Gaussian process regression for identifying variables and transients in ThunderKAT

Shih Ching Fu

Co-supervisors:

Dr Arash Bahramian, Dr Aloke Phatak

Associate Supervisors:

Dr James Miller-Jones, Dr Suman Rakshit

ICRAR
International Centre for Radio Astronomy Research

Curtin University

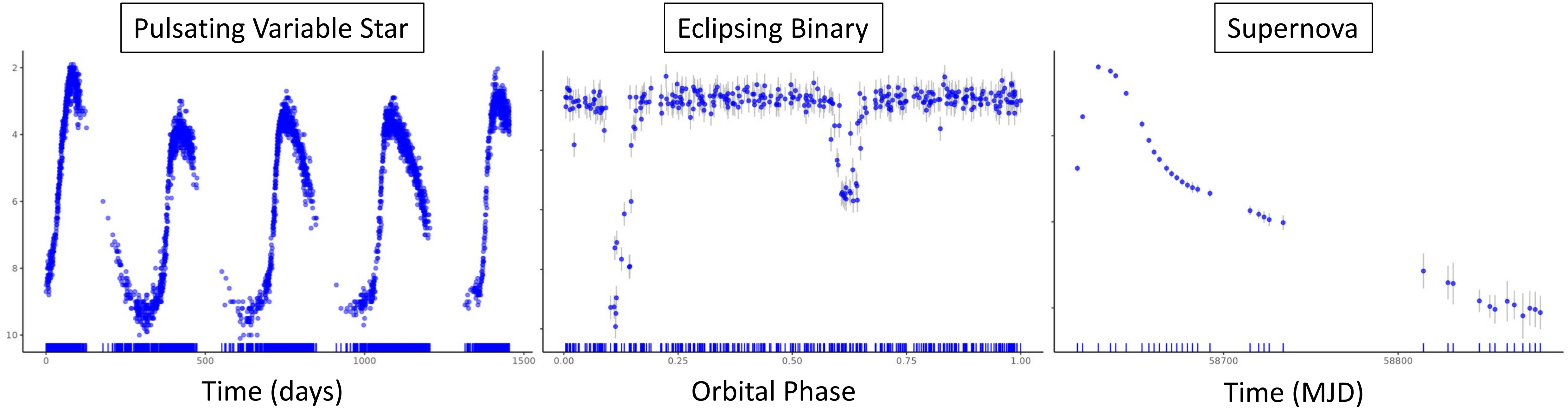
The University of Western Australia

Statistical Society of Australia

Light curves as stochastic processes

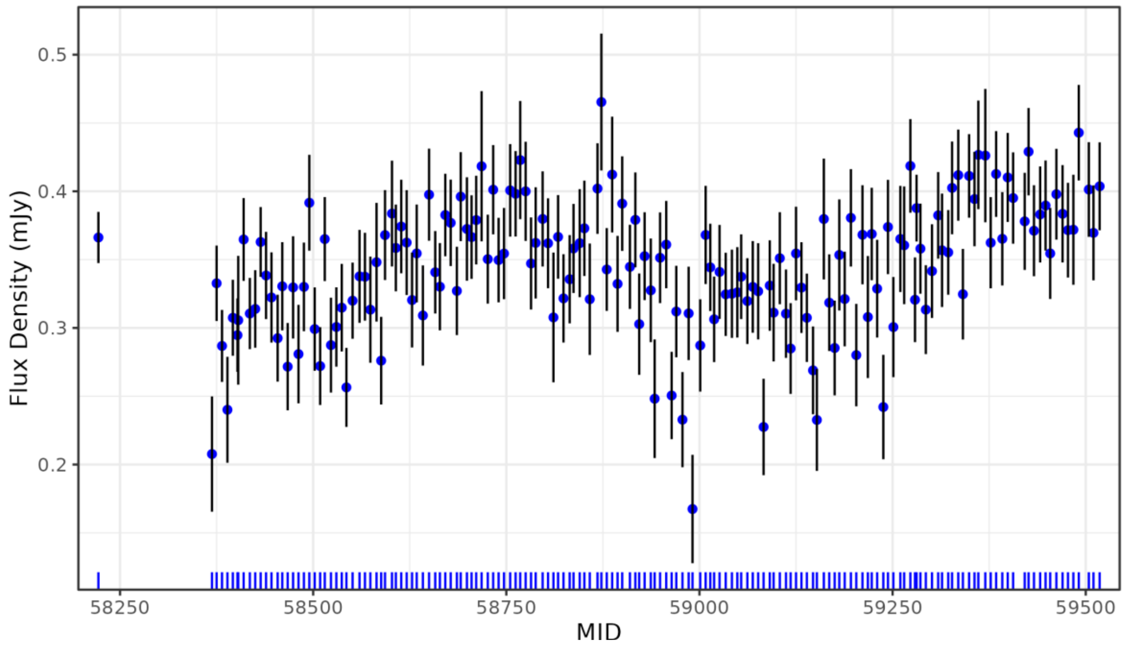
Light curves are time series describing the brightness of a source over time.

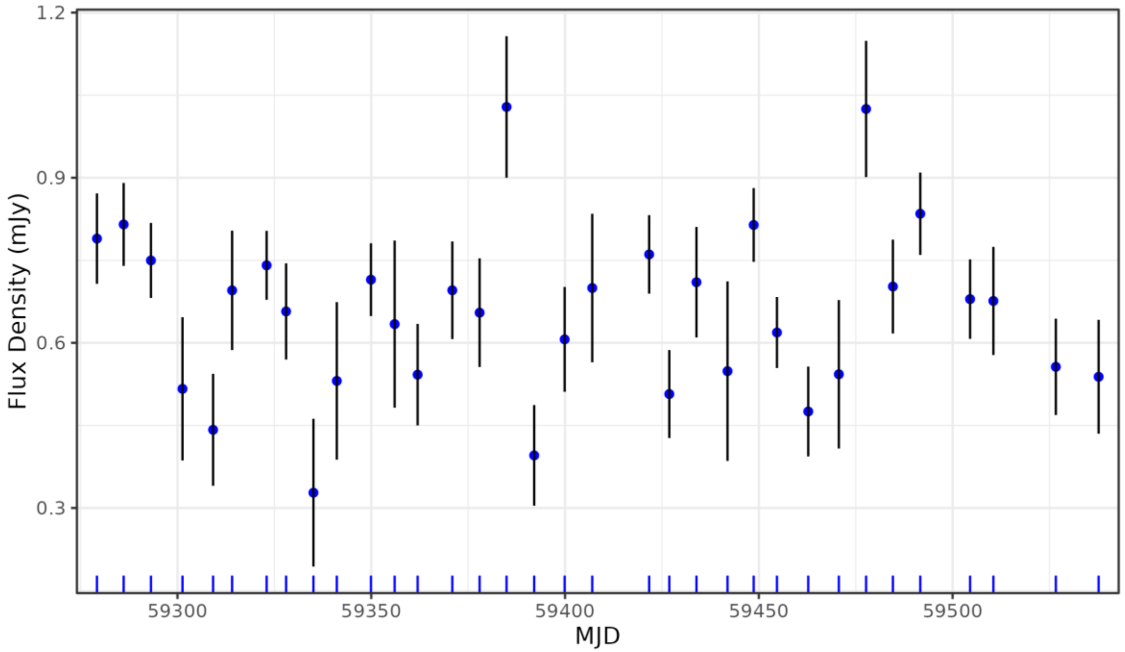
- Consider them a *stochastic process.*
- Random variables indexed in time with *autocovariance* structure.

