COMPARING DIFFERENT SOLAR FLARE PREDICTION METHODS

Where are we, and how far can we go?

Manolis K. Georgoulis[1] & Rami Qahwaji[2]

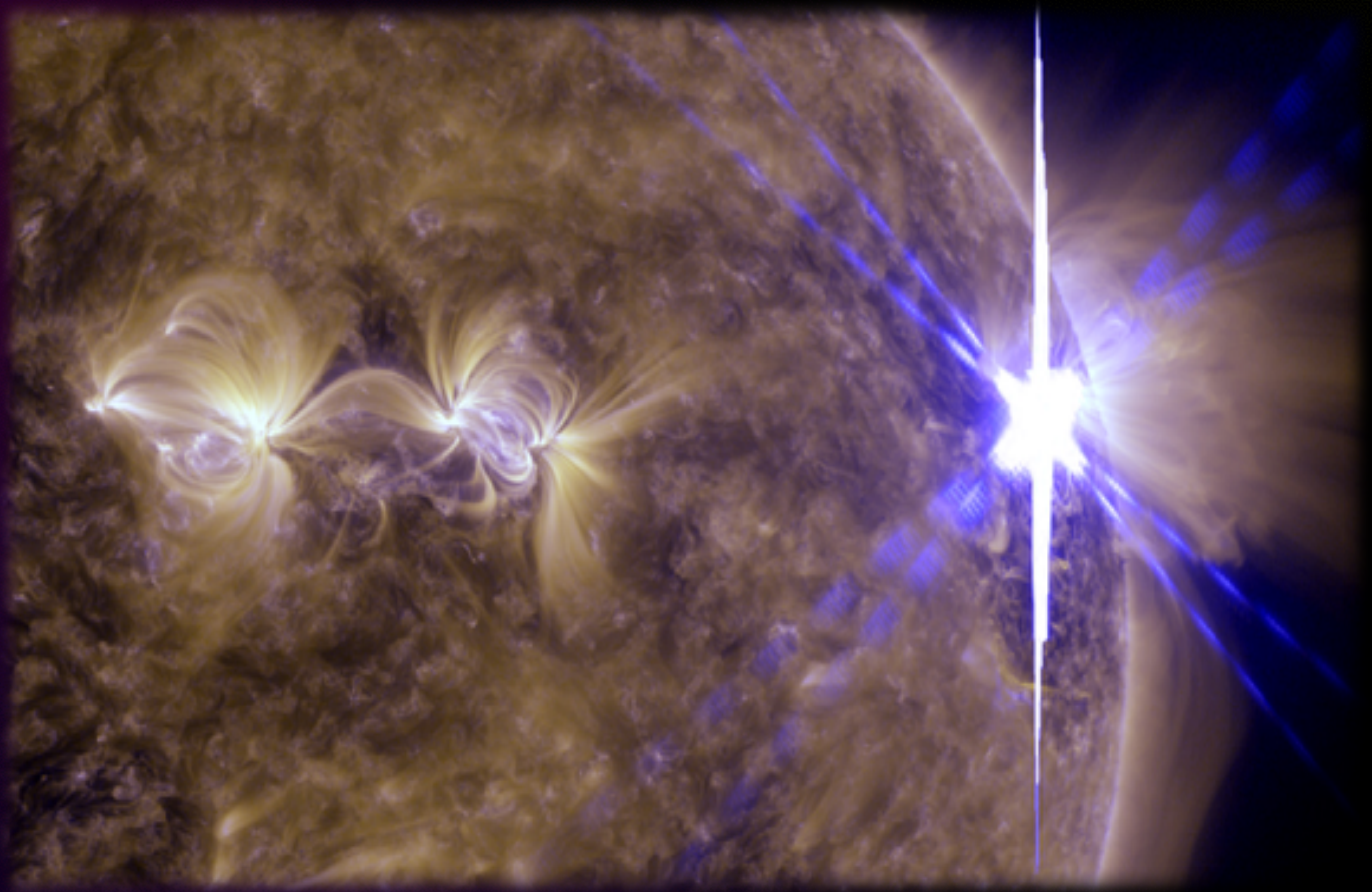
1 RCAAM of the Academy of Athens, Greece
2 Bradford University, United Kingdom

Oostende, Belgium, Nov 14 - 18, 2016

OUTLINE

- Why do we need flare prediction?
- The nature of flare occurrence - are flares random?
- Can flares be predicted? — different methods
- Recent trends in solar flare prediction
- Validation : process and intrinsics
- From a method to an operational forecasting service
- Conclusion

WHY PREDICT SOLAR FLARES?



WHY PREDICT SOLAR FLARES?

solar flare — t_0

arrival of "hard" (X-, γ)-ray photons — t_0+ 8 min

arrival of first flare-accelerated particles — ~ t_0+ 20 min

arrival of CME-shock-accelerated particles

arrival of CME — t_0+ 2-4 days

WHY PREDICT SOLAR FLARES?

solar flare — t_0

arrival of "hard" (X-, γ)-ray photons — t_0+ 8 min

arrival of first flare-accelerated particles — ~ t_0+ 20 min

arrival of CME-shock-accelerated particles

arrival of CME — t_0+ 2-4 days

Hard flare photons and non-thermal particulate (mostly protons >10 MeV) affect humans beyond LEO and on solar system bodies lacking an atmosphere. Damages in space-based electronics, radio blackouts, etc., can occur as a result of flares

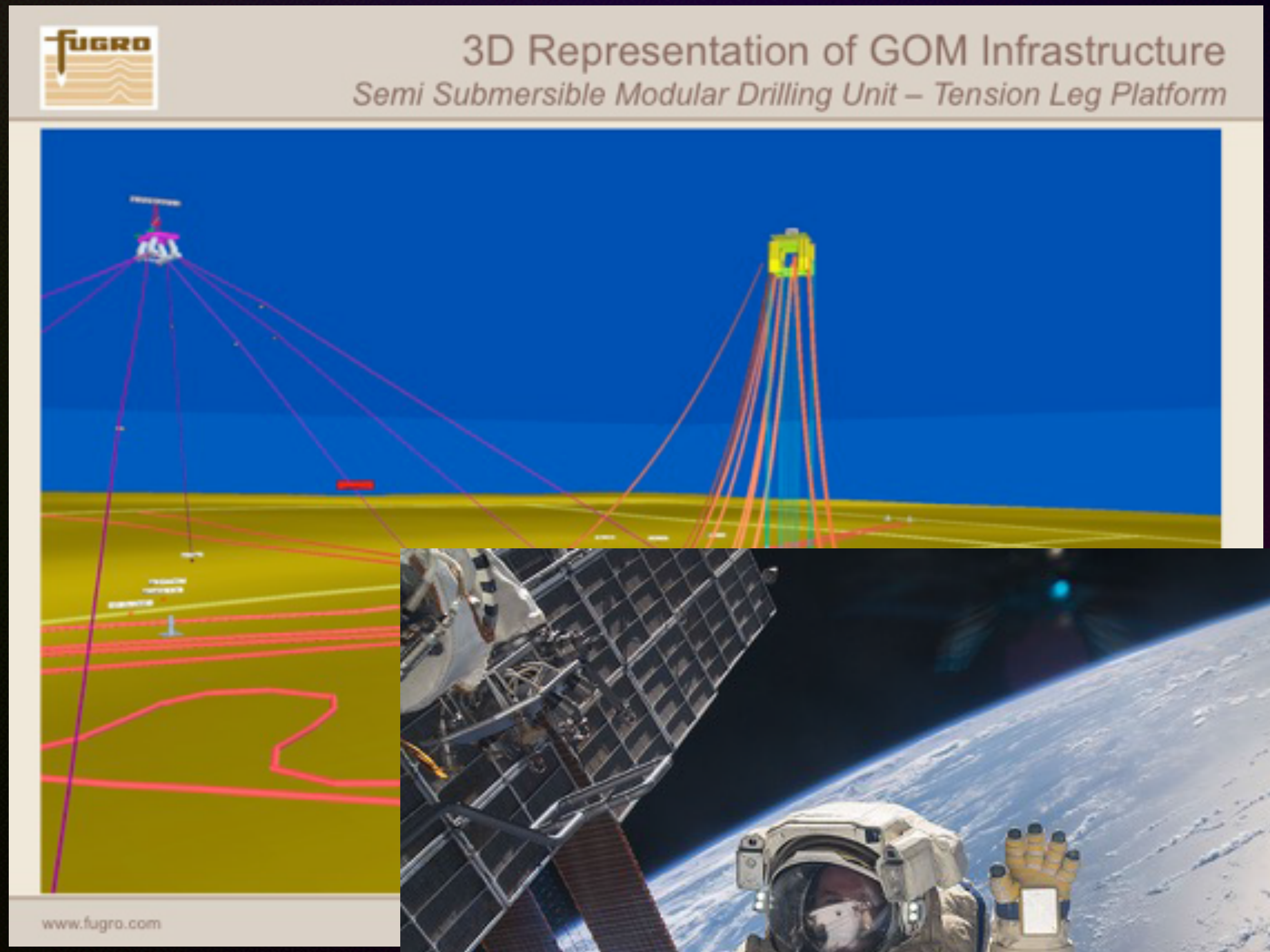
WHY PREDICT SOLAR FLARES?

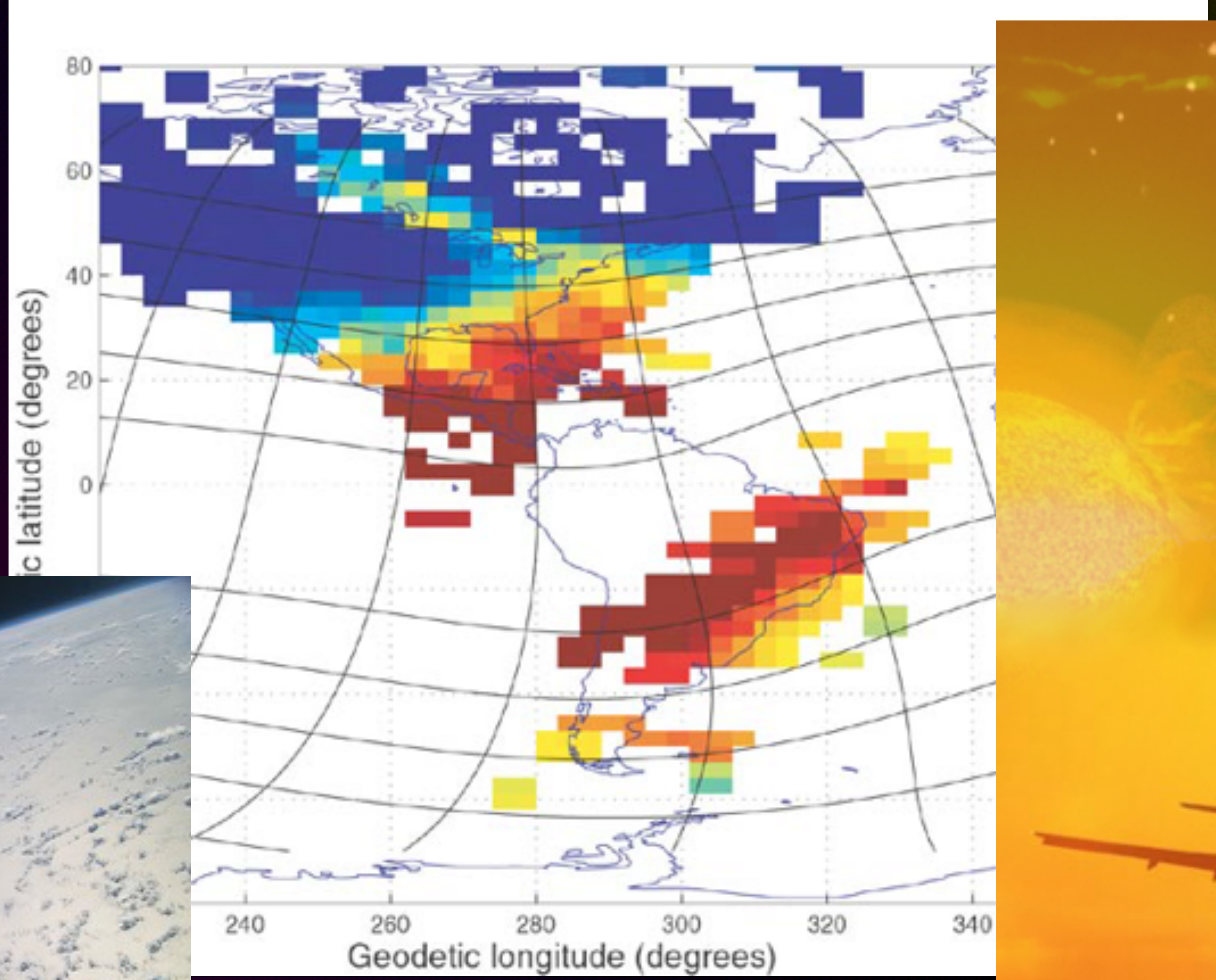
solar flare	arrival of "hard" (X-, γ)-ray photons	arrival of first flare-accelerated particles	arrival of CME-shock-accelerated particles	arrival of CME
t_0	t_0+ 8 min	~ t_0+ 20 min		t_0+ 2-4 days

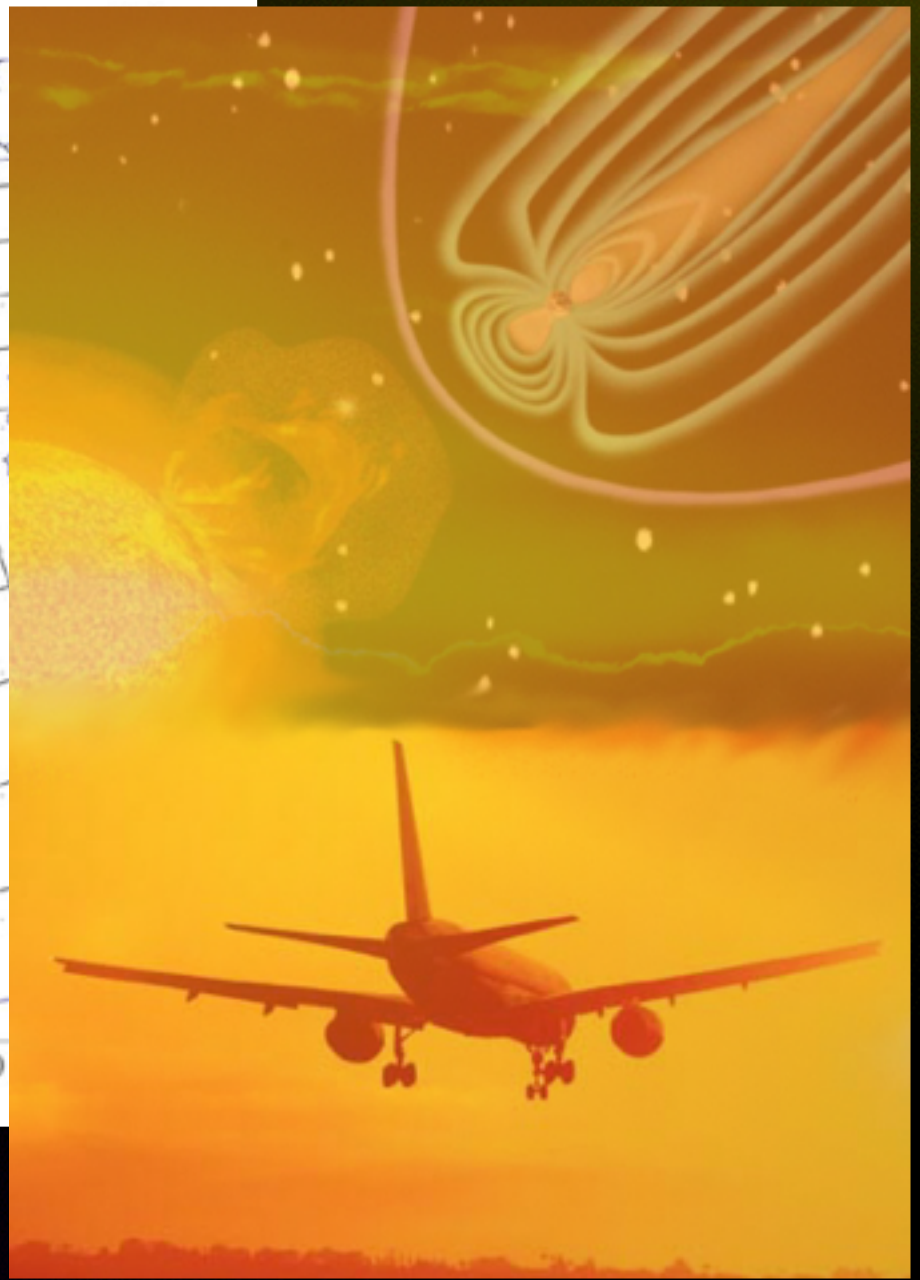
Hard flare photons and non-thermal particulate (mostly protons >10 MeV) affect humans beyond LEO and on solar system bodies lacking an atmosphere. Damages in space-based electronics, radio blackouts, etc., can occur as a result of flares

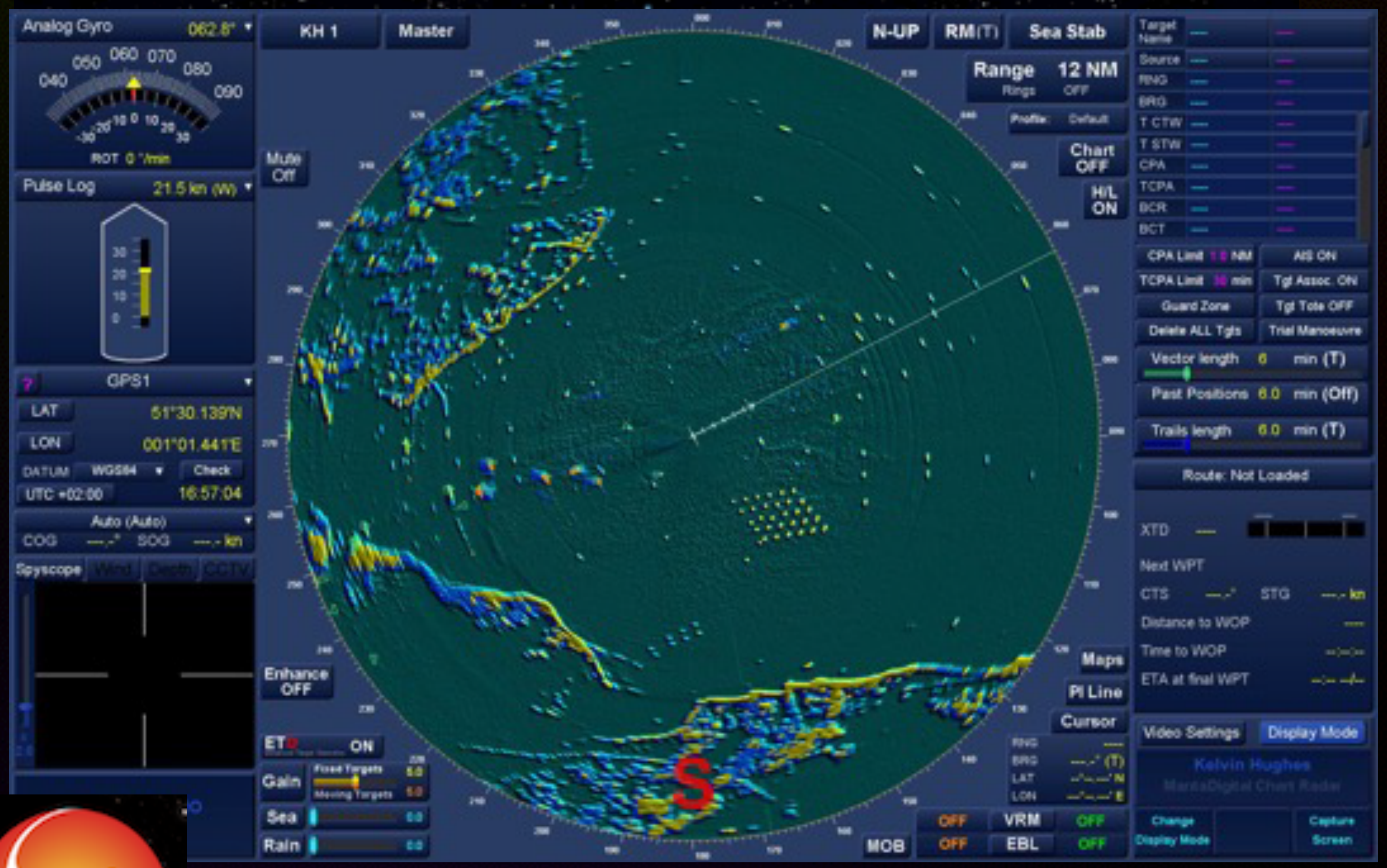
No early warning time for flare photons - slim window for particulate in worst case!

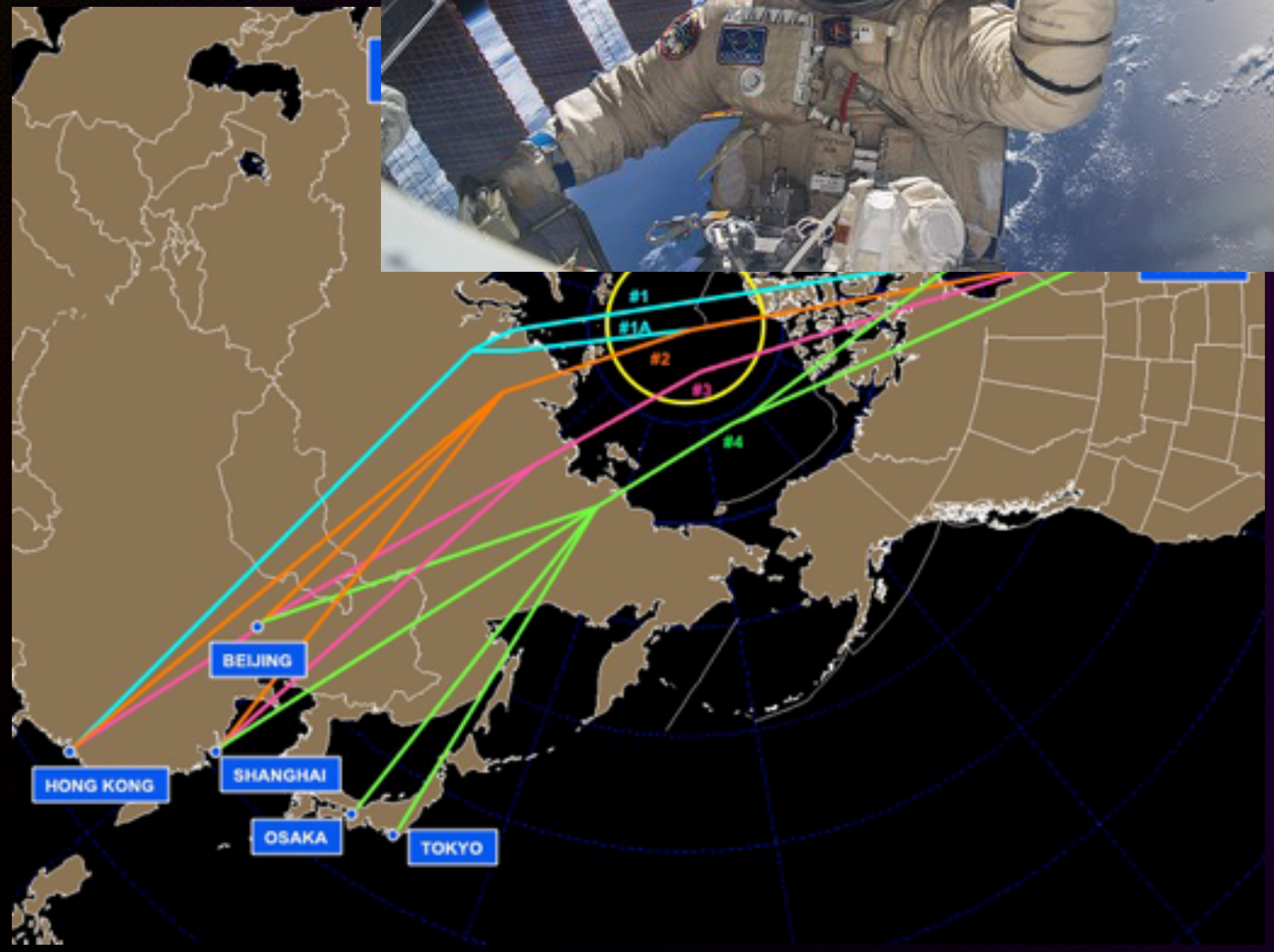
MAJOR FLARE REPERCUSSIONS: EVERYTHING UNDER THE SUN



