PERFORMANCE VERIFICATION OF SOLAR-FLARE PREDICTION MODELS

from Climatology to Skill and from Forecast Probabilities to Certainty

Manolis K. Georgoulis
RCAAM of the Academy of Athens

Work partially supported by:
* A-EFFort (ESA/SSA): a-effort.academyofathens.gr
* FLARECAST (EC/H2020): flarecast.eu

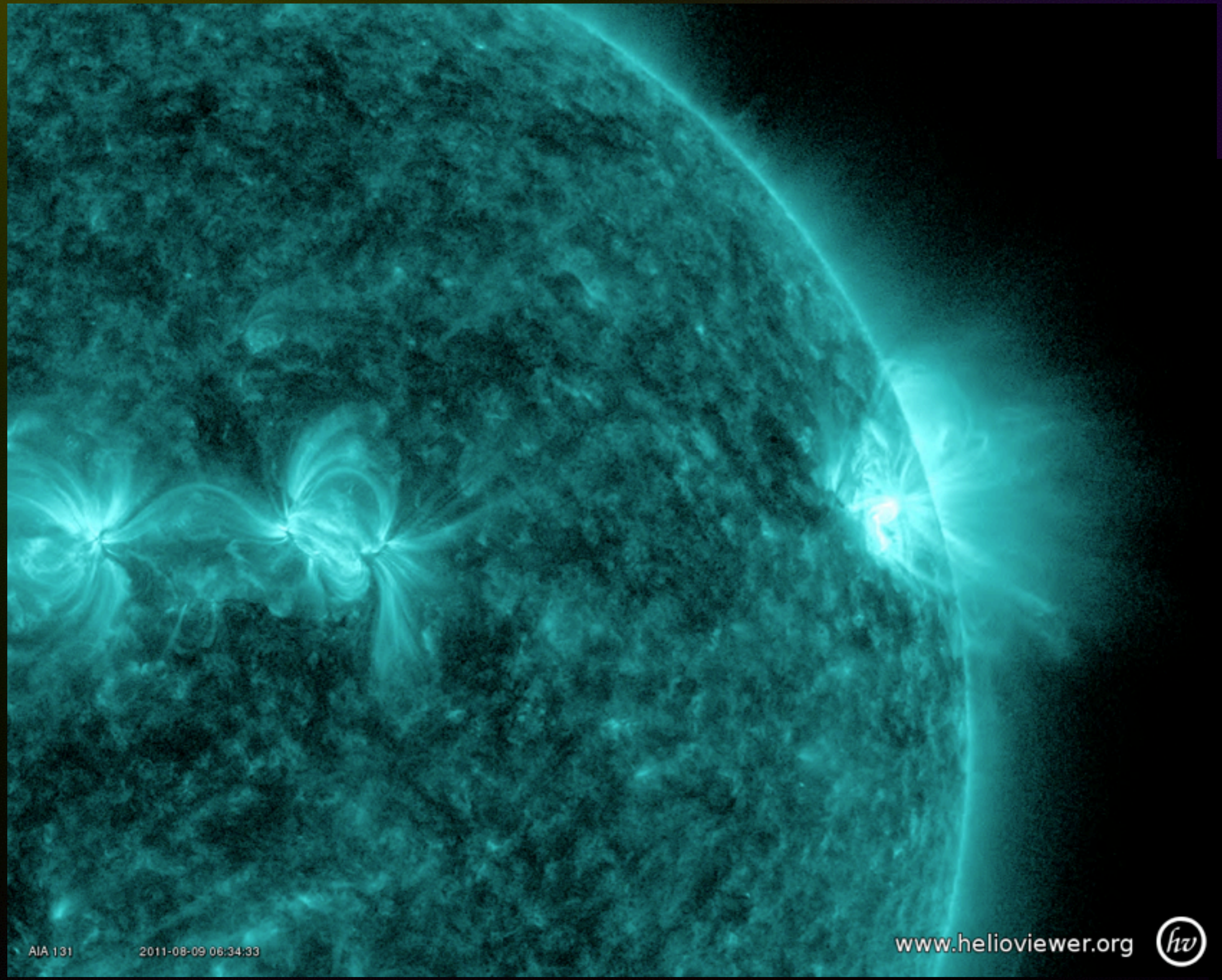
SCIENCE FOR SPACE WEATHER

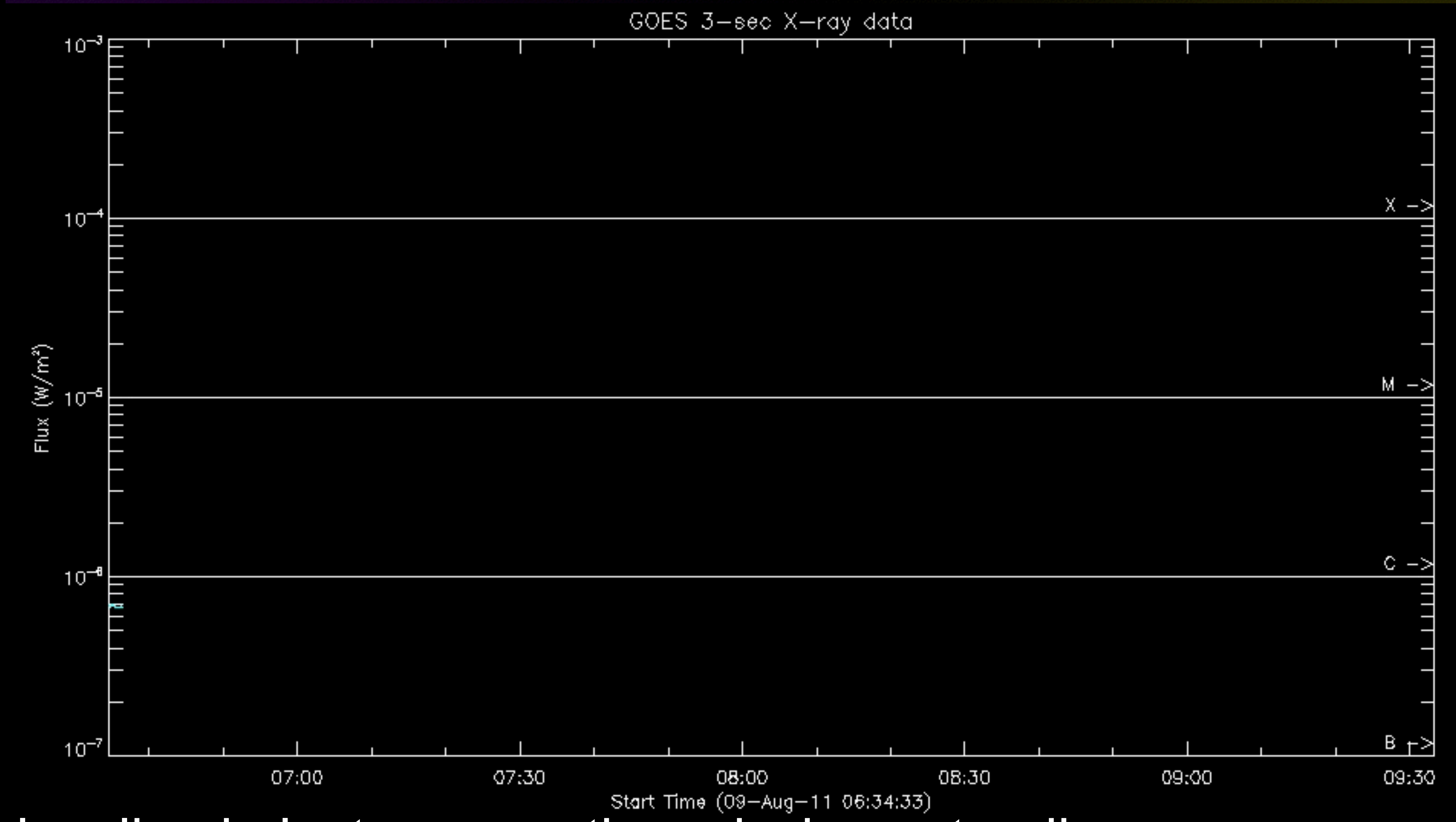
Goa, India, Jan 24 - 29, 2016

OUTLINE

- Solar flares and the prediction challenge
- Validation of flare prediction methods
  - *Dichotomous validation*
  - *Probabilistic validation*
- Tailoring prediction methods to the customers' needs
  - *Multi-variable forecasting*
  - *Ensemble forecasting*
- Conclusion

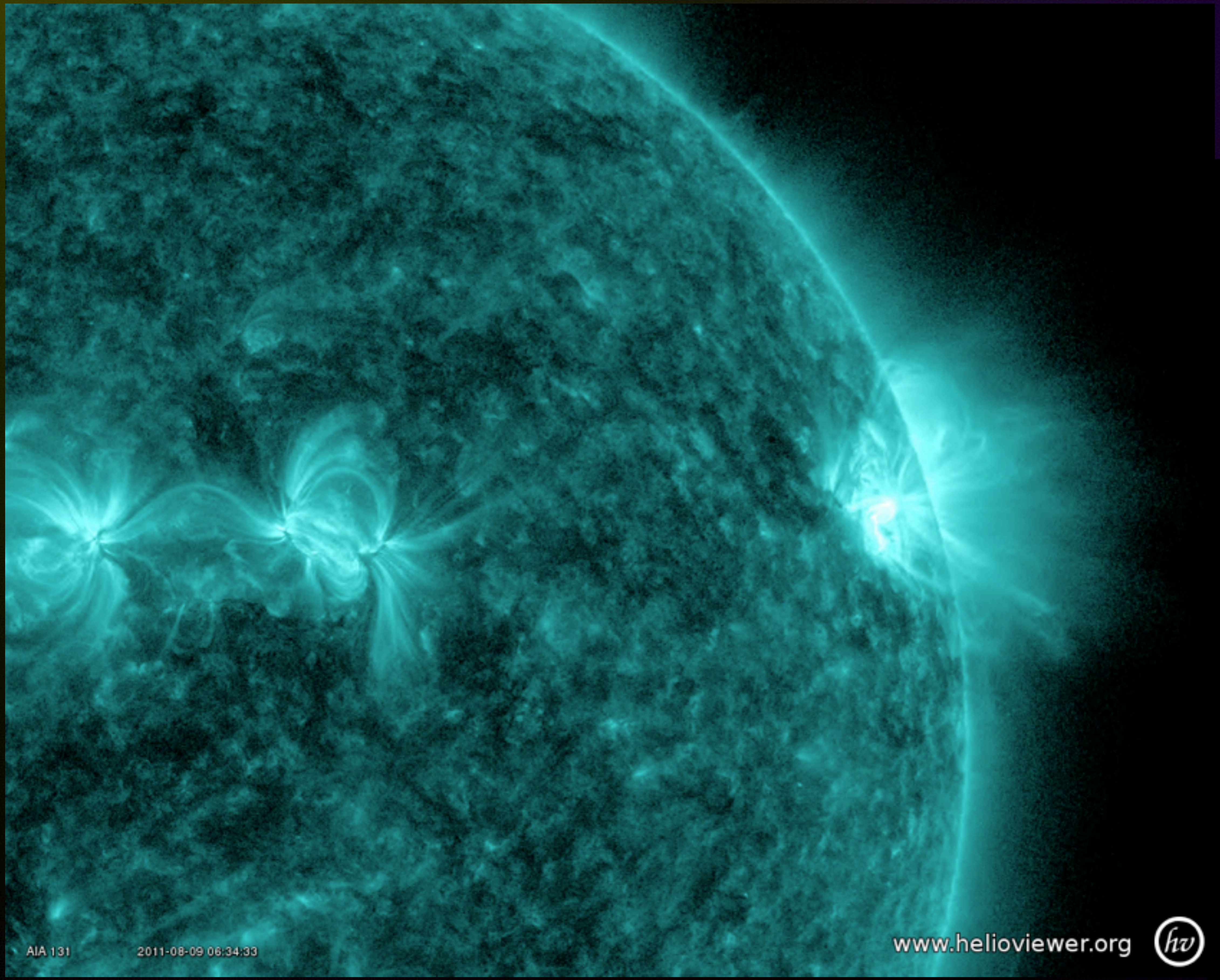
SOLAR FLARES: THE SINGLE …

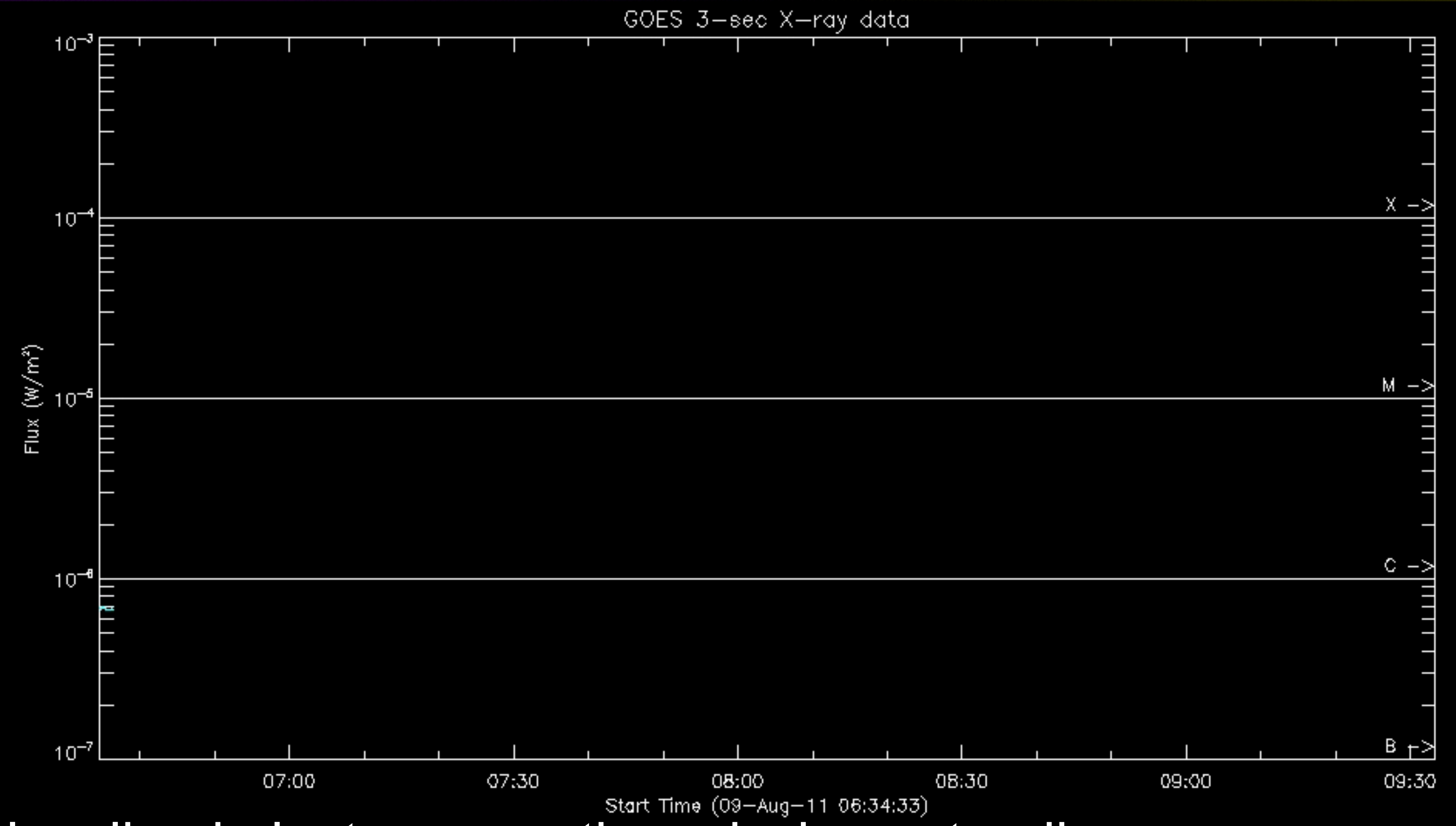




A sudden commencement of enhanced, localized electromagnetic emission extending over practically the entire range of the electromagnetic spectrum, from γ-ray to radio wavelengths

