ARCANE: An Operational Framework for Automatic Realtime ICME Detection in Solar Wind In Situ Data

Hannah T. Rüdisser[a,b], Gautier Nguyen[c], Justin Le Louëdec[a], Christian Möstl[a]

a) Austrian Space Weather Office, GeoSphere Austria, Graz, Austria
b) Institute of Physics, University of Graz, Graz, Austria
c) ONERA/DPHY, Université de Toulouse, F-31055, Toulouse, France

1. Introduction

Interplanetary coronal mass ejections (ICMEs) are among the primary drivers of space weather disturbances Earth, affecting both technological systems and human activities. Automatically detecting these events in solar wind in situ data is critical for early warning systems. Various approaches have been employed to identify these structures in time series data from in situ solar wind observations (e.g., Nguyen et al. 2019; Rüdisser et al. 2022; Pal et al. 2024). However, significant challenges remain in developing robust detection methods, especially in real-time applications.

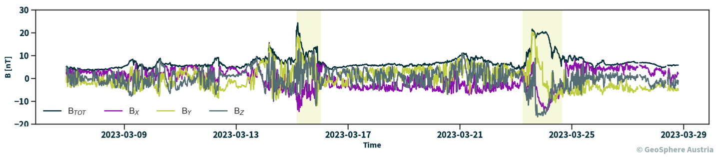

Fig.1: Solar wind magnetic field data from the NOAA/DSCOVR spacecraft at L1 in GSM coordinates. ICMEs are indicated by the shaded regions.

- Development of a **modular framework to evaluate methods** for early detection of ICMEs in realtime solar wind in situ data.
- Assessing models under **realistic operational conditions**. Operational deployment of a first prototype at https://helioforecast.space.
- **Reliable detection of high impact events** in a realtime setting and acceptable performance on low impact events.

2. Data & Methodology

Real-Time Solar Wind (RTSW) data

- Made available by the Space Weather Prediction Center (NOAA SWPC)
- Archived dataset
- Measurements from various spacecraft located at the L1 Lagrange point
- Tracked by the Real-Time Solar Wind Network of ground stations.
- Available from 1998 onward.
- NOAA/DSCOVR satellite serving as the primary operational RTSW spacecraft since July 2016.

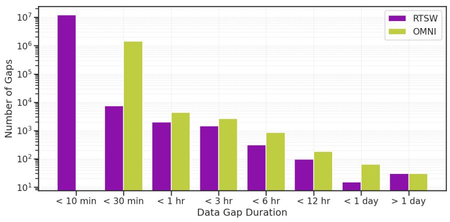

Fig.2 Duration of Data Gaps for both the RTSW and OMNI Dataset.

ICME Catalog

- Aggregated catalog based on several other catalogs (Möstl, Richardson and Cane, Nguyen, ...)
- 784 events between 1998 and 2023

We resample the data to a 30 minute frequency to reduce data gaps and eliminate the uncertainties on event boundaries introduced through the slightly different positions around L1 between spacecraft. We directly train our model on the RTSW dataset to help it learn how to account for the reduced data quality compared to OMNI data.

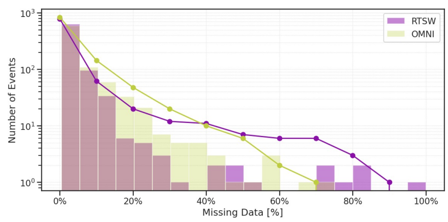

Fig.3 Percentage of missing Data in Events for both the RTSW and OMNI Dataset.

ARCANE (Automatic Realtime deteCtion ANd forEcast) serves as a highly modular and adaptable machine learning framework, created to address the complexities of time series event detection tasks. Its primary goal is to streamline workflows by offering integrated modules and tools for data preprocessing, model training, testing, evaluation and visualization.

The framework is built on Hydra (Yadan, 2019), which provides a **flexible and modular setup**, making it easy to configure and manage experiments. The configurable components are organized into eight main categories: Datasets, Boundaries, Callbacks, Collates, Models, Modules, Samplers, and Schedulers. Each module can be adjusted directly through configuration files, allowing for a quick modification of setups without altering the core code. These modules integrate with available scripts, which handle key tasks such as training, testing, analysis and prediction. The framework also includes routines specifically designed to download and process the RTSW data. While the current version of ARCANE comprises a **ResUNet++ model** as introduced in Rüdisser et al. 2022, the modularity of the framework allows for easy integration of newer models, as the field advances.

Automatic detection methods generally process time series data and generate event lists, which can then be **compared to ground truth catalogues**. Instead of evaluating the entire event after it has passed, we **identify the earliest point at which an event is detected** in a streaming context and compare it to the true event start.

Comparison between two different early detection classifiers:

1. Threshold Classifier:

- $B_{max} \geq 8$ nT
- $\beta \leq 0.3$
- $T_p \leq 4.3 \times 10^4$ K
- Following Lepping et al. 2005

2. ARCANE Classifier:

- Adapted from Rüdisser et al. (2022), where a ResUNet++ was used for time series segmentation
- Modified post processing routines
- Trained and evaluated on RTSW data

3. Results

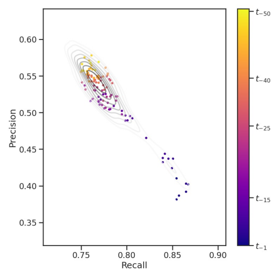

Fig.4: Precison vs. Recall for different waiting times.

