3rd URSI AT-AP-RASC, Gran Canaria, 29 May – 3 June 2022

Study of the Solar wind - magnetosphere coupling with machine-learning methods

Gautier Nguyen*(1)(2), Nicolas Aunai(1), and Dominique Fontaine (1)
(1) CNRS, Ecole polytechnique, Sorbonne Université, Univ Paris Sud, Observatoire de Paris, Institut Polytechnique de Paris, Université Paris-Saclay, PSL Research Univsersity, Laboratoire de Physique des Plasmas, Palaiseau, France
(2) Direction Générale de l'Armement, Paris, France

Abstract

Decades of in-situ data measurement by missions focused on the study of the solar wind and its relation with the near-Earth environment allowed the study of the Sun-Earth coupling from a statistical point of view. Nevertheless, these studies are limited by the manual selection of the events of interest in the data that is still a subjective, fastidious and hardly reproducible task. Using machine learning algorithms, we elaborate automatic detection methods of events from in-situ data measurement that outcome the performances of those based on manual, empirical thresholds and pave the way for statistical studies of in-situ measured events with an important number of samples.

1 Introduction

The solar wind and the Earth magnetosphere form a complex duet which dynamics is ruled by a multitude of physical processes at every steps of this coupling. Upstream, large-scale solar events such as Interplanetary Coronal Mass Ejections (ICMEs) transport important quantities of plasma and magnetic field which entry in the magnetosphere generates geomagnetic storms with a high impact on space-borne and ground-based technological systems. At the Earth proximity, the interaction of two different types of plasma define different regions of the near-Earth environment, delimited by two main boundaries, the bow shock and the magnetopause. Varying solar wind conditions generate small-scale physical processes such as magnetic reconnection that rule the entire dynamics of the system by modifying the location and shape of the boundaries or by allowing the transfer of mass and momentum between the two parts of the couple.

The study of the different actors of this coupling can be done through the statistical analysis of the in-situ data measurement by spacecraft orbiting the Earth or the Sun. Nevertheless, these studies often rely on a small number of samples usually selected after a time-consuming, ambiguous and poorly reproducible manual selection of events. This necessarily restricts the resulting statistical vision we can have on the different events of interest and spoils the potential of the decades of accumulation of spacecraft in-situ data measurement considered altogether. Improving the automatic event detection methods then appears as a necessity in the frame of the construction of a global, statistically representative vision of the different actors of the Sun-Earth relation.

In this paper, we we take a step further in the direction of a global, statistically representative vision of the solar wind-magnetosphere coupling by applying supervised machine learning algorithms to the automatic detection of the in-situ signatures of three of its main actors: ICMEs, the near-Earth environment regions and the magnetopause plasma jets.

2 Automatic detection of ICMEs

ICMEs are the interplanetary counterpart of Coronal Mass Ejections, spectacular manifestations of the solar activity which are responsible of the expulsion at large velocities of large quantities of plasma and magnetic field. For their capacity to trigger geomagnetic storms, ICMEs are often considered as the most geoeffective solar events. For space weather purposes, a global, statistically representative vision of such events would then be the opportunity to better understand their nature and how they interact with the magnetosphere and affect the human activity. Decades of studies have suggested several criteria to detect ICMEs in time series from in situ spacecraft measurements. Among them, the most common are an enhanced and smoothly rotating magnetic field, a low plasma beta, a declining velocity profile and a low proton thermal velocity as shown by the four first panels of Figure 1. However, these features are not all observed for each ICME due to their strong variability. This lack of consensus on the typical in-situ signature of ICMEs associated to the fact these catalogs were elaborated after a manual selection of events resulted in incomplete, ambiguous and hardly reproducible lists which using masks the statistical vision we can have on such events.

To cope with it, we used the data of the solar monitor WIND between 1997 and 2016 and trained an ensemble of Convolutional Neural Networks (CNNs) to estimate the extent at which a sliding window of data is similar to the typical in-situ signature of an ICME through the prediction of a so-called *similarity parameter* [1].

