British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Session: 5.1
Date: Monday 26 August

YUCATÁN MEXICO 2013 IAGA meeting

Correcting observatory data prior to scientific analyses

SUSAN MACMILLAN (smac@bgs.ac.uk), Ewan Dawson, Thomas Humphries, Sarah Reay, Jane Exton, and Nils Olsen*
British Geological Survey, West Mains Road, Edinburgh EH9 3LA, United Kingdom
*** DTU Space, Technical University of Denmark**

Summary

Using geomagnetic data submitted to the World Data Centre (WDC) for Geomagnetism in Edinburgh (part of the ICSU World Data System) we have undertaken two campaigns to identify potential data problems which may have an impact on the outcome of subsequent scientific analyses.

The first campaign was motivated by studies of magnetic storms using extreme value statistics. In these studies any large spikes in the minute mean values can significantly alter the results.

The second campaign was motivated by secular variation studies and also the requirement to produce high quality observatory datasets to complement the data from recent (Ørsted and CHAMP) and forthcoming (Swarm) satellite missions.

In some cases, where the data are indisputably incorrect, the files available from the WDC (**www.wdc.bgs.ac.uk**) have been updated and the corrections noted in the metadata. All original submitted datasets are retained offline, reflecting the WDC's role as a long-term repository as well as a place where quality-controlled data are made available. Reprocessed datasets from institutes operating the observatories should still be submitted to the WDC.

Identification of large spikes in minute means

- For each year plot the maximum rate of change of total intensity (nT/minute)
- Identify outliers using normal pattern of high values in auroral zones and low values elsewhere and high level of spatial coherence (map below)
- Get date from file of maximum values ordered by observatory latitude (Table 1 and graph)
- Inspect data using INTERMAGNET data viewer (below)
- Assign bad datum as missing

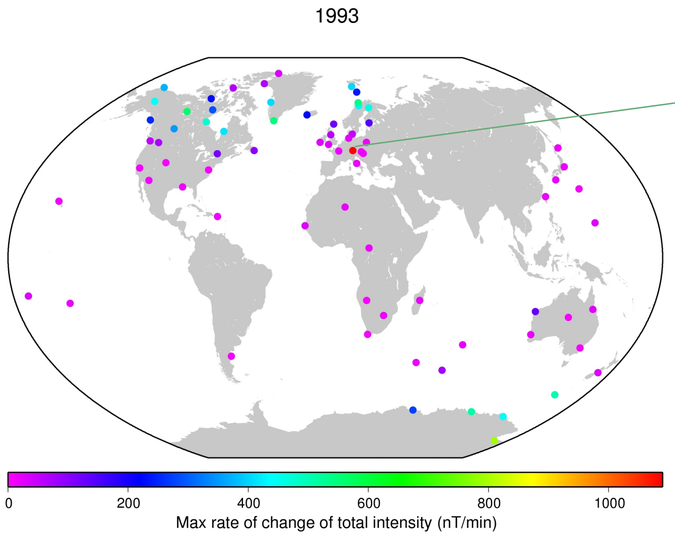
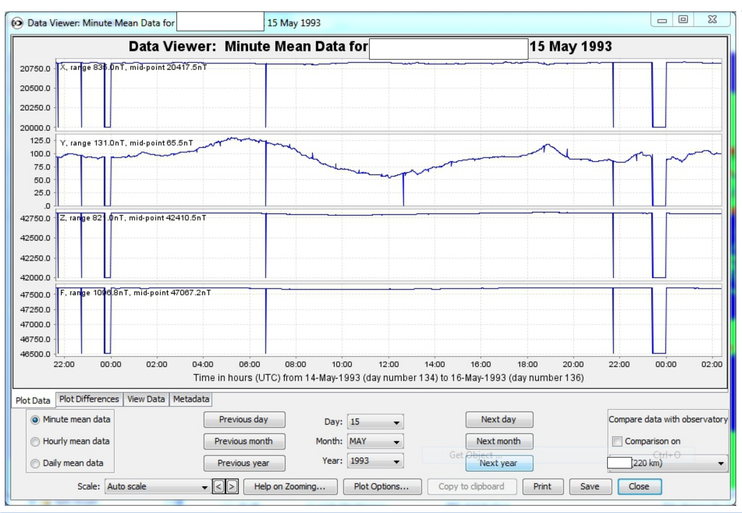


