Analiza i predviđanje utjecaja koroninih izbačaja na svemirske vremenske prilike

Dumbović, Mateja

Doctoral thesis / Disertacija

2015

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:464093

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-28**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb



University of Zagreb

FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

Mateja Dumbović

ANALYSIS AND FORECASTING OF CORONAL MASS EJECTION SPACE WEATHER EFFECTS

DOCTORAL DISSERTATION

Zagreb, 2015.



Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

Mateja Dumbović

ANALIZA I PREDVIĐANJE UTJECAJA KORONINIH IZBAČAJA NA SVEMIRSKE VREMENSKE PRILIKE

DOKTORSKI RAD

Zagreb, 2015.



University of Zagreb

FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

Mateja Dumbović

ANALYSIS AND FORECASTING OF CORONAL MASS EJECTION SPACE WEATHER EFFECTS

DOCTORAL DISSERTATION

Supervisor: dr. sc. Bojan Vršnak

Zagreb, 2015.



Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

Mateja Dumbović

ANALIZA I PREDVIĐANJE UTJECAJA KORONINIH IZBAČAJA NA SVEMIRSKE VREMENSKE PRILIKE

DOKTORSKI RAD

Mentor: dr. sc. Bojan Vršnak

Zagreb, 2015.

Information on the Supervisor

Bojan Vršnak is employed at the Hvar Observatory, Faculty of Geodesy, University of Zagreb since 1981. He graduated theoretical physics in 1980, obtained a Master of Science degree in the field of Atomic and Molecular physics in 1983, and obtained his Ph.D. degree in the field of Astrophysics (Solar Physics) in 1987 at the Faculty of Science in Zagreb. In 2000 he was elected to the permanent position of scientific adviser.

His main research area is solar activity, namely eruptive processes in the solar atmosphere and their influence on the heliosphere and magnetosphere of the Earth. He published over 300 scientific papers (141 in the peer reviewed journals indexed in "Current Contents", CC) which were cited more than 3700 times. He published a graduate-studies textbook "Temelji fizike plazme" (eng. Introduction to plasma physics), three astronomy textbooks for primary and highschool students, written thirteen chapters in various books, and authored and co-authord a number of scientific-popular articles.

He currently gives lectures within a post-graduate studies at the Faculty of Science in Zagreb. His previous educational experience includes lectures within graduate studies at the Faculty of Science in Zagreb and lectures as a guest-professor at the Karl-Franzens University in Graz, as well as supervising a number of graduate, master and PhD thesis. He reviewed more than a hundred scientific papers in various international peer-reviewed journals and a number of foreign research project proposals. He is a member of the Editorial board of the scientific journals "Solar Physics" (Kluwer), "The Scientific World Journal" (Hindawi), "Central European Astrophysical Bulletin", and scientific-popular journals "Čovjek i Svemir" and "Bolid". He is a member of International Astronomical Union (IAU), American Geophysical Union (AGU), European Astronomical Society (EAU), Committee on Space Research (COSPAR), Joint Organization for Solar Observation (JOSO), Community of European Solar Radio Astronomers (CESRA), Croatian Astronomical Society, and Croatian Physical Society.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, dr.sc. Bojan Vršnak, for giving me the opportunity to start my scientific career and teaching me (almost) everything I know.

I would also like to thank all of my other colleagues at Hvar Observatory: Roman, Hrvoje, Domagoj, Davor, Jaša, Tomislav, Nikša and Toni for providing the best possible working environment, in both scientific and humanly aspects.

Special thanks to my colleagues and co-authors of my scientific papers from Zagreb Observatory (Dragan, Darije, Damir, Ivan), University of Graz (Astrid and Manuela), Space Research Institute in Graz (Christian and Tanja), University of Kiel (Bernd), Technical University of Denmark (Susanne and Kristoffer), and Royal Observatory of Belgium (Laure, Luciano and Andy).

I gratefully acknowledge the European Comission FP7 project No. 263252 "COronal Mass Ejections and Solar Energetic Particles" (COMESEP), Croatian Science Foundation project No. 6212 "Solar and Stellar Variability" (SOLSTEL) and Croatian Science Foundations' "Young researchers' career development project - training of new doctoral students" funding program for financially supporting the making of this thesis.

My special appreciation goes to my family. Without their continuing support and love, none of this would be possible.

Abstract

Coronal mass ejections (CMEs) are most powerful eruptions in the solar system. They are driven by the energy explosively released from the coronal magnetic field and are often associated with solar flares, representing a dissipative energy release that causes a wide range of electromagnetic emission at different wavelengths, from radio waves to gamma rays. CMEs have strong impact on space weather - they can cause severe problems in the modern human technology and represent a significant factor in human space-born missions planning. Therefore, they are an important element of space weather forecast, which is based on a numerous ground-based and space-born observations, as well as a variety of modeling and empirical forecast methods. Namely, CMEs drive the most intense geomagnetic storms and largest short-term depressions in galactic cosmic ray (GCR) flux, so called Forbush decreases. Both of these are direct consequences of the near-Earth interplanetary conditions due to CME passage over the Earth. Currently, probabilistic forecast methods turned out to be the most efficient procedure for predicting the geomagnetic storm strength and Forbush decrease magnitude based on the remote solar observations. The presented statistical analysis reveales that both geomagnetic storms and Forbush decreases are stronger for faster and wider CMEs, associated with stronger flares originating closer to the center of the solar disc, especially when they are involved in a CME-CME interaction. Statistical relationships are employed in empirical statistical modeling based on the geometric distribution, which can provide forecast of the CME related geo- and GCR-effectiveness (*i.e.* geomagnetic storm strength and Forbush decrease magnitude). The evaluation reveales that the forecast is less reliable if it is more specific, and gives a relatively good prediction whether or not strongest storms and significant Forbush decreases will occur. The main advantage is in the early warning, based on the input parameters that are not necessarily satellite-dependent. Based on the presented research, two online forecast tools have been developed, available at Hvar Observatory web page. In addition, geomagnetic forecast model has been implemented in the "COMESEP alert system", which is the first fully automatic system for detection of CMEs and solar flares, forecasting the CME arrival as well as their potentially hazardous impact.

Key words: Sun – Space weather – Coronal mass ejections (CME) – Cosmic rays – Forbush decreases – Geomagnetic storms

Prošireni sažetak na hrvatskom jeziku

1. UVOD

Koronini izbačaji i Sunčevi bljeskovi su najsilovitiji eruptivni procesi na Suncu te se nerijetko smatraju glavnim pokretačima svemirskih vremenskih prilika. Praćenjem i predviđanjem svemirskih vremenskih prilika, odnosno stanja u međuplanetarnom prostoru, bliskoj okolici Zemlje te njenoj magnetosferi, ionosferi i termosferi bavi se svemirska prognostika (eng. "space weather"). Iako je to relativno novo područje istraživanja, usko vezano uz razvoj ljudske tehnologije (posebice svemirskih letjelica), može se tvrditi da je njen razvoj započeo davno prije "doba satelita" sa prvim opažanjima Sunčeve aktivnosti. Ljudska tehnologija napredovala je značajno u posljednjem stoljeću te je postala i osjetljivija na Sunčevu aktivnost. Živimo u doba satelita, aviona, elektroenergetskih sustava i svemirskih misija, koje izravno mogu biti pod (negativnim) utjecajem Sunčevih eruptivnih procesa. Stoga je shvaćanje i predviđanje takvih događaja te njihovih učinaka neophodno za moderno društvo.

1.1. Koronini izbačaji

Naziv koronini izbačaji dolazi od engleskog naziva "Coronal mass ejection" (u daljnjem tekstu CME), što je povijesni naziv budući su njihova prva opažanja bila koronagrafima u vidljivom dijelu spektra kao velike količine mase koja je izbačena u međuplanetarni prostor. CME-ovima su često pridruženi Sunčevi bljeskovi te eruptivne prominencije. Sunčevi bljeskovi su disipativni procesi u kojima se oslobađa energija u praktički cijelom spektru elektromagnetskog zračenja - od radiovalnih duljina do gama zračenja. Prominencije čini hladnija i gušća kromsferska plazma, koju unatoč gravitaciji magnetsko polje zadržava u toplijim i rjeđim višim slojevima Sunčeve atmosfere (koroni). Iako ne postoji jedan-na-jedan povezanost između CME-ova, bljeskova i prominencija, široko je prihvaćeno stajalište da su to usko povezane manifestacije jedinstvenog fizikalnog procesa, kojeg pokreću nestabilnosti magnetskog polja. Njihov nastanak opisuje se tzv. "standardnim modelom bljeska".

Prema standardnom modelu bljeska, magnetska arkada eruptira uslijed gubitka ravnoteže, te biva izbačena velikom brzinom u međuplanetarni prostor. Pritom, uslijed "razvlačenja" silnica u okolini arkade, dolazi do njihovog "prespajanja", što u konačnici rezultira preustrojem magnetske strukture. Proces prespajanja silnica uzrokuje impulzivno zagrijavanje plazme te stvaranje čestičnih snopova koji u interakciji s okolinom zrače u gotovo

svim područjima elektromagnetskog spektra (Sunčev bljesak). Unutar magnetske arkade može ostati "zarobljena" plazma nižih slojeva Sunčeve atmosfere, koja je stoga hladnija i gušća od okoline (eruptivna prominencija). Gibanje magnetske arkade pak gomila plazmu ispred sebe, što u koronagrafu vidimo kao nakupinu mase koja se giba u smjeru suprotnom od Sunca (CME). Standardni model bljeska je, uz odgovarajuće reference, detaljnije opisan u poglavlju 1.1.2.

Općenito opažanja podupiru model standardnog bljeska. Iako se opažaju CME-i bez popratnih bljeskova te bljeskovi bez popratnih CME-a, CME-i najveće energije gotovo uvijek su popraćeni snažnim bljeskovima, a vrlo često i eruptivnim prominencijama kao i raznim drugim poremećajima vidljivim u Sunčevoj koroni (detaljnije u poglavlju 1.1.1). Ova opažanja korisna su s aspekta svemirske prognostike, budući da pridruženi Sunčev bljesak može dati dodatne informacije o CME-u koje nisu dostupne iz koronagrafskih opažanja. Naime, budući je u koronagrafskim opažanjima Sunčev disk zasjenjen, nemoguće je odrediti područje na Sunčevom disku iz kojeg je erupcija krenula, što u konačnici otežava i određivanje smjera kretanja CME-a.

Međuplanetarni CME-ovi (eng. Interplanetary coronal mass ejection, ICME) uobičajeno se opažaju u *in situ* mjerenjima kao poremećaji niza parametara Sunčevog vjetra i međuplanetarnog magnetskog polja. Povezivanje *in situ* mjerenja ICME-a s daljinskim opažanjima CME-ova na Suncu nije jednostavan zadatak, budući da uključuje kompleksnu i nedovoljno razjašnjenu kinematičku evoluciju CME-a. Na većim udaljenostima od Sunca ($\approx 20R_{sun}$) propagacija ICME-a pod utjecajem je aerodinamičkog otpora koji prilagođava brzinu CME-a Sunčevom vjetru. Aerodinamični otpor kvalitativno vrlo uspješno opisuje propagaciju ICME-a, međutim kvantitativno je ograničen parametrima CME-a i Sunčevog vjetra, čije je određivanje vrlo zahtjevno i nedovoljno precizno. Kinematiku i propagaciju ICME-a dodatno kompliciraju interakcije dva ili više CME-a, određivanje smjera CME-a te odstupanja od originalnog smjera gibanja, što u konačnici može dovesti do pogrešnog određivanja vremena naleta ICME-a. Detaljniji opis *in situ* i propagacijskih svojstava ICME-a, s odgovarajućim referencama, dan je u poglavlju 1.1.3.

1.1. Utjecaj koroninih izbačaja na svemirske vremenske prilike

Prilikom heliosferske propagacije ICME-ovi interagiraju s magnetskim poljima i nabijenim česticama koje susreću. Međudjelovanje ICME-a sa Zemljinim magnetskim poljem uzrokuje geomagnetske oluje. Ti poremećaji geomagnetskog polja mogu uzrokovati mnoge negativne posljedice na ljudsku tehnologiju. Geomagnetske oluje mogu nastati ukoliko je orijentacija magnetskog polja ICME-a povoljna za magnetsko prespajanje sa geomagnetskim poljem, odnosno ako postoji jaka južna komponenta magnetskog polja. Uslijed magnetskog prespajanja oslobađa se energija te nabijene čestice iz Sunčevog vjetra ulaze duboko u Zemljinu magnetosferu formirajući električne struje u magnetosferi i ionosferi.

Formirane struje uzrokuju lokalne geomagnetske poremećaje koji se kvantificiraju tzv. indeksima geomagnetske aktivnosti, kao što je npr. Dst indeks (eng. Disturbance storm time index), koji mjeri poremećaje horizontalne komponente na dipolnom ekvatoru. ICME-i mogu uzrokovati jake ili slabe geomagnetske oluje, odnosno biti jako ili slabo geo-efektivni, međutim također ne moraju uopće biti geo-efektivni. Geo-efektivnost ICME-a posljedica je magnetskog prespajanja sa geomagnetskim poljem te stoga ovisi o konvektivnom električnom polju $E_y = v \cdot B_s$, gdje je B_s južna komponenta magnetskog polja ICME-a, a v brzina Sunčevog vjetra. Svemirske letjelice u L1 lagrangeovoj točki omogućuju direktna mjerenja B_s i v, međutim samo ≈ 1 sat unaprijed, što uvelike ograničava "vrijeme reagiranja". Budući nas trenutno razumijevanje CME-a i ICME-a ograničava u predviđanju parametara ključnih za određivanje geo-efektivnosti ICME-a, nameće se statistički pristup - pridruživanje svojstava CME-a opaženih tijekom erupcije na Suncu geomagnetskom odzivu na Zemlji. Detaljniji opis nastanka geomagnetskih oluja, indeksa geomagnetske aktivnosti te geo-efektivnosti ICME-ova i CME-ova, uz odgovarajuće reference, nalazi se u poglavlju 1.2.1.

Međudjelovanje ICME-a s galaktičkim kozmičkim zračenjem uzrokuje kratkotrajna smanjenja toka kozmičkog zračenja koja nazivamo Forbushevim smanjenjima. Forbusheva smanjenja mogu biti indikacija prolaska ICME-a kada druga mjerenja nisu dostupna (npr. prije doba satelita) te su zanimljiva s aspekta svemirskih putovanja. Nadalje, velike geomagnetske oluje gotovo su uvijek popraćene i intenzivnim smanjenjem kozmičkog zračenja, stoga predviđanje Forbushevih smanjenja može unaprijediti predviđanje geomagnetskih oluja. Forbusheva smanjenja se mogu mjeriti detektorima na Zemlji (npr. neutron monitorima) i na svemirskim letjelicama u čitavom međuplanetarnom prostoru, kao i na drugim planetima (npr. Marsu). Mjerenja na Zemlji otežana su zbog geomagnetskog polja i međudjelovanja kozmičkog zračenja s atmosferom, međutim prikladnija su za mjerenja vrlo intenzivnih događaja, za razliku od mjerenja svemirskih letjelica (detaljnije u poglavlju 1.2.2). Modulacija kozmičkog zračenja u heliosferi se može opisati transportnom jednadžbom koja opisuje četiri različita doprinosa: (1) difuziju zbog fluktuacija magnetskog polja, (2) drift zbog nehomogenosti magnetskog polja, (3) konvekciju Sunčevim vjetrom te (4) gubitak energije zbog ekspanzije magnetskog polja (odnosno sustava gibanja čestica). Isti fizikalni mehanizmi modulacije primjenjivi su i za opis Forbushevih smanjenja, gdje se u okviru konvektivno-difuzijskog koncepta očekuje ovisnost amplitude smanjenja o magnetskom polju i brzini ICME-a, što je i potvrđeno statističkim studijama. Međutim, slično kao i u slučaju geomagnetskih oluja, predviđanje Forbushevih smanjenja temeljem mjerenja letjelica u L1 lagrangeovoj točki nije dovoljno rano. Nadalje, zbog sličnih ograničenja i ovdje se nameće statistički pristup, odnosno pridruživanje svojstava CME-a opaženih tijekom erupcije na Suncu odzivu kozmičkog zračenja na Zemlji. Detaljniji opis modulacije kozmičkog zračenja ICME-ima i CME-ima, uz odgovarajuće reference, dan je u poglavlju 1.2.2.

2. PODACI I METODE MJERENJA

Za potrebe ovog istraživanja, prikupljen je veliki uzorak događaja. CME-ovima opaženima koronagrafima su pridruženi Sunčevi bljeskovi, a potom geomagnetski i odziv kozmičkog zračenja na Zemlji. Parametri CME-a preuzeti su iz SOHO LASCO CME kataloga (opisanog u poglavlju 1.1.1), dok su parametri Sunčevih bljeskova preuzeti iz NOAA kataloga Sunčevih bljeskova detektiranih u X-zračenju (poveznica dana u poglavlju 2). Promatrani su događaji u vremenskom periodu od 10. Siječnja 1996. do 30. Lipnja 2011. Sunčevi bljeskovi pridruženi su CME-ima automatskom metodom koristeći vremenski i prostorni kriterij (detaljnije opisano u poglavlju 2). Potom je izabrano 211 reprezentativnih CME-bljesak parova, gdje su različite brzine CME-a podjednako zastupljene u čitavom intervalu 400 $\mathrm{km\,s^{-1}} < v <$1500 $\mathrm{km\,s^{-1}}$ i uzeti su svi CME-i s brzinama $v >$ 1500 $\mathrm{km\,s^{-1}}$. Ovakav reprezentativni uzorak biran je umjesto slučajnog uzorka, jer bi slučajan uzorak mogao uključivati vrlo mali broj geomagnetskih oluja, koje su vrlo rijetki događaji (u usporedbi s brojem CME-a).

Koristeći dijagrame koji prikazuju mjerenja kinematike CME-a te vremenski niz mjerenja Dst indeksa, a koji je dostupan u sklopu SOHO LASCO CME kataloga, geomagnetski odziv je pridružen svakom CME-bljesak paru. Pri tom se koristila ekstrapolacija kinematičke krivulje do udaljenosti 214 radijusa Sunca, što je srednja udaljenost Zemlje od Sunca. Time je određeno približno vrijeme dolaska ICME-a na Zemlju. Zbog utjecaja aerodinamičkog otpora i smjera CME-a, geomagnetski odziv je tražen u vremenskom intervalu oko približnog dolaska ICME-a (detalji mjerenja opisani su u poglavlju 2 i prikazani na slici 2.1). Unutar tog vremenskog intervala tražena su smanjenja Dst indeksa, kao pokazatelja geomagnetske oluje te je mjerena amplituda smanjenja Dst indeksa u točki minimuma, gdje je kao referentna točka uzet početak geomagnetske oluje. Amplituda Dst indeksa, Dst, izražena je apsolutnom vrijednošću, dakle poprima pozitivne vrijednosti (iako u vremenskom nizu Dst indeksa prilikom geomagnetske oluje Dst indeks poprima negativne vrijednosti). Ukoliko unutar vremenskog intervala nije pronađena geomagnetska oluja, mjerena je prva izražena varijacija ($Dst >$ 10 nT) najbliža približnom vremenu dolaska ICME-a. Za napomenuti je da je kao relevantna geo-efektivnost korištena vrijednost $Dst >$ 100 nT, dakle događaji koji nisu uzrokovali značajnu geomagnetsku oluju ili uopće nisu stigli do Zemlje se ne smatraju geo-efektivnima.

Za svaki CME u uzorku određen je parametar interakcije, koji opisuje mogućnost interakcije s nekim drugim CME-om. U tu svrhu postavljena su tri kriterija CME-CME interakcije: (1) kinematički kriterij prema kojem dva CME-a mogu interagirati ukoliko im se kinematičke krivulje sijeku, (2) vremenski kriterij prema kojemu dva CME-a mogu interagirati ukoliko eruptiraju u "razumnom" vremenskom razmaku te (3) kriterij izvorišnog područja i širine prema kojem dva CME-a mogu interagirati ukoliko dolaze iz bliskih područja na vidljivom Sunčevom disku i/ili su relativno široki. Bitno je napomenuti da

ukoliko su sva tri kriterija zadovoljena, to ne znači da je zaista došlo do CME-CME interakcije, već da je CME-CME interakcija vrlo izgledna. Uzimajući u obzir ova tri kriterija interakcije, te da oni nisu usko specificirani, parametar interakcije, i, može poprimiti četiri vrijednosti: 1-nema interakcije, 2-interakcija nije izgledna, 3-interakcija je izgledna te 4-interakcija je vrlo izgledna. U slučajevima kada je $i = 3, 4$ u uzorku je zastupljen samo jedan događaj, kojeg karakteriziraju parametri najbržeg interagirajućeg CME-a te mu je pridružena širina najšireg CME-a. Određivanje parametra interakcije detaljno je opisan u poglavlju 2 te slikom 2.2.

Uzorak od 211 CME-bljesak-Dst događaja (u daljnjem tekstu Dst lista) nadopunjen je događajima koji opisuju odziv kozmičkog zračenja na Zemlji. U tu svrhu korištena su mjerenja neutron monitora (NM) na površini Zemlje, korigirana prema utjecaju atmosferskog tlaka i normirana na "mirni" period kada nema velikih promjena u toku kozmičkog zračenja. Na taj način amplitudu Forbushevog smanjenja, FD, moguće je mjeriti u postocima. Odziv je tražen u vremenskom periodu 5 dana prije te 15 dana nakon zabilježene Dst anomalije. Kako bi se smanjio utjecaj dnevnih varijacija toka kozmičkog zračenja uprosječena su mjerenja 3-4 NM stanice (ovisno o dostupnosti podataka) sličnog rigiditeta, ali smještenih na različitim geografskim duljinama (detalji metode nalaze se u poglavlju 1.2.2). Ova metoda ne uklanja dnevne varijacije u potpunosti, stoga je kao relevantni odziv kozmičkog zračenja korištena vrijednost amplitude $FD > 1\%$. Sukladno terminu "geo-efektivnost" za geomagnetski odziv, za odziv kozmičkog zračenja koristit će se termin "GCR-efektivnost" (eng. Galactic cosmic rays, GCR). Za napomenuti je da postoje događaji gdje su dvije uzastopne geomagnetske oluje razlučive jedna od druge, međutim opaža se samo jedno Forbushevo smanjenje. U takvim slučajevima dva su događaja spojena u jedan, kojem je pridružen parametar interakcije $i = 4$ te odgovarajući parametri bržeg, odnosno šireg CME-a (kao što je prethodno opisano). Također, u nekoliko slučajeva nije bilo odgovarajućih podataka za kozmičko zračenje, stoga uzorak CME-bljesak-Dst-FD sadrži 187 događaja (u daljnjem tekstu FD lista, detaljniji opis nalazi se u poglavlju 2).

3. STATISTIČKA ANALIZA

Statistička analiza fokusirana je na specifične CME-bljesak parametre, koji su povezani s geo- i GCR-efektivnošću u prijašnjim studijama različitih autora (detaljnije opisano u poglavljima 1.2.1 i 1.2.2). To su početna brzina CME-a, v, širina CME-a, w, vršna vrijednost intenziteta X-zračenja pridruženog Sunčevog bljeska, f, položaj pridruženog bljeska na Sunčevom disku (udaljenost od centra diska), r, te parametar interakcije, i. Niti jedan od promatranih CME/bljesak parametara ne pokazuje snažnu korelaciju s amplitudom Dst indeksa, Dst, niti amplitudom Forbushevog smanjenja, FD. Stoga je upotrijebljen probabilistički pristup. U tu svrhu korištene su Dst i FD raspodijele, gdje su Dst i FD vrijednosti grupirane u četiri odgovarajuća razreda, koja predstavljaju četiri različite razine geo- tj. GCR-efektivnosti. Dst razredi su: $|Dst| < 100$ nT, 100

nT $< |Dst| <$ 200 nT, 200 nT $< |Dst| <$ 300 nT te $|Dst| >$ 300 nT. FD razredi su: $FD < 1\%$, $1\% < FD < 3\%$, $3\% < FD < 6\%$ te $FD > 6\%$. CME/bljesak parametri su također podijeljeni u razrede. Podjela u razrede za neke je parametre očigledna, budući su diskretni (npr. parametar interakcije). Kontinuirani parametri podijeljeni su u razrede koji otprilike sadrže jednak broj događaja, dakle nisu ekvidistantni. Stoga se podjela CME/bljesak parametara u slučaju analize Dst i FD raspodjela donekle razlikuju. Za svaki razred CME/bljesak parametra načinjena je odgovarajuća Dst tj. FD raspodjela te je određena njena srednja vrijednost. Srednjoj vrijednosti svake raspodjele pridružena je (srednja) vrijednost odgovarajućeg razreda CME/bljesak parametra te je tražena korelacija na ovaj način uprosječenih vrijednosti (detaljnije objašnjenje dano je u poglavlju 3 te na slikama 3.1–3.4). Kao mjera raspršenja unutar pojedinog razreda korištene su standardne devijacije. Također je testirana razina statističke signifikantnosti korištenjem t-testa (eng. two-sample t-test, 2stt), gdje je kao razina signifikantnosti korištena vrijednost 0.05 (95% signifikantnost). Kako bi dodatno potvrdili dobivene rezultate, koristili smo metodu preklapajućih razreda. Naime, koristeći isti uzorak, ustrojena su dva različita skupa razreda (originalni i alterantivni), stoga se pojedini razredi ta dva skupa preklapaju. Kako bi rezultati statističke analize provedene na originalnom skupu bili potkrijepljeni, alternativni skup mora pokazivati iste tj. vrlo slične rezultate. Originalni i alternativni skupovi CME/bljesak razreda dani su u tablicama 3.1 i 3.2 (poglavlje 3).

Statističkom analizom provedenom u poglavljima 3.1 i 3.2 utvrđeno je da postoji određena povezanost promatranih CME/bljesak parametara i Dst, odnosno FD amplitude u skladu sa prijašnjim istraživanjima (opisanima u poglavljima GMS i FD). U oba slučaja (Dst i FD) za sve CME/bljesak parametre zamjećuje se određeni trend koji prate podaci i originalnih i alternativnih razreda. Međutim, standardne devijacije su vrlo velike, što ukazuje na veliko raspršenje podataka i kompleksnu ovisnost Dst odnosno FD o CME/bljesak parametrima. To potvrđuju i rezultati 2stt, gdje je vidljivo da su Dst odnosno FD raspodjele vrijednosno udaljenih razreda signifikantno drukčije, dok za bliske razrede to ne mora biti slučaj. Na temelju provedene analize zaključeno je da su CME-i koji imaju veću početnu brzinu, koji su širi, čiji pridruženi bljeskovi imaju veću vršnu vrijednost intenziteta X-zračenja i bliže su centru vidljivog diska te kod kojih je izglednija CME-CME interakcija jače geo- i GCR-efektivni. Nadalje, u svrhu predviđanja geo- tj. GCR-efektivnosti opaženog CME-a povezanost CME/bljesak parametara sa Dst, odnosno FD amplitudom je kvantificirana krivuljama koje najbolje opisuju trend podataka (slike 3.1–3.4).

4. EMPIRIJSKI STATISTIČKI MODELI

Rezultati statističke analize iskorišteni su za izradu modela raspodjele vjerojatnosti geo- i GCR-efektivnosti CME-a kojemu je pridružen set CME/bljesak parametara (v, w, r, f, i). Raspodjele Dst i FD amplituda (Dst i FD), prikazane u poglavljima 3.1 i 3.2 vrlo su

asimetrične i brzo padajuće, stoga je za njihovu matematičku rekonstrukciju korištena geometrijska raspodjela. Budući su promatrane Dst raspodjele asimetričnije nego FD raspodijele, modeli se donekle razlikuju - za konstrukciju Dst raspodjele korištena je regularna, a za FD raspodjelu pomaknuta geometrijska raspodjela (jednadžbe 4.1 i 4.2). Obje raspodjele jednostavno je konstruirati ukoliko je poznata srednja vrijednost raspodjele (jednadžbe 4.3 i 4.4), koju možemo ocijeniti na temelju relacija dobivenih statističkom analizom (slike 3.1–3.4). Budući je regularna geometrijska raspodjela definirana za razrede $k = 1, 2, 3, 4, ...$ potrebno je pridruživanje razreda $k \longleftrightarrow Dst$: $k = 1 \longleftrightarrow Dst < 100$ nT, $k = 2 \longleftrightarrow 100$ nT$< Dst < 200$ nT, $k = 3 \longleftrightarrow 200$ nT$< Dst < 300$ nT, $k = 4 \longleftrightarrow Dst > 300$ nT. Slično, razredi FD pridruženi su različitim vrijednostima k za pomaknutu geometrijsku raspodjelu: $k = 0 \longleftrightarrow FD < 1\%$, $k = 1 \longleftrightarrow 1\% < FD < 3\%$, $k = 2 \longleftrightarrow 3\% < FD < 6\%$, $k = 3 \longleftrightarrow FD > 6\%$. Raspodjele Dst i FD su modelirane i uspoređene sa stvarnim raspodjelama. Uočeno je da pomaknuta geometrijska raspodjela dobro opisuje stvarnu FD raspodjelu. Istovremeno, postoje odstupanja regularne geometrijske raspodjele od Dst raspodjele, stoga su za Dst raspodjelu uvedene dodatne korekcije (detaljno opisano u poglavlju 4.1).

Jedinstvena Dst/FD raspodjela konstruirana je pomoću raspodjela vjerojatnosti za pojedini CME/bljesak parametar (jednadžba 4.8), uz pretpostavke da se CME/bljesak parametri međusobno ne isključuju te da su nezavisni jedan od drugog. Iako posljednja pretpostavka nije sasvim točna (opisano u poglavlju 1.1.1), značajno pojednostavljuje postupak. U oba slučaja uočavamo da su raspodjele vrlo asimetrične te da je najveća vjerojatnost da CME neće biti niti geo- niti GCR-efektivan, što je i očekivano s obzirom na stvarnu raspodjelu geo- i GCR- efektivnosti korištenog uzorka. Stoga, kako bi se raspodjela vjerojatnosti upotrijebila za predviđanje Dst odnosno FD amplitude, potrebno je naći granične vrijednosti koje mogu definirati određenu geo- tj. GCR-efektivnost. U tu svrhu, za svaki događaj s Dst odnosno FD liste, prema CME/bljesak parametrima, izračunata je raspodjela vjerojatnosti tj. dobivene su relativne frekvencije $F_r(k)$ koje odgovaraju četirima različitim vrijednostima razreda k. Svaka relativna frekvencija $F_r(k)$ prikazana je u dijagramu za odgovarajući opaženi Dst odnosno FD razred (slike 4.3 i 4.6), gdje su Dst i FD razredi iskazani pomoću vrijednosti k (kao što je objašnjeno prethodno). Budući su k razredi diskretni, različite vrijednosti $F_r(k)$ raspršene su duž linija konstantnog k. Gustoća raspršenih $F_r(k)$ prikazana je korištenjem percentila te je korištena kao smjernica za određivanje graničnih vrijednosti koje odvajaju različite razine geo- te GCR-efektivnosti (detaljno objašnjeno u poglavljima 4.1 i 4.2). Zbog izrazito malog broja najintenzivnijih geomagnetskih oluja u uzorku, pokazalo se nemogućim odvojiti posljednja dva razreda geo-efektivnosti, stoga su ta dva razreda spojena u jedan. Primjenom uvjeta graničnih vrijednosti na Dst i FD raspodjelu moguće je dobiti procjenu specifičnog razreda geo- tj. GCR-efektivnosti opaženog CME-a s pridruženim bljeskom, kao što je demonstrirano primjerima u poglavljima 4.1 i 4.2.

5. EVALUACIJA MODELA

Modeli su evaluirani korištenjem trening-uzorka, tj. uzorka na kojem je model "treniran" (Dst i FD liste) kako bi se testirala pouzdanost s obzirom na korištene aproksimacije. Potom je evaluacija izvedena na testnom uzorku, tj. novom, neovisnom uzorku dodatno izabranih i izmjerenih događaja. Testni uzorak sadrži CME-bljesak-Dst-FD događaje u vremenskom periodu 1998-2012 koji se ne nalaze u trening-uzorku, a dobiven je istom metodom mjerenja i pridruživanja kao i trening-uzorak (tj. Dst i FD liste). Predviđanje modela evaluirano je usporedbom sa stvarnim rezultatima koristeći verifikacijske mjere za binarne događaje, koje su definirane prema tablici slučajeva sa četiri moguća ishoda: pogodak, lažno upozorenje, promašaj, te točno odbacivanje (detaljnije opisano u poglavlju 5). Pronađeno je da modeli imaju manje-više istovjetnu razinu točnosti predviđanja za trening- i testni uzorak. Dakle, točnost predviđanja neovisna je o korištenom uzorku te je uglavnom pod utjecajem korištenih aproksimacija. Nadalje, za oba modela utvrđeno je da je predviđanje manje pouzdano što je specifičnije. Model za geomagnetske oluje najuspješniji je u predviđanju vrlo intenzivnih geomagnetskih oluja kada je $Dst > 200$ nT, dok model za Forbusheva smanjenja najuspješnije predviđa da li će biti značajnijeg efekta, odnosno da li se očekuje $FD > 3$ %. Usporedbom s drugim modelima utvrđeno je da modeli daju dobra predviđanja s obzirom na korištene ulazne parametre i uspješnost predviđanja drugih modela.

6. KRATAK PREGLED I ZAKLJUČAK

Cilj predstavljenog istraživanja je predvidjeti geo- i GCR-efektivnost opaženog CME-a. Statističkom analizom utvrđeno je da su Dst i FD amplitude veće za brže i šire CME-e, kojima su pridruženi snažniji bljeskovi blizu centra Sunčevog diska, te za koje je izglednija CME-CME interakcija. Pronađene statističke veze upotrebljene su za uspostavu empirijskog modela, koji se bazira na geometrijskoj raspodjeli. Evaluacijom modela utvrđeno je da su modeli manje pouzdani kada se koriste za specifičnije predviđanje te im pouzdanost raste kada se rade "grublje" procjene. Glavne prednosti modela su: (1) ulazni parametri bazirani su na daljinskim opažanjima CME-a i Sunčevih bljeskova, što omogućuje rano upozorenje reda veličine ≈ 1 dan, te (2) opažanja CME-a i bljeskova potrebna za ulazne parametre ne moraju nužno biti vršena svemirskim letjelicama. Na temelju predstavljenog istraživanja izrađene su internet aplikacije za predviđanje svemirskih vremenskih prilika, koje su dostupne na stranicama Opservatorija Hvar. Nadalje, model predviđanja geomagnetskih oluja uključen je u tzv. "COMESEP sustav upozorenja", prvi u potpunosti automatizirani sustav detekcije CME-a i bljeskova te predviđanje njihovog vremena dolaska i potencijalno štetnih učinaka.

