# Identifying transient astronomical sources in MeerKAT light curves using Gaussian processes


Shih Ching Fu

Co-supervisors:

Dr Arash Bahramian, Dr Aloke Phatak

Associate Supervisors:

Dr James Miller-Jones, Dr Suman Rakshit



International Centre for Radio Astronomy Research

Curtin University

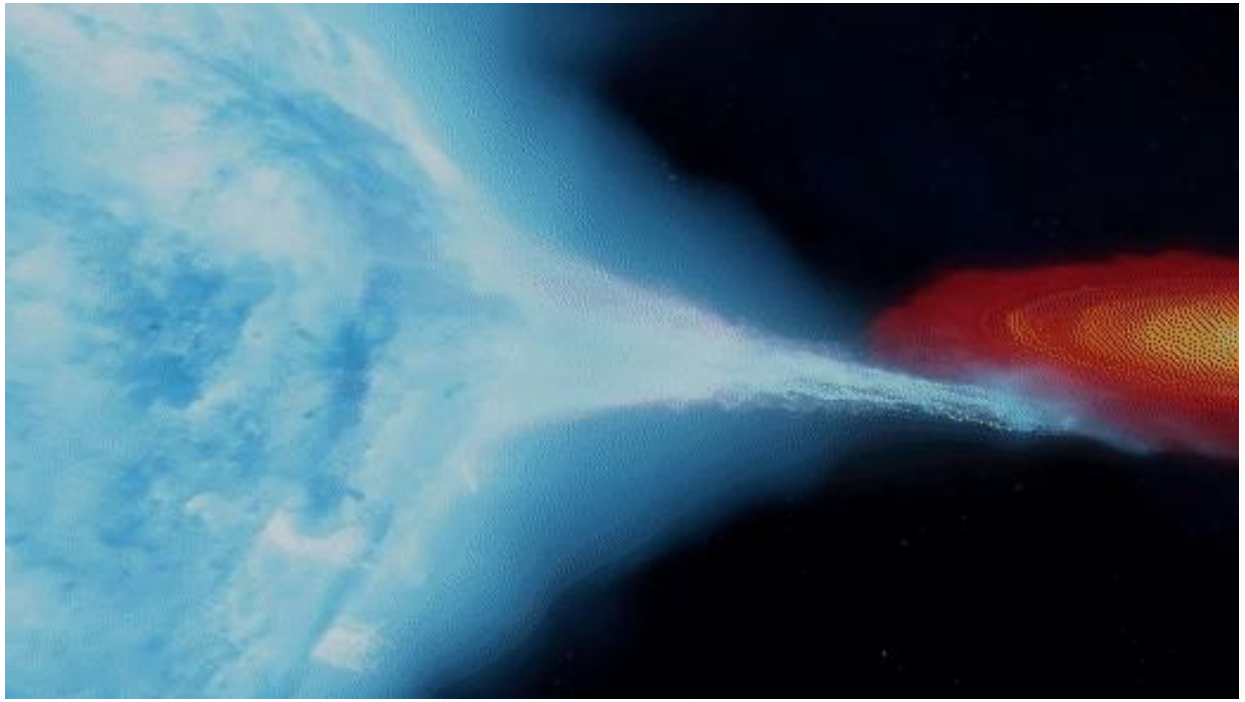
The University of Western Australia

Government of Western Australia
Department of the Premier and Cabinet
Office of Science

# Twinkle twinkle…

A *transient* is an astrophysical phenomenon whose brightness changes drastically over observable time.

- Supernovae
- Variable stars, e.g., pulsating,
- eclipsing binaries.
- Gamma-ray bursts (GRBs)
- Fast radio bursts (FRBs)
- Transiting planets
- Active galactic nuclei (AGN)
- Accreting blackholes
- and lots more…

