NWRA Flare-Forecasting Comparison Workshops: Goals, Participants, and Methodology

KD Leka (presenting), with the International Flare Forecasting Comparions Team (see below)

ABSTRACT

Two workshops have been held recently, in 2009 and 2013, to begin systematic comparisons of methods for forecasting solar flares. The first, also known as the "All-Clear Forecasting Workshop" was held jointly with NASA/Space Radiation Analysis Group and NOAA/Space Weather Prediction Center, had a focus on predicting "All-Clear periods" from the standpoint of major flares and solar energetic particle events. The second more recent workshop, held at NorthWest Research Associates in Boulder, CO, USA, focused on using and exploiting the recent data from the NASA/Solar Dynamics Observatory (SDO) Helioseismic and Magnetic Imager (HMI), particularly the vector magnetic field time-series now available. For both workshops, many researchers active in flare-forecasting research participated, and diverse methods were represented in terms of both how the methods characterize the Sun, and the statistical approaches used to create a forecast. We describe here the goals of both workshops, the participating methods, and the approaches developed for allowing standardized, testable comparisons between methods. Funding for the workshops and the data analysis was provided by NASA/Living with a Star contract NNH09CE72C and NASA/Guest Investigator contract NNH12CG10C.

Forecasting the ``All Clear"

hosted by
National Center for Atmospheric Research, NASA/Space Radiation Analysis Group, NOAA Space Weather Prediction Center, and NorthWest Research Associates
Boulder, CO 22—24 April 2009

Researchers:
KD Leka, G Barnes (NWRA)
D Falconer (NASA/MSFC)
M Wheatland (U. Sydney)
R T J McAteer, P Higgins, P Conlon, S Bloomfield (Trinity)
R Qahwaji, T Colak, O Wahmed (U. Bradford)
H Wang, V Abramenko, Y Yuan (NJIT/BBSO)
M Georgoulis (JHUAPL)
K Schrijver (LMSAL)
J Zhang (GMU)
D Alexander (Rice U.)
P Shea, D Smart (SSSRC)
B Welsch (U. Berkeley)
A Tylka (NRL)
C Fry (EXPI)
C St. Cyr, M Hesse (NASA/GSFC)
S Johnson (NASA/SRAG)
D Markiewicz, A Raphael (Max Planck)

Operations/End Users
C Balch, J Kunches, W Murtagh (NOAA/SWPC)
R van der Linden (SWENET/ESA)
M Stills (United Airlines)
B Jones (SolarMetrics)
J Sanders (AFWA)
R Turner (NASA/HQ)
D Fry, N Zapp, K I Lee, N Stoffle (NASA/SRAG)

