A 15 minute introduction to Bayesian data analysis

Greg Ashton

Definitions

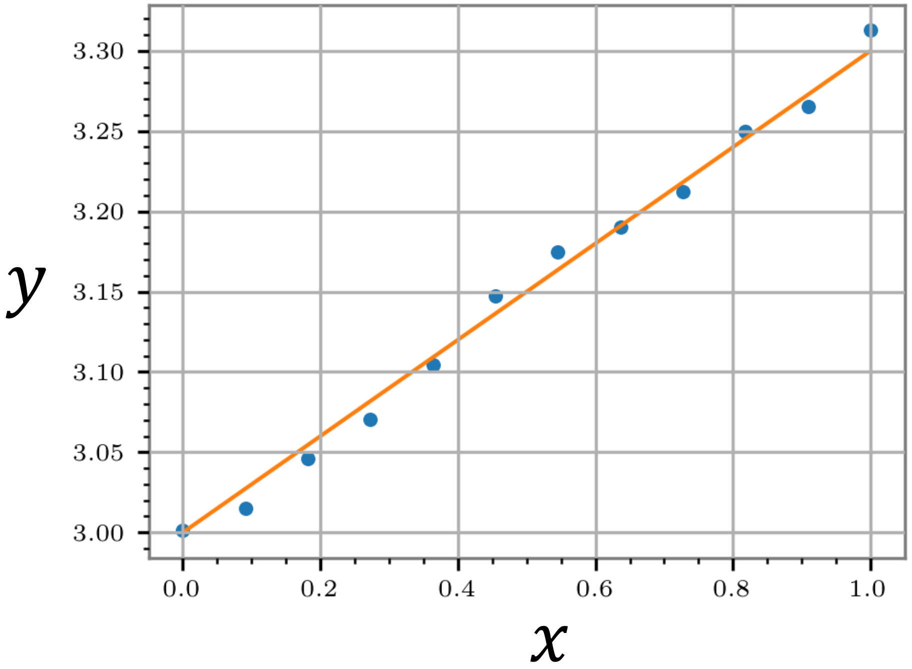
- We have some set of data $d = \{x_i, y_i\}$

- We have a model

- The model makes a prediction, e.g.,

$$y(x) = mx + c$$

- The m and c are model parameters which we denote by $\theta = \{m, c\}$



Questions

- How do we determine the model parameters, given the data?
- How do we decide if the model is a good fit to the data?

Parameter estimation / inference

For model M, how do we determine the model parameters, given the data?

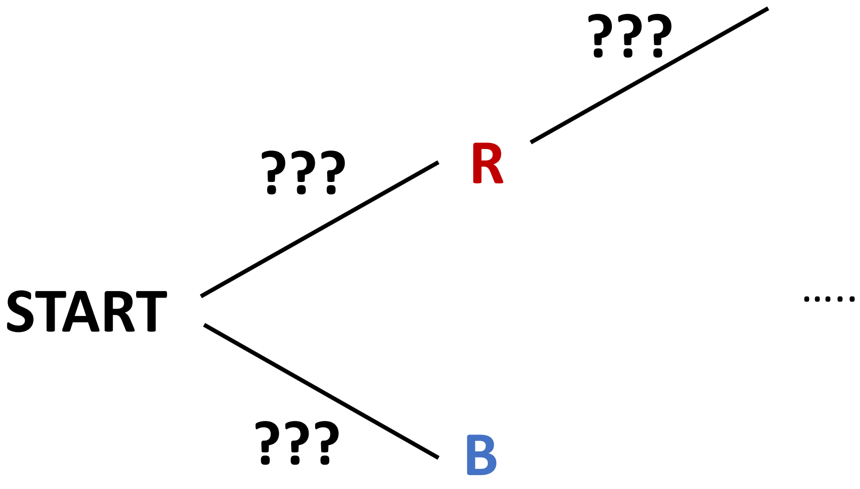
- In Bayesian analysis, this question is answered by calculating the **posterior distribution**

$$P(\theta|d, M)$$

- This distribution can be obtained by the application of conditional probability

