ESPM-16

6-10 September 2021,Online

www.aida-space.eu

European Commission

HORIZON 2020

This project AIDA receives funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No 776262.

AIDA: Artificial Intelligence Data Analysis

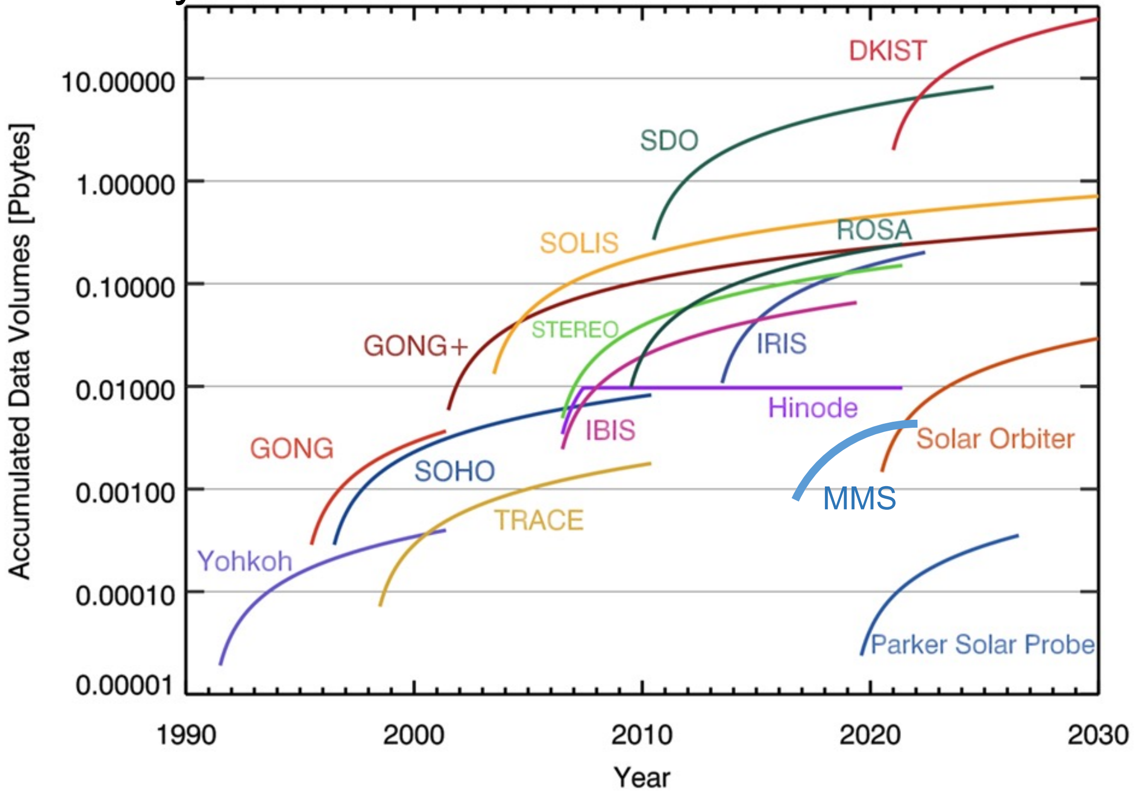
Giovanni Lapenta
(KULeuven)

AIDA Consortium

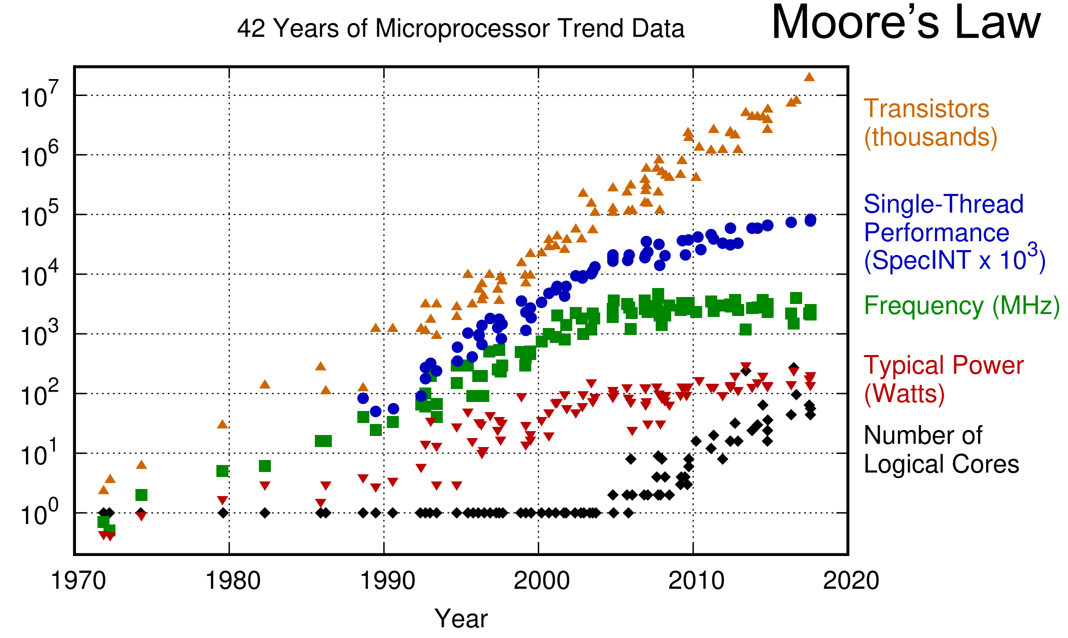
KU LEUVEN	KATHOLIEKE UNIVERSITEIT LEUVEN – Belgium (Coordinator)
UNIVERSITÀ DELLA CALABRIA	UNIVERSITA' DELLA CALABRIA – Italy
UNIVERSITÀ DI PISA	UNIVERSITA' DI PISA – Italy
CWI Centrum Wiskunde & Informatica	CENTRUM WISKUNDE & INFORMATICA – Netherlands
IRIDA LABS	SYSTHMATA YPOLOGISTIKIS ORASHS IRIDA LABS AE – Greece
CINECA	CINECA CONSORZIO INTERUNIVERSITARIO – Italy
Space Consulting International	SPACE CONSULTING INTERNATIONAL LLC – United States
cnrs	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS – France

