Time Series Analysis for the Multi-wavelength future

Tiana_Athriel
dhuppenkothen

Daniela Huppenkothen
NYU Center for Cosmology and Particle Physics
NYU Center for Data Science

How do we convert data into knowledge?

time series

How do we convert ~~data~~ into ~~knowledge~~?

black hole physics

1) How can we detect variability?

2) How can we classify variable sources?

3) How can we characterize variability?

machine learning

1) How can we detect variability?

2) How can we classify variable sources?

3) How can we characterize variability?

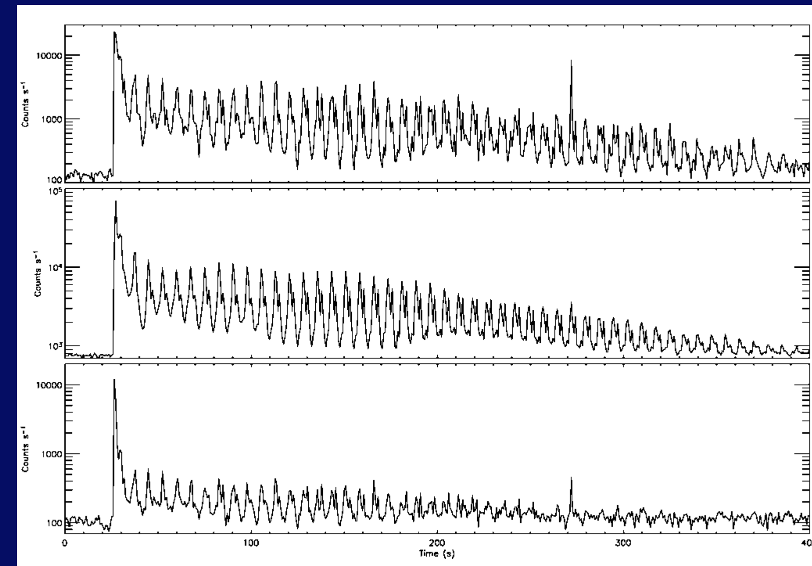
machine learning

1) How can we detect variability?

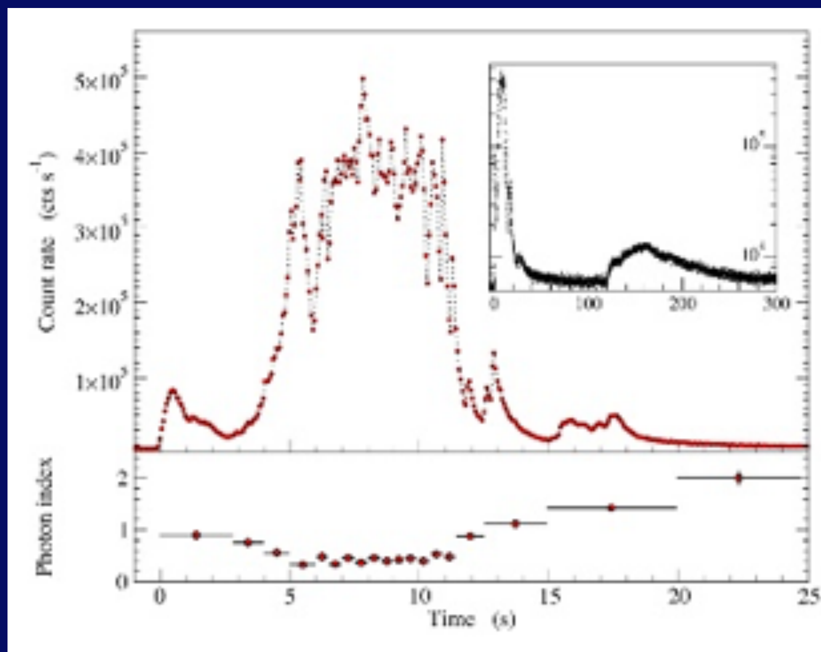
2) How can we classify variable sources?

3) How can we characterize variability?

Transients

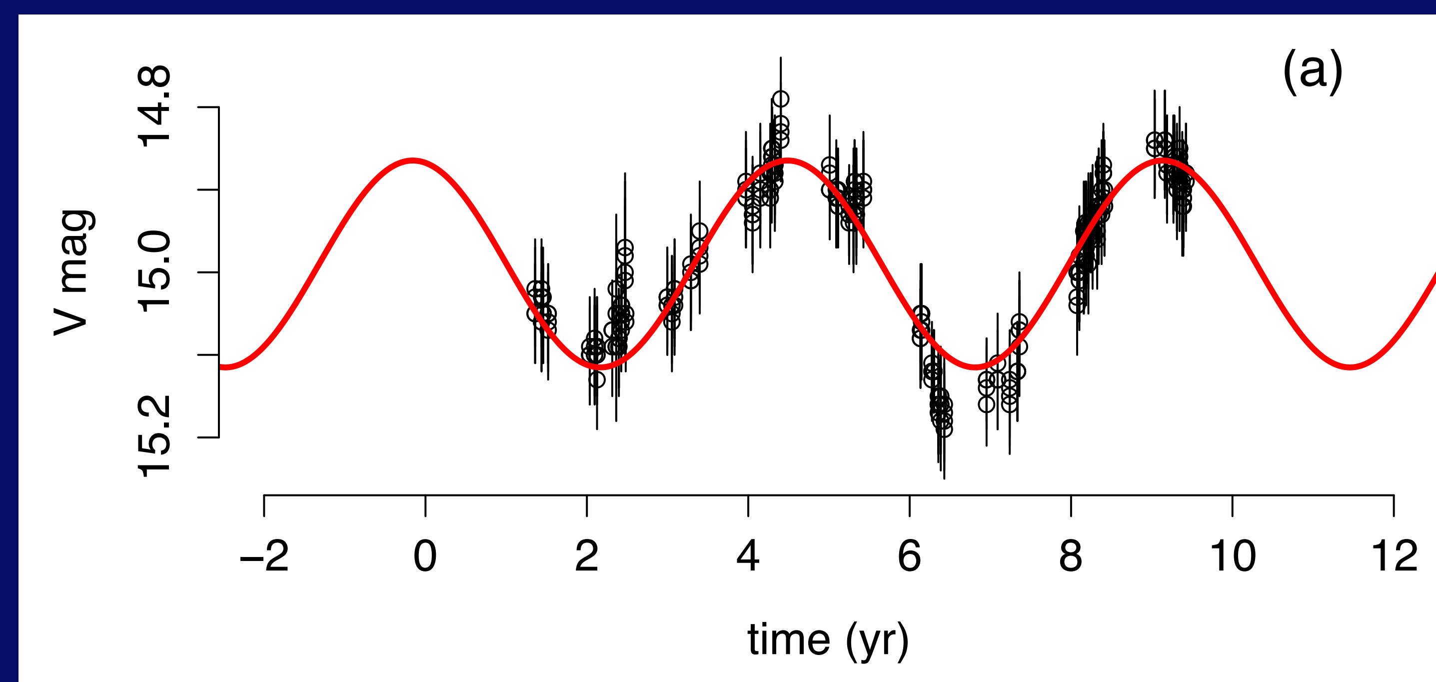


Watts & Strohmayer, 2005



Maselli et al, 2013

Vaughan et al, 2016

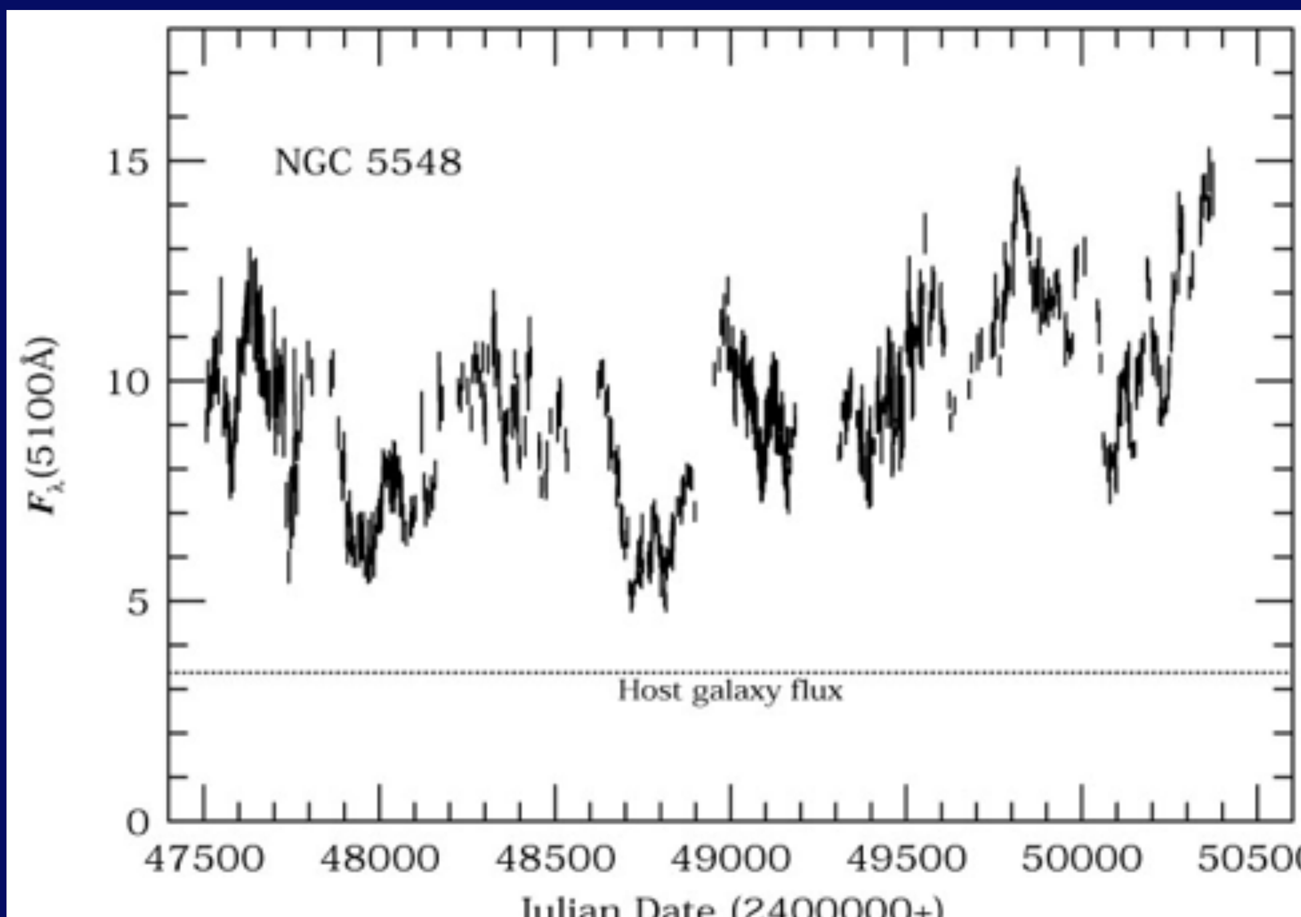


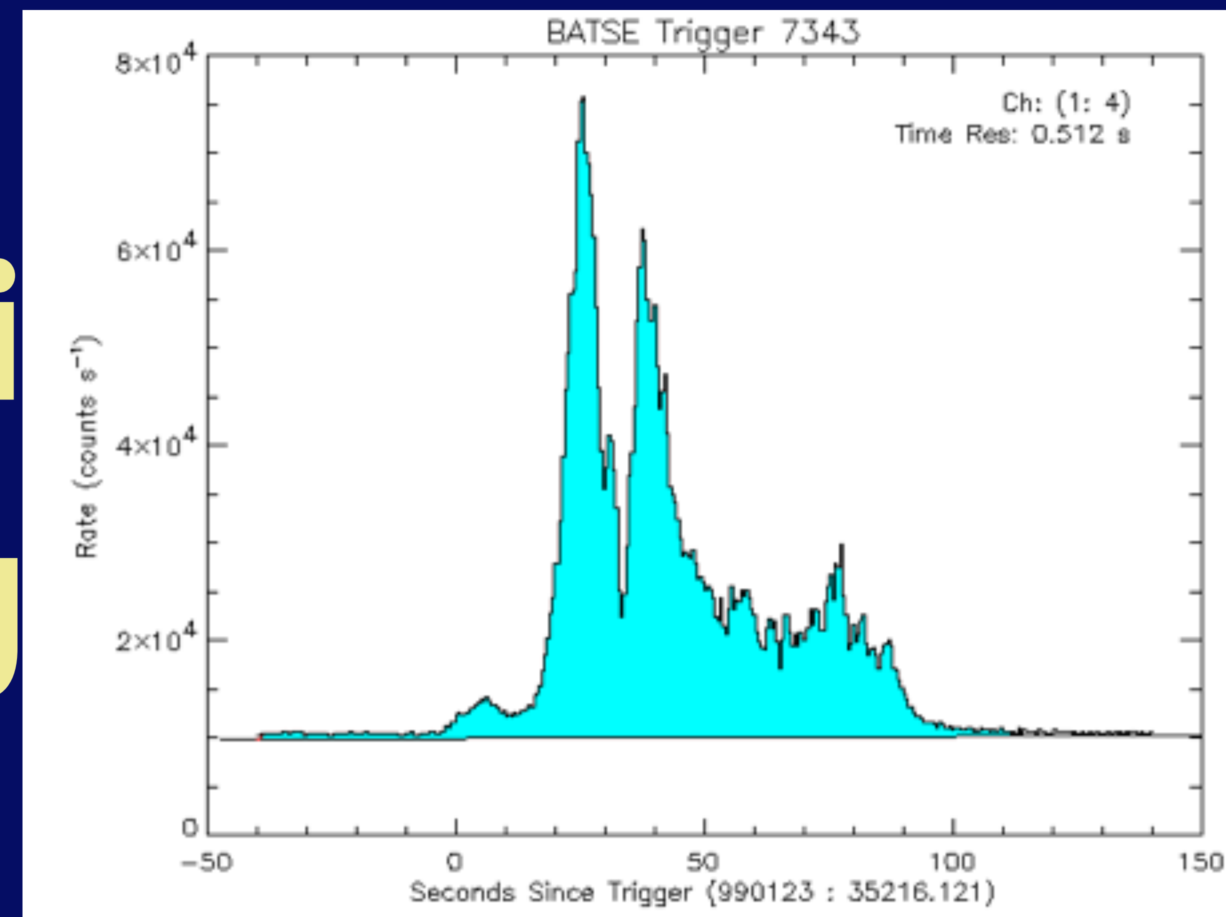
variable sources

Stationarity vs Non-Stationarity

“The joint probability distribution does not change over time.”