Stacked together, the predictions of this ensemble of CNNs return a 2D-similarity map where the intervals of data are likely to correspond to the in-situ signature of one or several events. This constitutes a first interesting multi-scale visual indicator in the frame of the manual selection. After a small post-processing, these maps can also be used to to rapidly generate reproducible catalogs of events that contain less false positives than the one obtained from based on manual, empirical thresholds [2]. This method currently result in the most exhaustive existing ICME list [1], directly usable for further statistical analysis.

Although less accurate, we show that this ensemble of CNNs works with one or several missing input parameter. We also that its performances improve by just increasing the amount of input data. As it requires no physical prior about ICMEs, this method is easily adaptable to other large-scale solar events such as co-rotating interaction regions of the sheath of ICMEs.

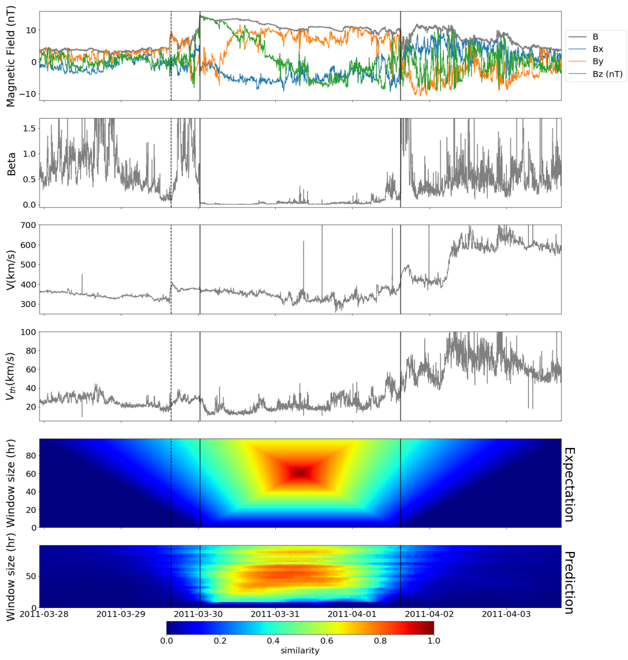

Figure 1. Solar wind observations during an ICME from the WIND spacecraft located at the Lagrangian point L1. The solid vertical lines delimitate the ICME while the dashed vertical line indicate the beginning of the sheath. From the top to the bottom are represented : the magnetic field amplitude and components, the plasma parameter β, the solar wind velocity, the thermal velocity, the artificial proxy we build to indicate the presence of an ICME in sliding windows of various sizes (from 1 hour to 100 hours) and the prediction of our method.

3 Classification of the Near-Earth regions

At first order, the magnetopause and the bow shock are the boundaries of three distinct regions of the near-Earth environment: the magnetosphere, the magnetosheath and the solar wind. By definition, the shape and location of these two boundaries do depend on the upstream solar wind conditions. The ever-growing quantity of near-Earth in-situ data allowed the realisation of statistical studies dedicated to the position, shape and dynamics of both the magnetopause and the bow shock that resulted in the development of numerous surface models for these two boundaries ([3] and [4] for reviews). The first step of both empirical modelling and statistical studies is still the same: establishing a consistent catalog of boundary crossings from the streaming in-situ data provided by missions of interest, which is here again a time-consuming, ambiguous and poorly reproducible task that should be automatized.

To cope with it, we developed an automatic classification method of the three near-Earth regions from their in-situ data measurement by multiple spacecraft [5]. Based on a gradient boosting classifier, the method was successfully applied to the data of 11 different spacecraft with different orbits from the missions Cluster, Double Star, MMS, ARTEMIS and THEMIS for which a typical prediction is shown in the Figure 2. This methods outperforms the previous existing methods based on empirical manual thresholds [6] and equals the performances of the most recently developed methods based on CNNs [7] with the advantage of being faster and lighter to train.

