Python tools for CDPP/AMDA and Machine Learning

Schulz Alexandre [1] Benjamin Renard [1] Nicolas André [1] Christian Jacquey [1] Vincent Génot [1] Myriam Bouchemit [1] Nicolas Dufourg [2] Illya Plotnikov [1] Nicolas Aunai [3] Alexis Jeandet [3]

[1]Centre de Données de la Physique des Plasmas, IRAP Toulouse [2]CNES, Toulouse
[3]Laboratoire de Physique des Plasmas, École Polytechnique, Palaiseau

May 9, 2022

Space Physics made EASY

speasy

speasy is a python package for downloading datasets from various space data providers.

- Latest stable version : **0.10.1**
- GitHub : `https://github.com/SciQLop/speasy.git`
- Documentation : `https://speasy.readthedocs.io/en/stable`
- Published on PyHC : `https://heliopython.org/projects/`
- Supported data providers include :
  - AMDA [**1**] : `http://amda.cdpp.eu`
  - CDAWeb : `https://cdaweb.gsfc.nasa.gov`
  - SSCWeb : `https://sscweb.gsfc.nasa.gov`
- Presented at:
  - PyHC 2022 summer school
  - Solar Orbiter 2022 summer school

Functionalities

- Downloading **time-series**, **time tables** and **catalogs**
- Navigating provider **inventory**: missions, instruments, datasets and components
- Local and proxy caching for fast data access
- Future developments:
  - Full CDAWeb support
  - Multidimensional data support

