INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

Received: 04-05-2024

Accepted: 16-01-2025

Published: 20-02-2025

Citation: Banerjee A, Bhattacharya R (2025) Solar Wind Speed Variations as Precursor Through X-Class Flare Intensity and Wind Density: A Radio Data Analysis. Indian Journal of Science and Technology 18(6): 403-414. https://doi.org/10.17485/IJST/v18i6.1506

* **Corresponding author**.

banerjee1984abhi028@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Solar Wind Speed Variations as Precursor Through X-Class Flare Intensity and Wind Density: A Radio Data Analysis

Abhijit Banerjee[1,2]*, Rina Bhattacharya[1]

1 Department of Physics, JIS University, 81 Nilgunge Road, Kolkata, 700109, India
2 Department of Computer Science, New Alipore College, Block L, New Alipore, Kolkata, 700053, India

Abstract

Objectives: This study aims to identify solar wind speed as a precursor to the giant class of solar flares, focusing on flare intensity and wind density. **Methods:** A three-tier analysis was conducted on radio data from X-class solar flares on January 9, 2023, and May 11, 2024. Forecasting involved analysing wind speed variations w.r.t flare intensity and wind density. Firstly the data processing included filtration, moving average regressions, Holt-Winters and ARIMA models to identify the best fit. Secondly, the forecasted data were scaled to solar parameter ranges, revealing binary time-scale dependencies. Boolean feature extraction was employed to determine wind parameter frequency. Final predictions were made using an ensemble Random Forest model. Temporal parameter variations of geomagnetic indices are also related to identifications of the perturbation. **Findings:** Machine learning models achieved R^2 scores of 0.503 ± 0.0002 and 0.7125 ± 0.0035 for the two dates. RMSE analysis showed a swing of $|-4.94 \pm 3.811|$ for May 11, 2024. Post-flare (for 1-18 hrs. passage) wind speed (kms^{-1}) and density ($particlesm^{-3}$) differences are $|\pm 1.882 \pm 0.2543|$ and $|\pm 0.257 \pm 0.01055|$ respectively. Boolean expressions revealed solar parameter time dependencies, with frequency analysis yielding a mean of 49.90 ± 0.26 and standard deviation of 28.45 ± 0.05. The Random Forest model yielded an R^2 of 0.8514 ± 0.0012. A Dst change of 18 nT was observed with post-flare enhanced solar wind. **Novelty:** This study introduces a state-of-the-art approach through a three-tier forecasting methods, combining model fitness analysis, Boolean function feature extraction and prediction. This framework captures the unique dependencies between solar wind speed, flare intensity and density highlighting the retention of increased average wind speed following a flare.

Keywords: Solar Flare; Solar Wind Speed; Regression; HW; ARIMA; Random Forest Model

1 Introduction

The solar plasma is an accumulation of charged particles that expands outward from the corona present in the photosphere forming the solar wind and the gravitational attraction of the star is unable to contain this because of its constant heating. It moves in alignment with the magnetic field lines of the Sun that radiate outwardly. Magnetic field lines form a massive spinning spiral above its polar regions during the Sun's twenty seven days period of one axial rotation and thereby producing a steady stream of "wind". These emissions typically instigate from the large luminous areas known as the "coronal holes". Solar wind (SW) connected with coronal mass ejections (CMEs) and solar flares (SFs) should have a finite time relationship among them therefore there will be a certain alteration in the density profile and speed in the SW heading towards the Earth [1,2]. The geomagnetic storms triggered through the CMEs and higher class of SFs can affect the operations of the electronics in the satellites and space-crafts by intervening with their magnetic fields and thereby the conduction system. Hence communication system of the Earth gets perturbed to a great extent. The storm alert system founded on the day-ahead prediction of the giant class of SFs and CMEs directed toward Earth by observing the solar wind speeds (SWSs) variation has been developing throughout the decades and some of them become highly effective [3,4]. Previous research revealed the link between various categories of flares and CMEs by analyzing the data collected from the deep space observatories placed at L1- Lagrangian point and has given light on the critical aspects of predictions. The classical modeling regarding the existence and properties of the SWS is termed as Eugene Parkar's model (1958) [5]. The empirical modeling of the SWS prediction describes its average behavior. The periodicity and variations of SWSs are found out through the Fourier and wavelet analysis on the curated time series data [6]. The magneto-hydrodynamic (MHD) simulations have come as a significant modeling tool to analyze the interaction of the SWS and magnetic field. With the advancement of artificial intelligence (AI) and machine learning (ML) along with deep learning techniques, SWS classification and prediction has shown significant progress. Some of these adopted models are Wang-Sheeley-Arge (WSA)-ENLIL, Physics-informed neural networks (PINN) and Real-time forecasting [7–9]. The WSA-ENLIL associates a semi-empirical close to Sun module and a 3-D MHD numerical model to forecast SWSs and CMEs. This model exhibits strong performance in resource specific, precise and prolonged forecasts. Assimilating the observational data with the fundamental physical principles in relation to the behavior of the SW the PINNs show enhanced accuracy in speed predictions. The real-time forecasting PREDSTORM model is updated every 10 minutes by the received data from space-crafts like DSCOVR and ACE and it's very effective for high-speed SW data [10,11]. In earlier studies, some important aspects of the SWS prediction system are shown. Applying the deep learning model on the extreme-ultraviolet ray imagery collected from the NASA, OMNI data while predicting the SWS at L1 was proven far better than the autoregressive and Naive model of prediction with the higher activation function in the vicinity of the dynamic coronal holes for high and slow wind, prior to the appearance of the SF and ejectras [12]. The coronal-intensity data correspond to the SWS at L1 while balliastically traced back to the Sun were obtained by applying a designed CNN architecture for training images data from the AIA (Atmospheric Imaging Assembly) at the wavelength of 193 Å by assuming the constant speed of propagation [13]. The classification based on time transition probability was obtained by the application of the Gaussian Process classifier algorithm over the hourly averaged SWS database [14]. The solar storms during the solar minimum cause the activities in the Earth's magnetic field. Ambient SW fluxes in interplanetary space represent the evolvement of the storms over the heliosphere before its arrival at the Earth. Thus an accurate forecasting of SW flow neighboring to the Earth is essential. Estimation in this respect was done by using the solutions of magnetic models of the corona by ML techniques. It was reported that an application of ML has classified the SWSs by finding most fittest combination out of 8191 cases using 13-SWS parameters with the comprehensive enumeration of an 8-dimensional scheme at the 1AU distance in the following categories viz. coronal hole origin plasma, sector-reversal-region plasma, streamer-belt-origin plasma and ejecta [15]. As per the above discussed literatures, it is realized that the research on the categorization of the SWSs depending on the parameters considered by utilizing various phenomenon, factors and data types was conducted. Considering these facts the present work is focused on a finer research gap i.e. to propose a new methodology to evaluate the fitness of the other solar parameters to the pre-categorized or leveled SWSs, thereby focusing on its capacity to indicate the occurrences of the perturbations at the distant solar photosphere likely flares and CMEs. Hence along with the regressions and ML models the present analyses introduce the binary dependencies in view of Boolean expression of the SWS with the flare intensity and wind density. In this present work, two major classes of solar flare with X-class intensity on January 9, 2023 and May 11, 2024 are considered. The radio data from the deep space observatories viz. wind speed, flare intensity, wind density and geomagnetic indices are collected. Data for both the flares area fitted in five different forecasting techniques are adopted here from the traditional regression model to supervised ML models. Upon comparing the statistical measures of those models, the most effective sets among them are selected. These forecasted data are considered in terms of time and logical states of their occurrences. Then the logical dependencies between the parameters using minimization techniques are analyzed. Based on the trends observed in the variation of SWS variables, the study maximizes the generation of Boolean datasets and establishes predictive capabilities using a Random Forest model of machine learning. A comprehensive visualization of these dependencies is also presented to enhance the understanding of the observed trends. Results obtained from the three phases of analysis when unfolded some extremely significant observations on the predictive property are found. These striking observations show significant advancement from their analytical view-points regarding the physics involved and the expressions developed in comparison to some of the previous benchmark research. Considering the fact that the ionospheric ring current in D-layer is perturbed at the intensive SFs and leads to the modification of the geomagnetic field pattern at the interaction of the charge particles with the ionospheric plasma [16], the correlations of the geomagnetic indices and SWS is also analyzed here to find the conformity to the obtained result of the solar parameters with respect to their time of variation. The present work is critically considered about two different giant solar flares that happened with an interval of 4 months 3 days and 6.7 hrs. and distributed in winter-evening dust to monsoonal-night time basis respectively for the diversified features and validations of the analysis of the one flare by the other.