Table 1

LAT	LONG	MAX dF/dT	MM	DD (in 1993)
82.500	297.650	33.3	10	25
77.483	290.833	67.7	11	4
77.000	15.550	402.0	3	3
74.683	265.100	66.0	1	10
74.500	19.200	238.1	11	6
71.300	203.383	371.4	12	3
69.667	18.950	536.1	12	1
69.250	306.467	405.4	11	4
69.117	254.967	226.6	3	15
68.350	18.817	427.4	12	1
67.367	26.633	456.8	3	11
64.867	212.167	438.2	8	27
64.333	263.967	286.6	1	11
64.183	338.300	225.6	4	5
62.483	245.517	562.1	10	25
61.167	314.567	546.5	5	10
60.500	24.650	147.3	4	4
60.133	358.817	123.4	9	13
58.783	265.917	477.2	11	6
57.067	224.667	251.1	4	5
55.633	11.667	47.3	4	4
55.317	356.800	61.6	3	9
55.283	282.250	413.3	9	13
54.617	246.650	350.1	5	10
53.750	9.067	25.3	3	11
51.933	349.750	26.9	4	4
51.833	20.800	14.0	3	9
51.000	355.517	23.7	3	8
48.517	236.583	55.7	9	13
48.267	242.867	74.0	3	9
48.167	11.283	1089.6	5	15
48.017	2.267	13.0	3	8
47.633	16.717	13.7	3	8
47.600	307.317	98.1	4	5
46.900	17.900	13.8	3	8
45.400	284.450	118.3	12	3
43.917	144.183	13.6	6	10
42.383	13.317	15.6	3	8
40.133	254.767	10.3	9	13
38.217	282.633	9.3	6	10
37.083	240.283	10.4	11	18
36.233	140.183	15.2	3	8
32.250	249.167	9.6	2	17
31.417	130.883	16.3	3	8
30.350	270.367	12.9	3	23
27.100	142.183	10.2	4	4
25.000	121.167	16.3	2	17
22.800	5.533	19.8	3	8
21.317	202.000	11.8	3	23
18.100	293.850	10.4	3	8
14.400	343.050	26.8	2	17
13.583	144.867	26.6	2	17
4.317	18.567	15.4	3	8
-13.800	188.233	31.0	1	20
-17.567	210.433	18.4	3	8
-18.917	47.550	20.2	3	8
-19.200	17.583	7.4	12	8
-20.083	146.267	26.7	11	4
-22.217	114.100	132.0	8	18
-23.767	133.883	9.4	3	8
-25.883	27.700	17.4	12	8
-31.783	115.950	17.9	4	18
-34.417	19.233	8.0	11	18
-35.317	149.367	7.2	4	4
-37.833	77.567	22.8	3	9
-43.267	294.617	7.3	12	6
-43.417	172.350	28.6	9	18
-46.433	51.867	14.6	2	17
-49.350	70.250	78.3	3	8
-54.500	158.950	504.3	3	11
-66.283	110.533	513.8	4	21
-66.667	140.017	440.7	6	24
-67.600	62.883	273.0	2	11
-77.850	166.783	793.5	2	17



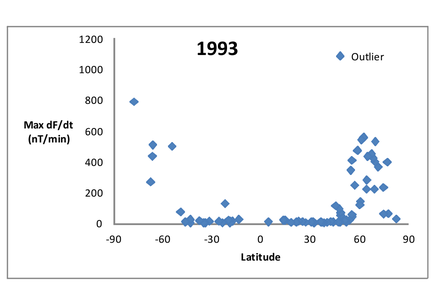


Table 2

YEAR	CODE	NUMBER SAMPLES FLAGGED
1977	*BRW*	*4*
1978	OTT	15
1978	SIT	11
1980	GDH	4
1984	BOU	17
1984	CMO	2
1985	THL	3
1989	LER	3
1990	KAK	1
1991	CBI	1
1991	FUR	12
1992	API	2
1992	FUR	10
1992	KNY	1
1992	LNP	1
1992	SBA	3
1993	FUR	7
1994	API	6
1994	SBA	11
1996	ABG	21
1999	ABG	84
2001	ABG	10
2002	LMM	2
2002	MLT	3
2003	KIR	10
2004	ABG	6
2011	ESA	2

