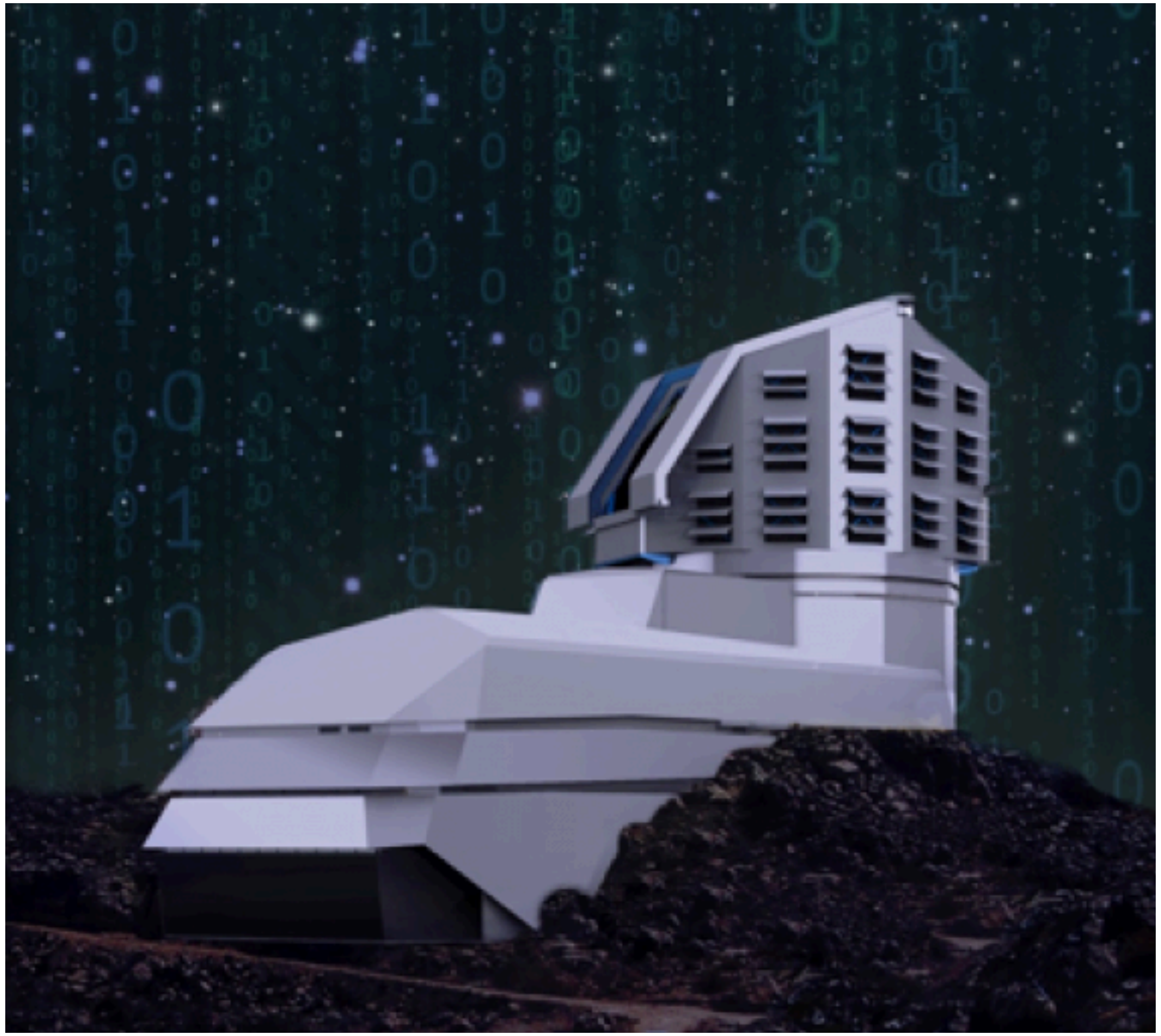
Real-time Detection of Anomalies in Transient Surveys