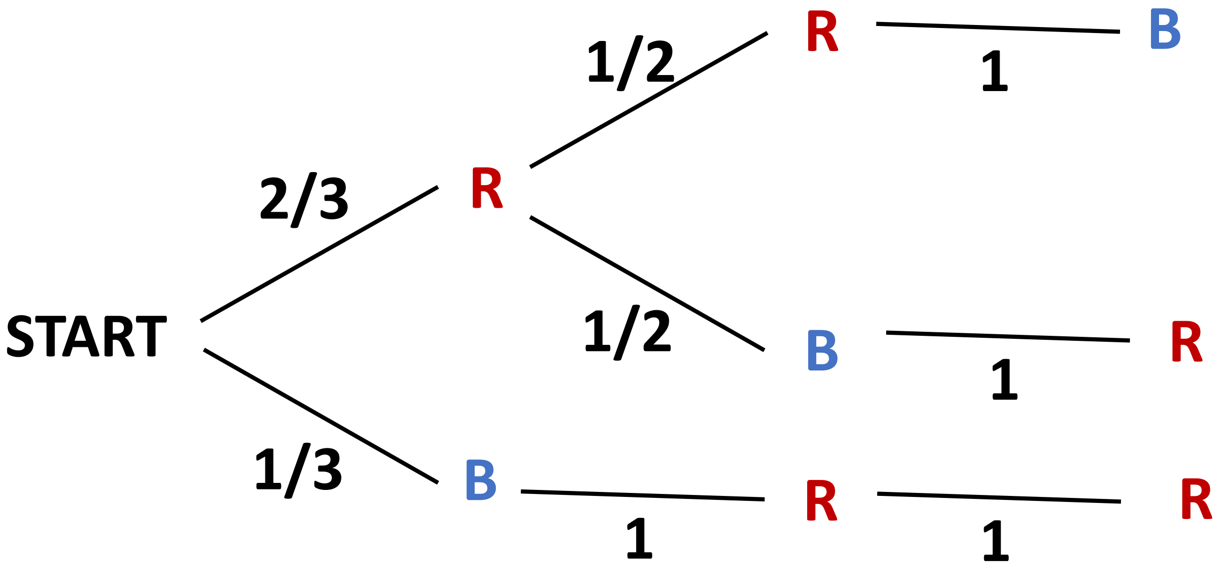
Conditional probability

- Bag containing 2 identical red balls and 1 blue ball
- Take one ball at a time and don't replace it
- Draw a tree diagram of the probabilities..



Conditional probability

- Bag containing 2 identical red balls and 1 blue ball
- Take one ball at a time and don’t replace it
- Draw a tree diagram of the probabilities..



On the first draw,

$$P(R) = 2/3$$
$$P(B) = 1/3$$

On the second draw,

$$P(R|R) = 1/2$$
$$P(B|R) = 1/2$$
$$P(B|B) = 0$$

And so on…

Conditional probability

- $P(A|B) \Rightarrow$ The probability of "A" given "B"

- Relation to the logical "AND"

$$P(A \text{ and } B) = P(AB) = P(A|B) \times P(B)$$

- E.g., draw 2 balls from the bag, what is the probability of a R and R?

$$P(RR) = P(R|R) \times P(R) = 1/2 \times 2/3 = 1/3$$

Bayes theorem

- We can write down two statements:

$$P(AB) = P(A|B)P(B)$$
$$P(AB) = P(B|A)P(A)$$

- Therefore

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

This is known as “Bayes theorem”

Why is this useful?

- **Question**: how do we determine the model parameters, given the data?
- **Answer**: the posterior distribution:

likelihood

prior

posterior

$$P(\theta|d, M) = \frac{P(d|\theta, M)P(\theta|M)}{P(d|M)}$$

evidence for model M

Calculating the posterior and evidence

- Calculating the posterior can be done a few ways
  1. Analytically, if the maths is easy enough
  2. Numerically on a **grid of points** (if in a small number of dimensions)
  3. Using **stochastic sampling** (MCMC/Nested Sampling)

- The evidence is a normalizing factor

$$P(d|M) = \int d\theta \, P(d|\theta, M) \, P(\theta|\, M)$$

- Which can be discarded if only the model parameters are of interest

How do we decide if the model is a good fit to the data?

- We can use conditional probability

evidence

model prior

$$P(M|d) = \frac{P(d|M)P(M)}{P(d)}$$

- If we can compute this, then it is exactly the probability of the model given the data

- Unfortunately, in practise computing $P(d)$ requires us to know **every** model which can generate the data!