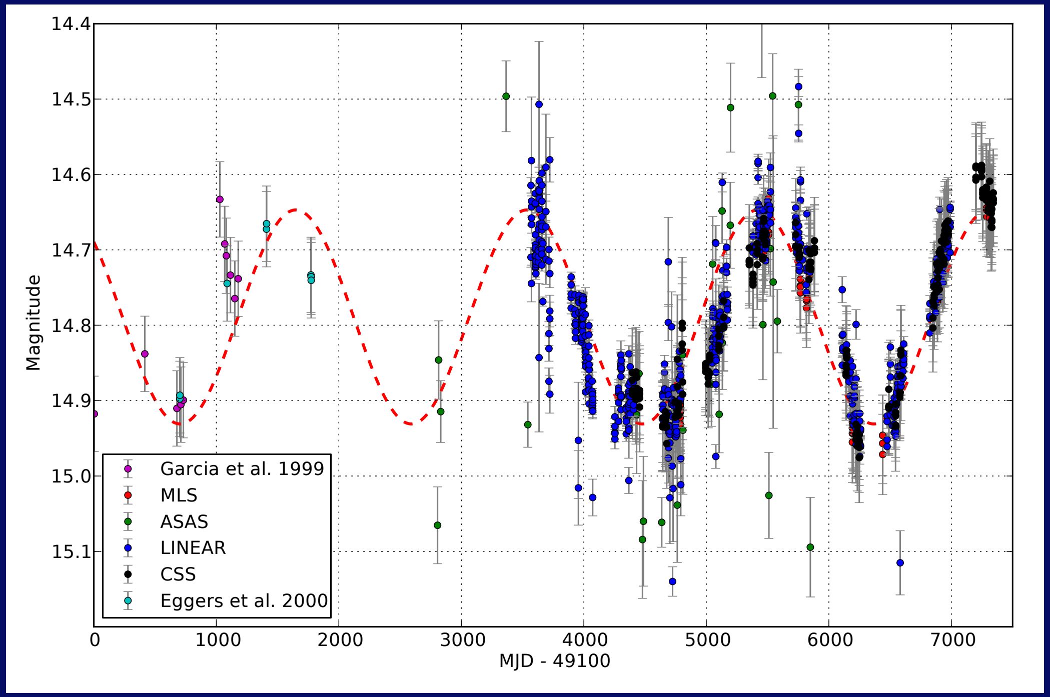
Stationarity vs Non-Stationarity



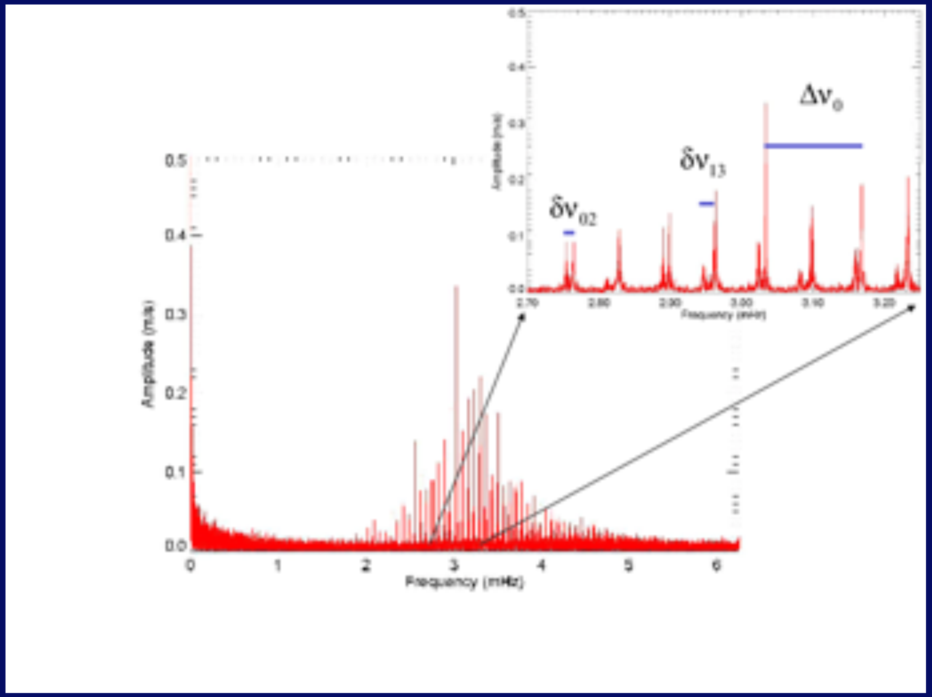


NGC 5548:
Peterson et al (1999)

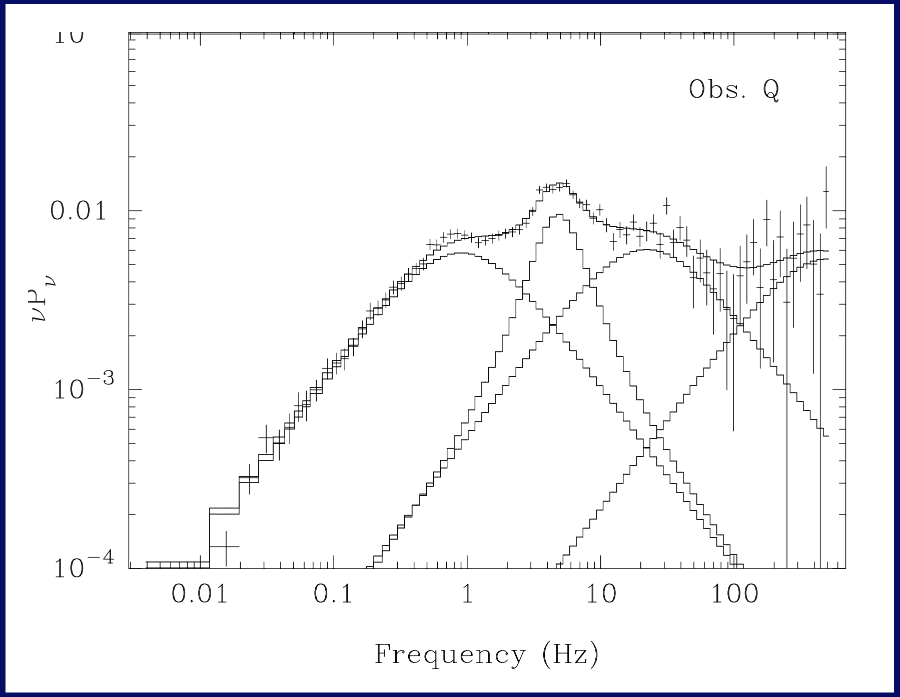
Periodicity Detection



Graham et al, 2015



Kjeldsen et al, 2009



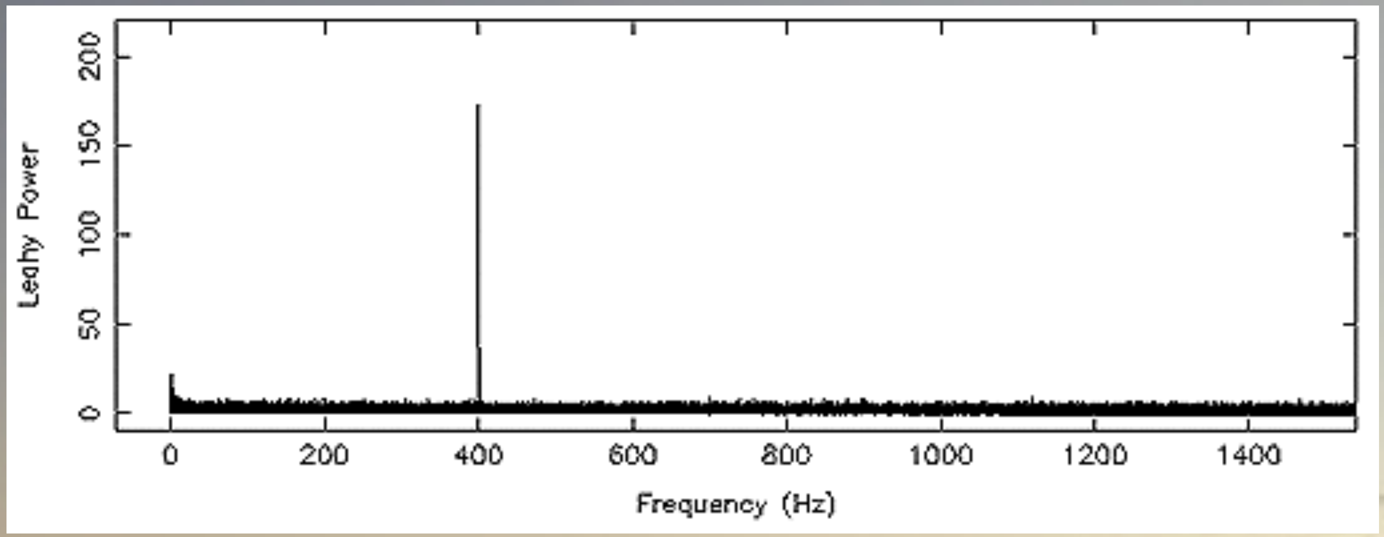
Belloni et al, 2002

Periodicity Detection I

1) Fourier Analysis
2) Lomb-Scargle Periodogram
3) Z^2 statistic, Rayleigh statistic, …
4) Gregory & Loredo 1992
5) Wavelets
6) …

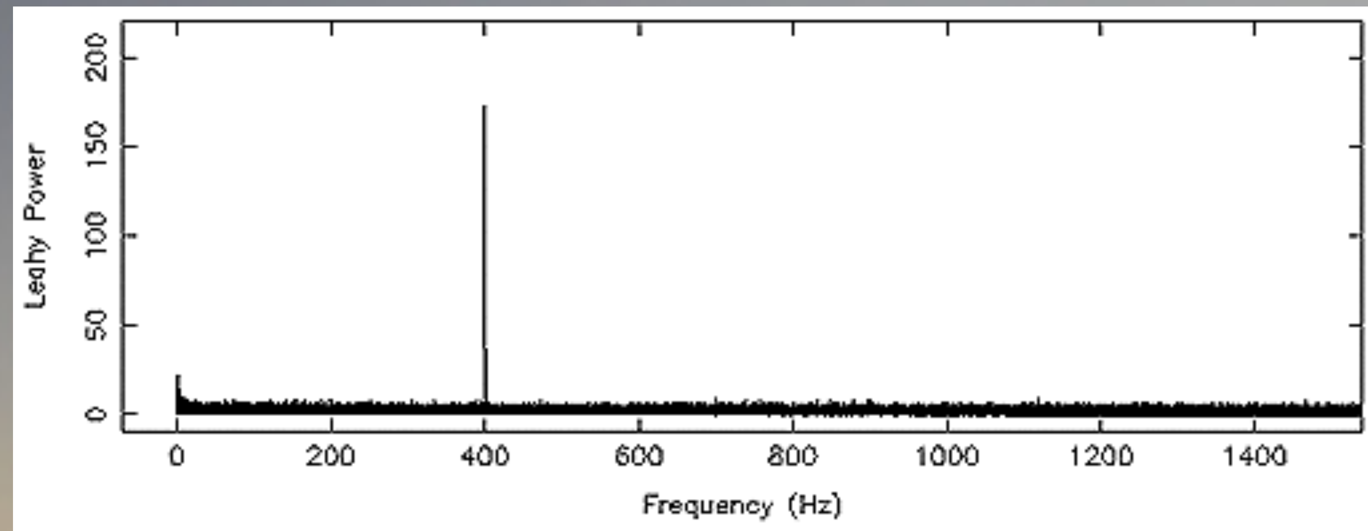
Almost all algorithms used for periodicity detection assume stationary processes and/or white noise

Wijnands + van der Klis 1998



credit: Richard Freeman,
flickr.com/photos/freebird710/, CC licensed

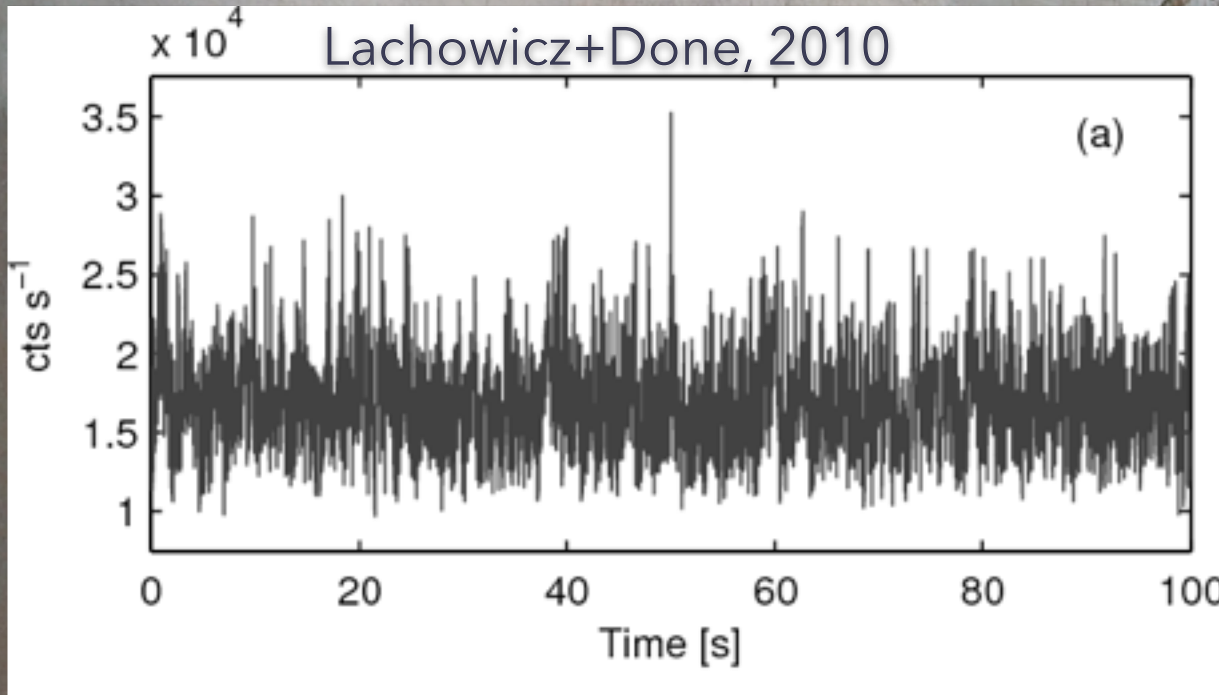
Wijnands + van der Klis 1998



constant background

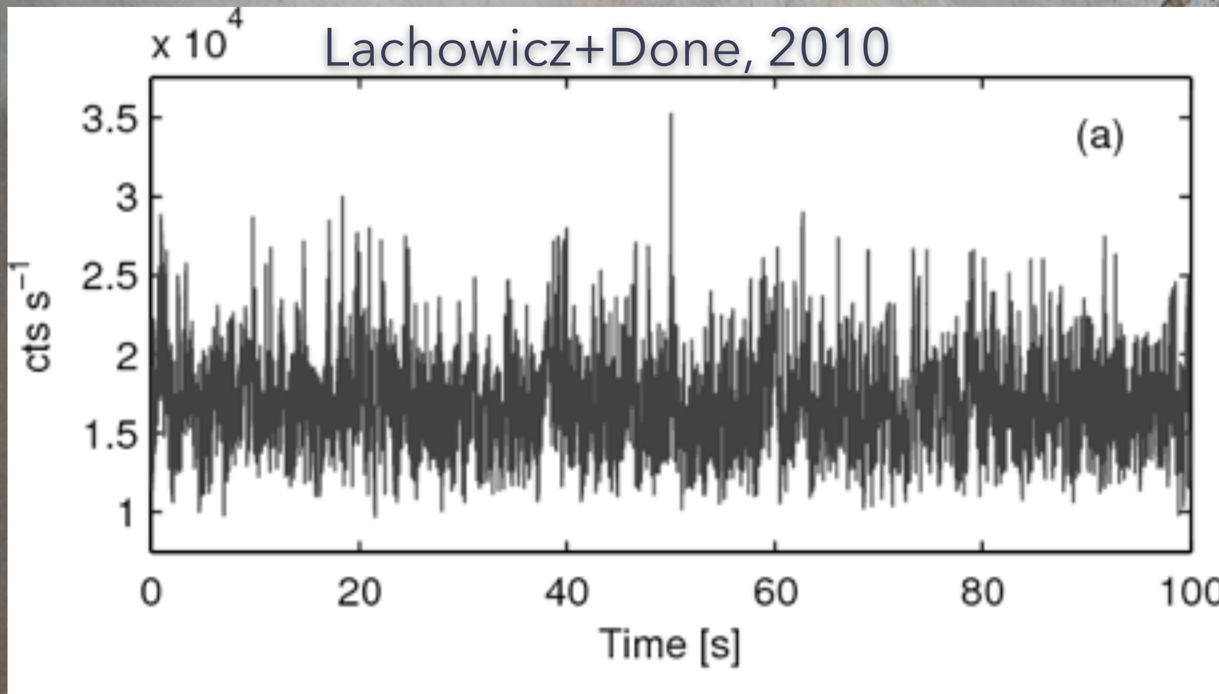
credit: Richard Freeman,
flickr.com/photos/freebird710/, CC licensed