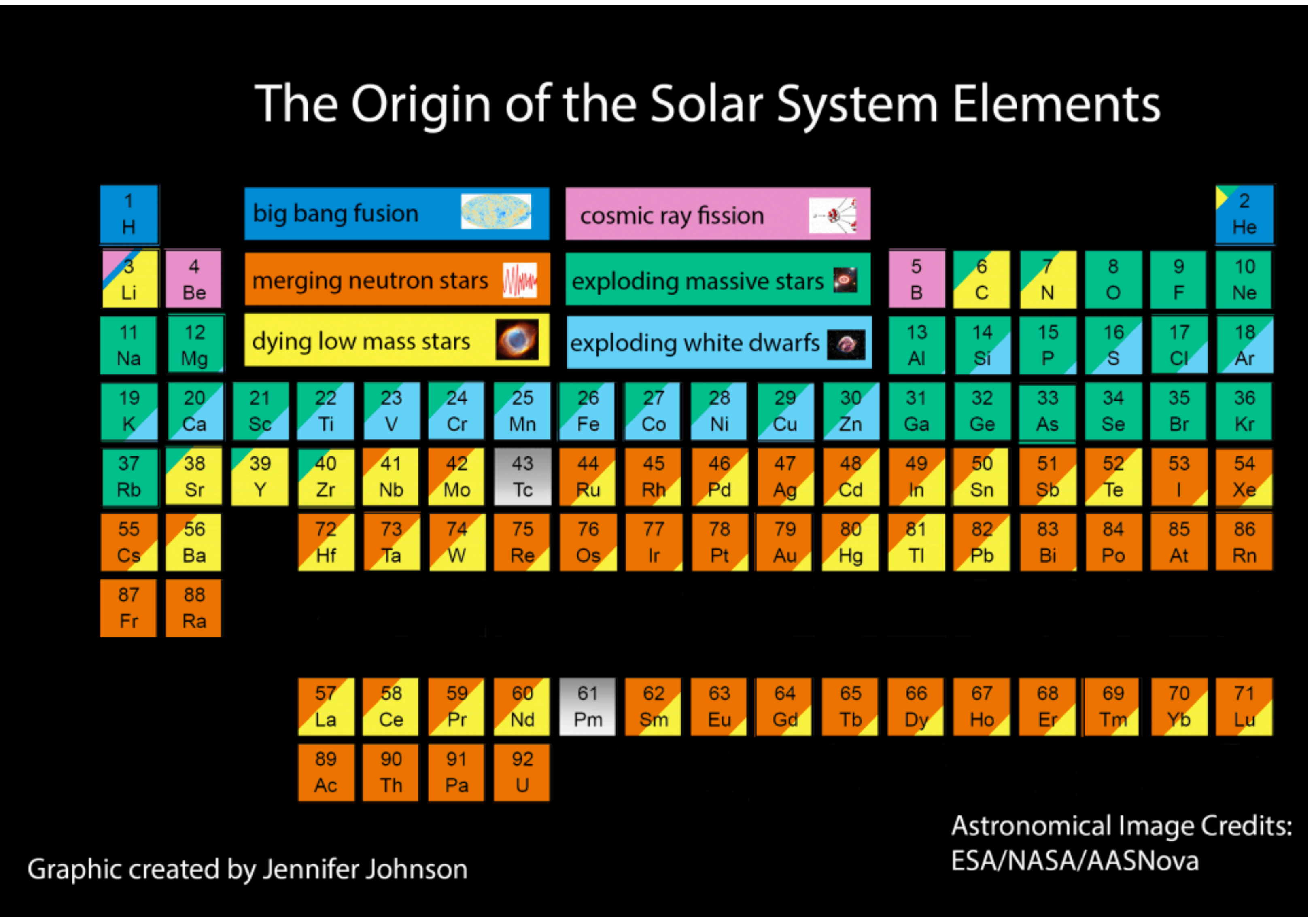
Daniel Muthukrishna
MIT

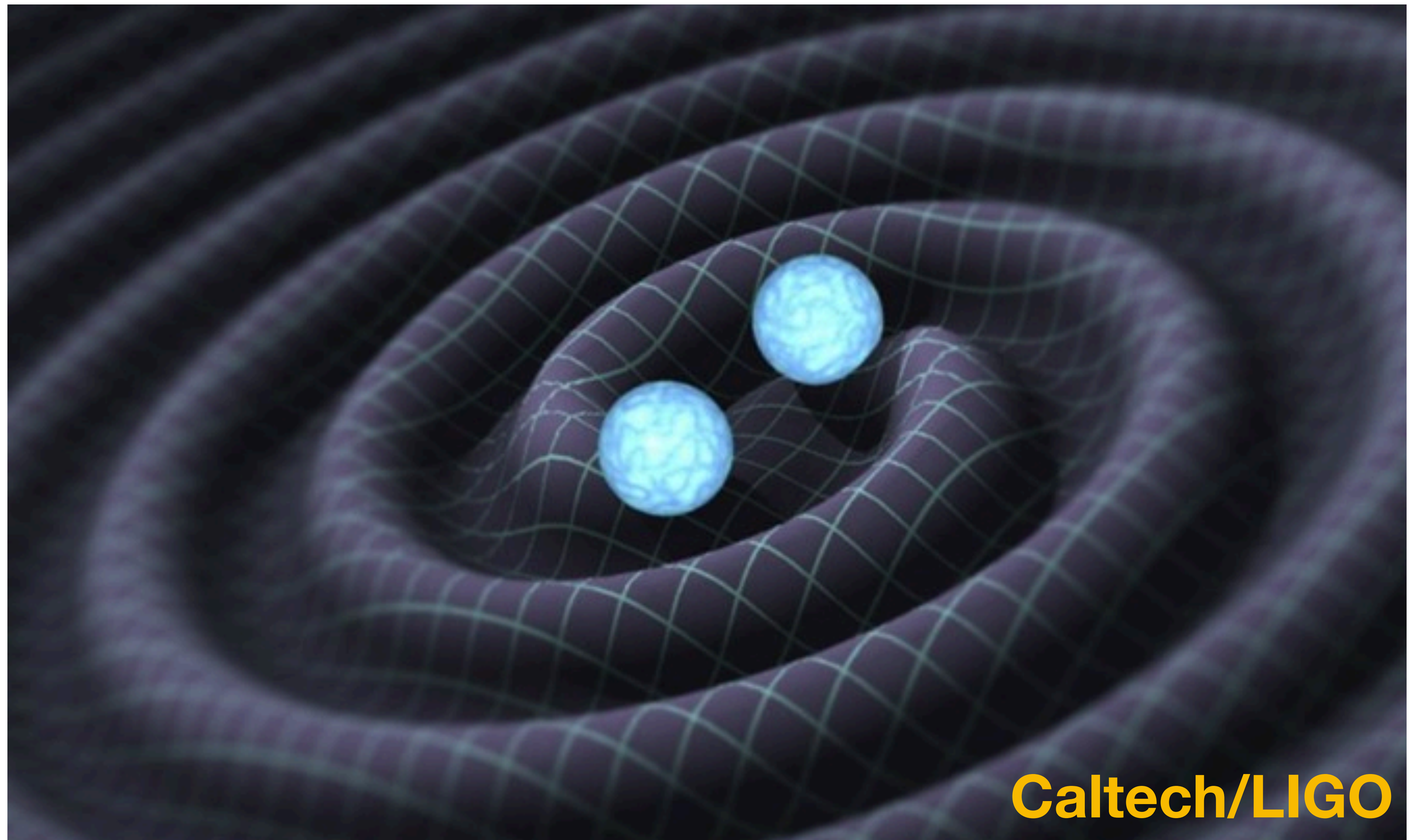
Collaborators:
Michelle Lochner (SARAO)
Kaisey Mandel (U. Cambridge)
Gautham Narayan (U. Illinois)
Sara Webb (Swinburne U.)

