Automatic detection of large-scale flux ropes and their geoeffectiveness with a machine learning approach

Sanchita Pal,[1] Luiz F. G. dos Santos,[2] Andreas J. Weiss,[1] Thomas Narock,[3] Ayris Narock,[4,5]
Teresa Nieves-Chinchilla,[5] Lan K. Jian,[5] and Simon W. Good[6]

[1] NASA Postdoctoral Program Fellow, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[2] SETI Institute, Mountain View, CA 94043, USA
[3] Center for Natural, Computer, and Data Sciences, Goucher College, Baltimore, MD, United States
[4] ADNET Systems Inc., 7515 Mission Drive, Lanham, MD 20706, USA
[5] Heliophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[6] Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland

ABSTRACT

Detecting large-scale flux ropes (FRs) embedded in interplanetary coronal mass ejections (ICMEs) and assessing their geoeffectiveness are essential since they can drive severe space weather. At 1 au, these FRs have an average duration of 1 day. Their most common magnetic features are large, smoothly rotating magnetic fields. Their manual detection has become a relatively common practice over decades, although visual detection can be time-consuming and subject to observer bias. Our study proposes a pipeline that utilizes two supervised binary-classification machine learning (ML) models trained with solar wind magnetic properties to automatically detect large-scale FRs and additionally determine their geoeffectiveness. The first model is used to generate a list of auto-detected FRs. Using the properties of southward magnetic field the second model determines the geoeffectiveness of FRs. Our method identifies 88.6% and 80% large-scale ICMEs (duration $\geq$ 1 day) observed at 1 au by Wind and Sun Earth Connection Coronal and Heliospheric Investigation (STEREO) mission, respectively. While testing with a continuous solar wind data obtained from Wind, our pipeline detected 56 of the 64 large-scale ICMEs during 2008 - 2014 period (recall= 0.875) but many false positives (precision= 0.56) as we do not take into account any additional solar wind properties than the magnetic properties. We found an accuracy of 0.88 when estimating the geoeffectiveness of the auto-detected FRs using our method. Thus, in space weather now-casting and forecasting at L1 or any planetary missions, our pipeline can be utilized to offer a first-order detection of large-scale FRs and geoeffectiveness.

Keywords: Sun: coronal mass ejection, flux rope, space weather, geomagnetic storm, magnetic field

1. INTRODUCTION

Magnetic flux ropes (FRs) are fundamental solar and heliospheric structures that are formed by twisted bundles of magnetic field lines. They can originate at the Sun or in the heliosphere. They are commonly associated with coronal mass ejections (CMEs; Crooker et al. 1997; Webb & Howard 2012), streamer blow-outs (SBOs; Vourlidas & Webb 2018; Nitta et al. 2021), small transients (STs; Kilpua et al. 2009; Yu et al. 2014, 2016) in solar wind, local small-scale magnetic flux rope structures near the heliospheric current sheet (HCS) (Moldwin et al. 1995, 2000) and magnetospheric flux transfer events (FTEs; Sun et al. 2020).

The large-scale FRs of interplanetary counterparts of CMEs (ICMEs) may have a significant contribution of flux from coronal reconnection (Qiu et al. 2007; Gopalswamy et al. 2017; Pal et al. 2017, 2018; Pal 2022) and thus may transport large quantities of solar mass, magnetic flux and helicity away from the Sun. Frequently these FRs are called magnetic clouds (MCs; Burlaga et al. 1981) which have an enhanced magnetic field strength, a smooth and coherent rotation in magnetic field lines (e.g. Burlaga et al. 1981; Klein & Burlaga 1982), declining velocity (e.g. Klein & Burlaga 1982; Russell & Shinde 2003), low proton temperature and low proton β (the ratio between proton thermal pressure and magnetic pressure) characteristics (e.g. Richardson & Cane 1995). At 1 au, ICME FRs have an average radial dimension of $\sim$0.25 au (Klein & Burlaga 1982). A duration of approximately 1 day, corresponds to a

characteristics dimension ∼0.25 au (Klein & Burlaga 1982; Kahler & Webb 2007; Lugaz et al. 2009). However, studies like Richardson & Cane (2004) show that there may be a lack of MCs in ICMEs observed at different heliocentric distances and different phases of the solar cycle. It may be possible that all ICMEs contain a well-defined flux rope close to their origin, but as ICMEs evolve on their way out from the Sun, their flux rope signatures may weaken (Crooker et al. 1997). Alternatively, ICMEs may always contain identifiable FRs but, when they are traversed far from their central FR axis, their FRs cannot be detected (e.g. Jian et al. 2006). Jian et al. (2006) and Nieves-Chinchilla et al. (2018) adopted a less restrictive term called magnetic obstacles (MOs) to address the magnetic structure inside ICMEs. The MOs are described as periods inside ICMEs having low proton β where in some cases, the magnetic field lines display smooth and monotonic rotations.

Apart from ICME FRs, small transients (STs) that may last from a few tens of minutes to a few hours are one of the important sources of FRs in the heliosphere. They are components of slow solar wind and may originate from the cusps of coronal streamers (Sheeley et al. 1997, 2008; Wang et al. 2000). Yu et al. (2014, 2016) have explained the ST characteristics in details. They differ from ICME FRs in a number of ways, including their duration (0.5–12 hours), the absence of a speed differential with respect to the slow background solar wind, and the lack of a discernible amplification of the heavy ion charge state within them. Some of the STs may have the geometry of small-scale flux ropes (SFRs) that are local structures mostly observed near HCS and are generated due to magnetic reconnection and a turbulence cascade (Cartwright & Moldwin 2008, 2010). Although SFRs show the typical magnetic configuration of FRs (i.e., smooth rotation) they may have proton β less than or greater than 1.

Decades of studies have been conducted to explore how solar magnetic fields affect the geospace environment through the solar wind and solar storms (Nandy et al. 2023). The major solar storms including ICMEs may shape the near-Earth heliospheric environment in a way that may trigger the most extreme forms of space weather at Earth. Even a moderate type geomagnetic storm is capable to adversely affect assets in geospace environment (Baruah et al. 2024). It is challenging to predict their impact on the Earth as they may continuously evolve in the heliosphere while propagating (Manchester et al. 2017; Luhmann et al. 2020; Pal et al. 2022, 2023). Their geo-effectiveness and capability to trigger magnetic storms have societal impacts, which cannot be disregarded in modern society. With high proton number density, increased proton speed and magnetic field component B_z anti-parallel to the geomagnetic field, ICMEs impact the magnetosphere and also allow solar energetic particles to drift in the magnetosphere. A simplified explanation of ICME - magnetosphere interaction (Dungey 1961) states that if an ICME B_z is directed opposite to the Earth's field, magnetic reconnection takes place at the day-side magnetosphere, and reconnected magnetic flux accumulates in the night side magnetotail region. Thus, the magnetic reconnection at the magnetotail leads to solar plasma injection toward the Earth on the night side. This results in the formation of aurora and a ring current around the Earth, which causes a reduction in the geomagnetic field measured by the disturbance storm time (Dst) index.

The ICME speed V and its southward magnetic field component $-B_z$ are of the most importance in creating geomagnetic storms (Fairfield & Cahill 1966; Cane et al. 2000; Wu & Lepping 2002). Certainly, a sufficiently long duration (δt) of $-B_z$ is also necessary (Wang et al. 2003). The product of the average value of $\overline{-VB_z}$ and δt defines the magnetic flux transferred from the interplanetary medium into the magnetosphere. Wang et al. (2003) performed a statistical study of the relationship between $\overline{-VB_z}$, δt and Dst_{min} that revealed the importance of $\overline{-VB_z}$ and its duration δt for generating geomagnetic storms. The study found an empirical formula $Dst_{min} = -19.01 - 8.43(\overline{-VB_z})^{1.09}(\delta t)^{0.3}$ that indicates a more prominent contribution of large $\overline{-VB_z}$ over δt to create a strong geomagnetic storm. They found a high anti-correlation with correlation coefficient (CC) of -0.95 between $(\overline{-VB_z})^{1.09}(\delta t)^{0.3}$ and Dst_{min}, whereas, a comparatively weak linear correlation with CC of -0.7 was found between $(\overline{-VB_z})(\delta t)$ and Dst_{min}. This suggests a non-linear relation between Dst_{min} and δt, whereas Dst_{min} and $\overline{-VB_z}$ have an almost linear relationship.

Visual inspection of the solar wind magnetic field and plasma environment via spacecraft data to identify ICMEs is widely practiced though it is quite time-consuming, ambiguous, subjective, and biased by the observer's interpretations (Shinde & Russell 2003). Some studies have begun to introduce the use of supervised machine learning (ML) algorithms that have the potential to learn the properties of labeled events by themselves. Their use in space physics is progressing since they can tackle a large and ever-growing database accumulated over decades. In solar physics, ML-based techniques have recently bloomed as a very attractive way of using our computing resources to extract conclusions from solar data (Asensio Ramos et al. 2023).

Nguyen et al. (2019) proposed a pipeline that used a Convolutional Neural Network (CNN) trained with 33 solar wind magnetic and plasma parameters obtained and derived from in situ spacecraft to automatically detect the start

and end times of ICMEs in in-situ data. Their method found 84% of the listed ICMEs collected from different existing catalogs where ICMEs are manually detected using Wind spacecraft data during 2010–2015. Out of 158 ICMEs found during the period using their pipeline, 25 ($\sim$16%) events were false positives. dos Santos et al. (2020) implemented a tool based on handwriting recognition models that uses a deep CNN to classify subsets of the ICMEs introduced in Nieves-Chinchilla et al. (2019). The weights of their model were trained with synthetic data obtained from a well-established physical flux rope model. They simulated the noise of synthetic training data using data points belonging to a normal distribution with standard deviations including 0.05 and 0.1. Their methodology classified 76% of 321 real cases (identified manually by Nieves-Chinchilla et al. (2019) using Wind spacecraft data) correctly during 1995–2015. Narock et al. (2022) attempted to predict the orientation of heliospheric magnetic flux ropes once they are identified in in-situ solar wind data using a model based on CNN architecture. Rüdisser et al. (2022) converted the ICME detection problem to a time series segmentation problem and proposed a pipeline for automatic detection of ICMEs using a variation of an advanced U-Net (Ronneberger et al. 2015) method that has recently proven successful in medical image segmentation. Using their method, 73% of the cataloged events observed by the Wind spacecraft during 1997-2015 were detected correctly. They found a total number of 254 ($\sim$40%) events as false positives during that period. A comparable method was proposed by Chen et al. (2022) that uses the residual U-net architecture and automatically detected 77% of ICMEs during the period 1997 October 1 to 2016 January 1 collected by the Wind spacecraft. Based on Grad-Shafranov reconstruction technique, several studies have developed automated SFR detection algorithms (Hu et al. 2018), and applied them to the spacecraft data at different heliocentric distances (Chen et al. 2019; Chen & Hu 2020, 2022). Using ML-based classification algorithm Farooki et al. (2024) compared SFRs, MCs, and ambient solar wind plasma properties.

