

# Investigating the Associations among Active Regions, Flares and CMEs using Machine Learning

R. Qahwaji, T. Colak and M. Al-Omari and S. Ipson  
University of Bradford

In this work, machine learning algorithms are applied to explore the characteristics of some active regions that are associated with significant flares and CMEs. The NGDC flares and sunspots catalogues and the SOHO/LASCO CMEs catalogue are processed automatically to extract the active regions that are closely associated with Flares and CMEs and to represent the characteristics of their sunspots in numerical training vectors. The sunspots characteristics (features) include numerical representations for the McIntosh and Mt Wilson classifications, timing information, etc. These training vectors are fed later to machine learning algorithms to extract the embedded knowledge and provide learning rules that can be used for creating reliable CMEs prediction models.

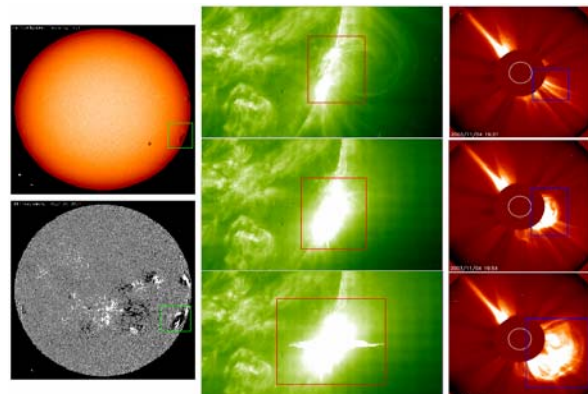
NGDC Sunspots Catalogue											
11031104 0812 S21W77 A	10486	RFX	5	121020	031025.5	031029.2	236	15VTC			
11031104 0812 M09W52 A	10487	RXX	1	3	30	031031.4	031031.4	237	CVTC		
11031104 0812 M09W5 A	10486	RXX	1	8	120	031025.1	031025.7	243	SVTC		
11031104 0812 S23W7 E	10495	DSO	4	170	0311	1.5	0311	1.3	244	SVTC	
11031104 1525 S18W6 BGC	10486	RFX	12	13	140	031025.2	031029.2	325	BCGL		
11031104 1525 M11W58 E	10487	CSC	2	3	10	031031.7	031031.4	328	BCGL		
11031104 1525 S22W5 E	10495	DSO	10	10	90	0311	1.2	0311	1.3	334	BCGL
11031104 1525 M10W42 E	10495	NRXC	3	6	10	0311	1.5	0311	1.8	340	BCGL
11031104 1525 M12E30 A	10495	DAXX	1	1		0311	6.9	0311	6.9	339	BCGL
11031105 0057 M11W61 A	10487	RXX	7	3	20	031031.4	031031.4	868	BCGL		

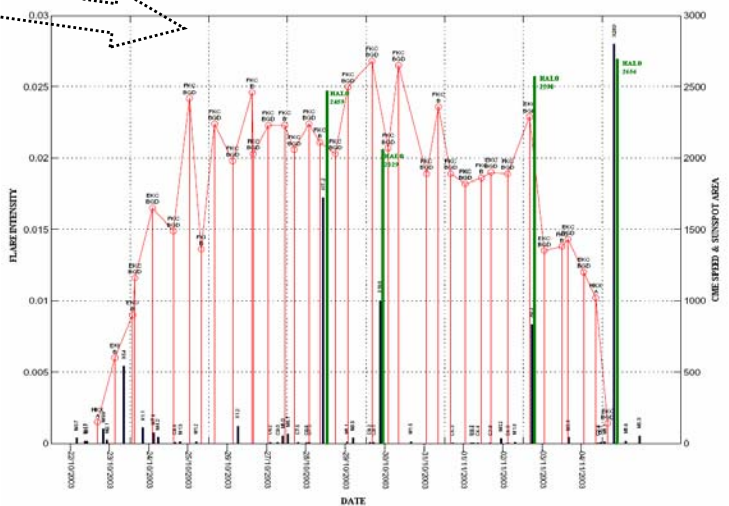
NGDC Flares Catalogue									
31777031104 0840 0807 0556	M 26	GOES	2.5E-02						
31777031104 0840 0950 0945	C 28	GOES	1.6E-03 1048E						
31777031104 1011 1033 1022	M 30	GOES	2.7E-02 1048E						
31777031104 1115 1125 1115	C 57	GOES	2.6E-03 1048E						
31777031104 1342 1401 1345	M 11	GOES	9.3E-03 1048E						
31777031104 1325 2006 1350 S19W33B	X 8.0	GOES	0.280E 1048E						
31777031105 0237 0245 0241 S19W92B	M 15	GOES	0.2E-03 1048E						
31777031105 0754 0803 0755	C 47	GOES	2.1E-03						
31777031105 1046 1056 1052 S10W95CF	M 53	GOES	1.7E-02 1048E						
31777031105 1629 1705 1645	C 19	GOES	1.8E-03						