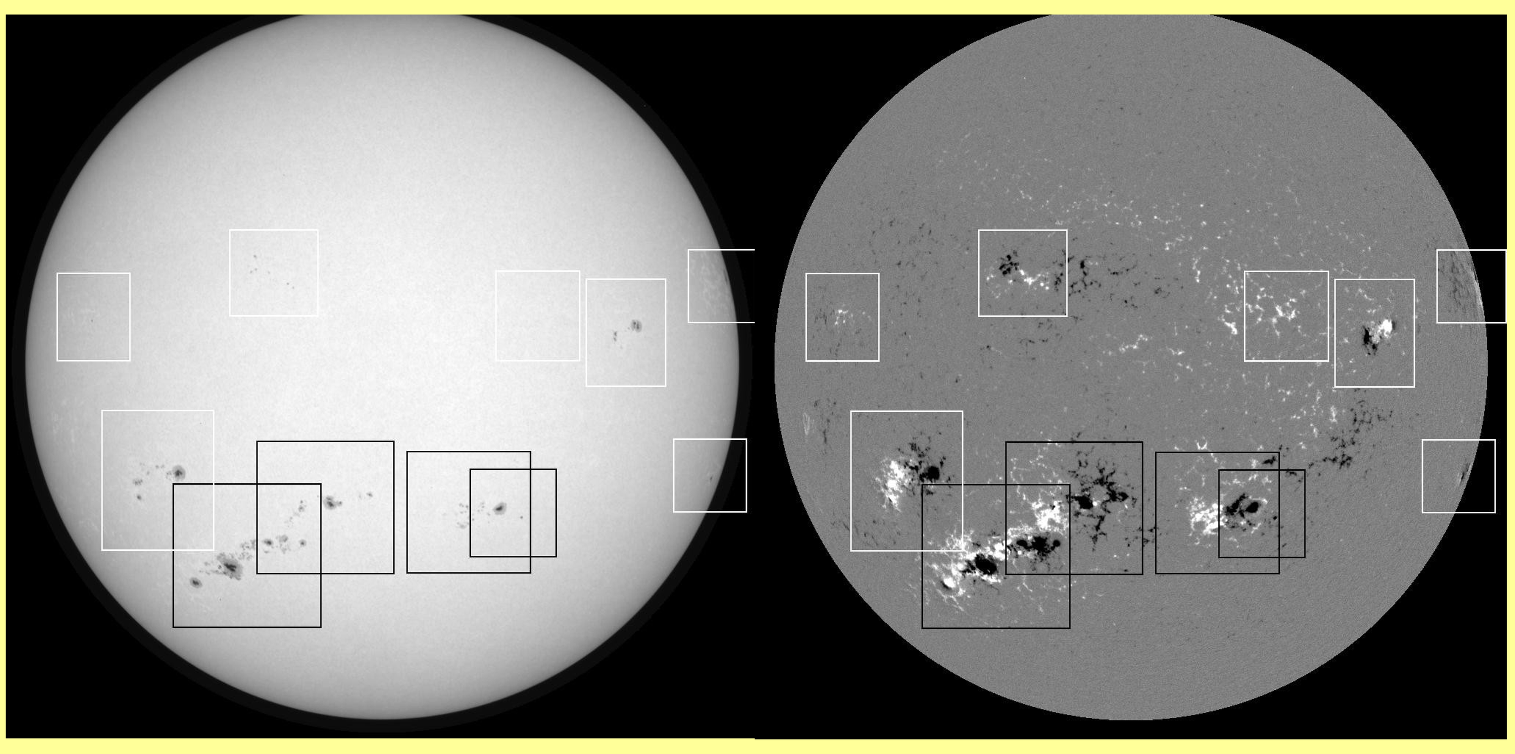
GOAL

To provide a much-needed opportunity for researchers and operational users to convene and discuss issues regarding the state of the science of solar event predictions important for space and terrestrial applications.
Additionally, the workshop served as a much-needed opportunity for the research community to understand operational forecasting needs and how ongoing research can support real-time operations. The focus was determining the length of time for which we can know that an operationally limiting CME/Flare/SEP event will NOT happen: i.e. an operational "All Clear Forecast". The following key points were discussed:

- Needed forecasting validity windows
- Needed forecast latency
- Model maturity: forecasting output, benchmarks.
- Transition to operational tools: required model input, model run time, complexity and software infrastructure.

DATA: Forecasting the All Clear

6 years (2000 – 2005 inclusive) of daily, line-of-sight magnetograms and intensity images from SoHO/MDI (below, example from 09 September 2001). Active Regions were "extracted", and prepared as FITS files. A method was developed to combine near-by active regions (black) with limits as to how many can be grouped as such to keep the data sets tenable. All numbered NOAA active regions were thus extracted and presented, with no further pre-selection according to size, location on the disk, or flaring history.



EVENT LISTS: Forecasting the All Clear

- C1.0 or greater flare within 24 hr after the observation
- M1.0 or greater flare within 12 hr after the observation
- M5.0 or greater flare within 12 hr after the observation

Table 12. All vs. Minimum-Common Dataset(s) Sample Sizes

Event List	"All Data" Flare	"All Data" No Flare	"All Data" Event Rate	Minimum Dataset #1 Flare	Minimum Dataset #1 No Flare	Minimum Dataset #1 Event Rate	Minimum Dataset #2 Flare	Minimum Dataset #2 No Flare	Minimum Dataset #2 Event Rate
C1.0+/24hr	2609	10356	0.201	879	4123	0.176	268	164	0.620
M1.0+/12hr	400	12565	0.031	103	2899	0.031	69	217	0.241
M5.0+/12hr	93	12872	0.007	26	3596	0.007	21	254	0.076

3 "common datasets" used:
All: 12965 as presented.
"Minimum Common Data Set #1": only those data for which all algorithms using the MDI data were run (hence excluding data for some methods did not perform due to observing angle, region size, etc).
"Minimum Common Data Set #2": MCD#1 plus restrictions of the Event Statistics Method, which required prior flare events to make a forecast. Sample sizes vary accordingly.

