Metrics for addressing UK Met Office user needs

Michael Sharpe, Sophie Murray*, David Jackson, Edward Pope, **Suzy Bingham**

CCMC Workshop, Florida, 3rd April 2017

* Trinity College Dublin

Outline

- Met Office Space Weather Operations Centre (MOSWOC)
- Verification in near real-time
  - Systems adapted from terrestrial weather
  - Geomagnetic storm / flare forecasts
- Flare forecast verification
- CME arrival time verification
- Summary

Met Office Space Weather Operations Centre (MOSWOC)

- 24/7 space weather monitoring service since 2014
- 2 forecasters on duty (1 dedicated to space weather)
- Fully integrated within Met Office Operations Centre
- Provide twice daily forecasts, & timely alerts & warnings
- 14 forecasters, 5 scientists,

3 Business, IT developers

- National capability supporting; government, military & critical sectors (National Grid, satellite operators, etc)

Public webpages:

http:// www.metoffice.gov.uk/publicsector/emergencies/space-weather

MOSWOC twice daily forecasts

Met Office

Space Weather Technical Forecast

Space Weather Technical Forecast (Ref: MO43)

Issued on Monday, 18 July 2016 at 01:31 local

This scientific guidance document provides a four day assessment of space weather events. The probabilities stated below are for reaching or exceeding the given levels. For more information about space weather impacts please see the Met Office Space Weather Scales http://www.metoffice.gov.uk/media/pdf/j/7/Space_weather_scales.pdf

Space Weather Forecast Headline: High energy electron fluence above Active threshold at times. Weak Earth directed CME observed.

Analysis of Space Weather Activity over past 24 hours

Solar Activity: Solar activity has been low with a number of C-class flares. A long duration event, from around 17/0600 UTC gave a peak of C1.4 at 17/0801 UTC. This originated from AR2567 and showed interaction with AR2565. The largest flare was a C6, which originated from the same area. AR2567 is the largest most complex sunspot region on the disc (Dhi Beta-gamma). A CME which originated in the region of AR2565 and AR2567 appeared in SOHO coronagraph imagery at 17/1230 UTC. A weak filament eruption also occurred in the southeast quadrant at around 17/1400 UTC, which may be a source of part of the CME activity observed. Analysis is ongoing and there is some uncertainty due to the different components of the CME.

Solar Wind / Geomagnetic Activity: The solar wind decreased from around 500 km/s to around 440 km/s through the period as the effects of a high speed stream from coronal hole 93 wane. The total IMF has been in the range 2-5nT. The Bz (north-south component) has been between -4 and +4nT. The Phi angle was largely in the negative (towards the sun) sector, changing to positive (away from the sun) from 17/1700 UTC. Geomagnetic activity was quiet.

Energetic Particles/Solar Radiation: The high energy proton flux (greater than 10 MeV) at background levels. The flux of high energy electrons (greater than 2 MeV) at geosynchronous orbit was at Moderate to High levels, with a peak flux of 5610pfu at 17/1650UTC. The corresponding 24-hour fluence was above the Active threshold (1e8) during the past 24 hours.

Four-Day Space Weather Forecast Summary

Solar Activity: Solar activity is expected to continue at mainly Low levels, with a slight chance (20%) of M-class flares throughout the forecast period.

Solar Wind / Geomagnetic Activity: Geomagnetic activity is expected to remain at quiet or unsettled levels on days 1 and 2 (16th and 19th), although with a slight chance of active periods later day 2 due to high speed stream effects. Active periods are possible on days 3 and 4 due to the arrival of a weak CME and the onset of another high speed stream with a slight chance of G1 minor storm periods.

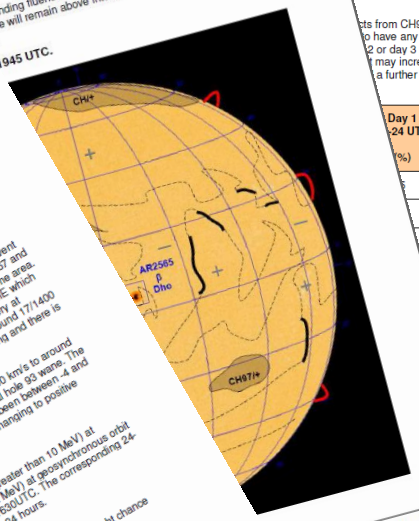
Energetic Particles / Solar Radiation: No solar radiation storms are expected during the period, with the high energy proton flux (greater than 10MeV) forecast to remain at background levels. The high energy electron flux (greater than 2MeV) is expected to continue at mainly Moderate to High

Space Weather Advisor : Helen Wade
Tel: 01392 886112 Email: moswoc@metoffice.gov.uk

Space Weather Technical Forecast

responding fluence likely to be above the Active threshold at these will remain above into day 3 and 4, however weak HSS on.

1945 UTC.



larity +/-, Coronal Holes: Grey
nd Zurich-McIntosh Axx etc.

Space Weather Technical Forecast

gion of AR2565 and AR2567 appeared in SOHO coronagraph
filament eruption also occurred in the southeast quadrant at
e a source of part of the CME activity observed. Analysis is
nty due to the different components of the CME. However, a
ly to arrive late day 3 (20th) or on day 4 (21st) with active
of G1 minor geomagnetic storms.

ts from CH93 HSS wane. Coronal hole 97 in the southern
o have any significant effects, although gives a slight
2 or day 3 (19th and 20th). Following this negative-polarity
may increase geomagnetic activity from day 4 (21st),
a further slight chance of G1.

Day 1 (00-24 UTC) (%)	Day 2 (00-24 UTC) (%)	Day 3 (00-24 UTC) (%)	Day 4 (00-24 UTC) (%)
	5	15	25
	1	1	1
	1	1	1
	1	1	1

ass Ejections

Estimated Speed	Estimated Arrival Time	Comments
497	20/2000 UTC	Low confidence with further analysis to take place.

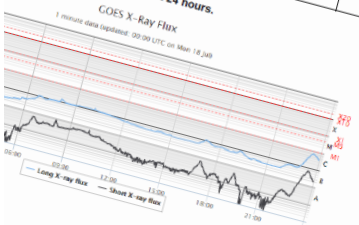
considered likely in the coming

Met Office Space Weather Technical Forecast

far side helioseismic data and STEREO-A EUVI imagery does not indicate any significant active regions rotate through the east limb in this forecast period.

Level	Past 24 Hours (Yes/No)	Day 1 (00-24 UTC) (%)	Day 2 (00-24 UTC) (%)	Day 3 (00-24 UTC) (%)	Day 4 (00-24 UTC) (%)
R1-R2 M Class	N	20	20	20	20
R3 to R5 X Class	N	1	1	1	1

lux from GOES 15 over the past 24 hours.



Energy Protons):

during the forecast period, with the high energy proton flux at background levels.

fice.gov.uk

served.

r Technical Forecast

Day 2 (00-24 UTC) (%)	Day 3 (00-24 UTC) (%)	Day 4 (00-24 UTC) (%)
	1	1
1	1	1

e at Moderate to High fluence is expected to o a decrease in flux ains a chance that

Day 4 (00-24 UTC) (%)
30
5

Summary for next 4 days

Solar analysis

CME arrival time at Earth predictions

4 day probability forecasts for: geomagnetic storms, X-ray flares, high energy protons & high energy electron events

Probabilistic forecasts to verify

- Probabilistic & multi-category (G-level / flare class)
- Truth data used for verification:
  - Geomag storms: daily max Kp (estimated Kp in near real-time, definitive Kp for forecasts older than ~1month)
  - Flare: GOES 1 min long wave radiation flux

Geo-Magnetic Storm Probability (Exceedance)	Level	Past 24 Hours (Yes/No)	Day 1 (00-24 UTC) (%)	Day 2 (00-24 UTC) (%)	Day 3 (00-24 UTC) (%)	Day 4 (00-24 UTC) (%)
Minor or Moderate	G1 to G2	N	5	5	40	30
Strong	G3	N	1	1	5	5
Severe	G4	N	1	1	1	1
Extreme	G5	N	1	1	1	1

Example MOSWOC geomagnetic storm forecast (above) & flare forecast (below).
Column 1 & 2: geomag storm/flare level. Column 3: identifies whether storm/flare has occurred in previous 24 h. Columns 4-7: probabilistic forecast for next 4 days.