A recent study by Nair et al. (2023) summarises the results obtained from an open data-science challenge called "MagNet: Model the Geomagnetic Field" to Predict Dst from solar wind Data. The challenge attracted 1,197 model submissions that used various ML approaches and predicted Dst in a real-time modeling environment. Hu et al. (2023) developed a 6-hour ahead Dst prediction model from solar wind observations using Gated Recurrent Unit networks. They used several solar wind parameters like proton density, velocity, magnetic field intensity, magnetic field vector in the z-direction, clock, and dipole tilt angle of the magnetic field, and Dst index and boosted the performance of the model by developing a multi-fidelity method. Xu et al. (2020) used an ensemble-learning algorithm that combines the artificial neural network, support vector regression, and long short-term memory network to predict the Dst index 1–6 hr in advance. They used solar wind's total magnetic field, B_z, total electric field, solar wind speed, plasma temperature, and proton pressure as inputs. Chandorkar et al. (2017) provided a method to generate a probabilistic prediction of Dst using a Gaussian method approach. Gruet et al. (2018) constructed the Dst index prediction model by combining the long short-term memory (LSTM) network with the Gaussian process approach. To forecast the Dst index during 80 intense geomagnetic storms ($Dst_{min} \leq 100$ nT) during 1995 - 2014, Lu et al. (2016) applied a support vector machine (SVM) and combined that with a distance correlation technique. They utilized 13 observed and derived plasma and magnetic parameters of the solar wind as inputs. Pricopi et al. (2022) developed binary classification models to predict the geoeffectiveness of CMEs, i.e. whether a CME is associated with a geomagnetic storm having $Dst_{min} \leq 30$ nT or not. They used supervised ML models like logistic regression, k-nearest neighbors, SVM, feed-forward artificial neural networks, and ensemble models trained on white-light coronagraph data sets of CMEs close to the Sun. They achieved an adequate hit rate with these models. Choi et al. (2012) found a better performance of SVM than the performance of simple classification based on constant classifiers, when they used SVM for the prediction of geoeffective CMEs. The SVM algorithm is proven to be efficient in a wide range of solar and space physics research, such as flare classification (e.g. Qahwaji & Colak 2007; Sinha et al. 2022), ICME arrival time prediction (e.g. Liu et al. 2018), Dst index prediction (e.g. Lu et al. 2016), and sunspot occurrence prediction (e.g. Rodríguez et al. 2022).

Previous studies reveal an impressive performance of ML-based methods in the field of automatic ICME detection and their geoeffectiveness prediction. Integrating automatic detection of both space weather driving events and their geoeffectiveness in a single pipeline is an efficient way for quickly assessing the real-time space weather and issuing alert. In this study, we present a pipeline that automatically detects large-scale ICME associated FRs from continuous spacecraft measurements at 1 au and additionally indicates their geoeffectiveness – whether they are minor to intensely geoeffective ($Dst_{min} \leq -50$ nT) or not. For the FR detection process, the pipeline uses CNN architecture trained with more realistic synthetic FR and non-flux rope (NFR) magnetic fields prepared using a physics-based FR model combined with power spectral densities of magnetic field fluctuations in FRs and NFRs derived from spacecraft data.

For indicating the geoeffectiveness of the automatically detected FRs, our pipeline uses another supervised binary classifier trained with geoeffective magnetic parameters of real events.

In Section 2, we present our proposed pipeline and describe our methodology of pre-processing the dataset, training the models, and post-processing techniques. Section 3 discusses the results and Section 4 summarises and concludes the article.

2. DESCRIPTION OF PIPELINE

We combine a supervised ML model 'M1' with our current understanding of ICME FR and non flux rope (NFR) solar wind segments and integrate this to another supervised ML model 'M2' to first detect the ICME FR intervals from solar wind data and then their geoeffectiveness. This work utilizes only the magnetic properties rather than other solar wind properties like composition and plasma. This is because in the future, we plan to make this pipeline usable in classification of real-time solar wind being observed by various planetary missions. Unlike magnetic field data, plasma and composition data are generally not available in real-time with reliable quality from planetary missions.

In the pipeline, the first model 'M1' is an image classification model that uses CNN architecture. It is trained on synthetic large-scale FRs and NFRs prepared by combining a well-established physical flux rope model and a sophisticated additive background solar wind profile model based on the power spectrum density (PSD) of solar wind magnetic field corresponding to real ICME FRs and NFRs. We refer to the model as $DeepFRi$ – **Deep** convolutional neural network **FR i**dentification)". The second model that detects the geoeffectiveness of FRs is named as $DicFR$ – Disturbance storm time (**D**st) **i**ndex **c**lassification of identified **FR**s. Note that the DicFR model aims to classify auto-identified FRs based on their geoeffectiveness, i.e. whether they are moderate to intensely geoeffective or not. It doesn't intend to predict the Dst index value (e.g. Hu et al. 2022). In the upper panel of Figure 1, using a flowchart we show the operations that we perform in this work to auto-identify FRs and their geoeffectiveness from solar wind data. As a final result, the pipeline generates a list that contains auto-identified FR start and end times, and their geoeffectiveness categories. Below we describe the pre-processing of data, the models, and the post-processing steps including boundary determination of auto-identified FRs and listing them with their predicted geoeffectiveness categories.

2.1. *Data Pre-processing*

As our first model DeepFRi is intended to classify solar wind into FRs and NFRs based on their magnetic characteristics, we use only processed and calibrated Magnetic Field Investigation (MFI Lepping et al. 1995) instrument data at 1 min resolution during 1995 - 2021 from the Wind spacecraft. The Wind/MFI instrument-measured solar wind magnetic field data in Geocentric Solar Ecliptic (GSE) coordinate system $(x_{gse}, y_{gse}, z_{gse})$. We first window the MFI data in 24-hour intervals having 24×60 data points and check for the data gaps. If there are data gaps within the window and the gaps cover no more than 10% of the window size the window is still considered as an acceptable input. The size of the window is chosen to be 24 hours because we focus here only to detect ICME-associated large-scale FRs and in general, at 1 au, the average ICME FRs are of $\sim$24-hour duration (Klein & Burlaga 1982; Bothmer & Schwenn 1998). To verify this, we consider 695 manually identified ICME FRs obtained from Wind ICME catalog (Nieves-Chinchilla et al. 2018) and STEREO ICME catalog (Jian et al. 2018) which are observed by both Wind and the Solar TErrestrial RElations Observatory (STEREO; Kaiser et al. 2008) spacecraft, respectively at 1 au and find their mean duration as 24.3 hours. Next, we sample each window into *npoint*, where each point corresponds to $\sim$ 5-min average of solar wind magnetic field data. After down-sampling the window, we normalize each magnetic field component with the solar wind magnetic field intensity in that window. Next, the normalized magnetic field components of the down-sampled window are plotted in three different planes $x_{gse} - y_{gse}$, $y_{gse} - z_{gse}$ and $z_{gse} - x_{gse}$, where each of them is in the -1 to 1 range and of pixel size (Width) 54 $\times$ (Height) 54. These are called magnetic field hodograms. After generating three different hodogram images, we merge them into a single image of pixel size (Width) 162 $\times$ (Height) 54. Thus, each window corresponds to a merged hodogram acceptable for our M1 model input.

If we consider a time interval (T_{dur}) starting at t_s hours and ending at t_e hours, we divide the interval in n number of windows – $W_0, W_1, ... W_{n-1}$, where n is an integer with a unit of hours. Thus, T_{dur} (in hours) can be expressed as,
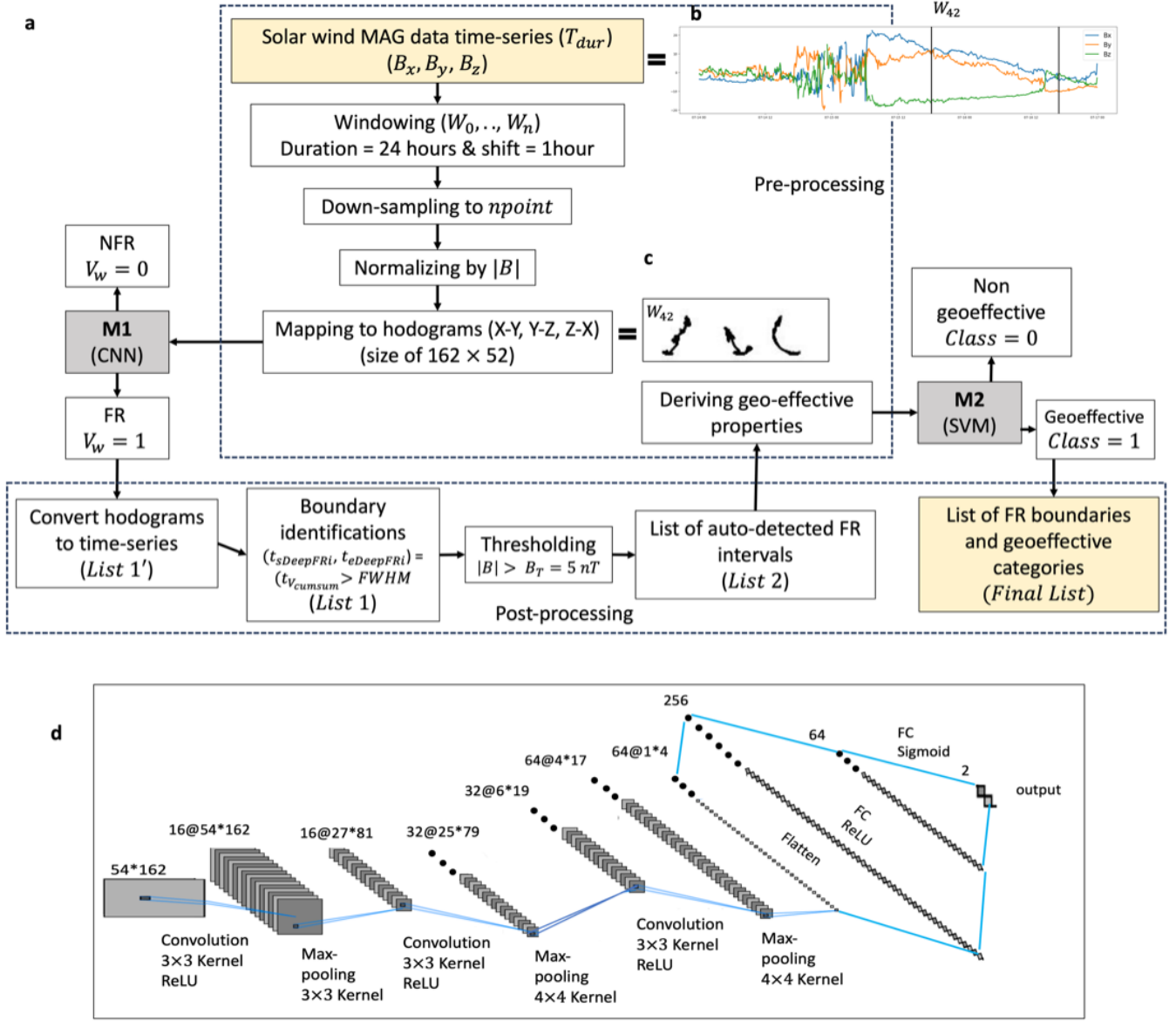
Figure 1. Upper panel- (a) A flowchart of our pipeline that shows the processes followed to convert raw data from the Wind spacecraft to a list of auto-detected FRs and their geoeffectiveness categories. The input and the final output of the flowchart are shown in yellow boxes. The models used are shown in grey boxes. The dashed boxes enclose the pre-processing and post-processing steps. As an example, (b) a 3-day interval Wind magnetic field vector data is shown on the right side where the intervals between black and green vertical lines indicate two consecutive windows (W_{42} and W_{43}) of 24 hours duration, respectively. The corresponding merged hodogram is shown at (c). Lower panel- schematic of M1 model architecture. The gray squares represent output of each layers. (d) Schematic of the CNN architecture used for M1 model. The squares represent the output of each layer. The architecture consists of three sets of convolutional and max-pooling layers using ReLU activation, followed by two fully connected layers with ReLU and sigmoid as the activation functions, respectively.

$$T_{dur} = \sum_{n=0}^{(t_e - t_s) + 24h} W_n,$$

$$W_n = \int_{t_s - 24h + n}^{t_s + n} dt.$$

(1)

Here, dt is in hour.