The number of minute mean values flagged in this manner, and by manual inspection of data in same year, is listed in Table 2. Hourly means have been recomputed and minute and hourly data and metadata have been updated. One exception is Alibag (ABG) where the hourly mean values come from a different measurement system than the minute mean values. In this case, as it was desirable to retain a homogeneous long time series of hourly data, the hourly mean values were not updated. In total 28 year-files of minute means and 24 of hourly means were updated.

Only large spikes that might have an impact on extreme statistics analyses (e.g. Thomson *et al,* 2011) have been removed. The data may still contain small spikes and other artefacts.

Identification of QC issues in hourly means

Many obvious errors of a typographical nature in the .wdc hourly mean files held at Edinburgh have already been corrected (Dawson *et al*, 2009, Reay *et al*, 2013).

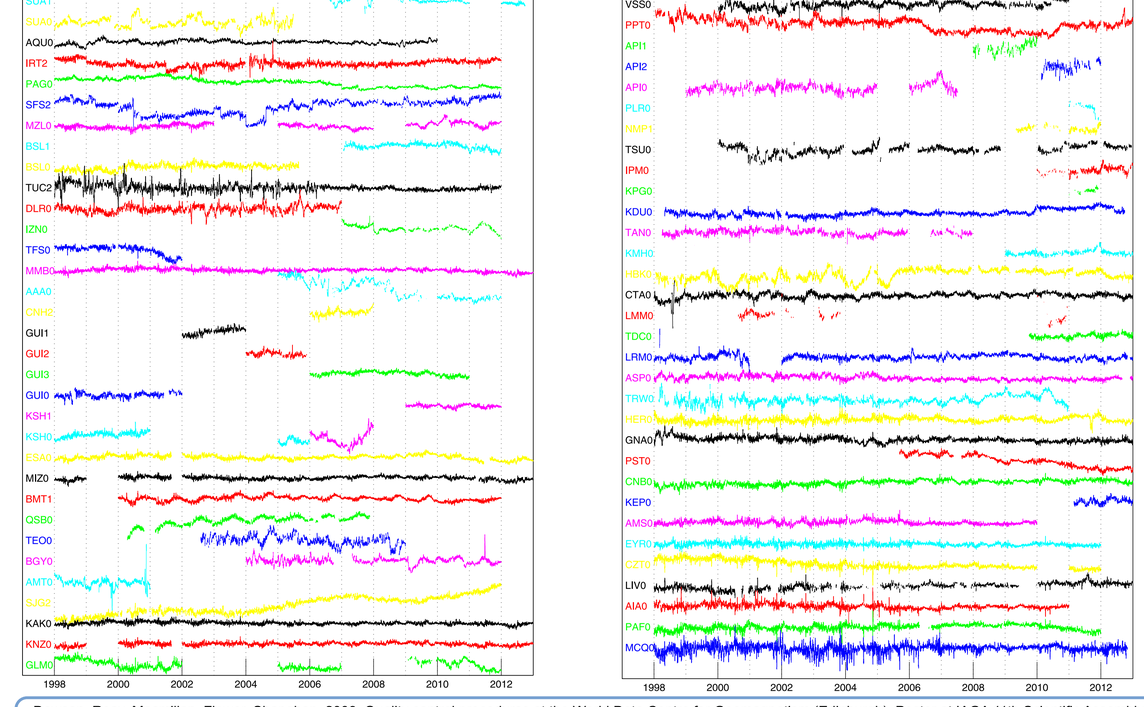
In preparation for the Swarm mission files of good quality observatory hourly means in a particular format and updated once every 3 months are required. These data should

- contain minimal measurement noise
- be as complete as possible (no data gaps in time)
- be corrected to absolute values over multi-year periods, i.e. drift-free
- be without discontinuities
- be in geocentric coordinate frame and time- and position-stamped

To meet these requirements misfits of spherical harmonic models can be inspected in the temporal and spatial domains. Pre-processing removes all known signals, i.e. core, crust and quiet-time ionosphere, and the modelling fits most of the remaining coherent field on an hourly basis. The misfits mainly represent measurement artefacts on the 0-10 nT scale (Macmillan and Olsen, 2013).

