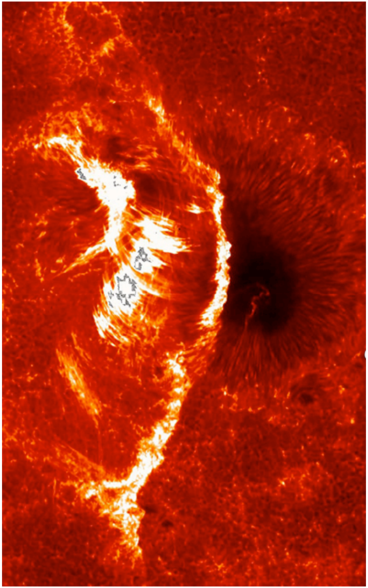
A baseline flare prediction using only event statistics

M.S. Wheatland

School of Physics
University of Sydney

Forecasting the Operational All-Clear
22-24 April 2009

The University of Sydney



Overview

Method

- *Persistence*
- *Flare statistics*
- *Event statistics method*

Workshop applications

- *Whole-Sun, active-region prediction of GOES flares*

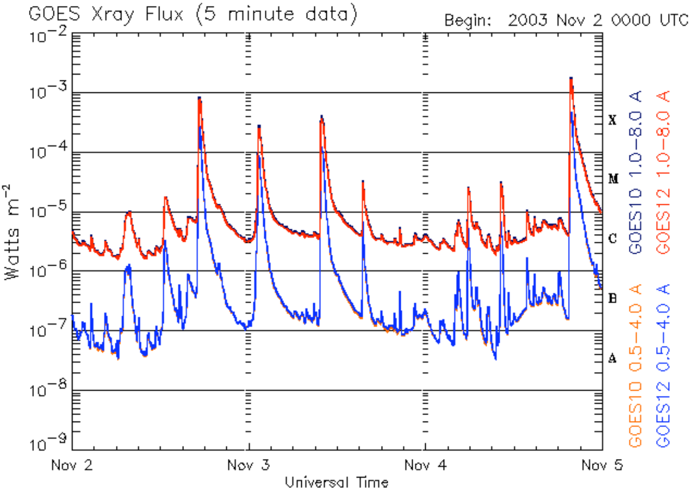
Validation

- *Accuracy of probabilistic forecasts*

Summary and comments

Persistence

- Past flaring history is an indicator of future flaring



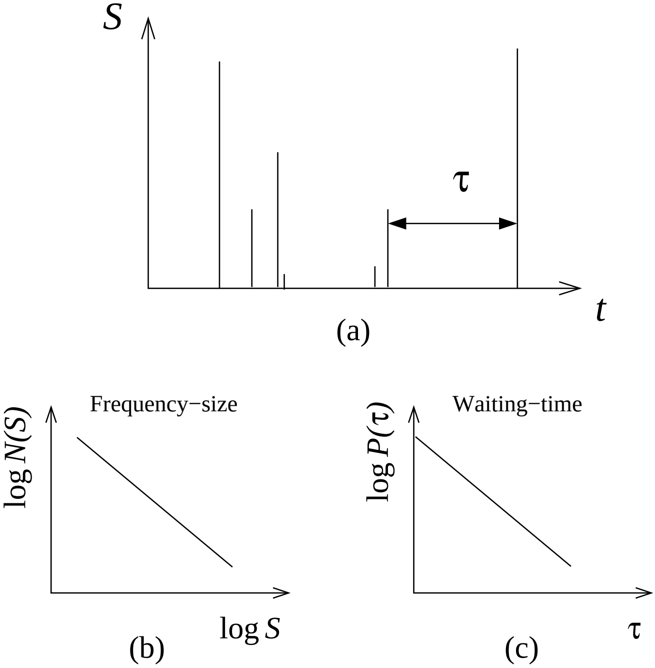
Flare statistics

- Flares obey a power-law frequency-size distribution (Drake 1971)