Willem van de Velde the Younger, “The Gust”

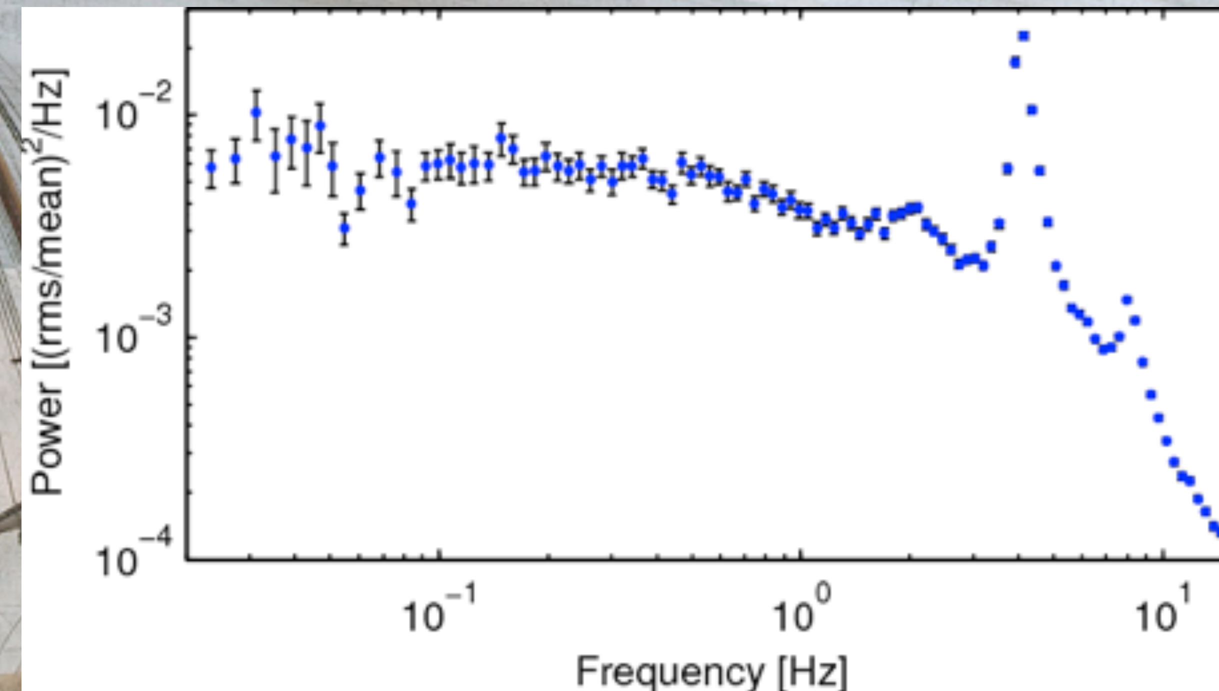


Lachowicz+Done, 2010

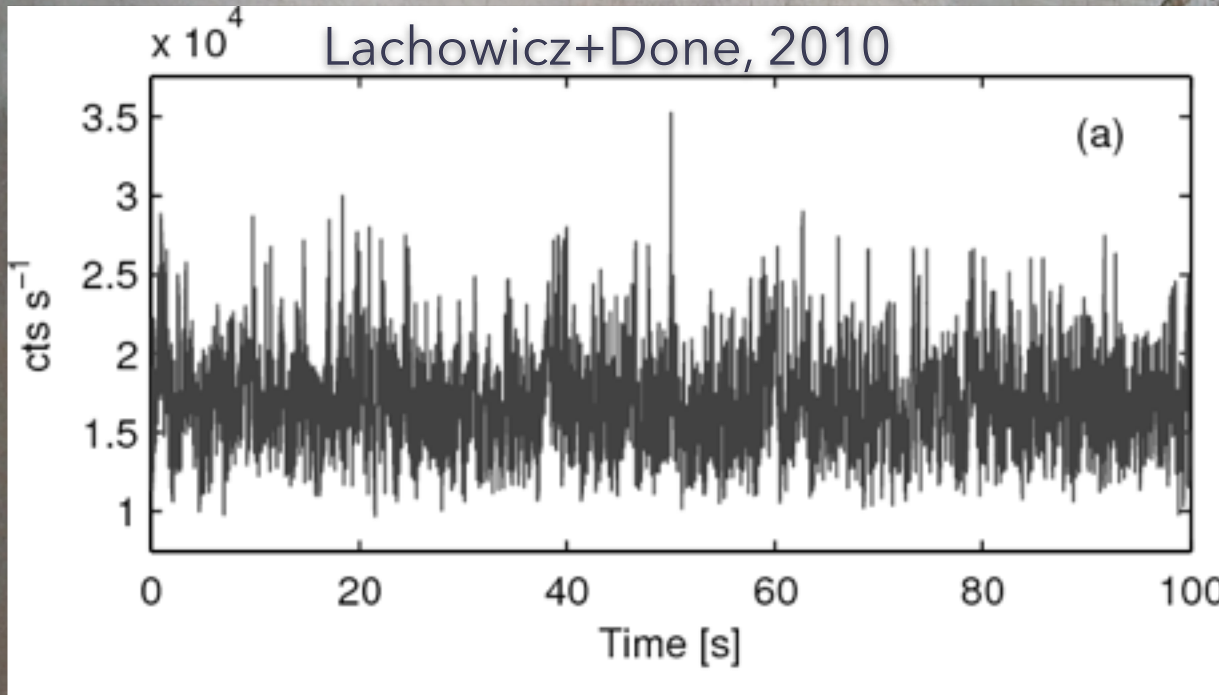
Willem van de Velde the Younger, “The Gust”

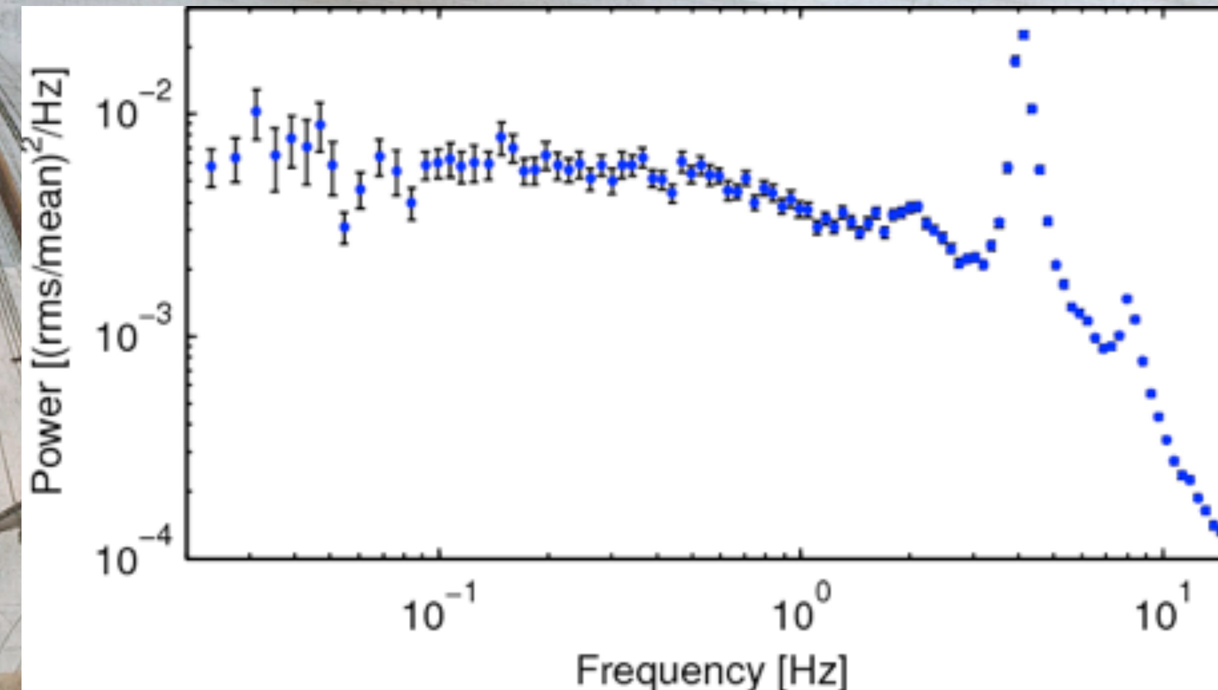


Lachowicz+Done, 2010



Willem van de Velde the Younger, “The Gust”





variable background!

Willem van de Velde the Younger, “The Gust”

The human brain is awesome at pattern recognition

*see: pareidolia

Corollary: the human brain is prone to overfitting* + see spurious patterns

(awesome for survival, less awesome for science) *see: pareidolia

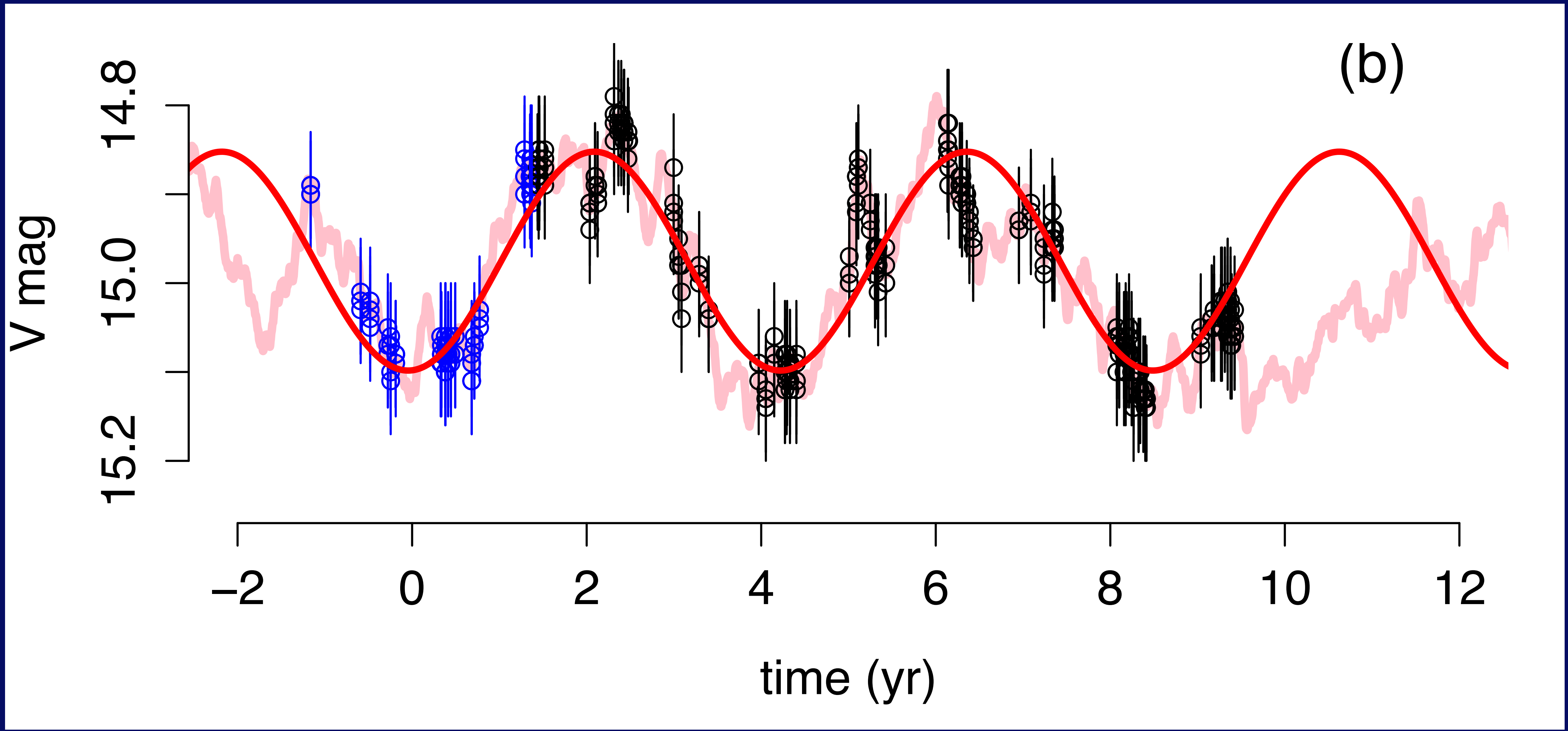
e hu

verfi

us p

By Viking 1, NASA [Public domain], via Wikimedia Commons

(awesome for survival, less awesome for science)