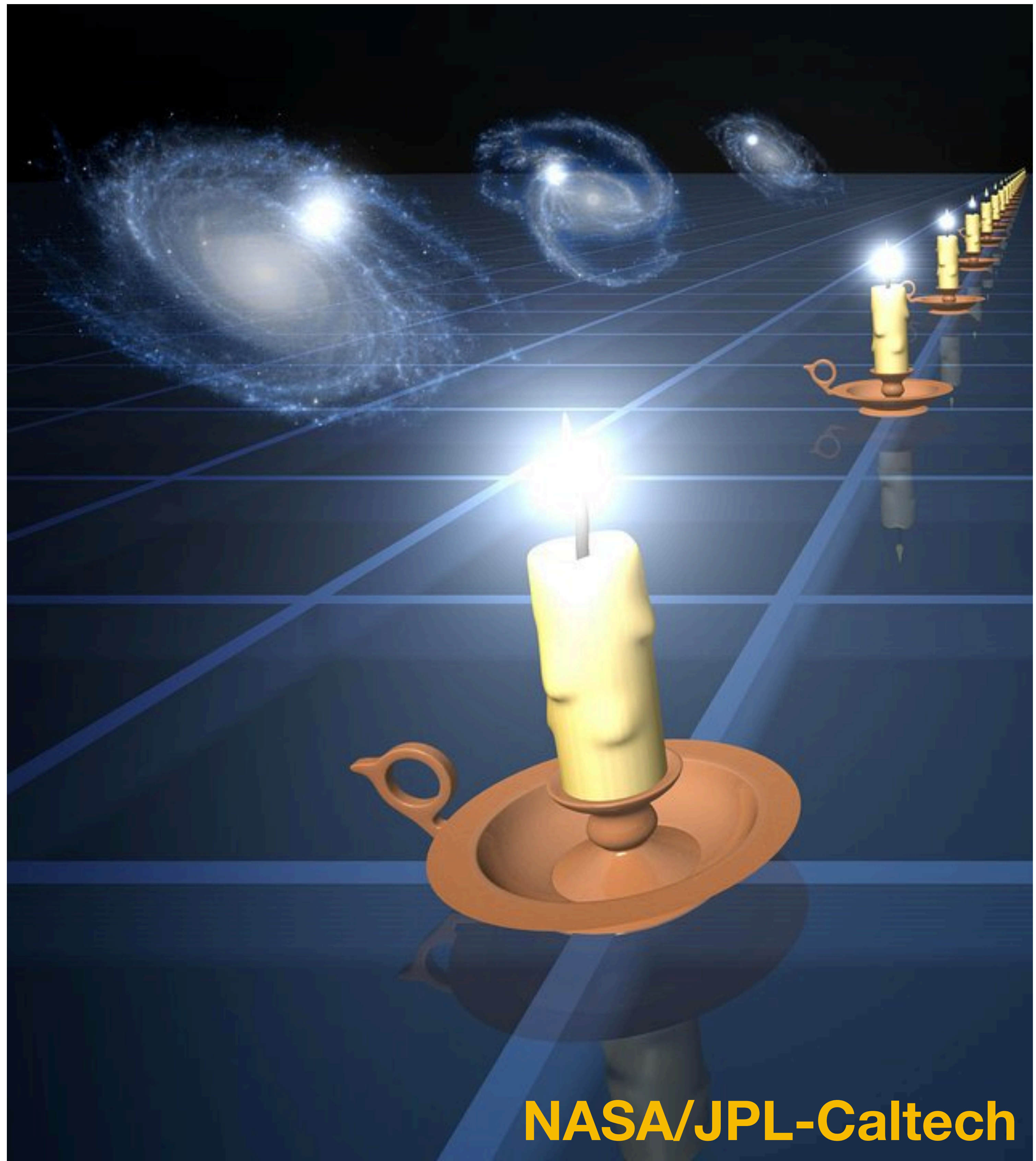


What have transients been useful for?

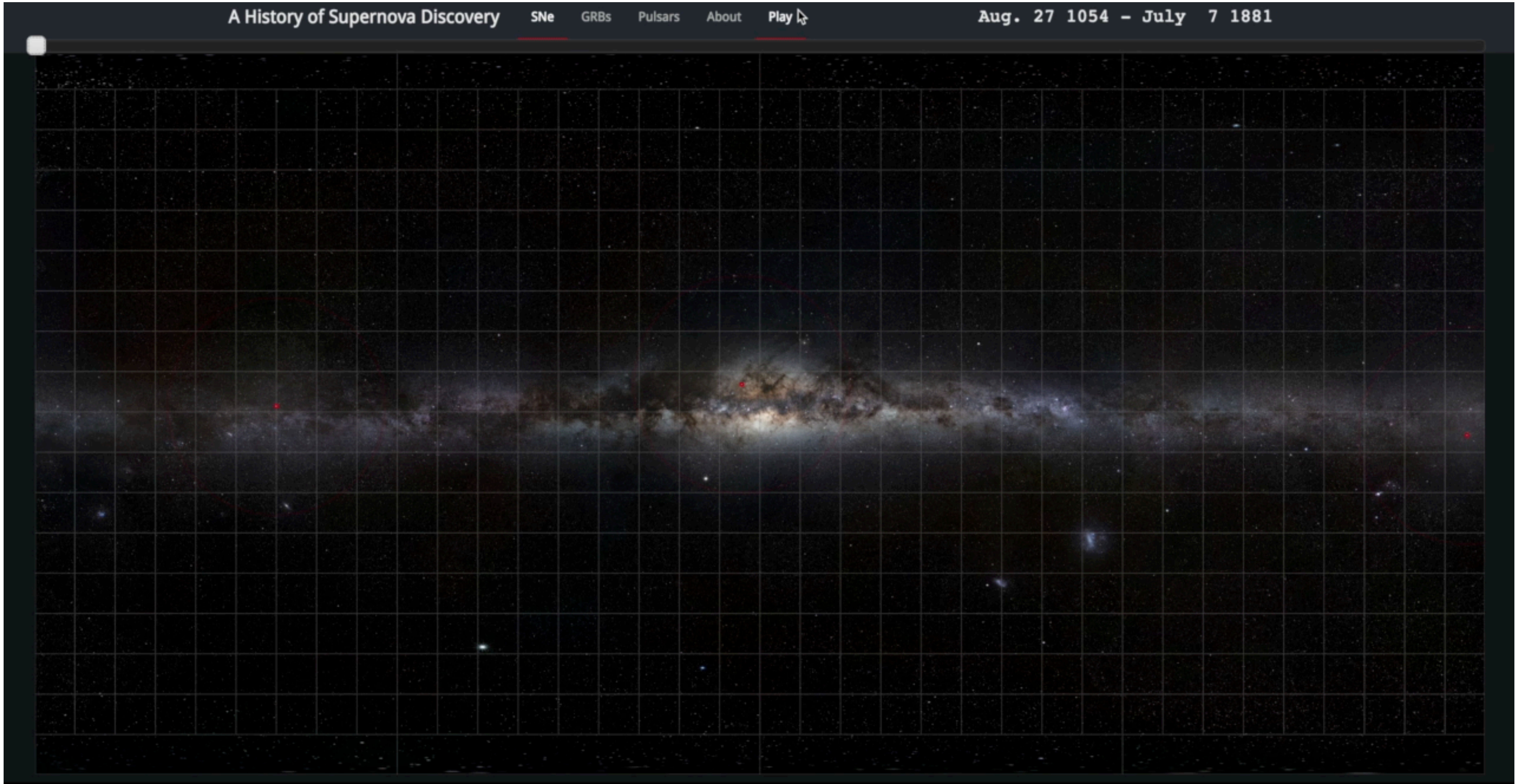
- Discovery of the accelerating expansion of the universe (Type Ia Supernova)
- Detection of gravitational waves (Kilonovae)
- Production of the universe’s heavy elements







12785 Supernovae in 15 seconds



Isaac Shivvers (UC Berkeley)