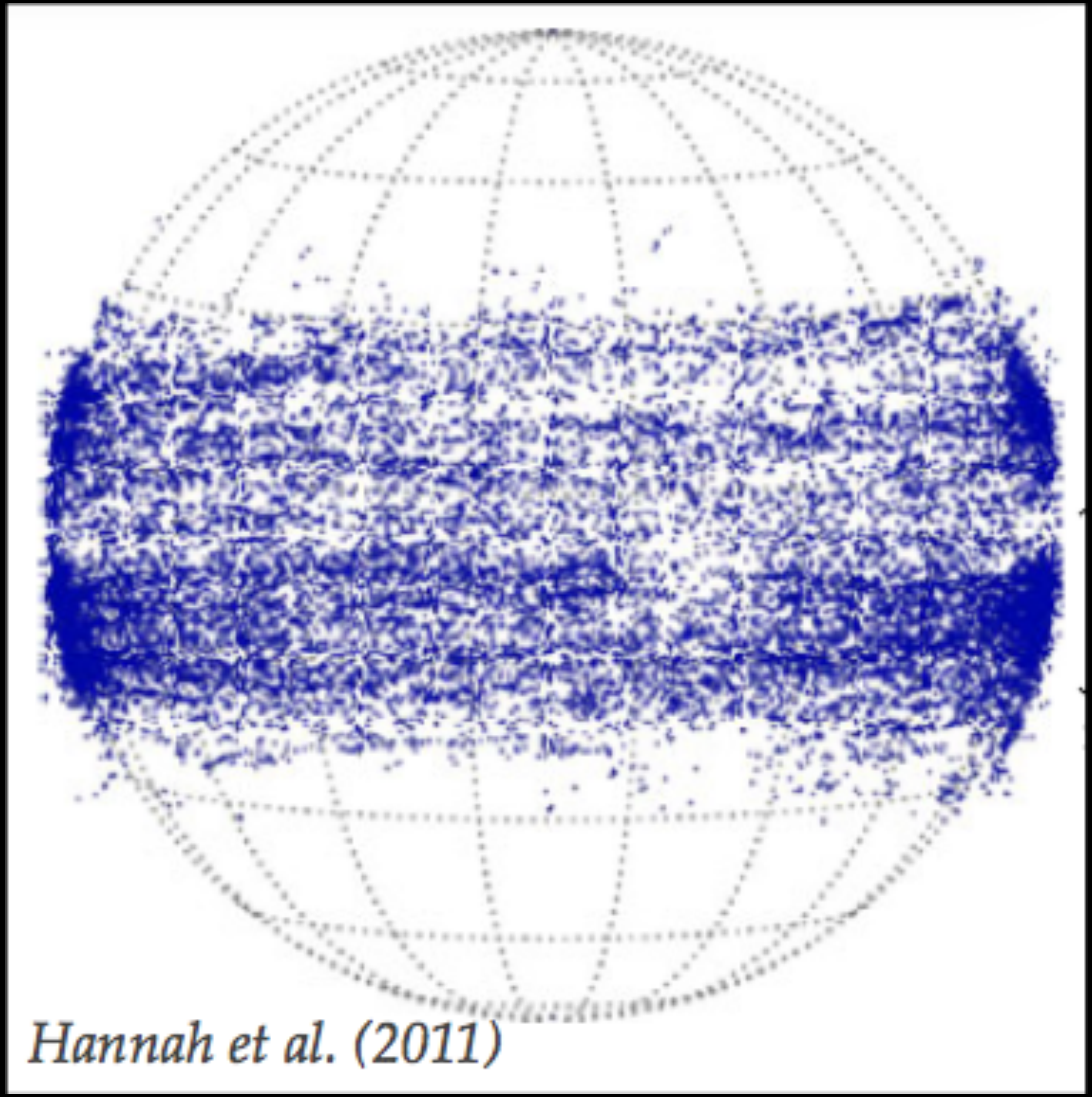
SOLAR FLARES: THE SINGLE …



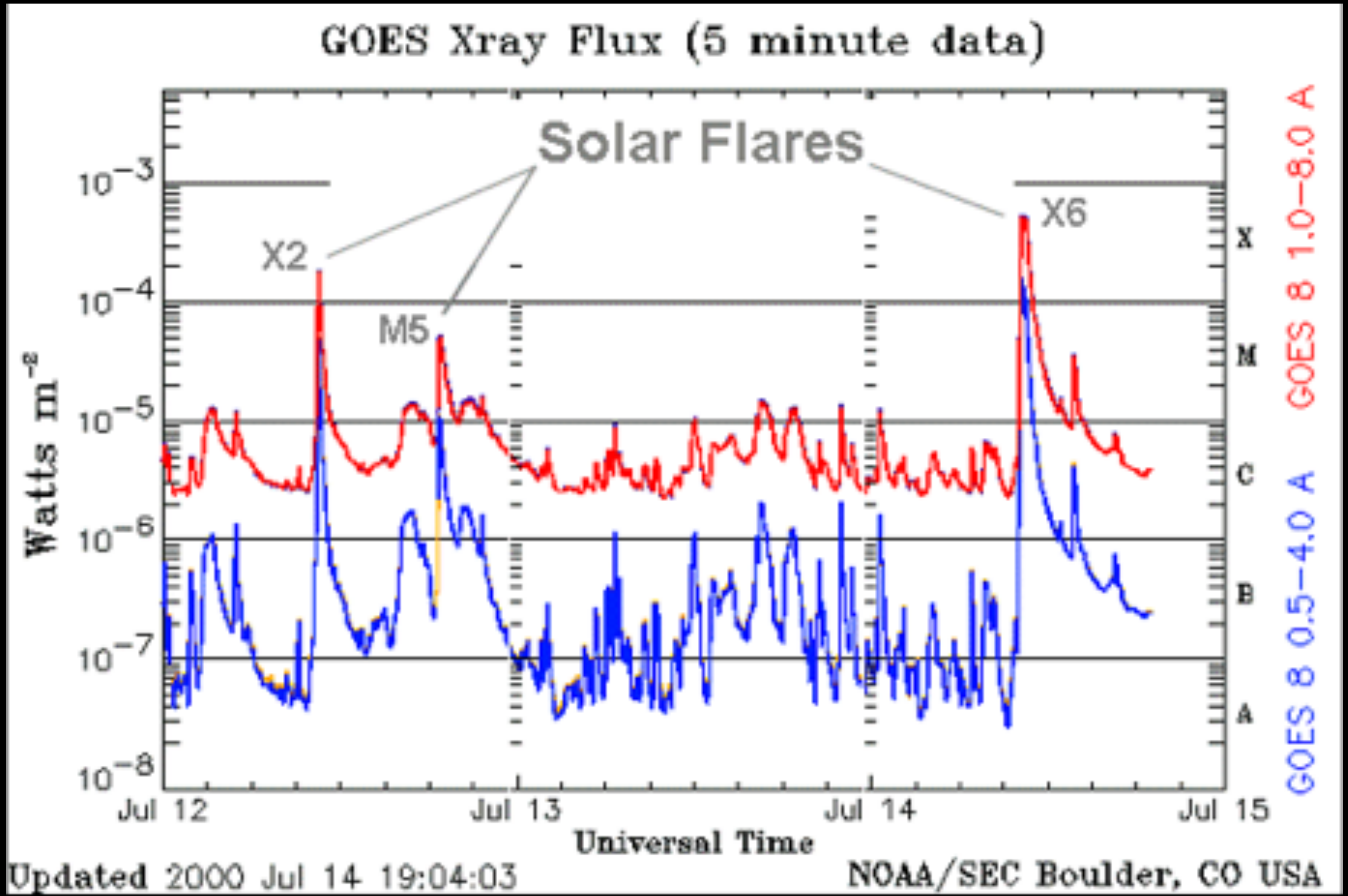


A sudden commencement of enhanced, localized electromagnetic emission extending over practically the entire range of the electromagnetic spectrum, from γ-ray to radio wavelengths

SOLAR FLARES: … AND THE PLENTY

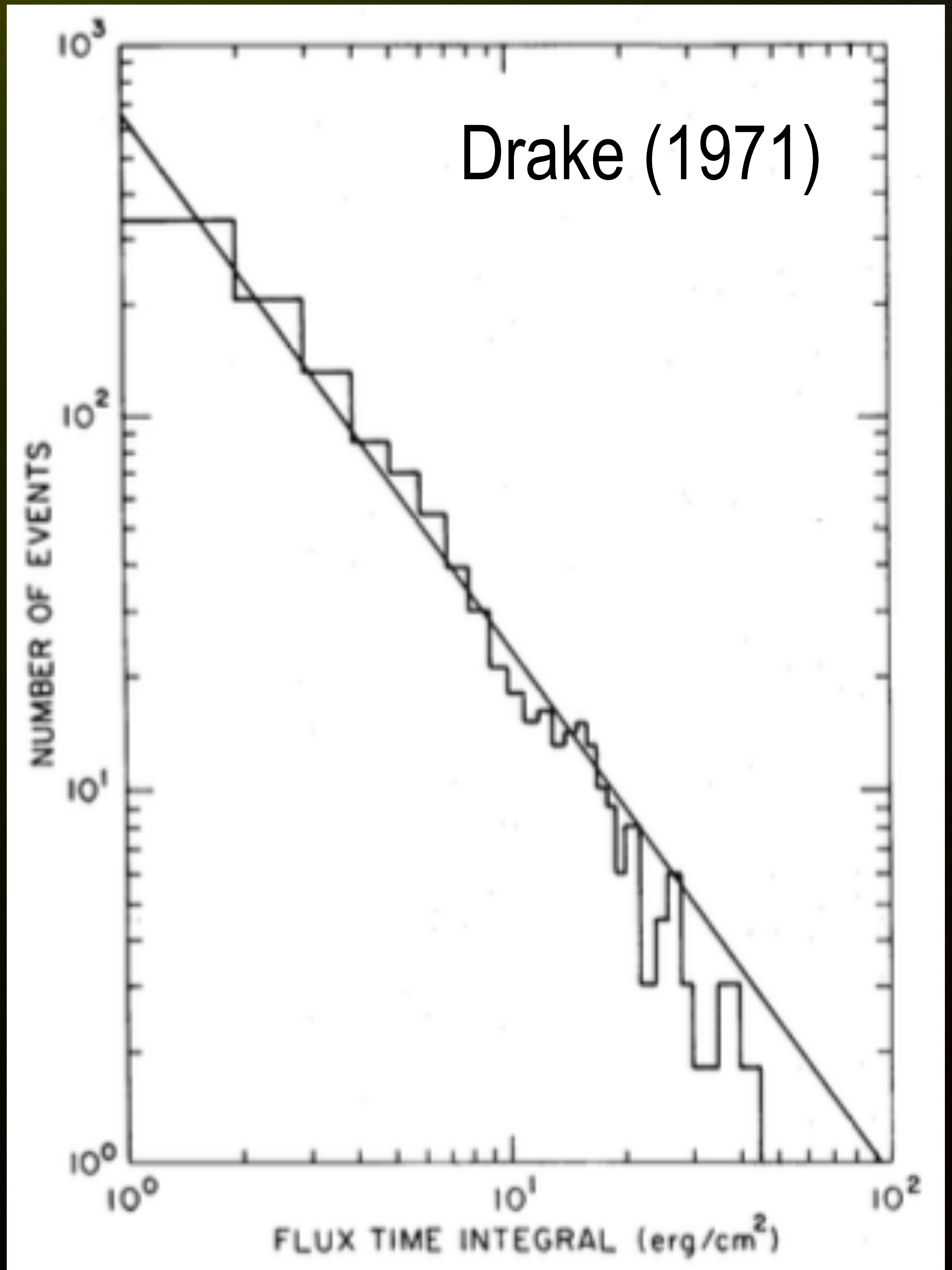


Location of some 27,000 flares from the RHESSI database



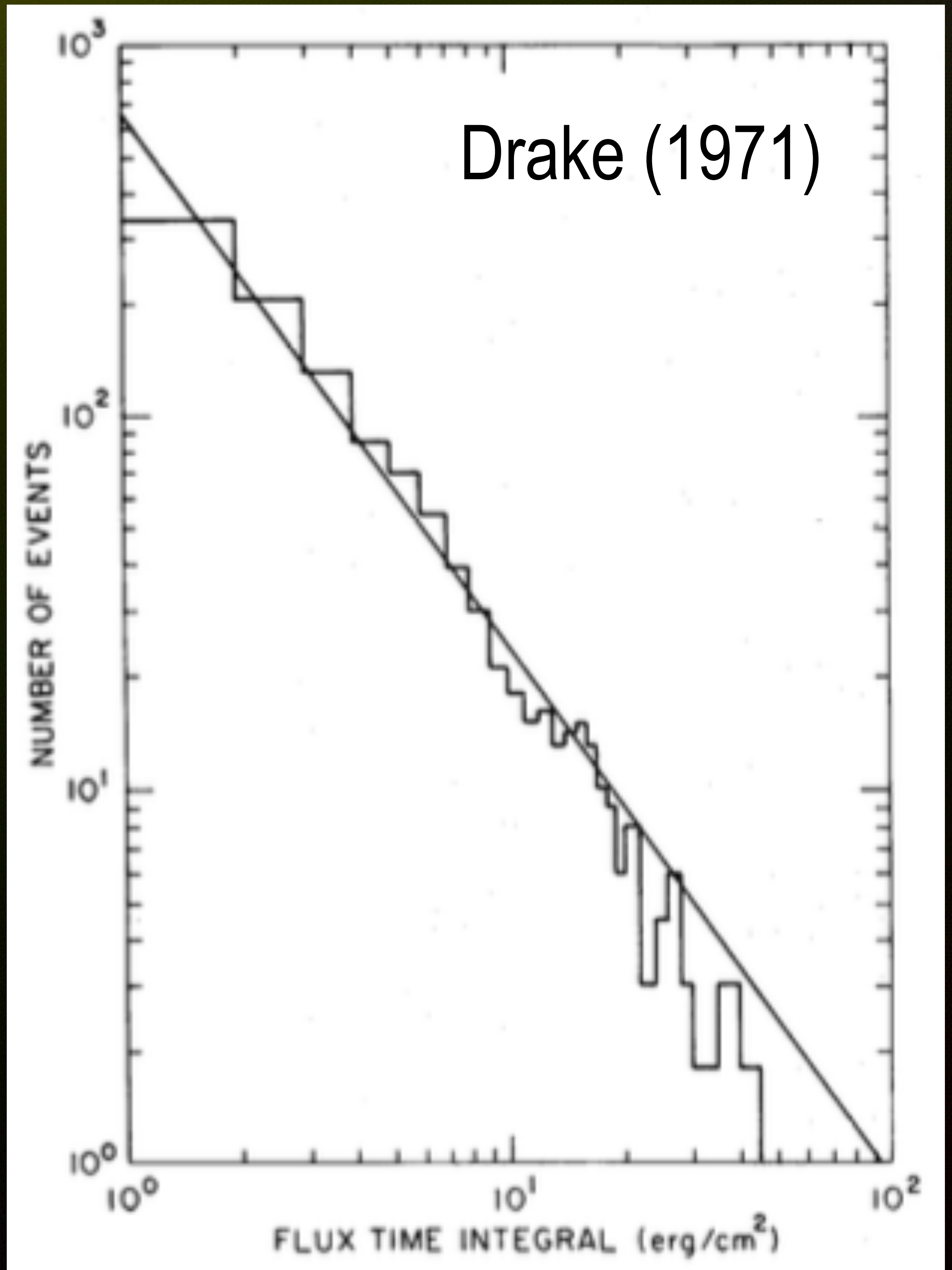
“Active” solar conditions over a 3-day period in July 2000

WHAT ARE SOLAR FLARES, PHYSICALLY AND STATISTICALLY?



Flare occurrence number vs. integrated photon flux

WHAT ARE SOLAR FLARES, PHYSICALLY AND STATISTICALLY?



Flare occurrence number vs. integrated photon flux

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

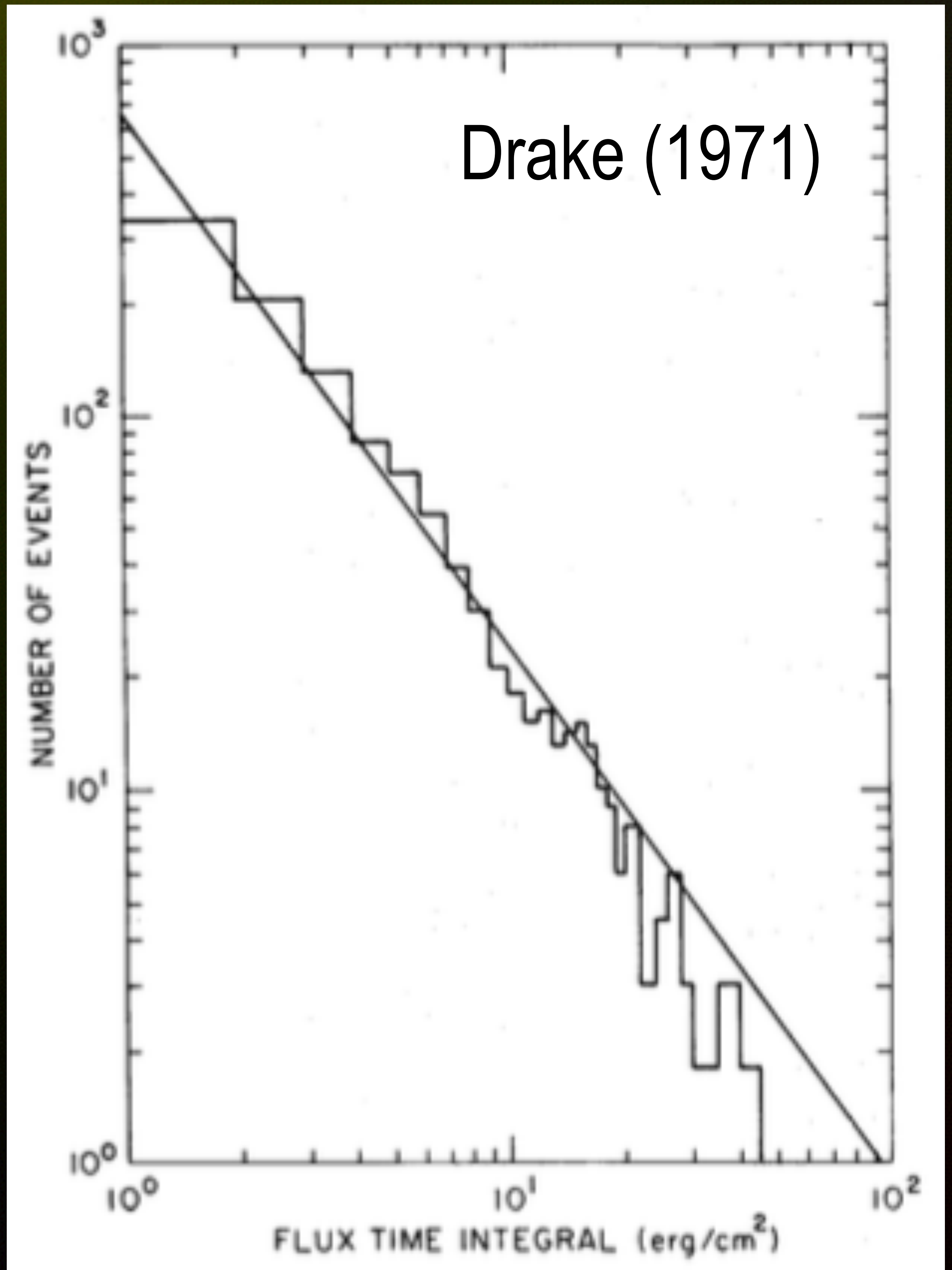
- Stochastic relaxation (storage and release) processes
- Physically uncoupled / independent
- Brief, comparing to intermediate times between flares

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

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

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

WHAT ARE SOLAR FLARES, PHYSICALLY AND STATISTICALLY?