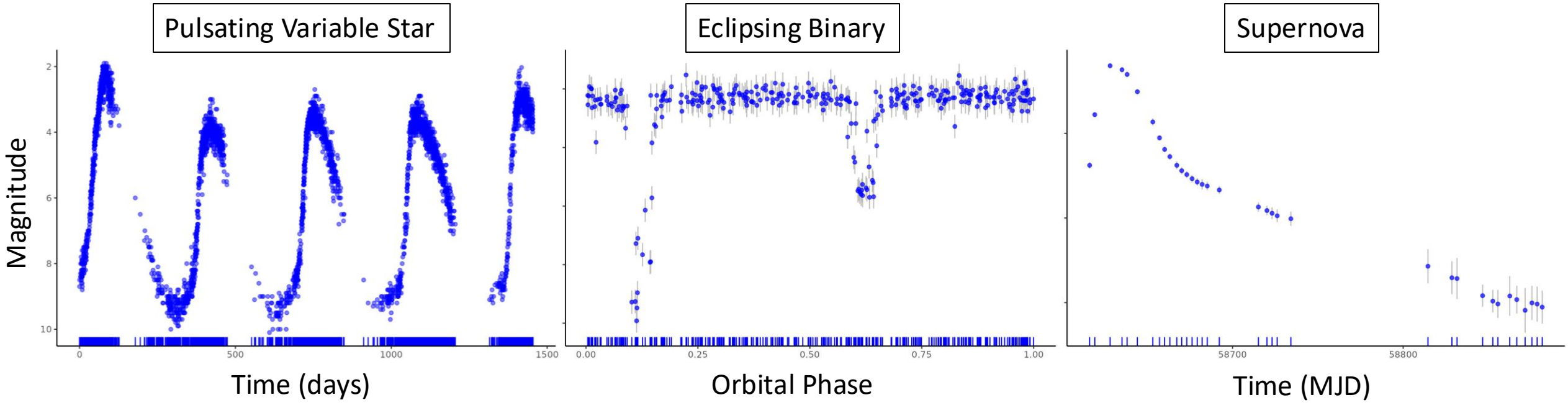


Artist's impression of the Cygnus X-1 system. Credit: ICRAR

# Light Curves

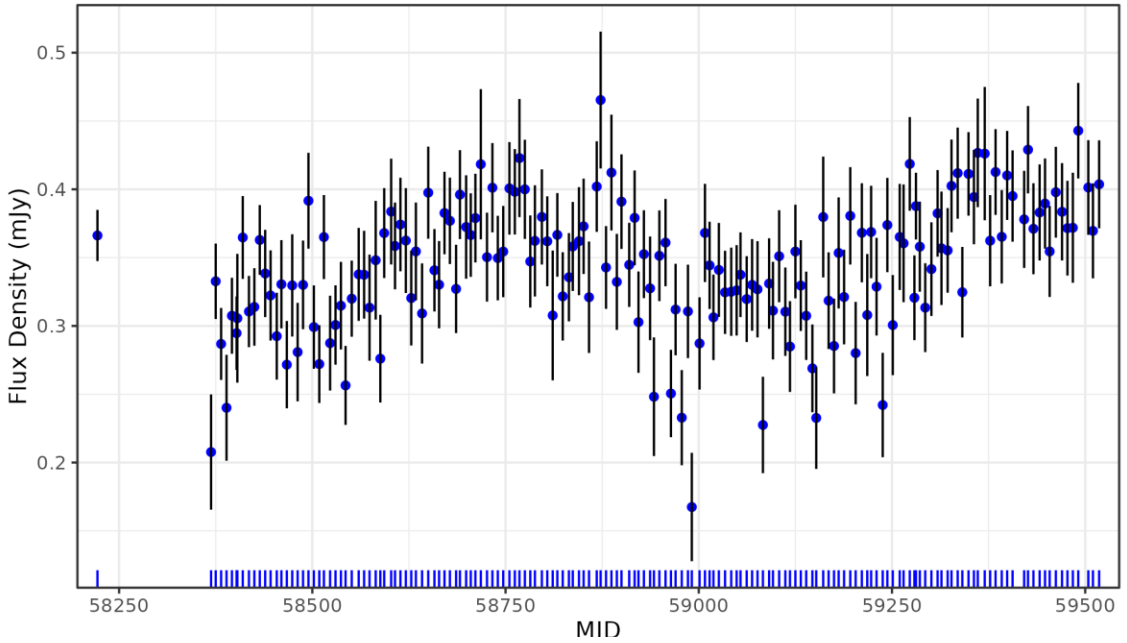
*Light curves* are time series describing the brightness of a source over time.

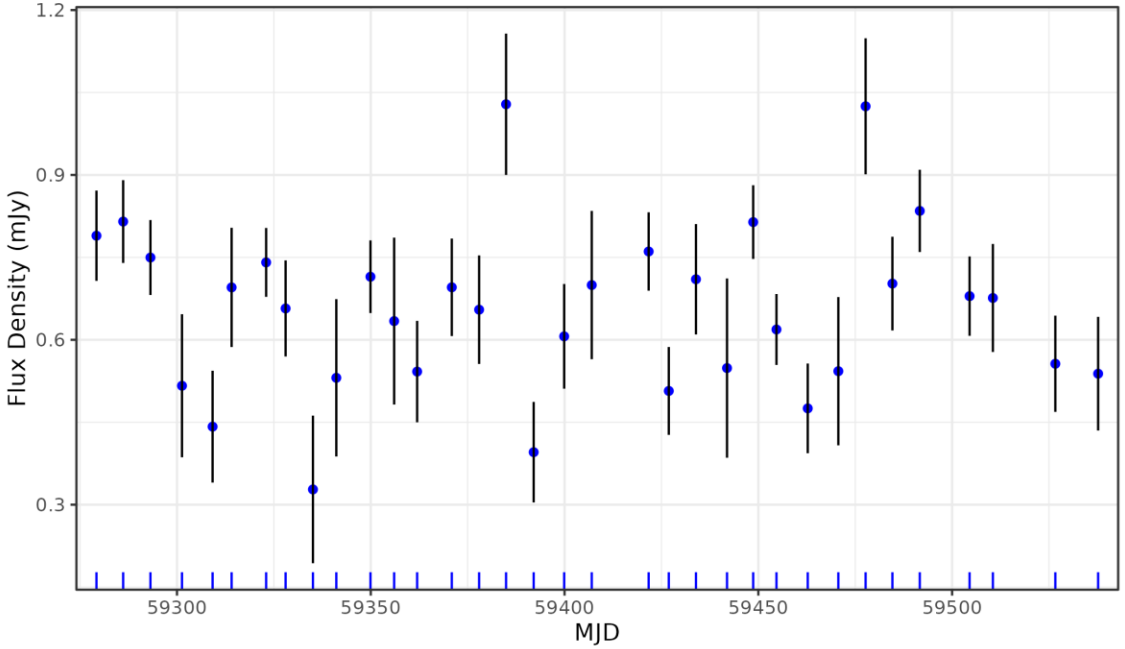
- The shape of a light curve can reveal the type of object or event.
- Variability in brightness can reveal information about the processes underlying the observed phenomenon.

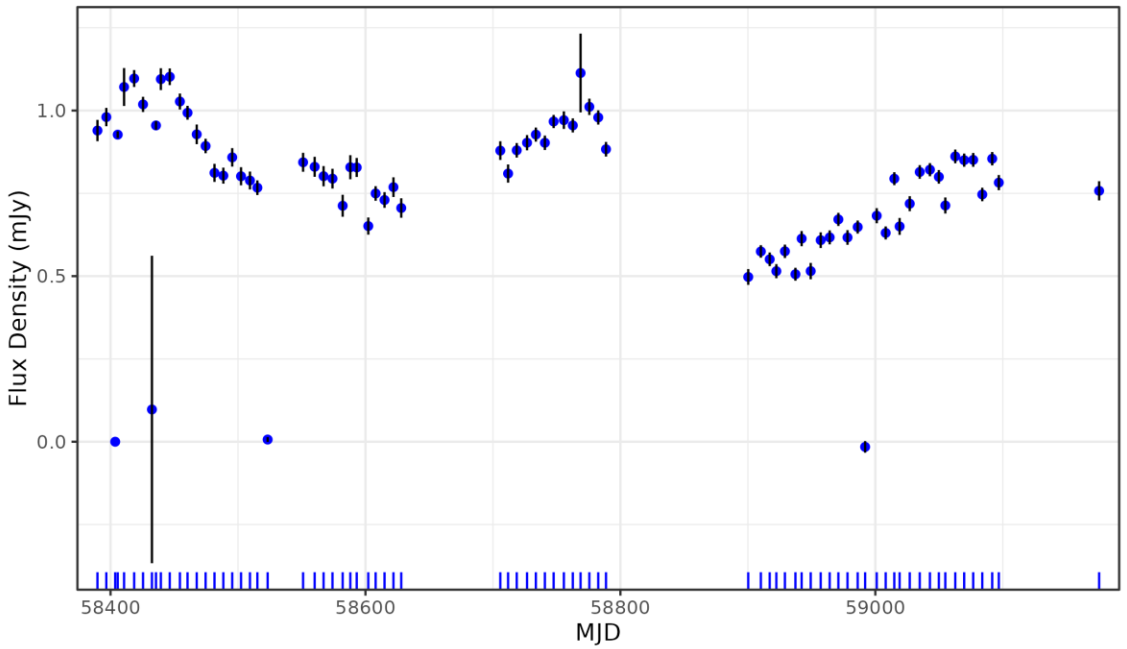


# Patchy Data

- Different cadences
- Sparse observations
- Irregular sampling
- Varying noise levels



# Characterising Light Curves

**Oversimplified** ⟷ **Overspecified**

- Fewer parameters
- Scales easily
- High information loss

- Many parameters
- High discriminatory power
- Overfitting

**Model light curves as a Gaussian Process (GP)**

# Gaussian Processes (GPs)

Extend multivariate Gaussian to 'infinite' dimensions.