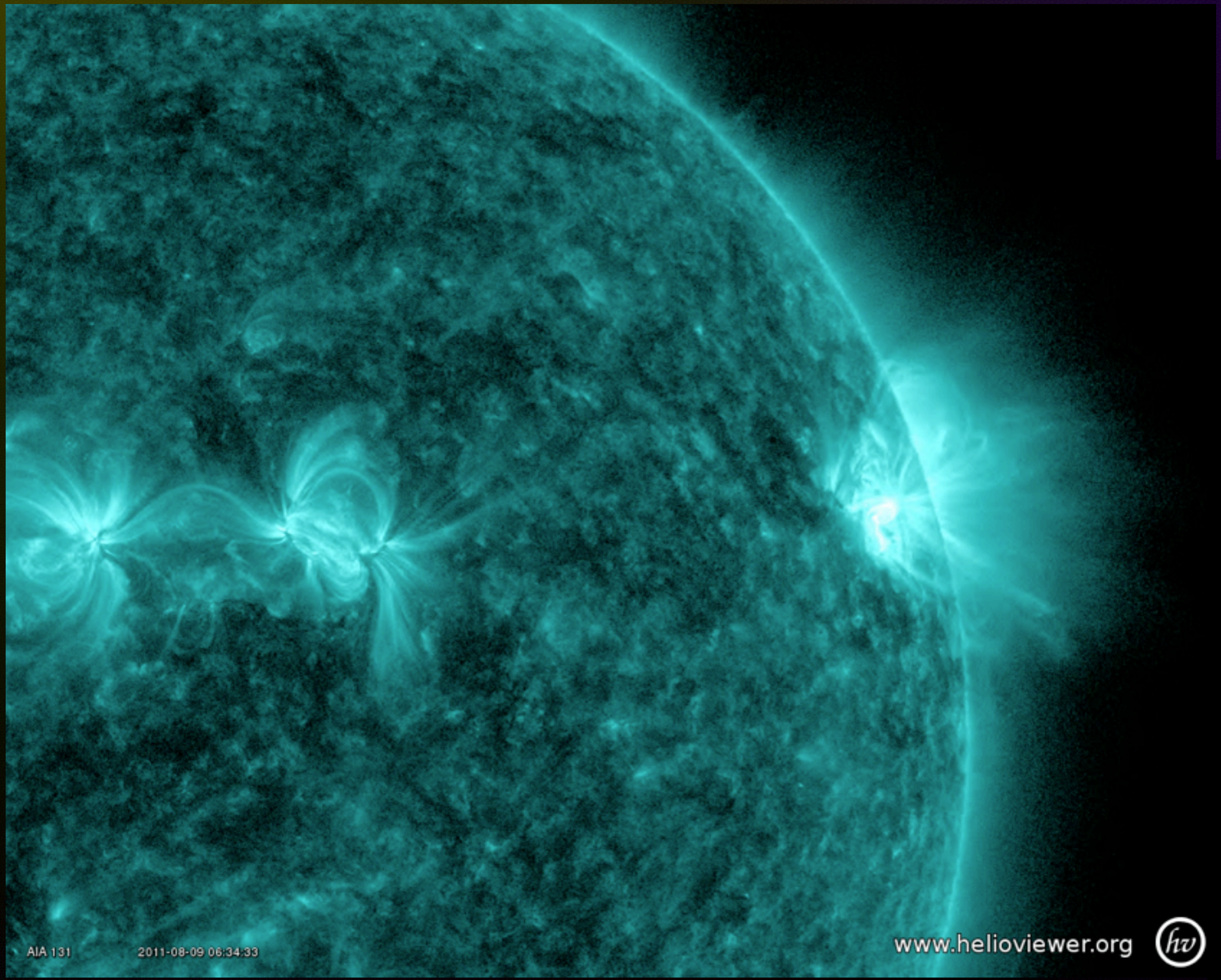


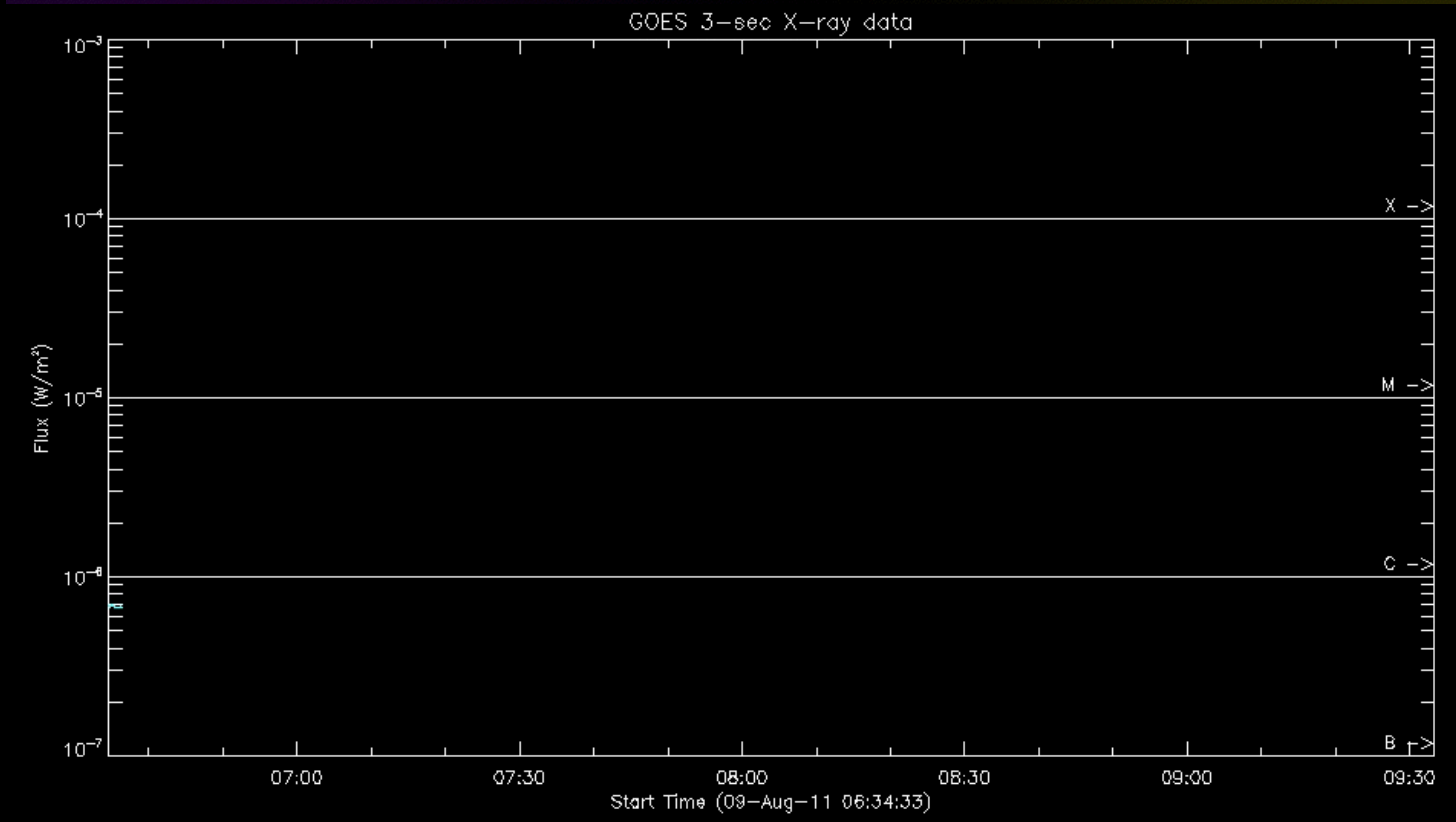






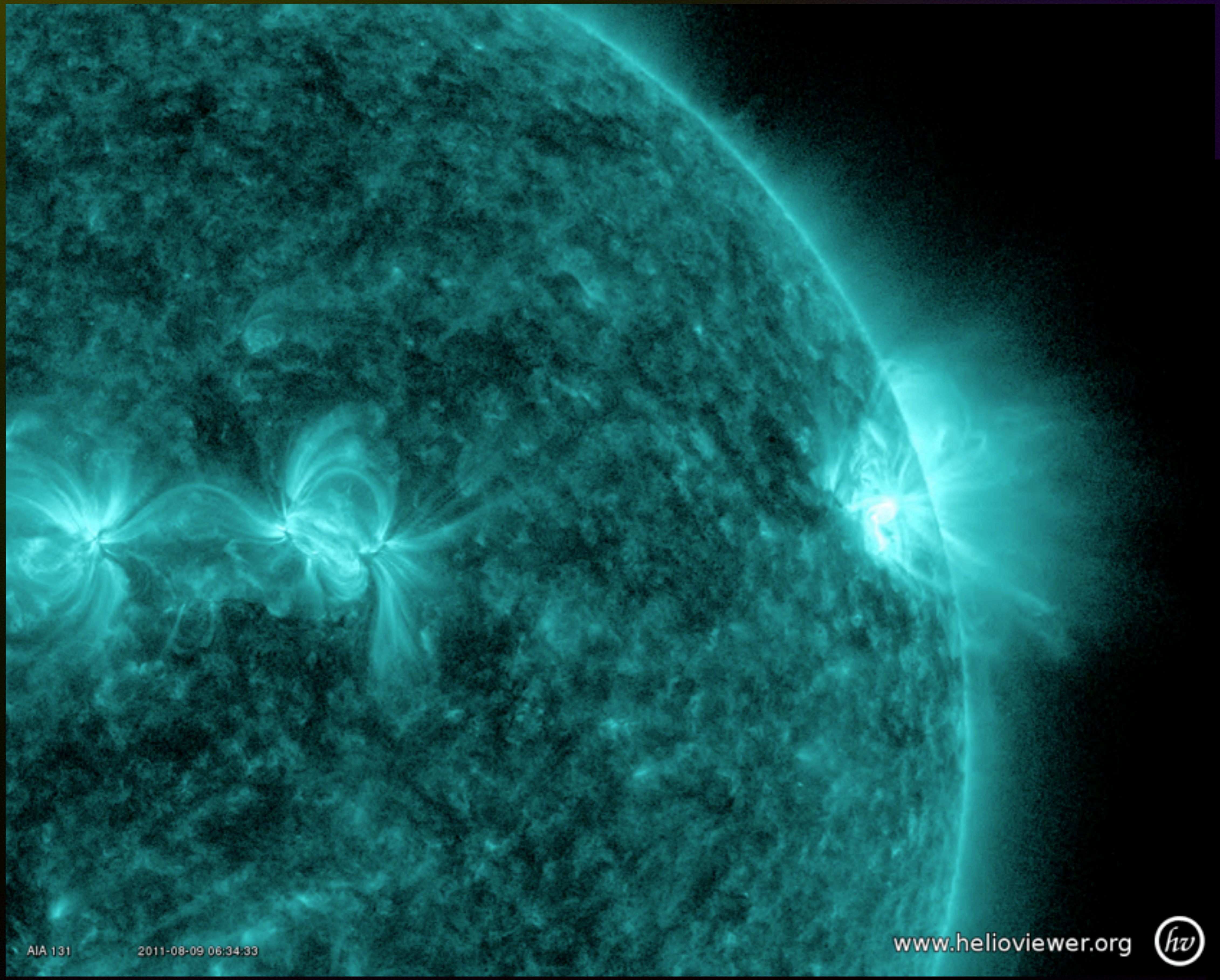
A PHENOMENOLOGY DEFINITION …

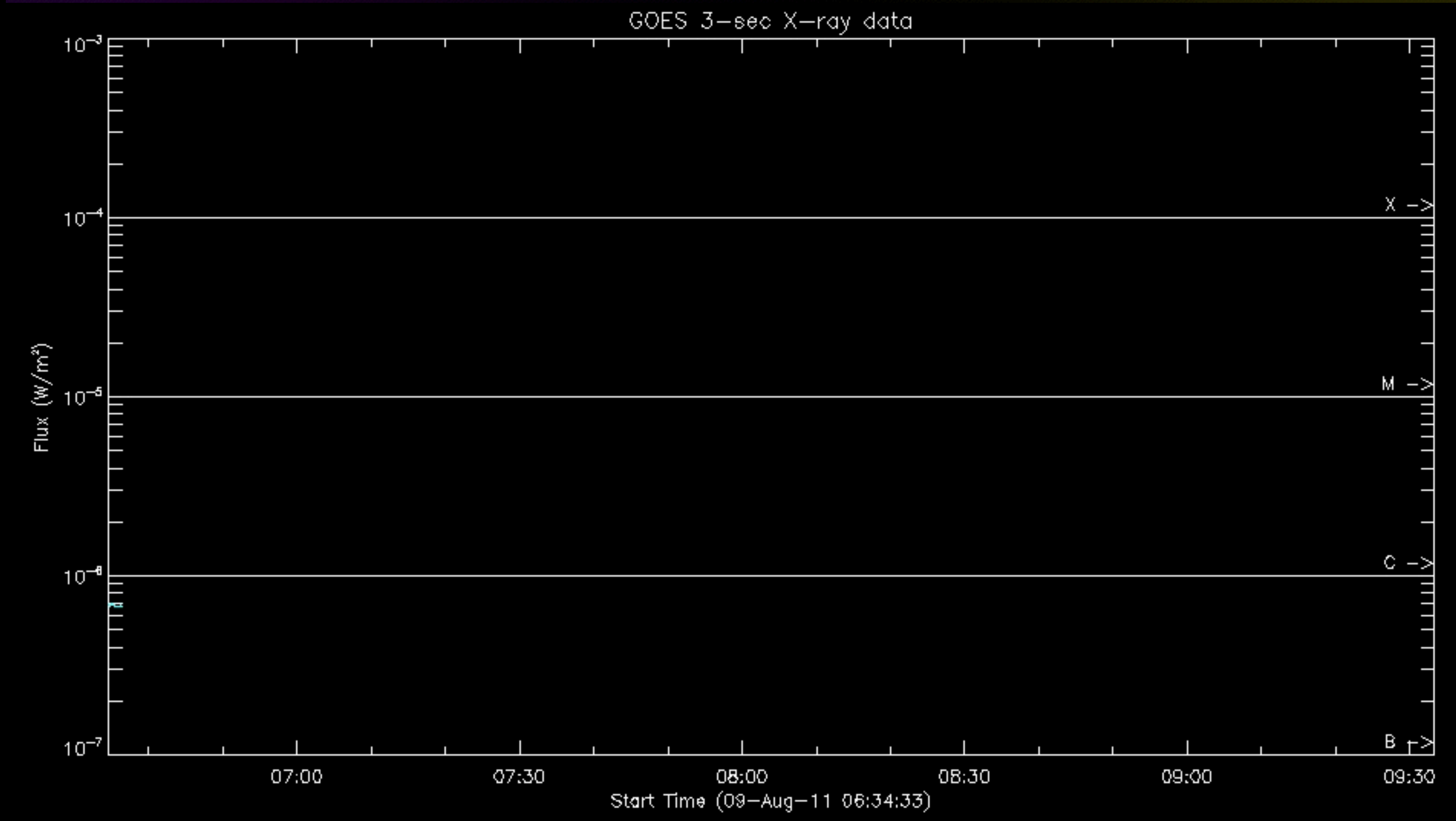




A sudden commencement of enhanced, localized electromagnetic emission extending over practically the entire range of the electromagnetic spectrum. Typically measured in 1 - 8 Å SXR

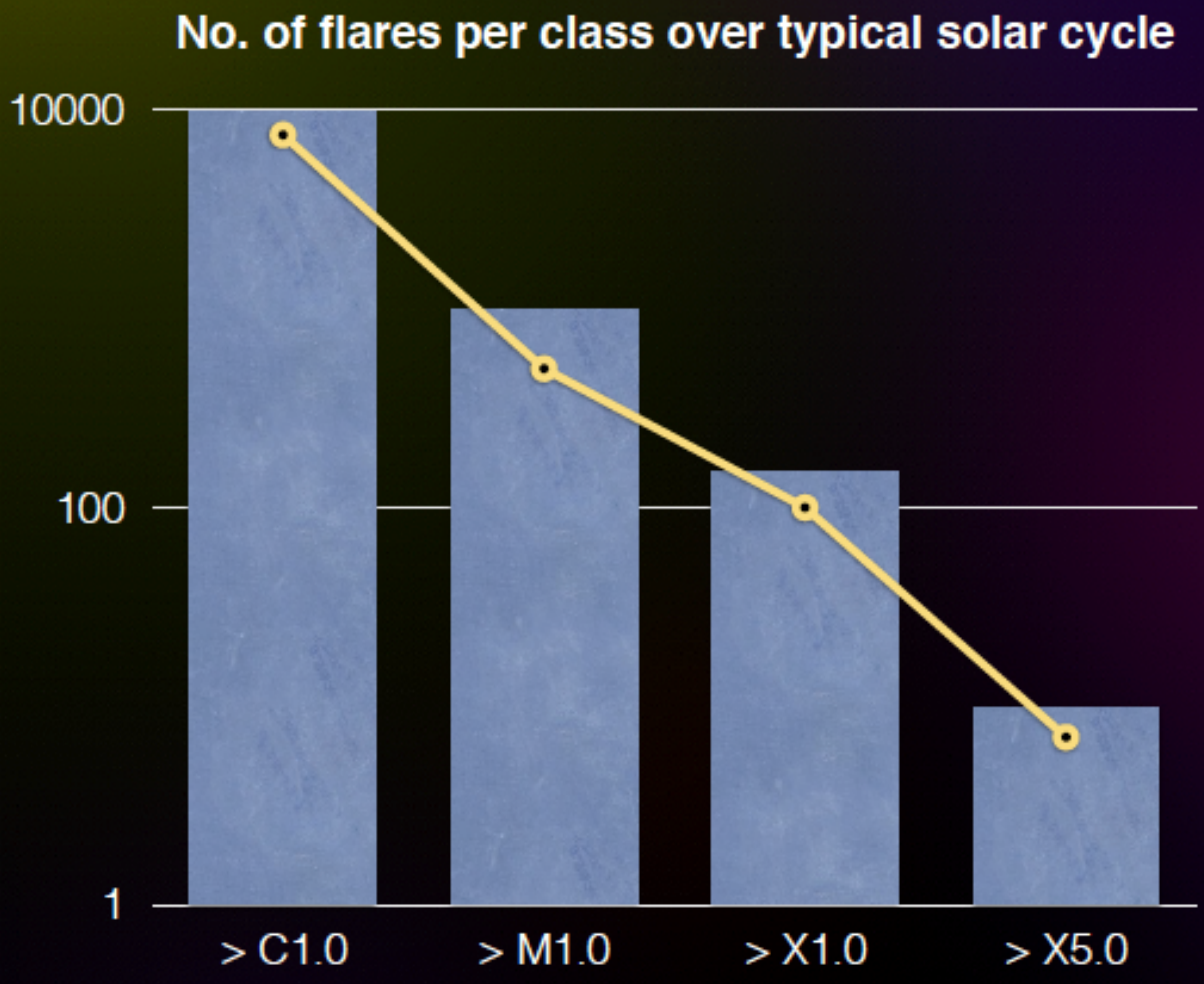
A PHENOMENOLOGY DEFINITION …

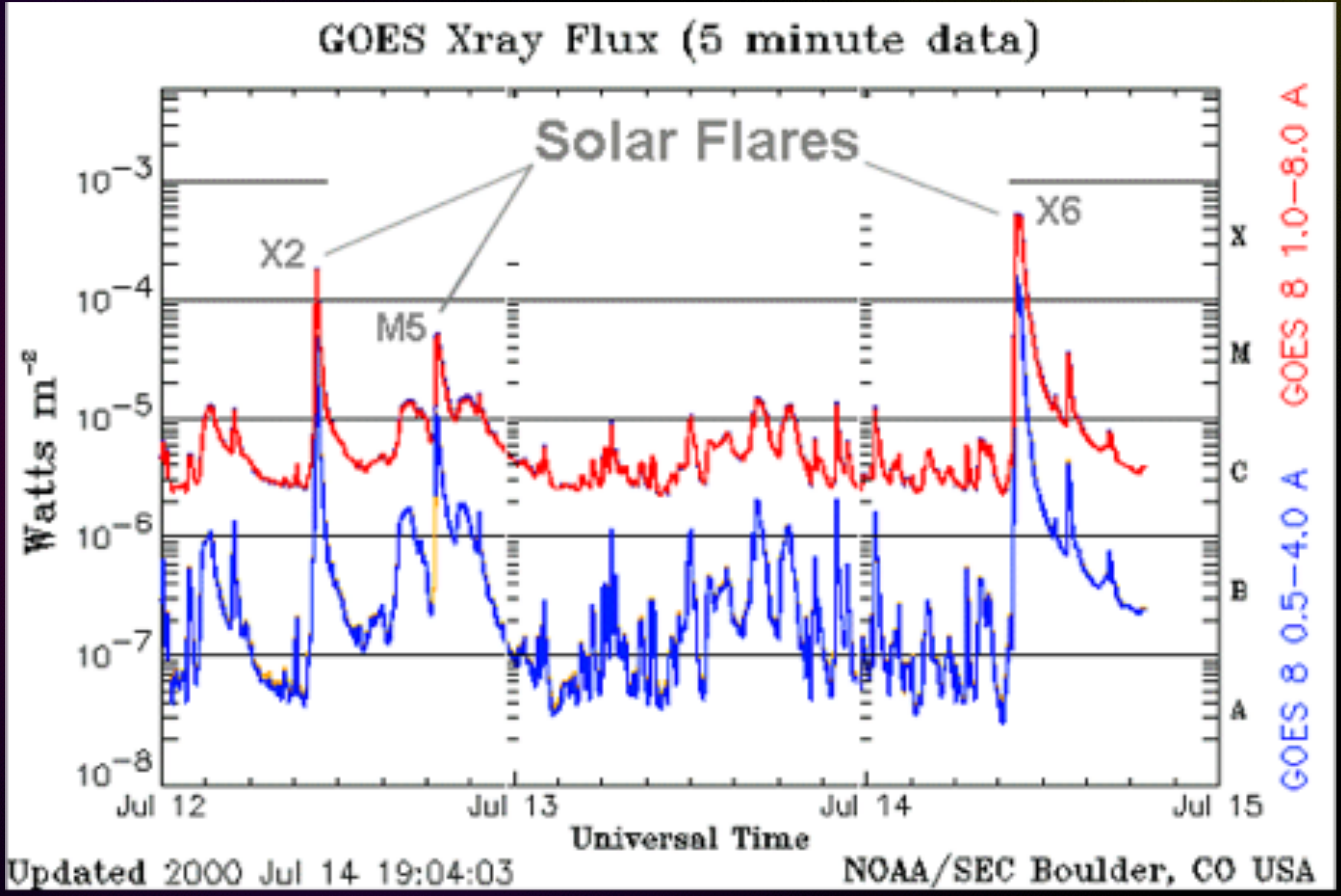




A sudden commencement of enhanced, localized electromagnetic emission extending over practically the entire range of the electromagnetic spectrum. Typically measured in 1 - 8 Å SXR

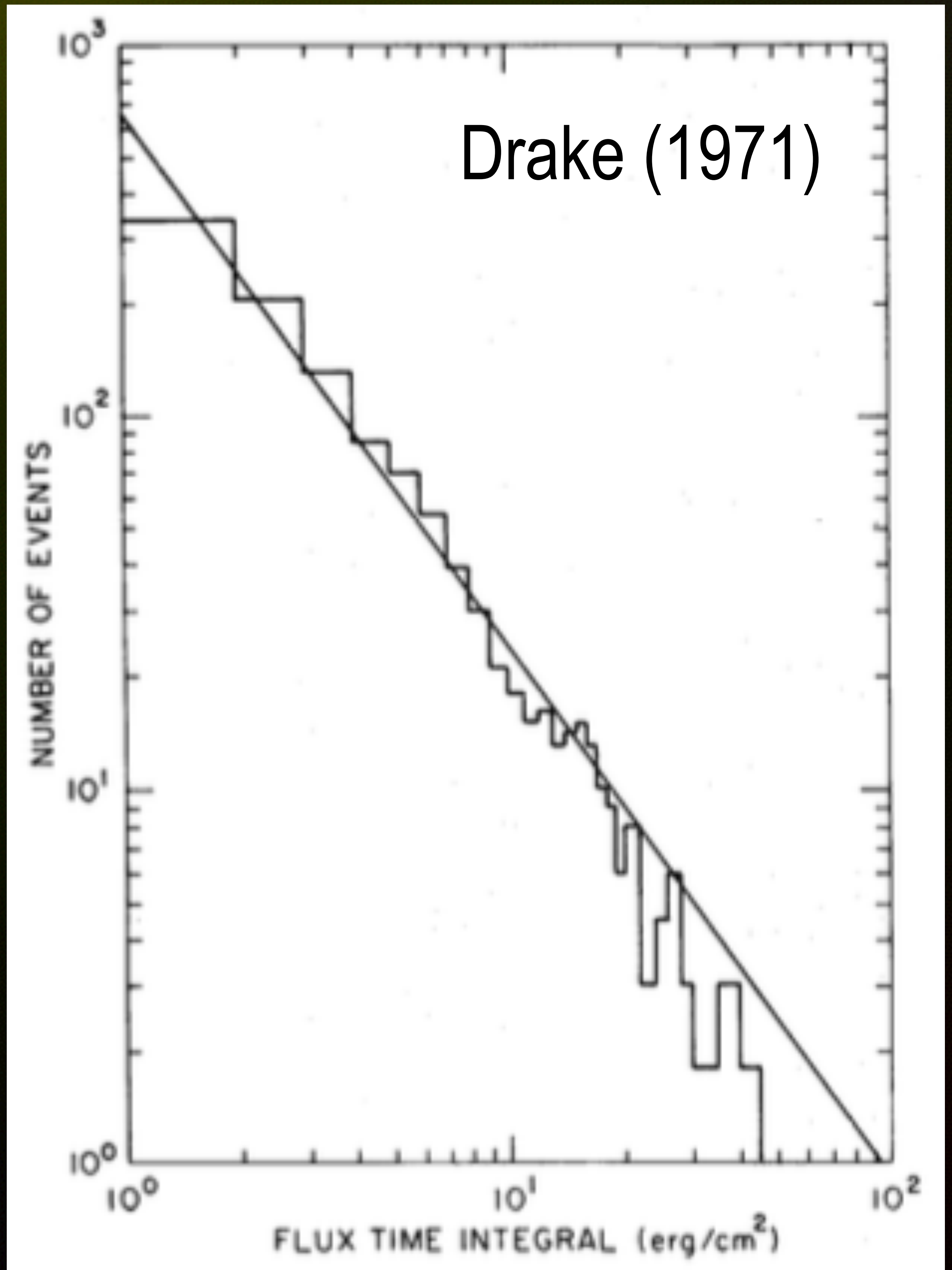
… AND STATISTICAL BEHAVIOR





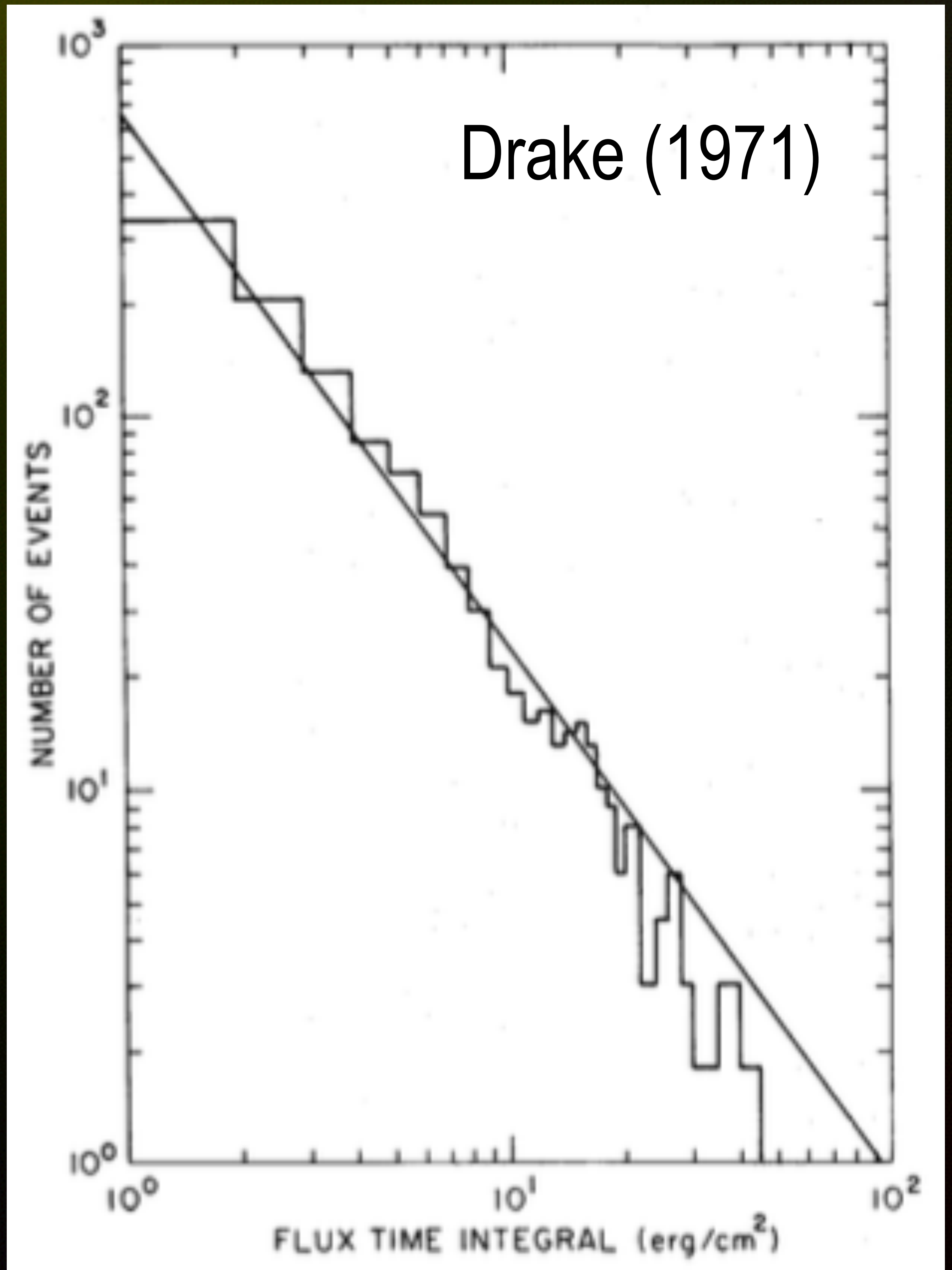
"Active" solar conditions over a 3-day period in July 2000

NATURE OF FLARE OCCURRENCE



Flare occurrence number vs. integrated photon flux

NATURE OF FLARE OCCURRENCE



Flare occurrence number vs. integrated photon flux

Flares are (**Rosner & Vaiana 1978**):

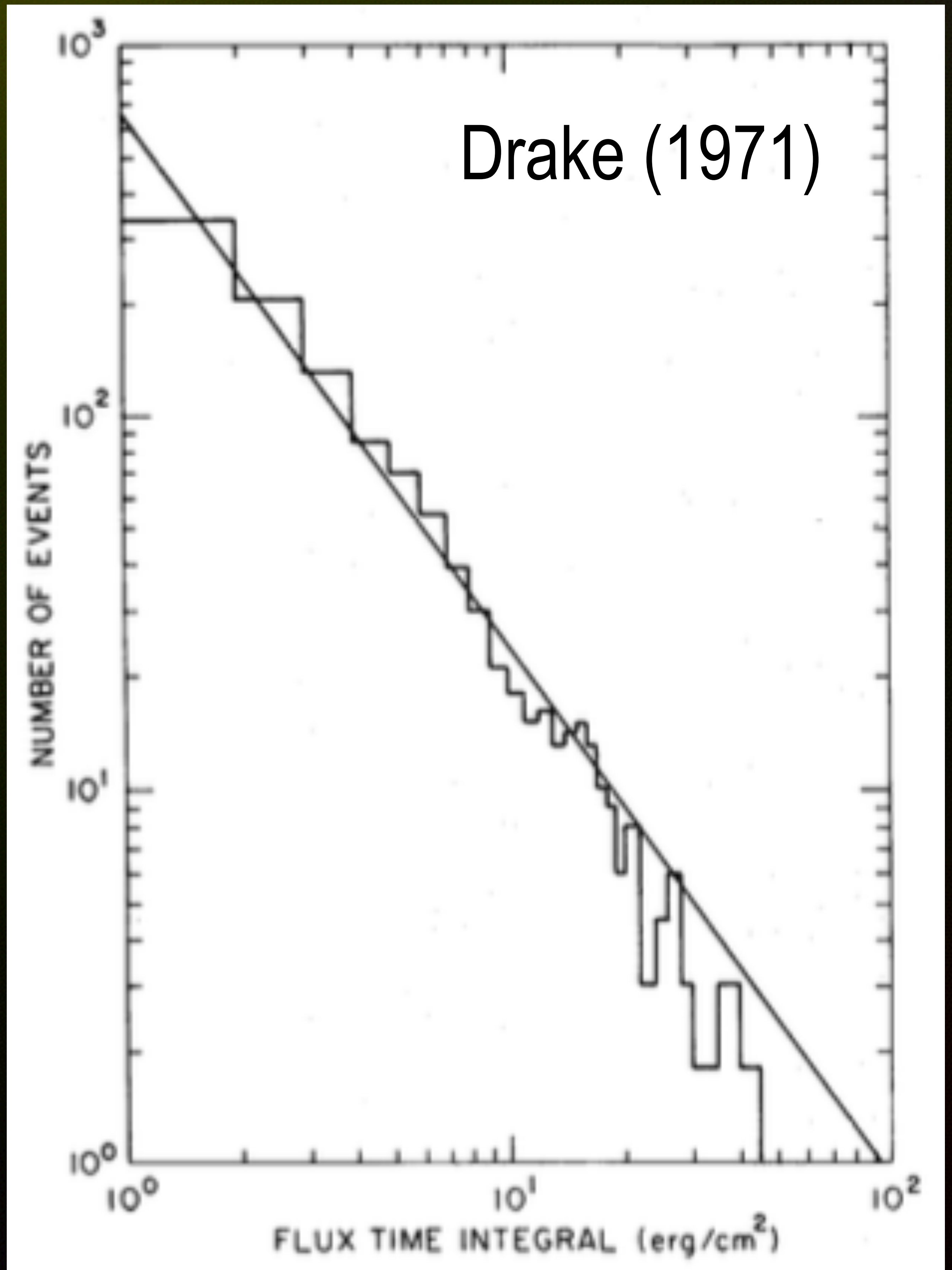
- Stochastic relaxation (storage and release) processes
- Physically uncoupled / independent
- Brief, comparing to intermediate times between flares

$$P(t) = \bar{\nu} e^{-\bar{\nu} t}$$

- Leading to a power-law occurrence frequency for flare energies

$$P(E) \sim \left(1 + {E}/{E_0}\right)^{-\gamma}$$

NATURE OF FLARE OCCURRENCE



Flare occurrence number vs. integrated photon flux

Flares are (**Rosner & Vaiana 1978**):

- Stochastic relaxation (storage and release) processes
- Physically uncoupled / independent
- Brief, comparing to intermediate times between flares

$$P(t) = \bar{\nu} e^{-\bar{\nu} t}$$

- Leading to a power-law occurrence frequency for flare energies

$$P(E) \sim \left(1 + {E}/{E_0}\right)^{-\gamma}$$

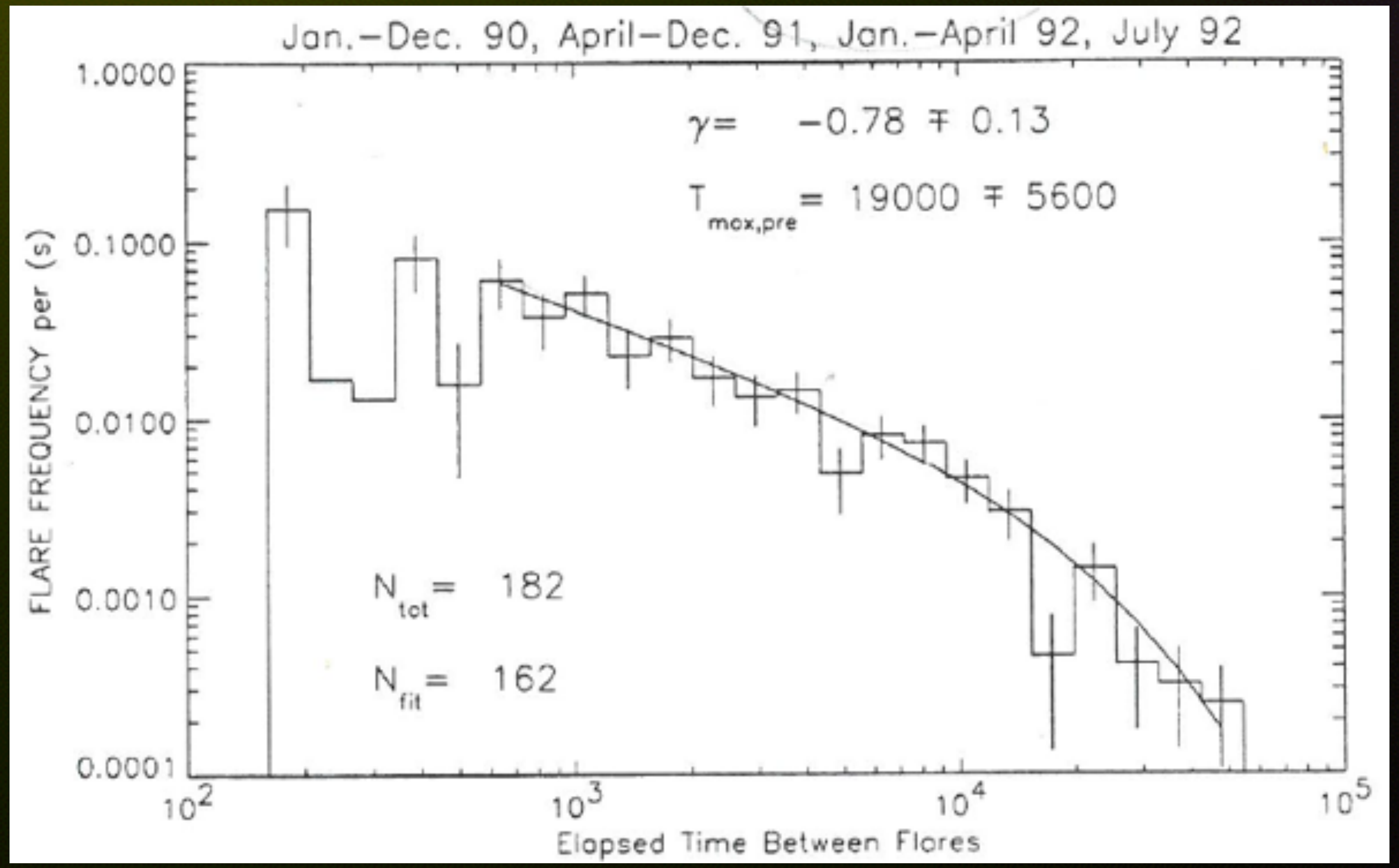
Power-law distribution of flare size later attributed to the concept of self-organized criticality (1990s)