X Ray Flares Probability	Level	Past 24 Hours (Yes/No)	Day 1 (00-24 UTC) (%)	Day 2 (00-24 UTC) (%)	Day 3 (00-24 UTC) (%)	Day 4 (00-24 UTC) (%)
Active	R1-R2 M Class	N	20	20	15	10
Very Active	R3 to R5 X Class	N	2	2	1	1

NRT forecast verification – Michael Sharpe

Probabilistic forecast (geomag storms / flares)

Treat as **multi**-category

Area Forecast Verification System (AFVS)

Assess forecast skill: ***Ranked Probability Score (RPS)***

Compare against a reference for benchmark of performance.
Geomag storm predictor: 6 months
Flare predictor: 8 months

Assess forecast performance compared to reference: ***Ranked Probability Skill Score (RPSS)***

Treat each category (storm/flare level) **separately**

Warnings Verification System (WVS)

ROC plots *(forecast resolution),* ***Reliability diagrams***

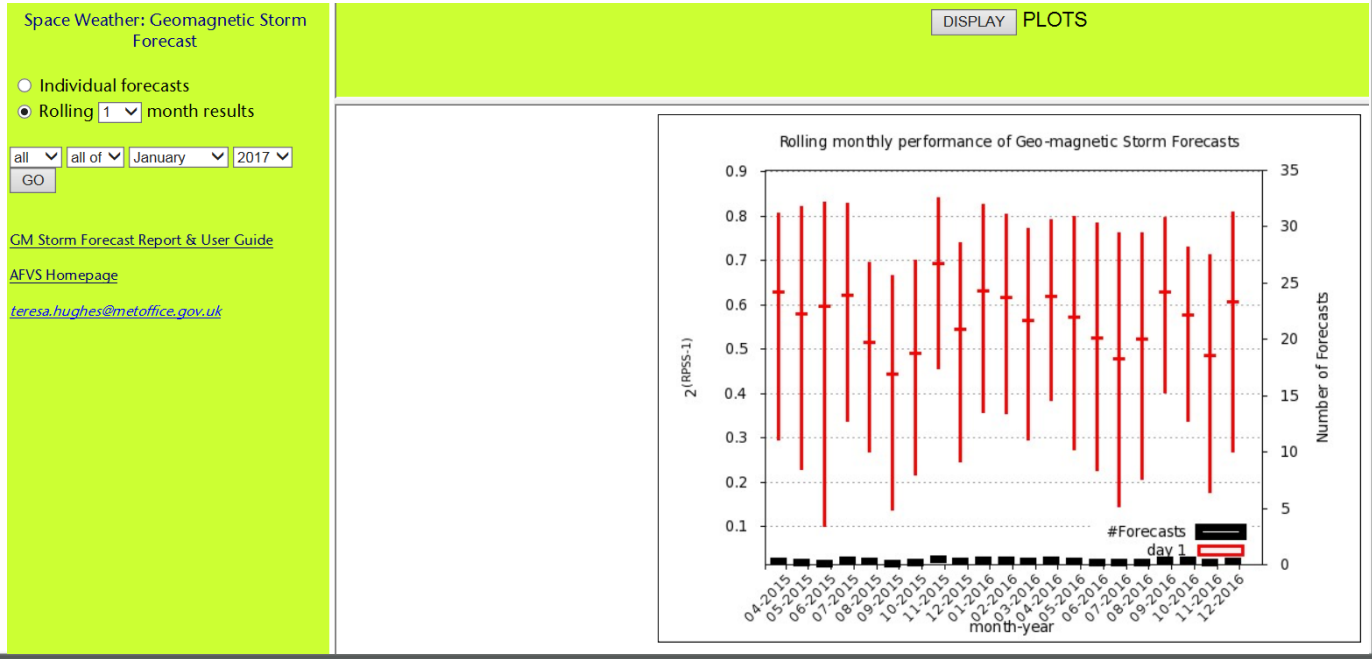
Area Forecast Verification System (AFVS)

Area Forecast Verification System: originally applied to marine products (e.g. shipping forecast)

Metric applied in AFVS: Ranked Probability Score (RPS)- commonly associated with multi-category probabilistic forecasts (Epstein 1969, Murphy 1971)

RPS is calculated for recent & archived forecasts – rolling 12 monthly performance plot is updated daily to monitor rolling skill compared to reference

Example of MOSWOC internal verification pages: AFVS showing rolling monthly geomag forecast performance



Sharpe, 2013, Verification of Marine Forecasts using an Objective Area Forecast Verification System, Meterol. Apps.

Epstein, 1969, A scoring system for probability forecasts of ranked categories, J. Applied Meteorology

Murphy, 1971, A note on the ranked probability score, Journal of Applied Meteorology

RPS and RPSS

Ranked Probability Score (RPS)
Shows how well the probability forecast predicts the category which the observation falls into.
Measures the sum of squared differences in cumulative probability space for a multi-category probabilistic forecast. Penalises more when probabilities are further from the observation (same as Brier Score for two category forecasts).

$$RPS = \sum_{n=0}^{5} (P(G_n) - O(G_n))^2$$

P(Gn): forecast probability that max storm level to be observed during 24 h is <=Gn (n = 0, 1/2, 3, 4 or 5)
O(Gn): 0 if max observed level is <Gn, & 1 otherwise
RPS range: 0 - 1 (where 0 = perfect)

Ranked Probability Skill Score (RPSS):
Skill score based on RPS values.
Shows relative improvement of probability forecast over reference forecast in predicting the category which the observation falls into. Benchmarks forecast performance by comparing against a reference.
For this verification: RPS is calculated separately for every day, for each forecast. Mean value ($\overline{RPS}$) is obtained by averaging RPS values calculated for a large number of forecasts.

$$RPSS = 1 - \frac{\overline{RPS}}{\overline{RPS}_{ref}}$$

RPSS range: -infinity – 1 (where 1 = perfect)
RPSS > 0 forecast is more skilful than reference.

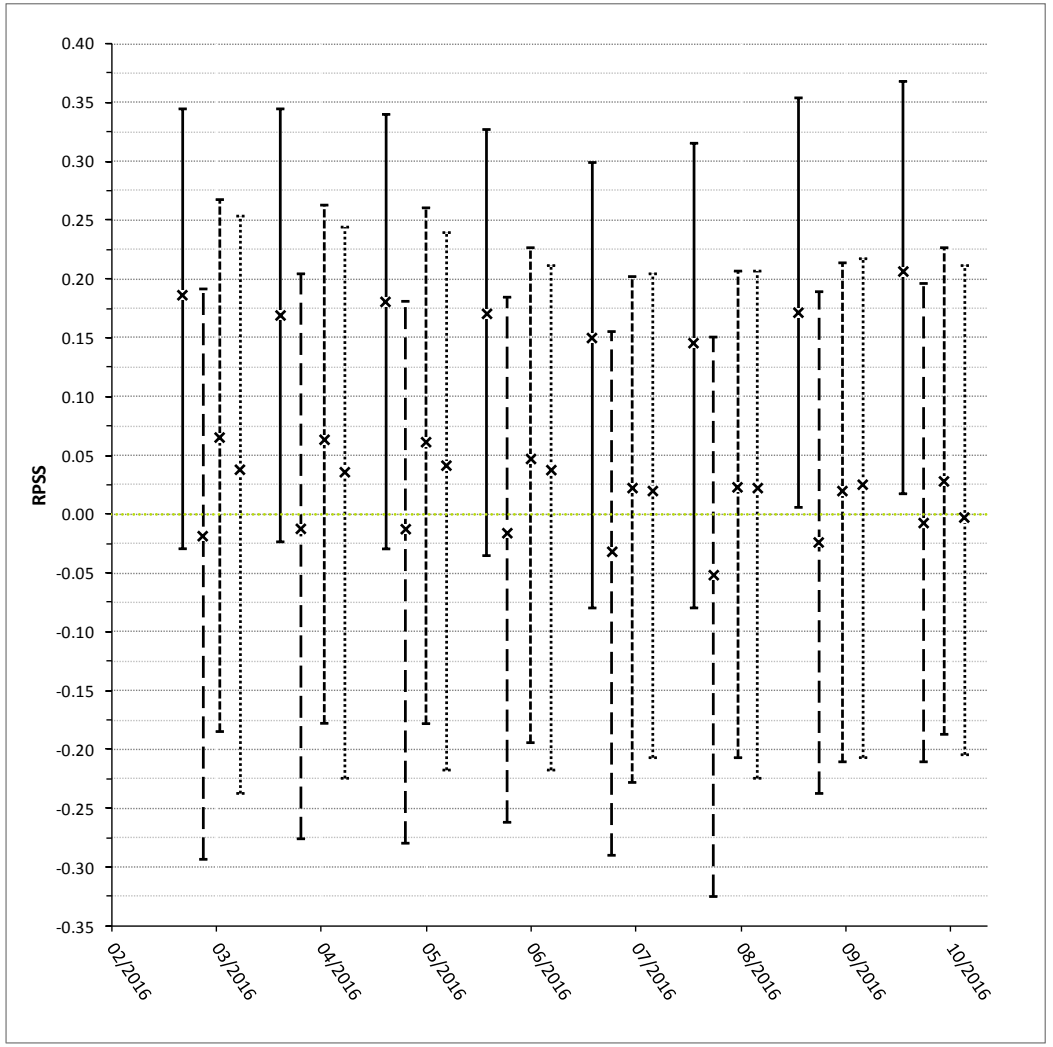
Geomagnetic storm forecast RPSS

- RPSS is greatest on day1 but associated Confidence Intervals (CIs) generally cross the green no-skill line.