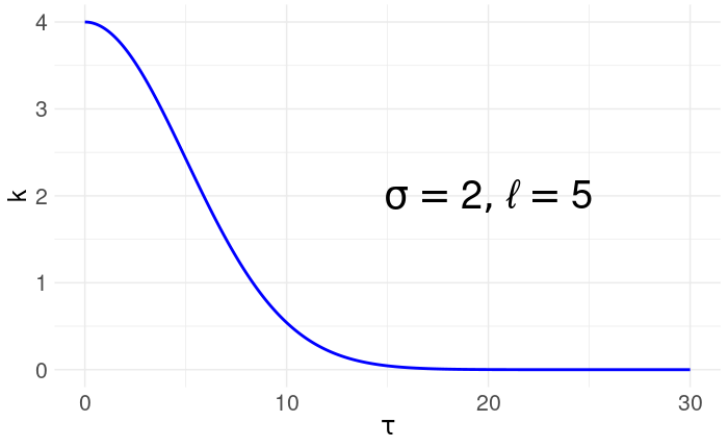
- Mean function, $\mu(t)$
- Covariance or **kernel function**, $\boldsymbol{\kappa(t, t)}$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \end{bmatrix} = \boldsymbol{Y} \sim GP(\mu(t), \boldsymbol{\Sigma})$$

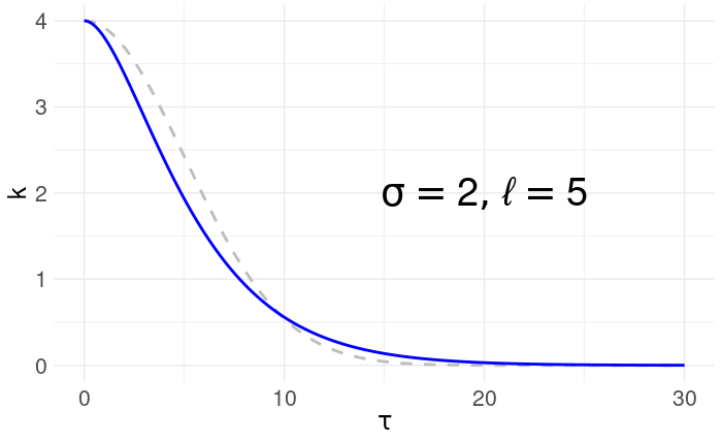
where $\boldsymbol{\mu} = \mu(t_i)$ and $\Sigma_{ij} = \boldsymbol{\kappa(t_i, t_j)}$, for $i, j = 1, 2, \ldots$

Rather than specifying a fixed covariance matrix with fixed dimensions, compute covariances using the kernel function.

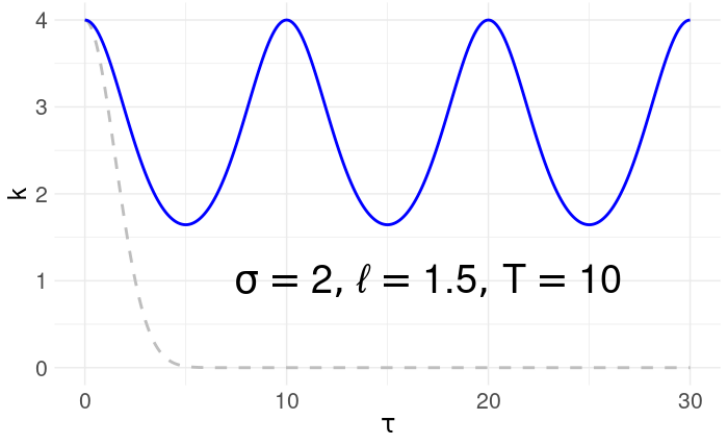
$$\tau = |t_r - t_c|;\ \sigma, \ell, T > 0$$




$$\kappa(\tau;\sigma,\ell) = \sigma^2 \exp\left\{-\frac{1}{2}\left(\frac{\tau}{\ell}\right)^2\right\}$$ Squared Exponential

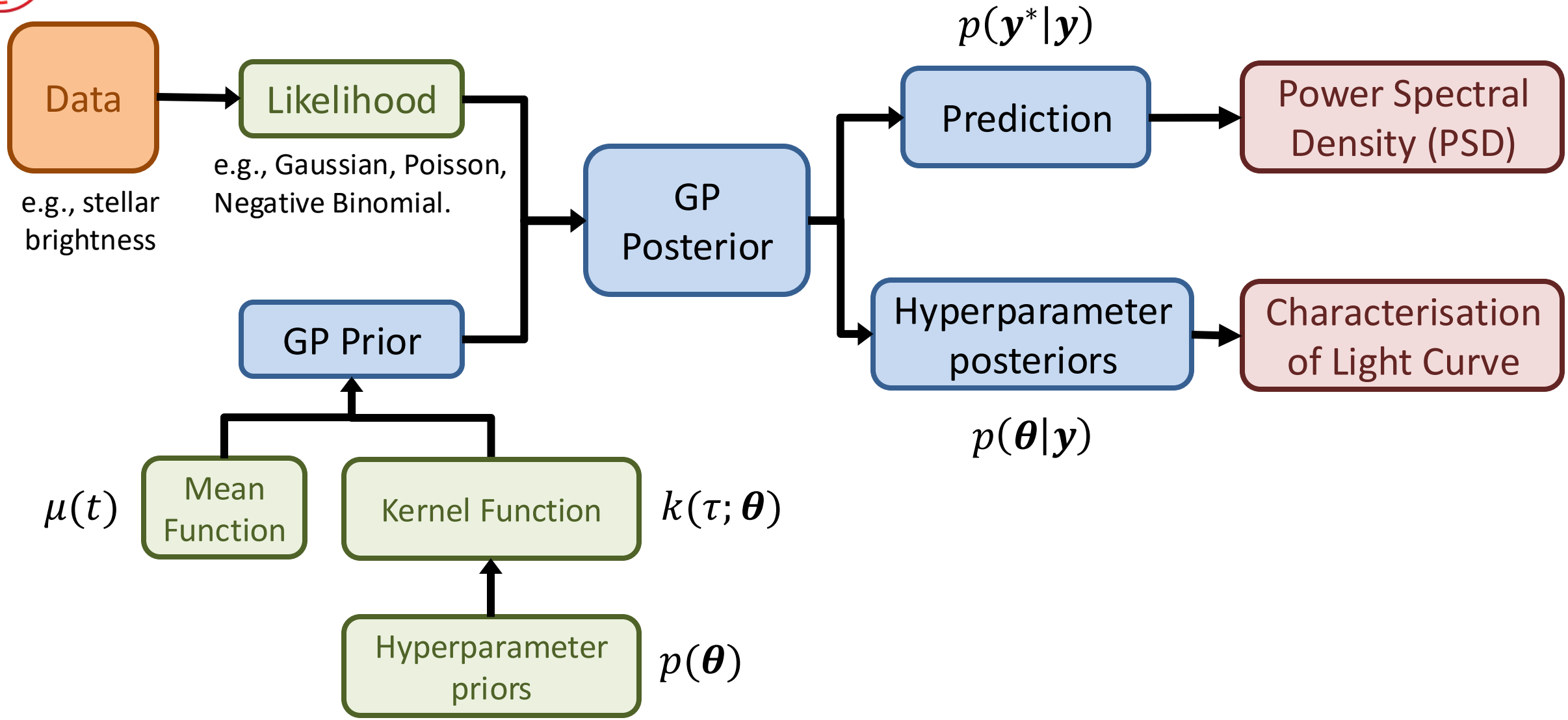



$$\kappa(\tau;\sigma,\ell) = \sigma^2 \left(1 + \sqrt{3}\frac{\tau}{\ell}\right) \exp\left\{-\sqrt{3}\frac{\tau}{\ell}\right\}$$ Matern 3/2




$$\kappa(\tau;\sigma,\ell,T) = \sigma^2 \exp\left\{-\frac{2}{\ell^2}\sin^2\left(\pi\frac{\tau}{T}\right)\right\}$$ Periodic