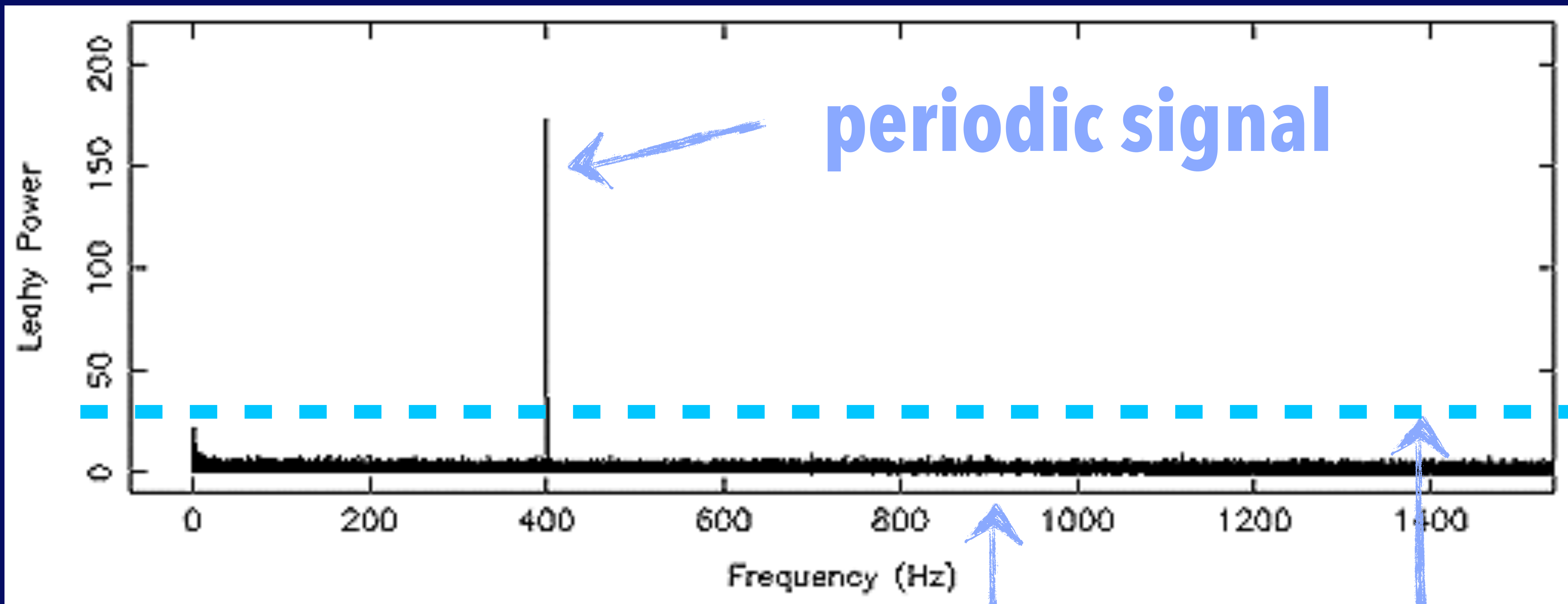
Vaughan et al, 2016

Hypothesis Testing

Hypothesis Testing

"What is the probability that I observe a frequency with power P_{obs} or higher, if my data is pure photon noise?"

Hypothesis Testing



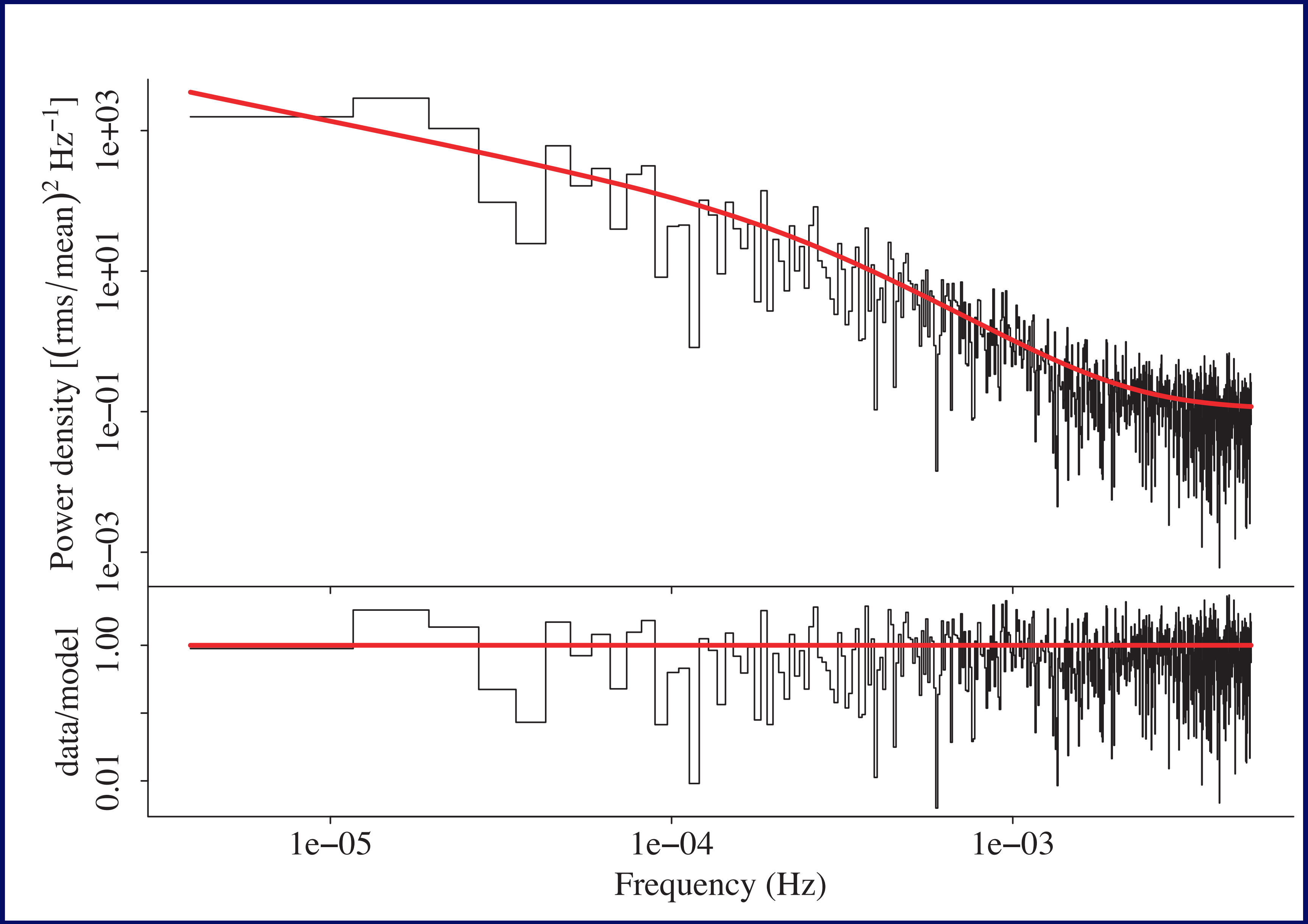
SAX J1808.4-3658:
Wijnands + van der Klis (1998)

Noise

Detection threshold

*unevenly sampled data: Lomb-Scargle periodogram

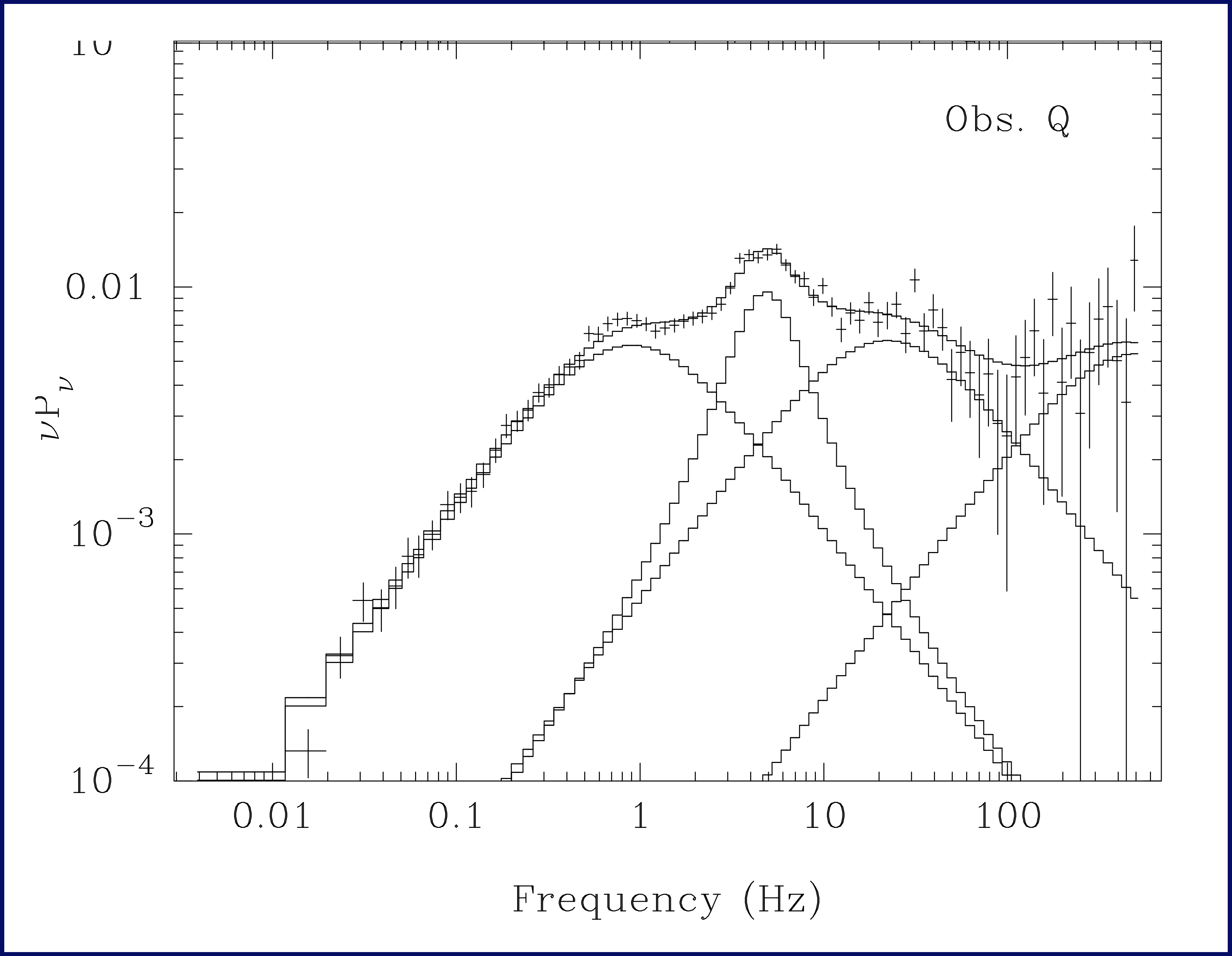
Hypothesis Testing II



Vaughan 2010

propagate model uncertainty!

Fit a Model to the Power Spectrum



Belloni et al, 2002

… but all my data is unevenly sampled!

