Heliospheric Data Center for Space Weather
ALTEC and OATo – INAF collaboration

ALTEC – Aerospace Logistics Technology Engineering Company
INAF – Italian National Astrophysics Institute
OATo – Turin Astrophysical Observatory
Following the excellent cooperation established for Gaia Italian Data Processing Center, as well as Metis Coronagraph calibration and test campaign, ALTEC and OATo are joining efforts to create the Heliospheric Data Center as a key element of a Space Weather Data Center.
Among other methods, ALTEC/OAT focuses on forecast by means of space observations of the outer solar corona and heliosphere.

- Identification of ejections of magnetic clouds, reaching the Earth magnetosphere:
  - **Long-term forecast (days)** with coronal observations,
  - **Short-term forecast (hours)** with in situ heliospheric observations at L1.

*Solar wind structure also allows predictions, days in advance, of minor geo-magnetic storm.*
Objectives

- Consolidate and evolve the **Heliospheric Data Center**, initially set up with the SOHO data coming from the ESA approved **SOLAR (SOho Long-term ARchive) archive**, in order to manage additional solar archives storing solar coronal and heliospheric data coming from ESA and NASA space programs.

- Develop a **Heliospheric Space Weather Center** for early forecast of impacts of solar disturbances on the heliosphere and the Earth magnetosphere (**forecast** instead of ‘now’cast).

Approach

- The Heliospheric Data Center is a joint effort between ALTEC and OATo.

- The first version of the data center has been developed by building on already existing OATo and ALTEC software tools, infrastructures and competences.
Space Weather Heliospheric Data Center

- Users/Institutions
- Public and private space mission operators
- Airlines companies
- Others

Space Weather National Authorities

Other Italian Data Centers

Data Access Services

Mirror Solar Data
Stereo Data
Metis (Solar Orbiter)
Asperics (Proba 3)
In Situ Li Data
Other Dataset (e.g., Mariner spacecraft)

Data Processing and Analysis Services (ASDP)

Solar Events Characteristics and Tools

IMAP/IMAP-Plus

Near Real-time and Deep-Real-time Tools

Data Retrieval Services

NASA Repositories
ESA Repository
Other Data Providers

Data Backup and Restore
➢ **Scientific capabilities**
  - Solar coronal events identification and characterization
  - Solar coronal events heliospheric propagation characterization
  - Prediction and validation with in situ data

➢ **Technical capabilities**
  - Near real time and off line solar data retrieval from ESA, NASA and others repositories
  - Data management of heterogeneous solar databases
  - Implementation and operation of complex automatic pipelines
  - Space weather prediction tools to compose pipelines and provide services

➢ Front end to Space Weather National Authority to provide forecast, data access, reporting and alerting services
➢ Integration with other space weather data centers to compose additional services
The Heliospheric Data Center has access to several solar databases of both remote sensing and in situ data.

- SOLAR MIRROR ARCHIVE (SOHO)
- STEREO ARCHIVE
- IN SITU L1 DATA (NASA Missions: SOHO, ACE, WIND, DISCOVER)
- OTHER DATASET (e.g. geomagnetic, ionospheric data)

The most relevant data for the current activities are coronagraphic data. At present there are only a very few space-based coronagraphs providing «real-time» images of the solar corona:

- SOHO/LASCO-C2 and LASCO-C3 (1 spacecraft in L1)
- STEREO/SECCHI/COR1 and COR2 (2 spacecraft orbiting around the Sun)
Main Research Fields: Physics of the solar corona and heliosphere, understanding the origin and evolution of the main drivers of geomagnetic storms:
- Solar Wind
- Coronal Mass Ejections (CMEs)

Competence on combined UV & VL coronagraphy and related diagnostics (Doppler dimming techniques)

Development of tools also applicable to Space Weather
- Coronal mass ejection directionality identification (UVCS-SOHO)
- Coronal Mass ejection flag (Metis - Solar Orbiter)
- Identification of coronal wind structure (to infer the heliospheric one) (UVCS-SOHO, Metis – Solar Orbiter)

The OATo Group has a leading role in the coronagraph instrument development and operation
- Development and operation of the UV Coronagraph Spectrometer (UCVS) on SOHO
- Coordination of the development of Metis, Solar Orbiter
Related ALTEC Competences

- Infrastructure designed to maintain PBs of Data (i.e DPCT)
- Integration to external super computer (e.g. CINECA)
- Cloud Environment and several Big Data cluster to develop new innovative project

- Usage of the most largely deployed big data processing frameworks: Hadoop & Spark, OODT etc
- Many years of experience in Relational and NoSQL databases
- Capability to develop an in-house distributed processing framework (ASDP).

