NASA Ames Research Center

Adventures in Modern Time Series Analysis

From the Sun to the Crab Nebula and Beyond

Jeffrey.D.Scargle@nasa.gov

Brad Jackson, Jay Norris, Jim Chiang, Tom Loredo, Joe Bredekamp

Alanna Connors

Applied Information Systems Research Program (NASA)
Keck Institute for Space Studies (Caltech)
Institute for Pure and Applied Mathematics (UCLA)
Center for Applied Mathematics and Computer Science (SJSU)
Statistical and Mathematical Sciences Institute (Duke)
Banff International Research Station

Outline

Time Series Analysis in the Age of Digital Astronomy

- Case Study: BATSE Gamma Ray Burst Data
- The Bayesian Blocks Algorithm
- Fermi Gamma Ray Space Telescope
- Activity in the Crab Nebula
- Edelson and Krolik DCF → ***Time Series Explorer***
- Active Galactic Nuclei with Kepler and Fermi
- Chromospheric Activity over 3+ Solar cycles
- *Multi-scale Structure of the Galaxy Distribution*

Astronomical Time Series Analysis Issues

- ✧ Uneven Sampling
- ✧ Data Gaps
- ✧ Observational Errors
- ✧ Variable Observational Errors - “heteroskedasticity”
- ✧ Exposure Variation
- ✧ Data Modes: Events, Bin Counts, Measurements, …
- ✧ Background
- ✧ Multivariate Time Series
- ✧ Data on the Circle

Time-Tags for Photons from BATSE Burst 0551

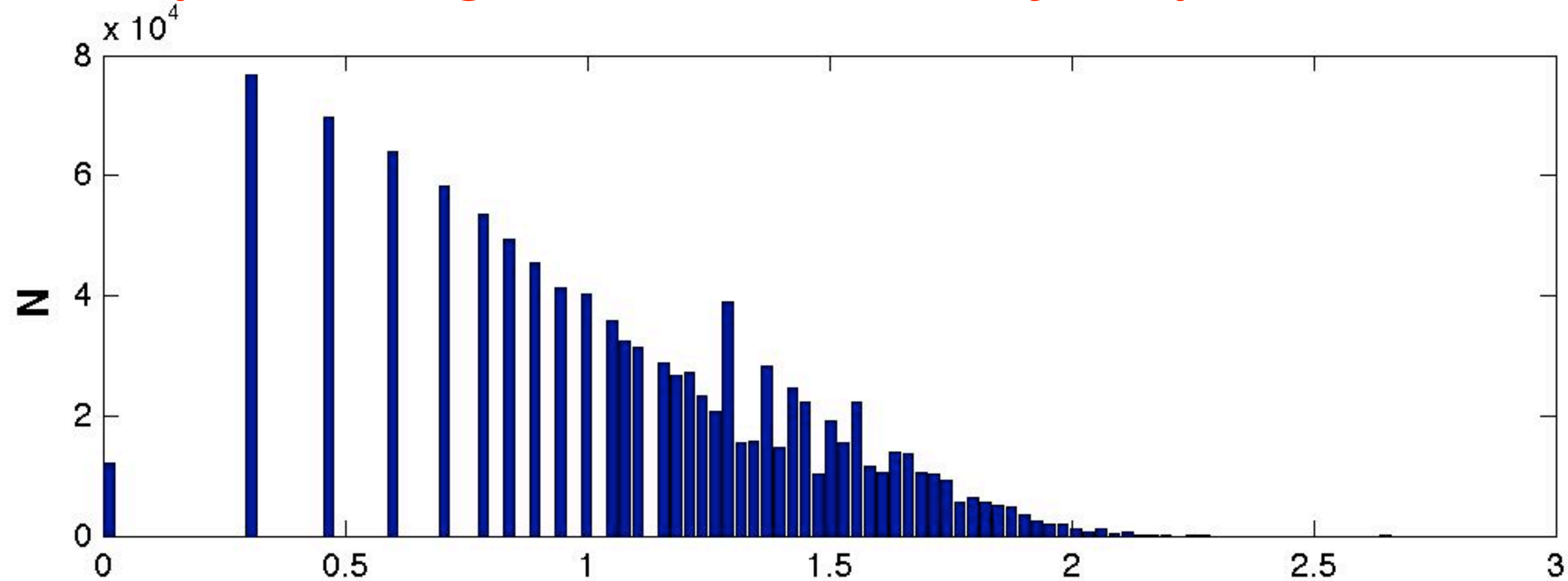
229	370	542	814	1025	1223
270	409	682	834	1033	1250
274	412	701	838	1074	1268
281	422	712	880	1103	1269
291	458	732	884	1127	1275
322	470	735	981	1156	1291

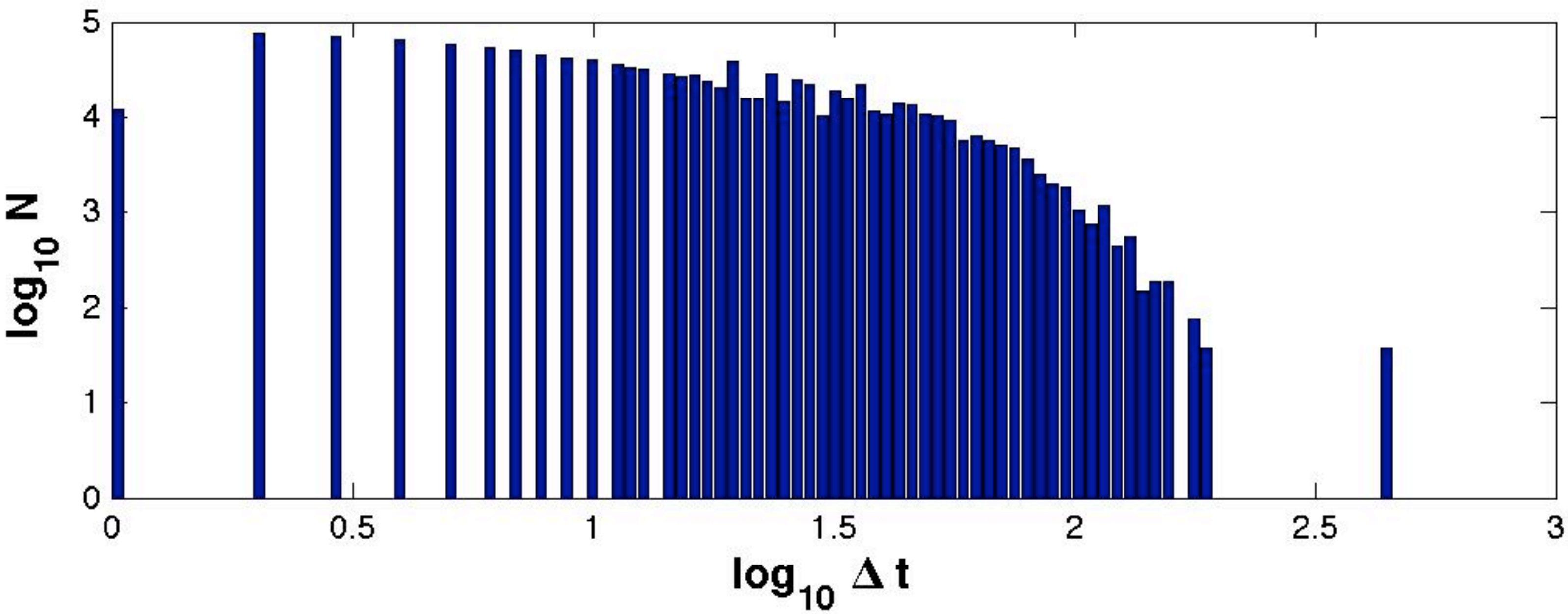
…

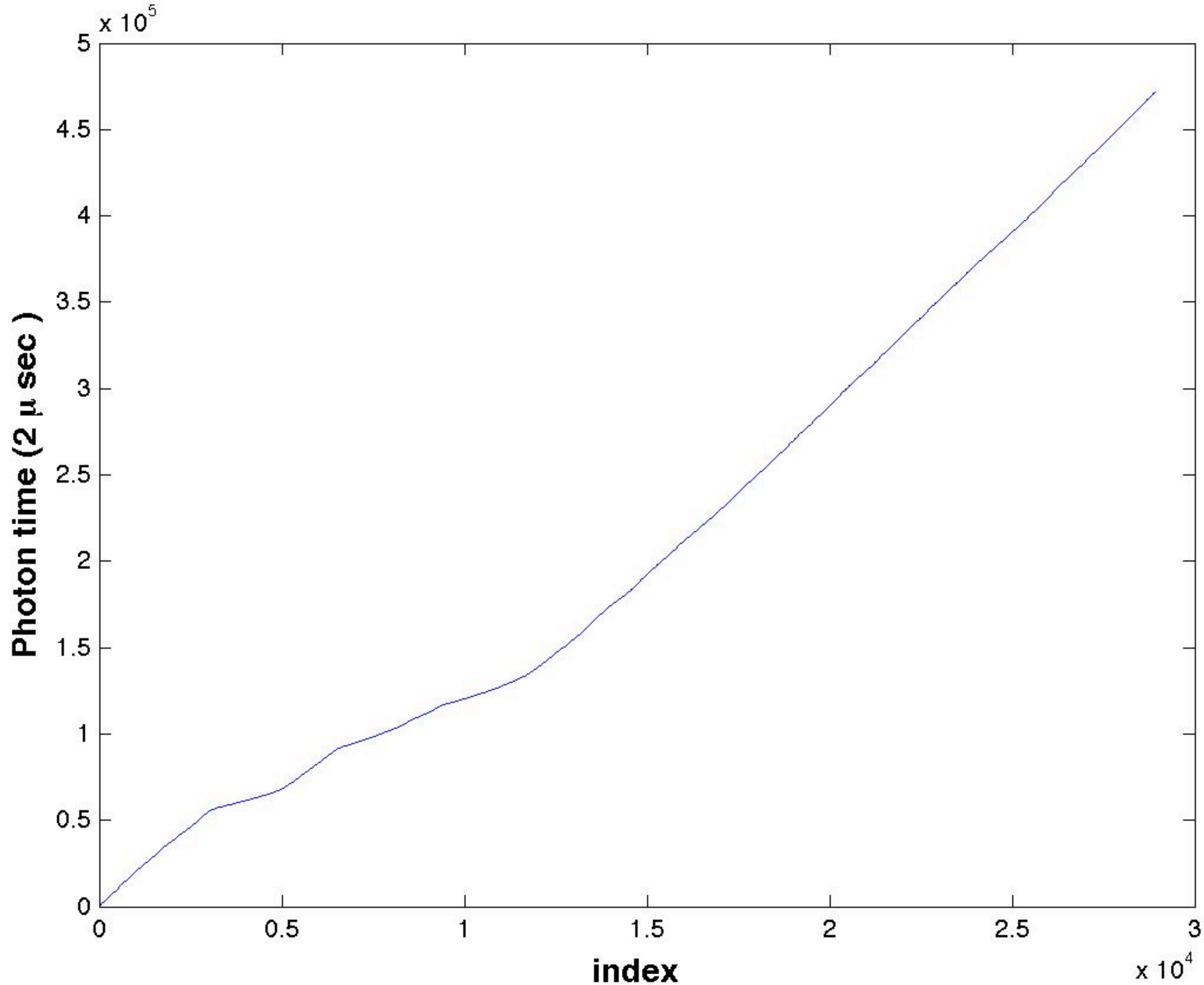
289454	289457	289486	289500	289582	289645
289652	289689	289691	289698	289738	289744
289758	289759	289776	289778	289779	289804

Unit: 2 microseconds   Origin: Trigger time

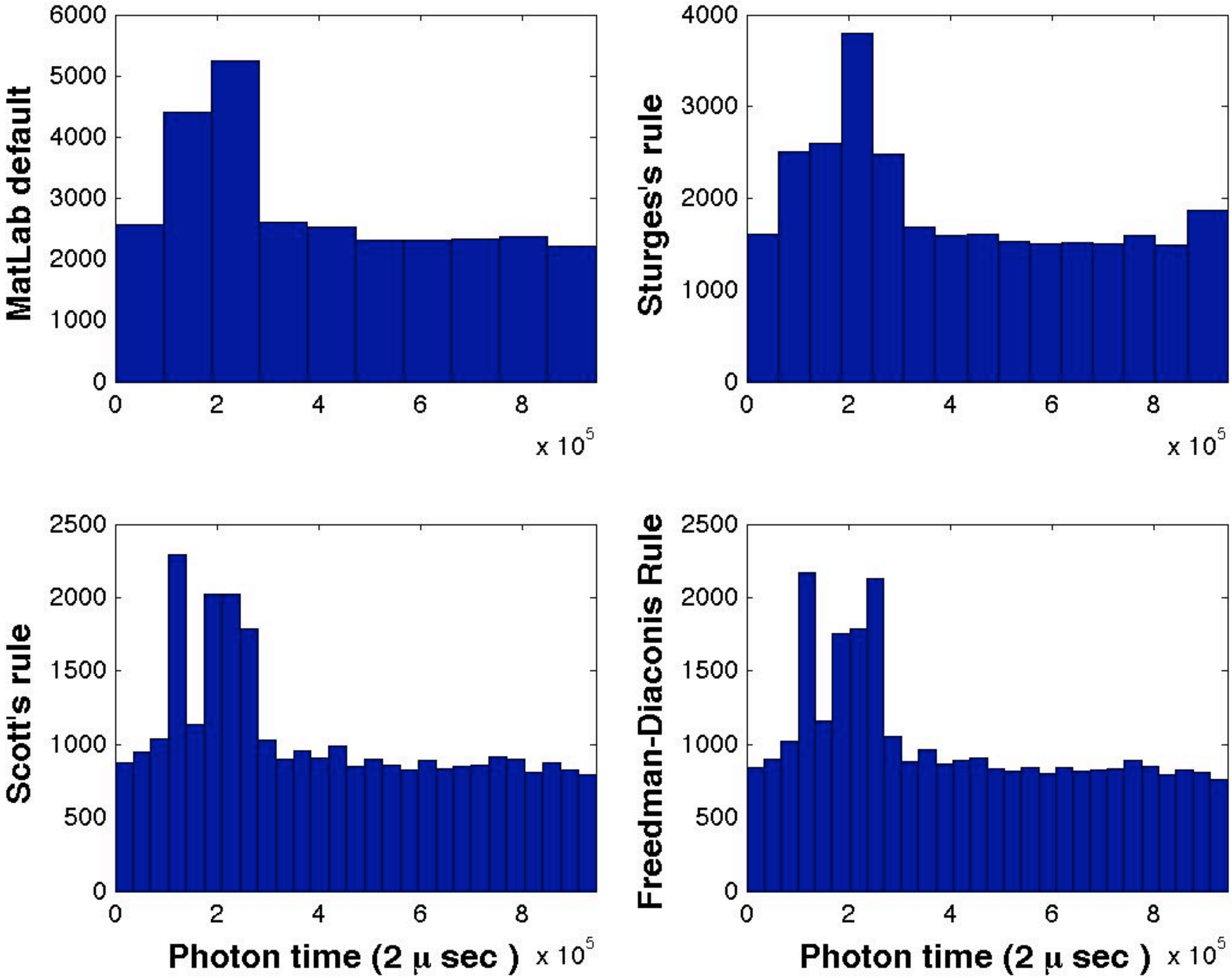
Step 0: Investigate the distribution of sample intervals