How do we decide if the model is a good fit to the data?

- Instead, we can make comparative statement via “the odds”
- Let M_a and M_b be two models, then

$$\frac{P(M_a|d)}{P(M_b|d)} = \frac{P(d|M_a)}{P(d|M_b)} \frac{P(M_a)}{P(M_b)}$$

odds

prior-odds

Bayes factor (BF)

$$BF = \frac{P(d|M_a)}{P(d|M_b)}$$

if $BF > 1$, model A is preferred

if $BF < 1$, model B is preferred

Rotational evolution of the Vela pulsar during the 2016 glitch

Greg Ashton, Paul Lasky, Vanessa Graber, and Jim Palfreyman

greg.ashton@monash.edu

The 2016 Vela glitch

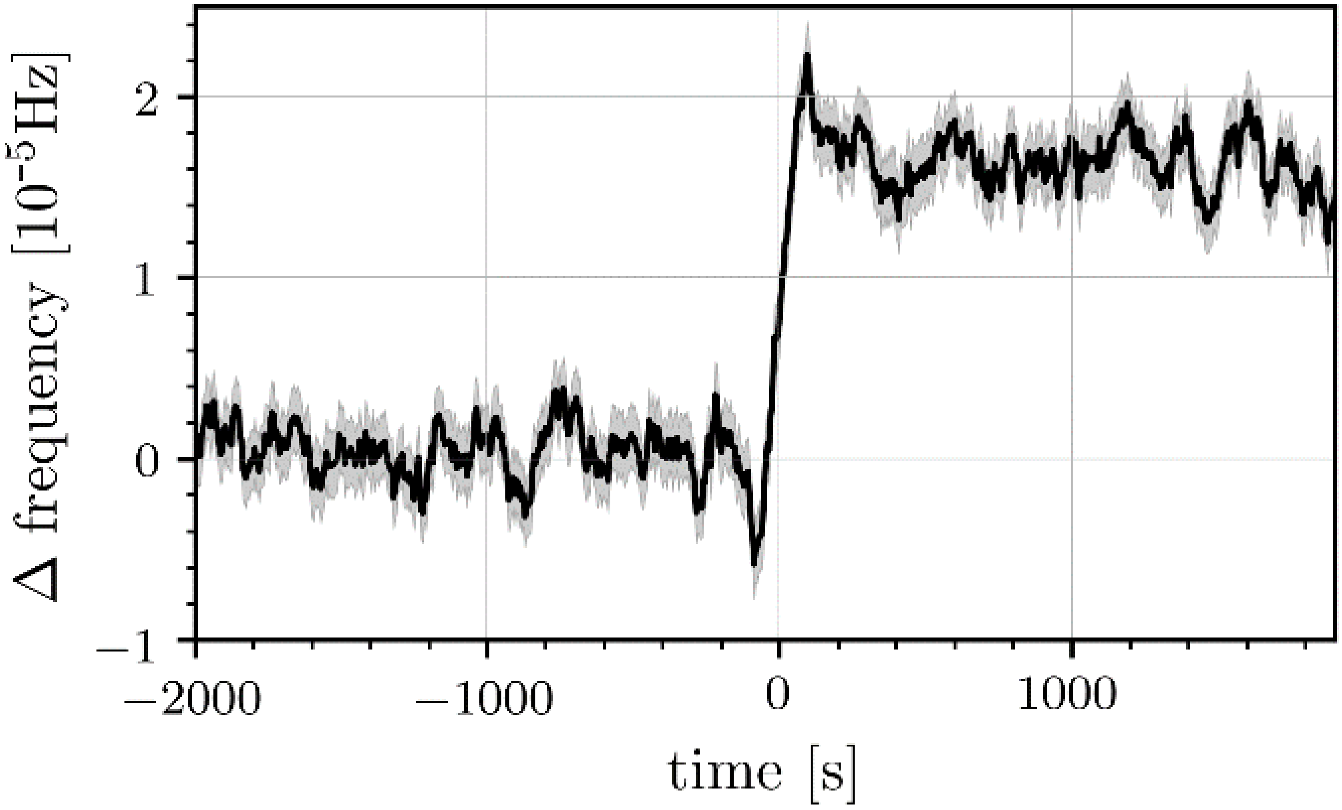
- Continuous monitoring by the 26m Mount Pleasant observatory in Hobart Tasmania (Palfreyman et al. (2018))
- First pulse-to-pulse data taken during a glitch
- New insights about the behaviour of the magnetosphere during the glitch
- Here, we focus on the dynamics:

What is happening to the rotation of the star during the glitch?