As an example, in Figure 1, we show a magnetic vector time series (T_{dur}). Inside T_{dur}, two windows – W_n and W_{n+1} (n = 42) of 24 hours duration are the intervals between the pairs of black and green vertical lines. A merged hodogram corresponding to W_{42} is also shown which can be directly used as input to M1.

For the second model M2 that is intended to classify the geoeffectiveness, we extract the geoeffective magnetic properties from Wind/MFI time series of the auto-detected FR intervals in Geocentric Solar Magnetospheric system (GSM). The model is trained using FR geoeffective properties and the minimum value of Dst index during the FR intervals obtained from OMNIWeb database. The geoeffective magnetic characteristics that we consider for this case are the minimum value of the magnetic field component in the $-z_{\mathrm{gsm}}$ direction ($-B_{z_{\mathrm{gsm}},min}$) and the duration of $-B_{z\mathrm{gsm}}$ with respect to the whole $B_{z\mathrm{gsm}}$ duration in a window. The second feature is named R_{Bz}. To obtain this, we consider only the hodogram in the $y_{\mathrm{gsm}} - z_{\mathrm{gsm}}$ plane and find the ratio of the number of points in a plane formed by $y_{\mathrm{gsm}} - (-z_{\mathrm{gsm}})$ and the total number of points in the whole $y_{\mathrm{gsm}} - z_{\mathrm{gsm}}$ plane. Thus, once we obtain an automated catalog of FRs from model M1, we process each identified FR to obtain $-B_{z_{\mathrm{gsm}},min}$ and R_{Bz} and provide them as input to the M2 model that is trained to classify FRs as geoeffective or non-geoeffective.

2.2. Models

In this Section, we describe model architectures, the method of training the models and their evaluations.

2.2.1. Model architectures

The model M1 uses a binary classification CNN architecture that is based on handwriting recognition models. A CNN is a deep learning architecture consisting of interconnected nodes or neurons having learnable filters constructing layers, whose objectives are to detect features and patterns present in a data set. Thus, a certain number of layers process information in a way similar to the human brain (LeCun et al. 2015; Goodfellow et al. 2016). In each layer, filters are convolved along the input data, and using a nonlinear activation function the layer indicates the presence (or absence) of the requested pattern. A complete description of the principles of CNN can be found in Géron (2017).

The architecture of M1 shown in the lower panel of Figure 1 is quite similar to that of the model used by dos Santos et al. (2020) which implemented a binary class handwritten digit-recognition model (Cireşan et al. 2011). The input to our model are merged hodograms of size 162×54 and the output is a two-element vector that describes the probability of the hodogram being an FR or NFR. We use three convolution layers with kernel size 3 × 3, two of which are followed by max-pooling layers with kernel size 2 × 2. The convolution layer expands the dimension, whereas the max-pooling layers contract them. Finally, the array is flattened to 256 nodes and the network completes with two fully connected layers. These last layers take the dimensions from 256 inputs to 64 outputs, then 64 inputs to 2 outputs. All the convolution layers and the first fully connected layer use a Rectified Linear Units (ReLU; Nair & Hinton 2010)) activation whereas the final layer uses a sigmoid activation function. The model and training are implemented with Tensorflow (Abadi et al. 2016) version 2.10.0 in a Python 3.10.8 environment. Once the model is constructed, it is trained using realistic simulated FR and NFR hodograms. The preparation of the synthetic hodograms is described in the next section.

The type of geoeffectiveness of a given FR is detected by M2. To classify the FR geoeffectiveness we consider the Dst index, which is an index of magnetic activity derived from a network of near-equatorial geomagnetic observatories that measures the decrease in the horizontal component of the Earth's field due to the ring current (Daglis 2001). We choose the Dst index threshold as -50 nT to include moderate to intense storms (Gonzalez et al. 1994), in which the contribution of ICME FRs is prominent. We classify the events having $Dst_{min} < -50$ nT and $Dst_{min} > -50$ nT as '1' and '0', respectively. The M2 is a supervised binary-classifier machine learning model. We choose support vector machine (SVM) classifier – a widely applied ML approach in solar and heliospheric physics and was first proposed by Cortes & Vapnik (1995). A general architecture of SVM model can be found in Álvarez-Alvarado et al. (2021). SVMs can identify both linear and nonlinear relationships between features by using linear and nonlinear kernels that can map these features into a higher dimensional space and thus SVMs may potentially explore previously unknown properties in the data. From previous studies (e.g., Wang et al. (2003); Kane (2005)) it is well known that the duration of southward solar wind magnetic field component has a nonlinear relationship with the Dst_{min}, whereas the product of solar wind velocity and the minimum value of southward magnetic field intensity has a linear relationship with the Dst_{min} (Gonzalez & Echer 2005; Gopalswamy 2008). Here, we use both linear and nonlinear kernels in M2.

2.2.2. Preparation of training dataset

As the number of observed and cataloged events is not enough to train a neural network, in this work we generate several synthetic ICME FRs ($\mathbf{B_{FR}}(t)$) and NFRs ($\mathbf{B_{NFR}}(t)$) to train the model M1. The $\mathbf{B_{FR}}(t)$ is generated combining a physics-based FR model – namely the circular-cylindrical (CC; Nieves-Chinchilla et al. 2016) model output $\mathbf{B_m}(t)$, with additive background noise $\delta \mathbf{B_{FR}}(t)$. Here, $\mathbf{B_m}(t)$ consists of the time series of each magnetic field component – $B_{mx}(t), B_{my}(t), B_{mz}(t)$ obtained in the GSE coordinate system assuming several different spacecraft trajectories through the modeled CC flux ropes. The CC model provides a conventional smooth rotation in the FR magnetic field, while the additive noise $\delta B_{FR}(t)$ provides the type of fluctuations that are observed in real FRs. Our generation process for synthetic FR data is inspired by dos Santos et al. (2020) who used a modeled FR with additive Gaussian noise set called "5% noise" and "10% noise", where the noise values were drawn from normal distributions having standard deviations of 0.05 and 0.1, respectively. The study did not implement physics-based fluctuations to their synthetic data. In our study, the background noise is modeled using the PSD of solar wind magnetic field fluctuations at inertial-range frequencies ($10^{-3} - 10^{-2}$ Hz). This frequency range corresponds to the spatial scales of magnetohydrodynamic turbulence. Typically, in this frequency range, magnetic field fluctuation power is higher in the solar wind than within FRs. This trend can be seen, for example, in the upstream wind and ICME FR of an event observed by Parker Solar Probe (PSP; Fox et al. 2016) at 0.25 au (Good et al. 2020). In Figure 2a, we show the magnetic field components along with the magnitude of an ICME FR and its upstream and downstream solar wind conditions observed by Wind at 1 au. In Figure 2b, we show the wavelet PSD of the magnetic field fluctuations sampled at inertial range frequencies. The vertical lines show the boundaries of the FR.

We use 399 ICME events from the Wind ICME catalog during 1995–2021 to obtain the PSD of real ICME FRs and NFRs required to generate synthetic events. This catalog refers to the ICME FRs as magnetic obstacles (MOs) and notes their start and end boundaries as 'MO start time' and 'MO/ICME end time', respectively. The MOs having a sheath (a pile-up of solar wind commonly observed ahead of ICME FRs) are listed such that the sheath start time is listed as the 'ICME start time'. The catalog also lists the minimum Dst value Dst_{min} during each event and mean magnetic field intensity. We decided not to merge other catalog events to the *Wind ICME catalog* because the criteria to select event boundaries are not exactly similar for all available catalogs.

GENERATION OF SYNTHETIC FLUX ROPES

Synthetic ICME FRs are generated by considering the local configuration of FRs as cylindrical structures with circular cross-sections. The model does not impose any condition on the internal force distribution but makes assumptions about the axial and poloidal current density components. In this work, we keep the ratio between the poloidal and axial current density components as $C1 = 1$, which implies that the structure can behave as force-free structure (see Nieves-Chinchilla et al. (2016) for more details). We generate several different unique FR magnetic field time series by varying the model inputs –

- θ (the angle between the FR axis and ecliptic plane, $-80° \leq \theta \leq 80°$) and ϕ (an angle formed by FR axis and $+x_{gse}$, $5° \leq \phi \leq 355°$) are chosen from equi-distributed points on surface of a sphere with polar angle θ and azimuthal angle ϕ (Deserno 2004).

- Y_0 (the perpendicular distance between the spacecraft trajectory and FR center normalized to the FR radius) from 0 to 0.75 from either side of the FR center in steps of 0.05, and

- H (handedness of the FR) either -1 or 1.

This results in 161,000 simulated FR magnetic field vectors ($\mathbf{B_m}(t)$) normalized to the magnetic field intensity.

To obtain the additive background profile $\delta B_{FR}(t)$ for generating realistic simulated FRs, we followed Weiss et al. (2021, 2022) that generated random and unique noise-field time series with PSD $P_{fr}(k)$ similar to the PSD of real FR magnetic fields. This method is inspired by an analogous process used to generate primordial perturbations for dark matter N-body simulations (Hahn & Abel 2011). The $\delta \mathbf{B_{FR}}(t)$ is expressed as

$$\delta \mathbf{B_{FR}}(t) = \mathcal{F}^{-1}(\sqrt{P_{\mathrm{fr}}(k)}u(k)), \tag{2}$$

where $u(t)$ is a random Gaussian sequence with mean 0 and standard deviation 1 and its Fourier transform $u(k) = \mathcal{F}(u(t))$. As Gaussian noise has a flat power spectrum, multiplying it in Fourier space with $\sqrt{(P_{fr}(k))}$ and transforming back to real space (time domain) provides a noisy sequence $\delta \mathbf{B_{FR}}(t)$ with PSD the same as $P_{fr}(k)$. We compute $P_{fr}(k)$ for each magnetic field component (normalized to magnetic field intensity) by applying Welch's algorithm (Welch 1967)
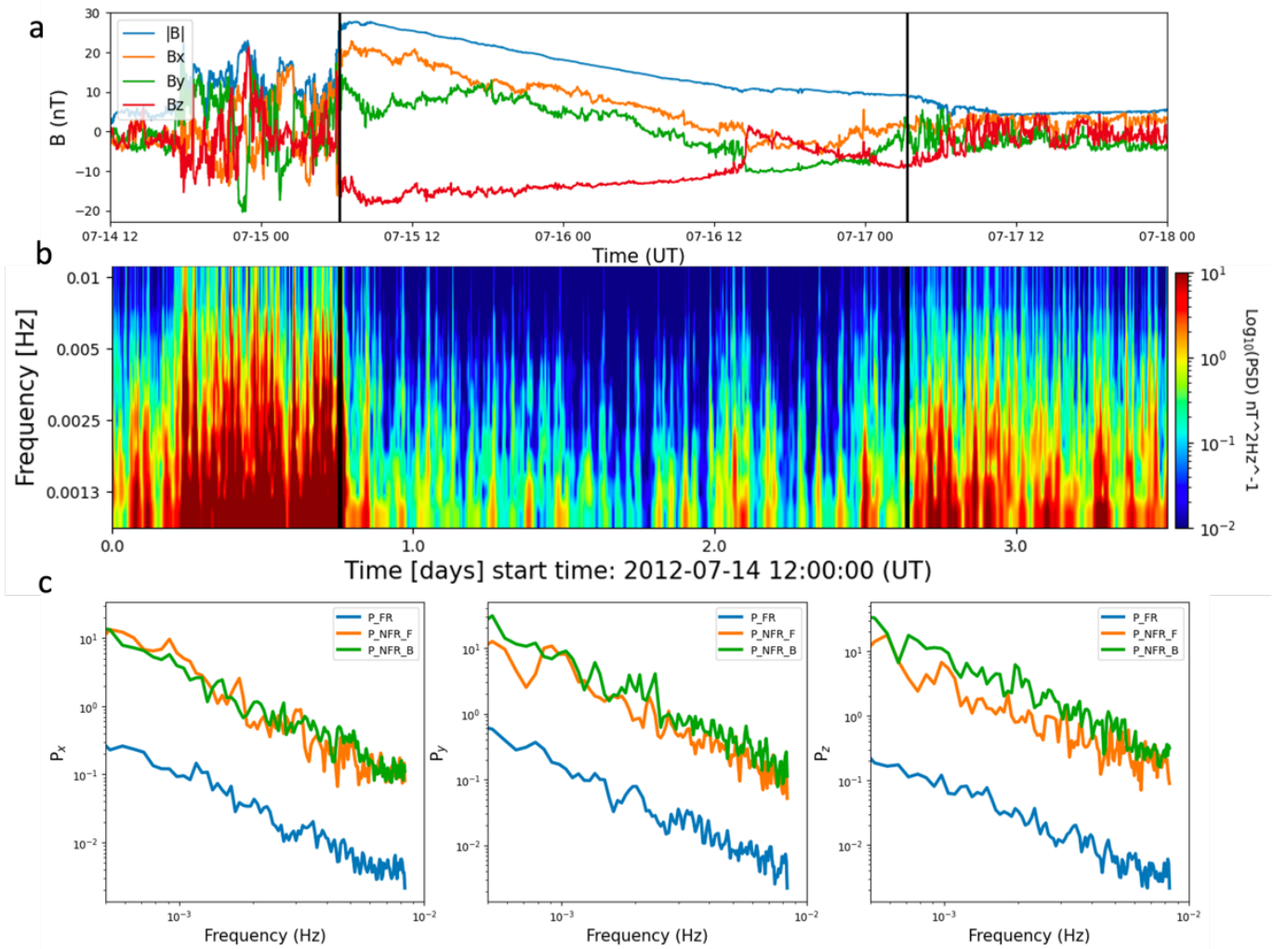
Figure 2. (a) Magnetic vector components (B_x, B_y, B_z) in GSE and magnetic field intensity ($|B|$) of solar wind containing an identifiable ICME FR bounded by vertical lines. (b) The wavelet PSD of **B** field fluctuations sampled at inertial range frequencies. (c) P_x, P_y, P_z are the PSDs of B_x, B_y, B_z field line fluctuations, respectively, where the field lines are normalized by maximum magnetic field intensity and belong to the ICME FR (blue), and non-FR segments (green and orange) before and after the ICME in 24-hours interval, respectively.