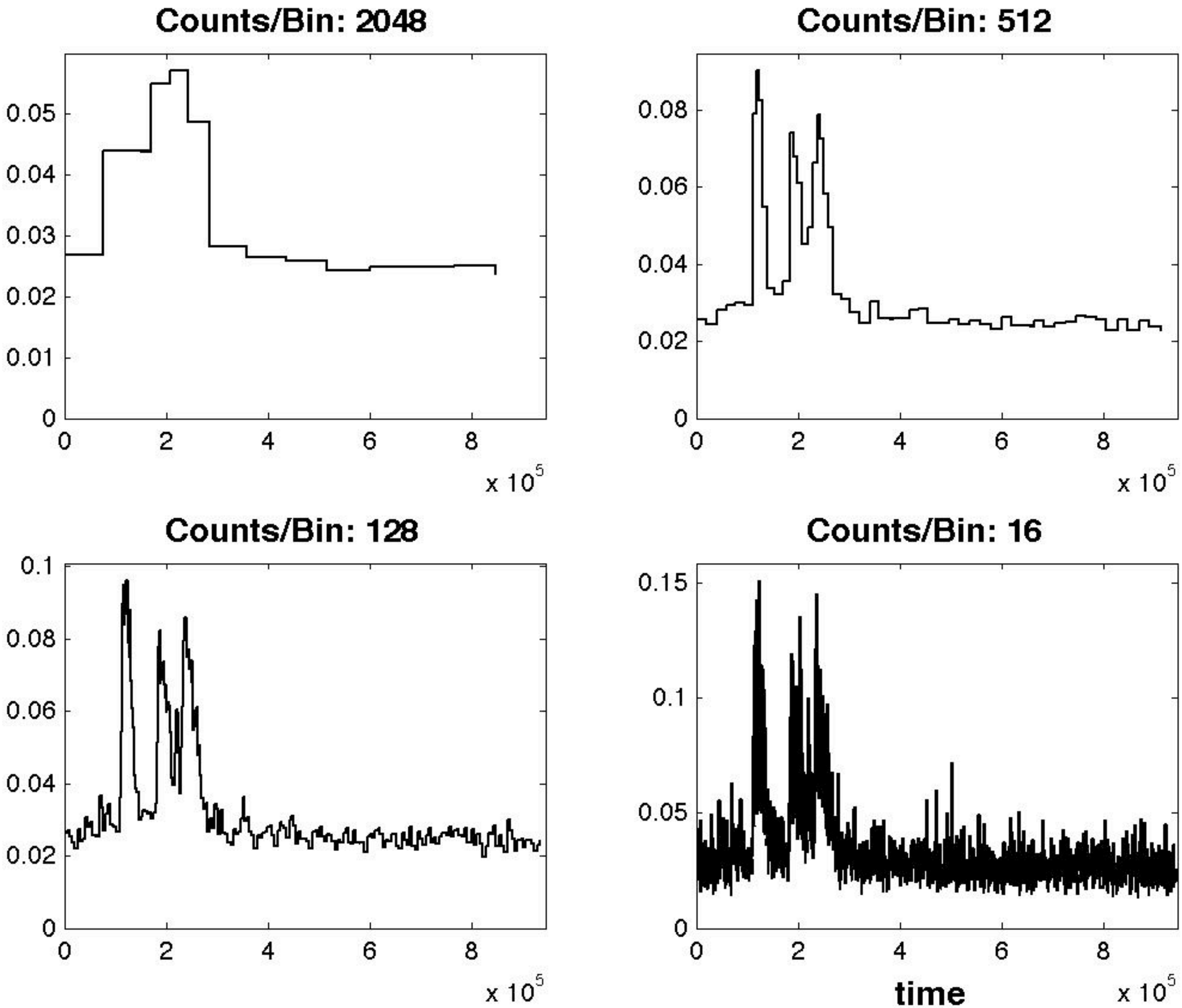






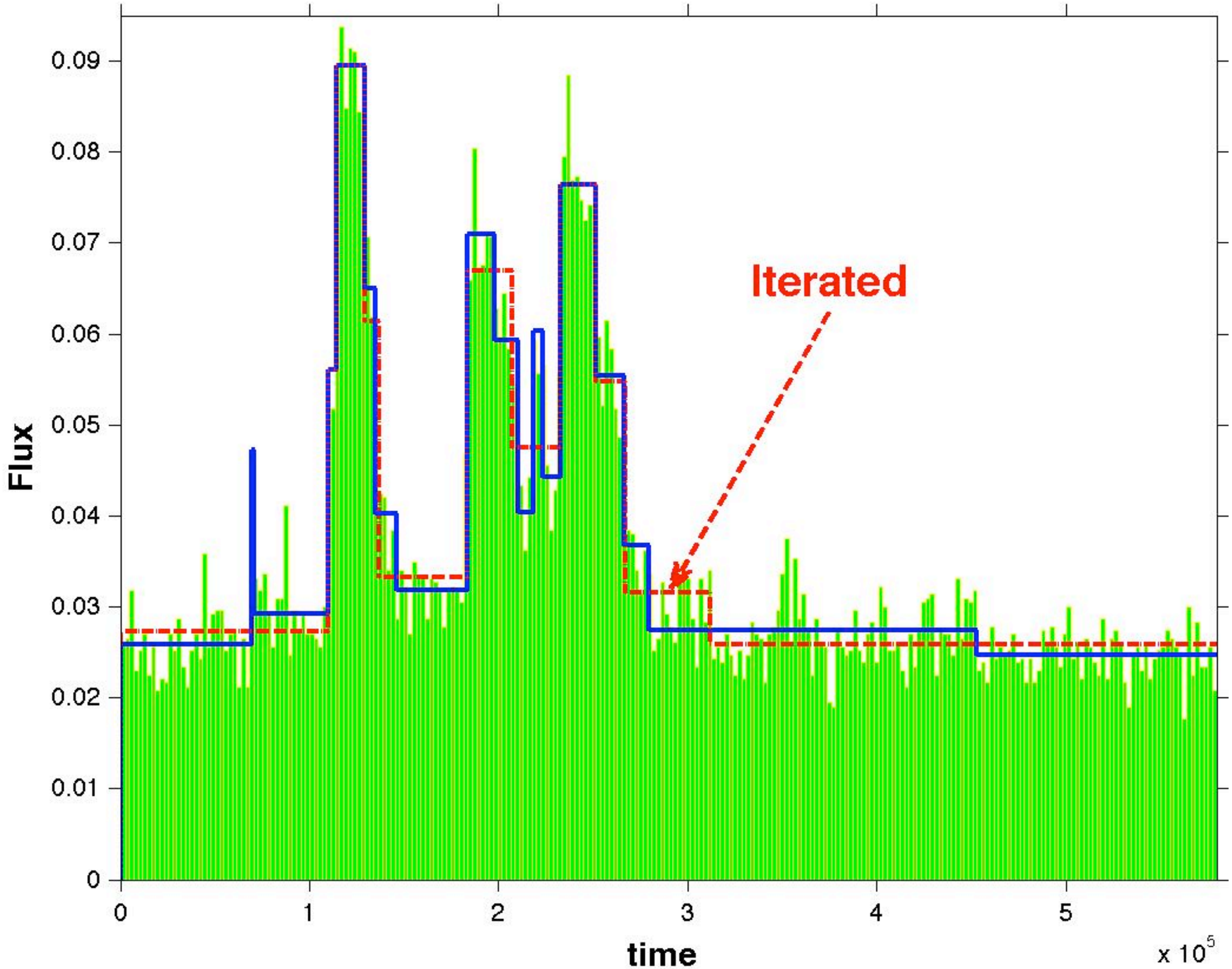
x 10^5
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Photon time (2 µ sec)
0
0.5
1
1.5
2
2.5
3
index
x 10^4

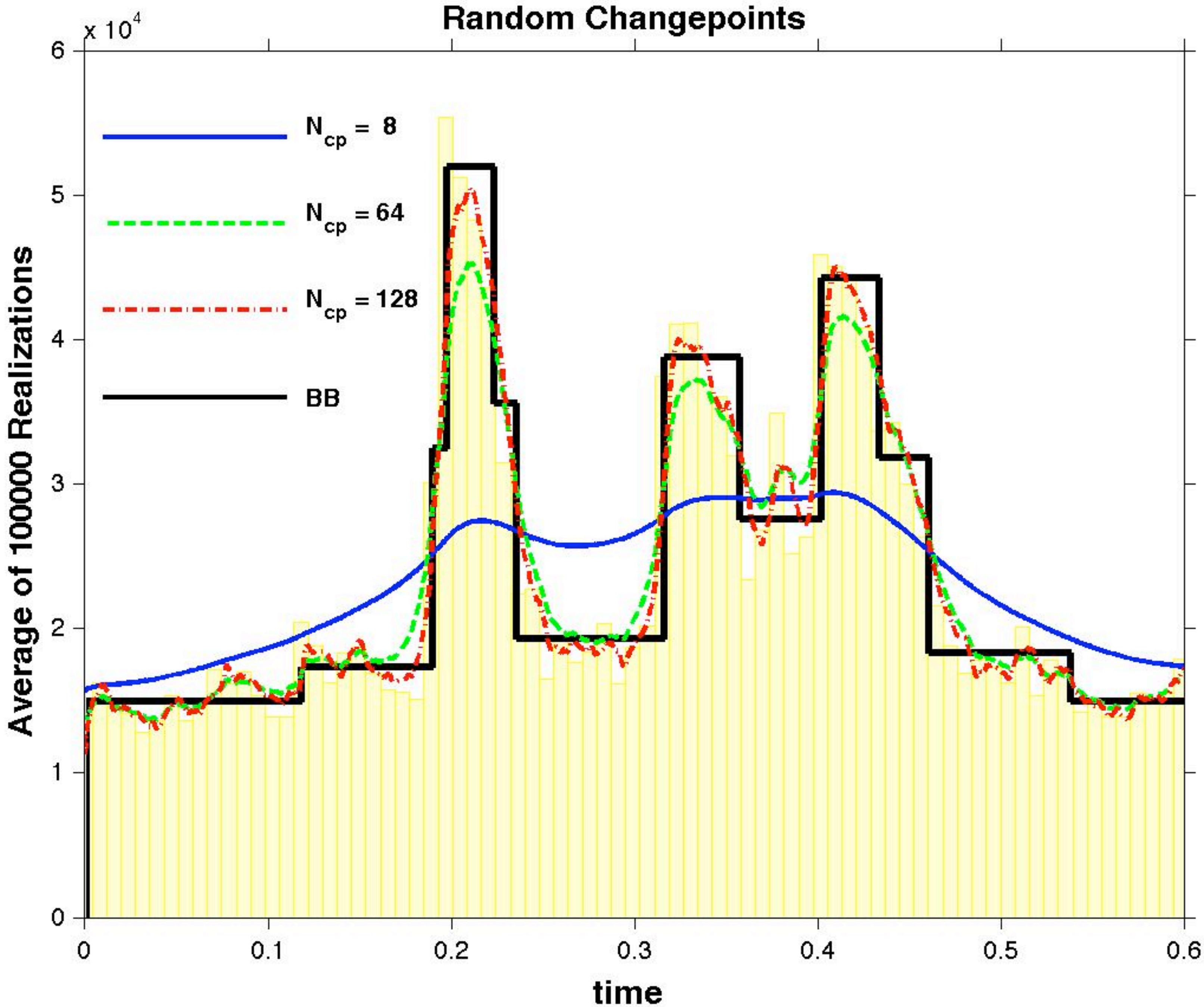

MatLab default
6000
5000
4000
3000
2000
1000
0
0
2
4
6
8
x 10^5
Sturges's rule
4000
3000
2000
1000
0
0
2
4
6
8
x 10^5
Scott's rule
2500
2000
1500
1000
500
0
0
2
4
6
8
Photon time (2 μ sec)
x 10^5
Freedman-Diaconis Rule
2500
2000
1500
1000
500
0
0
2
4
6
8
Photon time (2 μ sec)
x 10^5



Counts/Bin: 2048
0.05
0.04
0.03
0.02
0.01
0
0
2
4
6
8
x 10^5
Counts/Bin: 512
0.08
0.06
0.04
0.02
0
0
2
4
6
8
x 10^5
Counts/Bin: 128
0.1
0.08
0.06
0.04
0.02
0
0
2
4
6
8
x 10^5
Counts/Bin: 16
0.15
0.1
0.05
0
0
2
4
6
8
time
x 10^5

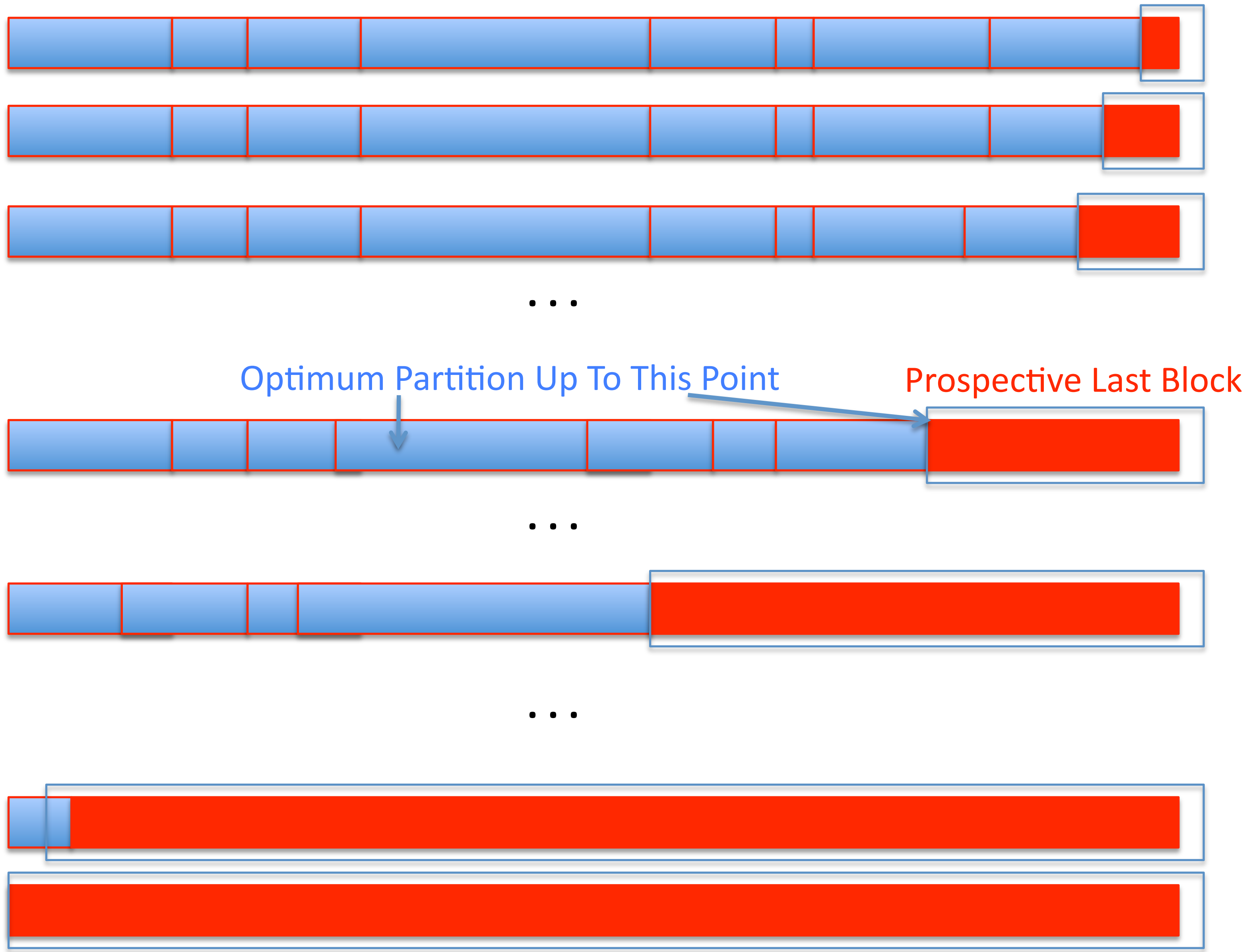
Bayesian Blocks




Random Changepoints
x 10^4
Average of 100000 Realizations
time
N_cp = 8
N_cp = 64
N_cp = 128
BB
0
1
2
3
4
5
6
0
0.1
0.2
0.3
0.4
0.5
0.6

Bayesian Blocks: The Optimizer

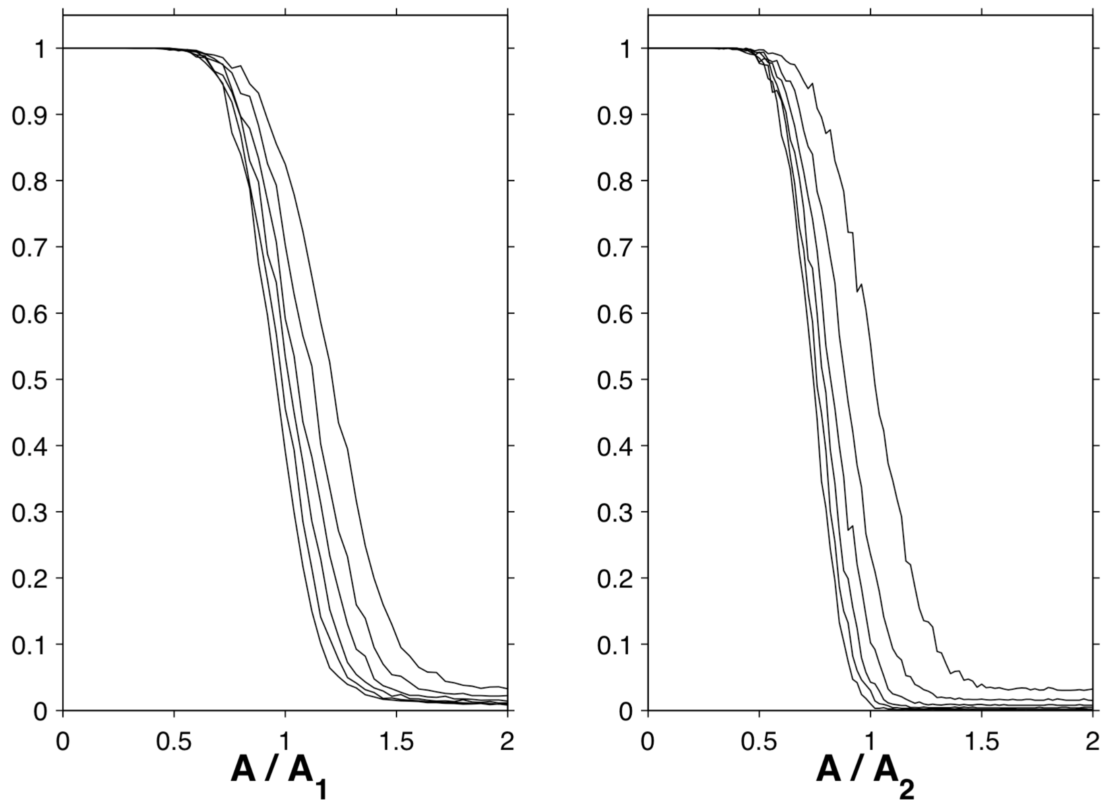
```
best = []; last = [];
for R = 1:num_cells
   [ best(R), last(R) ] = max( [0 best] + ...
        reverse( log_post( cumsum( data_cells(R:-1:1, :) ), prior,  type )  ) );

    if first > 0 & last(R) > first  % Option: trigger on first significant block
        changepoints = last(R); return
    end
end

% Now locate all the changepoints
index = last( num_cells );
changepoints = [];
while index > 1
   changepoints = [ index changepoints ];
    index = last( index - 1 );
end
```



Optimum Partition Up To This Point
Prospective Last Block

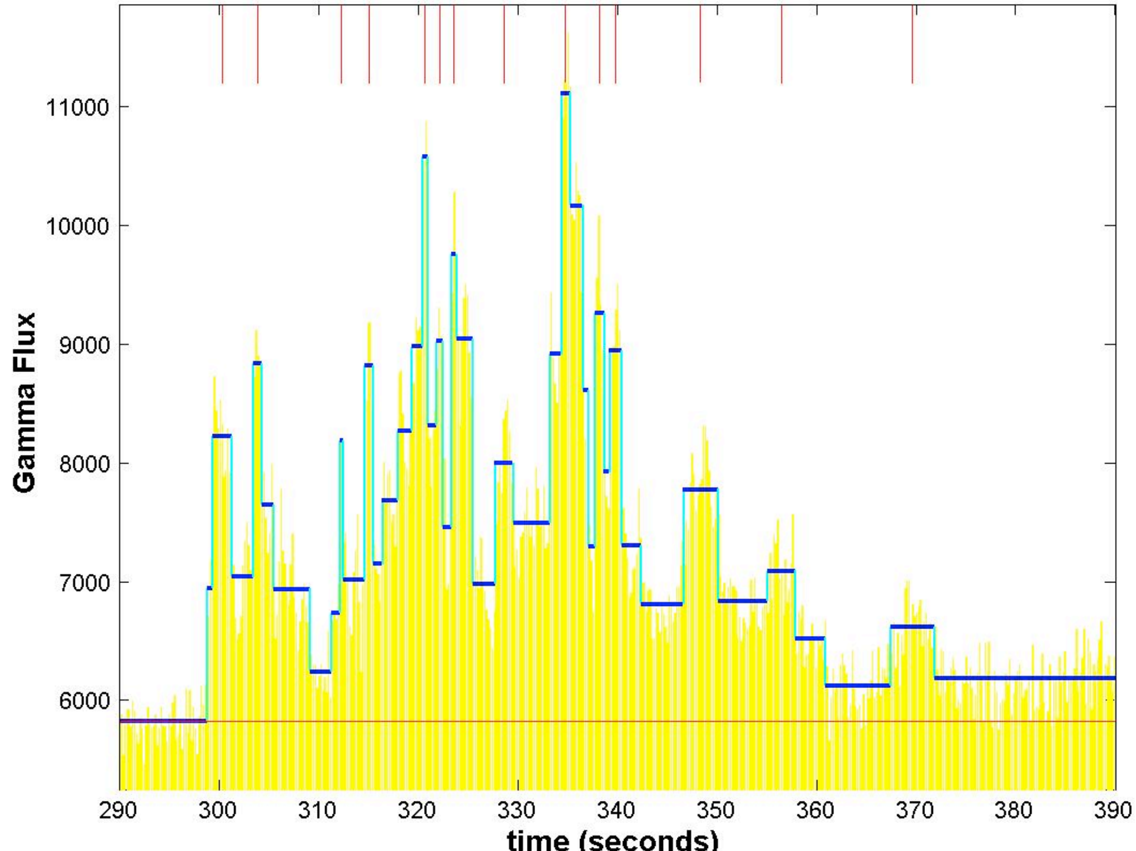
BB Nearly Achieves Theoretical Detection Limit

Detection error rate vs. signal amplitude in units of asymptotic result.