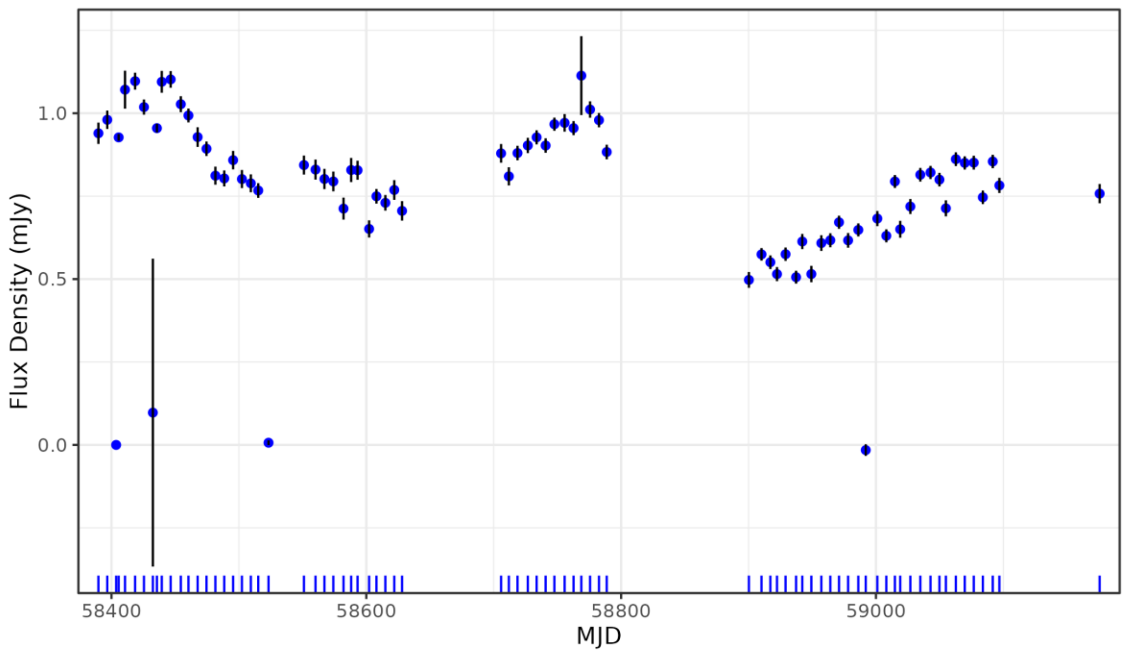


Data is heterogeneous

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



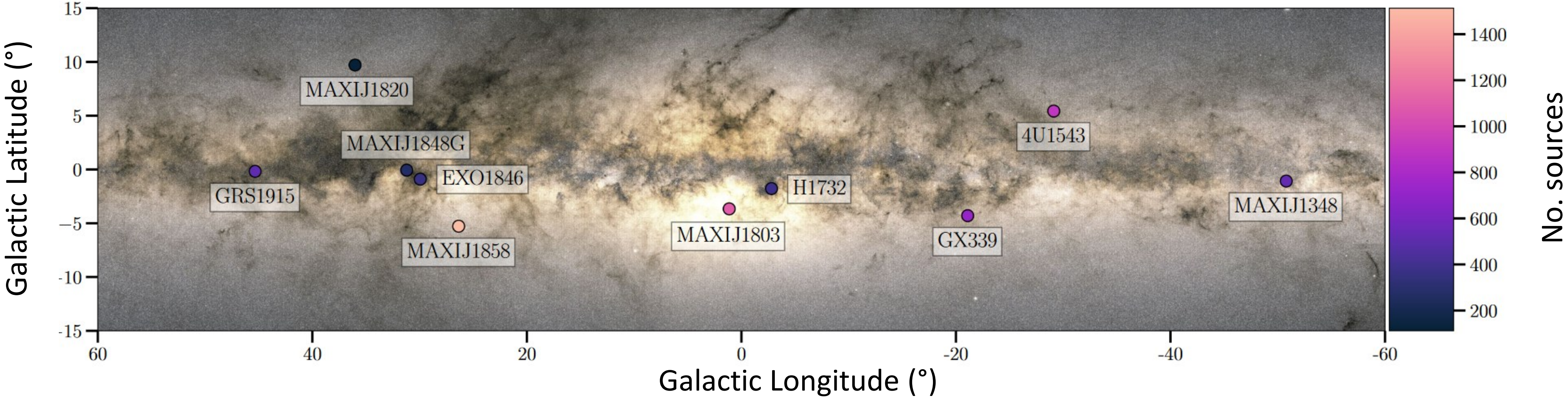




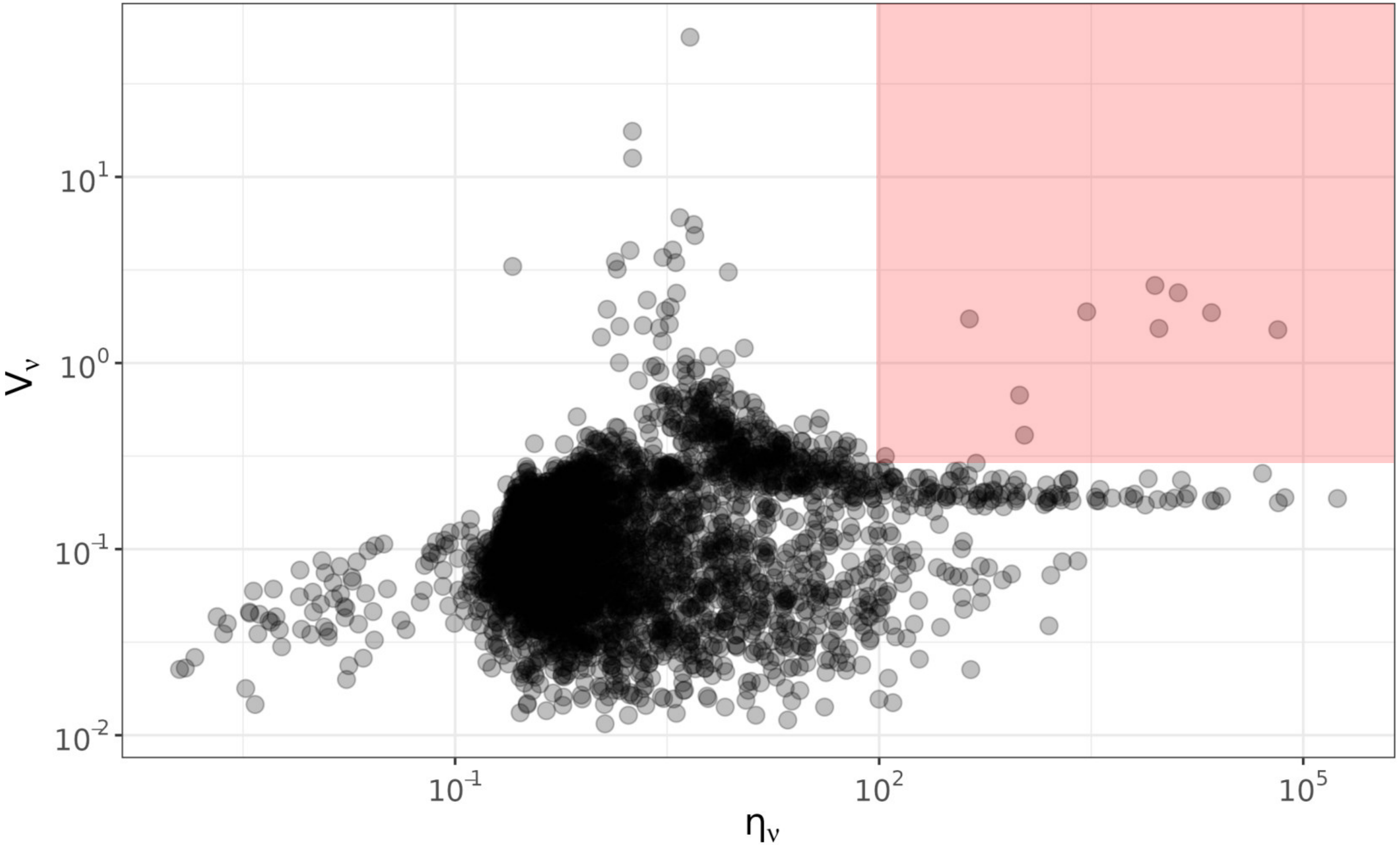
Dataset

- 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)



Variability Statistics: η_ν and V_ν



(Data courtesy of Andersson et al., 2023)

$$\eta_\nu = \frac{1}{N}\sum\left(\frac{\text{Obs.} - \text{Wt. Mean}}{\text{Std. Error}}\right)^2$$

$$\sim \chi^2_{N-1}$$

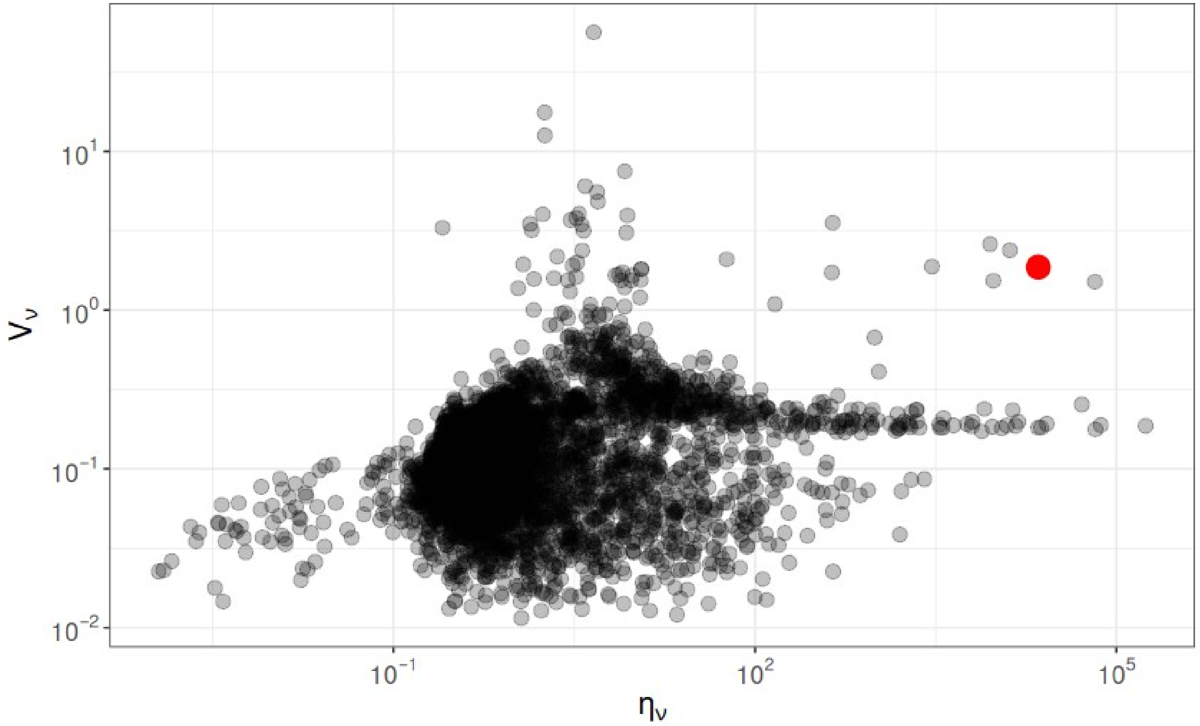
$$V_\nu = \frac{\text{Standard Deviation}}{\text{Mean}}$$

