Robust characterisation of transient radio light curves using Gaussian process regression

Shih Ching Fu

shihching.fu@postgrad.curtin.edu.au

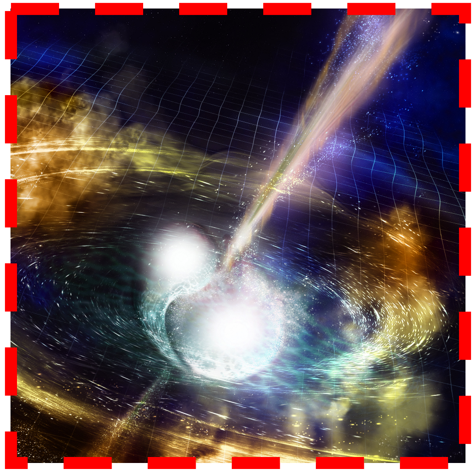
Supervisors:
Dr Arash Bahramian, Dr Aloke Phatak, Dr James Miller-Jones, Dr Suman Rakshit

International Centre for Radio Astronomy Research

Statistical Society of Australia

ICRAR is a joint venture between Curtin University and The University of Western Australia and receives support from the Western Australian and Australian Governments.

Twinkle twinkle little star...



Exotic phenomena

Large-scale survey

Raw Data Processing

10^3 to $>10^6$ sources

Identify

Transient candidates

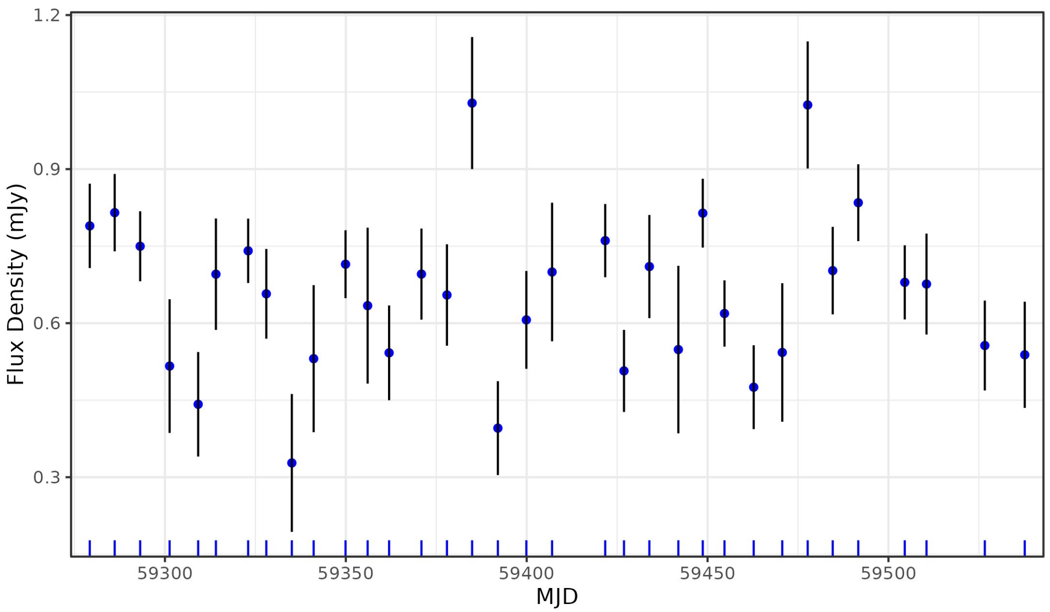
Classify

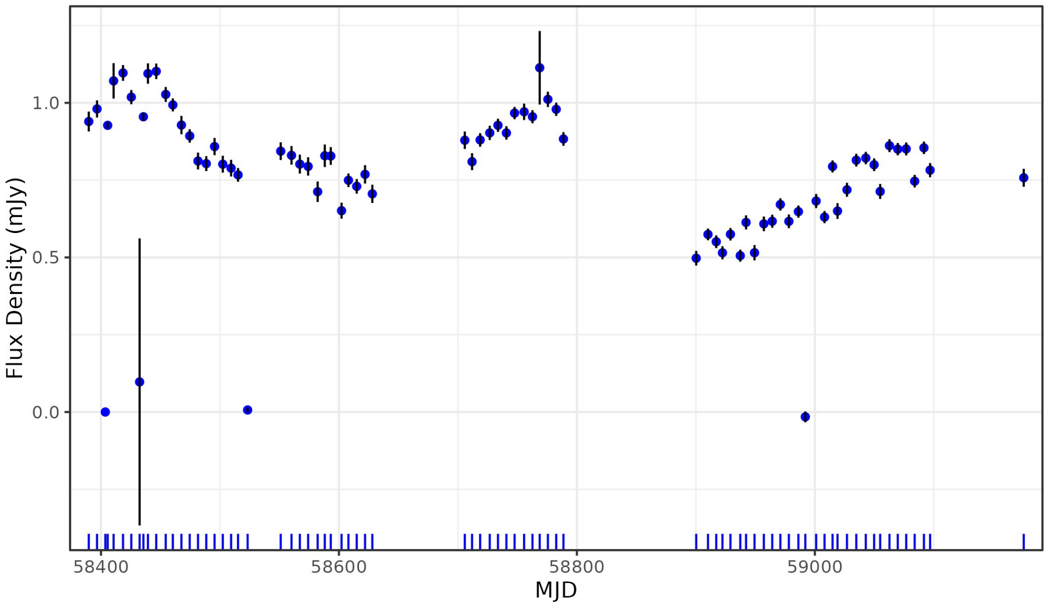
... how I wonder what you are?

