Contrib. Astron. Obs. Skalnaté Pleso **53/3**, 156 – 162, (2023)
https://doi.org/10.31577/caosp.2023.53.3.156

Forbush decrease events associated with coronal mass ejections: Classification using machine learning

M.R. Savić, N.B. Veselinović, A.L. Dragić, D.M. Maletić, R.M. Banjanac, D.R. Joković, D.Knežević, M.Travar and V.I. Udovičić

Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia (E-mail: msavic@ipb.ac.rs)

Received: September 27, 2023; Accepted: October 22, 2023

Abstract. In presented work we further explore previously indicated possibility of the existence of two classes of Forbush decrease events, established by the prior analysis of the correlation between the shape of energetic proton fluence spectra and Forbush decrease properties. In an attempt to increase statistical robustness of the analysis and potentially reduce the uncertainties, we have developed an alternative classification procedure that employs machine learning and utilizes space weather parameters as input variables. Based on the overall performance, efficiency and flexibility of different machine learning methods we selected the best performing algorithm and established the optimal boundary value of Forbush decrease intensity to be used for class separation. A subset of good input variables was selected based on their predictive power.

Key words: cosmic rays – Forbush decrease – coronal mass ejection – solar energetic particles

1. Introduction

The dynamic activity of the Sun's coronal magnetic field can give rise to complex space weather events. These events may include solar flares (SFs), coronal mass ejections (CMEs), their interplanetary counterparts known as interplanetary coronal mass ejections (ICMEs), the emission of solar energetic particles (SEPs), and similar phenomena (Kahler, 1992; Yashiro & Gopalswamy, 2008; Gopalswamy, 2022).

One such complex event can produce a number of effects in the heliosphere, one of which is the acceleration of solar wind particles. There is a distinction between particles accelerated by a SF in the lower Sun's atmosphere and those accelerated locally by the CME shock. The later are often referred to as energetic storm particles (ESPs) (Desai & Giacalone, 2016).

Additionally, the passage of a CME can affect the primary cosmic rays (CRs) potentially resulting in a sudden drop in the observed CR flux, followed by a recovery phase that takes place over the several following days. This effect is known as a Forbush decrease (FD) and can be observed by Earth-based CR detectors.

A previous study of the relationship between transient modulations in the fluxes of energetic protons and cosmic rays (measured near and at Earth respectively) indicated an existence of two classes of FD events (Savić et al., 2023). The main objective of this work is to expand this analysis and investigate whether a specific set of space weather (SW) parameters can be successfully used as input parameters for classification. The proposed procedure would aim to separate FD events into classes as indicated by the aforementioned analysis, while increasing the statistical significance and potentially the reliability of the analysis. Additional positive outcome of a successful classification would be the selection of a subset of SW parameters that prove to be good input variables. These variables could then be further used for the prediction of FD magnitudes utilizing some regression algorithm.

2. Motivation

As simultaneous ESP and FD events are very likely a consequence of the passage of an ICME, a relationship between them was assumed. To establish this possible connection, correlation of characteristics of proton fluence spectra and FD parameters was investigated (as described in more detail in Savić et al. (2023)).

The proton fluence spectra were calculated from in situ measurements at L1 by SOHO/ERNE instrument (Torsti et al., 1995), and fitted by a double-power law, as shown for one selected event on Figure 1.

Exponents obtained from these fits were used to parameterize the spectra shape, and some degree of correlation between these exponents and FD magnitudes was established. However, this analysis also indicated a possible existence of two classes of FD events, as illustrated in Figure 2. The plot shows the dependence of the FD magnitude corrected for the magnetospheric effect on one of the proton fluence spectra exponents. The green oval indicates a supposed class of events that exhibit a stronger correlation between these two variables, while the red oval indicates a class of events where this correlation is apparently weaker. One possible way to define the boundary between these two classes could be by introducing a cut on the intensity of the event.

Due to relatively low statistics of events where proton fluence can be reliably determined, one idea for extending this analysis is to try and utilize other space weather parameters in order to increase statistics and more strongly establish the assumed existence of two classes of FD events.

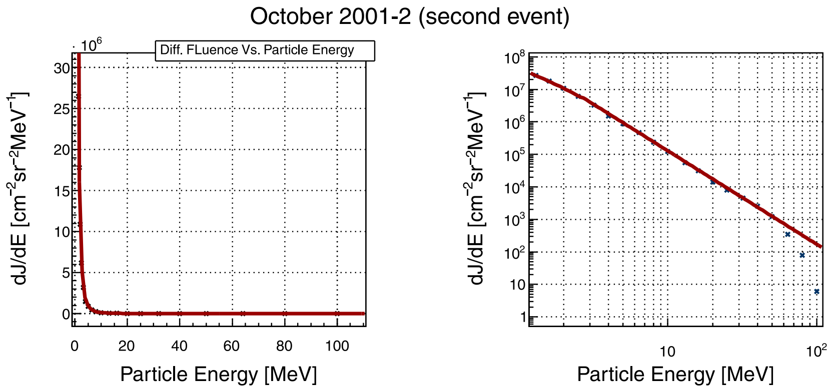

Figure 1. Proton fluence spectra at L1 for one event during October 2001, in linear (left) and logarithmic scale (right).