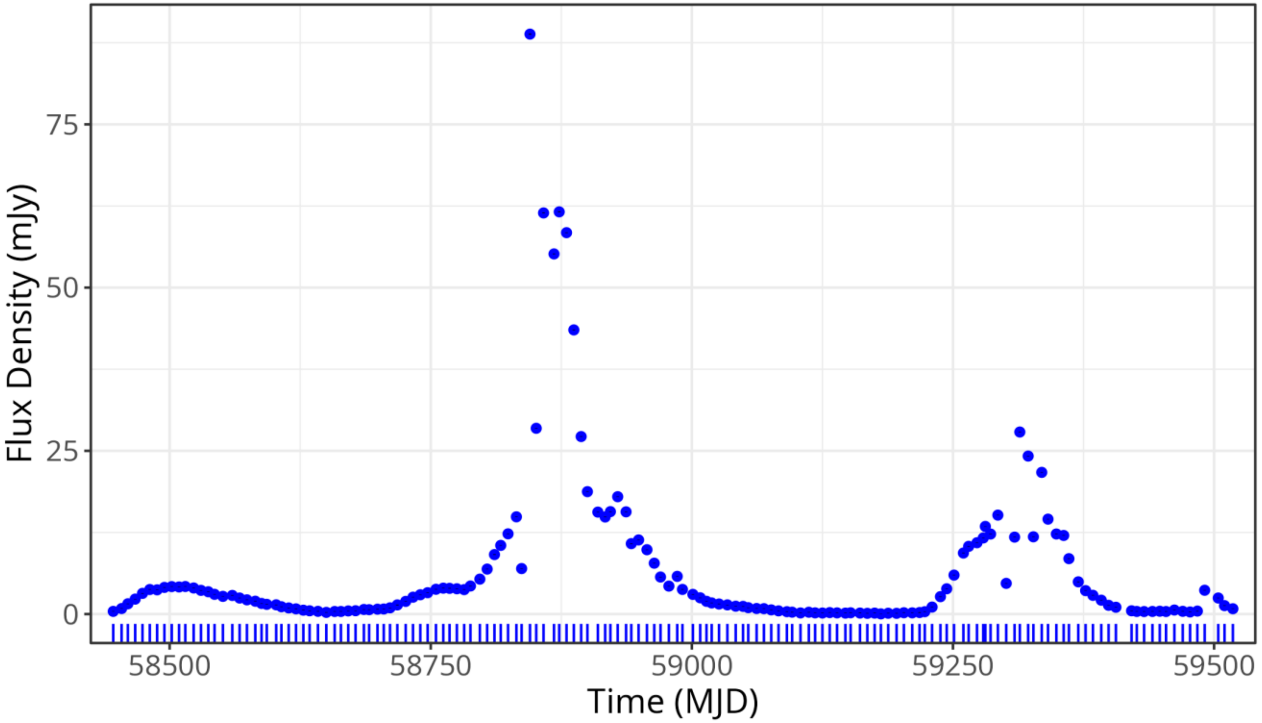
As $\eta_\nu \to \infty$ and $V_\nu \to \infty$

Source is likely transient

GX 339-4 in (η_ν, V_ν)-space

- Galactic LMXB and black hole candidate that flares





$$(\eta_\nu = 22427.6, V_\nu = 1.86)$$

Characterising Light Curves

Oversimplified

- Fewer parameters
- Scales easily
- High information loss

Overspecified

- Many parameters
- High discriminatory power
- Risks overfitting

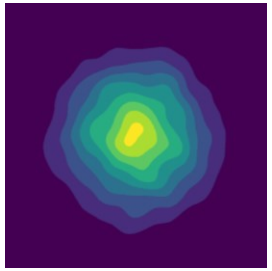
Model light curves as a Gaussian Process (GP)

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

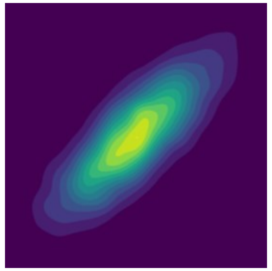
$\boldsymbol{Y}$ is a vector of n Gaussian random variables.

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

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 (GPs)

Extend multivariate Gaussian to 'infinite' dimensions.

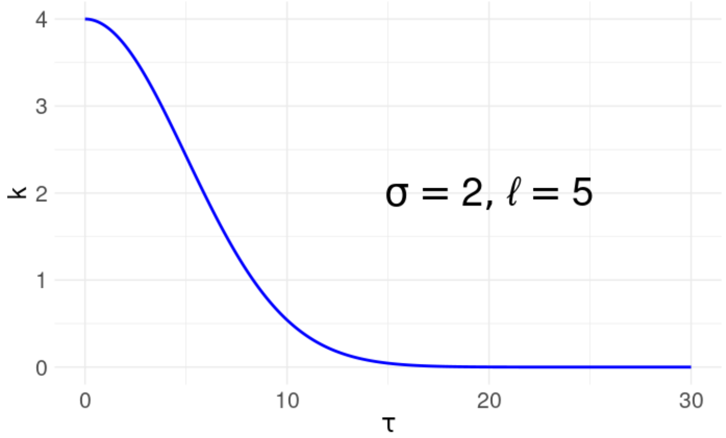
- Mean function, $\mu(t)$
- Covariance or **kernel function**, $\boldsymbol{\kappa}(\boldsymbol{t}, \boldsymbol{t})$

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

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

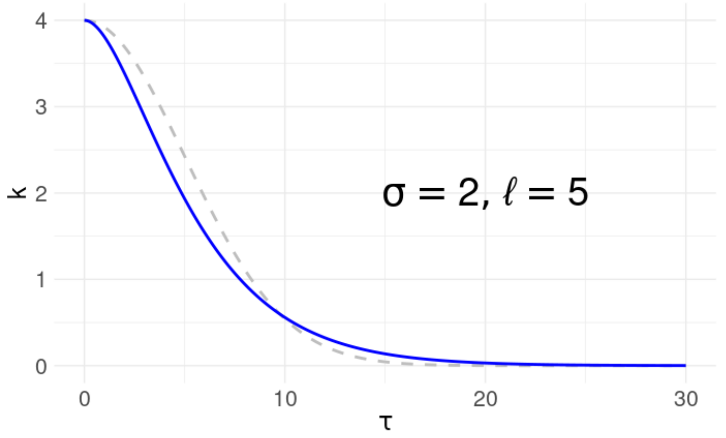
Rather than specifying a fixed covariance matrix with fixed dimensions, compute covariances using the kernel function.

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



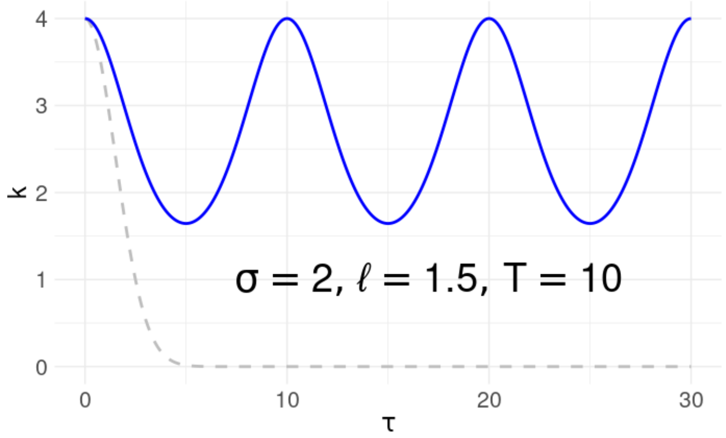
$$\kappa(\tau; \sigma, \ell) = \sigma^2 \exp\left\{-\frac{1}{2}\left(\frac{\tau}{\ell}\right)^2\right\}$$