Model-agnostic frequency-evolution

- Fit a constant-frequency model in a 200s sliding window

- Features are “smeared out” due to sliding window



Inference

- Using the time-of-arrival (TOA) data
- Model the frequency evolution
- Integrate to get phase, then invert to get the predicted TOA
- Use Bilby, Ashton et al. (2019), to perform parameter estimation and model selection

git.ligo.org/lscsoft/bilby

Step-glitch model

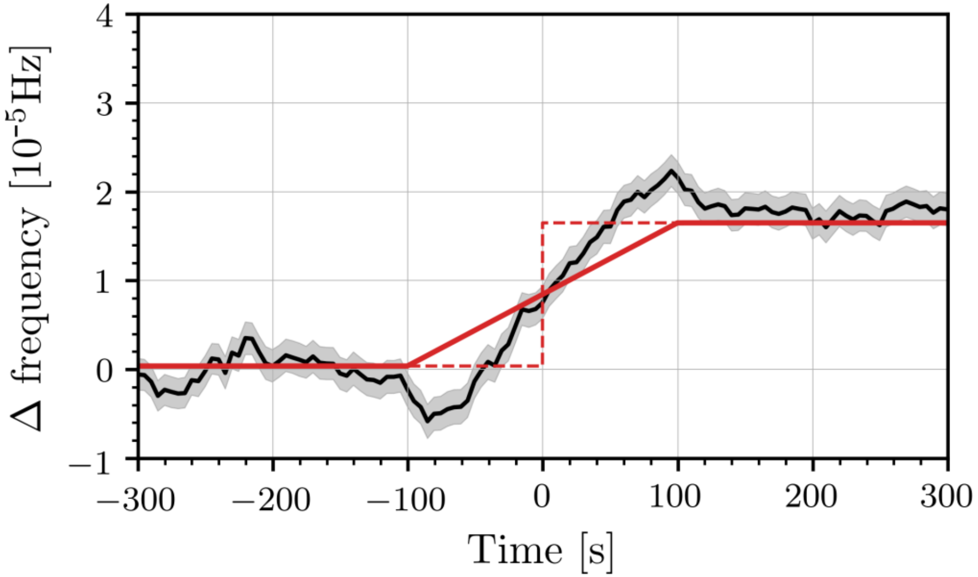
- The simplest model of a glitch
- Parameterise by the measureable glitch size Δf and glitch time

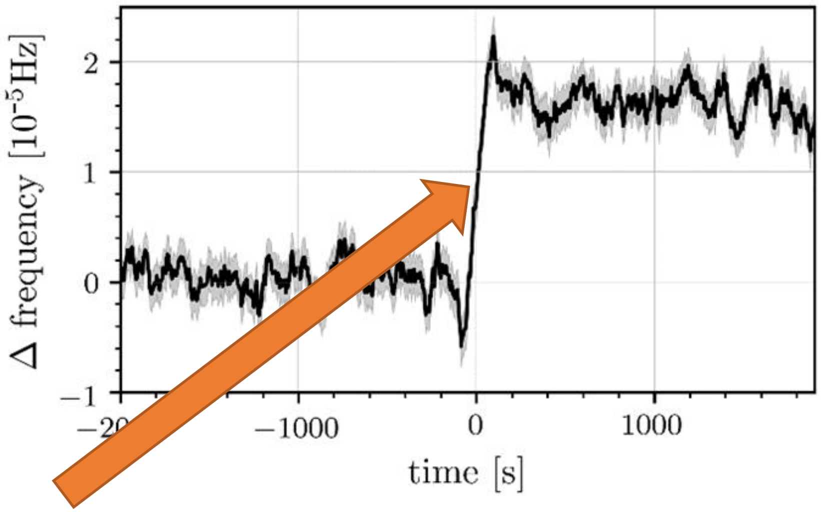
$$f(t) = f_0 + H(t - t_g)\Delta f$$

- Posterior on the glitch size agrees with Palfreyman et al (2016):

$$\Delta f \sim 16.11 \pm 0.04\ \mu\text{Hz}$$

- Time of the glitch consistent (to within 90% credible interval)
- Indicates we have handled the data consistently





