An interpretable machine learning method for forecasting the SYM-H Index

Daniel Iong^{1,1,1}, Yang Chen^{2,2,2}, Gabor Toth^{3,3,3}, Shasha Zou^{3,3,3}, Tuija I. Pulkkinen^{2,2,2}, Jiaen Ren^{4,4,4}, Enrico Camporeale^{5,5,5}, and Tamas I. I. Gombosi^{3,3,3}

¹University of Michigan, Ann Arbor ²University of Michigan ³University of Michigan-Ann Arbor ⁴University of Michigan - Ann Arbor ⁵University of Colorado

November 30, 2022

Abstract

In this work, we develop gradient boosting machines (GBMs) for forecasting the SYM-H index multiple hours ahead using different combinations of solar wind and interplanetary magnetic field (IMF) parameters, derived parameters, and past SYM-H values. Using Shapley Additive Explanation (SHAP) values to quantify the contributions from each input to predictions of the SYM-H index from GBMs, we show that our predictions are consistent with physical understanding while also providing insight into the complex relationship between the solar wind and Earth's ring current. We also perform a direct comparison between GBMs and neural networks presented in prior publications for forecasting the SYM-H index by training, validating, and testing them on the same data. We find that the GBMs have a comparable root mean squared error as the best published black-box neural network schemes and GBMs have better Heidke Skill Scores at predicting strong storms.

New Findings from Explainable SYM-H Forecasting using Gradient Boosting Machines

Daniel Iong¹, Yang Chen¹, Gabor Toth², Shasha Zou², Tuija Pulkkinen², Jiaen Ren², Enrico Camporeale^{3,4}, Tamas Gombosi²

¹Dept. of Statistics, University of Michigan, Ann Arbor, MI, USA ²Dept. of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA ³CIRES, University of Colorado, Boulder, CO, USA ⁴NOAA Space Weather Prediction Center, CO, USA

9	Key Points:		
10	• Spatish (BB M) fight Hit		
11	jabala.		
12	• Synthetic B AP by		
13	echich		
14	• Opbligstigh ligen		
15	h		

1

2

3

4

Corresponding author: Daniel Iong, **Bd**

- Abstract 16 Index contribution (CB M) fried B/4 17 Histophilip 18 spienie (IMF) pan ispan ab SA Hisr sU 19 igfatElp(SIP british 20 public Histo BM and Histo BM 21 isi jagi ki jigi kipa 22 jababasidaEba signa Inpa odda 23
- fissiehtse defetien 24 ide B Maddadi ide fe
- 25
- ight84 Hishiga, igt algebalandet & W 26
- fichter Mstachargestisje 27
- ka katakan 28
- Plain Language Summary 29
- Feijiblighfiele 30
- igendelisensiel Indew opt 31 attention Hitselinen
- 32 istilsid& ipiqfidhen (siSA Hisr el 33
- bie an and a state of the state
- 34
- 35 ġ6
- 36

1 Introduction 37

- (desoldeide) ister 38 ekkennetistetsi Digen 39 ietsn jefgebeEta steffebeidv 40 filajini higi higi da 41 tela, 2017). Högenisken Dord Star Highenia 42 iddidgikakatagiji da 43 en fiele (1982 Ken 1991). Hereit 44 deteinete keitideten 45 sistering N. F. 46 1989; N. A 🚛 1995, 2002a 2002b, stiplign 47 **jatilijijihtt** Lu 48 tha, 2020), spolitility 49 flux(Heath, 2015) ethetat(Bistelah, 2018). 50 & Huilinishish 51 facilei fa 52 **Geleterbeitelsbei**ch 53 SV4 Hist Casela (2010) al Bibably 2019) ala NabAu 54 Radia NAR Hatel ina 55 epüetSA HistobilepiSA Hist isidal 56 IMF perip. Catela (2010) idditate 57 67 juni 1998 t2006, to B keth (2019) std. 58 625 Hush2006 62013. NNGH 59
- jebish fela (2021) islan tayan LSM) 60 ata (CNN) itaziptetsia Historia 61 istplinf tache 134 Hasip Childhea 62
- (2021) beite te Historia by 63
- bebefectues Hereizue 64
- fan Bibbela (2021) elCkildela (2021) inder 65

< -100 fm fm 998 ts 2017

iddae25 gigs(n Ds 66 abbie 17 januar abbie 18 januar ab is abbie 18 januar abbie 18 67 jefijen ostesenetische 68 cepteh Fotetsfish, ødvætetsfæl (deh 69 jahan Cipatanish ta 1. 70 Millitate 71 bieisteebKriet Astrikite 72 bbbbDsiebbobbbbL bbb(V1994), 73 Geteta (1996), et al (1997).74 bitriperijitep 75 deDsistaskajaebtektelajas 76 engelikeite Deta (2006) et 77 efiski MF pesti petri de Dan 78 elsi isbilisi lini lini bi 79 isten Statisfistor (MF pad 80 shutshalfah (2021) elChi al 81 teh (2021). 82 Adjutitet 83 ishihay ishihadangi Cike 84 a (2017) jehen Gabrie Deiel Lu 85 tela (2016) pelestate (N) telav 86 Byzela (2011) del Nadel Jack Andrea 87 esta NARIAK dibishipata 1 - da 88 aDsafefisphefichyferhid i gen 89 pp Xeh (2020) Hitter Vites 90 abbeet ab Distantiable W 91 batelbeideide (eB a 92 elkka(1999) fælipkingsly. An 93 ffeisbaardefyste bildepi 94 tafalh Addition 95 **jaibbin Cip**(2019). 96 Districted addistantic to ge 97 infactibility bility 98 begildy Exfected for 99 iden testerintsidetet 100 ep(Meteh, 2020). Hydpha 101 generation in the second s 102 I je Ge Gelen and Andrea 103 fipilijan isiskolajaja 104 in a historic de la constant de la const 105 jon Cophan 2019). Ipiecientete 106 ela interactionalitate 107 bith the fisher backs by 108 etitefetig Gatip ten 109 ip iddala isin adatatia 110 geblepelegebe (Ceptera , 2020). 111 Elanded jata biasho 112 heth Ehrinden 113 heig iste attekten fyr he 114 isfarff distlyatifn 115 **bibili keşişisi**si 116 jarist Extended bisep iskel 117