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

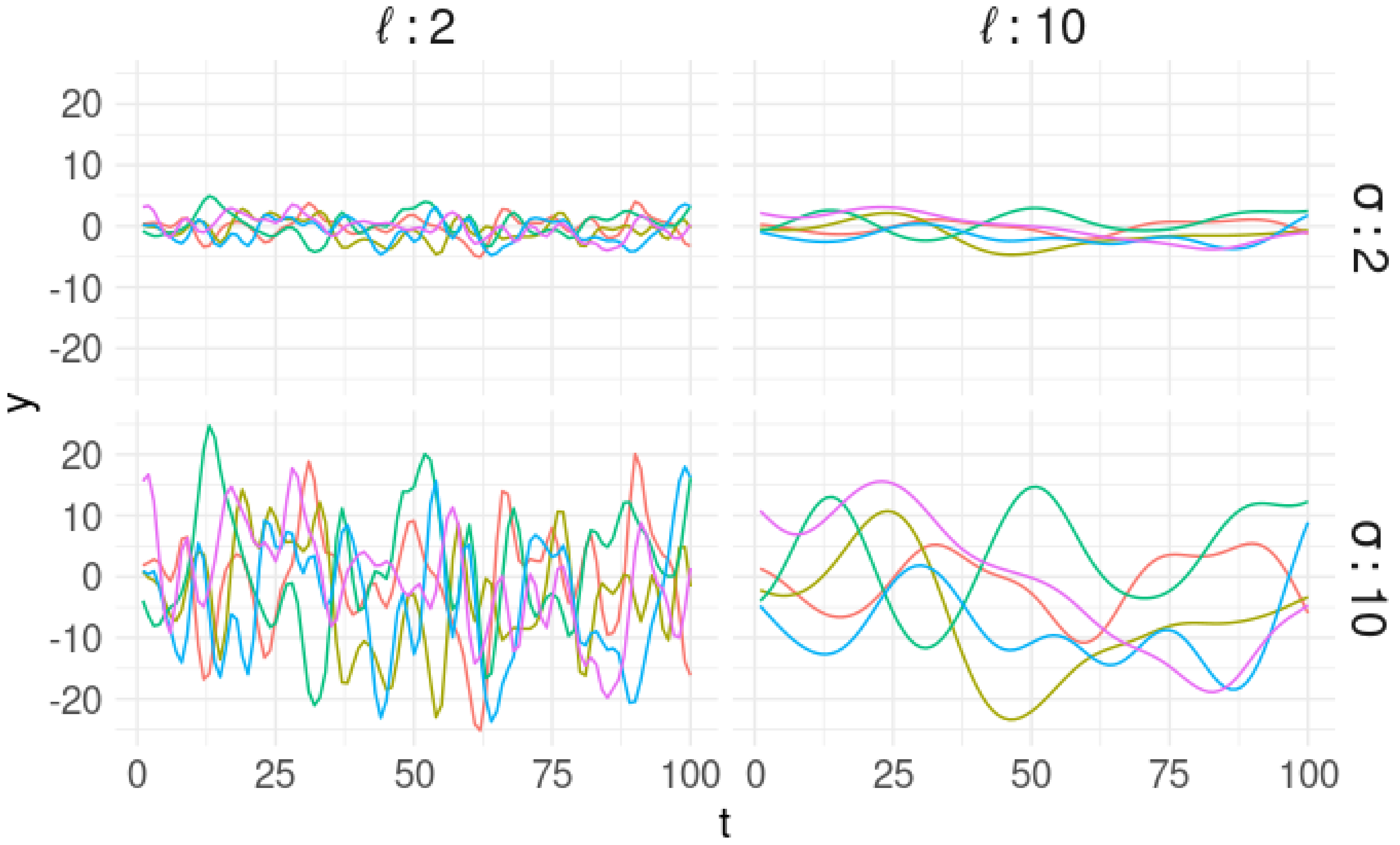


$$\kappa(\tau; \sigma, \ell, T) = \sigma^2 \exp\left\{-\frac{2}{\ell^2}\sin^2\left(\pi\frac{\tau}{T}\right)\right\}$$

Periodic

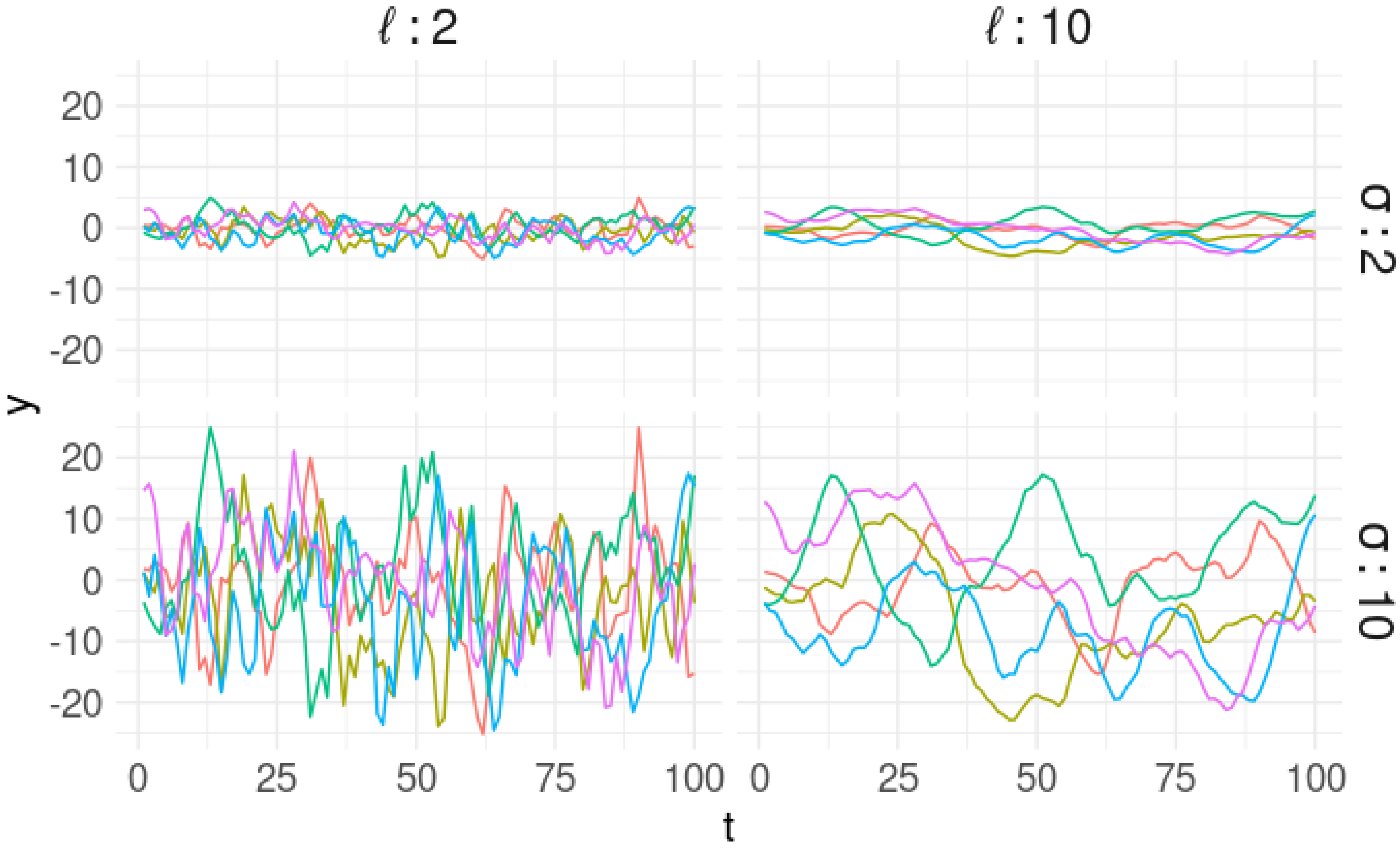
Squared Exponential Kernel

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



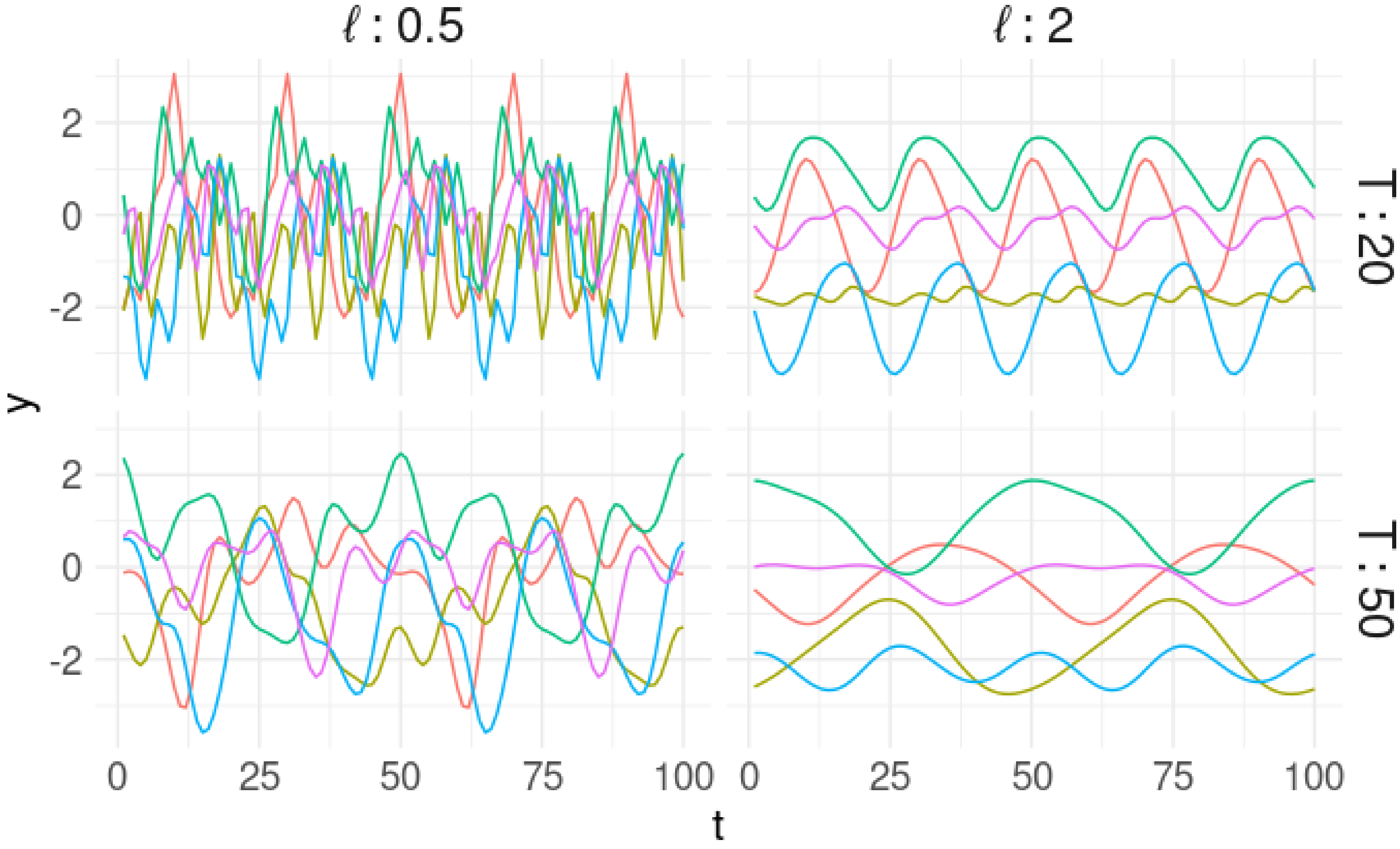
Matern 3/2 Kernel

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



Periodic Kernel

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



Bayesian Hierarchical Model

Data Model $\mathbf{Y} \sim \mathcal{N}_N(\mathbf{f}, \hat{e}^2)$ Observed Error

Process Model $\mathbf{f} \sim \mathcal{GP}(\mathbf{0}, \mathbf{\Sigma}_{N\times N})$