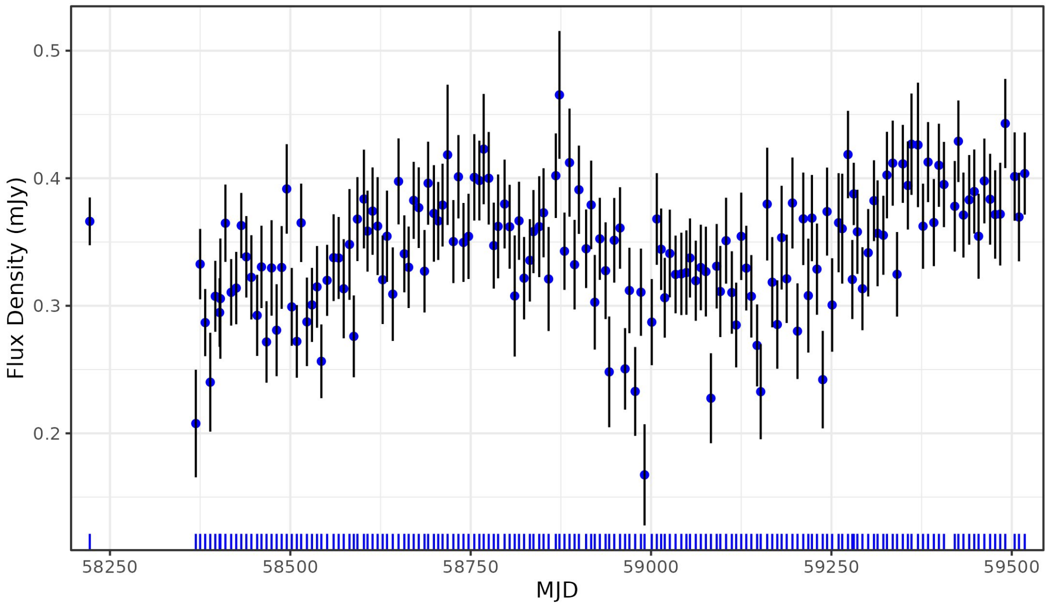
Encode the variability of a source in a way that is **concise**, **robust** and **descriptive**.

Data are highly varied

- Different cadences and sparsity
- Uneven sampling rates
- Varying noise levels
- Diverse types of objects



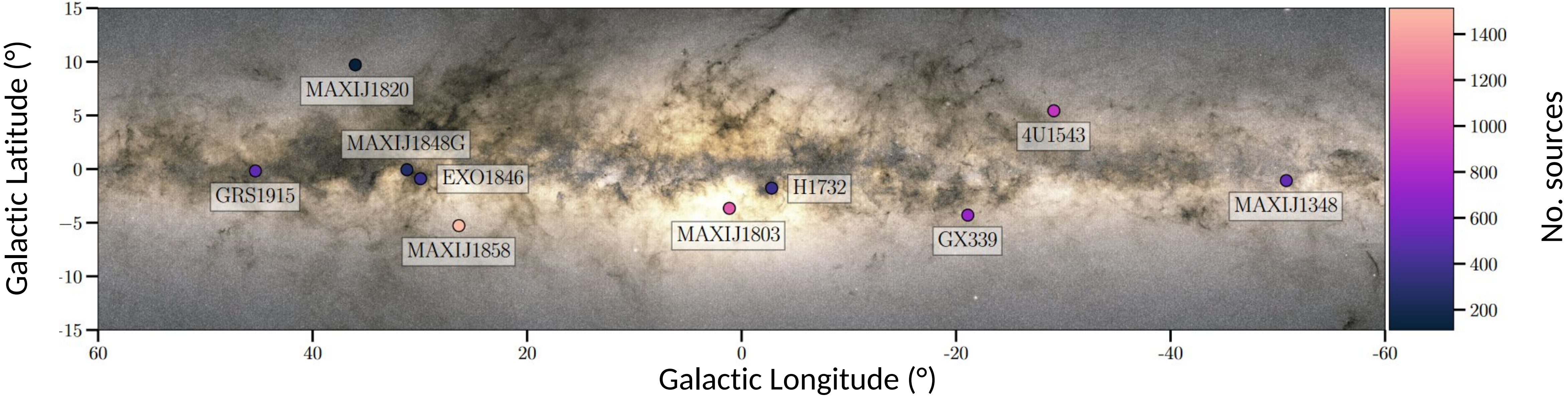




ThunderKAT Survey

- MeerKAT Image-domain transients survey
- Field of view of ≈1 square degree
- 6,394 radio light curves over 10 fields
- Flux densities and standard errors
- Credit: Andersson et al. (2023)

MeerKAT Radio Telescope (Credit: SARAO)