The Second Flare-Forecasting Comparison Workshop

hosted by
NorthWest Research Associates
Boulder, CO, 2—4 April 2013

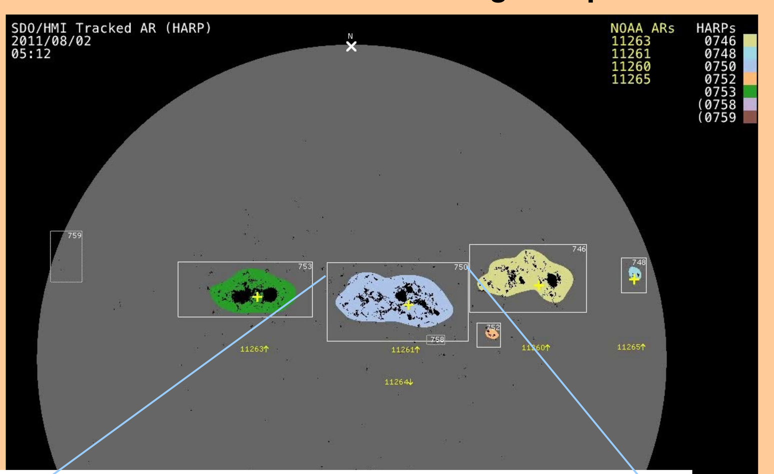
KD Leka, G Barnes (NWRA)
V Abramenko (Big Bear Solar Observatory)
C Balch (NOAA/SWPC)
S Bloomfield (Trinity College, Dublin)
D Falconer (MSFC/University of Alabama/Huntsville)
M Georgoulis (Academy of Athens)
X Huang (National Astronomical Observatories, Chinese Academy of Sciences)
P Higgins (Trinity College, Dublin; now LMSAL)
B Welsch, M Kazachenko (University of California, Berkeley)
M Alomari (ASU/Jordan, formerly University of Bradford, UK)
R T J McAteer (New Mexico State University)
S Wang (New Jersey Institute of Technology)

GOAL

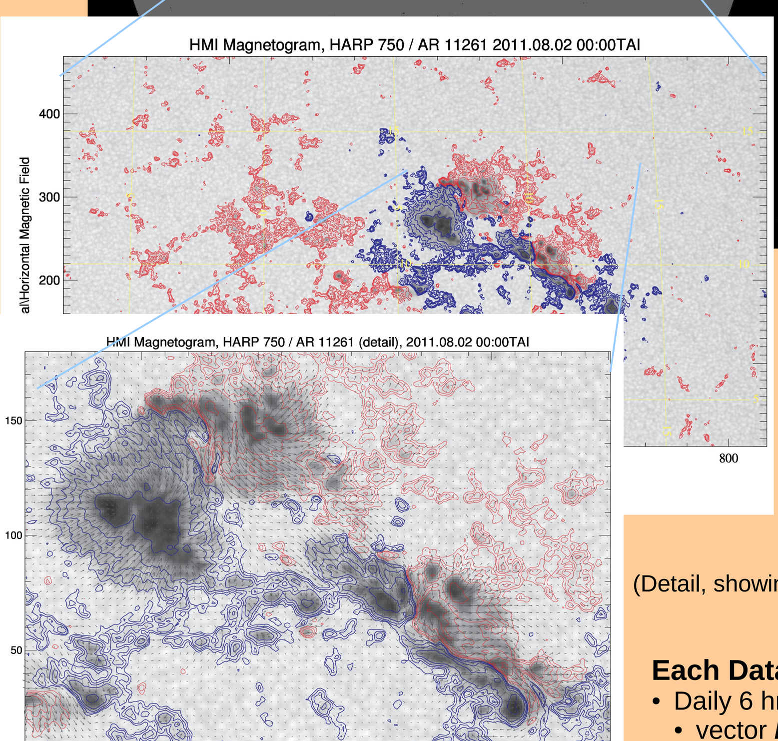
To provide a forum for frank and open comparison of the state-of-the art of solar flare forecasting methods, to share ideas for improving both the science and the validation of methods under development, and develop a framework within which established and new methods can be tested and understood – all so as to improve the understanding and effectiveness of solar flare forecast research. In this second workshop, only researchers gathered (plus a NOAA/SWPC representative and validation/verification expert).
The following key points were discussed:

- Refinements to the common dataset.
- Training sets vs. Testing sets.
- Skill Scores, sample sizes, and computing the relevant uncertainties.

DATA: Second Flare-Forecasting Comparison Workshop



One year of vector magnetograms from SDO/HMI (2011.08 – 2012.07). The HMI Active Region Patches (HARPs) were used, without regard to size, flaring history, *or association with NOAA active regions.* In some cases, HARPs had no NOAA number, in some cases a single HARP represented multiple NOAA active regions.



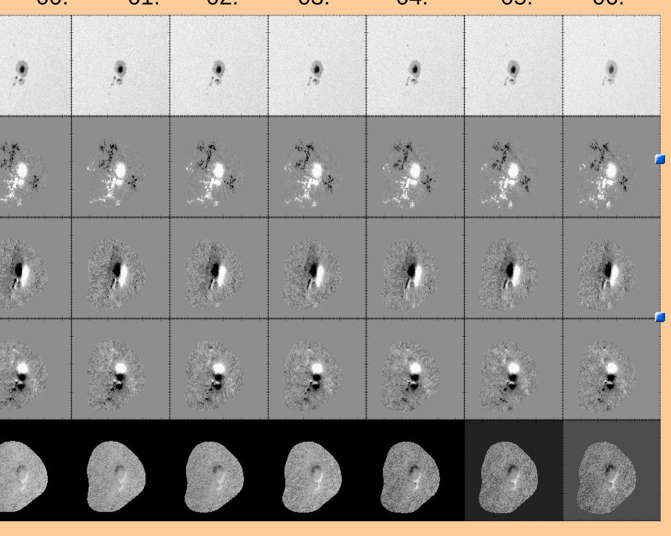
(A HARP region example, NOAA AR 11261, from HMI. Red/Blue are positive/negative polarity of the B_r magnetic field component.)

(Detail, showing vectors at every 3rd point.)

Each Data Set:

- Daily 6 hr time-series of HMI data:
  - vector ***B***, and uncertainties.
  - spherical components provided
  - Doppler velocity from inversion
  - Intensity image.
  - All data presented on the image grid.
- > 3,300 datasets presented.

Ref. Time (hr) -6, -5, -4, -3, -2, -1, 0
TAI time 00:, 01:, 02:, 03:, 04:, 05:, 06:



EVENT LISTS: FFC2

- C1.0 or greater flare within 24 hr after time-series end.
  - 484 flares, 2855 quiet. Event rate: 0.145.
- M1.0 or greater flare within 24 hr after time-series end.
  - 93 flares, 3246 quiet. Event rate: 0.028.

METHODS
(1): used in Forecasting All Clear Workshop
(2): used in Second Flare Forecasting Workshop

Magnetic Energy Spectrum (2)
V. Abramenko, Y. Yurchyshyn

- Parametrize character of spectrum of energy input vs. energy dissipation.
- Non-Flaring ARs display Kolmogorov-type spectrum, Flaring ARs have steeper spectrum.
- Integral of the Energy Spectrum over wavenumber also correlates with flaring history.
- Forecasts: Discriminant Analysis.

MultiFractal Spectrum, Generalized Correlation Dimensions (1,2)
R. T. James McAteer

- Parametrizes complexity of magnetic field based on spatial gradients.
- Spectral Index has been shown to steepen prior to flaring episodes.
- Forecasts: Discriminant Analysis.