For similar analysis of flare forecasts:

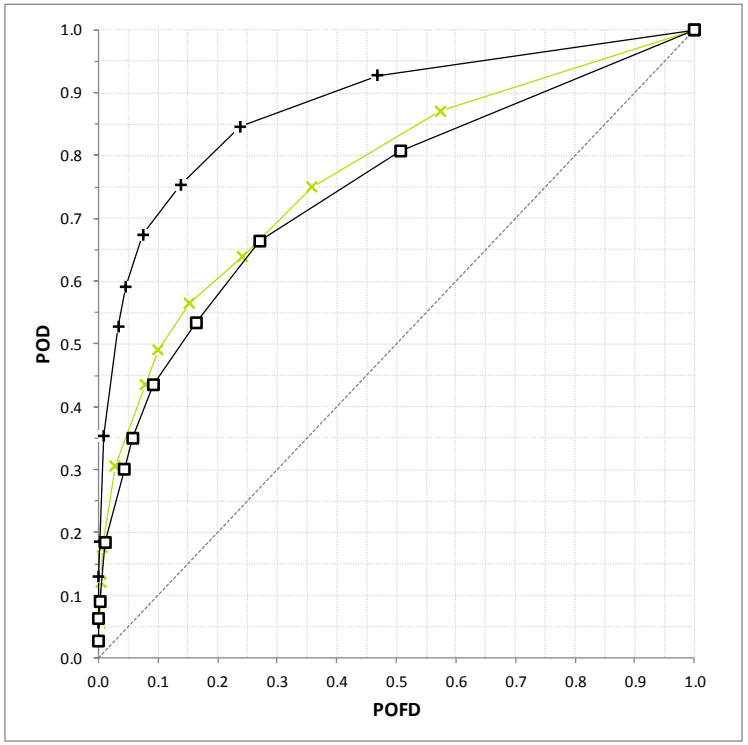
- Little evidence to suggest that forecast has more skill than predictor for forecasting max daily flare class.

- No evidence in CIs to suggest any forecast day is more skilful than another but some suggestion from RPSSs that day 1 is more accurate.



Rolling 12-monthly RPSS values (x) with 90% bootstrapped CIs for each day of the geomag storm forecast (Mar-Oct 2016).
Day 1,2,3 & 4 are indicated by solid, long dashed, short dashed and dotted lines, respectively.

Warnings Verification System (WVS)

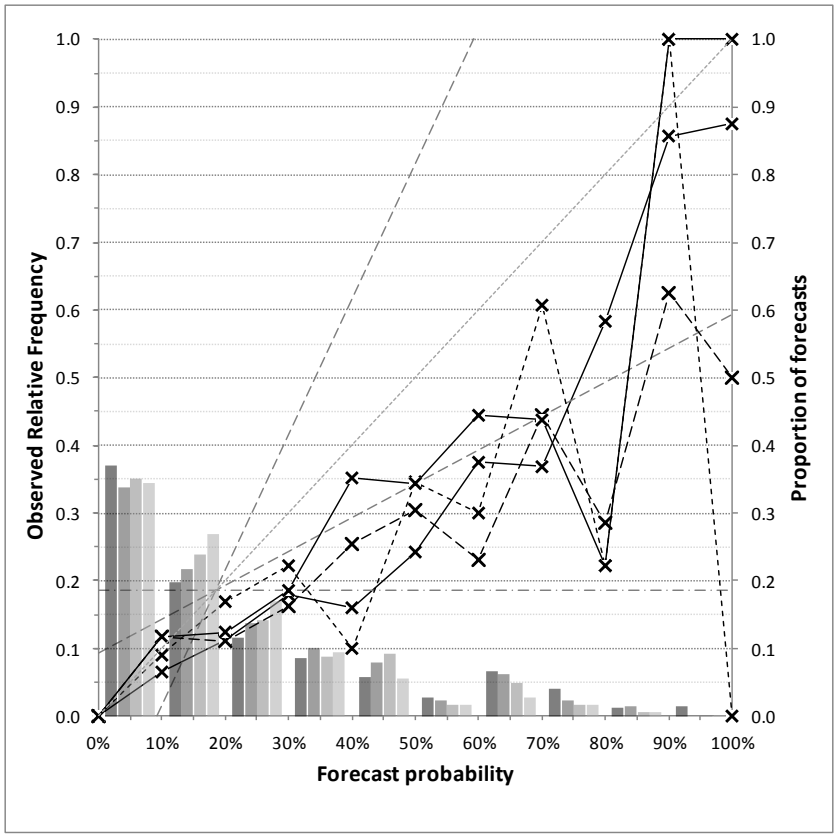


ROC-plot generated using the (x) un-flexed, the flexed including low-misses (☐) and the flexed excluding low-misses (+) technique for day 1 only of geomag storm forecasts for G1, issued between April 2015 and October 2016. Grey diagonal line is no-skill. POFD=probability of false detection. POD=probability of detection.

Relative Operating Characteristic (ROC) plot:

- Measures forecast discrimination
- Conditioned on observations
- Gives info on hit rates & false alarm rates expected from using different probability thresholds to trigger advisory action
- Can use ROC plot to select trigger threshold for an event that provides best balance between hit rate & false alarm rate for a particular decision
- Points are all above the grey diagonal no-skill line, thus indicating that geomag storm forecast has skill at discriminating events of G1 or above.

NRT WVS verification – geomagnetic storm forecast



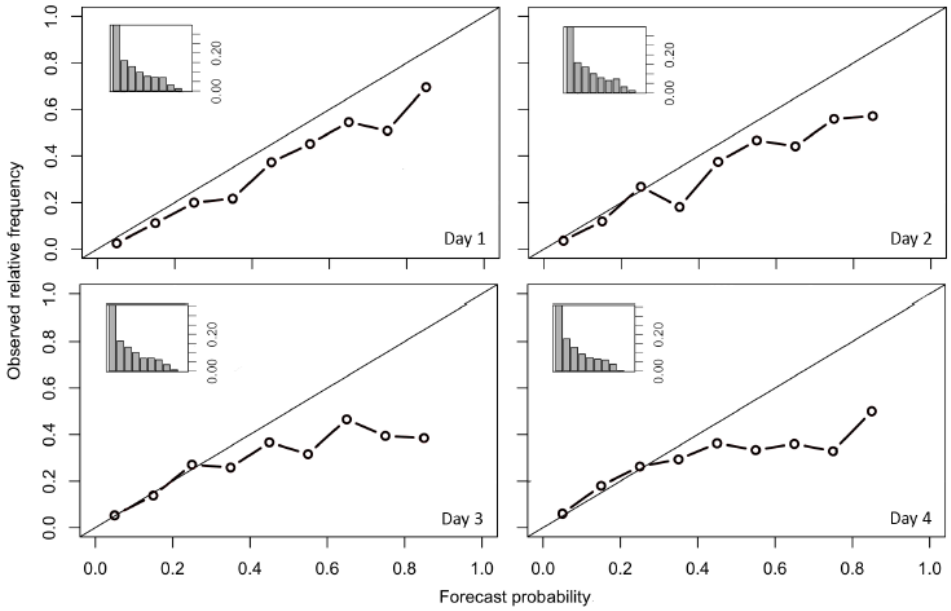
Reliability diagram:

- Measures how closely forecast probabilities of an event correspond to actual chance of observing event
- Conditioned on forecasts
- Plot of frequency of the observations v forecast probability.

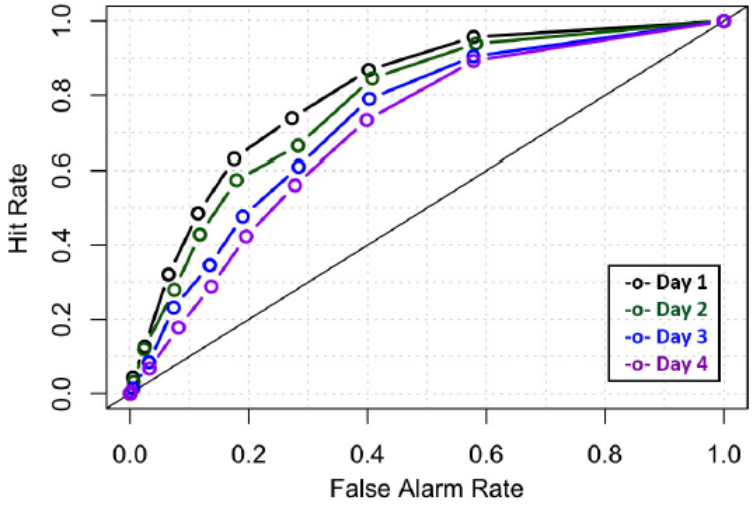
Reliability diagrams for geomag storm forecasts of G1 issued between April 2015 and October 2016 on: day 1 (solid/dark grey), day 2 (long-dashed/mid-dark grey), day 3 (short-dashed/mid-grey) and day 4 (dotted/light grey); when verified against daily maximum Kp values of at least 4-.