2 Methodology

The radio data of X-class solar flare on January 9, 2023 (Test Day-1: TD-1) and May 11, 2024 (Test Day-2: TD-2) for flare intensity, wind speed and density are collected and the analyses are executed by following processes.

I. Initial processing is carried out through listing and sorting, deletion of noisy data and insertion of data through random sampling. This processed data is considered the actual data set and is utilized for the following methods.

The following expression shows smoothening of data through simple moving average filtering model [17].

$SMA_{i+1} = \frac{1}{k}\sum_{i=n-k+2}^{n+1} p_i = SMA_i + \frac{1}{k}\left(p_{n+1} - p_{n-k+1}\right)$

The running mean is carried out over the last consecutive 'k' number of entries of the data; 'n' total data set entries, 'p' values of data points. Here moving average filtration is taken for three periods 5 (MA-P5), 10 (MA-P10) and 15 (MA-P15) respectively.

II. The weighted moving average model [17] is shown by the following expression

$\hat{y}_{i+1} = \frac{\sum_{i=1}^{n} W_i \times y_i}{n}\,; \sum_i^n W_i = 1$

Where, W_i = weight, n = number of observation, y_i = actual values, $\widehat{y_i}$ = predicted values. Four set of combinations of weights are considered here viz. (i) 0.4,0.3,0.2,0.1 (WMA-S1),(ii) $2^7,2^6,2^5,2^4,2^3,2^2,2^1,2^0$(WMA-S2) and (iii) $2^0,2^1,2^2,2^3,2^4,2^5,2^6,2^7$(WMA-S3).

III. The Exponential moving average model [17] is shown by the following expression

$\hat{y}_{i+1} = \hat{y}_i + \alpha(y_i - \hat{y}_i)$

n = number of observation, y_i = actual values, $\widehat{y_i}$ = predicted values. Here three set of α

values are considered viz. α = (i) 0.2; (EMA-S1), (ii) 0.3 (EMA-S2) and (iii) 0.4 (EMA-S3).

IV. Here two supervised machine learning models exponential Halt-Winter (HW) [18] and autoregressive-integrated-moving-average prediction (ARIMA) [19,20] are considered. The model training and testing of the processed data set and the estimation of forecasted data sets are carried out.

Forecast expression for HW model:

$\hat{y}_{t+h} = L_t + T_t + S_{t-m+h\%m}$

L_t, T_t, S_t, y_t and m_t are the level, trend, seasonal component, observed value at time t respectively and $\hat{y}_{t+h}$ forecasted value at time $t+h$, where h is the number of periods ahead to forecast, % denotes the modulo (returns remainder only) operation.

Forecast expression for ARIMA model:

$\emptyset(B)(1-B)^d Y_t = c + \theta(B)\epsilon_t$

p, d and q are the order of autoregressive part, differencing, moving average part respectively. $\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p$; Autoregressive operator

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$; Moving average operator

$By_t = y_{t-1}$; Back-shift operator

V. The models from sections 2.I to 2.IV with better performances are selected depending on their estimated values of the statistical parameters viz. Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD) and the regression parameters from the Analysis of Variance (ANOVA) [21] of the multivariate regression analysis viz. R^2, Significance-F, coefficient and p-value of the time and forecasted values of SFI, SWS and SWD.

$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\widehat{y}_i - y_i)^2}{n}}; MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\left(\frac{\widehat{y}_i - y_i}{y_i}\right)\right| \times 100; MAD = \frac{\sum_{i=1}^{n}|\widehat{y}_i - \mu|}{n};$

$R^2 = 1 - \frac{\sum_{i-1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i-1}^{n}(y_i - \mu)^2}; SignificanceF = \frac{{S_1}^2}{{S_2}^2}; S^2 = \frac{\sum(x_i - \bar{x})^2}{n}; pvalue = 1 - F_{(df_1, df_2)}F$

n = number of observation, y_i = actual values, $\widehat{y_i}$ = predicted values, μ = mean, x_i=sample value, $\bar{x}$=sample mean, S^2= standard deviation, df_1, df_2= degrees of freedom.

VI. On the basis of the prediction ability as derived from their best fitted plots further selection of the models is carried out.

VII. Scales of classifications of solar parameters viz. solar wind speed (SWS), solar flare intensity (SFI) and solar wind density (SWD) are shown in Table 1 . The maximum, minimum, mean and standard deviations of the dataset are considered as the basis of these selected ranges.

Table 1. Ranges of distribution

Ranges	Solar wind speed		Solar flare intensity		Solar wind density	
	Symbols	Values	Symbols	Values	Symbols	Values
1	SWSR1	≥250 : ≤350	SFR1	≥10 : ≤50	SWDR1	≥0 : ≤2
2	SWSR2	≥350 : ≤450	SFR2	≥50 : ≤100	SWDR2	≥2 : ≤5
3	SWSR3	≥450 : ≤500	SFR3	≥100 : ≤150	SWDR3	≥5 : ≤10
4	SWSR4	≥500 : ≤600	SFR4	≥150 : ≤350	SWDR4	≥10 : ≤15
5		-NR-		-NR-	SWDR5	≥15 : ≤20

":" for ranges,"-NR-"for no ranges specified

VIII. Depending on the ranges of solar parameters as depicted in Table 1 the discrete (binary) dependencies on the time of occurrence of wind speed w.r.t. the flare intensity and density are evaluated for the selected models.

IX. The SWS, SFI and SWD as presented in the respective ranges and their occurrences as obtained from section 2.VIII are combined in pair viz. SWS and SFI and SWS and SWD through four and five variables Karnaugh's (K) map respectively. The map is constrained by the bias that "no formation of group is allowed to close the min-terms among the rows" i.e. $\forall\ K_{ij}K_{i'j'} \in G, \left|i-i'\right| > 1$. Where i and i' are row indices of the cells in the same group, G represents any group of cells in the Karnaugh's map K. Parameters in pairs with their co-existence are considered as the "truth-truth" logical condition and denoted by the min-terms i.e. logic-1(occurrence function: $f_{ranges\ of\ SWS,SFI}$) and also other logical conditions with the non-existence of either and both of the parameters are considered as zero. The position of the combinations of the pairs in the K-map denoted by variable w and z (row-wise) for SWS and y and x for SFI (column-wise) and where w and x are the MSB and LSB respectively. Similarly, d and c for SWS (row-wise) and b and a for SWD (column-wise) are the variables to signify the position of the parameters (occurrence function: $f_{ranges\ of\ SWS,SWD}$) and where d and a are the MSB and LSB respectively. From these minimized canonical Boolean expressions to represent the dependencies are estimated.

