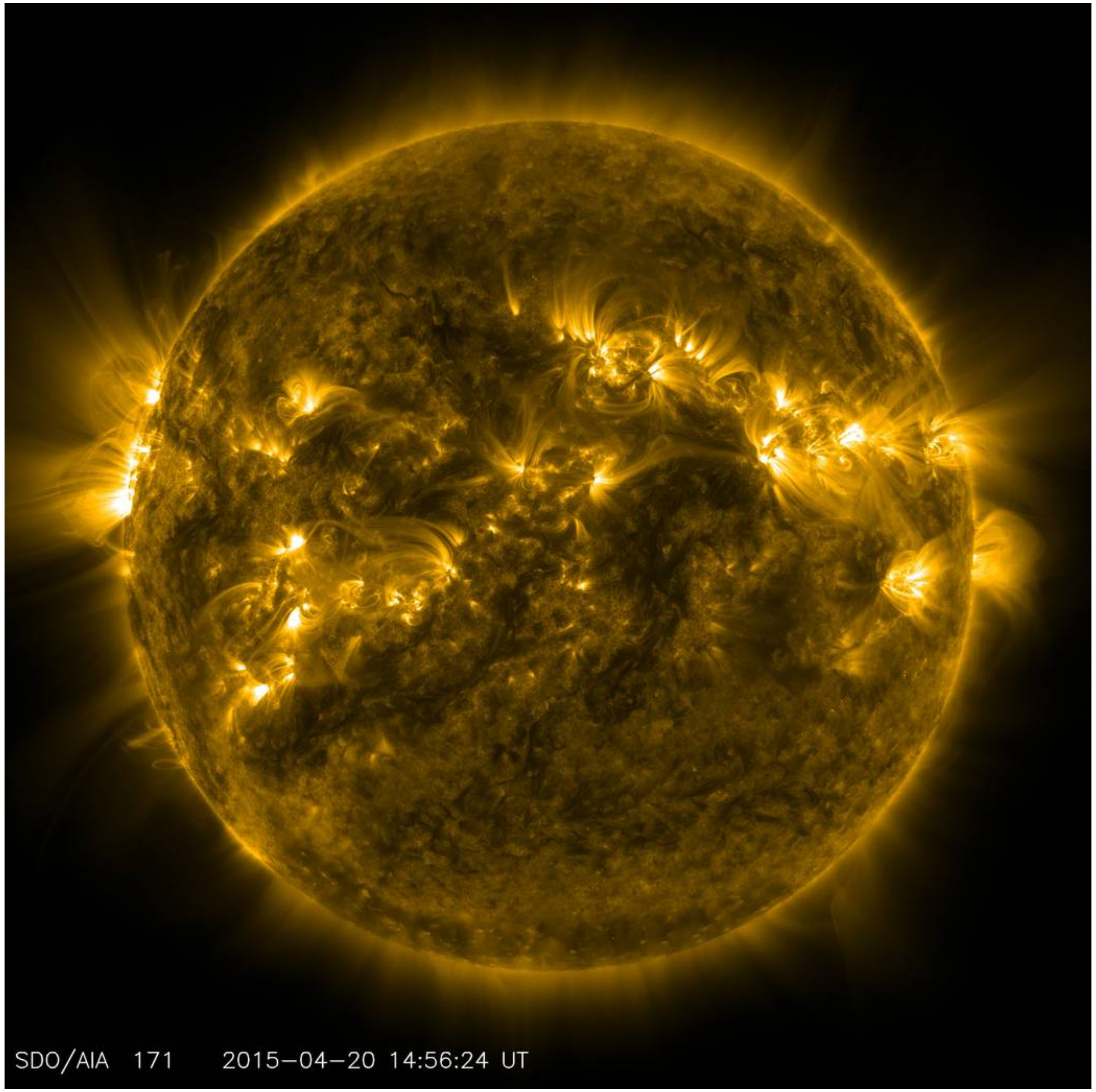
International Study of Earth-affecting Solar Transients 2017 Workshop
International Convention Center JEJU
18 September 2017

Development of a Daily Solar Major Flare Occurrence Probability Model Based on Vector Parameters from SDO/HMI

Daye Lim[1], Yong–Jae Moon[1,2], Jongyeob Park[3], Kangjin Lee[1,4], and Jin–Yi Lee[2]

[1]School of Space Research, Kyung Hee University, Korea
[2]Department of Astronomy & Space Science, Kyung Hee University, Korea
[3]Korea Astronomy and Space Science Institute, Korea
[4]Electronics and Telecommunication Research Institute, Korea

Solar Active Regions

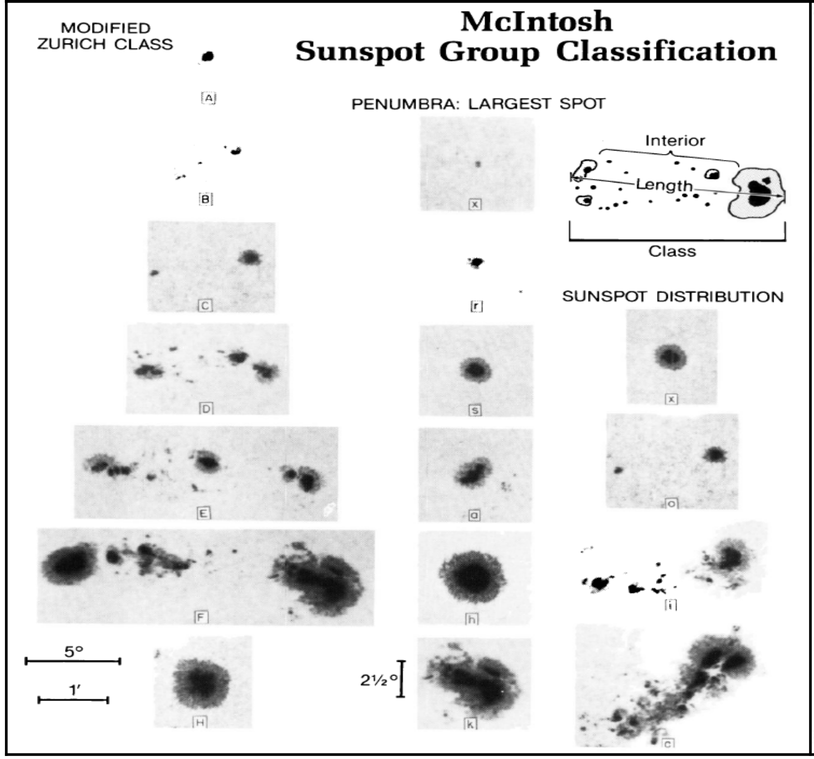


- Solar active regions are areas of intense and complex magnetic field.

- Most solar energetic events such as solar flares blast forth from active regions.

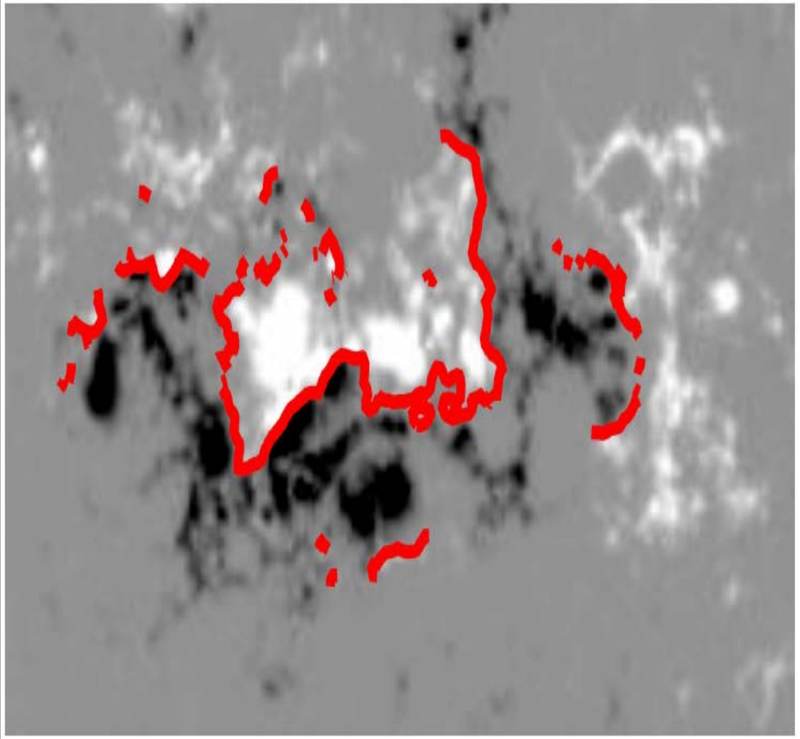
Previous Flare Forecasting Studies Based on Photospheric Magnetic Fields