A RATHER GRAPHIC EXAMPLE OF MARGINAL STABILITY

A RATHER GRAPHIC EXAMPLE OF MARGINAL STABILITY

HOWEVER, ARE FLARES RANDOM? - DISTRIBUTION OF WAITING TIMES

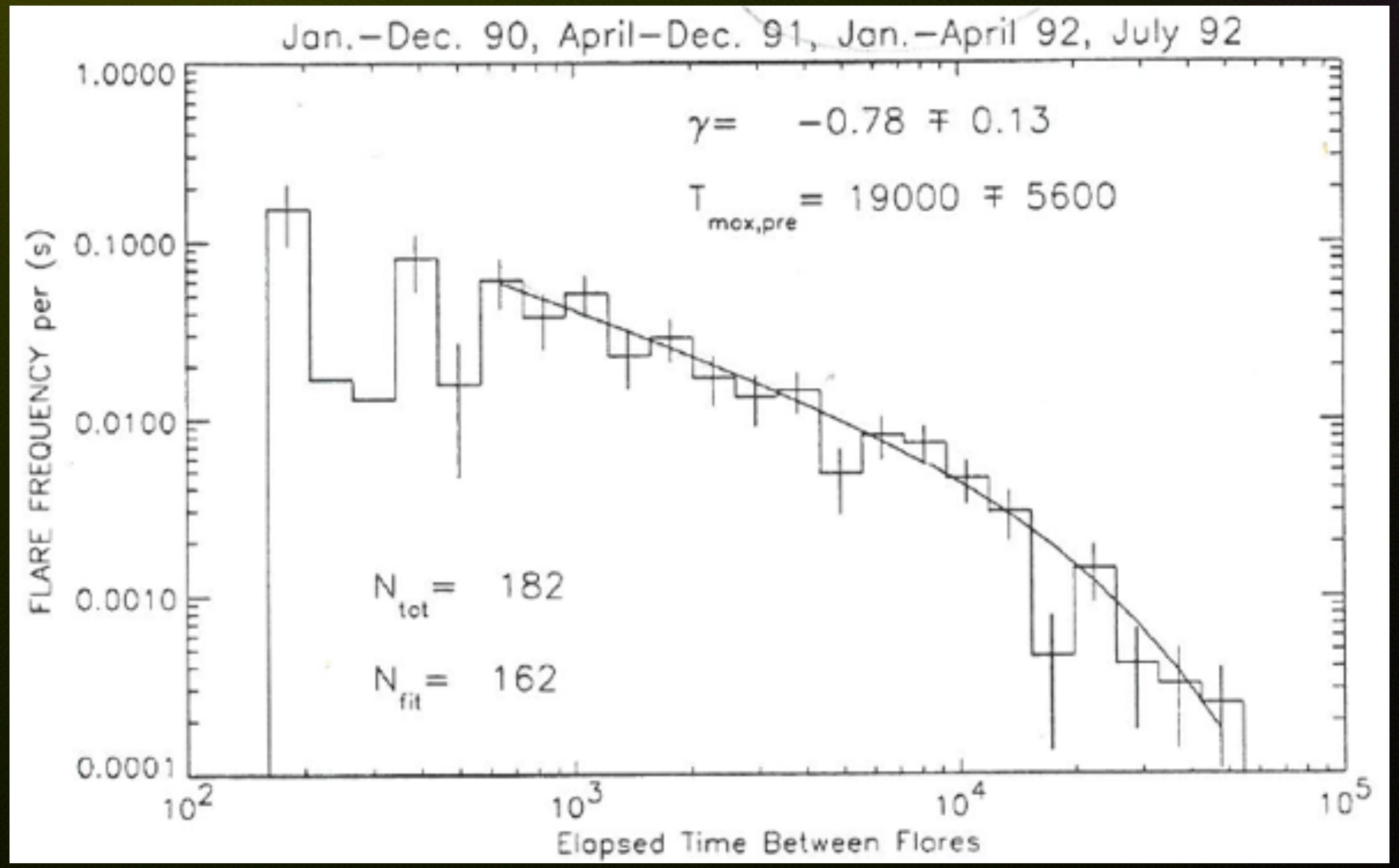
HOWEVER, ARE FLARES RANDOM? - DISTRIBUTION OF WAITING TIMES



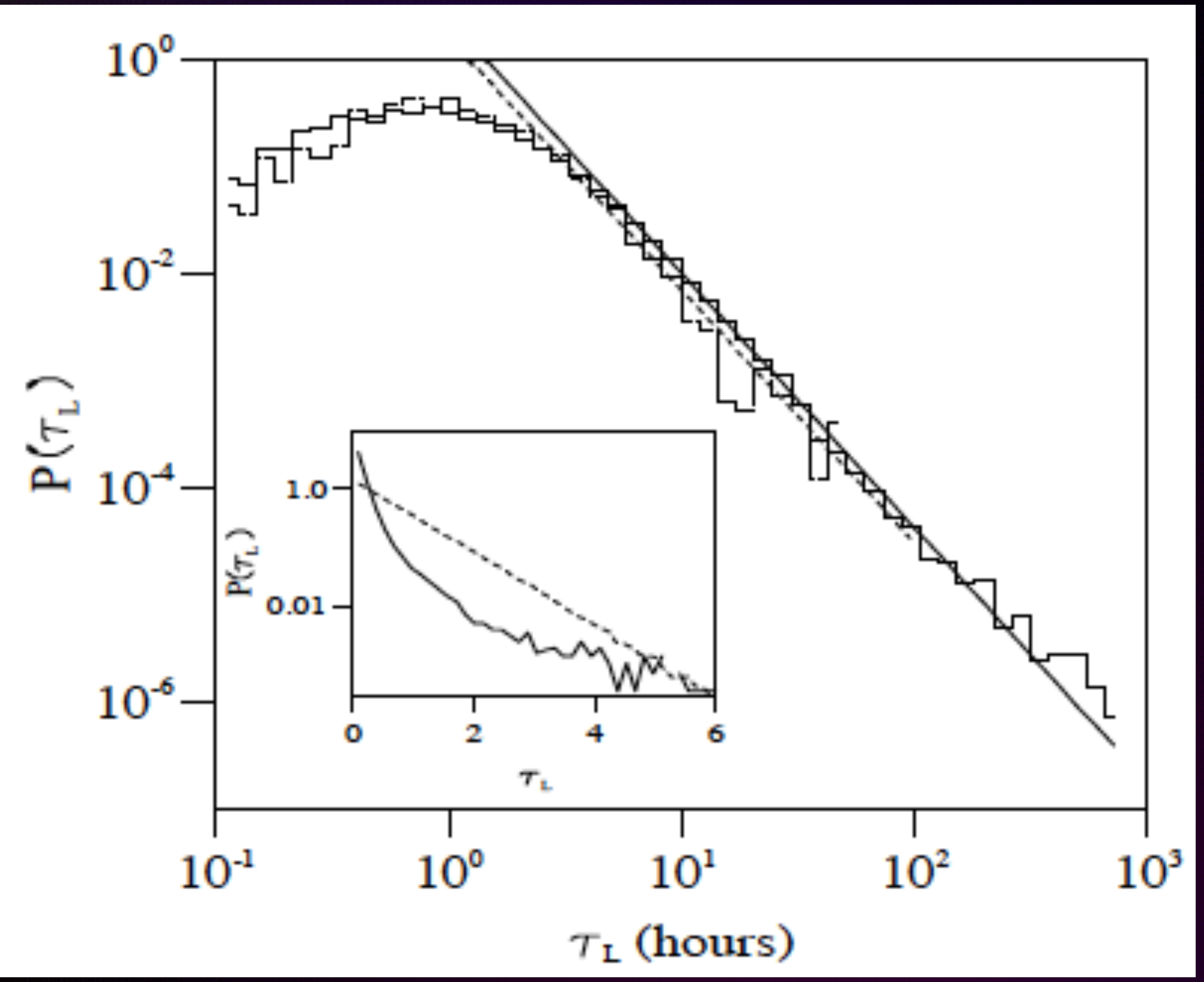
Crosby, PhD Thesis (1996)

Exponential law of waiting times: a totally random, memoryless flare occurrence along the classical self-organized criticality concept

HOWEVER, ARE FLARES RANDOM? - DISTRIBUTION OF WAITING TIMES



Crosby, PhD Thesis (1996)

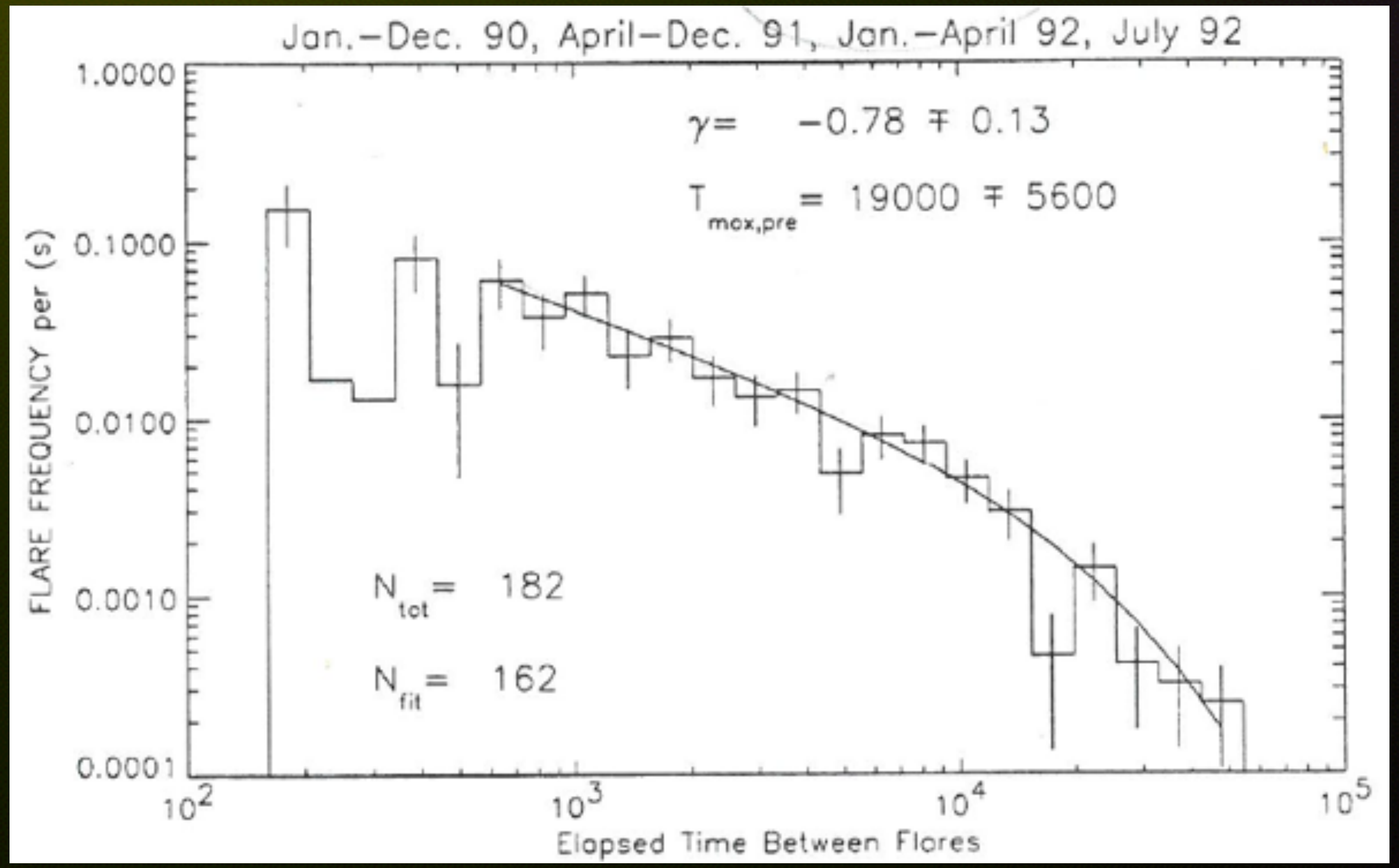


Bofetta et al., (1999)

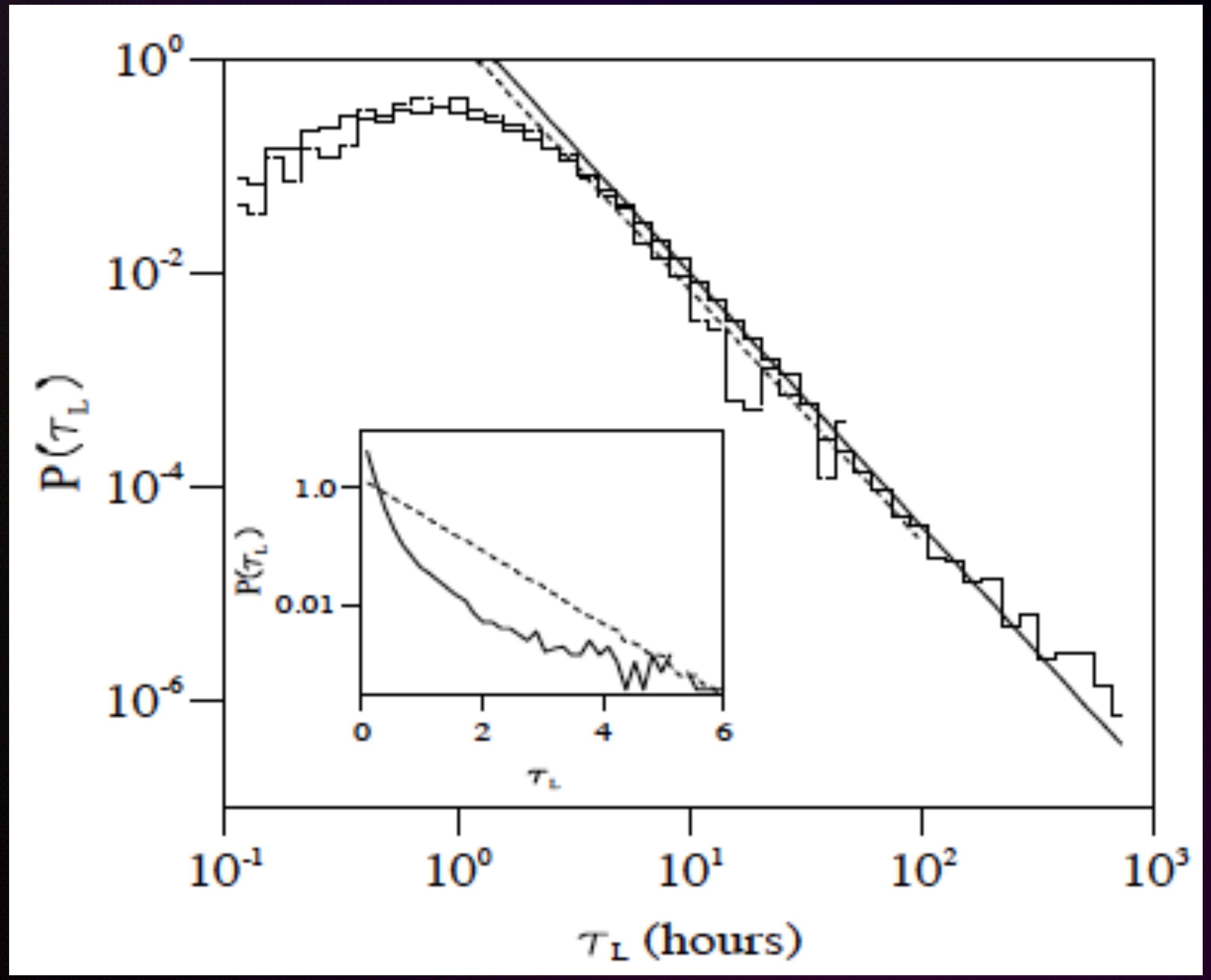
Exponential law of waiting times: a totally random, memoryless flare occurrence along the classical self-organized criticality concept

Robust power-law of waiting times: a system perfectly keeping a memory in giving flares

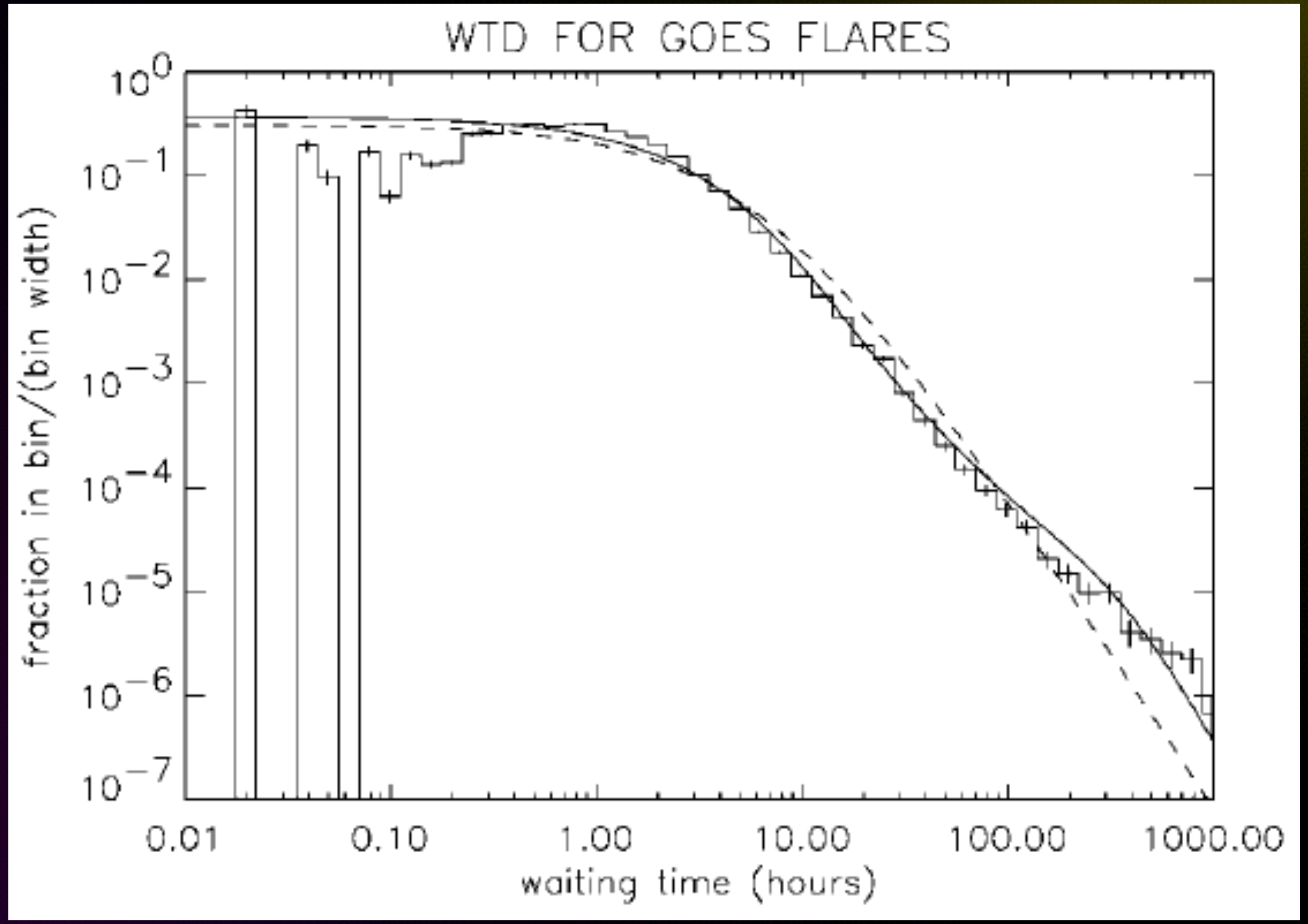
HOWEVER, ARE FLARES RANDOM? - DISTRIBUTION OF WAITING TIMES



Crosby, PhD Thesis (1996)



Bofetta et al., (1999)



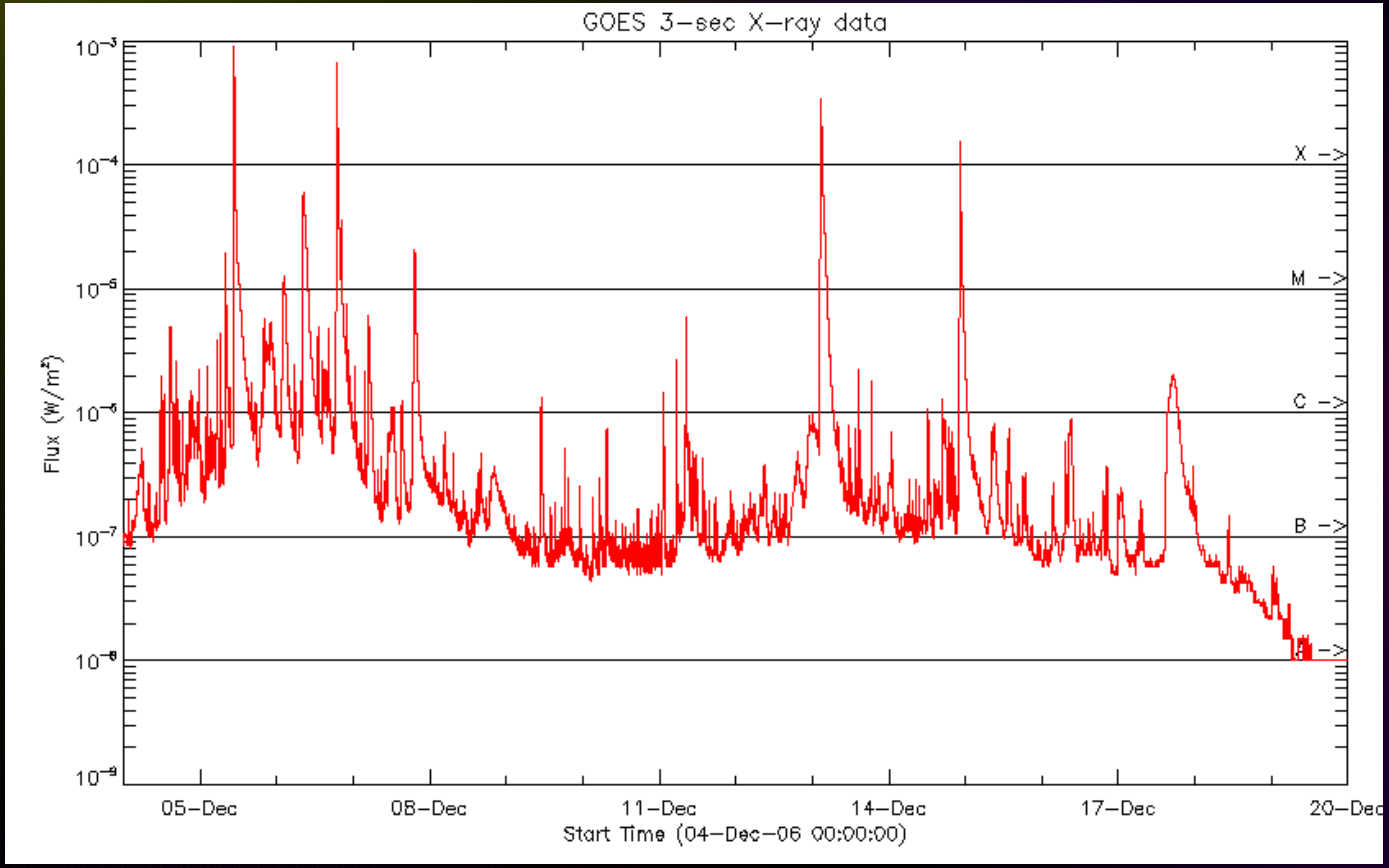
Wheatland (2000)

Exponential law of waiting times: a totally random, memoryless flare occurrence along the classical self-organized criticality concept

Robust power-law of waiting times: a system perfectly keeping a memory in giving flares

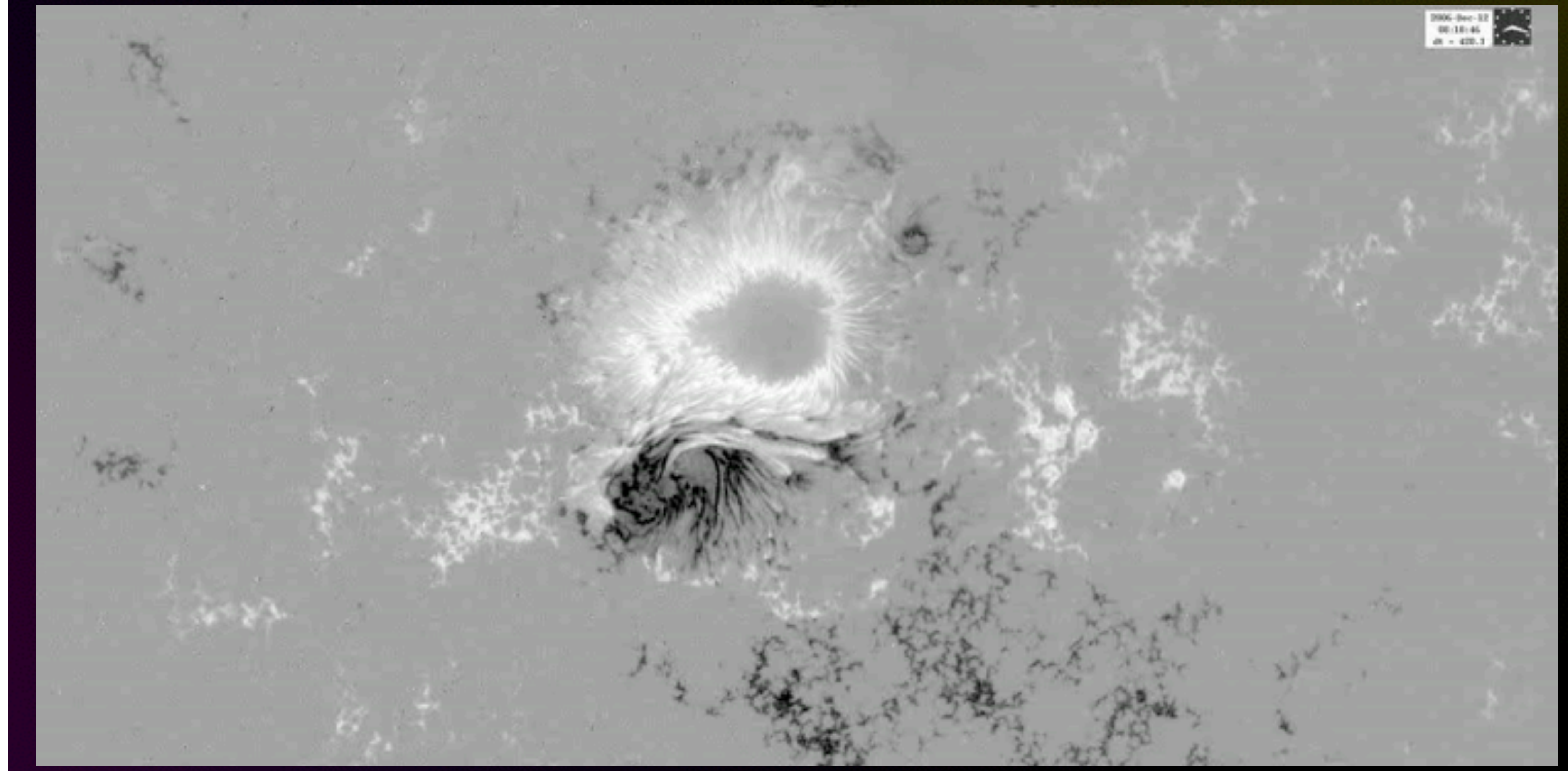
Time-dependent Poisson scaling in waiting times: some memory kept, with stochasticity demonstrated in an exponential distribution of different flaring rates

A MIX OF STOCHASTICITY AND MEMORY



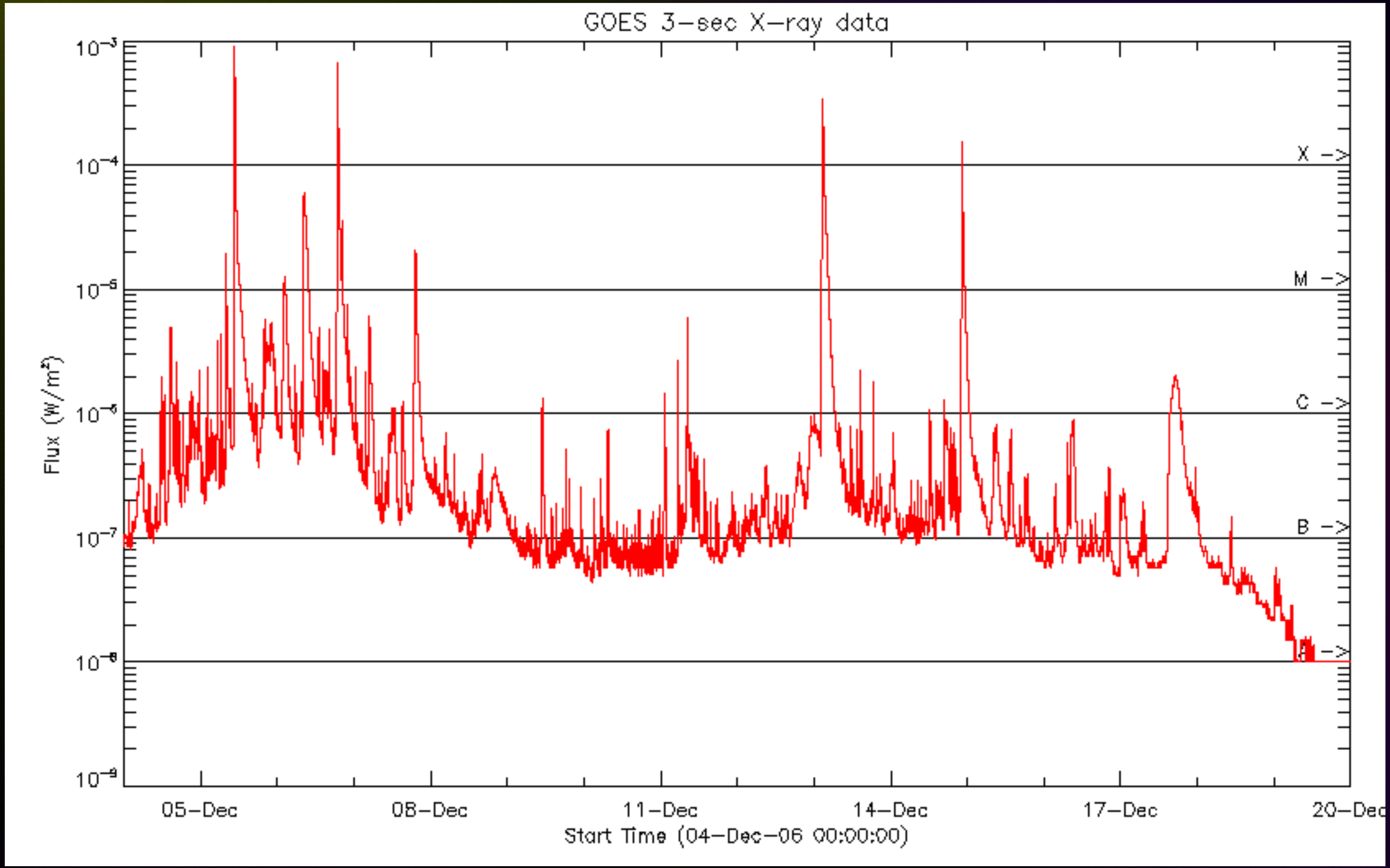
Response of NOAA AR 10930 over a two-week period in Dec 2006