on the real event intervals. By averaging several periodograms of overlapping segments that are shorter than the total length of the data, this approach averages out the noise in the data. The real event intervals are gathered from the *Wind ICME catalog*. We use the 'Welch' function available in the 'signal' library of the 'SciPy' package in Python, where $nperseg = 2^8$ is considered for the length of each overlapping segment at which the modified periodogram is computed. Before applying Welch's method, we detrended the normalized magnetic components with three different detrending time intervals ΔT – [10, 30, 60] minutes. The detrending of signals helps remove any kind of pattern at the detrending timescale in the data. In Figure 2c, we show $P_{fr}(k)$ as a function of spacecraft-frame frequency for the ICME FR shown in 2a, with the PSD for each of the three FR magnetic field components in blue. The ΔT used to create $P_{fr}(k)$ in this case is 30 min. Therefore, the y-axis ranges from $\sim 5 \times 10^{-4} - 10^{-2}$. The inertial-range spectral slope in the FR is found as ~ -1.6. The spectral index in the inertial range found in this case is consistent with various theories of Alfvénic MHD turbulence (Schekochihin 2022).

Using the *Wind ICME catalog*, we generate 399 PSDs of real FR events, randomly pick a PSD to generate a unique $\delta B_{FR}(t)$ and add it to each component of $\mathbf{B_m}(t)$ to finally obtain a synthetic FR magnetic field $\mathbf{B_{FR}}(t)$, where

$$\mathbf{B_{FR}}(t) = \mathbf{B_m}(t) + \delta\mathbf{B_{FR}}(t) \tag{3}$$
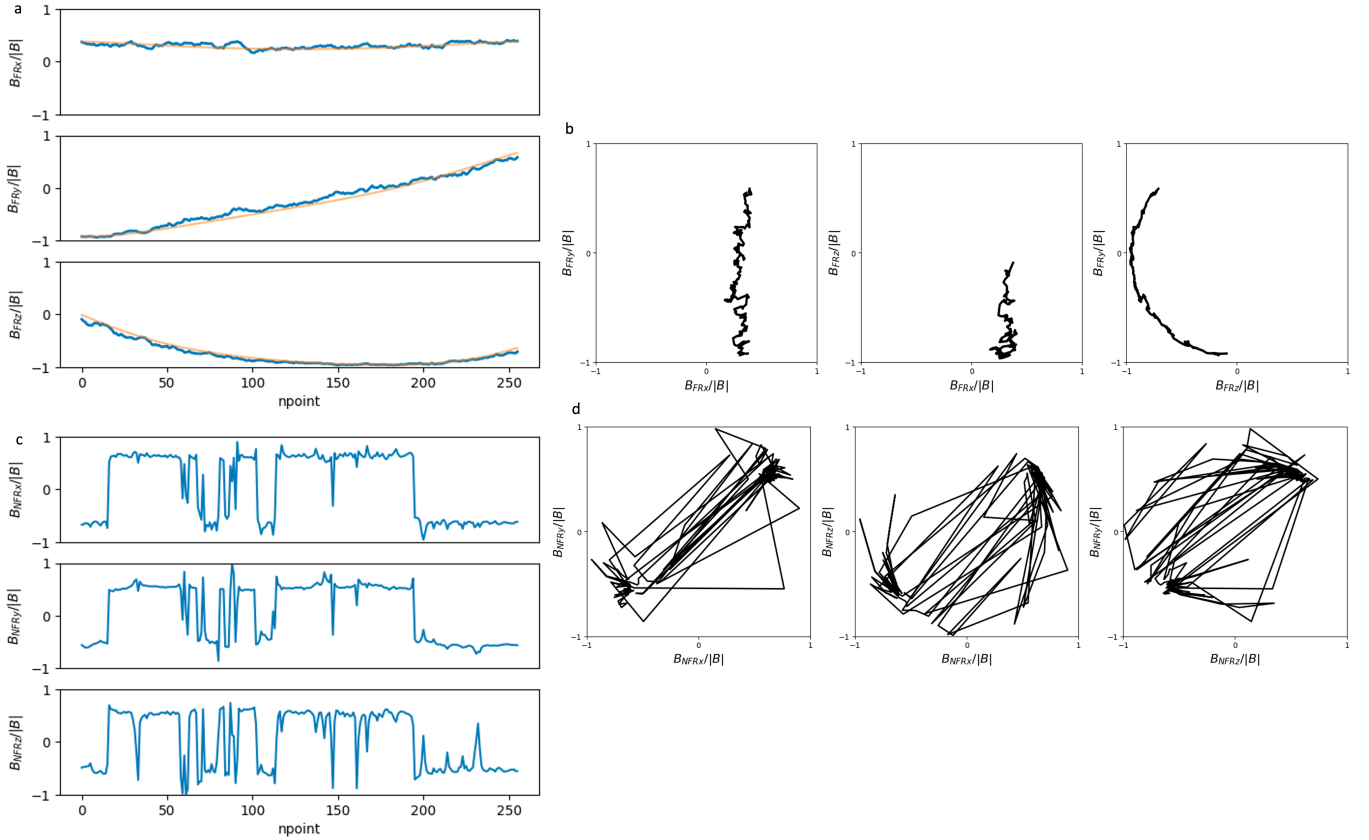
Figure 3. (a) Example of a simulated FR time series $\mathbf{B_{FR}}$ and (b) hodogram in $x_{gse} - y_{gse}$, $y_{gse} - z_{gse}$ and $z_{gse} - x_{gse}$ planes. The light-orange line over-plotted on (a) represents the $\mathbf{B_m}$ signal. (c) Example of a simulated NFR time series $\mathbf{B_{NFR}}$ and its (d) hodogram in three different planes. The magnetic field vectors are normalized with intensity $|B|$.

Note that here each of the $\mathbf{B_{FR}}(t)$ has *npoint* number of points in it. Thus, we simulate a total number of 161,000 unique synthetic large-scale FR magnetic field time series with each having underlying PSDs equal to one of the real events. As this work focuses on the geometry of magnetic structures rather than the magnitude, we plotted each $\mathbf{B_{FR}}(t)$ in three different planes $x_{gse} - y_{gse}$, $y_{gse} - z_{gse}$ and $z_{gse} - x_{gse}$ in the -1 to 1 range and generated three different images of resolution 54×54, where each of them are called hodograms. Finally, the hodograms of three different planes are merged into a single image of size (Width) $162 \times$ (Height) 54. In Figure 3a and b we show a simulated FR time series and hodogram, respectively, where the orange curves on Figure 3a represents $\mathbf{B_m}(t)$. The CC model inputs θ, ϕ, Y_0 and H and the detrending factor ΔT for generating this particular synthetic FR is used as 60 deg, 32 deg, 0.65, 1 and 30 min, respectively.

GENERATION OF SYNTHETIC NON-FLUX ROPES

Since we do not anticipate any typical rotation in NFRs, we only include the process of generating the solar wind background profile $\delta\mathbf{B_{NFR}}(t)$ to create synthetic NFRs. To create $\delta\mathbf{B_{NFR}}(t)$ we compute PSDs of the normalized magnetic field components of real NFRs, $P_{nfr}(k)$. From the Wind/MFI data, we randomly selected 2200 real unique NFR intervals having durations of 24 hours. A 24-hour duration for the NFRs was selected because this is equal to the average ICME duration at 1 au. Each of the NFRs starts at least 2 hours after the 'ICME/MO end time' and ends at least 2 hours before 'MO start time' mentioned in *Wind ICME catalog*. Additionally, we obtain 303 sheath intervals from the *Wind ICME catalog* which are bounded by 'ICME start time' and 'MO start time'. Thus we generate a total number of 2503 $P_{nfr}(k)$s, each of them corresponding to a real NFR that does not have any overlap with ICME MOs cataloged in the *Wind ICME catalog*. In Figure 2c, we show with green and orange lines $P_{nfr}(k)$ as a function of spacecraft-frame frequency for the three different magnetic field components of two 24-hour NFR segments ahead of and after the ICME FR shown in Figure 2a. The spectral slope is found to be ~ -1.8 for both upstream and downstream NFR segments. The ΔT used to create $P_{nfr}(k)$ in this case, is 30 min. As no rotation in the NFRs'
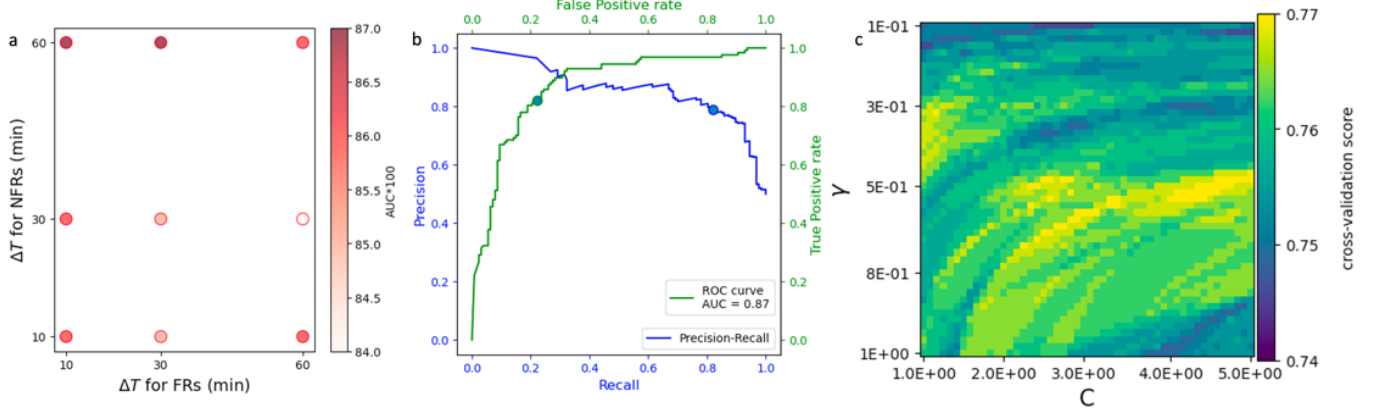
Figure 4. (a) The Area Under the Curve (AUC) values of nine M1 models with similar architecture trained with nine different combinations of simulated FR and NFR data. The detrending factors used to generate the simulated training data are 10, 30, and 60 mins. The maximum AUC value achieved as 0.87. (b) The ROC (green) and precision-recall (blue) curves corresponding to the M1 having maximum AUC accepted as the final model named DeepFRi. (c) Color map showing cross-validation scores obtained for different values of the hyperparameters C [1-$10^{0.7}$] and γ [$10^{-1} - 1$] of the RBF kernel of SVM model in M2.