118 **illiptvyfi** (Loth, 2013). **Ph**

jatsh tosta obstratisticka 119 celsen skienelsbille. 120 tiddigh Fraisfielden 121 el eMa(2019). 122 philipid Addeta (2016) plaNa 123 hand English NAR holdsbecking 124 kkoliki (2019) 125 phpNACh6hAE isbbih 126 曲 127 Index extended 1087 Hist 128 telpn telftet idpektigstigs 129 blen Inta bestelbestation ble 130 og bish diskiyanta bishisha 131 SANETA SIAP dilitication 132 dbibliolithe(Lbi)(Le 2017). SAP 133 skijstasjefe debide 134 ficfikkseich Lighta, 2018), kypie (Kkre 135 h, 2020, Μaδω L μα 2021), fine(B bath, 2020, Mateha, 2019), 136 depie (Beta, 2020). 63. 2 istaid 137 c**halsb6472**inbl 138 Heitherights In 62, ¢deb 139 tekentreligget Infe3, etabligt 140 tekken jaga aktivation 141 In 64, eijdeficija pildelja 142 iteiste etsten textstantefen 143 idda 88765 beautifield 144 **Gible i** 145 2 Data 146 ar Date (D) is is the first (in the first of 147 tebenber (Mgt 1980) onekab 148 149 Distribution 150 ebber (5) 1964). Astrikkefikk 151 ighinian interior afficiationa 152 **jepisbleDsielesle**v $-50\,\mathrm{fr}$ 153 8634 Hisisala indischie Diet d 154 sistep1-teel(Iign 1990, SAVE For 2006). HBNA H 155 156 igbiodeitichtige enden edele 157 kteidet oddeide gener 158 teteriteti Gesetid 159 biddeS/A Hisr teh -100< **S**A H < -50 T, Ġ(-250T160 < **S**A H < -100) T, also 184 H < -200161 ANEXA HINDRED HINDC 162 (https://spdf.gsfc.nas in the second states in the second sanf py (Kag 163 2005, AB (Q 2020) . Adder 2 bidger (p 164 jeniek (IMF) prinkak CjaEks (ACE) po 165 finten ASERDaFte https://cdaweb.gsfc.nasa 166 .gov/index.ht)mski/dsh autombi/MFm 167 stibute Matthewise Machine iso field , 1998) 168

```
taa 1.6- selal seksetyisi kepel sebestan
169
                                              tea64 det Inteltos
         e 1898 Mis (Mc letela , 1998) ,
170
         hidpedIMF (pn shelid), itp
171
         desigentefijet sovjetobliste
172
         peripen Nach , 2007).
173
                 Enden iselSHP isstanciples
174
         ia destrictionia
175
         itehtietele ogelSA Hist bidv
176
         ell MF pate5 itte Stady Chi &
177
         teh (2021); Heth (2021).
178
                 Frighter B Ministria 1, 4842 teg
179
         gesighet 998 to 2018 kelenin MA Hin
180
         elefilsh
                                          -100\,\mathrm{fr} lähtseigitet al2
181
         886 filling (sis3. 1) ist
182
         ierentistik kinstereitie
183
         igish Djasti gulstorigi isA
184
         dA
185
```

Table 1. Storms used to train GBMs. These storms are identical to the ones used to train and validate models in Collado-Villaverde et al. (2021).

6##	Sed Edd Mn SV4 H (Spr	
1	1998-02-14 1998-02-22	- 119
2	1998-08-02 1998-08-08	- 168
3	1998-09-19 1998-09-29	- 213
4	1999- 02-16 1999- 02-24	- 127
5	$1999 - 10 \ 15 \ 1999 - 10 \ 25$	- 218
6	2000 07-09 2000 07-19	- 335
7	2000 08- 06 2000 08- 16	- 235
8	2000 09-15 2000 09-25	- 196
9	2000 11- 01 2000 11- 15	- 174
10	2001-03-14 2001-03-24	- 165
11	2001-04-06 2001-04-16	- 275
12	2001-10-17 2001-10-22	- 210
13	201-10-31 201-11-10	- 313
14	2002-05-17 2002-05-27	- 113
15	203-11-15 203-11-25	- 488
16	2004 07-20 2004 07-30	- 208
17	2005 05 10 2005 05 20	- 302
18	2006-04-09 2006-04-19	- 110
19	$1998-12 \ 09 \ 1998-12 \ 19$	- 206
20	2012 08-01 2012 08-11	- 149
21	1998-04-28 1998-08-05	- 268
22	1999- 09- 19 1999- 09- 26	- 160
23	203-10-25 203-11-03	- 427
24	2015 06-18 2015 06-28	- 207
25	2017-09-01 2017-09-11	- 144

186 deb A HA	t ističien
---------------------	------------

187 Vielfichficheite 3. Aliel

188 is Δ t formula MASA Hisish his

189 filipeleskip diligefilde

t, **Etta**

 $y(t{+}\Delta t)$, w

6#	Sehn Eden Ma SNAH (S)T	
26	1998-06-22 1998-06-30	- 120
27	1998-11-02 1998-11-12	- 179
28	1999-01-09 1999-01-18	- 11
29	1999-04-13 1999-04-19	- 122
30	2000 01-16 2000 01-26	- 101
31	2000 04 02 2000 04 12	- 315
32	2000 05-19 2000 05-28	- 159
33	2001-08-26 2001-04-04	- 434
34	203-05-26 203-06-06	- 162
35	203-07-08 203-07-18	- 125
36	2004 01-18 2004 01-27	- 137
37	2004 11- 04 2004 11- 14	- 393
38	2012 09-10 2012 10 05	- 138
39	2013-05-28 2013-06-04	- 134
40	2013-06-26 2013-07-04	- 110
41	2015 03-11 2015 03-21	- 233
42	2018-08-22 2018-09-03	- 205

Table 2.Storms used to test GBMs. These storms are identical to the ones used to test models in Collado-Villaverde et al. (2021).

190	tetebetata fielSA Hisistel to30 ka	feBA	
191	Hisishi Hiji Hili (2021),		
192	ti hefteften	R^2 Hystatis	
193	idi hitiba di 1990 di 1		
194	beSA Heistelehelsenja ≸uta		R^2 t
195	telddylgae B0 ha ffel Mariana		
196	blep MMMA Hødsbep be		R^2 both
197	tybef90to180innahan bay	R^2 sheltin	
198	kkeph 180 km		

Table 3.Features used as input into our models.

Fita High (i)ta	
1 1 1	
$IMF: B_x, B_y, B_z$ ()	2 n 30
Solar ivnd V_x (phip), ρ (phim) 3), T (K)	2 n30
De ri e d quanti ti e βV_x^2 (\mathbf{P} , $E_s = \mathbf{a}0$, $ V_x B_z$) (\mathbf{W}/m)	2 n 30

9	Hiteschietspelike 4 Aufte	
	süsteriketikete	
	édic <mark>ipada</mark> Heteo	
	kägkul MF peripäävi H (tate	I_1 , bet stan
	FoIMF petity tisken2% filis	
	jebachjan Hetv stepiske(tep9%) 6	
	kiljilipa Ayrpililiji	
	h detrijejeje (es	
	io3. 4 i CheiCia (2016) fibil.	

209	Idiahadala	I_3 d I_4 by
210	stýdettjá Inþa ada	
211	(jp ρV_x^2 epigetasjin	
212	SA H. Hiter Construction and the second seco	
213	tsjon. Akjetetiska	<i>y</i> ந
214	the field $E_y = V_x B_z$ the field is	1
215	igfl wight bir Nich	$V_x < 0$ ihiga ha iga
216	(CN) ibyeh Heisebid	$E_s = \mathbf{x}_0 0$, E_y) is the
217	$\mathbf{E}_{y} \texttt{kel} MF \texttt{kell}_{y}$	$B_z < 0$, the
218	tslighe defide MF tighe	
219	statelisters (Batela, 1975). Ital	E_s
220	distrikitige B n	
221	ың ТРОВ њи Мар 2000, ТРОВ ње 2002).	
222	ölikidel MF jeftejildev	
223	ejfetM Her einsteine	I_2 d I_4 help
224	SMI Hbs(sEMI).	

 Table 4.
 Various sets of features used as inputs to train our models.