LSST TAKES 20TB OF IMAGES PER NIGHT

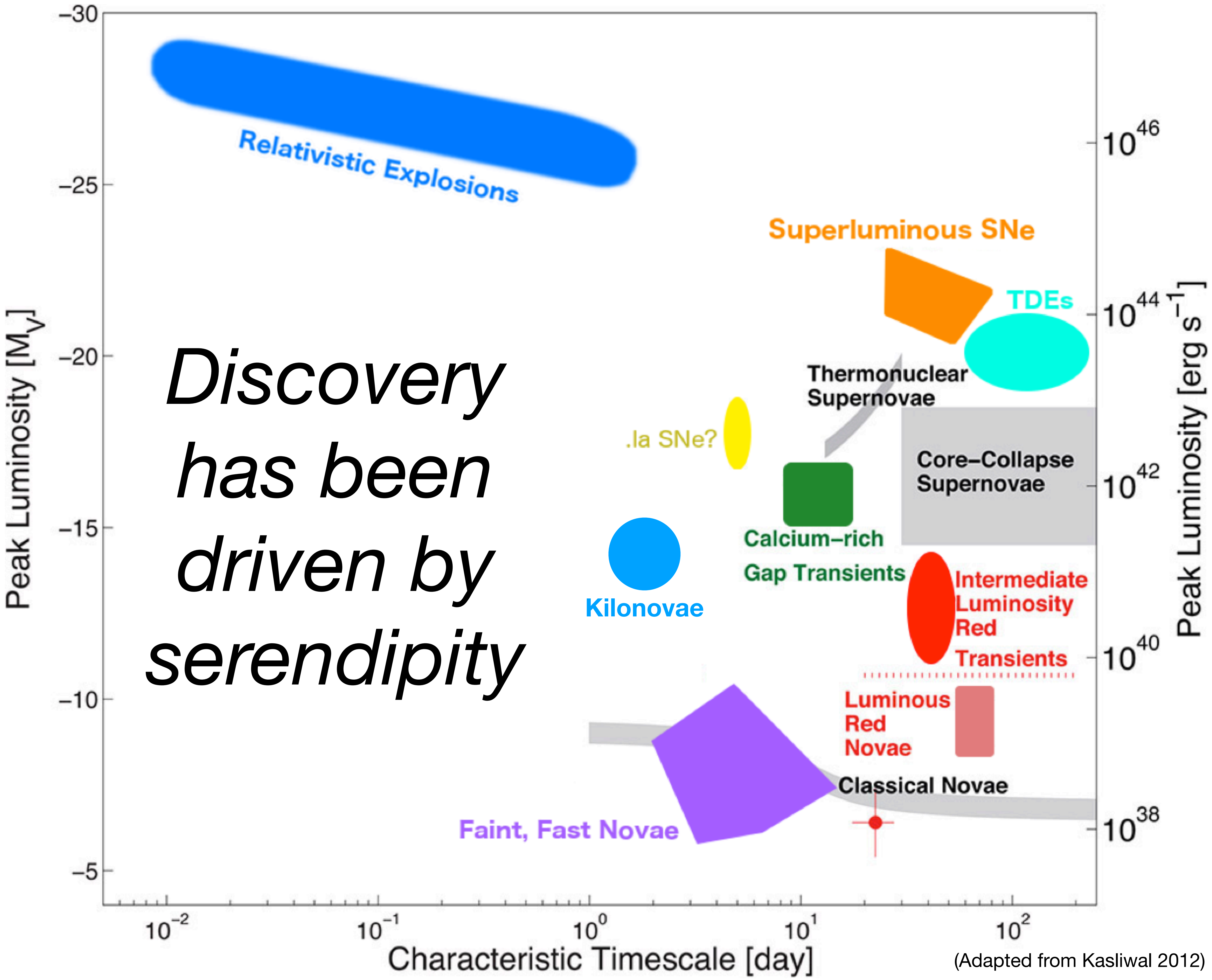
10 million transient alerts per night!

How can we expect to get lucky with such high volumes of data?

Transient searches have relied on human eyes for alerts

The known transient universe

- The transient universe remains largely mysterious
- New surveys will observe an unprecedented number of transients - new and known
- Need to prioritise follow-up based on class and epoch
- Automated, fast, and early classifications are required



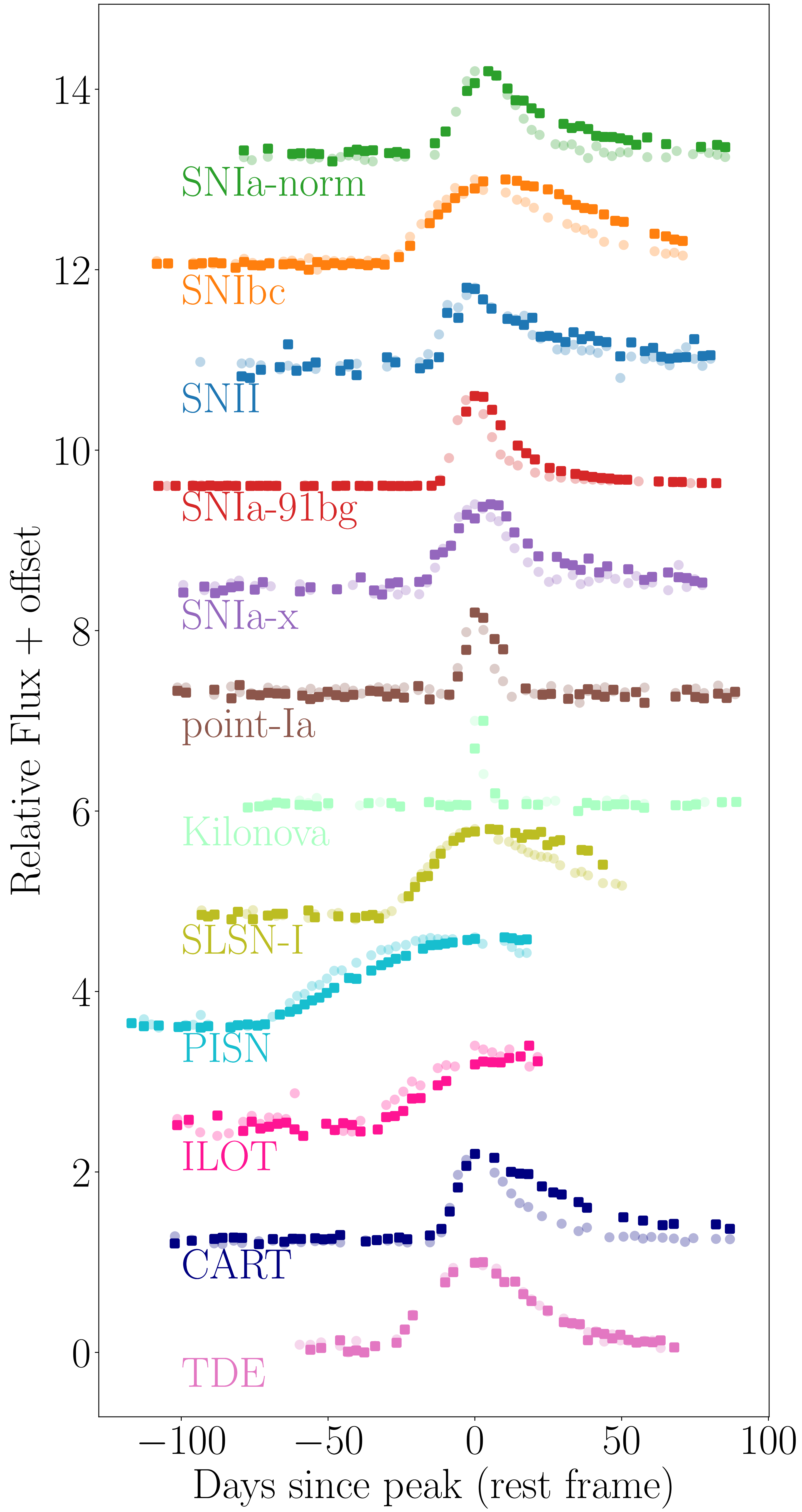
(Adapted from Kasliwal 2012)