Ključne riječi: Sunce – Svemirske vremenske prilike – Koronini izbačaji (CME) – Kozmičko zračenje – Forbusheva smanjenja – Geomagnetske oluje

Abbreviations and acronyms

2stt Two sample t-test
ACE Advanced Composition Explorer
AU Astronomical Unit
BIAS Bias
CIR Corotating Interaction Region
CME Coronal Mass Ejection
COMESEP Coronal Mass Ejections and Solar Energetic Particles
CR Cosmic Ray
DBM Drag Based Model
Dst Disturbance storm time (index)
EIT Extreme ultraviolet Imaging Telescope
EUV Extreme UltraViolet
FAR False Alarm Ratio
FD Forbush Decrease
GCRs Galactic Cosmic Rays
GOES Geostationary Operational Environmental Satellite
HI Heliospheric Imager
HSS Heidke Skill Score
ICME Interplanetary Coronal Mass Ejection
IMF Interplanetary Magnetic Field
LASCO Large Angle Spectroscopic COronagraph
MAG Magnetometer
MFI Magnetic Field Instrument
NM Neutron Monitor
NOAA National Oceanic and Atmospheric Administration
OSO Orbiting Solar Observatory
POD Probability Of Detection
SIR Stream Interface Region
SOHO SOlar and Heliospheric Observatory
SPIDR Space Physics Interactive Data Resource
SSC Sudden Storm Commencement
STEREO Solar TErrestrial RElations Observatory

SWE Solar Wind Experiment
SWEPAM Solar Wind Electron, Proton and Alpha Monitor
UV UltraViolet

Contents

1. Introduction **1**
1.1. Coronal mass ejections (CMEs) 2
1.1.1. Observational characteristics of CMEs and solar flares 3
1.1.2. CME initiation and the standard flare model 6
1.1.3. Interplanetary Coronal Mass Ejections (ICMEs) 7
1.2. CME-related Space weather effects 10
1.2.1. Geomagnetic storms 10
1.2.2. Forbush decreases 14

2. Data selection, measurement method and event list **18**

3. Statistical analysis **24**
3.1. The relation between CMEs/flares and geomagnetic storms 27
3.2. The relation between CMEs/flares and Forbush decreases 32

4. Empirical statistical models for space weather forecast **37**
4.1. Empirical statistical model for geomagnetic storms 40
4.2. Empirical statistical model for Forbush decreases 45

5. Space weather forecast evaluation **52**
5.1. Geomagnetic storm forecast evaluation 54
5.2. Forbush decrease forecast evaluation 56

6. Conclusion and summary **59**

Bibliography **62**

1. Introduction

Coronal mass ejections (CMEs) and solar flares are the most violent eruptive processes on the Sun and are often qualified as the main drivers of **space weather**: "... a field of research that will provide new insights into the complex influences and effects of the Sun and other cosmic sources on interplanetary space, the Earth's magnetosphere, ionosphere, and thermosphere, on space- and ground-based technological systems, and beyond that, on their endangering affects to life and health" [Bothmer and Daglis, 2007]. Although space weather is a relatively new research area, tightly connected to the increasing development of the human technology, one may argue that early studies of the solar activity already gave birth to space weather, long before the satellite era. The first recognized space weather event was the so-called Carrington event in 1859 when an intensive white-light solar flare was observed for the first time and followed by an intense and broad-range terrestrial responses [Cliver and Svalgaard, 2004]. These terrestrial responses included low-latitude aurorae, as well as the arcing from the induced currents in the telegraph wires in USA and Europe [Tsurutani et al., 2003, and references therein]. The human technology has advanced remarkably in the last century, making us more vulnerable to the solar activity. We live in the era of satellites, airplanes, electrical power grids and space travel, all of which can be affected by solar eruptive phenomena [see *e.g.* Feynman and Gabriel, 2000]. Therefore, understanding and forecasting of the solar eruptive phenomena and their space weather effects is of great importance for the modern human society.

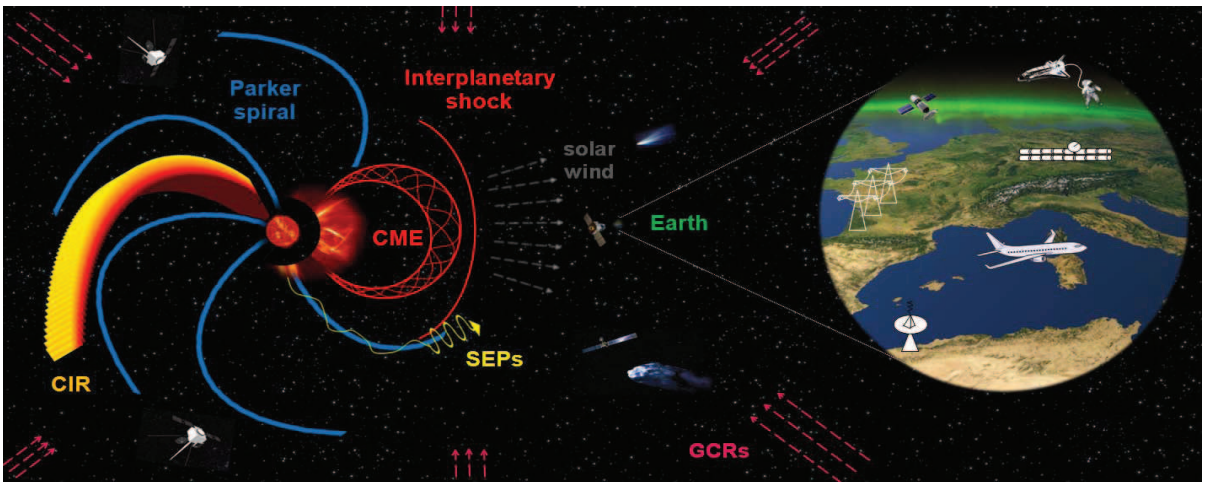


Figure 1.1.: A schematic overview of the space weather related topics

1.1. Coronal mass ejections (CMEs)

Coronal Mass Ejection (CME) is a historical term for a white-light coronagraphic signature of the mass moving away from the Sun. CMEs were first detected in early 1970-ties by the first space-borne coronagraph launched on the satellite Orbiting Solar Observatory, OSO-7. However, their existence was suspected earlier, especially after the discovery of the solar wind (*i.e.* outflow of mass from the Sun) with in situ measurements of the Mariner 2 probe to Venus in 1962 [see *e.g.* Howard, 2006, Foukal, 2004].

CMEs are often associated with solar flares and prominence eruptions. Solar flare is a dissipative energy release that causes a wide range of electromagnetic emission at different wavelengths, from radio waves to gamma rays. Prominences are cool and dense chromosperic material supported by magnetic field against the gravity in the hotter and tenuous corona. Although there is no one-to-one relationship between CMEs, flares, and prominences, it is generally accepted that these are closely related and are different manifestations of a single magnetically-driven physical process [Priest and Forbes, 2000]. Therefore, they are described by a unified model called a "standard" flare model, shortly described in Section 1.1.2. There are many observations which support this unified model, as described in Section 1.1.1. In the sense of the space weather forecast, associating CMEs and flares or prominences is a very useful approach, because it provides additional information on CMEs, which cannot be obtained from the white-light coronagraphic observations, as described in Section 1.1.1.

CMEs are observed in the interplanetary space by remote heliospheric imaging methods, such as *Heliospheric Imager* instruments [HI1 and HI2, Eyles et al., 2009] onboard *Solar-Terrestrial Relations Observatory* spacecraft [STEREO A and B, Kaiser et al., 2008]. Heliospheric imaging of CMEs with STEREO HI instruments theoretically enables tracking the CME from the Sun to Earth; however, this utility is limited by the STEREO's heliocentric orbit and observational constrains due to projection effects and optically thin medium [see *e.g.* Vourlidas and Howard, 2006, Rollett et al., 2012, and references therein]. Therefore, the best indication of the CME passage at a certain point are still *in situ* interplanetary plasma and magnetic field measurements, such as *Magnetic Field Instrument* [MFI, Lepping et al., 1995] and *e.g. Solar Wind Experiment* [SWE, Ogilvie et al., 1995] instruments onboard *Wind* spacecraft located at L1 Lagrangian point near Earth. CME *in situ* properties, as well as their propagation characteristics are shortly described in Section 1.1.3. This internal probing of the interplanetary CME (ICME) provides a unique insight into the CME structure; however, due to the fact that the CME propagation and evolution are still open problems, models and methods are needed to associate remote CME observations at the Sun with the *in situ* measurements near Earth.

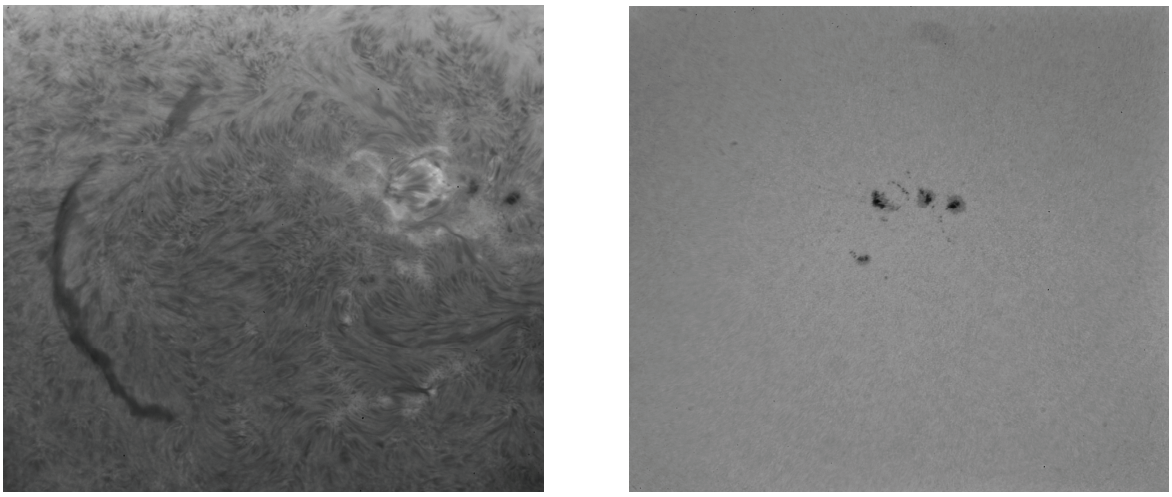


Figure 1.2.: H-alpha (left) and corresponding white-light (right) observation of the active region AR1271 and a nearby filament with a double solar telescope of Hvar Observatory (22 August 2011).

1.1.1. Observational characteristics of CMEs and solar flares

The eruptive solar phenomena that influence the space weather originate in the solar atmosphere and are driven by the energy stored in the magnetic field. The magnetic structure of the solar atmosphere consists of the "open" and closed magnetic fields and these two types of magnetic structures are associated to different types of phenomena we observe. Coronal holes, *i.e.* regions of open magnetic field lines are the sources of the high-speed flows in the heliosphere that are of relatively low density [Krieger et al., 1973, Gosling and Pizzo, 1999] and can later form the so-called corotating interaction regions (CIRs, see Section 1.1.3). CMEs originate from the regions of closed magnetic field lines, usually from active regions and quiescent-filament regions. Active regions are structures in the solar atmosphere above sunspots, with enhanced and structured magnetic field, and increased activity compared to the surrounding area. Filaments consist of the colder and denser plasma suspended in the warmer and tenuous solar atmosphere by the magnetic field, appearing darker than the surrounding medium. A filament can be observed in the H-alpha spectral line (Fig 1.2), which forms in the solar chromosphere. When they are observed at the limb, they appear brighter than the dark background and are then referred to as prominences. Quiescent filaments, i.e. prominences can be stable over a time period of months. Eruptive prominences, a special subset of active prominences and often associated with CMEs, can erupt into the high corona in the matter of minutes [Foukal, 2004].

CMEs can be observed directly using white-light coronagraphs, which image the solar corona. These are special telescopes where the bright solar photosphere is occulted (imitating a total eclipse) which detect photospheric light scattered by coronal electrons. The observed intensity is therefore determined by the line-of-sight column density so the bright

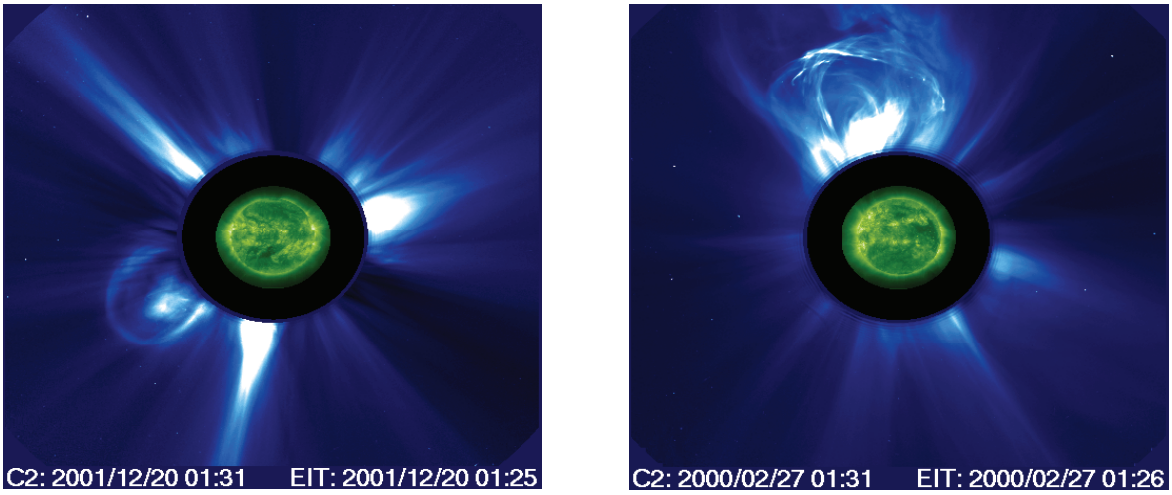


Figure 1.3.: White light image of a CME observed by the SOHO LASCO/C2 coronagraph (with a superposed image recorded by EIT/SOHO UV imager in 195 Å) on December 20th 2001 (left) and on February 26th 2000 (right). Both CMEs display a typical "three-part" structure (*Credit*: SOHO LASCO CME Catalog).

features mowing away from the Sun (CMEs) are interpreted as outward moving density structures [see *e.g.* Hudson et al., 2006, , and references therein]. Observed CMEs display a variety of morphologies: ranging from narrow jets to wide, seemingly global eruptions [Howard et al., 1985, Webb and Howard, 2012]. It is important to note that CME observations suffer from projection effects, *i.e.* they are 3D structures projected in two dimensions in an optically thin medium. This introduces distortions in their appearance and complicates the determination of their properties [Burkepile et al., 2004, Hudson et al., 2006]. Therefore, CMEs with large apparent angular width, especially so-called HALO CMEs (with apparent width of 360 degrees), are not actually global eruptions, but are directed close to the Sun-observer line. Around one third of the observed CMEs appear as a "three-part" structure (see Fig 1.3), with a bright leading edge, followed by a dark cavity and a bright core [Illing and Hundhausen, 1985]. These observations support the standard CME-flare model (see Section 1.1.2) where the bright leading edge is interpreted as a coronal plasma pileup, the cavity as the magnetic field dominated region and the bright core as the eruptive prominence. This configuration is therefore often viewed as a "standard CME" in both observational and theoretical studies [see *e.g.* Gopalswamy et al., 2006, , and references therein].

The occurrence of CMEs follows the solar cycle in both phase and amplitude and varies by an order of magnitude over the cycle, from $\approx$ 1 per day in the solar minimum to $\approx$ 5 per day in the solar maximum [Webb and Howard, 1994, Schwenn et al., 2006, Webb and Howard, 2012]. CMEs detected by white-light coronagraphs are characterized by kinematic properties (speed), apparent angular width and a central position angle in the sky plane, measured counter-clockwise from solar north. Measured speeds range from a few tens of km/s to nearly 3000 km/s, with an average value which

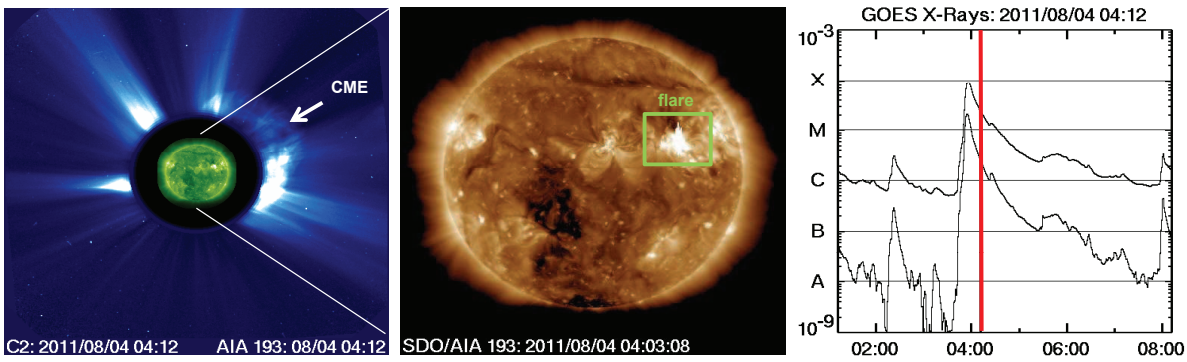


Figure 1.4.: CME detected by the white light coronagraph LASCO/C2 onboard SOHO on August 4th 2011(left), associated solar flare detected in EUV wavelength 193 Å with AIA instrument onboard SDO (left and middle), and the corresponding Soft X-ray flux measured by GOES satellite (right). EUV detection in the middle image is given at the time of the peak in the Soxt X-ray flux seen on the right image. Red line marks the time of the first detection of the CME in C2, as seen in the left image (*Credit*: SOHO LASCO CME Catalog).

is slightly higher than the ambient slow solar wind speed [$\approx$ 300 and 500 km/s in the solar minimum and maximum, respectively, see Howard et al., 1985, St. Cyr et al., 2000, Yashiro et al., 2004, Schwenn et al., 2006, Hudson et al., 2006]. The apparent angular width of CMEs ranges from a few degrees to more than 120 degrees, with an average value 40-70 degrees [different studies give different values within this range, see Howard et al., 1985, St. Cyr et al., 2000, Yashiro et al., 2004, Schwenn et al., 2006, Webb and Howard, 2012]. CME observations are obtained either by visual inspection of coronagraph images or by automated detection software, resulting in a variety of publically available CME catalogs [listed in Webb and Howard, 2012]. One of the most widely used CME catalogs is SOHO LASCO CME catalog [Yashiro et al., 2004] available at `http://cdaw.gsfc.nasa.gov/CME_list/`. The catalog provides CMEs detected in the field of view of the Large Angle Spectroscopic Coronagraph [LASCO, Brueckner et al., 1995] onboard Solar and Heliospheric Observatory [SOHO, Domingo et al., 1995]. The primary measurements provided by this catalog are performed "manually" on each CME and include the apparent central position angle, the angular width in the sky plane, and the height (heliocentric distance) as a function of time.

Due to the occulting disc, the coronagraph observations of CMEs do not provide information on the CME source region on the solar disc, which would allow to determine the CME direction. CMEs are often associated with a number of phenomena whose signatures can be seen on disc in various parts of the electromagnetic spectrum. These include most notably solar flares, eruptive filaments, waves and dimmings seen in extreme ultraviolet (EUV) imagers and radio bursts [see *e.g.* Webb and Howard, 2012, , and references therein]. One of the most common associations is the one between solar flares and CMEs, relying on the observation that the most energetic CMEs occur in close association with

powerful flares [*e.g.* Yashiro et al., 2006]. Furthermore, observational and statistical studies have shown that the early kinematic evolution of the CME is related to the energy release in the solar flare [*e.g.* Kahler et al., 1988, Zhang et al., 2001, Moon et al., 2003, Burkepile et al., 2004, Vršnak et al., 2004a, Maričić et al., 2007]. Therefore, the two can be associated using independent CME and solar flare measurements [*e.g.* Vršnak et al., 2005], providing the information on the CME source region. An example of the CME and associated solar flare is shown in Figure 1.4.

1.1.2. CME initiation and the standard flare model

In the pre-eruption stage a closed magnetic structure with non-potential magnetic field is generally considered, with stored free energy needed to describe the energy release during the CME/flare event. The most common magnetic structure employed in modeling is a flux rope, a cylindrical plasma structure with magnetic field draped around the central axis [Lepping et al., 1990]. There are observational evidences for flux ropes seen in the early CME initiation phase [Schmieder et al., 2015, and references therein], and their topology is often seen in the interplanetary CME counterparts (see Section 1.1.3). One of the unresolved questions is whether the flux rope is formed below the photosphere and emerges, or is formed above the photosphere by shearing motions or some other physical mechanisms introducing the free energy into the system [*e.g.* Forbes et al., 2006]. Once emerged or formed, the flux rope either evolves through a series of quasi-equilibrium states or undergoes an abrupt magnetic reconfiguration, both scenarios leading to the loss of equilibrium and the eruption of the coronal magnetic structure, *i.e.* flux rope [Schmieder et al., 2015, and references therein].

The eruption triggers magnetic reconnection of the overlying coronal magnetic field, a process which can be described as a "...topological restructuring of a magnetic field caused by a change in the connectivity of its field lines." [Priest and Forbes, 2000]. Reconnection can occur above the ejection, or bellow. In both cases the reconnection removes the overlying flux reducing the magnetic tension of the overlying field and enabling a fast outward expansion of the flux rope. When reconnection occurs below the flux rope, it feeds it with additional magnetic flux. On the other hand, reconnection releases both thermal and non thermal energy, producing a number of effects, which are all generally described as the solar flare [see *e.g.* Priest and Forbes, 2002]. The flux rope can support plasma, in which case also an eruptive prominence or filament can be observed. As the erupting flux rope moves away from the Sun, it can produce the disturbances in the local medium, which are observed as coronal waves (extreme ultraviolet, EUV waves) and can form a shock in front of its leading edge [see *e.g.* Warmuth, 2007]. The whole process (also known as the standard flare model) is schematically presented in Figure 1.5. This is a very general view of the CME initiation and many CMEs are not associated with some or even any of the forementioned phenomena (flares, prominences, waves), *i.e.* low-coronal

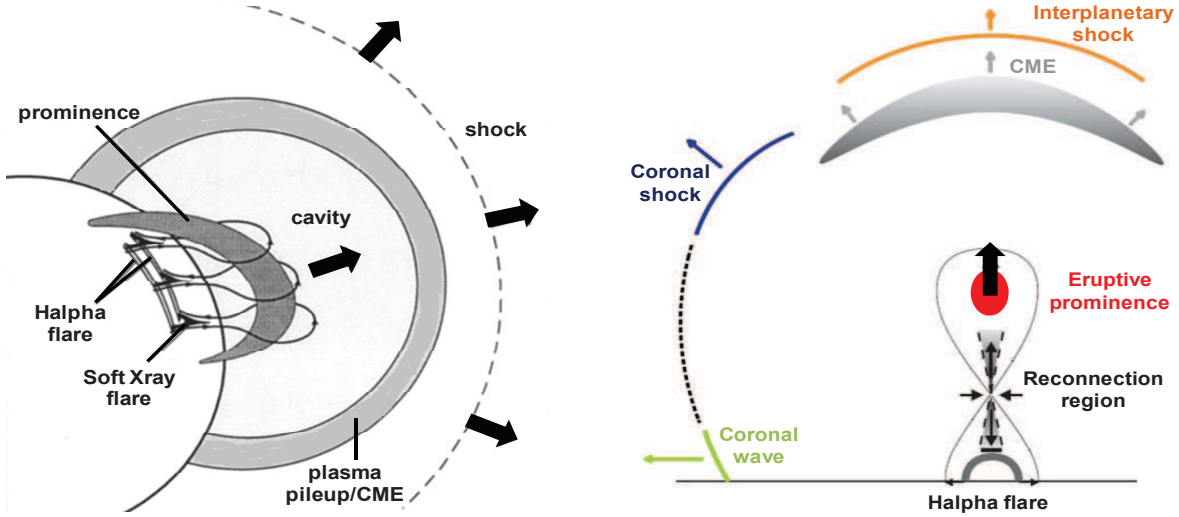


Figure 1.5.: A schematic overview of the standard model in 3D view [left, addapted from Forbes, 2000] and side-view [right, addapted from Warmuth, 2007].

signatures. However, even these so-called "stealth" CMEs fit well in the standard model if they are regarded as less-energetic CMEs erupting in the regions of weak overlying field, which then reconfigurates higher up in the corona, where the low density makes the observation of plasma heating challenging [Robbrecht et al., 2009, D'Huys et al., 2014].

1.1.3. Interplanetary Coronal Mass Ejections (ICMEs)

CMEs observed in the interplanetary space using remote heliospheric imaging or *in situ* measurements are called Interplanetary CMEs (ICMEs). The *in situ* measurements used for the identification of ICMEs usually include a number of changes in the solar wind and interplanetary magnetic field (IMF) parameters, but highly depend on how the ICME boundaries are defined. A historical approach is very common, where ICME includes the whole disturbance: the shock (if existing), the sheath region, and "driver" or ejecta [see *e.g.* Rouillard, 2011]. A shock is the discontinuity formed at the leading edge of the CME, when the CME is faster than the surrounding solar wind magnetosonic speed. If present, it is followed by a turbulent and heated sheath region, usually characterized by high plasma density and higher magnetic field strength. Finally, the "driver" or ejecta [which some authors refer to as ICME, *e.g.* Richardson and Cane, 2010] usually shows some or all of the following signatures: magnetic field enhancement, rotation of the magnetic field, low magnetic field fluctuations, low proton temperature, low proton density, low proton beta parameter (ratio of magnetic and kinetic pressure), monotonic speed decrease, enhanced alpha to proton ratio, elevated oxygen charge states, enhanced Fe charged states, bidirectional electron streaming [see *e.g.* Zurbuchen and Richardson, 2006, Rouillard, 2011, , and references therein]. A schematic of the three-dimensional structure of an ICME relating magnetic field, plasma, and particle signatures is given in Figure 1.6. Magnetic flux ropes are a special subset of ejecta which have magnetic field enhancement and smooth

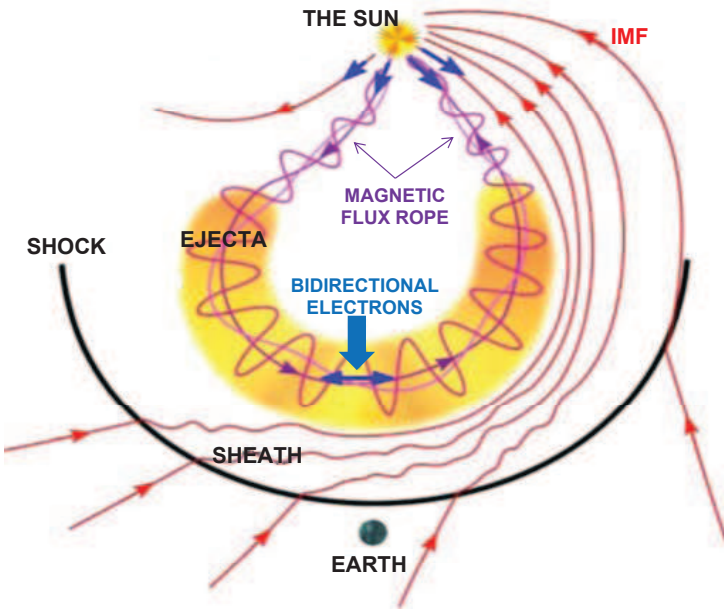


Figure 1.6.: A schematic of the three-dimensional structure of an ICME and two-dimensional structure of the interplanetary magnetic field (IMF) [adapted from Zurbuchen and Richardson, 2006]

rotation of the magnetic field, whereas magnetic flux ropes with low proton temperature and low proton beta parameter are called magnetic clouds [Burlaga et al., 1981].

ICMEs often do not have perfectly clear signatures and thus are often not easily identified in the *in situ* measurements. Form Figure 1.6 it is quite obvious that depending on the trajectory of the spacecraft through an ICME, different regions will be encountered. If the spacecraft passes through the flank of the ICME, only shock signatures will be observed. On the other hand, if the ICME does not form a shock, only ejecta signatures will be observed. Identification of the ICMEs in the in situ measurements is additionally hampered by CME-CME interaction and stream interaction regions (SIRs). SIRs are regions where a fast solar wind stream interacts with the slow solar wind. Often this region is persistent through couple or even several solar rotations resulting in a so-called corotating interaction region (CIR). A fast solar wind component originates from coronal holes, whereas the slow component originates from regions of closed magnetic field in the solar atmosphere (streamer belts). The spatial variability in the coronal expansion and solar rotation can cause solar wind flows of different speeds to become radially aligned and compressive interaction regions are produced where high-speed wind runs into slower plasma ahead [Gosling and Pizzo, 1999]. A defining structure within the CIR is the stream interface, which separates originally kinetically cool, dense, and slow solar wind from what was originally hot, tenuous, and fast solar wind. It is characterized by an abrupt drop in density, a similar increase in temperature, and a small increase in speed [Burlaga, 1974, Crooker et al., 1999]. CIRs can cause similar space weather effects as ICMEs (see Section 1.2), but often to a smaller degree. Therefore, it is important to distinguish between the

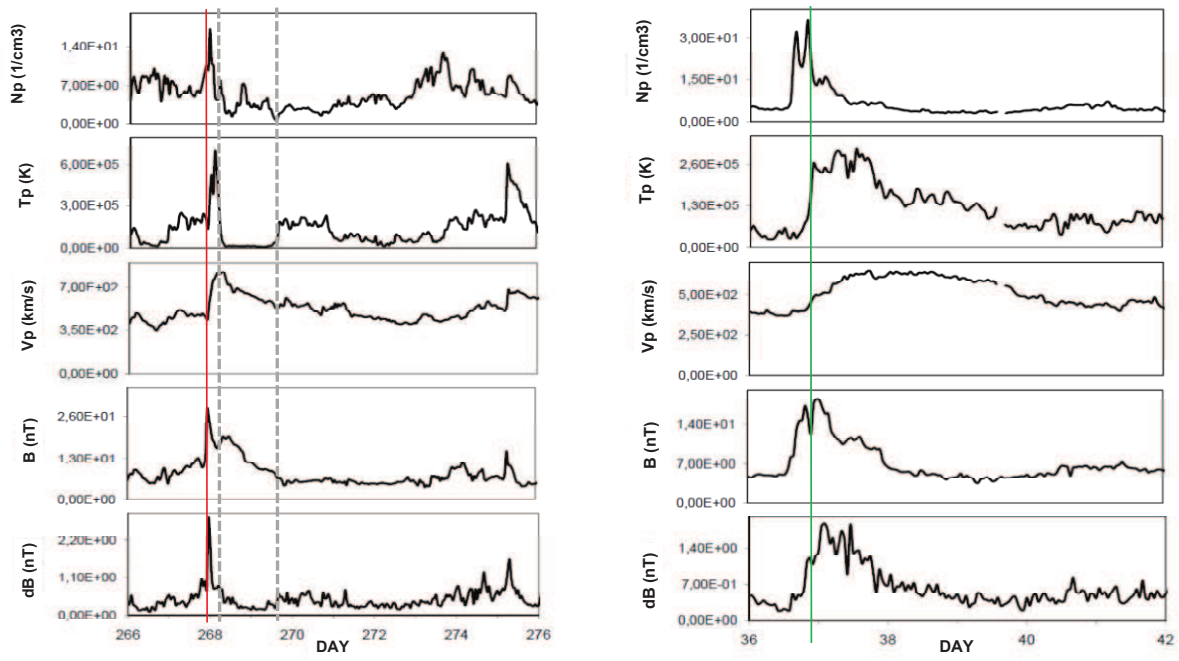


Figure 1.7.: *In situ* measurements of 24/25 September 1998 ICME (left) and 5 February 2000 CIR [right, both adapted from Dumbović et al., 2012b]. The panels show (top to bottom): solar wind proton density, temperature, and flow speed, magnetic field strength, and magnetic field fluctuations. The red line marks the shock, whereas the gray dashed lines mark the beginning and the end of the ejecta (left). The green line marks the stream interface (right).

two using *in situ* measurements. An example of ICME and CIR identification from the *in situ* measurements is shown in Figure 1.7.

Although ICMEs are undoubtedly related to their solar sources (CMEs), the association of the two still remains an important scientific issue. Many authors associated CMEs and ICMEs using a variety of different methods, providing CME-ICME lists [*e.g.* Zhang et al., 2003, Schwenn et al., 2005, Manoharan, 2006, Richardson and Cane, 2010]. However, it should be noted that there are many examples where the associations between different authors disagree. Relating the *in situ* measurements of ICMEs and their solar sources is not a straightforward task, since it involves a quite complex and not yet fully understood CME kinematic evolution. The CME kinematic evolution is often divided into three phases [Zhang et al., 2001]: the initiation phase, characterized by a slow rise of the magnetic structure; the acceleration phase, characterized by gradual or impulsive acceleration; and propagation phase, showing an almost constant speed in coronagraphic field of view. The forces governing these early kinematic phases are the Lorentz force, gravity and aerodynamic drag; however, at larger distances ($\approx 20 R_{\rm sun}$) the drag becomes dominant [Cargill, 2004, Vršnak et al., 2004b, 2013]. The drag depends on the CME properties as well as the properties of the ambient solar wind and acts to adjust the CME speed to the speed of the ambient solar wind. Therefore, CMEs faster than the solar wind tend to decelerate, whereas CMEs slower than the solar wind tend to accelerate. A drag based model of

the heliospheric propagation of ICMEs has been established [DBM, Vršnak et al., 2013] which qualitatively describes ICME propagation quite successfully, but quantitatively it is limited by parameters such as CME speed, width and mass as well as ambient solar wind density and speed, which are not easily derived. ICME propagation is furthermore complicated by the CME–CME interaction [*e.g.* Temmer et al., 2012], determination of the CME direction and possible deflections of the ICME from the original direction in the corona [*e.g.* Yashiro et al., 2008, Gui et al., 2011, Möstl et al., 2015], or in the interplanetary space [*e.g.* Wang et al., 2004, 2006]. The CME direction determines whether or not the ICME will arrive, and whether it will hit with an apex (frontally) or with a flank. The arrival time difference between the ICME apex and flank can be even 2 days [Möstl and Davies, 2013]. Other ICME propagation models are faced with similar challenges and in general the reliability of the propagation models in deriving ICME arrival times is around 10 hours [*e.g.* Siscoe and Schwenn, 2006, , and references therein].