AIDA builds on 3 current trends in science:
1. Growth in data (from observations and simulations)

Courtesy of NSF



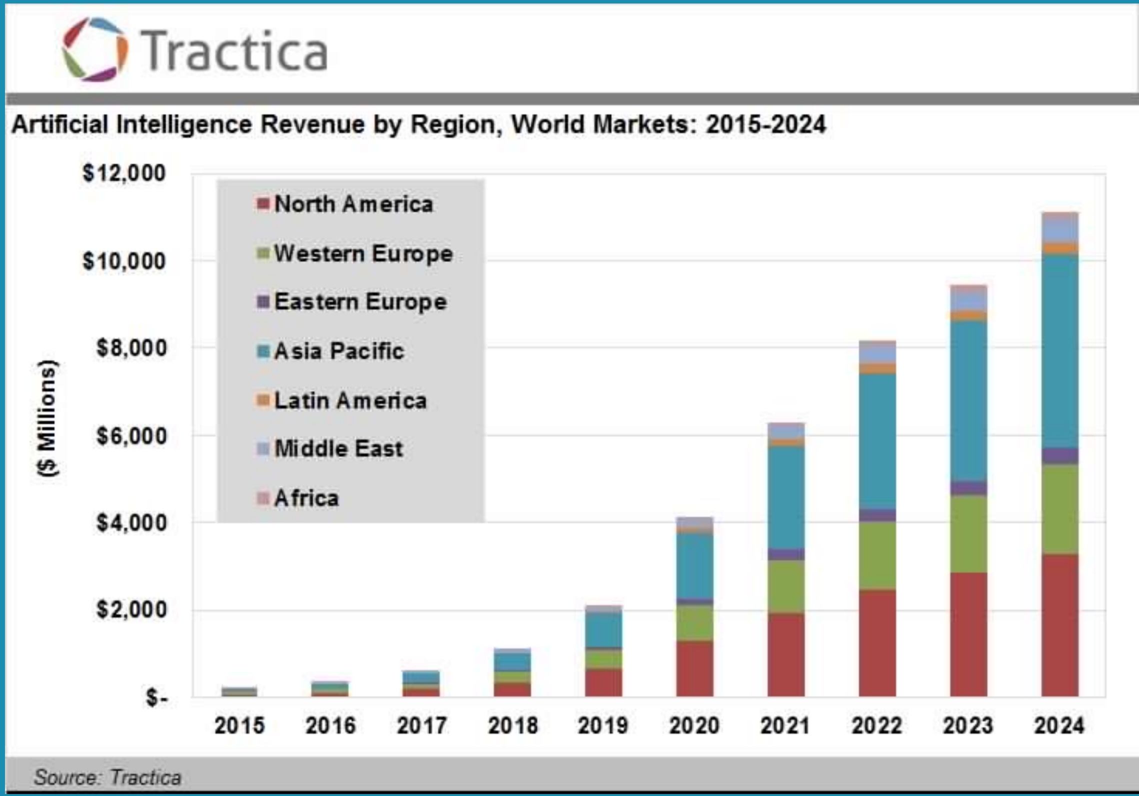
Moore's Law



AIDA provides tools to access, organise and analyse a wide range of data sources

AIDA builds on 3 current trends in science:
2. Artificial Intelligence

- Recent EU projects: Flarecast, AIDA
- US projects: Solstice
- Growth in publications, conferences