Simulated dataset

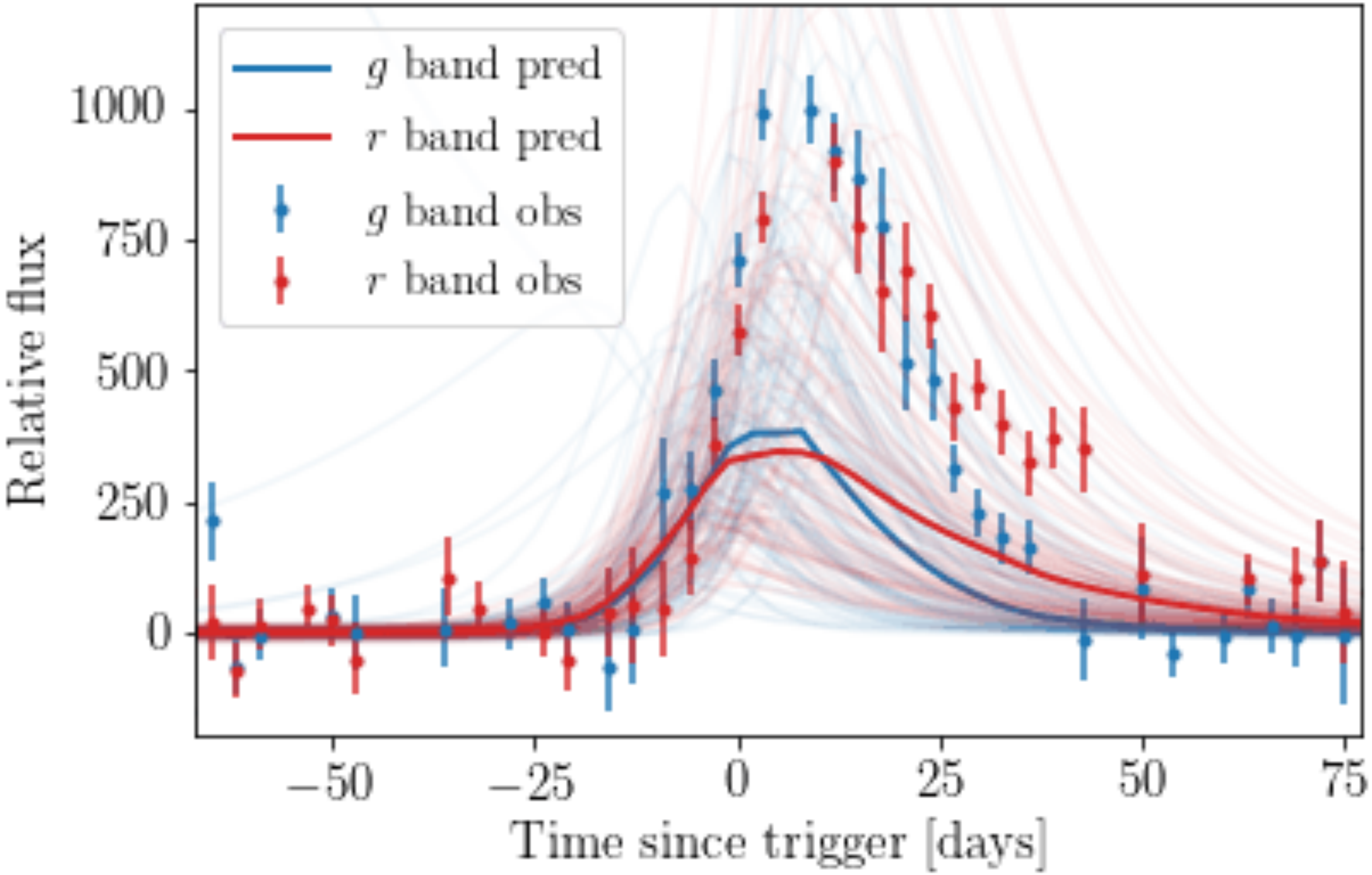
- We simulated transients with the observing properties of the Zwicky Transient Facility (ZTF) using PLAsTiCC software

- 10,000 ZTF light curves for each transient classes

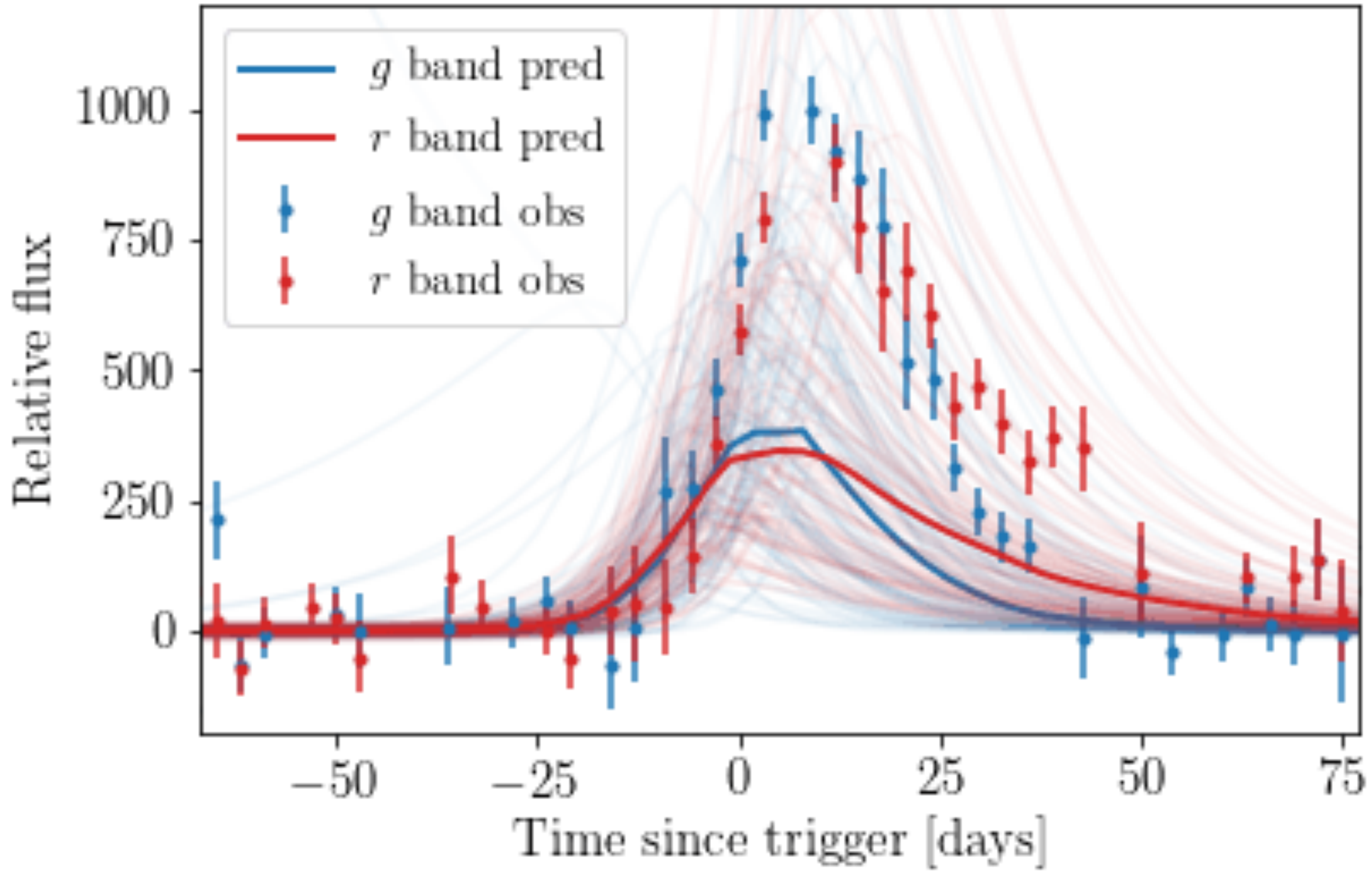
ZWICKY TRANSIENT FACILITY



Build regressive model of common transients



Build regressive model of common transients



Observed light curves
(up to time T)

