) > 0.8 \ \& \ \text{param2} < 0$

Averaging/Interpolation

Sampling time step
60 secs

Treat data absence as gap

Time interval greater than
5 × data sampling time

Start Time

Year / Mon / Day Hour : Min : Sec
2001 / 03 / 15 02 : 00 : 00

Time Interval

Day / Hour : Min : Sec
001 : 00 : 00

Reset

Generate Table To... AnisotropyTable

Generate Table From ... SearchTable

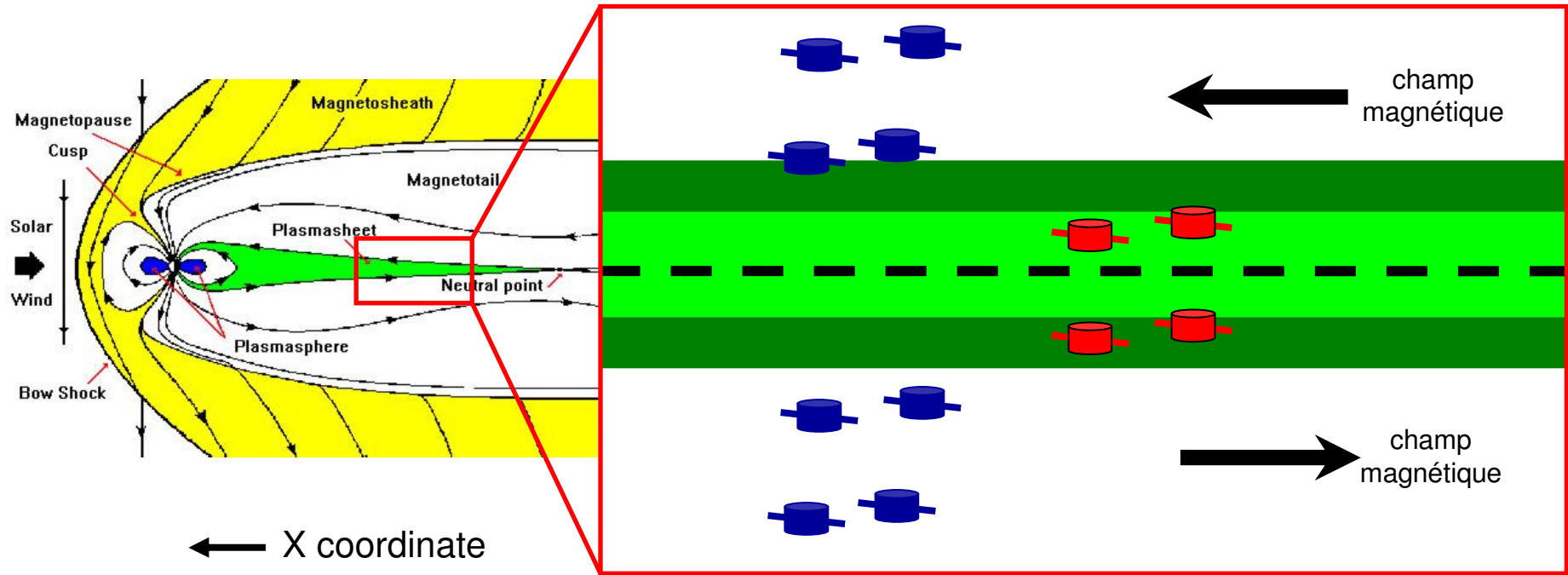
To new Time Table ... NewSearchTable

Load Condition From... [v]

Annotations:

- clics**: Points to the parameter selection tree.
- $T_{\perp}/T_{\parallel} > 1$: Points to the search condition.
- Editer une condition avec des opérateurs/fonctions mathématiques**: Points to the search condition editor.
- Definir un intervalle de temps ou utiliser une autre table d'événements**: Points to the Time Interval and Generate Table From options.