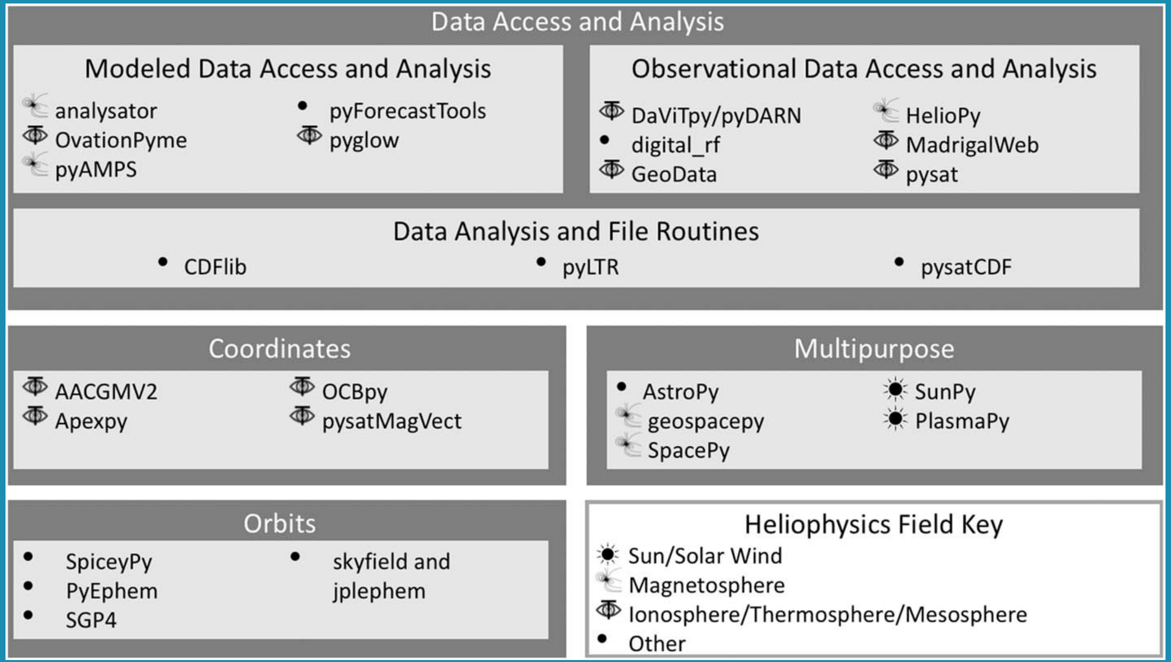


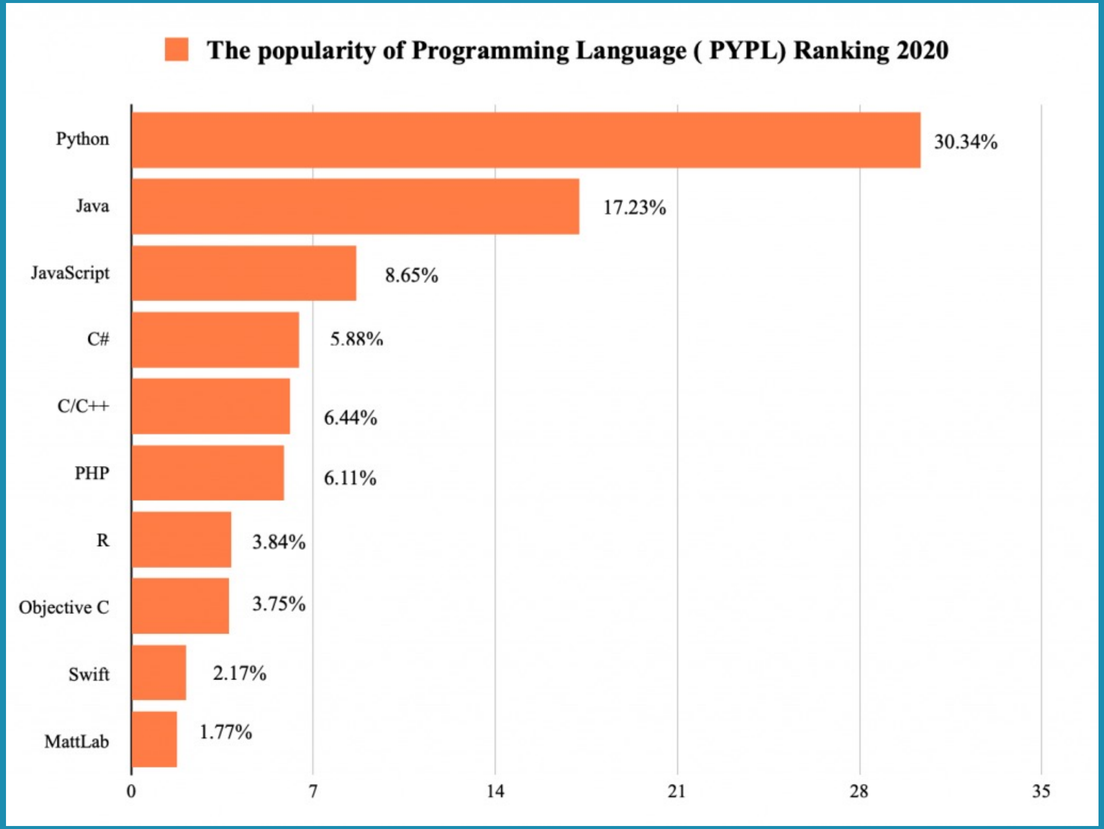
AIDA built several new space science ML tools, open source available to all.
AIDA provides simplified access and training examples of ML tools.
AIDA makes powerful ML software accessible to the space community.

KU LEUVEN

AIDA builds on 3 current trends in science:
3. affirmation of python

Burrell, Snakes on a Spaceship, JGR, 2018





AIDA collaborates with many other space-related python tool developments and contributes its own data management, statistical and ML tools.

The main tools produced by AIDA

AIDApy

- Python based tools for
  - *Data analysis*
  - *Machine learning*
  - *Deep Learning*
  - *Virtual spacecrafts*
  - *Real spacecraft data*
  - *Data assimilation*
  - *Validation and Verification*
- Interoperability with ongoing efforts
- Interoperability between virtual (simulation) and real (space mission) observations

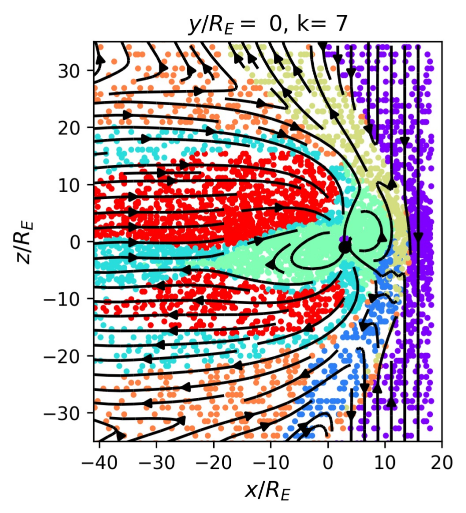
AIDAdb

- Database that includes:
  - *Simulations*
  - *Trained neural networks*
  - *Lists of events*
  - *Training sets*
  - *Examples and tutorial exercises (e.g. School 1 taped and School 2 TBD)*
- Based on EUDAT and iRODS
- Link with public databases via AIDApy
- Metadata specifications compatible with other efforts in the world (e.g. NASA planetary data server)