Effective Connected Magnetic Field Strength B_{eff} (1,2)
M. Georgoulis

- Determine minimum-length connectivity of magnetic sources.
- Compute distance-normalized total flux in connections. Emphasizes short (e.g., neutral-line) connections.
- Forecasts: probability of events as f(threshold) fit by a sigmoid function.
- Initial evaluation: works better than Total Flux.

$$B_{eff} = \sum_{i=1}^{p}\sum_{j=1}^{n}\frac{\Phi_{ij}}{L_{ij}^2}$$

$$P_{flare}(T) = A_2 + \frac{A_1 - A_2}{1+\exp\left(\frac{\log(T_c/T_0)}{W}\right)}$$

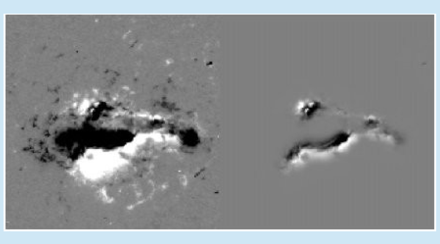
Machine Learning Techniques (1)
BBSO/NJIT -- Y. Yuan

- Three parameters describe active region:
  - total unsigned magnetic flux,
  - length of the strong-gradient neutral line,
  - total magnetic energy dissipation
- Forecasts: Ordinal logistic regression and supporting vector machines.

Space Weather Research Lab/NJIT (2)
J Jing, H Wang, S Wang

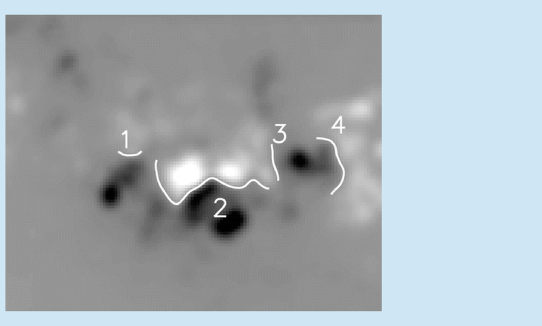
- Time-series of B_{los} or B_r used to compute helicity injection rate and Total magnetic flux.
- Forecasts: Machine Learning Techniques.

Magnetic Flux Near High-Gradient Polarity Inversion Lines (1,2)
K. Schrijver
Parameter "R" is a proxy for the emergence of current-carrying magnetic flux.
Computed from the line-of-sight (or radial) magnetic field maps:
• Dilate bitmaps of the magnetograms where the positive or negative flux density exceeds a threshold (150 Mx cm−2)
• Define high-gradient polarity-separation lines as areas where the bitmaps overlap.
• Convolve the resulting high-gradient polarity-separation line bitmap with a Gaussian to obtain a "weighting map"
• Obtain the parameter *R* by multiplying the weighting map by the unsigned field (B_{los} or B_r) and finding the total.
• Forecasts: R, log(R) fed into Discriminant Analysis.



Event Statistics (1, hopefully 2)
M. Wheatland

- Hypothesis: solar flares obey a power-law frequency/size distribution. Model flare rate as a Poisson process: $P(\tau) = \lambda \exp(-\lambda\tau)$.
- Baysian statistical method to predict flaring probability for different flare sizes according to the flaring history of the observed active regions.
- Since uses only flare history as input, results serve as "baseline" for magnetogram-data methods.



Total Non-Potentiality of Active Regions (1,2)
D Falconer

- Construct a proxy for the total free magnetic energy: $^{L}WL_{SG2} = \int(\nabla B_{LOS})^2\,dl$.
- Combine $^{L}WL_{SG2}$ with total Flux.
- Forecast: power-law fit produces event rate.

NWRA Magnetic Parametrization (1,2)
KD Leka, G Barnes

- Computed Quantities:
  - Magnetic Field strength, direction††
  - Horizontal Gradients of field vector††
  - Vertical current density**
  - Magnetic twist parameter(s)**
  - Current helicity density**
  - Magnetic free energy proxy**
  - Schrijver's R parameter††
- Parametrization:
  - 4th order moment analysis, plus total, net, best-fits.
  - Time-series data: linear slope and intercept at *t=0*
- Forecasts: Discriminant Analysis.