Exemple: recherche des événements de courants de queue “encerclés” par CLUSTER



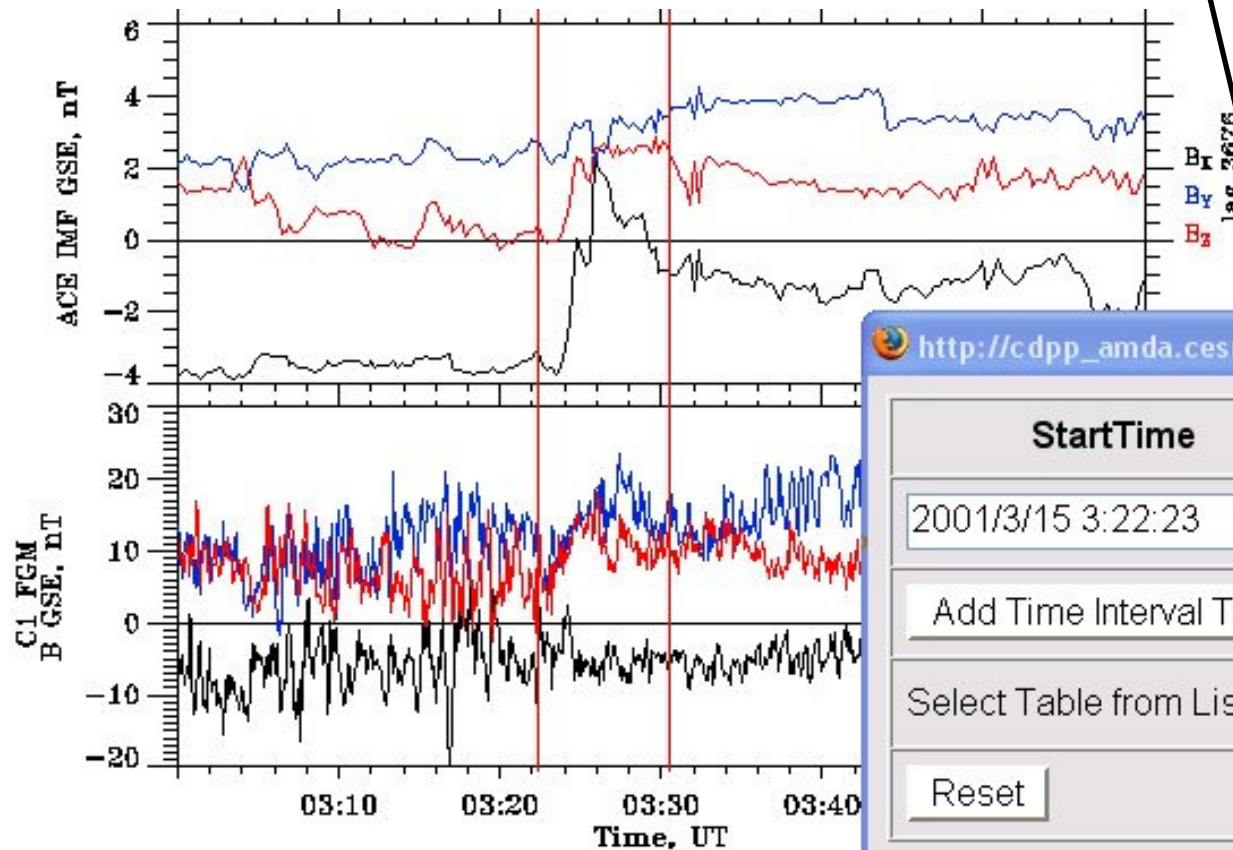
- position : dans la queue magnétosphérique
- 2 satellites sont au dessus de la couche de courant et 2 autres en dessous

Test:

- $X_1 < -10 \text{ Re}$
- $BX_1 * BX_2 * BX_3 * BX_4 > 0$
- $\min([BX_1 \ BX_2 \ BX_3 \ BX_4]) < 0$
- $\max([BX_1 \ BX_2 \ BX_3 \ BX_4]) > 0$

Créer une table d'événements : 2- Inspection visuelle

Les intervalles intéressants sont sélectionnés par clics



http://cdpp_amda.cesr.fr AMDA - Mozilla ...