Flare forecast verification – Sophie Murray

- Verified M-class flare forecasts between 2014-July 2016
- Show forecaster added-value to issued forecasts
- Forecast skill is best at shorter forecasting periods



Reliability diagrams for 4 day forecasts. Forecasts issued with probabilities >30% appear to over-predict flares. Day 1 forecasts are more reliable.



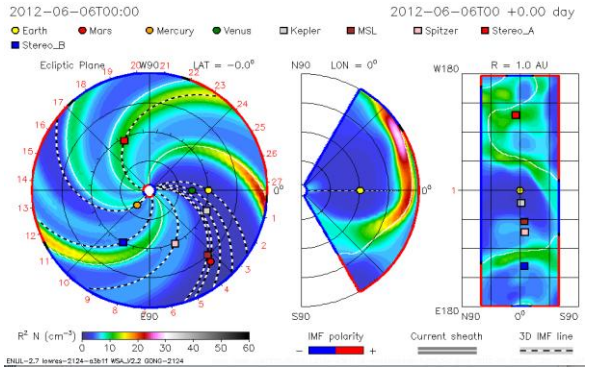
ROC plot for 4 day forecasts. Day 1 forecasts show greater skill.

Sophie Murray, et al., 'Flare forecasting at the Met Office Space Weather Operations Centre', Sp. Weather accepted Mar. 2017

CME forecast verification

Date/time 21.5R (UTC)	Halo: Full or Partial	Source	Source Location	Estimated Speed	Estimated Arrival Time	Comments
12/1920Z	Partial	Filament eruption	20S35W	600	16/0200	nil

Example MOSWOC CME forecast



- Forecast method: WSA Enlil + CME Analysis Tool
- Verification method:

  Compared observed CME arrival times (identified using ACE data) with MOSWOC forecasts:

  - Use verification statistics derived from 2x2 contingency table, e.g. hit rate, false alarm rate, Heidke/Peirce skill scores, etc
  - Bootstrap contingency table to get 90% confidence interval for each derived quantity
  - Compared MOSWOC performance against CCMC Scoreboard average

Forecast \ Observed		
	Hit	False alarm
	Miss	Correct rejection

MOSWOC v CCMC average CME arrival time verification

Category	Metric	MOSWOC	CCMC	90% conf. ints.overlap?
Accuracy	Proportion Correct	0.73	0.75	
	Threat Score	0.69	0.69	
Bias	Bias	0.93	1.44	N
Reliability	False Alarm Ratio	0.15	0.31	N
Discrimination	Hit Rate	0.79	1.00	N
	False Alarm Rate	0.46	0.57	N
Skill	Heidke	0.30	0.45	
	Peirce	0.32	0.43	
	Equit. Threat Score	0.18	0.30	

- **Hit rate**: CCMC average always predicts a hit; false alarm rate and ratio are also higher
- **Bias**: MOSWOC 0.9 - slight under-prediction of events

  CCMC 1.4 - over-prediction of events (consistent with the high hit/false alarm rate)
- **Equitable Threat Score** and **Heidke Skill Scores** are comparable
- Overall, results suggest broadly comparable performance of MOSWOC and CCMC average, CME forecasts despite slightly different approaches

Summary

Metrics applied:

- RPS – shows how well the probability forecast predicts the category which the observation falls into

- RPSS - shows relative improvement of probability forecast over reference forecast in predicting the category which the observation falls into

- ROC plot - measures forecast discrimination

- Reliability diagram - measures how closely forecast probabilities of an event correspond to actual chance of observing event

- Contingency table skill scores - e.g. Bias, False Alarm Ratio, Heidke, Peirce.

Verification methods/definitions: http://www.cawcr.gov.au/projects/verification/

Summary

Forecast verification undertaken:

- Near real-time probabilistic geomagnetic storms & flares
  - Forecasts were skilful at identifying minor geomag storms & M-class flares but both were over-forecast (Apr 2015-Oct 2016)
  - Rolling prediction periods of 6 and 8 months provide the most skillful forecasts for Geo-magnetic Storm and X-ray flare forecasts respectively during the 10year period between 2006 and 2015.
  - Rolling 12 month analysis using RPSS indicates day 1 geomag storm forecasts are more skilful than 6 month predictor reference forecast. No consistent evidence for flare forecasts (reference: 8 month predictor).
- Archived M-class probabilistic flares (2014- July 2016):
  - Showed forecaster added skill to issued forecasts (compared to model output)
  - Forecasts issued with probabilities >30% appear to over-predict flares
  - Day 1 forecasts are more skilful than forecasts on later days
- CME arrival times at Earth:
  - MOSWOC & Scoreboard Average – comparable in skill

Summary

Collaboration is key:

- NASA CCMC Flare/SEP scoreboards (Sophie Murray / Mike Marsh)

- International Space Environment Services
  - Forecast standardisation
  - Internationally consistent verification

- FLARECAST project:
  - Automated ensemble forecasting system will be compared with current forecasting methods.
  - Met Office involvement with verification and dissemination.

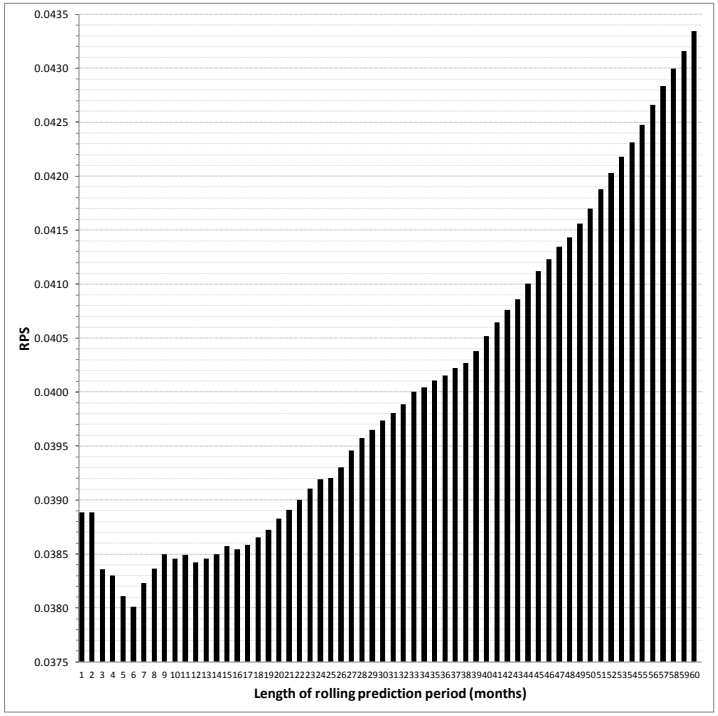
Met Office

Thank you

www.metoffice.gov.uk

Reference forecast

‘Climatology’ / predictor periods used: geomag storms - previous 6 months , flares - previous 8 months

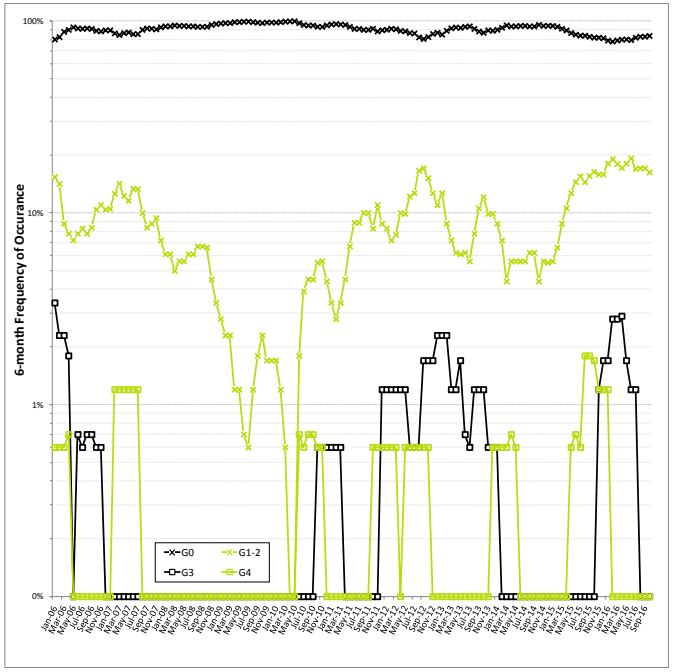


Determined best performing rolling predictor period (between 1 & 60 months) to use as a reference forecast.

Used geomag storm observations over 10 y period (’06-’15).

Looked at RPS for each rolling prediction period.

Found that for geomag storm, RPS was lowest for 6 month predictor period.