White light



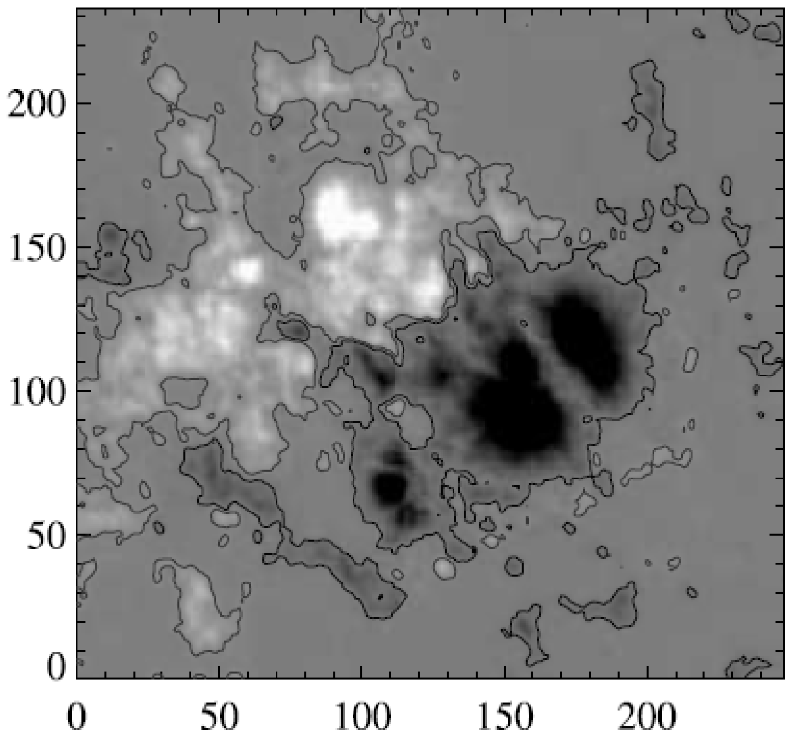
McIntosh (1990)

Line-of-sight magnetic field



Falconer *et al.* (2011)

Vector magnetic field



Leka and Barnes (2003a)

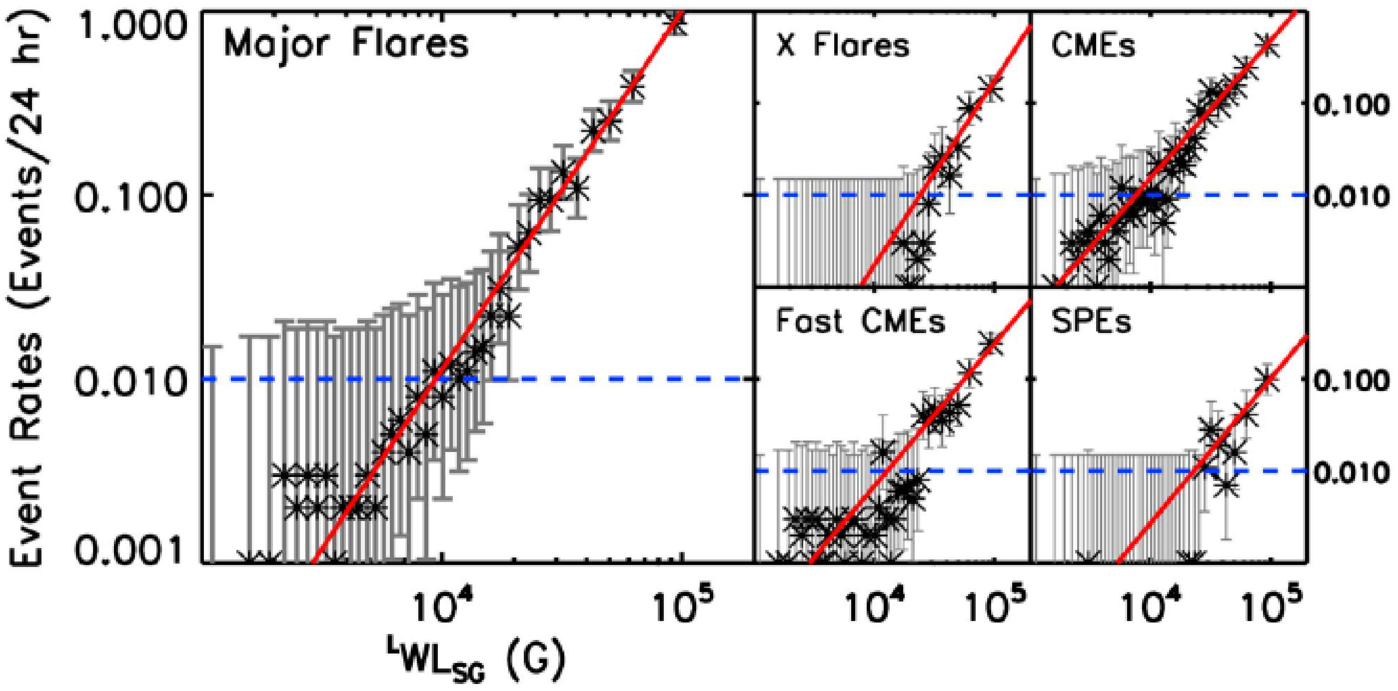
- Observed solar flares well correlate with the **size** and **non-potentiality** of active regions.

Empirical Relationship between Flare Occurrence Rates and Magnetic Parameters

- Proxy of active-region free magnetic energy

$$^{L}\mathrm{WL}_{SG} = \int |\nabla_{\perp} B_{\mathrm{los}}| dl$$

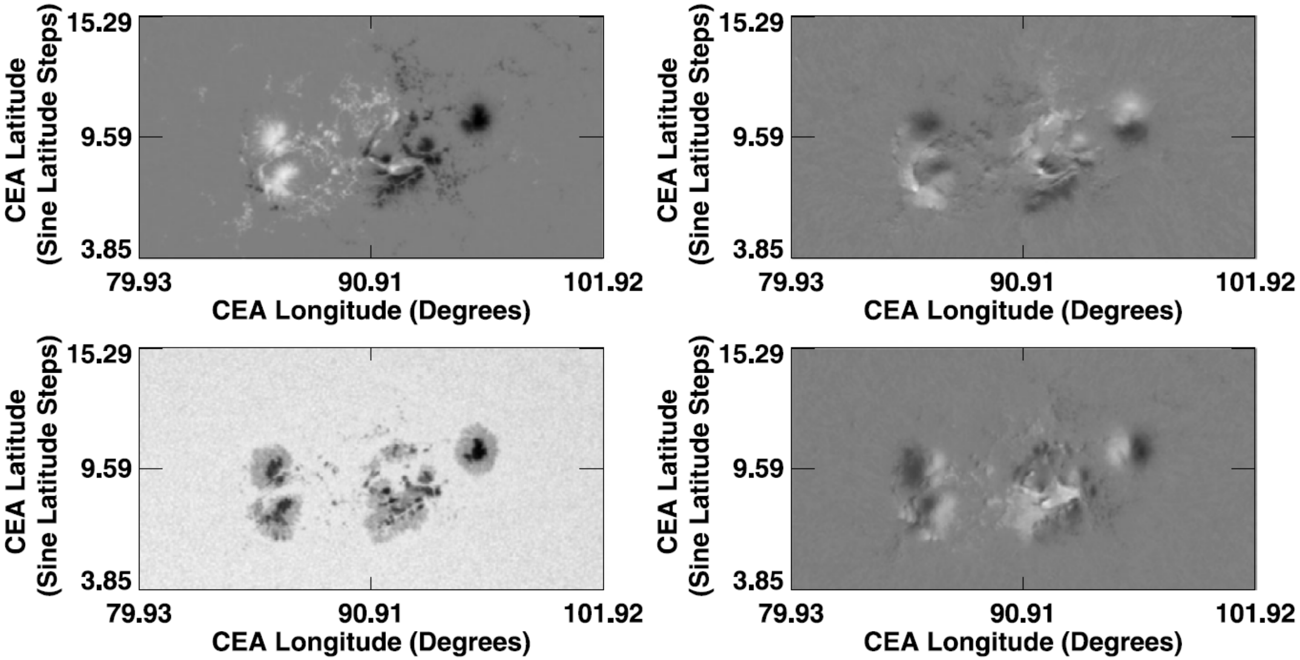
This parameter is the weighted length of the strong-gradient neutral line obtained by integrating the gradients of the **line-of-sight magnetic field** along the neutral lines.