AIDA services for space weather

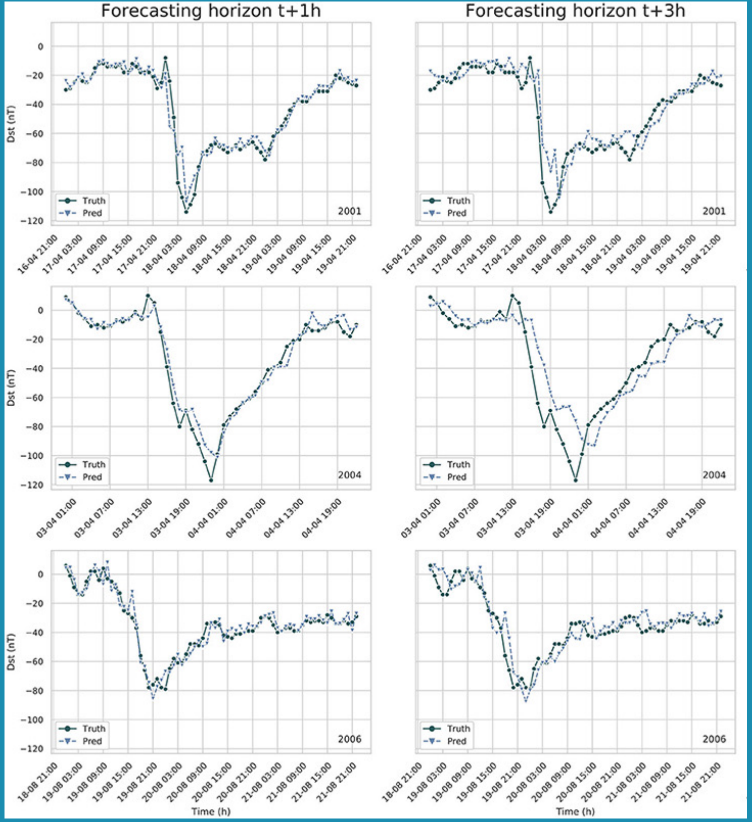
SOM classification of OpnGGM data

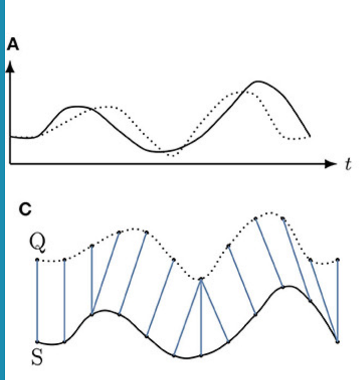
Millas et al.
Front. Astron. Space Sci., 08 October 2020
https://doi.org/10.3389/fspas.2020.571286

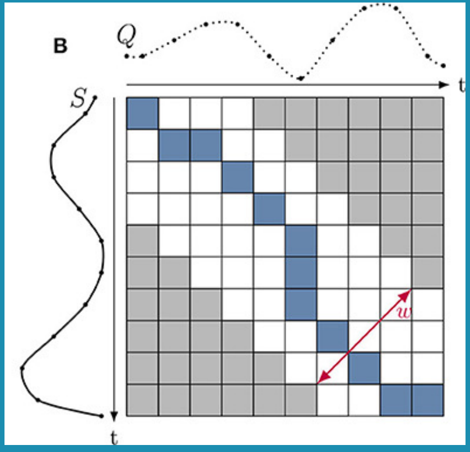


Class	Region
0	Pristine solar wind
1	Magnetosheath (DS BS)
2	Boundary layer
3	Plasmasphere
4	Magnetosheath (DS BS)
5	Magnetosheath
6	Lobes

- *Supervised Classification* of Plasma Regions in **Near-Earth Space**: applied to **MMS data** using CNN
- *Unsupervised Classification* of **incoming solar wind using Dimensionality reduction** and **Self Organizing maps (SOM)**
- *Unsupervised Classification* of Plasma Regions in **Near-Earth Space**: applied to **OpenGGCM** simulations using SOM
- *Prediction of DST index* and **time-warping methods** to establish the accuracy of **predicting storm times**.
- *Data Assimilation* methods based on Kalman filters: application of **representer technique to OpenGGCM and EUHFORIA**
- *Solar image segmentation with NN:* identification of coronal holes for space weather prediction







DST Iundex Prediction and Time Warping

Laperre et al., Front. Astron. Space Sci., 22 July 2020
https://doi.org/10.3389/fspas.2020.00039