Flare occurrence number vs. integrated photon flux

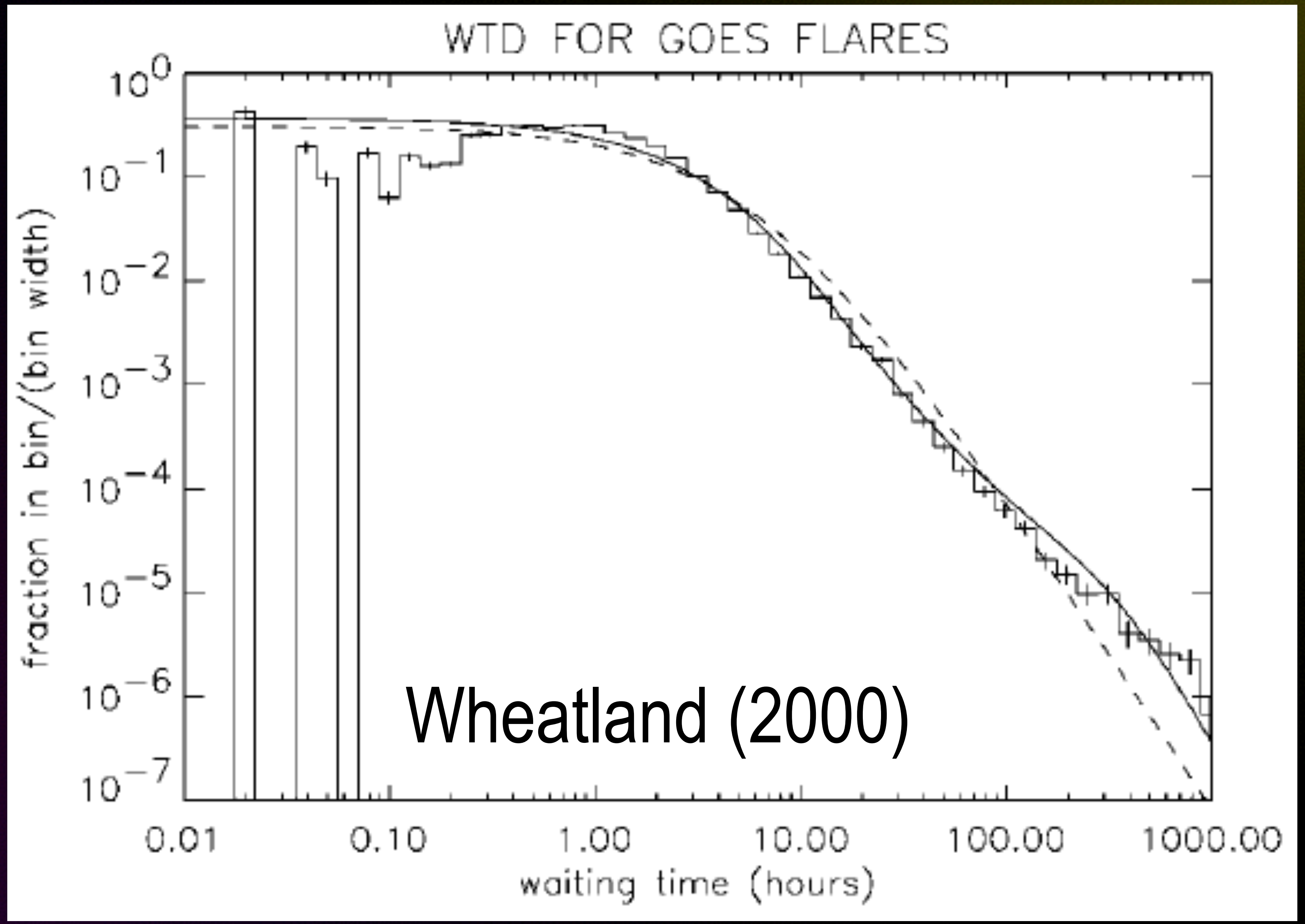
Flares are (**Rosner & Vaiana 1978**):

- 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 enegies

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

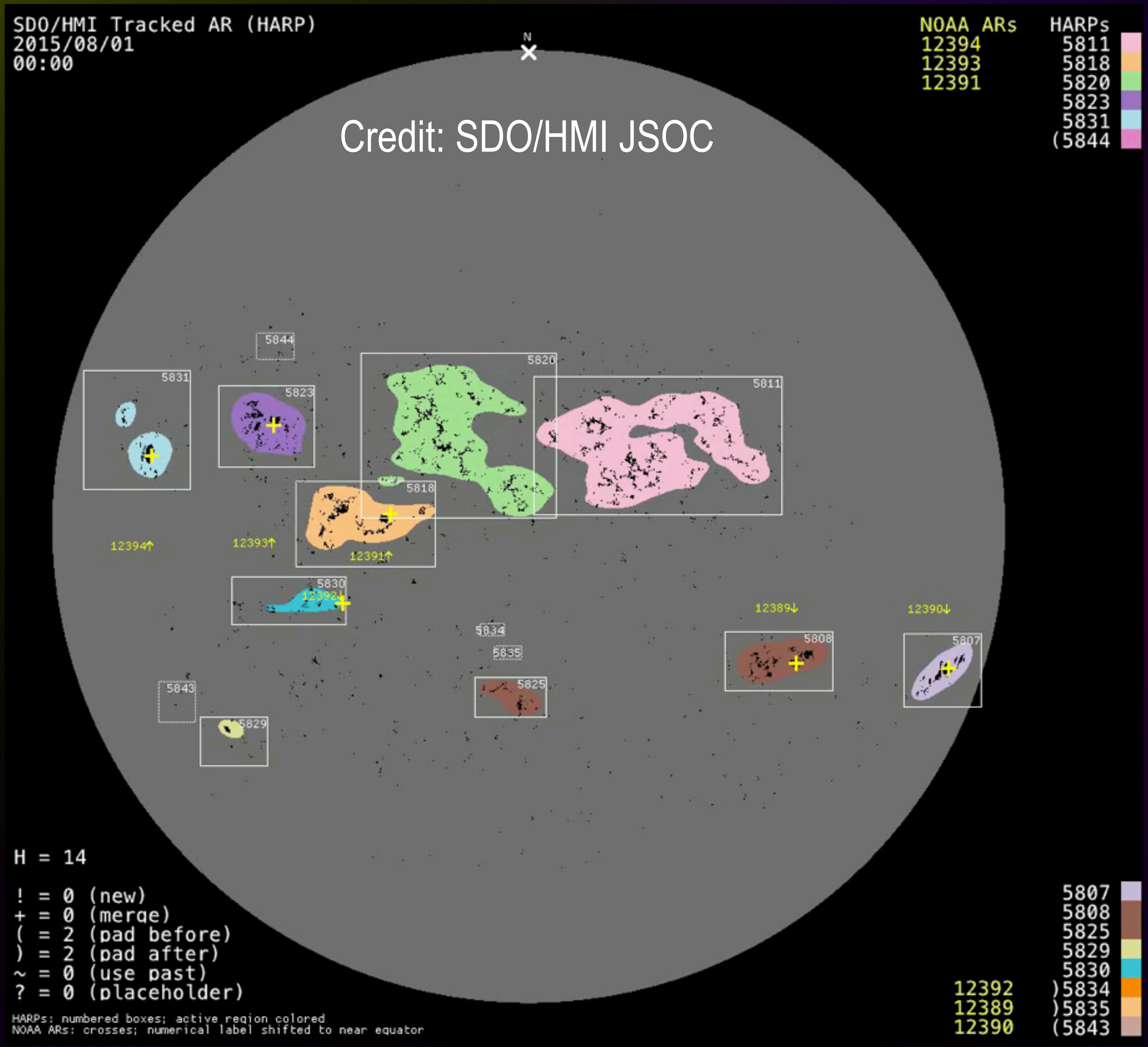


Flare occurrence is a time-dependent Poisson process (also Wheatland & Litvinenko 2002). This can explain power laws in the flare distribution functions.

$$P(k) = \frac{\bar{\nu}^k}{k!} e^{-\bar{\nu}}$$

Gallagher et al. (2002)

THE FLARE-PREDICTION CHALLENGE …

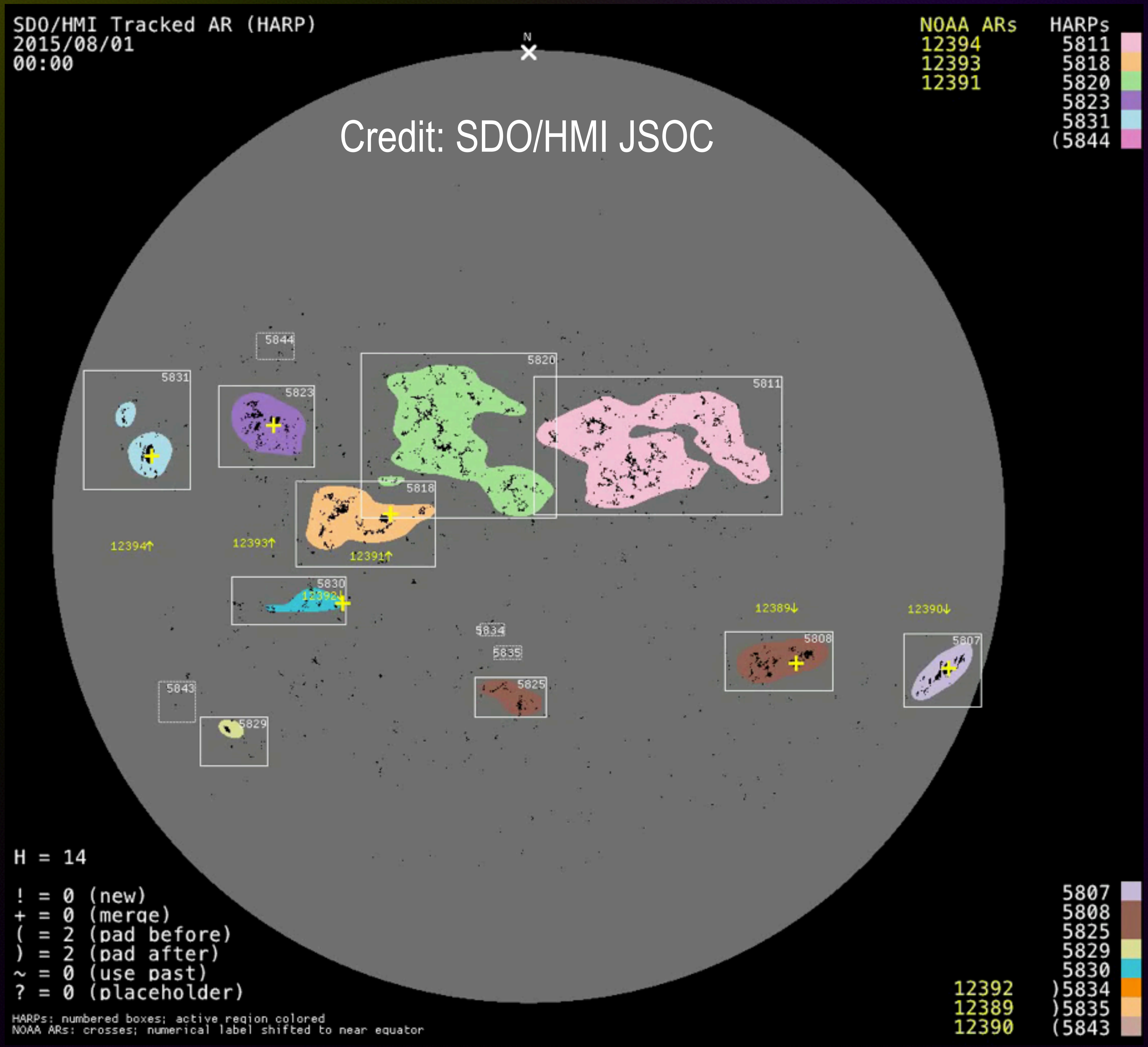


Objective: predict solar flares from near real-time (NRT) observations of solar evolution

- Prediction typically incolves solar photospheric (LOS or vector) magnetic field measurements. SDO/HMI is the most prominent source of these data
- Predictive parameters are inferred locally
- Observational cadence:
  - — 45 s, for full-disk LOS data
  - — 720 s, for full-disk vector data

Reference: SOHO/MDI full-disk LOS data: 5,760 s

THE FLARE-PREDICTION CHALLENGE …

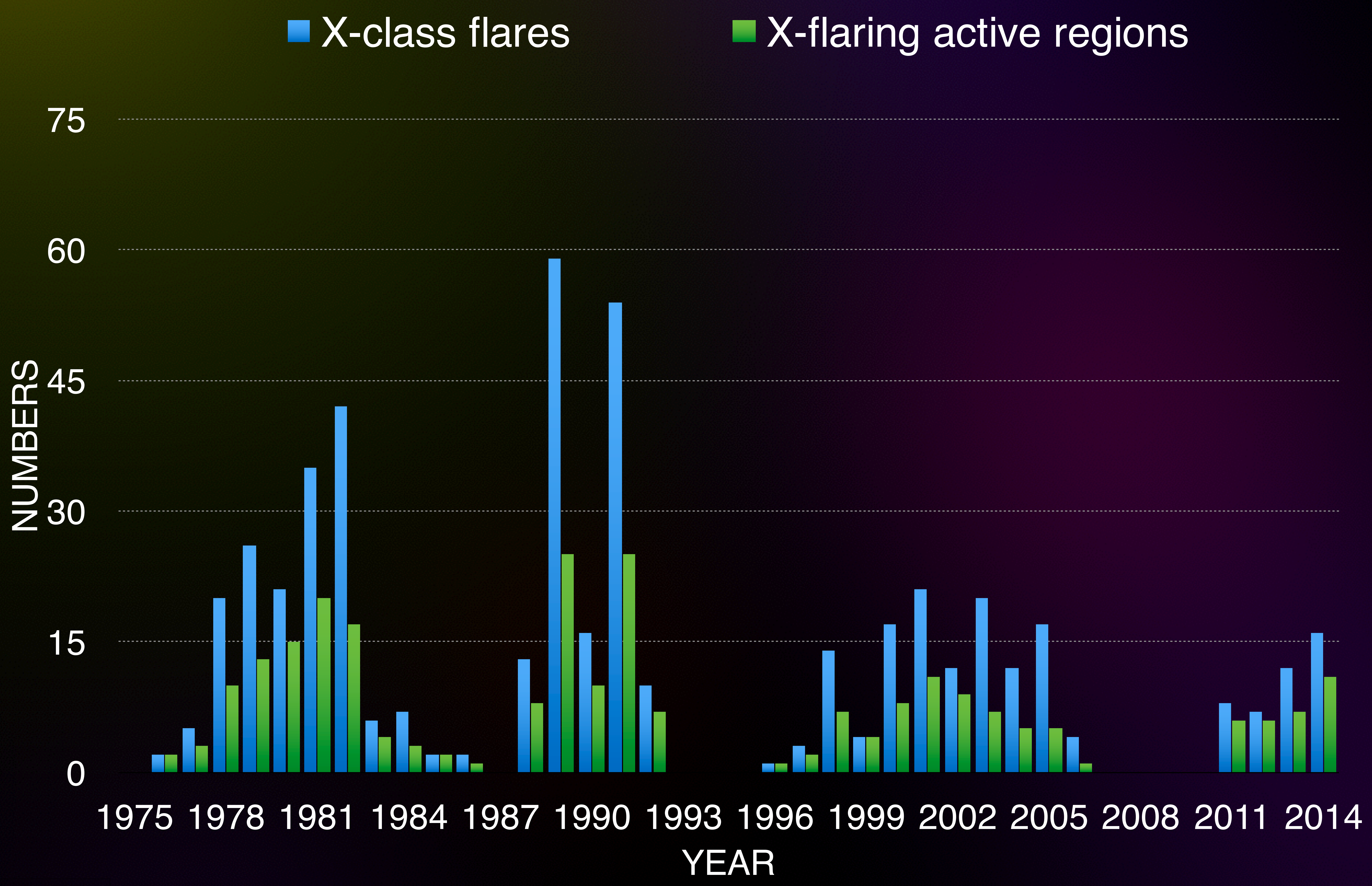


Objective: predict solar flares from near real-time (NRT) observations of solar evolution

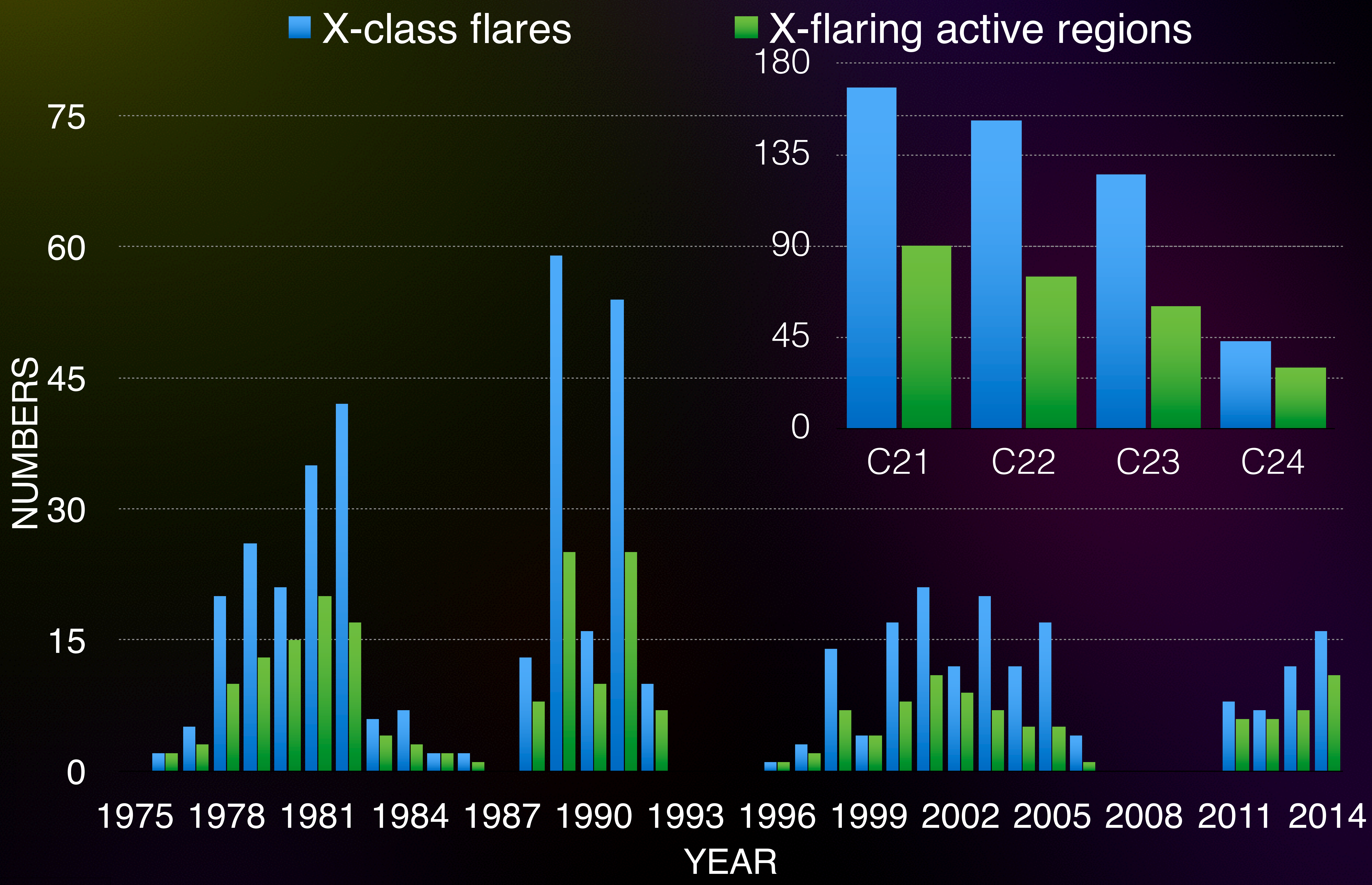
- Prediction typically incolves solar photospheric (LOS or vector) magnetic field measurements. SDO/HMI is the most prominent source of these data
- Predictive parameters are inferred locally
- Observational cadence:
  - — 45 s, for full-disk LOS data
  - — 720 s, for full-disk vector data