V_ν and η_ν

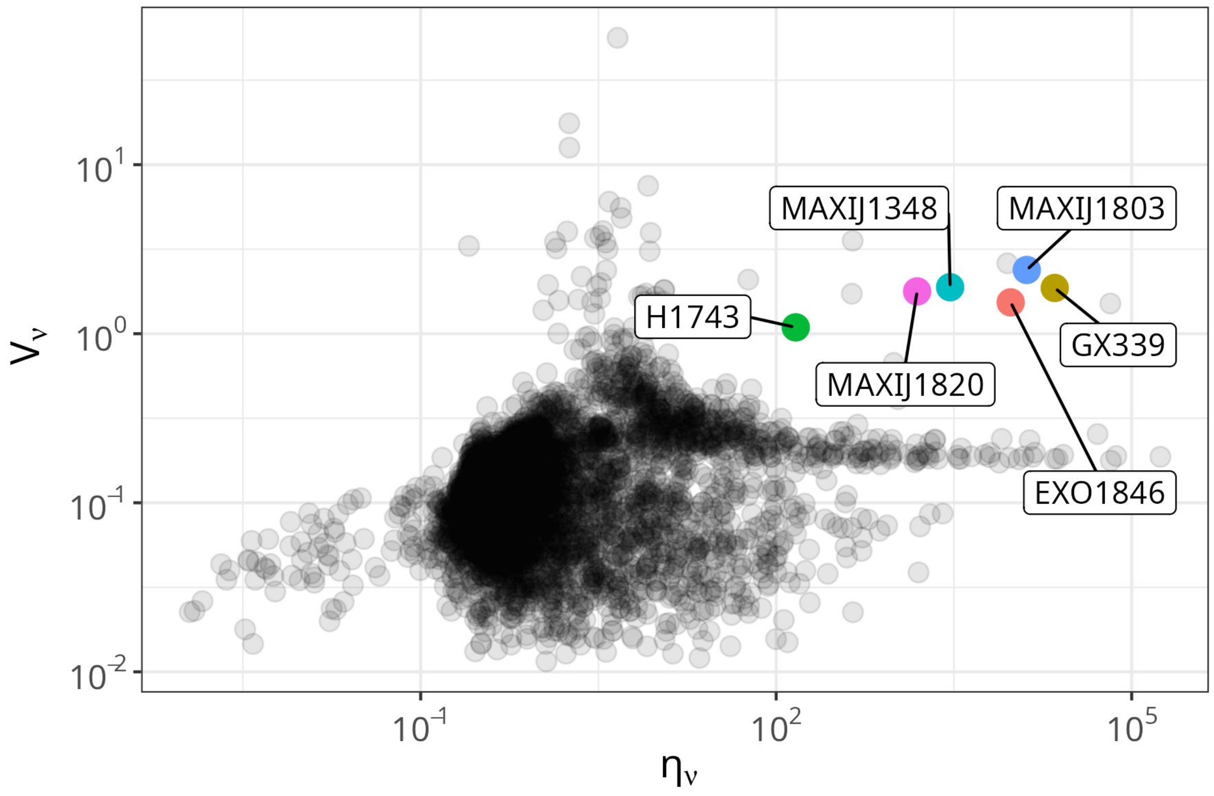
- Coefficient of Variation

$$V_\nu = \frac{s}{\overline{S_\nu}}$$

- Reduced χ^2 statistic of variability

$$\eta_\nu = \frac{1}{n-1}\sum_{i=1}^{n}\frac{\left(S_{i,\nu} - \overline{S_\nu}^*\right)^2}{\sigma_{i,\nu}^2} \sim \chi_{n-1}^2$$

$$\overline{S_\nu}^* = \frac{\sum_{i=1}^{n} w_{i,\nu} S_{i,\nu}}{\sum_{i=1}^{n} w_{i,\nu}} \qquad w_{i,\nu} = 1/\sigma_{i,\nu}^2 \qquad i = 1, \ldots, n.$$



Characterising Light Curves

Oversimplified

- Fewer parameters
- Scales easily
- High information loss

Overspecified

- Many parameters
- High discriminatory power
- Overfitting

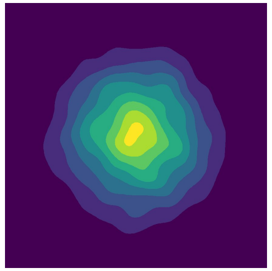
Model a light curve as a **Gaussian Process (GP)**

Multivariate Normal $\boldsymbol{Y} \sim N_n(\mathbf{0}, \boldsymbol{\Sigma}_{n\times n})$

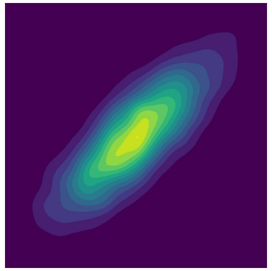
$\boldsymbol{Y}$ is a vector of n Gaussian random variables.

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{n\times n}), \qquad \boldsymbol{\Sigma}_{n\times n} = \begin{bmatrix} \Sigma_{11} & \cdots & \Sigma_{1n} \\ \vdots & \ddots & \vdots \\ \Sigma_{n1} & \cdots & \Sigma_{nn} \end{bmatrix}$$

where $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)$ and $\boldsymbol{\Sigma}$ is a $n \times n$ covariance matrix.



$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$

- Symmetric, positive semi-definite matrix.
- Linear combinations of covariance matrices are also valid covariance matrices.

Gaussian Processes

Extend multivariate Gaussian to ‘infinite’ dimensions.

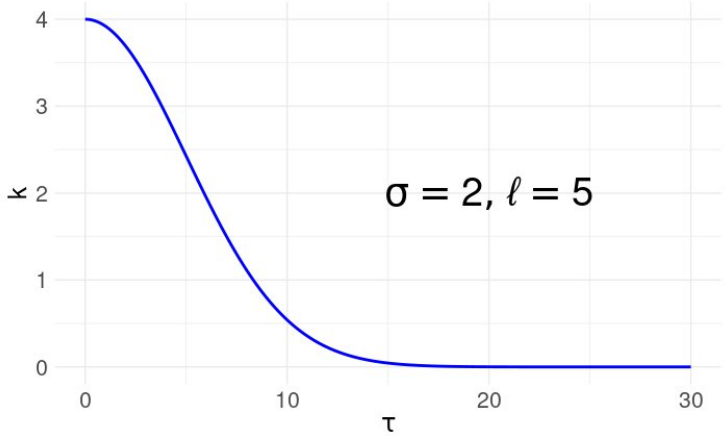
- Mean function, $\mu(t)$
- Covariance or **kernel function**, $\boldsymbol{\kappa(t, t)}$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \end{bmatrix} = \boldsymbol{Y} \sim GP\left(\boldsymbol{\mu}, \boldsymbol{K}\right)$$