Falconer *et al*. (2011)

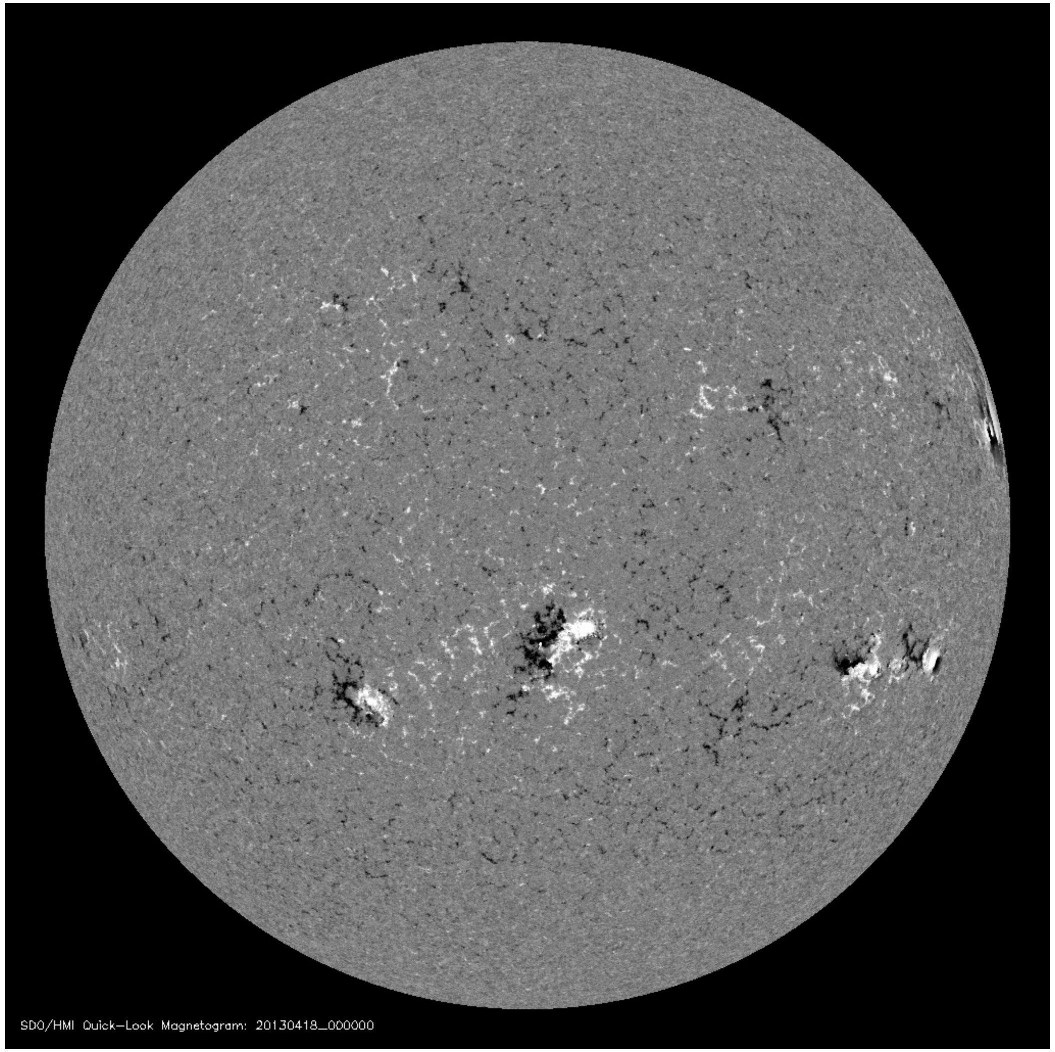
Solar Dynamics Observatory's Helioseismic and Magnetic Imager

B_r B_θ



Continuum B_φ

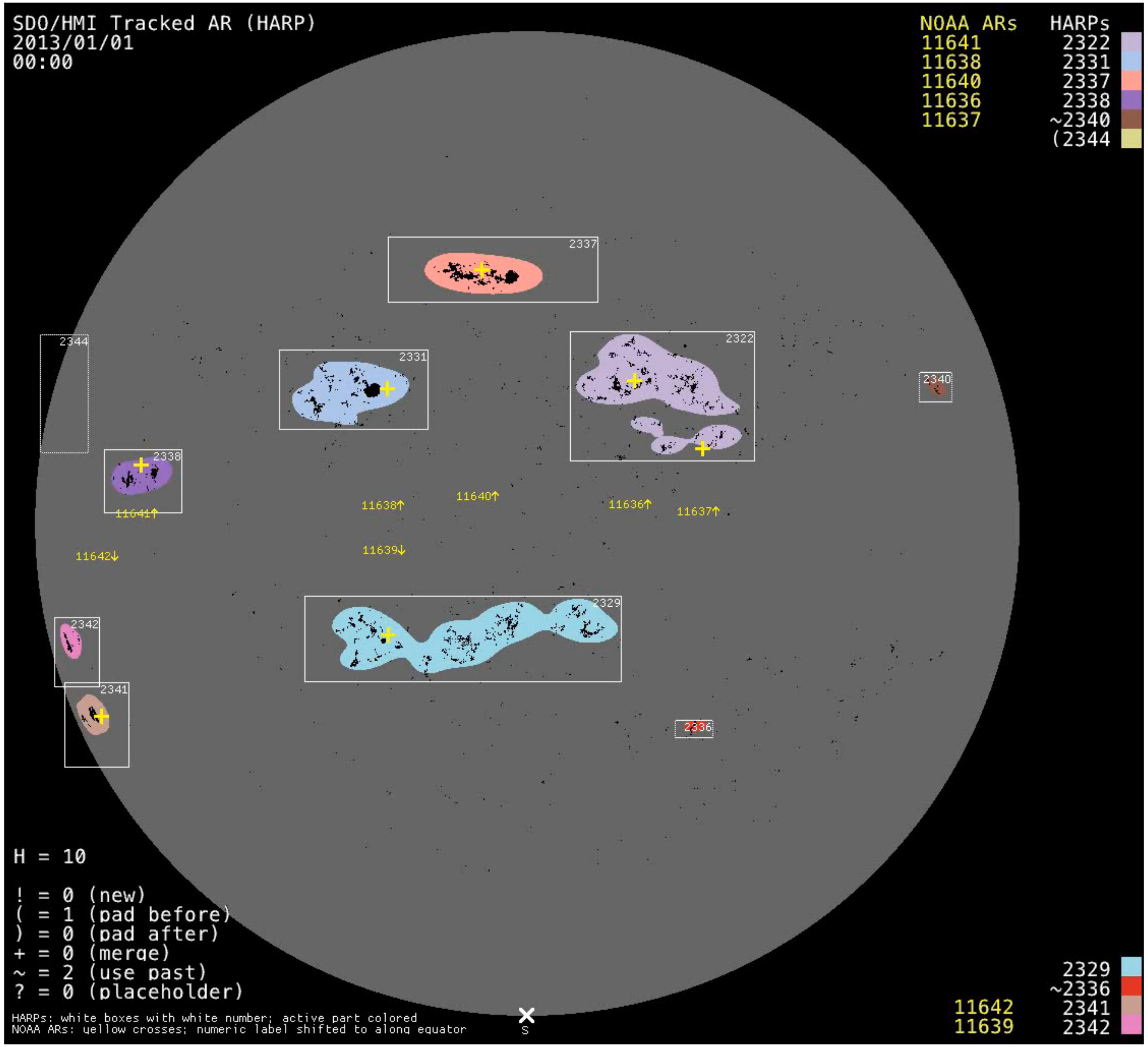
Bobra *et al.* (2014)



Vector magnetic field

Line-of-sight magnetogram

HMI Active Region Patch (HARP) and Space-Weather HMI Active Region Patch (SHARP)



- **HARPs** are automatically identified magnetic structures at the size scale of a solar active region.
- A HARP may include zero, one, or multiple NOAA active regions.
- **SHARPs** provide several parameters that characterize the magnetic field distribution and its deviation from a potential field.
- SHARP parameters are calculated per patch and are available on a twelve-minute cadence.