X. Extracting the features of the derived Boolean expressions (vide 2.IX) in terms of w,z (dependence of SWS on SFI) and d,c (dependence of SWS on SWD) is carried out (i.e. the trends for the combinations w∈{0,1}, z∈{0,1} and c∈{0,1}, d∈{0,1}are extracted). A Boolean datasets of 5000 samples are generated with the optimised variations by supervised machine learning from the extracted features. The selection of the sample number is based on the consideration of average number of major class of solar flare happened on solar maximum (5.5 years in 11 years cycle) for 2 years (2023-2024). That average number comes out as approximately 14,600 [22]. Where our sample size stands significant by covering 34% of that set. Here these samples are assigned by a number named as Method Index.

XI. For predicting the occurrence of SWS values based on input combinations of SFI and SWD variables. On the generated data set (vide 2.X) for SFI and SWD combinations, along with their corresponding target values for in w,z and d,c (scaled between 0 and 100) a supervised learning algorithm Random Forest Regressor model [23] is applied to train separately for the predictions of w,z and d,c. Where the input features consist of SFI and SWD combinations.

XII. 3-Dimentional (3-D) plotting with colour mapping to visualize the frequency trends (vide 2.X) and the predictions of the occurrences (vide 2.XI) of the two sets of variables are developed. Thus illustrating how the occurrence of w,z and d,c varies as a function of the SFI and SWD combinations.

XIII. The model's accuracy metrics viz. R^2, MAE and RMSE are calculated for a quantitative measure of the prediction accuracy and model fit.

XIV. Microsoft Advance Excel (processing, forecasting, statistical analysis, optimization and conditional searching from sections 2.I to 2.IV, 2.VI to 2.VIII and 2.IX) and Python programming (involving the activation of Graphics Processing Unit) for the machine learning algorithms for forecasting from the sections 2.IV, 2.V for the models HW, ARIMA, the sections from 2.X to 2.XIII for the extraction of features (supervised machine learning), Random Forest Regressor model (supervised machine learning), model accuracy metrics and 3-D colour plotting respectively are utilized in the present work.

XV. Correlations and the variations of the flare intensity, wind speed and geomagnetic indices for the actual and forecasted set of data have been estimated.

3 Data sources

Solar parameters: National Oceanic and Atmospheric Administration (NOAA), Advance Composition Explorer (ACE), Deep Space Climate Observatory, National Aeronautics and Space Administration (NASA) and European Space Agency (ESA). Geomagnetic indices: University of Kyoto, Helmholtz Center Potsdam German Geo Research Center.

4 Results and Discussion

The maximum variation of solar flare intensity on TD-1 and 2 are 198 (18.5UT) and 589 (1.22UT) Wm^{-2} respectively. Sharp growths and decay for X-class of flares are observed. The half power maxima and bandwidth of the X- class flares w.r.t maximum amplitude are of the values 140.007 (0.08UT) and 416.49 (0.114UT) Wm^{-2} and for TD-1 and 2 respectively. Apart from these 3 and 5 numbers of M-class flare with the highest amplitude of 54 (19UT) and 87.89 (15.12UT) Wm^{-2} are also observed on TD-1 and 2.

Table 2. Descriptive statistics of the forecasting

Methods	Para meters	RMSE (TD-1-TD-2)	MAD (TD-1-TD-2)	MAPE (TD-1-TD-2)	Para meters	P-value		R^2		Significance-F	
						TD-1	TD-1	TD-1	TD-2	TD-1	TD-2
	SFI	-21.96	-4.41	-5.82	Time	6.11×10^{-91}	6.11×10^{-91}				
MA-P5	SWD	-2.52	-0.93	2.53	SFI	2.72×10^{-08}	2.72×10^{-08}	0.74	0.40	1×10^{-91}	1.07×10^{-77}
	SWS	-12	-10.08	-0.55	SWD	3.9×10^{-08}	3.9×10^{-08}				
	SFI	-21.4	-4.92	-2.84	Time	1.55×10^{-97}	1.55×10^{-97}				
MA-P10	SWD	-2.78	-1.68	4.48	SFI	1.37×10^{-13}	1.37×10^{-13}	0.77	0.41	1.5×10^{-184}	3.34×10^{-81}
	SWS	-8.7	-8.3	-0.67	SWD	5.01×10^{-09}	5.01×10^{-09}				
	SFI	-21.8	-5.18	-1.97	Time	4.7×10^{-120}	4.7×10^{-120}				
MA-P15	SWD	-2.36	-1.13	-1.47	SFI	4.53×10^{-20}	4.53×10^{-20}	0.793	0.42	1.6×10^{-194}	4.8×10^{-20}
	SWS	-5.73	-3.59	-0.583	SWD	1.62×10^{-09}	1.62×10^{-09}				
	SFI	-26.17	-8.728	-115.1	Time	3.31×10^{-91}	3.31×10^{-91}				
WMA-S1	SWD	-2.31	-2.063	1.57	SFI	6.61×10^{-07}	6.61×10^{-07}	0.73	0.38	3.4×10^{-164}	5.4×10^{-74}
	SWS	-11.64	-25.23	-3.08	SWD	7.46×10^{-06}	7.46×10^{-06}				
WMA-S2	SFI	-23.92	-5.85	-47.51	Time	1.6×10^{-156}	1.6×10^{-156}				
	SWD	0.66	-1.828	-34.2	SFI	6.3×10^{-06}	6.3×10^{-06}	0.71	0.38	8×10^{-155}	1.9×10^{-70}
	SWS	15.84	-38.28	-0.77	SWD	0.32	0.32				
WMA-S3	SFI	-25.36	-6.46	-35.24	Time	0.074	0.074				
	SWD	1.58	1.257	-28.94	SFI	0.934	0.934	0.40	0.37	9.8×10^{-64}	1×10^{-70}
	SWS	41.47	23.3	0.224	SWD	5.1×10^{-65}	5.1×10^{-65}				
EMA-S1	SFI	-14.28	-3.219	-0.73	Time	9.51×10^{-88}	9.51×10^{-88}				
	SWD	-0.93	-0.35	0.08	SFI	0.00014	0.00014	0.703	0.35	1.9×10^{-153}	1.04×10^{-66}
	SWS	-4.04	-3.45	0.032	SWD	2.9×10^{-05}	2.9×10^{-05}				
EMA-S2	SFI	-7.35	0.38	101.54	Time	1.1×10^{-84}	1.1×10^{-84}				
	SWD	-3.78	-2.521	0.73	SFI	0.00067	0.00067	0.69	0.31	9×10^{-148}	1.61×10^{-57}
	SWS	-12.75	-7.638	0	SWD	0.0002	0.0002				
EMA-S3	SFI	-6.303	-1.49	-1.72	Time	8.6×10^{-88}	8.6×10^{-88}				
	SWD	-0.36	-0.16	0.513	SFI	0.00059	0.00059	0.702	0.33	4.2×10^{-153}	6.84×10^{-62}
	SWS	-2.45	-1.7	0.019	SWD	3.7×10^{-05}	3.7×10^{-05}				
	SFI	43.566	43.9	1859.8	Time	12.7	12.7				
HW	SWD	1.57	1.8	96.95	SFI	0.054	0.054	0.716	0.501	1.2×10^{-159}	1.98×10^{-8}
	SWS	90.04	94.63	40.7	SWD	-2.67	-2.67				
	SFI	51.7	51.74	2194.8	Time	11.9	11.9				
ARIMA	SWD	1.61	1.79	94	SFI	0.055	0.055	0.709	0.505	1.7×10^{-156}	1.04×10^{-9}
	SWS	90.01	95.63	40.93	SWD	-2.7	-2.7				