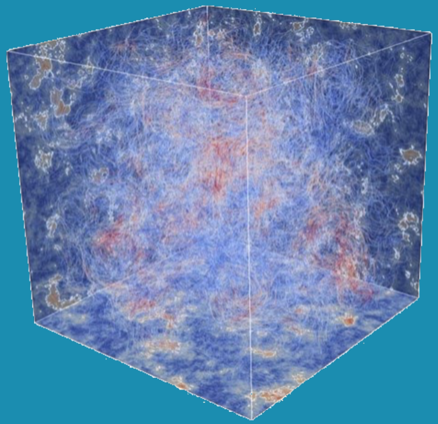
KU LEUVEN

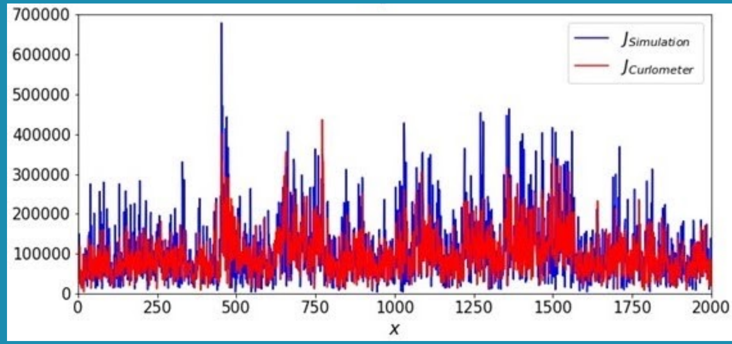
AIDA services for space missions

Supervised classification of MMS data

H. Breuillard et al.
Front. Astron. Space Sci.,
03 September 2020
https://doi.org/10.3389/fspas.2020.00055

- *Data retrieval tools:* avoiding any replication with other python tools, e.g. pySPEDAS, Heliopy.
- *Link with ongoing missions:* **SDO**,**MMS** and **PSP.**
- *Link with upcoming data:* strong initiative to provide AidaPy services to **SolO**
- *Virtual Mission Tool:* to create synthetic data from simulation as if it were created by a missiong under design
- *Linking simulation and observation:* so that the same analysis can be applied to both
- *ML tools to identify extreme events:* e.g. shock, reconnection.
- *ML-driven creation of lists of events*
- *Explainable ML tool for classification of in situ data:* exmple of MMS in previous slide.

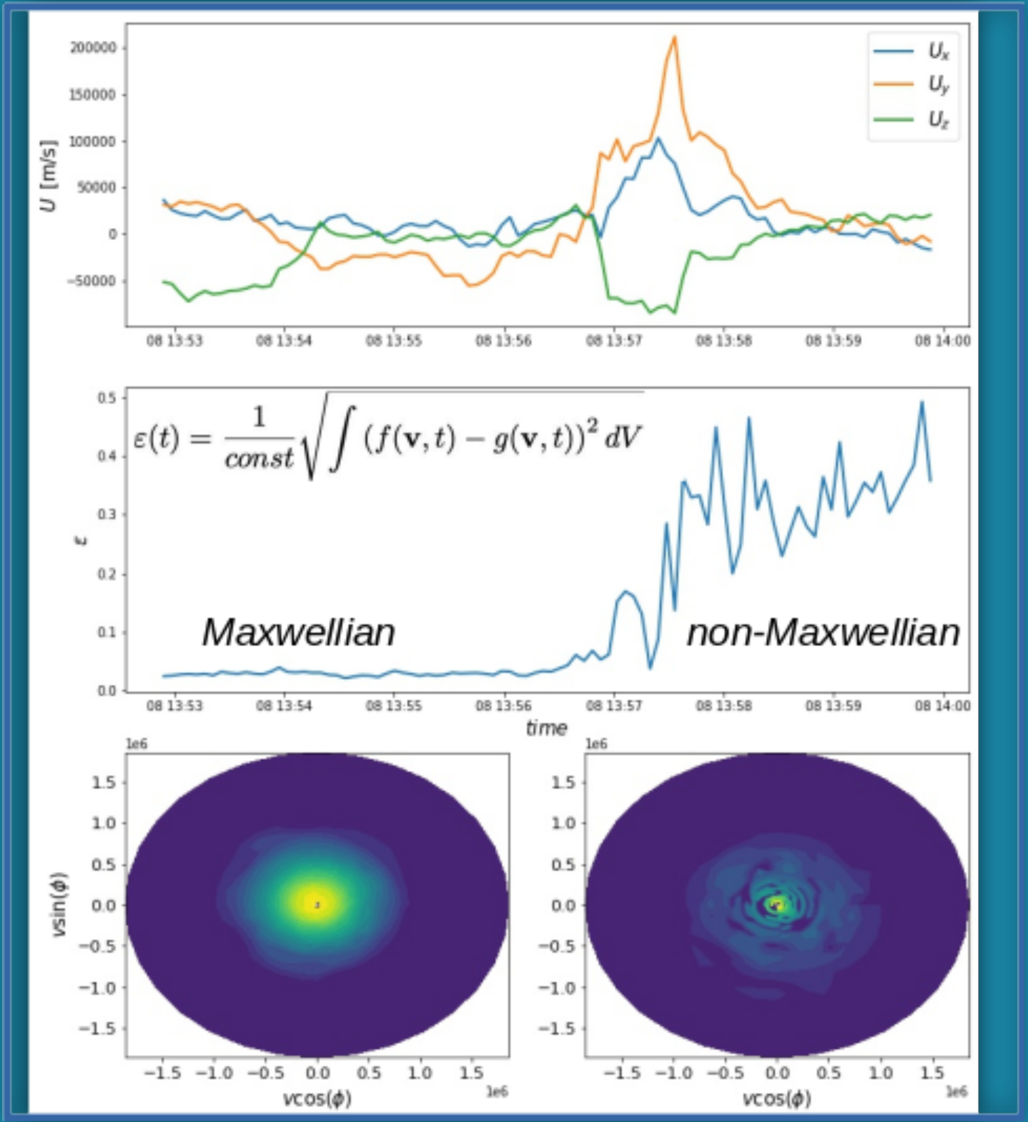




F. Valentini: *Curlometer technique applied to the time series of magnetic field from a tetrahedron of* ***virtual*** *spacecraft launched across the numerical box of a turbulence simulation.*