*this is **not** the rise-time

The glitch rise time

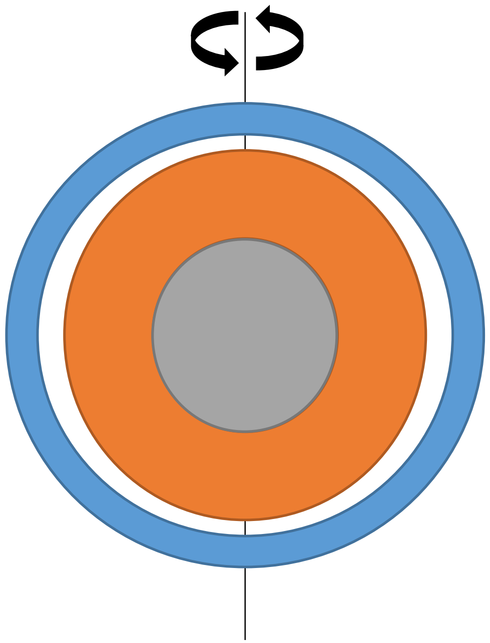
Rise-time model

- Parameterise by the measureable glitch size Δf and rise time τ_r

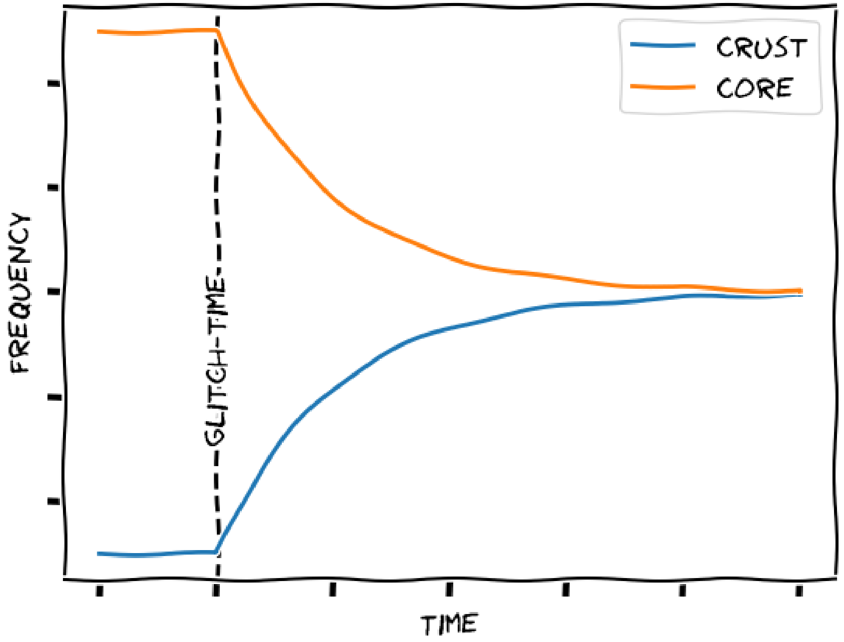
$$f(t) = f_0 + H(t - t_g)\left[\Delta f + \Delta f_r e^{-\frac{t-t_g}{\tau_r}}\right]$$

Physics:

- Simple two-component model
- Coupling torque proportional to the lag between the crust and core