Reference: SOHO/MDI full-disk LOS data: 5,760 s

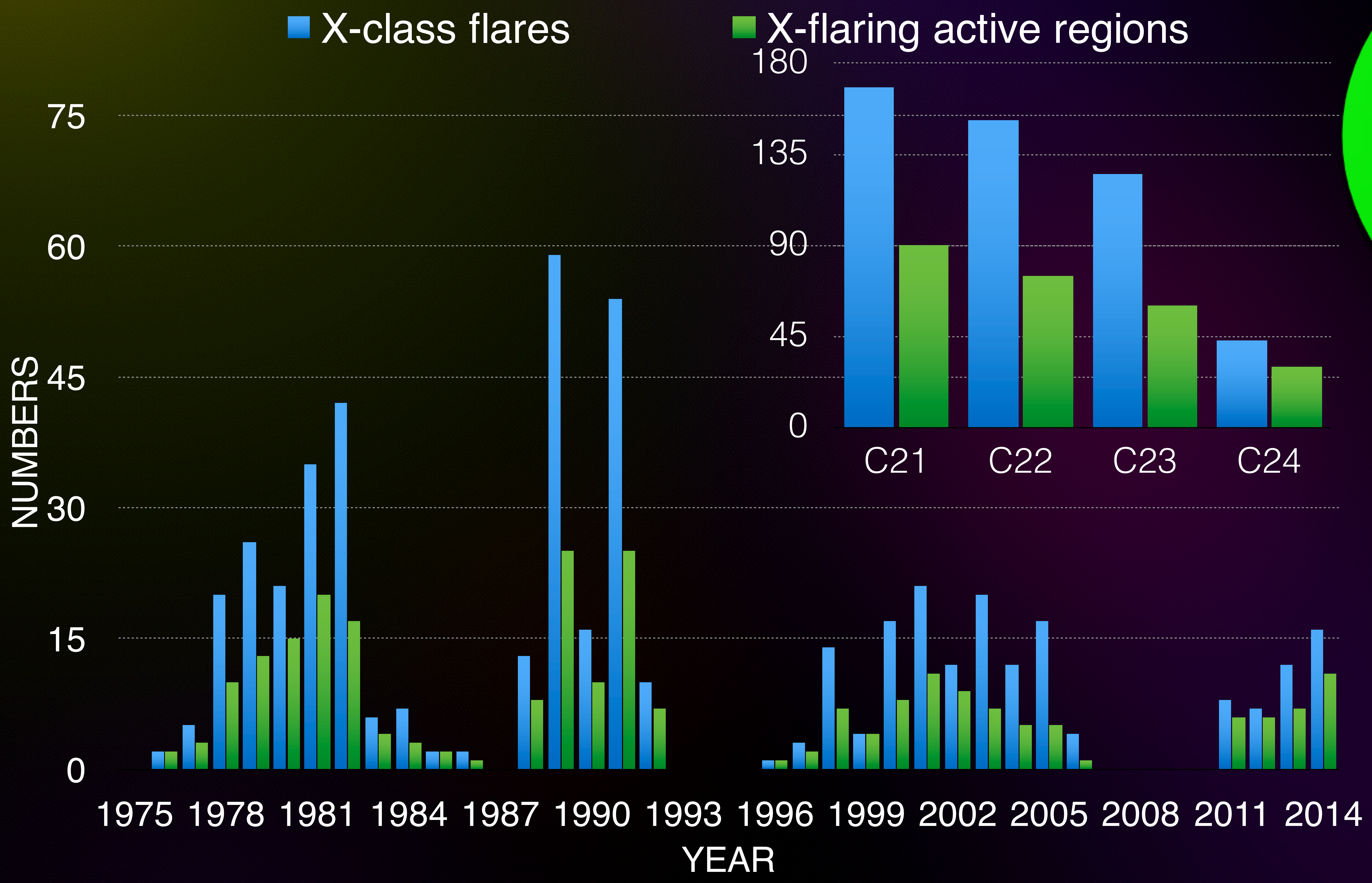
… AND ITS SPECIFICS

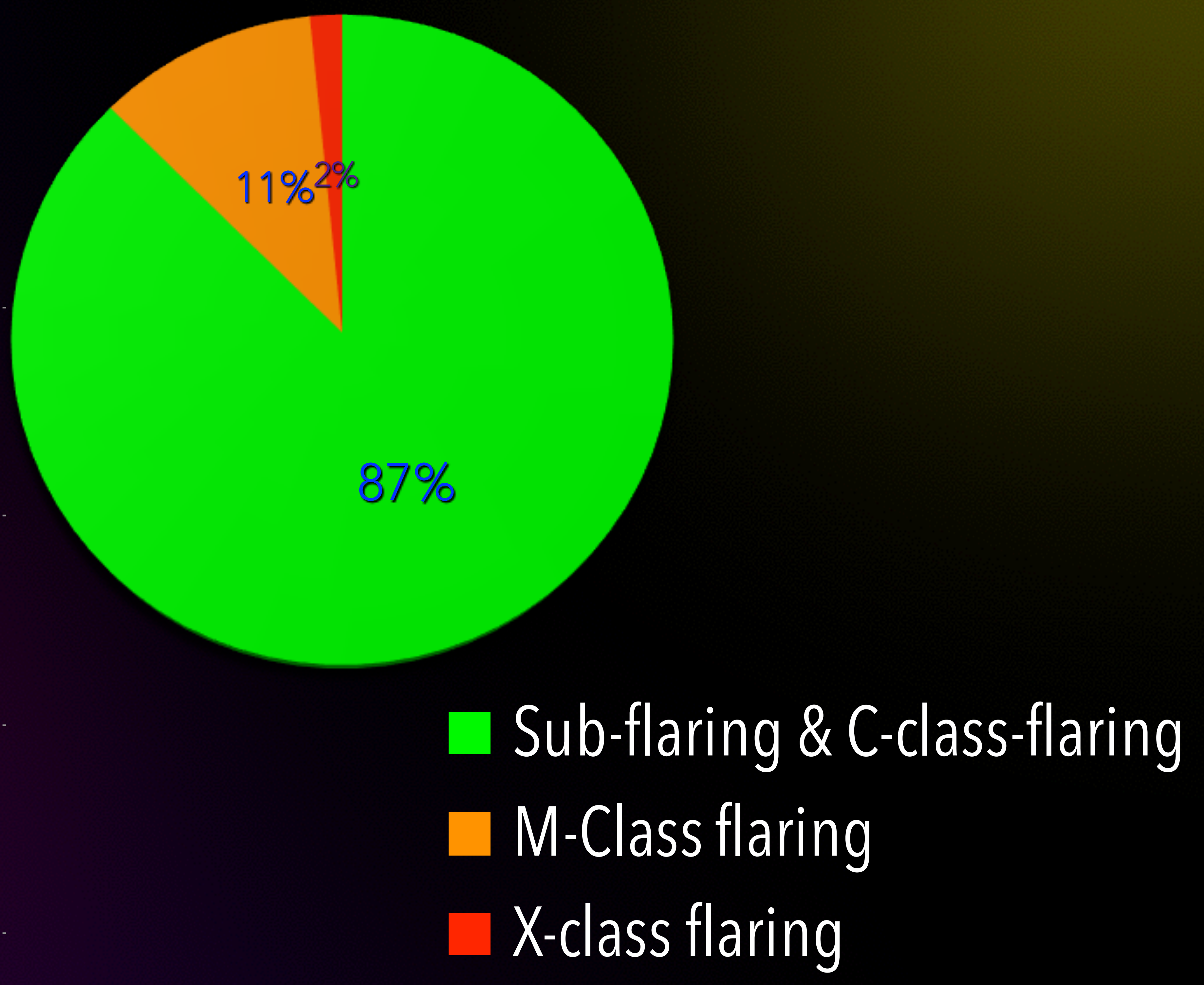


… AND ITS SPECIFICS



… AND ITS SPECIFICS

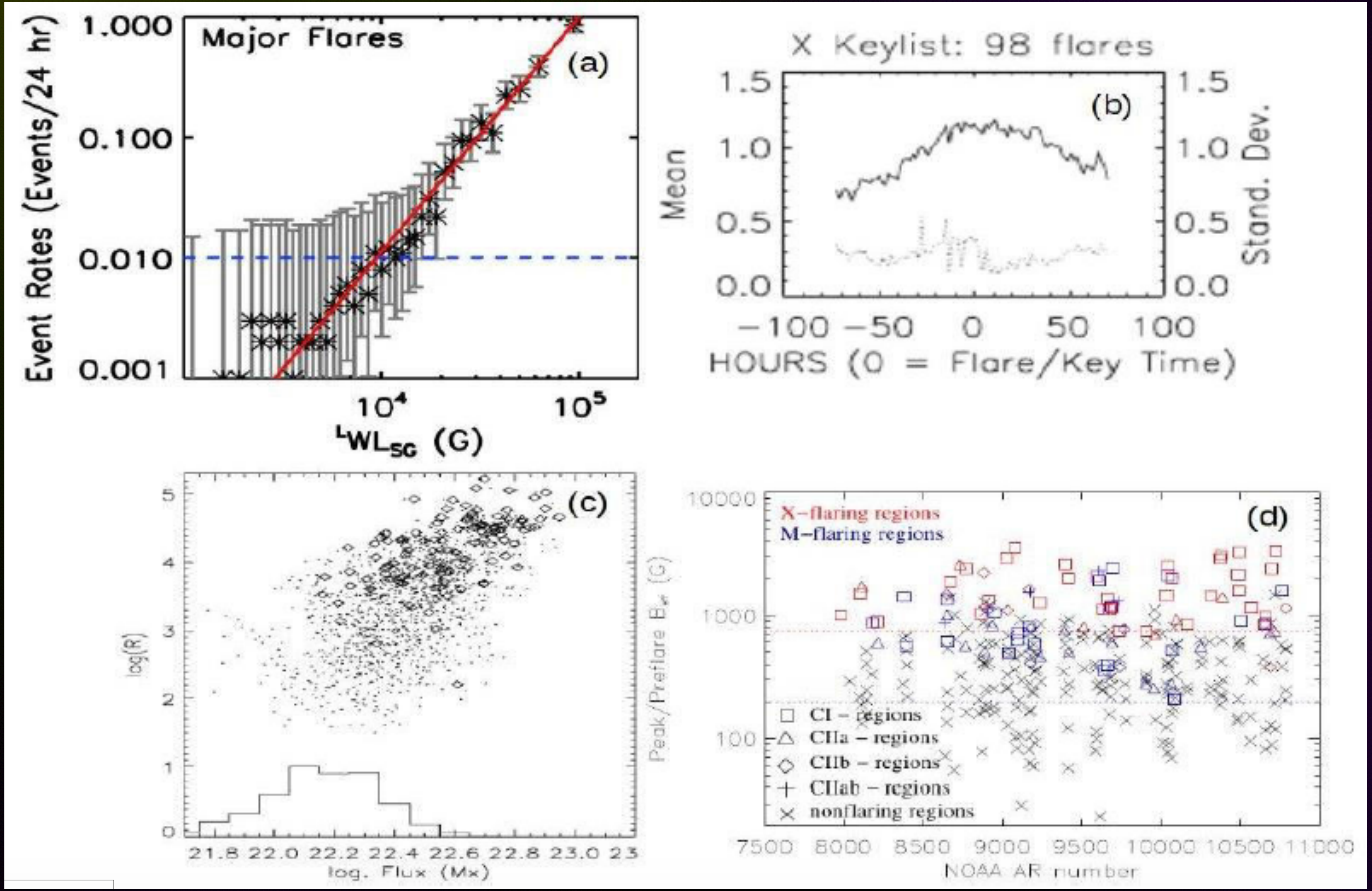




Only ~2% of active regions gave at least one X-class flare in C23!

DIFFERENT METHODS IN OPERATION WORLDWIDE

Mainly aiming to quantify the magnetic complexity of the host active regions



Upper left: WL_{SG} - Falconer et al. (2011)

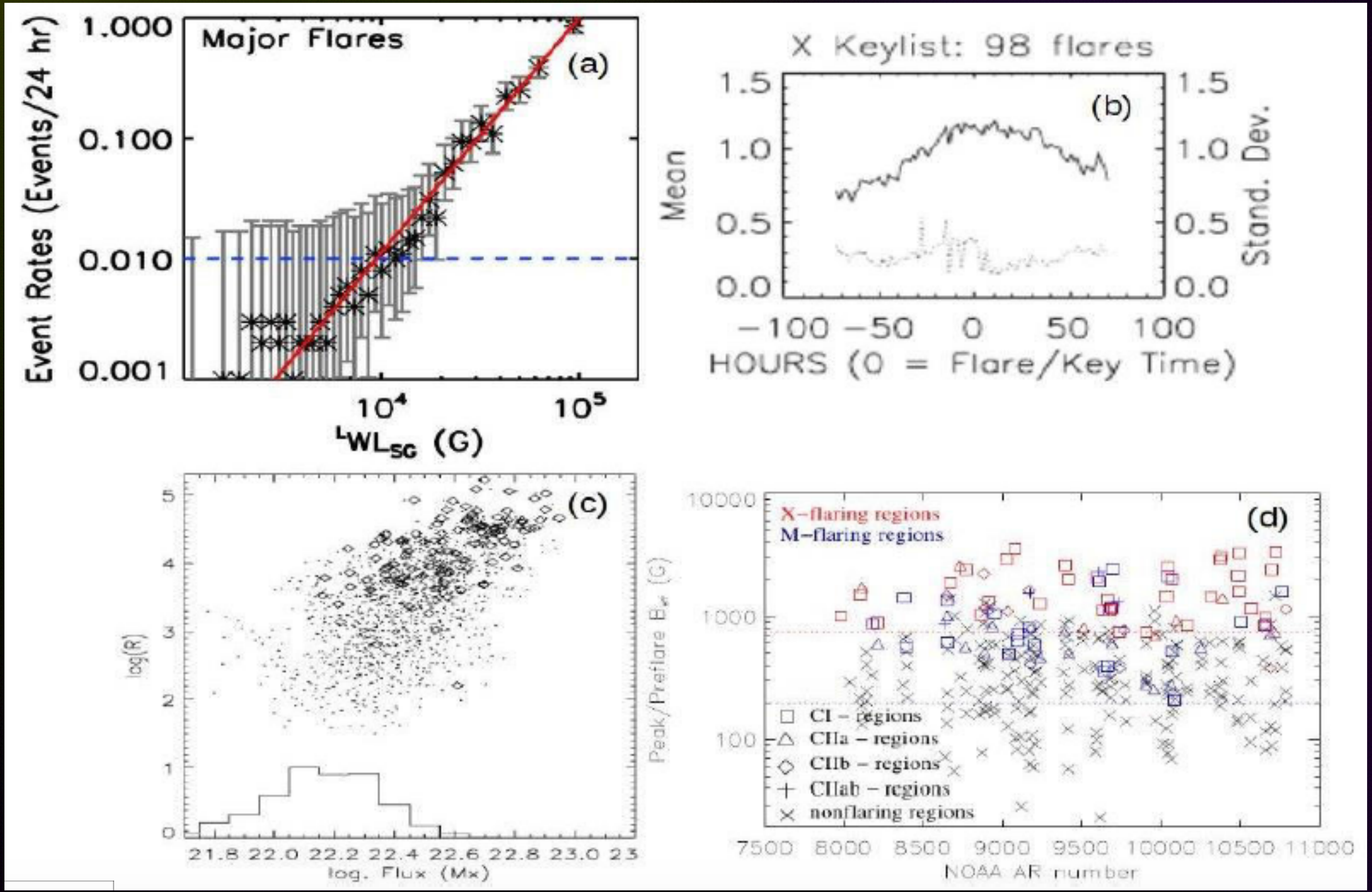
Upper right: GWILL 0- Mason & Hoeksema (2010)

Lower left: R - Schrijver (2007)

Lower right: B_{eff} - Georgoulis & Rust (2007)

DIFFERENT METHODS IN OPERATION WORLDWIDE

Mainly aiming to quantify the magnetic complexity of the host active regions



Upper left: WL_{SG} - Falconer et al. (2011)

Upper right: GWILL 0- Mason & Hoeksema (2010)

Lower left: R - Schrijver (2007)