Rolling 6-month frequency of occurrence of daily maximum geomag storm level

$\overline{RPS}$ *obtained using the frequency of occurrence of geomag storm forecasts during prediction periods between 1 and 60 months as the forecast of GMS activity during the subsequent month, analyzed over a the 10-year period from 2006 to 2015.*

Space weather verification: key aspects

- Why verify?
  - To monitor, improve & compare forecast quality
  - Understand strengths/limitations
  - To assess forecaster added value
  - For forecasters, modellers, users & stake-holders to understand skill/value
  - Near real-time verification for operational purposes

- Key aspects:
  - Often issued as categories, as probabilities
  - Interest is in extreme events which occur very rarely
  - Data records are short
  - Observations for comparison can be non-existent
  - Data are strongly modulated by 11-year solar cycle
  - Standardise verification procedures across centres to enable comparisons (working with International Space Environment Service http://www.spaceweather.org/)
  - Can adapt NWP verification methods but be aware of differences between space weather/terrestrial meteorology

Terrestrial weather verification resources

WMO Working Group under the World Weather Research Program (WWRP) & Working Group on Numerical Experimentation (WGNE)

Activities:

- Verification research
- Training
- Workshops & tutorials
- Publications on ‘best practices’

http://www.wmo.int/pages/prog/arep/wwrp/new/Forecast_Verification.html

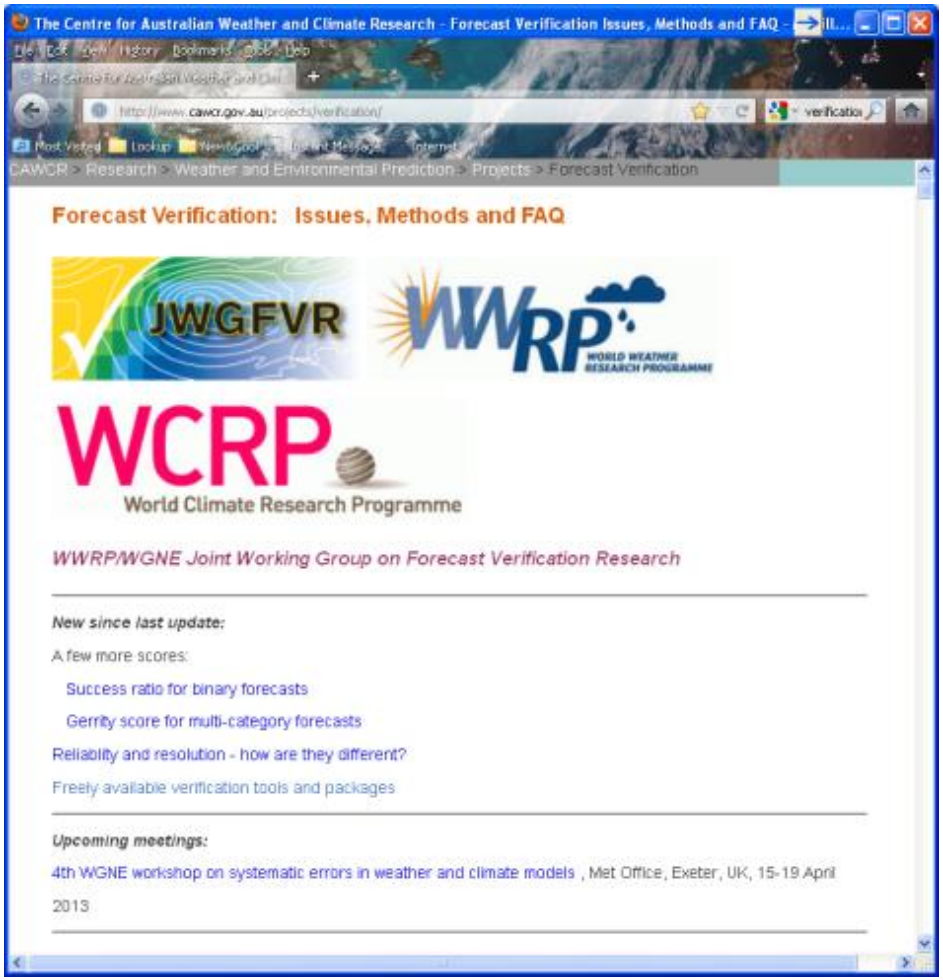
Info from: B. Brown. ISES Verification Workshop, Apr. 2015

Terrestrial weather verification resources

Website maintained by WMO verification Working Group includes:

- Methods (brief definitions)
- Verification issues
- FAQs
- Links and references
- Verification discussion group

http://www.cawcr.gov.au/projects/verification/

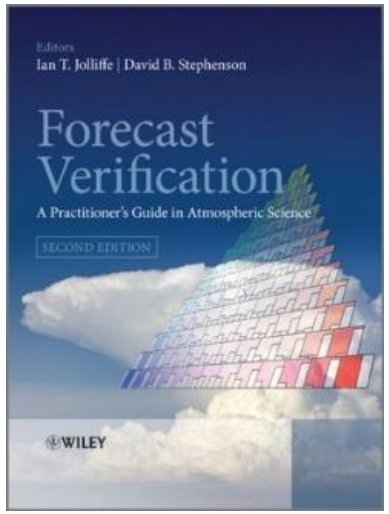


Info from: B. Brown. ISES Verification Workshop, Apr. 2015

Terrestrial weather verification resources

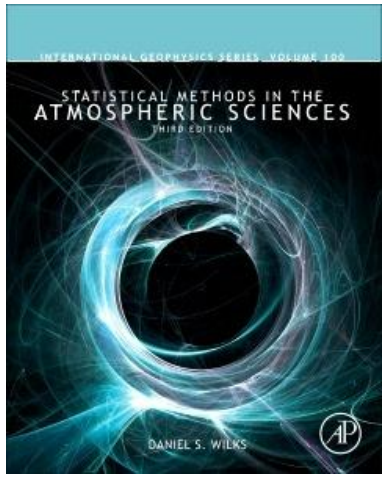
Papers:

- Casati et al. (2008), *Forecast verification: current status and future directions*, Meteorological Applications, **15**, 3-18.
- Ebert et al. (2013), *Progress and challenges in forecast verification,* Meteorological Applications, **20**, 130-139.



Books:

- Jolliffe and Stephenson (2012): *Forecast Verification: a practitioner's guide*, Wiley & Sons.
- Stanski, Burrows, Wilson (1989) *Survey of Common Verification Methods in Meteorology* (available at http://www.cawcr.gov.au/projects/verification/)
- Wilks (2011), *Statistical Methods in Atmospheric Science*, Academic press.

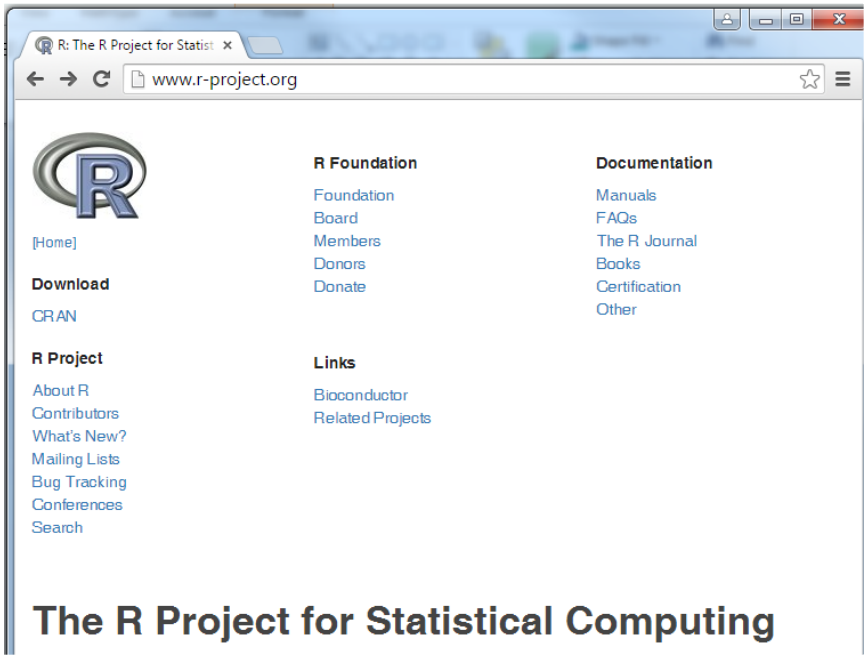


Info from: B. Brown. ISES Verification Workshop, Apr. 2015

Terrestrial weather verification tools: R verification libraries

R verification libraries:

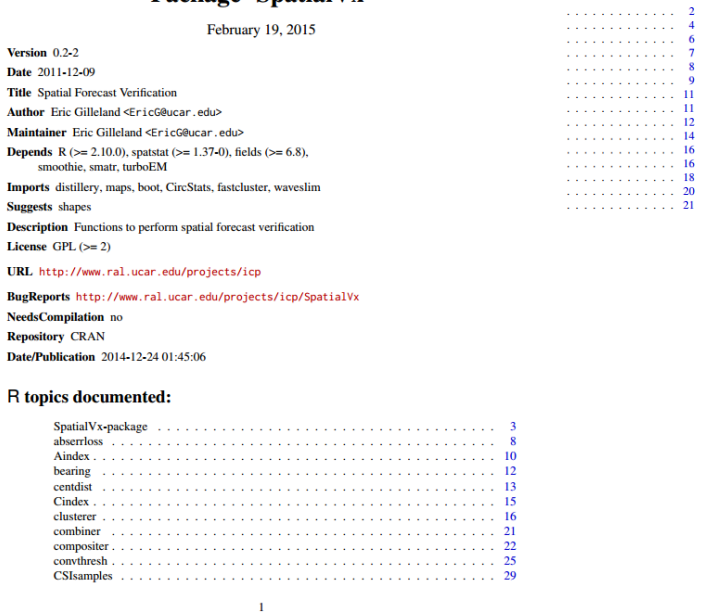
- Freely available statistics packages.
- http://www.r-project.org/
- Maintained & supported by NCAR.