$$N(S) = \lambda_1(\gamma - 1)S_1^{\gamma-1}S^{-\gamma} \tag{1}$$

  - size S: energy, or peak flux in X-ray,...
  - $N(S)$ is number of flares per unit time, per unit S
  - $\gamma = \frac{3}{2}$ to $\gamma = 2$ (depends on specific choice of S)
  - $\lambda_1 = \lambda_1(t)$ is total rate above size S_1
- Occurrence in time may be modelled as a Poisson process (e.g. Wheatland 2001)
  - if λ does not vary, distribution of waiting times τ:

$$P(\tau) = \lambda \exp(-\lambda\tau) \tag{2}$$


S
τ
t
(a)
Frequency−size
log N(S)
log S
(b)
Waiting−time
log P(τ)
τ
(c)

Event statistics method (Wheatland 2004, ApJ 609, 1134)

- S_1=size of "small" event, S_2=size of "big" event
- λ_1=observed rate above S_1; PL size distribution ⇒

$$\lambda_2 = \lambda_1 \left(S_1 / S_2 \right)^{\gamma - 1} \tag{3}$$

  - even if no big events have been observed
- Probability of at least one big event in time T_P is

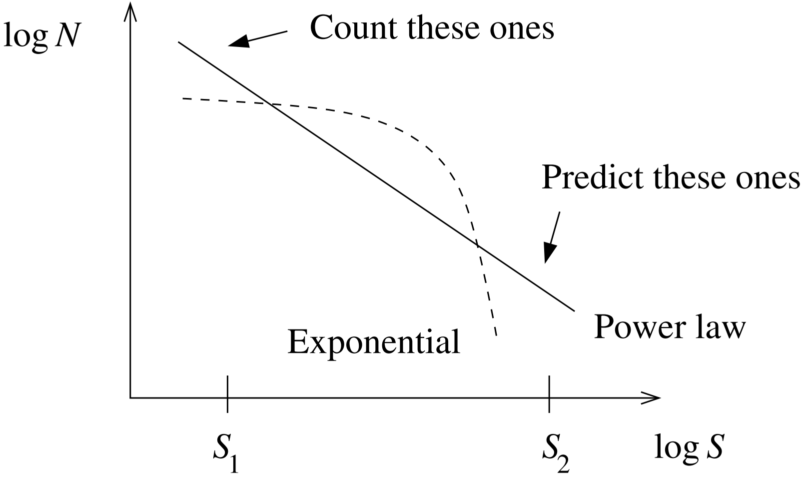
$$\begin{aligned} \epsilon &= 1 - \exp(-\lambda_2 T_P) \\ &= 1 - \exp\left[-\lambda_1 \left(S_1 / S_2 \right)^{\gamma - 1} T_P \right] \end{aligned} \tag{4}$$

  - assuming Poisson waiting times
- If M events are involved in inferring λ_1 then

$$\sigma_\epsilon / \epsilon \approx 1/\sqrt{M} \tag{5}$$

  - accurate if many small events observed

- Method well-suited to a power-law size distribution



Bayesian version

- Data D: events $s_1, s_2, ..., s_M$ at times $t_1 < t_2 < ... < t_M$
- Infer $P_\gamma(\gamma|D)$ and $P_1(\lambda_1|D)$
- Calculate

$$\begin{aligned} P_2(\lambda_2|D) &= \int_1^\infty d\gamma \int_0^\infty d\lambda_1 P_1(\lambda_1|D) P_\gamma(\gamma|D) \\ &\times \delta\left[\lambda_2 - \lambda_1 (S_1/S_2)^{\gamma-1}\right] \end{aligned} \tag{6}$$

and then

$$P_\epsilon(\epsilon|D) = P_2\left[\lambda_2(\epsilon)|D\right] \left|\frac{d\lambda_2}{d\epsilon}\right| \tag{7}$$

where

$$\lambda_2(\epsilon) = -\ln(1-\epsilon)/T_P \tag{8}$$

- Problem: infer power-law index and rate of small events

Inferring the power-law index γ

- Events power-law distributed, so (Bai 1993)

$$P(D|\gamma) \propto \prod_{i=1}^{M} (\gamma - 1)(s_i/S_1)^{-\gamma} \tag{9}$$

  - uniform prior over a range $\gamma_1 < \gamma < \gamma_2$ used
  - for $M \gg 1$, peaked around