: only for vector ***B*
††: computed from B_{los} and for B_z from a potential-field extrapolation, too.

Estimating Photospheric Electric Fields and Poynting Flux for Flare Forecasting (2)
M Kazachenko, B Welsch, G Fisher

- Poloidal Toroidal Decomposition: derive ***E*** from observed evolution of ***B***: $d\mathbf{B}/dt = c\,curl(\mathbf{E})$.
- Compute parameters based on:
  - magnetic energy flux $dU/dt = 1/4\pi \int (\mathbf{E} \times \mathbf{B})_z\, dA$
  - helicity flux $dH/dt = 1/4\pi \int (\mathbf{E} \times \mathbf{A})_z\, dA$
  - Poynting Flux: $S_z = c/4\pi\,(E_xB_y - E_yB_x)$
- Forecasts: Discriminant Analysis.

Solar Monitor Active Region Tracking (1,2)
P Higgins, S Bloomfield

- Automated Active Region detection
  - For All-Clear and FFC2, used provided data.
- Parametrize active regions:
  - Total, Net flux.
  - Length and orientation of polarity inversion lines.
  - Schrijver's *R*, Georgoulis' B_{eff}
- Forecasts:
  - Discriminant Analysis,
  - Cascade Correlation Neural Networks

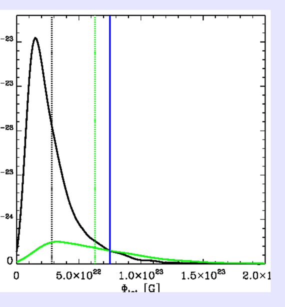
Automated Solar Activity Prediction (ASAP) (1,2)
R Qahwaji, T Colak, M Alomari

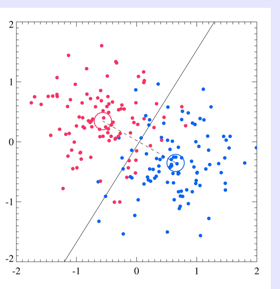
- Inputs:
  - Current sunspot area,
  - 48-hr rate of change of area.
  - Total X-ray flare index
  - McIntosh classification.
- Forecasts: Machine Learning module

Discriminant Analysis (1,2)
G Barnes, KD Leka

- **Statistical Method for Binary Forecasts**
  - Applied for NWRA parameters and many others.
- Using samples of parameters from two known populations →→ estimate the Probability Density Function.
  - Linear/Quadratic Parametric: assumes Gaussian(s).
  - Non-Parametric: Smoothing kernel, estimate PDF directly from data. Better for long-tailed distributions.
- Discriminant Boundary is where PDFs are equal, best separates two populations.
  - New observation is ``classified" as being one/other according to its PD for its observed parameter(s).
- **Extension to Probabilistic Forecasts:**
  - Assume the a priori probability of membership in a population is proportional to sample size(s). Probability of a measurement $x \in$ (flaring population) is:
  - n_j is sample size of [flaring, quiet] data, $f_j(x)$ is probability density function for j, j Є (f,q) for [flaring, quiet] populations.

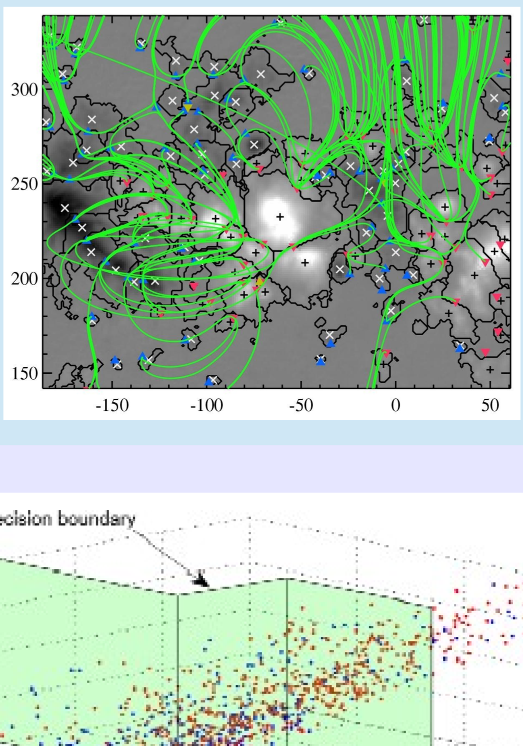
$$P_f(x) = \frac{n_f f_f(x)}{n_f f_f(x) + n_q f_q(x)}$$