- NOAA AR 10930
- Period observed: ~16 days



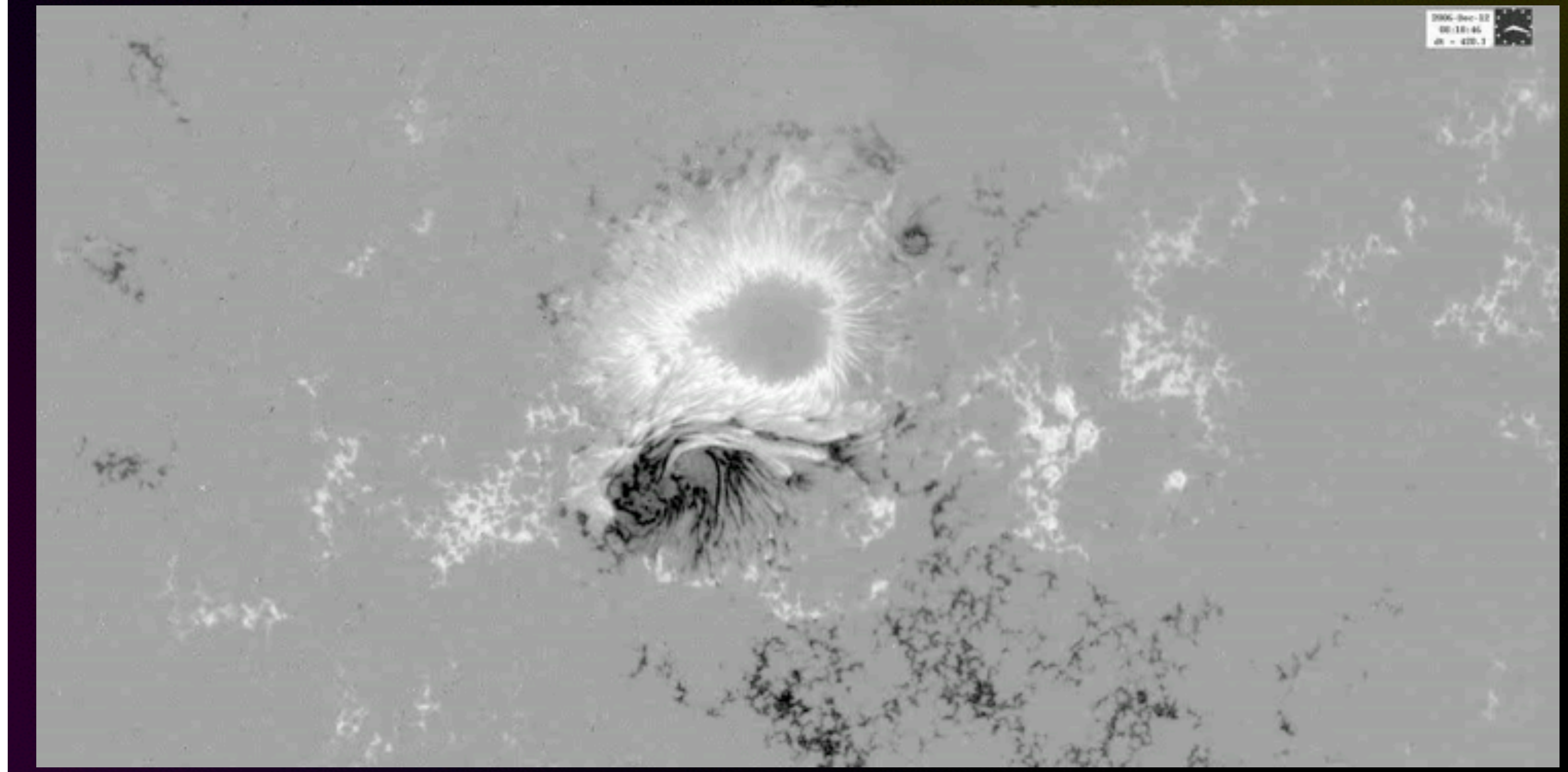
- Clustering of flares in a flaring active region
- Flaring features of active regions, i.e., complex magnetic PILs, continuously and consistently driven
- Typical situation of a pink-noise dynamical response timeseries

A MIX OF STOCHASTICITY AND MEMORY



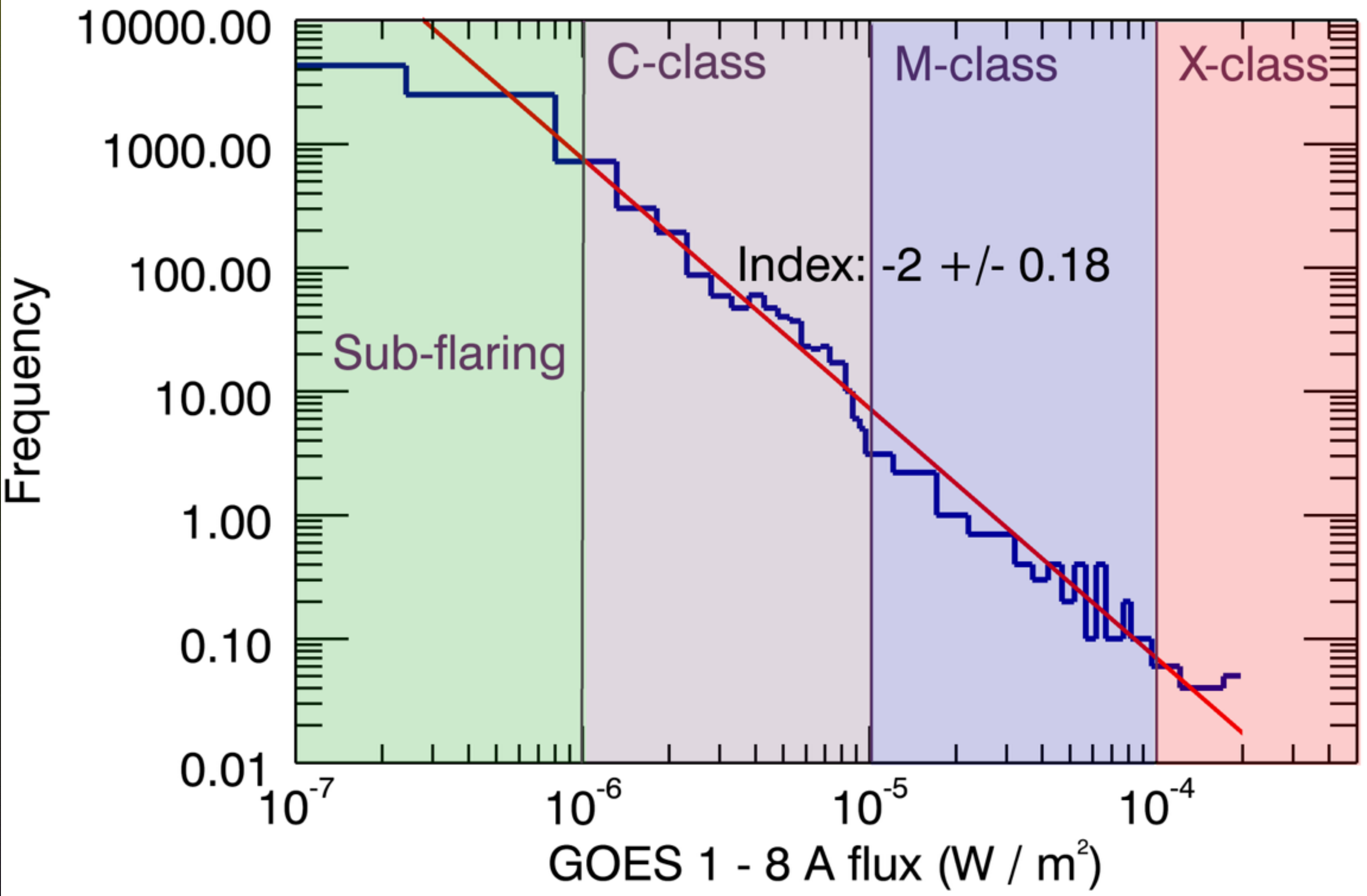
Response of NOAA AR 10930 over a two-week period in Dec 2006

- NOAA AR 10930
- Period observed: ~16 days



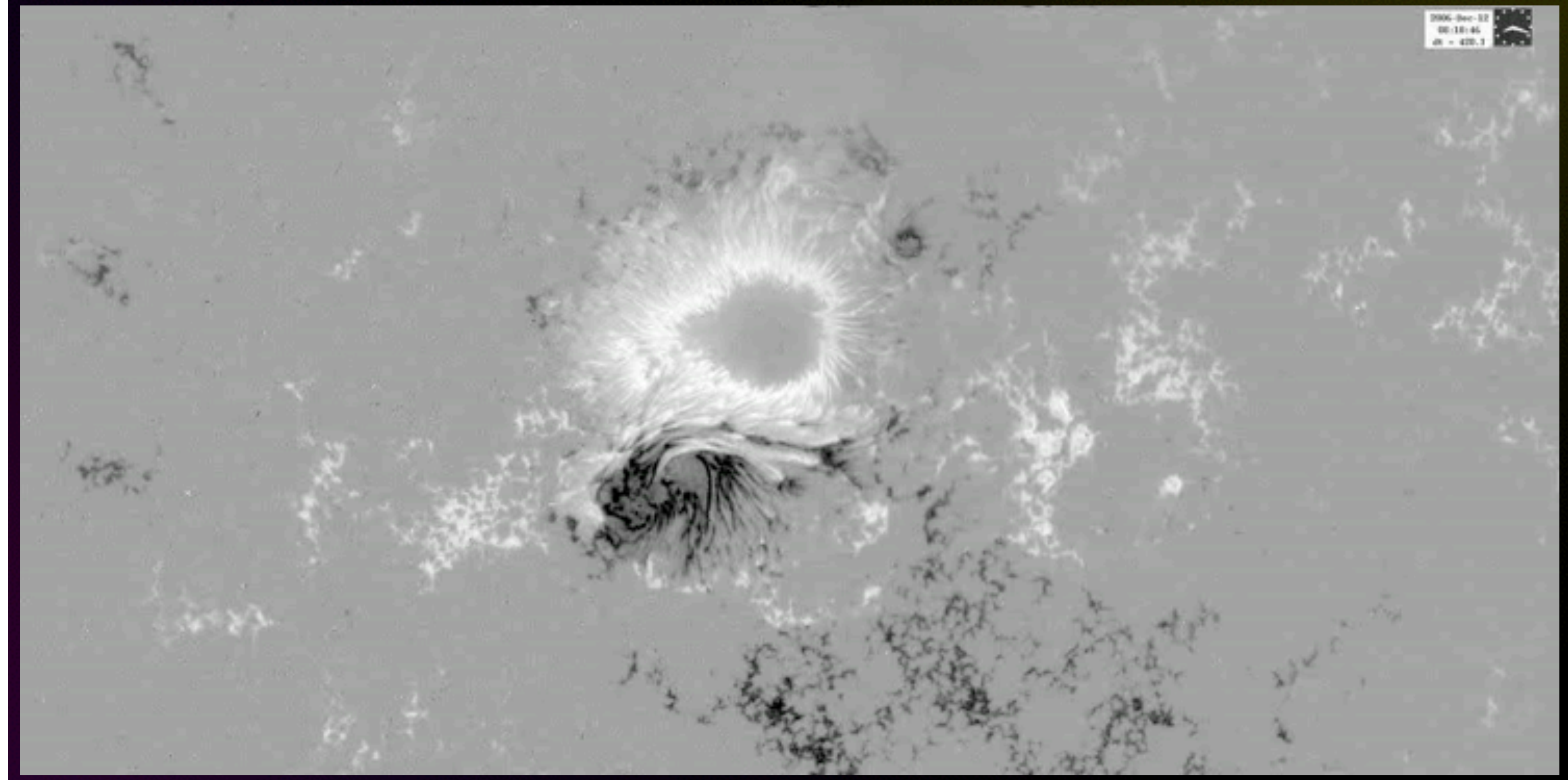
- Clustering of flares in a flaring active region
- Flaring features of active regions, i.e., complex magnetic PILs, continuously and consistently driven
- Typical situation of a pink-noise dynamical response timeseries

A MIX OF STOCHASTICITY AND MEMORY



Response of NOAA AR 10930 over a two-week period in Dec 2006

- NOAA AR 10930
- Period observed: ~16 days



- Clustering of flares in a flaring active region
- Flaring features of active regions, i.e., complex magnetic PILs, continuously and consistently driven
- Typical situation of a pink-noise dynamical response timeseries

QUALITATIVE COMPLEXITY CLASSIFICATION

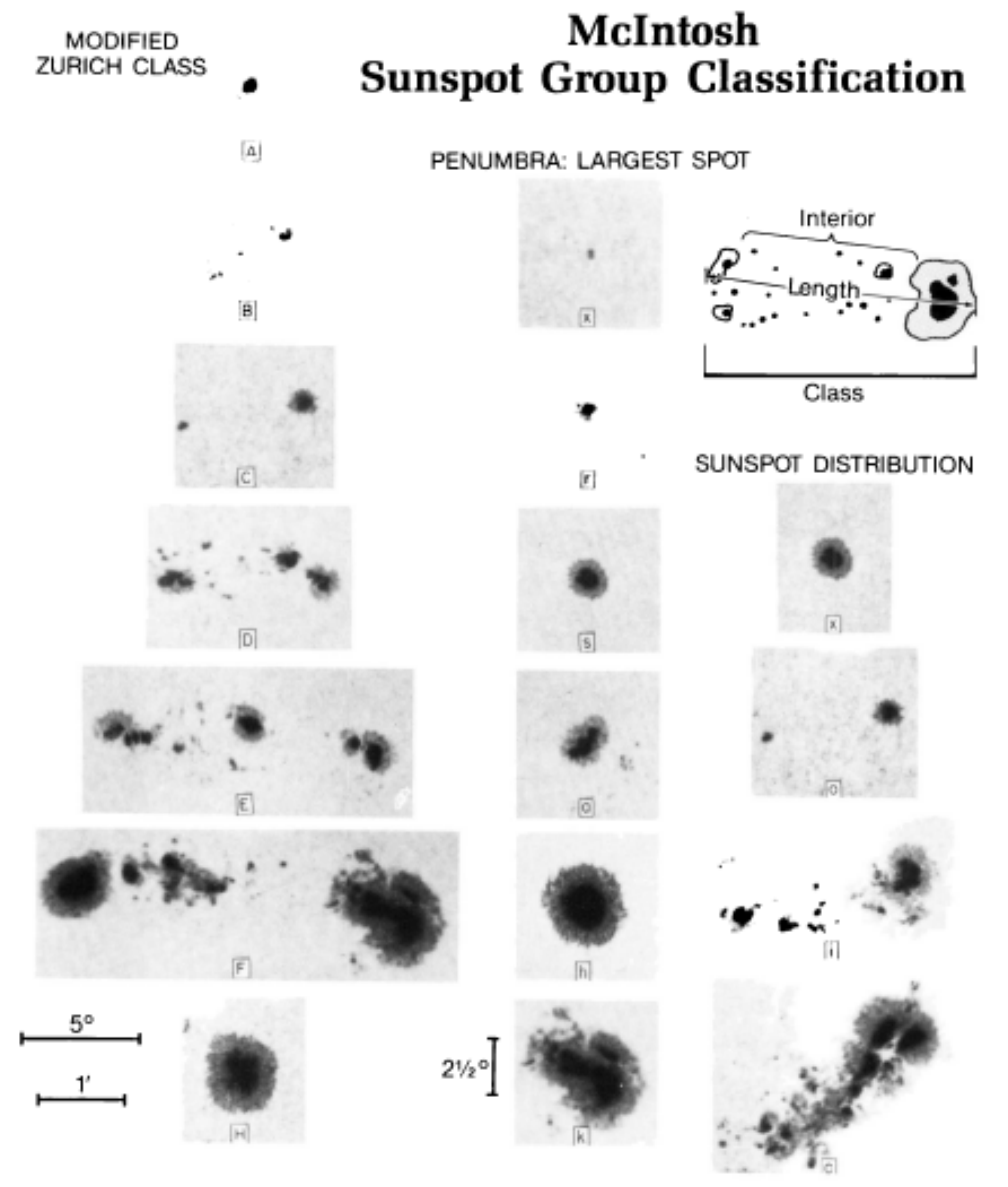


Fig. 1. The 3-component McIntosh classification, with examples of each category.

McIntosh (1990)

Mount Wilson classification

alpha: A unipolar sunspot group.

beta: A sunspot group having both positive and negative magnetic polarities (bipolar), with a simple and distinct division between the polarities.

gamma: A complex active region in which the positive and negative polarities are so irregularly distributed as to prevent classification as a bipolar group.

beta-gamma: A sunspot group that is bipolar but which is sufficiently complex that no single, continuous line can be drawn between spots of opposite polarities.

delta: A qualifier to magnetic classes(see below) indicating that umbrae separated by less than 2 degrees within one penumbra have opposite polarity.

beta-delta: A sunspot group of general beta magnetic classification but containing one (or more) delta spot(s).

beta-gamma-delta: A sunspot group of beta-gamma magnetic classification but containing one (or more) delta spot(s).

gamma-delta: A sunspot group of gamma magnetic classification but containing one (or more) delta spot(s).

Source: spaceweather.com

QUALITATIVE COMPLEXITY CLASSIFICATION

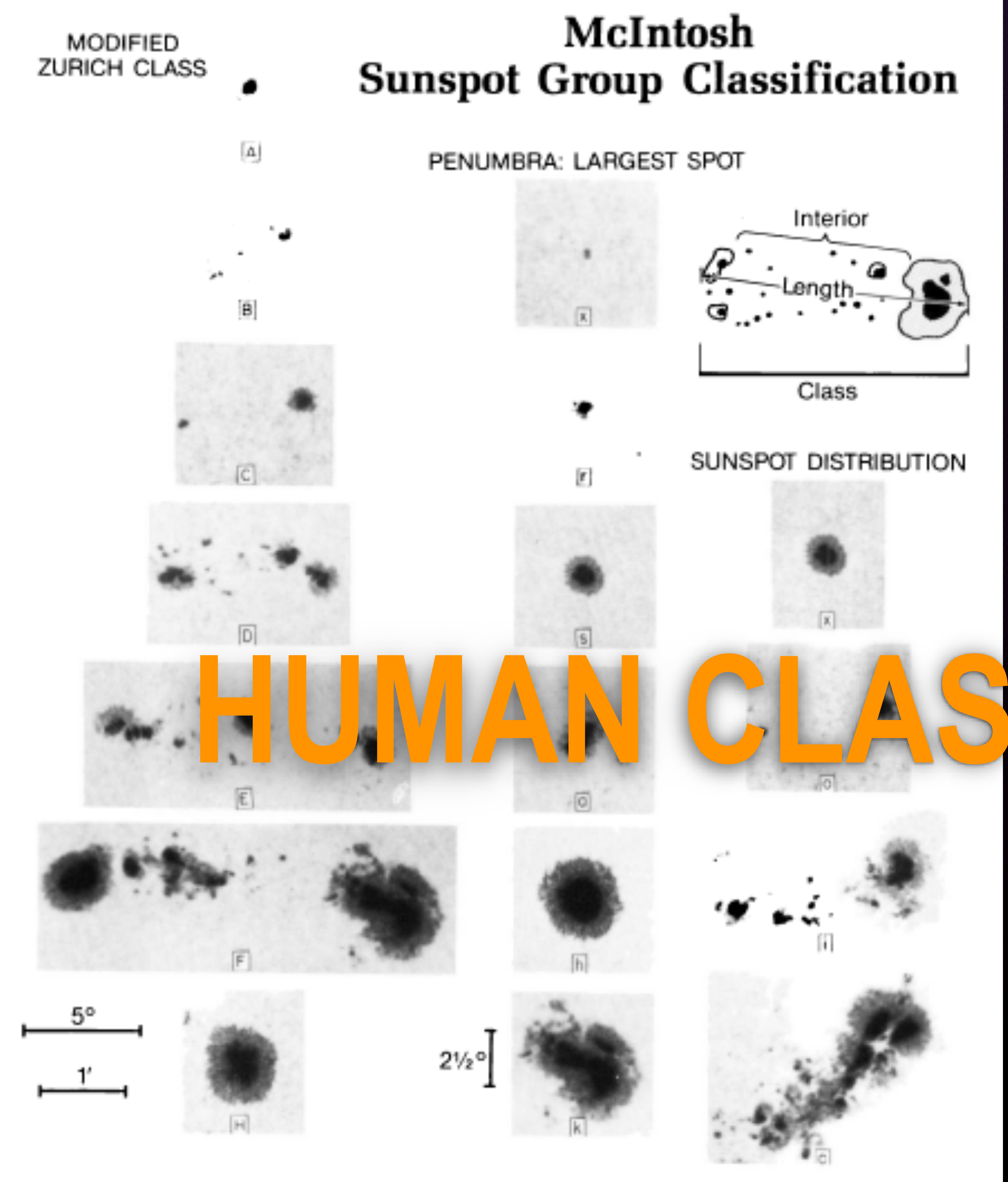


Fig. 1. The 3-component McIntosh classification, with examples of each category.

McIntosh (1990)

Mount Wilson classification

alpha: A unipolar sunspot group.

beta: A sunspot group having both positive and negative magnetic polarities (bipolar), with a simple and distinct division between the polarities.

gamma: A complex active region in which the positive and negative polarities are so irregularly distributed as to prevent classification as a bipolar group.

beta-gamma: A sunspot group that is bipolar but which is sufficiently complex that no single, continuous line can be drawn between spots of opposite polarities.

delta: A qualifier to magnetic classes (see below) indicating that umbrae separated by less than 2 degrees within one penumbra have opposite polarity.

beta-delta: A sunspot group of general beta magnetic classification but containing one (or more) delta spot(s).

beta-gamma-delta: A sunspot group of beta-gamma magnetic classification but containing one (or more) delta spot(s).

gamma-delta: A sunspot group of gamma magnetic classification but containing one (or more) delta spot(s).

Source: spaceweather.com

HUMAN CLASSIFICATION, SUBJECTIVE

QUANTITATIVE COMPLEXITY CLASSIFICATION

Numerous methods over the past 20 years. An effort to categorize them results in the following (*Georgoulis, 2012*):

QUANTITATIVE COMPLEXITY CLASSIFICATION

Numerous methods over the past 20 years. An effort to categorize them results in the following (*Georgoulis, 2012*):

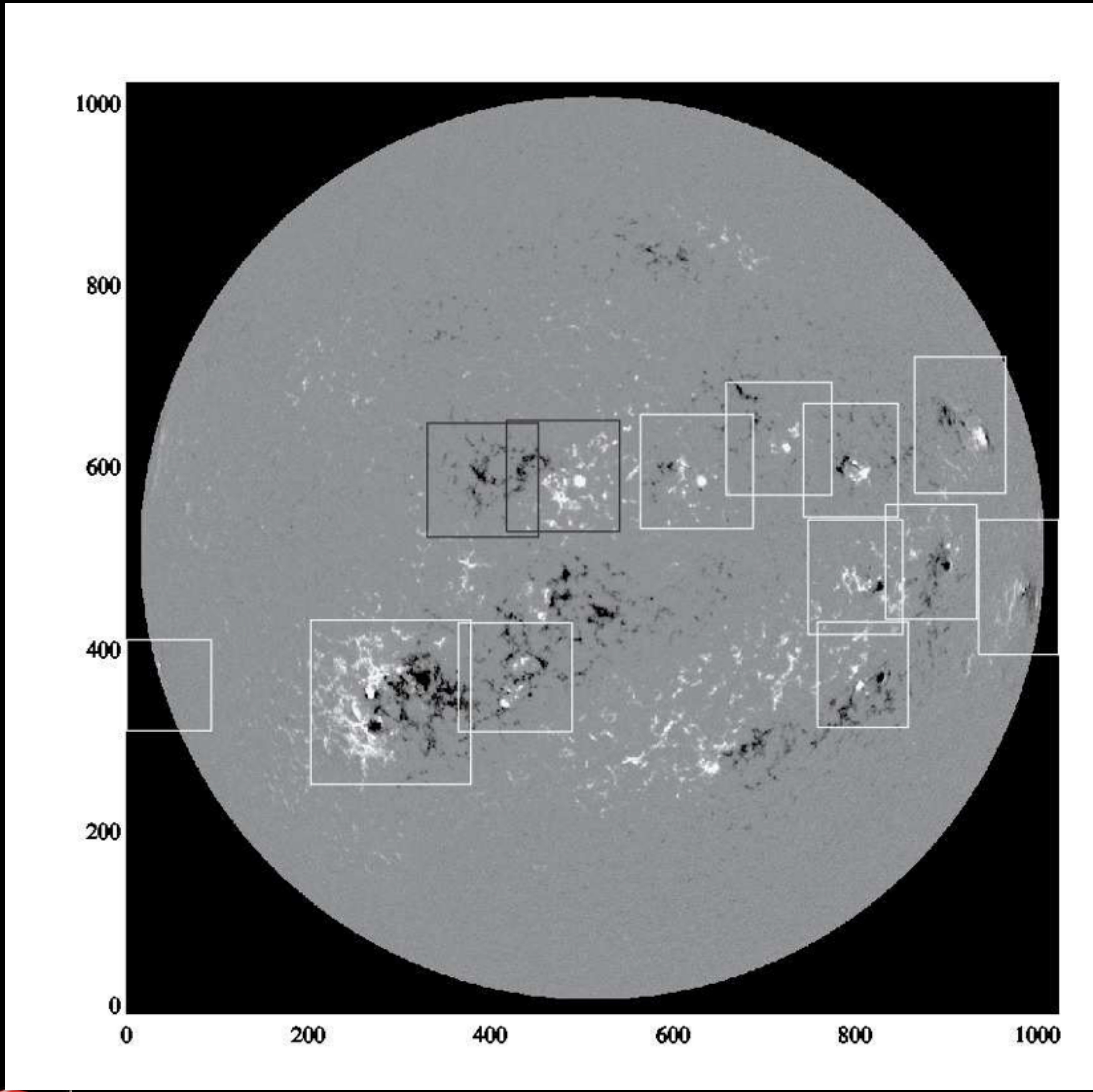
- Monoscale / multiscale methods
- Morphological methods
- Statistical methods (on historical & archived data)
- Machine-learning, combinatorial, & assimilation methods
- Analytical methods
- Local helioseismology methods
- Other (slightly exotic) methods

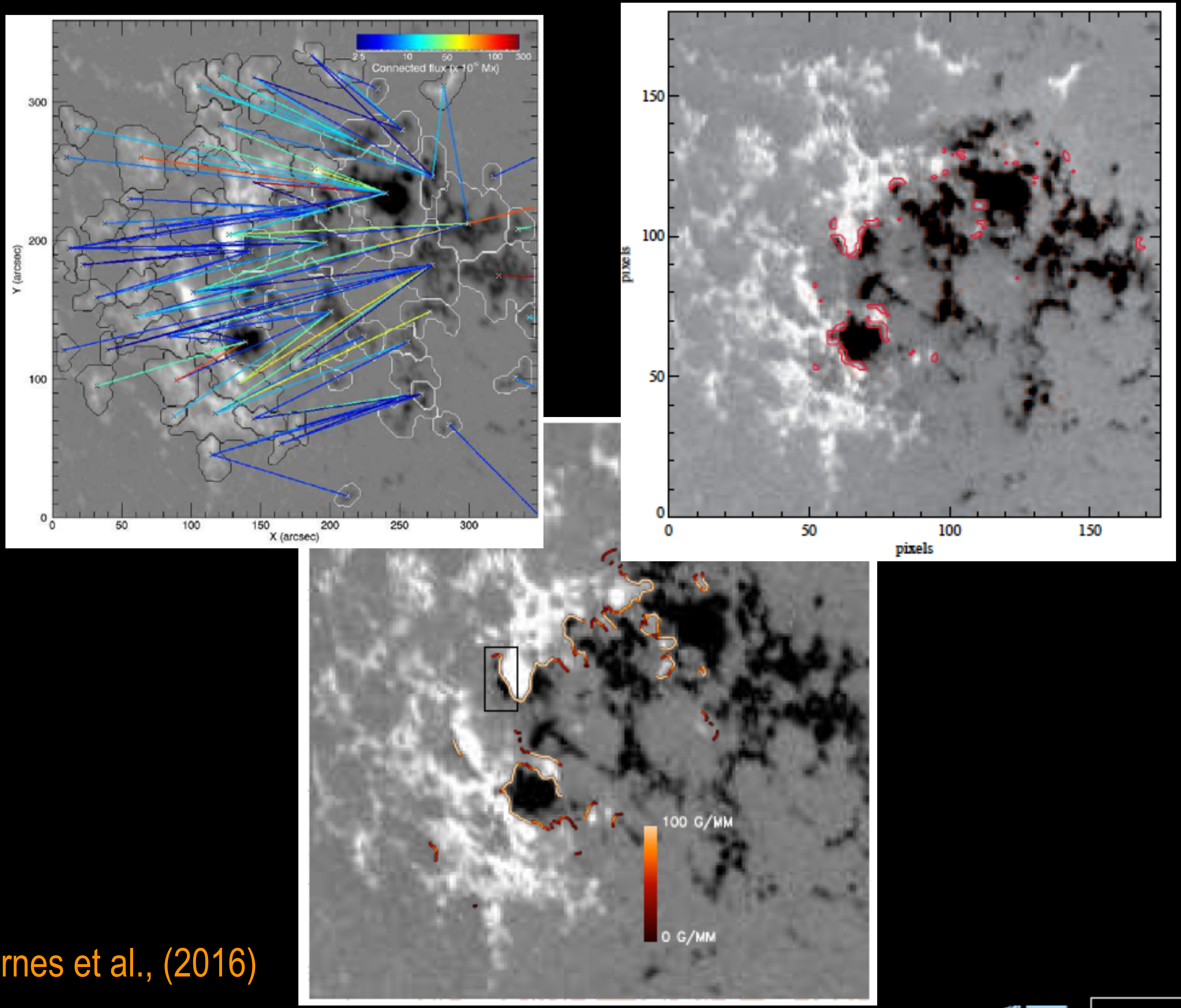
QUANTITATIVE COMPLEXITY CLASSIFICATION

Numerous methods over the past 20 years. An effort to categorize them results in the following (*Georgoulis, 2012*):