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

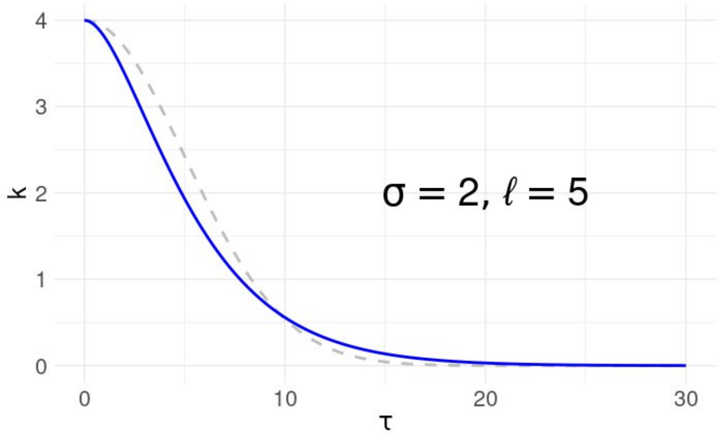
Rather than specifying a fixed covariance matrix with fixed dimensions, compute covariances using the kernel function.

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



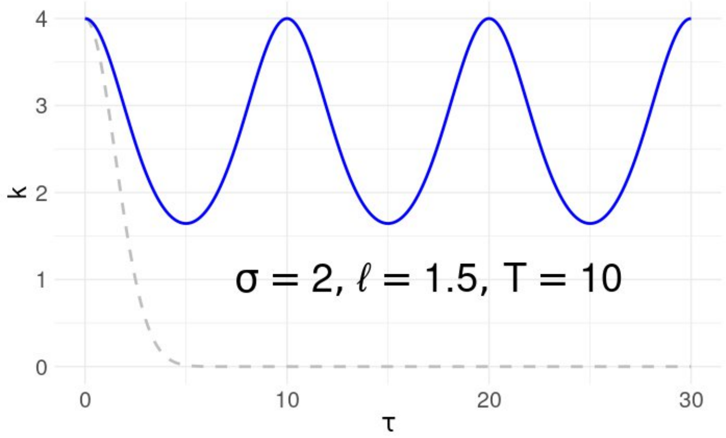
$$\kappa(\tau; \sigma, \ell) = \sigma^2 \exp\left\{-\frac{1}{2}\left(\frac{\tau}{\ell}\right)^2\right\}$$