Table 2 shows the derived results of the estimations of RMSE, MAPEand MAD for SFI, SWS and SWD on TD-1 and 2 along with the R^2, Significance-F, coefficient and p-value for SWS w.r.t time, SFI and SWD for all the models (vide 2.II to 2.V) from the forecasted datasets. The length of the dataset are utilized here are ranging as 720-726 for TD-1 and 590-595 for TD-2. Models with optimized performances from the five categories are selected on the basis of their R^2, Significance-F, p-value, MAPE and RMSE. For TD-1, EMA shows the lowest R^2 values among the others. Though WMA-S1 shows lesser R^2 than MA-P5 and MA-P10 but it has shown improved RMSE and MAPE values on TD-1. MAPE value shows few negative results that signify the outliers in the set of data. The best models for TD-2 are EMA-S1 and EMA-S2, which consistently exhibit low RMSE, MAD, and MAPE values across all the parameters. These models also show high R^2 values around 0.33 to 0.42 with statistically significant p-values (<0.05), indicating good model fit and reliability. HW models also perform well, especially in SFI and SWD with low RMSE and MAD. Notably, outliers can skew metrics like MAPE, especially in cases where the dataset contains extreme values. Models like EMA-S1 and EMA-S2, with low MAPE and may be more robust against outliers, as their predictions remain stable even with some irregular values. The ARIMA shows the best R^2 for TD-2. The Significance–F shows that for all the models the calculated value is greater than the critical value of F (2.62) for the 95% confidence interval. The p-values also shows the statistical significance with this. Hence the selected models are as follows viz. MA-P15, WMA-S1, HW and ARIMA for TD-1 and EMA-S1, EMA-S2, HW and ARIMA for TD-2.

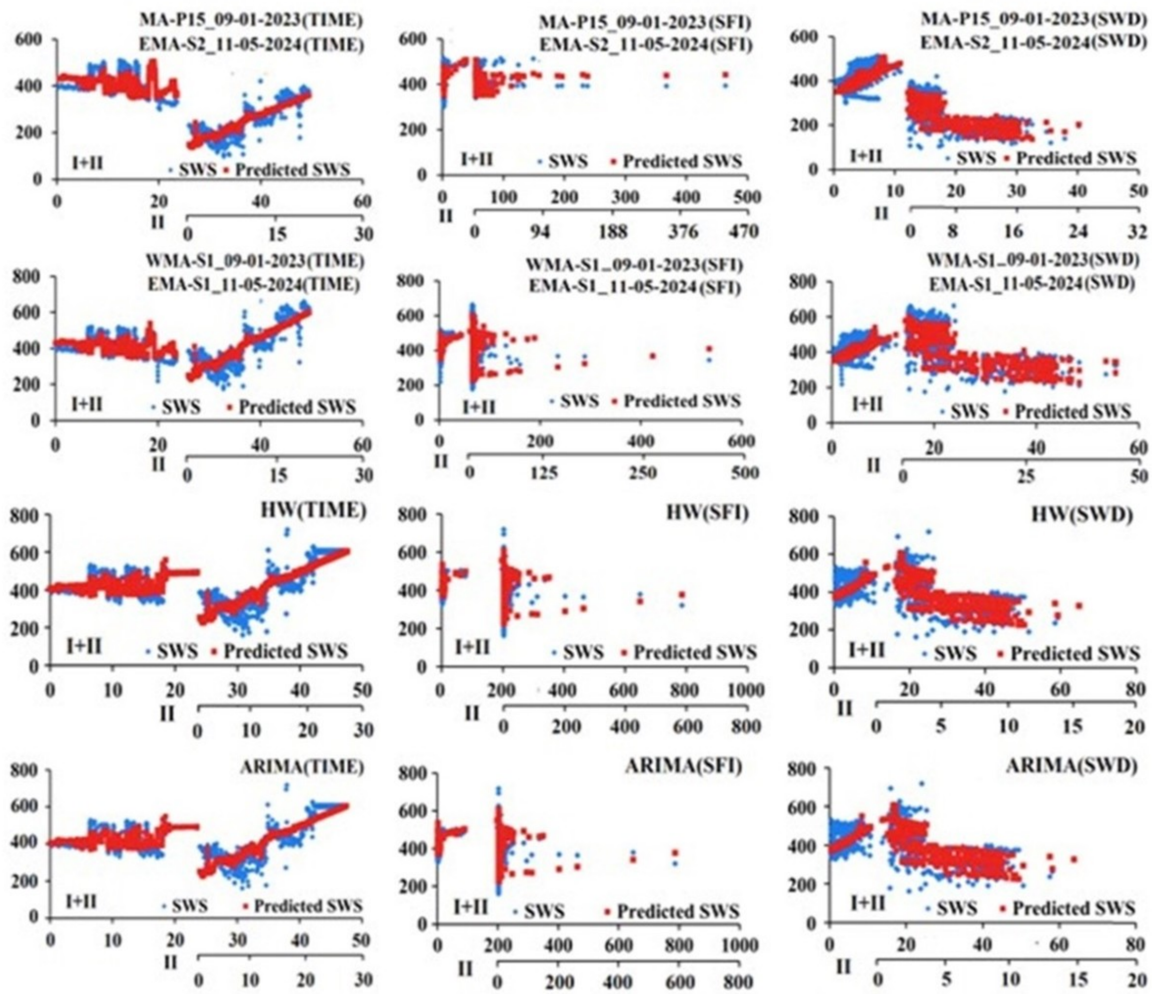

Fig 1. Fitted plot for predicted of SWS (ordinate) and actual SWS (ordinate) w.r.t independent variables [abscissae (TD-1: I and TD-2: II, combined primary abscissae: I+II and secondary abscissae: II)] 1st column: Time, 2nd column: SFI and 3rd column: SWD (estimated through 1st row: MA-P15 and EMA-S2, 2nd row: WMA-S1 and EMA-S1, 3rd row: HW and 4th row: ARIMA models)

Figure 1 shows the fitted plot for predicted and actual SWS w.r.t time, SFI and SWD as the independent variables for those selected models. From the previous part of analysis, the two models HW and ARIMA for TD-1 and 2 out of the four are retained for the remaining part of the analysis. It is considered on the facts that, these selected models show R^2>0.50 and F-values (~10^{-4})

falls in the region greater than the critical value of F-statistics along with their greater prediction ability than the others as more concentrated distributions of the predicted points is found surrounding the best fit (vide Figure 1). The forecasted data set from the selected models from the above part are fitted into the methods (vide 2.VII and 2.VIII) as depicted earlier. The variation of time of occurrences of the SWSs from the range 1 to 4, where each range corresponds to a column w.r.t SFIs from the range 1 to 4 in each frame and the similar distribution of the SWDs in place of SFIs with the same frame wise distribution of the SWSs of the actual, HW and ARIMA predicted data for TD-1 and 2 are shown in Figure 2. Binary logic values of these parameters at the bounds of the abscissae in all the frames as 0 and 1 for the non-occurrences and occurrences respectively are shown.