- Monoscale / multiscale methods — Abramenko et al. (2002, 2003); McAteer at al. (2005); Georgoulis (2005, 2012); Uritsky et al. (2007, 2013); Hewett et al. (2008); Conlon et al. (2010); Kestener et al. (2010), McAteer (2015)
- Morphological methods — Falconer et al. (2001, 2002, 2003, 2008, 2009, 2011); Georrgoulis & Rust (2007); Schrijver (2007); Mason & Hoeksema (2010); Leka & Barnes (2003a; b); Cabnfield et al. (1999); Barnes & Leka (2008), Korsos et al. (2015)
- Statistical methods (on historical & archived data) — Wheatland (2001); Moon et al. (2001); Gallagher et al. (2002); Wheatland (2004, 2005a, b)
- Machine-learning, combinatorial, & assimilation methods — Belanger et al. (2007); Qahwaji & Colak (2007); Colak & Qahwaji (2008, 2009); Qahwaji et al. (2008); Al-Omari et al. (2010); Yu et al. (2009; 2010a, b); Huang et al. (2010) ; Bobra & Couvidat (2014); Bobra & Ilonidis (2015); Boucheron et al., (2015); Nishizuka et al., (2016)
- Analytical methods — Wheatland & Glukhov (1998); Wheatland (2008)
- Local helioseismology methods — Reinard et al. (2010); Komm et al. (2011), etc.
- Other (slightly exotic) methods — Jenkins & Fischbach (2009); Javorsek et al. (2012); Strugarek & Charbonneau (2014)

ANALYSIS OF PHOTOSPHERIC ACTIVE-REGION MAGNETOGRAMS



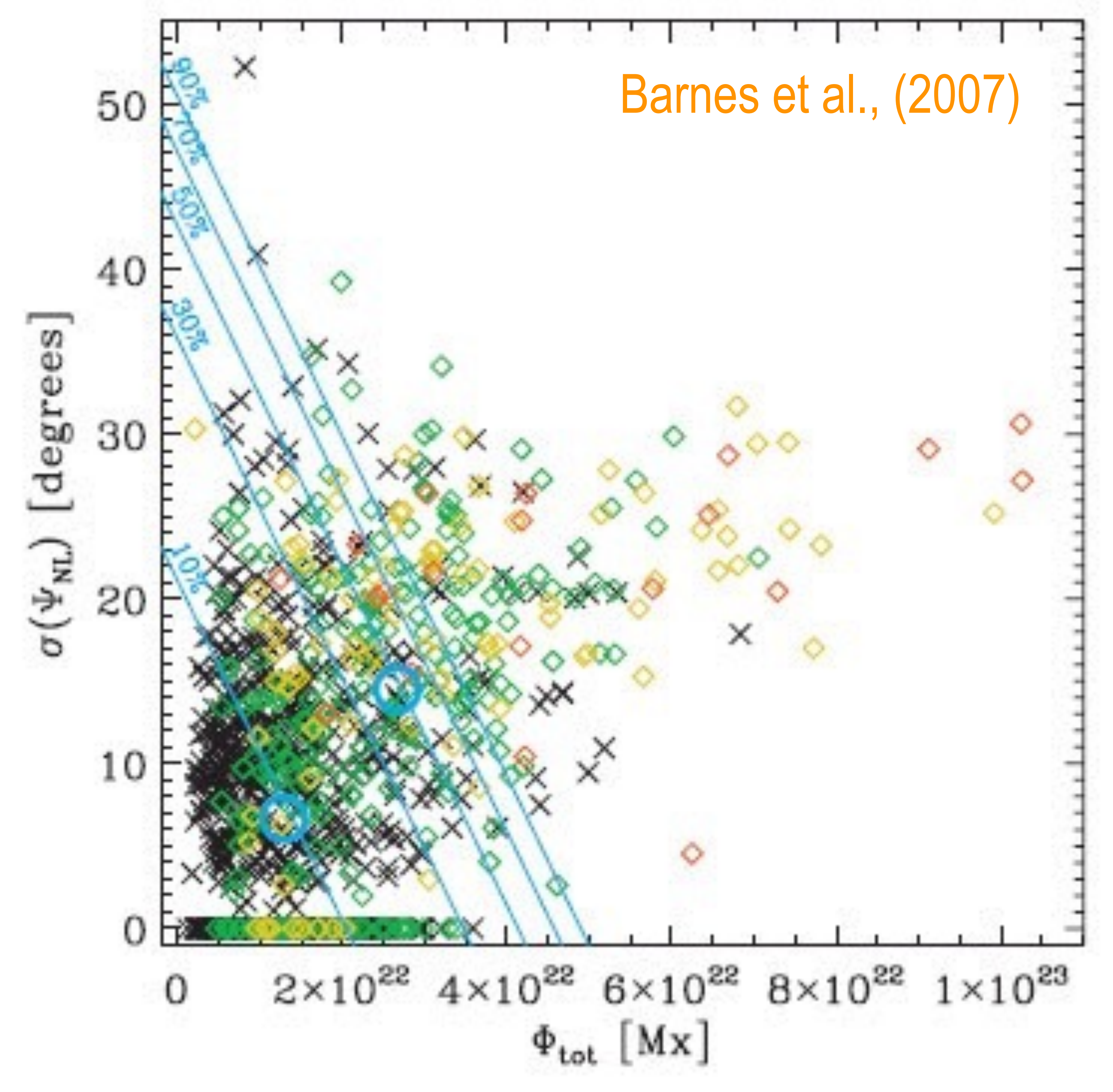


Barnes et al., (2016)

PROPERTIES TRANSLATED TO PREDICTIVE PROBABILITIES

Keyword	Description	Formula	F-Score	Selection
TOTUSJH	Total unsigned current helicity	$H_{c_{\text{total}}} \propto \sum \lvert B_z \cdot J_z \rvert$	3560	Included
TOTBSQ	Total magnitude of Lorentz force	$F \propto \sum B^2$	3051	Included
TOTPOT	Total photospheric magnetic free energy density	$\rho_{\text{tot}} \propto \sum \left(\boldsymbol{B}^{\text{Obs}} - \boldsymbol{B}^{\text{Pot}}\right)^2 dA$	2996	Included
TOTUSJZ	Total unsigned vertical current	$J_{z_{\text{total}}} = \sum \lvert J_z \rvert dA$	2733	Included
ABSNJZH	Absolute value of the net current helicity	$H_{c_{\text{abs}}} \propto \left\lvert \sum B_z \cdot J_z \right\rvert$	2618	Included
SAVNCPP	Sum of the modulus of the net current per polarity	$J_{z_{\text{sum}}} \propto \left\lvert \sum^{B_z^+} J_z dA \right\rvert + \left\lvert \sum^{B_z^-} J_z dA \right\rvert$	2448	Included
USFLUX	Total unsigned flux	$\Phi = \sum \lvert B_z \rvert dA$	2437	Included
AREA_ACR	Area of strong field pixels in the active region	Area $= \sum$ Pixels	2047	Included
TOTFZ	Sum of z-component of Lorentz force	$F_z \propto \sum (B_x^2 + B_y^2 - B_z^2) dA$	1371	Included
MEANPOT	Mean photospheric magnetic free energy	$\overline{\rho} \propto \frac{1}{N} \sum \left(\boldsymbol{B}^{\text{Obs}} - \boldsymbol{B}^{\text{Pot}}\right)^2$	1064	Included
R_VALUE	Sum of flux near polarity inversion line	$\Phi = \sum \lvert B_{LoS} \rvert dA$ within R mask	1057	Included
EPSZ	Sum of z-component of normalized Lorentz force	$\delta F_z \propto \frac{\sum (B_x^2 + B_y^2 - B_z^2)}{\sum B^2}$	864.1	Included
SHRGT45	Fraction of Area with shear $> 45°$	Area with shear $> 45°$ / total area	740.8	Included
MEANSHR	Mean shear angle	$\overline{\Gamma} = \frac{1}{N} \sum \arccos\left(\frac{\boldsymbol{B}^{\text{Obs}} \cdot \boldsymbol{B}^{\text{Pot}}}{\lvert \boldsymbol{B}^{\text{Obs}} \rvert \lvert \boldsymbol{B}^{\text{Pot}} \rvert}\right)$	727.9	Discarded
MEANGAM	Mean angle of field from radial	$\overline{\gamma} = \frac{1}{N} \sum \arctan\left(\frac{B_h}{B_z}\right)$	573.3	Discarded
MEANGBT	Mean gradient of total field	$\overline{\lvert \nabla B_{\text{tot}} \rvert} = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2}$	192.3	Discarded
MEANGBZ	Mean gradient of vertical field	$\overline{\lvert \nabla B_z \rvert} = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B_z}{\partial x}\right)^2 + \left(\frac{\partial B_z}{\partial y}\right)^2}$	88.40	Discarded
MEANGBH	Mean gradient of horizontal field	$\overline{\lvert \nabla B_h \rvert} = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B_h}{\partial x}\right)^2 + \left(\frac{\partial B_h}{\partial y}\right)^2}$	79.40	Discarded
MEANJZH	Mean current helicity (B_z contribution)	$\overline{H_c} \propto \frac{1}{N} \sum B_z \cdot J_z$	46.73	Discarded
TOTFY	Sum of y-component of Lorentz force	$F_y \propto \sum B_y B_z dA$	28.92	Discarded
MEANJZD	Mean vertical current density	$\overline{J_z} \propto \frac{1}{N} \sum \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)$	17.44	Discarded
MEANALP	Mean characteristic twist parameter, α	$\alpha_{\text{total}} \propto \frac{\sum J_z \cdot B_z}{\sum B_z^2}$	10.41	Discarded
TOTFX	Sum of x-component of Lorentz force	$F_x \propto -\sum B_x B_z dA$	6.147	Discarded
EPSY	Sum of y-component of normalized Lorentz force	$\delta F_y \propto \frac{-\sum B_y B_z}{\sum B^2}$	0.647	Discarded
EPSX	Sum of x-component of normalized Lorentz force	$\delta F_x \propto \frac{\sum B_x B_z}{\sum B^2}$	0.366	Discarded

Bobra & Couvidat (2014)



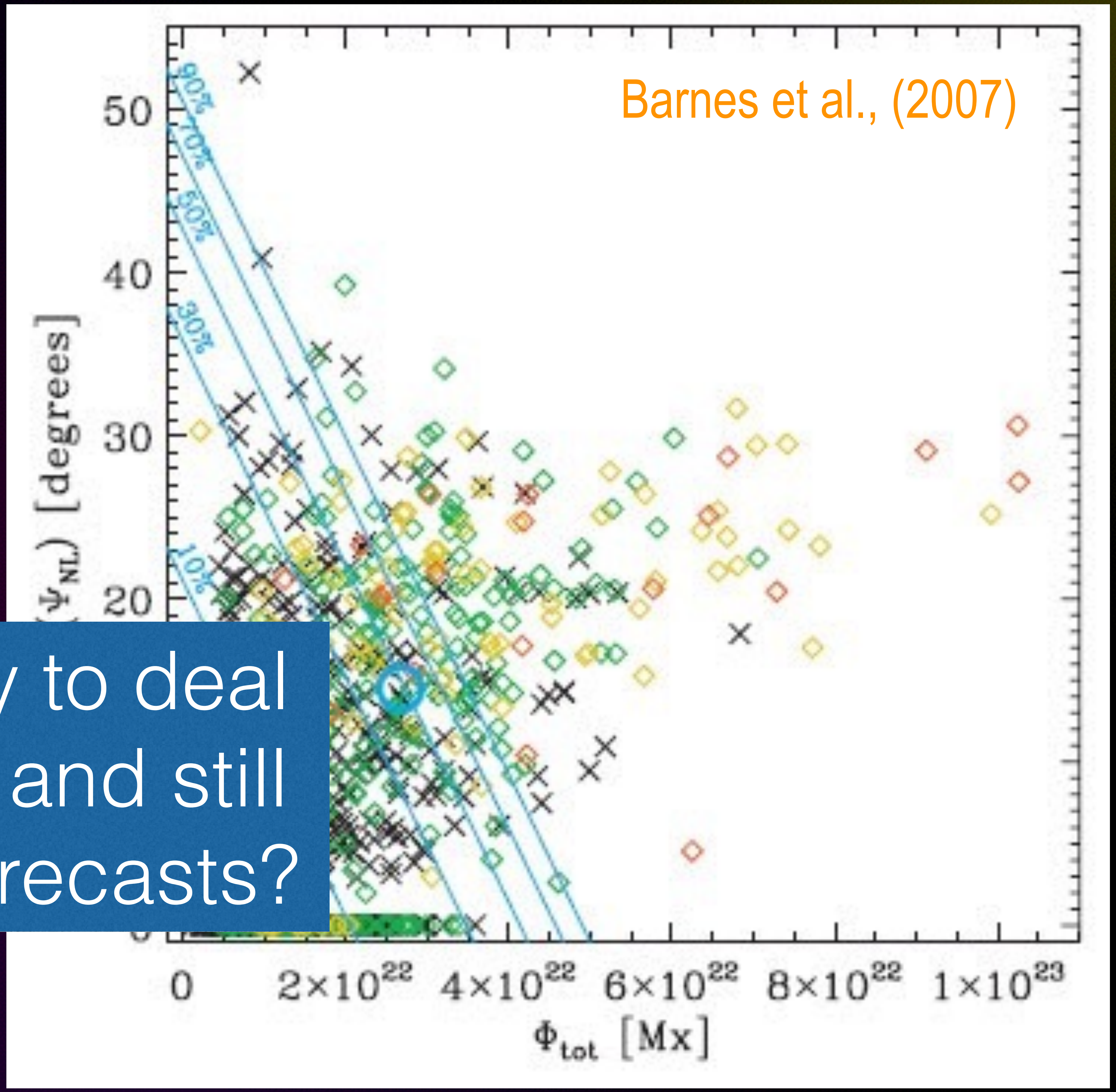
Discriminant analysis: Two-function, linear DA for four-class prediction (non-flaring, C, M, and X-class)

PROPERTIES TRANSLATED TO PREDICTIVE PROBABILITIES

Keyword	Description	Formula	F-Score	Selection
TOTUSJH	Total unsigned current helicity	$H_{c_{\rm total}} \propto \sum \lvert B_z \cdot J_z \rvert$	3560	Included
TOTBSQ	Total magnitude of Lorentz force	$F \propto \sum B^2$	3051	Included
TOTPOT	Total photospheric magnetic free energy density	$\rho_{\rm tot} \propto \sum \left(\boldsymbol{B}^{\rm Obs} - \boldsymbol{B}^{\rm Pot}\right)^2 dA$	2996	Included
TOTUSJZ	Total unsigned vertical current	$J_{z_{\rm total}} = \sum \lvert J_z \rvert dA$	2733	Included
ABSNJZH	Absolute value of the net current helicity	$H_{c_{\rm abs}} \propto \left\lvert \sum B_z \cdot J_z \right\rvert$	2618	Included
SAVNCPP	Sum of the modulus of the net current per polarity	$J_{z_{\rm sum}} \propto \left\lvert \sum^{B_z^+} J_z dA \right\rvert + \left\lvert \sum^{B_z^-} J_z dA \right\rvert$	2448	Included
USFLUX	Total unsigned flux	$\Phi = \sum \lvert B_z \rvert dA$	2437	Included
AREA_ACR	Area of strong field pixels in the active region	$\mathrm{Area} = \sum \mathrm{Pixels}$	2047	Included
TOTFZ	Sum of z-component of Lorentz force	$F_z \propto \sum (B_x^2 + B_y^2 - B_z^2) dA$	1371	Included
MEANPOT	Mean photospheric magnetic free energy	$\overline{\rho} \propto \frac{1}{N} \sum \left(\boldsymbol{B}^{\rm Obs} - \boldsymbol{B}^{\rm Pot}\right)^2$	1064	Included
R_VALUE	Sum of flux near polarity inversion line	$\Phi = \sum \lvert B_{LoS} \rvert dA$ within R mask	1057	Included
EPSZ	Sum of z-component of normalized Lorentz force	$\delta F_z \propto \frac{\sum (B_x^2 + B_y^2 - B_z^2)}{\sum B^2}$	864.1	Included
SHRGT45	Fraction of Area with shear $> 45°$	Area w[illegible]	[illegible]	[illegible]
MEANSHR	Mean shear angle	$\overline{\Gamma} =$ [illegible]	[illegible]	[illegible]
MEANGAM	Mean angle of field from radial	$\overline{\gamma}$ [illegible]	[illegible]	[illegible]
MEANGBT	Mean gradient of total field	$\overline{\lvert \nabla B_{\rm tot} \rvert}$ [illegible]	[illegible]	[illegible]
MEANGBZ	Mean gradient of vertical field	$\overline{\lvert \nabla B_z \rvert} =$ [illegible]	[illegible]	[illegible]
MEANGBH	Mean gradient of horizontal field	$\overline{\lvert \nabla B_h \rvert} =$ [illegible]	[illegible]	[illegible]
MEANJZH	Mean current helicity (B_z contribution)	$\overline{H_c} \propto \frac{1}{N} \sum B_z \cdot J_z$	46.73	Discarded
TOTFY	Sum of y-component of Lorentz force	$F_y \propto \sum B_y B_z dA$	28.92	Discarded
MEANJZD	Mean vertical current density	$\overline{J_z} \propto \frac{1}{N} \sum \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)$	17.44	Discarded
MEANALP	Mean characteristic twist parameter, α	$\alpha_{\rm total} \propto \frac{\sum J_z \cdot B_z}{\sum B_z^2}$	10.41	Discarded
TOTFX	Sum of x-component of Lorentz force	$F_x \propto -\sum B_x B_z dA$	6.147	Discarded
EPSY	Sum of y-component of normalized Lorentz force	$\delta F_y \propto \frac{-\sum B_y B_z}{\sum B^2}$	0.647	Discarded
EPSX	Sum of x-component of normalized Lorentz force	$\delta F_x \propto \frac{\sum B_x B_z}{\sum B^2}$	0.366	Discarded

What is the optimal way to deal with all this information and still achieve reliable NRT forecasts?

Bobra & Couvidat (2014)



Discriminant analysis: Two-function, linear DA for four-class prediction (non-flaring, C, M, and X-class)

RECENT TREDS IN FLARE PREDICTION

- Most (excluding machine-learning) methods use a univariate predictor.

RECENT TREDS IN FLARE PREDICTION

- Most (excluding machine-learning) methods use a univariate predictor.
- Multivariate forecasting can also be used in the form of :
  - — Synthetic predictors:

$$predictor = \omega_1\ predictor_1 + \omega_2\ predictor_2 + \ldots + \ \omega_n\ predictor_n$$

$$\omega_1, \ \omega_2 \ ,\ldots, \ \omega_n \quad \text{unrestricted}$$

  - — Ensemble forecasting:

$$P(flare) = \omega_1\ P_1(flare) + \omega_2\ P_2(flare) + \ldots + \ \omega_n\ P_n(flare)$$

$$\sum_{i=1}^{n} \omega_i = 1$$

RECENT TREDS IN FLARE PREDICTION

- Most (excluding machine-learning) methods use a univariate predictor.
- Multivariate forecasting can also be used in the form of :

  — Synthetic predictors:

$$predictor = \omega_1\ predictor_1 + \omega_2\ predictor_2 + \ldots +\ \omega_n\ predictor_n$$

$\omega_1,\ \omega_2\ ,\ldots,\ \omega_n$ unrestricted

  — Ensemble forecasting:

$$P(flare) = \omega_1\ P_1(flare) + \omega_2\ P_2(flare) + \ldots +\ \omega_n\ P_n(flare)$$

$$\sum_{i=1}^{n} \omega_i = 1$$

- **Task:** find ω_1, ω_2, …, ω_n such that validation results are optimized

RECENT TREDS IN FLARE PREDICTION

- Most (excluding machine-learning) methods use a univariate predictor.
- Multivariate forecasting can also be used in the form of :

  — Synthetic predictors:

$$predictor = \omega_1 \ predictor_1 + \omega_2 \ predictor_2 + \ldots + \ \omega_n \ predictor_n$$

$\omega_1, \omega_2, \ldots, \omega_n$ unrestricted

  — Ensemble forecasting:

$$P(flare) = \omega_1 \ P_1(flare) + \omega_2 \ P_2(flare) + \ldots + \ \omega_n \ P_n(flare)$$

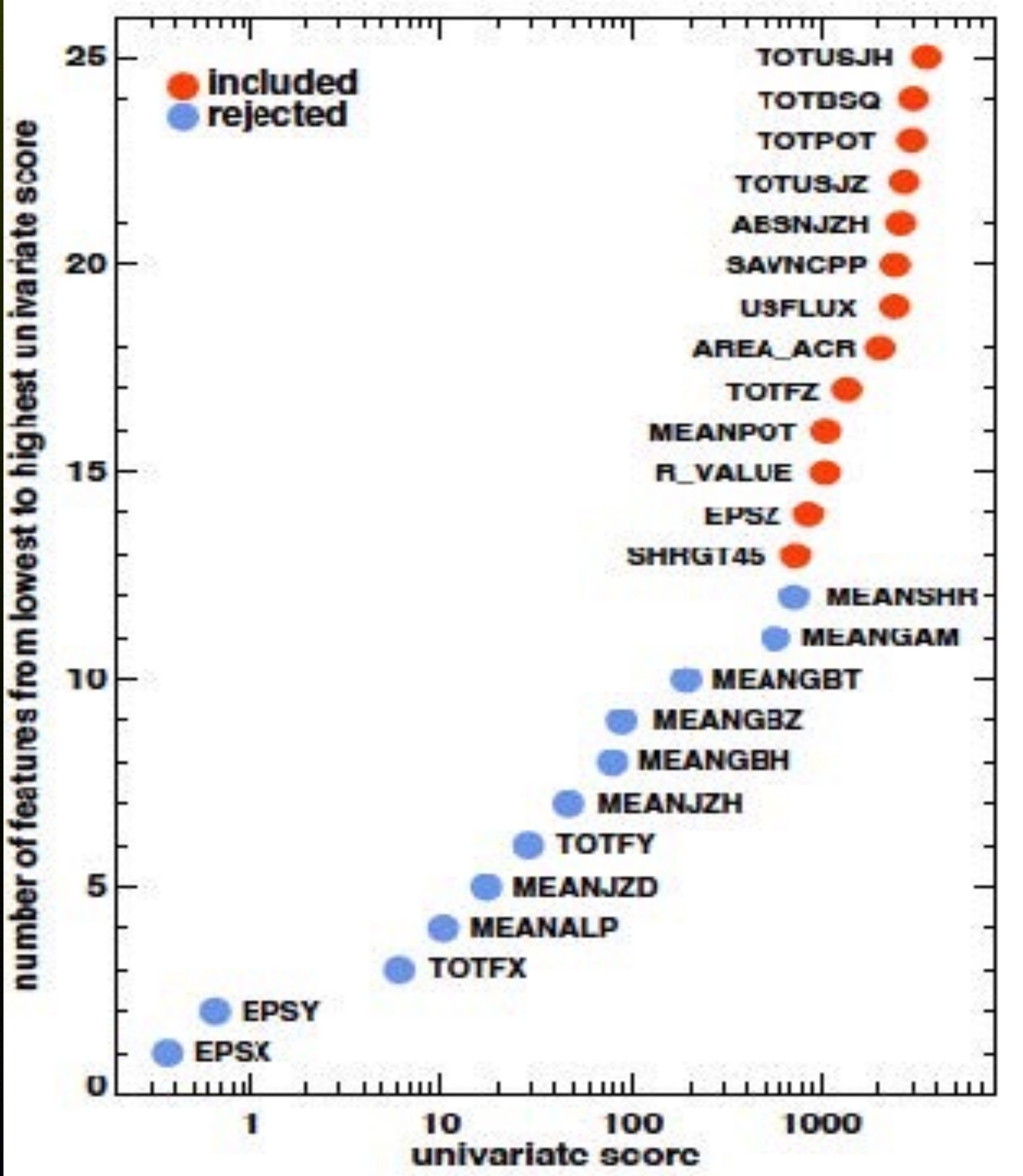
$$\sum_{i=1}^{n} \omega_i = 1$$

- **Task:** find ω_1, ω_2, …, ω_n such that validation results are <u>optimized</u>

However: optimization means different things to different communities!

INDICATIVE RESULTS

- Multivariate forecasting

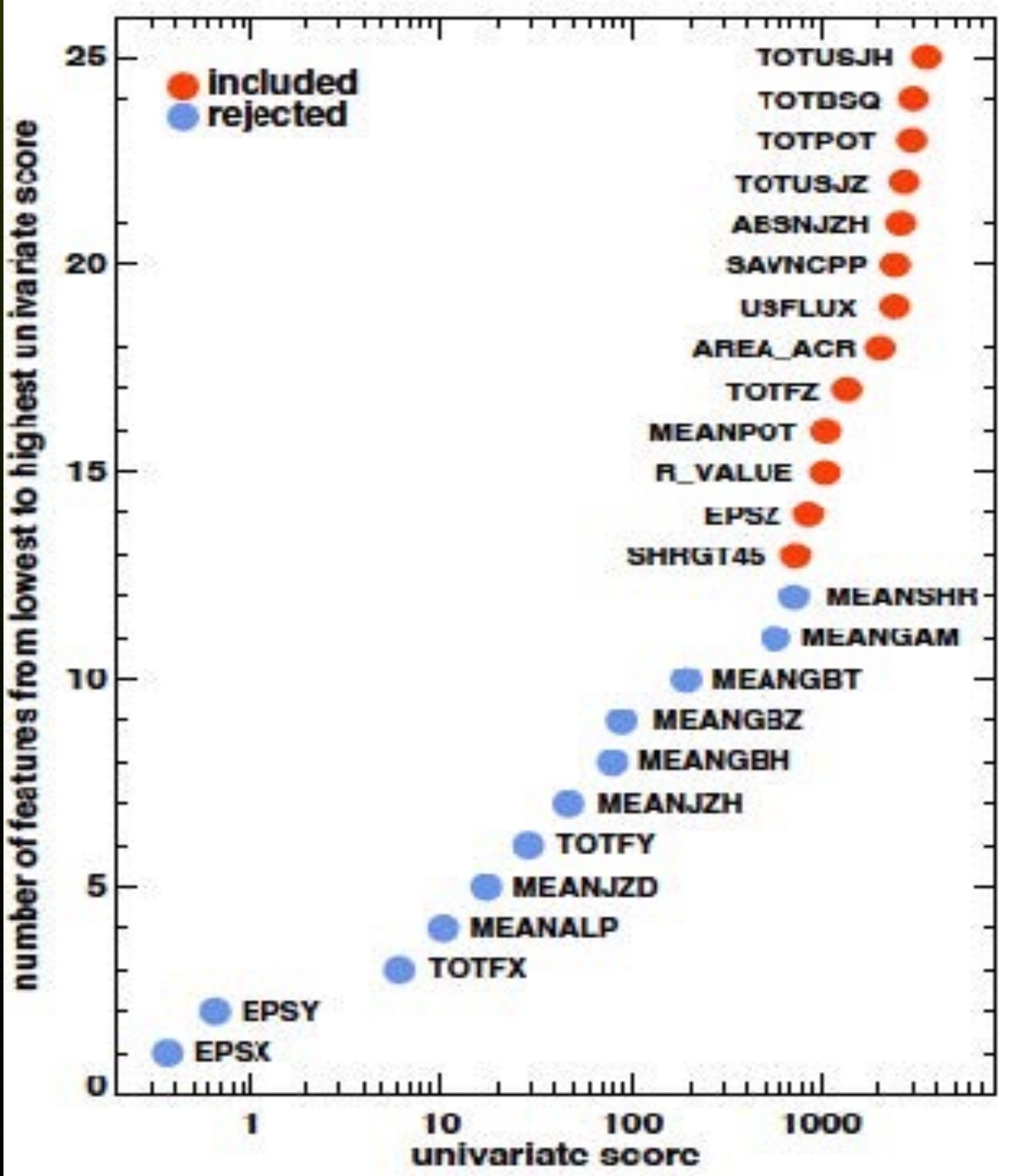


Bobra & Couvidat (2014)

- Ordering of predictors by means of a univariate Fisher ranking score
- Machine-learning classifiers adopted

INDICATIVE RESULTS

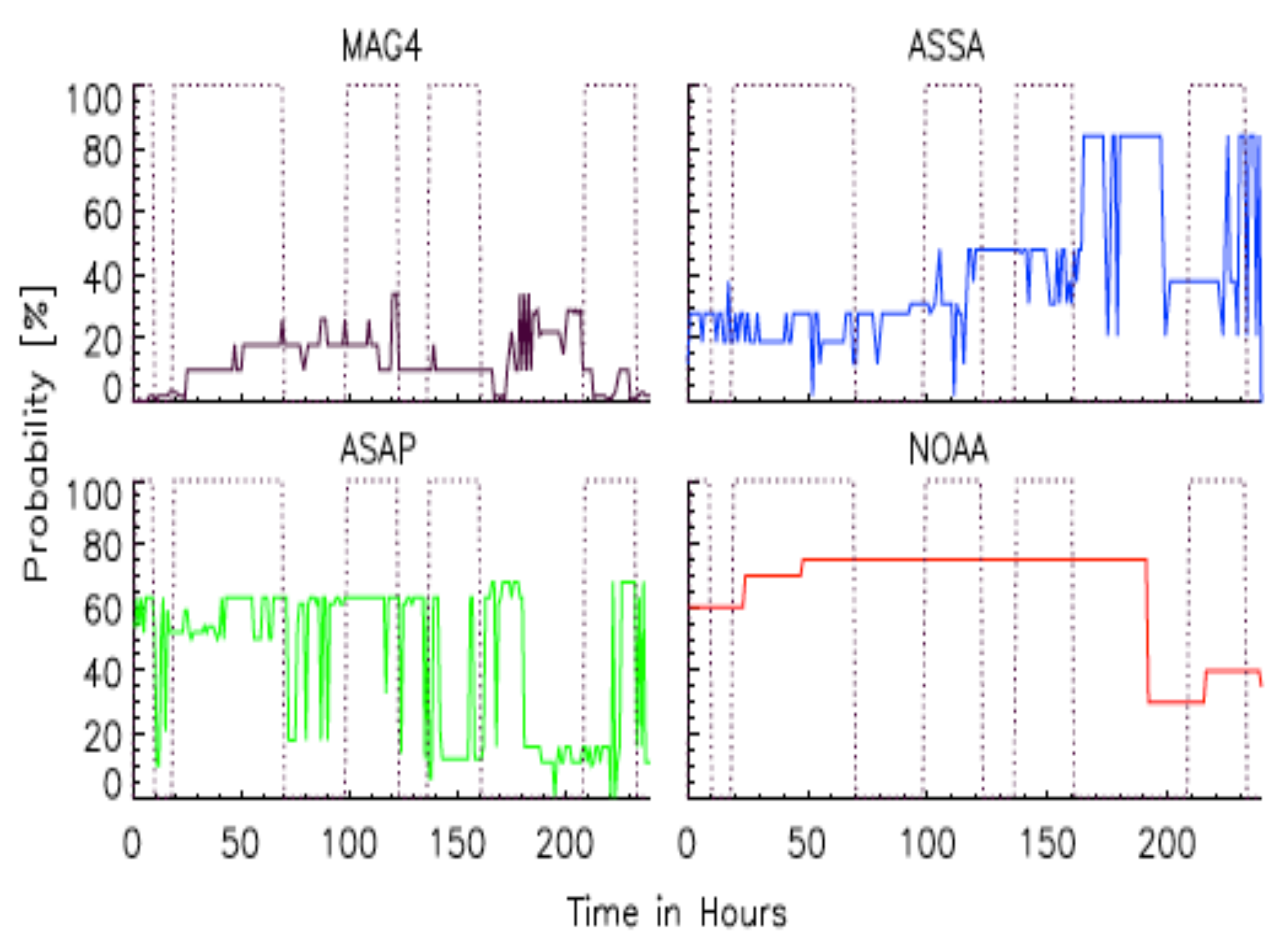
- Multivariate forecasting



Bobra & Couvidat (2014)