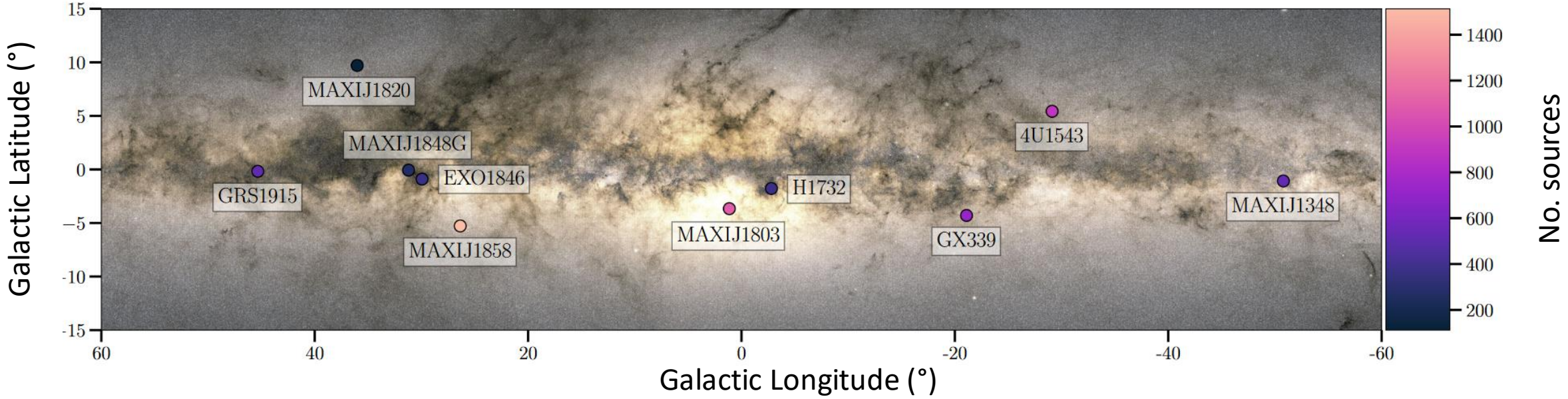
# Modelling Workflow



# ThunderKAT Survey

- The HUNt for Dynamic and Explosive Radio transients with MeerKAT
- Field of view of ≈ 1 square degree
- 6,394 radio light curves over 10 fields
- Flux density measurements + standard errors

MeerKAT Radio Telescope (Credit: SARAO)



# Gaussian Process Model

$$\mathbf{Y} \sim \mathrm{N}(f, \hat{\boldsymbol{e}}^2)$$

$$f \sim \mathrm{GP}(\mathbf{0}, \boldsymbol{K}_{N\times N})$$

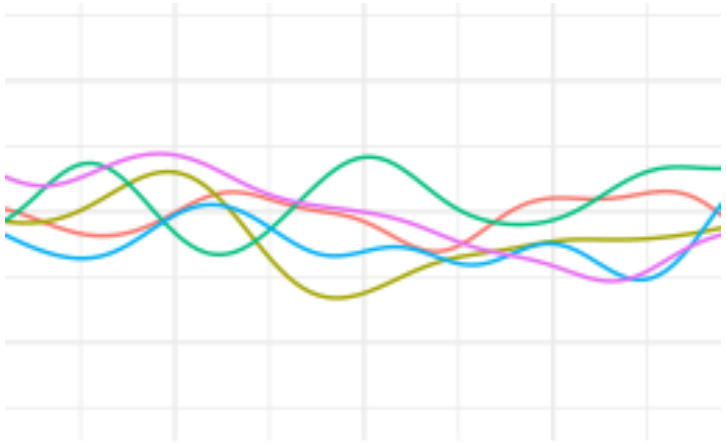
$$\boldsymbol{K}_{rc} = \kappa(t_r, t_c|\boldsymbol{\theta})$$

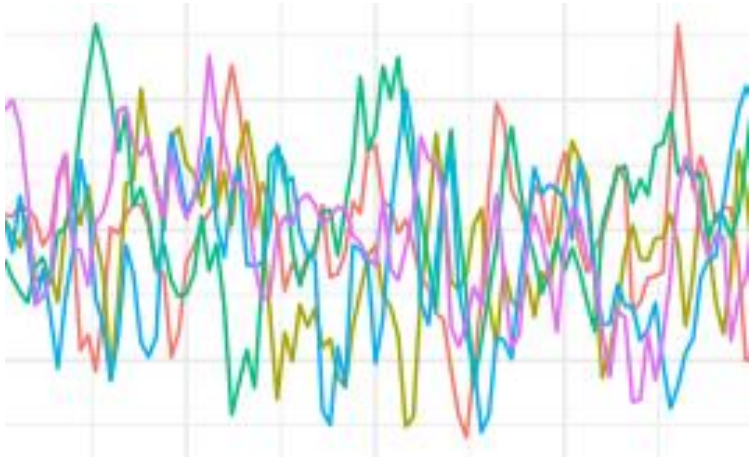
$$r, c = 1, \ldots, N.$$

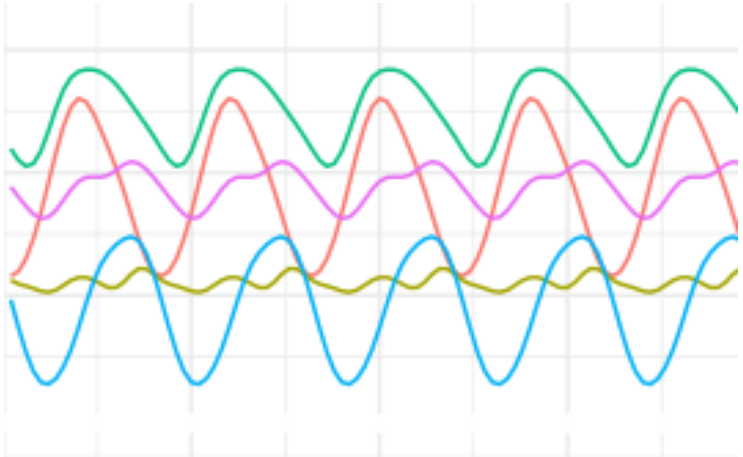
$$\boldsymbol{\theta} = (\sigma_{SE}, \ell_{SE}, \sigma_{M32}, \ell_{M32}, \sigma_P, \ell_P, T)$$

$$= \boxed{\kappa_1(\tau; \sigma_{SE}, \ell_{SE})}_{\text{Squared Exponential}} + \boxed{\kappa_2(\tau; \sigma_{M32}, \ell_{M32})}_{\text{Matern 3/2}} + \boxed{\kappa_3(\tau; \sigma_P, \ell_P, T)}_{\text{Periodic}}$$