Getting time-series

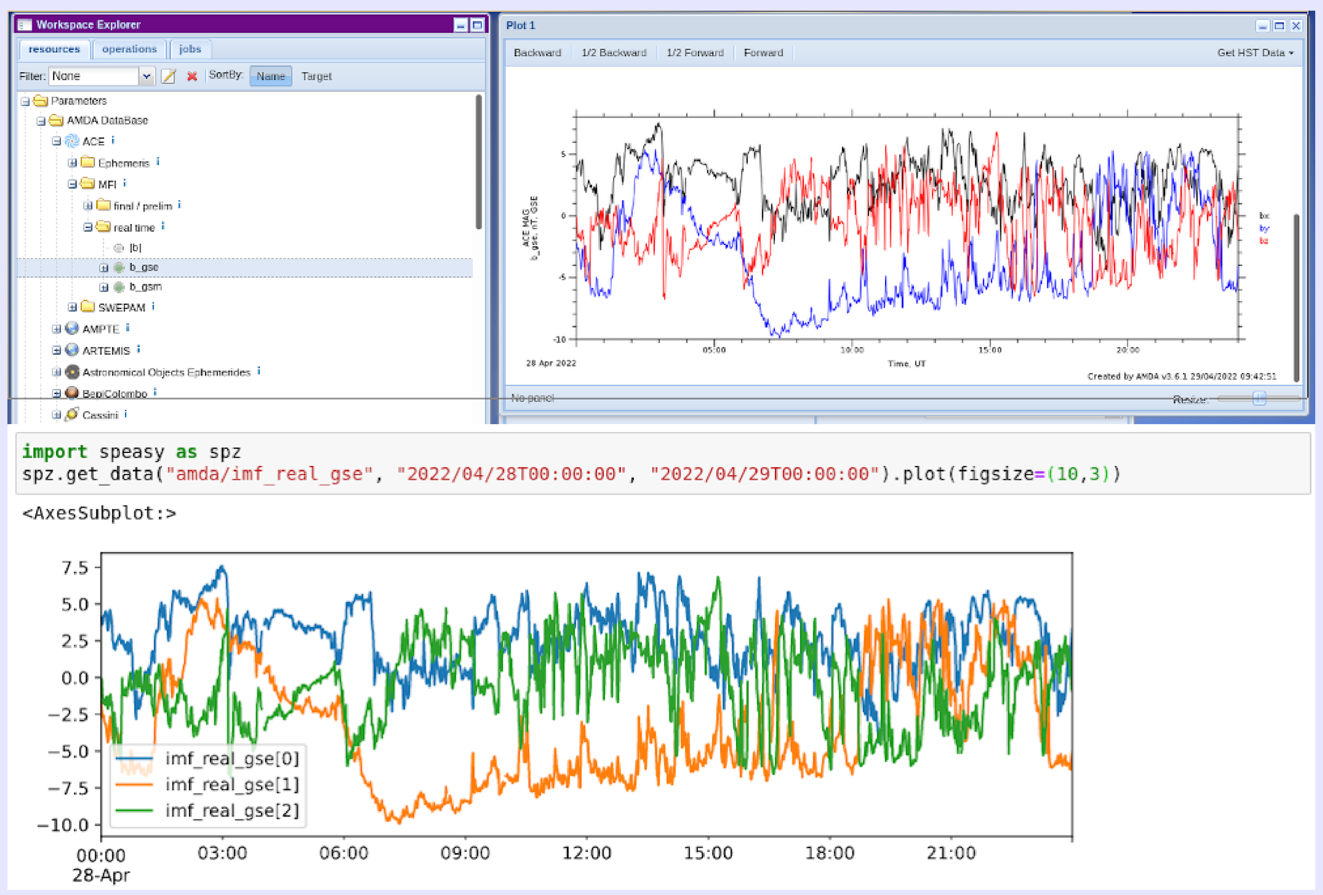


Figure: Plotting data with AMDA and plotting the same data with **speasy**

Caching

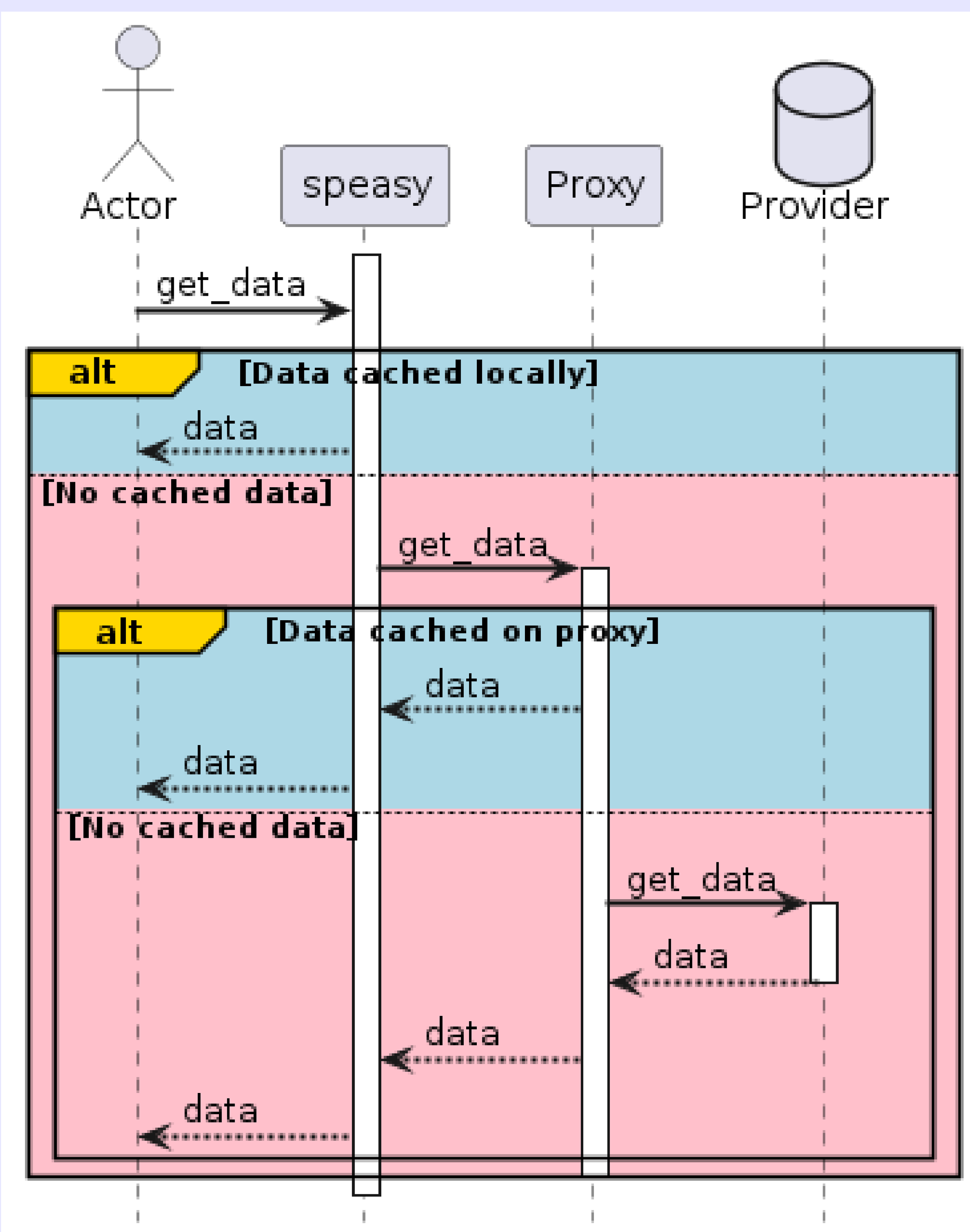


Figure: Caching mechanism flow

AMDA's Machine Learning Pipeline

Orchestra

Orchestra is a python module used by AMDA for managing Machine Learning modules.

- GitHub: `https://github.com/cdppirap/orchestra.git`
- Execution environment using **Docker**
- Integration of models from **Git** repositories
- Machine Learning models directly accessible in AMDA