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

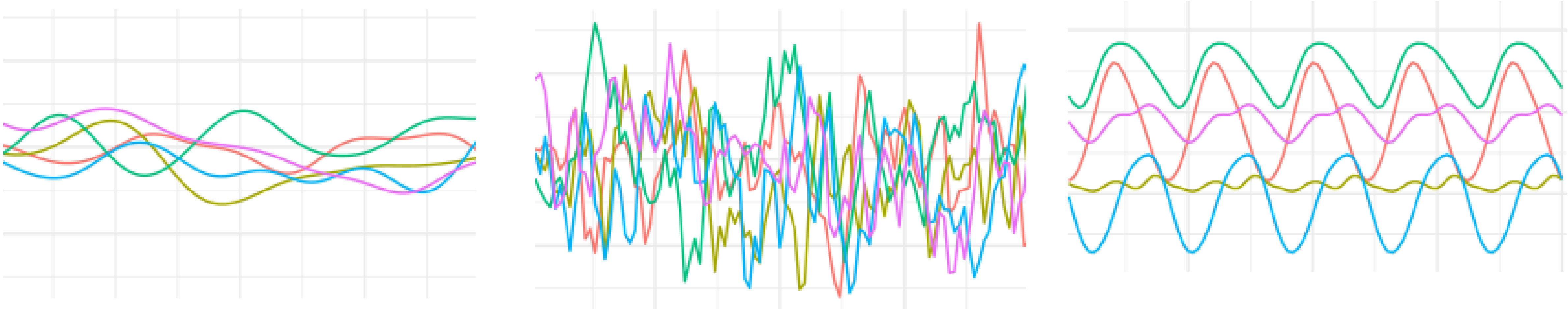
$$\boldsymbol{\theta} = \{\sigma_{SE}, \ell_{SE}, \sigma_{M32}, \ell_{M32}, \sigma_P, \ell_P, T\}$$

$$[\mathbf{\Sigma}]_{rc} = \kappa(t_r, t_c | \boldsymbol{\theta})$$

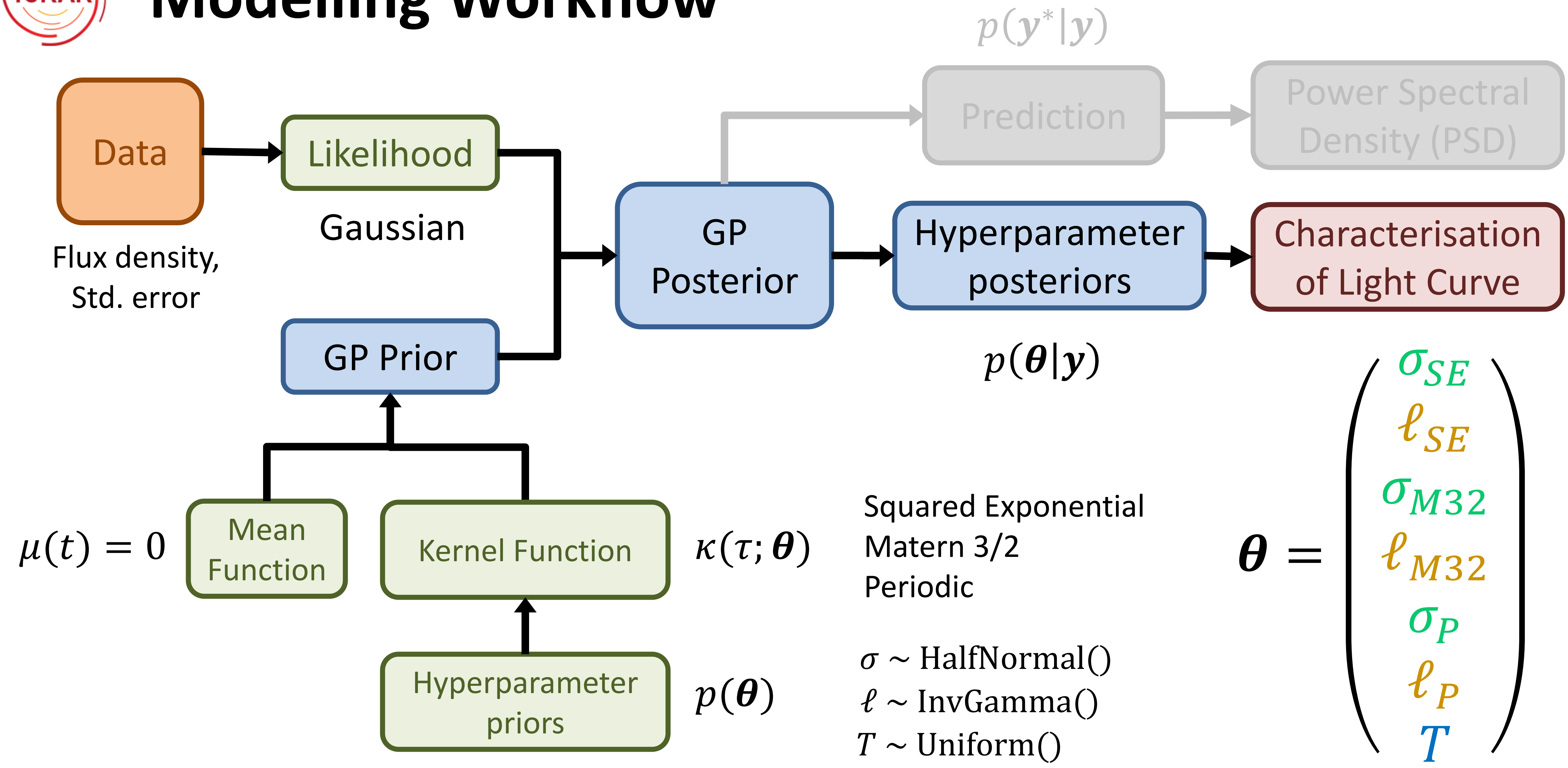
$$= \kappa_1(\tau; \sigma_{SE}, \ell_{SE}) + \kappa_2(\tau; \sigma_{M32}, \ell_{M32}) + \kappa_3(\tau; \sigma_P, \ell_P, T)$$