- Ordering of predictors by means of a univariate Fisher ranking score
- Machine-learning classifiers adopted

- Ensemble forecasting

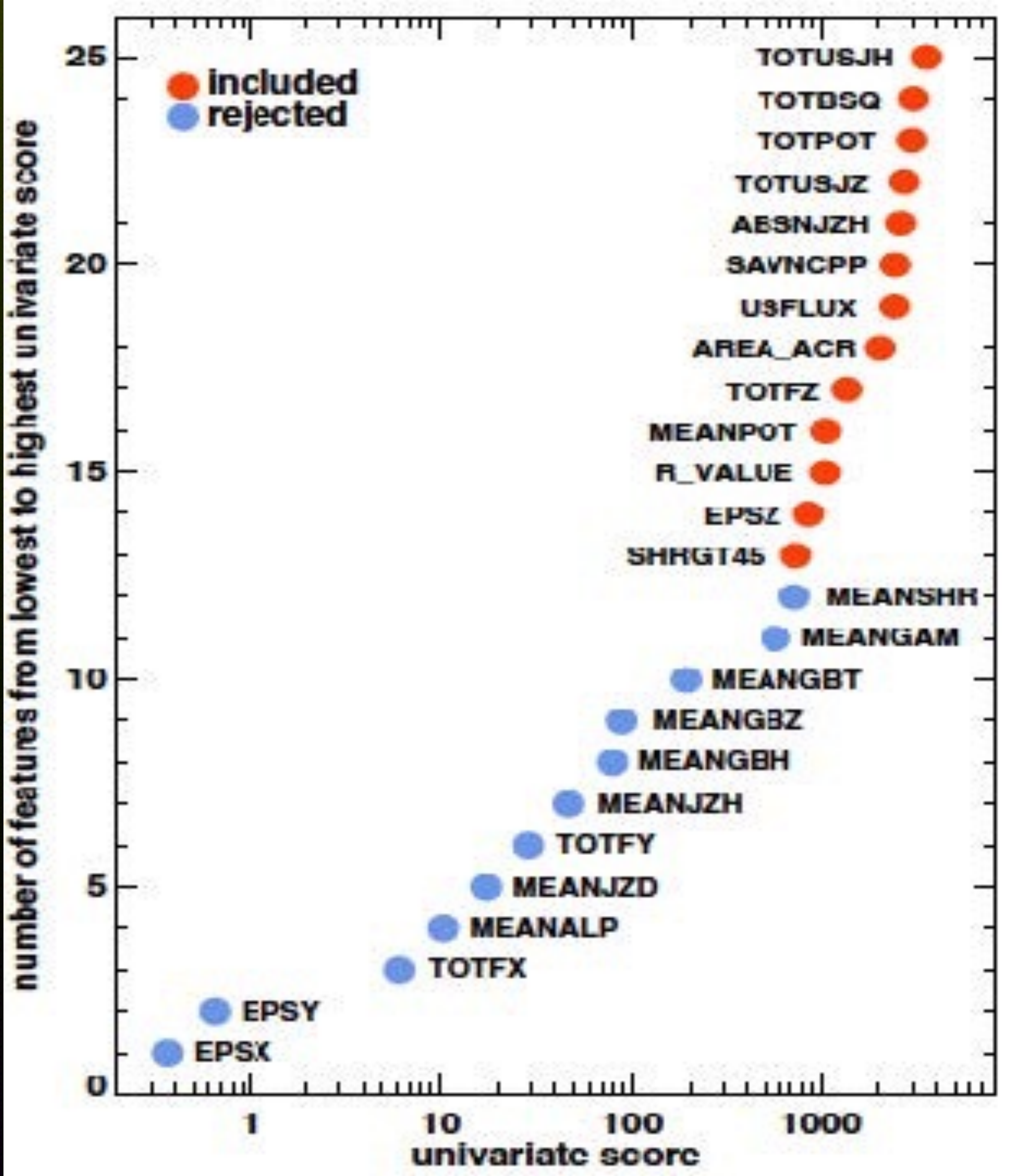


Guerra et al., (2015)

- Homogenizing the results of multiple flare prediction methods, using them with equal or non-equal weights for an ensemble forecasting

INDICATIVE RESULTS

- Multivariate forecasting

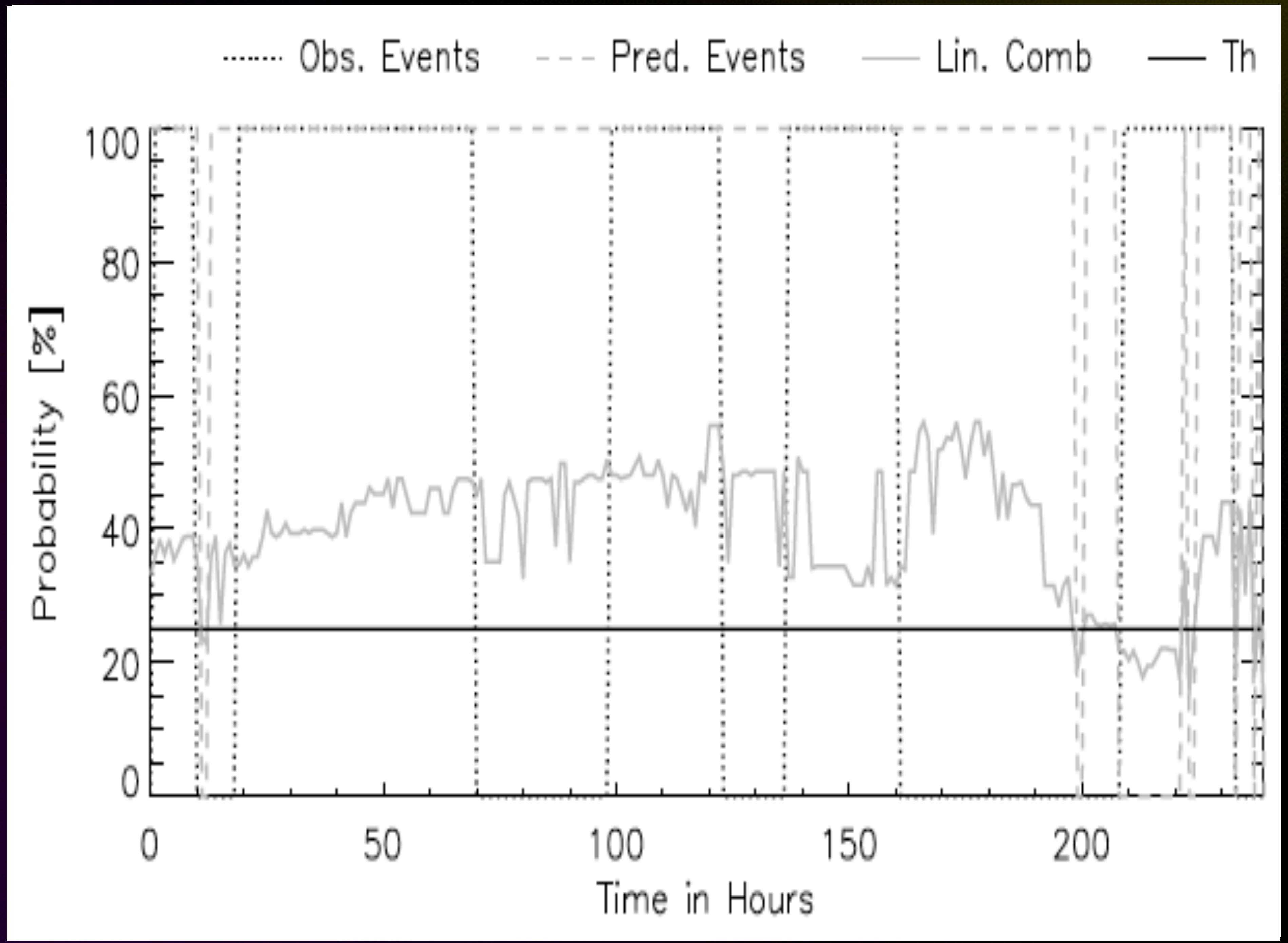


Bobra & Couvidat (2014)

- Ordering of predictors by means of a univariate Fisher ranking score
- Machine-learning classifiers adopted

- Homogenizing the results of multiple flare prediction methods, using them with equal or non-equal weights for an ensemble forecasting

- Ensemble forecasting



Guerra et al., (2015)

JUDGING WHICH METHODS WORK: VALIDATION

- Existing methods are borrowed from terrestrial weather forecasting
- Two types of validation
  - On binary (YES / NO) prediction output
  - On probabilistic $(0 < p < 1)$ prediction output
- **Both are used in flare prediction**

VALIDATION: BORROWED BY TERRESTRIAL WEATHER FORECASTING

Binary validation: Flare (YES) or No Flare (NO)

Tailoring according to different end user needs

	Forecast Flare	Forecast No-flare
Observed Flare	TP	FN
Observed No-flare	FP	TN

Table courtesy: Shaun Bloomfield

2 x 2 contingency table

- TP : true positives
- FN : false negatives
- FP : false positives
- TN : true negatives

- Generalized skill score:

$$SS = \frac{score - score_{reference}}{score_{perfect} - score_{reference}}$$

VALIDATION: BORROWED BY TERRESTRIAL WEATHER FORECASTING

Binary validation: Flare (YES) or No Flare (NO)

	Forecast Flare	Forecast No-flare
Observed Flare	TP	FN
Observed No-flare	FP	TN

Table courtesy: Shaun Bloomfield

2 x 2 contingency table

- TP : true positives
- FN : false negatives
- FP : false positives
- TN : true negatives

- Generalized skill score:

$$SS = \frac{score - score_{reference}}{score_{perfect} - score_{reference}}$$

Tailoring according to different end user needs

- Heidke skill score (ref: random prediction):

$$HSS = \frac{2(TP + TN) - N}{N}$$

- Appleman skill score (ref: climatology [v]):

$$ApSS = \frac{TP - FP}{N}$$

- True skill statistic (ref: weighting POD w. POFD):

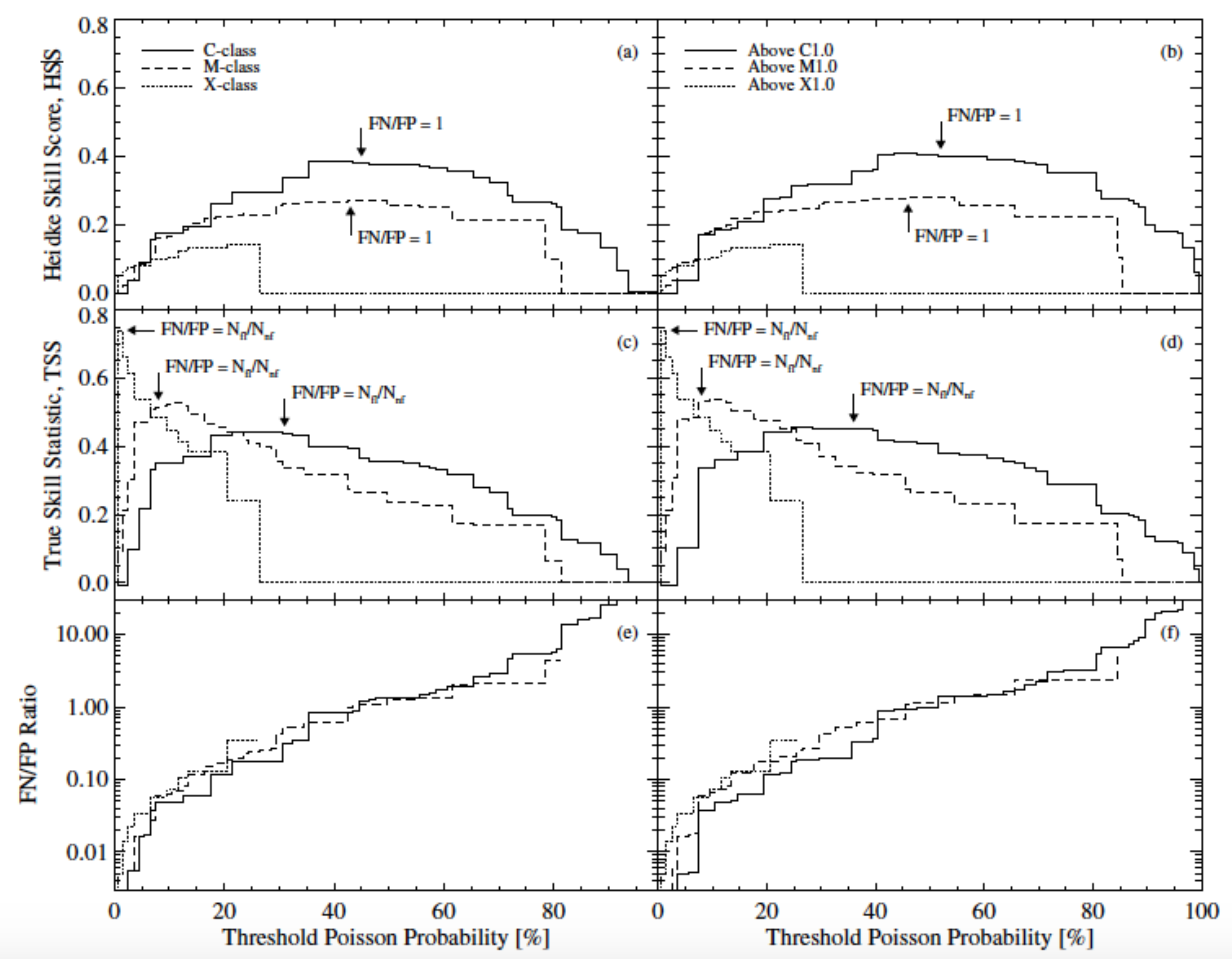
$$TSS = POD - POFD$$

(SOME) BINARY FORECAST VERIFICATION METRICS

Metric Name	Short Name	Format	Worst Score	"No skill" Score	Perfect Score
Accuracy	ACC	(TP + TN) / N	0	...	1
Probability of detection	POD	TP / (TP + FN)	0	...	1
Probability of false detection (false alarm rate)	POFD	FP / (FP + TN)	1	...	0
False alarm ratio	FAR	FP / (TP + FP)	1	...	0
True skill statistic	TSS	POD - POFD	-1	0	1
Heidke skill score	HSS	$(TP + TN - E_{random}) / (N - E_{random})$	-1	0	1

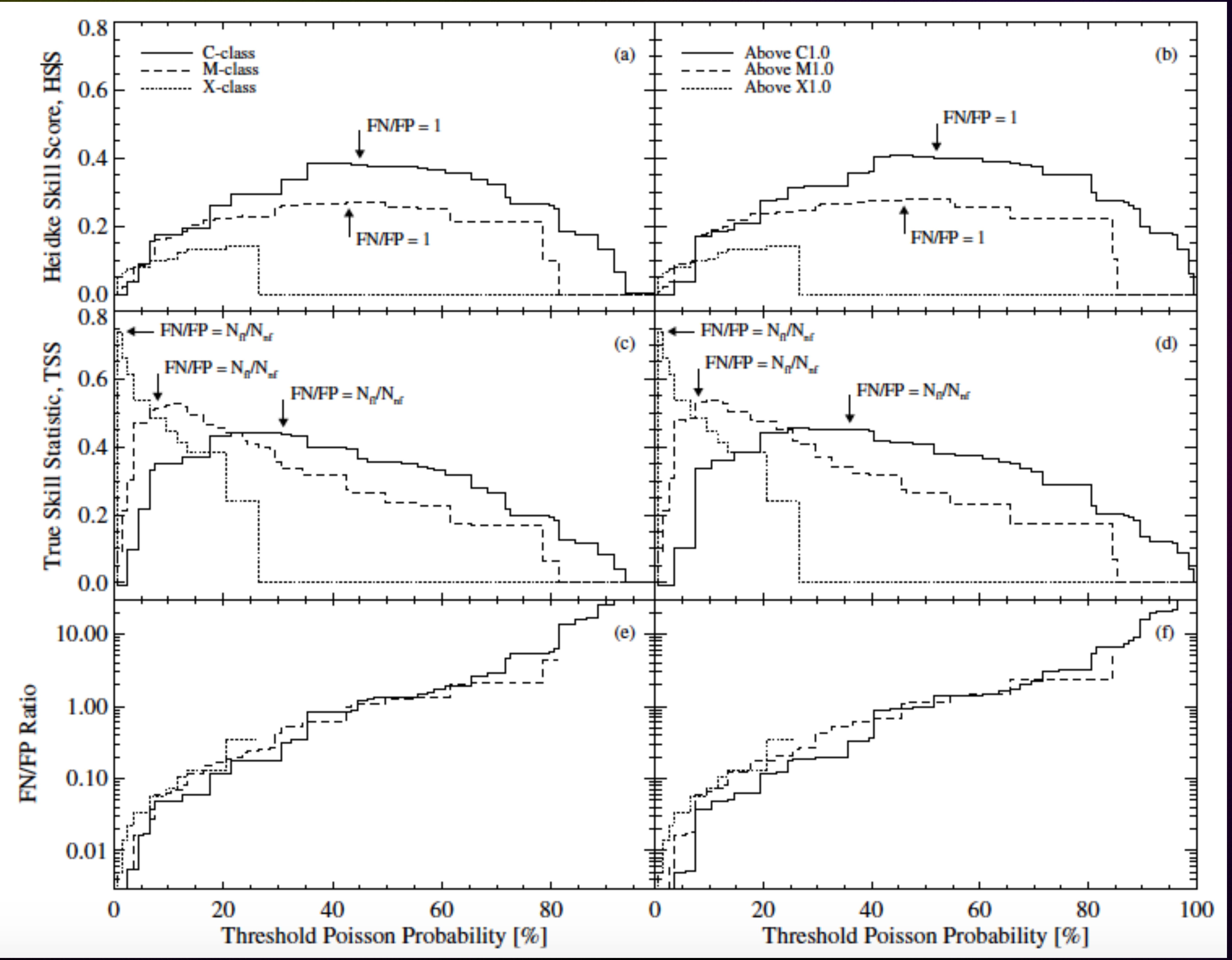
Slide courtesy: Shaun Bloomfield

SOME INDICATIVE RESULTS

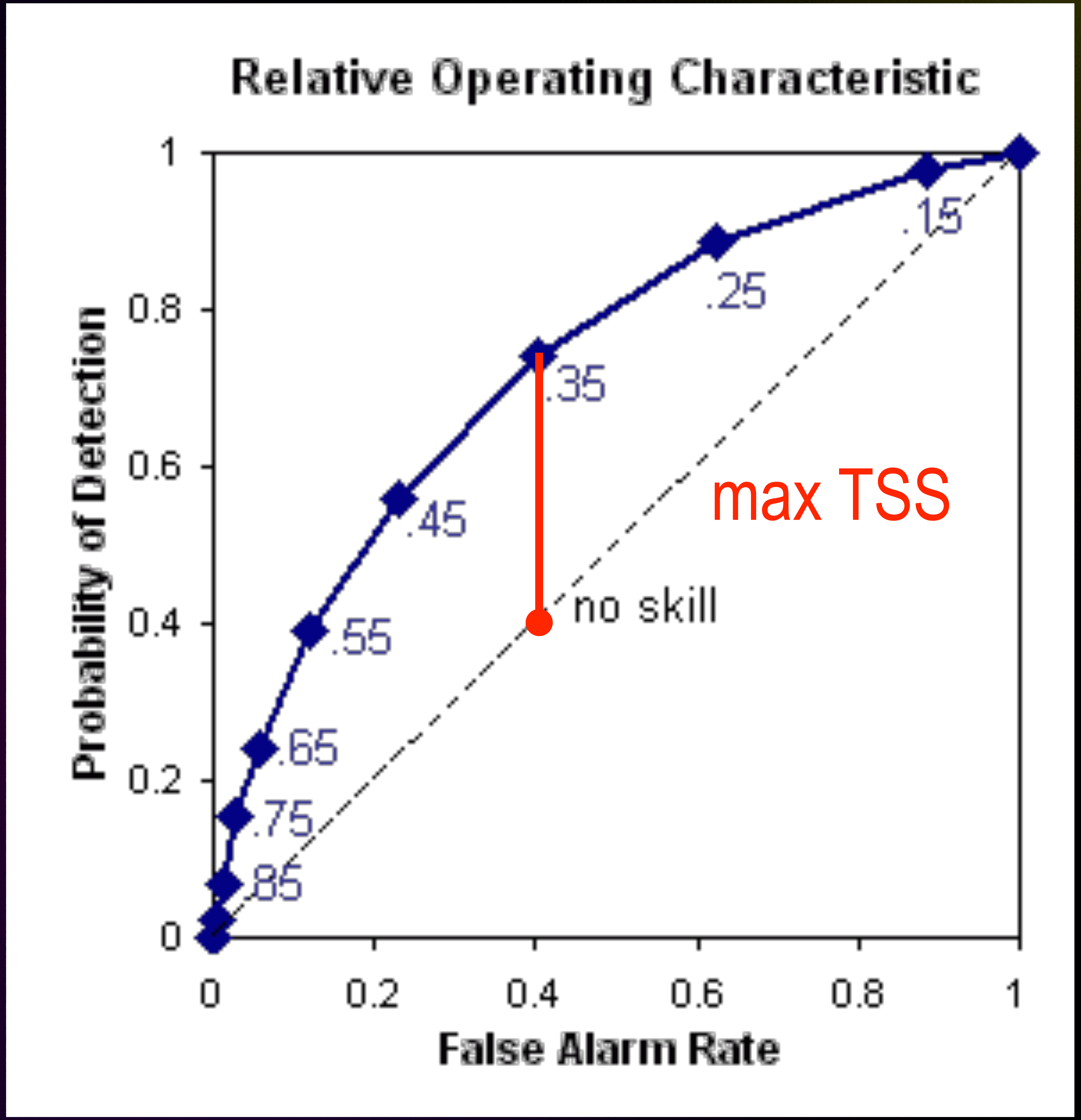


Bloomfield et el., (2012)

SOME INDICATIVE RESULTS



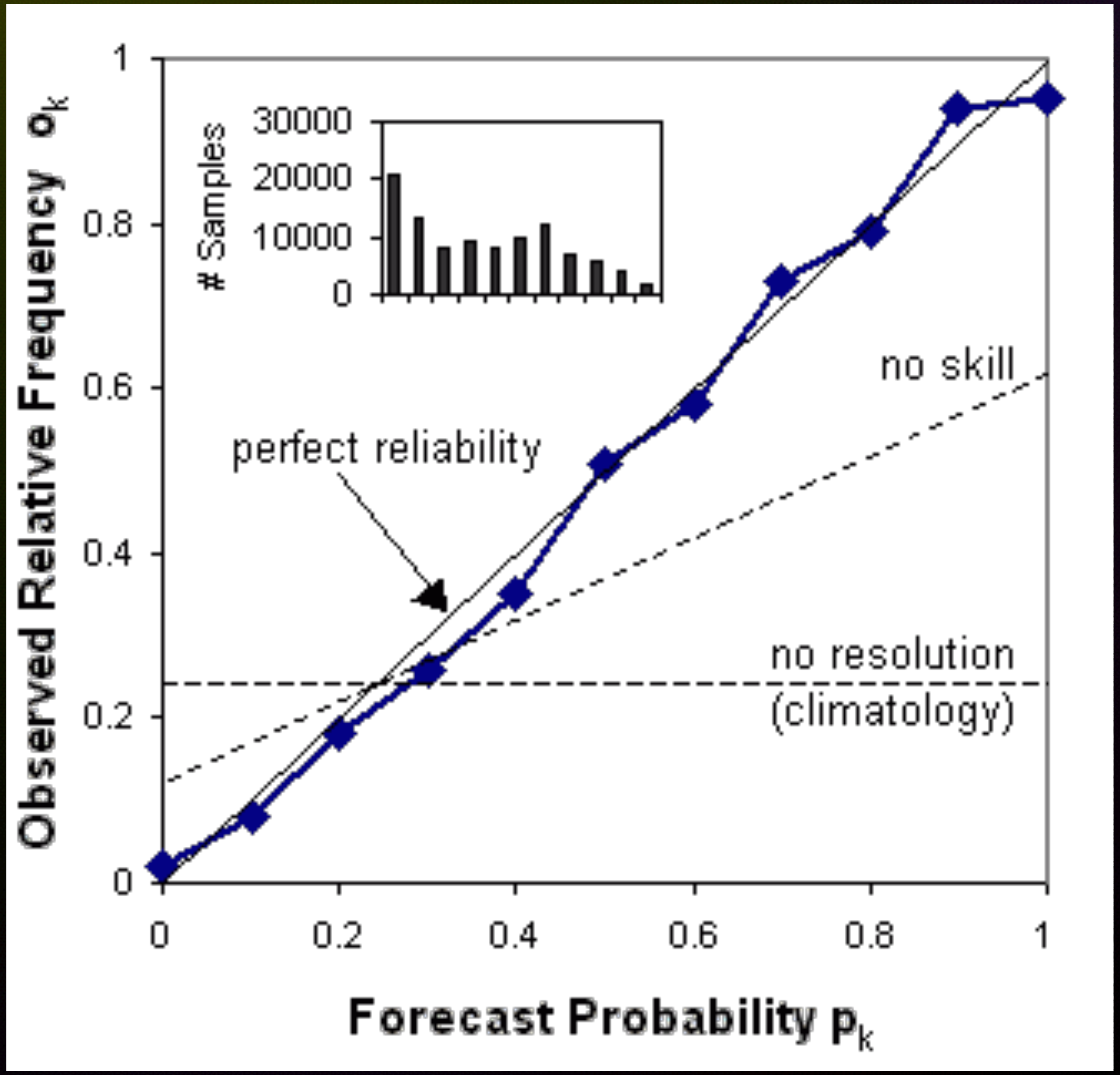
Bloomfield et el., (2012)



Example ROC curve (TSS vs. p_{thres})

PROBABILISTIC VALIDATION

Accept that a probability 0 < p < 1 is assigned to each prediction

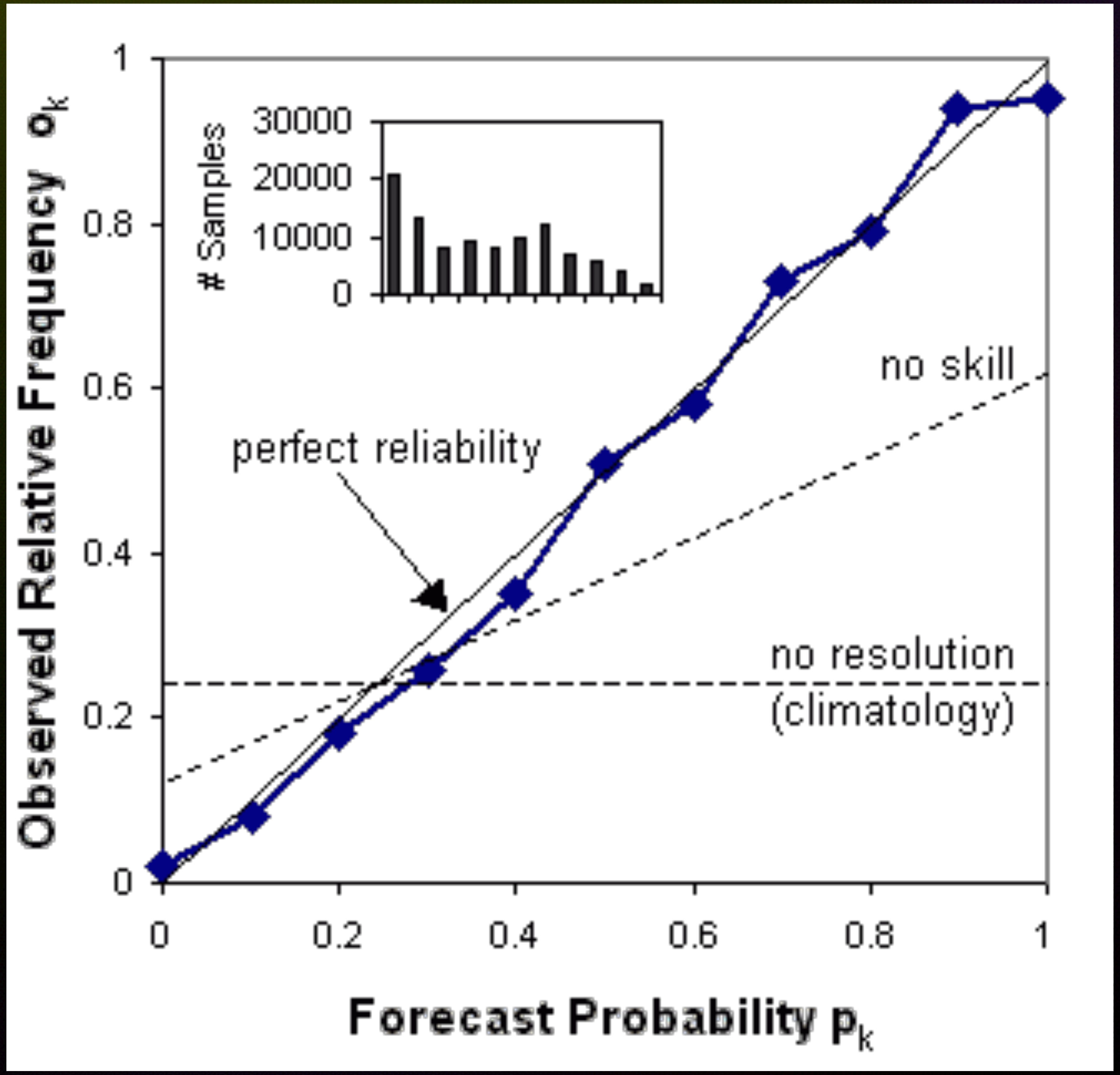


Reliabillity diagram

$$SS = 1 - \frac{MSE_{forecast}}{MSE_{reference}}$$

PROBABILISTIC VALIDATION

Accept that a probability 0 < p < 1 is assigned to each prediction



Reliabillity diagram

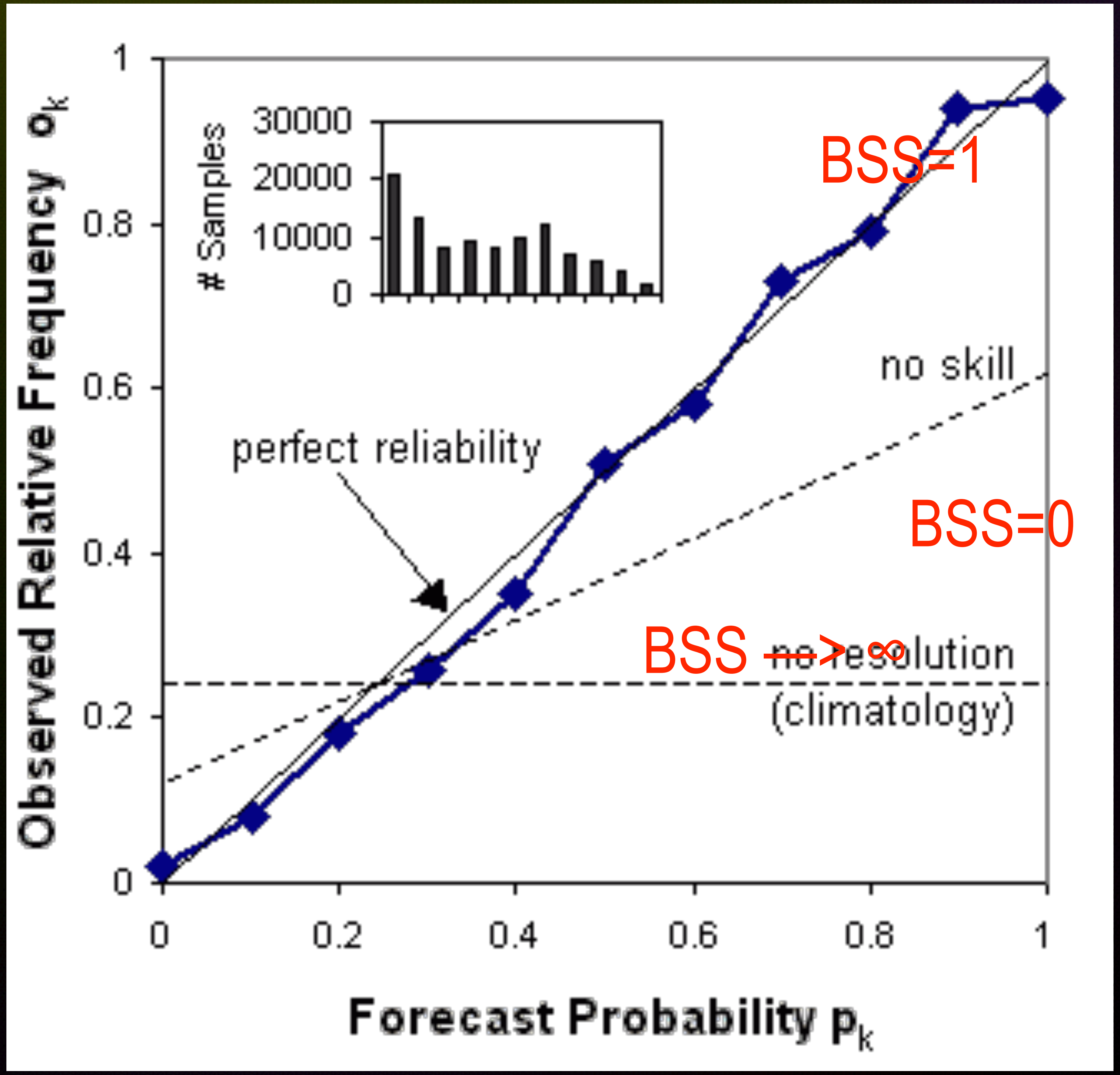
- Correlate forecast probability with observed frequency
- Compare your skill against climatology (mean flaring rate within forecast window)
- Generalized skill score:

$$SS = 1 - \frac{MSE_{forecast}}{MSE_{reference}}$$

$$MSE = \ < (o - p)^2 >$$

PROBABILISTIC VALIDATION

Accept that a probability 0 < p < 1 is assigned to each prediction



Reliabillity diagram

- Correlate forecast probability with observed frequency
- Compare your skill against climatology (mean flaring rate within forecast window)