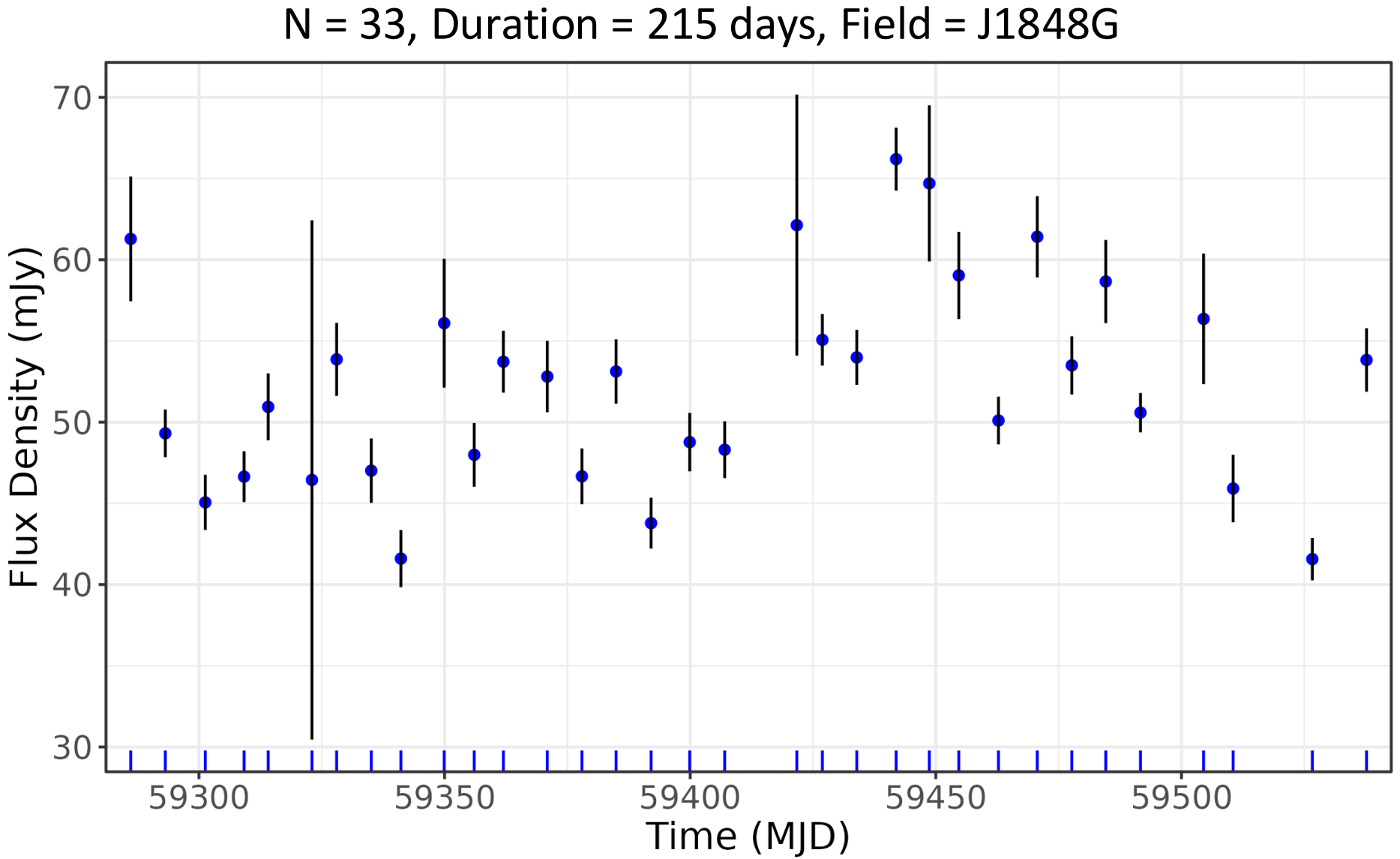
Covariance Kernel





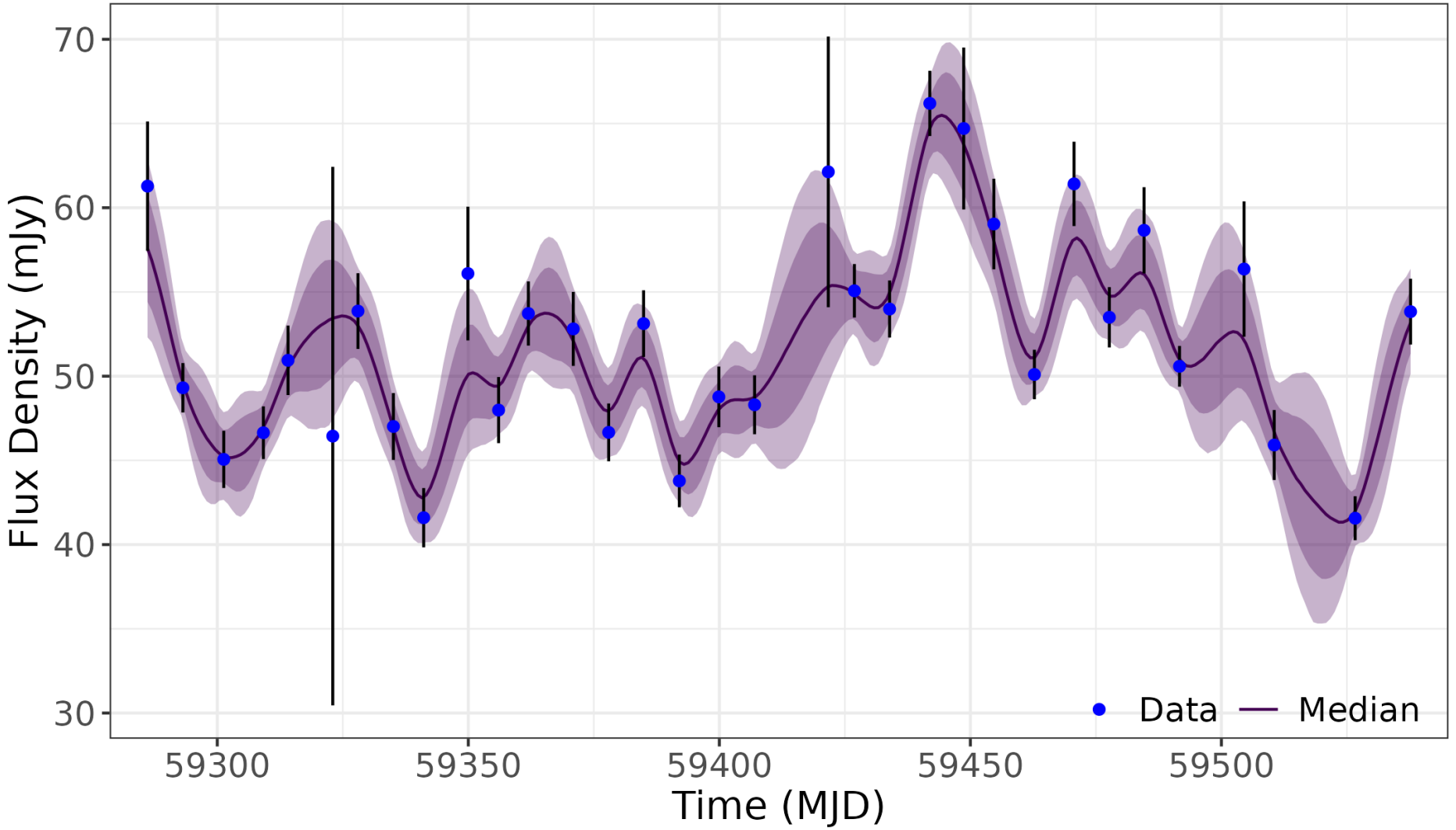


# GP Fitting Example



# Posterior Predictive Samples

N = 33, Duration = 215 days, Field = J1848G




$$\sigma_{\mathrm{SE}} = 0.35$$

$$\ell_{\mathrm{SE}} = 48.5$$

$$\sigma_{\mathrm{M32}} = 1.20$$