This method is then used to identify 15 062 magnetopause and 17 227 bow shock crossings in the 83 years of cumulated data provided by the mentioned missions. Those catalogs contain a reduced proportion of False Positives and can then be used for future massive multi-missions statistical studies of the near-Earth boundaries.

4 Automatic detection of magnetopause plasma flow

Magnetic reconnection occurring at the magnetopause surface is the dominant physical process when it comes to the transfer of momentum between the solar wind and the Earth magnetosphere. The merging and convection of the field lines erodes the magnetosphere in a location that varies with the orientation of the Interplanetary Magnetic Field (IMF), on the dayside when it is southward, at higher latitudes when it is northward. Because it is a key actor of the solar wind-magnetosphere system a precise knowledge about its location for a given set of upstream solar wind conditions is a key element of the understanding of the dynamics of this coupling. A solution that could address this question would stand in the collection of as many in-situ evidence of magnetopause plasma flow as we can and superimpose these events, together in the form of magnetopause plasma flow maps. Assuming it is the main process at the origin of the magnetopause flow and if they are steady enough, a global pattern would appear in the different produced flow maps and would follow the expected evolution of the location of magnetic reconnection.

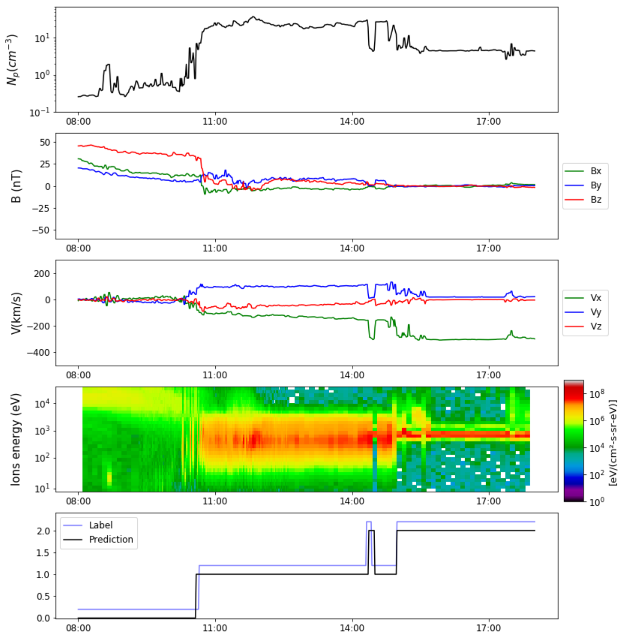

Figure 2. In-situ measurement provided by THEMIS B spacecraft on May 12, 2008. From the top to the bottom are represented: the ion density, the magnetic field components, the velocity components the omnidirectional differential energy fluxes of ions. The last bottom panel represents the evolution of the label (blue), intentionally shifted for visual inspection and the prediction made by the region classifier (black). 0 indicates the magnetosphere, 1 the magnetosheath and 2 the solar wind.

However, the selection of those events, that corresponds to plasma jets during a magnetopause crossings, is ambiguous because of its dependency on the external observer's decision and because of the variability that exists globally in the dataset. Once again we face a time-consuming problem which complexity is a serious bottleneck to these studies that often leads to poorly reproducible catalogs limited to the few most obvious events.

To cope with it, we combine the region classifier presented in the previous section to a peak detection and a second gradient boosting classifier to provide an automatic detection of the magnetopause plasma jets on the low-latitude, dayside magnetopause. First trained on THEMIS C data the method provides a detection with less false positives than the one obtained with manual thresholds [8]. It also works without loss of performance on the data of other spacecraft that have an equatorial orbit, the other THEMIS, Double Star and MMS and lead to the production of a catalog of magnetopause plasma flow evidence that contains 17 957 events, the most exhaustive and accessible existing list so far. Which paves the way for the elaboration of statistical studies of plasma flow induced by magnetic reconnection at the dayside magnetopause.