- Generalized skill score:

$$SS = 1 - \frac{MSE_{forecast}}{MSE_{reference}}$$

$$MSE = \; < (o-p)^2 >$$

- Brier skill score (reference: climatology):

$$BSS = 1 - \frac{\langle (o-p)^2 \rangle}{\langle (o-\bar{o})^2 \rangle}$$

$$BSS \in (-\infty, 1)$$

FIRST EFFORT TO COMPARE METHOD PERFORMANCES ON COMMON DATA SETS

Recently published (Barnes et al., 2016)

Parameter/ Method	Statistical Method	C1.0+, 24 hr ApSS	C1.0+, 24 hr BSS	M1.0+, 12 hr ApSS	M1.0+, 12 hr BSS	M5.0+, 12 hr ApSS	M5.0+, 12 hr BSS
B_{eff}	Bayesian	0.12	0.06	0.00	0.03	0.00	0.02
ASAP	Machine	0.25	0.30	0.01	-0.01	0.00	-0.84
BBSO	Machine	0.08	0.10	0.03	0.06	0.00	-0.01
WL_{SG2}	Curve fitting	N/A	N/A	0.04	0.06	0.00	0.02
NWRA MAG 2-VAR	NPDA	0.24	0.32	0.04	0.13	0.00	0.06
$\log(\mathcal{R})$	NPDA	0.17	0.22	0.01	0.10	0.02	0.04
GCD	NPDA	0.02	0.07	0.00	0.03	0.00	0.02
NWRA MCT 2-VAR	NPDA	0.23	0.28	0.05	0.14	0.00	0.06
SMART2	CCNN	0.24	-0.12	0.01	-4.31	0.00	-11.2
Event Statistics, 10 prior	Bayesian	0.13	0.04	0.01	0.10	0.01	0.00
McIntosh	Poisson	0.15	0.07	0.00	-0.06	N/A	N/A

Event List	Event	No Event	Event Rate	Event	No Event	Event Rate	Event	No Event	Event Rate
	AD			MCD#1			MCD#2		
C1.0+, 24 hr	2609	10356	0.201	789	3751	0.174	249	128	0.660
M1.0+, 12 hr	400	12565	0.031	102	3162	0.031	70	220	0.241
M5.0+, 12 hr	93	12872	0.007	26	3633	0.007	21	270	0.072

FIRST EFFORT TO COMPARE METHOD PERFORMANCES ON COMMON DATA SETS

Recently published (Barnes et al., 2016)

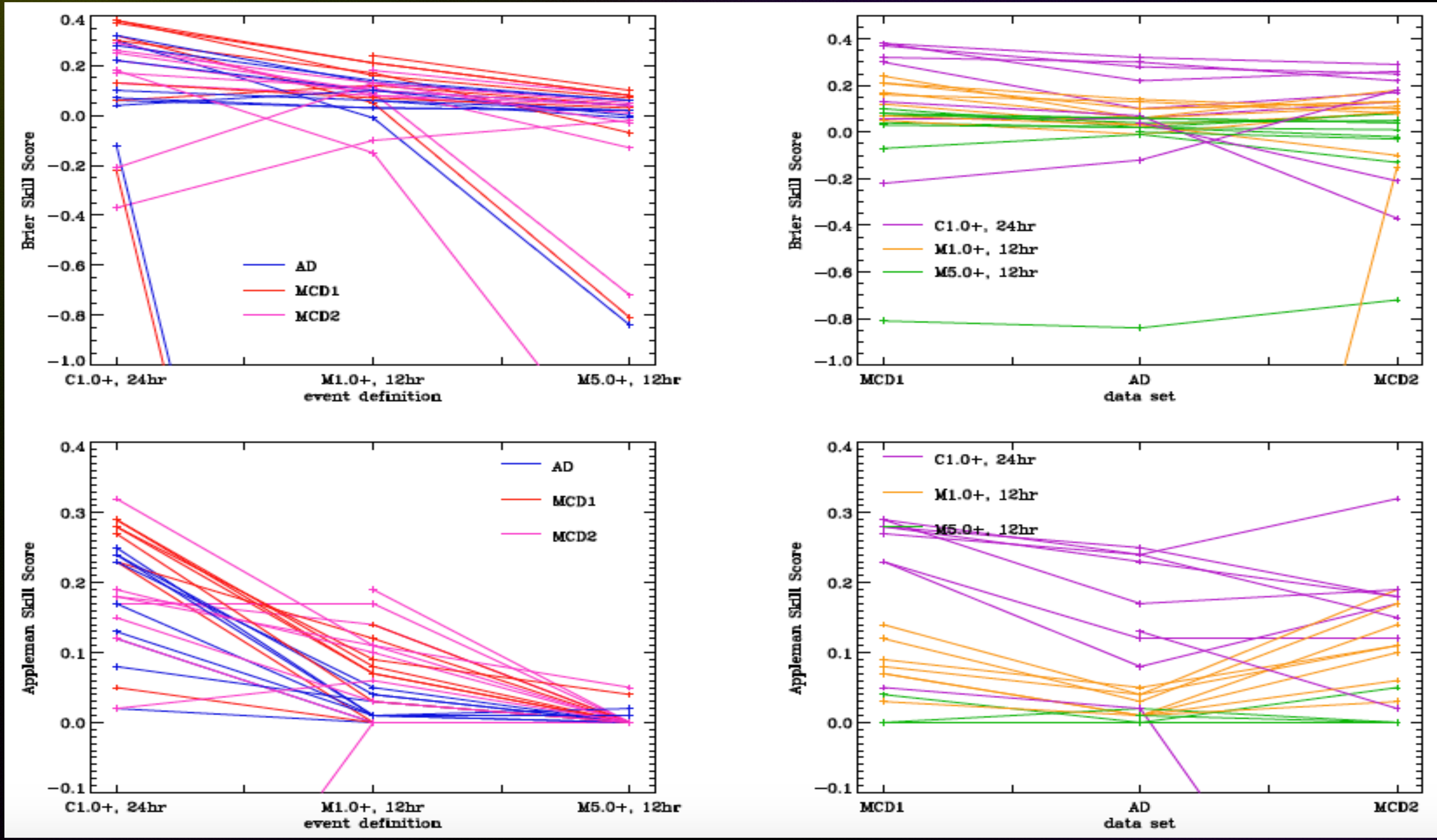
Parameter/ Method	Statistical Method	C1.0+, 24 hr ApSS	C1.0+, 24 hr BSS	M1.0+, 12 hr ApSS	M1.0+, 12 hr BSS	M5.0+, 12 hr ApSS	M5.0+, 12 hr BSS
B_{eff}	Bayesian	0.12	0.06	0.00	0.03	0.00	0.02
ASAP	Machine	0.25	0.30	0.01	-0.01	0.00	-0.84
BBSO	Machine	0.08	0.10	0.03	0.06	0.00	-0.01
WL_{SG2}	Curve fitting	N/A	N/A	0.04	0.06	0.00	0.02
NWRA MAG 2-VAR	NPDA	0.24	0.32	0.04	0.13	0.00	0.06
$\log(\mathcal{R})$	NPDA	0.17	0.22	0.01	0.10	0.02	0.04
GCD	NPDA	0.02	0.07	0.00	0.03	0.00	0.02
NWRA MCT 2-VAR	NPDA	0.23	0.28	0.05	0.14	0.00	0.06
SMART2	CCNN	0.24	-0.12	0.01	-4.31	0.00	-11.2
Event Statistics, 10 prior	Bayesian	0.13	0.04	0.01	0.10	0.01	0.00
McIntosh	Poisson	0.15	0.07	0.00	-0.06	N/A	N/A

Event List	Event	No Event	Event Rate	Event	No Event	Event Rate	Event	No Event	Event Rate
	AD			MCD#1			MCD#2		
C1.0+, 24 hr	2609	10356	0.201	789	3751	0.174	249	128	0.660
M1.0+, 12 hr	400	12565	0.031	102	3162	0.031	70	220	0.241
M5.0+, 12 hr	93	12872	0.007	26	3633	0.007	21	270	0.072

Apparently worse BSS for scarcer (i.e., increasing flare class) events

Typically a bit - but not much - better than climatology (> 0) / quite often worse than climatology (< 0)

FIRST EFFORT TO COMPARE METHOD PERFORMANCES ON COMMON DATA SETS



Generally, there is no method clearly outperforming the others

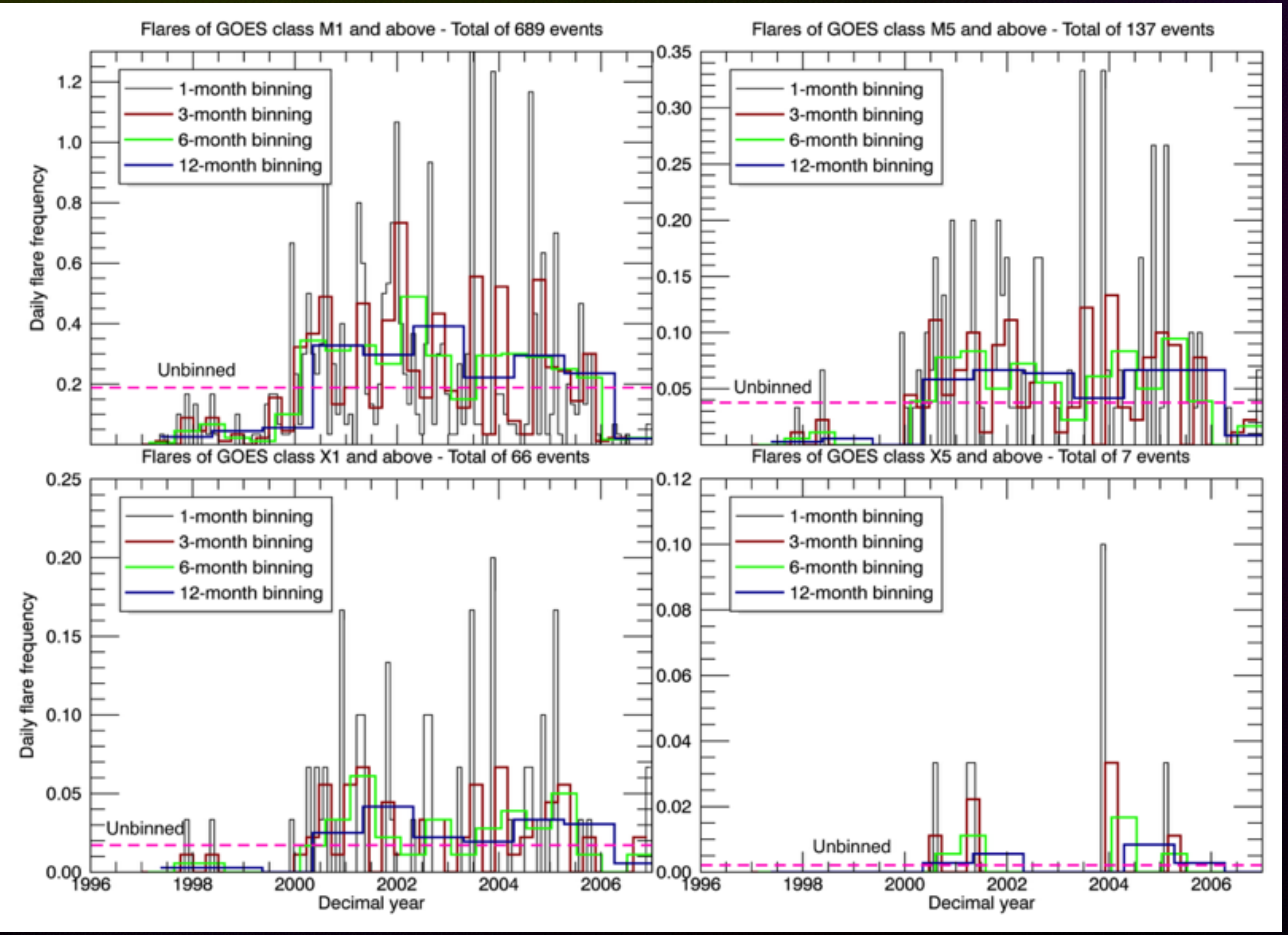
Barnes et al., (2016)

VALIDATION REQUIREMENTS

- Balanced dataset of flaring and non-flaring populations (correct flaring rates)

VALIDATION REQUIREMENTS

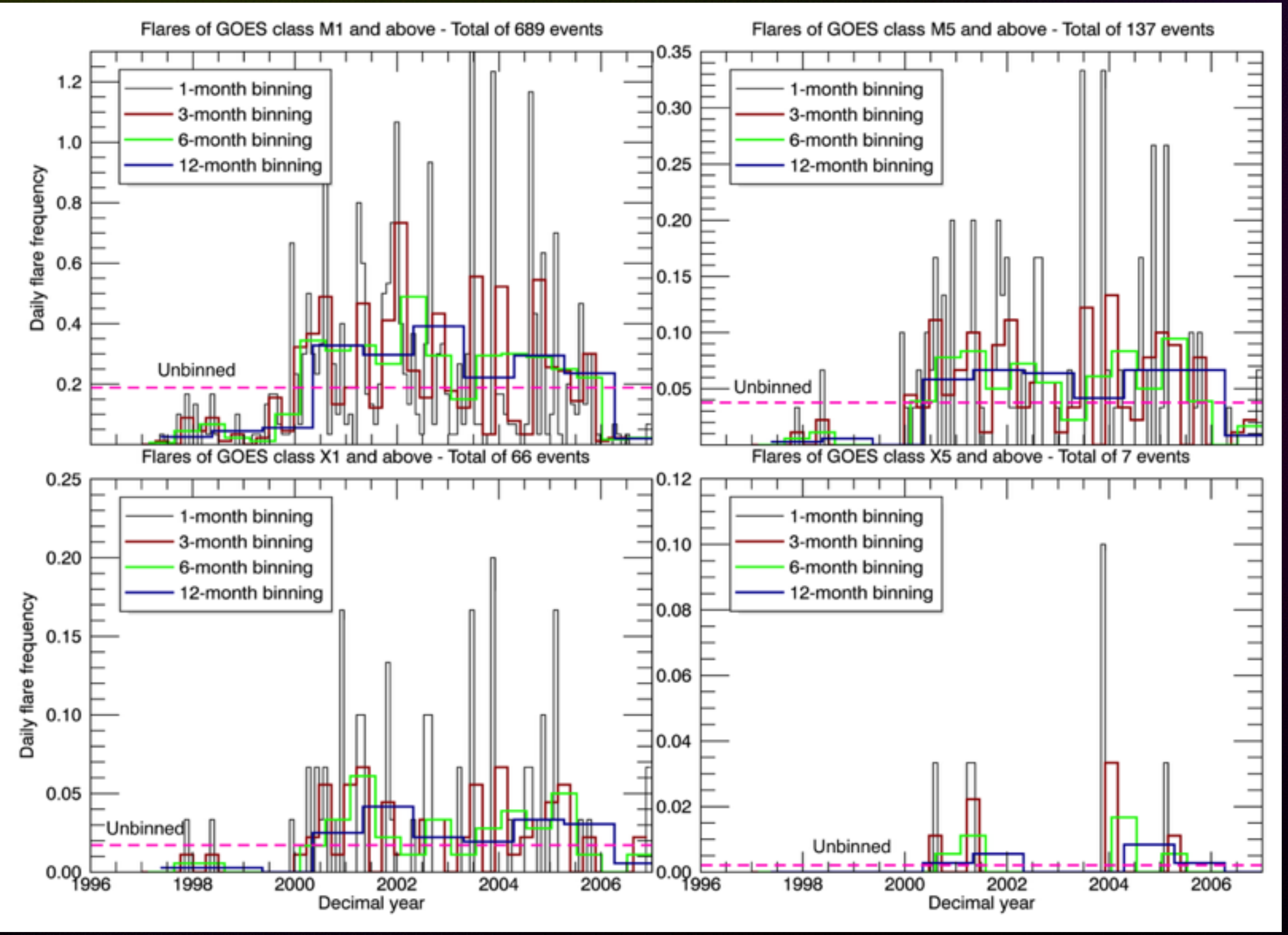
- Balanced dataset of flaring and non-flaring populations (correct flaring rates)



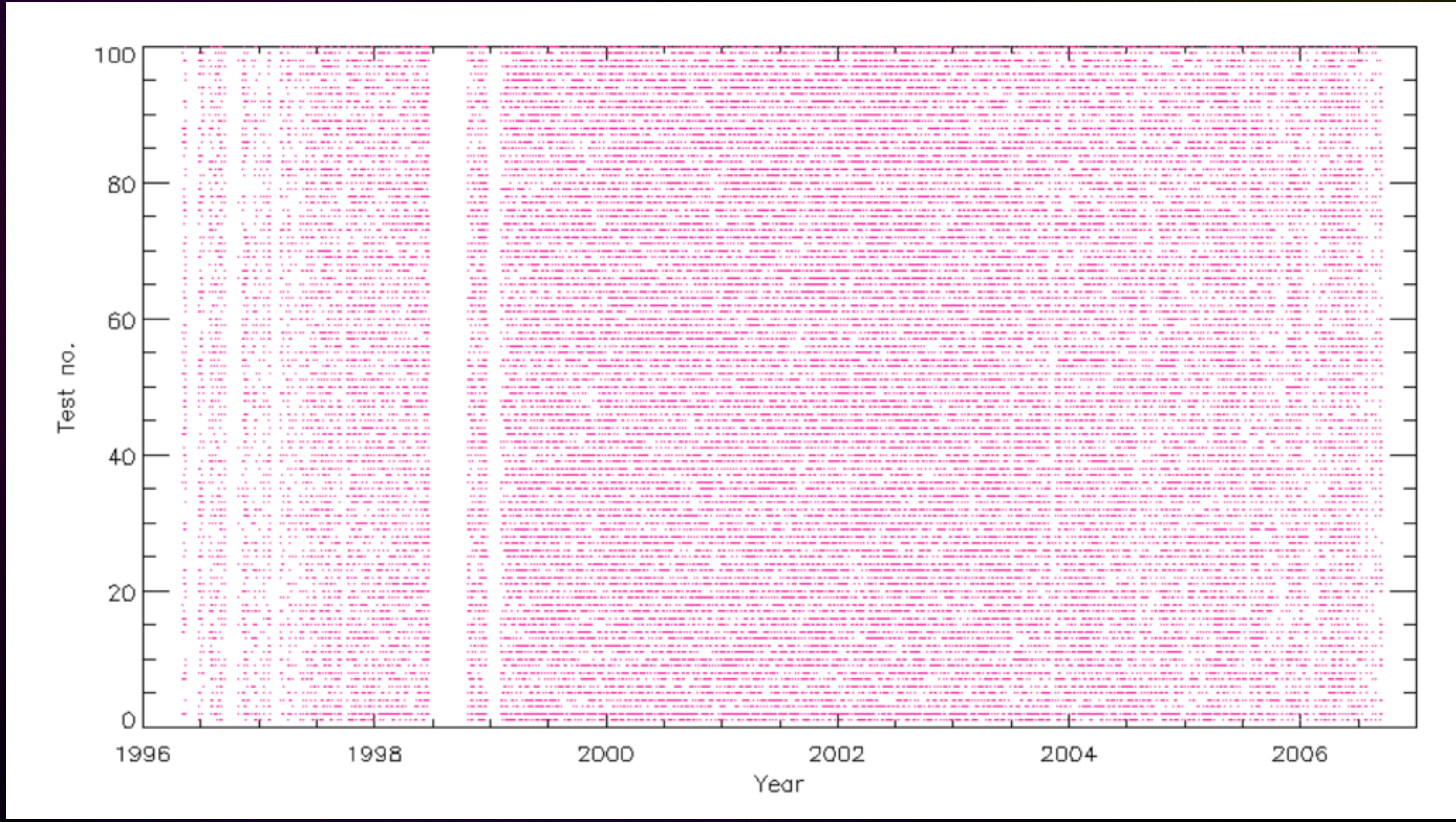
Flaring rates over solar cycle 23 (M1+, M5+, X1+, X5+)

VALIDATION REQUIREMENTS

- Balanced dataset of flaring and non-flaring populations (correct flaring rates)
- Large number of validation tests, using randomly chosen training and test sets



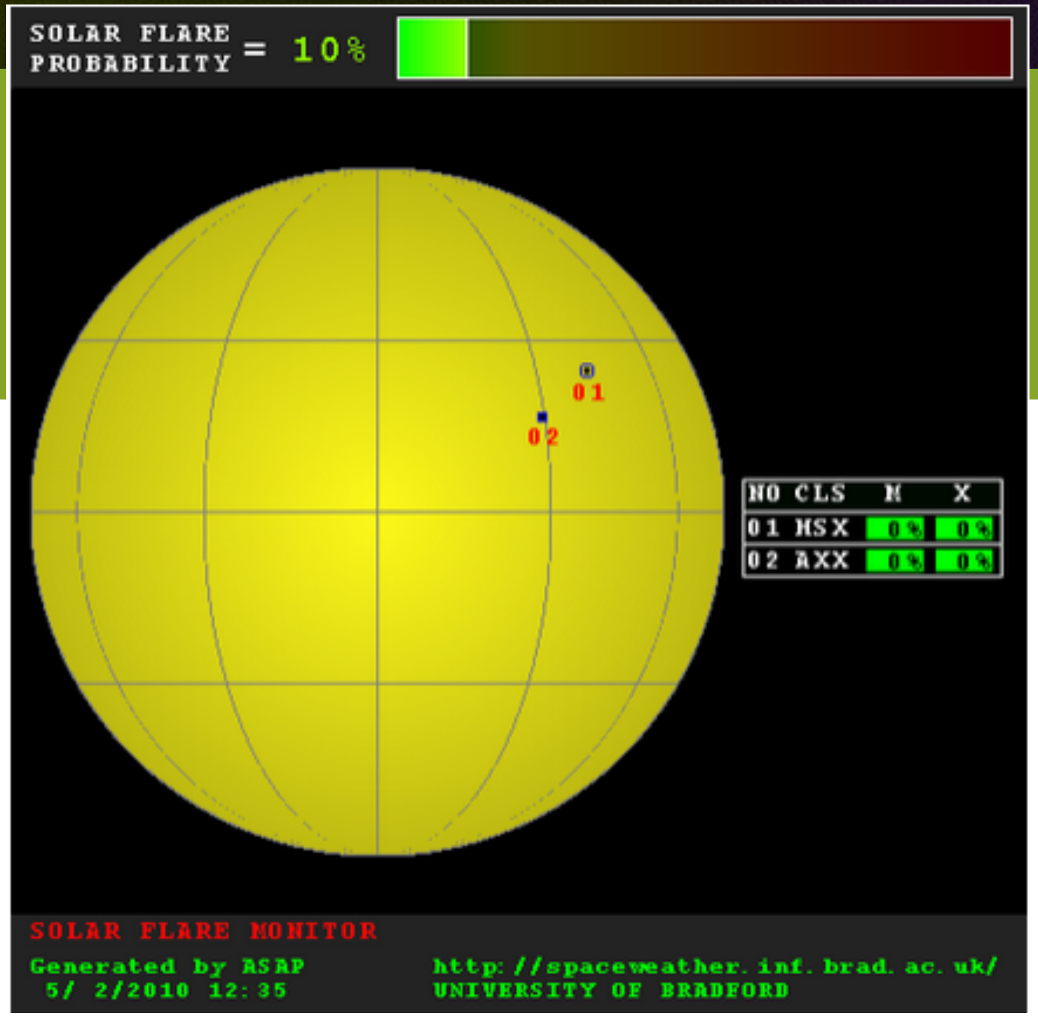
Flaring rates over solar cycle 23 (M1+, M5+, X1+, X5+)



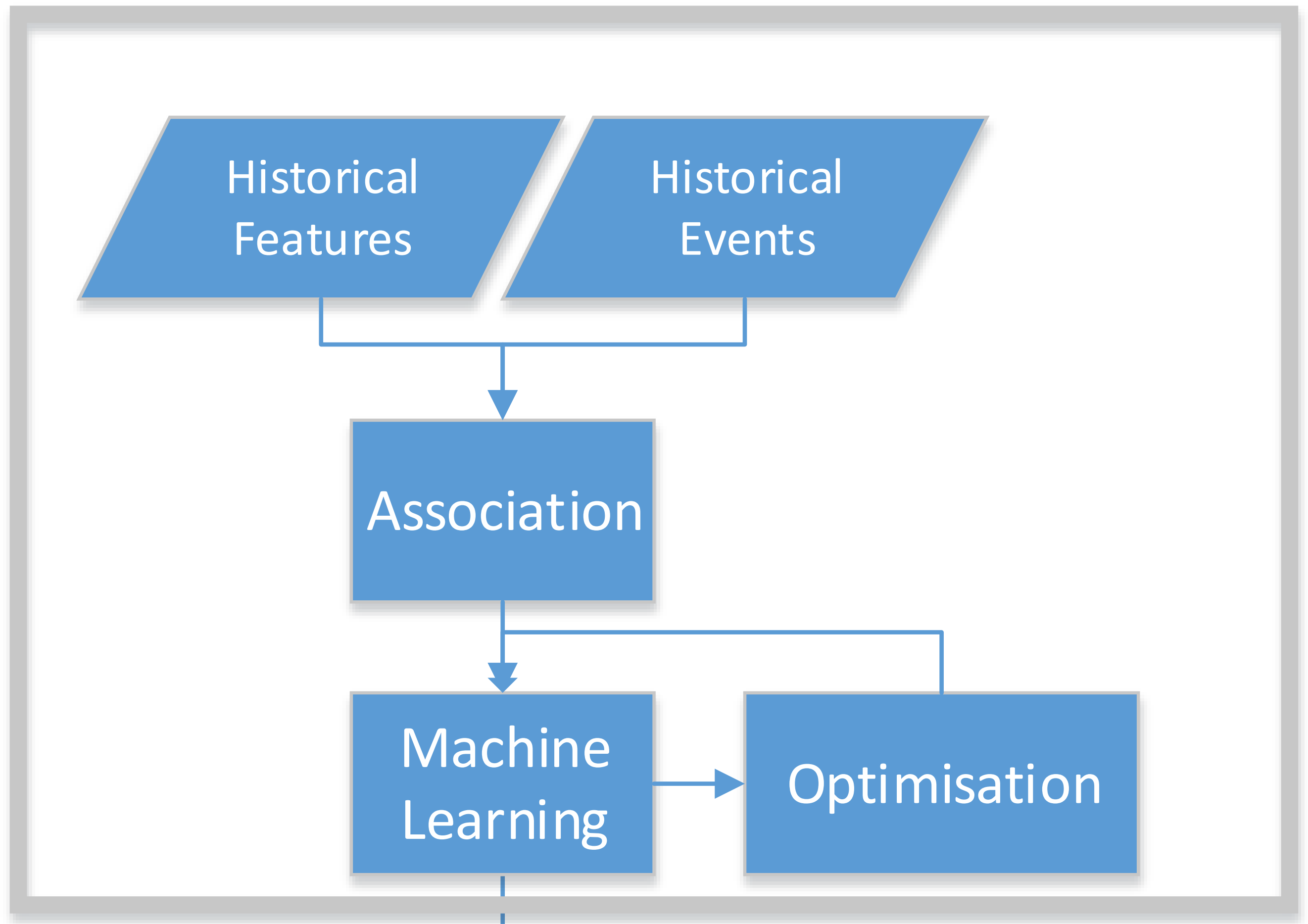
Random selection of training (white) and testing (red dots) subsets

FROM PREDICTION METHODS TO OPERATIONAL FLARE FORECASTING SERVICES

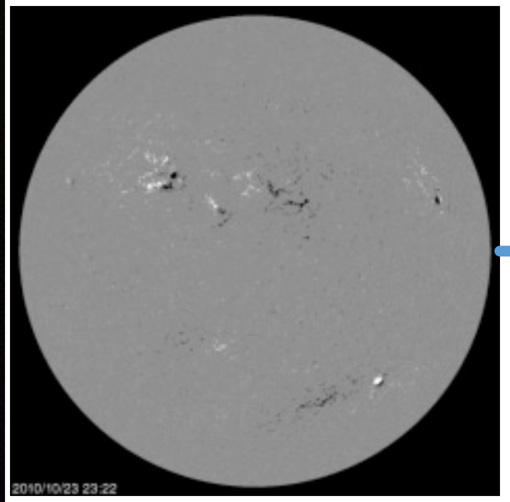
ASAP



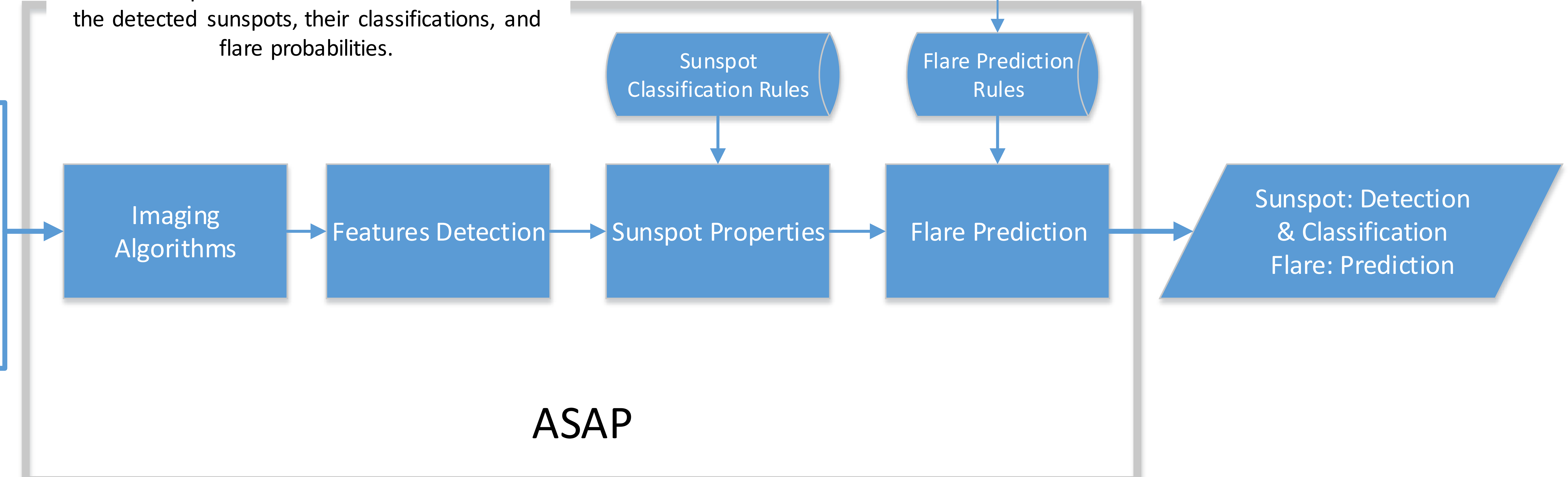
ASAP's Output: Flare Monitor. Visualisation of the detected sunspots, their classifications, and flare probabilities.