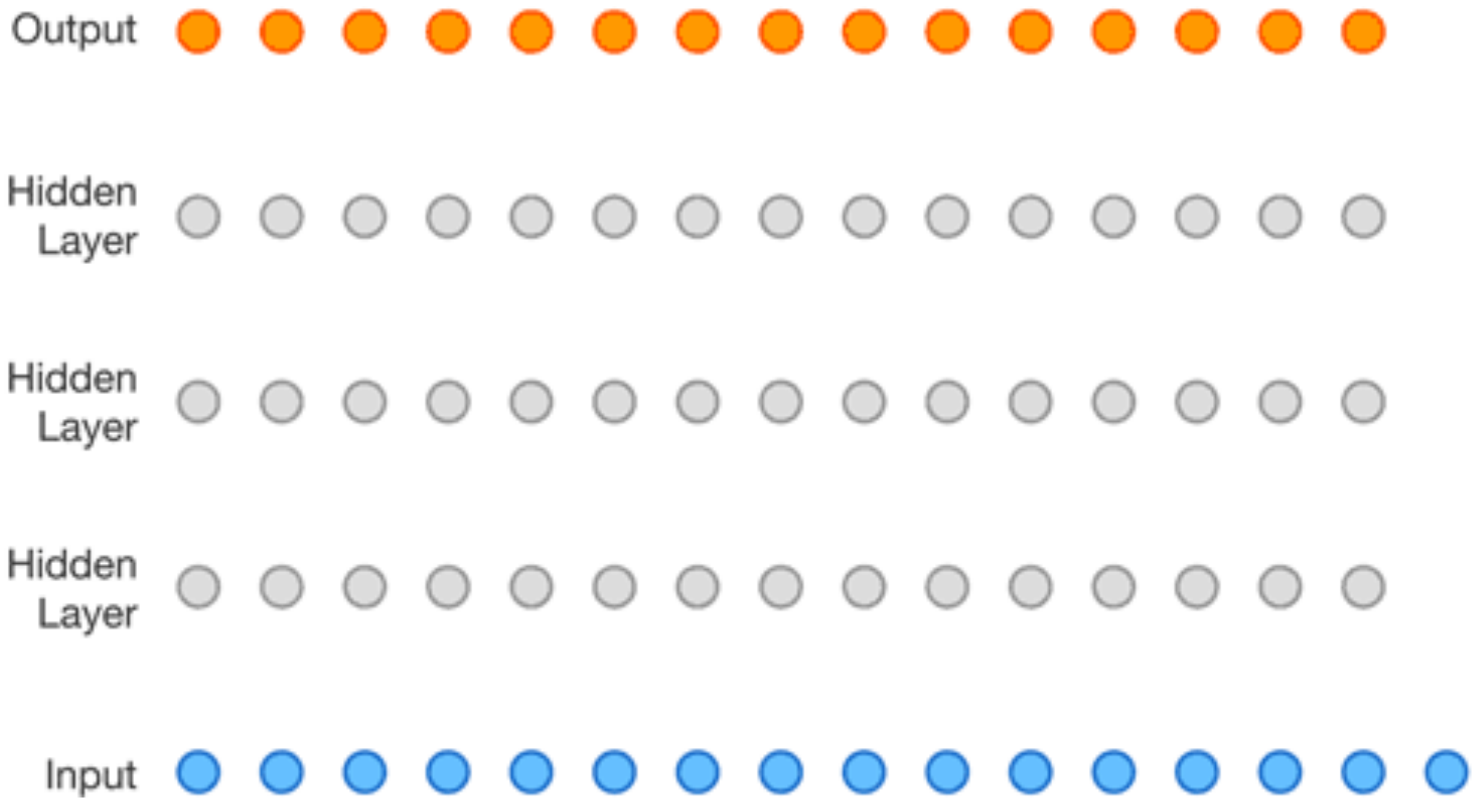
$\boldsymbol{D}_{sp(t\leq T)}$ → Model (DNN or Bazin) → $y_{sp(T+3)}$

$\boldsymbol{\sigma}_{D,sp(t\leq T)}$ → Model (DNN or Bazin) → $\sigma_{y,sp(T+3)}$

Next
predicted flux

Build regressive model of common transients

Model (1): Temporal Convolutional Neural Networks (probabilistic)