magnetic field components is expected, we generate the NFRs using Equation 2, where P_{fr} is replaced by P_{nfr}. Thus,

$$\mathbf{B_{NFR}}(t) = \delta\mathbf{B_{NFR}}(t) = \mathcal{F}^{-1}(\sqrt{P_{\mathrm{nfr}}(k)}u(k)). \qquad (4)$$

Once we generate a similar number of synthetic NFR time series as FRs, we obtain their hodograms in each plane and merge them into a single image of similar size as the synthetic FR hodograms. In Figure 3c and d, we show an example synthetic NFR time series and hodograms, respectively, where the ΔT for generating this particular synthetic NFR is used as 30 min. Finally, we generate nine sets of synthetic FR and NFR data, where each data set is prepared using a unique combination of ΔT values for both FRs and NFRs.

2.2.3. *Model training and evaluation*

Once we prepare the synthetic data set, we randomly select 30% of the training data for validation. To train the network weights, we use the Adam optimizer (Kingma & Ba 2014) with an initial learning rate of 0.001. We find that the accuracy and loss converge quickly for the validation set and infer that this may happen due to the simplicity of the classification problem in the simulation space. To avoid over-fitting we limit the training of the network to 50 epochs. Thus, we train nine M1 models of similar architecture using nine different sets of synthetic data. Using a system with a 10-core CPU processor and 16 GB RAM, we finish training each model in 58 minutes. After training each, we save their weights, evaluate each model with real cases, and derive the evaluation metrics. Finally, out of nine trained M1 models, we choose a model that provides the highest Area Under the Curve (AUC) value measuring the area under the 'Receiver Operating Characteristic' (ROC) curve and designate this M1 as the DeepFRi model.

For evaluating the performance of M1 models, we obtain 127 real ICME FRs having duration between 20-30 hours from *Wind ICME catalog* and an equal number of non-ICME windows of 24 hours randomly chosen from Wind/MFI data. We use these events to evaluate the M1 models and derive evaluation metrics – accuracy, precision, recall, F1 Score and AUC for each M1. These are the standard metrics for evaluating machine learning models. They are described using four terms, namely true positives (TP: cases where the predicted and original classification agree on being positive (i.e. FR)), true negatives (TN: cases where the predicted and original classification agree on being negative (i.e. NFR)), false positive (FP: if the prediction is positive (i.e. FR) but the ground truth is negative (i.e. NFR)), and false negative (FN: if the classification is negative (i.e. NFR) and the ground truth is positive (i.e. FR)). In Table 1, column 1, we describe the evaluation metrics.

In Figure 4a, we display the AUC values of nine M1 models with circles of varying shades of red colour. We notice that training the same model architecture with different data sets prepared using different detrending factors leads to varying AUC values. The maximum obtained AUC is 0.87. In Figure 4b, we show the ROC (in green) and precision-recall (in blue) curves of one of the highest performing models named as DeepFRi trained with a set of synthetic FRs and NFRs, whose underlying PSDs are prepared using the linear detrending scales of $\Delta T = 10$ and 60 minutes, respectively. The green and blue circles over-plotted to the curves corresponds to the decision threshold.

Table 1. Evaluation metrics of the DeepFRi and DicFR models.

Performance Metrics	Description	DeepFRi performance	DicFR performance
Accuracy	$\frac{TP+TN}{Total\ number\ instances}$	0.8	0.83
Precision	$\frac{TP}{TP+FP}$	0.78	0.85
Recall	$\frac{TP}{TP+FN}$	0.82	0.76
F1-Score	$\frac{2*Precision*Recall}{Precision+Recall}$	0.8	0.805

The evaluation scores including accuracy, precision, recall and F1 Score of the DeepFRi, are indicated in Column 3 of Table 1.

To train M2, we randomly select 75% of ICMEs listed in the *Wind ICME catalog.* We use the Python scikit-learn library to implement SVM in M2. SVM has been integrated into the Python scikit-learn library (Pedregosa et al. 2011). It has open-source access and a well-established documentation http://scikit-learn.org/stable/. We use the stratified k-fold cross-validation method with k=5. It is implemented in the cross-validation module of the scikit-learn library. The non-linear radial basis function (RBF) kernel of SVM has two hyper-parameters – ('C' and γ) that need to be tuned to achieve the highest model performance. The C is a regularization parameter that trades off correct classification versus maximizing the margin of the decision function. A low C value corresponds to a large margin and simple decision function and a high C value leads to overfitting. The γ is a kernel width parameter that affects the shape of the class-dividing hyperplane. While using the RBF kernel, we preliminary set the range of C and γ as $[10^{-2}, 10^2]$ (Pricopi et al. 2022) and tune the parameters in that range to obtain the highest cross-validation score. Based on the performance, we further shrink the range of C to $[10^0, 10^{0.7}]$ and γ to $[10^{-1}, 10^0]$ and find the best-performing model with the highest cross-validation score of 0.77 at $C = 4.428$, $\gamma = 0.543$, and random state = 42. In Figure 4c, we provide a color-map showing the variation of cross-validation scores with different values of C and γ. While using the linear kernel of SVM in M2, we find the cross-validation score to be 0.75. We compare our model performance with that of the K-nearest neighbor (KNN; Mucherino et al. 2009; Taunk et al. 2019) classifier. The KNN follows a similarity-based algorithm where the label of an event is assigned based on that of the majority of K most similar examples. Here similarity is expressed in terms of the generalized concept of distance. The maximum cross-validation score achieved in the case of the KNN classifier is 0.75 with $K = 12$. Therefore, we choose M2 with the SVM architecture and with a tuned RBF kernel to use as the DicFR model.

To evaluate the performance of DicFR, we use the remaining 25% cases from *Wind ICME catalog.* In Column 4 of Table 1, we provide the evaluation scores of the DicFR model.

2.3. *Post-processing*

The DeepFRi model classifies each input window as FRs or NFRs and assigns them a binary value $V_W = 1$ or 0, respectively. We first convert the hodograms to time-series data. Through a few post-processing steps, we derive a list of large-scale FR intervals containing the category of their geoeffectiveness. Below we describe the post-processing steps.

Step 1: First we note the start time of each window that is predicted as $V_W = 1$, by the model. If two consecutive windows (e.g. W_1 and W_2) with $V_W = 1$ have start times are more than 24 hours apart, we record both the windows' end time and start time as corresponding to the previous FR's end and the next FR's start boundaries, respectively. Thus we check whether the same condition holds for the next two consecutive windows (W_3 and W_4), and so on. If the condition is not satisfied, we continue examining the rest of the window start times until the requirement is satisfied. Following this process, we generate a list (*List 1'*) where each row contains a start and end time of an interval and each data points inside the intervals belong to windows with $V_W = 1$.

Step 2: In the next step, we compute the cumulative sum V_{cumsum} of V_W values of the windows within each interval in *List 1'*. Note that each interval may contain a single or multiple number of windows. Intervals containing multiple windows can have V_{cumsum} values varying over time. Once we derive the V_{cumsum} for each interval, we obtain its full width at half maximum (FWHM) duration and define it as an auto-detected FR interval. These intervals are noted in *List 1*. In Figure 5a, we show an interval containing solar wind magnetic field vectors which have been used as input to our model. The red-shaded rectangles indicate the times when $V_W = 1$. Using a black dashed curve, we
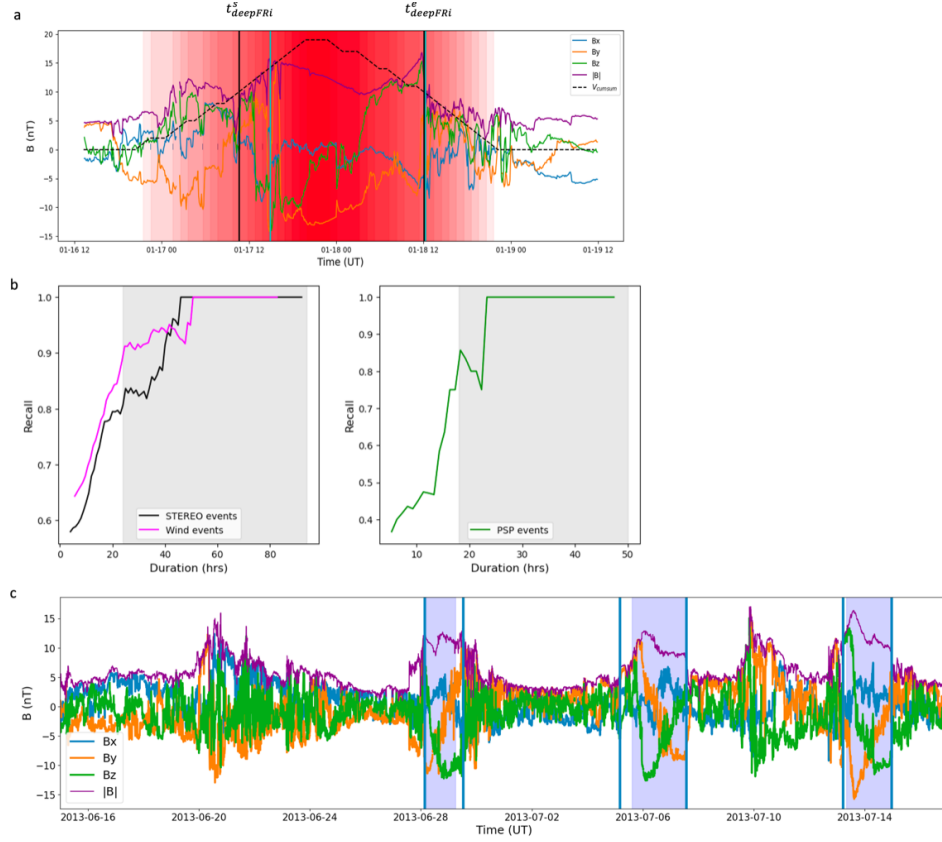
Figure 5. (a) A solar wind magnetic field interval during 2013/01/16 12:00 - 2013/01/19 12:00 UT. Each red-shaded rectangle has duration of 24 hours and $V_W = 1$. The black dashed line indicates V_{cumsum}. The black vertical lines mark the $t^s_{DeepFRi}$ and $t^e_{DeepFRi}$ interval. The cyan vertical lines show boundaries of the event obtained from the *Wind ICME catalog*. (b) Sensitivity of the DeepFRi model as a function of ICME FR duration. The purple, black and green colored lines correspond to the ICMEs observed by Wind, STEREO and PSP spacecraft, respectively. The recall values belonging to the gray-shaded regions are ≥ 0.8. (c) A 30-day interval of solar wind magnetic field vectors where the blue shaded regions indicate auto-detected geoeffective FR intervals. The solid vertical lines indicate FR start and end times obtained from RL.

show V_{cumsum}. We note the start and end times of the FR interval as $t^s_{DeepFRi}$ and $t^e_{DeepFRi}$ in black vertical lines, respectively. The cyan vertical lines show the FR boundaries noted in the *Wind ICME catalog*.