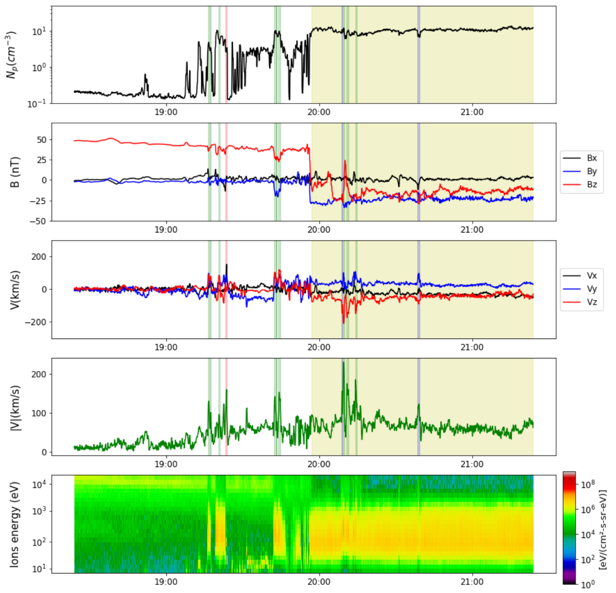

Figure 3. In-situ measurement provided by THEMIS C spacecraft during a magnetopause crossing on July 27, 2009. From the top to the bottom are represented the ion density, the magnetic field components, the velocity components and magnitude and the omnidirectional differential energy fluxes of ions. The yellow shading highlight the magnetosheath. The green intervals indicate the algorithm ground truth. The blue interval is a missed event. The red interval is a false positive.

5 Massive multi-missions statistical analysis of the Earth magnetopause and shape

For the three problems we mentioned in the previous sections, the method we developed resulted in events catalogs that are among the most exhaustive existing ones and that can be used for the statistical study of the events they refer to. For instance, we combine the magnetopause crossings catalog obtained in the section 3 to online accessible crossings from older missions (OGO, Hawkeye, IMP, Geotail, ISEE, Prognoz, Explorer, AMPTE) to provide a statistical study of the location and shape of this boundary [9].

On the one hand, this study confirms long-proven characteristics of the magnetopause: the earthward pushing for an increasing dynamic pressure or for a decreasing Geocentric Solar Magnetospheric (GSM) Z component of the IMF and the azimuthal asymmetry induced by the seasonal variations. On the other hand, we brought answer elements to questions that were still open on this subject. At first, the common expression of the magnetopause radial distance as an elliptic function was found to hold at lunar distances. Second, no particular influence of the IMF radial component B_x was found. Finally, through the evidence of the influence of the IMF clock angle, we showed that a changing IMF B_y could induce changes on the magnetopause shape, which is consistent with the expected effect of this component on the displacement of the reconnection sites on the dayside. These results are condensed into an analytical non-indented magnetopause model [10] for which a typical representation in the GSM (X-Z) plane is shown in the Figure 4 . Parameterized by the IMF components, the solar wind dynamic and magnetic pressure and by the Earth dipole tilt angle, this model offers a more precise description of the magnetopause than any other existing magnetopause model on the night side of the magnetosphere.

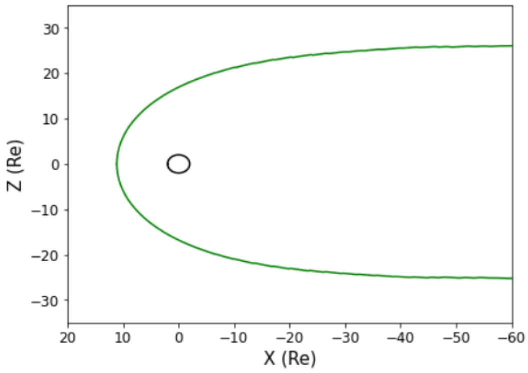

Figure 4. Projection in the GSM (X-Z) plane of the magnetopause model developed in [10]

Finally, we come back on the question of the near-cusp indentation of the magnetopause by fitting the same ensemble of magnetopause crossings to various indented and non-indented analytical models [11]. The results show that a non-indented model over-estimate the radial position of the near-cusp magnetopause thus confirming the actual indentation of the magnetopause. This indentation is then found to be consistent with the one observed in MHD simulations.