StartTime	StopTime
2001/3/15 3:22:23	2001/3/15 3:30:34
Add Time Interval To	
Select Table from List	AnisotropyTable
Reset	DONE

Mar 15 2001

Created by DD_SYSTEM(C)

Save Start-Stop

Zoom

<<<<

>>>>

DONE

AnisotropyTable

generated Thu Apr 19 12:44:21 2007

under conditions $MOM_C1(6)/MOM_C1(5) > 1$.

StartTime	StopTime
2001-03-15T02:00:00	2001-03-15T02:30:00
2001-03-15T02:32:00	2001-03-15T03:07:00
2001-03-15T03:08:00	2001-03-15T03:16:00
2001-03-15T03:17:00	2001-03-15T04:25:00
2001-03-15T04:26:00	2001-03-15T05:06:00
2001-03-15T05:11:00	2001-03-15T05:12:00
2001-03-15T05:22:00	2001-03-15T05:25:00

Les tables peuvent être exportées au format VOTable

```
<VOTABLE version="1.1" xsi:schemaLocation="http://www.ivoa.net/xml/VOTable/v1.1 C\
- <DESCRIPTION>
  List of shocks from: http://www.sp.ph.ic.ac.uk/~eal/shocks-staffsa-031022.txt This list has c
  from this time to get Start Time we add 5 minutes to this time to get Stop Time Creation Dat
</DESCRIPTION>
- <RESOURCE>
  <DESCRIPTION> SPACECRAFT=CLUSTER </DESCRIPTION>
- <TABLE>
  - <FIELD datatype="char" name="Start Time" ID="TimeIntervalStart" ucd="time.start":
    <DESCRIPTION>time tag for beginning of interval</DESCRIPTION>
  </FIELD>
  <FIELD datatype="char" name="Stop Time" ID="TimeIntervalStop" ucd="time.stop":
    <DESCRIPTION>time tag for end of interval</DESCRIPTION>
  </FIELD>
  - <DATA>
  - <TABLEDATA>
    - <TR>
      <TD>2000-12-22T08:19:49</TD>
      <TD>2000-12-22T08:29:49</TD>
    </TR>
    - <TR>
      <TD>2000-12-22T08:30:21</TD>
      <TD>2000-12-22T08:40:21</TD>
```

My Workspace

- MY PARAMETERS
 - bc1sbc2
 - AngleByBxDegres
 - valfRatio
 - TemperatureAnisotropy
- MY TIME TABLES
 - SearchTable
 - AnisotropyTable

Save to WS Reset

Construct Your Time Table

Table Name: AnisotropyTable

Date of Creation: Thu Apr 19 12:44:21 2007

Description: $MOM_C1(6)/MOM_C1(5) > 1$.

Source: AMDA Search

Number of Intervals: 12

StartTime - StopTime
yyy-mm-ddThh:mm:ss yyy-mm-ddThh:mm:ss

2001-03-15T02:00:00	2001-03-15T02:30:00	-- 1
2001-03-15T02:32:00	2001-03-15T03:07:00	-- 2
2001-03-15T03:08:00	2001-03-15T03:16:00	-- 3
2001-03-15T03:17:00	2001-03-15T04:25:00	-- 4
2001-03-15T04:26:00	2001-03-15T05:06:00	-- 5
2001-03-15T05:11:00	2001-03-15T05:12:00	-- 6
2001-03-15T05:22:00	2001-03-15T05:25:00	-- 7
2001-03-15T11:56:00	2001-03-15T11:57:00	-- 8
2001-03-15T16:33:00	2001-03-15T16:34:00	-- 9
2001-03-15T16:59:00	2001-03-15T17:00:00	-- 10
2001-03-15T17:16:00	2001-03-15T17:17:00	-- 11
2001-03-15T17:47:00	2001-03-15T17:48:00	-- 12