MDI Continuum



MDI Magnetogram



ASAP: The first flare prediction service utilizing machine learning methods

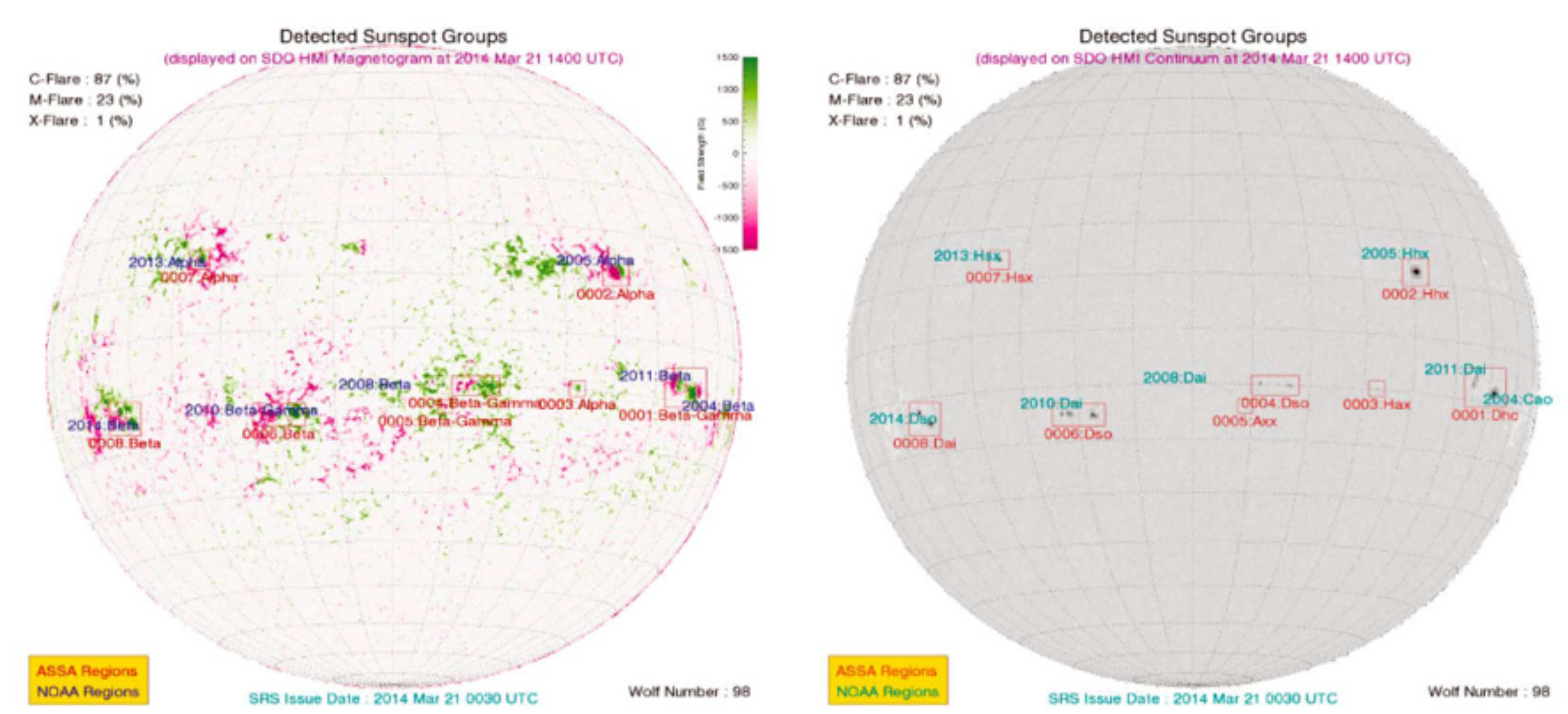
Besides the idea (prediction method), an infrastructure is also needed

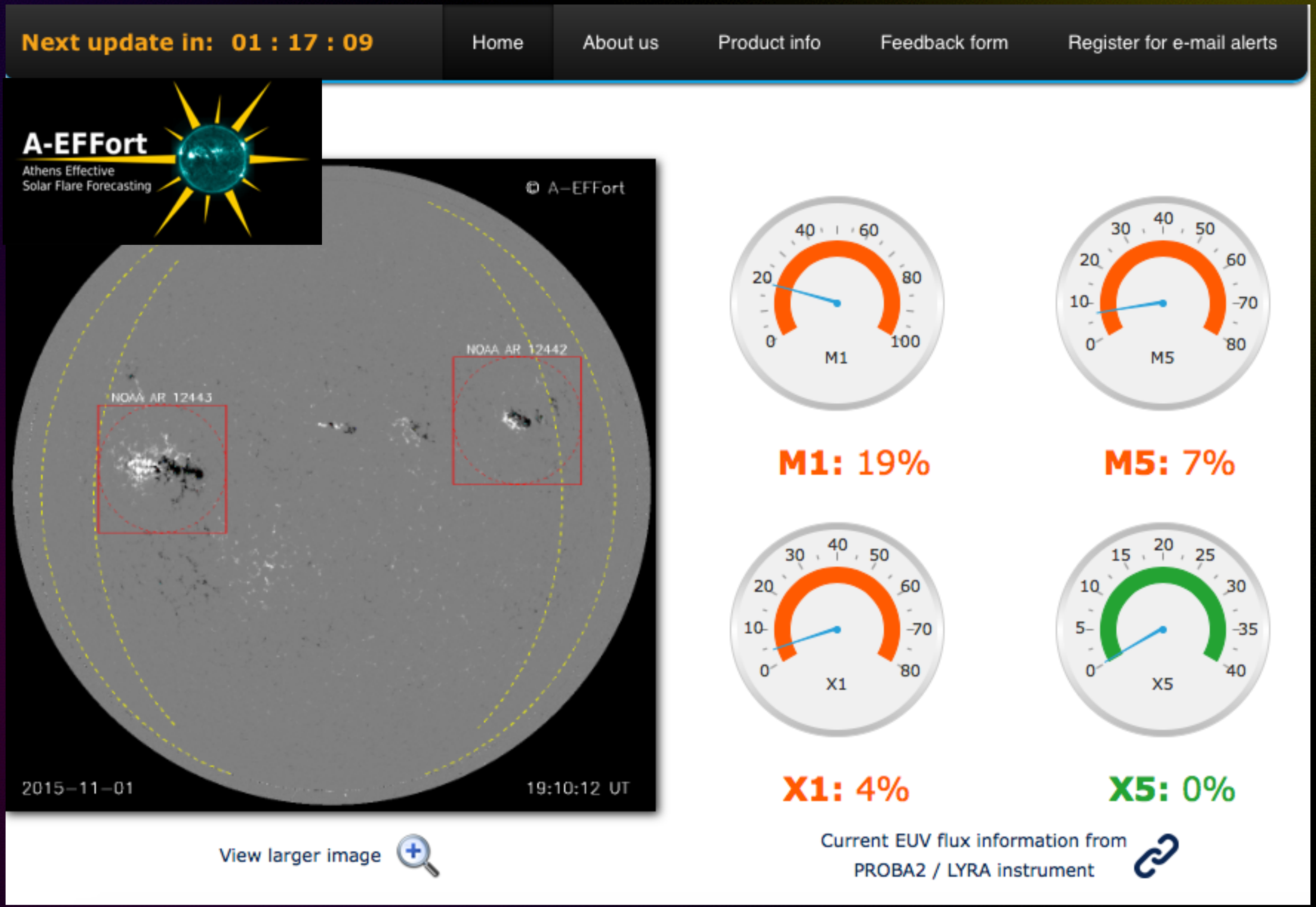
Complete automation means:

- Ease of calculations
- Ease of maintenance
- Resilience
- Modularity, for improvement

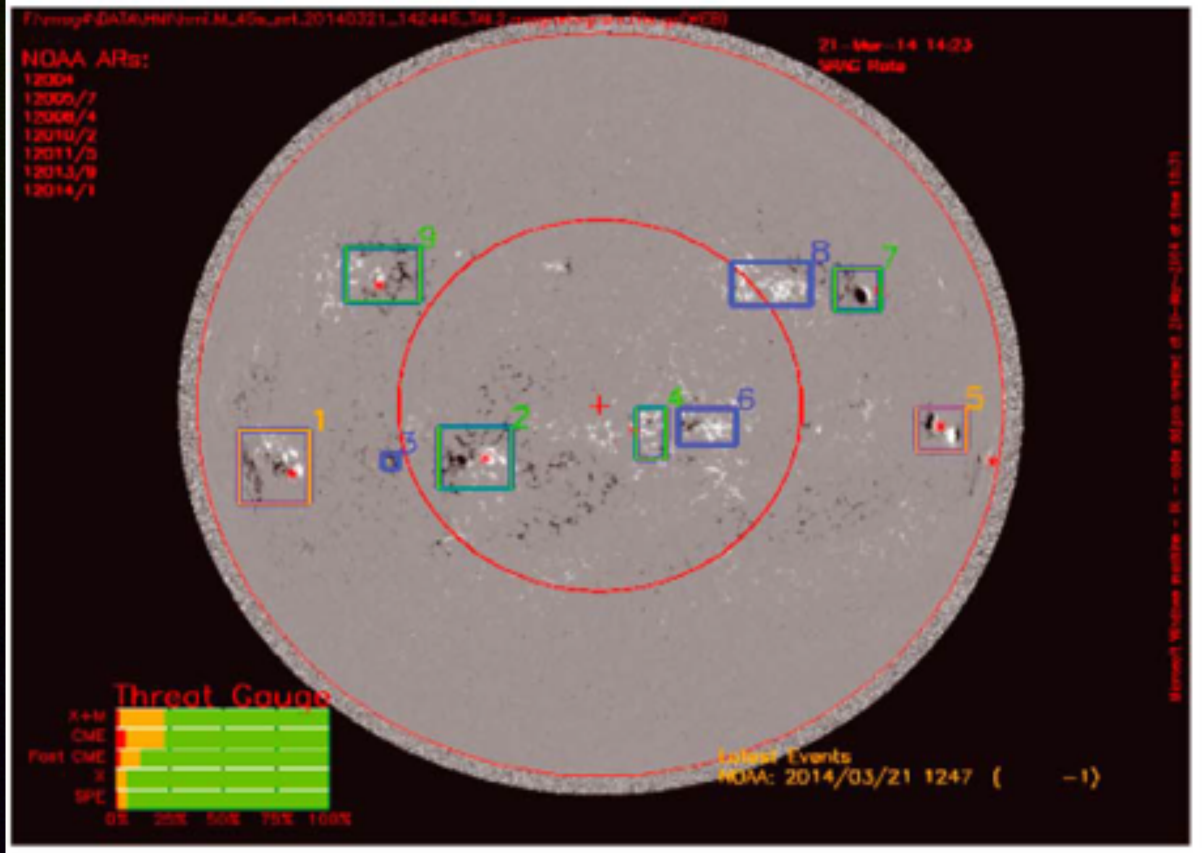
EXISTING FLARE PREDICTION SERVICES AROUND THE WORLD

ASSA

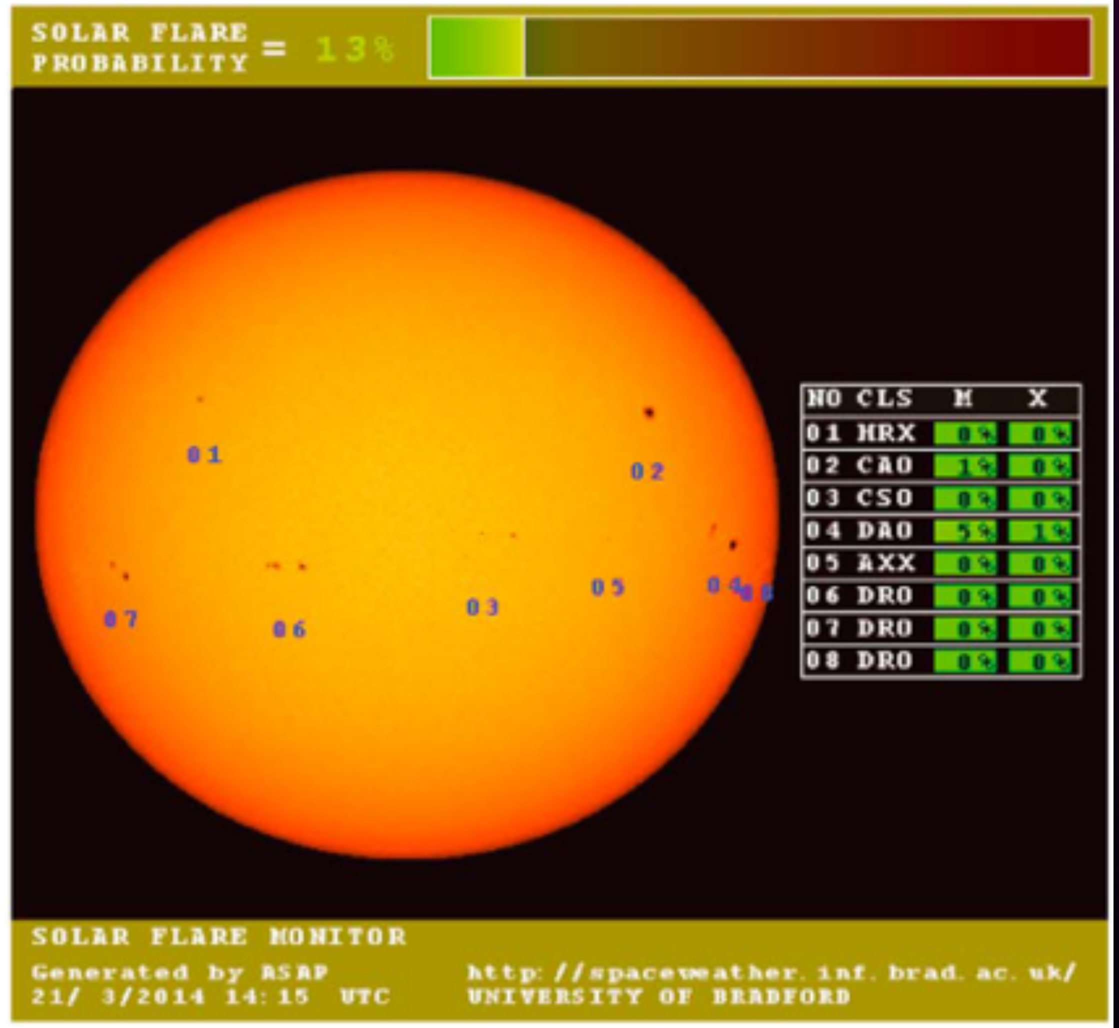




MAG4



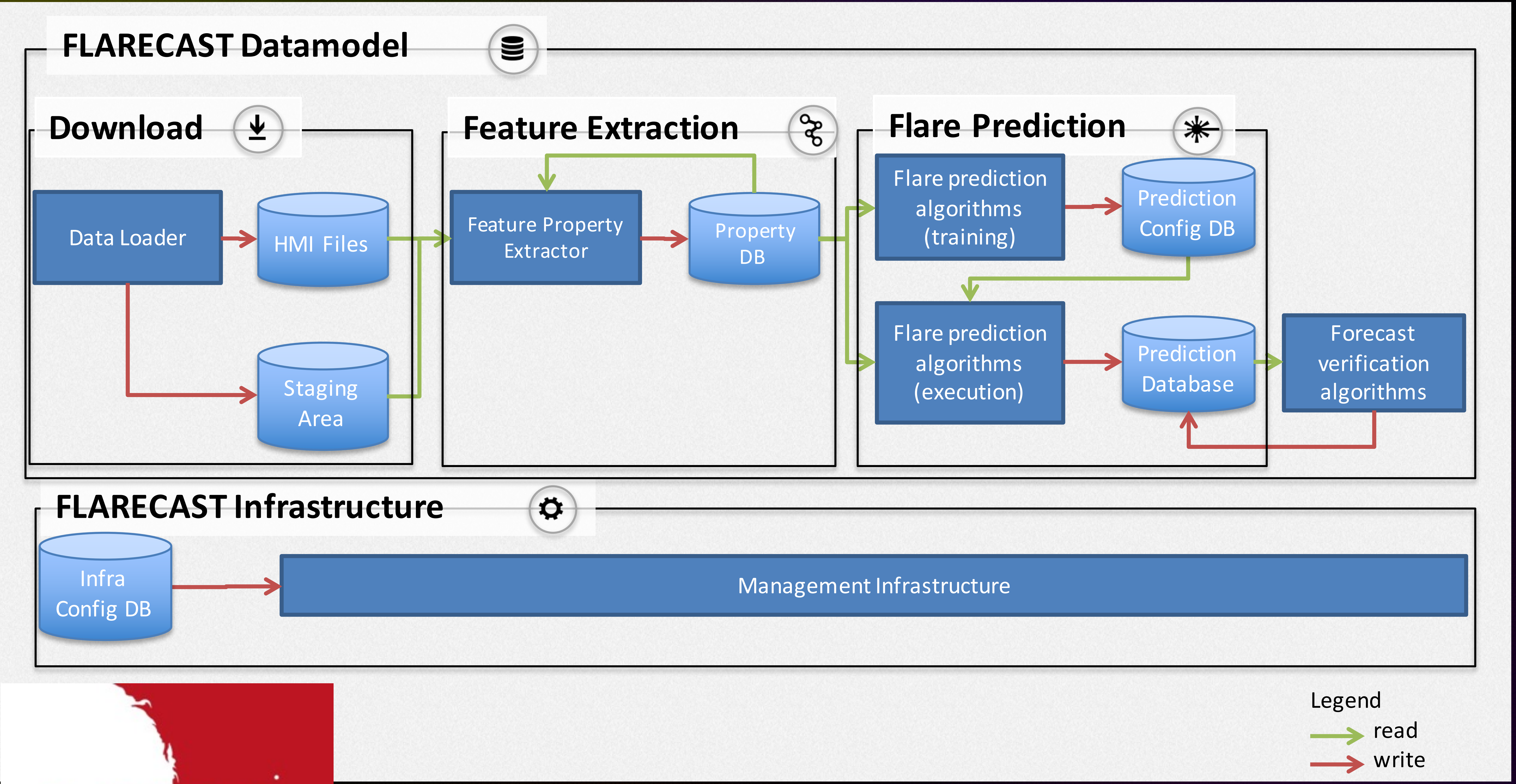
ASAP



www.SolarMonitor.org

Guerra et al., (2015)

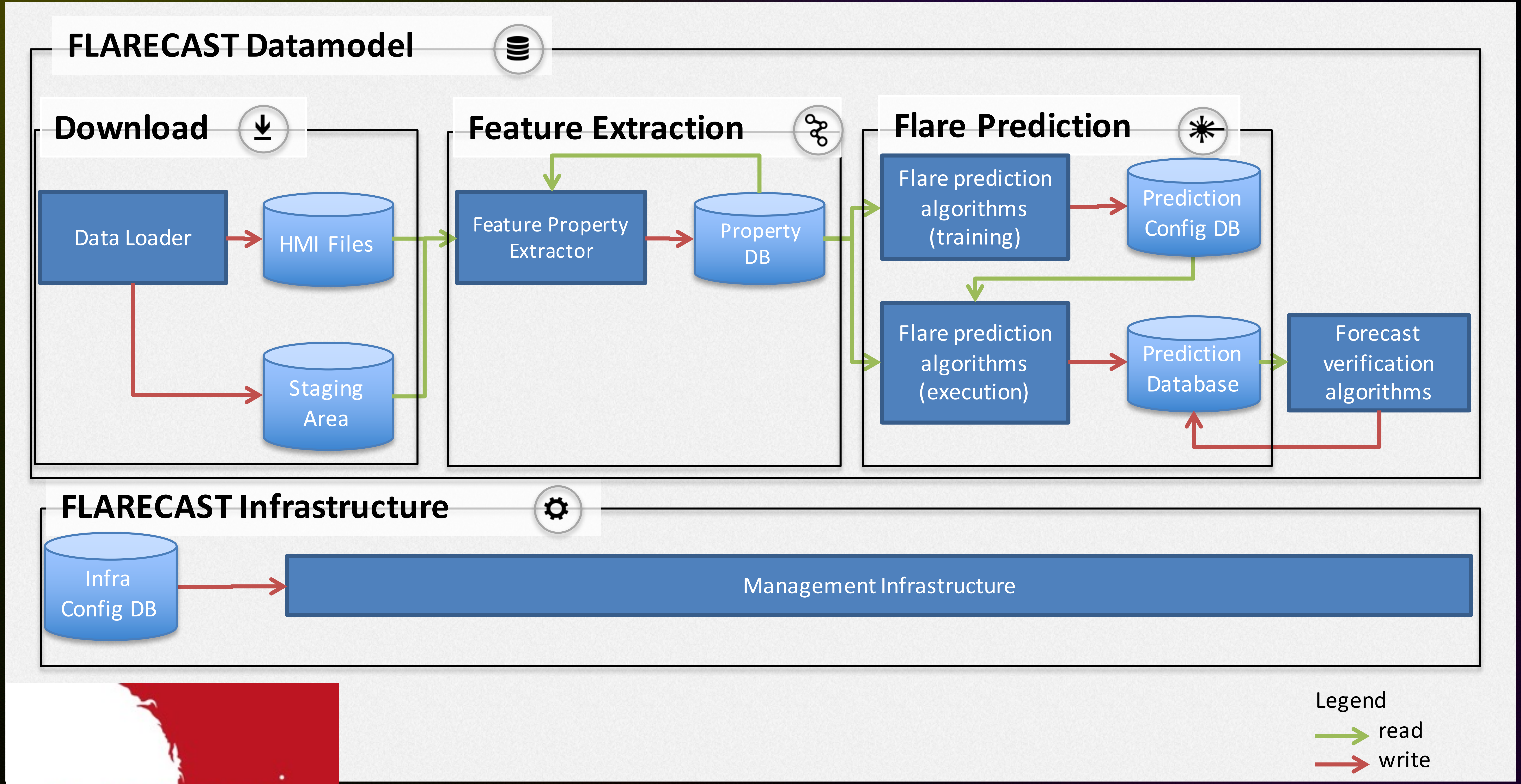
A NUMBER OF FORTHCOMING SERVICES



DAFFS

Europe, USA, Japan

A NUMBER OF FORTHCOMING SERVICES



DAFFS

PSTEP

FLARECAST

Europe, USA, Japan

- **Fact:** Virtually all planned services rely essentially on multivariate predictors
- **Aphorism:** validation, validation, validation …

CONCLUSIONS

- ★ Consensus that reliable, automated solar flare prediction should be an asset of our SWE forecasting efforts
- ★ Can flares be predicted, however?
  - Quite likely, major flare prediction will remain probabilistic in the future
  - But is this due to the nature of the problem, or due to lack of crucial information or a flawed approach?
  - How far along we can go remains TBS — the goal of various flare forecasting efforts is to bring probabilistic flare prediction as close as possible to a binary (YES / NO one)
- ★ Customized, but always <u>unbiased</u>, validation : its importance cannot be stressed enough
- ★ Multivariate forecasting, enabled by machine-learning and other methods (i.e., PCA, DA) seems to be the norm for future services — we can do it nowadays, can't we?
- ★ <u>However, we need to raise Occam's razor</u> : how many / which parameters do we need for a sufficient forecasting? The answer will drive developments in our physical understanding of flare triggering

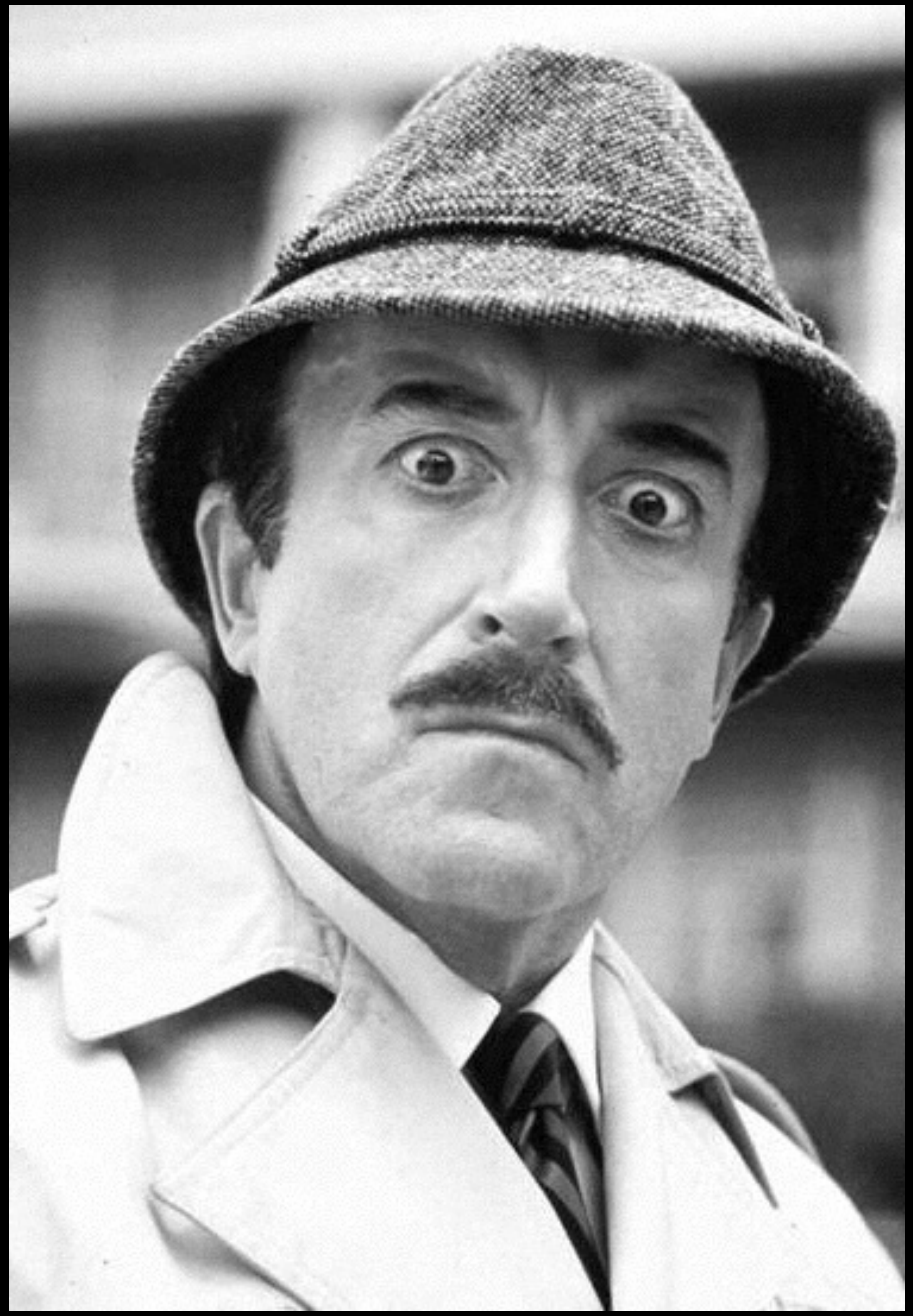
CONCLUSIONS

- Consensus that reliable, automated solar flare prediction should be an asset of our SWE forecasting efforts
- Can flares be predicted, however?
  - Quite likely, major flare prediction will remain probabilistic in the future
  - But is this due to the nature of the problem, or due to lack of crucial information or a flawed approach?
  - How far along we can go remains TBS — the goal of various flare forecasting efforts is to bring probabilistic flare prediction as close as possible to a binary (YES / NO one)
- Customized, but always <u>unbiased</u>, validation : its importance cannot be stressed enough
- Multivariate forecasting, enabled by machine-learning and other methods (i.e., PCA, DA) seems to be the norm for future services — we can do it nowadays, can't we?
- <u>However, we need to raise Occam's razor</u> : how many / which parameters do we need for a sufficient forecasting? The answer will drive developments in our physical understanding of flare triggering

Diverse expertise and ways of thinking are generally needed

All these issues and challenges referring to flare prediction …

All these issues and challenges referring to flare prediction …



… we haven’t even touched CME and SEP prediction yet!

BACKUP SLIDES