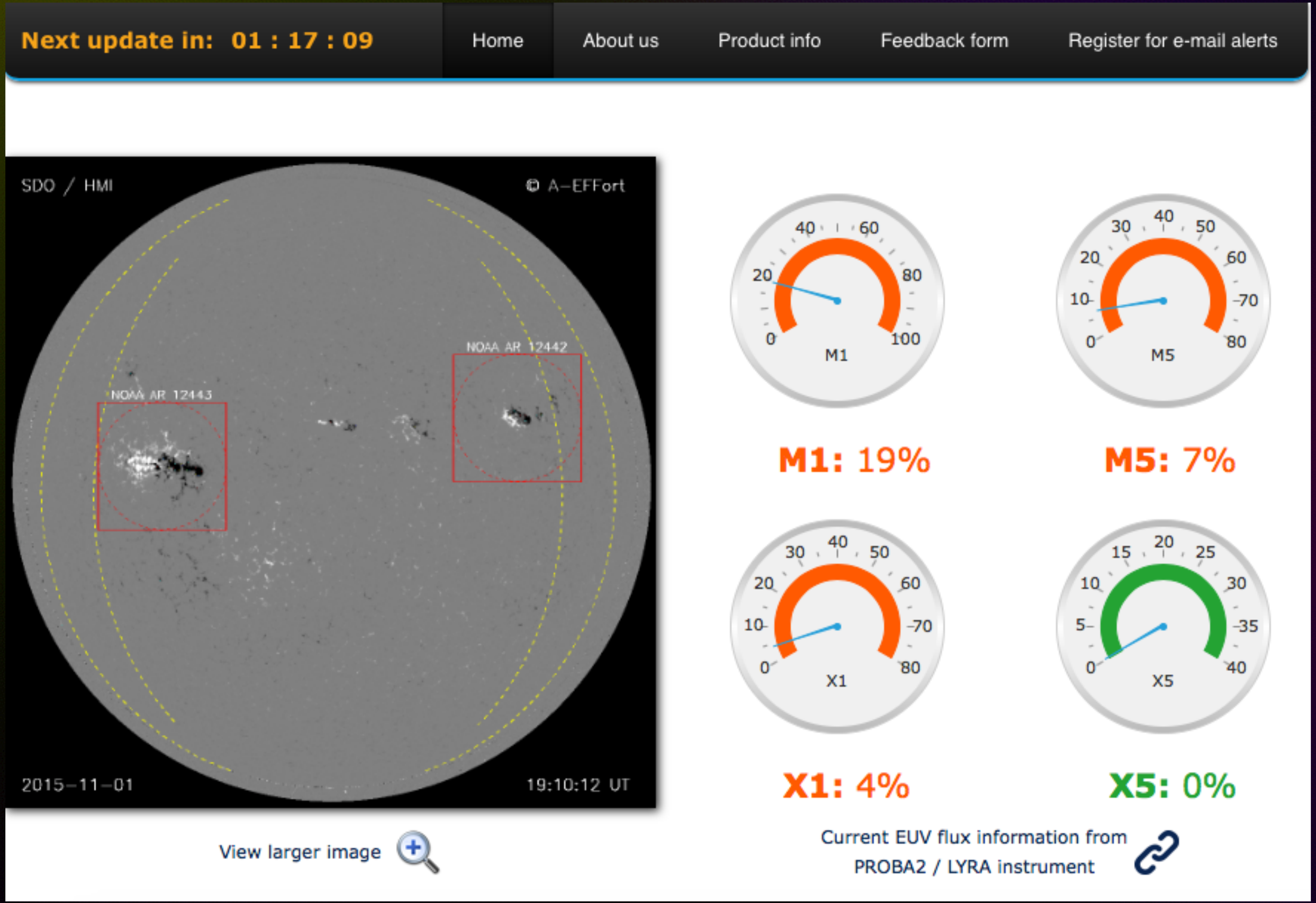
Lower right: B_{eff} - Georgoulis & Rust (2007)

Operational flare-prediction services:

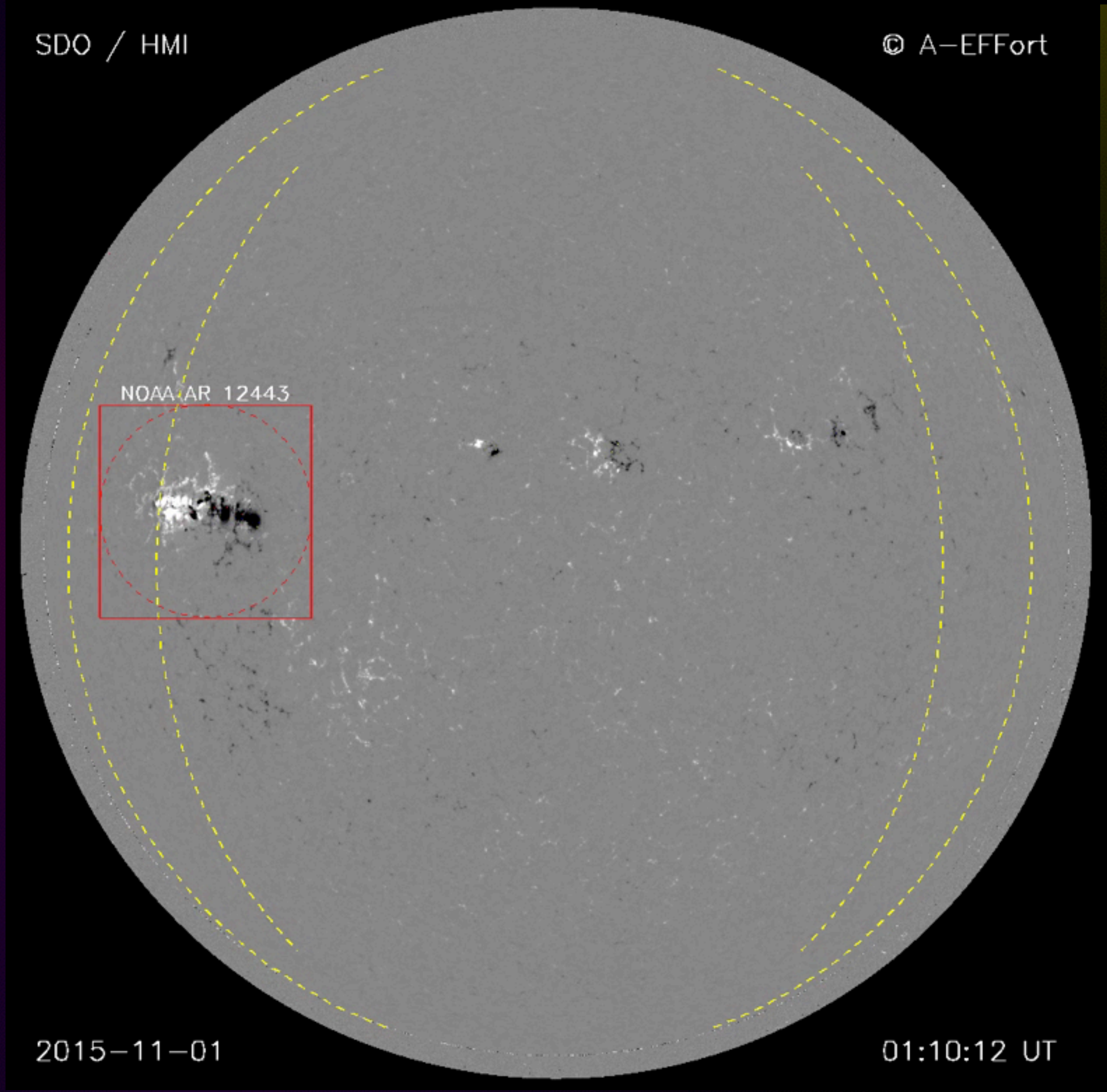
- NOAA / SWPC
- Max Millennium / Solar Monitor
- UK Met Office
- U. Bradford / ASAP
- NASA / SRAG / MAG4
- ASSA / KSPC
- ESA / A-EFFort

ESA's SSA / SWE A-EFFort METHOD

A-EFFort
Athens Effective
Solar Flare Forecasting

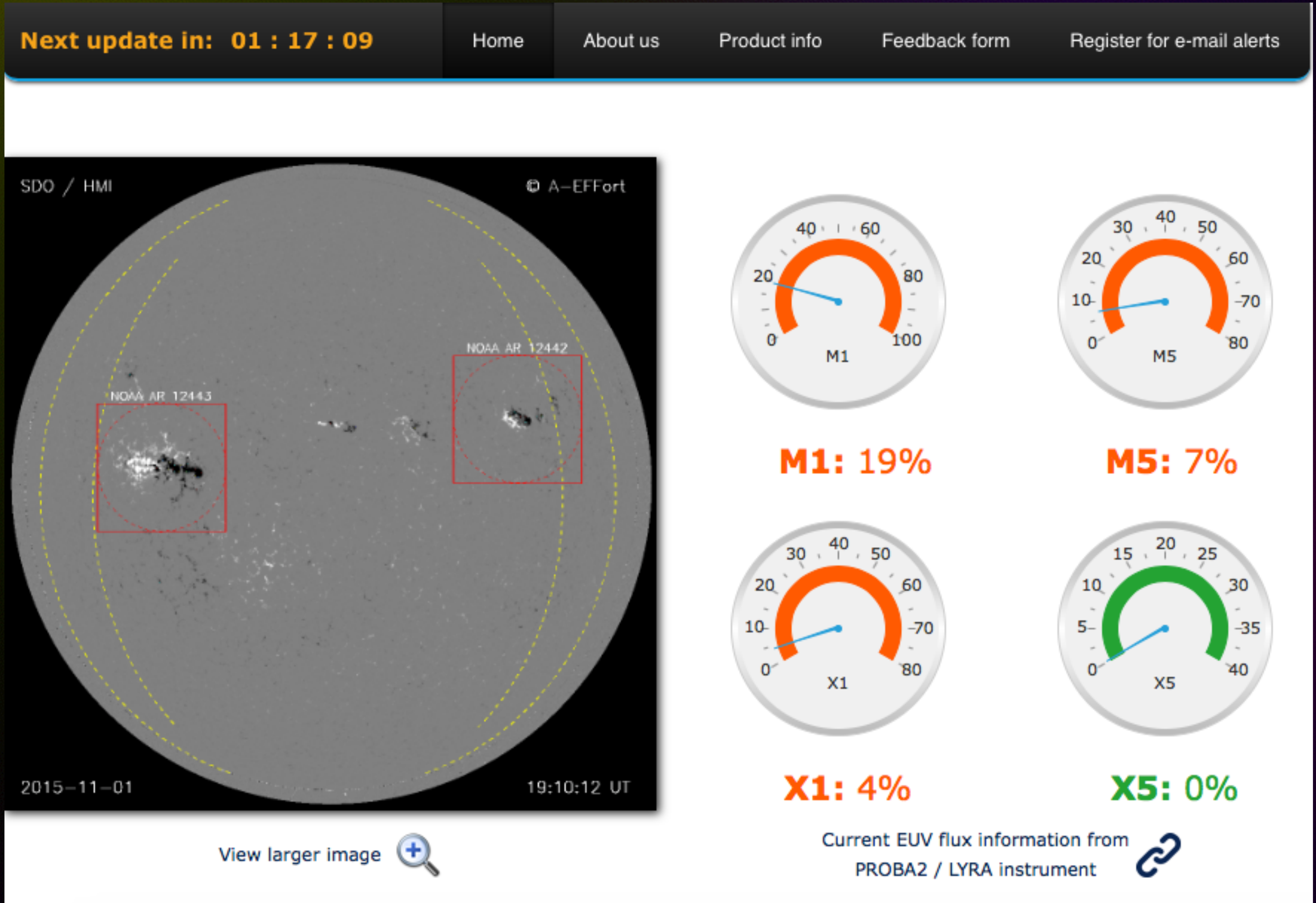


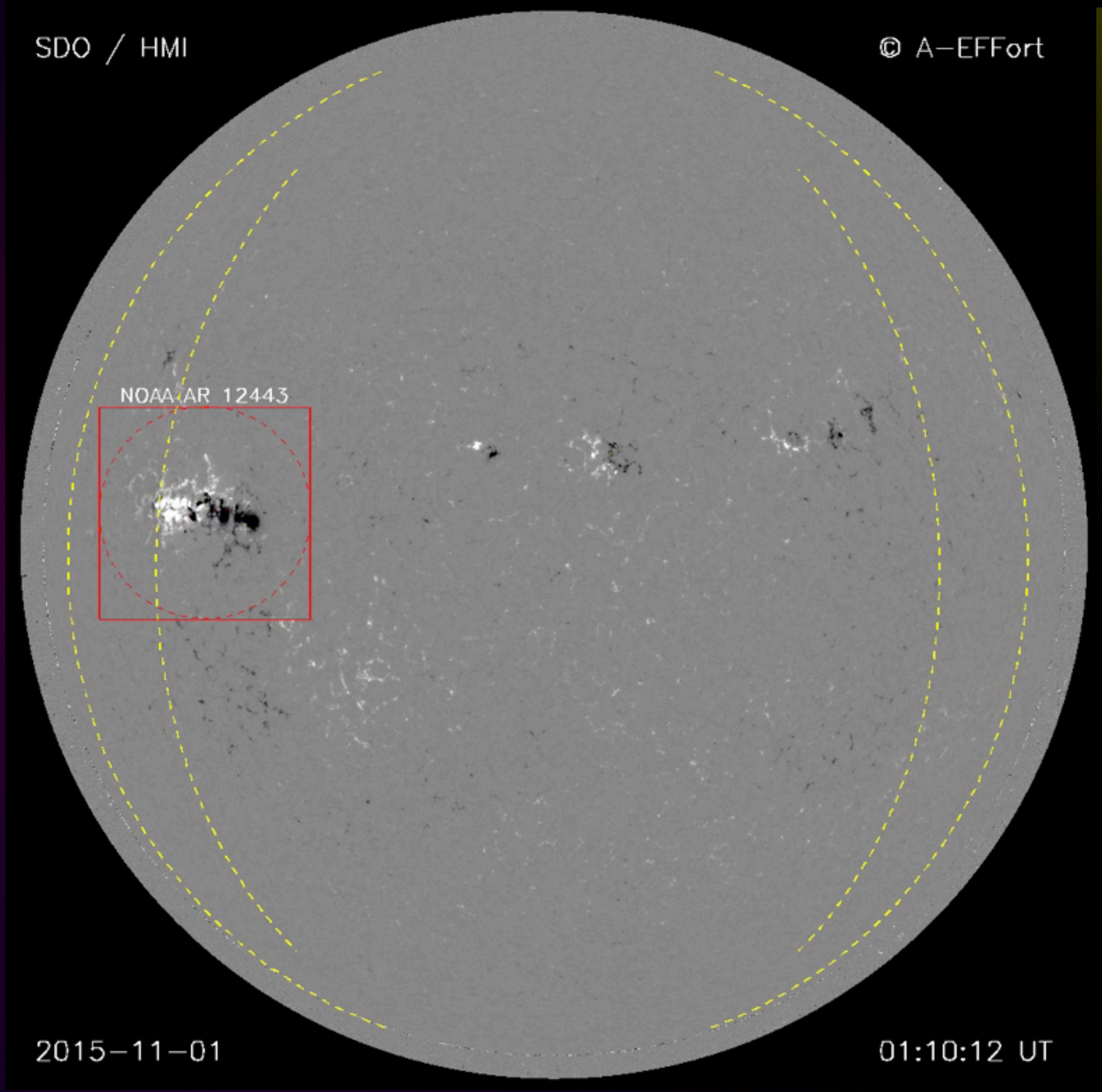
ESA's SSA/SWE A-EFFort interface a-effort.academyofathens.gr



ESA's SSA / SWE A-EFFort METHOD

A-EFFort
Athens Effective
Solar Flare Forecasting





ESA's SSA/SWE A-EFFort interface

a-effort.academyofathens.gr

IN FACT, MANY DIFFERENT METHODS…

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	$\text{Area} = \sum \text{Pixels}$	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 with shear > 45° / total area	740.8	Included
MEANSHR	Mean shear angle	$\overline{\Gamma} = \frac{1}{N} \sum \arccos\left(\frac{\boldsymbol{B}^{\rm Obs} \cdot \boldsymbol{B}^{\rm Pot}}{\lvert B^{\rm Obs} \rvert \lvert B^{\rm Pot} \rvert}\right)$	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	$\overline{\lvert \nabla B_{\rm tot} \rvert} = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2}$	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	$\alpha_{\rm total} \propto \frac{\sum J_z \cdot B_z}{\sum B_z^2}$	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

Bobra & Couvidat, ApJ, 2015

- plus many more:
  - * fractal / multifractal parameters
  - * other morphological parameters
  - * statistical / historical parameters

IN FACT, MANY DIFFERENT METHODS…

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

- plus many more:

  * fractal / multifractal parameters

  * other morphological parameters

  * statistical / historical parameters

What is the optimal way to deal with all this information and still achieve reliable NRT forecasts within a preset forecast window?