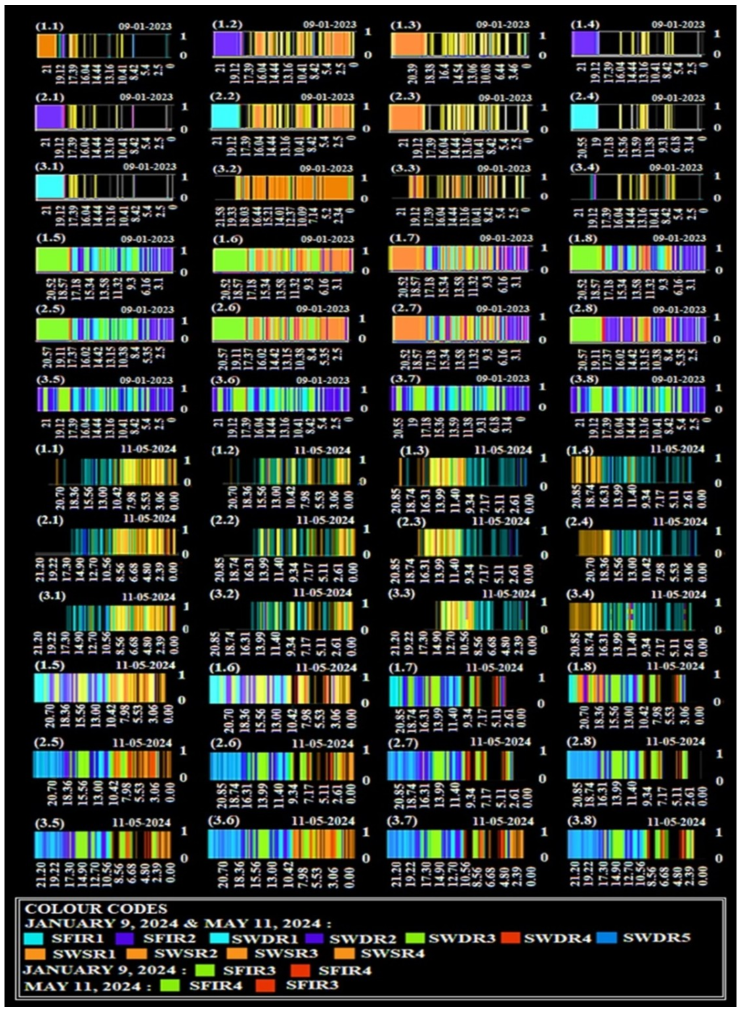

Fig 2. Time spectrum: Variation of time of occurrences (ordinate) of the solar wind speed [range 1 (1st column), range-2 (2nd column), range-3 (3rd column) and range-4 (4th column)] w.r.t solar flare intensity (range- 1 to 4) of the actual data (1st row), HW-predicted (2nd row) and ARIMA-predicted (3rd row) and the variation of time of occurrences (ordinate) of the solar wind speed [range-1 (1st column), range-2 (2nd column), range-3 (3rd column) and range-4 (4th column)] w.r.t solar wind density (range-1 to 5) of the actual data (4th row), HW-predicted (5th row) and ARIMA-predicted (6th row). Binary logic values for the occurrences at the bounds of the abscissae in all the frames (0= non-occurrences and 1 = occurrences) for TD-1 and 2. [Frame number (upper right corner of the each frame)-(Data types. Dependence), Data types: Actual-1, HW (prediction)-2 and ARIMA (prediction)-3; Dependence: 1 to 4- SWSR1 (SFIR1, SFIR2, SFIR3, SFIR4) to SWSR4 (SFIR1, SFIR2, SFIR3, SFIR4) respectively and 5 to 8 - SWSR1 (SWDR1, SWDR2, SWDR3, SWDR4, SWDR5) to SWSR4 (SWDR1, SWDR2, SWDR3, SWDR4, SWDR5) respectively]

Some significant observations from Figure 2 are as follows. SFIs are less frequent in SWSR1 across all datasets. On TD-1 this range includes most B, C and discrete M-class flares, while on TD-2 the maxima of X-class intensity falls in this range. In SWSR2, TD-1 shows C-class flares and background flux of 2.2 Wm^{-2}, while TD-2 includes B and C-class flares along with the background flux of 5.23 Wm^{-2}. SWSR3 on TD-1 contains higher C and most M-class flares. Less delay is seen between the speed attainment in this range and the occurrences of class of flare. On TD-2 it includes moderate-to-higher M-class flares and the edges of X-class flares. SWSR4 shows the maxima of X-class flares on TD-1. On TD-2, it contains mostly B-class flares, a few C-class and mild M-class flares. The half-power maxima's of X-class flares on TD-1 lie in SWSR3, while on TD-2 they lie in SWSR1. The HW model shows higher C-class flares in SWSR3 on TD-1 and in SWSR1 on TD-2. On TD-1, SWSR4 includes flares from ranges 1–4 but excludes X-class. On TD2, X-class flares are absent in the SWSR4. SWD ranges 1, 2, and 4 are distributed in SWSR1 for TD-1, while TD-2 shows ranges 2–4 for SWSR1. SWDR2 dominates the actual dataset on TD-1, HW and ARIMA on TD-2. SWDR4 distribution ascends as Actual, HW and ARIMA on TD-1 and Actual, ARIMA and HW on TD-2. SWDR5 is not observed in any range among all the datasets.

The following are the derived minimized canonical Boolean form of representation of the occurrence functions of SWS w.r.t SFI and SWD (vide 2.IX and 2.X).

For TD-1 from both the models HW and ARIMA:

$f_{SWSR2,SFI} = \neg w \wedge z \wedge \neg y \wedge \neg x$, $f_{SWSR3,SFI} = w \wedge \neg z \wedge \neg y$, $f_{SWSR4,SFI} = w \wedge z \wedge (\neg y \vee x)$.
$f_{SWSR1,SWD} = \neg d \wedge \neg c \wedge (\neg b \vee a)$, $f_{SWSR2,SWD} = \neg d \wedge c \wedge (\neg b \vee \neg a)$,
$f_{SWSR3,SWD} = d \wedge \neg c$, $f_{SWSR4,SWD} = d \wedge c \wedge (\neg b \vee \neg a)$.

For TD-2 from the model ARIMA:

$f_{SWSR1,SFI} = \neg w \wedge \neg z \wedge \neg y$, $f_{SWSR2,SFI} = \neg w \wedge z \wedge \neg y \wedge x$, $f_{SWSR3,SFI} = w \wedge z \wedge \neg y \wedge \neg x$, $f_{SWSR4,SFI} = w \wedge \neg z$. $f_{SWSR1,SWD} = \neg d \wedge \neg c$, $f_{SWSR2,SWD} = \neg d \wedge c$,
$f_{SWSR3,SWD} = d \wedge c \wedge (a \vee \neg b)$, $f_{SWSR4,SWD} = d \wedge \neg c \wedge (a \vee \neg b)$.

For TD-2 from the model HW:

$f_{SWSR1,SFI} = \neg w \wedge \neg z \wedge (\neg x \vee \neg y)$, $f_{SWSR2,SFI} = \neg w \wedge z \wedge \neg y$,
$f_{SWSR3,SFI} = w \wedge z \wedge \neg y \wedge \neg x$, $f_{SWSR4,SFI} = w \wedge \neg z$.
$f_{SWSR1,SWD} = \neg d \wedge \neg c$, $f_{SWSR2,SWD} = \neg d \wedge c$, $f_{SWSR3,SWD} = d \wedge c \wedge (a \vee \neg b)$,
$f_{SWSR4,SWD} = d \wedge \neg c$.