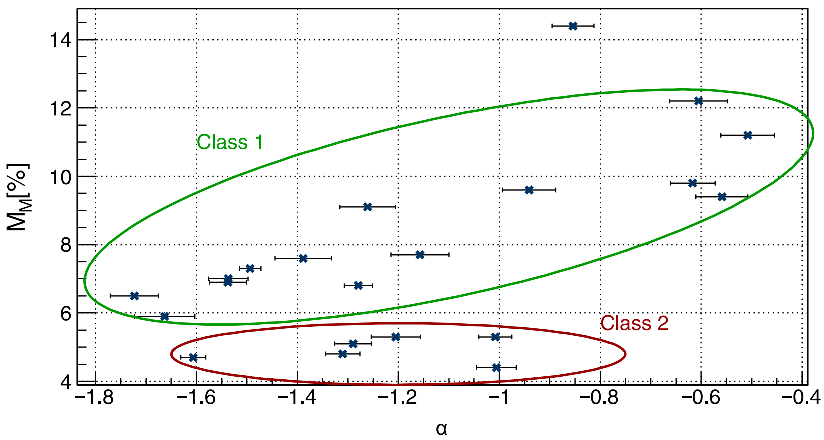

Figure 2. The dependence of the FD magnitude corrected for the magnetospheric effect (M_M) on one of the exponents used to parameterize the proton fluence spectra (α). Two assumed classes of FD events are indicated by the green and red ovals.

3. Methods and Results

IZMIRAN catalogue of Forbush effects (IZMIRAN, 2016) was used as the source of SW related data, as it contains an extensive list of FD events and associated SW parameters. The parameters selected from the IZMIRAN catalogue to be used in the analysis presented here fall into several cathegories: parameters describing the source (Otype, Stype) or the characteristics of the CME (Vmean, CMEwidth); solar wind parameters (Vmax, KTmax, KTmin); parameters describing interplanetary or geomagnetic field (Bzmin, Kpmax, Apmax, Dstmin); and parameters related to the associated solar flare (Xmagn, Sdur, SSN).

Several machine-learning-based classification methods implemented in the TMVA analysis network (Hoecker et al., 2007) were employed in order to establish the optimal FD magnitude for the separation of two classes (boundary criteria mentioned in Section 2), as well as to determine the optimal classification algorithm. Comparing the efficiency of various methods available in the TMVA (shown of Figure 3), it was found that the optimal separation between two classes is achieved with FD magnitude cut set to 6%, as separation efficiency seems to drop-off beyond that for most methods. Support vector machine (SVM) (Cortes & Vapnik, 1995) was identified as the overall best-performing algorithm.

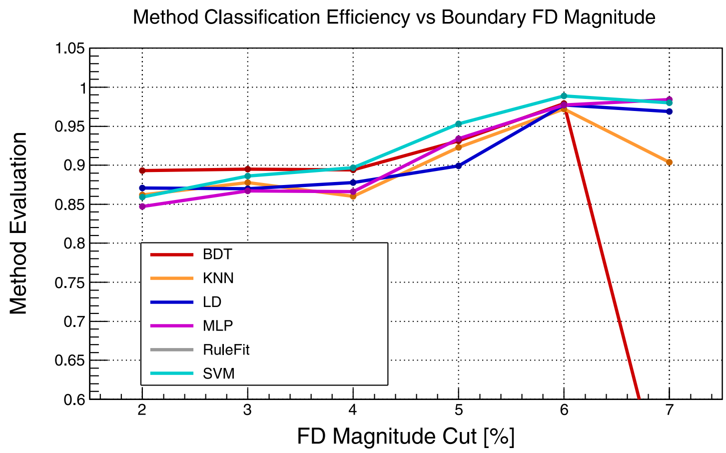

Figure 3. Comparison of the classification efficiency of various TMVA methods dependence on the FD magnitude cut used for class separation.

SVM implementation in the scikit-learn package (Pedregosa et al., 2011) was utilized to identify which of the SW parameters could reliably classify FD events. Third-degree polynomial kernel was found to have the most flexible and efficient performance.

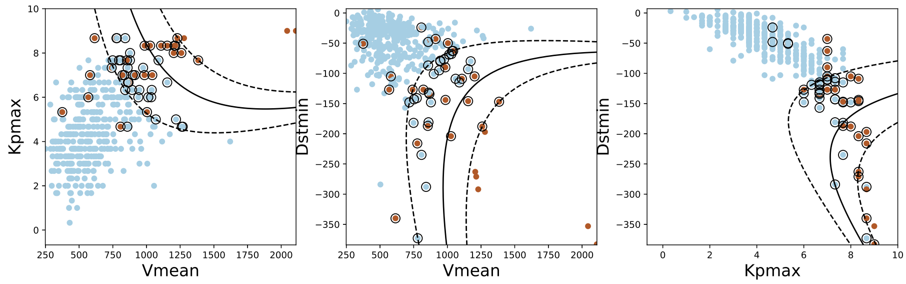

Figure 4. Example of SVM classification using some of SW parameters (mean CME velocity, maximum Kp index and minimal Dst index over the event's duration) that proved to be good input variables for FD classification.