1.2. CME-related Space weather effects

As they propagate through the heliosphere ICMEs can interact with magnetic field structures and charged particles they encounter. The interaction of ICMEs with the geomagnetic field drives geomagnetic storms, *i.e.* disturbances of the geomagnetic field. Geomagnetic storms are related to many of the previously mentioned harmful effects; therefore, their prediction is an important aspect of the space weather. On the other hand, the interaction of ICMEs with galactic cosmic rays produces short-term depressions in the galactic cosmic ray flux, called Forbush decreases. These depressions can be used as an indication of the ICME passage in the pre-satellite era, when interplanetary measurements were not available. Furthermore, they could be of relevance for the human space missions, where one of the most hazardous factors is a long-term exposure to galactic cosmic rays. Sections 1.2.1 and 1.2.2 shortly describe our current knowledge about geomagnetic storms and Forbush decreases and the physical processes behind them.

1.2.1. Geomagnetic storms

Geomagnetic storms are recorded by ground-based magnetometers for almost two centuries, but the explanation of how and why they occur depended on the discovery of the Earth's magnetosphere and its interaction with the magnetized solar plasma flow [see *e.g.* Akasofu, 2007, for historical overview]. Chapman and Ferraro [1931] were the first to introduce the concept of the Earth's magnetosphere. They suggested that the solar plasma flow forms a comet-like structure around the Earth, extending in the anti-solar direction and confining the Earth and its magnetic field in it. They also predicted the existence of the ring current - a westward flowing current system in the Earth's magnetosphere, responsible for the reduction in the horizontal component of the geomagnetic field during the storm. Dungey [1961] was the first to suggest that solar plasma flow is

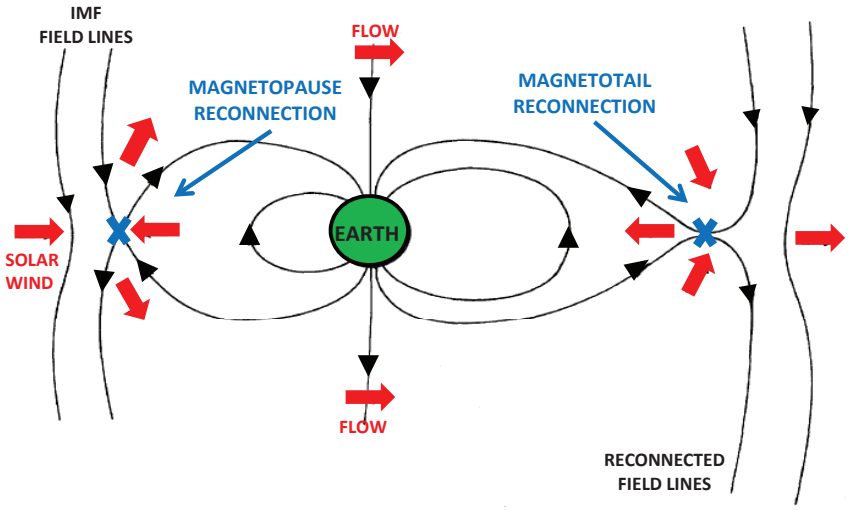


Figure 1.8.: The concept of the geomagnetic storm based on reconnection between the interplanetary magnetic field (IMF) lines and geomagnetic field lines at the magnetopause, followed by a subsequent reconnection in the magnetotail. For a more detailed explanaiton see the main text [adapted from Dungey, 1961].

magnetized and that there is connectivity between the geomagnetic and interplanetary magnetic field. He proposed that reconnection takes place on the dayside magnetosphere boundary (magnetopause) and that the newly connected field lines are then transported by the solar wind to the magnetotail (magnetosphere extended in the anti-solar direction). Subsequently, the field lines are reconnected there and then shrink towards Earth. This concept is shown in Figure 1.8.

The modern concept of the geomagnetic storm relies on the scheme presented in Figure 1.8. A geomagnetic storm occurs if the topology of the magnetic field in the ICME is favorable for reconnection, *i.e.* if there is a strong southward component of the magnetic field. As reconnection takes place and energy is released, charged particles originating from the solar wind enter deep into the magnetosphere. As a consequence currents are formed in the magnetosphere and ionosphere - ring current particle fluxes are increased introducing (westward flowing) partial ring currents, and particles are dumped into the high latitude (polar) regions of the Earth as field-aligned currents and (westward flowing) auroral electrojet currents [*e.g.* Campbell, 2001]. When reconnection stops so does the energy/particle feed as well, and charged particles gradually accumulate in the Earths radiation (Van Allen) belts. The contributions of these currents to ground-based magnetic recordings are seen as disturbances and are not the same throughout the entire Earth. At high latitudes the field aligned currents and auroral electroject currents dominate, whereas at low and equatorial latitudes the dominant contribution is from the ring current. At mid-latitudes both ionospheric and magnetospheric currents contribute to the magnetic recordings [Campbell, 2001]. Various measures of the magnetic activity, called geomagnetic indices, are used to describe these geomagnetic field variations and are

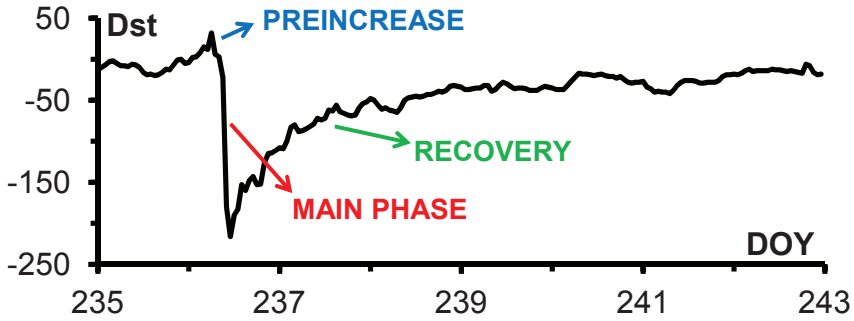


Figure 1.9.: Geomagnetic storm seen in the Dst index on August 24 2005 (data taken from Space Physics Interactive Data Resource, SPIDR).

correspondingly latitude dependent. Auroral electrojet index (AE) is used for the high latitudes, Planetary geomagnetic activity index (Ap) and its discrete equivalent Kp index are used for the mid-latitudes, whereas for low-latitudes Disturbance storm time index (Dst) is used [Campbell, 2001]. The Dst index is derived from the horizontal component recorded by four observatories located between -33 and $+30$ degree latitude and represents the axially averaged disturbance of the surface magnetic field at the dipole equator [*e.g.* Rostoker, 1972, Verbanac et al., 2011a]. The present day concept of geomagnetic storms, as measured by the Dst index, was first established by Cahampan and Bartels [1940]. The storm starts with a storm sudden commencement (SSC): a step-function-like increase in the horizontal component (*i.e.* Dst), and is followed by a main phase: a large and rapid decrease that follows the SSC. After reaching the maximum decrease during the main phase, the storm recovers slowly during the recovery phase. The SSC is caused by the impact of the ICME shock on the magnetosphere, whereas the main phase is caused by the formation of partial ring current. As the partial ring current slowly decays, the Dst index slowly recovers. An example of the geomagnetic storm as seen in the Dst index is presented in Figure 1.9.

ICMEs display a wide range of geo-effectiveness, *i.e.* may produce large or small geomagnetic storms or none at all. The enhanced geo-effectiveness is related to effective reconnection with the geomagnetic field and therefore with the southern component of the ICME magnetic field, B_s and the corresponding y component of the convective electric field, $E_y = v\dot{B}_s$ (where v is the solar wind speed). The relation between *in situ* properties of ICMEs and geomagnetic storms has been investigated in statistical studies considering different geomagnetic indices. Dst index was found to correlate with the B_s, a weaker correlation was found between v and Dst, and a strong correlation was found between Dst and E_y [*e.g.* Kane, 2005, Richardson and Cane, 2011b, Verbanac et al., 2013, , and references therein]. Most of the intense storms were found to be caused by magnetic clouds with shocks [*e.g.* Echer et al., 2008, Yermolaev et al., 2012]. In addition, it was found that faster ICMEs have stronger fields; therefore, faster ICMEs can enhance both crucial geo-effective factors, B_s and E_y [*e.g.* Gonzalez et al., 1998,

Koskinen and Huttunen, 2006, Verbanac et al., 2013]. The approximate threshold value for E_y needed to produce an intense storm with $Dst < -100$ nT was obtained empirically [$\approx E_y > 5$ mV/m for a duration of 2-3 hours, see *e.g.* Gonzalez and Tsurutani, 1987, Echer et al., 2008, Richardson and Cane, 2011b]. It should be noted that CIRs can also be geo-effective; however, they rarely cause intense geomagnetic storms [$Dst < -100$nT, *e.g.* Richardson et al., 1996, Verbanac et al., 2011b].

Measurements of B_s needed to estimate E_y are provided at L1 lagrangian point, *i.e.* ≈ 1 hour before the start of the disturbance (for typical ICME speed), providing very limited "response time" [*e.g.* Koskinen and Huttunen, 2006, Richardson and Cane, 2011b]. Our current knowledge restricts us from predicting the crucial B_s component of the ICME magnetic field at earlier times, *e.g.* from remote solar observations. There are studies trying to compare the magnetic field of the ICME to its solar source region magnetic fields in the initiation phase [*e.g.* Bothmer and Schwenn, 1994, Möstl et al., 2008]. However even if the original orientation of the magnetic field inside the CME would be known in the initiation phase, the prediction of B_s component at the Earth would be severely hampered by the fact that CMEs rotate while propagating [*e.g.* Vourlidas et al., 2011, Isavnin et al., 2013]. Nevertheless, several authors tried to relate remote solar observations of CMEs with geomagnetic storms, assuming that the solar sources of ICMEs (CMEs) must show some properties which can indicate a possible level of the associated geomagnetic activity. These studies led to the conclusion that the geo-effectiveness of CMEs is related to the following solar properties of CMEs and the associated solar flares: CME initial speed [*e.g.* Srivastava and Venkatakrishnan, 2004, Gopalswamy et al., 2007], apparent angular width [*e.g.* Zhang et al., 2003, Srivastava and Venkatakrishnan, 2004, Zhang et al., 2007], source region location [*e.g.* Zhang et al., 2003, Srivastava and Venkatakrishnan, 2004, Gopalswamy et al., 2007, Zhang et al., 2007, Richardson and Cane, 2010], the intensity of the CME-related flare [*e.g.* Srivastava and Venkatakrishnan, 2004], and occurrence of successive CMEs [*e.g.* Gopalswamy et al., 2007, Zhang et al., 2007]. However, most of the above studies have samples based on geomagnetic storms observed at Earth, which are then associated to CMEs at the Sun. They do not consider a sample of CMEs at the Sun and then relate them to geomagnetic activity observed at Earth (if there is any). Therefore, they do not take into account so called false and missing alarms. False alarms are CMEs apparently having favorable solar properties, which do not produce geomagnetic storms, whereas missing alarms are the geomagnetic storms produced by CMEs with apparently non-favorable solar properties [see *e.g.* Schwenn et al., 2005, Rodriguez et al., 2009]. There were several attempts to construct geomagnetic storm prediction-models based on the remotely-measured properties of CMEs [*e.g.* Srivastava, 2005, Valach et al., 2009, Kim et al., 2010, Uwamahoro et al., 2012]. The authors however point out a quite low success rate of the models, unless interplanetary conditions are also taken into account.

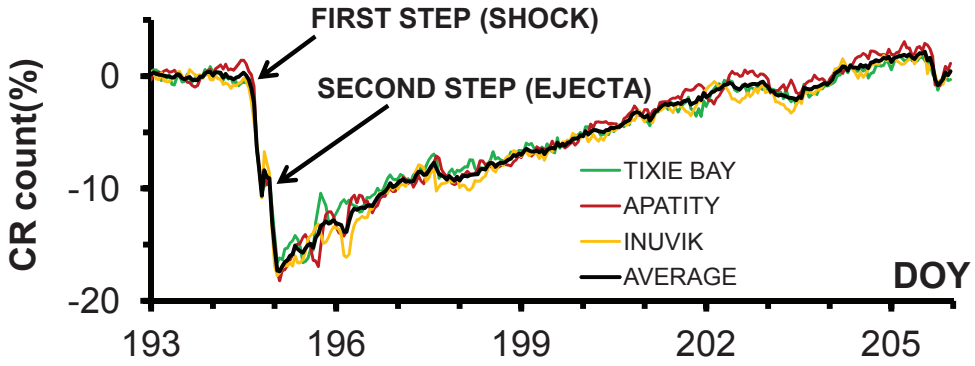


Figure 1.10.: Two-step Forbush decrease detected by ground-based neutron monitors at Earth on July 13 1982. Three high-latitude stations of the similar cutoff rigidity and spaced about equally in longitude are used to minimize daily variations (colored curves). The average of the stations is given by a black curve. Hourly averages of the relative particle counts are presented, normalized to the quiet 1 day period prior to the depression (data taken from Space Physics Interactive Data Resource, SPIDR).

1.2.2. Forbush decreases

Forbush decreases are short term depressions in the galactic cosmic ray (GCR) flux, first observed by Forbush [1937] and Hess and Demmelmair [1937]. There are two types, one caused by CIRs and the other caused by ICMEs. CIRs usually produce shallower and more symmetric depressions [*e.g.* Iucci et al., 1979, Richardson, 2004] and are often recurrent (due to corotating nature of their interplanetary sources). ICME-related depressions show a variety of shapes and magnitudes, which is generally thought to be related to the characteristics of the ICME part where the detector passes. Depending on the trajectory of the spacecraft through an ICME we expect to see different depressions, similarly as we would expect to see different ICME *in situ* measurements corresponding to different ICME regions [*e.g.* Cane et al., 1994, Cane, 2000, Blanco et al., 2013a]. If only ICME ejecta is intercepted, the decrease is confined within the duration of the ejecta, whereas the effect of the shock persists many days after the passage of the shock and causes a slow recovery [Cane et al., 1994]. If both the shock and ejecta are intercepted, a two-step decrease is expected, first step coming from the sheath, whereas the second depression is associated with the ejecta [*e.g.* Barnden, 1973, Cane, 2000, Richardson and Cane, 2011a]. Largest observed Forbush decreases show a two-step structure and the two regions are in average found to be roughly equal in magnitude [Richardson and Cane, 2011a]. However, two-step decreases are not very common and Forbush decreases generally show a diverse and complex structure, even in case when shock/sheath is followed by a single magnetic ejecta [Jordan et al., 2011]. An example of a two-step Forbush decrease is shown in Figure 1.10.

Forbush decreases can be measured in the interplanetary space and by ground based detectors at Earth. Due to the relatively small effect (several percent) a large statistics is

needed, *i.e.* large particle counts. These are easily provided by large ground based neutron monitors with $\approx 10^4$ counts per hour. However, ground based observations are influenced by several factors: (1) they do not detect primary GCRs, but secondary particles which are the product of GCR interacting with the atmosphere, (2) primary GCRs interact with the geomagnetic field before they enter the atmosphere and the point of their entrance is highly dependent on this interaction, and (3) the GCR flux exhibits daily variations which may represent noise in Forbush decrease measurements. Due to the geomagnetic effect there is a difference in the particle energies of primary GCRs which contribute to different neutron monitor measurements. Particles can enter the atmosphere more easily at poles and high latitudes (where the geomagnetic field is directed toward the atmosphere) than at the equator and low latitudes. Whether or not a particle can enter the atmosphere at a certain point depends on the rigidity, a quantity which depends on the magnetic field strength and particle energy. Depending on the latitude, different neutron monitors have different cutoff rigidities and even the stations close to the pole have cutoff rigidity > 0. Since Forbush decrease is rigidity dependant, *i.e.* it is more pronounced for low-energy particles [Lockwood, 1971, Cane, 2000] smaller depressions are observed at Earth than in the interplanetary space.

Another drawback of using ground based neutron monitor measurements are the daily variations of the detected particle counts. The daily variations are caused by the outward radial convection due to solar wind and an inward diffusion along the direction of the interplanetary magnetic field (see the transport theory description in the next paragraph). The balance between the two generates a small, but significant GCR spatial anisotropy which is observed as the daily variation in the ground-based measurements [$\approx 1\%$, Parker, 1964, Tiwari et al., 2012]. The daily variations can be reduced by averaging several stations located at approximately same latitude (which have nearly the same cutoff rigidity) and having different asymptotic viewing directions [which approximately correspond to different longitudes, see Dumbović et al., 2011, and Figure 1.10]. Spacecraft observations are not hampered by these effects, but most of them offer much less statistics due to their size and geometric factor. Single counters, which count all particles that enter from all directions, regardless of their energy, provide good statistics [*e.g.* Cane, 1993, Kühl et al., 2015]. However, these measurements are often contaminated by increased solar energetic particle flux from the ICME shock. Spacecraft measurements are therefore suitable for small Forbush decreases caused by ICMEs without shocks, whereas ground based measurements are more suited for large shock-associated effects.

The physical mechanism behind the modulation of cosmic rays can be described in general by a transport equation [Parker, 1965] which combines four different contributions: (1) diffusion across field lines due to magnetic field irregularities, (2) particle drifts, (3) convection by the solar wind, and (4) energy loss due to the expansion of the magnetic field. Parker [1965] proposed that the modulation of GCRs can be explained by their

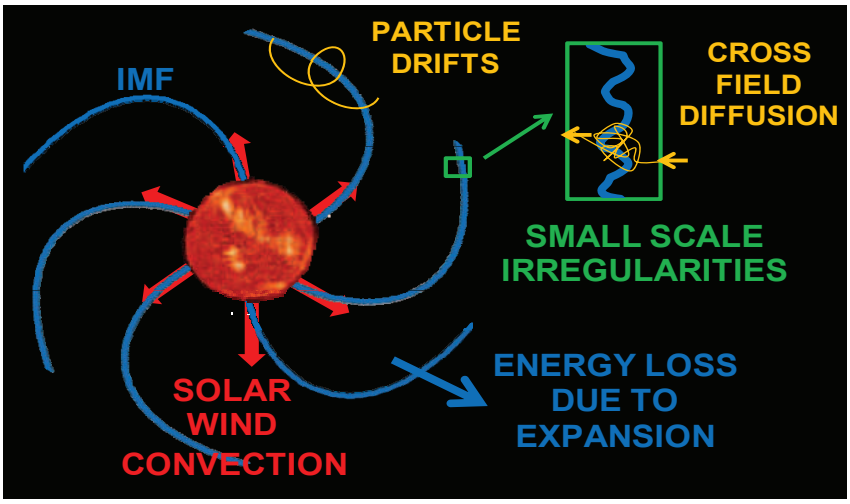


Figure 1.11.: A schematic overview of the physical processes governing modulation of cosmic rays based on the transport theory.

random walk in the frame of reference of the small-scale magnetic irregularities, which are known to be present in the interplanetary magnetic field. To describe the random walk he used the classical probability distribution of the particle whose change can be described by Fokker-Planck equation [see also Jokipii, 1971, Dröge, 2000]. In general case where the scattering frequency is not small compared to the gyration frequency of the particle, the motion of the particle in the frame of reference is described by random walk back and forth along a line of force (guiding center drifts) as well as with diffusion across the field lines. Since the interplanetary magnetic field is "frozen in" the solar wind (*i.e.* they move together) there is a collective movement of particles with the solar wind as seen outside of the frame of reference, *i.e.* convection by the solar wind. Another consequence of the "frozen in" condition is that the interplanetary magnetic field is expanding due to the movement of the solar wind. As their frame of reference is expanding, the particle momentum declines, *i.e.* particles lose their kinetic energy. The physical processes of the transport equation are shown in Figure 1.11.

The transport theory of GCR modulation was applied to explain Forbush decreases as well, where the distinction has to be made between the modulation of the shock/sheath region and ejecta of the ICME [*e.g.* Wibberenz et al., 1998, Cane, 2000]. The disturbances in the GCR distribution are treated as deviations from equilibrium caused by local variations in one or more transport parameters. The shock/sheath region can be regarded as the propagating diffusive barrier, where the decrease in the GCR flux starts at the shock boundary, but the recovery continues even after the passage of the barrier [*e.g.* Chih and Lee, 1986, Le Roux and Potgieter, 1991, Wibberenz et al., 1997, and references therein]. It was proposed that the primary reason for the depression caused by the magnetic ejecta is the closed magnetic field structure of the flux rope, which is assumed to be empty of GCRs close to the Sun. During its propagation it fills up slowly by GCRs entering the flux

rope by perpendicular diffusion, which can also include contribution from particle drifts. The decrease is confined within the borders of the magnetic ejecta [*e.g.* Cane et al., 1995, Kuwabara et al., 2009, Kubo and Shimazu, 2010, , and references therein]. All of these approaches are based on a convection-diffusion concept of the transport theory, where it is expected that the magnitude of the depression is related to diffusion coefficient (governed by the magnetic field strength and fluctuations) and convection by solar wind, *i.e.* the solar wind speed. Statistical studies confirm that the magnitude of the depression is related to magnetic field strength [*e.g.* Badruddin et al., 1986, Belov et al., 2001, Richardson and Cane, 2011a, Dumbović et al., 2012a,b, Blanco et al., 2013b], fluctuations [*e.g.* Badruddin et al., 1986, Dumbović et al., 2012a,b], and speed [*e.g.* Badruddin et al., 1986, Richardson and Cane, 2011a, Dumbović et al., 2012b, Blanco et al., 2013b]. However, the significance of the correlation for different parameters varies greatly study-to-study. Shock associated magnetic clouds are found to be most GCR-effective, *i.e.* produce the strongest depressions [*e.g.* Richardson and Cane, 2011a, Blanco et al., 2013a], analogously to the largest geomagnetic storms (see Section 1.2.1). Almost all major storms detected in the last 50 years were associated with Forbush decreases [Vennerstrom et al., 2015, Lefevre et al., 2015]. Therefore, the study and prediction of Forbush decreases may improve our knowledge and forecast of the geo-effects.

Correspondingly to geomagnetic storms, there are efforts to predict Forbush decreases based on the remote observations of CMEs. Since FD magnitude is dependent on the magnetic field and speed of the ICME, these should be derived from the initial CME properties during their liftoff. However, neither the magnetic field nor the true CME initial speed are directly observable. Chertok et al. [2013] used the magnetic flux at the photospheric level beneath EUV dimmings and post eruption arcades associated to CMEs as a measure of a CME magnetic field and obtained a good correlation with the FD magnitude. However, only a fraction of CMEs is associated to EUV dimmings; moreover, there are CMEs without any chromospheric or low coronal signatures (stealth CMEs, see Section 1.1.2). Recent studies have shown that using white light coronagraphic observations of CMEs can relate their properties to FD magnitudes. FD magnitude was found to be larger for faster CMEs [Blanco et al., 2013a, Belov et al., 2014], CMEs with larger apparent width [Kumar and Badruddin, 2014, Belov et al., 2014] and CMEs with larger mass [Belov et al., 2014]. In addition, Belov [2009] found that sources of the largest Forbush effects are usually located in the central part of visible solar disc. However, the relations between observational CME properties and FDs are much weaker compared to ICME-FD relations.

2. Data selection, measurement method and event list

For the purpose of this study, a large sample of events was compiled. CMEs were first associated to solar flares, and then to geomagnetic as well as cosmic ray response at the Earth. In this chapter a detailed explanation of the event data selection, association methods and measurements is given. Since this research has already been published in a peer-reviewed journal *Solar Physics*, this chapter abundantly contains citations from Dumbović et al. [2015a] and Dumbović et al. [2015b]).

The CME data was taken from the SOHO LASCO CME Catalog (see Section 1.1.1). The solar flare data was taken from the NOAA X-ray solar flare list available at `ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features` (under "/solar-flares"). CMEs were associated with solar flares in the time period 10 January 1996 – 30 June 2011 (hereafter "the SOHO era") using an automated method based on temporal and spatial criteria as described in Vršnak et al. [2005]. The temporal criterion is used to associate a CME with all the flares within the ± 1 hour period of the CME liftoff time, where liftoff time is derived by back-extrapolation of the CME height-time plot (available in the SOHO LASCO CME Catalog) to the solar surface assuming a linear speed. The spatial criterion associates a CME with all flares that were located within the opening angle of a CME, where the CME opening angle is a projection of the CME apparent width on the solar disc, centred around the central position angle of the CME obtained from LASCO-catalog. Therefore, the spatial criterion could not be used for halo CMEs (due to their apparent width of 360 degrees) and solar flares for which the location was not reported. Starting with a total of 16824 CMEs and 25907 flares in the SOHO era (reported by LASCO-catalog and NOAA Xray solar flare list, respectively) we first applied a temporal criterion to associate CMEs and flares. Then, the spatial criterion was used for the applicable events, resulting in a sample of 1392 CMEs and 1617 associated flares, meaning that some CMEs were associated with more than one flare. For those cases, the associated flare of the strongest intensity was chosen, resulting in 1392 CME-flare pairs. All but 38 pairs had a source position identified on the visible side of the Sun, meaning that they were front sided events. The remaining 38 CMEs for which the source position was not available, are halo CMEs; therefore, the association with flares was taken from the HALO CME SOHO LASCO catalog available at

http://cdaw.gsfc.nasa.gov/CME_list/halo/halo.html.

The association of CME–flare pairs with ICME effects at Earth cannot be efficiently done using an automated method, as explained in Section 1.1.3. Associating 1392 CME–flare pairs with geomagnetic and cosmic ray responses would be a severly time-consiming process. A subsample was thus selected, suitable to describe a CME-flare ensamble. The subsample of CME–flare pairs consists of CMEs with speeds larger than 400 km s^{-1}. From all the CMEs, we selected 211 events in order to equally cover the range of velocities (from 400 km s^{-1} to the fastest CMEs, *i.e.* $v > 1500$ km s^{-1}). Equal sampling was used due to the fact that 78% of CMEs in the sample of 1392 CME–flare pairs have speed less than 800 km s^{-1} (53% of CMEs have speed less than 500 km s^{-1}). Furthermore, previous studies have shown that faster CMEs are more geo-effective (see Section 1.2.1). Therefore, using a random sample would include only a small number of large geomagnetic storms in the sample, *i.e.* most interesting events. For this purpose all fast CMEs ($v > 1500$ km s^{-1}) were taken, including a total of 53 events, whereas for CMEs with 400 km s^{-1} $< v <$ 1500 km s^{-1} approximately 30 CMEs were randomly selected per bin of $\Delta v = 200$ km s^{-1}. It should be noted that the cases when slower CMEs are likely to be overtaken by faster ones were also taken into consideration. However, these CMEs launched in quick succession were not treated as individual events (see "TRAIN" events below). CMEs with less than three height-time measurements were discarded, due to uncertainty of the speed estimate. This criterion was relaxed in the case of very fast CMEs ($v > 1500$ km s^{-1}), where only two height-time measurements are not unusual [see SOHO LASCO CME Catalog Yashiro et al., 2004].

Using plots available on the SOHO LASCO CME Catalog, which associate the CME height-time measurement and the *Dst* index, we link the *Dst* events with CME–flare pairs (see an example shown in Figure 2.1). An extrapolation to the distance of 214 solar radii (approximately the distance from the Sun to Earth) was performed using the CME "height-time" to derive a proxy time of arrival to the Earth. A *Dst* event was then sought in a specific time window, chosen to account for possible errors in the SOHO LASCO CME Catalog speed measurements, influence of the drag and geometrical effects (see Section 1.1.3). For CMEs in the speed range $v = 400 - 600$ km s^{-1} the time window starts 24 hours before and ends 36 hours after the proxy of the arrival time. For CMEs with speed $v > 600$ km s^{-1} the time window starts 6 hours before and ends 48 hours after the proxy of the arrival time. In this case a longer time beyond the time of estimated arrival was assumed because of the drag-decceleration effect and possible delayed impact of the flank, both of which depend on the speed of the CME (see Section 1.1.3 for a more detailed explanation of these effects). Within the time window, the *Dst* index was measured at the point where it reaches the minimum value (*Dst* timing). If there was no geomagnetic storm within the time window corresponding to a specific CME, any recognizable variation in the *Dst* index ($|Dst| \geq 10$ nT) closest to the proxy of arrival time (within the time

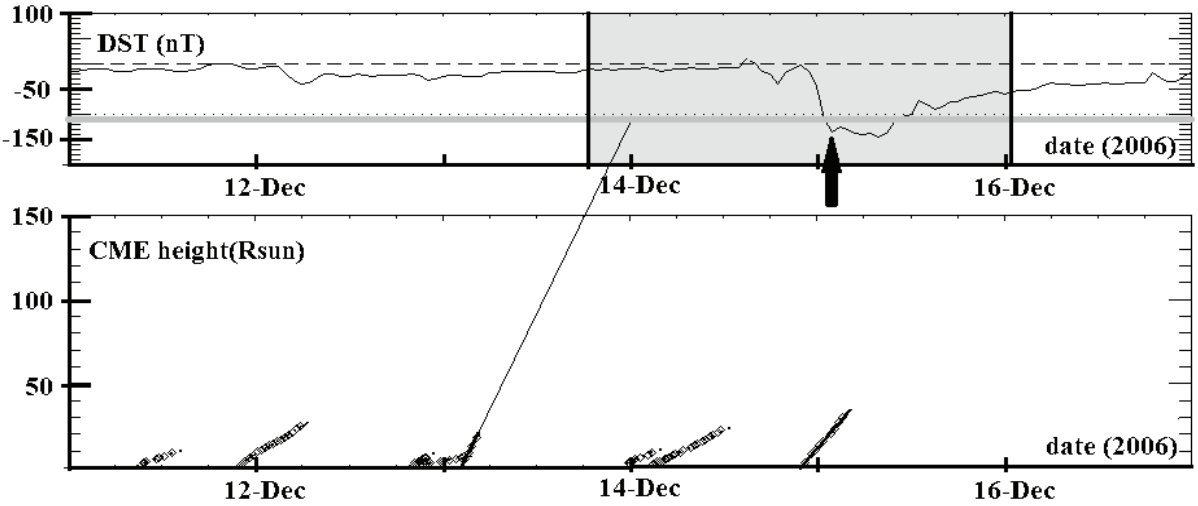


Figure 2.1.: Association of a flare-related CME (first LASCO-C2 appearance 13 December 2006, 02:54 UT) with a *Dst* event at Earth. The lower panel shows CME height-time measurements, whereas the upper panel shows *Dst*–time plot. The CME height-time curve (black solid line) is extrapolated to 1 AU (marked by gray solid line in the upper panel). The shaded area represents the time window in which a *Dst* event was sought (6 hours before and 48 hours after the proxy of arrival at Earth). Black arrow denotes the time at which the *Dst* level is measured [Figure taken from Dumbović et al., 2015a].

window) was taken as the associated *Dst* level. The *Dst* timing in those cases is not a reliable parameter; therefore, the temporal aspect of geomagnetic storms (*e.g.* duration) is not included in the analysis. If there was no variation in *Dst* index throughout the time window which could be associated to a specific CME, the value of the *Dst* index at the proxy of arrival time was taken as the associated *Dst* level. In the *Dst* – time plot a geomagnetic storm is seen as a decrease in the *Dst* index, where the intensity of the storm is given by the magnitude of the decrease. The magnitude of this decrease was measured from the reference value at the start of the storm and this value, *Dst* magnitude, was used for the study.

For each CME in the subsample a level of interaction with other CMEs was determined based on the following criteria:

- *the kinematic criterion* – interacting CMEs are associated with flares originating at the visible side of the Sun and their extrapolated kinematic curves cross or meet each other;

- *the timing criterion* – the liftoff of interacting CMEs is within a reasonable time window ($\approx$ 2 days);

- *the source position/width criterion* – interacting CMEs originating from the same or neighbouring source region, *i.e.* have close locations (unless halo and partial halo CMEs are involved, in which case this criterion was relaxed due to the fact that they have similar directions, *i.e.* they are presumably Earth-directed).

It should be noted that the listed criteria do not mean that CMEs necessarily interacted, they are used only to characterise the CMEs which are likely to interact. The kinematic criterion is based on the linear extrapolation of observed kinematic curves, without considering the drag effect. Furthermore, for simplicity we consider only flare-associated CMEs, for which the source location on the visible side of the Sun is identified. The timing criterion is introduced to prevent the unrealistically long chains of possibly interacting CMEs (*e.g* a "CME1" kinematically interacts with a "CME2" that was launched a day before, which interacted with a "CME3" that started a day prior to "CME2", *etc.*). Finally, a source position/width criterion resolves cases where, *e.g* two narrow CMEs from opposite limbs satisfy both kinematic and timing criterion, although they are unlikely to interact due to their different propagation directions.

These criteria in many cases do not clearly indicate a possible interaction therefore we introduce the "interaction parameter" by which we specify four levels of "interaction probability":

- "SINGLE" (S) events - no interaction;
- "SINGLE?" (S?) - interaction not likely;
- "TRAIN?" (T?) - probable interaction;
- "TRAIN" (T) - interaction highly probable.