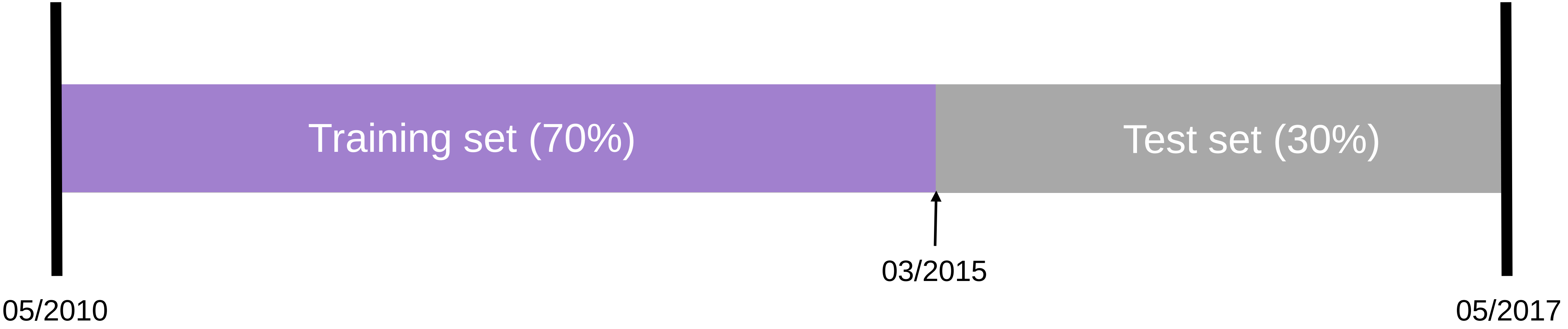
Relationship between Major Flare Occurrence Rates and Vector Magnetic Parameters

- Among the SHARP parameters, we consider **six SHARP vector magnetic parameters** with high F-scores as useful predictors of flaring activity from Bobra and Couvidat (2015).

Keyword	Description	Formula
TOTUSJH	Total unsigned current helicity	$H_{c_{\mathrm{total}}} \propto \sum \lvert B_z J_z \rvert$
TOTPOT	Total photospheric magnetic free energy density	$\rho_{tot} \propto \sum (\boldsymbol{B}^{\mathrm{Obs}} - \boldsymbol{B}^{\mathrm{Pot}})^2 dA$
TOTUSJZ	Total unsigned vertical current	$J_{z_{\mathrm{total}}} = \sum \lvert J_z \rvert dA$
ABSNJZH	Absolute value of the net current helicity	$H_{c_{\mathrm{abs}}} \propto \lvert \sum B_z J_z \rvert$
SAVNCPP	Sum of the net current emanating from each polarity	$J_{z_{\mathrm{sum}}} \propto \left\lvert \sum^{B_z^+} J_z dA \right\rvert + \left\lvert \sum^{B_z^-} J_z dA \right\rvert$
USFLUX	Total unsigned magnetic flux	$\Phi = \sum \lvert B_z \rvert dA$

Data Set

- We use hourly SHARP parameters when longitudes of HARPs are within ± 60 heliographic degrees of disk center from May 2010 to April 2017.

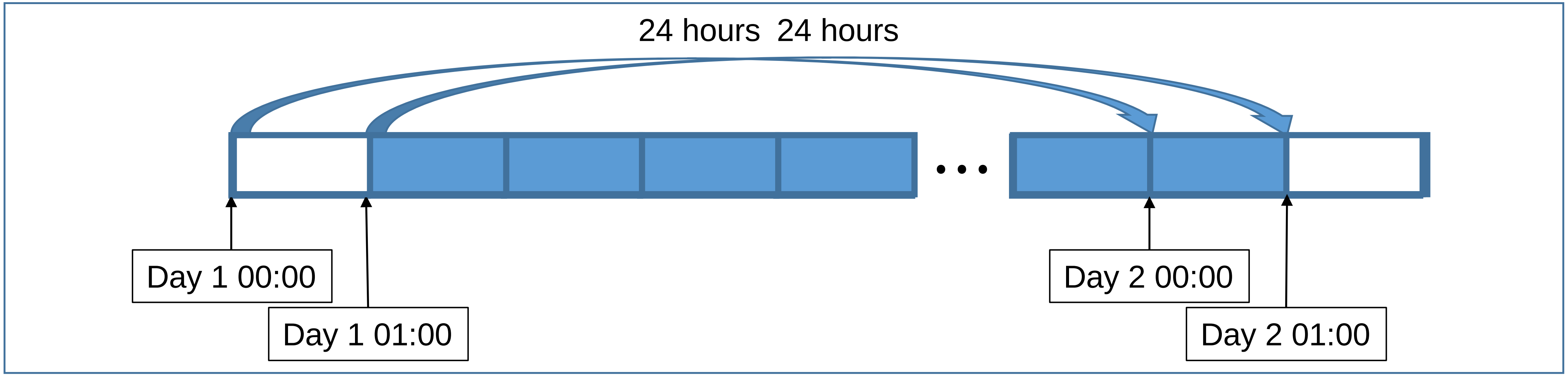


Training data (70%)	Test data (30%)
• From May 2010 to March 2015 • 251431 HARPs • 350 M-class and 23 X-class flares from LMSAL GOES SXR flare list	• From March 2015 to April 2017 • 107757 HARPs • 73 M-class and 1 X-class flares from LMSAL GOES SXR flare list

- Training data are divided into 100 groups having equal number for reasonable statistics per group.

Data Set

- We use hourly SHARP parameters when longitudes of HARPs are within ± 60 heliographic degrees of disk center from May 2010 to April 2017.



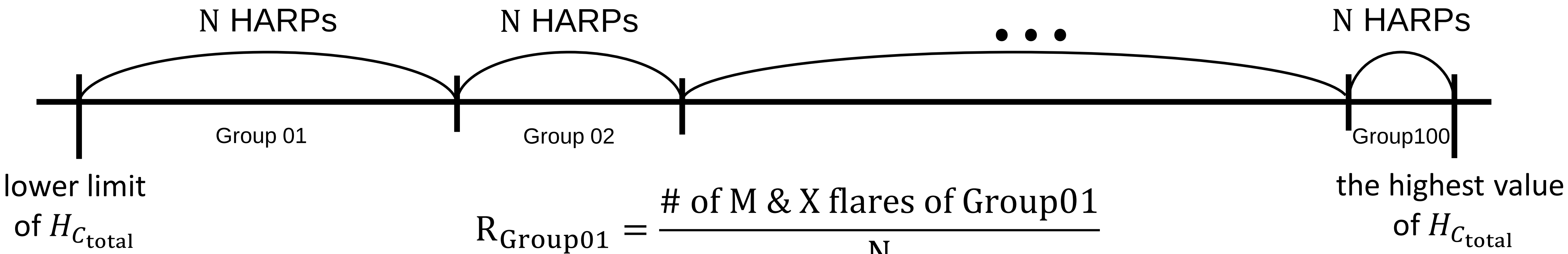
Training data (70%)	Test data (30%)
• From May 2010 to March 2015 • 251431 HARPs • 350 M-class and 23 X-class flares from LMSAL GOES SXR flare list	• From March 2015 to April 2017 • 107757 HARPs • 73 M-class and 1 X-class flares from LMSAL GOES SXR flare list

- Training data are divided into 100 groups having equal number for reasonable statistics per group.

Data Set

- We use hourly SHARP parameters when longitudes of HARPs are within $\pm$ 60 heliographic degrees of disk center from May 2010 to April 2017.

- For **the total unsigned current helicity**,



$$\mathrm{R_{Group01}} = \frac{\#\ \text{of M \& X flares of Group01}}{\mathrm{N}}$$

$$\vdots$$

$$\mathrm{R_{Group100}} = \frac{\#\ \text{of M \& X flares of Group100}}{\mathrm{N}}$$

- Training data for each SHARP parameter are divided into 100 groups having equal number for reasonable statistics per group.