The procedure was tested with 1998-2012 data from WDC Edinburgh, and quasi-definitive data from INTERMAGNET and other sources. The plots show misfits after data selection for geomagnetic B_φ, B_θ and B_r components, ordered by geomagnetic latitude (<64°). Straight lines close to zero reflect high quality data except at high latitudes. The importance of good long-term observing and data-processing practice is highlighted. Cleaned-up data are available at

ftp://ftp.nmh.ac.uk/geomag/smac/AUX_OBS_2/



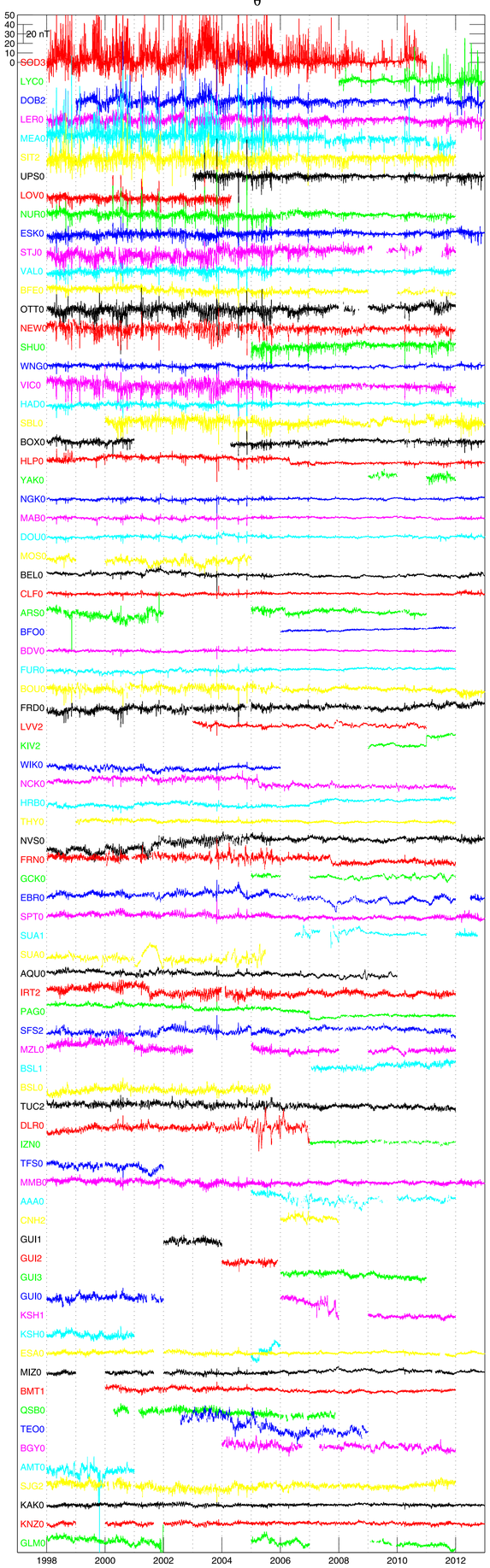
Dawson, Reay, Macmillan, Flower, Shanahan, 2009. Quality control procedures at the World Data Centre for Geomagnetism (Edinburgh). Poster at IAGA 11th Scientific Assembly, Sopron. (Unpublished) http://nora.nerc.ac.uk/11740/.