The determination of the interaction parameter is illustrated in Figure 2.2. The fastest CME (CME1, first appearance in LASCO-C2 15 June 2000, 07:54 UT) was a partial halo CME launched from N16W55; its proxy arrival time is marked with a black dot. It is preceded by three slower flare-related CMEs launched from source positions (chronologically backwards) N23W90 (CME2), N22W74 (CME3), and N21W69 (CME4), within a period of ≈ 2 days prior to the liftoff of the CME1. The extrapolated kinematic curve of CME1 crosses those of CME2 and CME4, but not of CME3. On the other hand, the extrapolated kinematic curves of CME3 and CME4 cross each other, whereas kinematic criterion for CME4 and CME2 is not met. Furthermore, CME2 is a narrow CME with source position at the limb, so we associate CME1 with an interaction level "T?" (interaction likely). The interaction parameter is assigned to each CME in the subsample of 211 events. We note that the whole CME train is then treated as one event that is characterized by solar parameters (*e.g.* speed, width, flare association, *etc.*) of the fastest CME within a train.

In situ plasma and magnetic field measurements were associated with *Dst* events. For this purpose we used the ICME list by Richardson and Cane [2010] available at: `http://www.srl.caltech.edu/ACE/ASC/DATA/level3/icmetable2.htm`.
In addition, we use *in situ* data from *Advanced Composition Explorer* satellite [ACE; Stone et al., 1998]) *Magnetometer* [MAG; Smith et al., 1998]) and *Solar Wind Electron, Proton, and Alpha Monitor* [SWEPAM; McComas et al., 1998]) instruments available at:

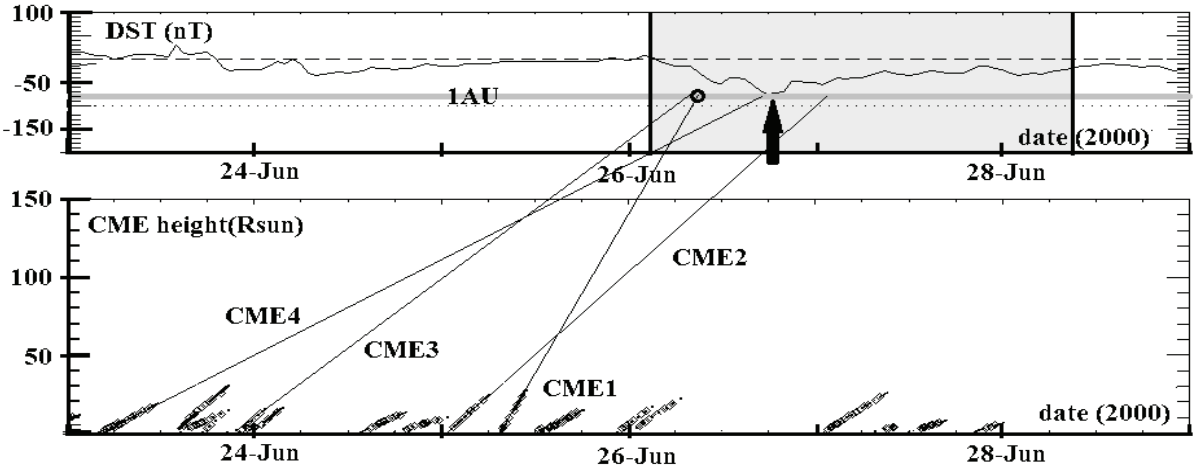


Figure 2.2.: Association of a group of flare-related CMEs with a *Dst* event at Earth. We associate the fastest of the CMEs (CME1) with an interaction parameter "T?" (interaction likely) due to possible interaction with CMEs 2-4 based on the interaction criteria (for details see the main text). The *Dst* level is estimated as in Figure 2.1 [Figure taken from Dumbović et al., 2015a].

`http://www.srl.caltech.edu/ACE/ASC/level2/lvl2DATA_MAG-SWEPAM.html`.
Due to occasional data gaps we supplement *ACE* measurements with measurements from *Magnetic Field Investigation* [MFI; Lepping et al., 1995]) and *Solar Wind Experiment* [SWE; Ogilvie et al., 1995]) instruments onboard *Wind* satellite available at: `http://wind.gsfc.nasa.gov/mfi_swe_plot.php`.
In this way we also checked if some of the geomagnetic storms with $|Dst| > 100$ nT were caused by a corotating interaction region (CIR) (see Section 1.1.3).

In the following analysis, CMEs associated with $|Dst| < 100$ nT are considered as non-relevant events of low geo-effectiveness. They either missed the Earth or did not produce a major storm. Although some of them are in fact associated with CIRs, from the prediction point of view, it is only relevant that they did not produce a geomagnetic storm with $|Dst| > 100$ nT, which is considered as the threshold for relevant strong geomagnetic activity. In our sample of 211 CME–flare pairs the majority of the events were associated with ICMEs (57%), whereas 41% of events could not be associated with clear ICME signatures, *i.e.* they were either CIRs, complex ejecta, or there was no *in situ* event at all. For 2% of events *in situ* data were not available due to measurement gaps. Out of 41% of events that were not associated with clear ICME signatures, only one had $|Dst| > 100$ nT; however, we did not discard it because it does not have clear CIR signatures as well.

Next we supplement the list with cosmic ray responses at Earth. For each event on the list we searched for a corresponding response in the relative pressure-corrected co-

smic ray (CR) count measured by the ground-based neutron monitor (NM) data taken from the *Space Physics Interactive Data Resource* (SPIDR) database available at `http://spidr.ngdc.noaa.gov/spidr/`.

We searched the time interval spanning 5 days before and 15 days after the reported *Dst* anomaly to find a response in the CR count (if there is one). To reduce the effect of daily variations we used the average of 3-4 different mid-latitude NM stations (depending on data availability) at different asymptotic longitudes, but of similar rigidity (Novosibirsk, Calgary, Kiel, and Magadan, with vertical cutoff rigidity 2.91 GV, 1.09 GV, 2.29 GV, and 2.10 GV, respectively; for method description see Section 1.2.2). This method reduces daily variations, but does not remove them completely. Therefore, a threshold of 1% (comparable to the daily variation amplitude) is chosen for the cosmic-ray response. In analogy with geo-effectiveness of CMEs, a measure of geomagnetic response, we adopt the term GCR-effectiveness [used by Kumar and Badruddin, 2014]) as a measure of the cosmic-ray response. If a clear depression in the CR count with a magnitude >1% (from the onset point to the time of maximum decrease) is observed around the reported *Dst* timing (*i.e.* the time of the minimum *Dst* is within the FD duration interval), the event is regarded as GCR-effective, otherwise it is not regarded as GCR-effective. Thus we treat small FDs (<1%) in the same way as "missing" FDs (*i.e.* when there is no event because CME did not arrive at the Earth). The CR counts were normalized to the CR count in the quiet period before any disturbance. In some cases, where two consecutive geomagnetic storms could be identified separately, but only one Forbush decrease is observed, the two events are merged into one event, which is then regarded as interacting CMEs event (interaction parameter, $i = 4$). In such cases involved CMEs are treated as one event, characterized by solar parameters of the fastest CME involved in the interaction and with the apparent width of the widest involved CME. This, in addition to data gaps for several events, resulted in a new list of 187 CME-flare-Dst-FD associations (FD list in further reading), which is used for the Forbush decrease study, whereas the list of 211 CME-flare-Dst associations is used for the geomagnetic storm study (Dst list in further reading). These lists are available online as a merged CME-flare-Dst-FD list at: `http://oh.geof.unizg.hr/FDFT/fdft.php`.

3. Statistical analysis

The following analysis is focused on specific remote CME–flare parameters, which were previously found relevant for geomagnetic and cosmic-ray response (see Sections 1.2.1 and 1.2.2). These are the initial CME speed, v, the apparent angular width, w, the solar flare soft X-ray intensity peak value, f, the source position (radial distance from the center of the solar disc), r, and finally a level of CME-CME interaction, i. None of these parameters show a strong correlation with Dst or FD magnitude (Dst and FD, respectively). Therefore, a probabilistic approach is utilized. It should be noted again that this research has already been published in a peer-reviewed journal *Solar Physics*. Therefore, this section abundantly contains citations from Dumbović et al. [2015a] and Dumbović et al. [2015b]).

Distributions are used as a statistical tool for the analysis of geomagnetic and cosmic-ray response, where Dst and FD magnitudes are grouped into four corresponding discrete bins, each bin representing a different level of geo-effectiveness and GCR-effectiveness, respectively.

Dst magnitudes were grouped into following levels of geo-effectiveness:

- $|Dst| < 100$ nT (not geo-effective);
- 100 nT $< |Dst| <$ 200 nT (moderatly geo-effective);
- 200 nT $< |Dst| <$ 300 nT (strongly geo-effective);
- $|Dst| >$ 300 nT (intensly geo-effective);

FD magnitudes were grouped into following levels of GCR-effectiveness:

- $FD < 1\%$ (not GCR-effective);
- $1\% < FD < 3\%$ (moderately GCR-effective);
- $3\% < FD < 6\%$ (strongly GCR-effective);
- $FD > 6\%$ (intensly GCR-effective).

Table 3.1.: CME/flare parameters binning and corresponding number of events based on the Dst list

CME/flare parameter	original bins		alternative bins	
	bin	Number of events	bin	Number of events
CME apparent width, w			$< 70°$	29
	$< 120°$	59	$70° - 130°$	32
	$120° - 360°$	35	$130° - 360°$	33
	$360°$	117	$360°$	117
CME initial speed, v	$400 - 600$ km s^{-1}	36		
	$600 - 800$ km s^{-1}	34	$400 - 700$ km s^{-1}	52
	$800 - 1000$ km s^{-1}	35	$700 - 1000$ km s^{-1}	54
	$1000 - 1200$ km s^{-1}	35	$1000 - 1500$ km s^{-1}	52
	$1200 - 1700$ km s^{-1}	41	> 1500 km s^{-1}	53
	> 1700 km s^{-1}	30		
CME/flare source position distance from the center of the solar disc, r			< 0.35 R$_{\mathrm{SUN}}$	35
	< 0.4 R$_{\mathrm{SUN}}$	45	$0.35 - 0.5$ R$_{\mathrm{SUN}}$	38
	$0.4 - 0.6$ R$_{\mathrm{SUN}}$	53	$0.5 - 0.65$ R$_{\mathrm{SUN}}$	37
	$0.6 - 0.8$ R$_{\mathrm{SUN}}$	53	$0.65 - 0.78$ R$_{\mathrm{SUN}}$	33
	> 0.8 R$_{\mathrm{SUN}}$	60	$0.78 - 0.92$ R$_{\mathrm{SUN}}$	37
			> 0.92 R$_{\mathrm{SUN}}$	31
solar flare soft X-ray flux peak value, f			$< 2.5 \cdot 10^{-6}$Wm^{-2}	31
	$< 10^{-5}$Wm^{-2}	98	$2.5 - 5 \cdot 10^{-6}$Wm^{-2}	40
	$10^{-5} - 10^{-4}$Wm^{-2}	74	$5 - 12 \cdot 10^{-6}$Wm^{-2}	34
	$\geq 10^{-4}$Wm^{-2}	39	$12 - 30 \cdot 10^{-6}$Wm^{-2}	35
			$30 - 100 \cdot 10^{-6}$Wm^{-2}	36
			$> 100 \cdot 10^{-6}$Wm^{-2}	35
interaction parameter, i	$i = 1$	98	$i = 1$ & $i = 2$	132
	$i = 2$	34	$i = 2$ & $i = 3$	62
	$i = 3$	28	$i = 3$ & $i = 4$	79
	$i = 4$	51		

The selected CME/flare parameters were binned as well. For some parameters the binning was obvious (*e.g.* interaction parameter) as they are already discrete parameters. For continuous parameters all the bins contain approximately the same number of events; therefore, these bins are not equidistant. Thus, the binnings used in Sections 3.1 and 3.2 slightly differ (the two lists do not contain the same number of events).

Table 3.2.: CME/flare parameters binning and corresponding number of events based on the FD list

CME/flare parameter	original bins		alternative bins	
	bin	Number of events	bin	Number of events
CME apparent width, w	$< 120°$	56	$< 60°$	20
	$120° - 360°$	32	$60° - 100°$	22
	$360°$	99	$100° - 140°$	25
			$140° - 360°$	21
			$360°$	99
CME initial speed, v	$400 - 650$ km s^{-1}	31		
	$650 - 800$ km s^{-1}	27	$400 - 700$ km s^{-1}	40
	$800 - 1000$ km s^{-1}	31	$700 - 1000$ km s^{-1}	49
	$1000 - 1200$ km s^{-1}	33	$1000 - 1500$ km s^{-1}	50
	$1200 - 1700$ km s^{-1}	34	> 1500 km s^{-1}	48
	> 1700 km s^{-1}	31		
CME/flare source position distance from the center of the solar disc, r	< 0.35 R$_{\mathrm{SUN}}$	28		
	$0.35 - 0.5$ R$_{\mathrm{SUN}}$	30	< 0.45 R$_{\mathrm{SUN}}$	49
	$0.5 - 0.6$ R$_{\mathrm{SUN}}$	25	$0.45 - 0.65$ R$_{\mathrm{SUN}}$	44
	$0.6 - 0.75$ R$_{\mathrm{SUN}}$	34	$0.65 - 0.85$ R$_{\mathrm{SUN}}$	41
	$0.75 - 0.9$ R$_{\mathrm{SUN}}$	32	> 0.85 R$_{\mathrm{SUN}}$	53
	> 0.9 R$_{\mathrm{SUN}}$	38		
solar flare soft X-ray flux peak value, f			$< 2.5 \cdot 10^{-6}$Wm^{-2}	28
	$< 10^{-5}$Wm^{-2}	84	$2.5 - 5 \cdot 10^{-6}$Wm^{-2}	33
	$10^{-5} - 10^{-4}$Wm^{-2}	67	$5 - 15 \cdot 10^{-6}$Wm^{-2}	34
	$> 10^{-4}$Wm^{-2}	36	$15 - 50 \cdot 10^{-6}$Wm^{-2}	34
			$50 - 150 \cdot 10^{-6}$Wm^{-2}	33
			$> 150 \cdot 10^{-6}$Wm^{-2}	25
interaction parameter, i	$i = 1$	83	$i = 1$ & $i = 2$	108
	$i = 2$	25	$i = 2$ & $i = 3$	46
	$i = 3$	21	$i = 3$ & $i = 4$	79
	$i = 4$	58		

The corresponding Dst and FD distribution mean values are then calculated, which can be correlated with the change in the mean value of the (discrete) CME/flare parameter. As a measure of the scatter within each bin, standard deviations are calculated. We expect to observe large scatter within each bin, *i.e.* a large event-to-event variability, because strong

correlations were not found. This would imply a complex relation to a number of solar parameters, and therefore motivates a probabilistic approach. The statistical significance of results was tested using two-sample t-test (2stt) at the 0.05 level (95% significance) of the test samples. 2stt is based on the normality assumption; however, due to agreeable sample sizes, it is applicable even for non-gaussian distributions [in accordance with the central limit theorem, *e.g.* Dekking et al., 2005].

To support/substantiate the statistical analysis we use the method of overlapping bins. With this method, in addition to the original binning, an alternative binning is used and the results for both are then compared. The alternative binning should lead to the same/similar results as the original binning. Both original and alternative bins, as well as the corresponding number of the events are given in Table 3.1 for Dst list and Table 3.2 for FD list, which are used for the statistical analysis presented in Sections 3.1 and 3.2, respectively.

3.1. The relation between CMEs/flares and geomagnetic storms

The first analyzed parameter is CME apparent width, w. The events in our data set were categorized first into three different CME apparent width bins, following the categorization from the SOHO LASCO CME Catalog (see Section 1.1.1) into non-halo, partial halo and halo CMEs (original bins in first panel of Table 3.1). Due to the fact that (possibly) interacting CMEs are regarded as one entity ("TRAIN" and "TRAIN?" events, see Section 2), these events were associated with the width of the widest CME involved in (possible) interaction. Three Dst distributions were made (Figures 3.1a-c). The mean of each distribution was calculated (black dots in Figures 3.1a-c), as well as the standard deviation (horizontal error bars in Figures 3.1a-c). We see an obvious progression in the Dst distribution towards larger Dst as the apparent width of the CME increases. For non-halos we find one-bin distribution within $Dst < 100$ nT, for partial halos the distribution gains a small tail, whereas for halos a long tail is observed. The distribution mean has an increasing trend with larger widths, although it should be noted that the distribution mean is positioned in the first bin for all the cases (non-halos, partial halos and halos). This implies that halos are more likely to be geo-effective than partial halos and non-halos, but still, most of the halos are not geo-effective. The latter is also reflected by the increased standard deviation for halo CMEs. These results are confirmed with the two sample t-test (2stt), showing that non-halo, partial halo and halo CME associated Dst distributions are significantly different (see first panel of Table 3.3).

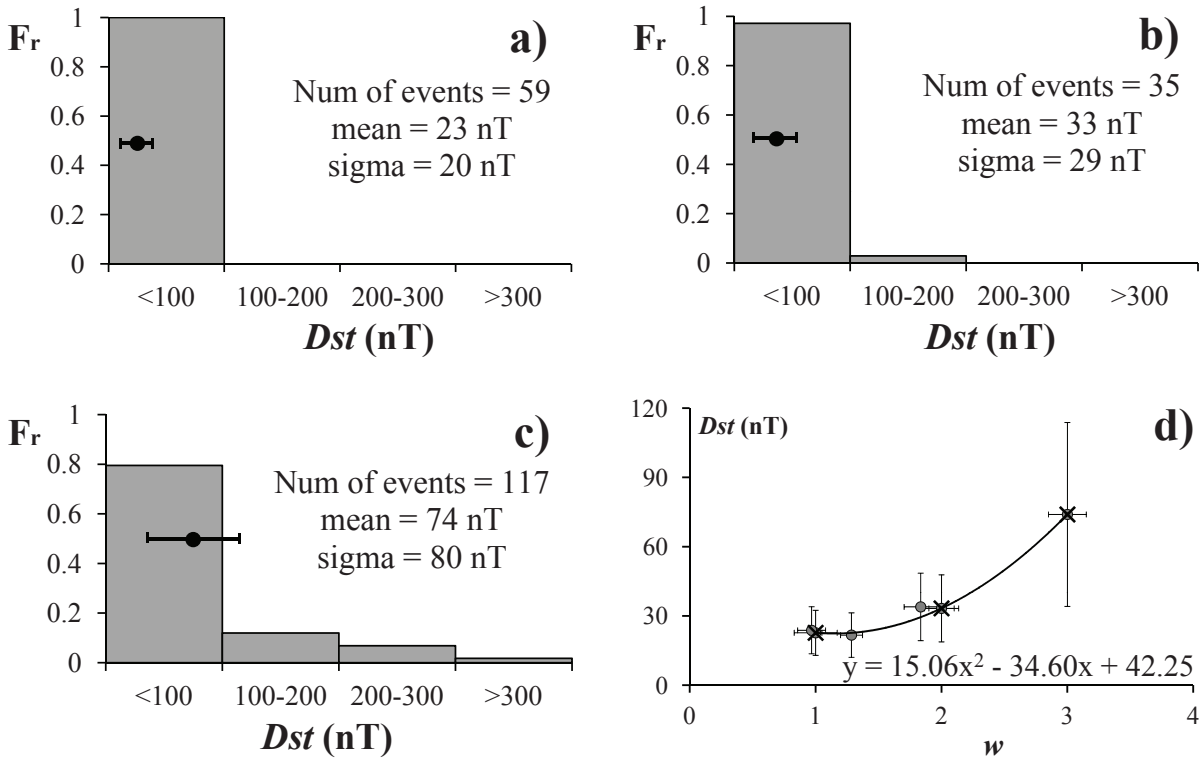


Figure 3.1.: (a-c): *Dst* relative frequencies (Fr) for different *Dst* and CME apparent width, *w* bins: (a) non-halo; (b) partial halo; (c) halo CMEs. A black dot with horizontal error bar marks distribution mean and standard deviation. (d): Dependence of the *Dst* distribution mean values on apparent width, where different width bins were associated with numerical values (non-halo with 1, partial halo with 2, and halo CME with 3). A best fit to original bins (marked by crosses) is presented, as well as the corresponding fitting parameters. Standard deviations are given by the error bars.

Figure 3.1d shows the dependence of the *Dst* distribution mean values on the apparent width for original bins (marked by crosses) and alternative bins (marked by grey circles). The numbers are associated to different width bins for quantitative reasons. The alternative bins show the same behavior as the original bins in Figure 3.1d, but there is a loss in the significance of sample differences between different bins (see first panel of Table 3.3). Halo CMEs remain significantly different from other width bins, whereas other width bins are not significantly different from each other. We attribute this loss of significance to "mixing" of different bin events due to introduction of more bins. We conclude that large apparent width is in general related to a greater geo-effectiveness (in accordance to previous studies, see Section 1.2.1), although a large event-to-event variability is possible. We propose a possible quantification of this relation, based on the quadratic function fitted to the original width bin data (Figure 3.1d).

Table 3.3.: Results of the two sample t-test for Dst distributions corresponding to different bins of CME/flare parameters. Results (p-values) showing statistical significance with more than 95% probability are colored blue.

CME/flare parametar	binning type		two sample t-test p-values (between corresponding bins)				
CME apparent width, w (degrees)	original bins		360	$120-360$			
		<120	0.0001	0.0377			
		$120-360$	0.0036				
	alternative bins		360	$130-360$	$70-130$		
		<70	0.0010	0.1234	0.6838		
		$70-130$	0.0003	0.0517			
		$130-360$	0.0053				
CME initial speed, v ($\mathrm{km\,s^{-1}}$)	original bins		>1700	$1200-1700$	$1000-1200$	$800-1000$	$600-800$
		$400-600$	0.0003	0.0214	0.3174	0.1795	0.3495
		$600-800$	0.0004	0.0323	0.7205	0.3981	
		$800-1000$	0.0217	0.2125	0.6621		
		$1000-1200$	0.0051	0.0989			
		$1200-1700$	0.4340				
	alternative bins		>1500	$1000-1500$	$700-1000$		
		$400-700$	0.0001	0.0393	0.0375		
		$700-1000$	0.0106	0.9873			
		$1000-1500$	0.0112				
CME/flare source position distance from the center of the solar disc, r ($\mathrm{R_{SUN}}$)	original bins		>0.8	$0.6-0.8$	$0.4-0.6$		
		<0.4	0.0004	0.0002	0.0048		
		$0.4-0.6$	0.2466	0.1020			
		$0.6-0.8$	0.8599				
	alternative bins		>0.92	$0.78-0.92$	$0.65-0.78$	$0.5-0.65$	$0.35-0.5$
		<0.35	0.0010	0.0094	0.0027	0.0007	0.0636
		$0.35-0.5$	0.0124	0.2508	0.0601	0.0187	
		$0.5-0.65$	0.5320	0.3313	0.5623		
		$0.65-0.78$	0.2703	0.5785			
		$0.78-0.92$	0.2041				
solar flare soft X-ray flux peak value, f ($\cdot 10^{-6}\mathrm{Wm^{-2}}$)	original bins		≥ 100	$10-100$			
		≤ 10	0.0001	0.1440			
		$10-100$	0.0263				
	alternative bins		≥ 100	$30-100$	$12-30$	$5-12$	$2.5-5$
		≤ 2.5	0.0008	0.0344	0.9995	0.0834	0.7274
		$2.5-5$	0.0001	0.0100	0.6697	0.0235	
		$5-12$	0.0701	0.3268	0.0519		
		$12-30$	0.0001	0.0218			
		$30-100$	0.6328				
interaction parameter, i	original bins		$i=4$	$i=3$	$i=2$		
		$i=1$	0.0004	0.0329	0.0572		
		$i=2$	0.4122	0.9152			
		$i=3$	0.4651				
	alternative bins		$i=3\,\&\,i=4$	$i=2\,\&\,i=3$			
		$i=1\,\&\,i=2$	0.0034	0.0718			
		$i=2\,\&\,i=3$	0.4688				

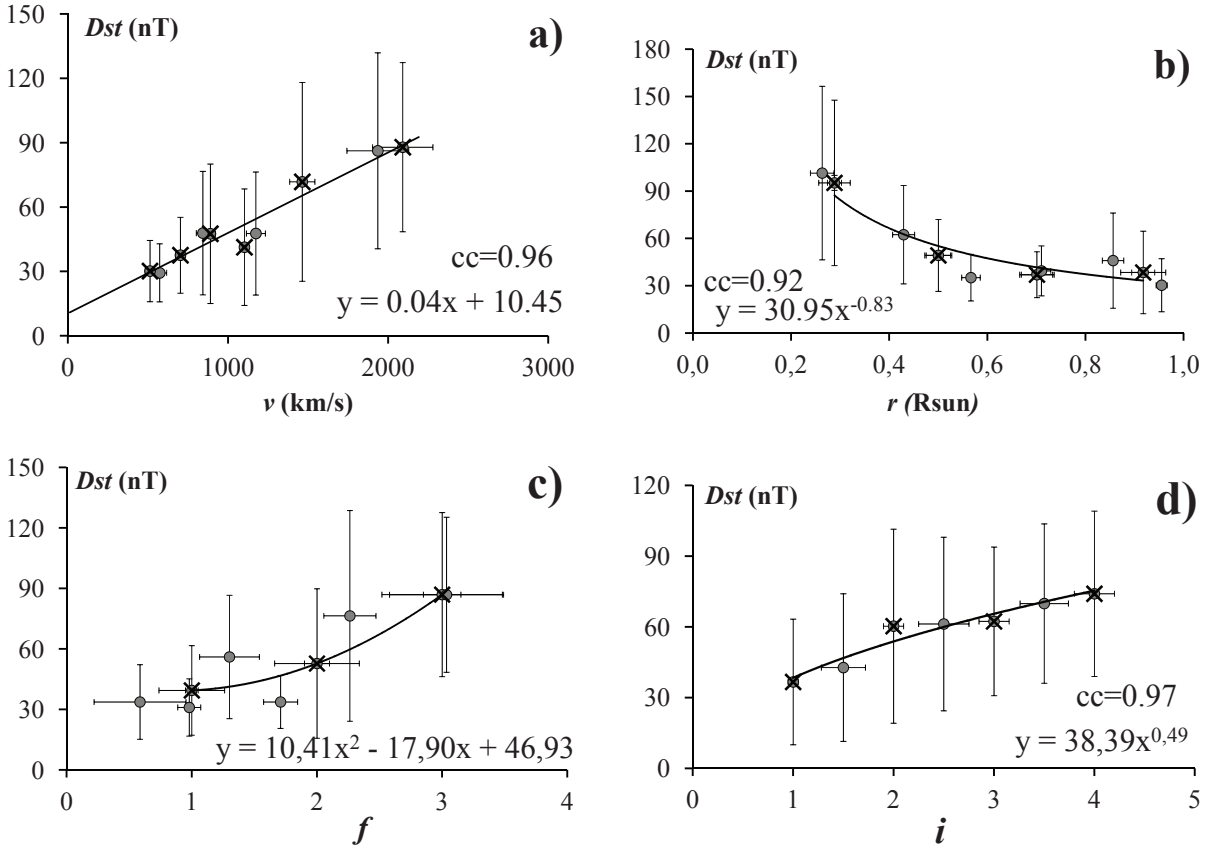


Figure 3.2.: *Dst* distribution mean as a function of the bin-averaged value for: (a) average value of the CME initial speed, v, within a specific bin; (b) average value of the source position distance from the solar disc center, r, within a specific bin; (c) numerical values associated to different flare class, f (B&C-class associated to 1, M-class associated to 2, X-class associated to 3); (d) interaction parameter, i. Best-fit to original binning is given for each of the solar parameter, as well as the corresponding fitting parameters and a correlation coefficient (cc, when applicable). The data corresponding to original bins are marked by crosses. Standard deviations are given by the error bars.

We next consider the CME initial speed, v. The same statistical method was applied as for the CME apparent width, w, and the results are presented in Figure 3.2a and Table 3.3 (second panel). The general trend shows that faster CMEs tend to produce larger *Dst*. However, like for the $w - Dst$ relation, the distribution mean is positioned in the first bin for all the cases indicating that although faster CMEs are more likely to be geo-effective, most of them are not geo-effective. This is also visible from the large standard deviations. The results of the 2stt for original bins show that there is no significant difference between two neighbouring speed bins (or several, as we go to lower speed bins). When the number of bins is reduced (see alternative bins in second panel of Table 3.3), the significance increases, correspondingly to what we observe for CME apparent width, w. Therefore, we conclude that faster CMEs are in general related to a stronger geo-effectiveness (in accordance to previous studies, see Section 1.2.1), although

a large event-to-event variability is possible. We propose a possible quantification of this relation, based on the linear fit to the original speed bin data (Figure 3.2a).

Next, a source distance from the solar disc center, r, was investigated, ranging from 0 to 1 (in units of solar radii). The same statistical method was applied as above (for width, w and speed, v) and the results are presented in Figure 3.2b and Table 3.3 (third panel). The general trend shows that CMEs originating closer to the center of the solar disc tend to produce larger Dst. Analogously to the discussions above (for w and v), the distribution mean is positioned in the first bin for all the cases, indicating that although central CMEs are more likely to be geo-effective, most of them are not geo-effective (also visible from large standard deviations). 2stt for the original bins shows significant differences only for the bin around the solar disc center ($r < 0.4$); however, loss of significance for other bins does not seem stochastic, since there is a decrease in significance as we go towards the near-limb source locations (third panel of Table 3.3). For alternative bins the loss of significance is even more pronounced, which can be explained by the increase of the number of bins and "mixture" of the event types (analogous to the analysis of w and v). We conclude that although a large event-to-event variability is possible, CMEs originating closer to the center of the solar disc are in general related to a stronger geo-effectiveness (in accordance to previous studies, see Section 1.2.1). We propose a possible quantification of this relation, based on the power-law fit to the original source-position bin data (Figure 3.2b).

The next analyzed parameter is the soft X-ray intensity peak value of the associated flare, f. We apply the same procedure as with other parameters and present the results in Figure 3.2c and Table 3.3 (fourth panel). The general trend shows that CMEs associated with stronger flares (*i.e.* flares with a higher soft X-ray intensity peak value) tend to produce larger Dst. Analogously to previous discussions, the distribution mean is positioned in the first bin for all the cases indicating that stronger flares are more likely to be geo-effective, but most of them are not geo-effective and there is a large event-to-event variability, as seen from large standard deviations. This is also reflected by the 2stt for original bins, showing that B & C class flares and M flares are not significantly different samples, but they are both significantly different from X flares (fourth panel of Table 3.3). Increasing the number of bins decreases the significance, as seen in other parameters (see alternative bins at fourth panel of Table 3.3). In spite of the possible large event-to-event variability, we conclude that CMEs associated with stronger flares are generally related to a stronger geo-effectiveness (in agreement with previous studies, see Section 1.2.1). We propose a possible quantification of this relation, based on the quadratic fit to the original flare bin data (where numerical values are associated to different flare class, see Figure 3.2c). We note that in Figure 3.2c we do not present the correlation coefficient, because the fit is based on only three data points (similar to Figure 3.1d).

Finally, the interaction parameter is analyzed using the same statistical method and

the results are presented in Figure 3.2d and Table 3.3 (fifth panel). A general trend shows that CME-CME interaction is related to larger Dst; however, again the distribution mean is positioned in the first bin for all the cases. This indicates that CMEs which are likely to interact are also more likely to be more geo-effective, but most of them are not geo-effective. Very large standard deviations are found, indicating that there is a large mixture between the bins and therefore a large event-to-event variability. This is also reflected in the 2stt for both original and alternative bins, presented in the fifth panel of Table 3.3 (note that the alternative bins are the mixture of the original bins). There is however a statistically significant difference between the interacting and non-interacting CMEs. Thus, we conclude that CMEs which are likely to interact are generally related to a stronger geo-effectiveness (in agreement with previous studies, see Section 1.2.1), although large event-to-event variability is possible. We propose a possible quantification of this relation, based on the power-law fit to the original interaction parameter data (Figure 3.2d).