Forecast Validation

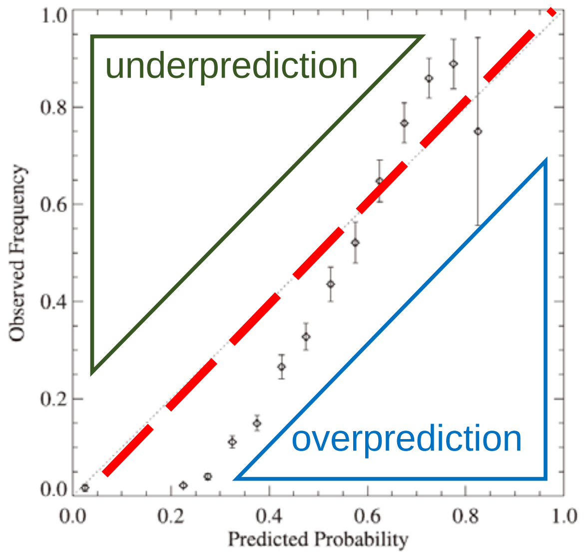
<Probabilistic Forecasts>

- The probability of observing one or more flares in any 24 hr interval is

$$P_\mu(N \geq 1) = 1 - \exp(-\mu),$$

where μ is the average flare rate.

- Reliability plot



Barnes et al. (2016)

A perfect forecasting occurs when all points lie on the diagonal line.

<Yes/No Forecasts>

- Contingency table

	Forecast	
Observed	**Flare**	**No flare**
Yes	True Positives (TP)	False Negatives (FN)
No	False Positives (FP)	True Negatives (TN)

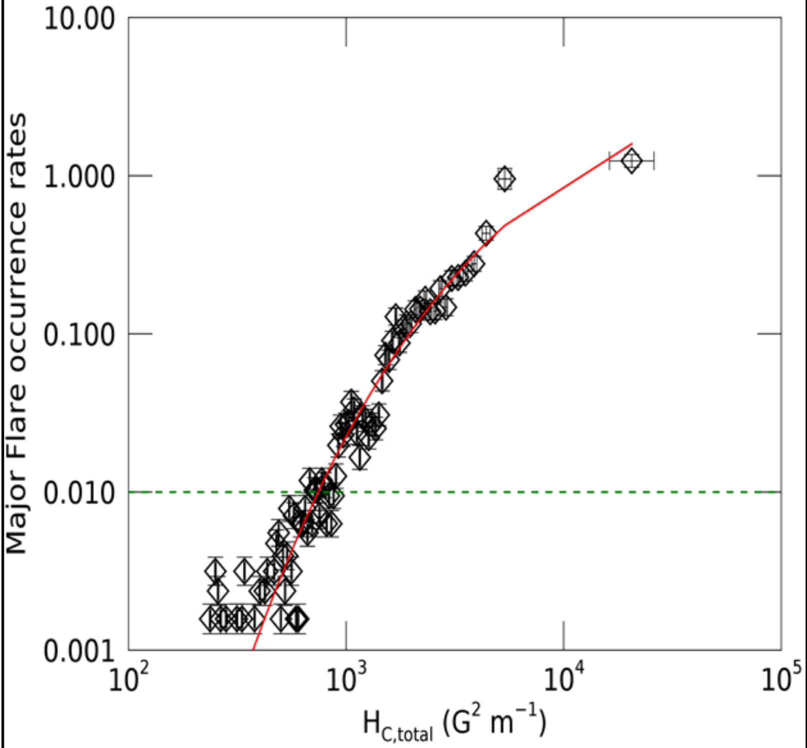
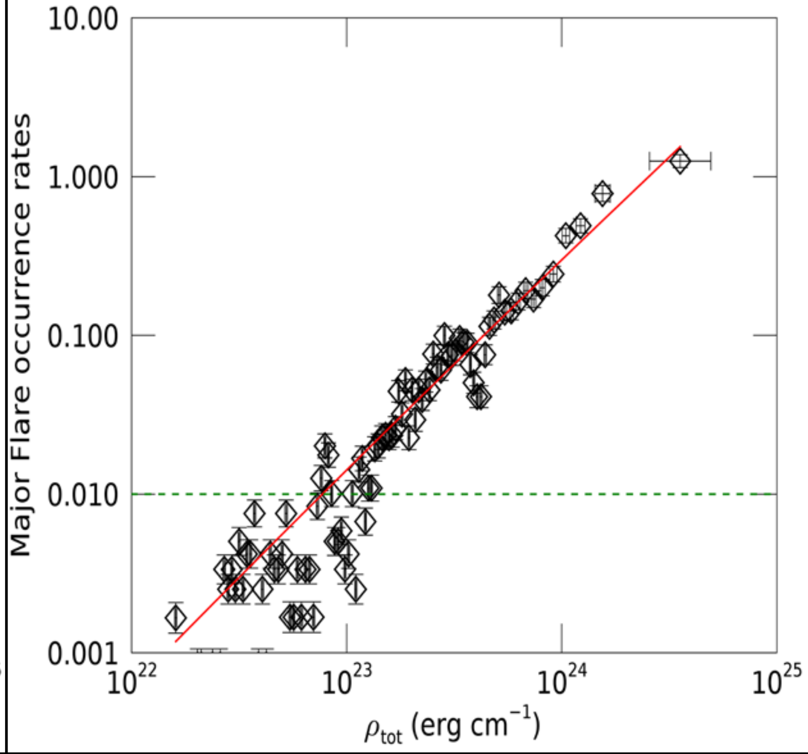
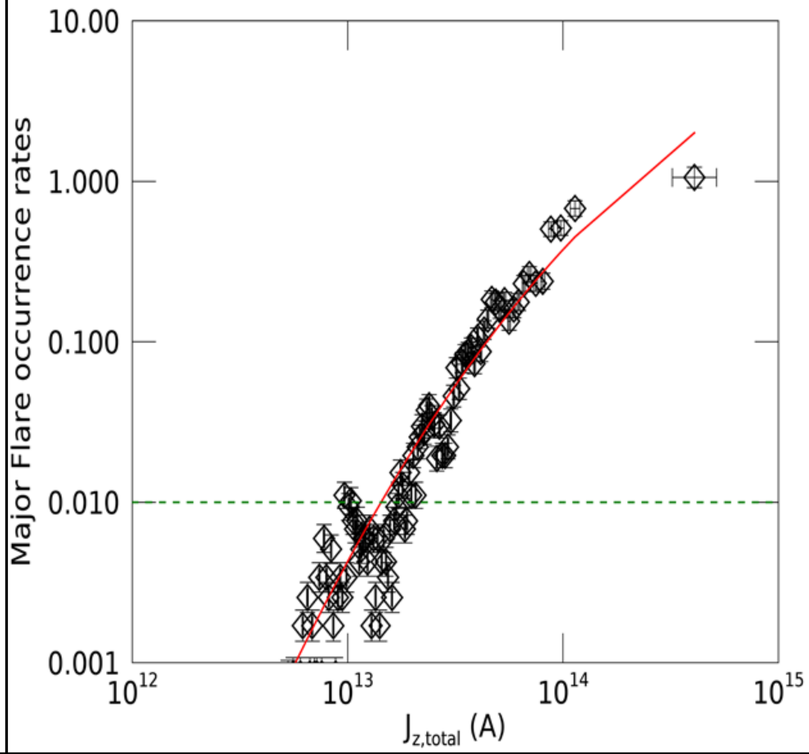
- True skill statistic (TSS)

$$\mathrm{TSS} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} - \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