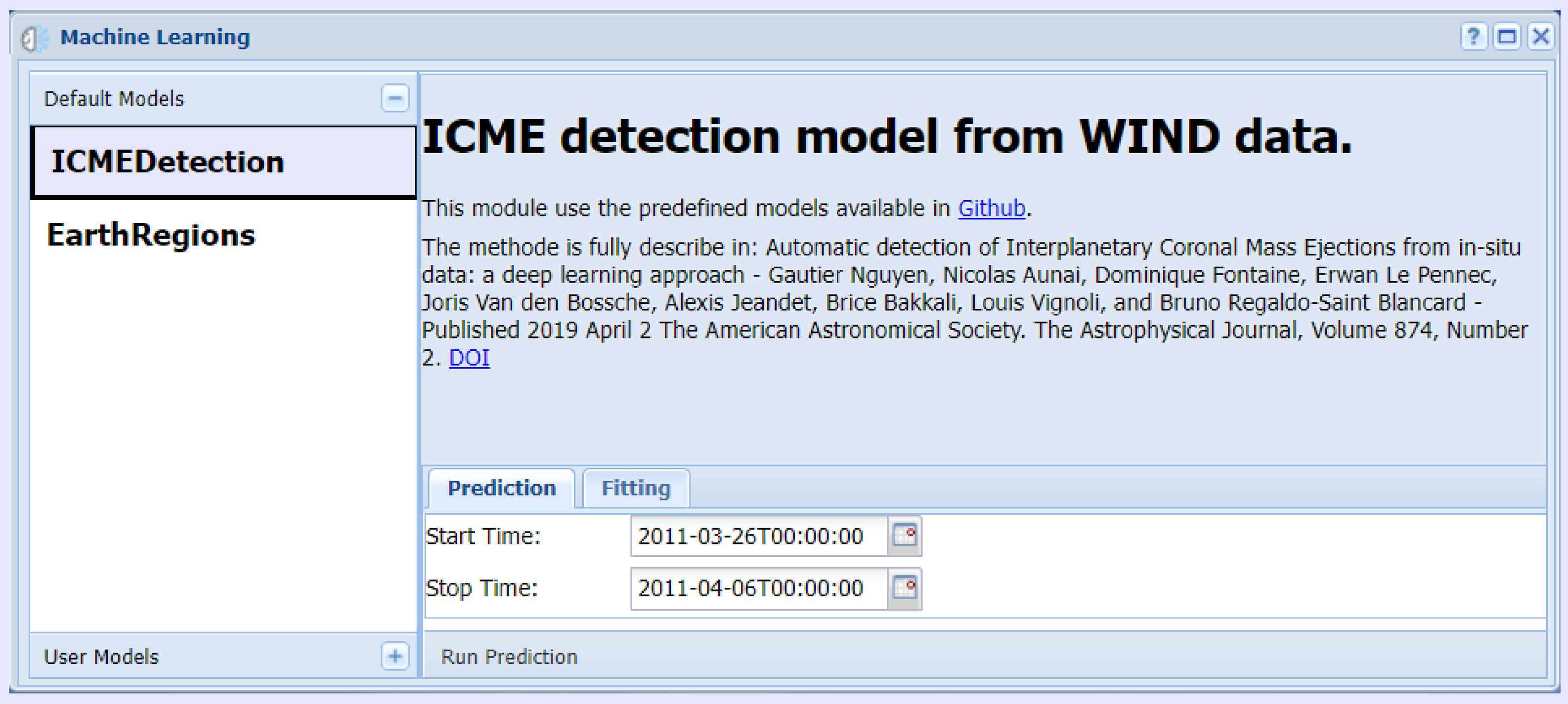
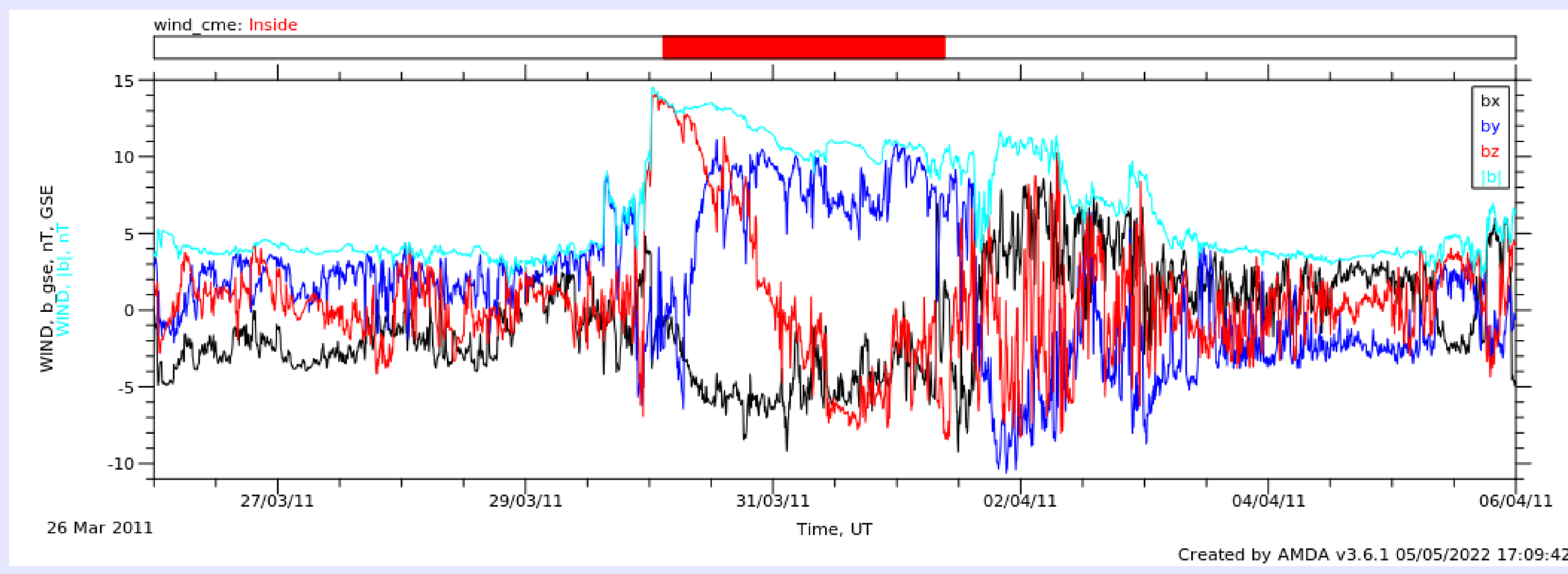