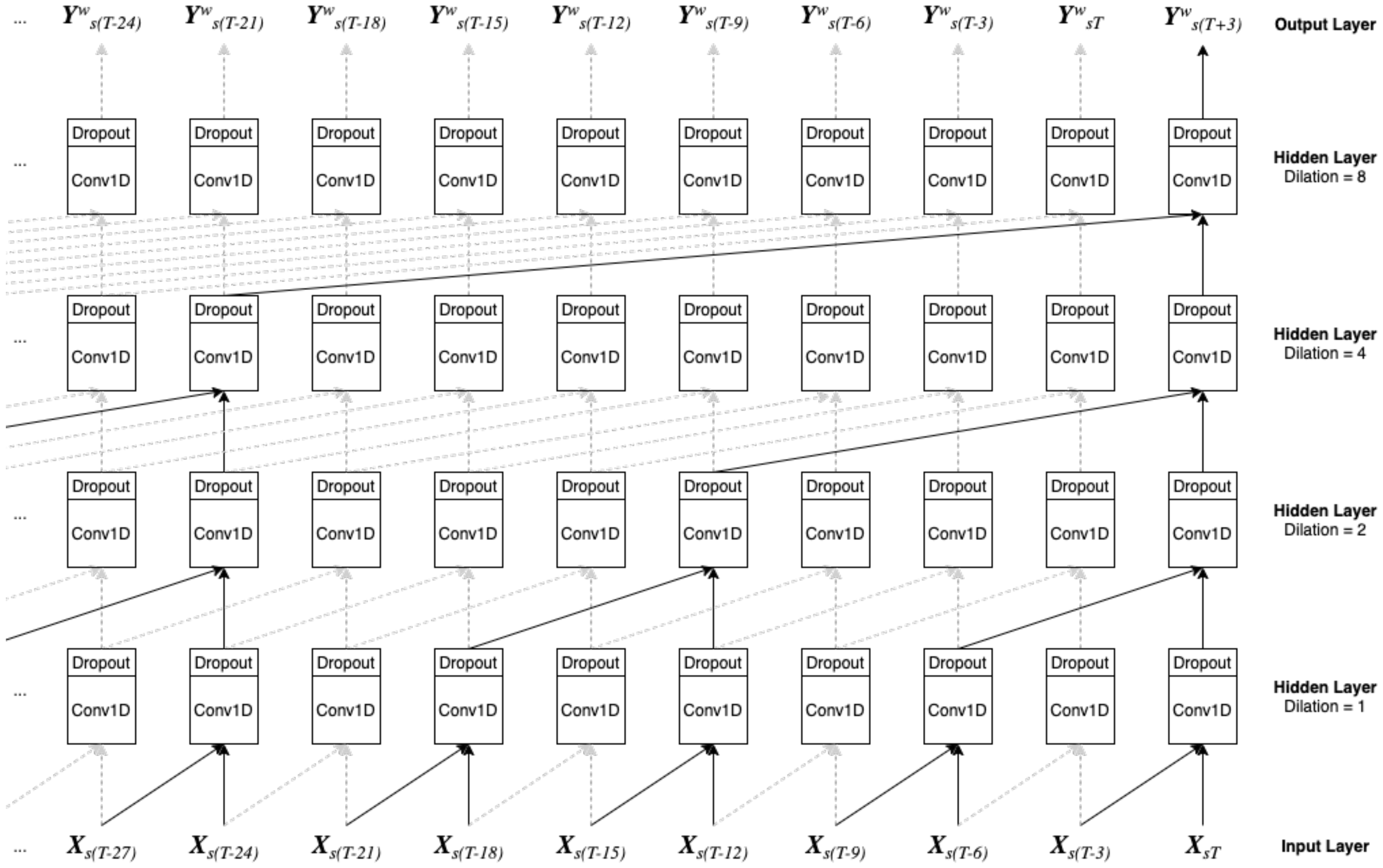


Model (2): Bayesian model based on the Bazin function (Bazin et al. 2009)

$$F(t) = A\frac{e^{-(t-t_0)/\tau_{\mathrm{fall}}}}{1+e^{-(t-t_0)/\tau_{\mathrm{rise}}}} + B + A\epsilon_{\mathrm{int}}(t)$$

Build regressive model of common transients

Model (1): Temporal Convolutional Neural Networks (probabilistic)



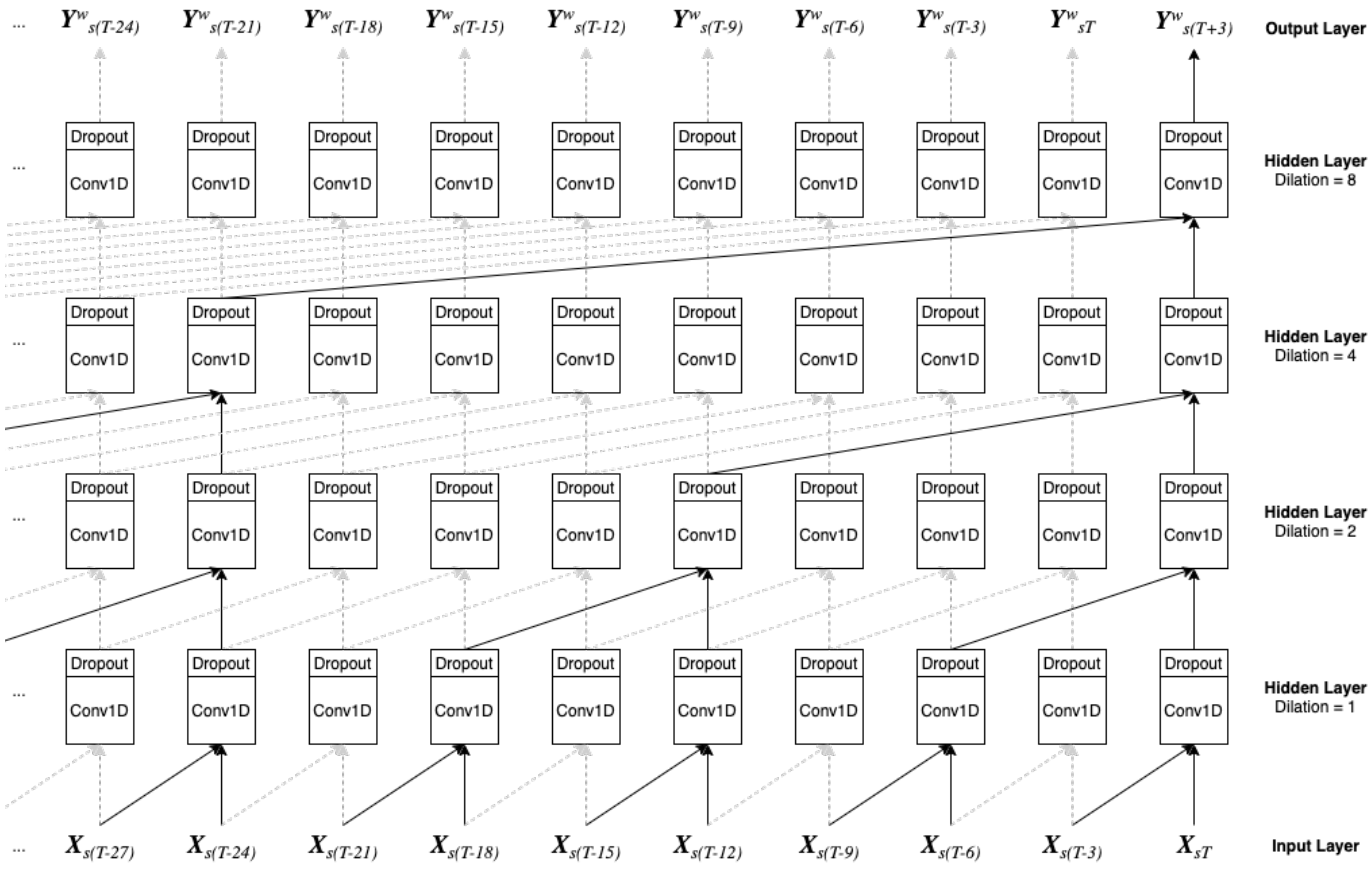
Model (2): Bayesian model based on the Bazin function (Bazin et al. 2009)

$$F(t) = A\frac{e^{-(t-t_0)/\tau_{\mathrm{fall}}}}{1+e^{-(t-t_0)/\tau_{\mathrm{rise}}}} + B + A\epsilon_{\mathrm{int}}(t)$$