Fit a Model to the Covariance

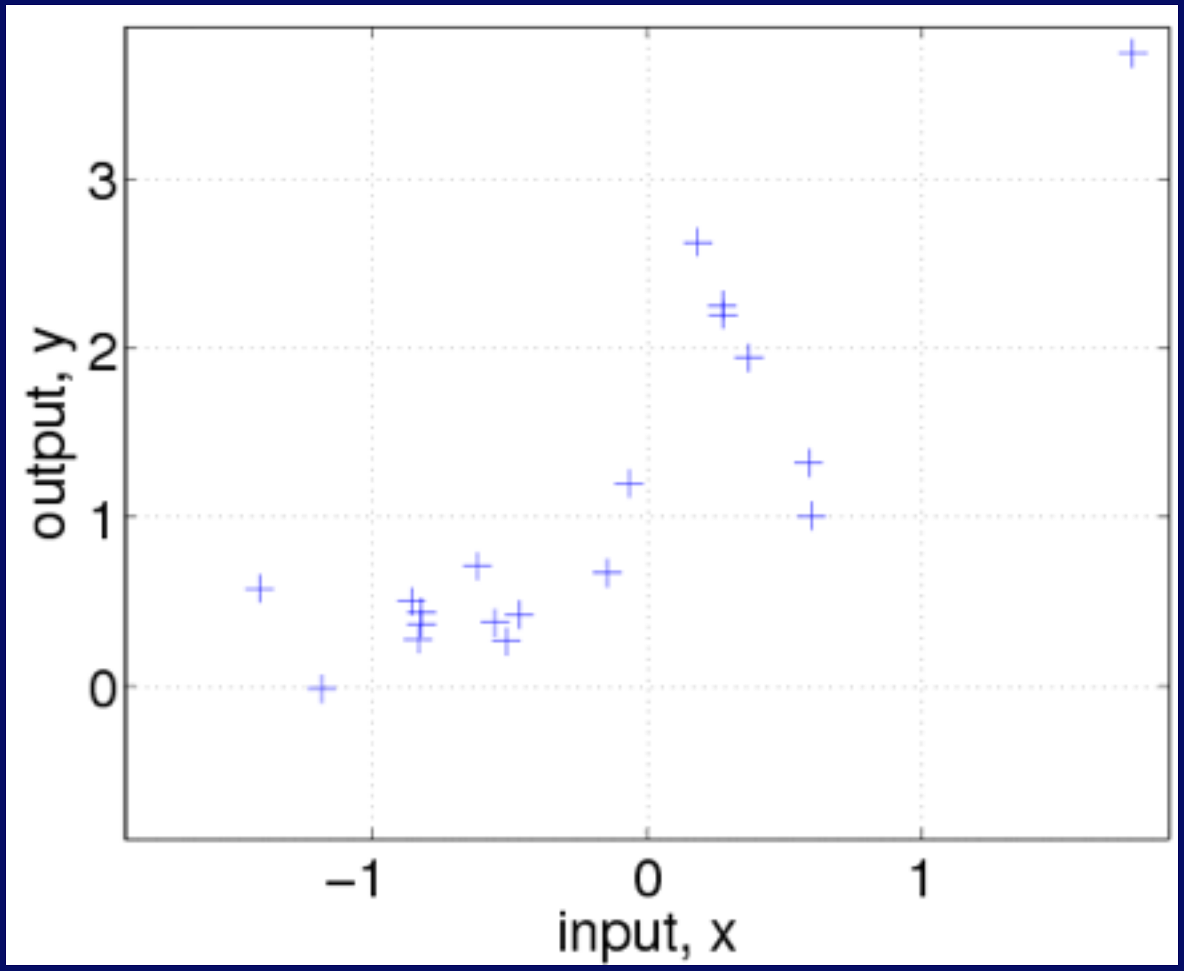
Gaussian Covariance

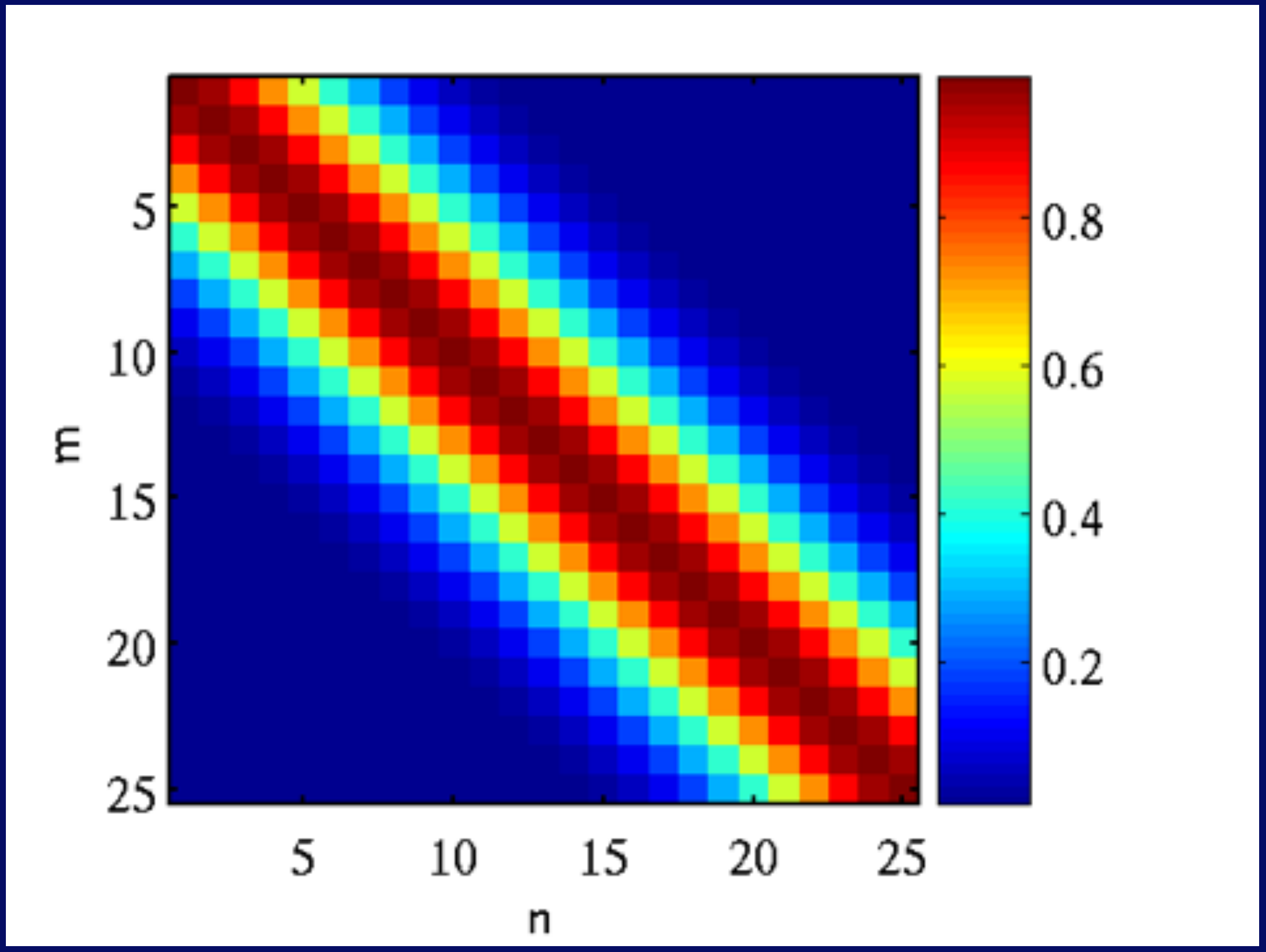
$$\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{f}_{\boldsymbol{\theta}}(t),\, K_{\boldsymbol{\alpha}}(t))$$