Squared Exponential | Matern 3/2 | Periodic

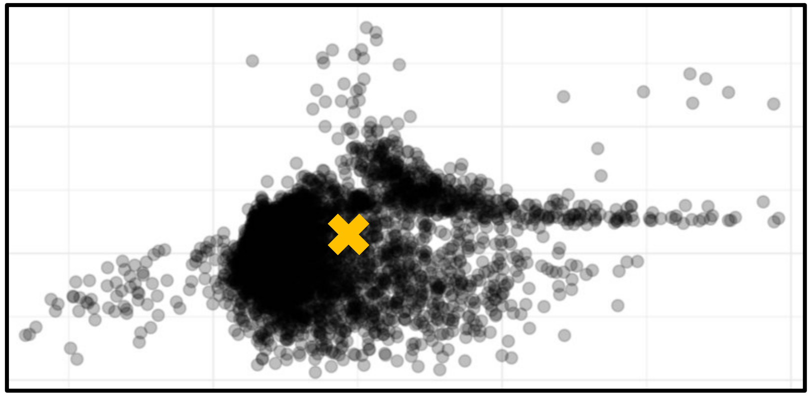
Covariance Kernel



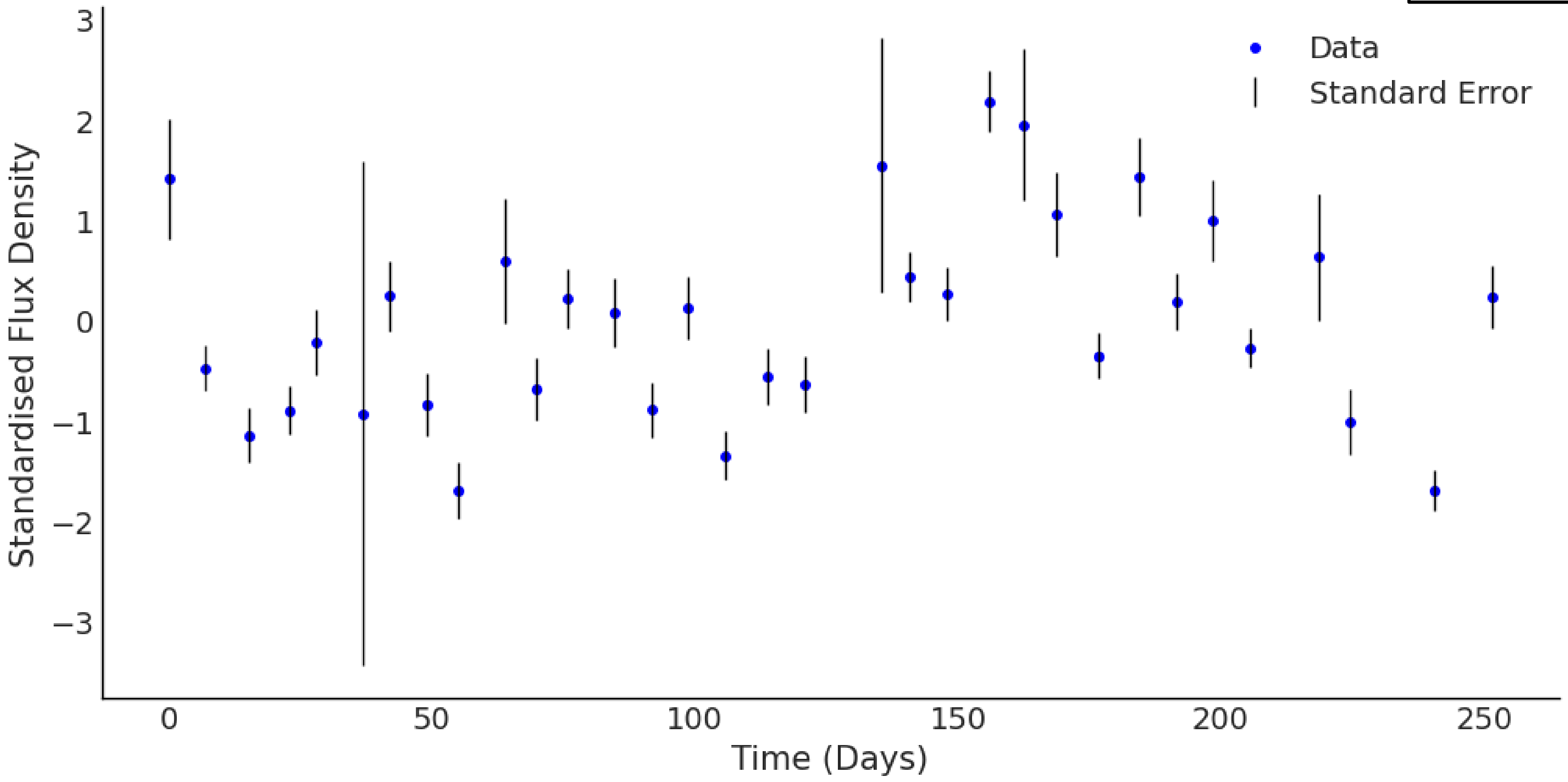
Modelling Workflow



Example



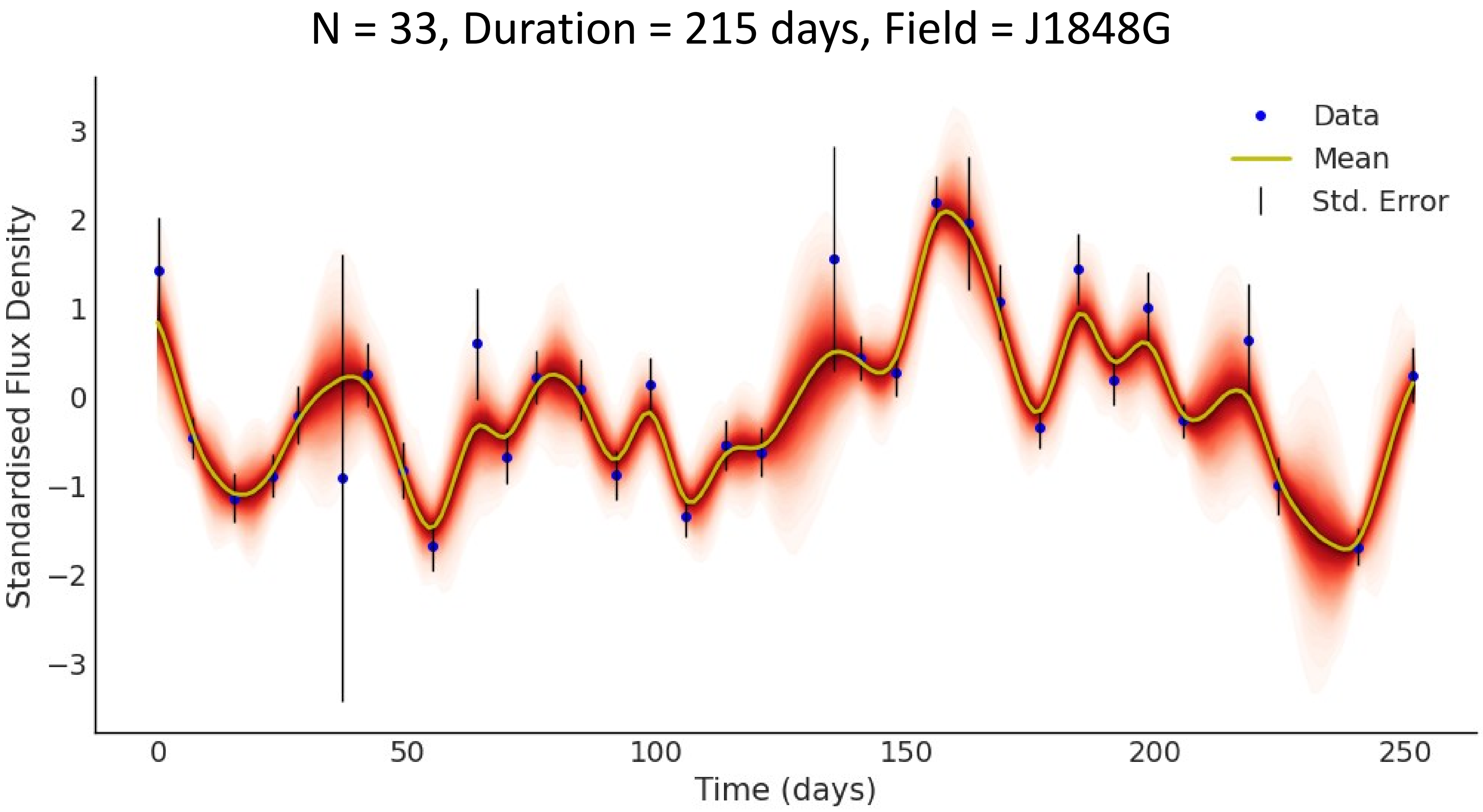
N = 33, Duration = 215 days, Field = J1848G