- Crust
- Core involved in the glitch
- The rest of the core

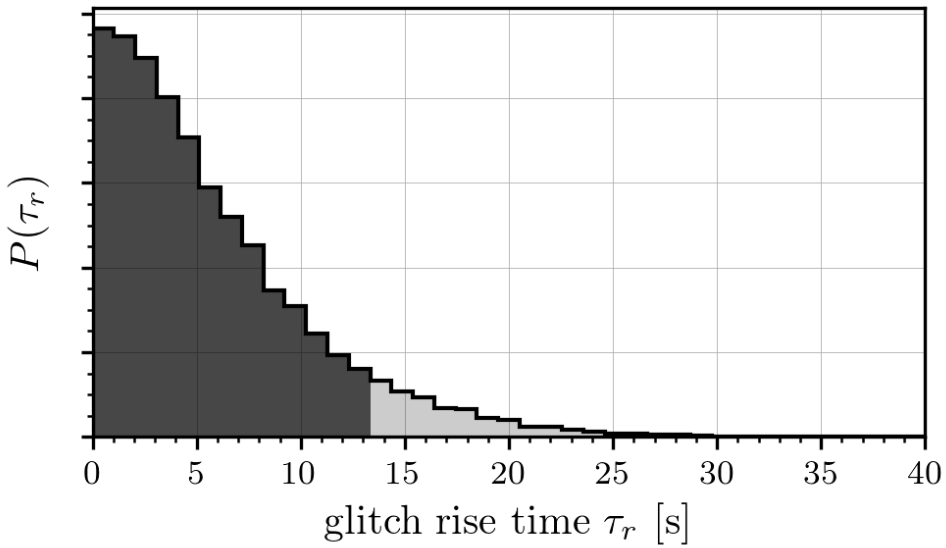


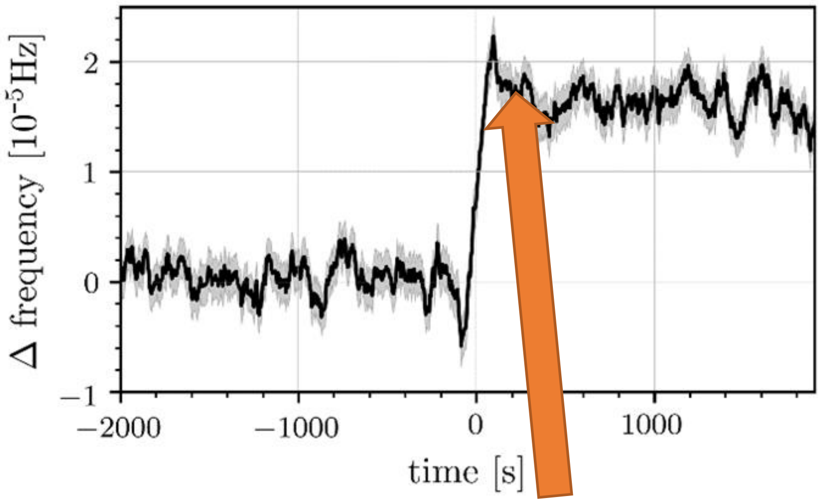
Rise-time model results

- Fit and compare model with a rise time to a step-glitch model:

$$BF = 10^{-1.7}$$

- In favour of step-glitch: we cannot resolve the rise-time
- Upper limit on the rise time of ~ 13s (90% confidence)
- Previous best upper limit was 30s ([Dodson et al. (2007)](#))
- Can use this to constrain the mutual friction coefficient, directly related to the superfluid vortex dynamics





The overshoot

Overshoot model

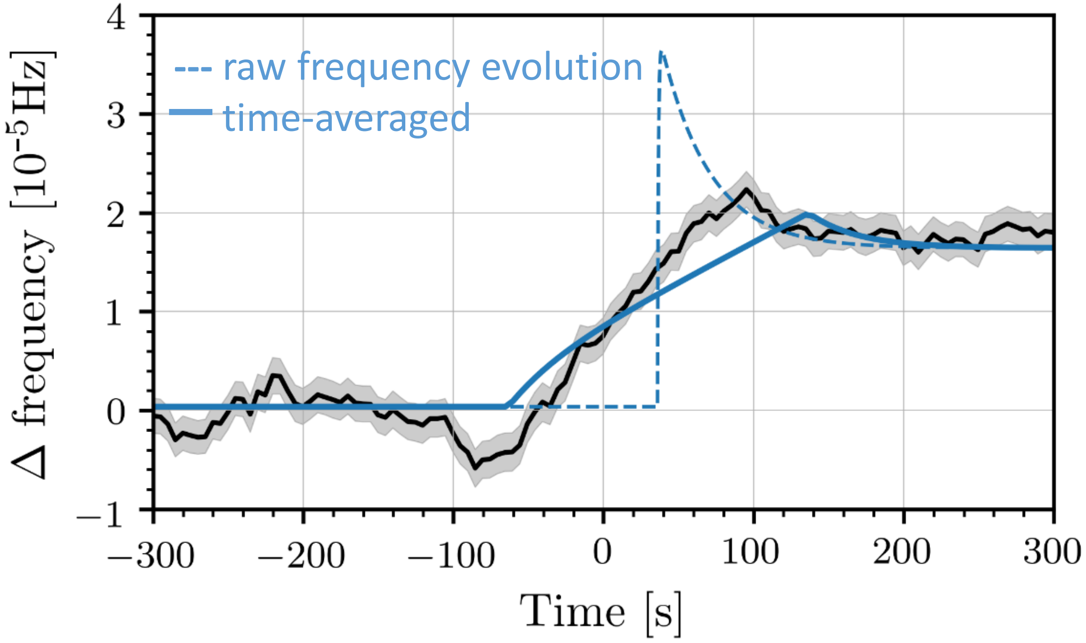
- Already a number of models in the literature for relaxation in glitches on timescales of order hours or longer (see, Haskell & Melatos (2015))
- The dynamics here are happening on order of seconds-minutes
- A few plausible known mechanisms:
  - Three-component model (Graber et al. (2018))
  - Two-component model (Haskell et al. (2012), Antonelli et al. (2017))
  - Ekman pumping (Van Eysden & Melatos (2015))