- ALTEC has being involved in space operations since its foundation.
- Operations execution of both space ground segment and science data center.
- Evaluate and manage mission operations impacts on data processing activities and services delivery.

- Know-how in several techniques: multidimensional analysis, full and partial cross-correlation, machine learning, classification, etc.
- Several data mining environment : R, IDL, Hadoop, Spark etc.
- Availability of huge datasets to extract information.

BIG DATA ARCHITECTURE
DATA PROCESSING AND MANAGEMENT
MISSION OPERATIONS
DATA MINING

www.altecspace.it
The automatic CME prediction pipeline is implemented by integrating known algorithms developed as modular building blocks:

- The **CME alert module** is based on the algorithm developed for the Metis CME flag (Bemporad et al. 2014).
- The **flare alert module** is based on work done by Caballero & Aranda (2013) for SOHO/EIT images.
- The first **CME kinematic module** is based on the CME cone model by Xue et al. (2005) where CMEs are approximated as circular cone and a hemispherical front on top.
- An alternative **CME kinematic module** will use the algorithm based on measurements of the polarized brightness will be available, it will be possible to apply the polarization ratio.

- The **module dedicated to model the background solar wind** uses radial speed profile based on the empirical model by Sheeley et al. (1997).
- The first implementation of the **CME propagation block** is based on the Drag Based Model + Empirical Solar wind Forecast (Vršnak, Temmer, Veronig, 2007; Rotter et al., 2015) that estimates of arrival time and impact factor with Earth.
- As for previous blocks, the **prediction module**, necessary to predict not only the CME speed and arrival time but also information on the plasma dragged by the CME will be based on empirical relationships derived on a statistical basis (e.g. Owens & Cargill 2004).

In order to prepare input data and manage the output of the CME pipeline several data management activities are executed:

- Near real time data retrieval from NASA and ESA repositories
- Data organization in data store suitable for the processing
- Output data distribution to authorized users and data archiving

www.altecspace.it
Future prospects
Magnetospheric Tools

SOLAR Module
- Flare/CME alert algorithm
- CME kinematic algorithm

HELIOSPHERIC Module
- Background solar wind model
- CME propagation tool
- CME prediction at 1AU

MAGNETOPHERIC Module
- Magnetosphere response
- Ionosphere response
- Geomagnetic storm prediction

Future prospects
Magnetospheric Tools
Start to manage additional data archives of existing missions in order to support new algorithms and improve validation of the existing ones.

- STEREO
- SDO

Set up the data center to manage the new data archives that will be available in the next years when the missions acquiring them will be in the operation phase.

- Metis – Solar Orbiter
- ASPIICS - PROBA-3
Future prospects
Solar phenomena impact on space asset

- Investigation of correlation and cause-effect relations among events occurred in different space position to improve design, operations and maintenance of space asset.
  - The solar data available in the Heliospheric Data Center are used to continue in the development of data analysis techniques to exploit the content of existing very large DBs (“big data”) for the forecast of solar phenomena impact on space assets.

- TECSEL2 (TEmporal Characterization of the remote SEnsors response to radiation damage in L2) is the first ALTEC research project with the goal of providing analyses of generic time series and comparisons between them.
  - This is scientifically interesting because it could provide new information about the not-well-known magnetic environment near the Lagrangian point L2, where Gaia operates.
  - Results showed a correlation between particles fluxes detected at L1 and charge flux and background detected on Gaia CCDs at L2 and processed by AIM. This correlation reaches its maximum with barely 4 hours of delay.
  - The project is developed using only Big Data technologies (NoSQL database, in memory time series processing ecc.)
Thanks for your attention!