AIDA services for scientific discovery via data analysis

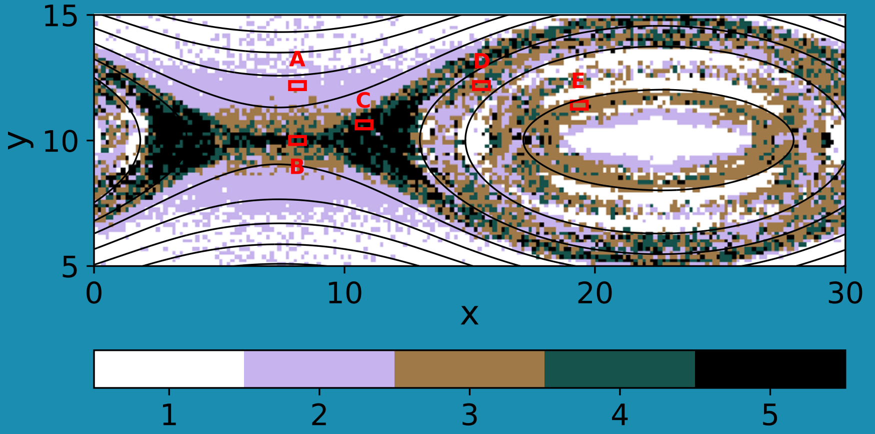
- *Reliance on unsupervised ML* tools that can discover unexpected features
- *AidaPy statistical tools package:* to standsrdise operations typically done in C, IDL
- *Statistical Analysis of VDF:* indicators of complexity, non maxwellianity, beams
- *ML analysis of VDF:* K-means, gaussian mixture method (GMM), SOMs, superposition of kappas.
- *Identification of Reconnection:* using simulation and observation together and using supervised ML trained on human-labelled events and using unsupervised methods.
- *Analysis of Turbulenct structure:* using unsupervised ML (DBSCAN)



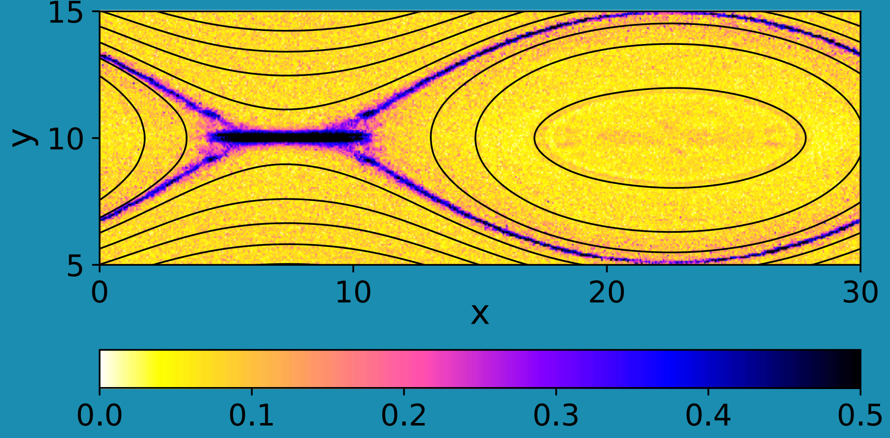
F. Sahraoui: Top: Ions bulk moments, revealing the encounter of a shock. Middle: non-Maxwellianity parameter. Bottom: 2D slices of the VDF in the Maxwellian (left) and non-Maxwellian (right) region.

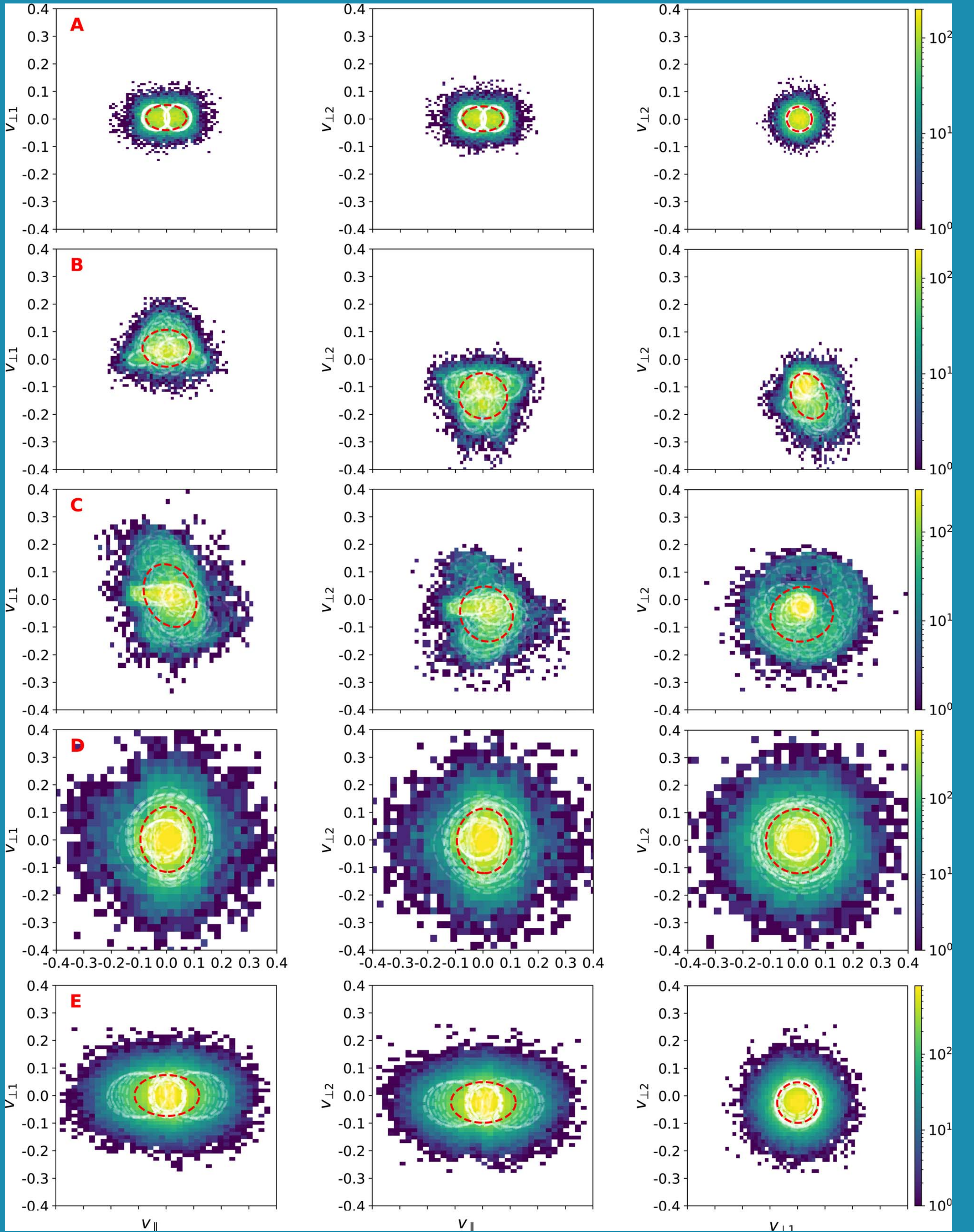
Using ML based on particle distributions