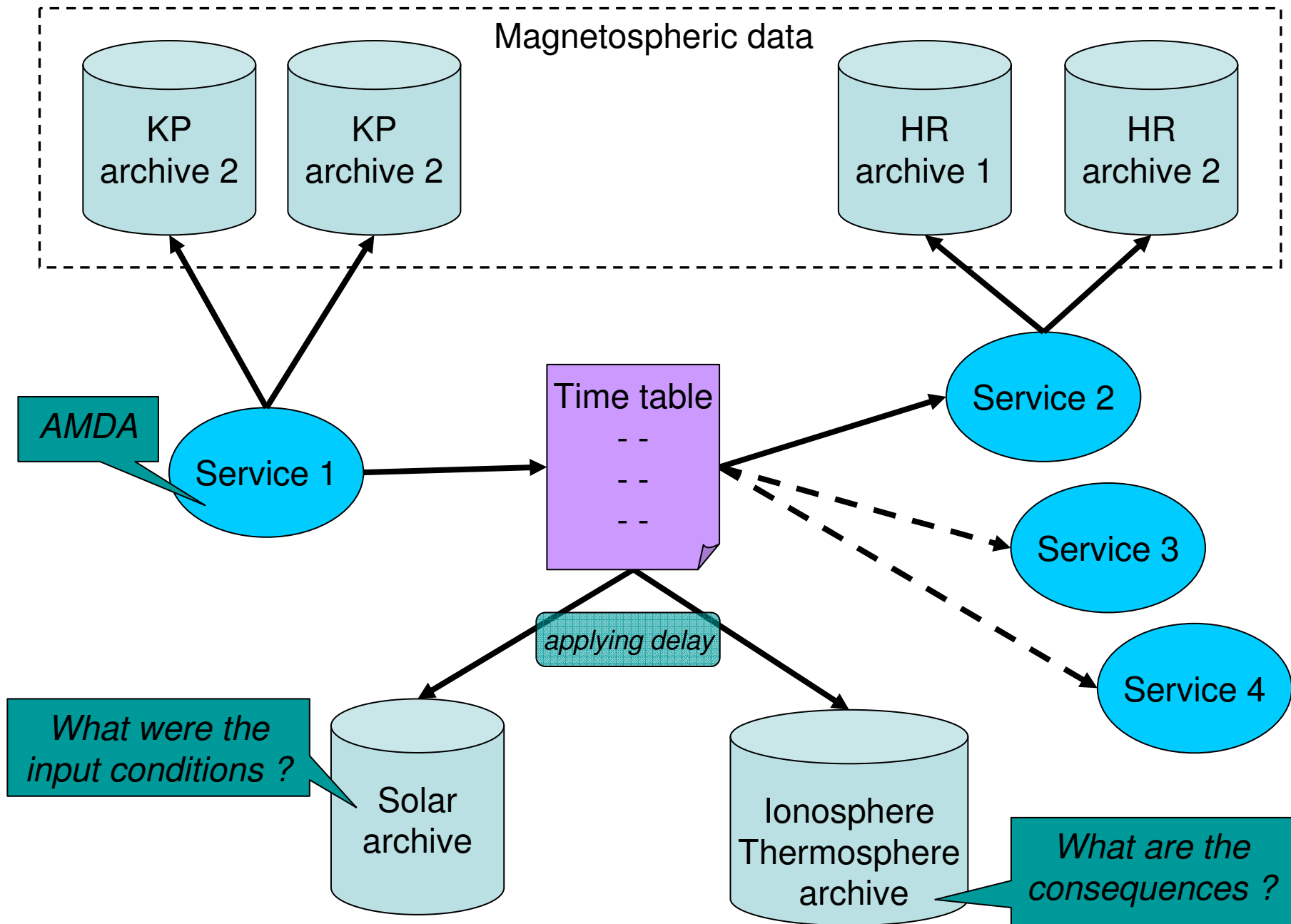
Les tables peuvent être éditées dans l'espace de travail