Ippe	Fall
$\overline{I_1}$	INF, sN A H
I_2	IMF
I_3	INF/bakinda panya H
I_4	INF/ kjá/ikk

225 3 Methods

226	3.1 Gradient Boosting Machines
227	Gilipha (BBM), kinedikis da
228	diki jimovigi i Na&
229	Kyb. 2013). Ev ZelAn (2021) Jefrigstv
230	gCB Ms filthli tsfyrið
231	fi ðrejslabel (B Msæjssiðe
232	ide in the second s
233	gistinining fp (Cells Ch. 2016). Intep
234	ia. (B Ma ddallytastrji an
235	the third of B beta , 2021) $de Det (Ma , 2020).$
236	In stelekterele len ide e
237	
238	jet skolet kan kan se
239	Eşişir bişişi n dav
240	Exatility i itinip
241	idijigi nish jarish bil
242	telegentelset (Figh 2001.). InterstCEBM\$ te
243	brige kerddelite
244	ighti hi i glaft fafte
245	idşhasha Fiya kaciye kacı BMş
246	eprijkejskigte
247	jefige Initiality jeich
248	
249	enter Bin 1996).



Figure 1. Structure of the first tree T_1 learned in a GBM trained with input set I_3 to predict the SYM-H index one hour ahead. The leaf nodes of the tree are shaded gray. The value in each leaf node is its corresponding leaf weight. Left splits correspond to the inequality in the previous node being true, and vice versa.



$$\hat{\mathbf{w}} \qquad y^{(m-1)}(t+\Delta t) = \sum_{k=1}^{m-1} T_k(I(t)) \quad \hat{\mathbf{w}}(T_j) = \gamma K_j + \frac{1}{2}\lambda \sum_{k=1}^{K_j} w_{j,k}^2. \tag{4}$$

253	Inq (4), K_j is defined as	$T_j; u$	$v_{j,k}$ 'sehibelyin	$T_j;$
254	al γ al λ and a set Ω is a set of μ			
255	istyfitigsjyldffisik			
256	ighting Ing	γ biddig		λ bib
257	blig Attligeistys			
258	biadpla Bialga	$\lambda,$	veletificti n	
259	ideler i signale de de la servicie d			
260	taalastagige ØB ø			
261	kaghtefahfrikele			
262	tetitsitetsit separati separati			
263	fahtste Agfyttelinen			
264	titis & CherCia (2016).			
265	alite and a			
266	hapan betar interfak ditetapan			
267	g, itily, abilitie (); . Afa			
268	ipn siltefin igskieligi			
269	förphel Net(fa k for 2018). The second sec			
270	blib(); , dz ajuski kita			
271	abblete fign ste			
272	hitahiji Hilda (BC N			
273	2005). Centre film in the second s			
274	5 te Asha onelasfindagi kiligi ha			
275				
276	legildt septraphility lanstrandel			
277	kielefikken affisciekanseigin			
278	bá			

Table 5.Hyperparameter values for training GBMs using the different input sets in table 4.

Ippe Hppan	₽V	
I_1, I_2	Lippa	0 072
	Mar teppi	4
	Ma Hagy	4
	Chang 0.78	
	# 6 8	84
I_3, I_4	Linge	0 147
	Mar teppi	3
	Ma hegy	2
	China 0 894	
	# 6 b	291

- 279 B Materia B Ms
- 280 de deig ditté
- 281 Éstistertite Apa
- 282 efisisisteri tatoiren
- 283 g CB Mschargerthebby
- 284 **Elighteeth Aife**
- 285 **GSTERGENARTIN**
- 286 **chitteride**((Hich., 2001)). **Here**ifig 2,
- 287 HERE MARKEN SA DE
- 288 **kp** (B M**sektopheis**) Asil
- 289 **Bigschlaught**

ta jej vista na sloven do s 290 selčja, Fija seide A el 2, definde 291 60611266Cld6i1206 292 (B) Machateria point describe 293 isthyphisteris (Killer Gil 294 Bh. 2015). Stablighter 295 jacke Sissisten tuster tylgeskelen. 296 ige dibbin wortefigsin 297 (B) Mathematical B b (2015). Date 298 eljähhav berehtisihen 299 ksepting fields , 2014). In the CBMs to 300 $y^{(m-1)}$ in (3) there are a standard by the second s bibio do î 301 időh őb**jatkoldálijati** 302 ender Feblikebleich (Kbhy 303 & Gel B ha. 2015). 304 3.2 Feature Importance 305 Maligen obe 306 307 igbigsbolthfaasthig (Chfa jaanst jaaith sidaatiin 308 jejst telefijesbytklijes 309 tabetet Ma factoristati 310 yteteh tetehieteh teeSKP tetene 311 is a state of the second secon 312 bish before bisting Bien 313 tela, 1984), þen Bipa 2001), dept(Che& Chpa 2016). In 314 kp elvépeltépeltén 315 biskowijelise 316 **Ny shipbi**te blb(nMa 2019), obtogbilg (SHP b efishbje(Light Le 2017). StARden phipper (pi 1953), throath konsineere teristets In befOSIAP by stillights SIAP het the California for ie**bielificitie** $g(\mathbf{z}) = \phi_0 + \sum_{i=1}^p \phi_i z_i,$ (5) $\mathbf{z} = (z_1 \ldots p)',$ ϕ_0 is a fix (e g) in; p isterije þ. 317 ϕ_i is the ϕ_i z_i satisfies *i* stead þ. 318 bffi i. Statista (5) bliste 319 ibbip baga jan ja Biba 320

abetesaistische Alle 324

321

- te (5) dependent (5) (2014), hv 325
- ið tilliggs SHARBSpillinsið 326

gaddaffefigap

- ingeleichtel 85 ingeleichteitigte 327
- 328
- ha Fan heffelden SPARen 329
- bebbel(inns, CBM\$). 330

- And SHARen inter beeling taken by 331 echeficte blbk B Msebathfac 332 **briges SARiad FA**(Lighta, 2019), 333 334 jabilitjav ig6KARkiefitja 335 biblydeffe bishefigisjert 336 chaffed fu aififHBh 337 n berphan et fan 338
- 338 **H** gripter Higher 339 **Harbe**(Ma 2019).
- 340 4 Results

Inisia eduphipific BMstala hajdeta (2021) et Cisi Meta (2021), es hittiget this is 2, eistap istatef fiest fisht bet ensign at hittiget CBMstright Mithia ensign temp(NES) effecting (6).

$$\mathbf{RE}(y,\hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$
(6)

- 341 HRE CHURCH
- 342 **B**SE**L**

diphNCE in skatablas(F\$) bl omelp(Mpi 1988) ighBtabijki∏ OBin alMc(2000) saballaba

$$\mathbf{FS} \quad y, \hat{y}, \ y_{\text{burton}}) = 1 - \frac{\mathbf{MS}(-y, \hat{y})}{\mathbf{MS}(-y, \ y_{\text{burton}})}, \tag{7}$$

 $y_{\rm burton}$ diplot tip ME(þ. 343 $(\hat{y}_i)^2$. all the table of table 344 ben seiseksikkelijä isaakset 345 sidijelisteride 346 jeje 15/4 Hist Heney (7) kebeji bibe 347 iskobbis IfFSiskon all (b)a, ba 348 shielifith Hy fFSsig h 349 ebieblésekbe 350

nRE chim

351

363

4.1 Comparison to existing methods

- In**isja optipligablera** 352 11111111111 (2021) (LSTM1/CNN1) dCh & 353 tela (2021) (LSTM2) del 7 tetride2 igleRE5 ten Chi 354 2021) isl-2 islip islie (2021) 355 juisil - ba Goldeta beta (2021) ististuel 356 the 1544 Historia the Children (2021) the 357 bhall H de CB Mible - 2 bekelv 358 tdeBA Historipatothighta 359 kish Advisefferiteit. Ŧ 360 RE kelikefetetetetete 361 kinhof to 9. Hab Chikk (2021), sheph 362