Bobra & Couvidat, ApJ, 2015

VALIDATION : BORROWED BY TERRESTRIAL WEATHER FORECASTING

Categorical (dichotomous) 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:

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

VALIDATION : BORROWED BY TERRESTRIAL WEATHER FORECASTING

Categorical (dichotomous) 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:

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

- 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):

$$TSS = POD - POFD$$

CATEGORICAL 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

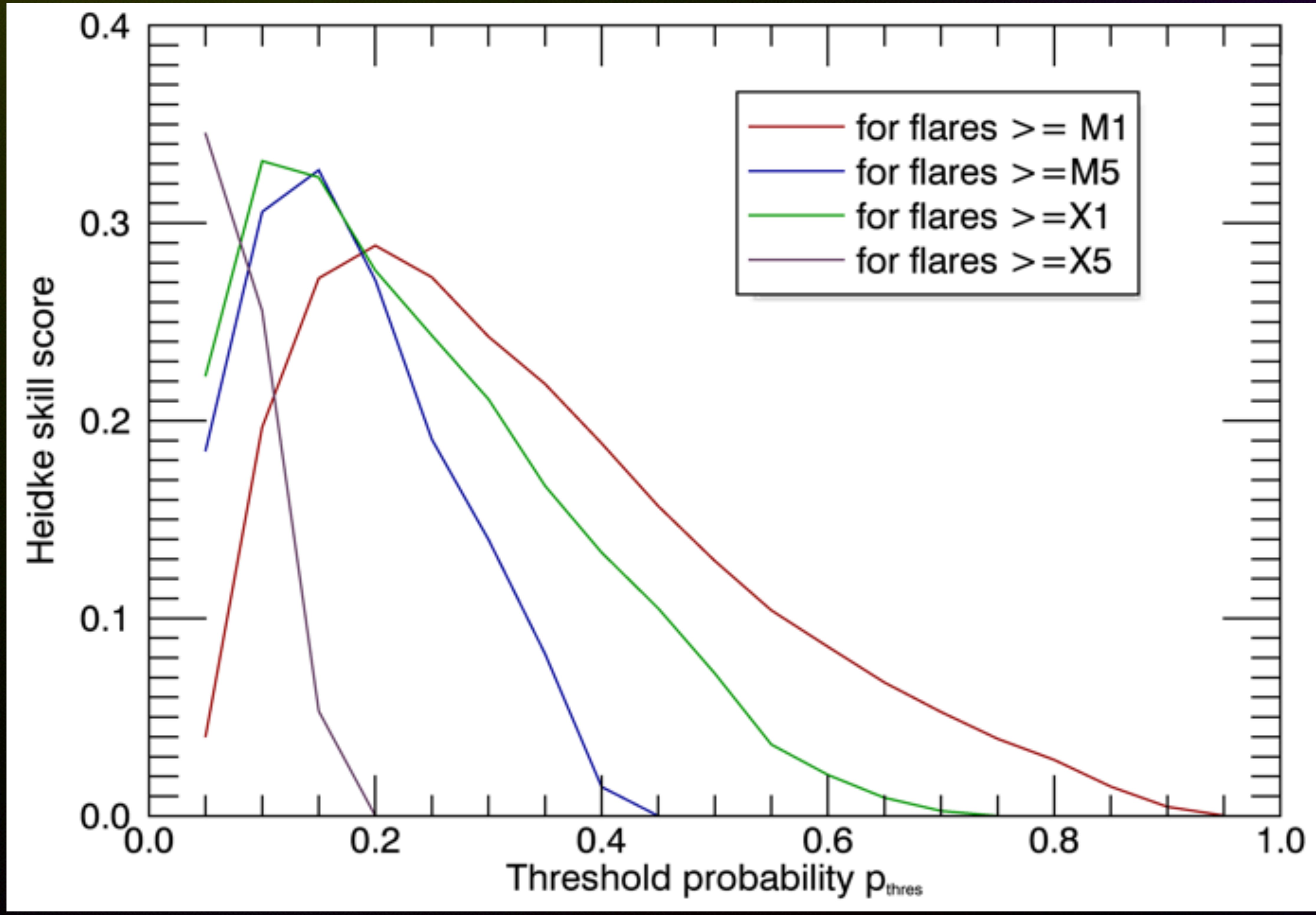
Slide courtesy: Shaun Bloomfield

CATEGORICAL VALIDATION IN A PROBABILISTIC FORECASTING?

One needs a threshold probability p_{thres} ($p \geq p_{thres}$ —> YES ; $p < p_{thres}$ —> NO)

CATEGORICAL VALIDATION IN A PROBABILISTIC FORECASTING?

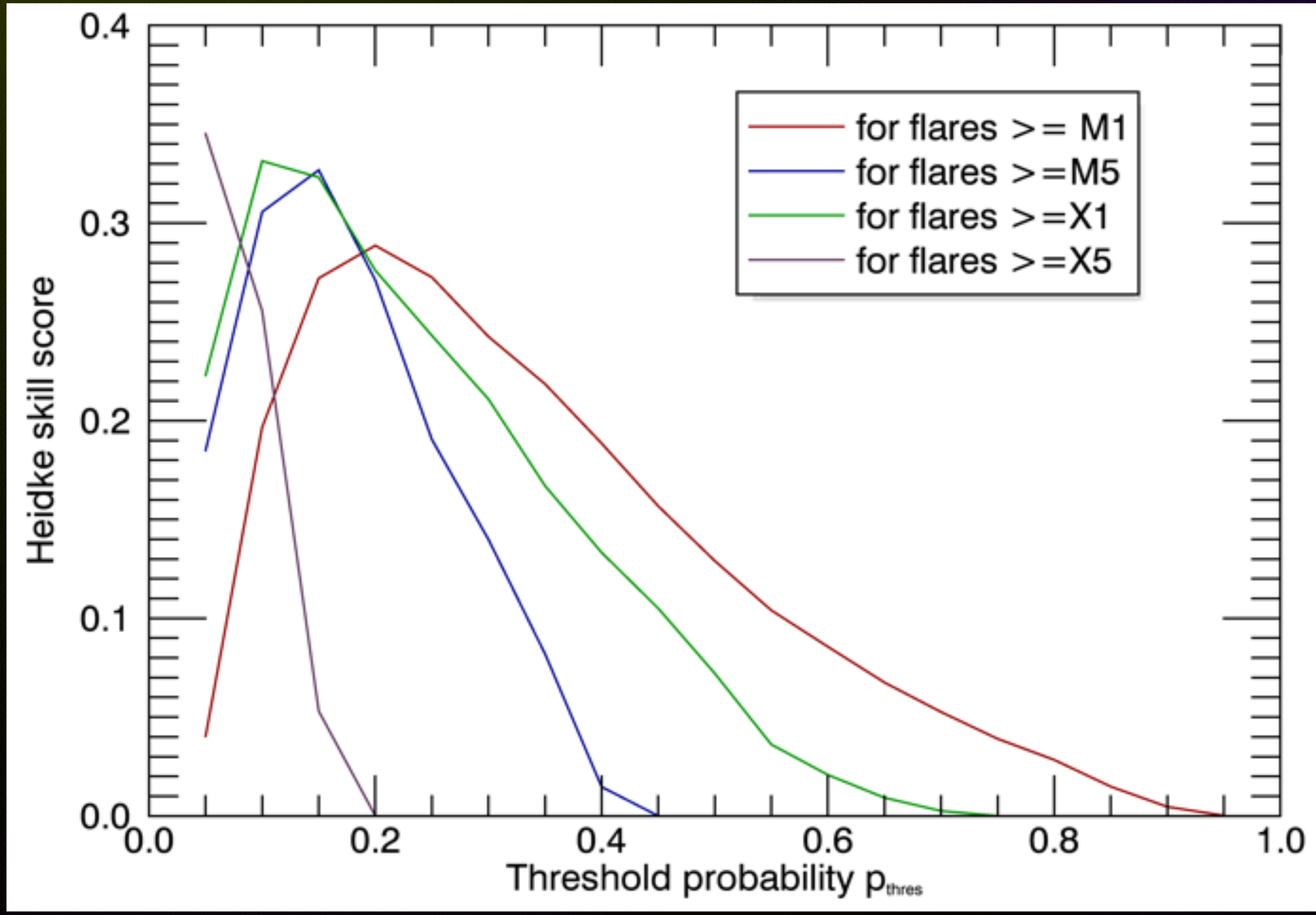
One needs a threshold probability p_{thres} ($p \geq p_{thres}$ —> YES ; $p < p_{thres}$ —> NO)



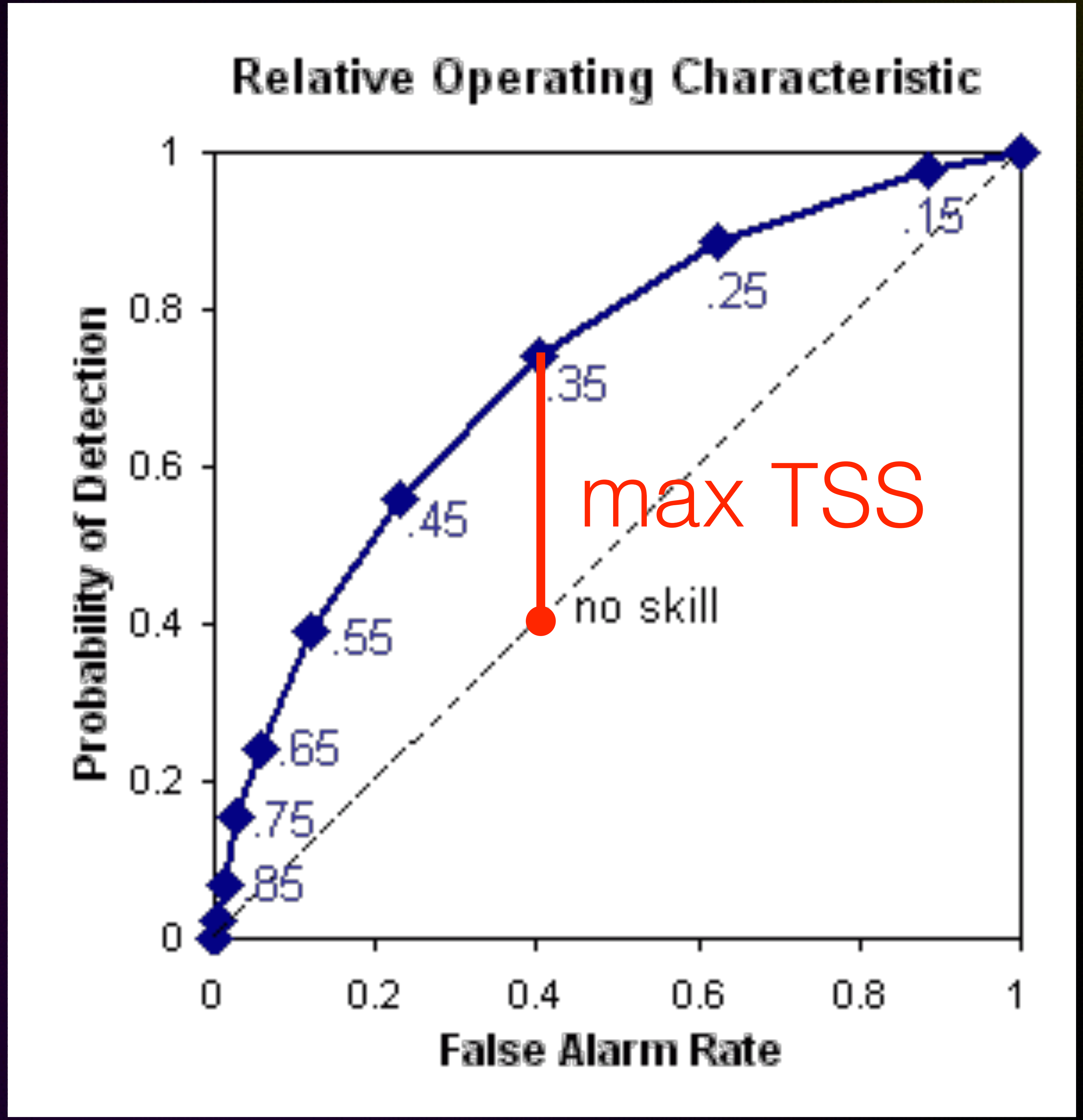
Example HSS vs. p_{thres} (A-EFFort)

CATEGORICAL VALIDATION IN A PROBABILISTIC FORECASTING?

One needs a threshold probability p_{thres} ($p \geq p_{thres}$ —> YES ; $p < p_{thres}$ —> NO)



Example HSS vs. p_{thres} (A-EFFort)

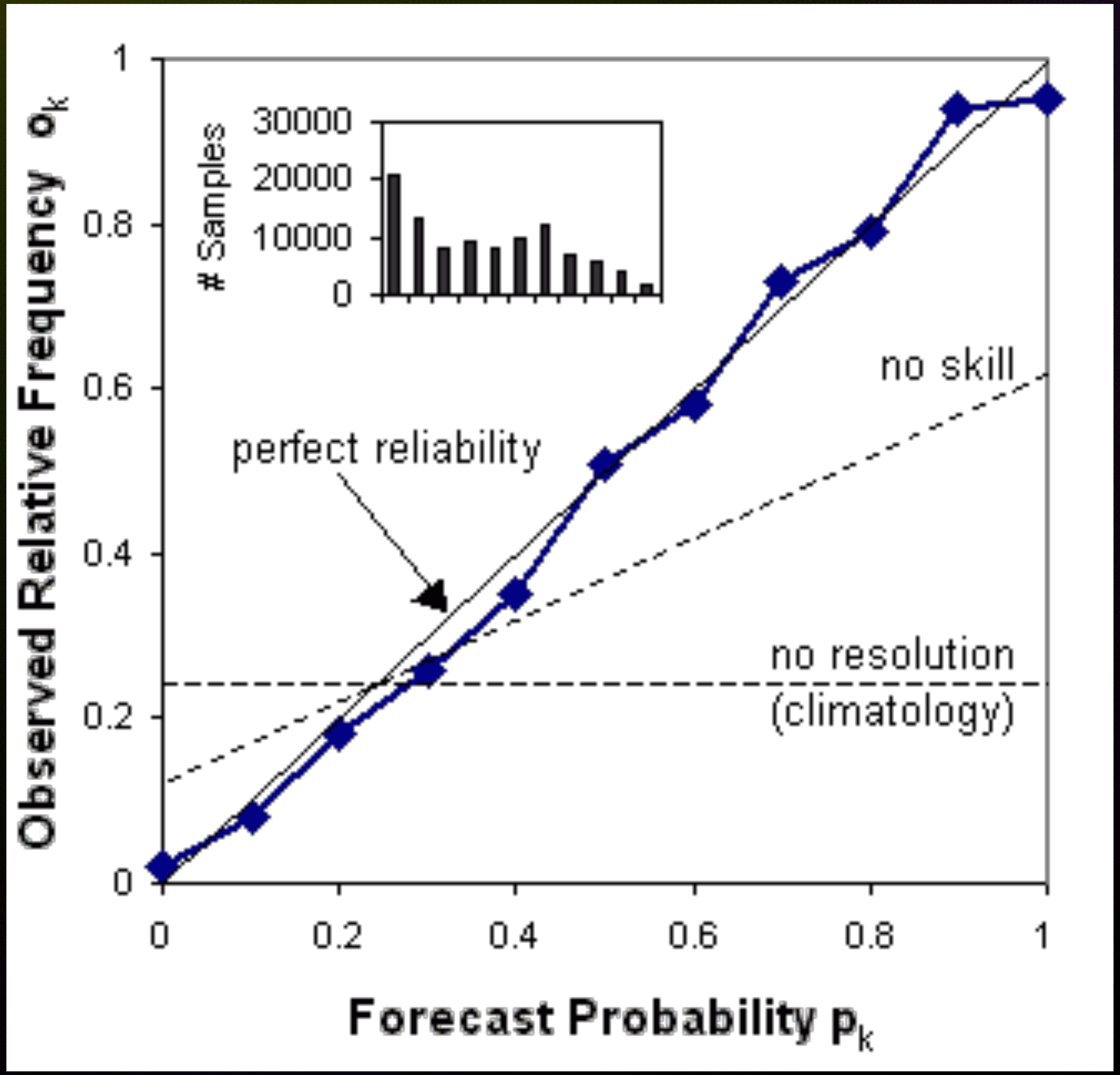


Example ROC curve (TSS vs. p_{thres})