Package 'verification'

February 20, 2015

Version 1.41
Date 2012-4-09
Title Weather Forecast Verification Utilities.
Author NCAR - Research Applications Laboratory
Maintainer Eric Gilleland <ericg@ucar.edu>
Depends R (>= 2.10), methods, fields, boot, CircStats, MASS, dtw
Description This package contains utilities for verification of discrete,continuous, probabilistic forecasts and forecast expressed as parametric distributions.
License GPL (>= 2)
LazyData yes
Repository CRAN
Date/Publication 2014-12-24 20:27:09



Package 'SpatialVx'

February 19, 2015

Version 0.2-2
Date 2011-12-09
Title Spatial Forecast Verification
Author Eric Gilleland <EricG@ucar.edu>
Maintainer Eric Gilleland <EricG@ucar.edu>
Depends R (>= 2.10.0), spatstat (>= 1.37-0), fields (>= 6.8), smoothie, smatr, turboEM
Imports distillery, maps, boot, CircStats, fastcluster, waveslim
Suggests shapes
Description Functions to perform spatial forecast verification
License GPL (>= 2)
URL http://www.ral.ucar.edu/projects/icp
BugReports http://www.ral.ucar.edu/projects/icp/SpatialVx
NeedsCompilation no
Repository CRAN
Date/Publication 2014-12-24 01:45:06

R topics documented:

SpatialVx-package . . . 3
abserrloss . . . 8
Aindex . . . 10
bearing . . . 12
centdist . . . 13
Cindex . . . 15
clusterer . . . 16
combiner . . . 21
compositer . . . 22
convthresh . . . 25
CSIsamples . . . 29

1

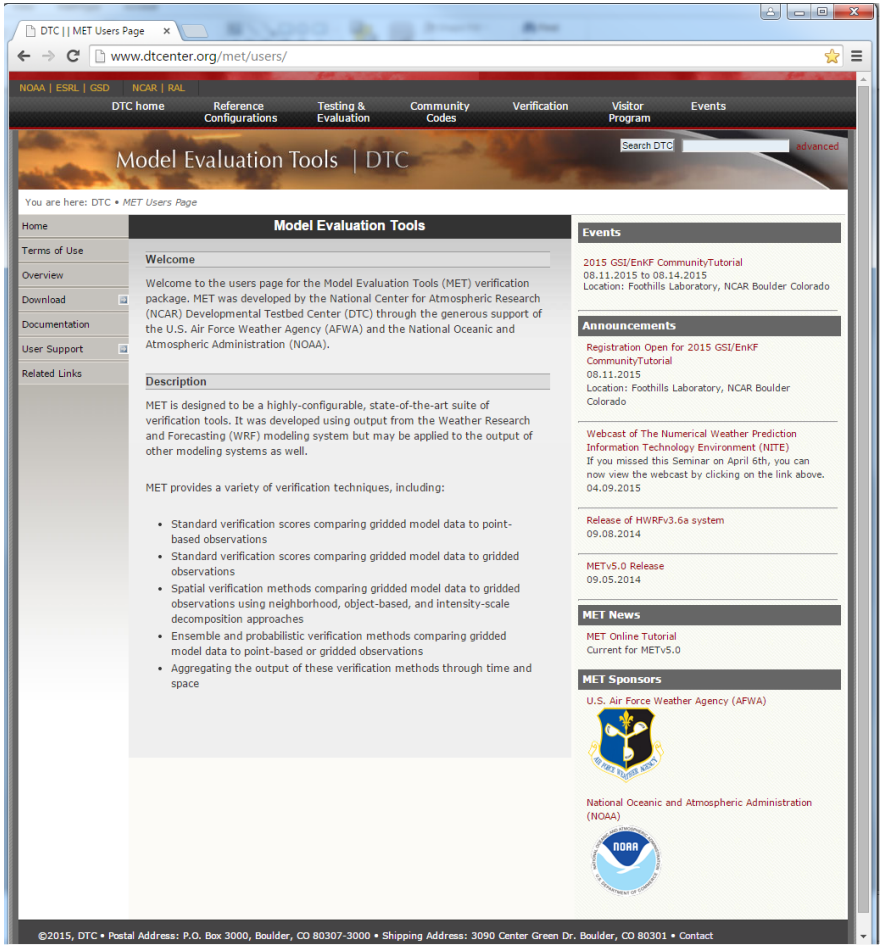
Info from: B. Brown. ISES Verification Workshop, Apr. 2015

Terrestrial weather verification tools: MET

http://www.dtcenter.org/met/users/

Model Evaluation Tools (MET):

- Forecast evaluation tools
- Implemented & supported by the Developmental Testbed Center (DTC) & Joint Numerical Testbed Program at NCAR/RAL
- Includes a suite of standard stats, non-traditional stats (e.g. spatial methods)
- Designed to undertake systematic evaluations
- Has a database & display system for aggregating & plotting data
- Provides a standardized evaluation platform for cross-institution comparisons
- Freely available
- Highly configurable
- Supported via the web & “live” user tutorials



Info from: B. Brown. ISES Verification Workshop, Apr. 2015

Terrestrial weather verification resources

- WMO World Weather Research Programme (WWRP):
  http://www.wmo.int/pages/prog/arep/wwrp/new/Forecast_Verification.html

WWRP/WGNE Joint Working Group on Forecast Verification Research

---

New: **Enter the** ***Challenge to Develop and Demonstrate the Best New User-Oriented Forecast Verification Metric***

The aim of this challenge is to promote user-oriented verification, that is, quantitative assessment of forecast quality in terms that are meaningful to particular forecast users. The scope includes all applications of meteorological and hydrological forecasts. The user-oriented verification metrics will help support the WWRP High Impact Weather Project.
Click here to find out more, or contact verifchallenge@ucar.edu.

Introduction - what is this web site about?

Issues:
Why verify?
Types of forecasts and verification
What makes a forecast good?
Forecast quality vs. value
What is "truth"?
Validity of verification results
Pooling vs. stratifying results

Methods:
Standard verification methods:
- Methods for dichotomous (yes/no) forecasts
- Methods for multi-category forecasts
- Methods for forecasts of continuous variables
- Methods for probabilistic forecasts

Scientific or diagnostic verification methods:
- Methods for spatial forecasts
- Methods for probabilistic forecasts, including ensemble prediction systems
- Methods for rare events
- Other methods

Sample forecast datasets:
- Finley tornado forecasts
- Probability of precipitation forecasts

Freely available verification tools and packages

- Website maintained by WMO verification Working Group, http://www.cawcr.gov.au/projects/verification/, includes:
  - Methods (brief definitions)
  - Verification issues
  - FAQs
  - Links and references
  - Verification discussion group