Arias-Castro, E., , Donoho, D., & Huo, X. 2003, Near-Optimal Detection of Geometric Objects by Fast Multiscale Methods
IEEE Transactions on Information Theory, 51, 2402-2425

Studies in Astronomical Time Series Analysis. VI. Bayesian Block Representations
Jeffrey D. Scargle, Jay P. Norris, Brad Jackson, James Chiang arxiv.org/abs/1207.5578
Listed in Astrophysics Source Code Library: http://asci.net

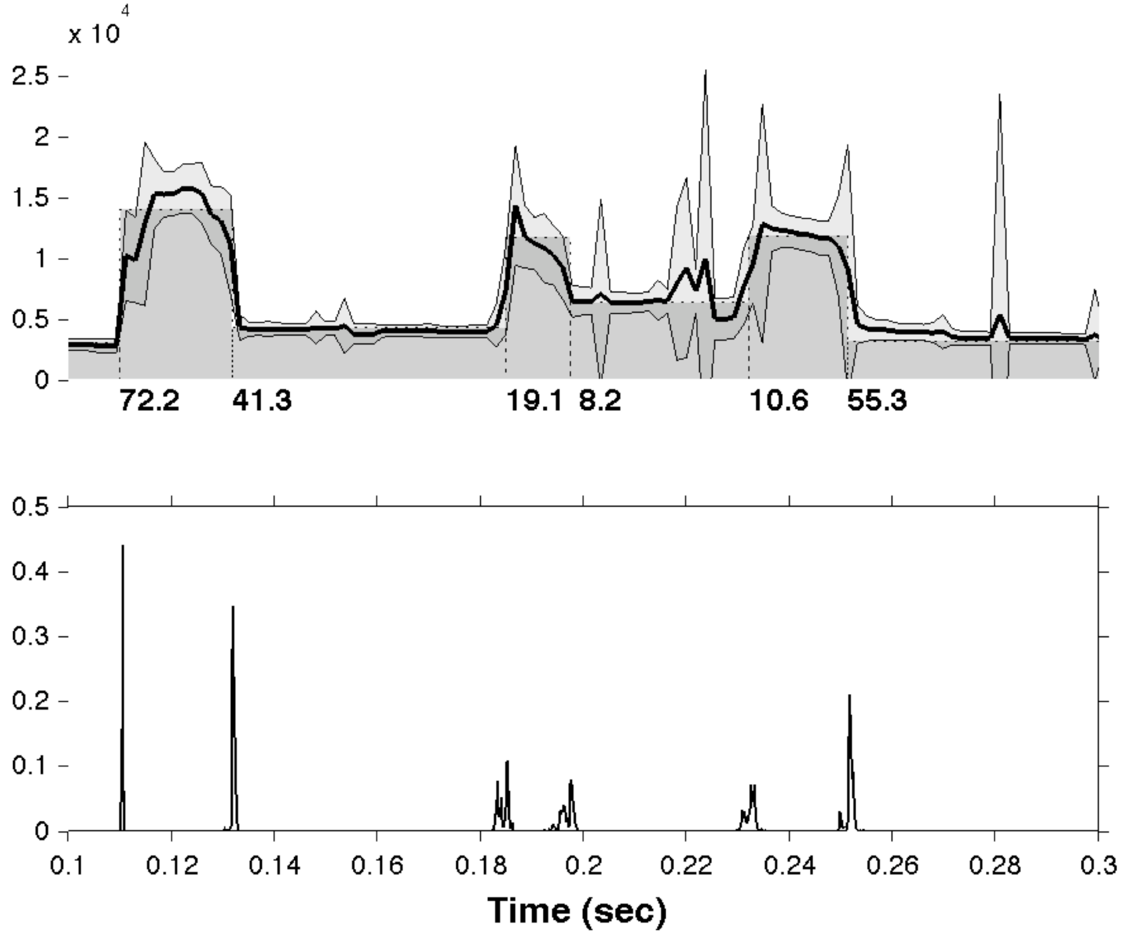


Bellman, R. 1961, On the approximation of curves by line segments using dynamic programming, Communications of the ACM, 4, 284.

15 Years of Reproducible Research in Computational Harmonic Analysis. Donoho, D. et al. 2009, Computing in Science and Engineering, 11, 8
stats.stanford.edu/~donoho/Reports/2008/15YrsReproResch-20080426.pdf

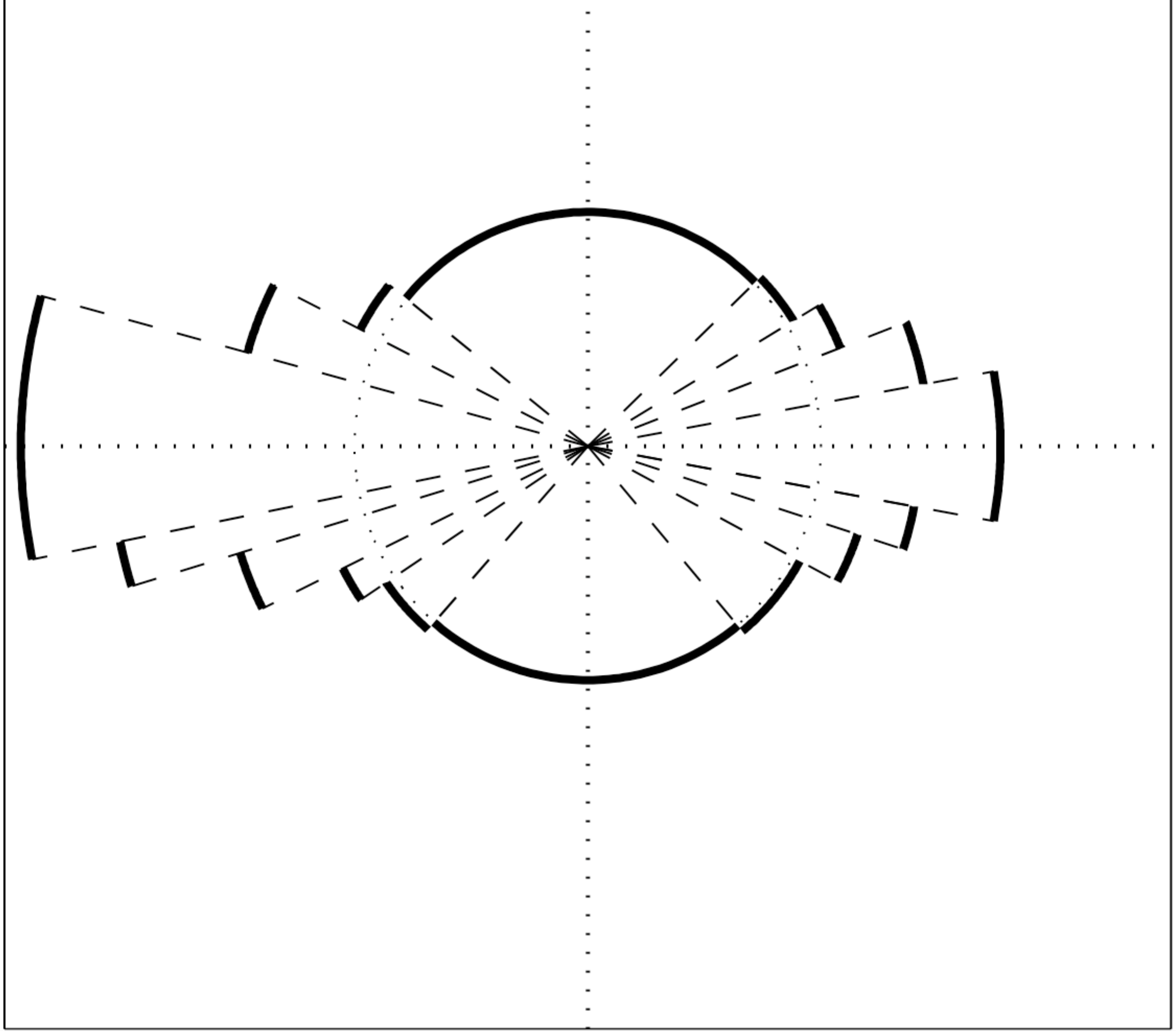
Analysis of Variance

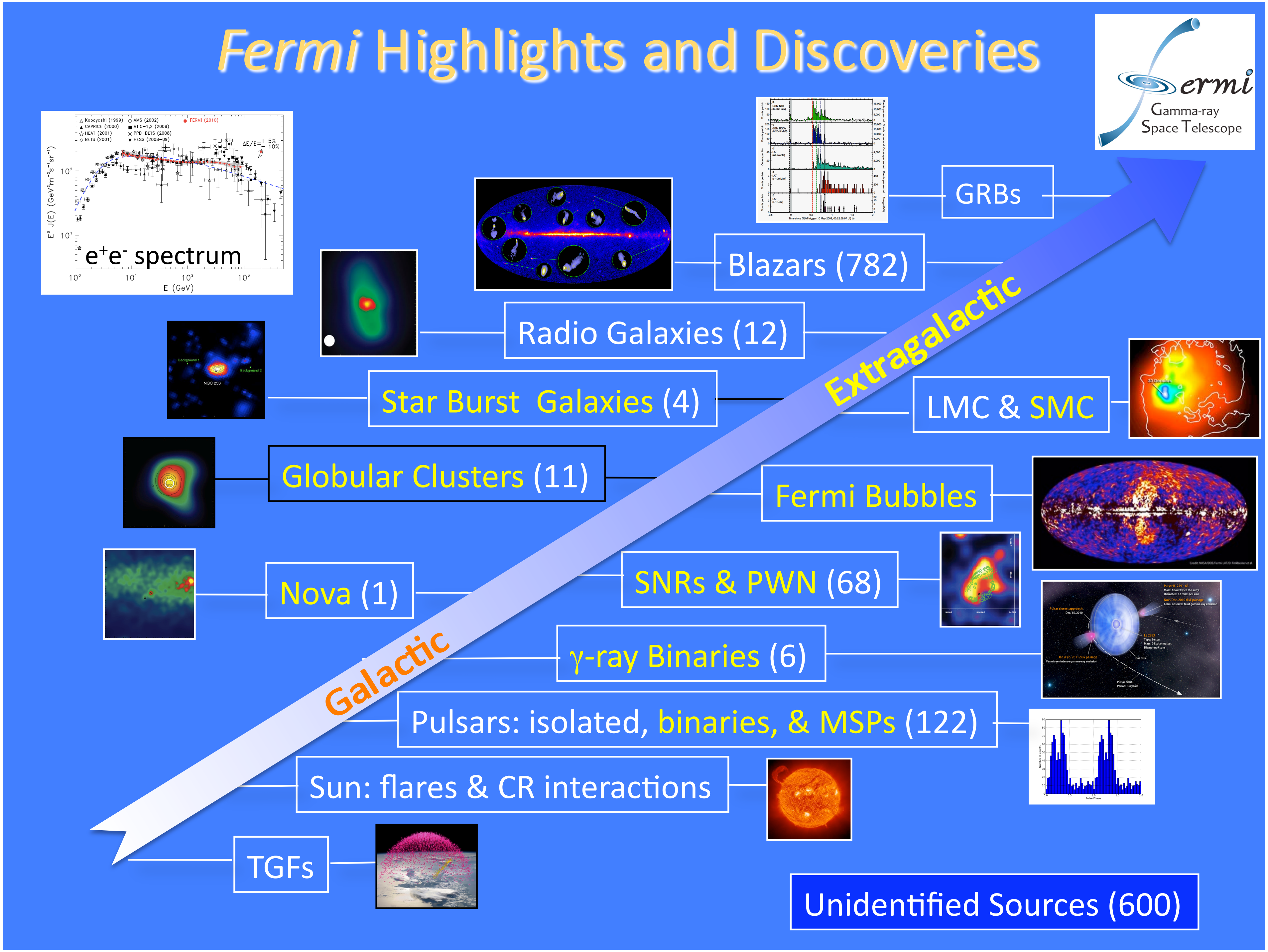
Bootstrap: (apologies to Tom Loredo
and possibly Aletheia)



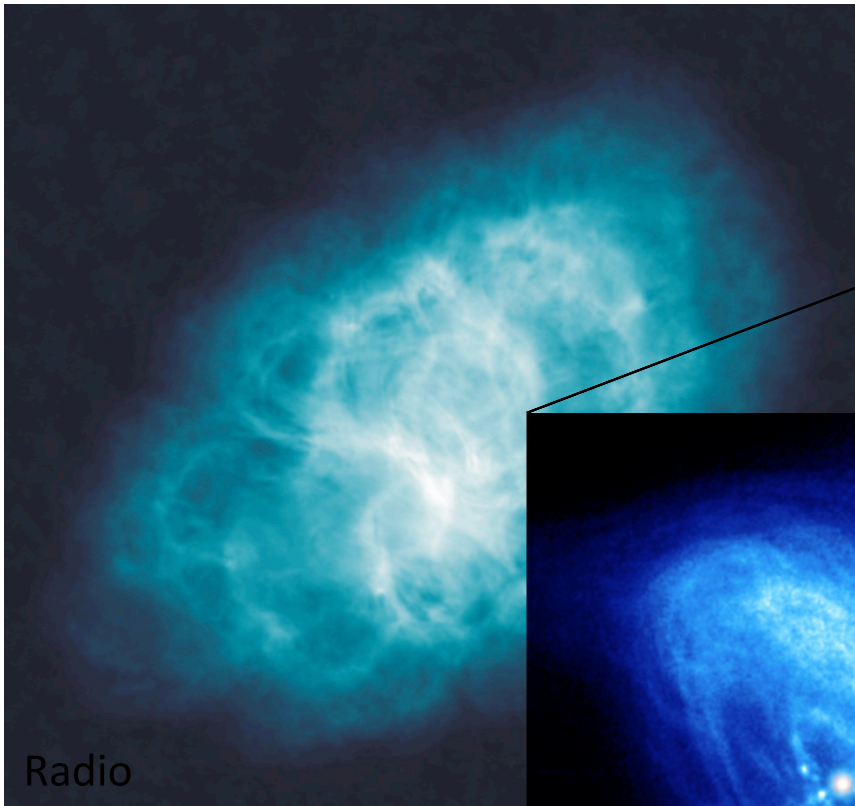
Single change-point likelihoods

Optimal Partitioning Data on the Circle

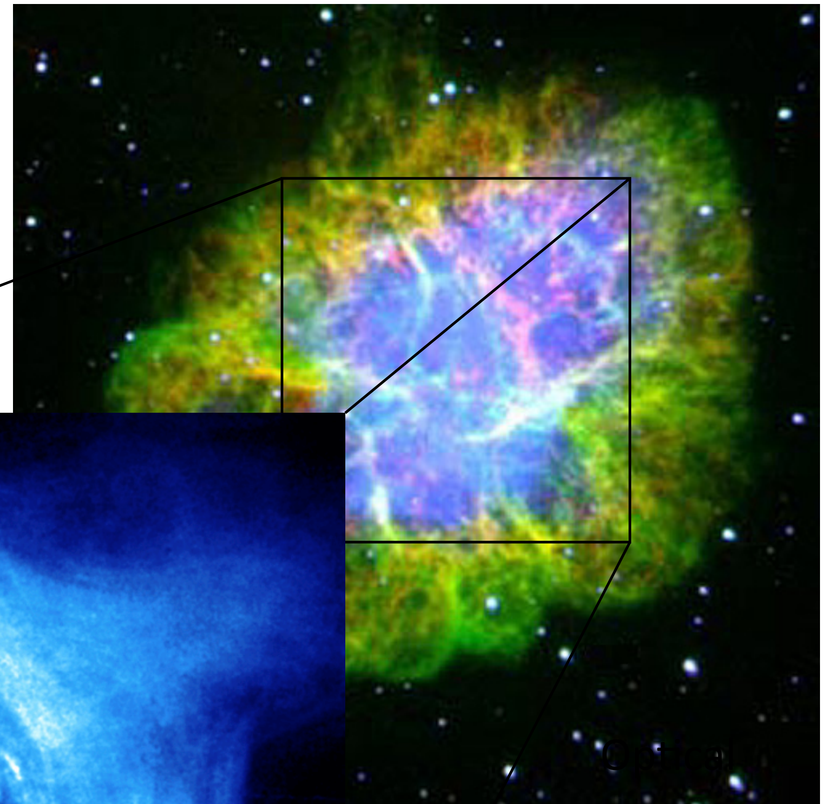



Fermi Highlights and Discoveries
Fermi Gamma-ray Space Telescope
e+e- spectrum
GRBs
Blazars (782)
Radio Galaxies (12)
Star Burst Galaxies (4)
LMC & SMC
Globular Clusters (11)
Fermi Bubbles
Nova (1)
SNRs & PWN (68)
γ-ray Binaries (6)
Pulsars: isolated, binaries, & MSPs (122)
Sun: flares & CR interactions
TGFs
Unidentified Sources (600)
Extragalactic
Galactic