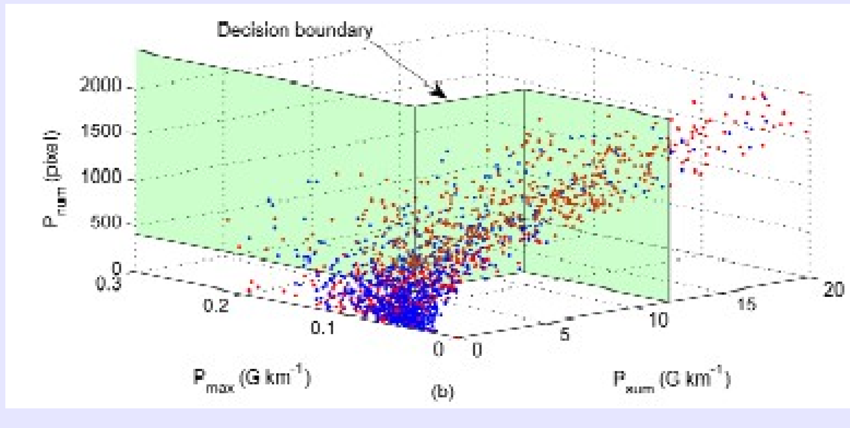
NWRA Magnetic Charge Topology (1,2)
G Barnes, KD Leka

- Use Photospheric ***B*** to investigate coronal ***B:***
  - Partition B_r (or B_{los}) maps, model as point sources
    - Potential-field extrapolation →→ connectivity matrix.
  - Characterize the magnetic connectivity, distribution of magnetic null points, magnetic separator field lines.
    - Parametrize by 4th order moment analysis, plus totals, net, as appropriate.
    - Time-series data: evolution characterized using slope, intercept at *t=0.*
- Forecasts: Discriminant Analysis.



NAO/Chinese Academy of Sciences (2)
X Huang, Huaning Wang

- Characterize Active Region using ***B:***
  - Vertical Current Density
  - Vertical Current Helicity Density
  - Twist parameter α
  - Free Magnetic Energy Density
  - Maximum Horizontal Gradient
  - Length of Neutral Line
  - Number of Singular Points
  - Tilt Angle
  - Magnetic Shear Angle
- Forecasts: Iteration of "Information Gain" training for each parameter to determine optimal location (based on accuracy) of decision boundary.

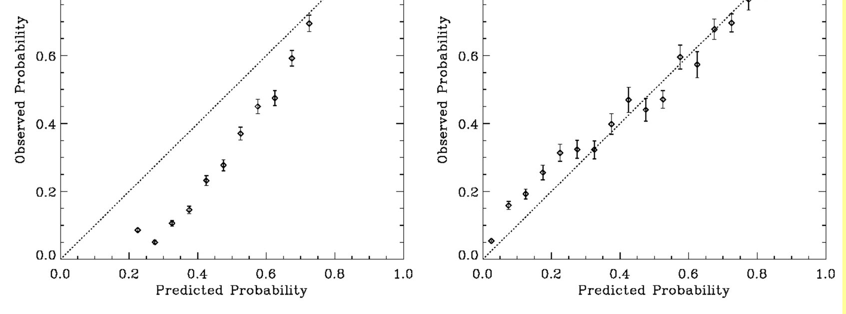


RESULTS: Forecasting the All Clear

Finally being compiled now.