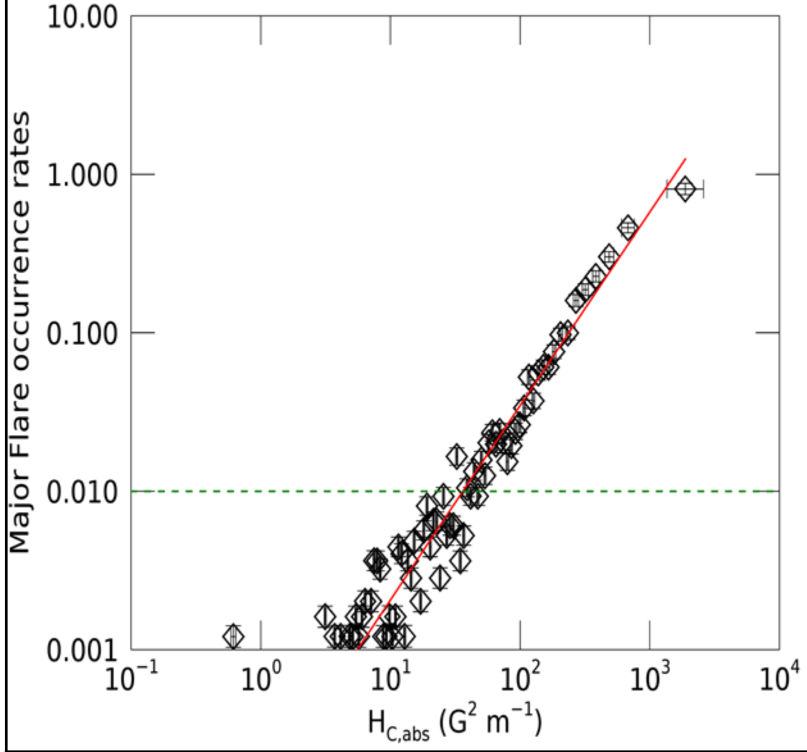
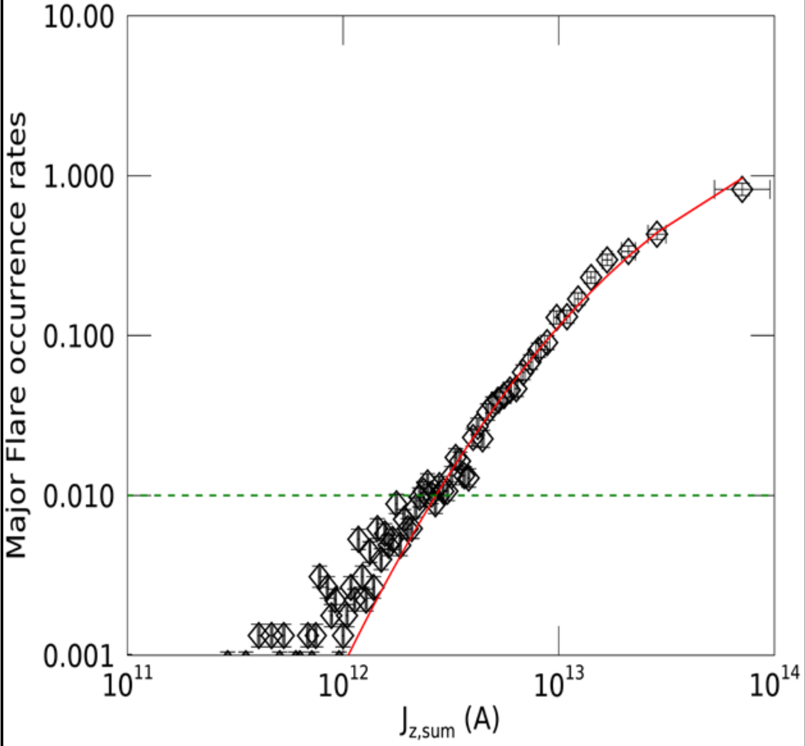
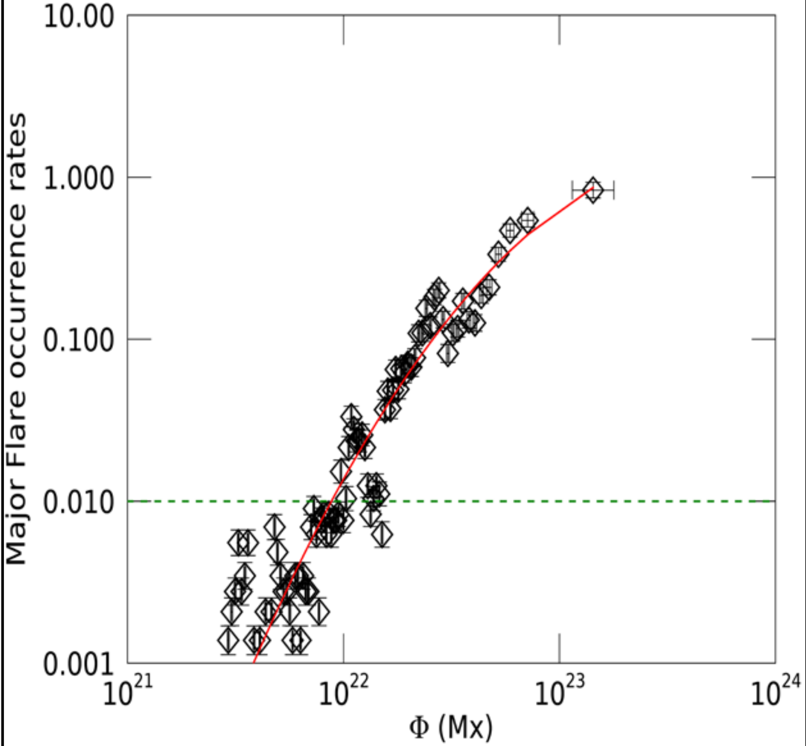
$\mathrm{TSS} = 1$ for perfect forecasts

- Bloomfield *et al*. (2012) found that for the Poisson method, **the best TSS** is typically produced by picking a threshold that depends on the ratio FN/FP, with **FN/FP** $\approx \boldsymbol{N}_{\textbf{event}}/\boldsymbol{N}_{\textbf{no event}}$.

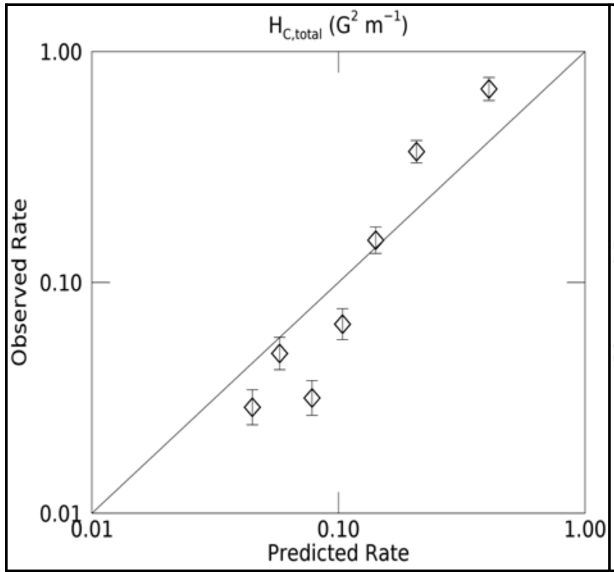
Major Flare Occurrence Rates

<table>
<tr><th colspan="2">Total unsigned current helicity</th><th colspan="2">Total photospheric magnetic free energy density</th><th colspan="2">Total unsigned vertical current</th></tr>
<tr><td colspan="2">
</td><td colspan="2">
</td><td colspan="2">
</td></tr>
<tr><td colspan="2">$y = -38(\log x)^2 + 54 \log x - 19$</td><td colspan="2">$y = 26(\log x)^2 - \log x - 50$</td><td colspan="2">$y = -416(\log x)^2 + 1001 \log x - 601$</td></tr>
<tr><td>CC = 0.94</td><td>RMSE = 0.13</td><td>CC = 0.96</td><td>RMSE = 0.14</td><td>CC = 0.94</td><td>RMSE = 0.17</td></tr>
</table>

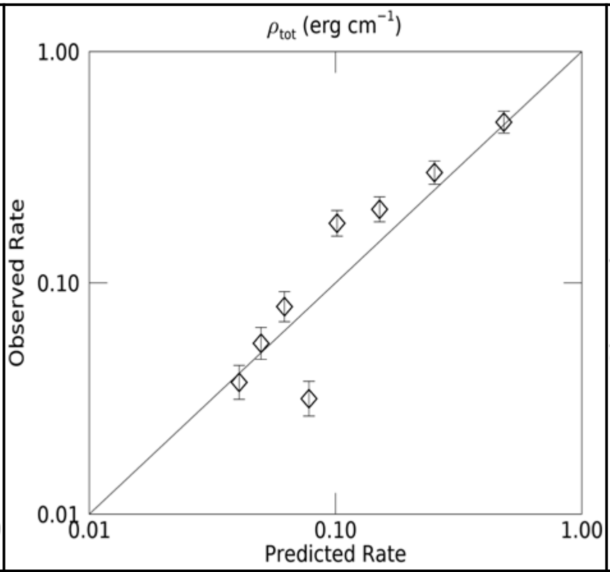
Major Flare Occurrence Rates

<table>
<tr><th colspan="2">Absolute value of the net current helicity</th><th colspan="2">Sum of the net current emanating from each polarity</th><th colspan="2">Total unsigned magnetic flux</th></tr>
<tr><td colspan="2">
</td><td colspan="2">
</td><td colspan="2">
</td></tr>
<tr><td colspan="2">$y = -9(\log x)^2 + 0.4 \log x - 2$</td><td colspan="2">$y = -387(\log x)^2 + 908 \log x - 532$</td><td colspan="2">$y = -1825(\log x)^2 + 5021 \log x - 3454$</td></tr>
<tr><td>CC = 0.98</td><td>RMSE = 0.11</td><td>CC = 0.98</td><td>RMSE = 0.08</td><td>CC = 0.92</td><td>RMSE = 0.18</td></tr>
</table>