We evaluate the model's performance by analyzing Precision and Recall at different waiting times δ. δ is defined as the time that passes before we classify a certain timestep as either event or non event. At a waiting time of 11 hours, we determine Precision, Recall and F1-Score and compare it to the performance of the Threshold Classifier, which are shown in Table 1.

	Precision	Recall	F1-Score
ARCANE Classifier	0.52	0.77	0.63
Threshold Classifier	0.50	0.59	0.54

Tab.1: Results at a waiting time of 11 hours.

Additionally, we define the Delay as the waiting time plus the error the model made on the starting time and show the resulting metrics in Table 2.

	Abs. Mean Delay	Rel. Mean Delay
ARCANE Classifier	4.0 hours	12.8 %
Threshold Classifier	7.3 hours	17.9 %

Tab.2: Results for the Delay analysis.

The Delay parameter aids in determining the model's ability to detect events early. We show the histogram of Delay Percentage in Figure 5.

To analyze which events our model performs worst on, we additionally show the Delay vs the duration of each event in Figure 6.

Finally, we plot the maximum value of the total magnetic field vs. the maximum value of the bulk velocity for true positives (TP), false positives (FP) and false negatives (FN) to estimate the severity of the events in each group.

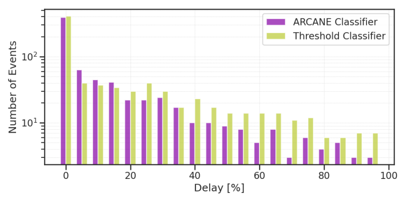

Fig.5 Delay in percentage of the duration shown as histogram.

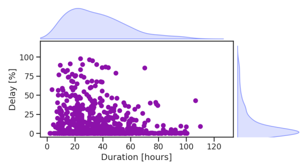

Fig.6: Delay vs duration for each event.

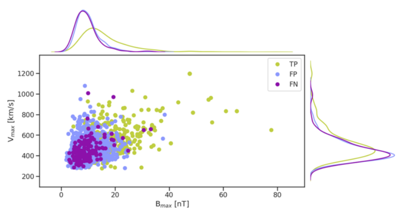

Fig.7.: Maximum value of B vs maximum value of V for each group of events.

4. Conclusion & Outlook

Key Features & Advancements

- **First operational ICME detection on streaming realtime solar wind data using machine learning.**
- **Outperforms traditional threshold-based detection** in both **precision and timeliness**, even on challenging datasets like RTSW.
- **Modular design:** Could easily integrate additional data sources (e.g., **STEREO-A**) or links with arrival time models to enhance detection.
- **Computational efficiency:** Low-cost retraining enables rapid adaptation to improved event catalogs and inference can be run on CPUs in short time.

Challenges

- **Current event catalogs lack severity labels**, limiting differentiation between high- and low-impact ICMEs.
- **Ideal training data would include full solar wind segmentation**, distinguishing **shocks, sheaths, flux ropes, HSSs and SIRs.**

Future Directions

- **Improved catalogs** or simultaneous **prediction of key ICME parameters** (e.g., min **Bz**, duration, velocity) to infer severity.
- **Integrating physical models** for real-time analysis of detected ICMEs.
- **Ensemble modelling approach** to improve detection reliability, robustness, and explainability.
- Include arrival time models to further improve performance.
- Integrate data from multiple sources.
- Apply the tool throughout the Heliosphere.

References

- Nguyen, G., Aunai, N., Fontaine, D., Pennec, E. L., Bossche, J. V. D., Jeandet, A., Blancard, B. R.-S. (2019). Automatic Detection of Interplanetary Coronal Mass Ejections from In Situ Data: A Deep Learning Approach. The Astrophysical Journal , 874 (2), doi: 10.3847/1538-4357/ab0d24178
- Rüdisser, H. T., Windisch, A., Amerstorfer, U. V., Möstl, C., Amerstorfer, T., Bailey, R. L., & Reiss, M. A. (2022). Automatic Detection of Interplanetary Coronal Mass Ejections in Solar Wind In Situ Data. Space Weather, 20(10), e2022SW003149, doi: 10.1029/2022SW003149
- Pal, S., Santos, L. F. G. d., Weiss, A. J., Narock, T., Narock, A., Nieves-Chinchilla, T., Good, S. W. (2024). Automatic Detection of Large-scale Flux Ropes and Their Geoeffectiveness with a Machine-learning Approach. The Astrophysical Journal , 972 (1), 94. doi: 10.3847/1538-4357/ad54c3184
- Yadan, O. (2019). Hydra - a framework for elegantly configuring complex applications. Github. https://github.com/facebookresearch/hydra
- Lepping, R. P., Wu, C.-C., & Berdichevsky, D. B. (2005). Automatic identification of magnetic clouds and cloud-like regions at 1 au: occurrence rate and other properties. Annales Geophysicae, 23 (7), 2687–2704. doi:10.5194/angeo-23-2687-2005653

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) P 36093.
Funded by the European Union (ERC, HELIO4CAST, 101042188). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Contact: Hannah T. Rüdisser, Austrian Space Weather Office, GeoSphere Austria | hannah.ruedisser@geosphere.at | www.geosphere.at