NGDC CMEs Catalogue										
2002/A1/03 10:06:05	293	103	1420	1474	1366	1375	-11.6°	1.3e+16	1.3e+32	301
2002/A1/03 15:21:43	109	26	641	560	722	948	-25.3°	2.0e+14	4.2e+29	109
2002/A1/04 12:06:06	NA1c	360	1208	1467	926	1090	-41.2°	1.1e+16*	8.3e+31*	84
2002/A1/04 14:54:05	263	72	605	185	1027	1038	44.0°	2.3e+15	4.2e+30	261
2002/A1/04 19:32:42	197	82	287	471	170		-146.2°	3.2e+15*	1.3e+30*	187
2002/A1/04 15:54:05	NA1c	360	2057	2031	3284	3731	434.8°	1.7e+16*	6.1e+32*	260
2002/A1/05 14:54:05	258	12	1075	1117	1038	804	-26.3°	3.0e+14*	1.7e+30*	256
2002/A1/06 03:54:05	127	12	536					7.9e+13*	1.1e+25*	133
2002/A1/06 04:06:05	247	23	643	799	490		-74.6°	3.9e+14	8.0e+29	252
2002/A1/06 04:30:05	310	21	301	323	281		-5.0°	3.8e+13*	1.7e+28*	311



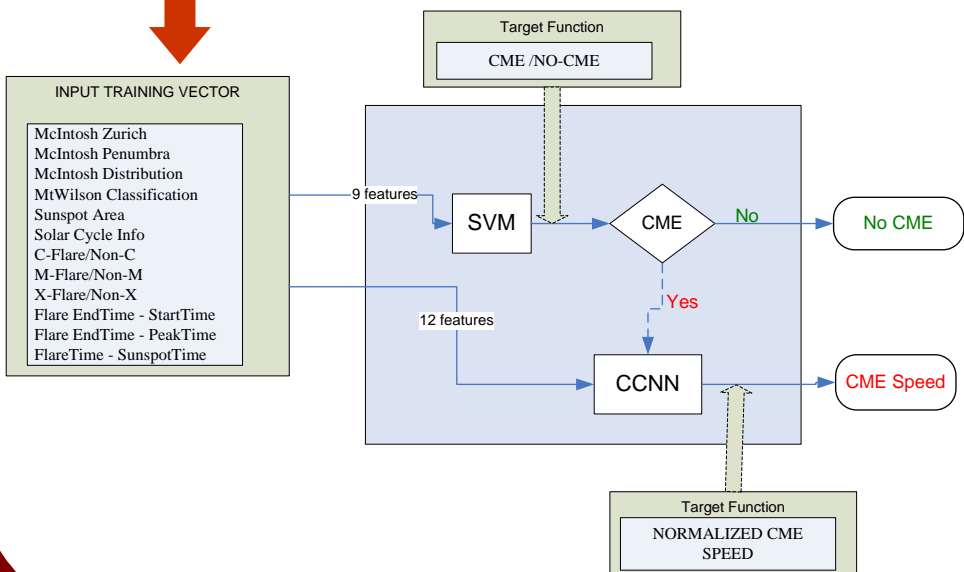
AR 10486 X-28 FLARE CME  
The Evolution of AR10486 and its Associated Flares and CMEs



## ASSOCIATION ALGORITHM

A C++ platform is created to automatically associate CMEs in the SOHO/LASCO CMEs catalogue with flares in the NGDC X-ray flares catalogue and the solar flares with sunspot groups from NGDC sunspot catalogues. The association is determined based on their timing information; the date and time for every CME is compared with date and time for every flare. And NOAA number and time of every flare is compared with sunspot groups [Qahwaji&Colak (2006), Solar Physics].

- Two criteria are used for comparison between:
- If there is not a CME recorded "α" minutes before or after a flare reach its peak time, then this flare is marked as not associated (NA) otherwise it is marked as possibly associated (PA).
  - If there is a CME recorded "β" minutes after a PA flare reaches its peak then this flare is marked as an associated (A) flare, other wise it is marked as NA. After finding all the associations, a numerical dataset is created for the machine learning algorithms using NA and A flares.



## PREDICTION RESULTS

	SVM CME PREDICTION	CCNN SPEED PREDICTION
1	64.4%	73.3%
2	64.6%	71.4%
3	63.3%	78.0%
4	63.1%	75.6%
5	64.3%	71.9%
6	67.1%	72.4%
7	64.5%	72.9%
8	64.8%	74.3%
9	63.1%	74.8%
10	64.4%	74.6%
Average	64.4%	73.9%