Step 3: Finally, we use a magnetic field magnitude threshold B_{Th} and check whether each interval's mean magnetic field intensity B_{mean} crosses B_{Th}. To obtain B_{Th}, we derive the cumulative distribution function $(F(X) = P[X \leq x])$, where X is the B_{mean} of real events from *Wind ICME catalog* with duration 24 hours or more. We assign B_{Th} such as $P(X \geq B_{Th}) = 0.9$. Once the auto-detected FR intervals satisfy the threshold criteria, they are put in the list *List 2*.

Step 4: We derive the geoeffective properties of the FR intervals required to provide input to the DicFR model. The DicFR model classifies the intervals as class 1 (geoeffective) or class 0 (non-geoeffective) and a *Final List* is created that contains $t^s_{DeepFRi}$, $t^e_{DeepFRi}$ and FR geoeffectiveness category (class 0 or 1).

3. RESULTS AND DISCUSSION

We now assess the performance of our pipeline by comparing its output to the existing manually detected open-source ICME lists. For this purpose, we use the list of ICME events observed by STEREO A and B during 2007–2020 available at *STEREO ICME catalog* and the reference list (RL) used by Nguyen et al. (2019). The RL is a union of different Wind ICME lists obtained from Jian et al. (2006); Lepping et al. (2006); Richardson & Cane (2010); Chi et al. (2016) and Nieves-Chinchilla et al. (2018) and compiled by Nguyen et al. (2019) available at https://github.com/gautiernguyen/Automatic-detection-of-ICMEs-at-1-AU-a-deep-learning-approach and a recently created list of ICMEs observed by PSP during 2018 - 2022 (Salman et al. 2024). For this purpose, we use magnetic field data derived from the In-Situ Measurements of Particles and CME Transients (IMPACT; Luhmann et al. 2008) instrument onboard STEREO twin spacecraft and Fields (FIELDS; Bale et al. 2016) investigation of PSP. It is

important to note that any existing ICME FR catalogs are likely not exhaustive and the event start and end times are based on the interpretation of data by human expertise. We first checked how our model identifies the listed events in the catalogs. We then chose a seven-year interval (2008–2014) of continuous Wind/MFI data, prepared the *Final List*, compared that with the open-source lists, and obtained the recall and precision values for both models in the pipeline. In Figure 5b, we show the sensitivity of DeepFRi as a function of duration of ICME FRs obtained from different spacecraft. In the left panel, we show how the model's recall value changes while testing it on real events with duration ranging from ∼5-90 hours observed at 1 au. The gray-shaded region corresponds to recall ≥ 0.8. The high recall values in the shaded region indicate the better performance of DeepFRi while testing it with large-scale FRs. We notice that the recall becomes 0.886 while testing the model with events having ≥ 24 hours interval obtained from RL list. However, we observe that even if we used 24-hour window in training purpose, our model identified 75% of the ICME FRs observed by Wind having duration ≥15 hours.

We provided 2008-2014 Wind/MFI data to our pipeline. The data is first pre-processed and provided to the DeepFRi model following the steps described in Section 2.1. The DeepFRi model classified 639 windows as class 1, where each window lasts ∼24 hours and two consecutive window start times are minimum 1 hour apart. By following *Step 1 – Step 3* described in Section 2.3 we generated a *List 1* with 296 FRs and *List 2* with 110 FR intervals having mean intensity greater than B_{Th}. In Figure 5c, we show a small 30-day subset of the whole interval 2008 - 2014. The blue-shaded regions show the intervals identified by our pipeline as geoeffective FRs presented in the *Final List*. The vertical solid lines are FR boundaries noted in RL.

We compared our auto-generated FR lists to the RL events and obtained the number of TPs, FPs, TNs, and FNs. We derived model precision and recall following the process described in Table 1. Out of 64 events with duration ≥ 24 hours obtained from RL during 2008-2014, we found 56 events that are detected by our pipeline and presented in *List 1*. We computed the recall as 0.875. while comparing *List 2* to the RL events, we find the recall as 0.8 and precision as 0.56. The high number of FPs suggests that the DeepFRi model identifies a number of rotations in solar wind magnetic field that may not correspond to ICME FRs. Including additional solar wind features e.g. temperature, proton β, velocity etc in detection process may help reducing the FPs.

In Figure 6a and b we summarise the result that we found while applying the DeepFRi model to detect ICME FRs during the 2008 - 2014 time period. The blue bars in Figure 6a indicate the number of FRs detected by DeepFRi each year. The over-plotted gold bars show the number of events that belong to the RL and are also detected by our DeepFRi model. The number of events listed in RL for each year between 2008 and 2014 with duration more than or equal to 24 hours is displayed in Figure 6b by the orange bars. Black bars are superimposed on the orange ones to display the events listed in RL detected by DeepFRi.

Nguyen et al. (2019) used 33 different magnetic field and plasma parameters directly measured and derived from in situ observations to automatically detect ICME FRs using a CNN architecture. During 2010 – 2015, using Wind spacecraft data, they obtained a maximum recall of 0.84 and average precision of 0.697. When they considered only the magnetic field magnitude and component data, they found an average precision of 0.593 that is similar to what we obtain. Using Deep Residual U-Net model and training with the same 33 FR parameters used by Nguyen et al. (2019), a study by Rüdisser et al. (2022) found the recall and precision to be 0.67 and 0.7, respectively. Our pipeline finds the start of an ICME present in the *Wind ICME catalog* with a mean absolute error (MAE) of ∼ 7 hours, and the end time with an MAE of ∼ 6 hours. In both Nguyen et al. (2019) and Rüdisser et al. (2022), the MAE values for the ICME start and end are less than that of our result because while training, they used a more sophisticated set of features and multiple sliding window sizes. Moreover, they optimized the decision threshold such that their model can determine the ICME intervals with high precision.

With careful manual observation of solar wind plasma and magnetic data, Salman et al. (2024) identified ICMEs observed by PSP during 2018 - 2022 at heliocentric distances in the range of 0.23 - 0.83 au. We notice that the DeepFRi model can detect more than 80% of listed PSP ICMEs when their duration is ≥ 18 hours. In the right panel of Figure 5b, we show the change of recall values while the model performance is tested on PSP events whose duration ranges from ∼6-48 hours. Note that the number of FNs increases because we train our model using 24 hours magnetic field data window and the PSD of fluctuations at 1 au. Good et al. (2023) studied the turbulence properties of ICMEs in the inner heliosphere and found that with increasing heliocentric distance the power of magnetic field fluctuation decreases. Using PSP, Chen et al. (2020) showed the variation of solar wind magnetic field PSD at different heliocentric distances. They found a significant radial evolution of the power of magnetic field fluctuations and the

inertial range spectral slope of magnetic field PSDs. Therefore, for achieving better results at distances less than 1 au using our model, the preparation of the training dataset should include magnetic field PSDs at those distances.

Now, we compare the geoeffectiveness of auto-detected FRs obtained from the pipeline and available in the *Final List* with the real minimum Dst index during these intervals and RL events. During 2008 - 2014, we found 110 auto-detected FRs with the boundaries identified by $t^s_{DeepFRi}$ and $t^e_{DeepFRi}$ using our pipeline. We derive R_{Bz} and $-B_{z_{\mathrm{gsm}},min}$ for each interval and provide these as inputs to the DicFR model. The model provides a binary output that indicates if the interval contains a $Dst_{min} < -50$ nT or not. Once the geoeffectiveness is predicted, the *Final List* is generated following *Step 4* described in Section 2.3. In our local system, it took $\sim$30 minutes to generate the *Final List* from the 2008 - 2014 interval of raw Wind/MFI data. After comparing with the geoeffectiveness category obtained from real Dst_{min} values of the auto-detected FR intervals, we find the accuracy of the model to be 0.88, recall to be 0.8, precision to be 0.77 and F1-Score to be 0.78. The *Final List* may include non-ICME FRs, whereas the DicFR model is trained only on ICME FRs. Furthermore, certain events might have R_{Bz} and $-B_{z_{\mathrm{gsm}},min}$ parameter values similar to those of geoeffective FRs, but their pressure and velocity might not be similar to those of geoeffective ones. This may lead to the generation of more false positive events.

In Figure 6c and d we summarise the DicFR model output using histograms. The green and red bars in the plots indicate the number of DeepFRi-detected FR intervals classified as non-geoeffective (class 0) and geoeffective (class 1), respectively. In Figure 6c, the solid and dashed-line bars superimposed on the red and green bars indicate the number of DeepFRi detected events that are geoeffective and non-geoeffective in reality, respectively. In Figure 6d, the black solid and dashed-line bars indicate the number of geoeffective and non-geoeffective events as identified in the RL. The red and green bars show the number of events correctly classified by DicFR.

4. SUMMARY AND CONCLUSION

This work provides a procedure to automatically detect heliospheric large-scale flux ropes (FRs) in solar wind data obtained from spacecraft at 1 au and classify their geoeffectiveness. We establish a pipeline that accepts solar wind magnetic field data from spacecraft and produces a list that contains auto-detected FR boundaries and information on whether the FR interval can produce moderate to intense geomagnetic storms or not. To accomplish this operation, the pipeline uses two models called DeepFRi and DicFR which are based on two ML techniques – convolutional neural network (CNN) and support vector machine (SVM), respectively. The DeepFRi model is trained with simulated FR (positive class) and non-FR (negative class) events prepared by combining a modeled magnetic topology of FRs and power spectral densities of the magnetic field fluctuations in the solar wind. The DicFR model is trained with two geoeffective characteristics – maximum southward magnetic field intensity and duration of southward magnetic field with respect to the whole FR intervals. Furthermore, these two models only use solar wind magnetic field parameters.

The pipeline we propose here introduces a way of automatically now-casting FRs including their geoeffective characteristics. The pipeline can be used for space weather operations with further improvements. Currently, the pipeline uses calibrated and processed magnetic field data. To utilise this in auto-flagging space weather events in real-time and issuing alert, additional data pre-processing steps are required. This includes the preparation and calibration of raw data obtained straight from instruments. We use the existing knowledge of FRs' generic structure, its duration, and FR and NFR magnetic field fluctuations properties at 1 au to train the models in our pipeline. In this work, we mainly focus on detection of 1-au ICME FRs and their geoeffectiveness. While applying this pipeline at distances less than 1 au, it under-performs, which reflects the changing FR duration and solar wind fluctuations with radial distance from the Sun. Therefore, before using this pipeline at other radial distances, re-training of the models based on the local FR and NFR properties is required.