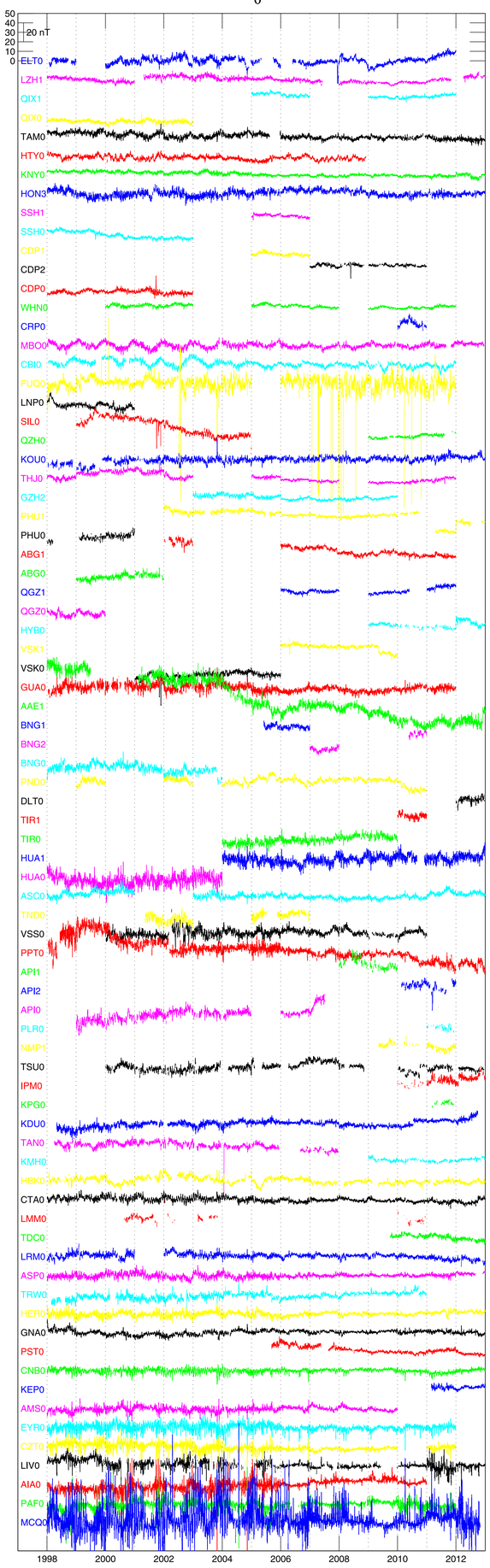
Macmillan and Olsen, 2013. Observatory data and the Swarm mission. Accepted for *Earth, Planets and Space* Swarm special issue.

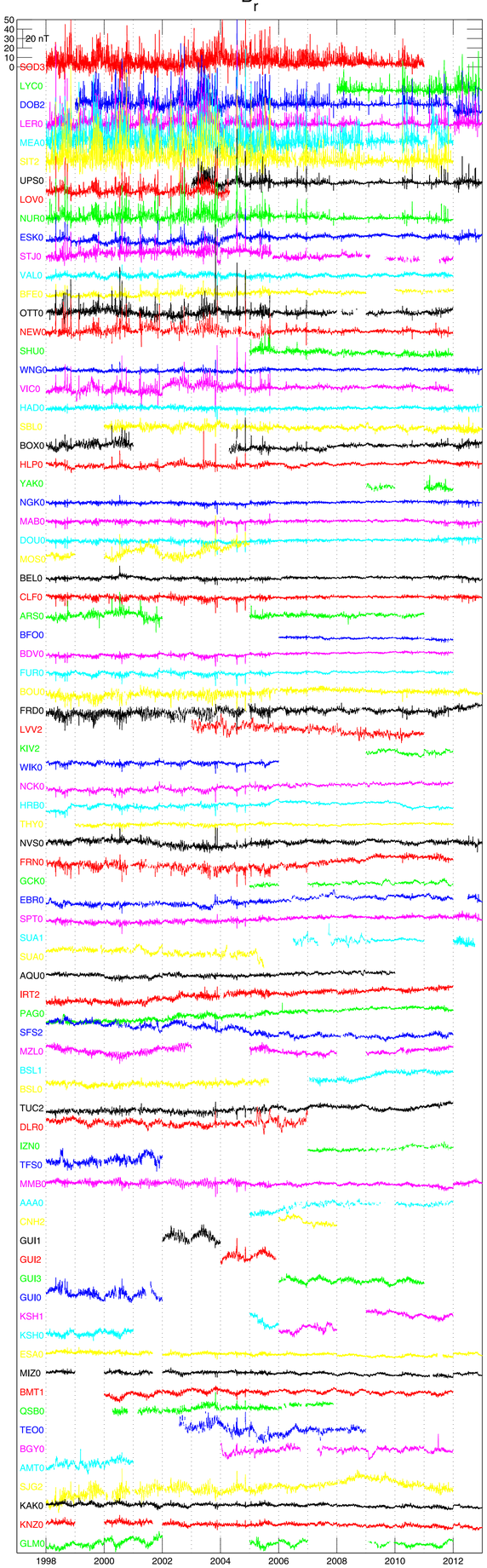
Reay, Clarke, Dawson, Macmillan, 2013 Operations of the World Data Centre for Geomagnetism, Edinburgh [in special issue: Proceedings of the 1st WDS Conference in Kyoto 2011] *Data Science Journal*, 12. WDS47-WDS51. 10.2481/dsj.WDS-005.