Types of distrubutions(GMM)



Agyrotropy



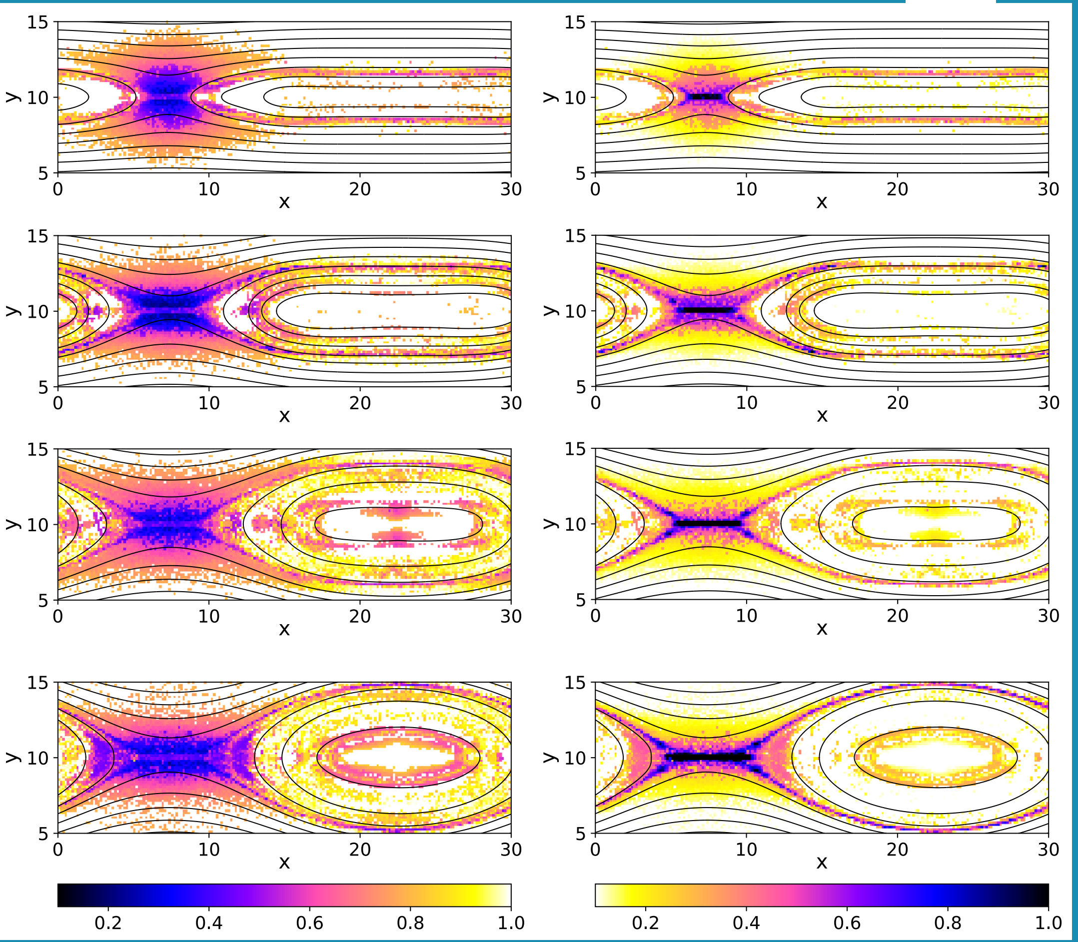


Dupuis, R., et al(2020). ApJ, *889*(1), 22.