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

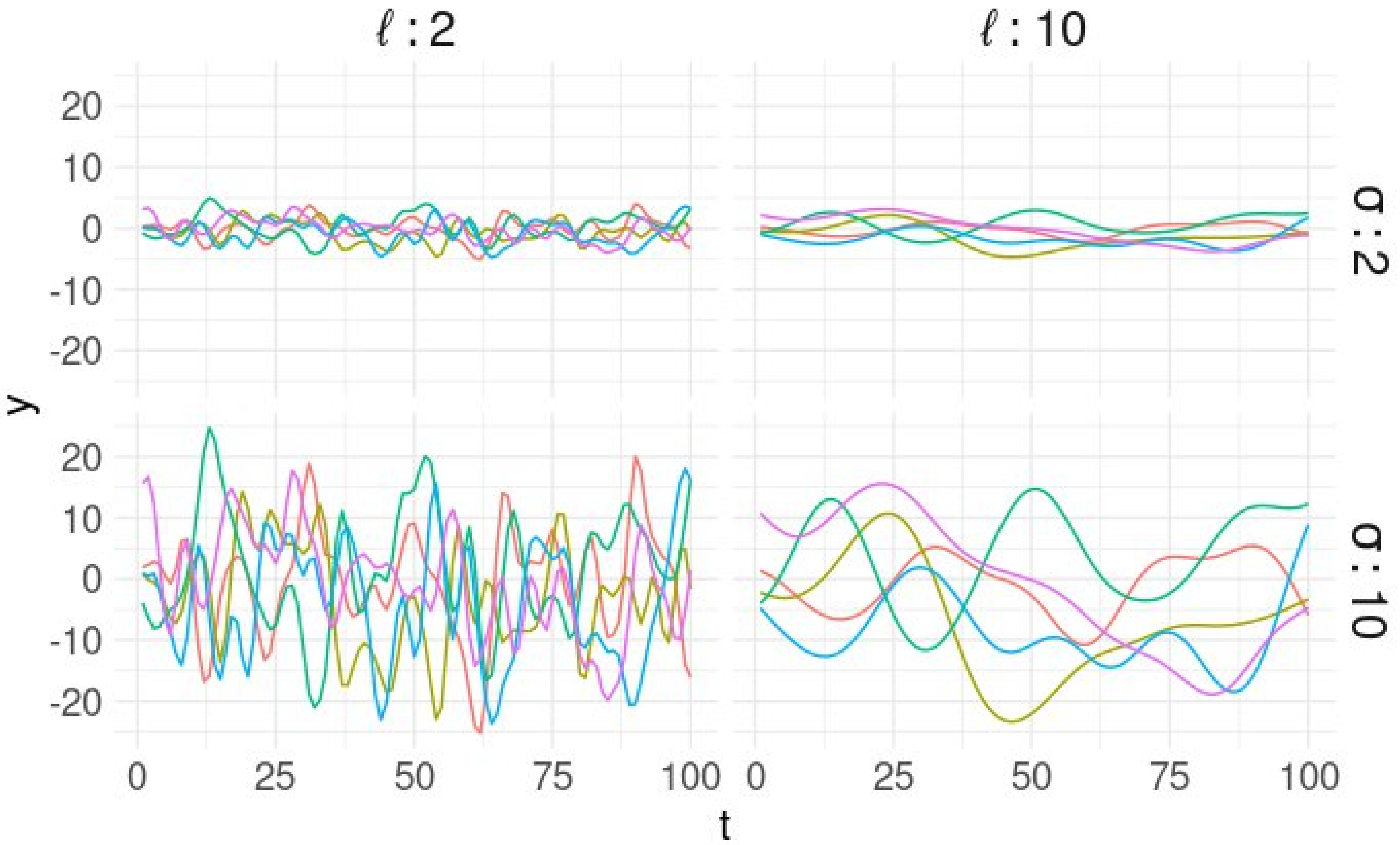


$$\kappa(\tau; \sigma, \ell, T) = \sigma^2 \exp\left\{-\frac{2}{\ell^2}\sin^2\left(\pi\frac{\tau}{T}\right)\right\}$$

Periodic

Squared Exponential Kernel

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



Bayesian Hierarchical Model

$S_i \sim N\left(f_i, \hat{e}_i^2\right)$ Observed flux density

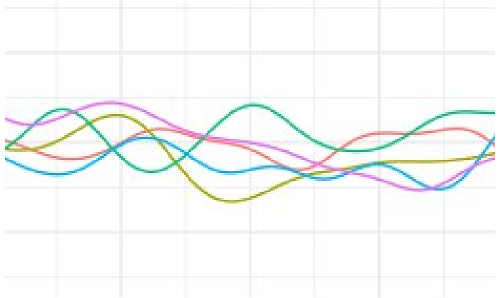
$i, r, c = 1, \dots, n.$

$\boldsymbol{f} \sim GP\left(\boldsymbol{0}, \boldsymbol{K}_{n\times n}\right)$ Gaussian process prior

$\tau = |t_r - t_c|$

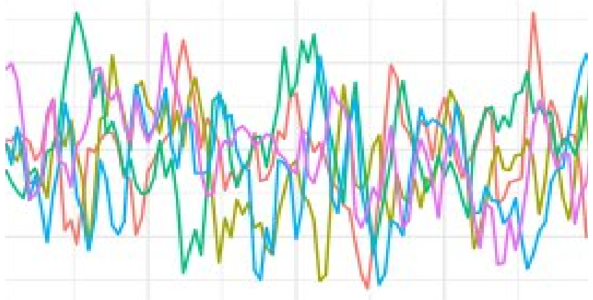
$[\boldsymbol{K}]_{rc} = \kappa\left(t_r, t_c | \boldsymbol{\theta}\right)$ Covariance Kernel

$$= \kappa_1\left(\tau | \sigma_{SE}, \ell_{SE}\right) + \kappa_2\left(\tau | \sigma_{M32}, \ell_{M32}\right) + \kappa_3\left(\tau | \sigma_P, \ell_P, T\right)$$