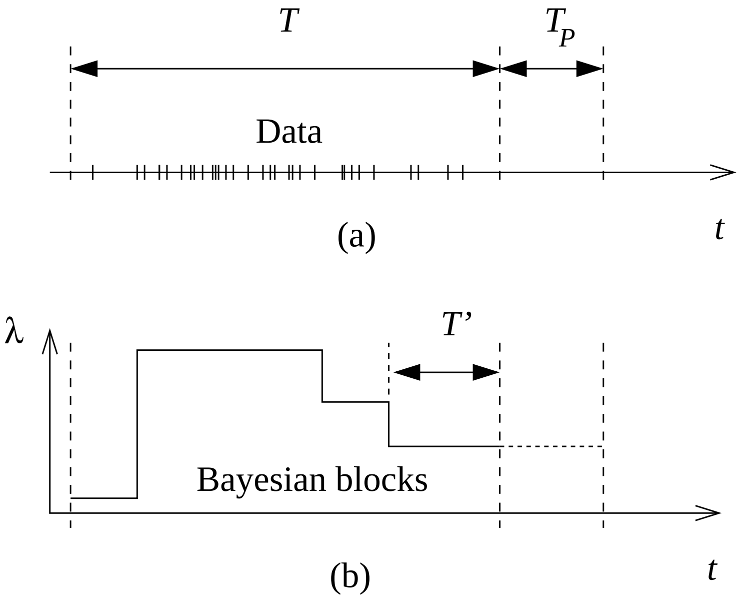
$$\gamma_{\mathrm{ML}} = \frac{M}{\ln \pi} + 1 \quad \text{where} \quad \pi = \prod_{i=1}^{M} \frac{s_i}{S_1} \tag{10}$$

  - use $P_\gamma(\gamma|D) = \delta(\gamma - \gamma_{\mathrm{ML}})$

Inferring the rate of small events λ_1

- Complicated by time variation
- Bayesian blocks used (Scargle 1998)
  - iterative comparison of one- versus two-rate Poisson models
  - decomposition into piecewise-constant Poisson process
  - data D' in last piece (block): M' events in time T'
  - most recent interval with constant rate

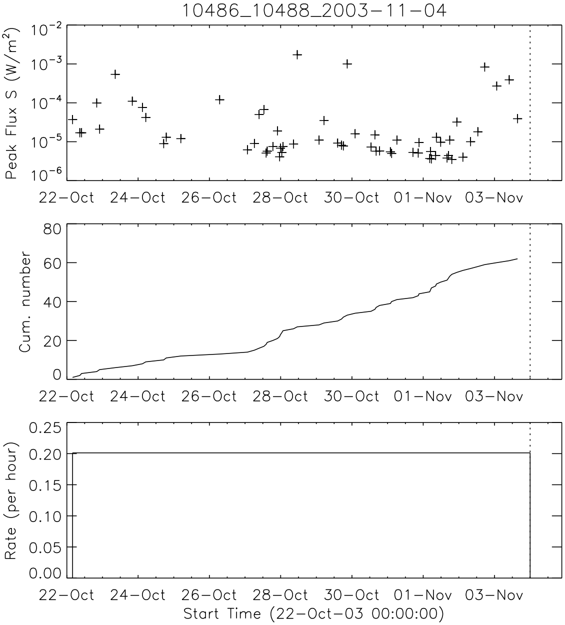
$$P_1(D'|\lambda_1) \propto \lambda_1^{M'} e^{-\lambda_1 T'} \tag{11}$$