364	Folejja oficija stubilen			
365	ffeiRBsessigiftee5% igeb			
366	Ağlı sekelçi çelçi kaşalar			
367	jali berk ji kete n It ste gn			
368	ana pina tadah tingka tanga sanahan			
369	bRBstiffebilesty			
370	4.1.1 1 - H			
371	86 e17 kleRSE kerlikter 61 hille			
372	ista®A HiebenjageCBM, LSML, LSM2, ede			
373	işindin Inise, cCB Müddinendin			
374	REE ghilth 602B Mithell 448 ff $(5, 7\%)$ direct			
375	bL \$72 , al. 138 ff 13. 3%) bRE bL \$71 , b al. 942 ff 20	8%)		
376	terre high Fign to B Matterre			
377	fighted 514 td 517 tet (26 - 32, 35 - 38, 41, 42) . Fig 2 have the figure of the fig			
378	el deplínces Mell \$92 glines			
379	þíður H	< -300 f 31 ,	33,	37) betw
380	tigte Hillitisty			
381	bridetstan FederMa2001 te(netperfig 2), ta			
382	CB Masterigitin Hfrat 400 fillesv			
383	Hende: 00 to 1 2 00 Wende 31 to the state of			
384	Hahfihini M.			

Table 6. RMSEs for 1-hour ahead prediction over the test storm set with our GBM model, LSTM1 (Siciliano et al., 2021) and LSTM2 (Collado-Villaverde et al., 2021) neural networks, Burton equation (T. O'Brien & McPherron, 2000) and simple persistence. Here, the GBM, LSTM1, and LSTM2 were trained with past SYM-H and IMF parameters as inputs. The lowest RMSE for each row is shown in **bol d**.

6#	CBM L S72	L S71 B t	P		
26	5.863	6. 630	6. 700	6. 839	7. 63
27	7.729	8.913	8.900	7.954	9.623
28	4.281	5 858	5 400	5 697	5 814
29	5.833	6.683	7.200	6. 511	7.174
30	4.927	5 200	$5\ 600$	4.614	4 810
31	8.277	8. 584	$10 \ 700$	8. 838	10 429
32	6.841	7.259	8. 300	9.487	10 528
33	14.492	13.340	$16. \ 300$	$16.\ 630$	21. 167
34	$10 \ 190$	10.034	11. 300	10 888	$10 \ 913$
35	7.154	7.693	8. 500	7.918	8. 011
36	8.512	9.525	8.700	9.082	9.708
37	14.548	$15 \ 184$	17.500	$15 \ 713$	19.698
38	3.886	4 080	4 200	4.572	4 842
39	5 901	6. 431	5.600	6.663	7.597
40	4 976	4.673	5 500	$5 \ 371$	5 057
41	7.558	7.882	9.000	8. 358	9.984
42	5.030	$5\ 669$	5 900	5 549	6. 086
Ma	7.412	7.860	8. 550	8. 276	9.354
Ma	6.841	7.259	8. 300	7.918	8. 011
Mh	3.886	4 080	4 200	4 572	4 810
Mat	14.548	$15 \ 184$	17.500	16.630	21. 167
8 D	0 763 0	713 0	901 0	840	1. 131

Table 7. Forecast skill scores (using the Burton equation (T. O'Brien & McPherron, 2000) as the baseline) for 1-hour ahead prediction over the test storm set with our GBM model, LSTM1 (Siciliano et al., 2021) and LSTM2 (Collado-Villaverde et al., 2021) neural networks. Here, the GBM, LSTM1, and LSTM2 were trained with past SYM-H and IMF parameters as inputs. The highest skill score for each row is shown in **bol d**.

6#	CBM L S12	L SA	
26	0.143	0 081	0 020
27	0.028	- 0 120	- 0 119
28	0.249	- 0. 028	0 052
29	0.104	- 0. 026	- 0. 106
30	- 0 068	- 0. 127	- 0 214
31	0.063	0 029	- 0 211
32	0.279	$0 \ 235$	$0 \ 125$
33	0 129	0.198	0.020
34	$0 \ 064$	0.078	- 0 088
35	0.096	0.028	- 0. 074
36	0.063	- 0. 049	0.042
37	0.074	$0 \ 084$	- 0 114
38	0.150	$0 \ 108$	$0 \ 081$
39	$0 \ 114$	$0 \ 085$	0.160
40	0.074	0.130	- 0 024
41	0.096	0 057	- 0. 077
42	0.094	- 0 022	- 0 063

385 4.1.2 2 - http://www.sec.up/action/actio

 is all is all

392 4.1.3 **34**

393	kwitel Historic till - till cCBM
394	fab. SVI ebbefesCNN1. For O blance
395	61- bole 61 M L SAL, eiCNN1 ei2- bole
396	1566 BM 60 BM 16 Be 3. 5 T (15 4%) 16 BE 16 S1
397	deal. 6 50(7.7%) deal REE & CNN1. Fign & CB Midela
398	HARE 611 to 617 totsn Herr CNN1 kenter E
399	61b3 (gene (133, 37, 40).

4.1.4 **b**ja

400

401	ET 1 kitep kelipit ekkitesja	
402	füst 1. Folstpo errestelerifisin RE etsm	
403	ka Maliki dekiz	
404	fittea5% tijek (pkr	≤ 0 (05) page 2 heights
405	L \$\$2	

Table 8.RMSEs for 2-hour ahead prediction over the test storm set with our GBM model,
the LSTM2 neural network (Collado-Villaverde et al., 2021), Burton equation (T. O'Brien &
McPherron, 2000) and persistence. Here, the GBM and LSTM2 model were trained with past
SYM-H and IMF parameters as inputs. The lowest RMSE for each row is shown in **bold**.

6#	CBM L SΩ	Ba P		
26	8.285	8. 989	10 690	12 374
27	11.585	13. 418	$12 \ 465$	$15 \ 387$
28	5.650	5 877	8. 858	9.331
29	8.826	$9.\ 314$	9.776	11. 415
30	7.280	7.288	6.266	7.416
31	$12 \ 613$	12.436	13.604	$17.\ 193$
32	9.927	8.937	$13.\ 766$	15 282
33	24 519	18.481	25 729	33. 9 <i>2</i> 7
34	13.736	13.941	14 695	$15 \ 109$
35	9.504	9.932	10 586	11. 211
36	$12 \ 668$	12.058	$13.\ 117$	$14 \ 687$
37	22 327	21.084	24, 446	30 582
38	5.153	5 213	6.546	7.353
39	7.391	6.798	$10 \ 159$	$12 \ 322$
40	5 633	5.281	6. 032	$6.\ 373$
41	12 121	11.707	$12 \ 622$	$15 \ 437$
42	7.976	8. 273	8. 877	$10 \ 130$
Ma	10 858	10.530	12 249	14 443
Mel	9.504	9.314	$10 \ 690$	$12 \ 374$
Mh	5.153	5 213	6. 032	$6.\ 373$
Ma	24 519	21.0840	25 729	33.927
8 Ð	1. 310 1.	077 1.	338	1.808

406	4.2 Explaining predictions		
407	In kia optitasitut hs		
408	işiri kiri isa. 2 dili kiri		
409	hib3, aphlyth		
410	Fig8 Matsch1- Hylindea		
411	abdebY4 H bp) lepbY4 Hbojah boela		
412	SVI Hishstal- bida Sibeblyp		
413	islisty: Histophia		B_z , b
414	s est es ja i glyt tog		
415	stiligen telesteliky		V_x de
416	fidefiel E_s displayed SA He		
417	dv = -100)). Infa tilispäisisjap		ϕV_{x}^{2} .
418	Girije de tije s a		
419	tebapèkajint844 H Iaga èka		
420	$\delta \rho V_x^2$ el B_z epstjepstar H, potesn		
421	phan by schichephy Hel	B_z absph	
422	ta Dja ja daljende defideren		
423	by Histórefia kaa B_x, B_y eripping	ajan 🛛 👘	
424	b Nathfiel E_s is a splane in the second		B_z det
425	ija istigerijation		
426	ejeb.		