Figure: AMDA's Machine Learning prediction form

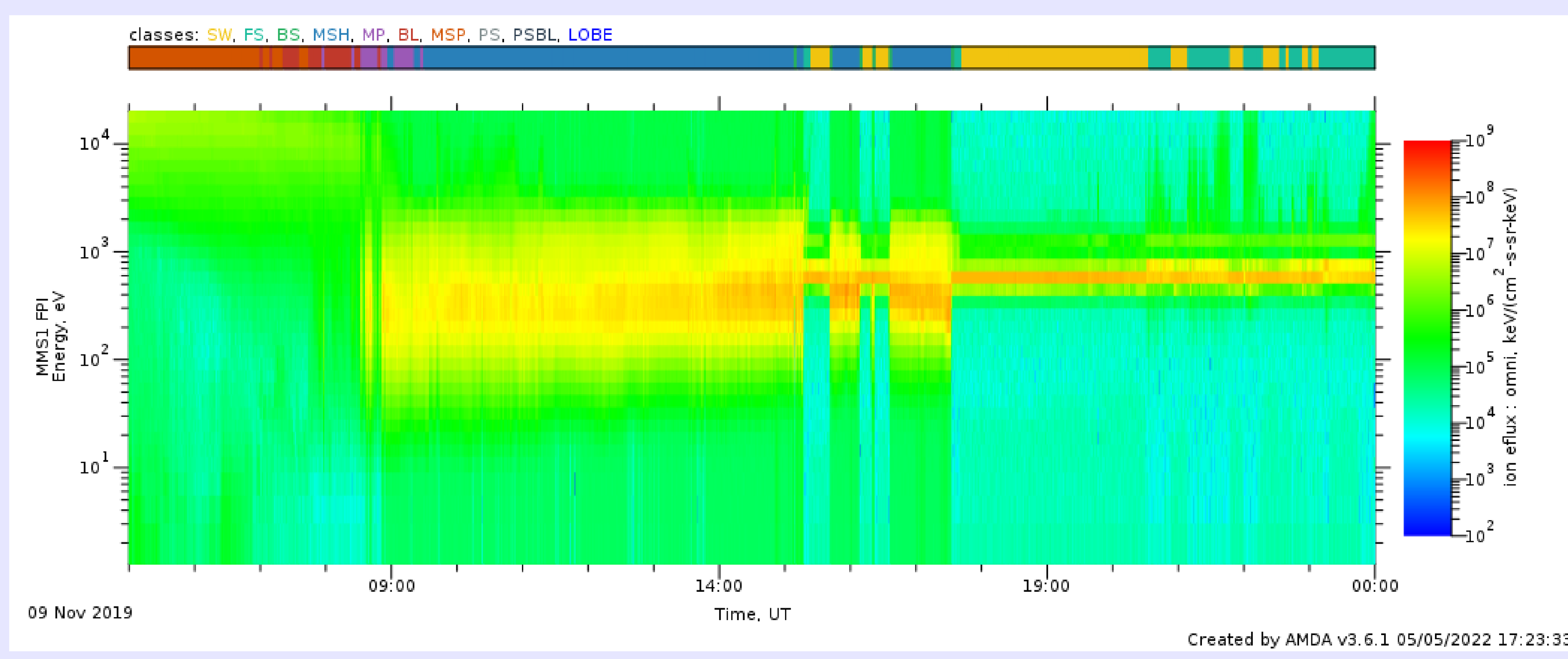
ICME detection

ICME prediction based on the model defined in [**2**] by Nguyen et al. using WIND data.



Earth plasma region detection

Prediction of Earth's plasma region obtained by the model defined in [**3**] by Breuillard et al. using MMS data.



Future developments

- Added support for training models on new datasets
- Include SPIDER project models:
  - Mercury boundary detection models
  - ICME detection models

References

[1] V. Génot et al., *Automated Multi-Dataset Analysis (AMDA): An on-line database and analysis tool for heliospheric and planetary plasma data*, Planetary and Space Science, vol. 201, `https://doi.org/10.1016/j.pss.2021.105214`

[2] G. Nguyen et al., *Automatic Detection of Interplanetary Coronal Mass Ejections from In Situ Data: A Deep Learning Approach*, The Astrophysical Journal, `https://doi.org/10.3847/1538-4357/ab0d24`

[3] H. Breuillard et al., *Automatic Classification of Plasma Regions in Near-Earth Space With Supervised Machine Learning: Application to Magnetospheric Multi Scale 2016–2019 Observations*, Frontiers in Astronomy and Space Sciences, vol. 7, `https://doi.org/10.3389/fspas.2020.00055`