T
T_P
Data
(a)
t
λ
T'
Bayesian blocks
(b)
t

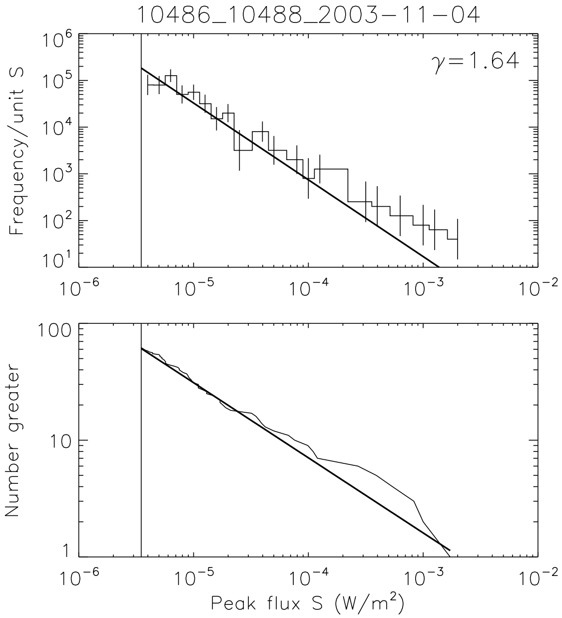
Whole-Sun, active-region prediction of GOES flares

- Method applied to to whole-Sun prediction of GOES events (Wheatland 2005, Space Weather 3, S07003)
  - this workshop: predictions for times of MDI observations
  - magnetogram information not used
- Applied here also to specified active regions
- Illustration for 10486 and 10488
  - largest soft X-ray flare: 4 Nov 2003 20:06UT in 10486
  - 10486, 10488: flare-productive active regions on disk
  - D: previous events from these active regions (62 events)
  - $S_1 = 3.6 \times 10^{-6}\,\mathrm{W\,m^{-2}}$ (smallest event)
  - inference of γ: ML approximation not made
  - predictions for C1.0 (24 hr), and M1.0, M5.0 (12 hr) at 00:00UT

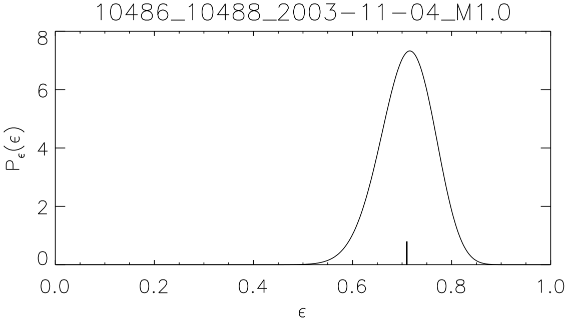
- Bayesian blocks procedure



- Frequency-size distribution of events



- Posteriors for M1.0, M5.0



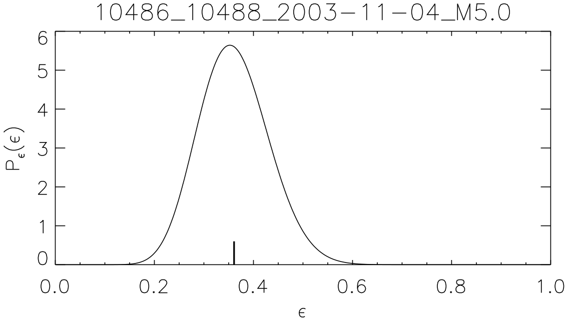


Table: Predictions for 10486, 10488 at 4 Nov 2003 00:00UT

Event	T_P	ε
C1	24 hrs	1.00 ± 0.00
M1	12 hrs	0.71 ± 0.05
M5	12 hrs	0.36 ± 0.07

- In fact five events $\geq$C1 within 24 hrs, one $\geq$ M1, zero $\geq$M5 within 12 hrs
- Workshop: applied to ARs listed in each MDI file
  - magnetogram information not used