Tables d'événements 2/2

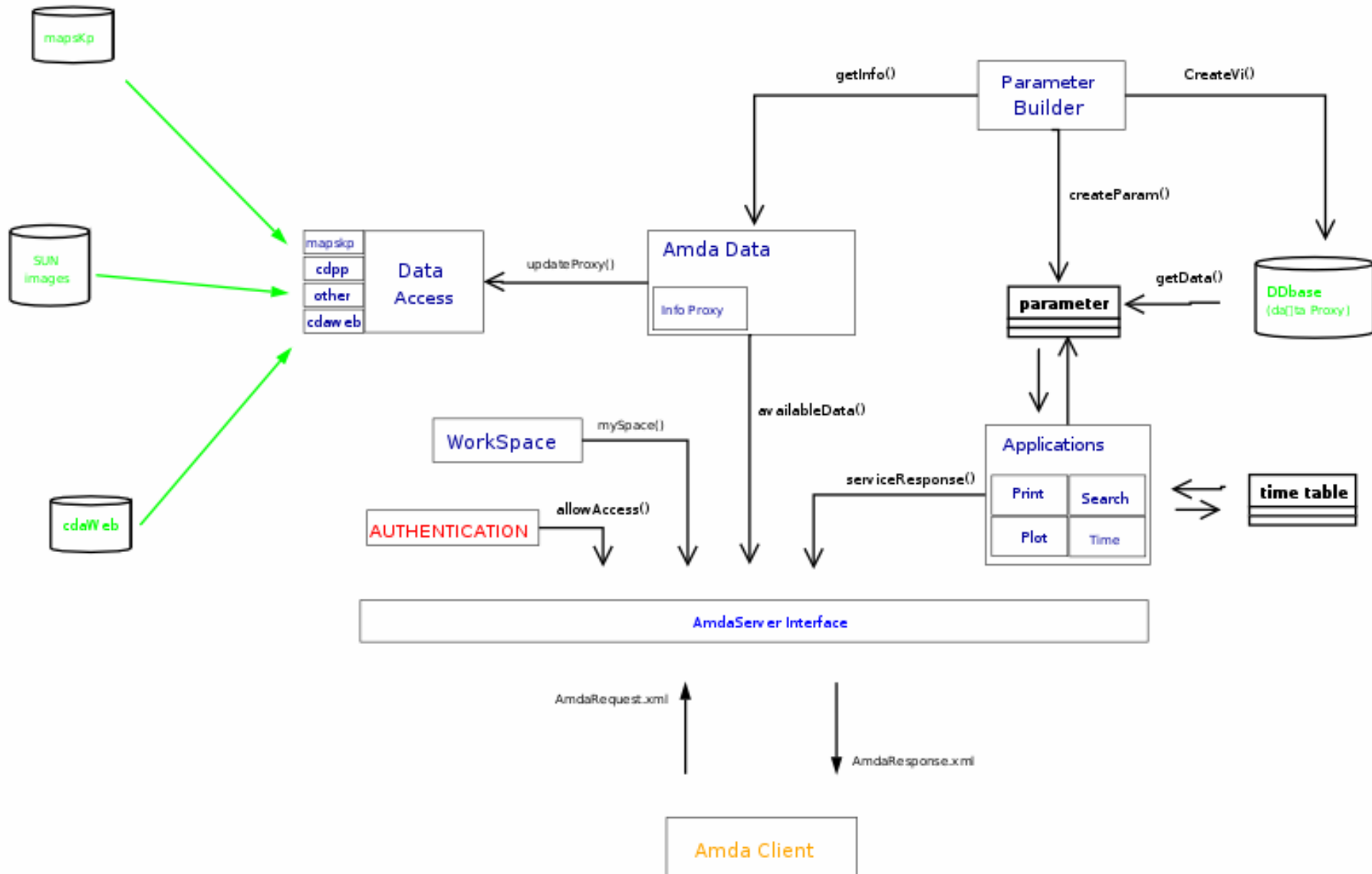
- La description complète d'une table d'événements peut être compliquée
 - sources de données
 - critères de sélection
 - paramétrage des méthodes utilisées (filtre temporel par ex.)
 - ...