3.2. The relation between CMEs/flares and Forbush decreases

In the statistical analysis of the relation between CMEs/flares and Forbush decreases the same procedure is used as in Section 3.1. The first analyzed parameter is CME apparent width. Three FD distributions were made (Figures 3.3a-c). The mean of each distribution was calculated, as well as the standard deviation, but unlike in Figure 3.1, they are not marked in the Figure because the distribution bins are not equidistant. They are presented in Figure 3.3d. We see that unlike for Dst, FD distribution already shows a long tail for non-halo CMEs. The distribution mean slightly increases for partial halos; however, there is no prominent change in the shape of the distribution. For halo CMEs both the change in the distribution mean and in the shape of the distribution is prominent. This is also reflected by the w vs FD plot in Figure 3.3d. We see an obvious progression of the FD mean as the apparent width of the CME increases, but the standard deviations are large. Unlike for Dst, the distribution mean is not positioned in the first bin for all the cases, indicating that CMEs are more likely to be GCR-effective than geo-effective. It should be noted though, that this relation between GCR- and geo-effectiveness depends highly on the scale by which it was defined. If the threshold would be 3% instead of 1%, we would observe similar positioning of the distribution mean with respect to bins. The two sample t-test (2stt) for original bins shows that halo CMEs are significantly different from non-halo and partial halo CMEs, but non-halos and partial halos are not significantly different (see first panel of Table 3.4). There is a further loss in significance when the number of the bins is increased (see alternative binning on the first panel of Table 3.4). Therefore, we can conclude that halo CMEs are more GCR-effective than CMEs with

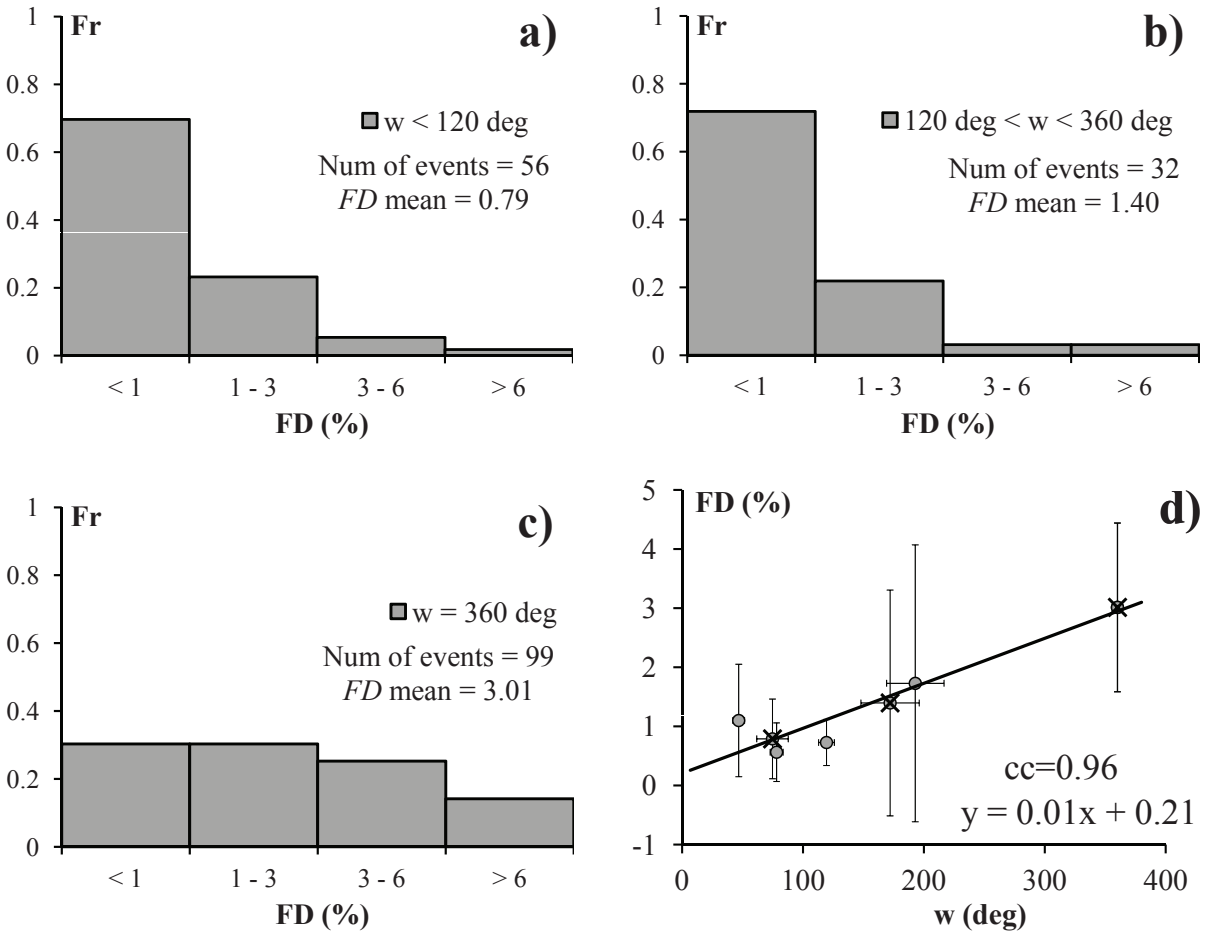


Figure 3.3.: FD relative frequencies (Fr) for different FD and CME apparent width (w) bins (a-c), and FD distribution mean as a function of the bin-averaged value for the CME apparent width (d). A linear fit to all of the data obtained by the method of overlapping bins is given in d) with fitting parameters and a correlation coefficient (cc). The data corresponding to original bins (used for distributions in a-c) are marked by crosses. Standard deviation is given by the error bars [taken from Dumbović et al., 2015b].

smaller apparent width and there is some indication that the level of GCR-effectiveness might be related to CME apparent width (in agreement with previous studies, see Section 1.2.2). We propose to quantify this relation based on the linear fit through all of the data (both original and alternative bins), as shown in Figure 3.3d. Unlike with Dst analysis, we use all of the data in the fit instead of only the original bins due to large scatter (large standard deviations) and reduced level of 2stt significance. In addition, we do not use apparent width as a discrete, but rather a continuous parameter.

Next, we analyze the 1st order (linear) CME speed, v, using the same procedure and present the results in Figure 3.4a and Table 3.4 (second panel). A linear least square fit to all of the data in Figure 3.4a (for both original and alternative bins) shows a strong correlation, indicating that FD magnitude is larger for faster CMEs, although large standard deviations imply a large event-to-event variability. 2stt reveals that fast CMEs

Table 3.4.: Results of the two sample t-test for FD distributions corresponding to different bins of CME/flare parameters. Results (p-values) showing statistical significance with more than 95% probability are colored blue.

CME/flare parametar	binning type	two sample t-test p-values (between corresponding bins)				
CME apparent width, w (degrees)	original bins		360	$120-360$		
		<120	0.0001	0.2823		
		$120-360$	0.0119			
	alternative bins		360	$140-360$	$100-140$	$60-100$
		<60	0.0049	0.5801	0.3778	0.2521
		$60-100$	0.0001	0.2603	0.5281	
		$100-140$	0.0001	0.2978		
		$140-360$	0.1014			
CME initial speed, v ($\mathrm{km\,s^{-1}}$)	original bins		>1700	$1200-1700$	$1000-1200$	$800-1000$	$600-800$
		$400-650$	0.0001	0.0012	0.0587	0.2942	0.0295
		$650-800$	0.0038	0.0480	0.9337	0.2224	
		$800-1000$	0.0001	0.0052	0.2904		
		$1000-1200$	0.0021	0.0314			
		$1200-1700$	0.5900				
	alternative bins		>1500	$1000-1500$	$700-1000$	
		$400-700$	0.0001	0.0423	0.8126	
		$700-1000$	0.0001	0.0164		
		$1000-1500$	0.1934			
CME/flare source position distance from the center of the solar disc, r ($\mathrm{R_{SUN}}$)	original bins		>0.9	$0.75-0.9$	$0.6-0.75$	$0.5-0.6$	$0.35-0.5$
		<0.35	0.0356	0.0096	0.3359	0.4407	0.7598
		$0.35-0.5$	0.0054	0.0006	0.1791	0.2157	
		$0.5-0.6$	0.1671	0.0406	0.7641		
		$0.6-0.75$	0.4299	0.2345			
		$0.75-0.9$	0.5389				
	alternative bins		>0.85	$0.65-0.85$	$0.45-0.65$	
		<0.45	0.0008	0.1486	0.3316	
		$0.45-0.65$	0.0105	0.5004		
		$0.65-0.85$	0.2034			
solar flare soft X-ray flux peak value, f ($\cdot 10^{-6}\mathrm{Wm^{-2}}$)	original bins		≥ 100	$10-100$			
		≤ 10	0.0001	0.0013			
		$10-100$	0.0264				
	alternative bins		≥ 150	$50-150$	$15-50$	$5-15$	$2.5-5$
		≤ 2.5	0.0004	0.0032	0.3226	0.0833	0.1005
		$2.5-5$	0.0001	0.0001	0.0679	0.0027	
		$5-15$	0.0189	0.1754	0.8445		
		$15-50$	0.0409	0.2137			
		$50-150$	0.2266				
interaction parameter, i	original bins		$i=4$	$i=3$	$i=2$	
		$i=1$	0.0029	0.1603	0.9296	
		$i=2$	0.0195	0.2013		
		$i=3$	0.5392			
	alternative bins		$i=3\,\&\,i=4$	$i=2\,\&\,i=3$		
		$i=1\,\&\,i=2$	0.0012	0.3588		
		$i=2\,\&\,i=3$	0.0881			

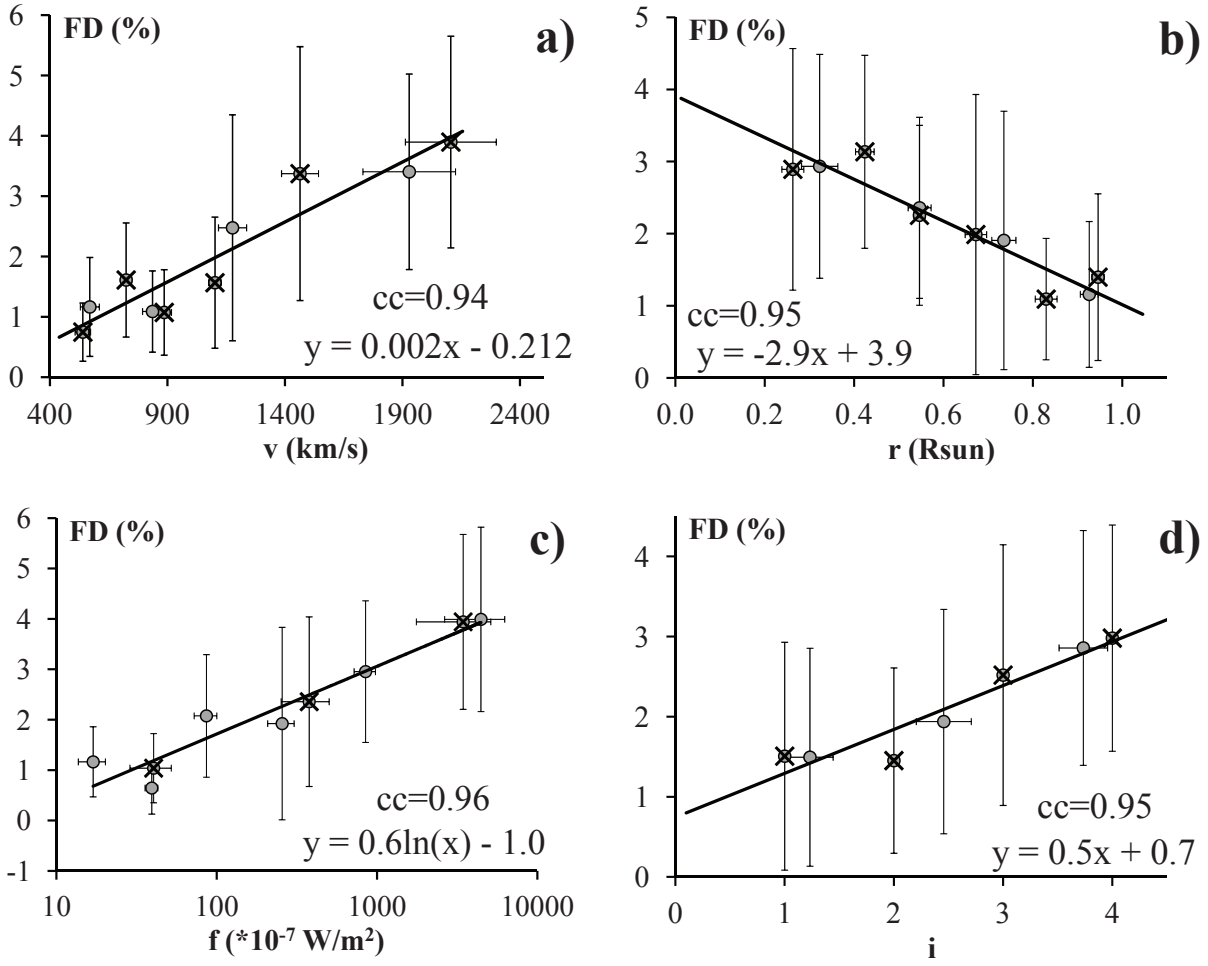


Figure 3.4.: FD distribution mean as a function of the bin-averaged value for: (a) CME initial speed, v, (b) CME/flare source position, r, (c) solar flare Soft X-ray peak intensity (in logarithmic scale), f, and (d) CME-CME interaction level, i. The best-fit to all data obtained by the method of overlapping bins is given for each solar parameter, as well as the corresponding fitting parameters and a correlation coefficient (cc). The data corresponding to original bins are marked by crosses. Standard deviations are given by error bars [taken from Dumbović et al., 2015b].

are significantly different than medium-speed and slow CMEs, whereas there is a loss of significance between medium-speed and slow CMEs. The significance is improved when the number of bins is reduced (see alternative bins in the second panel of Table 3.4). We therefore conclude that faster CMEs are in general more GCR-effective in agreement with previous studies (see Section 1.2.2), although large event-to-event variability is expected. We propose to quantify this relation defined by linear fit through all data (both original and alternative bins), as shown in Figure 3.4a.

A similar analysis was repeated for the CME/flare source position, *i.e.* for the radial distance of the CME/flare source position from the center of the solar disc, expressed in solar radii, r. The results are presented in Figure 3.4b and Table 3.4 (third panel). The linear least square fit to all of the data in Figure 3.4b results in a strong correlation,

but again large standard deviations are observed. 2stt shows that central CMEs are significantly different than non-central CMEs; however, there is a loss of significance between other bins. When the number of bins is reduced, again there is only significant difference for the central CMEs; however, the loss of significance for other bins does not seem to be stochastic, since there is a decrease in significance as one goes towards the near-limb source locations (third panel of Table 3.4). Therefore, we conclude that CMEs originating close to the center of the solar disc are more GCR-effective than CMEs further away from the center and there is indication that the level of GCR-effectiveness is related to CME distance from the solar disc center (in agreement with previous studies, see Section 1.2.2). This relation is quantified using linear fit through all data (both original and alternative bins), as shown in Figure 3.4b.

Next, FD magnitude was related to associated flare strength, *i.e.* soft X-ray flux peak value, f, using the same procedure and results are presented in Figure 3.4c and Table 3.4 (fourth panel). The best fit to all of the data, in spite of the large standard deviations within bins, reveals a logarithmic dependence (3.4c), where FD is found to be larger for stronger flares. 2stt shows that original bins are significantly different, but there is a loss in significance when the number of bins increases (alternative bins, see fourth panel of Table 3.4). Therefore, we conclude that CMEs associated with stronger flares are more GCR-effective and we quantify this relation using linear fit through all of the data (both original and alternative bins), as shown in Figure 3.4c.

Finally, we relate the FD magnitude, FD, to the CME–CME interaction parameter and present the results in Figure 3.4d and Table 3.4 (fifth panel). A linear least square fit to all data in Figure 3.4d (for both original and alternative bins) shows a strong correlation, but, as with all the previous solar parameters, the standard deviations are again large within the bins. Similarly to results for the Dst, 2stt show statistically significant difference between the interacting and non-interacting CMEs, but other bins are not significantly different, probably due to large mixture between the bins. We conclude that CMEs that are likely to interact are generally related to a greater GCR-effectiveness, although a large event-to-event variability is possible. We propose a possible quantification of this relation defined by linear fit through all data (both original and alternative bins), as shown in Figure 3.4d.

4. Empirical statistical models for space weather forecast

The results of the statistical analysis presented in Section 3 are used to construct the distribution of Dst and FD magnitudes for a specific set of remote solar observations of a CME and associated flare which will be used to forecast Dst and FD levels. Again, this section abundantly contains citations from Dumbović et al. [2015a] and Dumbović et al. [2015b].

It was shown in Sections 3.1 and 3.2 (Figures 3.1 and 3.3, respectively) that the Dst and FD distributions are asymmetric, rapidly descending discrete distributions. Therefore, as a mathematical tool, the geometric distribution will be used [see *e.g.* Stirzaker, 2003]. Since the Dst distribution is more asymmetric than the FD distribution, the two procedures are slightly different. For constructing a Dst distribution a regular geometric distribution is used:

$$P(X = k) = p \cdot (1 - p)^{k-1}\,, \tag{4.1}$$

where $P(X = k)$ is the probability that the kth trial is a first success and p is the probability of the success in each trial (k=1,2,3,... is the number of trials). For constructing an FD distribution, a shifted geometric distribution is used:

$$P(X = k) = p \cdot (1 - p)^{k}\,, \tag{4.2}$$

where $P(X = k)$ is the probability that there will be k trials with a failure before the first trial with a success, and p is the probability of the success in each trial (k=0,1,2,... is the number of trials). Both regular and shifted geometric distribution can be simply mathematically reconstructed if the distribution mean is known. The probability of the success in each trial, p, for regular geometric distribution is given by:

$$p = \frac{1}{m}, \tag{4.3}$$

where m is the distribution mean. For shifted geometric distribution p is given by:

$$p = \frac{1}{1 + m}. \tag{4.4}$$

For the same distribution mean, m, the probability of success in each trial, p, is larger for regular, than for the shifted geometric distribution, *i.e.* the regular distribution will be more asymmetric.

In order to use geometric distribution, Dst, *i.e.* FD levels were associated with number of trials, k. The association between the number of trials and the Dst bins was performed in the following way:

- $k = 1 \longleftrightarrow Dst < 100$ nT;
- $k = 2 \longleftrightarrow 100$ nT$< Dst < 200$ nT;
- $k = 3 \longleftrightarrow 200$ nT$< Dst < 300$ nT;
- $k = 4 \longleftrightarrow Dst > 300$ nT.

In this way, the conversion of the Dst distribution mean, m_{Dst}, into the geometric distribution mean, m_{GD}, can be done in a simple way ($m_{GD} = 1 + m_{Dst}$ [nT]/100). Similarly, FD magnitude ranges were associated with the number of trials, k (where FD distribution mean, m_{FD}, can be used as the shifted geometric distribution mean, m_{GD}):

- $k = 0 \longleftrightarrow FD < 1\%$;
- $k = 1 \longleftrightarrow 1\% < FD < 3\%$;
- $k = 2 \longleftrightarrow 3\% < FD < 6\%$;
- $k = 3 \longleftrightarrow FD > 6\%$.

It was shown in Section 3.1 (see Figures 3.1d and 3.2) that the trend of the change in the Dst distribution mean, m_{Dst}, with a specific solar parameter can be fitted by a corresponding function. Similarly in Section 3.2 (see Figures 3.3d and 3.4) the change in the FD distribution mean, m_{FD}, with a specific solar parameter was fitted by a corresponding function. Therefore, for each solar parameter a Dst, *i.e.* FD distribution can be constructed, using Equations 4.1 and 4.3, *i.e.* 4.2 and 4.4, respectively. The obtained empirical distributions are treated as probability distributions. For a specific solar parameter they provide the information on the probability for associating it with a specific value of k, *i.e.* a corresponding Dst/FD level. To combine the effect of the solar parameters, *i.e.* to obtain a joint probability distribution, the key parameters are treated as mutually non-exclusive events, for which the following formula applies:

$$P(\mathrm{A} \cup \mathrm{B}) = P(\mathrm{A}) + P(\mathrm{B}) - P(\mathrm{A} \cap \mathrm{B}). \tag{4.5}$$

In general, $P(\mathrm{X})$, where X = A,B, is the (marginal) probability of the event X, $P(\mathrm{A} \cup \mathrm{B})$ is the probability that either event A or event B or both occur, and $P(\mathrm{A} \cap \mathrm{B})$ is their

joint probability. Specifically, in our case $P(\mathrm{X}) = P(X = k)$ is the probability that for a specific solar parameter X a specific Dst/FD level, k, will be observed. It should be noted that since a particular solar parameter is tied to the same event, they should be regarded as mutually non-exclusive. In general, joint probability is given by:

$$P(\mathrm{A} \cap \mathrm{B}) = P(\mathrm{A}|\mathrm{B}) \cdot P(\mathrm{B}), \tag{4.6}$$

where $P(\mathrm{A}|\mathrm{B})$ is the conditional probability, *i.e.* the probability for event A given that the event B occured. Assuming that the events are independent of each other, combining Equations (4.5) and (4.6) one gets:

$$P(\mathrm{A} \cup \mathrm{B}) = P(\mathrm{A}) + P(\mathrm{B}) - P(\mathrm{A}) \cdot P(\mathrm{B}). \tag{4.7}$$

This assumption is not fully valid, due to the fact that not all key solar parameters are independent of each other (see Section 1.1.1). Since the constructed geometric distribution directly depends on the key solar parameter for which it is obtained, the relation between two solar parameters directly leads to a relation between two constructed geometric distributions. Moreover, positively correlated parameters will lead to conditional probability greater than the marginal probability for larger k ($k = 2, 3, 4$ in case of Dst, and $k = 1, 2, 3$ in case of FD) and *vice versa* for first bin ($k = 1$ in case of Dst, and $k = 0$ in case of FD). Consequently, the assumption of independence redefines parameter space in a way that it will at worst underestimate the joint probability $P(\mathrm{A} \cap \mathrm{B})$, *i.e.* overestimate the probability $P(\mathrm{A} \cup \mathrm{B})$ for larger k and underestimate the probability for smaller k. Therefore, the constructed probability distribution will to some extent increase the number of false alarms.

Finally, the probability of observing the Dst/FD in a specific bin k for a set of solar key parameters is then given by the formula derived from Equation (4.7):

$$\begin{aligned} P(X = k) = \sum_{\alpha} P_\alpha - \sum_{\alpha \neq \beta} P_\alpha \cdot P_\beta + \sum_{\alpha \neq \beta \neq \gamma} P_\alpha \cdot P_\beta \cdot P_\gamma - \\ - \sum_{\alpha \neq \beta \neq \gamma \neq \delta} P_\alpha \cdot P_\beta \cdot P_\gamma \cdot P_\delta + \sum_{\alpha \neq \beta \neq \gamma \neq \delta \neq \epsilon} P_\alpha \cdot P_\beta \cdot P_\gamma \cdot P_\delta \cdot P_\epsilon \,, \end{aligned} \tag{4.8}$$

where X are Dst and FD, respectively, P_α=$P(\alpha)$ represents the probability of Dst/FD level k ($k = 1, 2, 3, 4$ for Dst, and $k = 0, 1, 2, 3$ for FD) for a specific solar key parameter α (the five solar parameters related to Dst and FD are described in Sections 3.1 and 3.2). Based on the Equation 4.8, probabilities of Dst/FD levels can be calculated for a specific set of solar parameters, as explained in more detail in Sections 4.1 and 4.2, respectively.

4.1. Empirical statistical model for geomagnetic storms

Using statistical relations obtained in Section 3.1, *i.e.* Figures 3.1d and 3.2, the association of different Dst levels with values of $k = 1, 2, 3, 4$ (see the introductory part of the Section 4 above), and Equations 4.1 and 4.3, the corresponding geometric distribution was constructed for each solar parameter associated to different original bins (see Section 3). For continuous solar parameters bin-averaged values are used. The solar parameters are CME apparent width, w, CME speed, v, CME/flare source distance from the center of the solar disc, r, flare class, f, and CME–CME interaction level, i. We note that the CME speed, v, and the CME source distance from the center of the solar disc, r, are regarded as continuous parameters in the ranges of $v \geq 400$ km s^{-1} and $0 < r \leq 1$, respectively. The range of v is determined based on the limitations of the sample, whereas the range of r is restricted by the mathematical singularity of the power-law function ($r = 0$) and the physical boundary ($r = 1$, *i.e.* the solar limb). The other three solar parameters, the apparent width, w, the associated flare class, f, and the level of interaction, i, are considered as discrete parameters associated with integers 1–3 and 1–4, respectively (1 meaning least significant, *i.e.* the lowest interaction parameter, width, and flare class).

Table 4.1.: The constants added to geometric distribution to obtain the adjusted distribution, for different Dst bins, k, and different solar parameters: CME speed, v, CME source position distance from the center of the solar disc, r, the apparent width, w, the associated flare class, f, and the interaction parameter, i [taken from Dumbović et al., 2015a].

k	v	r	w	f	i
1	0.13	0.12	0.14	0.15	0.15
2	-0.10	-0.12	-0.12	-0.12	-0.13
3	-0.03	0	-0.02	-0.02	-0.01
4	0	0	0	-0.01	-0.01

The mathematically obtained geometric distribution underestimates the observed Dst distribution for $k = 1$ and overestimates it for $k = 2$. This can be seen in Figure 4.1, where the two are compared for a number of different cases. Therefore, new "adjusted" distributions for each of the key solar parameters are obtained by adding a specific constant to each bin to best fit the observed distribution in all the ranges, *i.e.* for all the values of key solar parameters. These constants are added so that the new distribution is also normalized (Table 4.1) and are different for different Dst bins, k and different solar parameters. It can be seen in Figure 4.1 that the empirical distribution still slightly underestimates the observed Dst distribution for $k = 1$. However, the agreement between the two distributions for higher values of k is substantially improved.

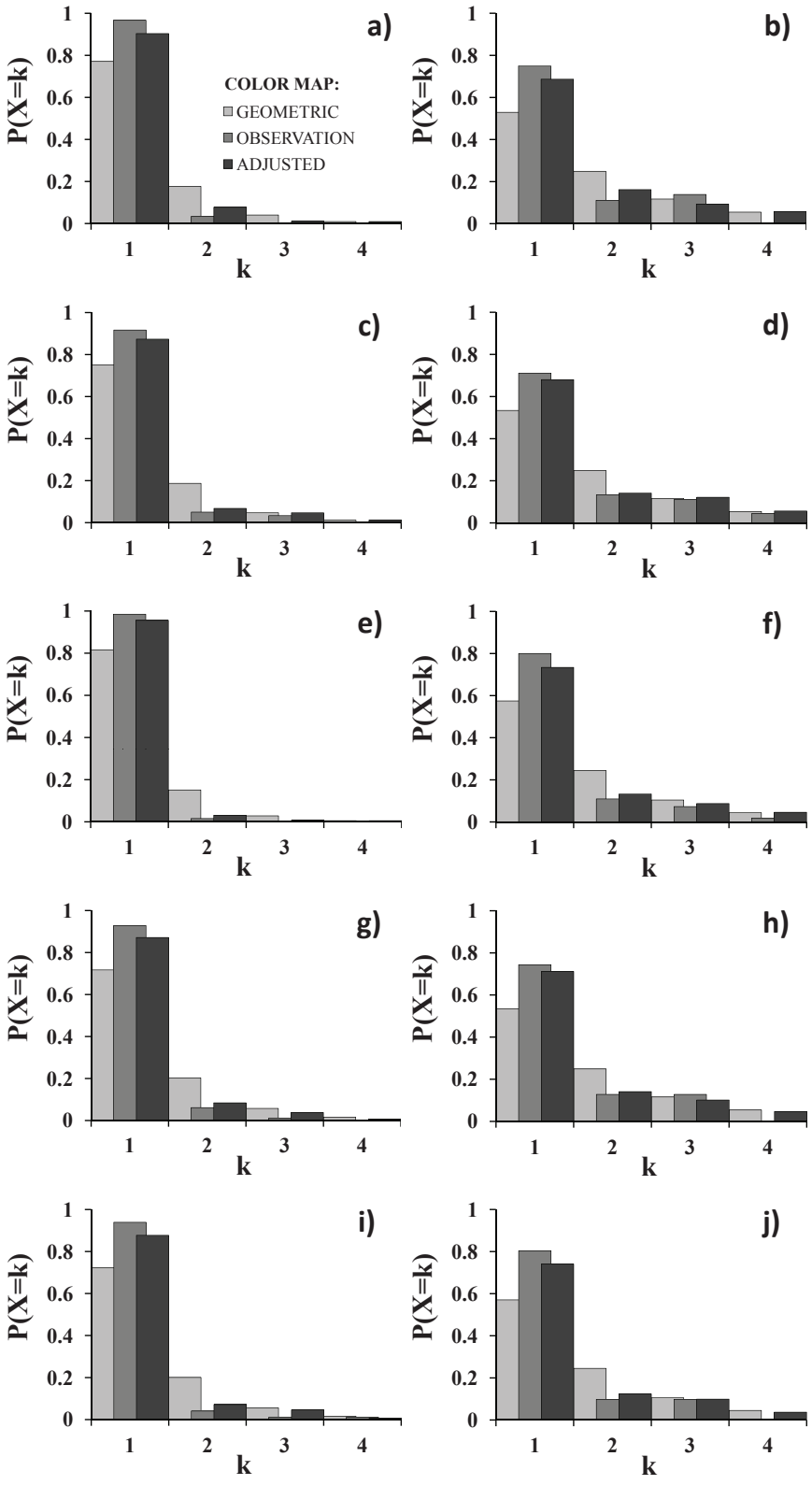


Figure 4.1.: Geometric, observational and adjusted distributions for different ranges/values of key solar parameters: (a-b) for the CME speed in ranges $400 - 600$ and $v >$1700 km s^{-1}; (c-d) for the CME source distance from the center of the solar disc in ranges $r > 0.8$ and $r < 0.4$; (e-f) for non-halo and halo CMEs; (g-h) for B & C-class and X-class associated flares; (i-j) for the lowest and highest interaction parameter, $i = 1$ and $i = 4$, respectively [taken from Dumbović et al., 2015a].

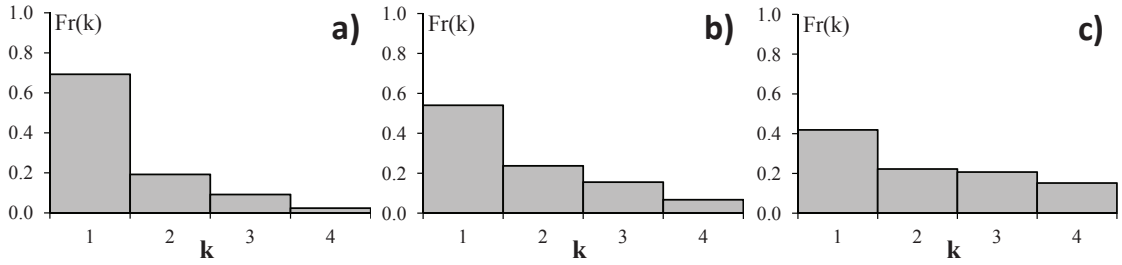


Figure 4.2.: Relative frequencies, $F_r(k)$ for observing Dst in a specific Dst bin, k for different sets of solar parameters: a) $v = 400$ km s^{-1}; $r = 1\mathrm{R_{sun}}$; $w = 1$ (non-halo); $f = 1$ (B or C-class flare); $i = 1$ (no interaction); b) $v = 800$ km s^{-1}; $r = 0.3\mathrm{R_{sun}}$; $w = 2$ (partial halo); $f = 2$ (M-class flare); $i = 3$ (interaction probable); c) $v = 2000$ km s^{-1}; $r = 0.01\mathrm{R_{sun}}$; $w = 3$ (halo); $f = 3$ (X-class flare); $i = 4$ (interaction highly probable). Adapted from [Dumbović et al., 2015a].

In Figure 4.2 we present three different Dst distributions obtained using Equation (4.8), and adding a specific constant to each bin (based on Table 4.1) to obtain "adjusted" distributions. The three distributions correspond to three different solar parameter sets v, r, w, f, i and we can see that the three distributions are different. The probabilities of large geomagnetic storms are higher for faster and wider CMEs which originate near the disc center, are associated with more energetic flares and are likely to be involved in a CME–CME interaction. However, in all three distributions the highest probability is that the event will not be geo-effective, *i.e.* that Dst will be $Dst < 100$ nT ($k = 1$). This depicts the general behavior of CMEs - a large majority of CMEs will never reach the Earth and/or will not have a favorable magnetic field orientation. Therefore, although the model produces a probability distribution, it does not give a straightforward prediction of whether or not (and how strong) a geomagnetic storm will occur. The level of geo-effectiveness needs to be obtained by imposing some criteria (thresholds) on the probability distribution.

Figure 4.2a displays an event with solar parameters shown to be related to low geo-effectiveness and it can be seen that the Dst distribution has a higher value of relative frequency for $Dst < 100$ nT than those in Figures 4.2b and c (which show events of higher geo-effectiveness). In addition, we observe a much higher value of relative frequency for $Dst > 300$ nT in Figure 4.2c (which shows an event with solar parameters related to a high geo-effectiveness). Therefore, thresholds on the value of the relative frequency for a certain bin can be established to enclose certain geo-effectiveness.

To empirically derive the thresholds, we use the Dst list (see Section 2) and calculate Dst distribution for each of the events in the list, based on the corresponding CME/flare parameters. For each event we obtain four different relative frequency values, $F_r(k)$, corresponding to four different distribution bins, $k = 1, 2, 3, 4$. For each relative frequency, $F_r(k)$, we produce a scatterplot against the observed Dst value, where Dst is expressed as one of the four possible Dst ranges associated to four different k (see Section 4).

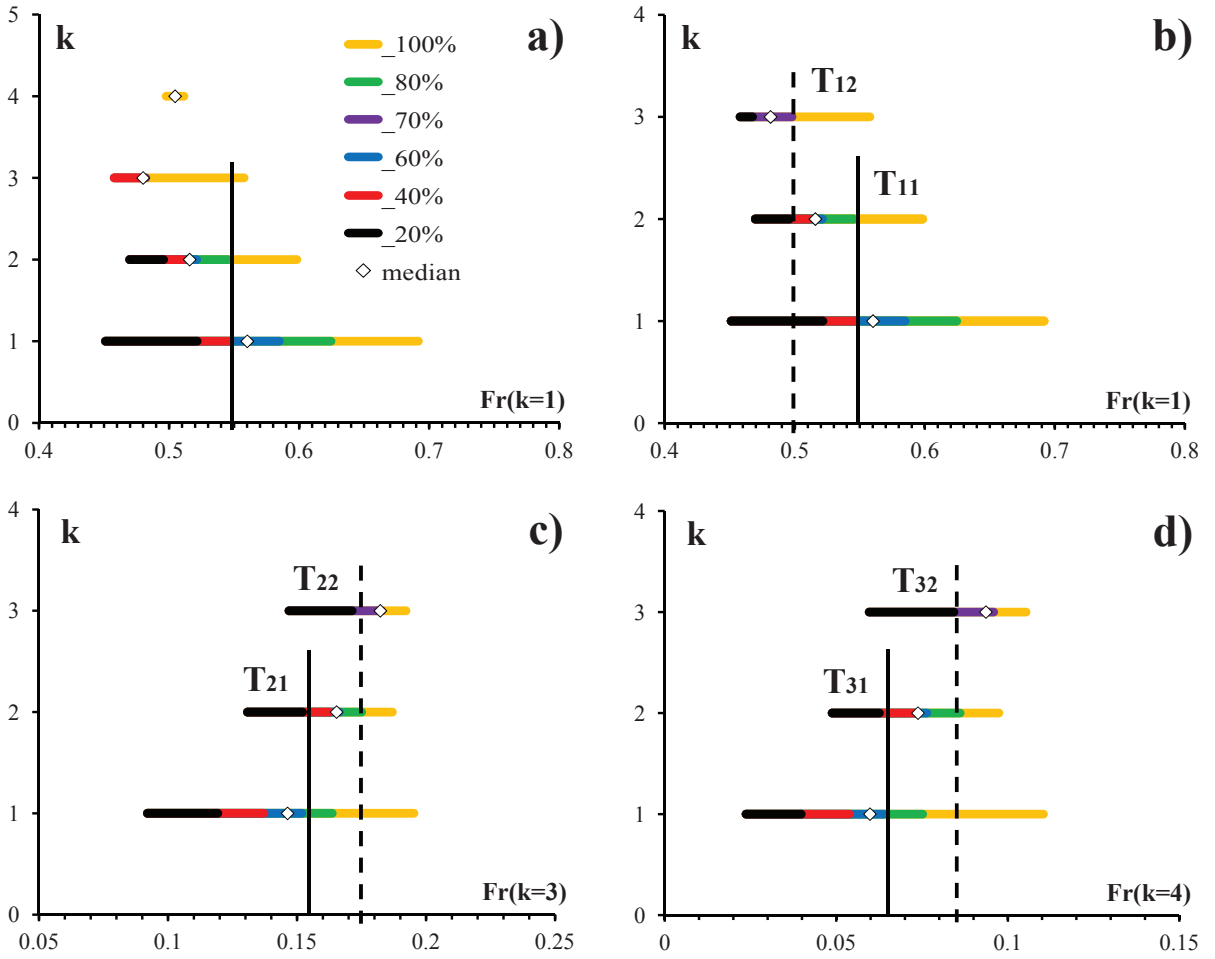


Figure 4.3.: Density plots representing data scatter of the calculated relative frequencies, $F_r(k)$, against the observed Dst ranges, k, for the Dst list (see Section 2). The density of the data points is expressed by differently colored percentiles. White diamond marks median, whereas black solid and dashed lines mark the established thresholds (for explanation see the main text).