6 Acknowledgements

This work has been funded by the French Ministry of Defence. The authors thank Benoit Lavraud, Bayane Michotte de Welle, Alexis Jeandet and Erwan Le Pennec for their collaboration for this work.

References

[1] G. Nguyen, N.Aunai, D.Fontaine, E. Le Pennec, J. Van den Bossche, A. Jeandet, B. Bakkali, L. Vignoli and B. Regaldo "Automatic Detection of Interplanetary Coronal Mass Ejections from In Situ Data: A Deep Learning Approach" *The Astrophysical Journal*, **874**, 2, April 2019, pp. 145–163, doi:10.3847/1538-4357/ab0d24.

[2] R. P. Lepping, C. C. Wu, and D. B. Berdichevsky, "Automatic identification of magnetic clouds and cloud-like regions at 1 AU: occurrence rate and other properties" *Annales Geophysicae*, **23**, 7, October 2005, pp. 2687–2704, doi: 10.5194/angeo-23-2687-2005.

[3] Z. Němeček, J. Šafránková, and J. Šimůnek, "An Examination of the Magnetopause Position and Shape Based Upon New Observations" *Dayside Magnetosphere Interactions*, **8**, 20, February 2020, pp. 135–151, doi: 10.1002/9781119509592.ch8.

[4] O. Kruparova, V. Krupar, and J. Šafránková, "Statistical Survey of the Terrestrial Bow Shock Observed by the Cluster Spacecraft" *Journal of Geophysical Research (Space Physics)*, **124**, 3, March 2019, pp. 1539–1547, doi: 10.1029/2018JA026272.

[5] G. Nguyen, N. Aunai, B. Michotte de Welle, A. Jeandet, B. Lavraud and D. Fontaine, "Massive multi-mission statistical study and analytical modeling of the Earth's magnetopause: 1 - A gradient boosting based automatic detection of near-Earth regions" *Journal of Geophysical Research (Space Physics)* 14, December 2021, doi: 10.1029/2021JA029773.

[6] K. Jelinek, Z. Němeček, and J. Šafránková, "A new approach to magnetopause and bow shock modeling based on automated region identification" *Journal of Geophysical Research (Space Physics)*, **117**, A5, May 2012, doi: 10.1029/2011JA017252

[7] H. Breuillard, R. Dupuis, A. Retino, O. Le Contel, J. Amaya and G. Lapenta, "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*, **7**, 3, September 2020, doi: 10.3389/fspas.2020.00055

[8] G. Paschmann, I. Papamastorakis, W. Baumjohann, N. Sckopke, C. Carlson, B. Sonnerup, H. Luhr "The magnetopause for large magnetic shear: AMPTE/IRM observations" *Journal of Geophysical Research*, **91**, A10, October 1986, 11099–11115 doi: 10.1029/JA091iA10p11099

[9] G. Nguyen, N. Aunai, B. Michotte de Welle, A. Jeandet, B. Lavraud and D. Fontaine, "Massive multi-mission statistical study and analytical modeling of the Earth's magnetopause: 2 - Shape and location" *Journal of Geophysical Research (Space Physics)* 16, December 2021, doi: 10.1029/2021JA029774.

[10] G. Nguyen, N. Aunai, B. Michotte de Welle, A. Jeandet, B. Lavraud and D. Fontaine, "Massive multi-mission statistical study and analytical modeling of the Earth's magnetopause: 3- An asymmetric non indented magnetopause analytical model" *Journal of Geophysical Research (Space Physics)* 16, December 2021, doi: 10.1029/2021JA030112.

[11] G. Nguyen, N. Aunai, B. Michotte de Welle, A. Jeandet, B. Lavraud and D. Fontaine, "Massive multi-mission statistical study and analytical modeling of the Earth's magnetopause: 4- On the near-cusp magnetopause indentation" *Journal of Geophysical Research (Space Physics)* 16, December 2021, doi: 10.1029/2021JA029776.