$\eta_\nu = 2.91$

$V_\nu = 0.12$

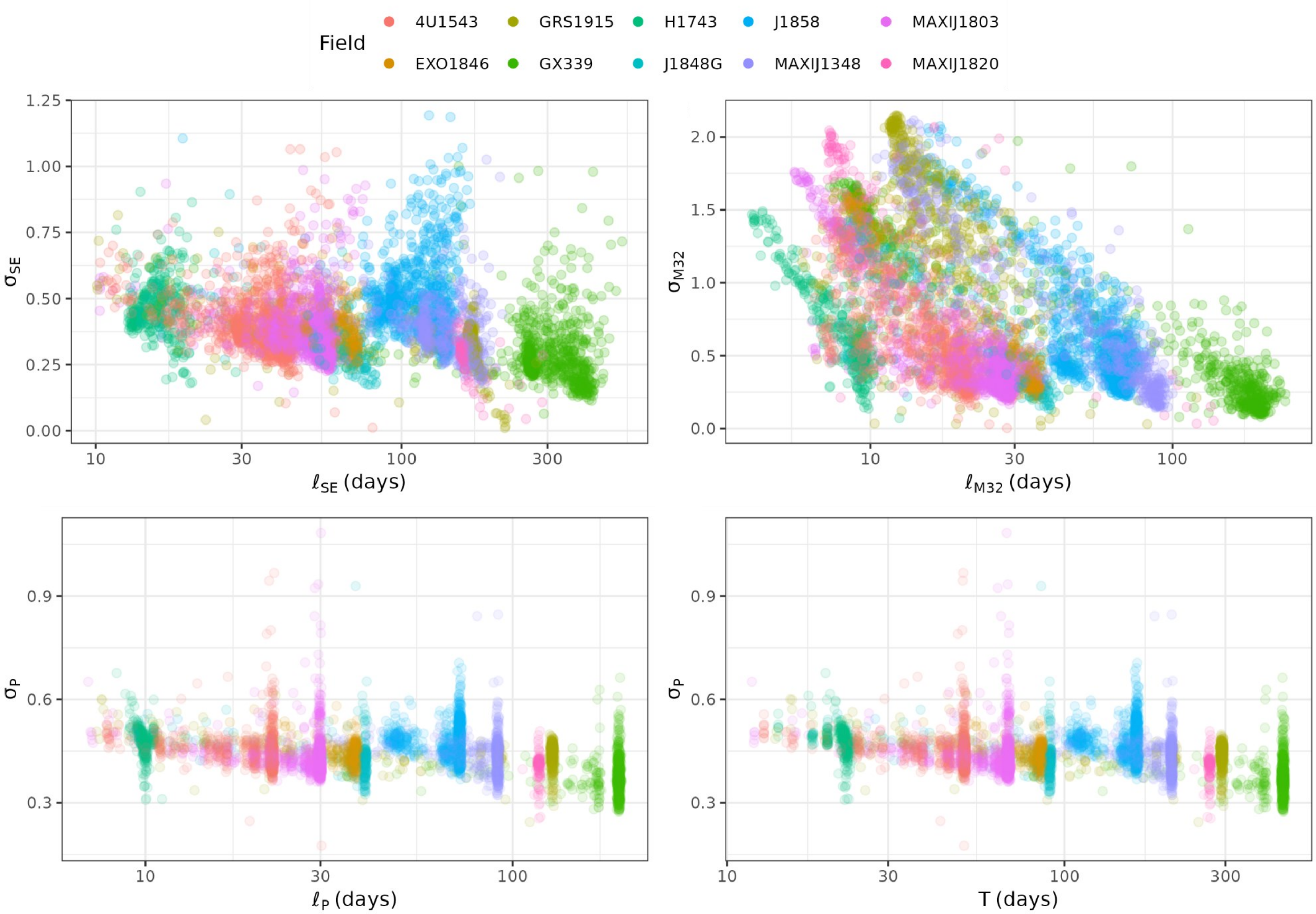
Posterior Predictive Curves



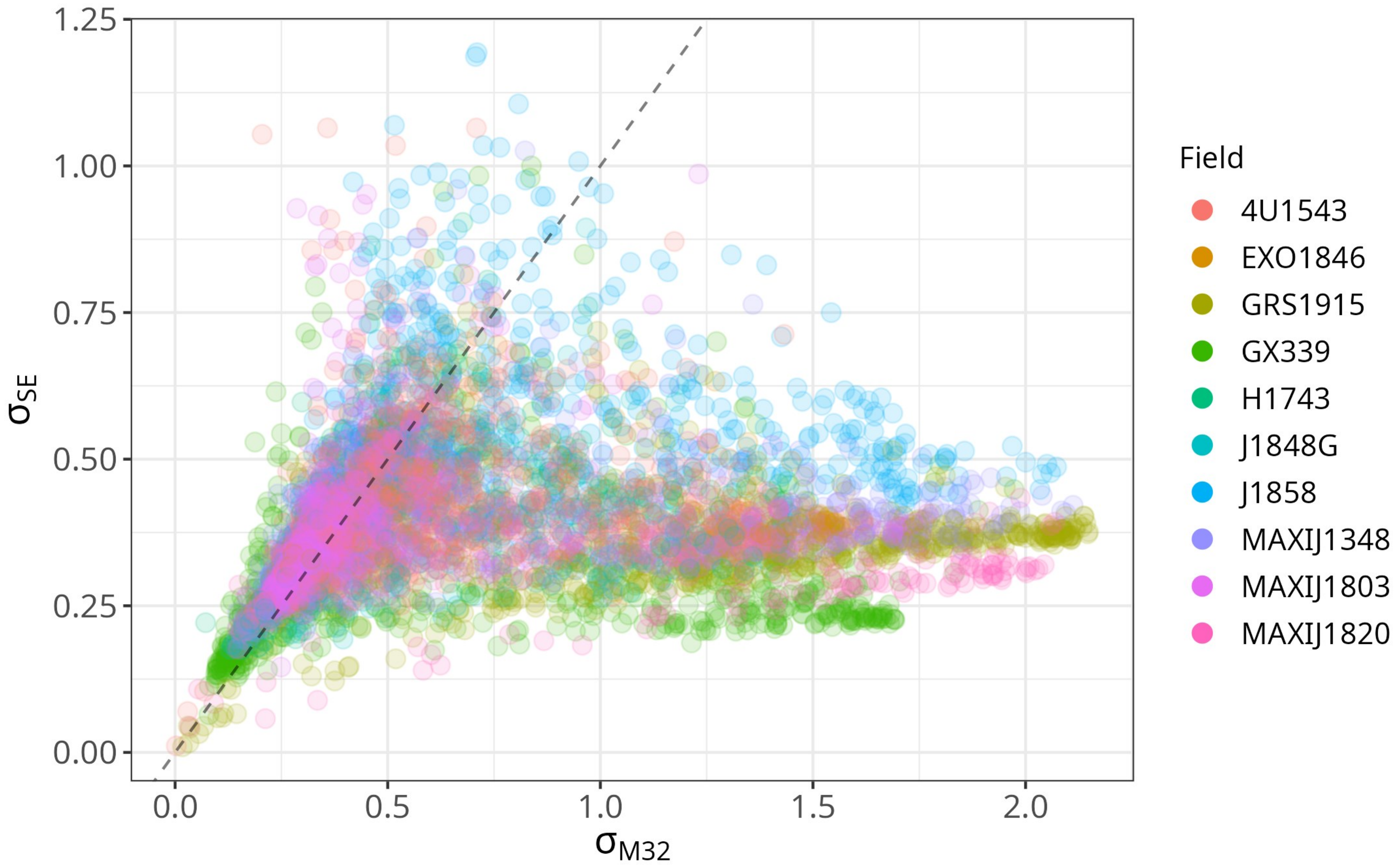
Posterior Medians

$\sigma_{SE} = 0.39$
$\sigma_{M32} = 1.26$
$\sigma_{P} = 0.50$
$\ell_{SE} = 50.0$
$\ell_{M32} = 11.9$
$\ell_{P} = 46.7$
$T = 41.1$

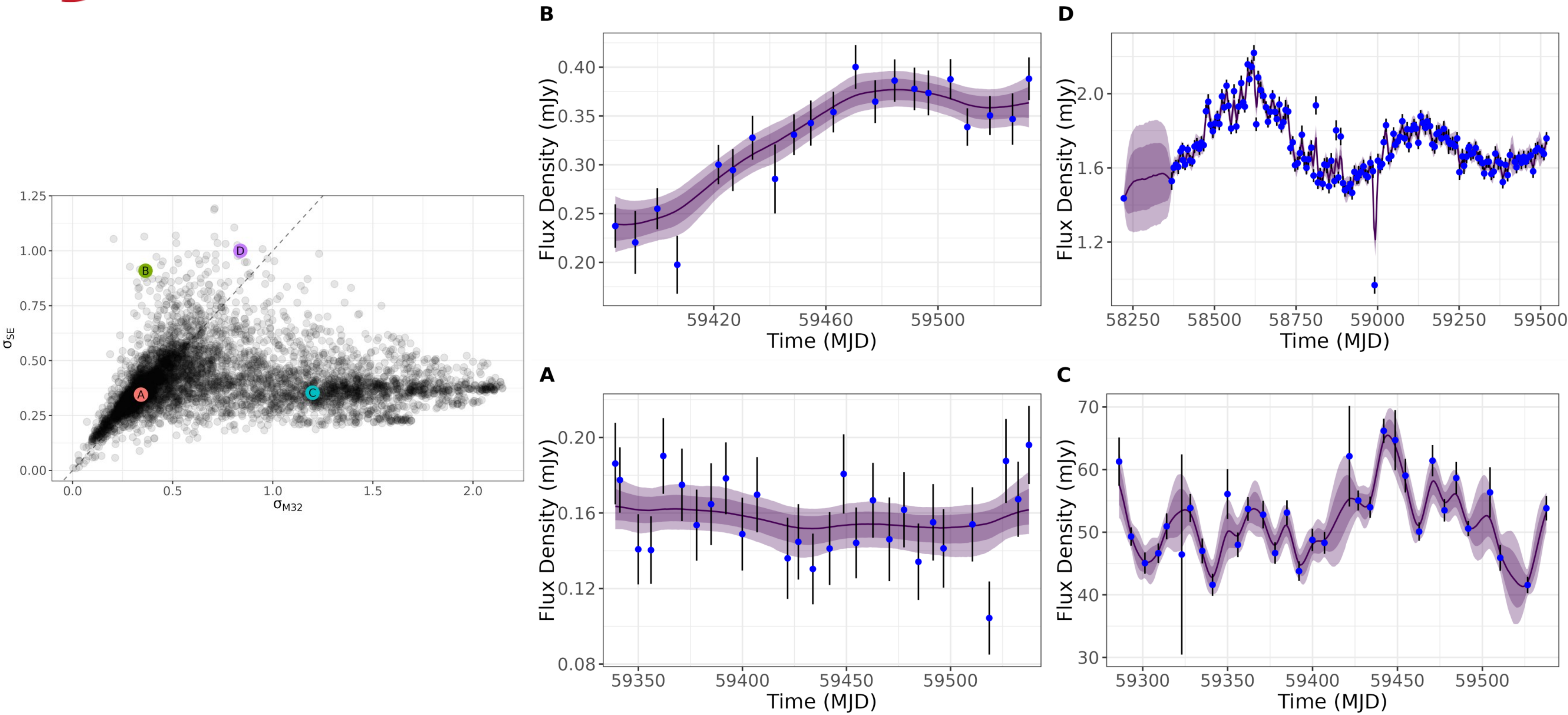
$\eta_{\nu} = 2.91$
$V_{\nu} = 0.12$


Field
4U1543
EXO1846
GRS1915
GX339
H1743
J1848G
J1858
MAXIJ1348
MAXIJ1803
MAXIJ1820
σ_{SE}
ℓ_{SE} (days)
σ_{M32}
ℓ_{M32} (days)
σ_P
ℓ_P (days)
σ_P
T (days)