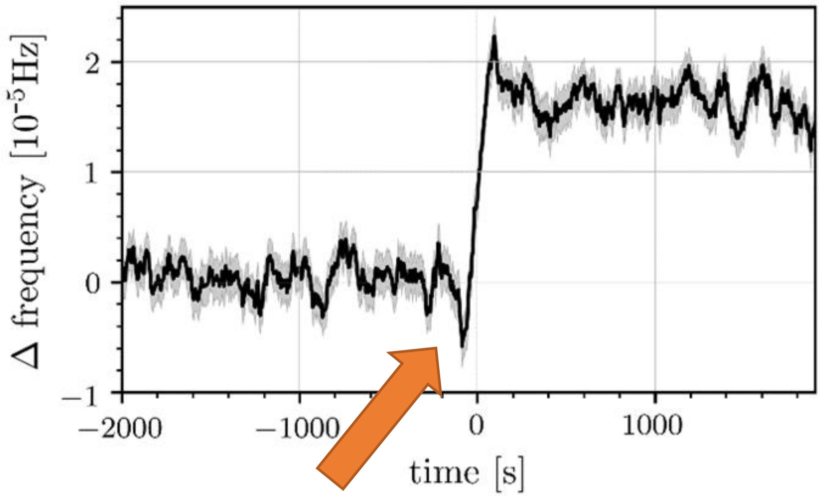
Overshoot model: results

- Compare rise-time + overshoot model with just rise-time

$$BF = 10^{2.1}$$

- Overshoot model preferred!
- Decay time ~ 1 min
- Dodson et al. (2002) found a similar component with a similar timescale





A precursor slowdown?

An “antiglitch”?