Generalized Sum of Products (SOP) expression for the feature extraction involves the Boolean functions for each combination of variables (SWS with SFI and SWD) for 5000 data points (vide 2.X). The generated SOP expression for each set of data points is represented as follows Generalized occurrence function of w and z variables symbolizing SWS dependence on SFI as,

$$f_{generalised} = \sum_{j=1}^{5000} \left[\left((\neg w_j \wedge \neg z_j \wedge \neg y_j) \vee (\neg w_j \wedge z_j \wedge \neg y_j) \vee (w_j \wedge z_j \wedge \neg y_j) \vee (w_j \wedge \neg z_j) \right) \right]$$

Generalized occurrence function of d and c variables symbolizing SWS dependence on SWD as,

$$f_{generalised} = \sum_{j=1}^{5000} \left[\left((\neg d_j \wedge \neg c_j) \vee (\neg d_j \wedge c_j) \vee (d_j \wedge c_j \wedge (a_j \vee b_j)) \vee (d_j \wedge \neg c_j) \wedge ((a_j \vee \neg b_j)) \right) \right]$$

The final form of common expression for occurrence function of extracted features involving both set of variables of w,z (SWS-SFI) and c,d (SWS-SWD) as,

$$f_{generalised} = \sum_{j=1}^{5000} \left[\left(\prod_{i \in \{w,z,y\}} (\neg w_j \wedge \neg z_j \wedge \neg y_j) \vee \prod_{i \in \{d,c\}} (\neg d_j \wedge c_j) \right) \right]$$

This captures all the combinations of the Boolean functions for each data point and aggregates these terms for all data points. This feature extraction mechanism considers all possible inputs for SFI and SWD following the constraints and conditions specified.

Table 3. Critical estimations of the matrices and parameters

Estimations	TD-1				TD-2			
R^2	I	0.37 ± 0.055	II	0.503 ± 0.0002	I	0.69 ± 0.2	II	0.7125± 0.0035
RMSE (SWS)		51.39		|+9.52 ± 9.66| #		51.38		|-4.94±3.811|#
RMSE (SWD)		1.91		|+0.26 ± 0.36|#		3.33		|0.08∓0.147|#

Continued on next page

Table 3 continued

RMSE (SFI)				9.33		$\|+4.32 \pm 12.23\|$#		28.06		$\|+0.86 \mp 4.325\|$#
SWS[1]	hrs.	1		62.51		19.8$\pm$0.5016		36.54		36.705 $\mp$ 0.317
		5		143.9		19.805$\pm$0.535		32.2		63.66 $\pm$ 0.00
		10		150.61		50.54$\pm$0.05		31.4		29.81 $\pm$ 0.00
		15		152.17		69.17$\pm$0.23		31.4		29.81 $\pm$ 0.00
		18		154.02		90.34$\pm$0.12		109.44		103.481 $\mp$ 0.19
SWD[1]	hrs.	1	III	1.98	IV	0.302 $\pm$0.0029	III	9.23	IV	9.066 $\pm$ 0.016
		5		3.82		0.308$\pm$0.0027		9.47		9.72 $\pm$ 0.00
		10		3.47		0.97$\pm$0.0029		9.24		9.26 $\pm$ 0.0056
		15		3.28		1.041$\pm$0.004		9.24		9.26 $\pm$ 0.0056
		18		3.1		1.21$\pm$0.008		9.447		9.46 $\pm$ 0.00
SWS[2]				69.23		69.23$\pm$ 0.00		$\|$-19.19$\|$		$\|$-33.1$\mp$5.23$|$
SWD[2]				7.407		7.407$\pm$ 0.00		$\|$-3.02$\|$		$\|$-1.8$\mp$-1.4$\|$

Overall models estimated I: excluding, II: including HW and ARIMA, # : deviation from overall models, [1]after- flare, estimated on III: actual data, IV: HW and ARIMA forecasted data, [2]before- flare

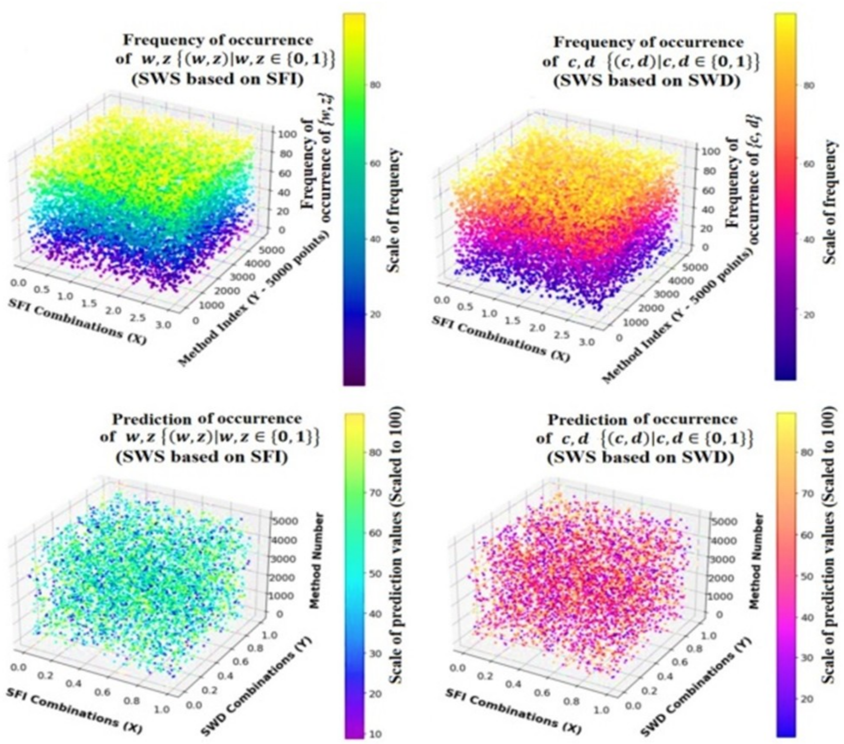

Fig 3. 3-D visualization of feature extraction (upper half): occurrence of w,z frequencies (Z-axis) for combinations of SFI (X-axis) (Viridis colour map- left) occurrence of d,c frequencies (Z-axis) (Plasma colour map-right) and of SWD (Y-axis) across Method Index (X-axis). 3-D visualization ofprediction Random Forest Model (lower half): predictions of w,z as Weighting Zone (Viridis colour map-left) and of d,c as Decision Criteria (Plasma colour map-right) based on SFI and SWD. Colour intensity represents frequency and normalized prediction values