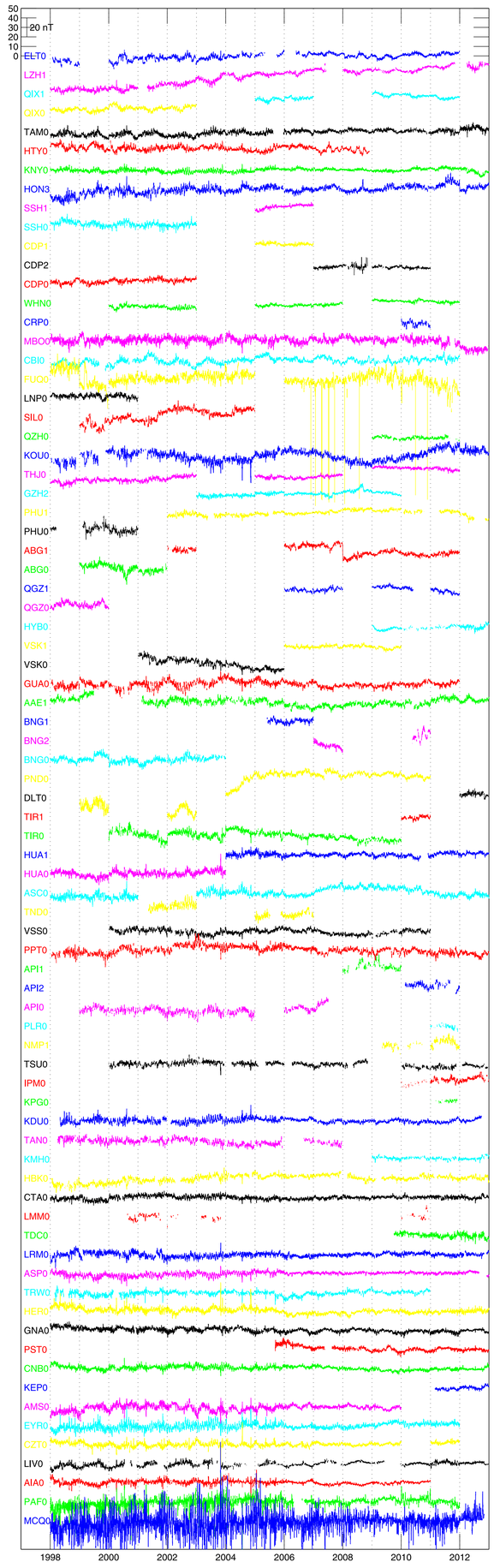
Thomson, Dawson and Reay, 2011. Quantifying extreme behaviour in geomagnetic activity. *Space Weather*, VOL. 9, S10001, doi:10.1029/2011SW000696.

The many organisations around the world running and supporting observatories are thanked for their efforts.