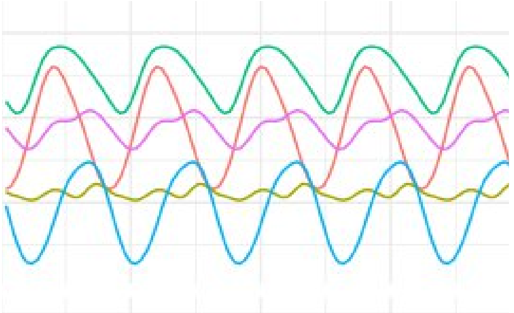
Squared Exponential

+



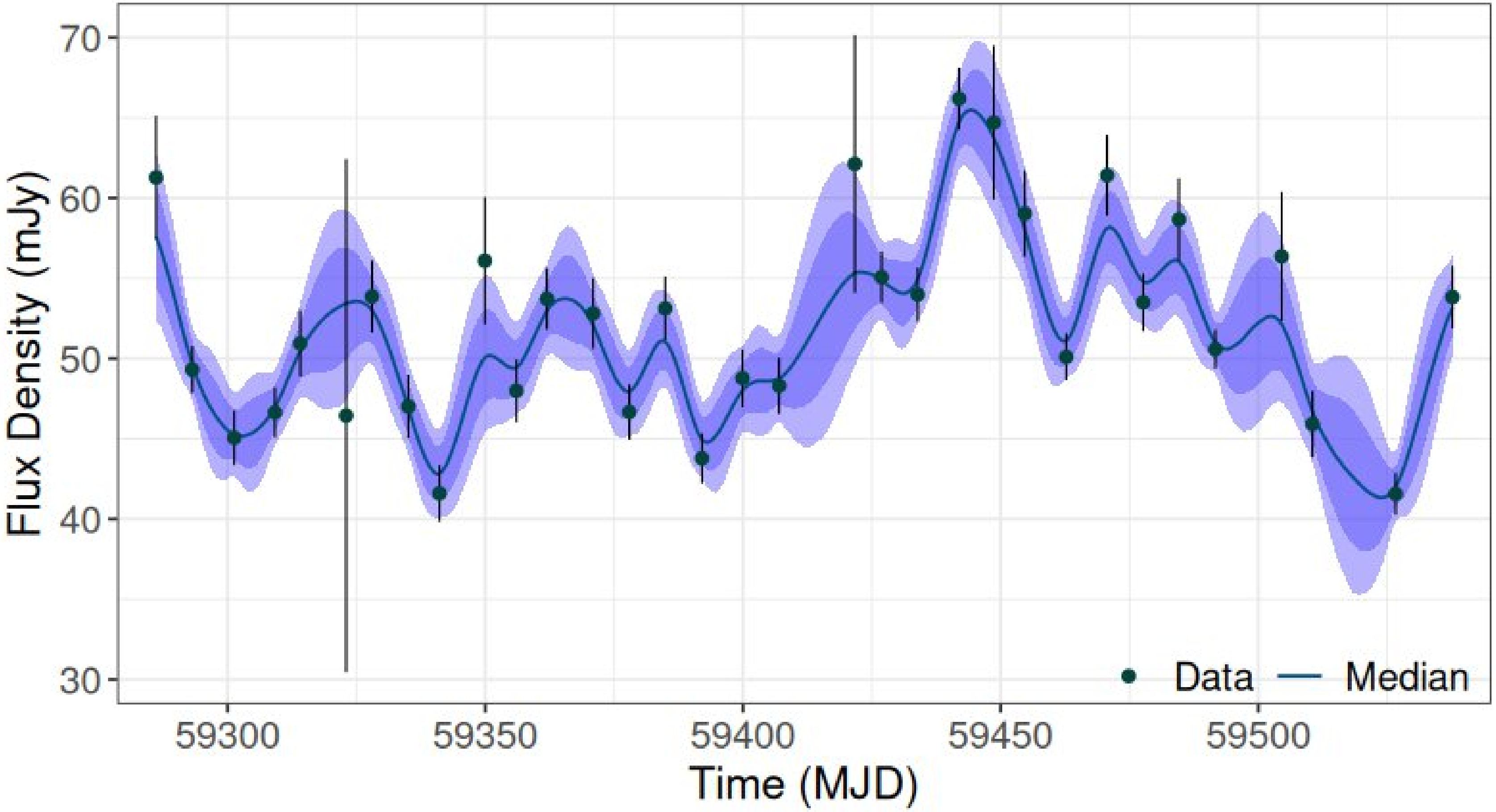
Matern 3/2

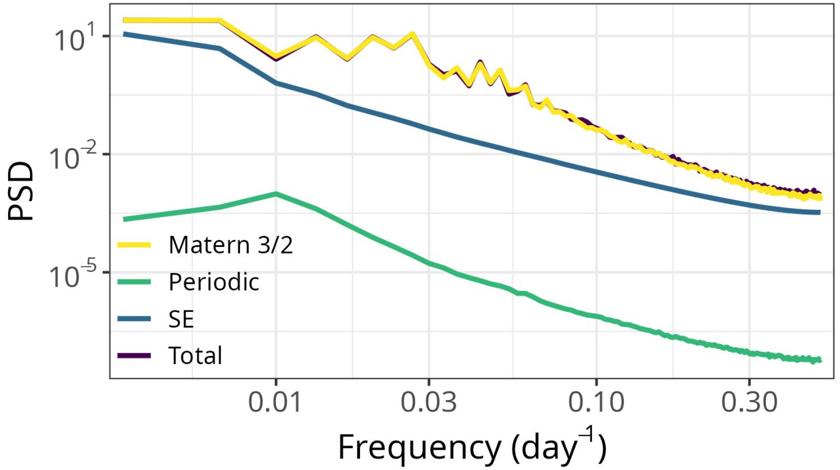
+



Periodic

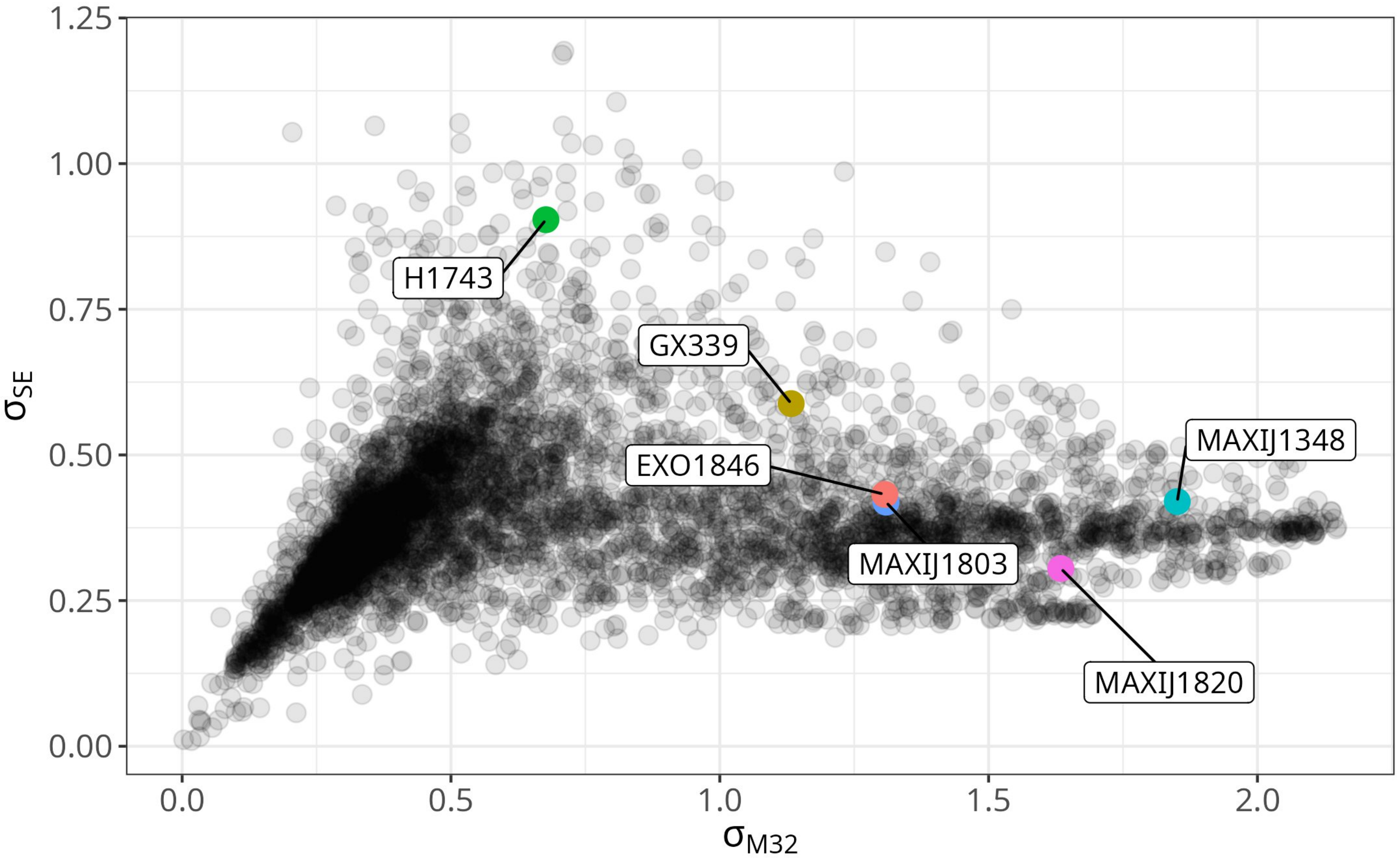