Obtained results appear to confirm the assumption regarding the existence of two classes of FD events. Furthermore, a subset of SW parameters that provide a more reliable classification of FD events was determined. These include mean CME velocity (Vmean) and geomagnetic indices (Kpmax, Apmax, Dstmin), with a possible inclusion of the solar wind speed (Vmax) and minimal hourly component of the interplanetary magnetic field (Bzmin). Decision boundaries between some pairs of mentioned good input variables are showed on Figure 4. Other SW variables proved to be less well suited for classification (as illustrated in Figure 5, for KTmin and KTmax).

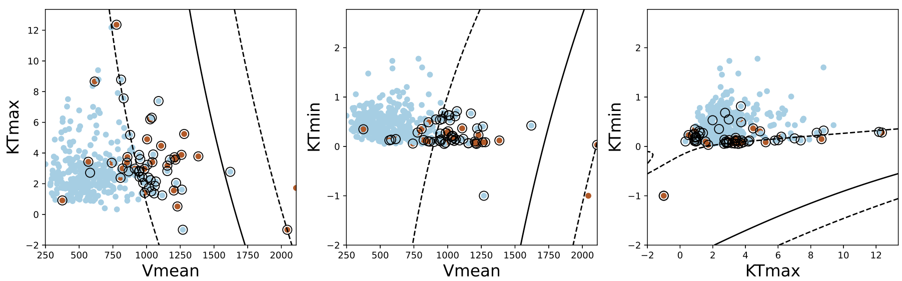

Figure 5. Example of SVM classification using some of SW parameters (KTmax, KTmin) that proved to be less well suited input variables for FD classification.

The identified good variables could prove useful in a potential future extension of the analysis. More specifically, they could serve as an input for a regression procedure that would potentially allow the prediction of FD magnitudes.

This prediction would provide either estimates of FD magnitude as measured by Earth-based detectors or, more importantly, estimates of FD magnitudes corrected for the magnetospheric effect.

4. Conclusions

The potential existence of two classes of FD events was investigated. To increase statistical robustness and reduce uncertainties, the analysis was expanded to include a wider set of various space weather parameters. Machine learning techniques were employed in an attempt to separate FD events into two assumed classes, using a number of selected SW parameters as input variables. We compared the efficiency of different machine learning algorithms, and established the optimal boundary value of FD intensity to be used for class separation. The SVM algorithm was selected for the analysis based on its overall performance, efficiency and flexibility, and used to select a subset of space weather variables to be used for reliable classification of FD events. This subset of good variables variables could prove useful for a future extension of the analysis, where they would provide an input for a regression procedure used to predict FD magnitudes.

Acknowledgements. This work was funded by the Institute of Physics Belgrade, University of Belgrade, through a grant by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia.

We kindly acknowledge the usage of data from the SOHO experiment, as well as data from the catalogue of Forbush effects and interplanetary disturbances provided by the IZMIRAN Space Weather Prediction Center.

References

Cortes, C. & Vapnik, V., Support Vector Networks. 1995, *Machine Learning*, **20**, 273

Desai, M. & Giacalone, J., Large gradual solar energetic particle events. 2016, *Living Reviews in Solar Physics*, **13**, 3, DOI: 10.1007/s41116-016-0002-5

Gopalswamy, N., The Sun and Space Weather. 2022, *Atmosphere*, **13**, DOI: 10.3390/atmos13111781

Hoecker, A., Speckmayer, P., Stelzer, J., et al., TMVA - Toolkit for Multivariate Data Analysis. 2007, *arXiv e-prints*, physics/0703039, DOI: 10.48550/arXiv.physics/0703039

IZMIRAN. 2016, Space weather prediction center (IZMIRAN), http://spaceweather.izmiran.ru/eng/index.html

Kahler, S. W., Solar flares and coronal mass ejections. 1992, *Annual Review of Astron and Astrophys*, **30**, 113, DOI: 10.1146/annurev.aa.30.090192.000553

Pedregosa, F., Varoquaux, G., Gramfort, A., et al., Scikit-learn: Machine Learning in Python. 2011, *Journal of Machine Learning Research*, **12**, 2825

Savić, M., Veselinović, N., Dragić, A., et al., New insights from cross-correlation studies between solar activity indices and cosmic-ray flux during Forbush decrease events. 2023, *Advances in Space Research*, **71**, 2006, DOI: 10.1016/j.asr.2022.09.057

Torsti, J., Valtonen, E., Lumme, M., et al., Energetic Particle Experiment ERNE. 1995, *Solar Physics*, **162**, 505, DOI: 10.1007/BF00733438

Yashiro, S. & Gopalswamy, N., Statistical relationship between solar flares and coronal mass ejections. 2008, *Proceedings of the International Astronomical Union*, **4**, DOI: 10.1017/S1743921309029342