In the earlier studies, various findings in context to the solar wind speed are observed. Out of those few relevant observations are discussed in context to the present work. The study carried out by Upendran et al. [12] in 2020 using spectral data, a best fit correlation of 0.55$\pm$0.03 with the observed data was achieved. Although the high value of activation function was achieved at the coronal holes but that was applicable to both the high and low speed winds that leads to a certain uncertainty at the time of interpretation. Another interesting forecasting scheme was predicted by Raju et al. [13] in 2021, where the SWSs with an RMSE (kms^{-1}) of 76.3$\pm$1.87 and an overall correlation coefficient of 0.57$\pm$0.02 by processing the solar dataset over the year 2018 were shown. That scheme significantly outperformed the contemporary benchmark models. A critical analysis was implemented by Camporeale et al. [14] in 2015 by applying Gaussian classifier and had showed a median accuracy of greater than of 90% over all the categories of wind speed. Even using a small set of data for training that scheme was able to classify the events as undecided and also found the solar wind baselines. In the work of Li et al. [15] in 2020 using the KNN classifier model obtained the highest overall classification accuracy of 92.8%. Although the correlation coefficients are estimated by considering all the models in this work which are falling apart from the previous studies but the coefficient for selected models are interestingly matching with them, especially for TD-2. At the incidence of X-class solar flare radiation intensity was captured at the L1-Lagrangian point by the instrument AIA-131Å of the SDO (ESA) and the simultaneous SWD value were 509.849 kms^{-1} and 336.96 kms^{-1} forTD-1 and 2 respectively. The radiation takes the time to reach the point (1.5×10^{11} meter distant from the Sun) is about 8.33 minutes, as it is traversed by the speed of light ($2.99\times10^{8} ms^{-1}$). The SW consists of the sea of protons and other elementary particles and having finite masses. So if there would be any backward extrapolation of the velocity of the solar wind from the values as received at the Lagrangian point, then the time taken by the wind to reach the point would be unexpectedly large. Therefore, it is evident that there should be a certain delay in the wind particles w.r.t the radiation intensity to reach that point. Interestingly as per the literatures, the delay in solar wind particles to reach the 1 AU distance from the solar corona does not cross a specific limit of approximately 4.34 days at the speed of 400 kms^{-1} and for the 60º of solar rotation [24]. In this respect, solar wind velocity may be considered to be entangled with the propagation velocity (in phase) of the magnetic field of the interplanetary space as generated from CMEs in connection to the solar flares along with the consideration of the magnetic field switch-backing phenomenon as a limiting factor as suggested by the earlier studies [16,25,26]. Although from the actual data, it is observed (vide Table 3) that there is a tendency of increase in the value of solar wind speed especially after the completion phase of the solar flare but this is not sufficient proof to recommend that the solar wind velocity may act as an indicator to the perturbation happened. Hence different prediction methodologies for this problem are adopted and most optimal of them are selected from which it is clearly found (vide Table 3) that there is a certain enhancement in the SWS for a sustainable period. That point towards a definite relationship between intensity and wind density variation at the Lagrangian point. The frequency plot in Figure 3 provides a detailed visualization of the relationship between SWS, SWD, and SFI, illustrating their complex, nonlinear interaction. This 3-D scatter plot uses the X-axis to represent combinations of SFI values, the Y-axis to represent the Method Index and the Z-axis to depict the frequency of occurrences. With dense clustering of data points, it's indicating a significant interaction between these parameters during the flare events. The prediction plot in Figure 3 further refines this view by showing the predicted occurrences of SWS and SWD based on given values of SFI and SWD. Predictions are scaled between 0 and 100. Statistical analysis of $\{|(w,z)\ , z\in\{0,1\}\}$ and $\{|(c,d)\ , d\in\{0,1\}\}$ for the frequency of occurrences shows that mean, median, mode and standard deviation as 49.90∓0.26, 49.5∓0.5, 31.5∓1.5 and 28.45∓0.05 respectively. Random Forest Model shows R^2-score for the prediction of w,z and d,c as 0.8514∓0.0012. Prediction accuracy metrics for w,z and d,c come as MAE:9.38$\pm$0.01 and RMSE:11.085$\pm$0.035 respectively. The plot reveals that SWS and SWD do not follow simple linear patterns but instead exhibit variable behavior influenced by SFI. The clustering of points suggests that SWS is more responsive to SFI fluctuations, particularly during high-energy flare events. The w,z plot, representing SWS prediction, reveals that SWS is a reliable and dynamic indicator of SFI, with high values corresponding to stronger flare events. The d,c plot, suggests that SWD also responds to SFI but to a lesser degree. The statistical analysis confirms these findings, showing a balanced frequency distribution with a moderate variability for SWS's relationship with both SFI and SWD. The Random Forest model, with high predictive accuracy (R^2=0.8514) further supports the significant role of SWS as a reliable predictor of solar flare events.

Figure 4 shows the full day variation of Dst and K_p indices with the SFI and SWS values estimated from all three data sets. From the figure it is noted that $R^2_{Dst,SFI-\ Kp,SFI}$=0.0128 on TD1 and |-0.0465| on TD2. Therefore Dst is more correlated with SF than K_p on TD-1 and the opposite is noted on TD-2. Correlations between the Dst-SFI and SWS ($R^2_{Dst,SFI-\ Dst,SWS}$=0.012: TD-1 and 0.3312: TD-2) is noted higher for TD-2. The ease of geomagnetic storm estimation can be done through values of Dst and to some extent by the K_p index. Thus observations of unsettled field during the passage of the giant X- class at 18.5 UT on TD-1 at 1.22 UT on TD-2 are expected. Here the maximum increase in K_p value is observed during 15-20 UT for TD-1 and 12.10-15.10 UT for TD-2. Noticeable observation for the ring current in this regard is obtained during 10-20 hrs. on TD-1 and 2-3 hrs. on TD-2, where 18 nT on TD-1 (with zero crossing) and |-412| nT on TD-2 (no zero crossing) change in Dst value is observed. Although the period of disturbances in lower ionospheric ring current is in agreement with the time of enhanced SWS value on and after the passage of X class-flare for TD-1 but on TD-2 some delay (transient state response: ~0.78 hr. and steady state response: ~1.28 hr.) is observed between the time of peak of the X-class flare and the maximum value of ring current.

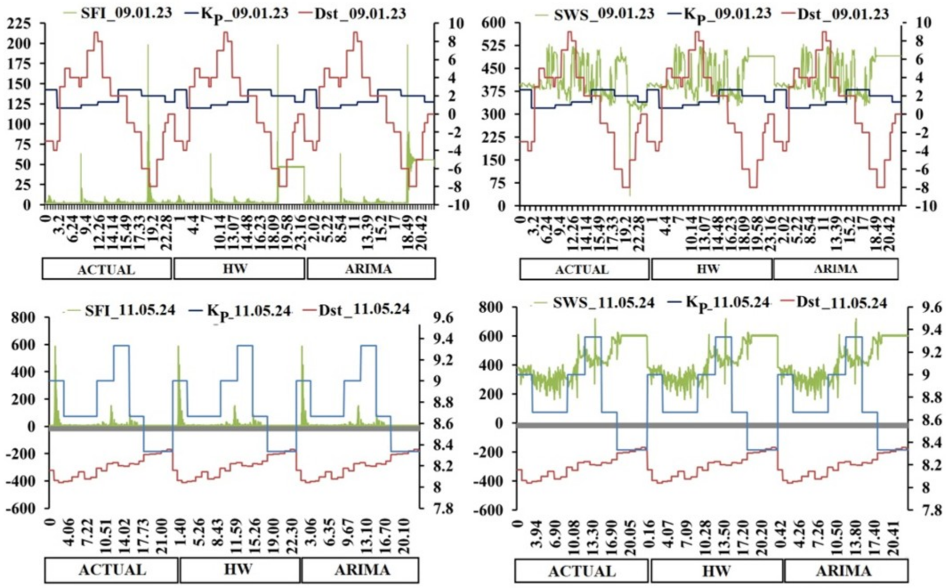

Fig 4. Whole day variation of K_p and Dst indices (left handed ordinate) w.r.t solar flare intensity (upper: TD-1and lower: TD-2 right panel in right handed ordinate) and solar wind speed (upper: TD-1 and lower: TD-2 left panel in right handed ordinate) for actual, HW and ARIMA predicted data

5 Conclusions

By generating 5000 random Boolean expressions based on extracted features, the study creates a diverse dataset that reflects a wide range of potential scenarios. Therefore systematic to explore the parameter space and identify critical patterns in the data. The 3-D scatter plots highlight the density and clustering of specific parameter combinations. By applying the Random Forest method to predict the combinations of variable based on this dataset, the study demonstrates the model's ability to uncover nonlinear relationships and provide reliable predictions. Random Forest, with its ensemble learning approach, ensures robust and accurate predictions even in the presence of noise and complex interactions become significant for this analysis. The funnel like structure of prediction starting from the fundamental regression models to advanced machine learning models with the introduction of newly designed binary prediction model along with the built-in-physics makes this analysis a model one. Although this analysis was carried out for two flares of X-class intensity approximately at five month intervals one in winter (evening dust) and another in the monsoonal season (night time) yields valuable extracted features when considered together. Again it is noteworthy to mention that this study shows few significant validations and anomalies in the flare characteristics between winter time and monsoonal occurrences. There would be a more accurate prediction of the speed variation if the analysis by the inclusion of large number flares and CMEs with several categories for a long period and distributed on seasonal and geo-spatial basis is carried out.