The reason for considering only magnetic field information in this work is mainly the lower availability of continuous solar wind plasma and composition information from probes at different locations. For example, during the cruise phase, a planetary mission Juno spent 5 years in the solar wind between 1 - 5.4 au and allowed us to study only the magnetic properties of ICMEs (Davies et al. 2021) using its Magnetic Field Investigation instrument (MAG; Connerney et al. 2017). Another planetary mission, BepiColombo (Benkhoff et al. 2021), which aims to understand Mercury, provides magnetic data but no solar wind plasma information. Therefore, our framework is capable of detecting FRs from spacecraft that provide only magnetic data, and also from Solar Orbiter (Müller et al. 2020) and PSP, even when one or several plasma measuring instruments malfunction and are unable to provide data coverage. It can also be applicable to other spacecraft like Cluster, Helios, Ulysses, Venus Express, etc. that have measured the solar wind magnetic field over the past two decades throughout the heliosphere. However, training with only
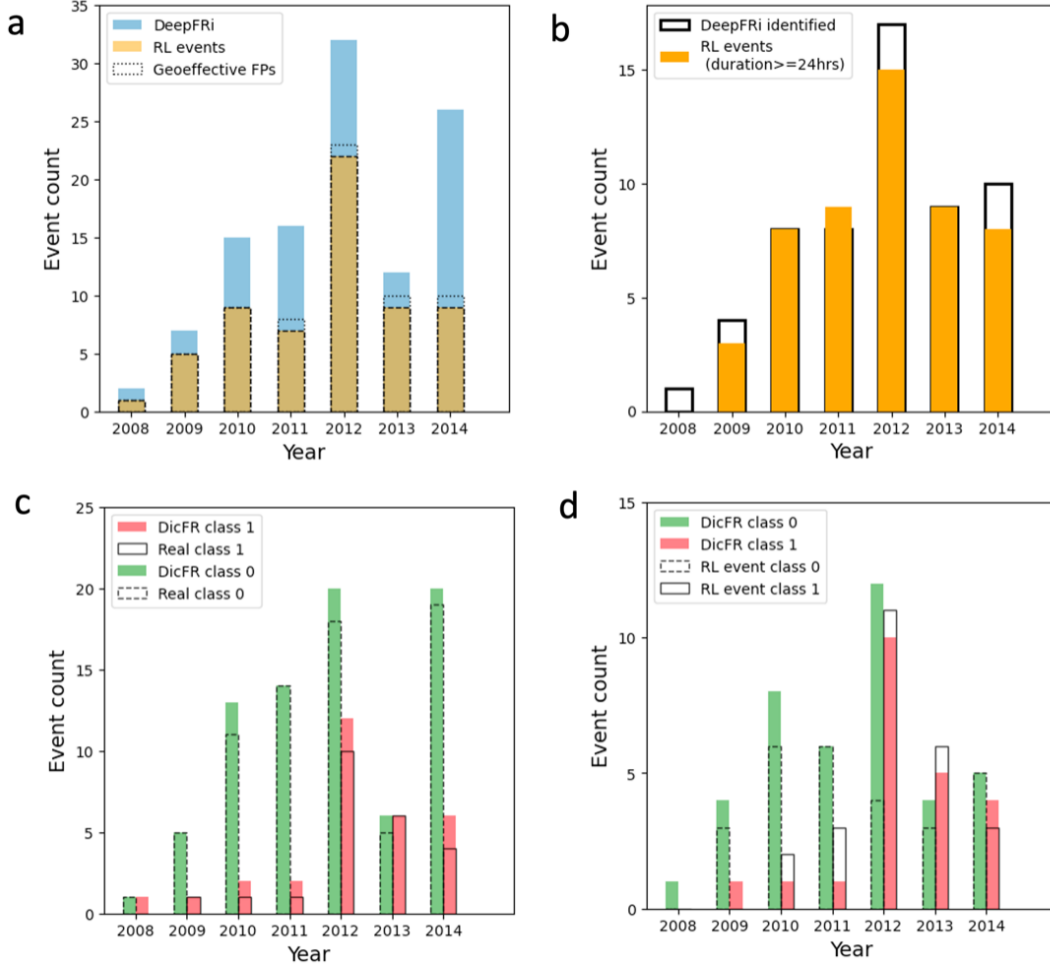
Figure 6. (a) Histogram with blue bars showing the number of events identified by the DeepFRi model during 2008-2014. The listed number of ICMEs detected by DeepFRi is shown using gold bars. The false positive events that have $Dst_{min} < -50$nT are represented by the region between dotted and dashed lines at the bars of 2011–2014. (b) The number of ICMEs listed in RL each year between 2008 and 2014 with duration ≥ 24 hours are shown with orange bars. The solid line bars represent DeepFRi-identified ICMEs belonging to the RL. (c) DicFR model-classified (class 1 (red): geoeffective and class 0 (green): non-geoeffective) FRs obtained from DeepFRi each year between 2008-2014. The solid and dashed-line bars represent the real geoeffective and non-geoeffective classes found from the Dst index available at the OMNIWeb database. (d) Solid and dashed line bars represent class 1 and class 0 events listed in the RL. Green and red bars represent DicFR-identified classes of the events.

the magnetic field features leaves us with a number of false positive cases. While our pipeline may be helpful for space weather event identification purposes, we recommend manually verifying each auto-detected event utilizing its plasma and composition properties for event-based scientific study. After examining the false negative events, we have determined that our pipeline is unable to identify events that are too close to each other to be distinguished separately, have a short duration, or lack any rotation. However, an ensemble of training dataset with different windows may lead to success in the identification of short-duration ICME FRs and thus may decrease the number of false negative events. In future, we intend to modify the models in our pipeline to facilitate the identification of short-duration ICMEs.

Due to the lack of a very large FR database, we use synthetic data to train our model based on the architecture of a handwriting recognition model. For preparing the training dataset, we utilize a simplistic physics-based FR model that is unable to provide complicated magnetic configurations of distorted FRs. That may lead to the misclassification of solar wind data. This means that creating intricate FR models with more observable features is necessary, as is incorporating the models to produce better training data. Recently a study by Nieves-Chinchilla et al. (2023) has established a technique that may enable the modeling of realistic distortion of FR cross-sections. This work only aims

to classify solar wind in flux rope and non-flux rope segments. To make our framework more comprehensive, we intend in the future to automatically classify the detected FRs in different categories based on their rotation, distortion and more complex structures. This may call for more advanced and complicated FR models to prepare training dataset.

The performance of the second model, DicFR, which detects the geoeffectiveness of the DeepFRi-identified FRs in the pipeline, could be improved by including parameters like solar wind velocity and proton density during the FR interval. As we focused only on magnetic field information, we decided not to involve any plasma parameters in this particular study. We employed ICME-related FRs to train both the DeepFRi and DicFR models since they are more likely to cause geomagnetic storms than non-ICME ones. Integrating the second model into our pipeline makes it more favorable to space weather applications. Our tool could substantially contribute to space weather now-casting and forecasting processes.

The trained ML model weights, codes that are used to apply the pipeline to the solar wind data and *List 1*, *List 2* and *Final List* derived during the test period 2008 - 2014 are made available in https://github.com/spal4532/DeepFRi-and-DicFR. Users may download our pipeline in their local system and run on real 1-au solar wind data to prepare auto-detected FR lists and their geoeffectiveness.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., et al. 2016, arXiv e-prints, arXiv:1603.04467, doi: 10.48550/arXiv.1603.04467

Álvarez-Alvarado, J. M., Ríos-Moreno, J. G., Obregón-Biosca, S. A., et al. 2021, Applied Sciences, 11, 1044

Asensio Ramos, A., Cheung, M. C. M., Chifu, I., & Gafeira, R. 2023, Living Reviews in Solar Physics, 20, 4, doi: 10.1007/s41116-023-00038-x

Bale, S. D., Goetz, K., Harvey, P. R., et al. 2016, SSRv, 204, 49, doi: 10.1007/s11214-016-0244-5

Baruah, Y., Roy, S., Sinha, S., et al. 2024, Space Weather, 22, e2023SW003716, doi: 10.1029/2023SW003716

Benkhoff, J., Murakami, G., Baumjohann, W., et al. 2021, SSRv, 217, 90, doi: 10.1007/s11214-021-00861-4

Bothmer, V., & Schwenn, R. 1998, Annales Geophysicae, 16, 1, doi: 10.1007/s00585-997-0001-x

Burlaga, L., Sittler, E., Mariani, F., & Schwenn, R. 1981, J. Geophys. Res., 86, 6673, doi: 10.1029/JA086iA08p06673

Cane, H. V., Richardson, I. G., & St. Cyr, O. C. 2000, Geophys. Res. Lett., 27, 3591, doi: 10.1029/2000GL000111

Cartwright, M. L., & Moldwin, M. B. 2008, Journal of Geophysical Research (Space Physics), 113, A09105, doi: 10.1029/2008JA013389

—. 2010, Journal of Geophysical Research (Space Physics), 115, A08102, doi: 10.1029/2009JA014271

Chandorkar, M., Camporeale, E., & Wing, S. 2017, Space Weather, 15, 1004, doi: 10.1002/2017SW001627

Chen, C. H. K., Bale, S. D., Bonnell, J. W., et al. 2020, ApJS, 246, 53, doi: 10.3847/1538-4365/ab60a3

Chen, J., Deng, H., Li, S., et al. 2022, ApJS, 259, 8, doi: 10.3847/1538-4365/ac4587

Chen, Y., & Hu, Q. 2020, ApJ, 894, 25, doi: 10.3847/1538-4357/ab8294

—. 2022, ApJ, 924, 43, doi: 10.3847/1538-4357/ac3487

Chen, Y., Hu, Q., & le Roux, J. A. 2019, ApJ, 881, 58, doi: 10.3847/1538-4357/ab2ccf

Chi, Y., Shen, C., Wang, Y., et al. 2016, SoPh, 291, 2419, doi: 10.1007/s11207-016-0971-5

Choi, S.-H., Moon, Y.-J., Vien, N. A., & Park, Y.-D. 2012, Journal of Korean Astronomical Society, 45, 31, doi: 10.5303/JKAS.2012.45.2.31

Cireşan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. 2011, in Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two, IJCAI'11 (AAAI Press), 1237–1242

Connerney, J. E. P., Benn, M., Bjarno, J. B., et al. 2017, SSRv, 213, 39, doi: 10.1007/s11214-017-0334-z

Cortes, C., & Vapnik, V. 1995, Machine learning, 20, 273

Crooker, N., Joselyn, J. A., & Feynman, J. 1997, Geophysical Monograph Series, 99, doi: 10.1029/GM099

Daglis, I. A. 2001, SSRv, 98, 343, doi: 10.1023/A:1013873329054

Davies, E. E., Forsyth, R. J., Winslow, R. M., Möstl, C., & Lugaz, N. 2021, ApJ, 923, 136, doi: 10.3847/1538-4357/ac2ccb

Deserno, M. 2004, If Polymerforshung (Ed.), 99, 1

dos Santos, L. F. G., Narock, A., Nieves-Chinchilla, T., Nuñez, M., & Kirk, M. 2020, SoPh, 295, 131, doi: 10.1007/s11207-020-01697-x

Dungey, J. W. 1961, PhRvL, 6, 47, doi: 10.1103/PhysRevLett.6.47

Fairfield, D. H., & Cahill, L. J., J. 1966, J. Geophys. Res., 71, 155, doi: 10.1029/JZ071i001p00155

Farooki, H., Abduallah, Y., Noh, S. J., et al. 2024, ApJ, 961, 81, doi: 10.3847/1538-4357/ad0c52

Fox, N. J., Velli, M. C., Bale, S. D., et al. 2016, SSRv, 204, 7, doi: 10.1007/s11214-015-0211-6

Géron, A. 2017, Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems (Sebastopol, CA: O'Reilly Media)

Gonzalez, W. D., & Echer, E. 2005, Geophys. Res. Lett., 32, L18103, doi: 10.1029/2005GL023486

Gonzalez, W. D., Joselyn, J. A., Kamide, Y., et al. 1994, J. Geophys. Res., 99, 5771, doi: 10.1029/93JA02867

Good, S. W., Kilpua, E. K. J., Ala-Lahti, M., et al. 2020, ApJL, 900, L32, doi: 10.3847/2041-8213/abb021

Good, S. W., Rantala, O. K., Jylhä, A. S. M., et al. 2023, ApJL, 956, L30, doi: 10.3847/2041-8213/acfd1c