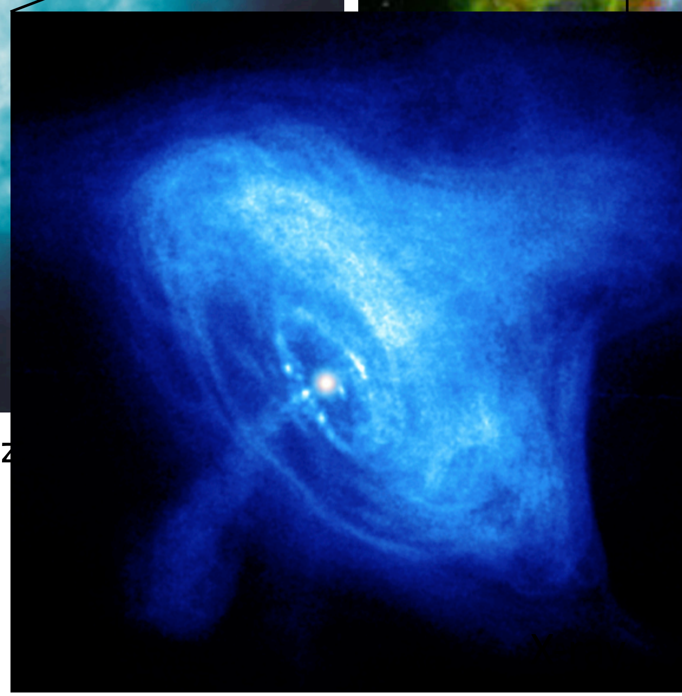
The Crab Nebula



NRAO/AUI and M. Bietenholz



SU/J.Hester & A.Loll)

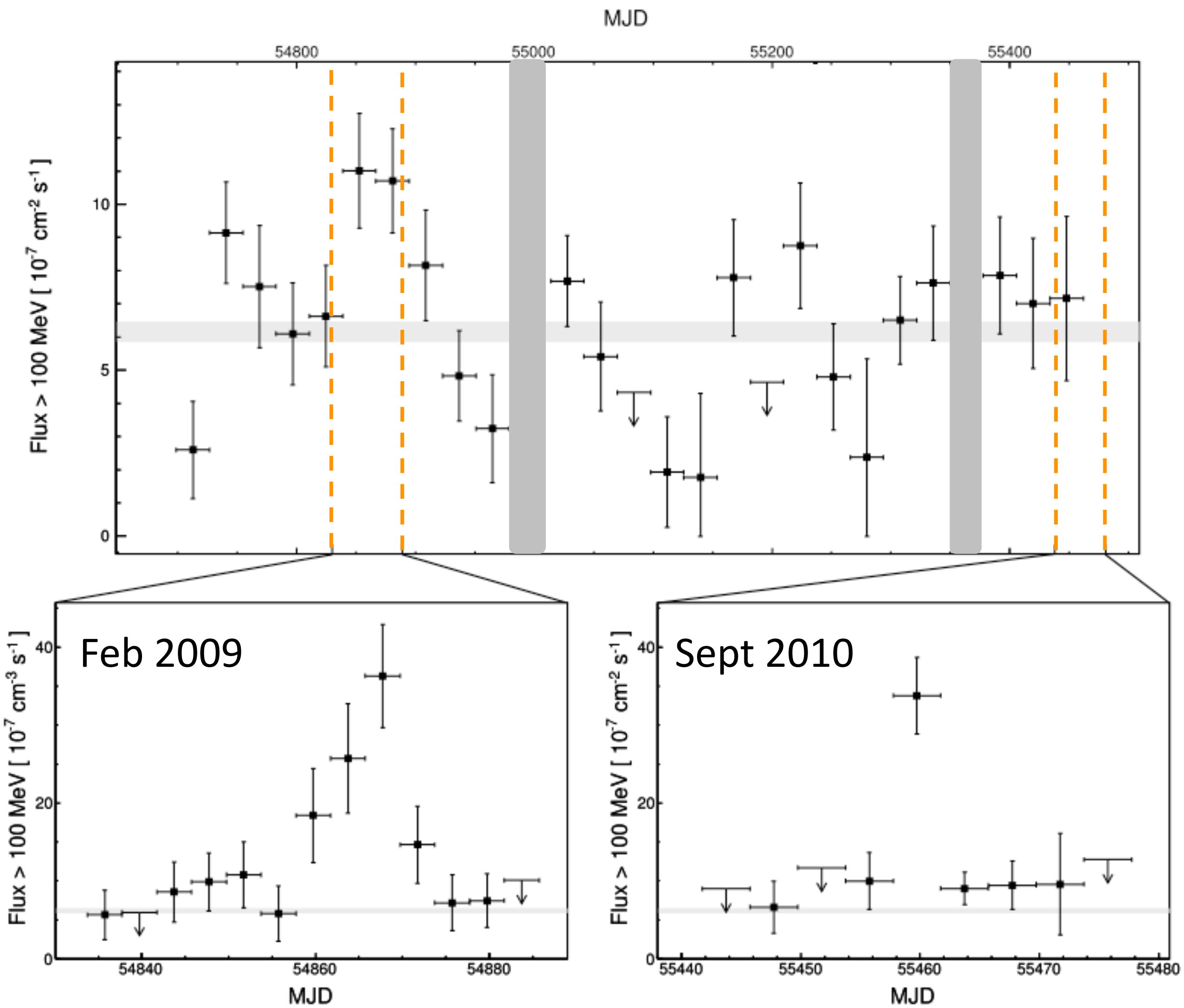


NASA/CXC/ASU/J.Hester et al.

See review
Hester, J. J., 2008,
ARA&A, 46, 127

Fermi Gamma-ray Space Telescope

Previous Flares viewed by Fermi LAT



No variability found in pulsar or high energy (IC) LAT component

4 week intervals

Sun passages excluded

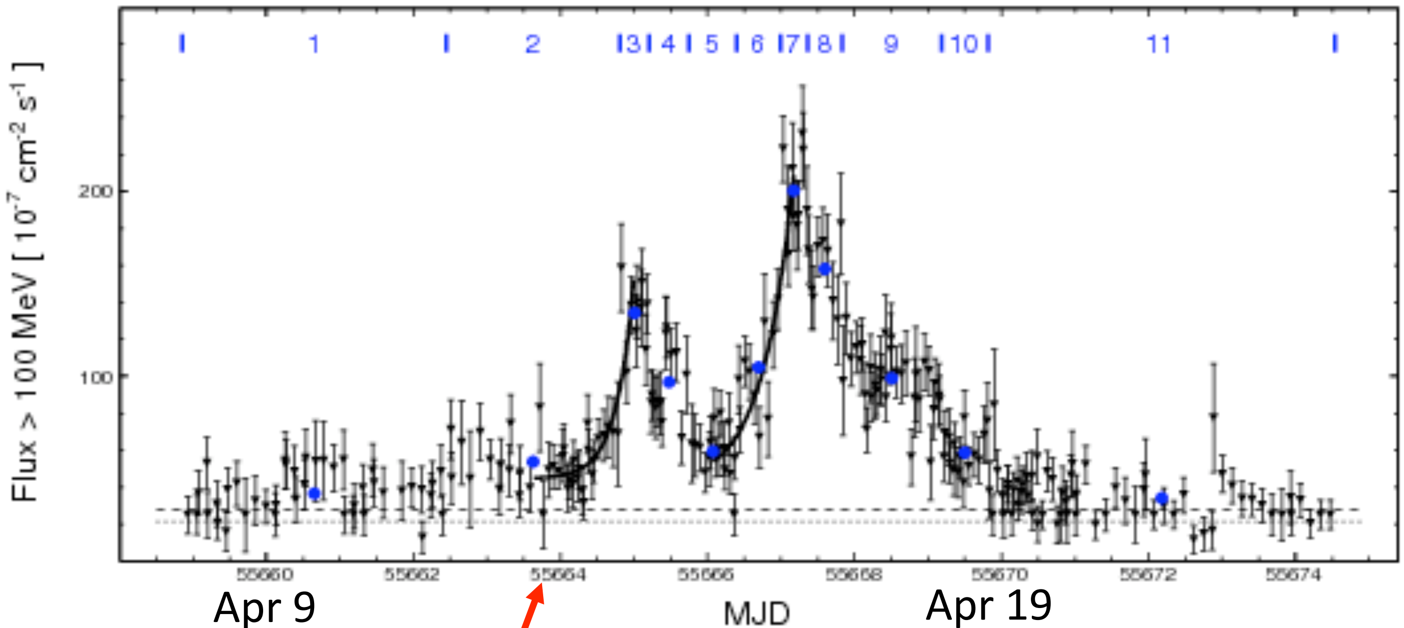
4 day intervals covering flare periods

Abdo et al. 2011, Science, 331, 739

Fermi Gamma-ray Space Telescope

April 2011 Flare

Lightcurve in bins of equal exposure (mean 9 minutes!)



Apr 9

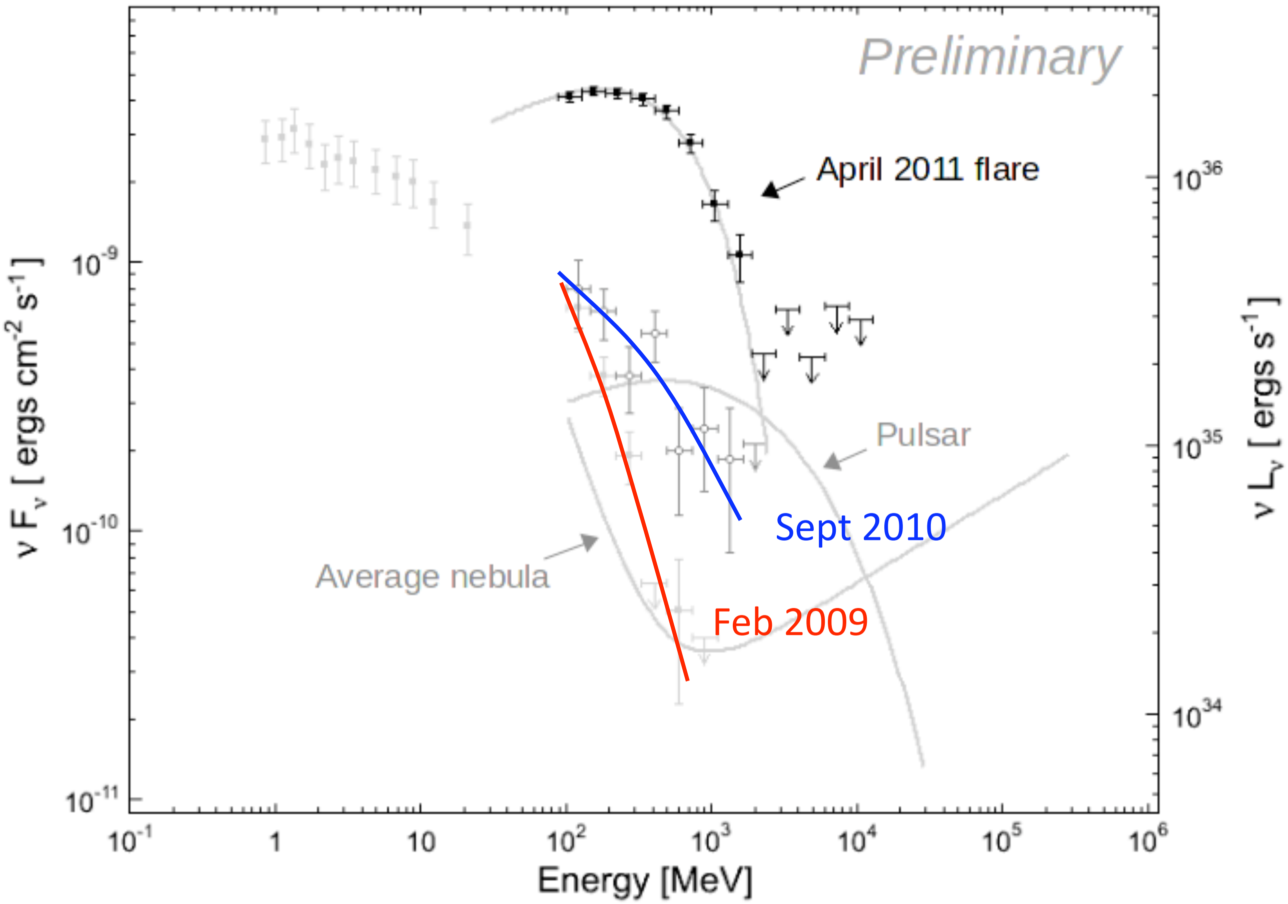
Apr 19

Beginning of LAT TOO

Flux doubling in 8 hours constrains emission region size <0.0003 pc

Buehler, R. et al. 2011, ApJ

Spectrum during the April 2011 flare

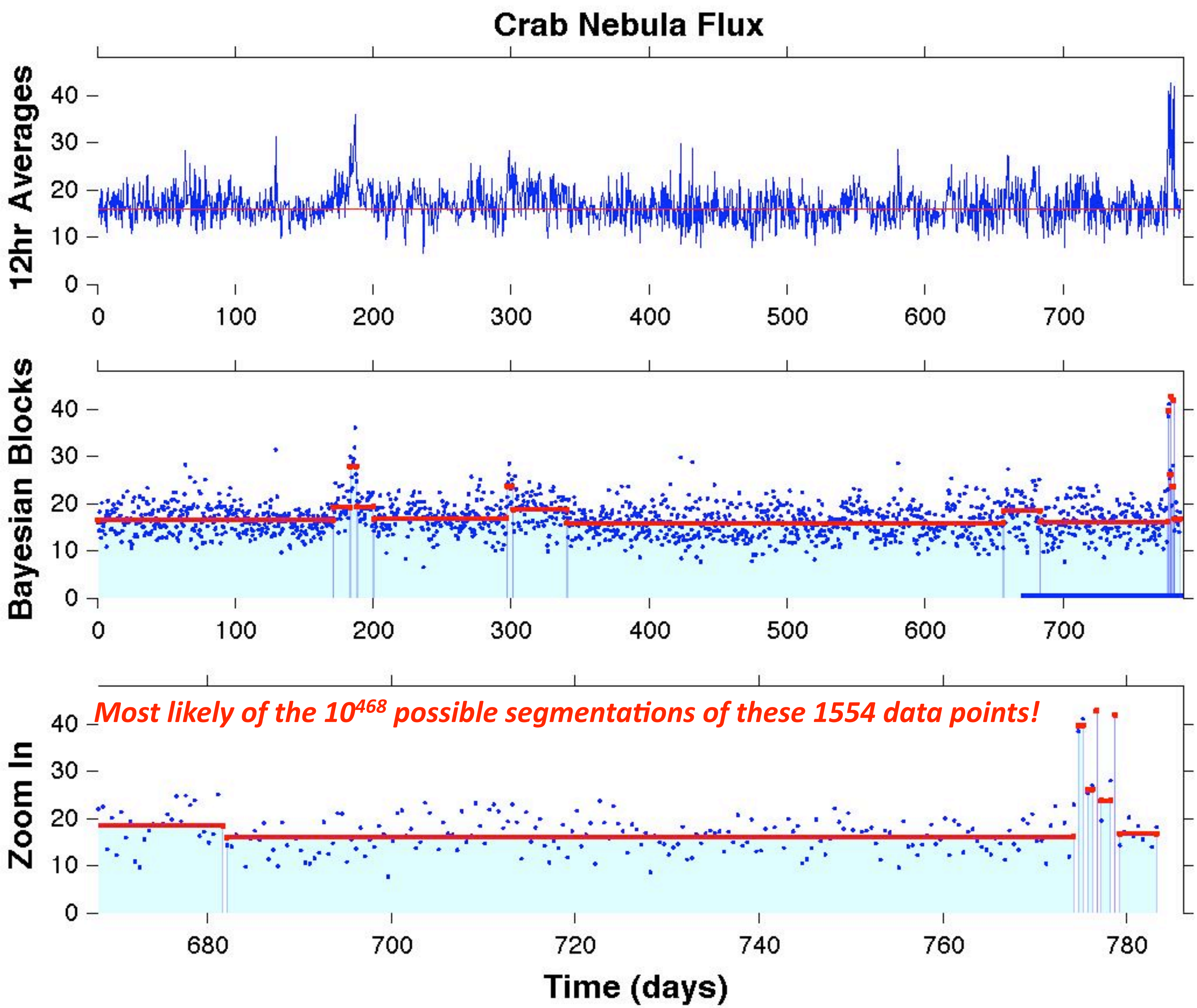


Flare is well described by additional power-law component that cuts off with energy