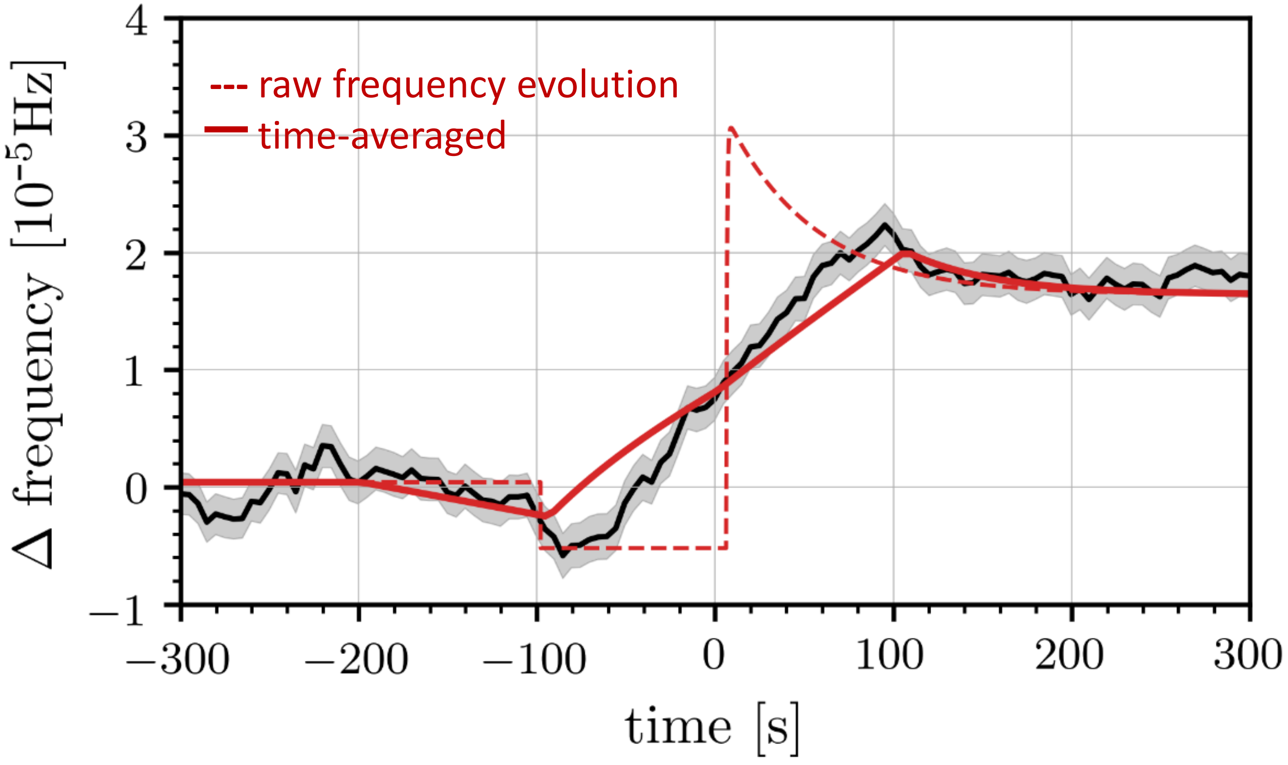
Precursor slowdown

- Not predicted in the literature

- Substantially favoured over any of the other models tested

$$BF > 10^{2.5}$$

- Magnitude $\sim 5 \times 10^{-6}$ Hz

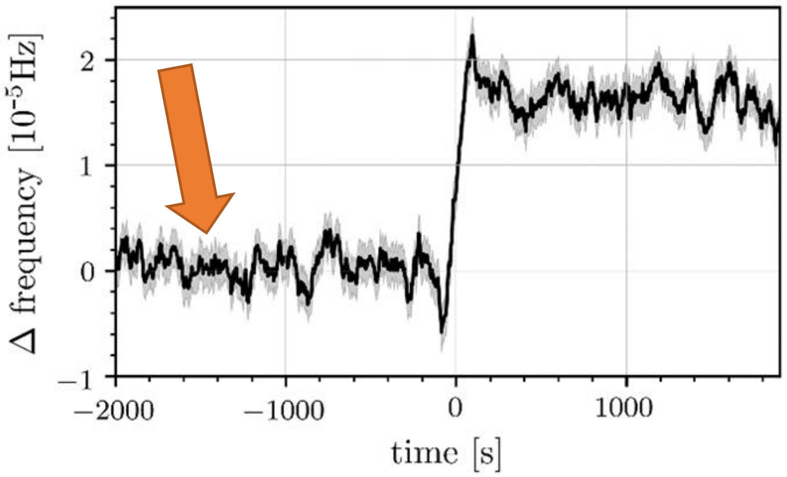


Physics of the preceding slowdown

- Model is very coarse with no physics: opportunity for better physics input
- While this may be the first confirmed observation of a slowdown prior to a glitch, it is hinted at in Dodson et al (2002)

Speculation

- Stochastic frequency noise oscillations are $\sim 10\%$ of the glitch size
- Perhaps a glitch is "triggered" by a particularly large (and negative) noise event
- We can test this using "off-glitch" data



Summary

- Lots of interesting dynamics to pull out of the 2016 Vela glitch
- Improved constraints on the rise time
- Strong support for an overshoot
- Support for a pre-cursor slowdown, but lacking a physical model
  - Corroborated by previous 2000 Vela glitch
  - We speculate the slow-down may trigger the glitch
  - The slow-down could be a large stochastic frequency noise event
  - Or, something else entirely…
- Paper arXiv:1907.01124 accepted in Nature Astronomy