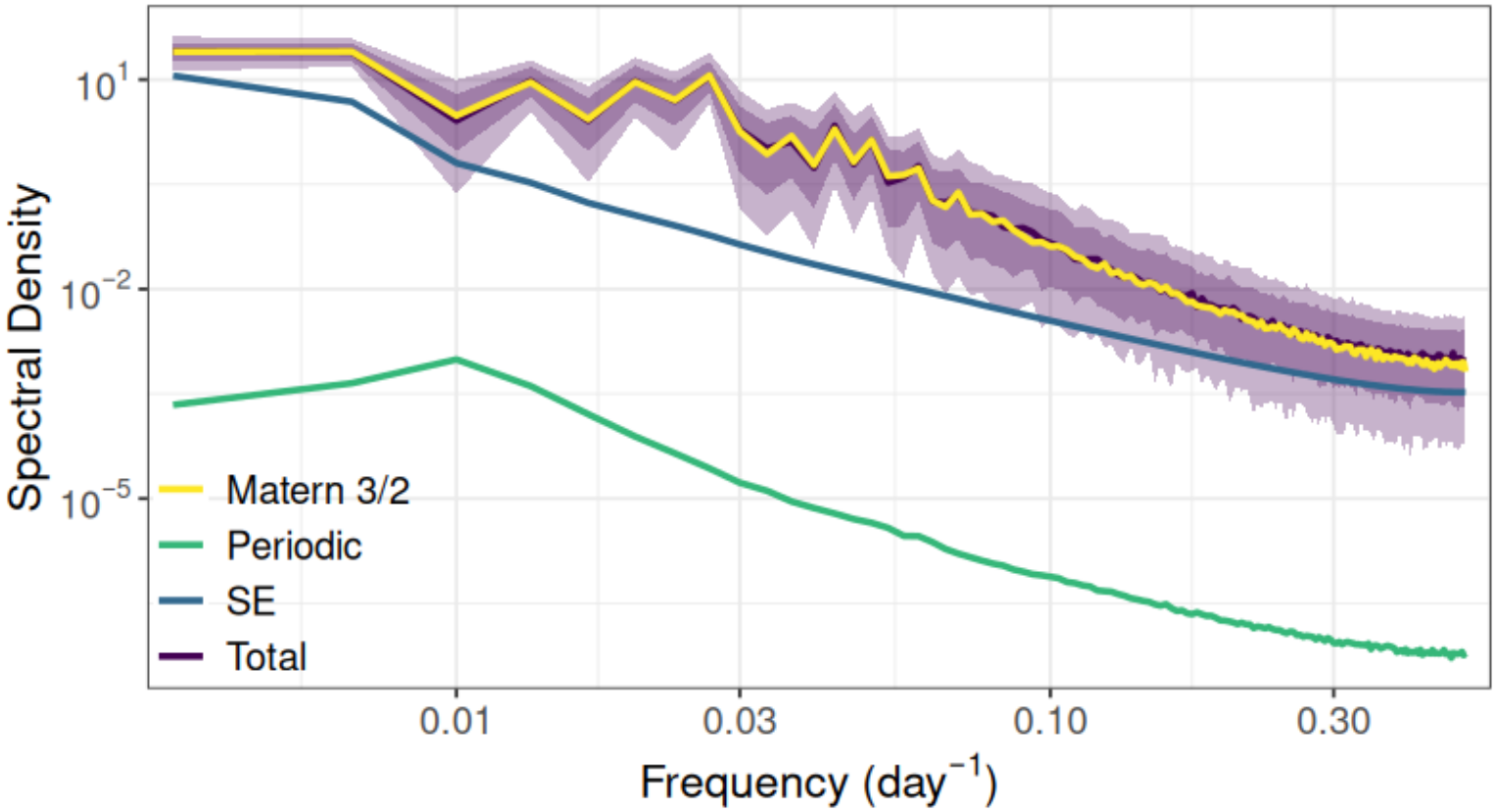
$$\ell_{\mathrm{M32}} = 12.5$$

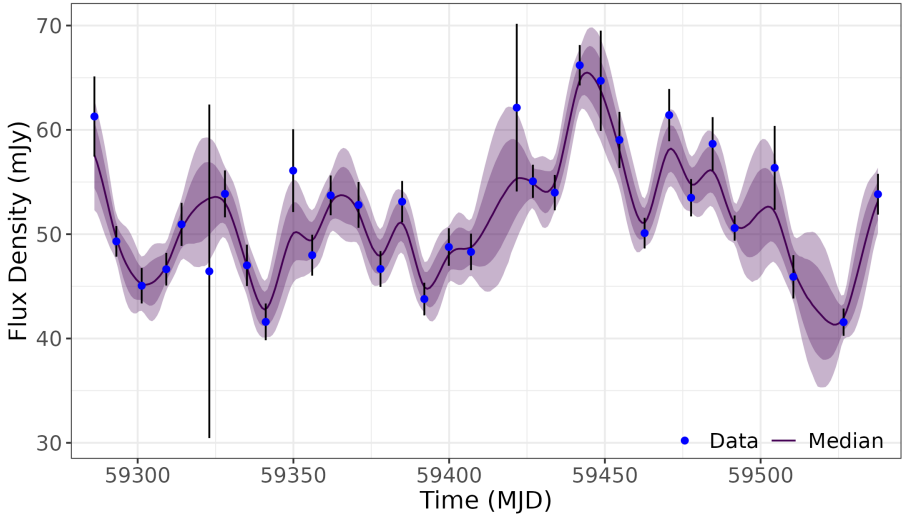
$$\sigma_{\mathrm{P}} = 0.45$$

$$\ell_{\mathrm{P}} = 37.6$$

$$T = 85.6$$

# Power Spectral Density (PSD)

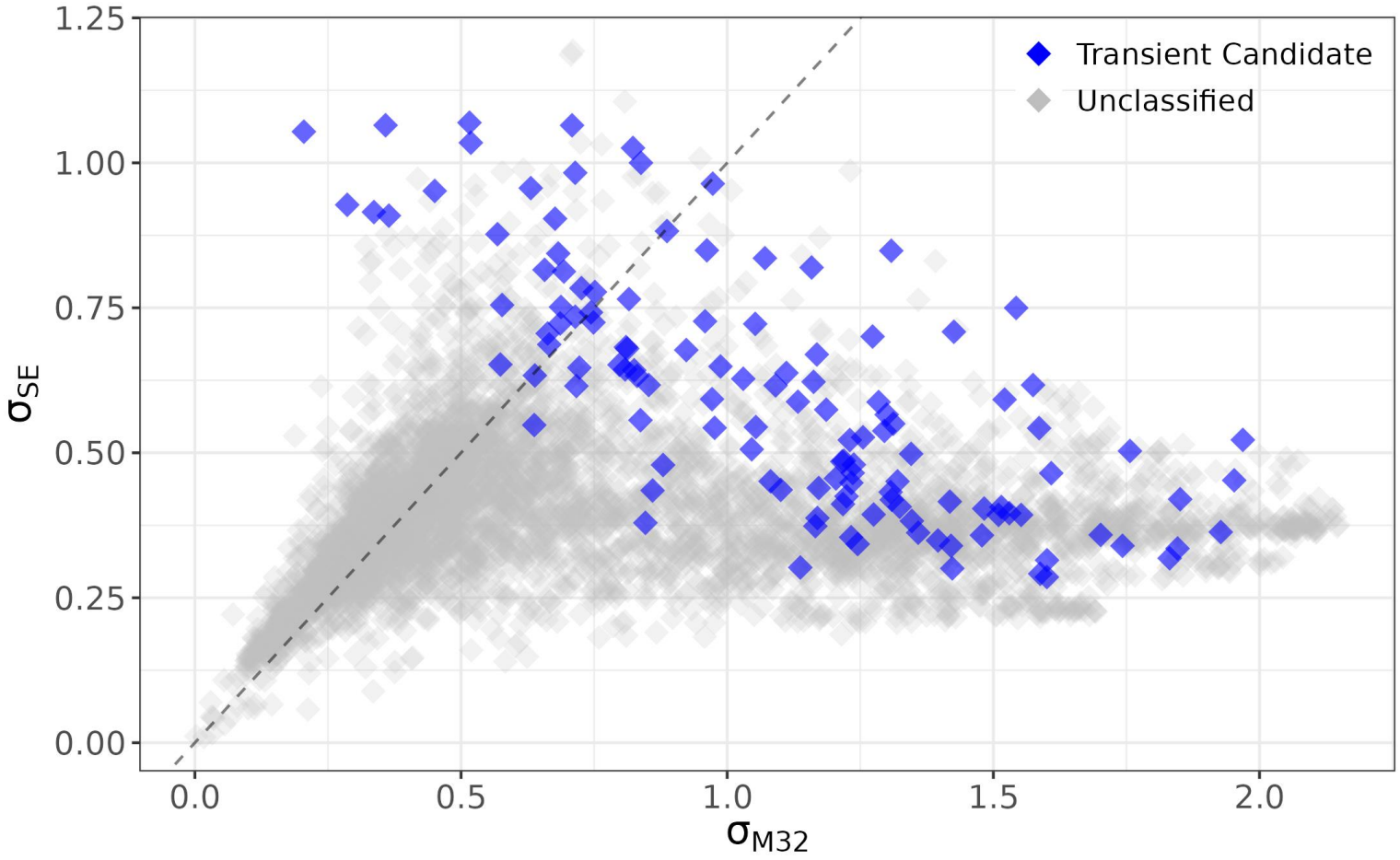



- Matern term dominates