Mean

Image Credit: Dan Foreman-Mackey

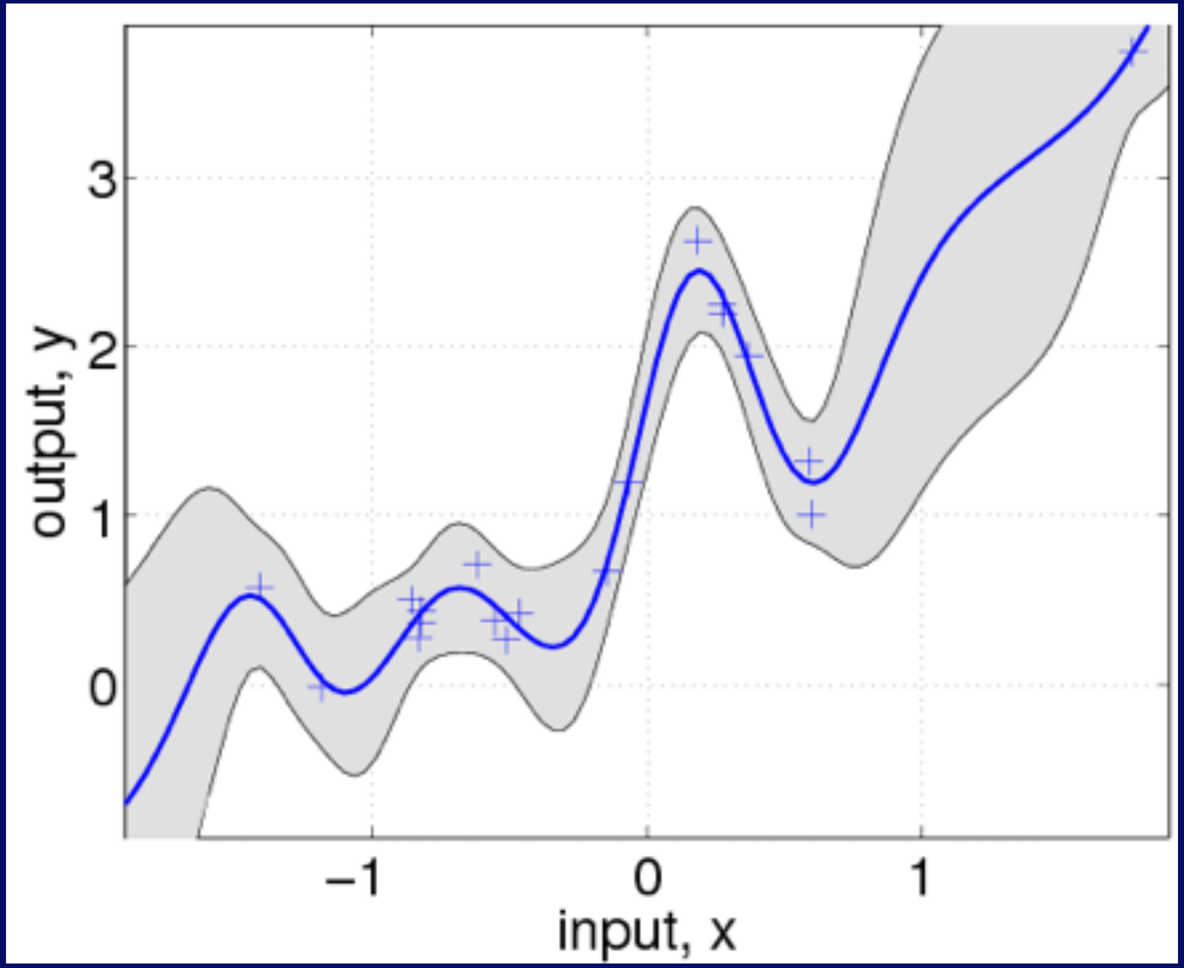
Fit a Model to the Covariance

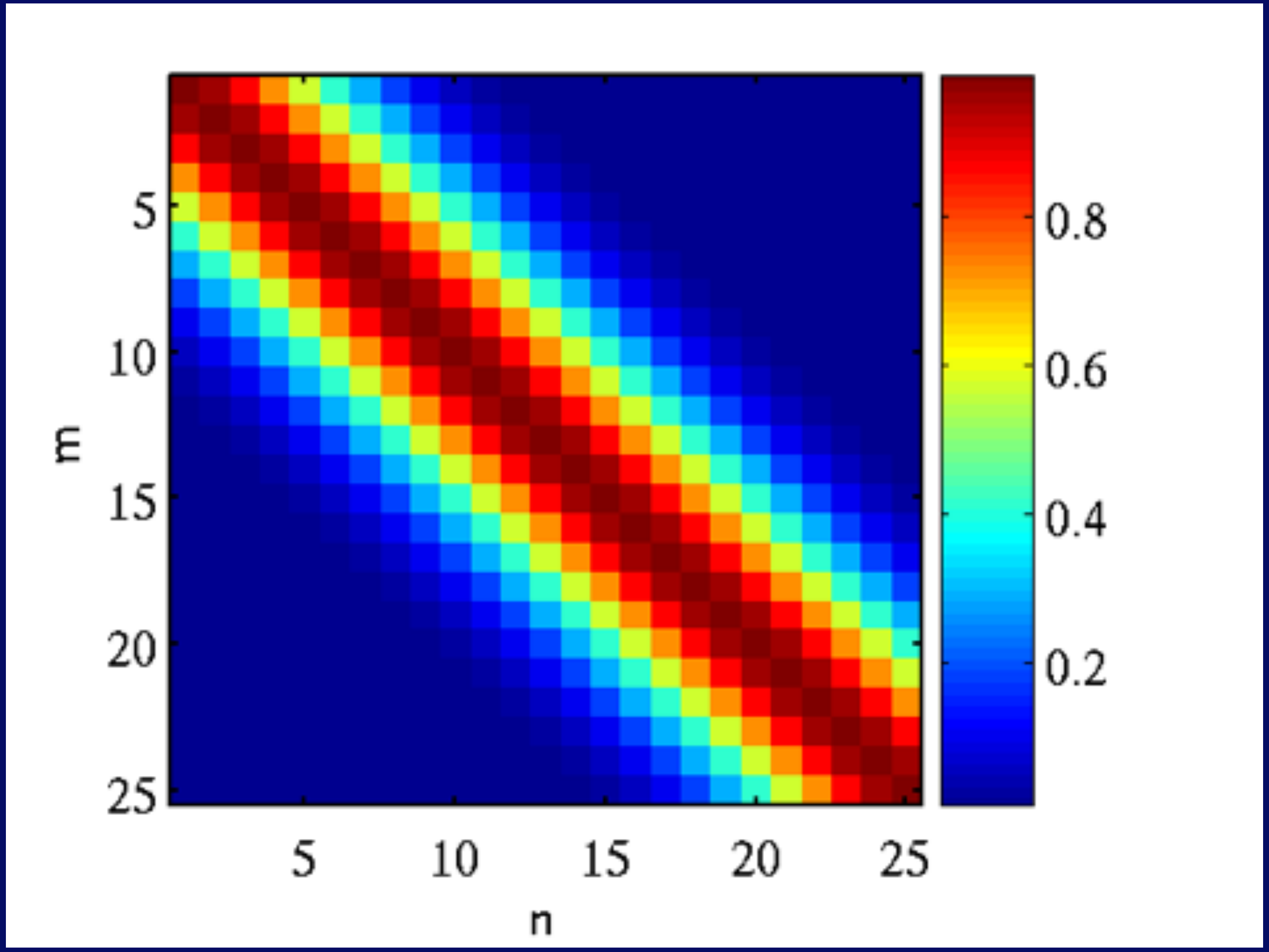




Rasmussen + Williams

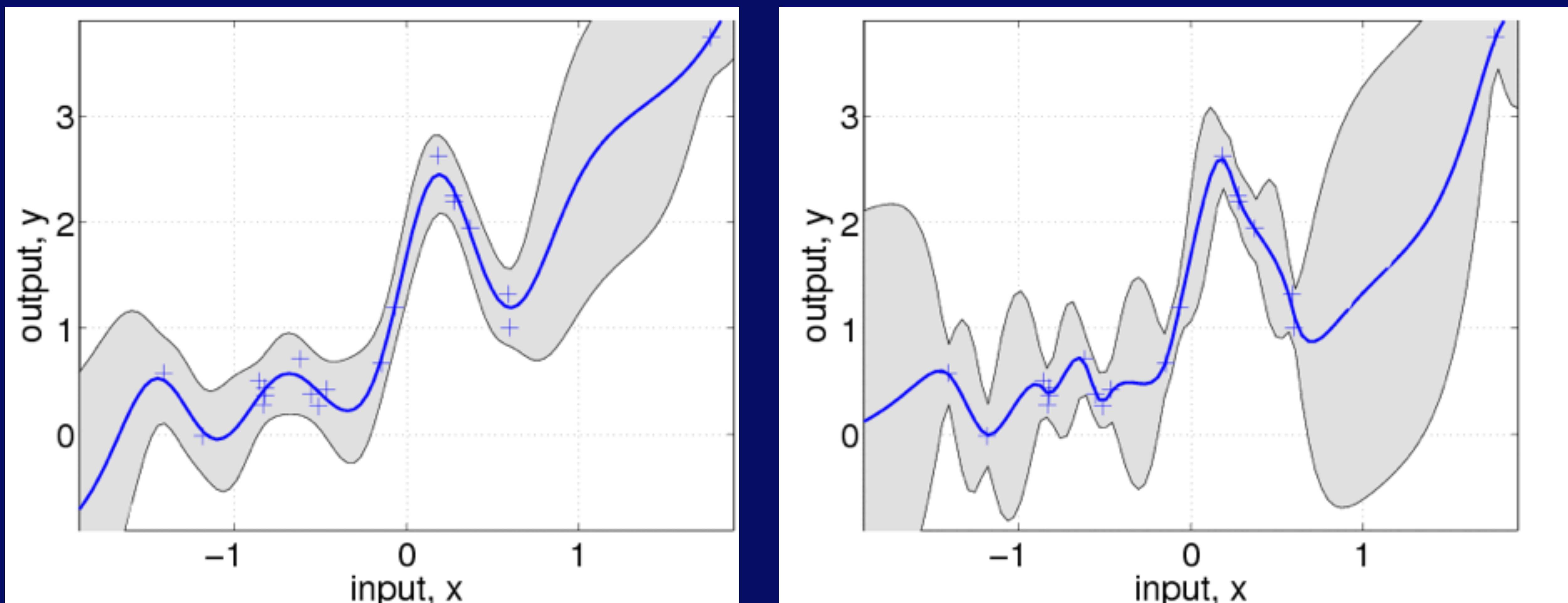
Fit a Model to the Covariance





Rasmussen + Williams

Fit a Model to the Covariance



Rasmussen + Williams

+ works for stochastic processes

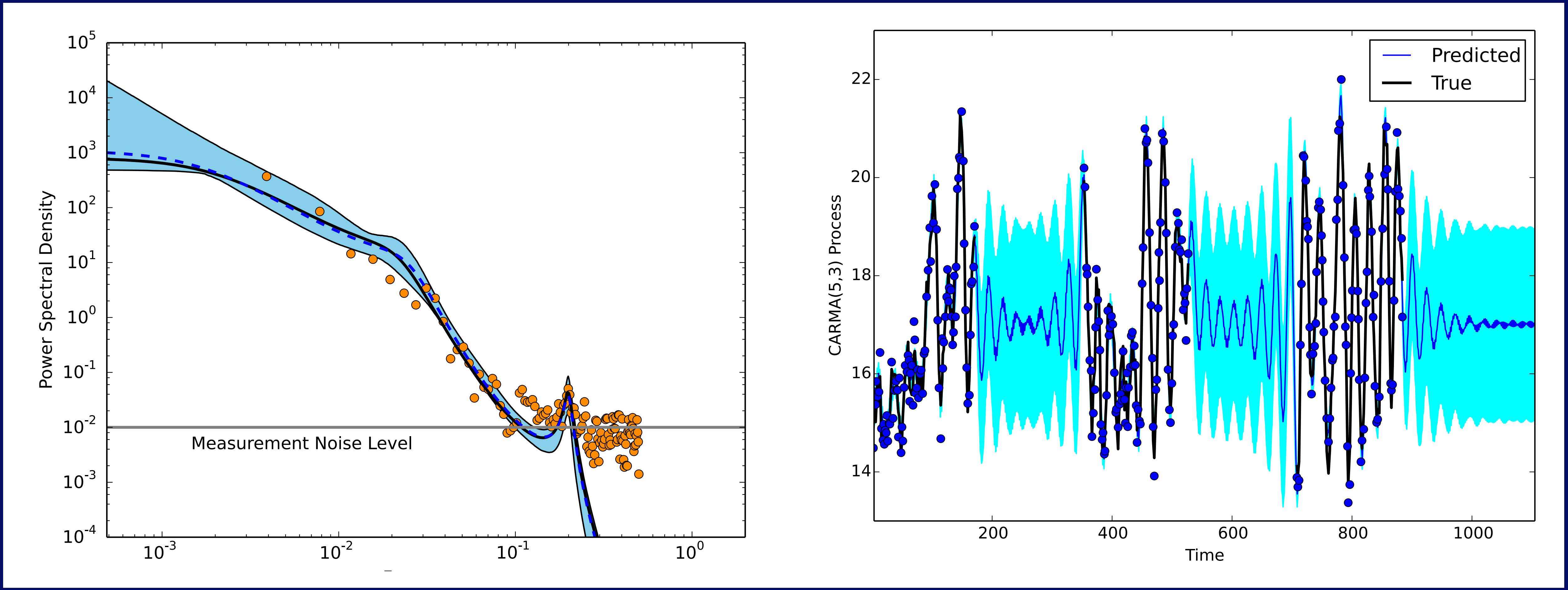
+ probabilistic generative model

– computationally expensive

CARMA

Continuous-time AutoRegressive Moving Average Processes

CARMA

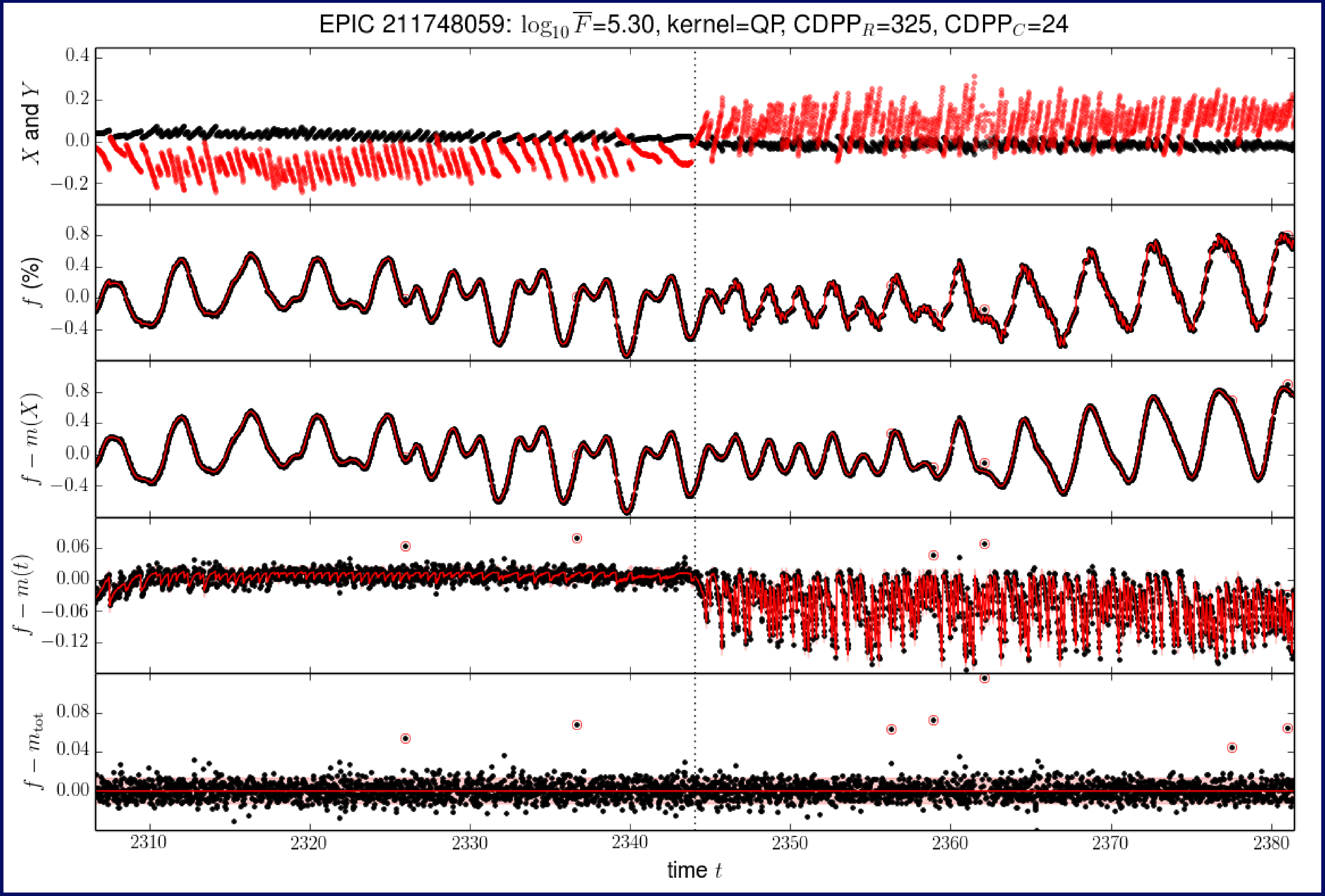


Kelly et al, 2014

CARMA

+ fast
– assumes specific underlying process

Beyond CARMA: Gaussian Processes



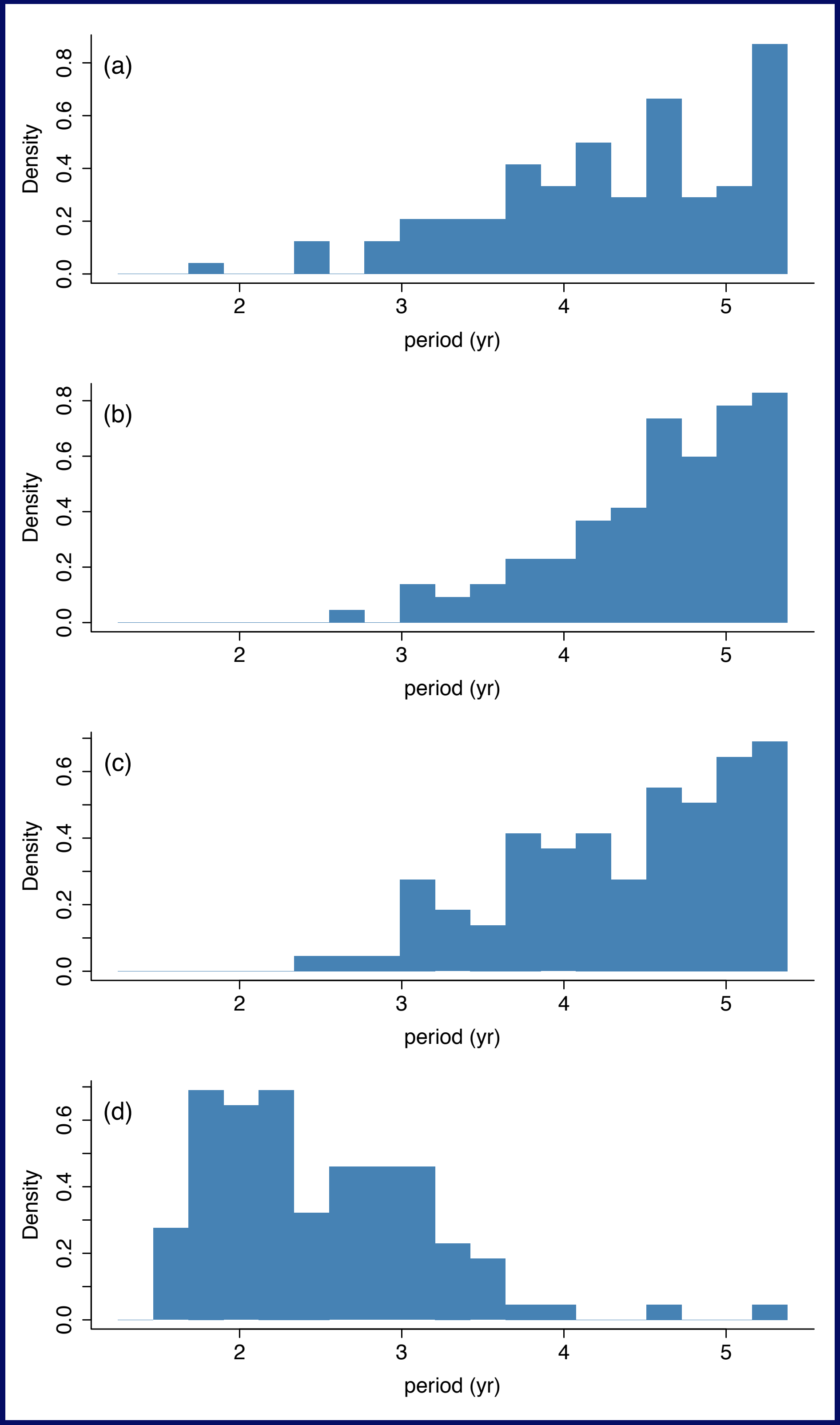
Aigrain et al, 2016

Sample Studies

easier to "go wide" than "go deep"