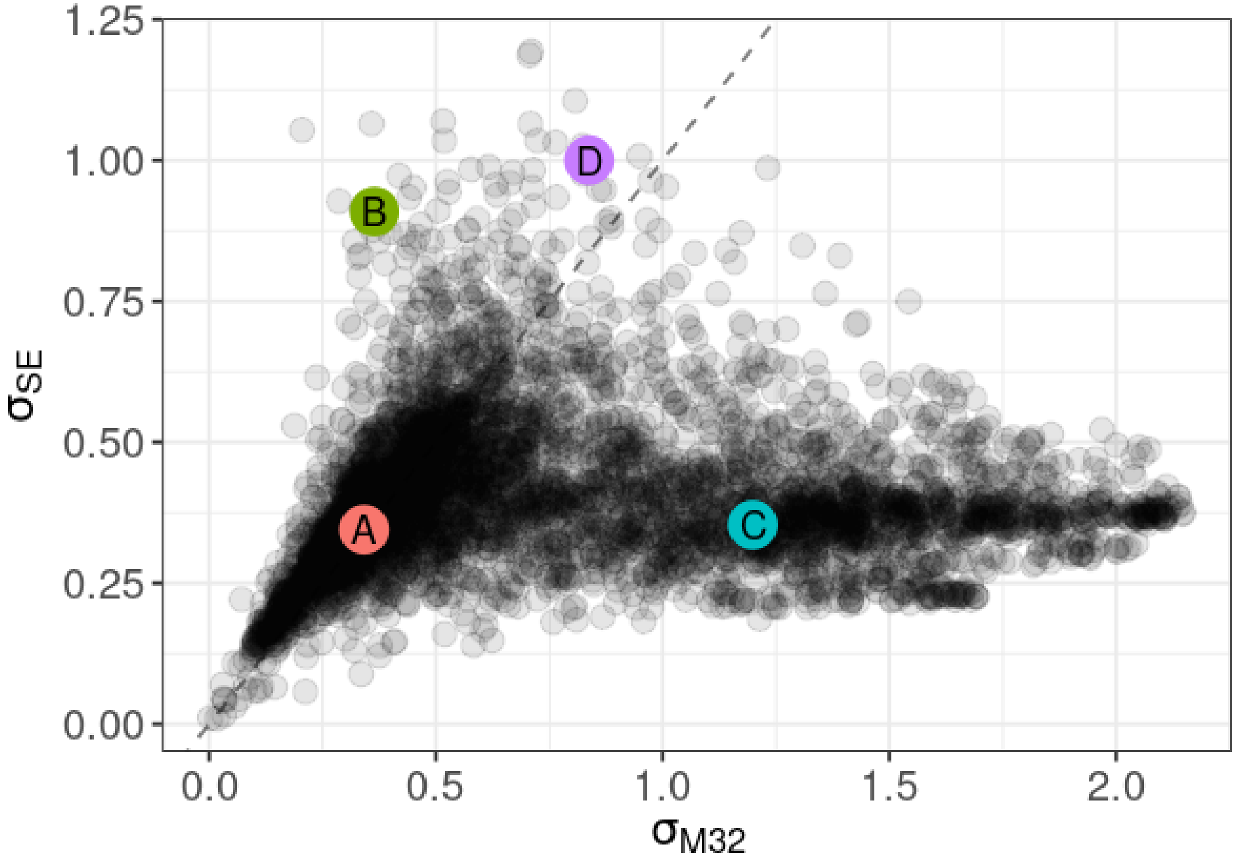
Amplitude Hyperparameter

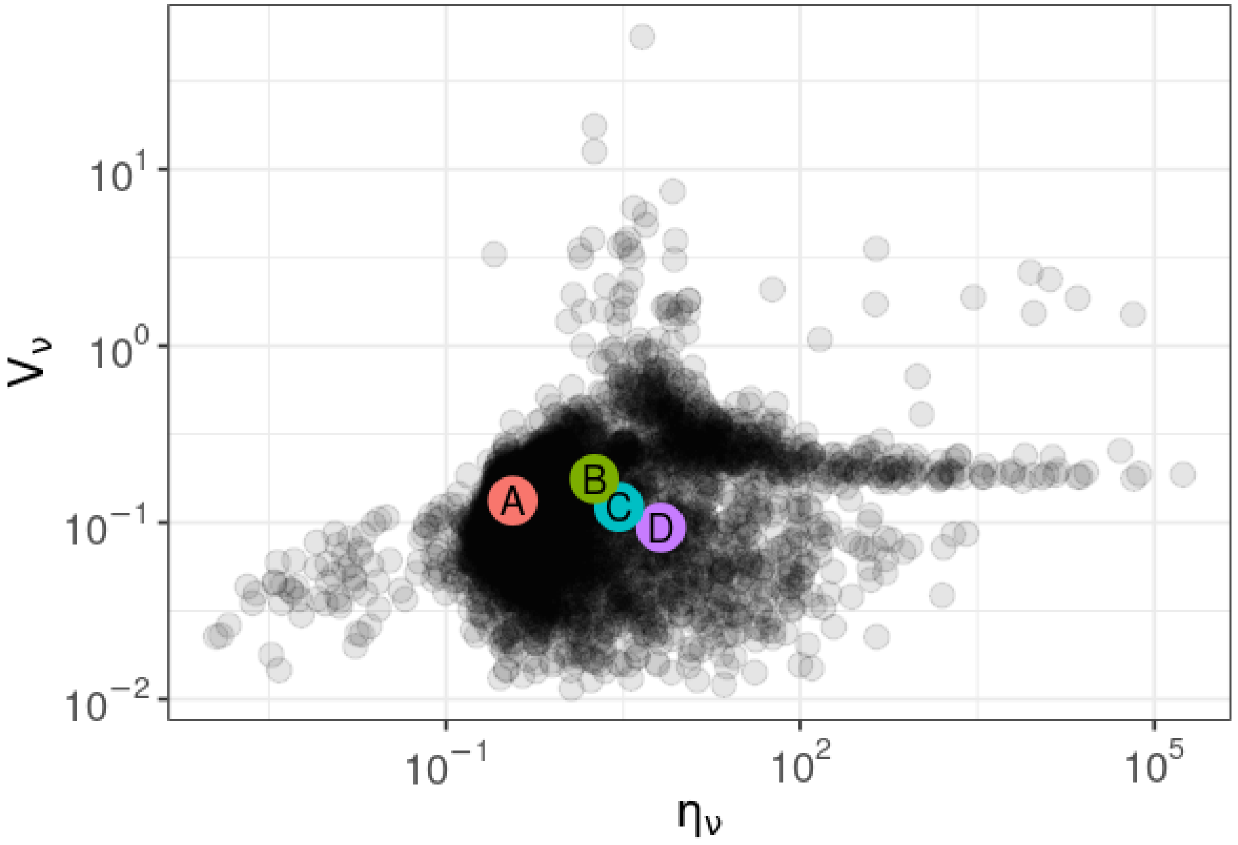


Explore the amplitude space



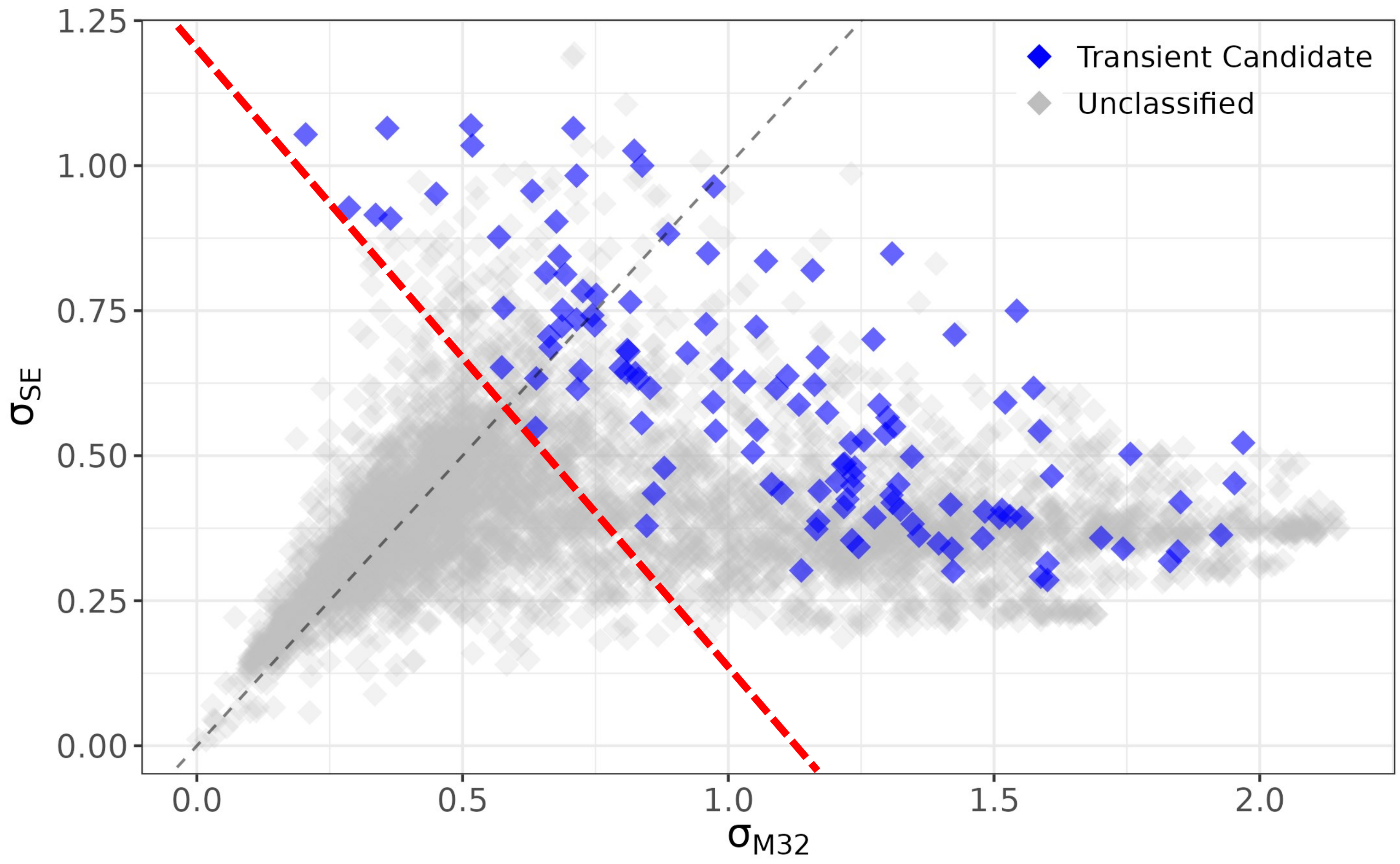
Comparison with (η_ν, V_ν)





Interpreting Amplitude as Transcience

- Transience seems to manifest as large values in **amplitude**, σ.
- 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.)

Summary

- Developed models and code suitable for fitting univariate GPs to the light curves of a large radio survey, i.e., ThunderKAT.
- GPs can be used to perform hyperparameter inference as well as interpolation in time-domain astronomy.
- Amplitude hyperparameter, σ, is an interpretable descriptor of variability; more discriminatory than (η_ν, V_ν).
- Next: extend to multi-band light curves.

Twinkle twinkle little star… a Gaussian process is what you are!