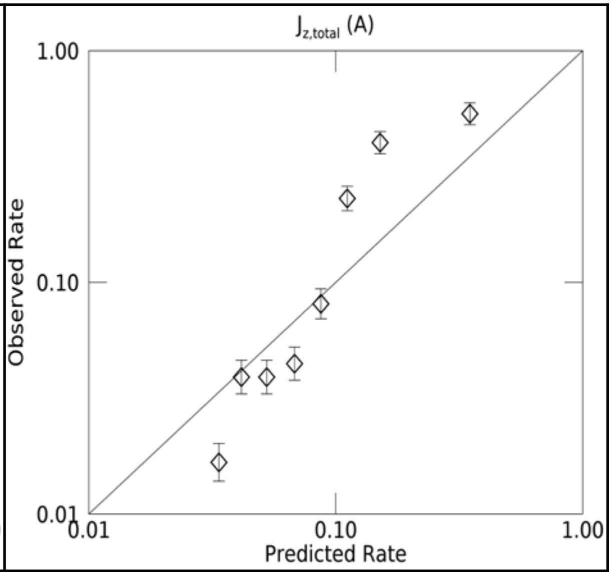
Reliability Plot



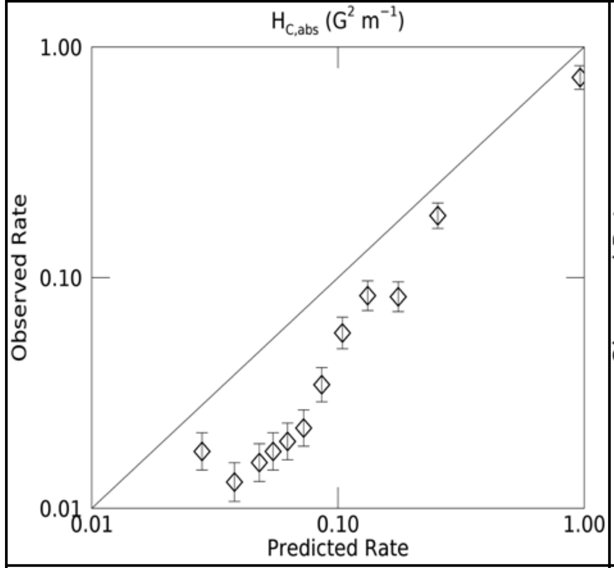
RMSE = 0.22



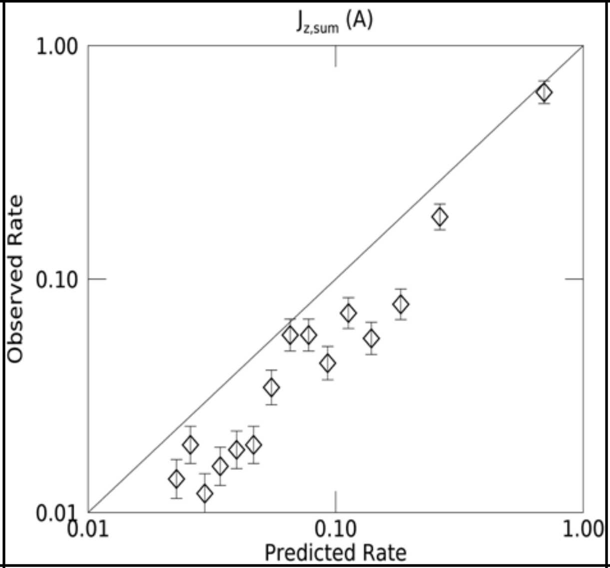
RMSE = 0.18



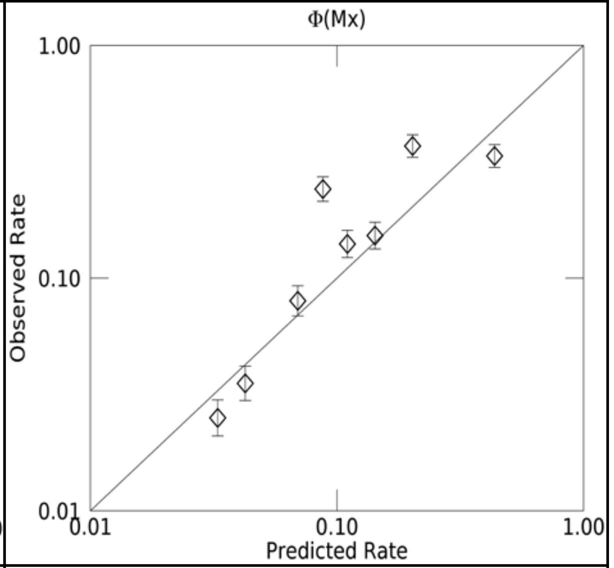
RMSE = 0.24



RMSE = 0.37



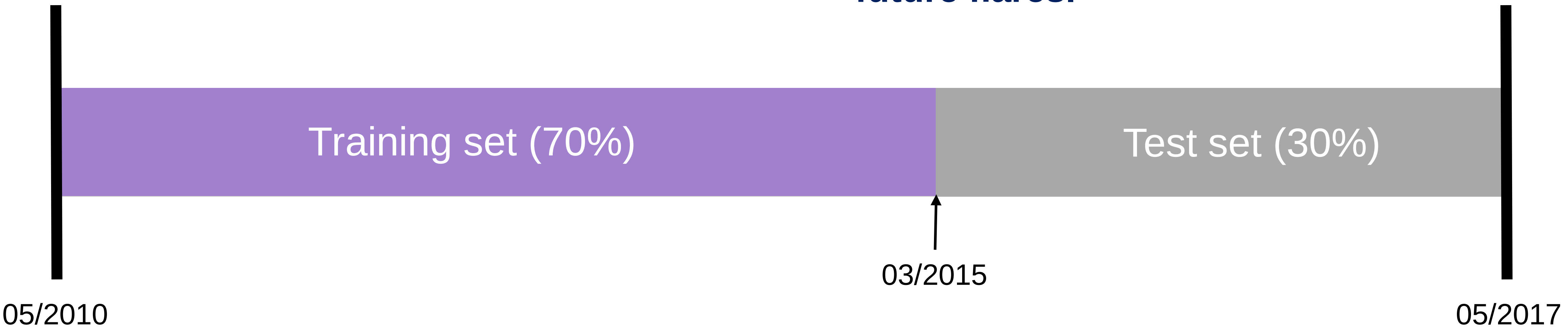
RMSE = 0.27



RMSE = 0.2

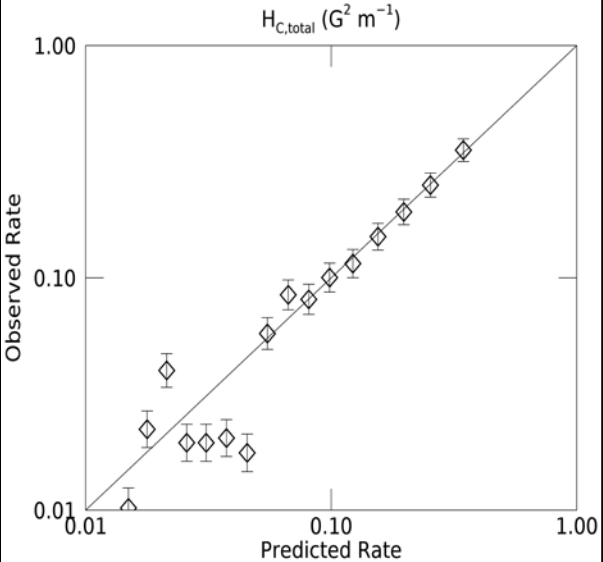
Data Set

- Chronologically selected data set ➡ **This data set is appropriate for forecasting future flares.**