Photon index ~-1.3

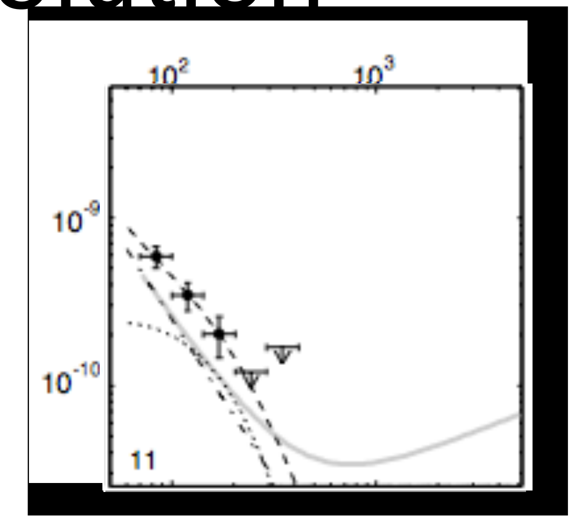
Energy of peak ~375 MeV

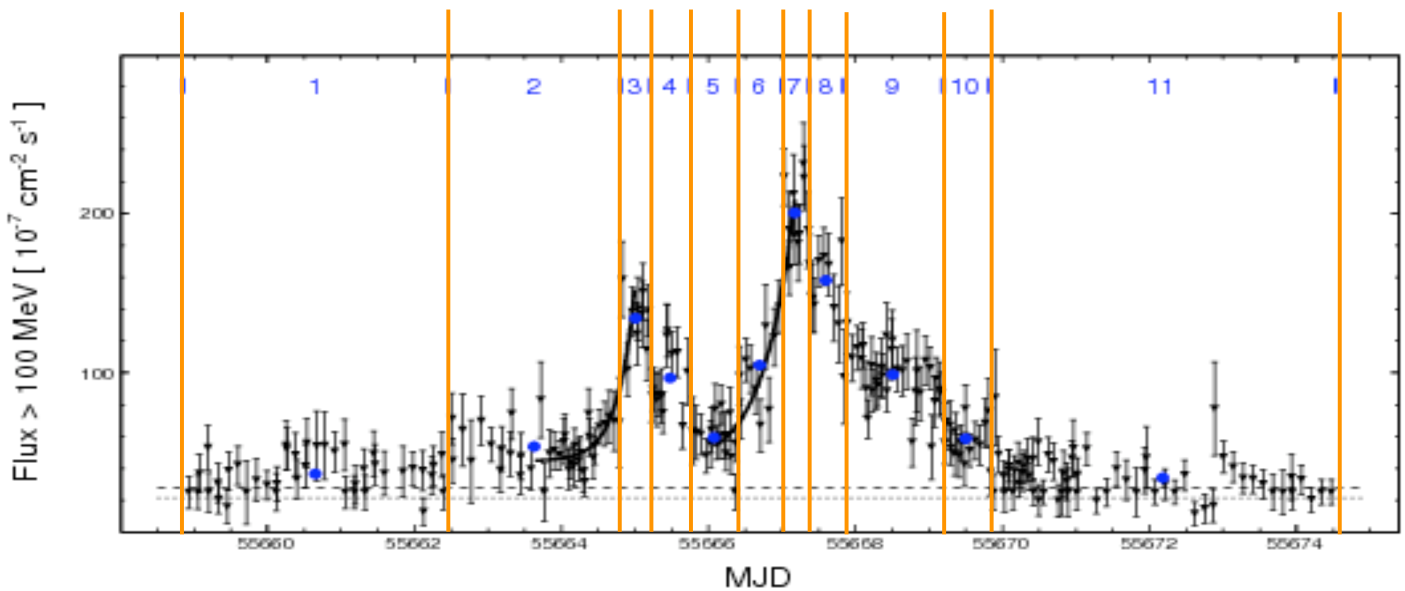
Energies above ~250 MeV exceed expected maximum for the electron synchrotron process


Crab Nebula Flux
12hr Averages
Bayesian Blocks
Zoom In
Most likely of the 10468 possible segmentations of these 1554 data points!
Time (days)

Spectral Evolution

"Sufficient statistics to partition the spectral fit. Bins of constant flux defined by Bayesian Blocks analysis."

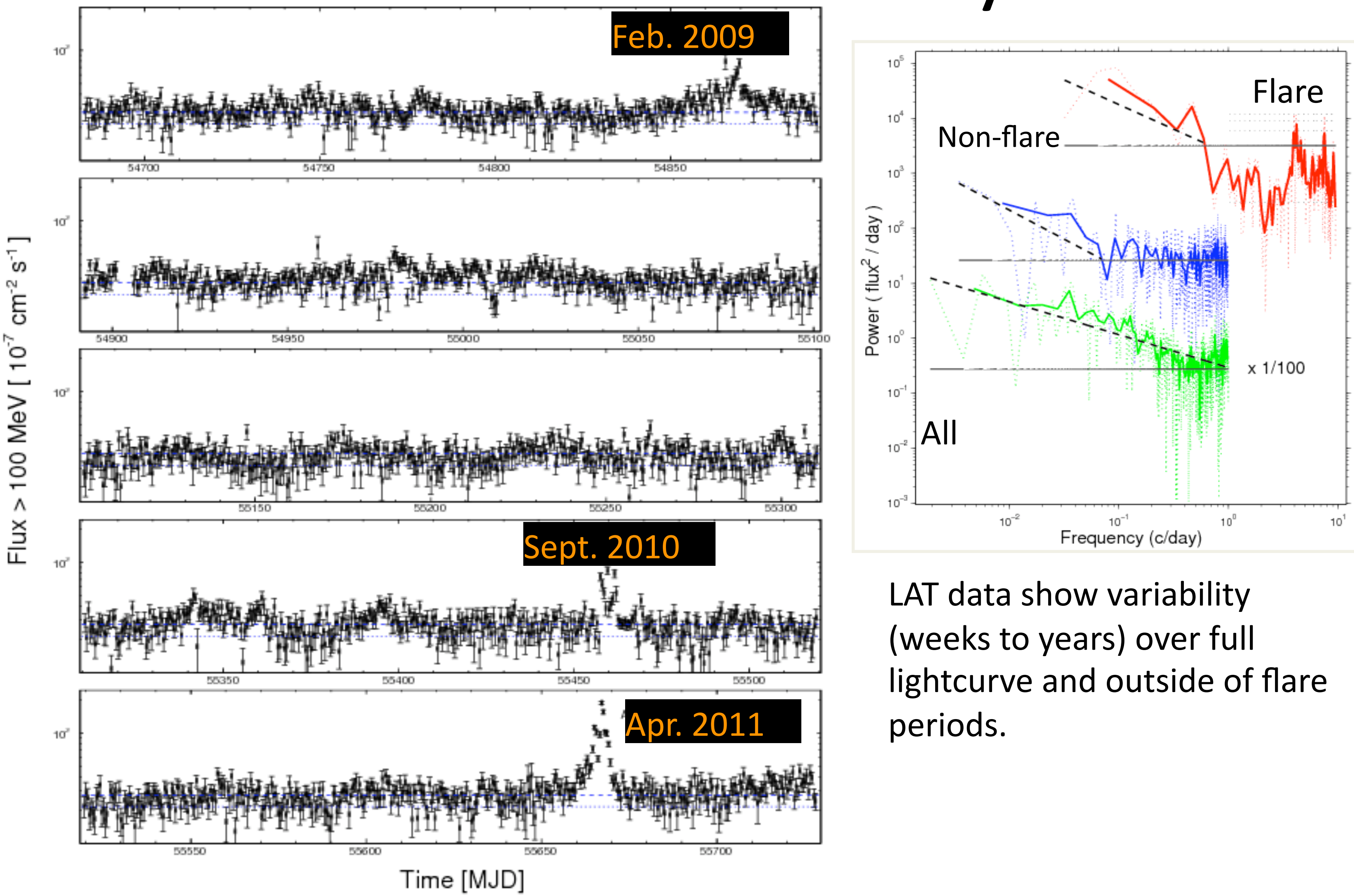




Searching for the Emission Region Sept 2010

No corresponding variability found in radio, optical, infrared, soft and hard X-rays around time of Sept. 2010 flare.

Persistent Variability



LAT data show variability (weeks to years) over full lightcurve and outside of flare periods.

Summary

- Fast, high-energy flares from the Crab are a largely gamma-ray phenomena
- Electrons reach very high energy in a very short time (8 hour flux doubling time)
- Where do the gamma rays originate in the nebula?
  - No pulsations -> outside the pulsar light cylinder
  - 0.0003 pc constraint is smaller than the termination shock region
  - Despite good coverage, no correlated variations or changes in features yet found at other wavelengths
- Rapid variability and high energy suggest relativistic beaming of electron synchrotron
- Time scale and small region imply electrons accelerated through electrostatic acceleration or magnetic reconnection
- Nebula is highly dynamic over the spectrum, but not clear how to connect features and timescales
- Future observations of large flares may help further pinpoint the emission site

Algorithm Categories ...

- Time Domain (local)
  - Nonparametric models
  - Bayesian Blocks (piecewise constant, linear, or exponential)
  - Symbolic Representation
  - Parametric model fitting

- Frequency/Scale Domain (global)
  - Power Spectrum
  - Correlation Function
  - Structure Function
  - Wavelet Spectrum

- Time-Frequency/Time Scale Domain (hybrid)
  - Time-frequency distribution
  - Wavelet transform

Cross- and Auto- Correlation Functions for unevenly spaced data

Edelson and Krolik: "The Discrete Correlation Function: a New Method for Analyzing Unevenly Sampled Variability Data"
Astrophysical Journal 333 (1988) 646

$\rho_{xy}(\tau) = (1/N_k)\ \Sigma\, X(t_n)\, Y(t_m)$ for $t_n - t_m$ in τ bin

Data Mode

- Photon events
- Time-to-Spill
- Counts in bins
- Flux measurements
- Any Mode/Sampling!

Universal Time Series Analysis Machine

Auto-

- Correlation Function
- Fourier Power Spectrum
- Fourier Phase Spectrum
- Wavelet scalgram
- Wavelet scaleogram
- Structure Function
- Time-Frequency Distribution
- Time-Scale Distribution
- …

Extension of Edelson & Krolik
Algorithm for Correlation Function
of Unevenly Sampled Data

Jeff Scargle

Data Mode

- Photon events
- Time-to-Spill
- Counts in bins
- Flux measurements
- Any Mode/Sampling!

any combination

- Photon events
- Time-to-Spill
- Counts in bins
- Flux measurements
- Any Mode/Sampling!

Universal Time Series Analysis Machine

Extension of Edelson & Krolik Algorithm for Correlation Function of Unevenly Sampled Data

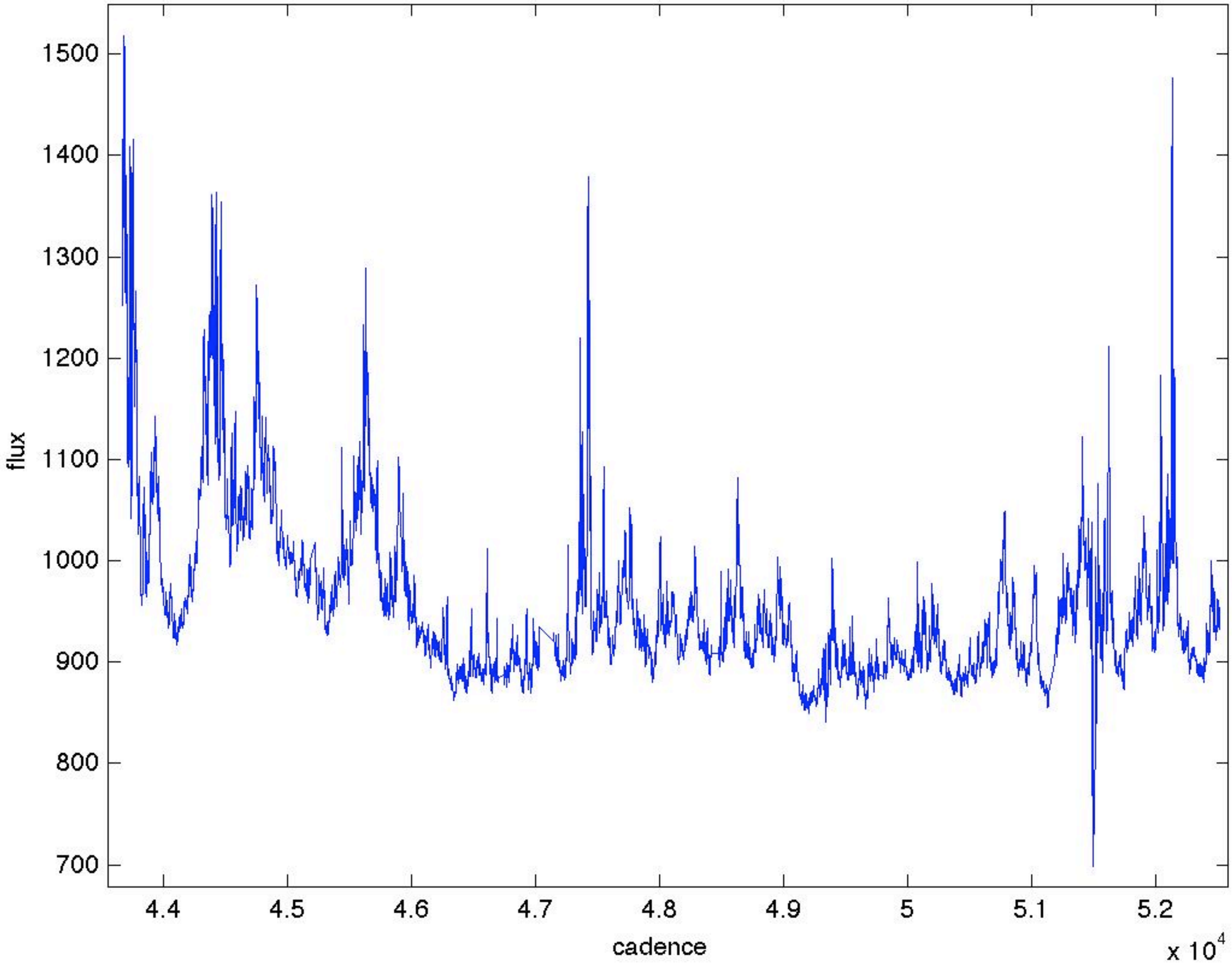
Cross-

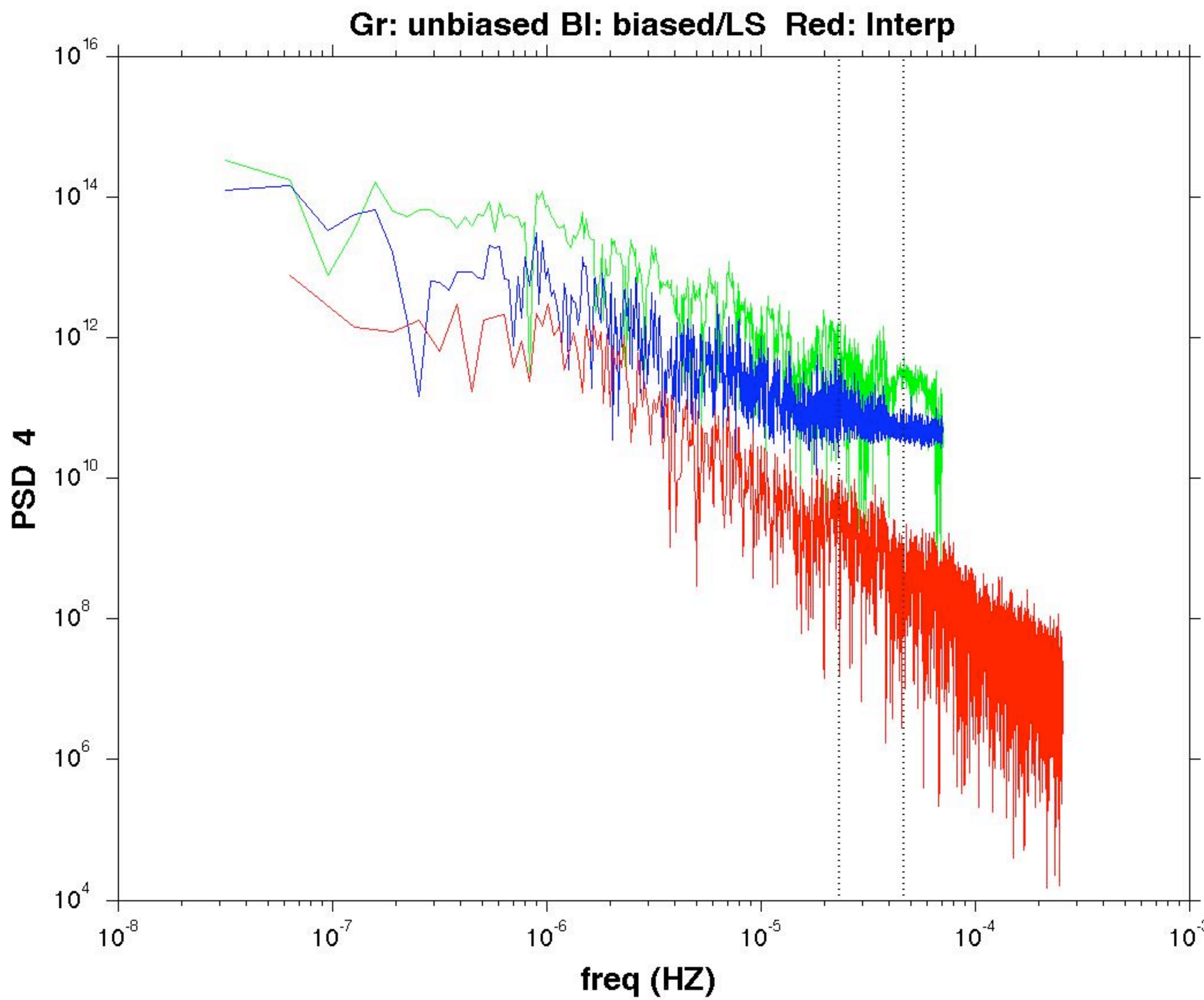
- Correlation Function
- Fourier Power Spectrum
- Fourier Phase Spectrum
- Wavelet scalgram
- Wavelet scaleogram
- Structure Function
- Time-Frequency Distribution
- Time-Scale Distribution
- …