Acknowledgements

The authors sincerely acknowledge the National Aeronautics and Space Administration (NASA) and National Oceanic and Atmospheric Administration (NOAA), University of Kyoto and GFZ Potsdam for partly using their archived relevant data. Authors also acknowledge JIS University for constant support and encouragement for this work.

References

1) Gopalswamy N. History and development of coronal mass ejections as a key player solar terrestrial relationship. *Geoscience Letters*. 2016;3. Available from: https://geoscienceletters.springeropen.com/articles/10.1186/s40562-016-0039-2.
2) Gombosi TI, Van Der Holst B, Manchester WB, Sokolov IV. Extended MHD modeling of the steady solar corona and the solar wind. *Living Reviews in Solar Physics*. 2018;15(4). Available from: https://link.springer.com/article/10.1007/s41116-018-0014-4#:~:text=The%20main%20assumption%20in%20this,plasma%20flow%20considerably%20slows%20down.
3) Curto JJ. Geomagnetic solar flare effects: A review. *Journal of Space Weather and Space Climate*. 2020;10(27):1–15. Available from: https://www.swsc-journal.org/articles/swsc/full_html/2020/01/swsc190079/swsc190079.html.

4) Gopalswamy N. The Sun and Space Weather. *Atmosphere*. 2022;13(11). Available from: https://www.mdpi.com/2073-4433/13/11/1781.
5) Neugebauer M, Steiger RV. The solar wind. In: Bleeker JAM, Geiss J, Huber MCE, editors. The Century of Space Science. Springer. 2001;p. 1115–1140. Available from: https://link.springer.com/chapter/10.1007/978-94-010-0320-9_47.
6) Katsavrias C, Papadema PP, Moussas X, X. Wavelet Analysis on Solar Wind Parameters and Geomagnetic Indices. *Solar Physics*. 2012;280(2). Available from: https://www.researchgate.net/publication/224934397_Wavelet_Analysis_on_Solar_Wind_Parameters_and_Geomagnetic_Indices.
7) Ribas AG. Cone ensemble modeling of an Earth-directed ICME: Comparison with in-situ observations by Solar Orbiter and WIND at L1. 2023. Available from: https://uu.diva-portal.org/smash/get/diva2:1778148/FULLTEXT01.pdf.
8) Johnson R, Boubrahimi SF, Bahri O, Hamdi SM. Physics-Informed Neural Networks for Solar Wind Prediction. In: JJ R, B K, editors. Pattern Recognition, Computer Vision, and Image Processing;vol. 13645. Springer. 2023;p. 273–286. Available from: https://link.springer.com/chapter/10.1007/978-3-031-37731-0_21.
9) Son J, Sung SK, Moon YJ, Lee H, Jeong HJ. Three-day Forecasting of Solar Wind Speed Using SDO/AIA Extreme-ultraviolet Images by a Deep-learning Model. *The Astrophysical Journal Supplement Series*. 2023;267(45). Available from: https://iopscience.iop.org/article/10.3847/1538-4365/ace59a.
10) Angryk RA, Martens PC, Aydin B, Kempton D, Mahajan SS, Basodi S, et al. Multivariate time series dataset for space weather data analytics. *Scientific Data*. 2020;7(227). Available from: https://www.nature.com/articles/s41597-020-0548-x.
11) Bailey RL, Möstl C, Reiss MA, Weiss AJ, Amerstorfer UV, Amerstorfer T, et al. Prediction of Dst during solar minimum using in situ measurements at L5. *Space Weather*. 2020;18(5). Available from: https://doi.org/10.1029/2019SW002424.
12) Upendran V, Cheung MC, Hanasoge S, Krishnamurthi G. Solar wind prediction using deep learning. *Space Weather*. 2020;18(9). Available from: https://doi.org/10.1029/2020SW002478.
13) Raju H, Das S. CNN-Based Deep Learning Model for Solar Wind Forecasting. *Solar Physics*. 2021;296(134). Available from: https://link.springer.com/article/10.1007/s11207-021-01874-6.
14) Camporeale E, Carè A, Borovsky JE. Classification of solar wind with machine learning. *Journal of Geophysical Research: Space Physics*. 2017;122(11):910–920. Available from: https://ui.adsabs.harvard.edu/abs/2017JGRA..12210910C/abstract#:~:text=We%20present%20a%20four%2Dcategory,and%20sector%20reversal%20origin%20plasma.
15) Li H, Wang C, Tu C, Xu F. Machine learning approach for solar wind categorization. *Earth and Space Science*. 2020;7(5). Available from: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019EA000997.
16) Banerjee A, Bhattacharya R. On the Transient Response of the Troposphere and Ionosphere during Annular Solar Eclipse Using Radio Signal Analysis. *Geomagnetism and Aeronomy*. 2023;62:S142–S158. Available from: https://link.springer.com/article/10.1134/S0016793222100048.
17) Box GEP, Jenkins GM, Reinsel GC, Ljung GM. Time Series Analysis: Forecasting and Control. 5th ed. and others, editor;Willy Publications. 2015.
18) Plenos MCF. Time Series Forecasting Using Holt-Winters Exponential Smoothing: Application to Abaca Fiber Data. *Scientific Journal Warsaw University of Life Sciences*. 2022;22(2):17–29. Available from: https://www.researchgate.net/publication/364511952_Time_Series_Forecasting_Using_Holt-Winters_Exponential_Smoothing_Application_to_Abaca_Fiber_Data.
19) Abotaleb M, Makarovskikh T, Yonar A, Badr A, Mishra P, Williams AJ, et al. State of the art in wind speed in England using BATS, TBATS, Holt's Linear and ARIMA model: Wind speed in England using BATS, TBATS, Holt's Linear and ARIMA model. *Mausam*. 2022;73(1):129–138. Available from: https://mausamjournal.imd.gov.in/index.php/MAUSAM/article/view/598.
20) Buyuksahin UC, Ertekin S. Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. *Neurocomputing*. 2019;361:151–163. Available from: https://doi.org/10.1016/j.neucom.2019.05.099.
21) Guerrero DCB, Mora SBS, MJIB. Statistical Analysis of the Global Solar Radiation in Cucuta using the ANOVA Model. *Revista Tecnura*. 2021;25(70):16–31. Available from: https://www.redalyc.org/journal/2570/257070315002/html/.
22) The Solar FAQ. . Available from: https://hesperia.gsfc.nasa.gov/rhessi3/public-outreach/the-solar-faq/index.html.
23) Breiman L. Random Forests. *Machine Learning*. 2001;45:5–32. Available from: https://link.springer.com/article/10.1023/A:1010933404324.
24) Suess ST, Tsurutani BT. SOLAR SYSTEM/SUN, ATMOSPHERES, EVOLUTION OF ATMOSPHERES | Solar Winds. *Encyclopedia of Atmospheric Sciences*. 2015;p. 189–194. Available from: https://colab.ws/articles/10.1016/b978-0-12-382225-3.00373-x.
25) Banerjee A, Bhattacharya R. On the solar dynamics from radio signal communication by recent solar probes. *Journal of Physics: Conference Series*. 2022;2161. Available from: https://dx.doi.org/10.1088/1742-6596/2161/1/012037.
26) Sokolov IV, Zhao L, Gombosi TI. Stream-aligned Magnetohydrodynamics for Solar Wind Simulations. *The Astrophysical Journal*. 2022;926(1):102–102. Available from: https://dx.doi.org/10.3847/1538-4357/ac400f.