Effect on the definition of thermal energy

$$E_{\text{drop}} = \frac{E^{(K)}_{\text{thermal}}}{E_{\text{thermal}}}.$$

$$E^{(K)}_{\text{dev}}$$



- Fluid thermal energy:

$$E_{\text{thermal}} = \frac{1}{N_p} \sum_{i=1}^{3} \left[\sum_{p} (\boldsymbol{V_p} - \langle \boldsymbol{V_p} \rangle)^2 \right]_i, \text{ with } \langle \boldsymbol{V_p} \rangle = \sum_p \frac{\boldsymbol{V_p}}{N_p}.$$

- Multibeam thermal energy

$$E^{(K)}_{\text{thermal}} = \frac{1}{2} \sum_{i=1}^{3} \sum_{k=1}^{K} w_k^2 [\boldsymbol{\sigma}^2_{\boldsymbol{k}}]_i.$$

- Drop in thermal energy

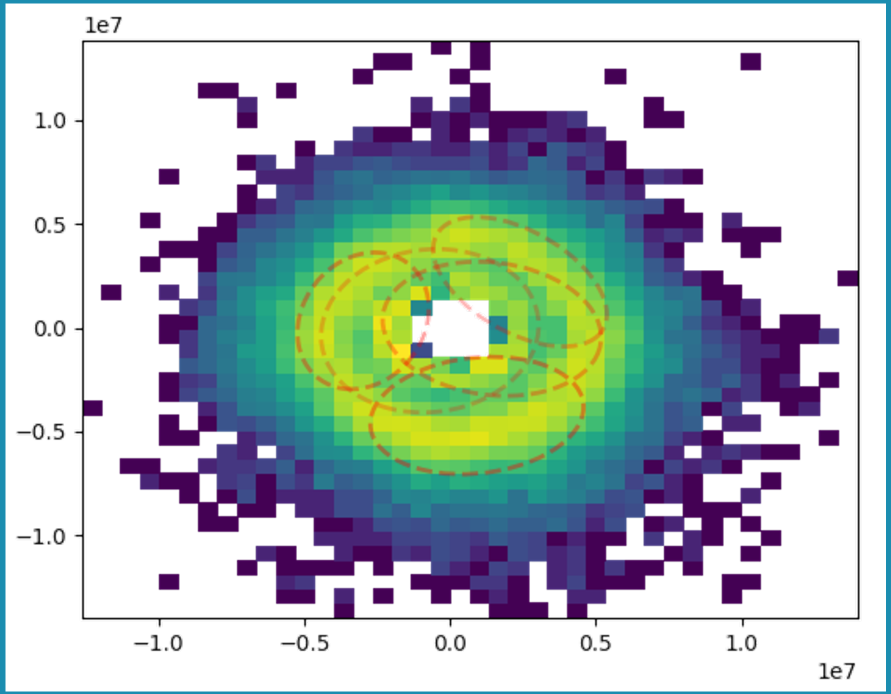
$$E_{\text{drop}} = \frac{E^{(K)}_{\text{thermal}}}{E_{\text{thermal}}}.$$

- Pseudo (False) thermal energy

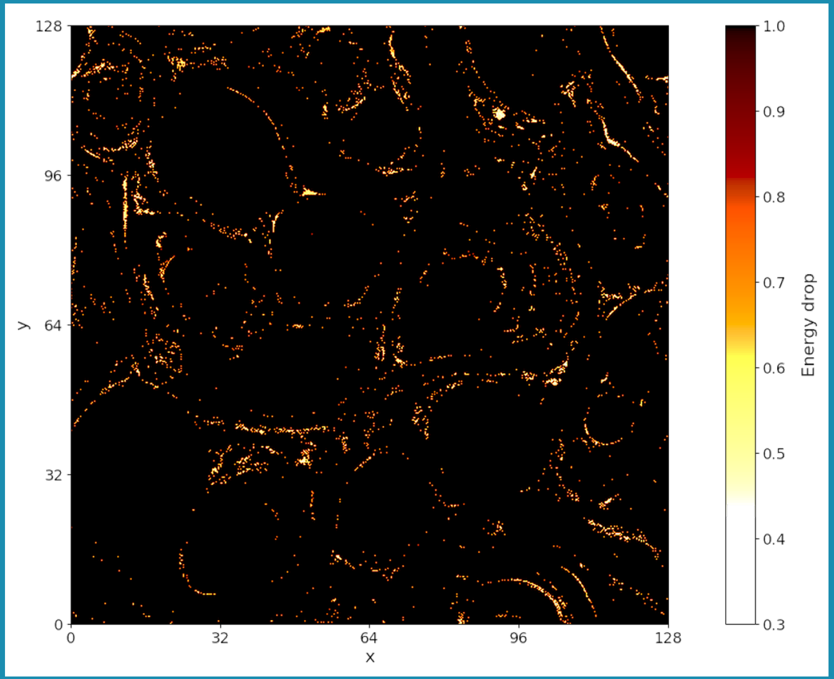
$$E^{(K)}_{\text{dev}} = \sum_{i=1}^{3} \left[\sum_{k=1}^{K} w_k (\boldsymbol{\mu_k})^2 - \left(\sum_{k=1}^{K} w_k (\boldsymbol{\mu_k}) \right)^2 \right]_i.$$

Dupuis, R., et al.(2020). ApJ, *889*(1), 22.
Goldman, M. V., et al. *JGR* 125.12 (2020): e2020JA028340.

Extension to MMS observations and turbulence simulations



MMS Reconnection event:
16 October 2015-13:07:02.235



Regions of interest from turbulent simulations

To know more about AIDA

- ✓ Web site: http://www.aida-space.eu
- ✓ Talks and Publications: http://www.aida-space.eu/publications
- ✓ AIDApy and AIDApy: https://gitlab.com/aidaspace