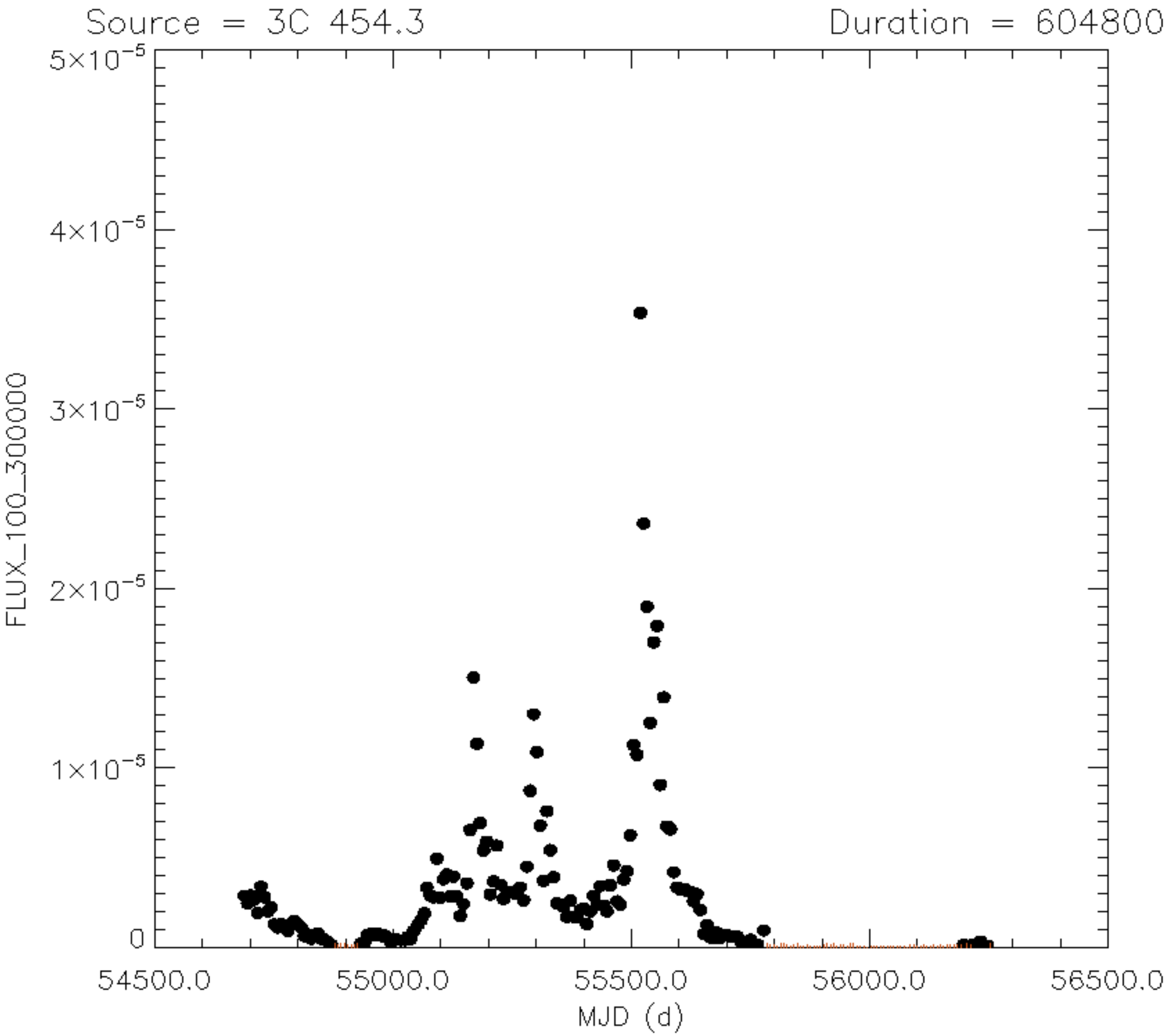
Jeff Scargle

Kepler observations of rapid optical variability in the BL Lac object W2R1926+42

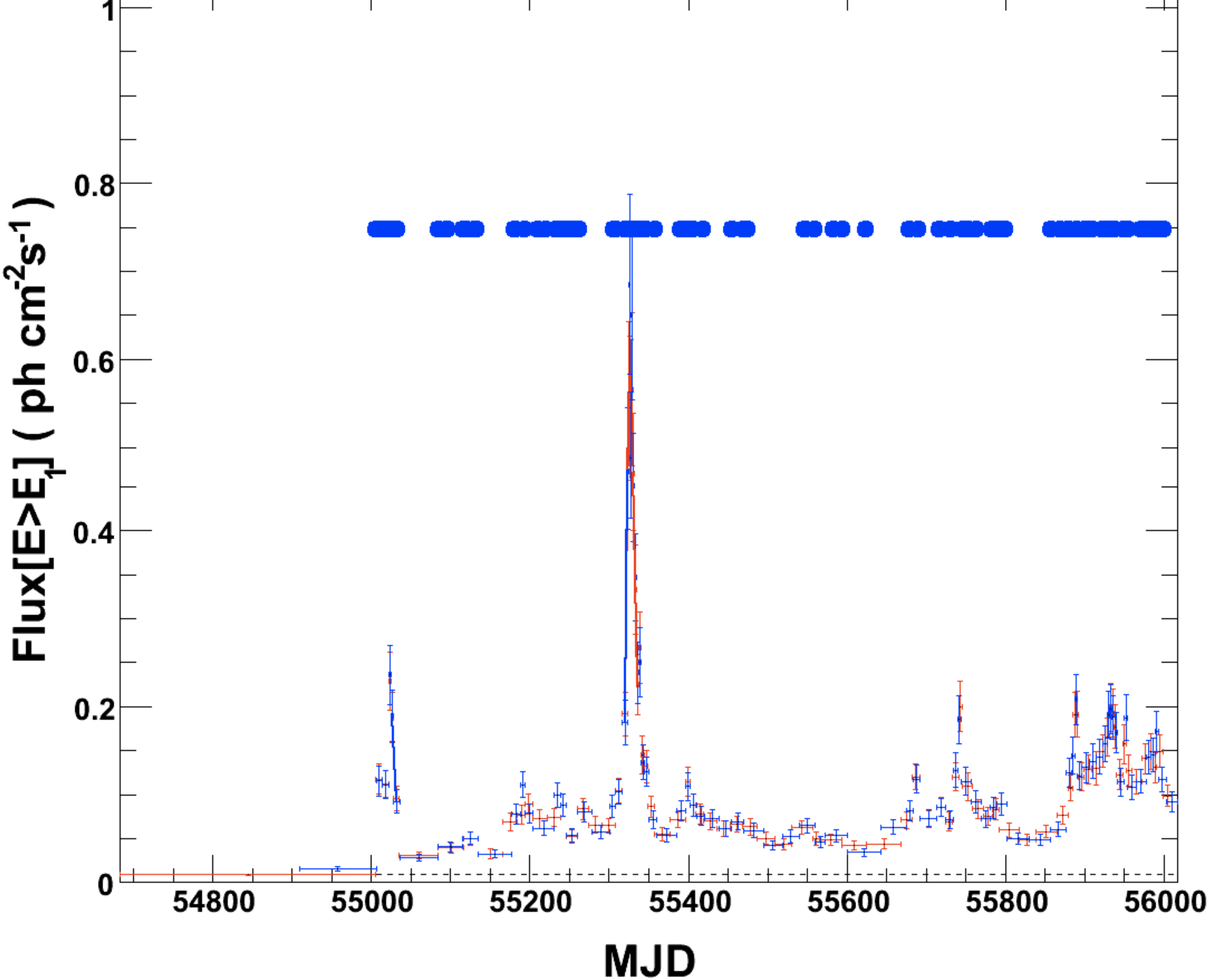
R. Edelson, R. Mushotzky, S. Vaughan, J. Scargle, P. Gandhi, M. Malkan, W. Baumgartner




Gr: unbiased Bl: biased/LS Red: Interp
PSD 4
freq (HZ)
10^{16}
10^{14}
10^{12}
10^{10}
10^{8}
10^{6}
10^{4}
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}



Source = 3C 454.3
Duration = 604800.
FLUX_100_300000
MJD (d)
5×10⁻⁵
4×10⁻⁵
3×10⁻⁵
2×10⁻⁵
1×10⁻⁵
0
54500.0
55000.0
55500.0
56000.0
56500.0



Flux[E>E_1] (ph cm^-2 s^-1)
1
0.8
0.6
0.4
0.2
0
54800
55000
55200
55400
55600
55800
56000
MJD

Solar Cycle Variability and Surface Differential Rotation from Ca II K-Line Time Series Data

Jeff Scargle, Ames Research Center
Stephen Keil, National Solar Observatory
Pete Worden, Ames Research Center

Submitted to Astrophysical Journal

Ca II K-line Monitoring Program

- National Solar Observatory/Sacramento Peak disk-integrated Calcium II K-line data
- November 1976 – December 2012
- Data: http://nsosp.nso.edu/data/cak_mon.html
- Keil, Henry and Fleck, Synoptic Solar Physics, ASP Conference Series (1998)

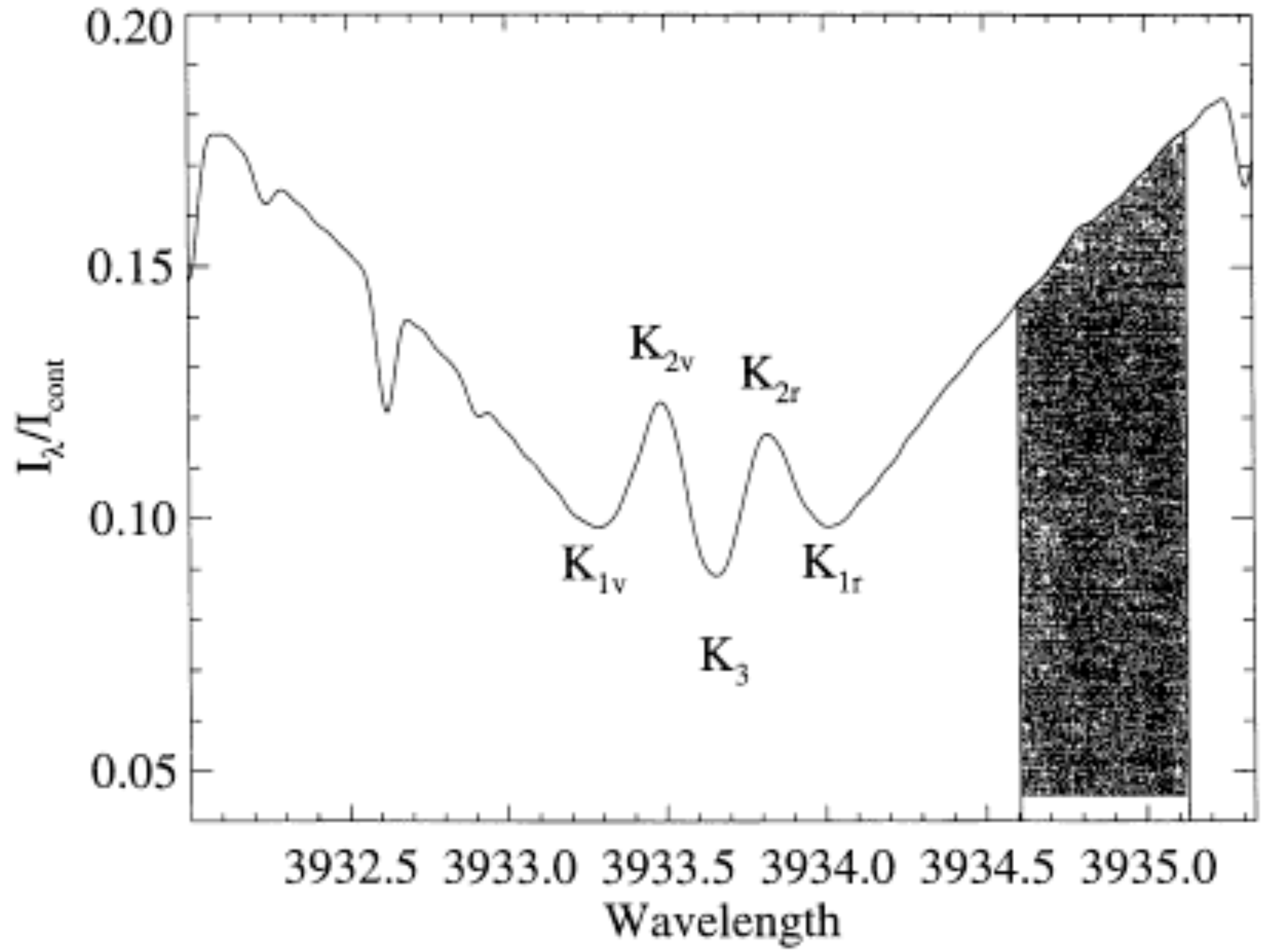
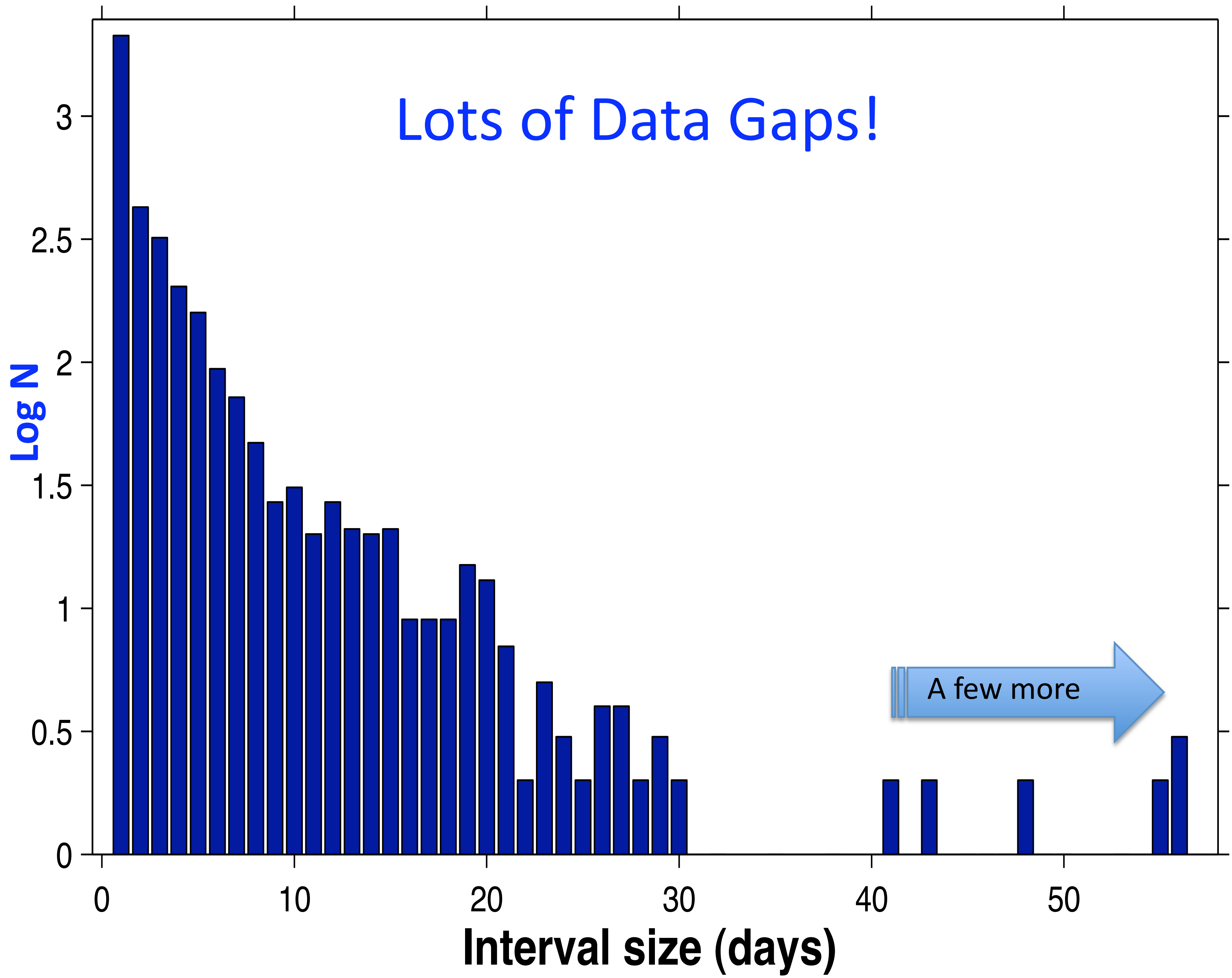


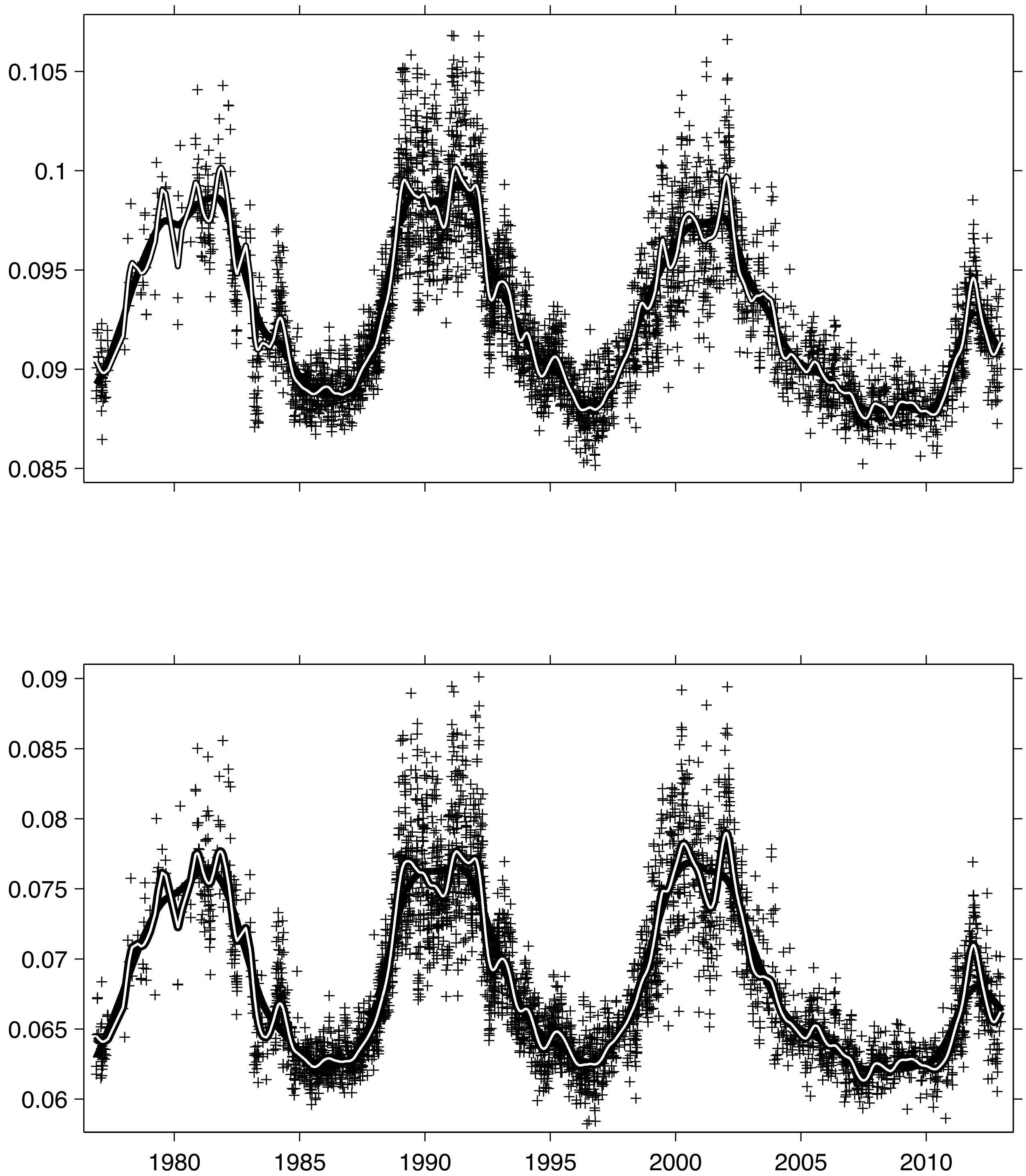
Fig. 1. A representative K-line profile showing the emission features and the 0.528 Å flux calibration passband defined by White and Livingston (1981) and Keil and Worden (1984).