Sample Studies

Vaughan et al, 2016



Graham et al, 2015 sample

DRW simulations

red noise simulations

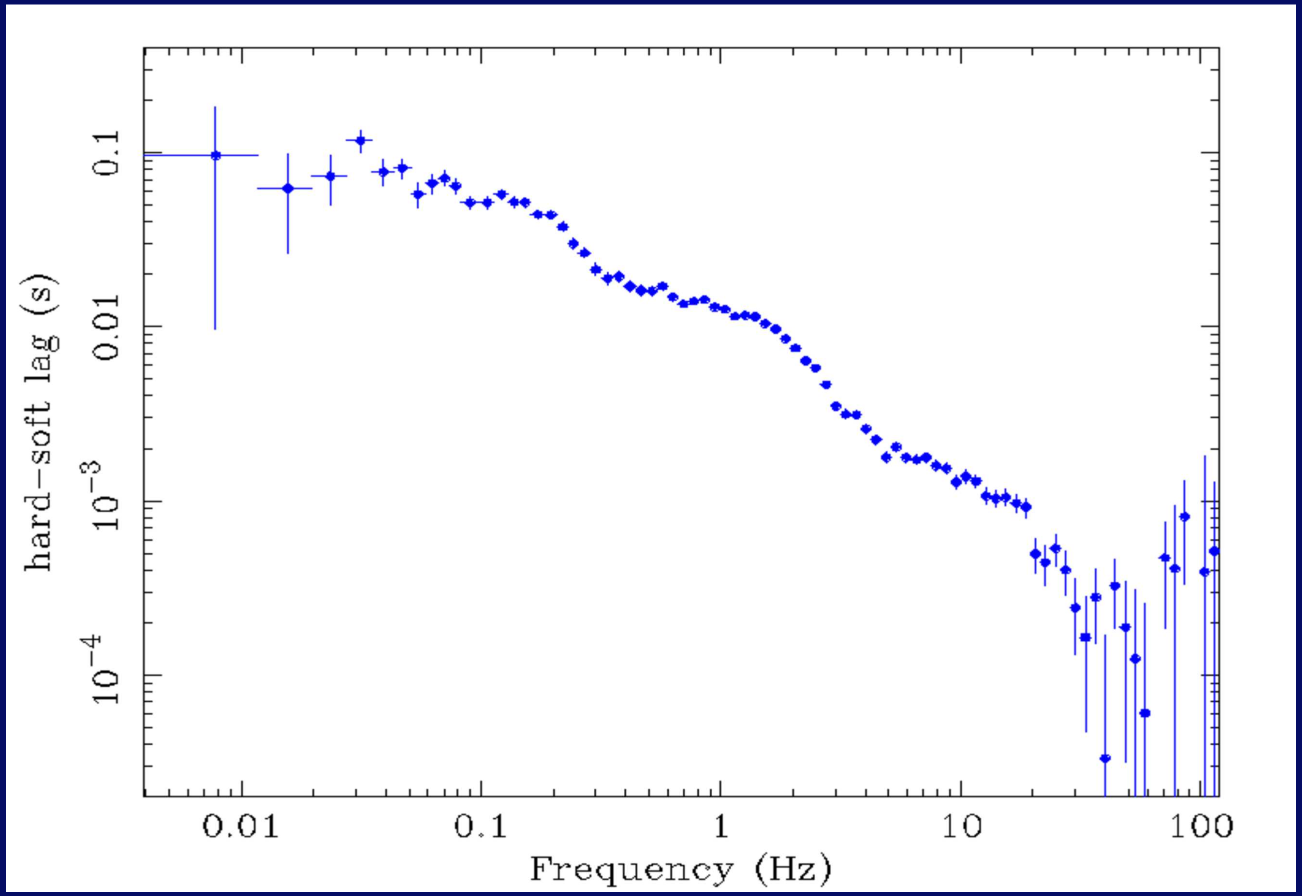
periodic signal simulations

Spectral Timing

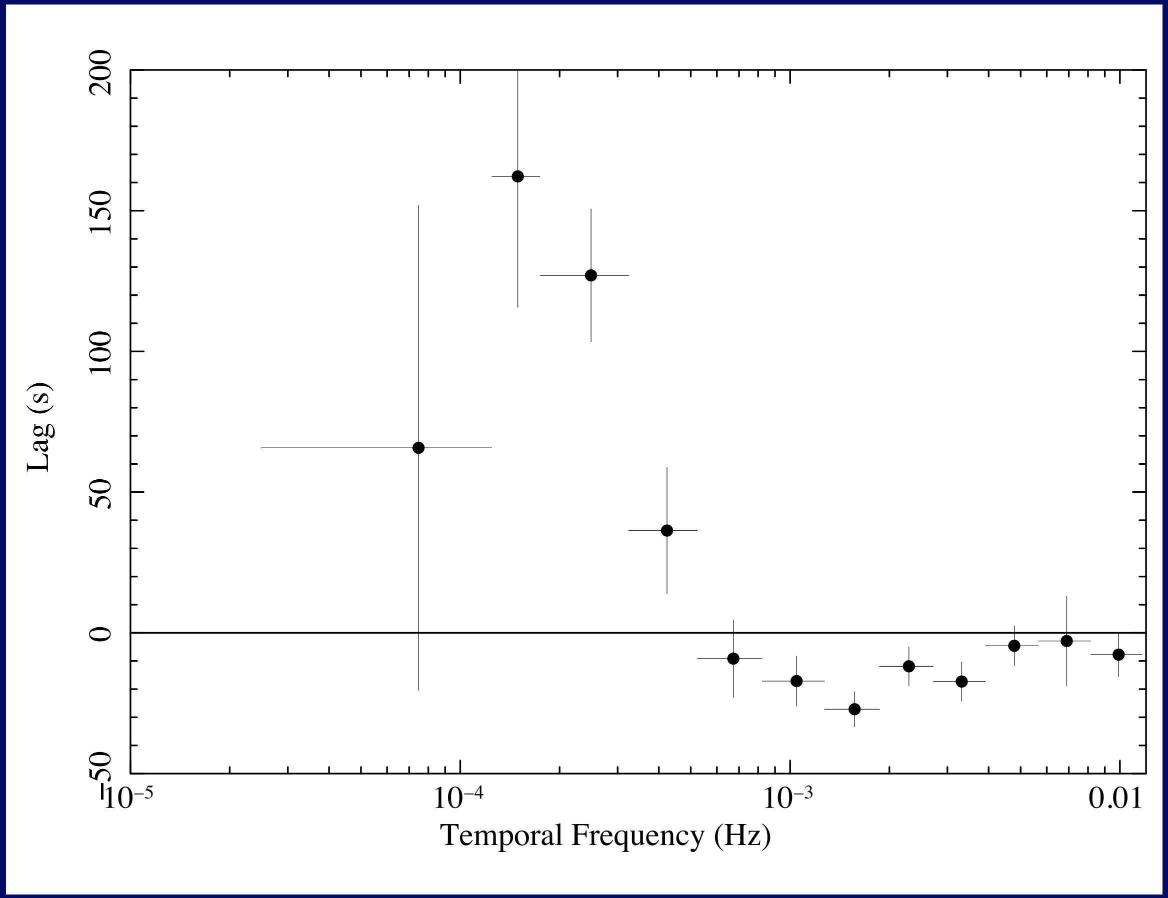
= consider time and energy information at the same time*

*with evenly sampled data

Time Lags

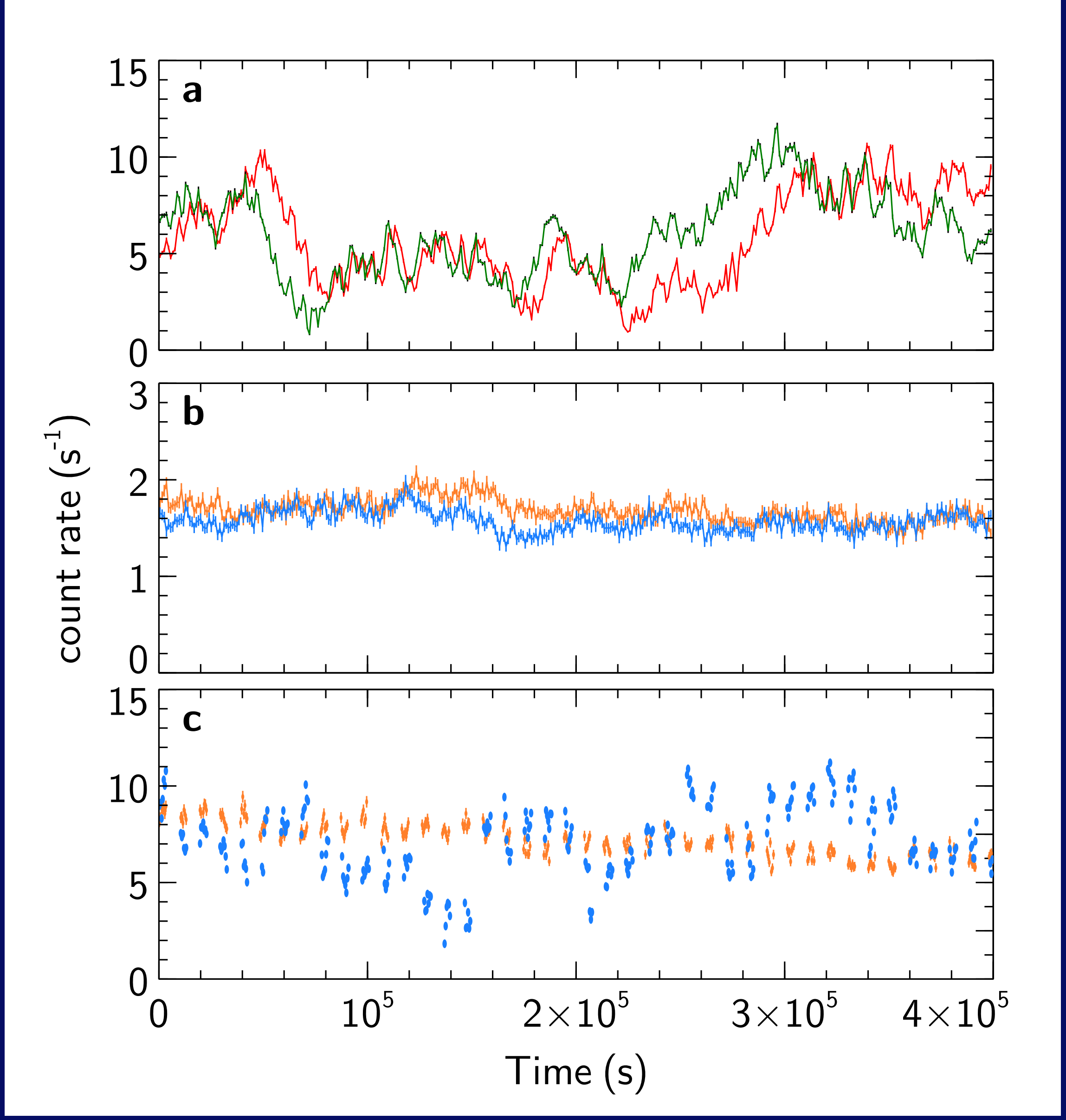


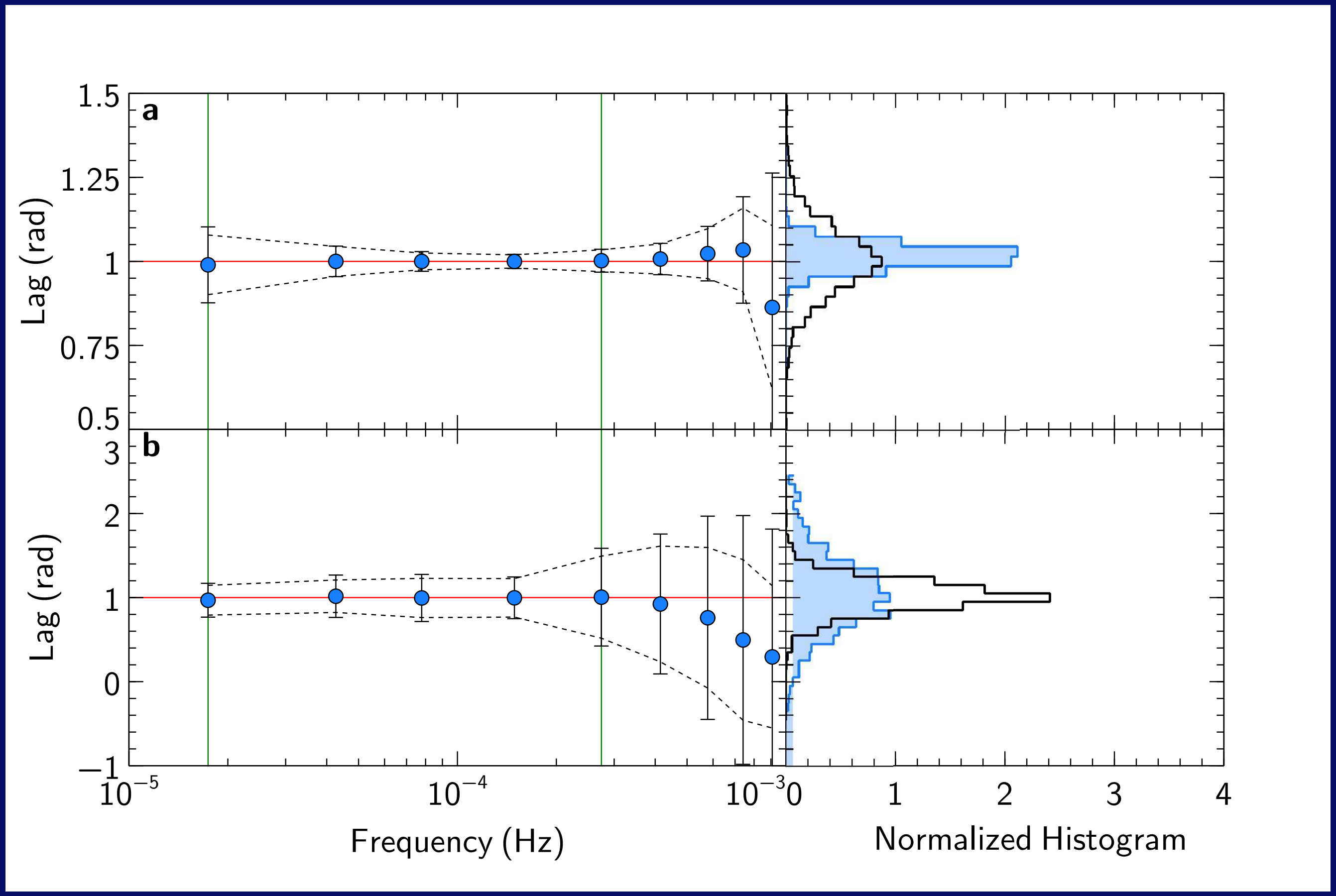
Cygnus X-1: Nowak, 2000



1H0707-495: Uttley et al, 2014

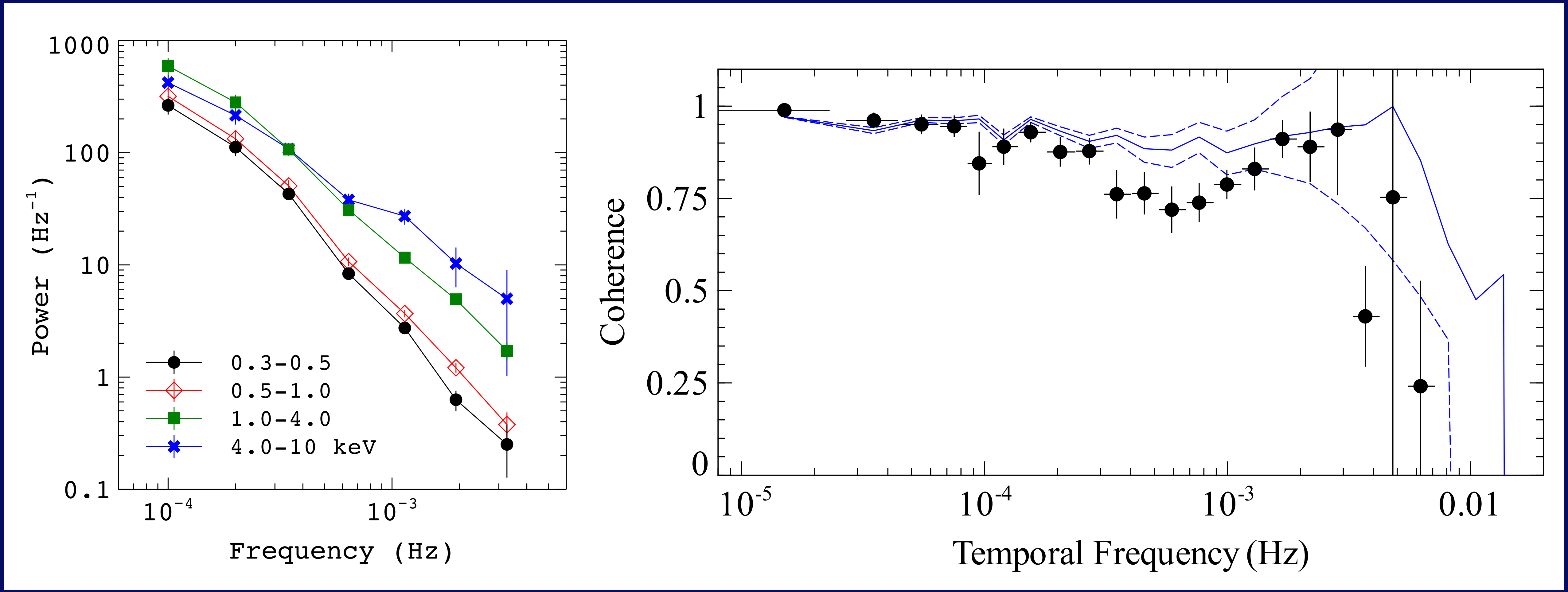
Time Lags II





Zoghbi et al, 2013

Coherence



1H0707-495: Zoghbi et al, 2011

+ lag-energy spectra, lag-frequency spectra, covariance spectra, bispectra, …

Future Challenges

- Multi-wavelength CARMA/GPs?
- Generative models for time-energy data sets
- additional dimensions: polarization

… but my data is non-stationary!