- Very weak periodic term

# Interpreting Amplitude as Transcience

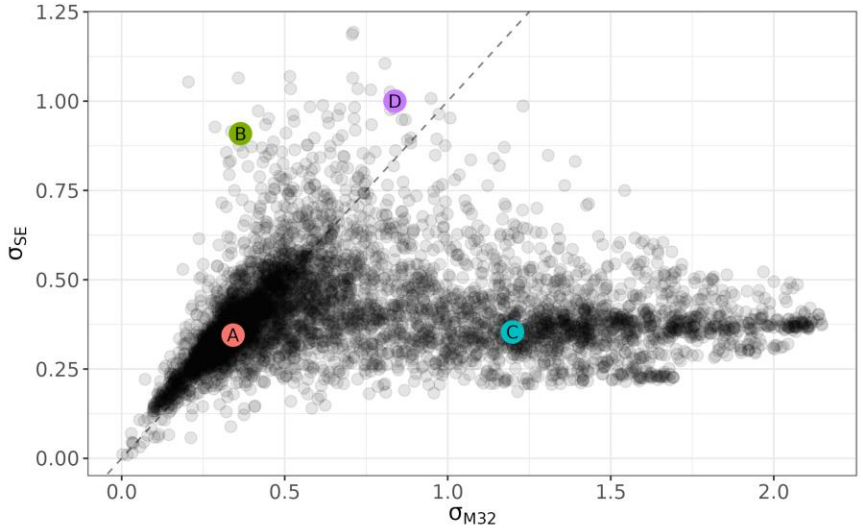
- Transience seems to manifest as large values in **amplitude**, $\sigma$.
- Previously identified transient candidates all seem to lie the upper right of this parameter space.