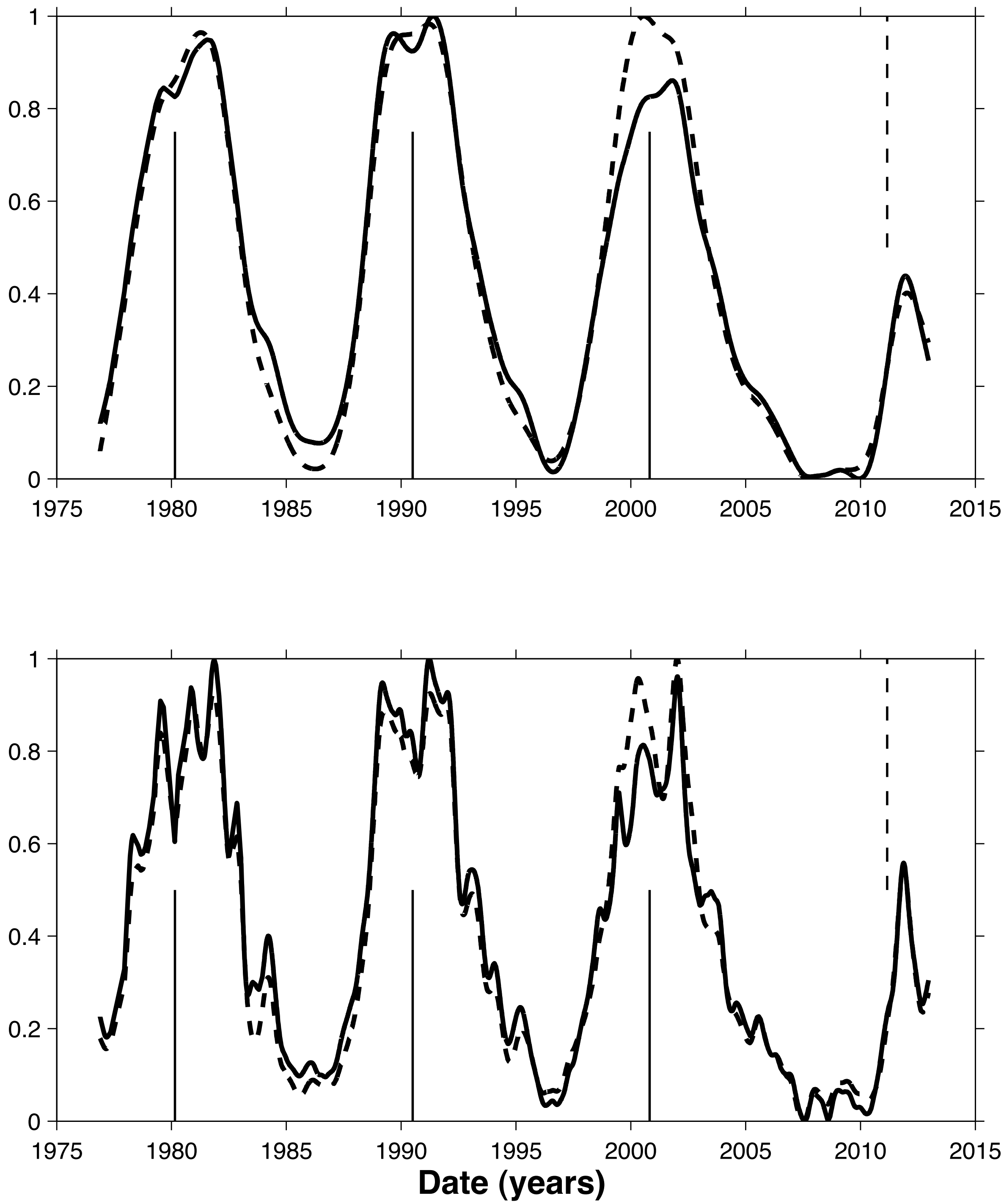
Measured K-line Parameters

- EM : Emission Index: equivalent width in 1Å band centered on core
- K_3 : Intensity in the core
- K_{2V}/K_3 : Relative strength of blue K_2 emission peak relative to K_3
- K_{2V}-K_{2R} : Separation of blue and red K_1 emission maxima
- K_{1V}-K_{1R} : Separation of blue and red K_1 minima
- K_{2V}/K_{2R} : Line asymmetry, ratio of blue/red emission peaks
- WB : Wilson-Bappu parameter, width between outer edges of emission peaks

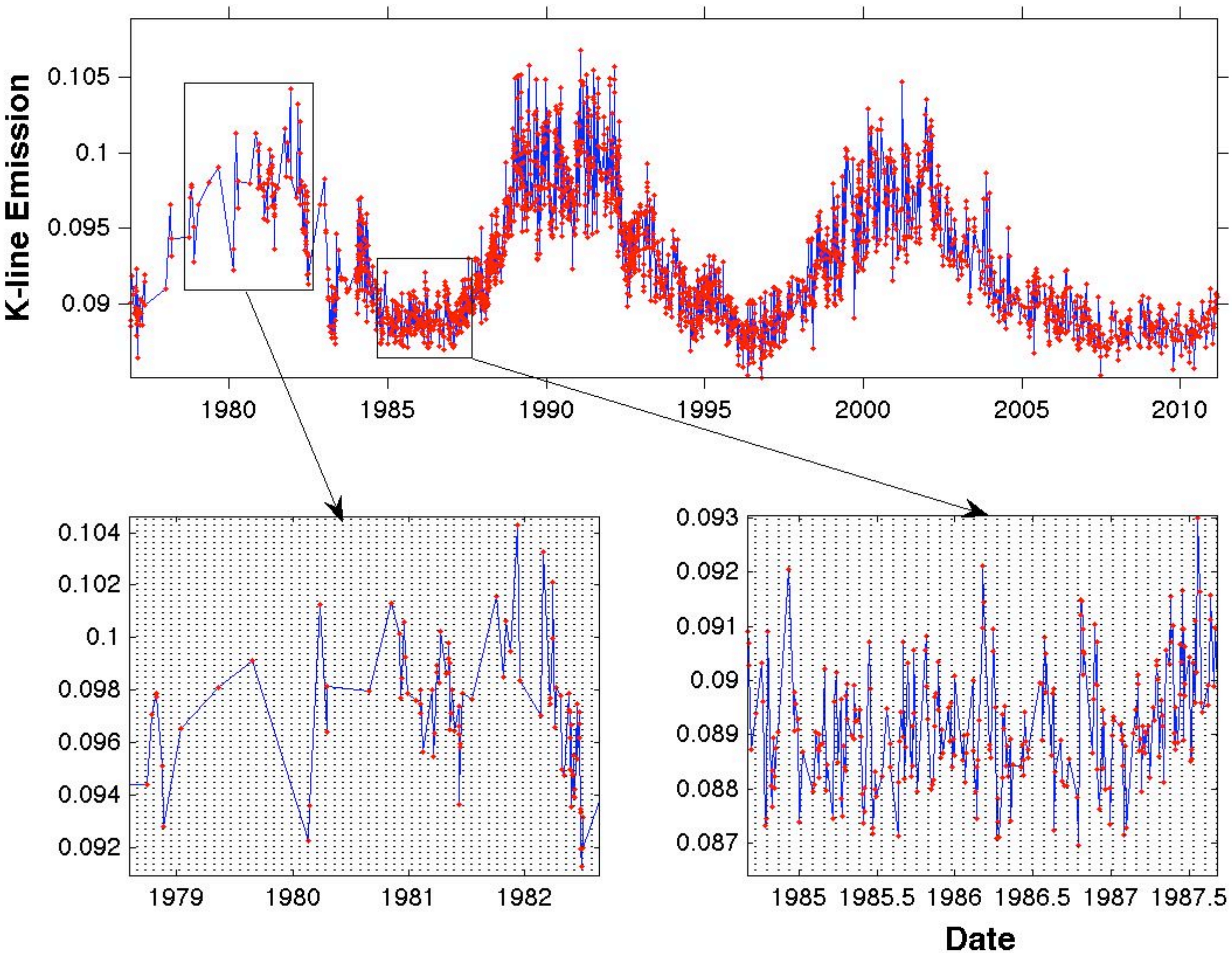


Lots of Data Gaps!
A few more
Log N
0
0.5
1
1.5
2
2.5
3
0
10
20
30
40
50
Interval size (days)

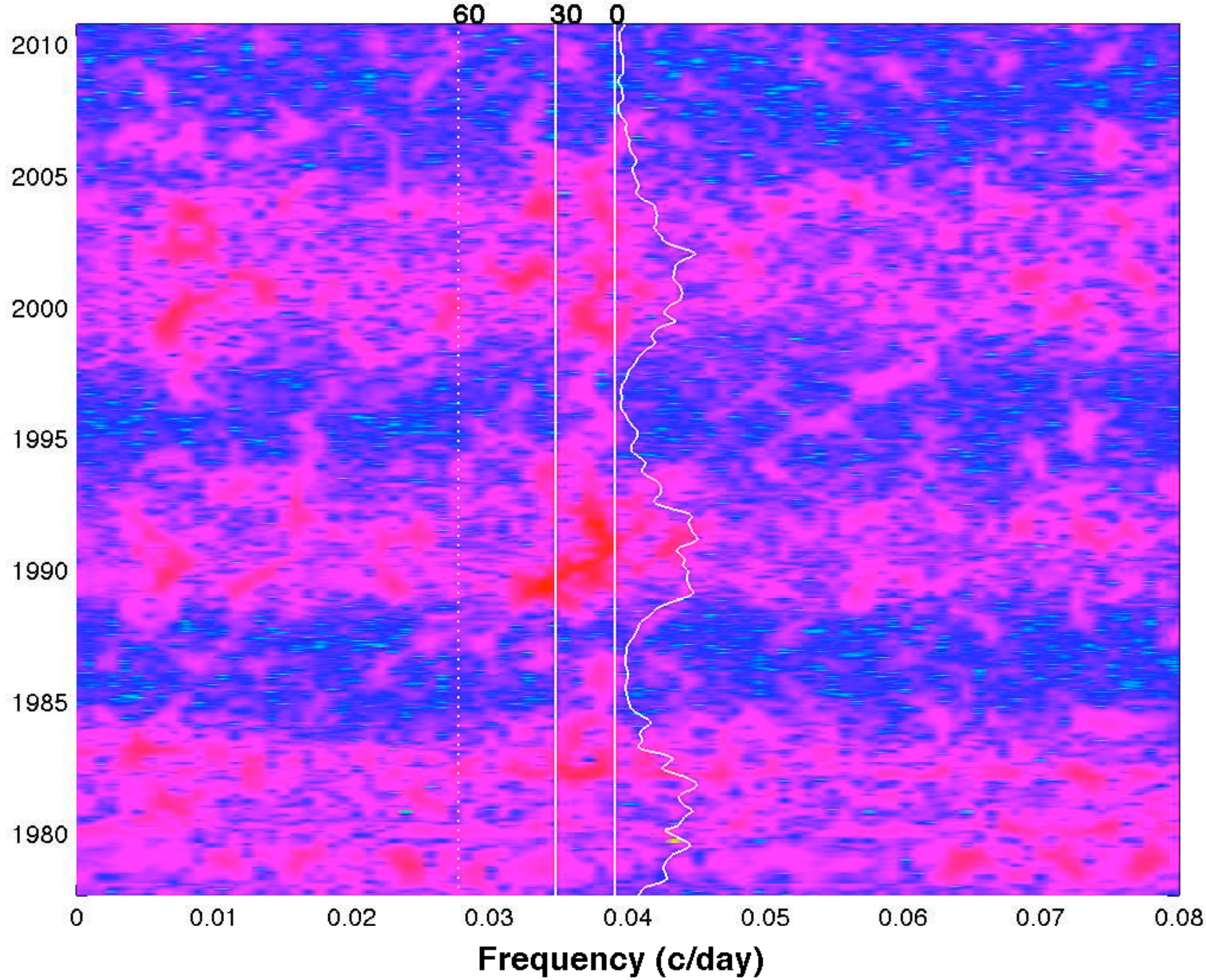

0.105
0.1
0.095
0.09
0.085
0.09
0.085
0.08
0.075
0.07
0.065
0.06
1980
1985
1990
1995
2000
2005
2010



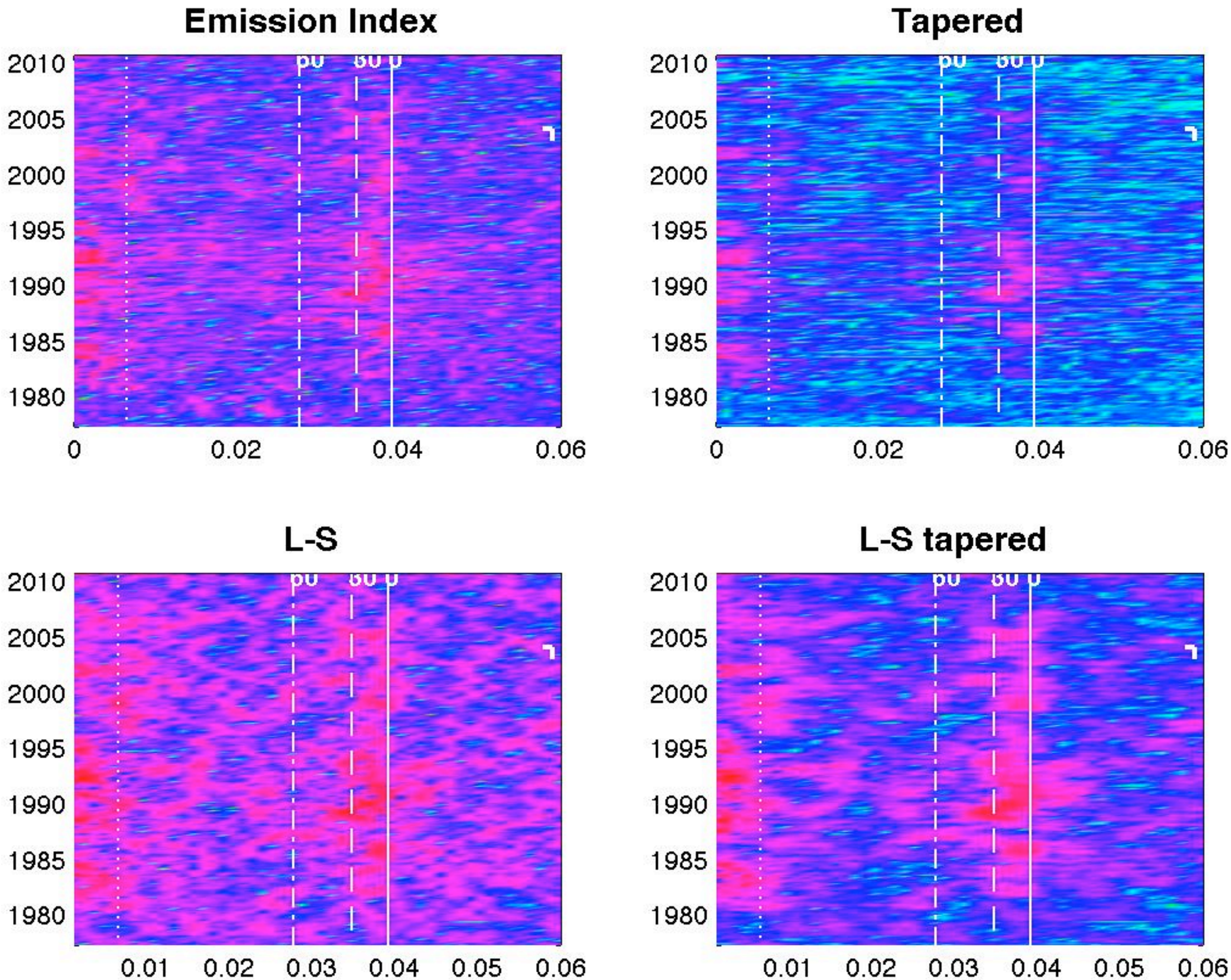
1
0.8
0.6
0.4
0.2
0
1975
1980
1985
1990
1995
2000
2005
2010
2015
Date (years)