Goodfellow, I., Bengio, Y., & Courville, A. 2016, Deep Learning (MIT Press)

Gopalswamy, N. 2008, Journal of Atmospheric and Solar-Terrestrial Physics, 70, 2078, doi: 10.1016/j.jastp.2008.06.010

Gopalswamy, N., Yashiro, S., Akiyama, S., & Xie, H. 2017, SoPh, 292, 65, doi: 10.1007/s11207-017-1080-9

Gruet, M. A., Chandorkar, M., Sicard, A., & Camporeale, E. 2018, Space Weather, 16, 1882, doi: 10.1029/2018SW001898

Hahn, O., & Abel, T. 2011, MNRAS, 415, 2101, doi: 10.1111/j.1365-2966.2011.18820.x

Hu, A., Camporeale, E., & Swiger, B. 2023, Space Weather, 21, e2022SW003286, doi: 10.1029/2022SW003286

Hu, A., Shneider, C., Tiwari, A., & Camporeale, E. 2022, Space Weather, 20, e2022SW003064, doi: 10.1029/2022SW003064

Hu, Q., Zheng, J., Chen, Y., le Roux, J., & Zhao, L. 2018, ApJS, 239, 12, doi: 10.3847/1538-4365/aae57d

Jian, L., Russell, C. T., Luhmann, J. G., & Skoug, R. M. 2006, SoPh, 239, 393, doi: 10.1007/s11207-006-0133-2

Jian, L. K., Russell, C. T., Luhmann, J. G., & Galvin, A. B. 2018, ApJ, 855, 114, doi: 10.3847/1538-4357/aab189

Kahler, S. W., & Webb, D. F. 2007, Journal of Geophysical Research (Space Physics), 112, A09103, doi: 10.1029/2007JA012358

Kaiser, M. L., Kucera, T. A., Davila, J. M., et al. 2008, SSRv, 136, 5, doi: 10.1007/s11214-007-9277-0

Kane, R. P. 2005, Journal of Geophysical Research (Space Physics), 110, A02213, doi: 10.1029/2004JA010799

Kilpua, E. K. J., Luhmann, J. G., Gosling, J., et al. 2009, SoPh, 256, 327, doi: 10.1007/s11207-009-9366-1

Kingma, D. P., & Ba, J. 2014, CoRR, abs/1412.6980. https://api.semanticscholar.org/CorpusID:6628106

Klein, L. W., & Burlaga, L. F. 1982, J. Geophys. Res., 87, 613, doi: 10.1029/JA087iA02p00613

LeCun, Y., Bengio, Y., & Hinton, G. 2015, Nature, 521, 436, doi: 10.1038/nature14539

Lepping, R. P., Berdichevsky, D. B., Wu, C. C., et al. 2006, Annales Geophysicae, 24, 215, doi: 10.5194/angeo-24-215-2006

Lepping, R. P., Acũna, M. H., Burlaga, L. F., et al. 1995, SSRv, 71, 207, doi: 10.1007/BF00751330

Liu, J., Ye, Y., Shen, C., Wang, Y., & Erdélyi, R. 2018, ApJ, 855, 109, doi: 10.3847/1538-4357/aaae69

Lu, J. Y., Peng, Y. X., Wang, M., Gu, S. J., & Zhao, M. X. 2016, Planet. Space Sci., 120, 48, doi: 10.1016/j.pss.2015.11.004

Lugaz, N., Vourlidas, A., & Roussev, I. I. 2009, Annales Geophysicae, 27, 3479, doi: 10.5194/angeo-27-3479-2009

Luhmann, J. G., Gopalswamy, N., Jian, L. K., & Lugaz, N. 2020, SoPh, 295, 61, doi: 10.1007/s11207-020-01624-0

Luhmann, J. G., Curtis, D. W., Schroeder, P., et al. 2008, SSRv, 136, 117, doi: 10.1007/s11214-007-9170-x

Manchester, W., Kilpua, E. K. J., Liu, Y. D., et al. 2017, SSRv, 212, 1159, doi: 10.1007/s11214-017-0394-0

Moldwin, M. B., Ford, S., Lepping, R., Slavin, J., & Szabo, A. 2000, Geophys. Res. Lett., 27, 57, doi: 10.1029/1999GL010724

Moldwin, M. B., Phillips, J. L., Gosling, J. T., et al. 1995, J. Geophys. Res., 100, 19903, doi: 10.1029/95JA01123

Mucherino, A., Papajorgji, P. J., & Pardalos, P. M. 2009, k-Nearest Neighbor Classification (New York, NY: Springer New York), 83–106, doi: 10.1007/978-0-387-88615-2_4

Müller, D., St. Cyr, O. C., Zouganelis, I., et al. 2020, A&A, 642, A1, doi: 10.1051/0004-6361/202038467

Nair, M., Redmon, R., Young, L.-Y., et al. 2023, Space Weather, 21, e2023SW003514

Nair, V., & Hinton, G. E. 2010, in International Conference on Machine Learning. https://api.semanticscholar.org/CorpusID:15539264

Nandy, D., Baruah, Y., Bhowmik, P., et al. 2023, Journal of Atmospheric and Solar-Terrestrial Physics, 248, 106081, doi: 10.1016/j.jastp.2023.106081

Narock, T., Narock, A., Dos Santos, L. F. G., & Nieves-Chinchilla, T. 2022, Frontiers in Astronomy and Space Sciences, 9, 838442, doi: 10.3389/fspas.2022.838442

Nguyen, G., Aunai, N., Fontaine, D., et al. 2019, ApJ, 874, 145, doi: 10.3847/1538-4357/ab0d24

Nieves-Chinchilla, T., Hidalgo, M. A., & Cremades, H. 2023, ApJ, 947, 79, doi: 10.3847/1538-4357/acb3c1

Nieves-Chinchilla, T., Jian, L. K., Balmaceda, L., et al. 2019, SoPh, 294, 89, doi: 10.1007/s11207-019-1477-8

Nieves-Chinchilla, T., Linton, M. G., Hidalgo, M. A., et al. 2016, ApJ, 823, 27, doi: 10.3847/0004-637X/823/1/27

Nieves-Chinchilla, T., Vourlidas, A., Raymond, J. C., et al. 2018, SoPh, 293, 25, doi: 10.1007/s11207-018-1247-z

Nitta, N. V., Mulligan, T., Kilpua, E. K. J., et al. 2021, SSRv, 217, 82, doi: 10.1007/s11214-021-00857-0

Pal, S. 2022, Advances in Space Research, 70, 1601, doi: 10.1016/j.asr.2021.11.013

Pal, S., Balmaceda, L., Weiss, A. J., et al. 2023, Frontiers in Astronomy and Space Sciences, 10, 1195805, doi: 10.3389/fspas.2023.1195805

Pal, S., Gopalswamy, N., Nandy, D., et al. 2017, ApJ, 851, 123, doi: 10.3847/1538-4357/aa9983

Pal, S., Nandy, D., & Kilpua, E. K. J. 2022, A&A, 665, A110, doi: 10.1051/0004-6361/202243513

Pal, S., Nandy, D., Srivastava, N., Gopalswamy, N., & Panda, S. 2018, ApJ, 865, 4, doi: 10.3847/1538-4357/aada10

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, the Journal of machine Learning research, 12, 2825

Pricopi, A.-C., Paraschiv, A. R., Besliu-Ionescu, D., & Marginean, A.-N. 2022, ApJ, 934, 176, doi: 10.3847/1538-4357/ac7962

Qahwaji, R., & Colak, T. 2007, SoPh, 241, 195, doi: 10.1007/s11207-006-0272-5

Qiu, J., Hu, Q., Howard, T. A., & Yurchyshyn, V. B. 2007, ApJ, 659, 758, doi: 10.1086/512060

Richardson, I. G., & Cane, H. V. 1995, J. Geophys. Res., 100, 23397, doi: 10.1029/95JA02684

—. 2004, Geophys. Res. Lett., 31, L18804, doi: 10.1029/2004GL020958

—. 2010, SoPh, 264, 189, doi: 10.1007/s11207-010-9568-6

Rodríguez, J.-V., Rodríguez-Rodríguez, I., & Lok Woo, W. 2022, PASP, 134, 124201, doi: 10.1088/1538-3873/aca4a3

Ronneberger, O., Fischer, P., & Brox, T. 2015, in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, ed. N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Cham: Springer International Publishing), 234–241

Rüdisser, H. T., Windisch, A., Amerstorfer, U. V., et al. 2022, Space Weather, 20, e2022SW003149, doi: 10.1029/2022SW003149

Russell, C. T., & Shinde, A. A. 2003, SoPh, 216, 285, doi: 10.1023/A:1026108101883

Salman, T. M., Nieves-Chinchilla, T., Jian, L. K., et al. 2024, ApJ, 966, 118, doi: 10.3847/1538-4357/ad320c

Schekochihin, A. A. 2022, Journal of Plasma Physics, 88, 155880501, doi: 10.1017/S0022377822000721

Sheeley, N. R., J., Herbst, A. D., Palatchi, C. A., et al. 2008, ApJ, 675, 853, doi: 10.1086/526422

Sheeley, N. R., Wang, Y. M., Hawley, S. H., et al. 1997, ApJ, 484, 472, doi: 10.1086/304338

Shinde, A. A., & Russell, C. T. 2003, in AGU Fall Meeting Abstracts, Vol. 2003, SH21B–0133

Sinha, S., Gupta, O., Singh, V., et al. 2022, ApJ, 935, 45, doi: 10.3847/1538-4357/ac7955

Sun, W.-J., Slavin, J., Smith, A., et al. 2020, in European Planetary Science Congress, EPSC2020–62, doi: 10.5194/epsc2020-62

Taunk, K., De, S., Verma, S., & Swetapadma, A. 2019, in 2019 International Conference on Intelligent Computing and Control Systems (ICCS), 1255–1260, doi: 10.1109/ICCS45141.2019.9065747

Vourlidas, A., & Webb, D. F. 2018, ApJ, 861, 103, doi: 10.3847/1538-4357/aaca3e

Wang, Y., Shen, C. L., Wang, S., & Ye, P. Z. 2003, Geophys. Res. Lett., 30, 2039, doi: 10.1029/2003GL017901

Wang, Y. M., Sheeley, N. R., Socker, D. G., Howard, R. A., & Rich, N. B. 2000, J. Geophys. Res., 105, 25133, doi: 10.1029/2000JA000149

Webb, D. F., & Howard, T. A. 2012, LRSP, 9, 3, doi: 10.12942/lrsp-2012-3

Weiss, A. J., Nieves-Chinchilla, T., Möstl, C., et al. 2022, Journal of Geophysical Research (Space Physics), 127, e2022JA030898, doi: 10.1029/2022JA030898

Weiss, A. J., Möstl, C., Davies, E. E., et al. 2021, A&A, 656, A13, doi: 10.1051/0004-6361/202140919

Welch, P. D. 1967, IEEE Trans. Audio & Electroacoust, 15, 70

Wu, C.-C., & Lepping, R. P. 2002, Journal of Geophysical Research (Space Physics), 107, 1314, doi: 10.1029/2001JA000161

Xu, S. B., Huang, S. Y., Yuan, Z. G., Deng, X. H., & Jiang, K. 2020, ApJS, 248, 14, doi: 10.3847/1538-4365/ab880e

Yu, W., Farrugia, C. J., Galvin, A. B., et al. 2016, Journal of Geophysical Research (Space Physics), 121, 5005, doi: 10.1002/2016JA022642

Yu, W., Farrugia, C. J., Lugaz, N., et al. 2014, Journal of Geophysical Research (Space Physics), 119, 689, doi: 10.1002/2013JA019115