Posterior Predictive Curve & PSD



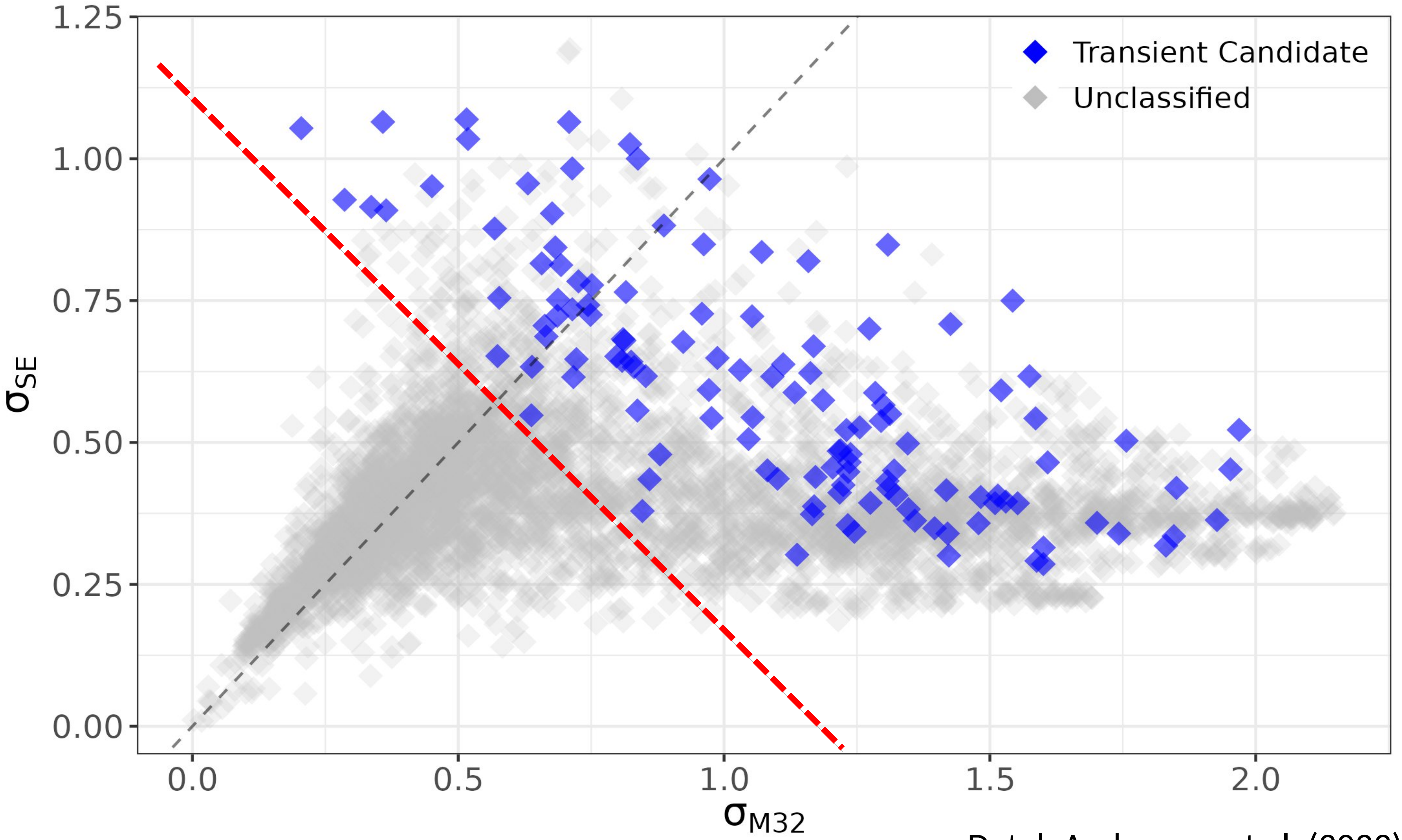


$$\sigma_{M32} = 1.20,\ \ell_{M32} = 12.5,\ \sigma_{SE} = 0.35,\ \ell_{SE} = 48.5,\ \sigma_P = 0.45,\ \ell_P = 37.6,\ T = 85.6$$