Since Dst is given by four discrete values, the data in these plots will be scattered in 4 "lines" at $k = 1, 2, 3, 4$. Each of the lines contains a number of data points which correspond to the number of observations of different Dst range (184 events with $k = 1$, 17 events with $k = 2$, 8 events with $k = 3$, and 2 events with $k = 4$). The scatterplots for each of these 4 lines are presented as density plots using percentiles. In that way, it is noticeable how many data points are encompassed into each $F_r(k)$. An example is given for $F_r(k = 1)$ in Figure 4.3a. It can be seen that a threshold line corresponding to the value $F_r(k = 1) = 0.55$ separates 60% $k = 1$ events (right of the threshold line, with $F_r(k = 1) > 0.55$) from 80% $k = 2$ events and almost all of the $k = 3$ and $k = 4$ events (left of the threshold line, $F_r(k = 1) < 0.55$). Similarly $k = 2$ events can be uncoupled from $k = 3$ and $k = 4$ events. However, the data corresponding to $k = 4$ cannot be uncoupled from the data corresponding to $k = 3$. We see that the median of the $k = 4$

Table 4.2.: The conditions for determing the geo-effectiveness level using thresholds (T_{ij}, $i, j = 1, 2, 3$) for relative frequencies of certain bins, $F_r(k)$ given in Figures 4.3b-c. Average of these conditions give a unique geo-effectiveness level $k = 1, 2, 3$ (*i.e.* $Dst <$ 100nT, 100nT$< Dst <$ 200nT, $Dst >$ 200nT).

CONDITION I

threshold	condition	result
$T_{11} = 0.55$ $T_{12} = 0.5$	$F_r(k=1) > T_{11}$	$k = 1$
	$F_r(k=1) < T_{11}$ $F_r(k=1) > T_{12}$	$k = 2$
	$F_r(k=1) < T_{12}$	$k = 3$

CONDITION II

threshold	condition	result
$T_{21} = 0.155$ $T_{22} = 0.175$	$F_r(k=3) < T_{21}$	$k = 1$
	$F_r(k=3) > T_{21}$ $F_r(k=3) < T_{22}$	$k = 2$
	$F_r(k=3) > T_{22}$	$k = 3$

CONDITION III

threshold	condition	result
$T_{31} = 0.065$ $T_{32} = 0.085$	$F_r(k=4) < T_{31}$	$k = 1$
	$F_r(k=4) > T_{31}$ $F_r(k=4) < T_{32}$	$k = 2$
	$F_r(k=4) > T_{32}$	$k = 3$

scatterplot is shifted to higher $F_r(k = 1)$ values than that of $k = 3$, which is contrary to what we expect. Based on Figure 4.2 higher geo-effectiveness should be related to smaller values of $F_r(k = 1)$, as discussed in the previous paragraph. Similar behavior is observed for other $F_r(k)$ $(k = 2, 3, 4)$. This is related to the small number of intense geomagnetic storms in the sample (which are generally rare events) - there are only 8 events with $k = 3$ and 2 events with $k = 4$. Therefore, we combine $k = 3$ and $k = 4$ events into one $k = 3$ category, which corresponds to $Dst > 200$ nT (as shown in Figures 4.3c-d). Using the density of data scatter as a guideline, we derive thresholds using $F_r(k)$ $(k = 1, 3, 4)$, where thresholds are values which best separate different geo-effectiveness.

We interprete the thresholds as values which encompass most of the events with a certain geo-effectiveness. For example, most of the $Dst > 200$ nT have a relative frequency for $k = 4$, $F_r(k = 4) > T_{32}$. Therefore, if $F_r(k = 4) < T_{32}$ we expect $k = 2$ or $k = 1$. Similarly, if $F_r(k = 4) < T_{31}$ we expect $k = 1$. Therefore, each threshold defines a condition to determine level of geo-effectiveness. Note that there are three thresholds corresponding to $F_r(k = 1)$, $F_r(k = 3)$, and $F_r(k = 4)$ which determine the same level of geo-effectiveness. A unique geo-effectiveness level can be obtained by averaging the three conditions, as presented in Table 4.2.

We apply the three conditions from Table 4.2 to examples given in Figure 4.2. For the event shown in Figure 4.2a all three conditions are in favor of $k = 1$, so the expected Dst level is $Dst < 100$nT. For the event shown in Figure 4.2b the first condition results in $k = 2$, the second codition yields $k = 1$, and the third one $k = 2$. The average of the three conditions gives $k = 2$, so the expected Dst level is 100nT$< Dst <$ 200nT. Finally, for the event shown in Figure 4.2c all three conditions are in favor of $k = 3$, so the expected Dst level is $Dst > 200$nT. Therefore, starting from different solar CME/flare parameters we derive different Dst distributions, resulting in three different predictions of the geo-effectiveness level.

4.2. Empirical statistical model for Forbush decreases

Using the $k \leftrightarrow FD$ association (see introductory part of Section 4) we obtain the relative frequency distribution of FD for our sample of 187 events in the FD list (see Section 2). After calculating the distribution mean, $m = 2.07$, and using Equations (4.2) and (4.4) with renormalization (so that the total probability on all trials equals 1), we construct the shifted geometric distribution for the whole sample of 187 events.

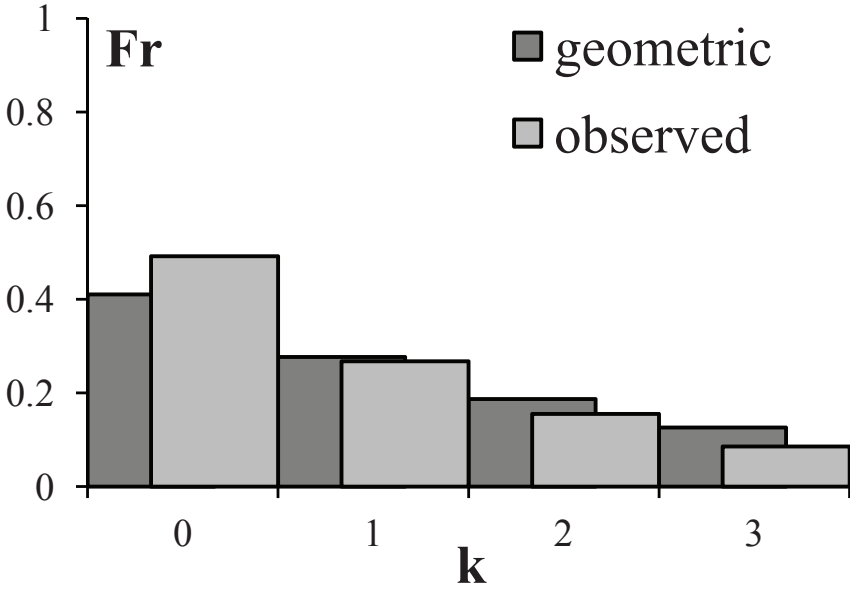


Figure 4.4.: Comparison of the observed FD distribution and calculated geometric distribution (FD relative frequencies, F_r for different FD magnitude ranges, k) for the whole sample of 187 events in the FD list (see Section 2 for details regarding the list; the graph is taken from [Dumbović et al., 2015b]).

In Figure 4.4 the distribution for our sample reconstructed using shifted geometric distribution is compared with the observed distribution and a reasonable agreement between the two can be seen (note that a similar distribution of the observed FDs was obtained by Belov [2009]). Since our sample in general follows the shifted geometric distribution,

Table 4.3.: CME/flare input parameters, α and corresponding calculated geometric distribution parameters, $m(\alpha)$ and $p(\alpha)$ for two extreme events (taken from [Dumbović et al., 2015b]).

EVENT 1			EVENT 2		
α	$m(\alpha)$	$p(\alpha)$	α	$m(\alpha)$	$p(\alpha)$
$v = 2000$ kms $^{-1}$	3.79	0.2089	$v = 450$ kms $^{-1}$	0.69	0.5924
$w = 360°$	3.81	0.2079	$w = 50°$	0.71	0.5848
$r = 0.05$ R_{SUN}	3.76	0.2103	$r = 0.99$ R_{SUN}	1.03	0.4929
$f = 5000 \cdot 10^{-7}$ Wm $^{-2}$	4.11	0.1957	$f = 10 \cdot 10^{-7}$ Wm $^{-2}$	0.38	0.7238
$i = 4$	2.70	0.2703	$i = 1$	1.20	0.4545

we assume that shifted geometric distribution can describe the probability distribution of FD magnitude, FD, and no additional changes to the distribution are needed, in contrast to the procedure described in Section 4.1.

It was demonstrated in Figures 3.3d and 3.4 in Section 3.2 that the trend of the change in the FD distribution mean with a specific solar parameter can be fitted by a corresponding function. Therefore, based on the relationships between FD and solar parameters, a corresponding shifted geometric distribution can be obtained employing Equations (4.2) and (4.4) for each solar parameter. We treat the obtained empirical distribution as a probability distribution for a specific solar parameter α, where $\alpha = v, w, r, f, i$ (*i.e.* initial CME speed, v, CME apparent width, w, CME/flare source position distance from the center of the solar disc, r, flare strength, f, and interaction parameter, i). The probability distribution for a specific parameter provides the information on the probability for associating it with a specific value of k, *i.e.* FD magnitude range. Equation 4.8 combines the effect of solar parameters and yields a joint probability distribution for a spacific set of solar parameters.

We calculate the probability distribution for two extreme events, EVENT 1: a very fast and wide CME, involved in a CME–CME interaction and associated with a strong X-class flare close to the center of the solar disc (presumably intensly GCR-effective), and EVENT 2: a slow and narrow CME, which is not involved in a CME–CME interaction and is associated with a weak B-class flare near the limb of the solar disc (presumably not GCR-effective). The input CME/flare parameters for both of these extreme events is given in Columns 1 and 4 in Table 4.3, respectively. Using the relationships between FD and solar parameters from Figures 3.3d and 3.4 in Section 3.2, we obtain the distribution mean for each of the solar parameters, $m(\alpha)$ (Columns 2 and 5 in Table 4.3). It can be seen that $m(\alpha)$ attains smaller values for EVENT 2, as expected (the distribution is shifted towards smaller FD magnitudes). With Equation (4.4) we derive a corresponding probability of the success in each trial for each of the solar parameters, $p(\alpha)$ (Columns 3 and 6 in Table 4.3), where the shift of the FD distribution for the EVENT 2 towards

smaller FD magnitudes is reflected in the increased values of $p(\alpha)$.

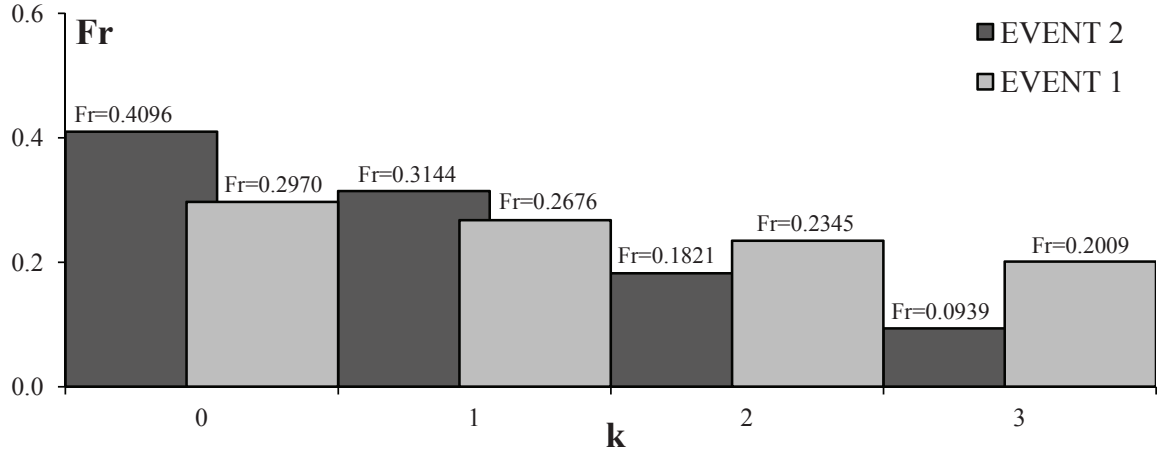


Figure 4.5.: Joint probability distribution for EVENT 1 (light grey) and EVENT 2 (dark grey). Relative frequencies are given above the corresponding k bin (taken from [Dumbović et al., 2015b]).

Using Equation (4.2) the relative frequency for each trial, k ($k = 0, 1, 2, 3$), and each solar parameter, α ($\alpha = v, w, r, f, i$), can be calculated for each of the two extreme events. Finally, using Equation (4.8), we calculate the joint probability distribution, *i.e.* the relative frequency for a given set of solar parameters $\{v, w, r, f, i\}$ for each trial, k, and renormalize it so that the total probability equals 1 ($\sum_{k=0}^{3} P(k) = 1$). The resulting distribution represents the joint probability distribution of the expected FD magnitude, FD, in a specific range $\{FD < 1\%,\ 1\% < FD < 3\%,\ 3\% < FD < 6\%,\ FD > 6\%\} \leftrightarrow \{k = 0,\ k = 1,\ k = 2,\ k = 3\}$ for a CME/flare event with a specific set of solar parameters $\{v,\ w,\ r,\ f,\ i\}$. The joint probability distributions for EVENT 1 and EVENT 2 are shown in Figure 4.5.

It can be seen in Figure 4.5 that the distribution for the two extreme events is different and that the probability for higher FD magnitudes is larger for EVENT 1 that represents faster and wider CMEs that originate near the disc center, are associated with more energetic flares, and are likely to be involved in a CME–CME interaction. However, in both distributions the highest probability is that the event will not be GCR-effective, *i.e.* that FD magnitude will be $FD < 1\%$ ($k = 0$) (similarly as found in Section 4.1). The probability distribution changes with CME/flare parameters, but it is always highly asymmetric with the highest probability of not being GCR-effective. Analogously to CME geo-effectiveness, this is a consequence of the general behavior of CMEs seen in Figure 4.4: a large majority of CMEs will never reach the Earth and/or will not be very GCR-effective. Therefore, the probability distribution does not give a straightforward prediction of whether or not (and how strong) Forbush decrease will be and the level of GCR-effectiveness needs to be obtained by imposing criteria (thresholds) on the probability

distribution, analogous to the procedure described in Section 4.1.

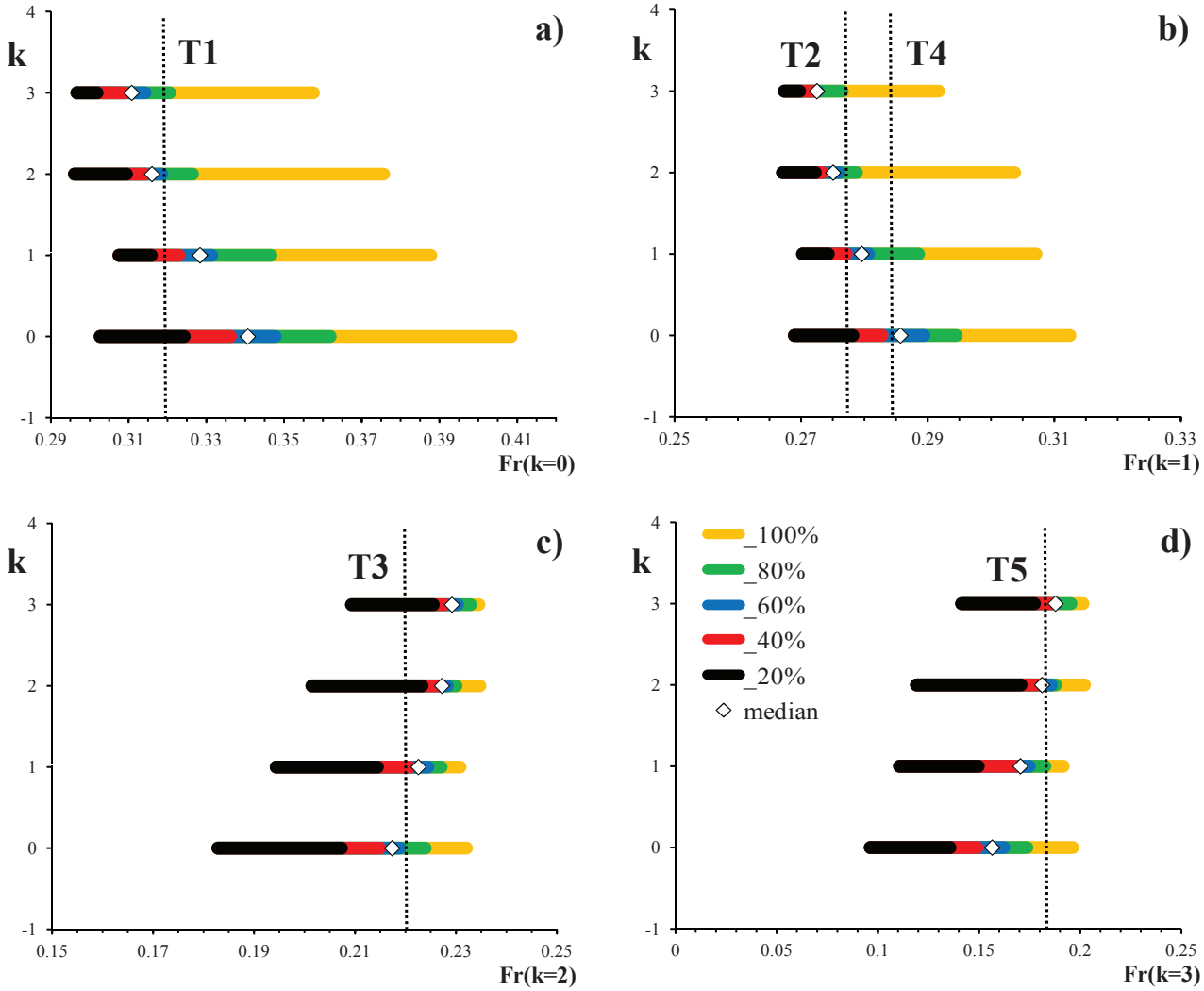


Figure 4.6.: Density plots representing data scatter of the calculated relative frequencies, $F_r(k)$, against the observed FD magnitude ranges, k, for 187 events from the FD list (see Section 2). The density of data points is expressed by differently colored percentiles. White diamond marks median, whereas black dotted lines mark established thresholds T1-T5 (for explanation see main text, figure is taken from [Dumbović et al., 2015b]).

Figure 4.5 shows that for low GCR-effectiveness (EVENT 2) one can expect a much higher value of relative frequency for $k = 0$, $F_r(k = 0)$ than for the highly GCR-effective event (EVENT 1). Reversly, we expect a much higher value of relative frequency for $k = 3$, $F_r(k = 3)$, for a highly GCR-effective event (EVENT 1) than for a low GCR-effectiveness (EVENT 2). Therefore, thresholds on the value of the relative frequency for a certain bin can be established to enclose certain GCR-effectiveness. These thresholds are derived empirically. For that purpose we use the FD list of 187 events (described in Section 2) and calculate FD magnitude distribution for each of the events in the list, based on the corresponding CME/flare parameters. Therefore, for each event we obtain

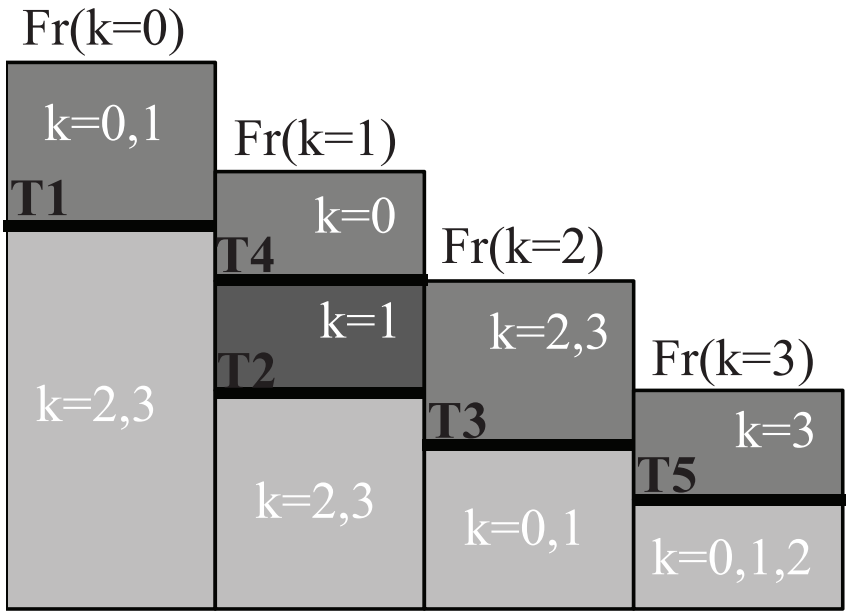


Figure 4.7.: Schematic of thresholds for relative frequencies of certain bins, $F_r(k)$: $T_1 = 0.32$, $T_2 = 0.277$, $T_3 = 0.222$, $T_4 = 0.285$, and $T_5 = 0.183$. Possible GCR-effectiveness level k is given for values above/below the corresponding threshold (taken from [Dumbović et al., 2015b]).

four different relative frequency values, $F_r(k)$, corresponding to four different distribution bins, $k = 0, 1, 2, 3$. For each relative frequency, $F_r(k)$, we produce a scatterplot against the observed FD value, analogously to the procedure described in Section 4.1. Since FD is given by four discrete values, the data in these plots will be scattered in 4 "lines" at $k = 0, 1, 2, 3$. Each of the lines contains a number of data points that corresponds to the number of observations of different FD magnitude range (92 events with $k = 0$, 50 events with $k = 1$, 29 events with $k = 2$, and 16 events with $k = 3$). The scatterplot for each of these 4 lines is presented as a density plot using percentiles. Using the density of data scatter as a guideline, we derive thresholds T1-T5 as values which best separate different GCR-effectiveness. These density plots, representing data scatter of the calculated relative requencies, $F_r(k)$, against the observed FD magnitude, as well as thresholds T1-T5 are presented in Figure 4.6.

It can be seen in Figure 4.6a that almost 80% of ($k = 3$) events and 60% of ($k = 2$) events have $F_r(k = 0) < 0.32 \equiv T1$, whereas more than 80% of non GCR-effective events ($k = 0$) and more than 60% of moderately GCR-effective events ($k = 1$) have $F_r(k = 0) > T1$. Therefore, we establish $T1$ as a threshold separating $k = 0, 1$ events from $k = 2, 3$ events. Similarly, we obtain thresholds $T2$ and $T3$ in Figures 4.6b and 4.6c, respectively. Finding a threshold between $k = 0$ and $k = 1$ events is more challenging, since the difference in the data density is less pronounced compared to that separating $k = 0, 1$ and $k = 2, 3$ events. In Figure 4.6b a threshold $T4$ is shown, which separates $k = 0$ data (more than 50% events have $F_r(k = 1) > T4$) from $k = 1$ data (more than 60%

Table 4.4.: Conditions for determing the GCR-effectiveness level using thresholds (T_i, $i = 1, 2, 3, 4, 5$) for relative frequencies of certain bins, $F_r(k)$ given in Figure 4.6. Combination of these conditions give a unique GCR-effectiveness level (taken from [Dumbović et al., 2015b]).

condition based on thresholds	result (if satisfied)	result (if not satisfied)	description of the conditions
$F_r(k=3) < T_1$	$k = 0, 1$	$k = 2, 3$	the combination of the first
$F_r(k=2) < T_2$	$k = 0, 1$	$k = 2, 3$	three conditions determines
$F_r(k=0) > T_3$	$k = 0, 1$	$k = 2, 3$	whether $k = 0, 1$ or $k = 2, 3$
$F_r(k=1) > T_4$	$k = 0$	$k = 1$	once established that $k = 0, 1$ this condition determines whether $k = 0$ or $k = 1$
$F_r(k=1) > T_5$	$k = 2$	$k = 3$	once established that $k = 2, 3$ this condition determines whether $k = 2$ or $k = 3$

events have $F_r(k = 1) < T4$). The difference in the data density is even less pronounced in separating $k = 2$ and $k = 3$ events. In Figure 4.6d a threshold $T5$ is shown, which separates $k = 2$ data (more than 50% events have $F_r(k = 3) < T5$) from $k = 3$ data (more than 50% events have $F_r(k = 3) > T5$).

We interprete the thresholds as values which divide most of the events with a certain GCR-effectiveness. For example, most of the $k = 3$ events have a relative frequency for $k = 3$, $F_r(k = 3) > T5$. Therefore, we expect that if $F_r(k = 3) < T5$ the event will have $k = 3$, otherwise it will be less GCR-effective and have $k < 3$. A schematic of the thresholds for relative frequencies of certain bins is given in Figure 4.7. Conditions for some of the thresholds immediately give the information on the expected GCR-effectiveness level. However, for some thresholds there still remains a set of possible GCR-effectiveness levels. Combining conditions for different thresholds, a unique GCR-effectiveness level can be obtained. The conditions for determing the GCR-effectiveness level using thresholds is given in Table 4.4.

For example, when we apply first three conditions from Table 4.4 to the joint probability distribution for EVENT 1 (Figure 4.5), we derive the following: $F_r(k = 3) > T_1$, $F_r(k = 2) > T_2$, and $F_r(k = 0) < T_3$. All three conditions are in favor of $k = 2, 3$; therefore, we apply the final condition from Table 4.4 and find that $F_r(k = 1) < T_5$, which means that the expected FD magnitude is $k = 3 \leftrightarrow FD > 6\%$. We repeat the calculation for EVENT 2 from Figure 4.5, where the first three conditions from Table 4.4 result in $F_r(k = 3) < T_1$, $F_r(k = 2) < T_2$, and $F_r(k = 0) > T_3$ being in favor of $k = 0, 1$. We then apply the fourth condition from Table 4.4 and find that $F_r(k = 1) > T_4$, which means that the expected FD magnitude is $k = 0 \leftrightarrow FD < 1\%$. Therefore, starting from extremely different solar CME/flare parameters we derive two extremes of GCR-effectiveness level.

The model is empirical and based on the remote solar CME/flare observations of the sample used; therefore, the model input has certain limitations. CME speed, v, is a continuous parameter given in $\mathrm{km\,s^{-1}}$ in the range $v > 106\,\mathrm{km\,s^{-1}}$, restricted by the x-intercept in Figure 3.4a. The CME/flare source distance from the center of the solar disc, r, is also a continuous parameter given in units of solar radii, with the range restricted by the physical boundaries, $0 \leq r \leq 1$ (*i.e.* the center of the solar disc and the solar limb). The apparent width, w, is a continuous parameter restricted to the range $0° < w \leq 360°$, determined by observational boundaries ($w = 0°$ means a CME was not detected, $w = 360°$ is a halo CME). The flare strength parameter, f, *i.e.* flare soft X-ray peak intensity is a continuous parameter given in units $10^{-7}\mathrm{Wm^{-2}}$ in the range $f > 5.3$ restricted by the x-intercept in Figure 3.4c. Finally, the interaction parameter, i, is a discrete parameter that can attain values $i = 1, 2, 3, 4$ based on the likeliness of the CME–CME interaction (Section 2).

5. Space weather forecast evaluation

In this section the succesfulness of the models presented in Sections 4.1 and 4.2 is validated. Parts of this research have already been published so the section abundantly contains citations from Dumbović et al. [2015b].

The prediction was first evaluated by using the **training set**, *i.e.* the sample used for the statistical analysis (Dst list for geomagnetic storms and FD list for Forbush decreases, respectively, see Section 2). The evaluation applied to the training set describes the succesfullness and the reliability of the prediction model with respect to the approximations used, since we assume that our sample represents the ensemble of possibilities for a certain event. Next we perform the evaluation using a **test set**, *i.e.* independent sample of additionally selected and measured events. We note that the two sets are conveniently named in analogy with neural network approach [see *e.g.* Valach et al., 2009, Uwamahoro et al., 2012, Sudar et al., 2015], that typically uses three different sets (training set, validating set, and test set), with the difference that in our case, the validating set is identical to the training set. The test set consists of events in the time period 1998 – 2012, which are not present in the training set. The method for CME-flare-GMS-FD association is the same as for the training set (described in Section 2). After 2011 cosmic ray data are no longer available at the SPIDR website and were taken from the *Neutron Monitor Database* event search tool (`http://www.nmdb.eu/nest/search.php`) (Kiel, Magadan, and Newkirk neutron monitor stations). As described in Section 2, the association of CME/flares with geomagnetic and cosmic ray response is not identical, thus the samples slightly differ: Dst test list consists of 43 events, whereas the FD test list consists of 42 events. These lists are available online as a merged CME-flare-Dst-FD list at: http://oh.geof.unizg.hr/FDFT/fdft.php.

Table 5.1.: Contingency table for a binary event

		Observation	
		YES	NO
Forecast	YES	a = number of hits, *i.e.* correctly forecasted events	b = number of false alarms, *i.e.* forecasts of an event while no event was observed
	NO	c = number of misses, *i.e.* events which were not forecasted	d = number of correct rejections, *i.e.* events which were not forecasted while indeed no event was observed

The forecast was validated by comparing the predicted value with the observed value using verification measures for binary events [see *e.g.* Devos et al., 2014]. The verification measures are defined by the contingency table (Table 5.1), which describes four possible outcomes (hit, false alarm, miss, and correct rejection). For the purpose of the evaluation we redefine the "event" as association of Dst, *i.e.* FD with a particular value. For example, we define $k = 0$ as an event. The event is classified as a "hit" when $k = 0$ was both observed and predicted; "false alarm" is when $k = 0$ is observed, while $k \neq 0$ was predicted; "miss" is when $k \neq 0$ was observed, while $k = 0$ was predicted; "correct rejection" is when $k \neq 0$ was both observed and predicted.

Using the values of a, b, c, and d defined in Table 5.1 we apply the following verification measures [for more details see Devos et al., 2014, and references therein]:

- *The Probability Of Detection* (POD) or hit rate, the ratio of the number of hits and the number of events, calculated as POD= $a/(a + c)$;
- *The False Alarm Ratio* (FAR), the ratio of the number of false alarms and the total number of forecasts, calculated as FAR= $b/(a + b)$;
- *Bias* (BIAS), the ratio of the number of forecasts of occurrence to the number of actual occurrences, calculated as B= $(a + b)/(a + c)$
- *Heidke Skill Score* (HSS), skill score taking into account the number of correct random forecasts, calculated as HSS= $(a + d - E)/(n - E)$,
  where $E = ((a + c)(a + b) + (c + d)(b + d))/n$ and $n = a + b + c + d$

Each of the verification measures gives an information on the quality of the prediction; however, none of them gives a full information on the quality of the forecast system. POD describes what fraction of the observed "yes" events were correctly forecast and ranges from 0 to 1, with perfect score POD=1 (all hits). It is sensitive to hits, but ignores false alarms; therefore, it should be used in conjunction with FAR. FAR describes how many of the predicted "yes" events were false alarms; however, it ignores misses and consequently has to be used in conjunction with POD. It ranges from 0 to 1, with perfect score FAR=0 (no false alarms). BIAS measures the ratio of the frequency of forecasts to the frequency of observations and ranges from 0 to ∞, with perfect score BIAS=1. It reveals whether the forecast has a tendency to underforecast (BIAS<1) or overforecast (BIAS>1) events. However, it tells nothing about how well the forecast corresponds to the observations. Finally, HSS estimates the accuracy of the forecast relative to that of random chance. It ranges from -∞ to 1, where HSS=1 is a perfect score, HSS=0 means that the forecast is no better than random, and HSS<0 means that the forecast is worse than random.