Table 9. Forecast skill scores (using the Burton equation (T. O'Brien & McPherron, 2000) as the baseline) for 2-hour ahead prediction over the test storm set with our GBM model and the LSTM2 neural network (Collado-Villaverde et al., 2021). Here, the GBM and LSTM2 model were trained with past SYM-H and IMF parameters as inputs. The highest skill score for each row is shown in **bol d**.

6 #	CBM L S22	
26	0.225	0 159
27	0.071	- 0. 076
28	0.362	$0 \ 337$
29	0.097	0.047
30	-0.162	- 0 163
31	0 073	0.086
32	0 279	0.351
33	0.047	0.282
34	0.065	0.051
35	0.102	$0 \ 062$
36	0 080	0.081
37	0 087	0.138
38	0.213	0 204
39	0 272	0.331
40	0 066	0.125
41	0.040	0.072
42	0.101	0 068

427	Fig4 kebőnéf teleteszág	
428	pað HAppel B_z betespel Nappel	
429	jalansigis operations and the second s	
430	H, et jeje state (viele	
431	ipetasp Y_1 H) . Heid E_s is a blue	
432	ter B_z . Et hjælt b	B_z
433	el V_x : B_z beneficiente finite for a finite field of the field o	
434	this the V_x periods V_x	-100
435	Freikepeffektygen Welstelle	
436	Es ikten enstitetsfien	
437	effekteneffekten istei	
438		
439	ilite Aliante	
440	SA H işciştiştirdili	
441	ighting Eb(Kath, 2017).	
442	4.2.1 BR 004 Br	
443	ANA A A A A A A A A A A A A A A A A A A	
444	Fig5 klauktivetat - b	
445	ei2- titipi1594 HighNe2004 igen	
446	ETEN Histo -400 TEN Historia de REC de	
447	301161-beal891162-beaute(p)w. He	
448	trikki i kysjej	
449	ben Finil 8:00 to 20 45 W (betel Chan SC) ,	
450	habba Hisip daisiga daisiga da da	
451	pl tightybe ba strap	

		1-		2 Ha	
6#	(BM	L SA	CNN1		ŒМ
26		12.6	18. 0	19. 8	12 9
27		20.1	16.8	23.4	20 9
28		12.7	18.6	144	$12 \ 4$
29		15.4	21.1	20 0	16. 7
30		17.0	24. 2	25.8	17.1
31		28.5	$32\ 5$	$32 \ 1$	29.6
32		21.8	23 . 4	18.9	21.9
33		35 7	33.8	26.7	38. 1
34		15.3	17.9	16. 6	$15 \ 5$
35		16.9	21.3	18.6	17.3
36		16.2	20 4	21.4	16.8
37		41.6	42 6	36.9	42 7
38		10.5	18. 6	13. 0	$10 \ 6$
39		13.0	20.3	16.5	12 8
40		$10 \ 9$	13.6	9.2	$10 \ 6$
41		23.2	27.3	25 4	23. 7
42		16.9	17.8	16.7	$17.\ 1$
Ma		19.3	22 8	20.9	19. 8
Ma		16.9	20.8	19.9	17.1
Ma		$10 \ 5$	13. 6	9.2	$10 \ 6$
Mat	4	41.6	42 6	36.9	42.7
8 Đ	2 284	1. 9	94 1.85	3	2 402

Table 10. RMSEs for 1- and 2-hour ahead predictions using only the IMF as input (I_2) with our GBM model and the LSTM1 and CNN1 models of Siciliano et al. (2021). For 1-hour ahead predictions, the lowest RMSE in each row is shown in **bol d**.



Figure 2. 1-hour ahead predictions for the 3 strongest geomagnetic storms in the test set during the main and recovery phases from our GBM (left column) and the LSTM2 developed by Collado-Villaverde et al. (2021) (right column). The observed SYM-H (black), the predicted SYM-H (blue) and the error (red) are shown for storms 31, 33, and 37 in the 3 rows, respectively.

```
isteisofäittisikkkyhelen
452
                                                                            tejobladiela alkaeli - bijbeela
 453
                                                                          S bliedebal bleisbebligsbeien
 454
                                                                          ikelephitetti 1 teEh. Kin
455
                                                                          istel - bizisiebisisie
456
                                                                            a abatabl MF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                B_z. Behated of
457
                                                                            shijsberchildisi th
458
                                                                          istate pieluktetation
459
                                                                          hidekteid
460
                                                                                                                                      Detaile(22:00 No7 t06:00 No8) filter to the Het
461
                                                                          hetsiken
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -400 TFightwhip
462
                                                                          bibbb
                                                                                                                                                                                                                                                                                                                                                                                       B_z \not\equiv 0 0 n N t 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d01 : 00 db
 463
                                                                          04 00 T Hispathik
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           B_z ships
 464
                                                                          - 50 filst affective and MF show
 465
                                                                                                                                                                                                                                                                                                                                                 titigetate (04 00 both
                                                                        bfm
                                                                                                                                                                                                                                                                                    B_z,
 466
                                                                          SA Higherfitzing and
 467
                                                                                                                                          the set level of the set level of the set of
                                                                          B_y,
 468
                                                                                                                                                                                                                                                            B_u isfitient to be a set of the set of t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 B_z is g
                                                                          izbici
469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B_y pattering this
                                                                          isteleten iste
470
```



Table 11.P-values from paired t-tests for null hypothesis that the mean difference in RMSEacross storms for GBM vs. competing methods is zero.

Figure 3. Scatter plot of percentage contributions (y-axis) against SYM-H (x-axis) for all the geomagnetic storms. The panels show the contributions of all considered features to the 1-hour ahead GBM prediction. Each prediction is represented as black dots. Kernel density estimates using a Gaussian kernel are shown in color with the corresponding color legend on the right of each scatter plot.

471	ins butfitiberit Fla			
472	identified Hit Ma Hitch			
472	title , tv	B_z		
474	Fe6 kk/fextfillip			
475	SV4 Histosi HARSE Mare 3 3 7763 7 7764 1 nd2- hep			
476	isl hø Fodel- biel RKE hækheð fr			
477	title - title RE state -			
478	sallifie2 tillSA Helikkelh			
479	effektstell i presidente ander and			
480	ka ikkazi ikk			
481	hap Andelsished - heistop			
482	kon Hielentige Stricture -			
483	hithithe A the second			
484	Juli			
485	Hile (b) istiplik			
486	jep In blaciji 18:00 to21:00 107 tababy H			
487	sip bitatilih 6504 Hegig			
488	B_z which		E_s , d	B_x (fol- hpp