- + Les tables d'événements peuvent être un outil performant de collaboration entre centres de données
 - *initiation d'une collaboration entre CAA, GAIA, CDAWeb, et CDPP*

Time table circulation in a multi-archive environment



Architecture d'AMDA



Connexion à des bases externes

1. Lorsqu'une base est connectée à AMDA toutes les données de la base sont visibles dans AMDA (WebService)
 2. L'utilisateur définit le groupe de paramètres de la base externe qu'il veut analyser (Interface de construction d'arbre)
 3. A la volée, les paramètres sont directement ingérés par AMDA (WebService) afin d'être manipulés comme des paramètres de la base locale
- Pas de notion de fichier pour l'utilisateur
 - Pas de problèmes de mise à jour des données
 - AMDA peut fonctionner sans base locale donc est intégrable dans un OV

Interface de construction de l'arbre de données extérieures

Bases externes

Arbre des données choisies par l'utilisateur

The image displays two side-by-side tree views. The left view, titled 'External Tree', shows a hierarchical structure of external data sources. The right view, titled 'User's Tree', shows the tree structure selected by the user, with a 'save tree' button at the top. A green arrow points from the 'Np' node in the 'External Tree' to the 'Np' node in the 'User's Tree'. On the far right of the 'User's Tree' panel, there is a vertical column of six red 'X' marks.

External Tree

- close all open all
- CDAWEB
 - ACE
 - EPM
 - MFI
 - SIS
 - SSC
 - SWE
 - AC_H2_SWE
 - Np
 - Vp
 - Tpr
 - alpha_ratio
 - V_GSE
 - V_RTN
 - V_GSM
 - SC_pos_GSE
 - SC_pos_GSM
 - AC_K0_SWE
 - AC_H0_SWE
 - AC_K1_SWE
 - SWI
 - ULE

User's Tree

save tree

- close all open all
- CDAWEB
 - ACE
 - SWE
 - AC_H0_SWE
 - Np
 - Vp
 - Tpr

Vertical column of red 'X' marks on the right side of the User's Tree panel.

Arbre de données extérieures

My Workspace Plot Data Download E

Select parameters to construct new Workspace parameter

open all | close all

- ⊕ Missions
- ⊕ Ground-based Indices
- Model Parameters along Orbit / Time Series
- Models along Orbit / Space

My External Data

- ⊖ close all ⊕ open all
- ⊖ CDAWEB
 - ⊖ ACE
 - ⊖ SWE
 - ⊖ AC_H0_SWE
 - Np
 - Vp
 - Tpr