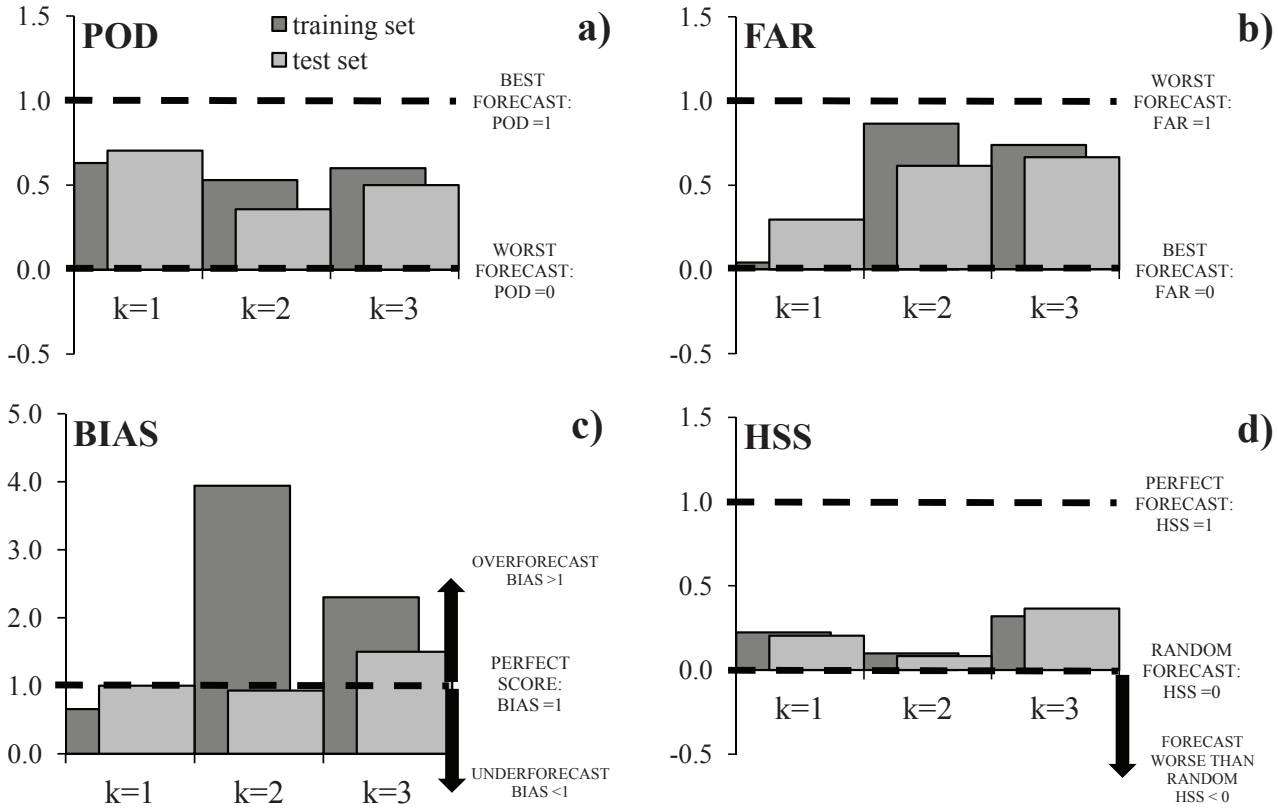


Figure 5.1.: Probability of Detection (POD), False Alarm Ratio (FAR), BIAS, and Heidke Skill Score (HSS) for the training and test sets for the forecast of the Dst range for a set of CME/flare solar parameters ($\{k = 1,\ k = 2,\ k = 3\} \leftrightarrow \{Dst < 100\text{nT},\ 100\text{nT} < Dst < 200\text{nT},\ Dst > 200\text{nT}\}$, see Section 4).

5.1. Geomagnetic storm forecast evaluation

The validation of the succesfulness of the empirical model for CME geo-effectiveness forecasting was done using the Dst training list (master list, described in Section 2) and Dst test list (described in the introductory part of Section 5). The number of possible outcomes based on the contingency table (Table 5.1), as well as the corresponding verification measures for both the training and the test set are given in Table 5.2 for the following "events": $k = 1$ ($Dst < 100$nT), $k = 2$ ($100\text{nT} < Dst < 200$nT), and $k = 3$ ($Dst > 200$nT). Verification measures for these events are also presented separately in Figure 5.1 for the training and test sets.

It can be seen from Table 5.2 and Figure 5.1 that there are differences in the verification measures between the training and test sets, especially for the BIAS for $k = 2$. Large value of BIAS for the training sample indicates that the model tends to overforecast $k = 2$ events, *i.e.* there are far more forecasts of the $k = 2$ events than observed, indicating that many $k = 1$ and $k = 3$ events were forecasted as $k = 2$ events. Since BIAS for $k = 1$ indicates underforecast and for $k = 3$ overforecast, it can be concluded that the large BIAS for $k = 2$ comes from $k = 1$ events. The False Alarm Ratio (FAR) for

Table 5.2.: Number of possible outcomes based on the contingency table (Table 5.1) and the corresponding verification measures for the Dst validation and test samples, for different events

	event	a	b	c	d	POD	FAR	BIAS	HSS
training	$k=1$	116	5	68	22	0.63	0.04	0.66	0.22
set	$k=2$	9	58	8	136	0.53	0.87	3.94	0.10
	$k=3$	6	17	4	184	0.60	0.74	2.30	0.32
test	$k=1$	19	8	8	8	0.70	0.30	1.00	0.20
set	$k=2$	5	8	9	21	0.36	0.62	0.93	0.08
	$k=3$	1	2	1	39	0.50	0.67	1.50	0.36

the considered $k = 2$ event is also highest, whereas the Probability of Detection (POD) assumes medium values. All together, this results in a low, but positive, Heidke Skill Score (HSS), indicating that the prediction for $k = 2$ events in the training set is still better than the random. The results for the test set $k = 2$ events are similar to those of the training set, except for the BIAS. The BIAS for the test set is close to 1 meaning that the prediction does not have tendency to overforecast or underforecast events.

In the training set, $k = 1$ events have highest POD and smallest FAR; however, as mentioned above, the BIAS implicates an underforecasting of $k = 1$ events (they are often forecasted as $k = 2$ events). The BIAS is improved in the test set, but the number of false alarms increases. The HSS in both cases is much better than for the $k = 2$ events indicating that the model prediction for $k = 1$ events is better than random. The forecast of $k = 3$ events has quite high FAR and medium values for POD, and a tendency to overforecast (in both the training and test samples). However, it also shows the best forecasting skill (highest HSS). Therefore, we can conclude that the model gives a most reliable prediction of whether or not there will be an intense geomagnetic storm ($Dst > 200$nT), whereas it has difficulties in discerning 100nT$< Dst <$ 200nT storms. The fact that the forecast of the intermediate bin is the least reliable aspect, implies that the forecast has a "resolution" problem, *i.e.* has difficulties in discerning between neighbouring bins.

Finally, we compare our results with the results of previous studies. Kim et al. [2010] found that their model is not able to reliably predict the Dst index based on remote solar parameters, but is quite reliable in predicting the occurence of the storm (*i.e.* whether or not $Dst > 50$nT). Their conclusion is based on the calculated critical success index, CSI=0.8 (CSI=$a/(a+b+c)$, where a, b, c are given by Table 5.1). Our model does not provide this information, but we calculate the CSI for occurence of $Dst > 100$nT and $Dst < 200$nT storms ($k > 1$ and $k < 3$, respectively). The calculated values of CSI are 0.2 and 0.9, respectively. This confirms that our model gives a most reliable prediction of whether or not there will be $Dst > 200$nT. Valach et al. [2009] used a neural network approach using the Kp index (see Section 1.2.1). They obtained a 48% successful forecast

for geomagnetic response occurence and 47% for the occurence of severe geomagnetic response (measured by POD), which are slightly lower values than the values resulting from our model. However, as seen in Table 5.2 and Figure 5.1, POD alone is not a good verification measure, FAR and BIAS must also be taken into account. Therefore, full comparison cannot be performed. It should be noted that Valach et al. [2009] used a different set of solar parameters, mostly related to the associated flare. Finally, best evaluation results are presented by Srivastava [2005] and Uwamahoro et al. [2012] (used POD as verification measure). A logistic regression model by Srivastava [2005] was able to predict all of the intense (100nT$< Dst <$ 200T) and 50% of super intense ($Dst >$ 200nT) storms in their sample. Therefore, the overall prediction ability of the model was estimated at 77.7%. The results of the neural network model by Uwamahoro et al. [2012] are even more promising, as it predicts all of the intense storms ($Dst >$ 100nT) and 75% of moderate storms (50nT$< Dst <$ 100nT) with the overall prediction ability of 86%. It should be noted though, that both Srivastava [2005] and Uwamahoro et al. [2012] used samples where all events are associated with storms. In addition, they used interplanetary parameters. The POD value calculated for our model is somewhat lower; however, we only use remote solar parameters. Therefore, based on the presented analysis we can conlude that our model provides some advantages compared to other forecast models. It is based exclusivelly on the remote solar parameters and gives a prediction ≈ 1 day in advance. It can provide the information of storm occurence in a specific Dst range. The prediction ability of the model is quite reliable for the strongest storms with $Dst >$ 200nT, which is certainly the most important aspect of forecasting. The drawbacks of the model is that it does not discern between no-storm events ($Dst <$ 50 nT) and moderate storms (50nT$< Dst <$ 100nT) and that the prediction capability for 100nT$< Dst <$ 200nT is quite low.

5.2. Forbush decrease forecast evaluation

Validation of the empirical model for the CME GCR-effectiveness forecasting was performed using the FD training list (master list, described in Section 2) and FD test list (described in the introductory part of Section 5). The number of possible outcomes based on the contingency table (Table 5.1), as well as the corresponding verification measures for both the training and the test set are given in Table 5.3 for the following "events": $k = 0$ ($FD < 1\%$), $k = 1$ ($1\% < FD < 3\%$), $k = 2$ ($3\% < FD < 6\%$), $k = 3$ ($FD > 3\%$), $k = 0, 1, 2$ ($FD < 6\%$), $k = 0, 1$ ($FD < 3\%$), $k = 1, 2, 3$ ($FD > 1\%$). We note that the first four "events" correspond to the four bins of the probability distribution presented in Sections 3.2 and 4.2, whereas the last three "events" represent a less specific forecast. For these two groups of events verification measures are also presented separately in Figure 5.2 for the training and test samples.

It can be seen from Table 5.3 and Figure 5.2 that there are differences in the verification

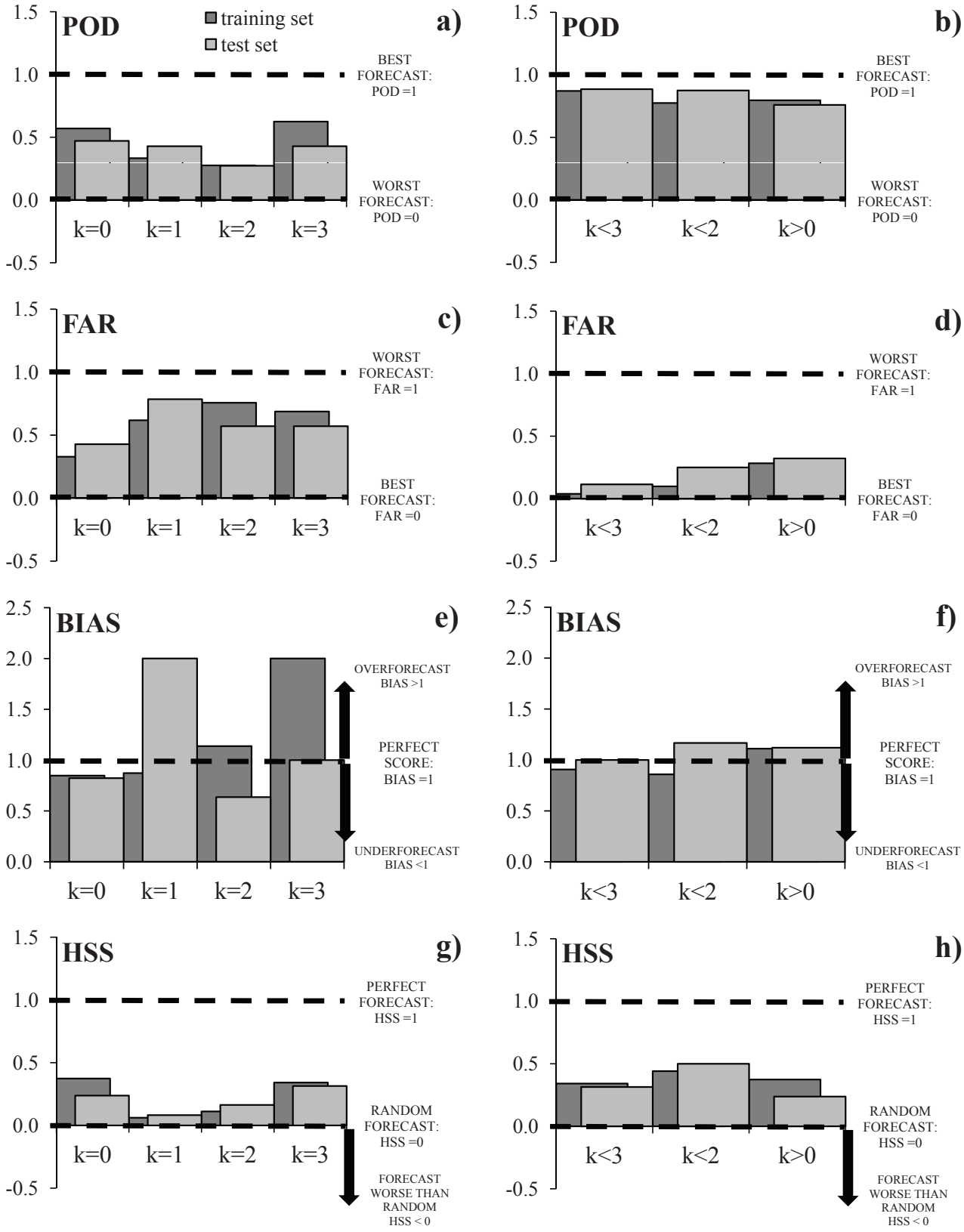


Figure 5.2.: Probability of Detection (POD), False Alarm Ratio (FAR), BIAS, and Heidke Skill Score (HSS) for the FD training and test samples for a more specific ($k = 0, 1, 2, 3$) and less specific ($k < 3$, $k < 2$, and $k > 0$) forecast (taken from [Dumbović et al., 2015b]).

Table 5.3.: Number of possible outcomes based on the contingency table and the corresponding verification measures for the validation and test samples, for different events (taken from [Dumbović et al., 2015b]).

	event	a	b	c	d	POD	FAR	BIAS	HSS
training set	$k = 0$	45	22	34	86	0.57	0.33	0.85	0.37
	$k = 1$	21	34	42	90	0.33	0.62	0.87	0.06
	$k = 2$	8	25	21	133	0.28	0.76	1.14	0.11
	$k = 3$	10	22	6	149	0.63	0.69	2.00	0.34
	$k < 3$	149	6	22	10	0.87	0.04	0.91	0.34
	$k < 2$	110	12	32	33	0.77	0.10	0.86	0.44
	$k > 0$	86	34	22	45	0.80	0.28	1.11	0.37
test set	$k = 0$	8	6	9	19	0.47	0.43	0.82	0.24
	$k = 1$	3	11	4	24	0.43	0.79	2.00	0.08
	$k = 2$	3	4	8	27	0.27	0.57	0.64	0.16
	$k = 3$	3	4	4	31	0.43	0.57	1.00	0.31
	$k < 3$	31	4	4	3	0.89	0.11	1.00	0.31
	$k < 2$	21	7	3	11	0.88	0.25	1.17	0.50
	$k > 0$	19	9	6	8	0.76	0.32	1.12	0.24

measures between the training and test sets, especially for the BIAS in case of a more specific forecast (*i.e.* for the forecast of a specific bin, $k = 0, 1, 2, 3$). However, the differences are not large and they are not systematic, indicating that the successfulness of the forecast mainly relies on the approximations used, and not on the sample itself. The forecast of the intermediate bins $k = 1, 2$ is least reliable, since we get the lowest number of hits and largest number of false alarms. This is also evident in the HSS, which gives lowest values, indicating that the forecast is only slightly better than random for these two bins. Analogously as in Section 5.1 we conclude that the forecast has a "resolution" problem, *i.e.* has difficulties in discerning between neighbouring bins. This is also supported by the fact that when less specific bins are regarded ($k < 3$, $k < 2$, and $k > 0$), POD is much higher, FAR is lower, BIAS is closer to perfect value (BIAS$\approx$ 1) and HSS has larger positive values, the latter indicating that the forecast has skill compared to random forecast (see Table 5.3 and Figures 5.2c and 5.2d). Therefore, we conclude that the Forbush decrease prediction is more reliable for less specific forecast, *i.e.* for predicting whether or not CME will be GCR-effective ($k > 0 \leftrightarrow FD > 1\%$), whether or not it will be strongly/intensly GCR-effective ($k < 2 \leftrightarrow FD < 3\%$) and whether or not it will be intensly GCR-effective ($k < 3 \leftrightarrow FD < 6\%$). Given the verification measures presented in Table 5.3 and Figure 5.2, the most reliable forecast (highest POD, lowest FAR, BIAS$\approx$ 1, and high HSS) is the prediction whether or not CME will produce $FD > 3\%$.

6. Conclusion and summary

A sample of CME-flare pairs detected remotely was compiled and associated to a geomagnetic and cosmic ray response at Earth, resulting in two interconnected lists - the Dst and FD lists (presented in Section 2). The aim of the study is to forecast the CME-associated Dst index and Forbush decrease magnitudes, which quantify geomagnetic and cosmic ray response, respectively. The advantage of the proposed approach is in the early forecast, since the travel time for a CME from Sun to Earth most often ranges from 1 to 4 days. To characterize CME/flare event we use the L1 coronagraphic CME observations, the EUV flare-position observation, as well as the flare Soft X-ray flux measurements. We note that some properties derived from these observations can also be obtained from ground-based measurements (*e.g.* proxy of the CME speed can be obtained from solar Type II radio bursts, the flare position can be determined from H_α observations). Therefore, the remotely observed CME/flare properties are not necessarily satellite-dependent.

The relationship between observed Dst and FD magnitudes at the Earth and remote observations of CMEs and associated solar flares is studied *via* statistical analysis. It was found for both Dst and FD magnitudes that they are larger for faster CMEs with larger apparent width, associated with stronger flares, originating close to the center of the solar disc and (possibly) involved in a CME–CME interaction. These relations are quantified through the change in the distribution of the Dst/FD magnitude, which is mathematically reconstructed using the (shifted) geometric distribution. The reconstructed distributions are used to obtain a joint probability distribution for a certain CME/flare event, where Dst and FD lists are used as an ensemble of possibilities for a certain event. The joint probability distributions for a certain CME/flare event behave differently when different CME/flare properties are used as input, depicting the behavior found by statistical analysis. However, distributions are always highly asymmetric with greatest probability that CME will not be geo- or GCR-effective, which is the general behavior of CMEs (a large majority of CMEs will never reach the Earth and/or will not be very geo- and GCR-effective). Probability distribution for Dst index is more asymmetric than the one for FD magnitude, indicating that a CME which is GCR-effective doesn't have to be geo-effective necesseraly. This is expected given that cosmic rays respond to the magnetic field strength of the CME, whereas for geomagnetic storms the magnetic field orientation plays a role, as well. Empirically optimized thresholds on the probability distributions were imposed, to obtain the estimation of the geo- and GCR-effectiveness for a CME/flare event

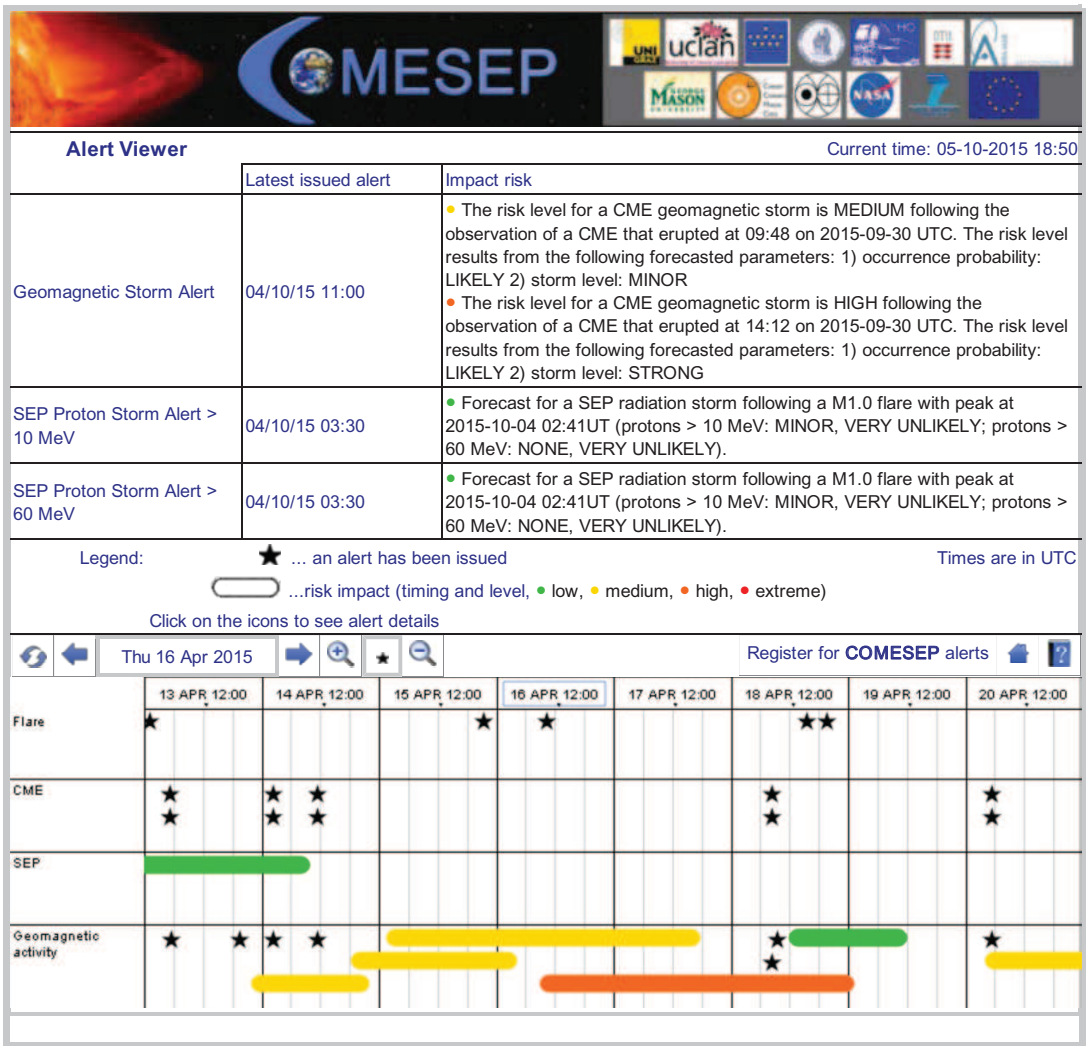


Figure 6.1.: The visual presentation of the COMESEP alert system (Credit: COMESEP project webpage, `http://comesep.aeronomy.be/alert/`)

of specific characteristics. In this way an empirical probabilistic model was developed in which selected remote solar observations of a CME and the associated solar flare are used as input providing the expected Dst/FD magnitude range as an output.

Validation of the forecast method is performed on the training set (the Dst and FD lists used for the statistical analysis) and test set (independent Dst and FD lists shown in Section 5). The validation procedure revealed that the forecast is less reliable when it is more specific, due to difficulties in discerning between neighbouring bins. It was found especially ineffective for prediction of intermediate Dst and FD magnitudes. However, when the forecast is less specific, the quality of the forecast improves. The Dst prediction is most reliable for the strongest storms with $Dst > 200$nT. The Forbush decrease prediction is found to be most reliable in predicting whether or not a CME will produce $FD > 3\%$.

Based on the research presented in this thesis, online applications for the models were developed, available at *Hvar Observatory* webpage. The prediction of geomagne-

tic storms, *i.e.* Dst index magnitude based on the remote solar observations of a CME and associated solar flare, "CME Geo-effectiveness Forecast Tool (CGeFT)" is available at: `http://oh.geof.unizg.hr/CGEFT/cgeft.php`. The prediction of Forbush decrease magnitude, "Forbush Decrease Forecast Tool (FDFT)" is availabale at: `http://oh.geof.unizg.hr/FDFT/fdft.php`. The corresponding publication, as well as the full training and test set list are also available at same webpages under "Documentation". In addition to the online forecast tools, the first version of the geomagnetic storm prediction model was used in the "COMESEP alert system", the first fully automatic system for detection of CMEs and solar flares, forecasting the CME arrival, as well as their potentially hazardous impact. The system runs fully automatically, *i.e.* without human intervantion. It was developed within a EU FP7 project "Coronal Mass Ejections and Solar Energetic Particles (COMESEP)". The geomagnetic storm prediction model presented here and described by Dumbović et al. [2015a] was used as one of the modules that constitute the "CME Geomagnetic Forecast Tool (CGFT)" of the COMESEP alert system, which is available at `http://comesep.aeronomy.be/alert/`.

The results of this research were disseminated at more than 10 international conferences and workshops. In addition, the results were published in a peer-review journal *Solar Physics* (Dumbović et al. [2015a] and Dumbović et al. [2015b]). This research is a continuation of the previous research of the space weather effects of the interplanetary coronal mass ejections published in a peer-reviewed papers Dumbović et al. [2011], Dumbović et al. [2012a], and Dumbović et al. [2012b].

Bibliography

S.I. Akasofu. *Exploring the secrets of the aurora (2nde).* Springer Science+Business Media, LLC, New York, 2007.

Badruddin, R. S. Yadav, and N. R. Yadav. Influence of magnetic clouds on cosmic ray intensity variation. *Sol. Phys.*, 105:413–428, June 1986. doi: 10.1007/BF00172057.

L. R. Barnden. The Large-Scale Magnetic Field Configuration Associated With Forbush Decreases. In *International Cosmic Ray Conference*, volume 2 of *International Cosmic Ray Conference*, pages 1277–+, 1973.

A. Belov, A. Abunin, M. Abunina, E. Eroshenko, V. Oleneva, V. Yanke, A. Papaioannou, H. Mavromichalaki, N. Gopalswamy, and S. Yashiro. Coronal Mass Ejections and Non-recurrent Forbush Decreases. *Sol. Phys.*, 289:3949–3960, October 2014. doi: 10.1007/s11207-014-0534-6.

A. V. Belov. Forbush effects and their connection with solar, interplanetary and geomagnetic phenomena. In N. Gopalswamy and D. F. Webb, editors, *IAU Symposium*, volume 257 of *IAU Symposium*, pages 439–450, March 2009. doi: 10.1017/S1743921309029676.

A. V. Belov, E. A. Eroshenko, V. A. Oleneva, A. B. Struminsky, and V. G. Yanke. What determines the magnitude of forbush decreases? *Adv. Space Res.*, 27:625–630, 2001. doi: 10.1016/S0273-1177(01)00095-3.

J. J. Blanco, E. Catalán, M. A. Hidalgo, J. Medina, O. García, and J. Rodríguez-Pacheco. Observable Effects of Interplanetary Coronal Mass Ejections on Ground Level Neutron Monitor Count Rates. *Sol. Phys.*, 284:167–178, May 2013a. doi: 10.1007/s11207-013-0256-1.

J. J. Blanco, M. A. Hidalgo, R. Gómez-Herrero, J. Rodríguez-Pacheco, B. Heber, R. F. Wimmer-Schweingruber, and C. Martín. Energetic-particle-flux decreases related to magnetic cloud passages as observed by the Helios 1 and 2 spacecraft. *Astron. Astrophys.*, 556:A146, August 2013b. doi: 10.1051/0004-6361/201321739.

V. Bothmer and I.A. Daglis. *Space Weather - Physics and Effects.* Springer-Verlag Berlin Heidelberg, Berlin, 2007.

V. Bothmer and R. Schwenn. Eruptive prominences as sources of magnetic clouds in the solar wind. *Space Sci. Rev.*, 70:215–220, October 1994. doi: 10.1007/BF00777872.

G. E. Brueckner, R. A. Howard, M. J. Koomen, C. M. Korendyke, D. J. Michels, J. D. Moses, D. G. Socker, K. P. Dere, P. L. Lamy, A. Llebaria, M. V. Bout, R. Schwenn, G. M. Simnett, D. K. Bedford, and C. J. Eyles. The Large Angle Spectroscopic Coronagraph (LASCO). *Sol. Phys.*, 162:357–402, December 1995. doi: 10.1007/BF00733434.

J. T. Burkepile, A. J. Hundhausen, A. L. Stanger, O. C. St. Cyr, and J. A. Seiden. Role of projection effects on solar coronal mass ejection properties: 1. A study of CMEs associated with limb activity. *J. Geophys. Res.*, 109:A03103, March 2004. doi: 10.1029/2003JA010149.

L. Burlaga, E. Sittler, F. Mariani, and R. Schwenn. Magnetic loop behind an interplanetary shock - Voyager, Helios, and IMP 8 observations. *J. Geophys. Res.*, 86:6673–6684, August 1981. doi: 10.1029/JA086iA08p06673.

L. F. Burlaga. Interplanetary stream interfaces. *J. Geophys. Res.*, 79:3717, 1974. doi: 10.1029/JA079i025p03717.

S. Cahampan and J. Bartels. *Geomagnetism.* Oxford University Press, London, 1940.

W. H. Campbell. *Earth magnetism.* A Harcourt Science and Technology Company, San Diego, 2001.

H. V. Cane. Cosmic ray decreases and magnetic clouds. *J. Geophys. Res.*, 98:3509–3512, March 1993. doi: 10.1029/92JA02479.

H. V. Cane. Coronal Mass Ejections and Forbush Decreases. *Space Sci. Rev.*, 93:55–77, July 2000. doi: 10.1023/A:1026532125747.

H. V. Cane, I. G. Richardson, T. T. von Rosenvinge, and G. Wibberenz. Cosmic ray decreases and shock structure: A multispacecraft study. *J. Geophys. Res.*, 99:21429, November 1994. doi: 10.1029/94JA01529.

H. V. Cane, I. G. Richardson, and G. Wibberenz. The Response of Energetic Particles to the Presence of Ejecta Material. *International Cosmic Ray Conference*, 4:377, 1995.

P. J. Cargill. On the Aerodynamic Drag Force Acting on Interplanetary Coronal Mass Ejections. *Sol. Phys.*, 221:135–149, May 2004. doi: 10.1023/B:SOLA.0000033366.10725.a2.

S. Chapman and V. C. A. Ferraro. A new theory of magnetic storms. *Terrestrial Magnetism and Atmospheric Electricity*, 36(2):77–97, 1931. ISSN 0096-8013. doi: 10.1029/TE036i002p00077. URL `http://dx.doi.org/10.1029/TE036i002p00077`.

I. M. Chertok, V. V. Grechnev, A. V. Belov, and A. A. Abunin. Magnetic Flux of EUV Arcade and Dimming Regions as a Relevant Parameter for Early Diagnostics of Solar Eruptions - Sources of Non-recurrent Geomagnetic Storms and Forbush Decreases. *Sol. Phys.*, 282:175–199, January 2013. doi: 10.1007/s11207-012-0127-1.

P. P. Chih and M. A. Lee. A perturbation approach to cosmic ray transients in interplanetary space. *J. Geophys. Res.*, 91:2903–2913, March 1986. doi: 10.1029/JA091iA03p02903.

E. W. Cliver and L. Svalgaard. The 1859 Solar-Terrestrial Disturbance And the Current Limits of Extreme Space Weather Activity. *Sol. Phys.*, 224:407–422, October 2004. doi: 10.1007/s11207-005-4980-z.

N. U. Crooker, J. T. Gosling, V. Bothmer, R. J. Forsyth, P. R. Gazis, A. Hewish, T. S. Horbury, D. S. Intriligator, J. R. Jokipii, J. Kóta, A. J. Lazarus, M. A. Lee, E. Lucek, E. Marsch, A. Posner, I. G. Richardson, E. C. Roelof, J. M. Schmidt, G. L. Siscoe, B. T. Tsurutani, and R. F. Wimmer-Schweingruber. CIR Morphology, Turbulence, Discontinuities, and Energetic Particles. *Space Sci. Rev.*, 89:179–220, July 1999. doi: 10.1023/A:1005253526438.

F. M. Dekking, C. Kraaikamp, H. P. Lopuhaa, and L. E. Meester. *A Modern Introduction to Probability and Statistics*. Springer-Verlag London, London, 2005.

A. Devos, C. Verbeeck, and E. Robbrecht. Verification of space weather forecasting at the Regional Warning Center in Belgium. *J. Space Weather Space Clim.*, 4(27):A29, October 2014. doi: 10.1051/swsc/2014025.

E. D'Huys, D. B. Seaton, S. Poedts, and D. Berghmans. Observational characteristics of coronal mass ejections without low-coronal signatures. *Astrophys. J.*, 795(1):49, 2014. URL `http://stacks.iop.org/0004-637X/795/i=1/a=49`.

V. Domingo, B. Fleck, and A. I. Poland. The SOHO Mission: an Overview. *Sol. Phys.*, 162:1–37, December 1995. doi: 10.1007/BF00733425.

W. Dröge. Particle Scattering by Magnetic Fields. *Space Sci. Rev.*, 93:121–151, July 2000. doi: 10.1023/A:1026588210726.

M. Dumbović, B. Vršnak, J. Čalogović, and M. Karlica. Cosmic ray modulation by solar wind disturbances. *Astron. Astrophys.*, 531:A91+, July 2011. doi: 10.1051/0004-6361/201016006.

M. Dumbović, B. Vršnak, and J. Čalogović. Solar influences on the short-term cosmic ray modulation. *Cent. European Astrophys. Bull.*, 36:65–70, 2012a.

M. Dumbović, B. Vršnak, J. Čalogović, and R. Župan. Cosmic ray modulation by different types of solar wind disturbances. *Astron. Astrophys.*, 538:A28, February 2012b. doi: 10.1051/0004-6361/201117710.

M. Dumbović, A. Devos, B. Vršnak, D. Sudar, L. Rodriguez, D. Ruždjak, K. Leer, S. Vennerstrøm, and A. Veronig. Geoeffectiveness of Coronal Mass Ejections in the SOHO Era. *Sol. Phys.*, 290:579–612, February 2015a. doi: 10.1007/s11207-014-0613-8.

M. Dumbović, B. Vršnak, and J. Čalogović. Forbush Decrease Prediction Based on Remote Solar Observations. *Sol. Phys.*, November 2015b. doi: 10.1007/s11207-015-0819-4.

J. W. Dungey. Interplanetary magnetic field and the auroral zones. *Phys. Rev. Lett*, 6:47–48, Jan 1961. doi: 10.1103/PhysRevLett.6.47. URL `http://link.aps.org/doi/10.1103/PhysRevLett.6.47`.

E. Echer, W. D. Gonzalez, B. T. Tsurutani, and A. L. C. Gonzalez. Interplanetary conditions causing intense geomagnetic storms (Dst $\leq$ -100 nT) during solar cycle 23 (1996-2006). *J. Geophys. Res.*, 113:A05221, May 2008. doi: 10.1029/2007JA012744.