Figure 4. Scatter plot of percentage contributions (y-axis) against SYM-H (x-axis) from solar wind and IMF parameters for 1-hour ahead prediction from GBM using only solar wind and IMF parameters as input. Each prediction is represented as black dots. Kernel density estimates using a Gaussian kernel are shown in color with the corresponding color legend on the right of each scatter plot.

```
à d
                                                                                   B_y (62-b)) bestelltigen (9)
489
                             SA Heby
                                                                                                                     -10017 Hom
                                                                                                                                                                                                                                                              B<sub>z</sub> bille
490
                             isteichtigte Hittistetten
491
                             61200 WN as by
                                                                                                                                                    B_z tajo Eadajo
                                                                                                                                                                                                                                                                                                                      B_z, by bin
492
                                                                                                                    B_z. Static
                                                                                                                                                                                                                                                                   E_s, typeth
                            bian
493
                            efilfip
                                                                                                                                     B_z, is shift in
494
                                                    4.2.2 January 2004 Storm
495
                                                   Nex esphere J ye 2004 block MA H fiber
496
                             - 1407 r stiszista taszek Ne2004 para
497
                                                                                                                                                                                                                                                                                                                                                                                    B_z fild
                             Astrifie7, sisasterbeldester
498
                             ibCME b( 00 00 U611:00 U a22) jabidada
499
                                                                                                                        Bz. Height 4. 22 Heil 9. 96 HRSE 6
500
                             蜔
                             tell- et2- tejsi ja tistegefitsetjen
501
                             ta Intel CME ta tentistensi Helleta
502
                             ġþ.
503
                                                    81 - the the state of the state
504
                             2000 listintipits stats
505
                             cipSA Haisial be etisiyabilayi
506
                             inpostilized 82- tipi tor
507
                             çe þisjáls A Har (stíghin
508
                             ected SA Hitze 2 telet, spottere
509
                                                    B col: 00 cll 1: 00 Weitschrift Hate
510
                             iden hv
                                                                                                                                  B_z ignicity A$11:00 \mathbb{T} ign
                                                                                                                                                                                                                                                                                                                                                                                B_z ta
511
                             hal bibliblib
512
                             fetsin bro2 kebfetskin jø 812-bro
513
                                                                                                                                                       B_y bit 0 all 2 00 \mathbb{T} Applica
                             idalahan
514
                             teteti di teteta setteti
                                                                                                                                                                                                                                                                                                                                                                   B_x o B_y
515
                                                                                                                                                                                                                                                     B_z biblic
                             yabaqabaqa jiying
516
                             þn
517
```

5 Discussion and conclusions 558 **Additional Additional Addition** 559 idSA Har bidy IMF, didential teleSA 560 Hiell to2 kbal. Impa jajogen (GBM) end 561 alapisedeler Reh delates 562 jelekten fote fotblekte 563 Update Synthesista Hist 564 Fablicitis anticistary 565 iliginging ingen 566 567 bidebénéEbasigen Impa enelepéN4 568 Ηđ B_z etspektelinget 569 ignitsipalisti Digitseljan 570 p\$14 H, pial by alterna signatefield 571 biilitija ASA Heidija 572 \$854 Hel 573 SHOW TO BE SERVICE 574 sidite jähkeeletetete 575 B_u here be **deetse**ten 576 h **didatti**ð B_z state 577 tid bid E_s . Here shapping 578 k (B ka 2018) yffedd SiAkshipi Alpa 579 the states and the second 580 ababba Fapin fatig 3 ald, seebbba 581 V_x iskap
SA Historip kisk ₫afn 582 ebbeeks Hel V_x . (F at the state of the 583 **tişbiri tehişbiştişi**en 584 tight teach Fastistiach 585 tightfielightfrieffe 586 **bfyffet**n 587 Adag ka skilleristeiler 588 **histobelighei**gen 589 isja Hjev **ipiski (1478**) ostaja 590 haite bisten stadtstatet. 591 enticidade (m. 1811). Si state (m. 1911). 592 heite Initia 593 fijeljiskljepin Clyr elvefa 594 fist**ige fig**n 595 596

- 597 Appendix A
- 598

612

A1 Graphical comparison with persistence model & Burton equation

599	FigA klad thiston (BM (tap84 H				
600	ell NF perpe ethylighteda				
601	spillestan H	<	300 H 31,	33,	37) b
602	brighte Affect photos				
603	(B) Malifikia (B)				
604	iel Folpen igbingefes 37, etgielen				
605	idela > 100 filteri idel¥i Hel				
606	ightip Mar CB Mightad 100 o				
607	40 fr brittightight H Fig.42				
608	klal blektor (B Mele B trige).				
609	singi Ining ta Menaphighten				
610	heite ta the heiter the states and t				
611	endersetetetetetetetetetetetetetetetetetete				

A2 Descriptive statistics of solar wind & IMF parameters

Table A1.Descriptive statistics for the solar wind and IMF parameters in the 25 storms usedfor training listed in table 1. The minimum temperature (MK) is most likely a measurementerror.

ħ	Ma 25% bQ Mal 75%	76 bij Mar			
B _x ()	- 43. 700	- 3. 131	0 340	3. 378	34 681
B _v ()	- 51. 968	- 2 901	0 221	3. 289	46.862
B₂ ()∏	- 77. 258	- 2 296	- 0 092	$2 \ 179$	38. 717
V _x (km)	- 1233. 693	- 539. 489	- 445 287	- 384 021	- 264 722
Dig (phon	²) 0.041	2 912	5 027	8. 477	76. 239
ff(MK)	0 0082 0 00885	0 07 02	$0 \ 1262$	1. @83	

Table A2.Descriptive statistics for the solar wind and IMF parameters in the 25 test stormslisted in table 2. The minimum temperature (MK) is most likely a measurement error.

ħ	Ma 25	%bQ Mal 75	‰bQ Mar			
B _x ())	-	4 8. 717	- 2 868	0.221	3. 444	33. 827
Ву ())Г	-	48.963	- 2 816	- 0.205	2 855	54 563
B₂()∑	-	48. 585	- 2 357	- 0 084	1. 933	53.02
V _x (phys	- 8	87.784	-535 138	- 424. 304	- 373. 465	- 251. 481
Dja(pánm	3)	0 295	$2 \ 760$	4 424	7.643	113. 982
ff MK)	0 0052	$0 \ 087$	0 0658	$0 \ 122$	0.9909	

- 613 Acknowledgments
- 614 ANASB P. SQAB and Mal and De
- 615 6ChappelEg(CL) station
- 616 **CERTIS BASENNAS DRY**
- 617 ECHESONS20K1600, NAMMERONS19K064, NF R-
- 618 EXENSEL663800 aNS KgHY 2027 555. EC shall