Info from: B. Brown. ISES Verification Workshop, Apr. 2015

Terrestrial weather verification tools: MET & 'R' libraries

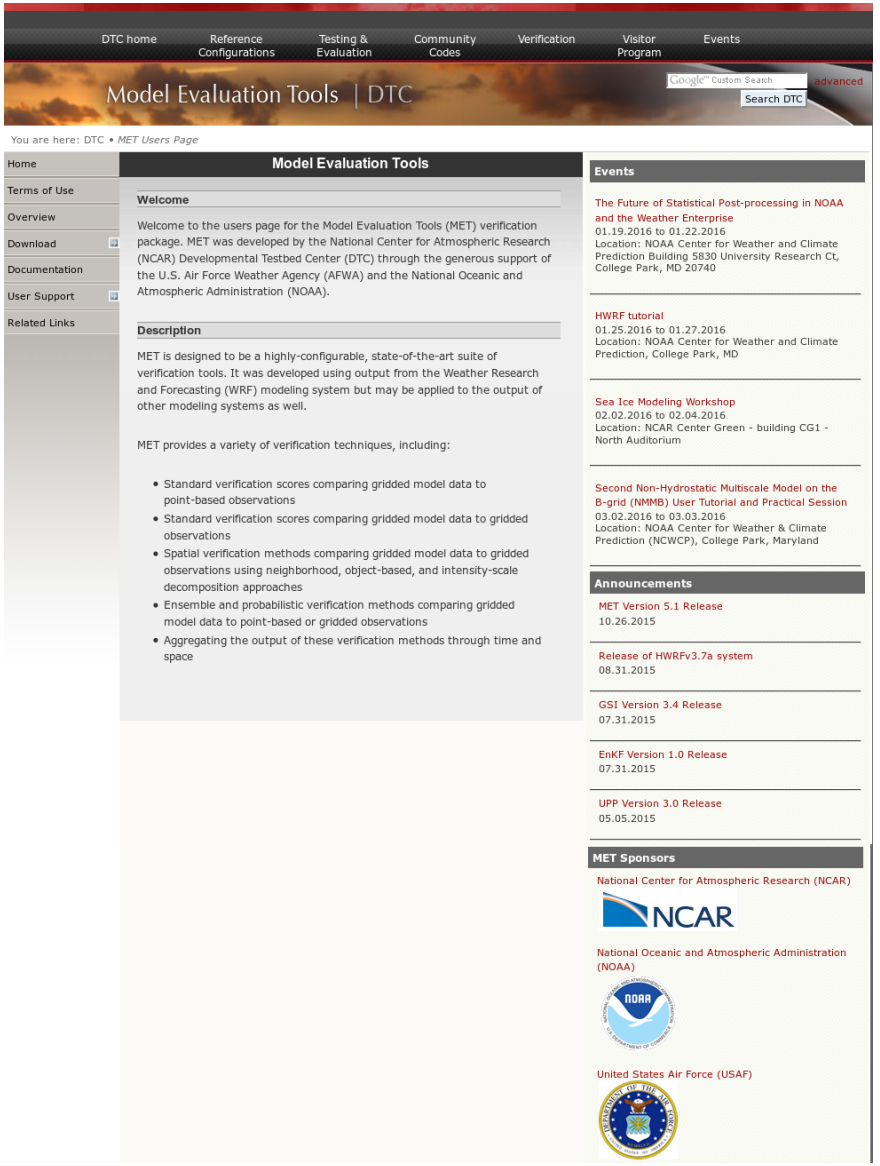
- Model Evaluation Tools (MET):

http://www.dtcenter.org/met/users/

  - Forecast evaluation tools
  - Implemented & supported by the Developmental Testbed Center (DTC) & Joint Numerical Testbed Program at NCAR/RAL
  - Includes a suite of standard stats & non-traditional stats (e.g. spatial methods)
  - Designed to undertake systematic evaluations
  - Has a database & display system for aggregating & plotting data
  - Provides a standardized evaluation platform for cross-institution comparisons
  - Freely available, highly configurable, "live" tutorials

- NCAR verification stats packages:

http://www.r-project.org/



Info from: B. Brown. ISES Verification Workshop, Apr. 2015

Markov chain persistence model

- When the geomagnetic field is disturbed, the Kp-index time series exhibits an almost instantaneous rise, followed by a decay which occurs over a period of 1-2 days
  - A one-step Markov chain provides an informative description:
    - Use time series of daily maximum Kp/G-index to generate a matrix of transition probabilities (T), i.e. $P_{ji} = P(X_{n+1} = j \mid X_n = i)$
  - Starting from the observed state on a given day, u (e.g. u = (0,1,0,0,0)), the forecast probabilities on the *n*th day are: $u_n = uT^n$
  - Quantify uncertainty in transition matrix (and forecast probabilities) by bootstrapping.
  - For N >=3, $T^n \sim P_{clim}$

Verification of Kp

To verify GM Storm forecast observations are needed in near real-time.

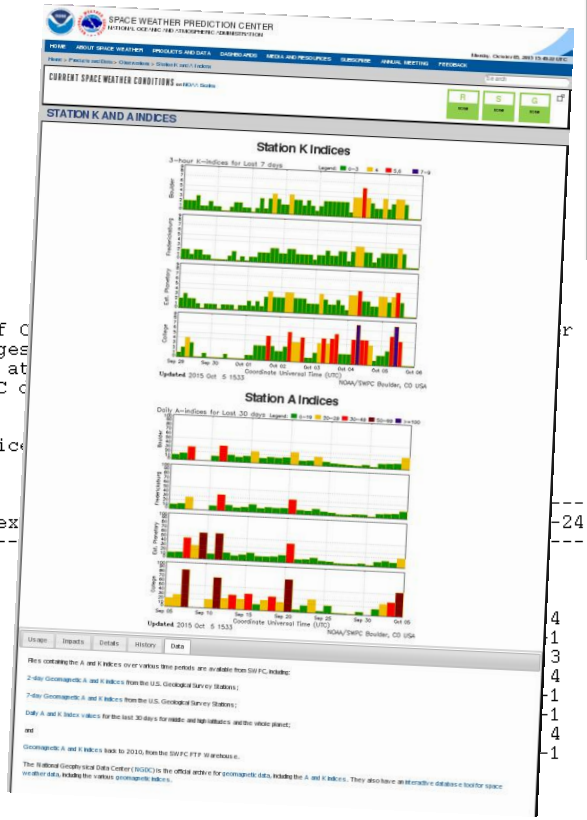
SWPC's 7day_AK.txt contains:

Data from the past 7 days

3-hourly values of…

- Kp
- 7 station K values

Files are extracted & processed every 3 hours




```
 1 :Product: Geomagnetic Data
 2 :Issued: 0933 UTC 04 Sep 2015
 3 #
 4 # Prepared by the U.S. Dept. of C
 5 # Please send comments and sugges
 6 # Updated every hour beginning at
 7 # Values shown as reported, SEC c
 8 # Missing Data: -1
 9 #
10 #      Geomagnetic A and K indic
11 #
12 #                Geomagnetic
13 #                  Dipole     A
14 # Station        Lat Long   Index
15 #--------------------------------
16
17 2015 Aug 28
18
19 Boulder           N49 W 42    32
20 Chambon-la-foret  N-- E---    -1
21 College           N65 W102    57
22 Fredericksburg    N38 W 78    28
23 Kergulen Island   S57 E130    -1
24 Learmonth         S22 E114    -1
25 Planetary(estimated Ap)       43
26 Wingst            N54 E 95    -1
27
28
29 2015 Aug 29
30
31 Boulder           N49 W 42    15     4     2     2     4     2     4     2     2
32 Chambon-la-foret  N-- E---    17    -1    -1     2     3     2     3     3     2
33 College           N65 W102    20     3     3     3     6     3     2     2     2
34 Fredericksburg    N38 W 78    13     4     2     2     4     3     2     2     2
35 Kergulen Island   S57 E130    -1    -1    -1    -1    -1    -1    -1    -1    -1
36 Learmonth         S22 E114    -1    -1    -1    -1    -1    -1    -1    -1    -1
37 Planetary(estimated Ap)       16     5     3     2     4     2     3     2     2
38 Wingst            N54 E 95    -1    -1    -1    -1    -1    -1    -1    -1    -1
39
40
41 2015 Aug 30
42
43 Boulder           N49 W 42     6     2     1     2     1     2     2     1     2
44 Chambon-la-foret  N-- E---     8     1     1     1     1     1     1     2     1
45 College           N65 W102     4     2     2     1     1     1     1     1     1
46 Fredericksburg    N38 W 78     5     1     1     2     2     2     2     1     1
47 Kergulen Island   S57 E130    -1    -1    -1    -1    -1    -1    -1    -1    -1
48 Learmonth         S22 E114    -1    -1    -1    -1    -1    -1    -1    -1    -1
49 Planetary(estimated Ap)       5     2     1     2     1     1     1     1     1
50 Wingst            N54 E 95    -1    -1    -1    -1    -1    -1    -1    -1    -1
51
52
53 2015 Aug 31
54
55 Boulder           N49 W 42     7     2     1     1     1     2     3     2     2
56 Chambon-la-foret  N-- E---    11     1     1     1     2     2     1     2     2
57 College           N65 W102     2     2     1     0     0     0     1     1     1
58 Fredericksburg    N38 W 78     5     1     1     1     2     2     1     2     2
59 Kergulen Island   S57 E130    -1    -1    -1    -1    -1    -1    -1    -1    -1
60 Learmonth         S22 E114    -1    -1    -1    -1    -1    -1    -1    -1    -1
```

