EPSC Abstracts
Vol. 13, EPSC-DPS2019-2008-1, 2019
EPSC-DPS Joint Meeting 2019

EPSC
EPSC-DPS Joint Meeting 2019
15–20 September 2019 | Geneva, Switzerland

TECSEL2: a machine learning system to exploit cross-mission datasets

Ruben De March (1), Deborah Busonero (2), Rosario Messineo (1), Alessandro Bemporad (2), Francesco Vaccarino (3), Angelo Fabio Mulone (1) and Mario Lattanzi (2)
(1) ALTEC S.p.A., Corso Marche 79, 10146 Torino, Italy (ruben.demarch@altecspace.it)
(2) INAF-OATo, Via Osservatorio 20, 10025 Pino Torinese, Torino, Italy (deborah.busonero@inaf.it)
(3) Politecnico di Torino, Dept. Math. Sciences, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
(francesco.vaccarino@polito.it)

Abstract

The project TECSEL2 was born at ALTEC S.p.A. for the usage of the Gaia AIM/AVU daily pipeline output and solar events data to characterize the response of detectors subjected to strong radiation damage within an environment not protected by the terrestrial magnetic field, the Lagrangian point L2. The project also aims at identifying anomalies in the scientific output parameters and relate them to detectors malfunctioning due to radiation damage issues correlating with solar events occurred in the same time range. This project is an example of how data that can be processed for other purposes in addition to those they were designed for. TECSEL2 actually designs and implements a system, based on state-of-the-art big data technologies, which goes beyond its initial goal related to the Gaia project, in fact it provides useful analysis processing for other ALTEC projects.

1. Introduction

Gaia is the ESA space mission launched on December 19, 2013, aiming at Global Astrometry at few μas, scanning continuously the whole sky in order to build the largest, most precise three-dimensional map of our Galaxy by surveying more than a thousand million stars [3], [7]. It is the first mission to operate within the Charge Couple Devices bandwidth in L2. Gaia focal plane array with its 106 CCDs is therefore an invaluable source of information about the CCDs behaviour within a strong radiation environment. The Astrometric Instrument Model (AIM) is one of the crucial components of the Astrometric Verification Unit (AVU), the verification counterpart operating independently from the data reduction chain of the Gaia baseline. The AIM system is devoted to the monitoring, diagnostic and calibration of the Gaia astrometric instrument response over the mission lifetime [1], [2]. AVU has its own dedicated Data Processing Centre in Turin, the DPCT, which is one of the six Gaia DPCs spread across Europe. DPCT is also designed to provide computation, storage, data access and operations services to Italian Gaia Science Community [6]. The availability of new big data technologies opens new scenarios in which science data collected inside a specific mission can be used to find new information and correlate it with datasets coming from different sources. The goal of TECSEL2 project is to study the possibility to use data acquired by remote sensors like CCDs to describe the environment in which the CCDs operate. The project starts with the preparation of data coming from Gaia CCDs in order to compare them with events related to the Sun.

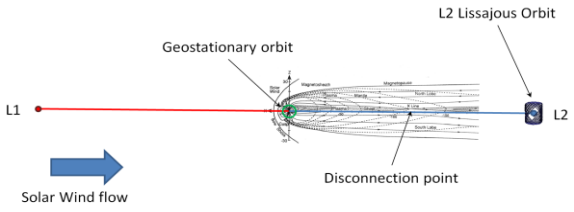

Figure 1: L2 location vs. Earth magnetosphere.

2. Data Processing

2.1 Architecture

TECSEL2 architecture is designed considering the described algorithms but also possible future integrations and other implementations. A high level description of the architecture can be seen in Figure 2. Other implementations of this schema were used for the participation in the Mars Express Power Challenge, promoted by ESA in 2016, and mainly the

analysis for the telemetry data coming from the PMM module of the ISS [5].

TECSEL2 system storage is populated through dedicated extraction/ingestion processes with solar datasets stored in files or with Oracle RDBMS used at DPCT for Gaia's data access and repository. Since the algorithms belong to time series area, the system has to foresee a data storage component that uses a structure suitable for time series and provides scalable data access services. In big data area, this kind of storage system is called "time series database" (TSDB). The TSDB has been constructed comparing the available solutions and taking benefit from studies already conducted on the space sector [4].

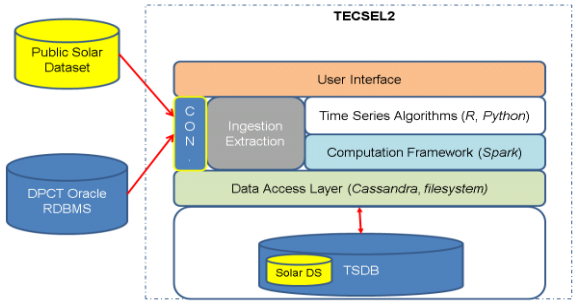

Figure 2: TECSEL2 high-level architecture.

2.2 Algorithms

Main computations involve cross-correlation and partial correlation between two paired samples (which are eventually pre-processed to treat missing values or other metadata anomalies), aiming at finding out if anomalous values of solar wind cause any effect on Gaia CCDs. Moreover, an incrementing time-shift is applied iteratively to one of the series to deduce, e.g., the delay between solar phenomena and their effect on Gaia. This delay depends on the time particles need to reach L2 point, but it can also depend on some additional time it takes for radiation damage to manifest itself.

3. Conclusions

TECSEL2 is an innovative project for several reasons. It is one of the first studies devoted to the monitoring and characterization of the behaviour of CCD detectors located in the L2 environment and it is therefore a key study for future space missions equipped with CCDs array similar to the one used on Gaia. As a service, TECSEL2 system is a powerful tool for efficient analysis of large and generic time series data, built with big data technologies, the state of art for the treatment of huge amount of data like those coming from the new generation of space and on-ground telescope.

Our results show 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. Further analyses are in progress to investigate this effect.

References

[1] Busonero, D.: Astrometric instrument modeling in the context of gaia astrometric verification: tasks and activities," Proc. SPIE 8449, Modeling, Systems Engineering, and Project Management for Astronomy V, 84490F, 25 September 2012.

[2] Busonero, D., et al.: Monitoring, diagnostic, and calibration of the gaia astrometric instrument response within the astrometric verification unit, Proc. SPIE 7731, Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave, 773133, August 10 2010.

[3] de Bruijne, J., et al.: Gaia: 1,000 million stars with 100 ccd detectors, Proc. SPIE 7731, Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave, 77311C, 5 August 2010.

[4] Giotis, I., et al.: Bridging big sensor data analytics from water management to space, in Proc. Big Data from Space (BiDS14). IEEE, 2014.

[5] Leuzzi, C., et al.: Innovative approach for PMM data processing and analytics, SpaceOps Conference, Marseille, France, 2018.

[6] Messineo, R., et al.: The italian dpc: infrastructure and operations for the italian contribution to the gaia data processing and analysis consortium, Proc. SPIE 8451, Software and Cyberinfrastructure for Astronomy II, 84510E, 24 September 2012.

[7] Perryman, M.A.C., et al.: Gaia: Composition, formation and evolution of the galaxy, A&A, vol. 369, pp. 339–363, 2001.