1/NASche 80NS21K1 580 a80NS21K1 27 5 HACE 1/2 al 619 620 https:// BEDAFFA S(BF) CHEDAAHWCDAWA 621 cdaweb.gsfc.nasa.gov/ . HSA Hitting F' sON-622 https://omniweb.gsfc.nasa.gov/ ₩ . Attended 623 624 DeB HDaff https://deepblue.lib.umich.edu/data), **bis**fo 625 hig 6 Mist bibliothic sig 626 the Belver Defender (DC) , by 627 testistist bittis 628 References 629 ABI SQ (2020). suf py **B**n https://gitlab.umich.edu/swmf 630 _software/swmfpy (is 2020 5) 631 J. R, WH-L., Bp RJ., WSN., & Bp SA 632 (2016). Meterinia Eta 633 K p is a NANG MOEL INGAND e Aeto 634 $S \ pace We \ ath, e1r \ 410$, 899916. RDICINO K p INDEX d 635 $10\ 1002/2016$ NM 463 636 Bkg, R.L., KR, M.A., Ag, C.N., My, C., Ko, M.J., Afon 637 UV, ... He J. (2021). **Heit**in 638 **dibili** arX i v 2 0 0 6.1 28 3 5 [astro-ph, 639 ph si c s: ph si c s640 Bpa E., & Kapa R (1999). AEb0infe/Clie 641 έαλξenΒig, Βig, α€V Machine Learni, $\mathfrak{B}q (61/2)$, 642 105±39. ii 10 1023/A1007515423169 643 (2019). BhaA, & NG FebS/Hel/Slieb 644 **jenfibl**y24 **je**S ∎P syd 2015 ten 645 Journal of Space Weather and Space Q.g.mate ie NANdav 646 **d** 10 1051/**s**/2019007 647 Bby K., Bban M., Jbp A., Kag M., Kapi S., & Ban 648 O (2020). CitChy intCheFinCisP 649 S S RNJ our nal. in EinfaMeLinAp d 650 10 2139/s 352659 651 B by J. E. (2018). Objective device 652 Ð Journal of Geophysical Research: Space, Physical 3, 2029. d 653 10 1002/2017JA024650 654 Bha J., Cha X, Ma Q, Li, W, Ka X, Ka R M, ... ha (2018). 655 AGHDAGH In Ma-656 $chine \ learning \ techniques for \ space \ (eqna t <math>270300)$. Ebr **d** 10 657 . 1016/B978-0 12 811788-0 00011-1 658 Boga R, Boha M, Boha S, Fan A, & Anan O (2011). Dat 659 identation In Anales q e oph spic $ae(\mathbb{W} 29, \mathbb{p} 965971)$. 660 Mach Lear, n24(2), 123440 d 10107/ Bipan L. (1996). Bippi 661 BF0058655 662 Machine Learni, 4g (51), 332 id 10 1023/ Binn L. (2001). Hone 663 A1010933404324 664 Bipin L., Fiph J. H., KD R.A., & K. C. J. (1984). C assi f c ati on665 and regression tree BB 666 B ta R K , MP R L ., & KR C .T (1975). Appl 667 Journal of Geophysical Research **djish**Ds 668 8 Q 4204 d 10 1029 / JA980B104204 669 Ca. L., Ma. S.Y., & Z.Y.L. (2010). H1844 Hidda 670 stoNANDSOMF abide Anal es Geo-671

physicae 28(2), 381393. d 10 5194/g 28-381-2010 672 Chan E. (2019). Heitheringher New 673 $S \ pace \ We \ ath, e1r \ (8), 1166+207.$ **d** 10 1029/ ġŧ 674 2018802061 675 Chan E., Cha M. D., fog H. J., Bha C. C., Hag, Z, & F. G. 676 AGt B «MebaleEboleCel (2020). 677 MAR ENANDAM 678 J. Geophys. Res. Space Physi, ds2(511). MAL ja d 679 $10\ 1029/2019$ J A27 684 680 Obl. M., Open E., & MANS (2017). Height 681 detected Address 682 S pace We ath, et [8], 1004-019. id 10 1002/2017 SNI 627 683 Ора Т, & Ора С. (2016). ØВ 16 АБАВВ јербл Pro-684 ceedings of the 22 nd AMSI GKDDI nternational Conference on Knowedge 685 Discoveryand Data Mining, 785794 di 10 1145/2939672 2939785 686 Chiel A, Marí P, & Cal C. (2021). DeNeNandaV 687 Cheel SYLL set SYL Hel/SYHFe S pace688 We ath e r id 10 1029 / 2021 8027 48 689 Gebien Aidem Fach J. H (2001). 690 An. $S tatist., 29(5), 1189\pm232$ id 10 1214/(d) 1013203451b 691 Chan, Jun A, & Lhen M (2017). **ENT** 692 d d d C d P S pace S c i e nc e Re, i 2 lo (23-4), 1315-693 1344 id 10 1007/\$1214 017-0412 2 694 Сран, Lpd, н, & Mar (1996). **∄g**n ∎ 695 An. Geophys., 147, **fib idde**iniddau 696 $679686. \quad \text{id} \ 10 \ \ 1007 / \texttt{505}85 \ \ 996\text{--} \ \ 0679\text{--} \ 1$ 697 GaY, WH-L., Bop RJ., WSN., & Bha MA (2019). (m) 698 ItianDa DiaFanta India 699 takter NeCel NASA Journal of Geophysical 700 Research: Space Physice 23(41), 248263. id 10 1029/2018 JA22957 701 Ha, T., Har R., & Fadah J. (2001). The elements of statistical learn-702 ing Noby NY BU DNoby n 703 Iogn T (1990). Snitgsffiihdd 704 Journal of geomagnetismand geoelectricity jefi iba 705 4 \$\mathcal{2}11), 1249+265 706 Kapa E. K. J., Baja A., Kapja R., & Liu, Y.D. (2017). Gefe 707 **BUBBLED** S pace S c i. 708 Re v, 21, 2, 1271 - 1314 id 10, 1007/ ± 1214 017 - 041 - 3 709 Kag J. H (2005). **Eidebi ciptiki**NV 710 Journal of Geophysical Research 711 dACE bebliebel 1 1 (02), 2104 **d** 10 1029 / 2004 JA010649 712 KHANY R, & CALB ha R (2015). D4RD per pha 713 InG L ho& S V N. ht/Ed), idese 714 Proceedings of the eighteenth international conference on artificial intelligence 715 and statistic (NV 38, p 489497). SDg Chi SURLR 716 Kha M, Cha U, Ban V, & Cha O (2020). Galato 717 HEREFERE 718 *I EEE & c e ss 8*, 187814-87823. fibler **d** 10 1109/ 719 ACCES 2020 3081477 720 Liu L., ZZ S, AYY, & AVZ (2020). Fieldalla 721 TEC HEL into S pace We ath, e1r & 11). d 722 10 1029/2020000000 723 La Y, Caa R, Caa J., & Has G (2013). And the design of the second secon 724 Proceedings of the 19 th AMSI GKDD interhiaia In 725 national conference on Knowledge discovery and data m'(pi 6q2631). 726