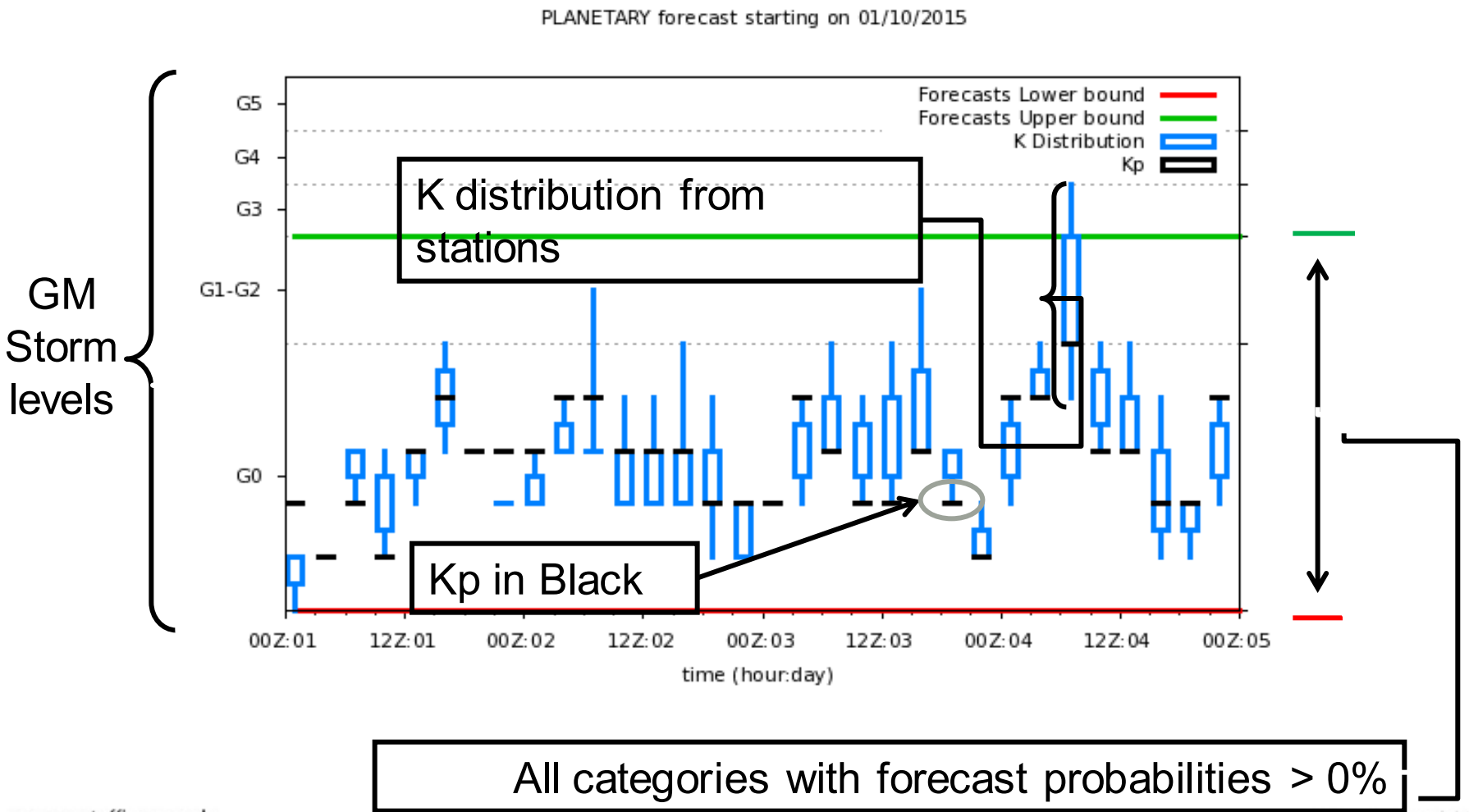
Verification of Kp

Probabilities are cumulative
Probability ≥G0 is always 100%
Min probability = 1%

The probability density function gives the probability each category will occur

Insignificant	G0	Y	100 85	100 70	100 70	100 70
Geo-Magnetic Storm Probability (Exceedance)	Level	Past 24 Hours (Yes/No)	Day 1 (00-24 UTC) (%)	Day 2 (00-24 UTC) (%)	Day 3 (00-24 UTC) (%)	Day 4 (00-24 UTC) (%)
Minor or Moderate	G1 to G2	N	15 15	30 30	30 30	30 30
Strong	G3	N	1 0	1 0	1 0	1 0
Severe	G4	N	1 0	1 0	1 0	1 0
Extreme	G5	N	1 0	1 0	1 0	1 0

Distribution of K observations and Kp from 1-4 Oct 2015.



RPS calculated for forecast on 1 Oct. ‘15

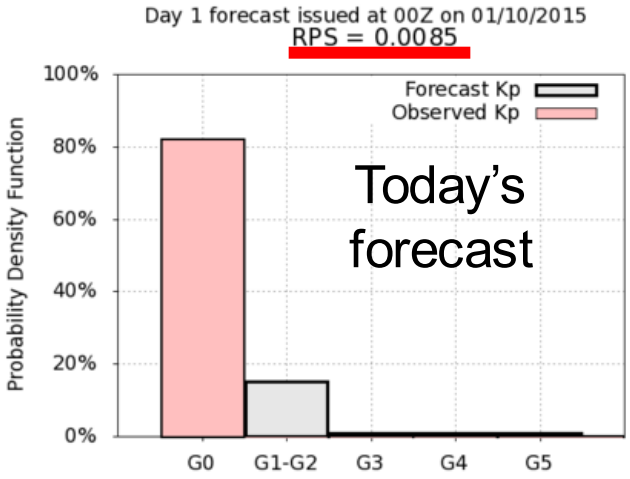
Probability density function Maximum daily Kp value

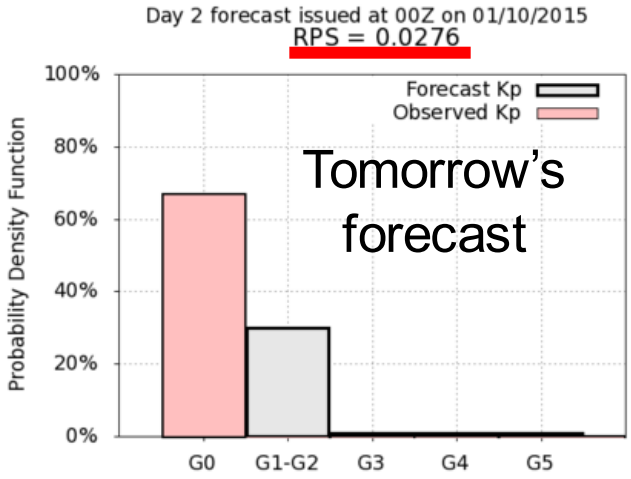
Day 1 RPS=0.01
Day 2 RPS=0.03
Day 3 RPS=0.03
Day 4 RPS=0.11

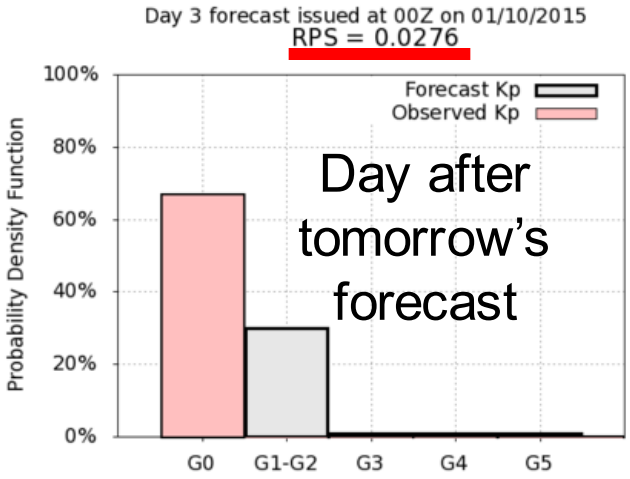
This particular forecast looks good

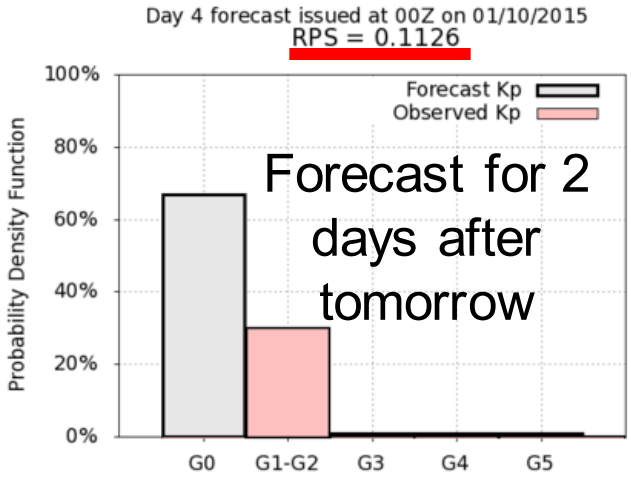
BUT

what is good?









Kp forecast v climatology