Gaussian Processes Revisited

$$\boldsymbol{y} \sim \overset{\text{Gaussian}}{\mathcal{N}}\left(\underset{\text{Mean}}{\boldsymbol{f}_{\boldsymbol{\theta}}(t)},\ \overset{\text{Covariance}}{K_{\boldsymbol{\alpha}}(t)}\right)$$

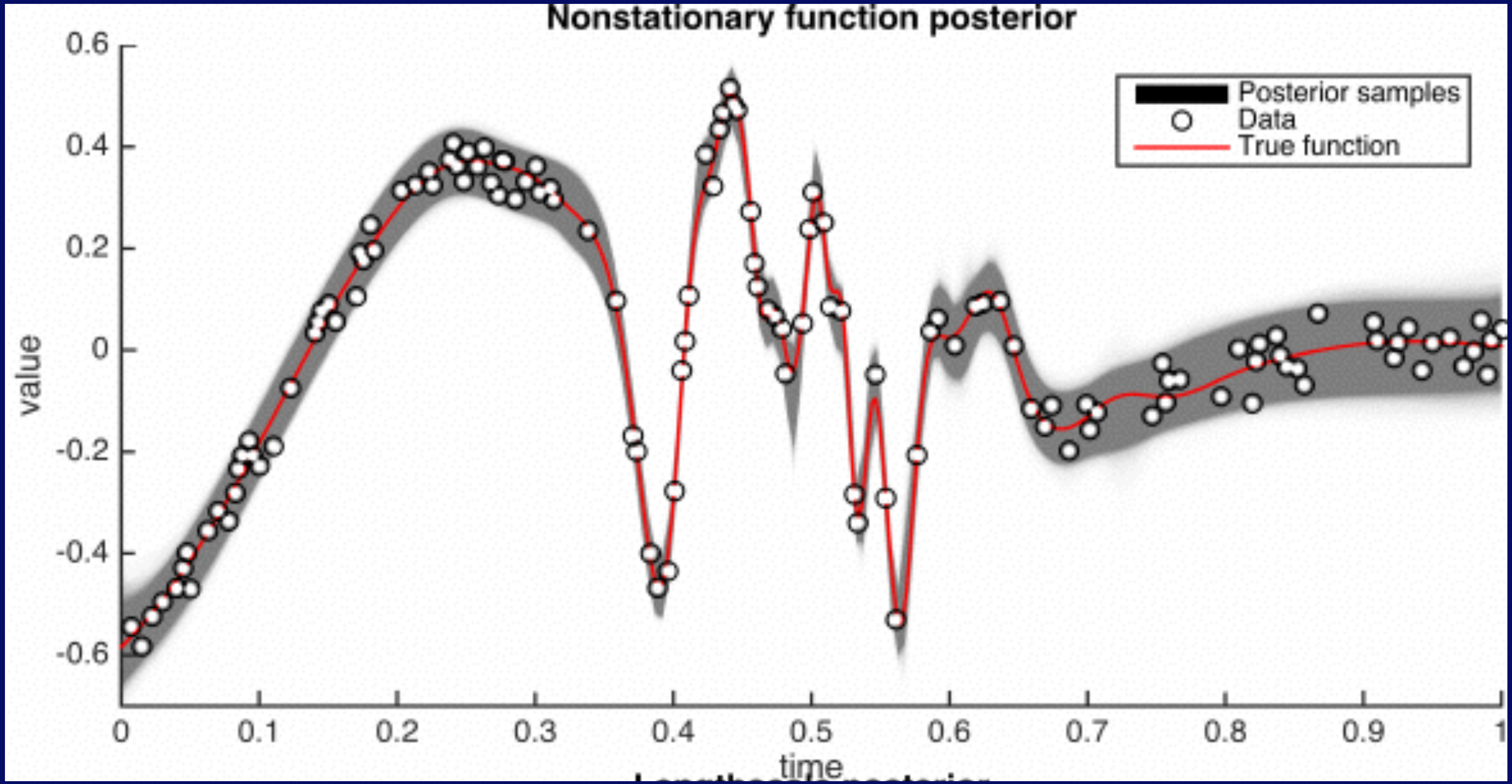
Credit: Dan Foreman-Mackey

Gaussian Processes Revisited

$$\boldsymbol{y} \sim \overset{\text{Gaussian}}{\mathcal{N}}\big(\underset{\text{Mean}}{\boxed{\boldsymbol{f}_{\boldsymbol{\theta}}(t),}}\ \overset{\text{Covariance}}{K_{\boldsymbol{\alpha}}}(t)\big)$$

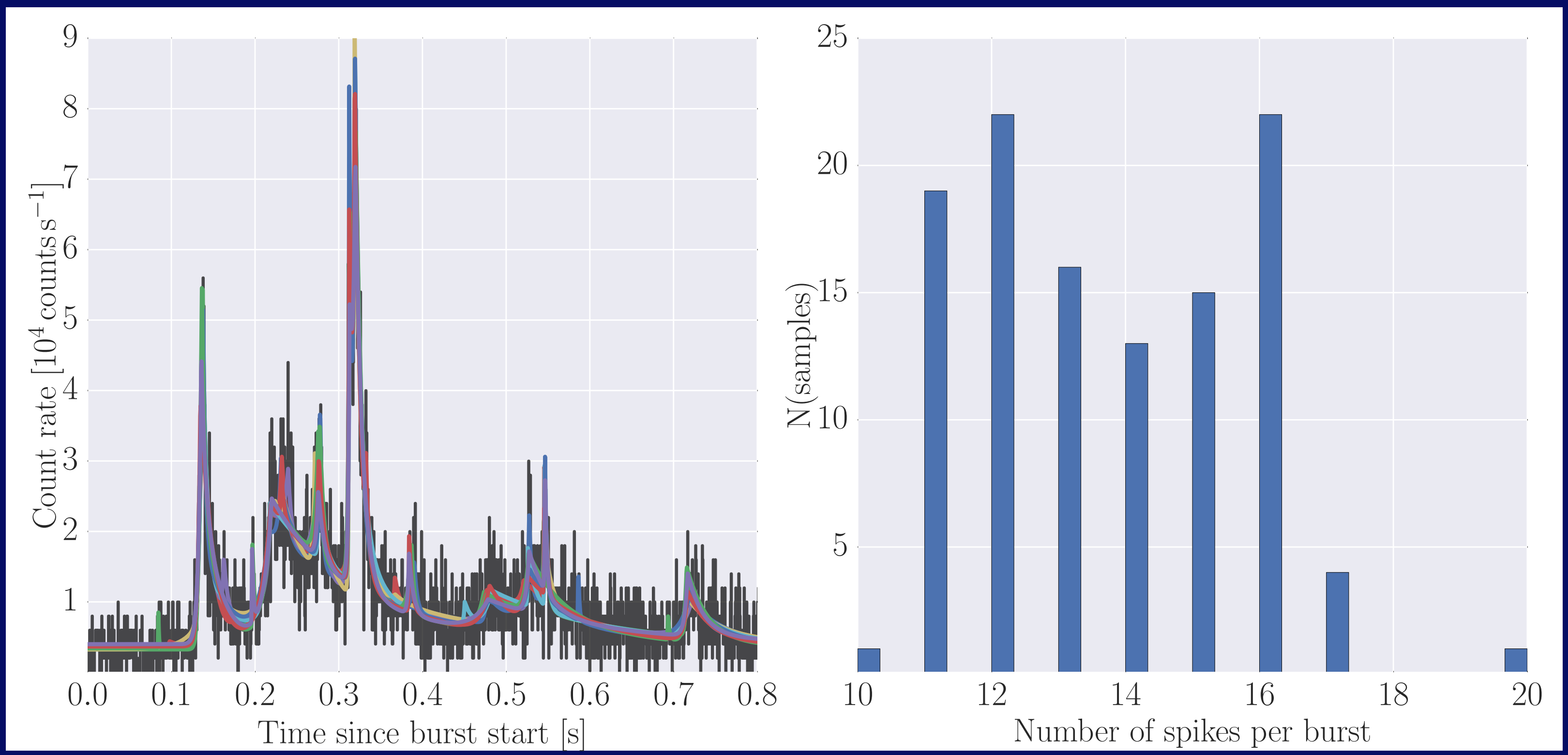
Credit: Dan Foreman-Mackey

Gaussian Processes Revisited

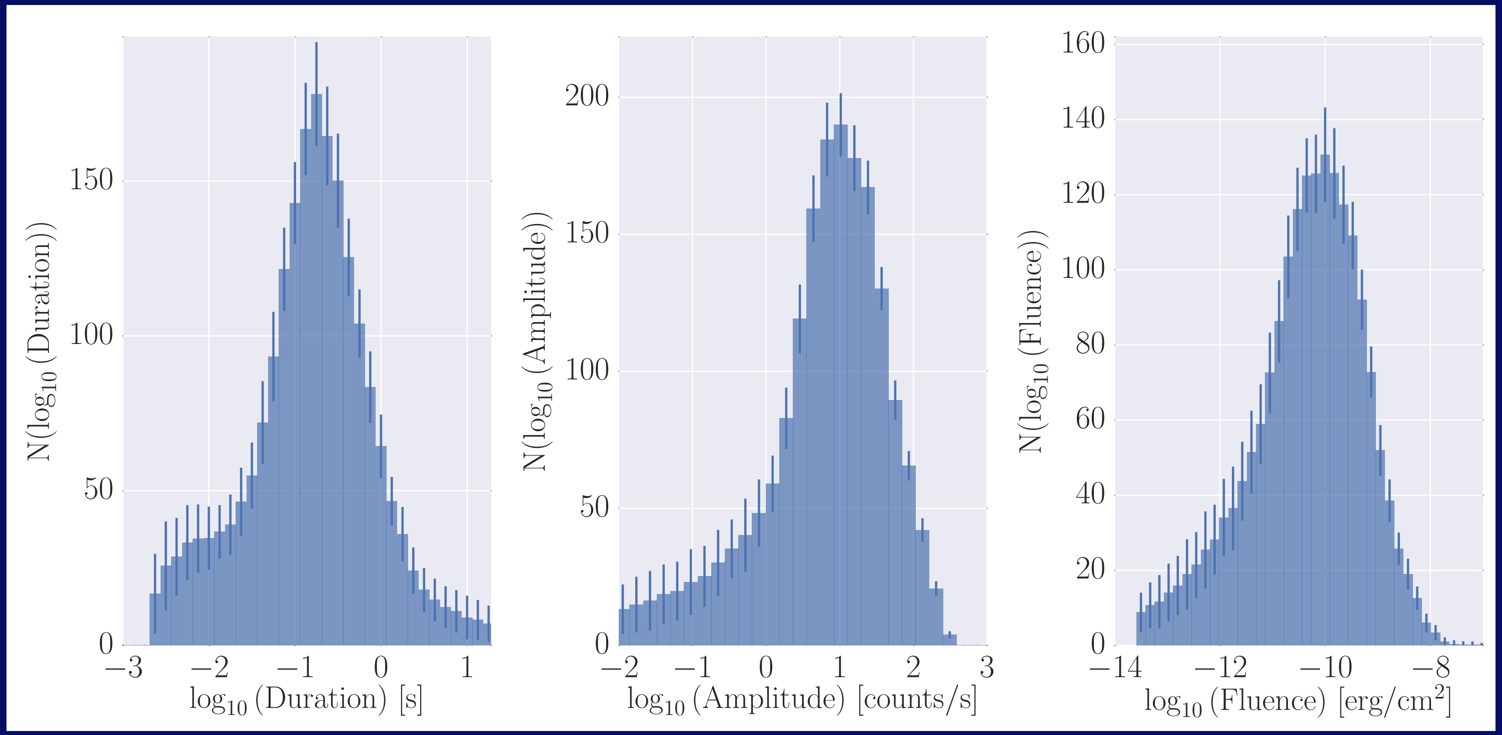


Heinonen et al, 2015

Hierarchical Flare Modeling



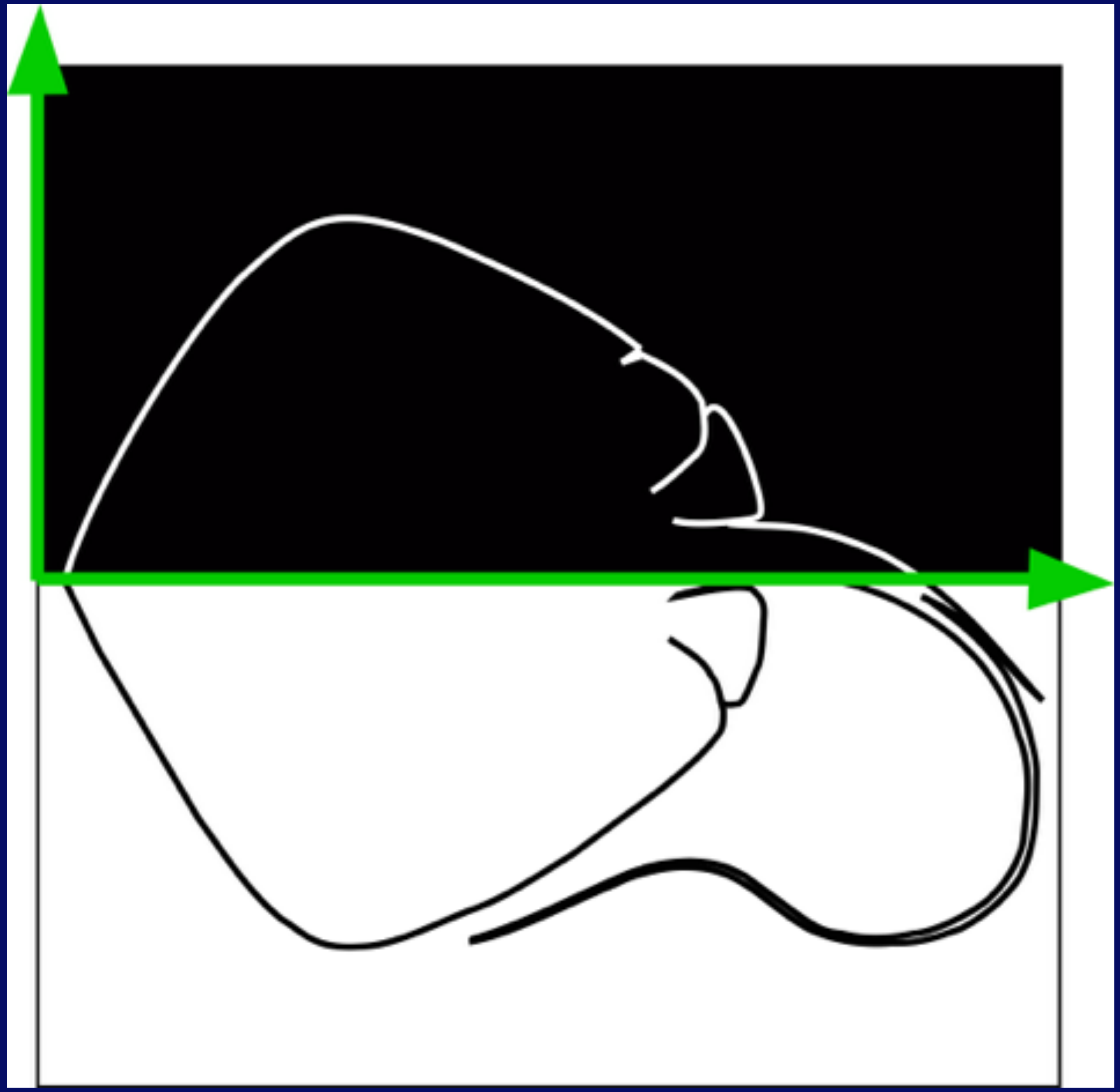
Huppenkothen et al, 2015



Software

- Lomb-Scargle: gatspy , http://www.astroml.org/gatspy
- time series features: extraction: FATS, https://github.com/isadoranun/tsfeat
- Machine learning on time series: http://cesium.ml
- time series analysis: stingray, https://github.com/StingraySoftware/stingray
- CARMA: carma-pack, http://ascl.net/1404.009
- CARMA in Julia: https://github.com/farr/CARMA.jl
- Gaussian Processes: george, http://dan.iel.fm/george/current/
- Hierarchical flare models: https://github.com/dhuppenkothen/magnetron2

Help us build Stingray!*



*https://github.com/StingraySoftware/stingray

Thank you!