Child ACM id 10 1145/2487575 2487579 727 Lu J., PY, WM, Gu S, & AM (2016). **fitte**n 728 **idailie**ftelijein 729 **tite**n Planetaryand Space Sci,elnc2e,0 4855. d 730 10 1016/j p 2015 11. 004 731 Ligh SM, Ein, GG, & Le, S-I. (2019). ChildFe 732 arX i v 1 8 0 2.0 3 8 8.8 [c s, stat] **taktéEle**n 733 LH SM, & Le S-I. (2017). AfidatoIteMa 734 **P** In **Alu** nces in Neural I of ormation Processing System (\mathbf{M} 30). 735 Calle In 736 Lbj S M , Na B ., by M S , Ho M , Ea; M J ., And T , 737 ... Le S-I. (2018). Epidenajaja filoso 738 Nat Bone d Eng , 2(10) , 749760 id 739 10 1088/41551-018-0804-0 740 Ltd H, & N P (1994). 741 Bitto An. Geophys., $1 \ (21)$, 1924 d 742 **ideldes**fe**tte**v 10 1007/80585-994-0019-2 743 Mgt P N. (1980). Head In Derivition, meaning, and use of geomag-744 ne ti c i nd i (p s 115 129). Acquir AQU. d 10 1002/ 745 9781118663837. **B** 746 MChan D. J., Bean S. J., Behar P, Febha W.C., Par J. L., Par 747 P, & Offe J. W (1998). **6NF/66**/ 748 Space Science Re, i e so (SNAM) field CipEles 749 8 6 563-612 d 10 1023/A1005040232597 750 Mar G, & Las H (2021). AI 751 5 660 60 B is 752 arX i v 2 1 0 8 . 0 4 0 58 [cs, stat] 753 Mar K. E., Høg B. P, & Base A. (2019). Infibie 754 ELTER In Proceedings of the 29 th annual international confer-755 ence on computer science and software engine (eprilog 6472). SI IBM 756 Cp 757 Ma C. (2019). Interpretable machine leader tops://christophm.github 758 .io /interpretable-ml)bo o k/ 759 Ma. C., Cha, G., & Bha, B. (2020). Iph/MaLig= 760 arX i v 2 0 1 0 . 0 9 3 3.7 [cs, stat] AB ÉHS ÉSÉ LA ANDA 761 Mar A H (1988). Self true obt 762 **jat habilita** Month ly Wé ath $e r \operatorname{Re} \dot{v} e \psi 1 1 (612)$, 763 2417 - 2424 **B**n https://journals.ametsoc.org/view/ 764 journals/mwre/116/12/1520-0493988_116_2417_ssbot2n0_co2.xml 765 **d** 10 1175/1520 0493(1988) 116 **h**2417: **B Q** i2 0 0**2** 766 Nha A, & Kha A (2013). Gübben abu Front. 767 Ne worob ot, 7. d 10 3389/b 2013. 0021 768 New PT, 65 T, Lia K, Mg, C. I., & R. F. J. (2007). Ah 769 bisk bifffin 0 en 770 itti k Journal of Geophysical Research, 2 01206. d 771 $10\ 1029/2006$ J AQI 2015 772 OBie T, & MPRL. (2000). Fieldetaikin 773 Journal of tAnospheric and Solar-Terrestrial Phose (cls4), 1295haim 774 https://www.sciencedirect.com/science/article/ 1299.**B**n 775 pii/S136468260000729 (polywy) d ps//d g 10 776 . 1016/\$364 6826(0) 0072 9 777 OB in TP (2002). EditorDation Journal of 778 Ge oph spi c al Re se a; cth 0 (7A1), 1341. d 10 1029 / 2021 A009 435 779 OB in TP, & MPR RL. (200). Athen 780 stietish Biddfindd 781

Journal of Geophysical Research: Space, Hh ga (54), 77077719. d 782 10.1029/1998JA00437 783 βR G, Apan E., Cha, G, Ma, M. F., & Bho, I. (2006). Ge 784 ja Disisten MF taka An. Ge oph sg. , 24(3), 785 989999. id 10 5194/ g 24 989- 206 786 **飛**J., & 好 O (2018). Ne ergrad - Agradient-free optimiztion platform 787 https://GitHub.com/FacebookResearch/NeGHorgrad 788 R. C. (2019). Constitution to the second sec 789 Nature Machine Intelligence 790 1(5), 26215 **d** 10 1088/**s**226-019-008-x 791 I nte rpre ti ng rand omf ore sts -Di i ng i nto da tith A (2014). 792 https://blog.datadive.net/interpreting-random-forests/ 793 БаК, Nan T, R, G, D., & Б. Н. Е. (2015). ₽ 794 795 Space Weath, elr \$12), 853867. B221-08-05, **ji**h 796 https://onlinelibrary.wiley.com/doi/10.1002/201550001254 797 **f**n 10 1002/2015501254 798 βa L. S (1953). Abbén kRûn ImH WKbo&A WkEΓ 799 (Est), Contributions to the Theory of Games (M4-28), Volume I3D7-800 318). 🗰 d 10 1515/9781400881970 018 801 fev7ZR, & Apan A. (2021). HDa: DeLigis NoAddY 802 Nel arX i v 2 1 0 6.0 3 2 53 [c s] 803 þã F., Cþa G., ø∕T R., Cþa M., Cþa F., & DeMþa 804 Ρ Fight Hist ACipB al g6 (2021). 805 S pace We ath, e r (2). id filige Clarks Nev 806 $10\ 1029/202089$ 807 £m C. W, L'Ha, J., №, Ν. F., anδũ M. H., Bha, L. F., & F. 808 J. (1998). HACE Marthetian Space Science Re, i & to 809 613-632 d 10 1023/A100092216668 810 🖆 N., Ha, G., Kajar A., 🛱 I., & 🛱 R 811 Dp Appropriately (2014). 812 Journal of Machine Learning Reset no 56), 1929+958. 813 Ég R, Chan J., Kha, S, Ha6h, M, Anal H, Fha, J., 814 ... For O (2020). Me te orol og ydri en ariabilityof air pollution (PMI) 815 revaled in the polainable machine le Perning AN FEMA 816 tadar bioto . d 817 10 5194/p 2020 469 818 **§** M (1964). **(**1. In Hourly also of equatorial dst for the $\mathbf{R}g y$ 819 ₽ 820 & Kian T (1991). Elaphil 957+1986, japh, 40 821 **A** M, b yA Brthelier and M. Menville (Int. Serv Geomagn. Indices Publ. Off., Saint 822 Maur, 1991) 823 F.N. (1989). Applietbergh 824 Pl and taryand S pace $S c i e \Re(e^{1})$, 520 \mathbf{d} 10 1016/0082 taba 825 -0633(89)90066-4826 N.A. (1995). MeleEla seperiel 827 Journal of Geophysical Research 6 dataten 828 $1 0(0^{\text{A}})$, 5599. **id** 10 1029/94JAB193 829 FIT N. A. (2002). Additionships the 830 spin. Min ANEW ACNERER MAG 831 NETIC FIELD MOEL, 1. Journal of Geophysical Research: Space, Physics 832 $1 \ 0 ($ **28**), **SP** 24**SP** 245 **id** 10 1029 / 2001 J A0021 9 833 N.A. (2002)b. Additional 834 Ha Philab ANEW 835 MACHERER MACHERIC FIELDMOEL, 2 Journal of Geophy-836

837	ical Research: Space Physics107(A8), SMP 10{1{SMP 10{17. Retrieved
838	2021-08-05, fromhttp://doi.wiley.com/10.1029/2001JA000220 doi:
839	10.1029/2001JA000220
840	Wanliss, J. A., & Showalter, K. M. (2006). High-resolution global storm index
841	Dst versus sym-h. Journal of Geophysical Research: Space Physics 11(A2).
842	Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/
843	10.1029/2005JA011034 doi: https://doi.org/10.1029/2005JA011034
844	Wu, JG., & Lundstedt, H. (1997). Geomagnetic storm predictions from solar wind
845	data with the use of dynamic neural networks. J. Geophys. Res, 102(A7),
846	14255{14268. doi: 10.1029/97JA00975
847	Xu, S. B., Huang, S. Y., Yuan, Z. G., Deng, X. H., & Jiang, K. (2020). Prediction of
848	the Dst Index with Bagging Ensemble-learning Algorithm. ApJS, 248(1), 14
849	doi: 10.3847/1538-4365/ab880e

Zhang, T., & Yu, B. (2005). Boosting with early stopping: Convergence and consistency. Ann. Statist., 33(4). doi: 10.1214/009053605000000255

manuscript submitted to Space Weather

Figure 5. 1-hour (left) and 2-hour (right) ahead predictions for the Nov. 2004 storm using GBM trained on all considered features. The rst row shows the observed (black) and predicted (blue) SYM-H values. Rows 2-9 show the contributions from each feature (left axis, colored) and its value (right axis, black). The percentage contributions are shown in the last row. The contribution from past SYM-H on predictions is omitted, but its percentage contribution is implicitly shown as the remaining white area in the last row.