C. J. Eyles, R. A. Harrison, C. J. Davis, N. R. Waltham, B. M. Shaughnessy, H. C. A. Mapson-Menard, D. Bewsher, S. R. Crothers, J. A. Davies, G. M. Simnett, R. A. Howard, J. D. Moses, J. S. Newmark, D. G. Socker, J.-P. Halain, J.-M. Defise, E. Mazy, and P. Rochus. The Heliospheric Imagers Onboard the STEREO Mission. *Sol. Phys.*, 254:387–445, February 2009. doi: 10.1007/s11207-008-9299-0.

J. Feynman and S. B. Gabriel. On space weather consequences and predictions. *J. Geophys. Res.*, 105:10543–10564, May 2000. doi: 10.1029/1999JA000141.

T. G. Forbes. A review on the genesis of coronal mass ejections. *J. Geophys. Res.*, 105: 23153–23166, October 2000. doi: 10.1029/2000JA000005.

T. G. Forbes, J. A. Linker, J. Chen, C. Cid, J. Kóta, M. A. Lee, G. Mann, Z. Mikić, M. S. Potgieter, J. M. Schmidt, G. L. Siscoe, R. Vainio, S. K. Antiochos, and P. Riley. CME Theory and Models. *Space Sci. Rev.*, 123:251–302, March 2006. doi: 10.1007/s11214-006-9019-8.

S. E. Forbush. On the effects in cosmic-ray intensity observed during the recent magnetic storm. *Phys. Rev.*, 51(12):1108–1109, Jun 1937. doi: 10.1103/PhysRev.51.1108.3.

P. V. Foukal. *Solar Astrophysics, 2nd, Revised Edition.* WILEY-VCH Verlag CmbH & Co. KCaA, Weinheim, 2004.

W. D. Gonzalez and B. T. Tsurutani. Criteria of interplanetary parameters causing intense magnetic storms (Dst of less than -100 nT). *Planet. Space. Sci.*, 35:1101–1109, September 1987. doi: 10.1016/0032-0633(87)90015-8.

W. D. Gonzalez, A. L. C. de Gonzalez, A. Dal Lago, B. T. Tsurutani, J. K. Arballo, G. K. Lakhina, B. Buti, C. M. Ho, and S.-T. Wu. Magnetic cloud field intensities and solar wind velocities. *Geophys. Res. Lett.*, 25:963–966, 1998. doi: 10.1029/98GL00703.

N. Gopalswamy, Z. Mikić, D. Maia, D. Alexander, H. Cremades, P. Kaufmann, D. Tripathi, and Y.-M. Wang. The Pre-CME Sun. *Space Sci. Rev.*, 123:303–339, March 2006. doi: 10.1007/s11214-006-9020-2.

N. Gopalswamy, S. Yashiro, and S. Akiyama. Geoeffectiveness of halo coronal mass ejections. *J. Geophys. Res.*, 112:A06112, June 2007. doi: 10.1029/2006JA012149.

J. T. Gosling and V. J. Pizzo. Formation and Evolution of Corotating Interaction Regions and their Three Dimensional Structure. *Space Sci. Rev.*, 89:21–52, July 1999. doi: 10.1023/A:1005291711900.

B. Gui, C. Shen, Y. Wang, P. Ye, J. Liu, S. Wang, and X. Zhao. Quantitative Analysis of CME Deflections in the Corona. *Sol. Phys.*, 271:111–139, July 2011. doi: 10.1007/s11207-011-9791-9.

V. F. Hess and A. Demmelmair. World-wide Effect in Cosmic Ray Intensity, as Observed during a Recent Magnetic Storm. 140:316–317, August 1937. doi: 10.1038/140316a0.

R. A. Howard. A Historical Perspective on Coronal Mass Ejections. *Washington DC American Geophysical Union Geophysical Monograph Series*, 165:7, October 2006.

R. A. Howard, N. R. Sheeley, Jr., D. J. Michels, and M. J. Koomen. Coronal mass ejections - 1979-1981. *J. Geophys. Res.*, 90:8173–8191, September 1985. doi: 10.1029/JA090iA09p08173.

H. S. Hudson, J.-L. Bougeret, and J. Burkepile. Coronal Mass Ejections: Overview of Observations. *Space Sci. Rev.*, 123:13–30, March 2006. doi: 10.1007/s11214-006-9009-x.

R. M. E. Illing and A. J. Hundhausen. Observation of a coronal transient from 1.2 to 6 solar radii. *J. Geophys. Res.*, 90:275–282, January 1985. doi: 10.1029/JA090iA01p00275.

A. Isavnin, A. Vourlidas, and E. K. J. Kilpua. Three-Dimensional Evolution of Erupted Flux Ropes from the Sun (2 - 20 R ?) to 1 AU. *Sol. Phys.*, 284:203–215, May 2013. doi: 10.1007/s11207-012-0214-3.

N. Iucci, M. Parisi, M. Storini, and G. Villoresi. High-speed solar-wind streams and galactic cosmic-ray modulation. *Nuovo Cimento C Geophysics Space Physics C*, 2: 421–438, August 1979. doi: 10.1007/BF02558283.

J. R. Jokipii. Propagation of cosmic rays in the solar wind. *Reviews of Geophysics and Space Physics*, 9:27–87, 1971. doi: 10.1029/RG009i001p00027.

A. P. Jordan, H. E. Spence, J. B. Blake, and D. N. A. Shaul. Revisiting two-step Forbush decreases. *J. Geophys. Res.*, 116:A11103, November 2011. doi: 10.1029/2011JA016791.

S. W. Kahler, R. L. Moore, S. R. Kane, and H. Zirin. Filament eruptions and the impulsive phase of solar flares. *Astrophys. J.*, 328:824–829, May 1988. doi: 10.1086/166340.

M. L. Kaiser, T. A. Kucera, J. M. Davila, O. C. St. Cyr, M. Guhathakurta, and E. Christian. The STEREO Mission: An Introduction. *Space Sci. Rev.*, 136:5–16, April 2008. doi: 10.1007/s11214-007-9277-0.

R. P. Kane. How good is the relationship of solar and interplanetary plasma parameters with geomagnetic storms? *J. Geophys. Res.*, 110:A02213, February 2005. doi: 10.1029/2004JA010799.

R.-S. Kim, K.-S. Cho, Y.-J. Moon, M. Dryer, J. Lee, Y. Yi, K.-H. Kim, H. Wang, Y.-D. Park, and Y. H. Kim. An empirical model for prediction of geomagnetic storms using initially observed CME parameters at the Sun. *J. Geophys. Res.*, 115:A12108, December 2010. doi: 10.1029/2010JA015322.

H. E. J. Koskinen and K. E. J. Huttunen. Geoeffectivity of Coronal Mass Ejections. *Space Sci. Rev.*, 124:169–181, June 2006. doi: 10.1007/s11214-006-9103-0.

A. S. Krieger, A. F. Timothy, and E. C. Roelof. A Coronal Hole and Its Identification as the Source of a High Velocity Solar Wind Stream. *Sol. Phys.*, 29:505–525, April 1973. doi: 10.1007/BF00150828.

Y. Kubo and H. Shimazu. Effect of Finite Larmor Radius on Cosmic-ray Penetration into an Interplanetary Magnetic Flux Rope. *Astrophys. J.*, 720:853–861, September 2010. doi: 10.1088/0004-637X/720/1/853.

P. Kühl, S. Banjac, B. Heber, J. Labrenz, R. Müller-Mellin, and C. Terasa. Extended measurement capabilities of the electron proton helium instrument aboard SOHO - understanding single detector count rates. *CEAB, In Press*, 2015.

A. Kumar and Badruddin. Interplanetary Coronal Mass Ejections, Associated Features, and Transient Modulation of Galactic Cosmic Rays. *Sol. Phys.*, 289:2177–2205, June 2014. doi: 10.1007/s11207-013-0465-7.

T. Kuwabara, J. W. Bieber, P. Evenson, K. Munakata, S. Yasue, C. Kato, A. Fushishita, M. Tokumaru, M. L. Duldig, J. E. Humble, M. R. Silva, A. Dal Lago, and N. J. Schuch. Determination of interplanetary coronal mass ejection geometry and orientation from ground-based observations of galactic cosmic rays. *J. Geophys. Res.*, 114:A05109, May 2009. doi: 10.1029/2008JA013717.

J. A. Le Roux and M. S. Potgieter. The simulation of Forbush decreases with time-dependent cosmic-ray modulation models of varying complexity. *Astron. Astrophys.*, 243:531–545, March 1991.

L. Lefevre, S. Vennerstrom, M. Dumbović, B. Vršnak, D. Sudar, R. Arlt, F. Clette, and N. Crosby. Detailed Analysis of solar data related to historical extreme geomagnetic storms: 1868-2010. *submitted to Sol. Phys.*, 2015.

R. P. Lepping, L. F. Burlaga, and J. A. Jones. Magnetic field structure of interplanetary magnetic clouds at 1 AU. *J. Geophys. Res.*, 95:11957–11965, August 1990. doi: 10.1029/JA095iA08p11957.

R. P. Lepping, M. H. Acũna, L. F. Burlaga, W. M. Farrell, J. A. Slavin, K. H. Schatten, F. Mariani, N. F. Ness, F. M. Neubauer, Y. C. Whang, J. B. Byrnes, R. S. Kennon, P. V. Panetta, J. Scheifele, and E. M. Worley. The Wind Magnetic Field Investigation. *Space Sci. Rev.*, 71:207–229, February 1995. doi: 10.1007/BF00751330.

John A. Lockwood. Forbush decreases in the cosmic radiation. *Space Sci. Rev.*, 12: 658–715, 1971. ISSN 0038-6308. URL `http://dx.doi.org/10.1007/BF00173346`. 10.1007/BF00173346.

P. K. Manoharan. Evolution of Coronal Mass Ejections in the Inner Heliosphere: A Study Using White-Light and Scintillation Images. *Sol. Phys.*, 235:345–368, May 2006. doi: 10.1007/s11207-006-0100-y.

D. Maričić, B. Vršnak, A. L. Stanger, A. M. Veronig, M. Temmer, and D. Roša. Acceleration Phase of Coronal Mass Ejections: II. Synchronization of the Energy Release in the Associated Flare. *Sol. Phys.*, 241:99–112, March 2007. doi: 10.1007/s11207-007-0291-x.

D. J. McComas, S. J. Bame, P. Barker, W. C. Feldman, J. L. Phillips, P. Riley, and J. W. Griffee. Solar Wind Electron Proton Alpha Monitor (SWEPAM) for the Advanced Composition Explorer. *Space Sci. Rev.*, 86:563–612, July 1998. doi: 10.1023/A:1005040232597.

Y.-J. Moon, G. S. Choe, H. Wang, Y. D. Park, and C. Z. Cheng. Relationship Between CME Kinematics and Flare Strength. *Journal of Korean Astronomical Society*, 36: 61–66, June 2003.

C. Möstl and J. A. Davies. Speeds and Arrival Times of Solar Transients Approximated by Self-similar Expanding Circular Fronts. *Sol. Phys.*, 285:411–423, July 2013. doi: 10.1007/s11207-012-9978-8.

C. Möstl, C. Miklenic, C. J. Farrugia, M. Temmer, A. Veronig, A. B. Galvin, B. Vršnak, and H. K. Biernat. Two-spacecraft reconstruction of a magnetic cloud and

comparison to its solar source. *Ann. Geophys.*, 26:3139–3152, October 2008. doi: 10.5194/angeo-26-3139-2008.

C. Möstl, T. Rollett, R. A. Frahm, Y. D. Liu, D. M. Long, R. C. Colaninno, M. A. Reiss, M. Temmer, C. J. Farrugia, A. Posner, M. Dumbović, M. Janvier, P. Démoulin, P. Boakes, A. Devos, E. Kraaikamp, M. L. Mays, and B. Vršnak. Strong coronal channelling and interplanetary evolution of a solar storm up to Earth and Mars. *Nature Communications*, 6:7135, May 2015. doi: 10.1038/ncomms8135.

K. W. Ogilvie, D. J. Chornay, R. J. Fritzenreiter, F. Hunsaker, J. Keller, J. Lobell, G. Miller, J. D. Scudder, E. C. Sittler, Jr., R. B. Torbert, D. Bodet, G. Needell, A. J. Lazarus, J. T. Steinberg, J. H. Tappan, A. Mavretic, and E. Gergin. SWE, A Comprehensive Plasma Instrument for the Wind Spacecraft. *Space Sci. Rev.*, 71:55–77, February 1995. doi: 10.1007/BF00751326.

E. N. Parker. Theory of streaming of cosmic rays and the diurnal variation. *Planet. Space. Sci.*, 12:735–749, August 1964. doi: 10.1016/0032-0633(64)90054-6.

E. N. Parker. The passage of energetic charged particles through interplanetary space. *Planet. Space. Sci.*, 13:9–+, January 1965. doi: 10.1016/0032-0633(65)90131-5.

E. Priest and T. Forbes. *Magnetic Reconnection*. Cambridge University Press, Cambridge, UK, 2000.

E. R. Priest and T. G. Forbes. The magnetic nature of solar flares. *Astron. Astrophys. Rev.*, 10:313–377, 2002. doi: 10.1007/s001590100013.

I. G. Richardson. Energetic Particles and Corotating Interaction Regions in the Solar Wind. *Space Sci. Rev.*, 111:267–376, April 2004. doi: 10.1023/B:SPAC.0000032689.52830.3e.

I. G. Richardson and H. V. Cane. Near-Earth Interplanetary Coronal Mass Ejections During Solar Cycle 23 (1996 - 2009): Catalog and Summary of Properties. *Sol. Phys.*, 264:189–237, June 2010. doi: 10.1007/s11207-010-9568-6.

I. G. Richardson and H. V. Cane. Galactic Cosmic Ray Intensity Response to Interplanetary Coronal Mass Ejections/Magnetic Clouds in 1995 - 2009. *Sol. Phys.*, 270:609–627, June 2011a. doi: 10.1007/s11207-011-9774-x.

I. G. Richardson and H. V. Cane. Geoeffectiveness (Dst and Kp) of interplanetary coronal mass ejections during 1995-2009 and implications for storm forecasting. *Space Weather*, 9:S07005, July 2011b. doi: 10.1029/2011SW000670.

I. G. Richardson, G. Wibberenz, and H. V. Cane. The relationship between recurring cosmic ray depressions and corotating solar wind streams at ≤ 1 AU: IMP 8 and Helios 1 and 2 anticoincidence guard rate observations. *J. Geophys. Res.*, 101:13483–13496, June 1996. doi: 10.1029/96JA00547.

E. Robbrecht, S. Patsourakos, and A. Vourlidas. No Trace Left Behind: STEREO Observation of a Coronal Mass Ejection Without Low Coronal Signatures. *Astrophys. J.*, 701:283–291, August 2009. doi: 10.1088/0004-637X/701/1/283.

L. Rodriguez, A. N. Zhukov, C. Cid, Y. Cerrato, E. Saiz, H. Cremades, S. Dasso, M. Menvielle, A. Aran, C. Mandrini, S. Poedts, and B. Schmieder. Three frontside full halo coronal mass ejections with a nontypical geomagnetic response. *Space Weather*, 7: S06003, June 2009. doi: 10.1029/2008SW000453.

T. Rollett, C. Möstl, M. Temmer, A. M. Veronig, C. J. Farrugia, and H. K. Biernat. Constraining the Kinematics of Coronal Mass Ejections in the Inner Heliosphere with In-Situ Signatures. *Sol. Phys.*, 276:293–314, February 2012. doi: 10.1007/s11207-011-9897-0.

G. Rostoker. Geomagnetic indices. *Reviews of Geophysics and Space Physics*, 10:935–950, 1972. doi: 10.1029/RG010i004p00935.

A. P. Rouillard. Relating white light and in situ observations of coronal mass ejections: A review. *J. Atmos. Sol. Terr. Phys.*, 73:1201–1213, June 2011. doi: 10.1016/j.jastp.2010.08.015.

B. Schmieder, G. Aulanier, and B. Vršnak. Flare-CME Models: An Observational Perspective (Invited Review). *Sol. Phys.*, June 2015. doi: 10.1007/s11207-015-0712-1.

R. Schwenn, A. dal Lago, E. Huttunen, and W. D. Gonzalez. The association of coronal mass ejections with their effects near the Earth. *Astrophys. J.*, 23:1033–1059, March 2005.

R. Schwenn, J. C. Raymond, D. Alexander, A. Ciaravella, N. Gopalswamy, R. Howard, H. Hudson, P. Kaufmann, A. Klassen, D. Maia, G. Munoz-Martinez, M. Pick, M. Reiner, N. Srivastava, D. Tripathi, A. Vourlidas, Y.-M. Wang, and J. Zhang. Coronal Observations of CMEs. Report of Working Group A. *Space Sci. Rev.*, 123:127–176, March 2006. doi: 10.1007/s11214-006-9016-y.

G. Siscoe and R. Schwenn. CME Disturbance Forecasting. *Space Sci. Rev.*, 123:453–470, March 2006. doi: 10.1007/s11214-006-9024-y.

C. W. Smith, J. L'Heureux, N. F. Ness, M. H. Acuña, L. F. Burlaga, and J. Scheifele. The ACE Magnetic Fields Experiment. *Space Sci. Rev.*, 86:613–632, July 1998. doi: 10.1023/A:1005092216668.

N. Srivastava. A logistic regression model for predicting the occurrence of intense geomagnetic storms. *Ann. Geophys.*, 23(9):2969–2974, 2005. doi: 10.5194/angeo-23-2969-2005. URL `http://www.ann-geophys.net/23/2969/2005/`.

N. Srivastava and P. Venkatakrishnan. Solar and interplanetary sources of major geomagnetic storms during 1996-2002. *J. Geophys. Res.*, 109:A10103, October 2004. doi: 10.1029/2003JA010175.

O. C. St. Cyr, S. P. Plunkett, D. J. Michels, S. E. Paswaters, M. J. Koomen, G. M. Simnett, B. J. Thompson, J. B. Gurman, R. Schwenn, D. F. Webb, E. Hildner, and P. L. Lamy. Properties of coronal mass ejections: SOHO LASCO observations from January 1996 to June 1998. *J. Geophys. Res.*, 105:18169–18186, 2000. doi: 10.1029/1999JA000381.

D. Stirzaker. *Elementary Probability*. Cambridge University Press, New York, 2003.

E. C. Stone, A. M. Frandsen, R. A. Mewaldt, E. R. Christian, D. Margolies, J. F. Ormes, and F. Snow. The Advanced Composition Explorer. *Space Sci. Rev.*, 86:1–22, July 1998. doi: 10.1023/A:1005082526237.

D. Sudar, B. Vršnak, and M. Dumbović. Predicting Coronal Mass Ejections transit times to Earth with neural network. *Mon. Not. R. Astron. Soc.*, November 2015.

M. Temmer, B. Vršnak, T. Rollett, B. Bein, C. A. de Koning, Y. Liu, E. Bosman, J. A. Davies, C. Möstl, T. Žic, A. M. Veronig, V. Bothmer, R. Harrison, N. Nitta, M. Bisi, O. Flor, J. Eastwood, D. Odstrcil, and R. Forsyth. Characteristics of Kinematics of a Coronal Mass Ejection during the 2010 August 1 CME-CME Interaction Event. *Astrophys. J.*, 749:57, April 2012. doi: 10.1088/0004-637X/749/1/57.

A. K. Tiwari, A. Singh, and S. P. Agrawal. Study of the Diurnal Variation of Cosmic Rays during Different Phases of Solar Activity. *Sol. Phys.*, 279:253–267, July 2012. doi: 10.1007/s11207-012-9962-3.

B. T. Tsurutani, W. D. Gonzalez, G. S. Lakhina, and S. Alex. The extreme magnetic storm of 1-2 September 1859. *J. Geophys. Res.*, 108:1268, July 2003. doi: 10.1029/2002JA009504.

J. Uwamahoro, L. A. McKinnell, and J. B. Habarulema. Estimating the geoeffectiveness of halo CMEs from associated solar and IP parameters using neural networks. *Ann. Geophys.*, 30:963–972, June 2012. doi: 10.5194/angeo-30-963-2012.

F. Valach, M. Revallo, J. Bochníček, and P. Hejda. Solar energetic particle flux enhancement as a predictor of geomagnetic activity in a neural network-based model. *Space Weather*, 7:S04004, April 2009. doi: 10.1029/2008SW000421.

S. Vennerstrom, L. Lefevre, M. Dumbović, N. Crosby, O. Malandraki, I. Patsou, F. Clette, A. Veronig, B. Vršnak, K. Leer, and T. Moretto. Extreme geomagnetic storms - 1868-2010. *submitted to Sol. Phys.*, 2015.

G. Verbanac, M. Mandea, B. Vršnak, and S. Sentic. Evolution of Solar and Geomagnetic Activity Indices, and Their Relationship: 1960 - 2001. *Sol. Phys.*, 271:183–195, July 2011a. doi: 10.1007/s11207-011-9801-y.

G. Verbanac, B. Vršnak, S. Živković, T. Hojsak, A. M. Veronig, and M. Temmer. Solar wind high-speed streams and related geomagnetic activity in the declining phase of solar cycle 23. *Astron. Astrophys.*, 533:A49, September 2011b. doi: 10.1051/0004-6361/201116615.

G. Verbanac, S. Živković, B. Vršnak, M. Bandić, and T. Hojsak. Comparison of geoeffectiveness of coronal mass ejections and corotating interaction regions. *Astron. Astrophys.*, 558:A85, October 2013. doi: 10.1051/0004-6361/201220417.

A. Vourlidas and R. A. Howard. The Proper Treatment of Coronal Mass Ejection Brightness: A New Methodology and Implications for Observations. *Astrophys. J.*, 642: 1216–1221, May 2006. doi: 10.1086/501122.

A. Vourlidas, R. Colaninno, T. Nieves-Chinchilla, and G. Stenborg. The First Observation of a Rapidly Rotating Coronal Mass Ejection in the Middle Corona. *Astrophys. J. Lett.*, 733:L23, June 2011. doi: 10.1088/2041-8205/733/2/L23.

B. Vršnak, D. Maričić, A. L. Stanger, and A. Veronig. Coronal Mass Ejection of 15 May 2001: II. Coupling of the Cme Acceleration and the Flare Energy Release. *Sol. Phys.*, 225:355–378, December 2004a. doi: 10.1007/s11207-004-4995-x.

B. Vršnak, D. Ruždjak, D. Sudar, and N. Gopalswamy. Kinematics of coronal mass ejections between 2 and 30 solar radii. What can be learned about forces governing the eruption? *Astron. Astrophys.*, 423:717–728, August 2004b. doi: 10.1051/0004-6361:20047169.

B. Vršnak, D. Sudar, and D. Ruždjak. The CME-flare relationship: Are there really two types of CMEs? *Astron. Astrophys.*, 435:1149–1157, June 2005. doi: 10.1051/0004-6361:20042166.

B. Vršnak, T. Žic, D. Vrbanec, M. Temmer, T. Rollett, C. Möstl, A. Veronig, J. Čalogović, M. Dumbović, S. Lulić, Y.-J. Moon, and A. Shanmugaraju. Propagation of Interplanetary Coronal Mass Ejections: The Drag-Based Model. *Sol. Phys.*, 285:295–315, July 2013. doi: 10.1007/s11207-012-0035-4.

Y. Wang, C. Shen, S. Wang, and P. Ye. Deflection of coronal mass ejection in the interplanetary medium. *Sol. Phys.*, 222:329–343, August 2004. doi: 10.1023/B:SOLA.0000043576.21942.aa.

Y. Wang, X. Xue, C. Shen, P. Ye, S. Wang, and J. Zhang. Impact of Major Coronal Mass Ejections on Geospace during 2005 September 7-13. *Astrophys. J.*, 646:625–633, July 2006. doi: 10.1086/504676.

A. Warmuth. Large-scale Waves and Shocks in the Solar Corona. In K.-L. Klein and A. L. MacKinnon, editors, *Lecture Notes in Physics, Berlin Springer Verlag*, volume 725 of *Lecture Notes in Physics, Berlin Springer Verlag*, page 107, 2007.

D. F. Webb and R. A. Howard. The solar cycle variation of coronal mass ejections and the solar wind mass flux. *J. Geophys. Res.*, 99:4201–4220, March 1994. doi: 10.1029/93JA02742.

D. F. Webb and T. A. Howard. Coronal Mass Ejections: Observations. *Living Reviews in Solar Physics*, 9:3, June 2012. doi: 10.12942/lrsp-2012-3.

G. Wibberenz, H. V. Cane, and I. G. Richardson. Two-step Forbush Decreases in thecInner Solar System and their Relevance for Models of Transient Disturbances. *International Cosmic Ray Conference*, 1:397, 1997.

G. Wibberenz, J. A. Le Roux, M. S. Potgieter, and J. W. Bieber. Transient Effects and Disturbed Conditions. *Space Sci. Rev.*, 83:309–348, January 1998.

S. Yashiro, N. Gopalswamy, G. Michalek, O. C. St. Cyr, S. P. Plunkett, N. B. Rich, and R. A. Howard. A catalog of white light coronal mass ejections observed by the SOHO spacecraft. *J. Geophys. Res.*, 109:A07105, July 2004. doi: 10.1029/2003JA010282.

S. Yashiro, S. Akiyama, N. Gopalswamy, and R. A. Howard. Different Power-Law Indices in the Frequency Distributions of Flares with and without Coronal Mass Ejections. *Astrophys. J. Lett.*, 650:L143–L146, October 2006. doi: 10.1086/508876.

S. Yashiro, G. Michalek, S. Akiyama, N. Gopalswamy, and R. A. Howard. Spatial Relationship between Solar Flares and Coronal Mass Ejections. *Astrophys. J.*, 673:1174–1180, February 2008. doi: 10.1086/524927.

Y. I. Yermolaev, N. S. Nikolaeva, I. G. Lodkina, and M. Y. Yermolaev. Geoeffectiveness and efficiency of CIR, sheath, and ICME in generation of magnetic storms. *J. Geophys. Res.*, 117:A00L07, May 2012. doi: 10.1029/2011JA017139.

J. Zhang, K. P. Dere, R. A. Howard, M. R. Kundu, and S. M. White. On the Temporal Relationship between Coronal Mass Ejections and Flares. *Astrophys. J.*, 559:452–462, September 2001. doi: 10.1086/322405.

J. Zhang, K. P. Dere, R. A. Howard, and V. Bothmer. Identification of Solar Sources of Major Geomagnetic Storms between 1996 and 2000. *Astrophys. J.*, 582:520–533, January 2003. doi: 10.1086/344611.

J. Zhang, I. G. Richardson, D. F. Webb, N. Gopalswamy, E. Huttunen, J. C. Kasper, N. V. Nitta, W. Poomvises, B. J. Thompson, C.-C. Wu, S. Yashiro, and A. N. Zhukov. Solar and interplanetary sources of major geomagnetic storms (Dst $\leq$ -100 nT) during 1996-2005. *J. Geophys. Res.*, 112:A10102, October 2007. doi: 10.1029/2007JA012321.

T. H. Zurbuchen and I. G. Richardson. In-Situ Solar Wind and Magnetic Field Signatures of Interplanetary Coronal Mass Ejections. *Space Sci. Rev.*, 123:31–43, March 2006. doi: 10.1007/s11214-006-9010-4.

Curriculum Vitae

Mateja Dumbović (born 1984 in Zagreb) is employed as a PhD student at Hvar Observatory (Faculty of Geodesy, Uni. Zagreb) since 2010, after finishing physics studies at the Uni. Zagreb with diploma work: "Influence of the interplanetary coronal mass ejections on the cosmic ray flux" under supervision of dr. sc. Bojan Vršnak. In the course of her scientific career she participated in 7 national, bilateral and EU FP7 scientific projects. She published 1 conference proceeding and 12 peer-reviewed papers (5 as first author) in international scientific journals, most notably in *Nature Communications*, *Astrophysical Journal Supplement series*, *Astronomy & Astrophysics*, and *Solar Physics*. 3 papers are currently submitted to international scientific peer reviewed journals. The research is mostly related to heliophysics and space weather, namely propagation of coronal mass ejections and their space weather effects. Her presence in the international scientific conferences includes 32 abstracts (6 oral and 8 poster presentations as the presenting author). In addition, she participated in three scientific workshops and 2 international schools and had a dozen of research stays and project meetings at prominent European solar physics and space weather related institutions in Austria, Belgium, Denmark, France, Greece and Germany. In 2014 she was awarded with a research grant for her project proposal at the Royal Observatory of Belgium and she won best student poster award at the 24th European Cosmic Ray Symposium. Her scientific community services include reviews for *Solar Physics*, *Journal of Atmospheric and Solar-Terrestrial Physics*, *Journal of Physics: Conference Series* and *Central European Astrophysical Bulletin*, as well as participation in the local organizing committee at 5 international scientific conferences and workshops. She is a member of Croatian Astronomical Society, Croatian Physical Society, and National Commission for Competition in Astronomy.

List of papers

Papers published in refereed journals:

1. Sudar, D., Vršnak, B., and **Dumbović, M.**, Predicting Coronal Mass Ejections transit time to Earth with neural network, *Monthly Notices of the Royal Astronomical Society, In Press*, 2015

2. Heber, B., Wallmann, C., Galsdorf, D., Herbst, K., Kühl, P., **Dumbović, M.**, Vršnak, B., Veronig, A., Temmer, M., Möstl, C., and Dalla, S., Forbush decreases associated to stealth coronal mass ejections, *Central European Astrophysical Bulletin*, Volume 39, pp. 75-82, 2015

3. **Dumbović, M.**, Vršnak, B., and Čalogović, J., Forbush decrease prediction based on the remote solar observations, *Solar Physics, online first*, 2015

4. Möstl, C., Rollett, T., Frahm, R. A., Liu, Y. D., Long, D. M., Colaninno, R. C., Reiss, M. A., Temmer, M., Farrugia, C. J., Posner, A., **Dumbović, M.**, Janvier, M., Démoulin, P., Boakes, P., Devos, A., Kraaikamp, E., Mays, M. L., and Vršnak, B., Strong coronal channelling and interplanetary evolution of a solar storm up to Earth and Mars, *Nature Communications*, Volume 6, id. 7135, 2015

5. **Dumbović, M.**, Devos, A., Vršnak, B., Sudar, D., Rodriguez, L., Ruždjak, D., Leer, K., Vennerstrom, S., and Veronig, A., Geoeffectiveness of Coronal Mass Ejections in the SOHO Era, *Solar Physics*, Volume 290, Issue 2, pp.579-612, 2015

6. Vršnak, B., Temmer, M., Žic, T., Taktakishvili, A., **Dumbović**, M., Möstl, C., Veronig, A. M., Mays, M. L., and Odstrcil, D., Heliospheric Propagation of Coronal Mass Ejections: Comparison of Numerical WSA-ENLIL+Cone Model and Analytical Drag-based Model, *The Astrophysical Journal Supplement*, Volume 213, Issue 2, article id. 21, 9 pp., 2014

7. Maričić, D., Vršnak, B., **Dumbović**, M., Žic, T., Roša, D., Hržina, D., Lulić, S., Romštajn, I., Busić, I., Salamon, K., Temmer, M., Rollett, T., Veronig, A., Bostanjyan, N., Chilingarian, A., Mailyan, B., Arakelyan, K., Hovhannisyan, A., and Mujić, N., Kinematics of Interacting ICMEs and Related Forbush Decrease: Case Study, *Solar Physics*, Volume 289, Issue 1, pp.351-368, 2014

8. Vršnak, B., Žic, T., Vrbanec, D., Temmer, M., Rollett, T., Möstl, C., Veronig, A., Čalogović, J., **Dumbović**, M., Lulić, S., Moon, Y.-J., and Shanmugaraju, A., Propagation of Interplanetary Coronal Mass Ejections: The Drag-Based Model, *Solar Physics*, Volume 285, Issue 1-2, pp. 295-315, 2013

9. **Dumbović, M.**, Vršnak, B., Čalogović, J., Župan, R., Cosmic ray modulation by different types of solar wind disturbances, *Astronomy & Astrophysics*, Volume 538, id.A28, 13 pp., 2012

10. Čalogović, J., **Dumbović, M.**, Novak, N., Vršnak, B., Brajša, R., Pötzi, W., Hirtenfellner-Polanec, W., Veronig, A., Hanslmeier, A., Klvana, M., and Ambrož, P., Solar Halpha and white light telescope at Hvar Observatory, *Central European Astrophysical Bulletin*, Vol. 36, p. 83-88, 2012

11. **Dumbović, M.**, Vršnak, B., and Čalogović, J., Solar influences on the short-term cosmic ray modulation, *Central European Astrophysical Bulletin*, Vol. 36, p. 65-70, 2012

12. **Dumbović, M.**, Vršnak, B., Čalogović, J., and Karlica, M., Cosmic ray modulation by solar wind disturbances, *Astronomy & Astrophysics*, Volume 531, id.A91, 17 pp., 2011

Papers published in proceedings:

1. Maričić, D., Bostasyan, N., **Dumbović, M.**, Chilingarian, A., Mailyan, B., Rostomyan, H., Arakelyan, K., Vršnak, B., Roša, D., Hržina, D., Romštajn, I., and Veronig, A., The Successive CME on 13th; 14th and 15th February 2011 and Forbush decrease on 18 February 2011, *Journal of Physics: Conference Series*, Volume 409, Issue 1, article id. 012158, 2013

Papers submitted to refereed journals:

1. Vršnak, B., Rollett, T., Veronig, A.M., Möstl, C., Amerstorfer, U.V., Leitner, M., Temmer, M., Ibsen, T., Vennerstrom, S., **Dumbović, M.**, Žic, T., Farrugia, C.J., and Isavnin, A., Heliospheric Evolution of Magnetic Clouds, *The Astrophysical Journal, submitted*, 2015

2. Lefevre, L., Vennerstrom, S., **Dumbović, M.**, Arlt, R., Vaquero, J.M., Vršnak, B., Sudar, D., Clette, F., and Crosby, N., Detailed Analysis of solar data related to historical extreme geomagnetic storms: 1868-2010, *Solar Physics, submitted*, 2015

3. Vennerstrom, S., Lefevre,L., **Dumbović, M.**, Crosby, N., Malandraki, O., Patsou, I., Clette, F., Veronig, A.;,Vršnak, B., Leer, K., and Moretto, T., Extreme geomagnetic storms - 1868-2010, *Solar Physics, submitted*, 2015