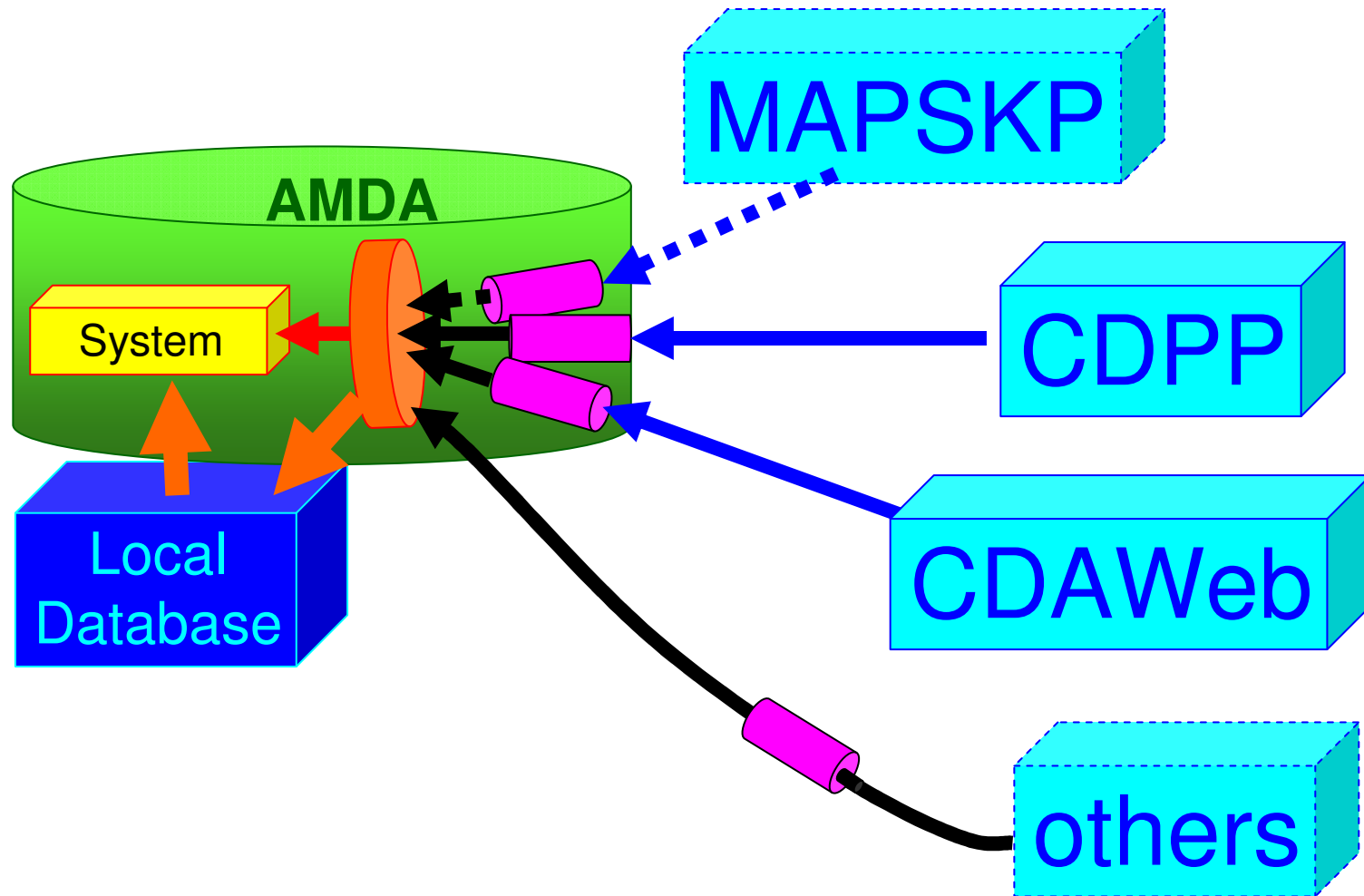
Base locale

Extrait de la base externe

Les paramètres sont disponibles pour des tracés, la recherche conditionnelle, ... à la volée, ce qui implique :

- le téléchargement des données (par bloc de 3j min)
- la transformation de format
- la création du paramètre AMDA

Connexion à des bases externes



Production scientifique avec AMDA

Analyses statistiques multi-missions

- Principe d'utilisation
 - Définition d'un critère sélectif sur les données
 - Critère mathématique
 - Critère visuel
 - Création de table d'événements
 - Itération avec possible combinaison de critères
 - Extraction des données et travail hors-ligne

- Exemples
 - Multi-mission study of plasma sheet conditions minutes before substorm onsets (*Jacquey et al.*)
 - The active plasma sheet: definition of 'events' and statistical analysis (*Louarn et al.*)
 - Statistical study of Alfvénic fluctuations in the Earth magnetosheath (*Alexandrova et al.*)
 - Statistical study of mirror mode fluctuations in the Earth magnetosheath (*Génot et al.*)

- Collaborations
 - Visites scientifiques au CDPP / Organisation d'ateliers
 - Mise à disposition de fonctionnalités évoluées non implémentées dans le prototype actuel (fonctions temporelles)

AMDA-NG

- AMDA “refactoring”
 - ⇒ Modularisation
 - ⇒ Documentation
 - ⇒ Fiabilisation
 - ⇒ Fonctionnement “universel” (indépendent de la plate-forme, de l’OS, du navigateur, ...)

- Mise en place de nouvelles fonctions
 - ⇒ Fonctions temporelles
 - ⇒ Time-Table Manager
 - ⇒ Fonctions Physiques
 - ⇒ ...

- Développement des capacités d’interopérabilité d’AMDA
 - ⇒ SPASE-based Connector
 - ⇒ IVOA-based connector
 - ⇒ Time-Table exchange