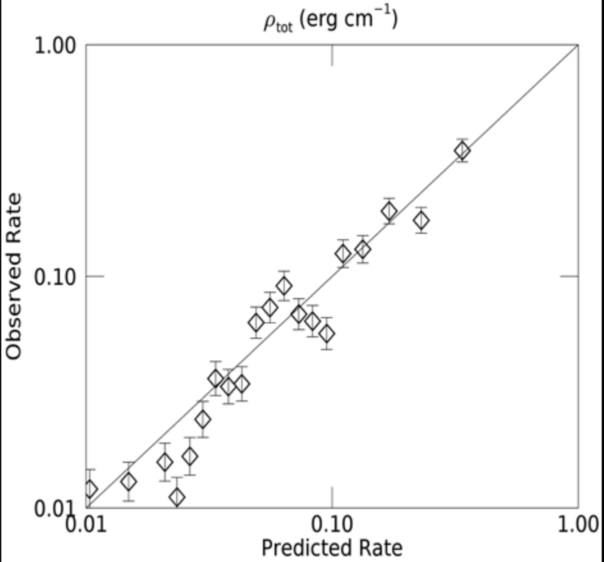


- Randomly selected data set ➡ **This data set is considered to reduce solar cycle phase effect.**

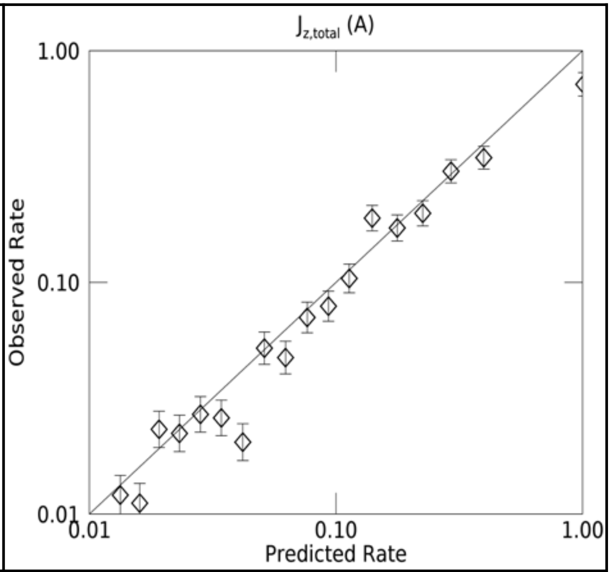
Reliability Plot (Randomly Selected Data)



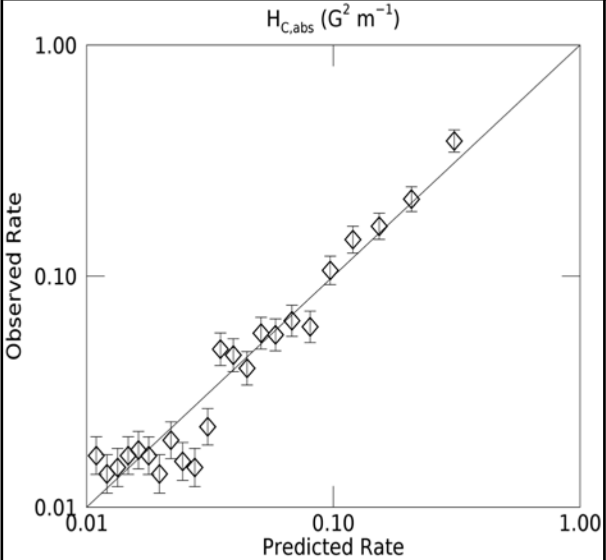
RMSE = 0.16



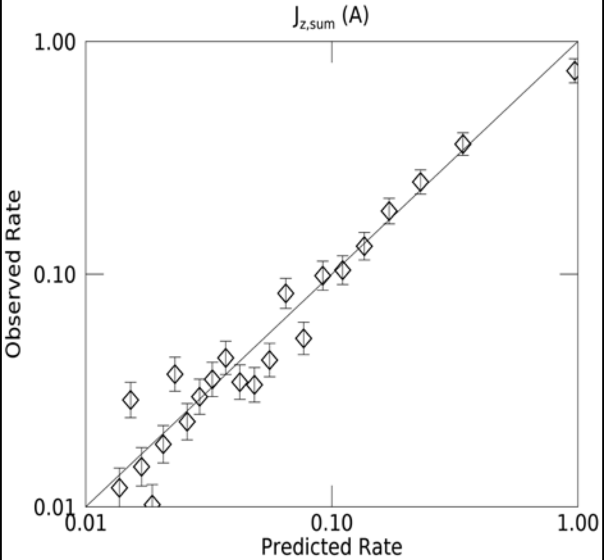
RMSE = 0.14



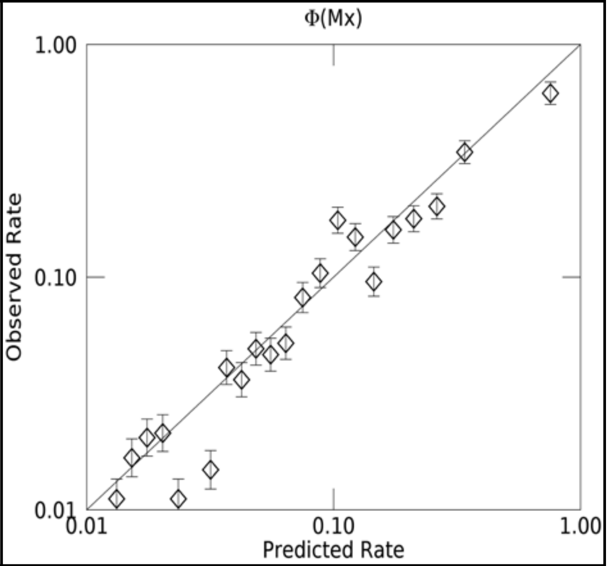
RMSE = 0.11



RMSE = 0.11



RMSE = 0.12



RMSE = 0.13

True Skill Statistic (TSS) Comparison

<table>
<tr><th colspan="2">Reference</th><th>Data</th><th>Forecast</th><th>Method</th><th>Period</th><th>Training/Test</th><th>TSS</th></tr>
<tr><td colspan="2">Bloomfield et al. (2012)</td><td>McIntosh Classification</td><td>Probability</td><td>Historical Poisson Statistic</td><td>Dec 1988 ~ Dec 2010</td><td>Chronological Selection</td><td>0.54</td></tr>
<tr><td colspan="2">Bobra and Couvidat (2015)</td><td rowspan="6">SDO/HMI</td><td>Yes/No</td><td>Support Vector Machine</td><td>May 2010 ~ May 2014</td><td>Random Selection</td><td>0.76</td></tr>
<tr><td colspan="2">Nishizuka et al. (2017)</td><td>Yes/No</td><td>K-Nearest Neighbor + UV Emission + Flare History</td><td>June 2010 ~ December 2015</td><td>Random Selection</td><td>0.91</td></tr>
<tr><td colspan="2">Liu et al. (2017)</td><td>Multi class</td><td>Random Forest</td><td>May 2010 ~ December 2016</td><td>Random Selection</td><td>0.53</td></tr>
<tr><td rowspan="3">This work</td><td>TOTUSJH</td><td rowspan="3">Probability</td><td rowspan="3">Empirical Relationship</td><td rowspan="3">May 2010 ~ April 2017</td><td rowspan="3">Chronological Selection</td><td>0.81</td></tr>
<tr><td>TOTUSJZ</td><td>0.8</td></tr>
<tr><td>USFLUX</td><td>0.79</td></tr>
</table>

Conclusion

- The major flare occurrence rates (M- and X-class) are well correlated with six SHARP magnetic parameters.
- The occurrence rate ranges from 0.001 to 1 for M- and X-class flares.
- The slopes between the logarithmic values of six magnetic parameters and flaring rates tend to decrease as the values of parameters increase.
- The test shows that the total photospheric magnetic free energy density gives the minimum RMS error between observed flare rates and predicted ones.
- Among six parameters, the total unsigned current helicity, the total unsigned vertical current, and the total unsigned magnetic flux have higher TSS values than the other parameters.