Accuracy of probabilistic forecasts

- f = forecast, x = observation (0 or 1)
- Mean square error

$$\mathrm{MSE}(f,x) = \langle (f-x)^2 \rangle \tag{12}$$

- "Climatological skill score":

$$\mathrm{SS}(f,x) = 1 - \mathrm{MSE}(f,x)/\mathrm{MSE}(\langle x \rangle, x) \tag{13}$$

  - improvement over forecasting the average
- Reliability plots
  - observed frequencies versus forecast probabilities

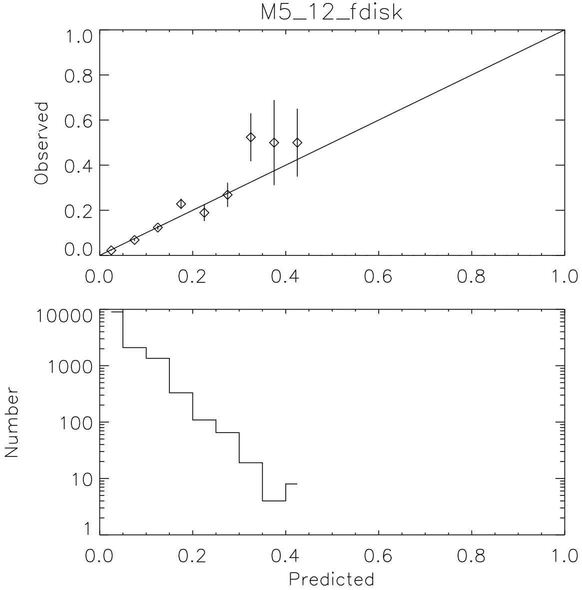
Table: Whole-Sun predictions for workshop

Event	T_P	$\langle f \rangle$	$\langle x \rangle$	$SS(f, x)$
C1	24 hrs	0.94	0.90	0.13
M1	12 hrs	0.28	0.23	0.16
M5	12 hrs	0.049	0.049	0.064

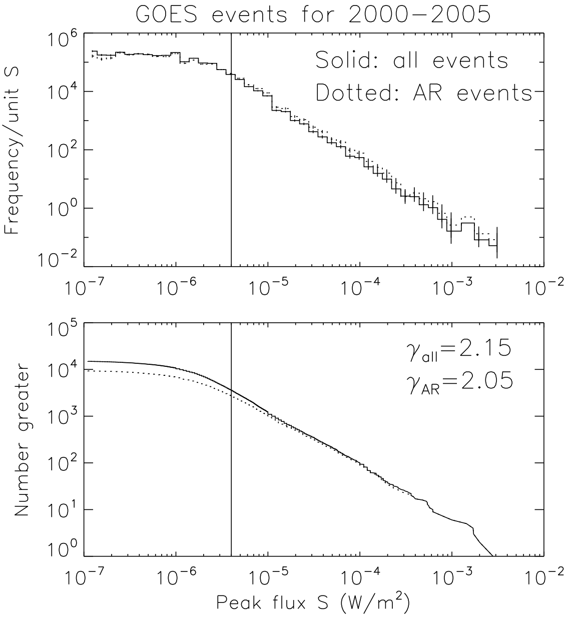
Table: Active-region predictions for workshop ($\geq$ 10 events)

Event	T_P	$\langle f \rangle$	$\langle x \rangle$	$SS(f, x)$
C1	24 hrs	0.89	0.57	-0.27
M1	12 hrs	0.23	0.15	0.088
M5	12 hrs	0.0835	0.046	-0.022

- Reliability plot for whole-Sun, M5–12 hr prediction



- Problem of missed small events



Summary and comments

- Bayesian method of solar flare prediction
  - uses only event statistics
  - exploits power-law size distribution
  - quantifies "persistence"
  - applications to whole-Sun, AR prediction of GOES events
  - provides "baseline" predictions: does not use the MDI data
  - very simple to apply
- Method could be improved by:
  - more complete observations of small events
  - use of data less affected by event selection problems
  - inclusion of additional information (e.g. Wheatland 2006)