B θ
20 nT
50
40
30
20
10
0
SOD3
LYC0
DOB2
LER0
MEA0
SIT2
UPS0
LOV0
NUR0
ESK0
STJ0
VAL0
BFE0
OTT0
NEW0
SHU0
WNG0
VIC0
HAD0
SBL0
BOX0
HLP0
YAK0
NGK0
MAB0
DOU0
MOS0
BEL0
CLF0
ARS0
BFO0
BDV0
FUR0
BOU0
FRD0
LVV2
KIV2
WIK0
NCK0
HRB0
THY0
NVS0
FRN0
GCK0
EBR0
SPT0
SUA1
SUA0
AQU0
IRT2
PAG0
SFS2
MZL0
BSL1
BSL0
TUC2
DLR0
IZN0
TFS0
MMB0
AAA0
CNH2
GUI1
GUI2
GUI3
GUI0
KSH1
KSH0
ESA0
MIZ0
BMT1
QSB0
TEO0
BGY0
AMT0
SJG2
KAK0
KNZ0
GLM0
1998
2000
2002
2004
2006
2008
2010
2012


B θ
20 nT
50
40
30
20
10
0
ELT0
LZH1
QIX1
QIX0
TAM0
HTY0
KNY0
HON3
SSH1
SSH0
CDP1
CDP2
CDP0
WHN0
CRP0
MBO0
CBI0
FUG0
LNP0
SIL0
QZH0
KOU0
THJ0
GZH2
PHU1
PHU0
ABG1
ABG0
QGZ1
QGZ0
HYB0
VSK1
VSK0
GUA0
AAE1
BNG1
BNG2
BNG0
PND0
DLT0
TIR1
TIR0
HUA1
HUA0
ASC0
TND0
VSS0
PPT0
API1
API2
API0
PLR0
NMP1
TSU0
IPM0
KPG0
KDU0
TAN0
KMH0
HBK0
CTA0
LMM0
TDC0
LRM0
ASP0
TRW0
HER0
GNA0
PST0
CNB0
KEP0
AMS0
EYR0
CZT0
LIV0
AIA0
PAF0
MCQ0
1998
2000
2002
2004
2006
2008
2010
2012


B r
20 nT
50
40
30
20
10
0
SOD3
LYC0
DOB2
LER0
MEA0
SIT2
UPS0
LOV0
NUR0
ESK0
STJ0
VAL0
BFE0
OTT0
NEW0
SHU0
WNG0
VIC0
HAD0
SBL0
BOX0
HLP0
YAK0
NGK0
MAB0
DOU0
MOS0
BEL0
CLF0
ARS0
BFO0
BDV0
FUR0
BOU0
FRD0
LVV2
KIV2
WIK0
NCK0
HRB0
THY0
NVS0
FRN0
GCK0
EBR0
SPT0
SUA1
SUA0
AQU0
IRT2
PAG0
SFS2
MZL0
BSL1
BSL0
TUC2
DLR0
IZN0
TFS0
MMB0
AAA0
CNH2
GUI1
GUI2
GUI3
GUI0
KSH1
KSH0
ESA0
MIZ0
BMT1
QSB0
TEO0
BGY0
AMT0
SJG2
KAK0
KNZ0
GLM0
1998
2000
2002
2004
2006
2008
2010
2012


B r
20 nT
50
40
30
20
10
0
ELT0
LZH1
QIX1
QIX0
TAM0
HTY0
KNY0
HON3
SSH1
SSH0
CDP1
CDP2
CDP0
WHN0
CRP0
MBO0
CBI0
FUG0
LNP0
SIL0
QZH0
KOU0
THJ0
GZH2
PHU1
PHU0
ABG1
ABG0
QGZ1
QGZ0
HYB0
VSK1
VSK0
GUA0
AAE1
BNG1
BNG2
BNG0
PND0
DLT0
TIR1
TIR0
HUA1
HUA0
ASC0
TND0
VSS0
PPT0
API1
API2
API0
PLR0
NMP1
TSU0
IPM0
KPG0
KDU0
TAN0
KMH0
HBK0
CTA0
LMM0
TDC0
LRM0
ASP0
TRW0
HER0
GNA0
PST0
CNB0
KEP0
AMS0
EYR0
CZT0
LIV0
AIA0
PAF0
MCQ0
1998
2000
2002
2004
2006
2008
2010
2012