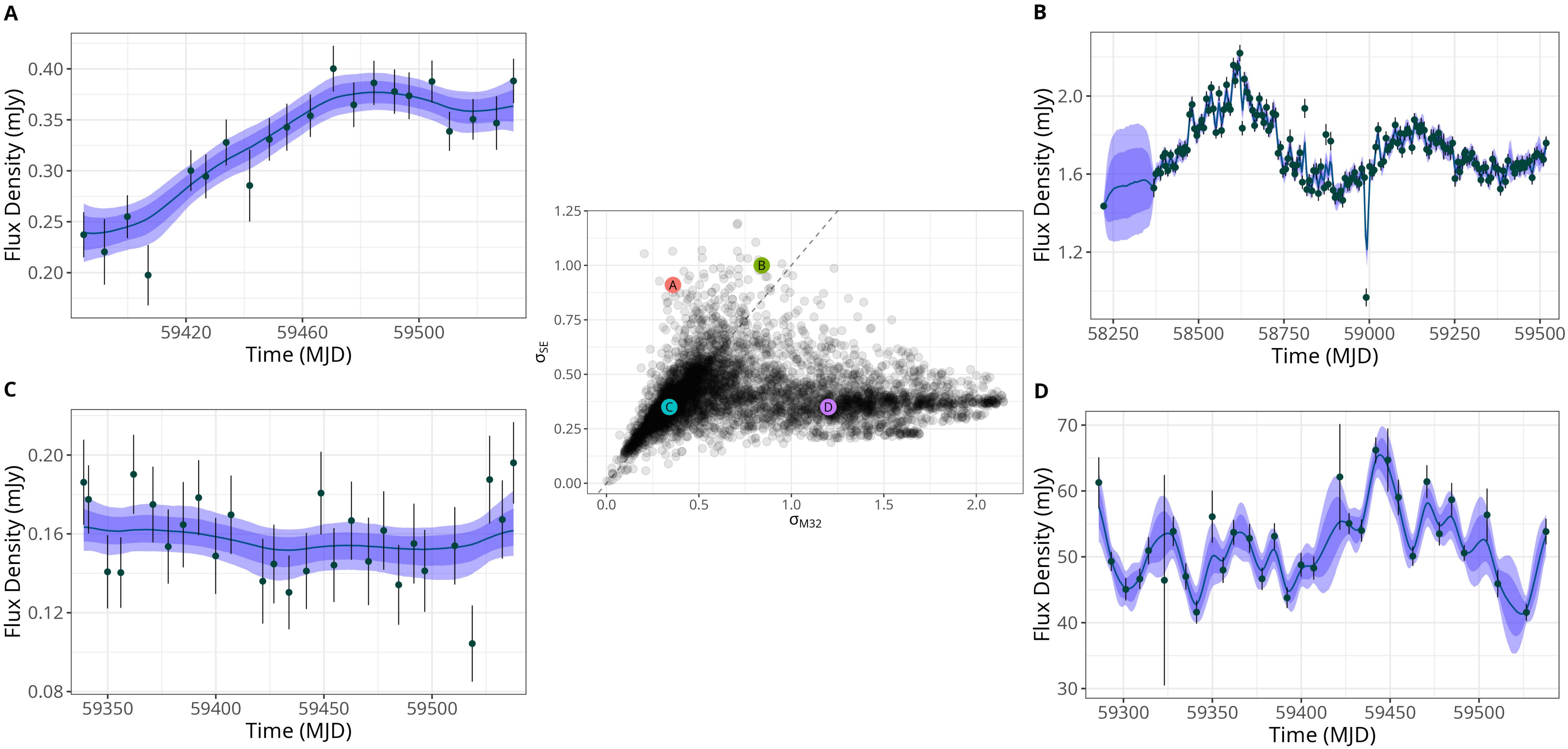
Amplitude Hyperparameters $(\sigma_{M32}, \sigma_{SE})$



$(\sigma_{M32}, \sigma_{SE})$ as a transience metric



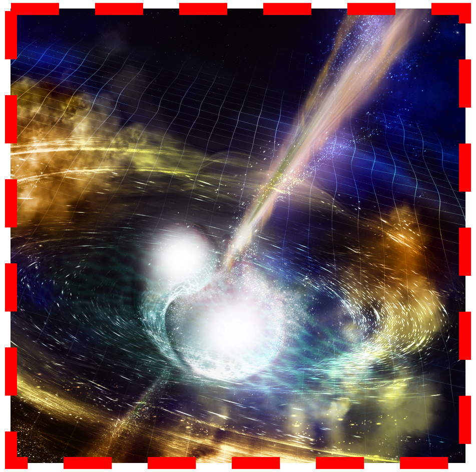
Datal: Andersson et al. (2023)


A
Flux Density (mJy)
0.40
0.35
0.30
0.25
0.20
59420
59460
59500
Time (MJD)
B
2.0
1.6
1.2
58250
58500
58750
59000
59250
59500
C
0.20
0.16
0.12
0.08
59350
59400
59450
59500
D
70
60
50
40
30
59300
59350
59400
59450
59500
σ_{SE}
σ_{M32}
1.25
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
2.0

Summary

- To help identify transients in large surveys we want a statistical representation that is **concise**, **robust** and **descriptive**.
- Gaussian process models can handle situations where data quality is varied and object types are unknown *a priori*.
- We can tune multi-term kernels for different behaviours and scales.
- Kernel **amplitude** hyperparameters (σ_{M32}, σ_{SE}) seem to encode the transience of sources; consistent with the judgement of volunteer citizen scientists.
- This model needs to be tested on data from other surveys.
- Upcoming: extend from single- to multi-band light curves.

Twinkle twinkle little star…



Raw Data Processing → Identify → Classify

Exotic phenomena

Large-scale survey

10^3 to $>10^6$ sources

Transient candidates

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

… a Gaussian Process is what you are!