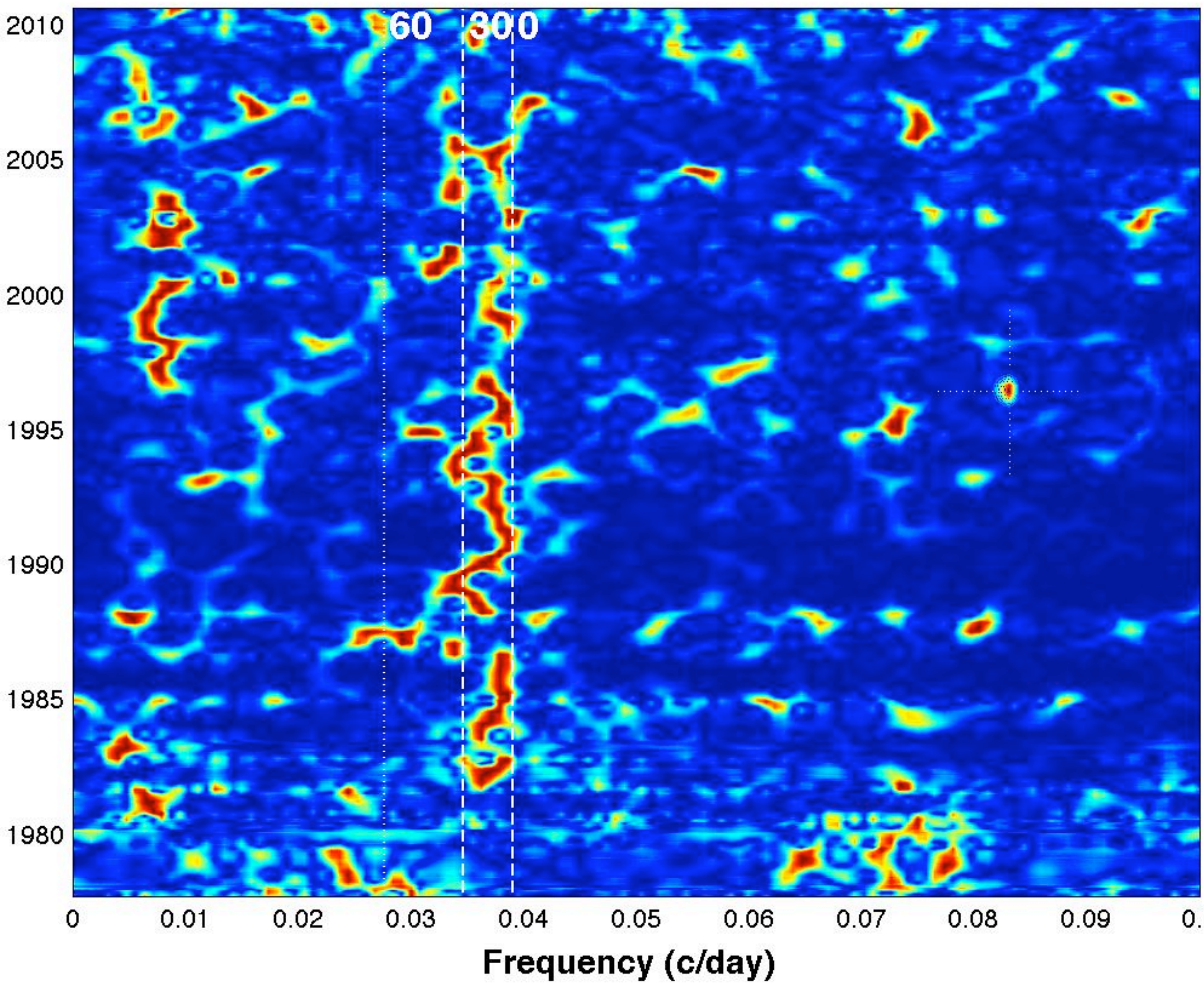


K-line Emission
0.105
0.1
0.095
0.09
1980
1985
1990
1995
2000
2005
2010
0.104
0.102
0.1
0.098
0.096
0.094
0.092
1979
1980
1981
1982
0.093
0.092
0.091
0.09
0.089
0.088
0.087
1985
1985.5
1986
1986.5
1987
1987.5
Date



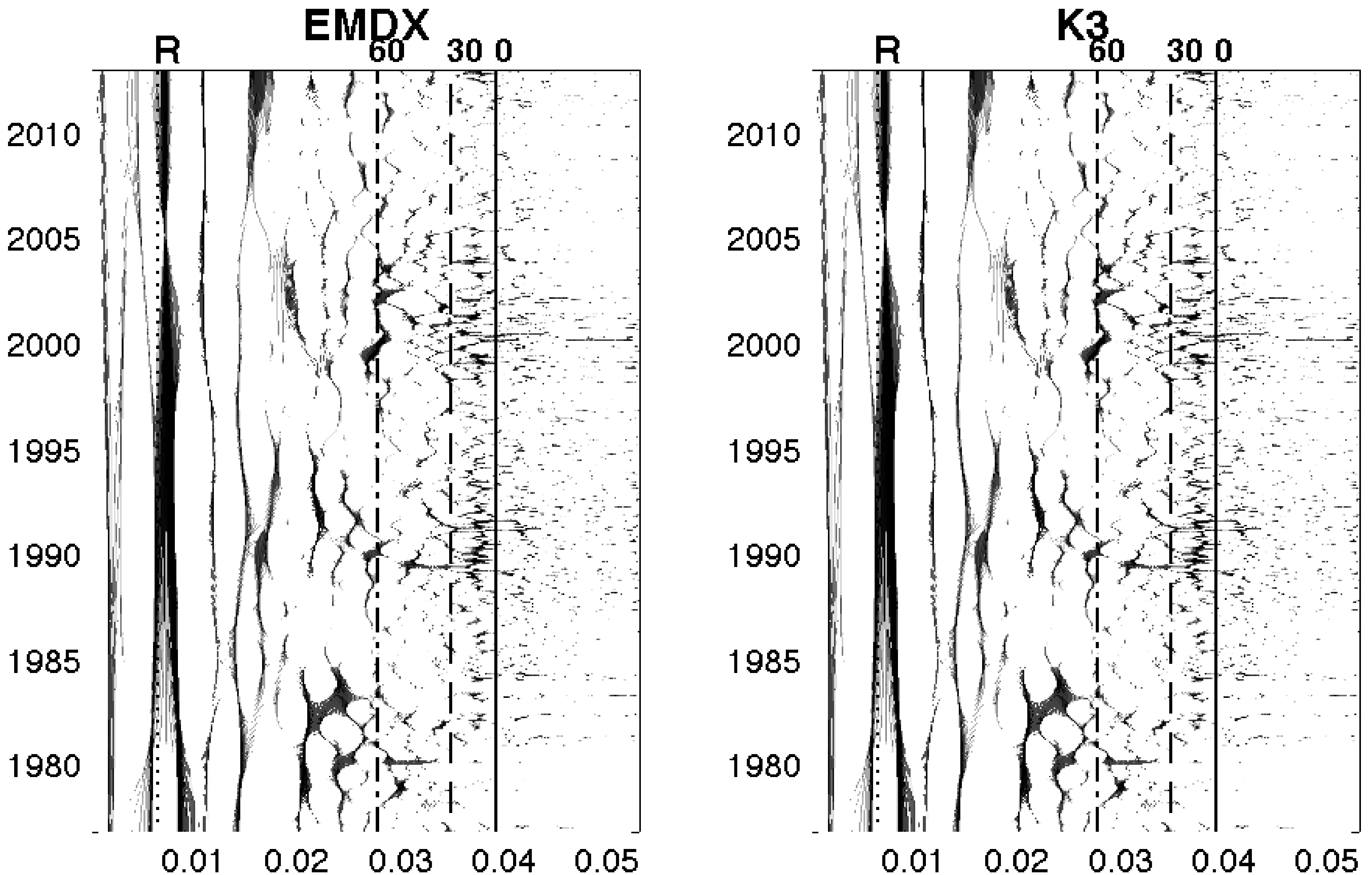
60
30
0
2010
2005
2000
1995
1990
1985
1980
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Frequency (c/day)


Emission Index
Tapered
L-S
L-S tapered
2010
2005
2000
1995
1990
1985
1980
0
0.01
0.02
0.03
0.04
0.05
0.06



60
30
0
2010
2005
2000
1995
1990
1985
1980
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1
Frequency (c/day)

Synchrosqueezing Algorithm: Brevdo and Daubechies

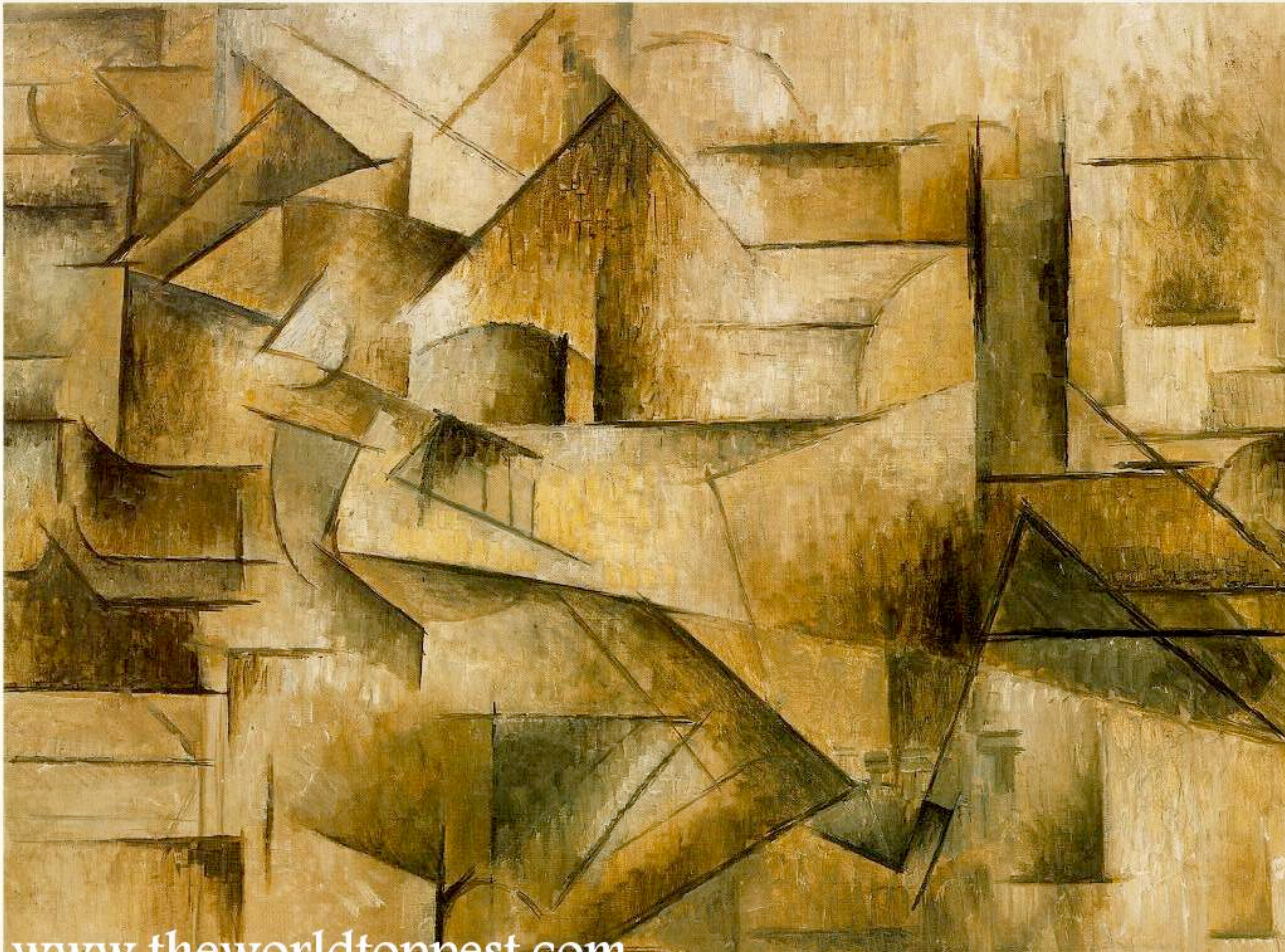


Processes Underlying K-line Variations

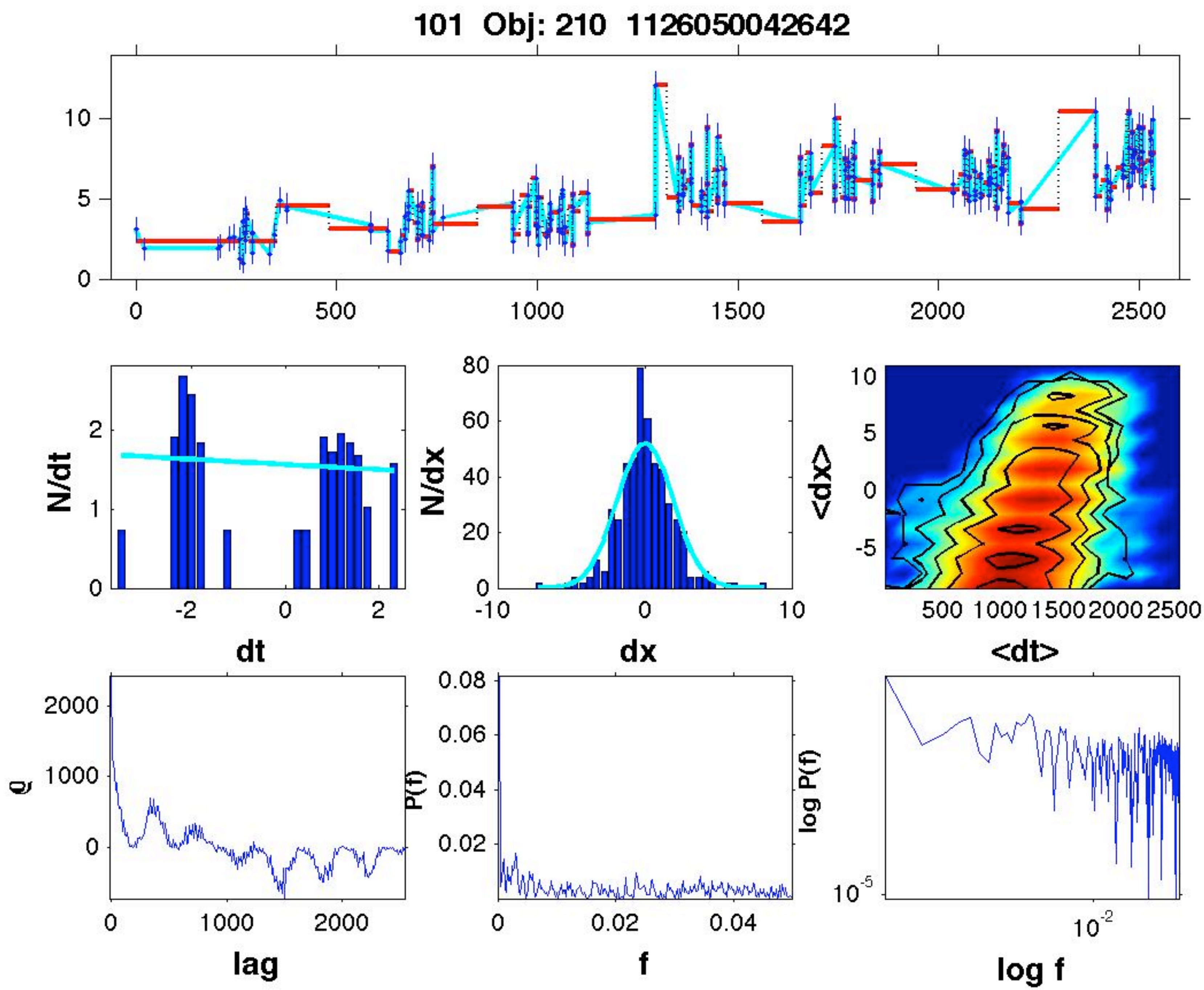
	Amplitude	Time Scale	Nature
Solar Cycle	Large	Long (~ 11 years)	Deterministic
Rieger Periodicity etc.	Medium	Intermediate (~ 100 days)	Quasi-Periodic
Flicker Noise	Small	Short (<~ year)	Random
(Differential) Rotational Modulation	Medium	Short (~ 27 days)	Quasi Periodic
Observation Errors	Small	Instantaneous	Random


www.theworldtoppest.com


www.theworldtoppest.com


www.theworldtoppest.com

Function	Domain	Range	Auto-	Cross-	Physical Interp
Bayesian blk. Light Curve	Time	Flux	✔	✔ multivar. BB	Flares, events etc.
Scatter Plot	Flux 1	Flux 2		✔	Dependency (not just cor.)
Correlation	Lag	<X^2> <XY>	✔	✔	Correlated behavior/lags
Spectrum	Frequency	Power	✔	✔	Periodicity 1/f noise …
		Phase	✔	✔	Shifts, lags
Structure	Lag	<X^2> <XY>	✔	✔	Correlated behavior/lags
Scalogram	Scale/Time	Power	✔	✔	Dynamic behavior
Scalegram	Scale	Power	✔	✔	1/f noise QPOs
Distribution	Time/scale/ frequency	Power	✔	✔	Dynamic behavior



101 Obj: 210 1126050042642
N/dt
dt
N/dx
dx
<dx>
<dt>
lag
P(f)
f
log P(f)
log f

Bayesian Blocks Bibliogrpahy

Studies in Astronomical Time Series Analysis.
VI. Bayesian Block Representations
Scargle, J., Norris, J., Jackson, B. and Chiang, J.
The Astrophysical Journal, 764, 167 (2013)
Our Blog: http://bayesianblocks.blogspot.com/

Jake Vanderplas' Blog ***Dynamic Programming in Python: Bayesian Blocks***
http://jakevdp.github.com/blog/2012/09/12/dynamic-programming-in-python/

Starship Asterisk* APOD and General Astronomy Discussion Forum
Bayesian Blocks: Detecting local variability in time series
http://asterisk.apod.com/viewtopic.php?f=35&t=29458

An algorithm for optimal partitioning of data on an interval
Jackson, Scargle, Barnes, Arabhi, Gioumousis, Gwin, Sangtrakulcharoen, Tan, Tun Tao Tsai
IEEE Signal Processing Letters, 2005, 12, 105

Studies in Astronomical Time Series Analysis.
V. Bayesian Blocks, a New Method to Analyze Structure in Photon Counting Data
Scargle, 1998, Astrophysical Journal, 504, 405