Source: WMO Forecast Verification Research

PROBABILISTIC VALIDATION

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

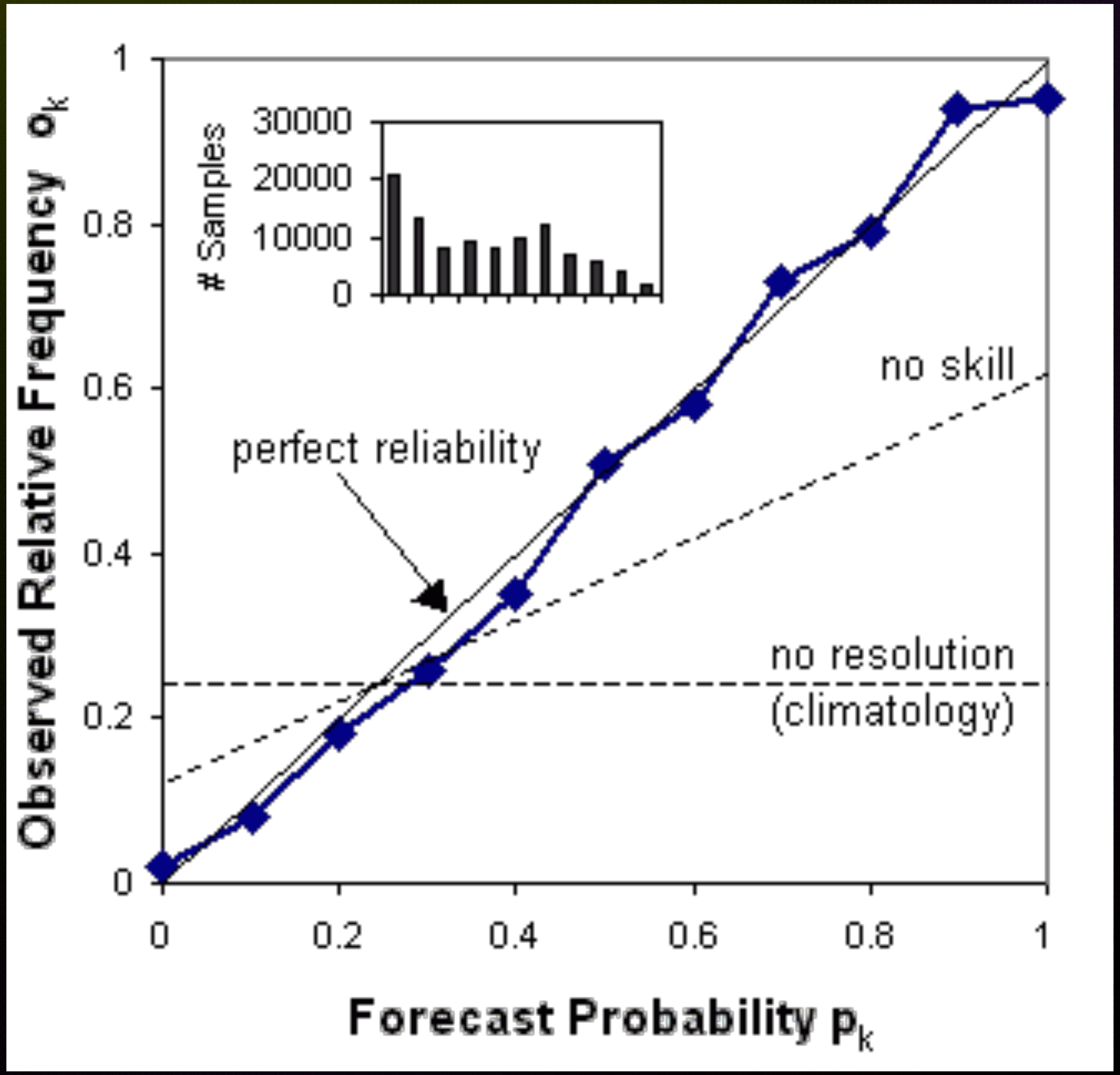


Reliabillity diagram

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

PROBABILISTIC VALIDATION

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



Reliabillity diagram

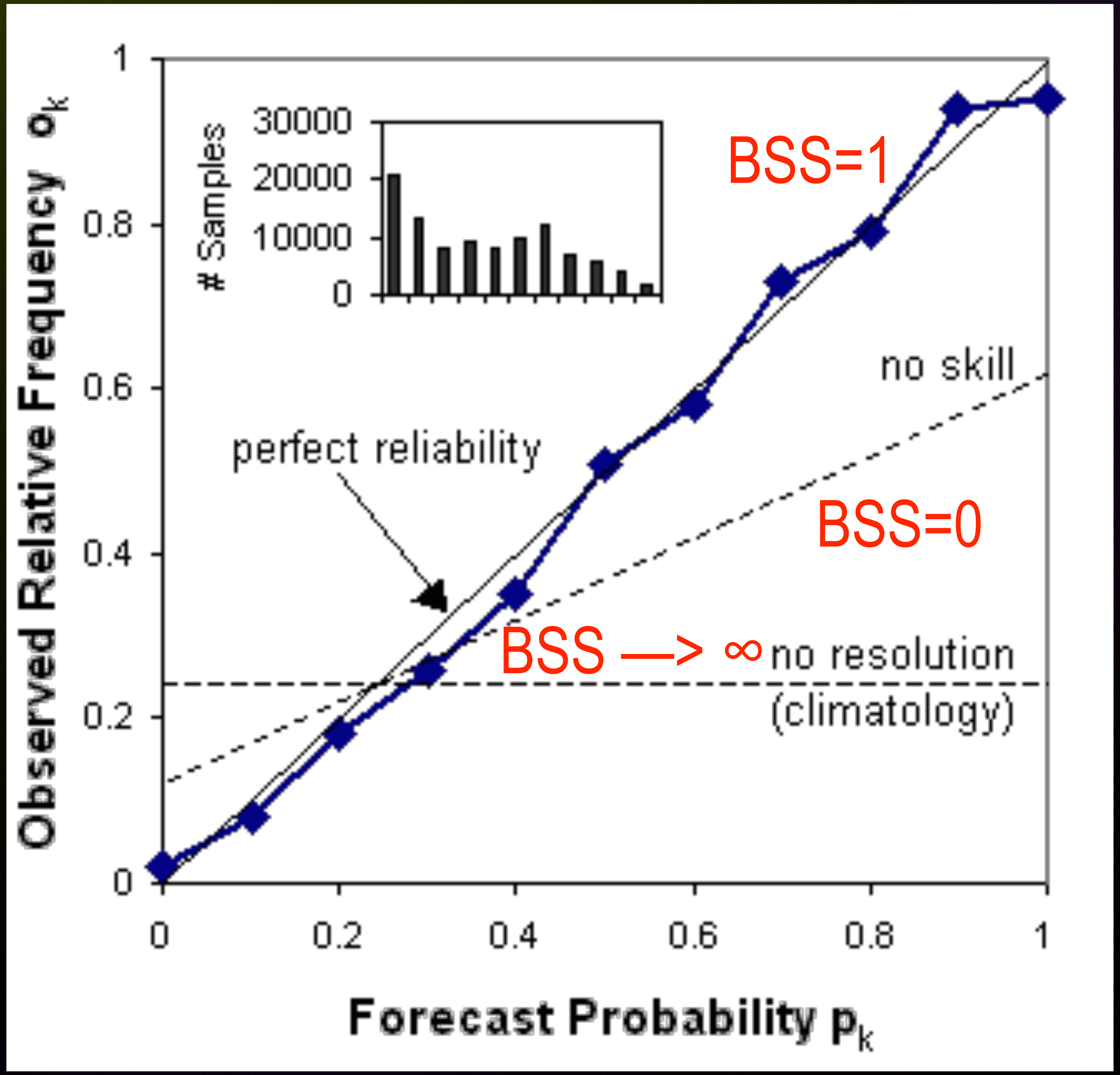
- Correlate forecast probability with observed frequency
- Compare your skill against climatology (mean flaring rate within forecast window)
- Generalized skill score:

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

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

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:

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

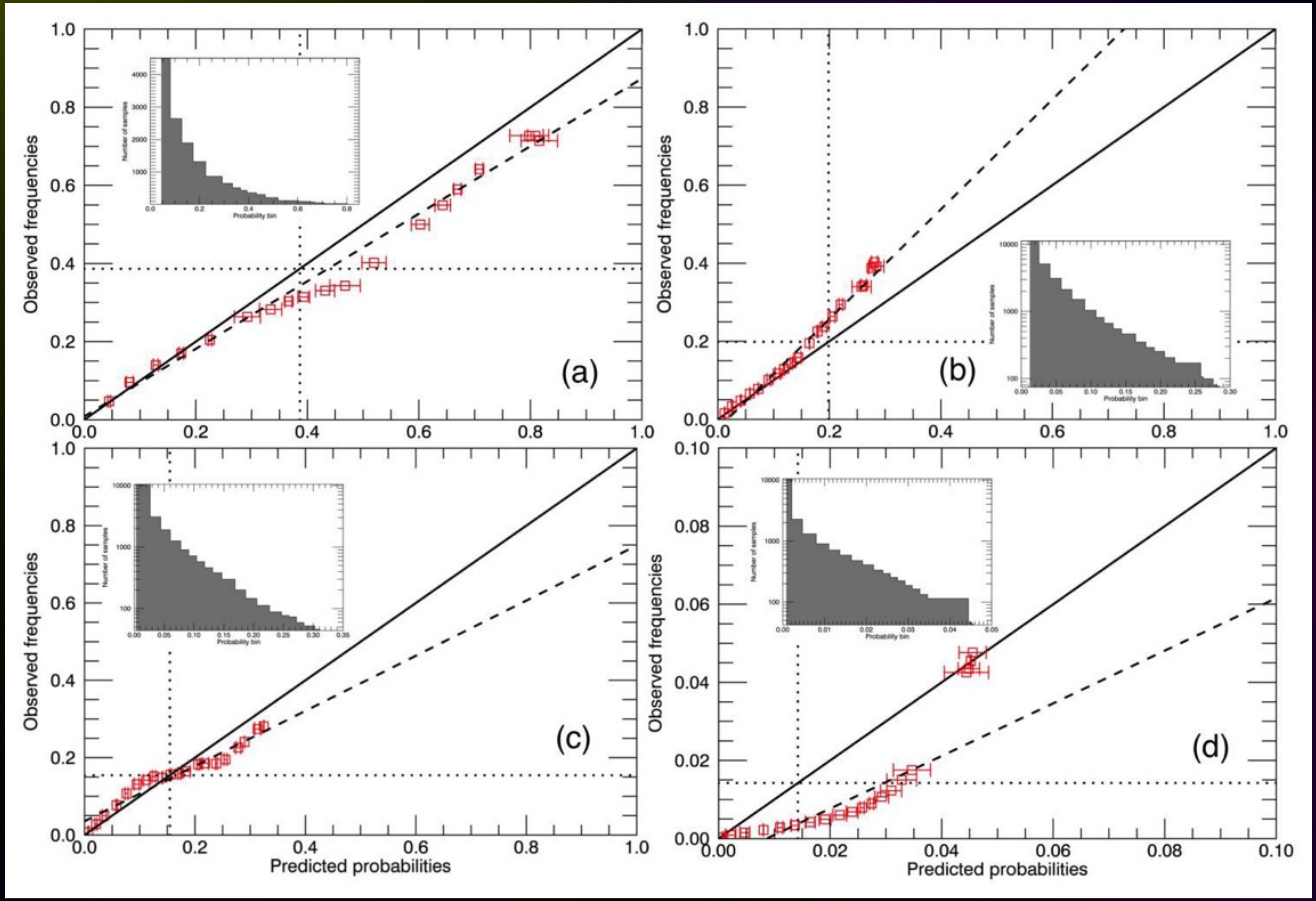
$$MSE = \; <(o-p)^2>$$

- Brier skill score (reference: climatology):

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

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

PROBABILISTIC VALIDATION



Example probabilistic validation (A-EFFort):

(a) M1 and above: BSS = 0.88

(b) M5 and above: BSS = 0.78

(c) X1 and above: BSS = 0.80

(d) X5 and above: BSS = 0.38

TAILORING FLARE PREDICTION METHODS TO THE CUSTOMER NEEDS

- Univariate forecasting is what all (but U. Bradford's ASAP) automated operational methods use

TAILORING FLARE PREDICTION METHODS TO THE CUSTOMER NEEDS

- Univariate forecasting is what all (but U. Bradford's ASAP) automated operational methods use
- Multivariate forecasting can also be used in the form of :
  - — Multi-variable predictors:

$$predictor = \omega_1 \; predictor_1 + \omega_2 \; predictor_2 + ... + \; \omega_n \; predictor_n$$

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

  - — Ensemble forecasting:

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

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

TAILORING FLARE PREDICTION METHODS TO THE CUSTOMER NEEDS

- Univariate forecasting is what all (but U. Bradford's ASAP) automated operational methods use
- Multivariate forecasting can also be used in the form of :

— Multi-variable predictors:

$$predictor = \omega_1 \; predictor_1 + \omega_2 \; predictor_2 + ... + \; \omega_n \; predictor_n$$

$\omega_1, \omega_2, ..., \omega_n$ unrestricted

— Ensemble forecasting:

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

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

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

TAILORING FLARE PREDICTION METHODS TO THE CUSTOMER NEEDS

- Univariate forecasting is what all (but U. Bradford's ASAP) automated operational methods use
- Multivariate forecasting can also be used in the form of :

— Multi-variable 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:

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

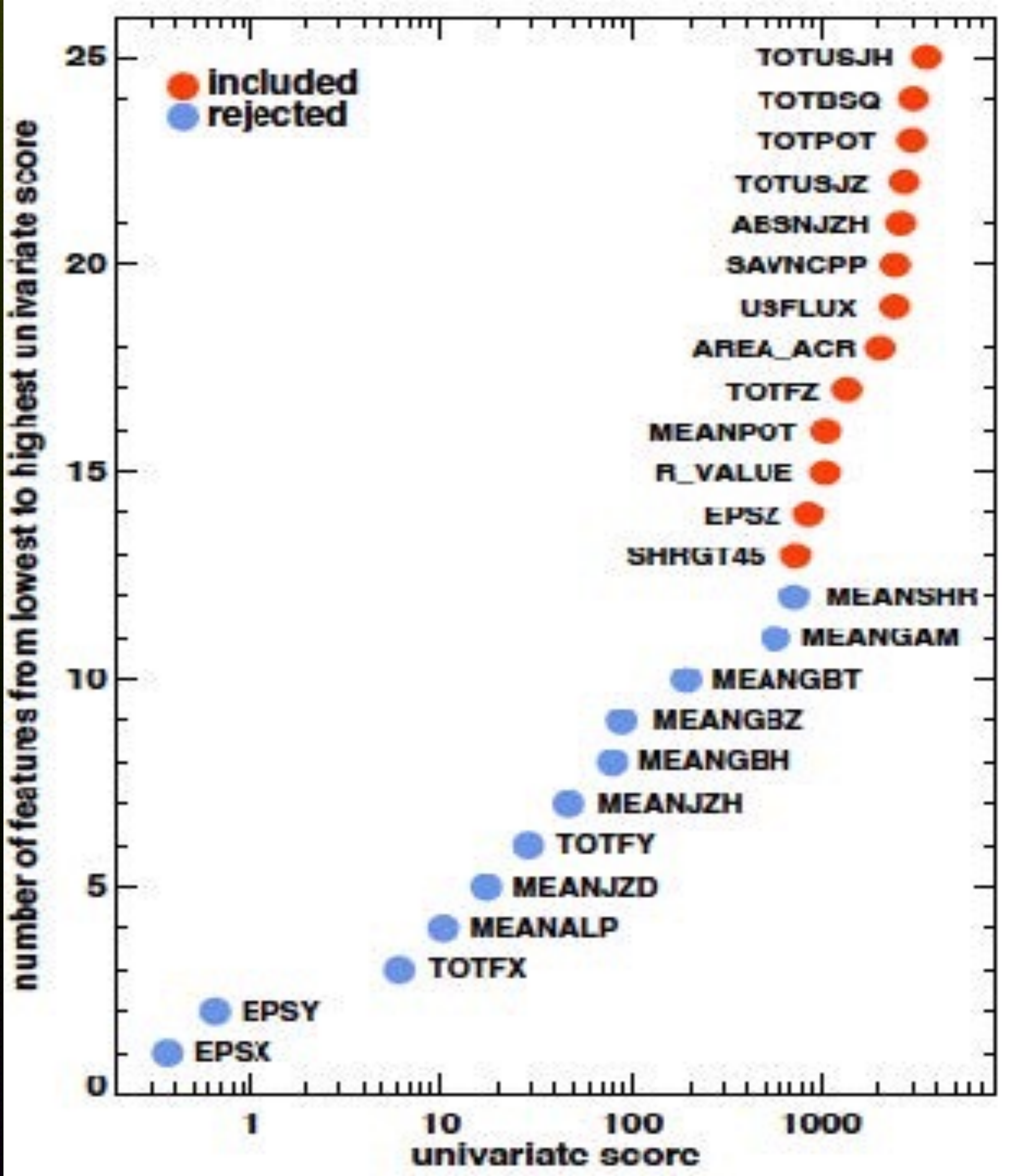
$$\sum_{i=1}^{n} \omega_i = 1$$

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

However: optimization means different things to different customers!

SOME RECENT, PRELIMINARY EXAMPLES

- Multivariate forecasting

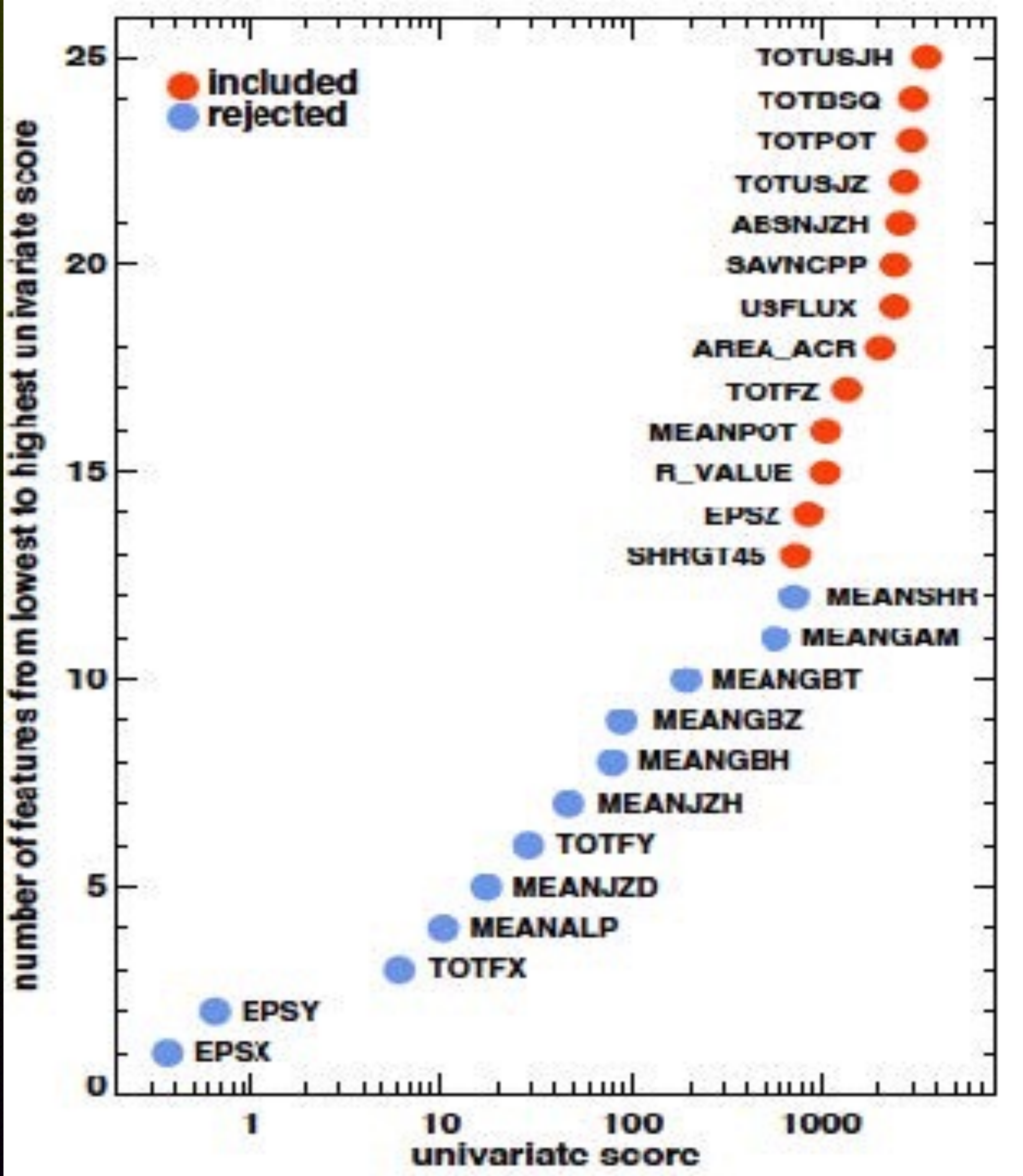


Bobra & Couvidat (2015)