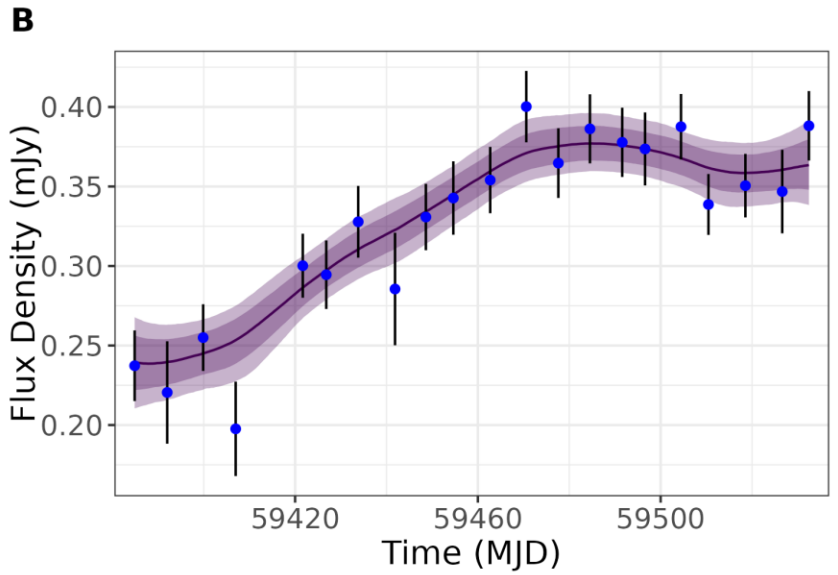


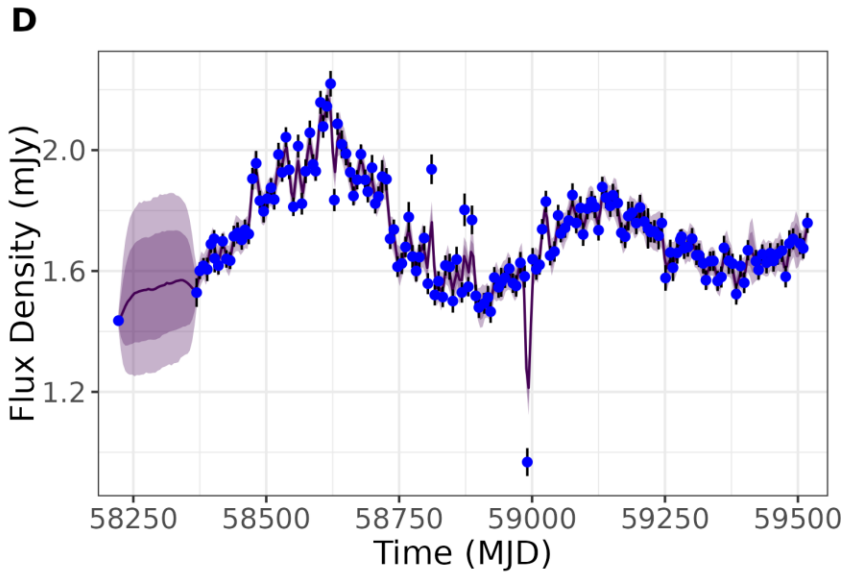

Data: Andersson et al. (2023)

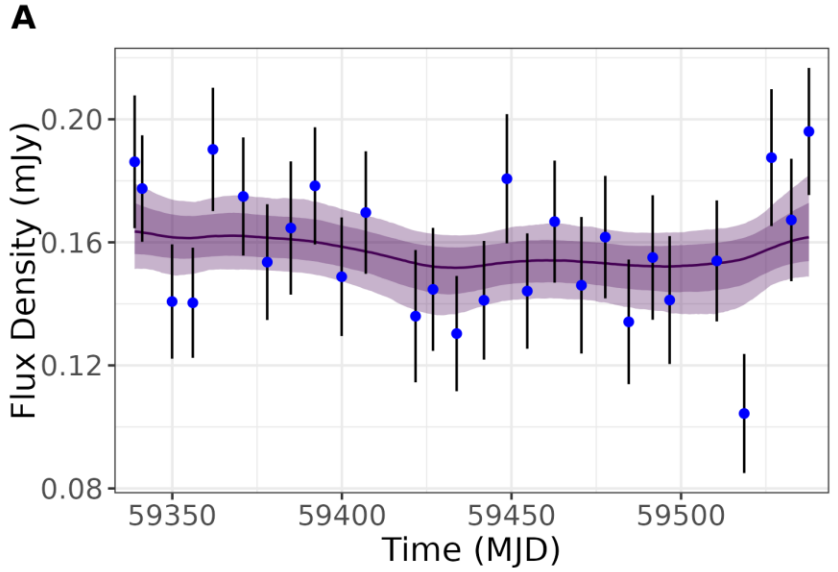
Figure: Fu et al. (in prep.)

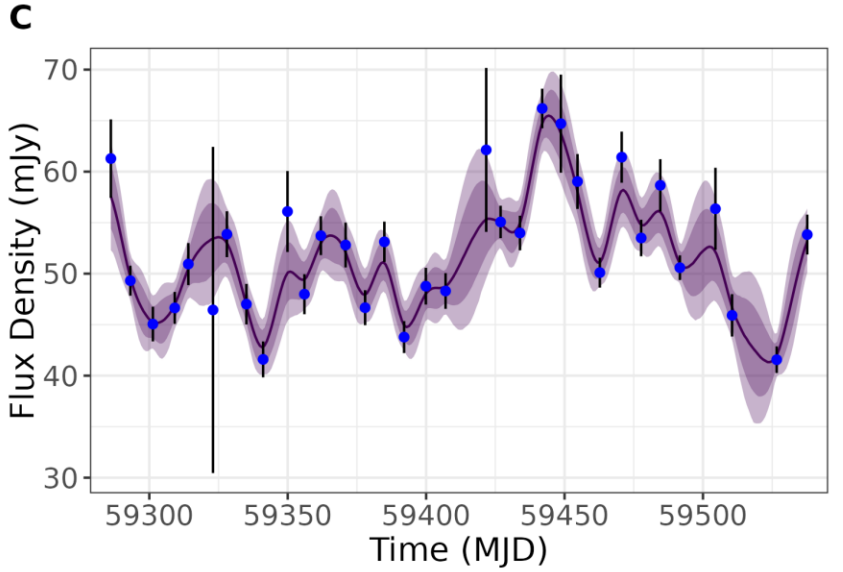
# Explore the hyperparameter space



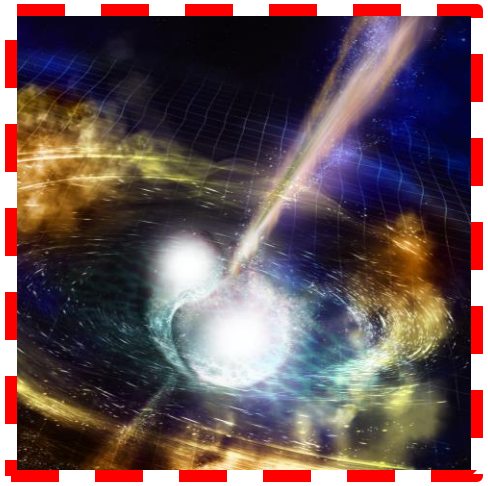



# Twinkle twinkle little star…



Raw Data Processing → Identify → Classify

Exotic phenomena

Large-scale survey

$10^3$ to $10^6$ light curves

Transient candidates

Black holes, supernova, eclipsing binary, GRB, FRB, AGN, etc, …

**… a Gaussian Process is what you are!**

# Tools

- Implemented in `Python`[1] (v3.10) and `PyMC`[2] (v3.5.2)
  - Accessible to astronomers
  - Probabilistic programming framework
  - Well-maintained open-source software
- Repeated analyses in `R`[3] (v4.3.1) and `Stan`[4] (v2.34)
- Also considered: `celerite2`[5], `george`[6].

1. https://www.python.org
2. https://www.pymc.io
3. https://cran.r-project.org/
4. https://mc-stan.org/
5. https://celerite2.readthedocs.io/en/latest/
6. https://george.readthedocs.io/en/latest/