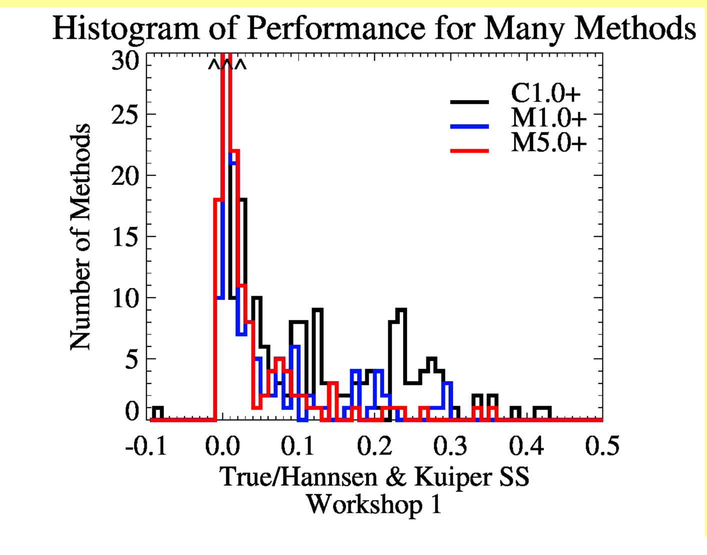
- Quantitative Skill Scores on methods acting on targeted, consistent data sets is crucial.
  - Skill score use (how much information is present vs. a "default" forecast) is now common. ☺
  - Relative *ranking* of methods within a particular SS is fairly consistent. ☺
  - Different Skill Scores emphasize different things, answer different questions. No single skill score tells the whole story. ☹

Example:	Method 1	Method 2	
True/HK SS	0.43	0.42	← **similar**
Brier SS	0.04	0.31	← **different**

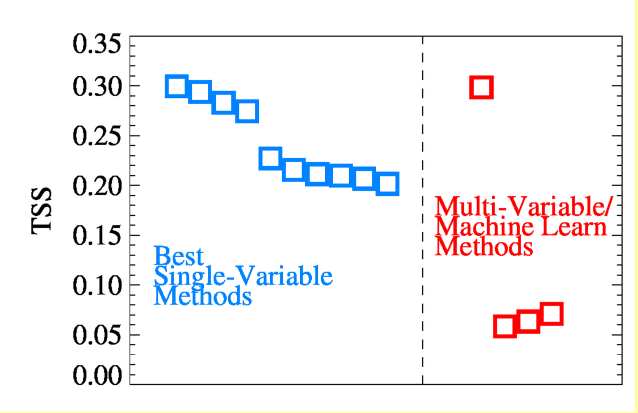


Difference reflected in Reliability plots.

- Many methods perform fairly similarly, scoring 0.2—0.4 on a variety of skill-score tests, even for M5.0+ flare events.



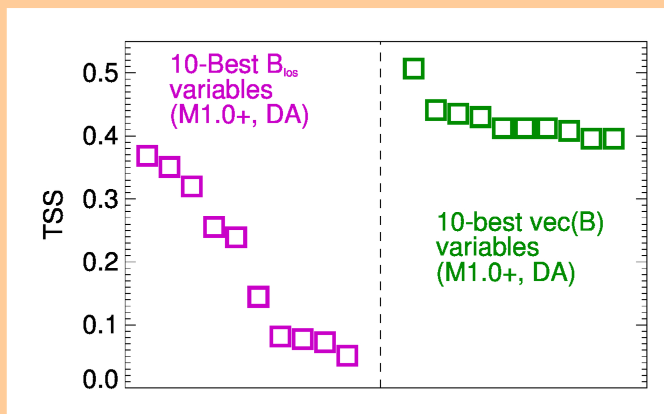
- Multiple parameters + sophisticated computer-learning algorithms do *not* necessarily perform better than single-variable DA/NPDA results.



RESULTS: FFC2

Minimal. Do not have results from most groups yet.

- Confirmed: Vector Magnetic Field data adds information (*c.f.* B_{los} data).



- Overall, somewhat higher Skill Scores coming from FFC2 data and methods.
  - Why?
    - Data?
    - Algorithm improvement?
- *(Find out next year....when we've analyzed it.)*

Conclusions:

- **The group has worked extremely hard to establish infrastructure for this kind of required systematic testing.**
- **We will work over the next year toward more definitive results.**