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

SOME RECENT, PRELIMINARY EXAMPLES

- Multivariate forecasting

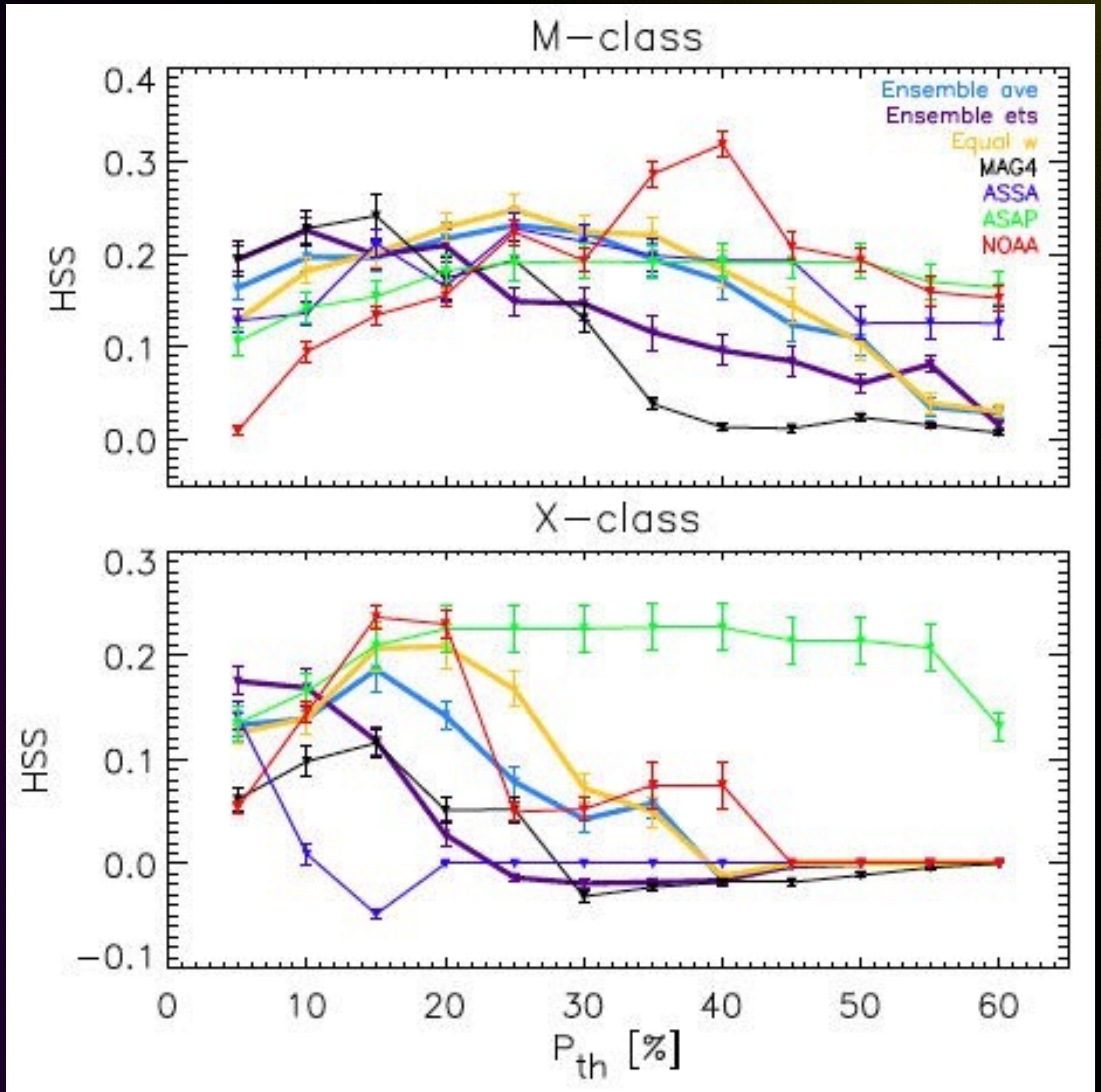


Bobra & Couvidat (2015)

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

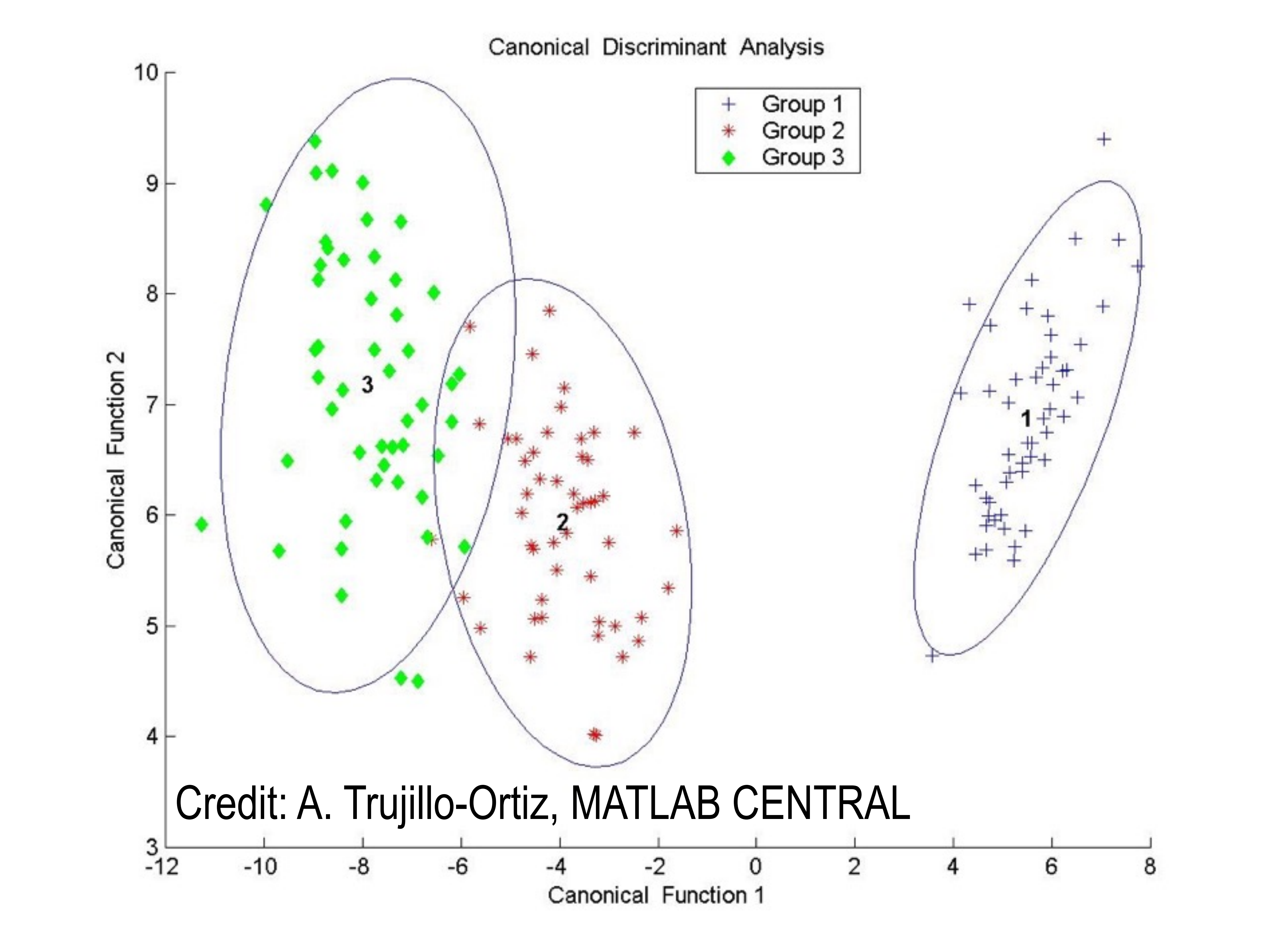
- Homogenizing the results of multiple flare prediction methods, using them with equal or non-equal weights for an ensemble forecasting

- Ensemble forecasting

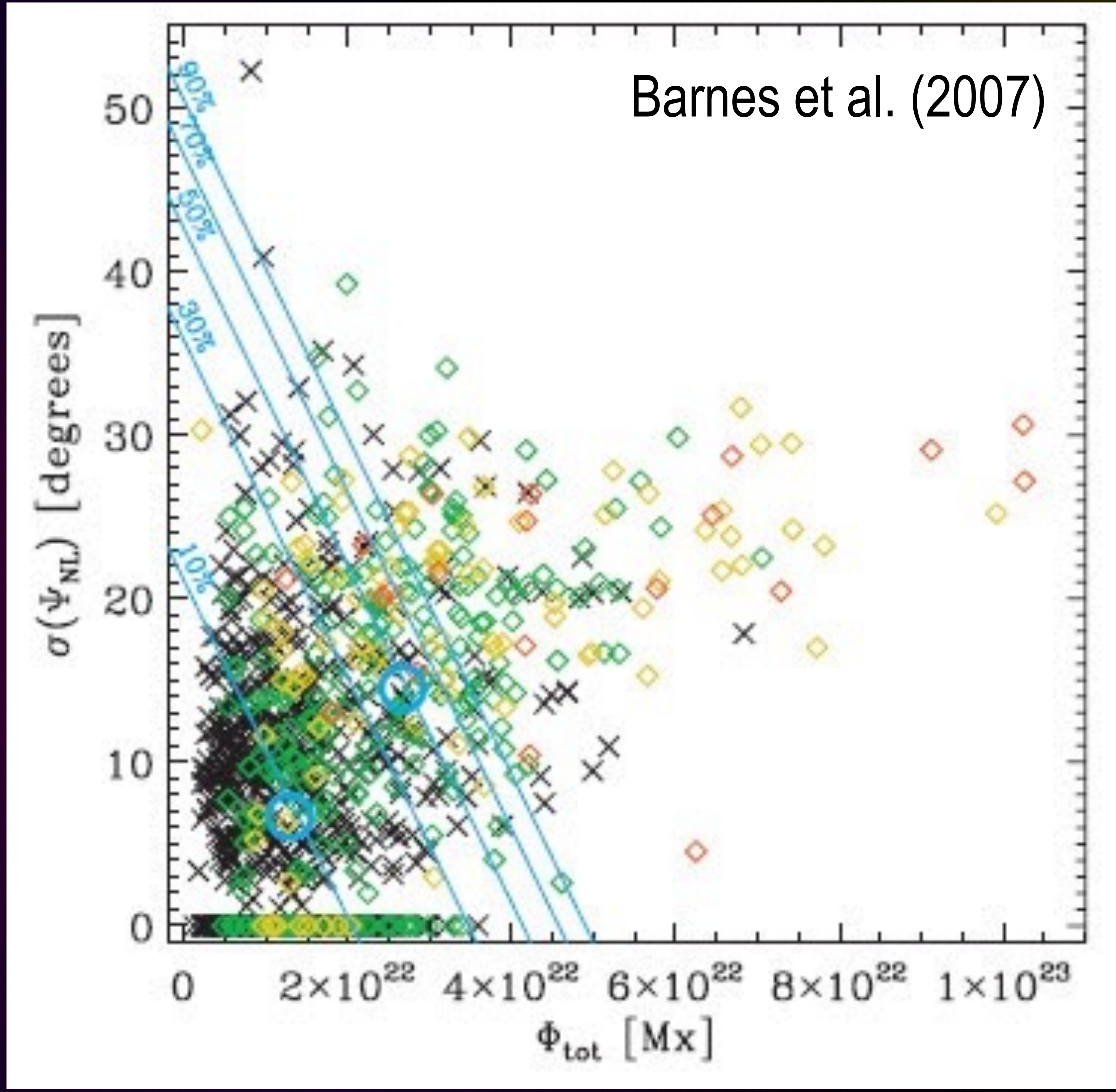


Guerra et al. (2015)

TAMING THE (LARGE) PARAMETER SPACE: DISCRIMINANT ANALYSIS and/or PCA



Example of two-function, three-group canonincal DA



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

FLARECAST: SYNTHESIS IN ACTION

FLARECAST is a European research project aiming to develop an automated solar-flare forecasting system with unmatched accuracy compared to existing facilities.

FLARECAST, using diverse European expertise, will:

- ★ use or reproduce already available predictors
- ★ classify predictors with respect to predictive ability
- ★ validate results in different ways

Project Partners:

- Academy of Athens, *Greece*
- Trinity College Dublin, *Ireland*
- Universita Degli Studi Di Genova, *Italy*
- Consiglio Nazionale Delle Ricerche, *Italy*
- Centre National de la Recherche Scientifique, *France*
- Université Paris-Sud, *France*
- Fachhochschule Nordwestschweiz, *Switzerland*
- Met Office, *United Kingdom*

FLARECAST: SYNTHESIS IN ACTION

FLARECAST is a European research project aiming to develop an automated solar-flare forecasting system with unmatched accuracy compared to existing facilities.

FLARECAST, using diverse European expertise, will:

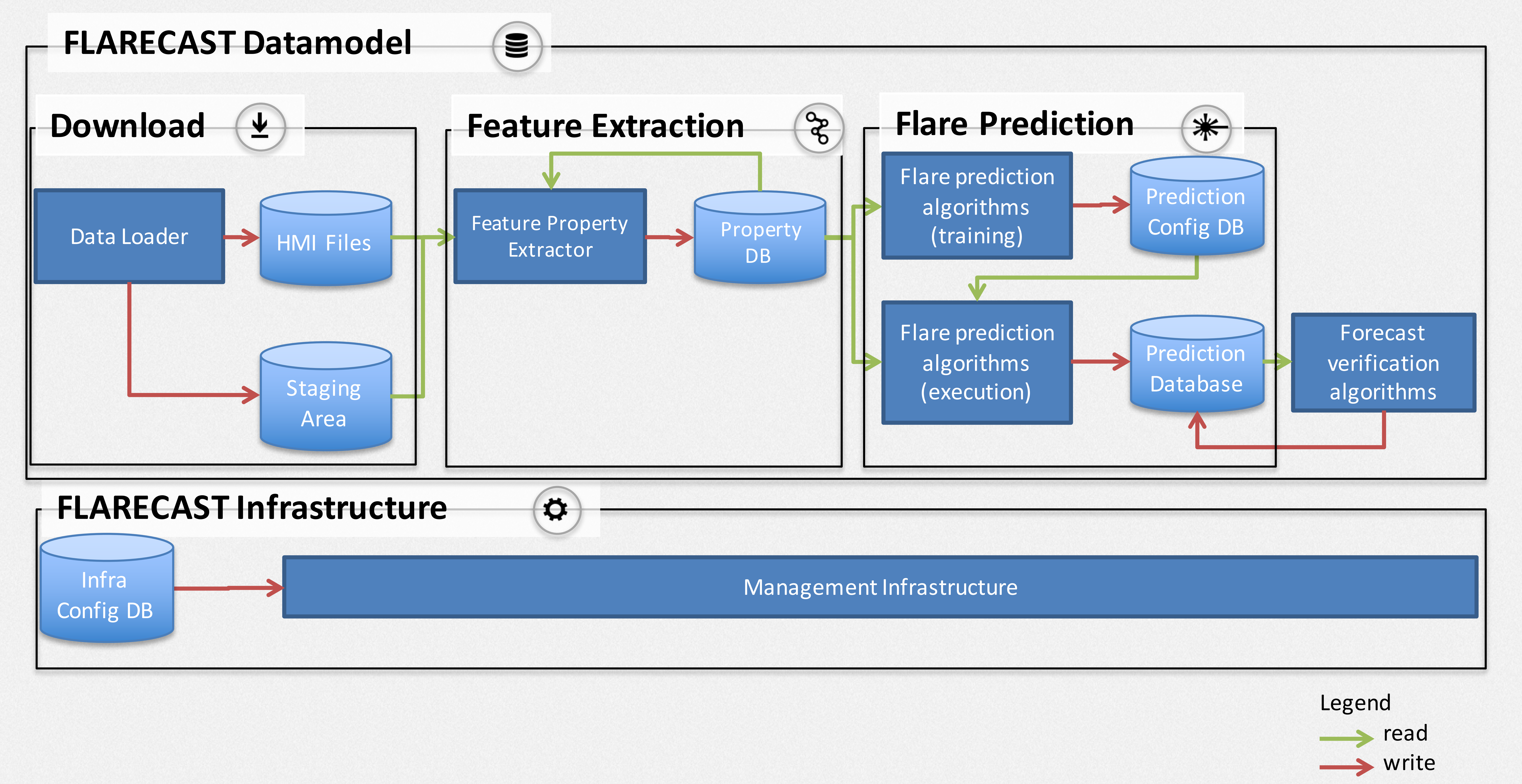
- ★ use or reproduce already available predictors
- ★ classify predictors with respect to predictive ability
- ★ validate results in different ways

Project Partners:

- Academy of Athens, *Greece*
- Trinity College Dublin, *Ireland*
- Universita Degli Studi Di Genova, *Italy*
- Consiglio Nazionale Delle Ricerche, *Italy*
- Centre National de la Recherche Scientifique, *France*
- Université Paris-Sud, *France*
- Fachhochschule Nordwestschweiz, *Switzerland*
- Met Office, *United Kingdom*

The FLARECAST forecasting system will be openly accessible, featuring open-source software that will allow end users to perform their own tests. In this way FLARECAST will aim to both revamp solar flare prediction and contribute to a better understanding of the drivers of flare activity.

FLARECAST DATAMODEL & INFRASTRUCTURE : COLLECTIVE EXPERTISE



CONCLUSIONS

- Solar flare prediction: consensus that it should be an asset of our space-weather forecasting toolbox
- Validation: a vital task, done in multiple ways as per the customer's needs
- We should understand how to enhance various validation metrics against others
- Our ultimate task should be to bring a purely probabilistic prediction (due to stochasticity of the process) as close as possible to a categorical (YES / NO) one
- The solution <u>to this will not be unique</u> - <u>however, the used methods should be</u>
- Customized forecasts should rely on multivatiate or ensemble modeling
- <u>Standard datasets</u> could also be created for the validation of all methods

CONCLUSIONS

- Solar flare prediction: consensus that it should be an asset of our space-weather forecasting toolbox
- Validation: a vital task, done in multiple ways as per the customer's needs
- We should understand how to enhance various validation metrics against others
- Our ultimate task should be to bring a purely probabilistic prediction (due to stochasticity of the process) as close as possible to a categorical (YES / NO) one
- The solution <u>to this will not be unique</u> - <u>however, the used methods should be</u>
- Customized forecasts should rely on multivatiate or ensemble modeling
- <u>Standard datasets</u> could also be created for the validation of all methods

All these tasks are being investigated by the FLARECAST Consortium

BACKUP SLIDES