Build regressive model of common transients

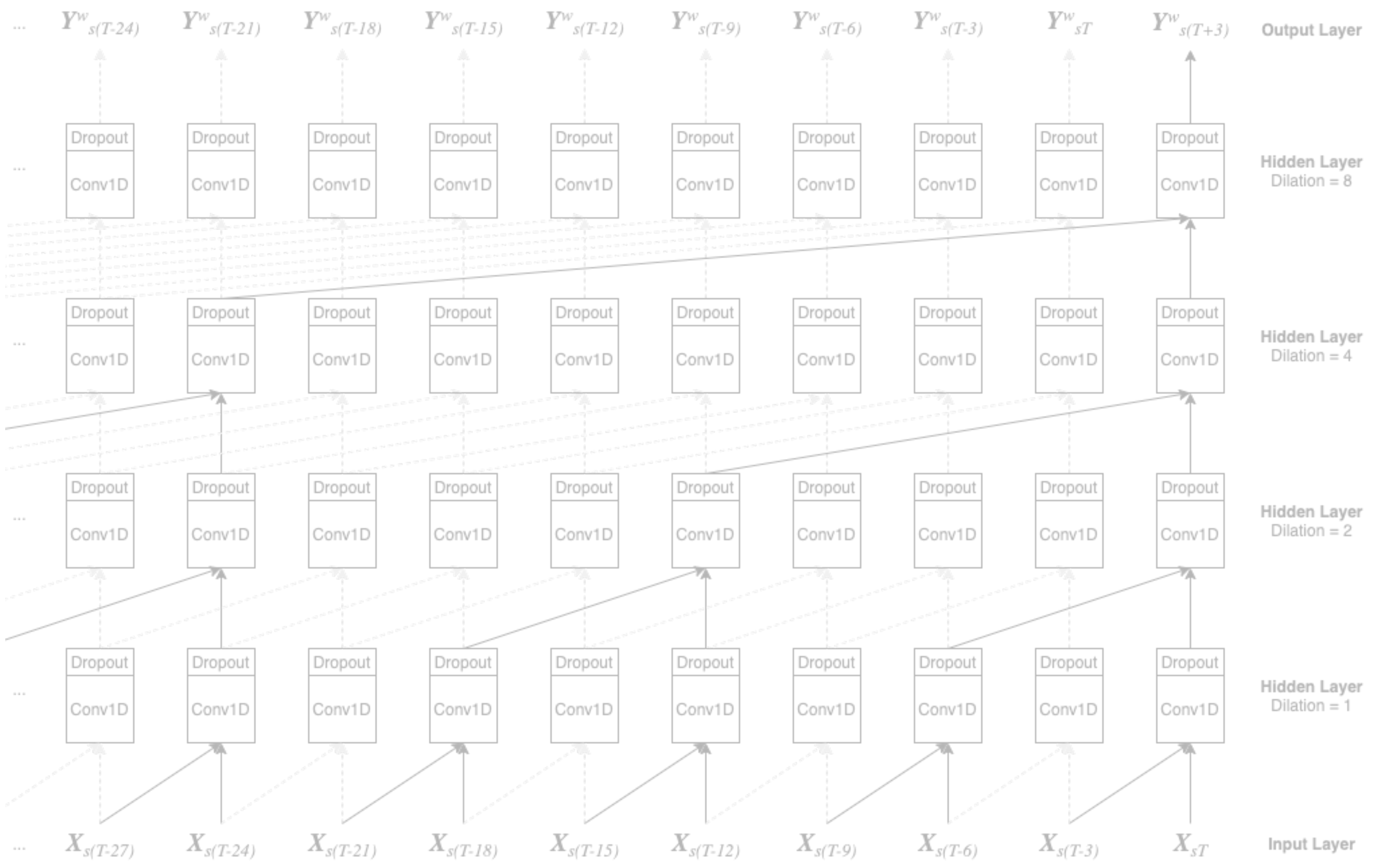
Model (1): Temporal Convolutional Neural Networks (probabilistic)

- Output parameterised as a Normal distribution using a probabilisitic neural network
- Include flux and predictive uncertainties in the loss function
- Bayesian Neural Network using MCDropout (Gal & Ghahramani 2015)



Build regressive model of common transients

Model (1): Temporal Convolutional Neural Networks (probabilistic)

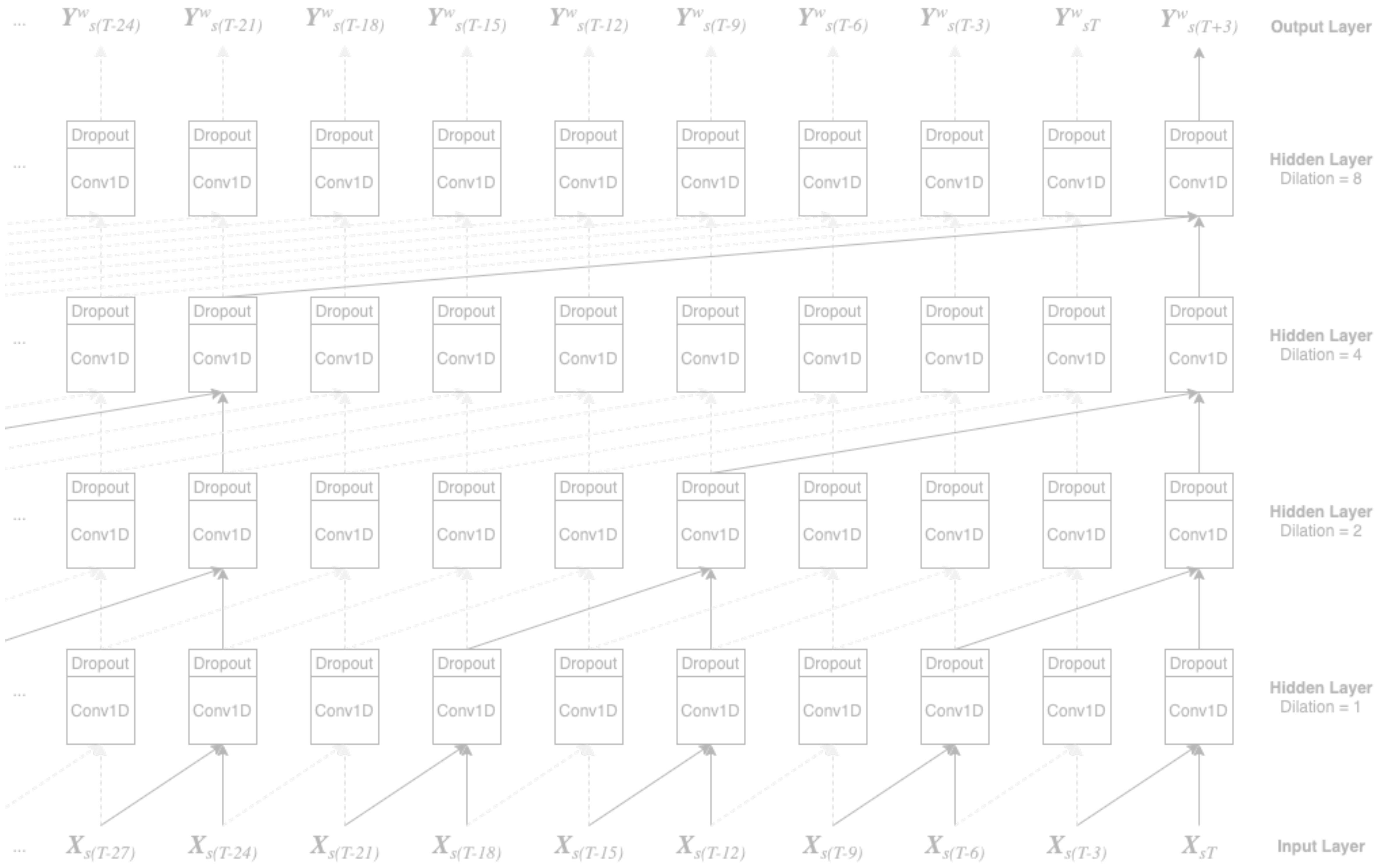


Model (2): Bayesian model based on the Bazin function (Bazin et al. 2009)

$$F(t) = A\frac{e^{-(t-t_0)/\tau_{\rm fall}}}{1+e^{-(t-t_0)/\tau_{\rm rise}}} + B + A\epsilon_{\rm int}(t)$$

Build regressive model of common transients

Model (1): Temporal Convolutional Neural Networks (probabilistic)



Model (2): Bayesian model based on the Bazin function (Bazin et al. 2009)

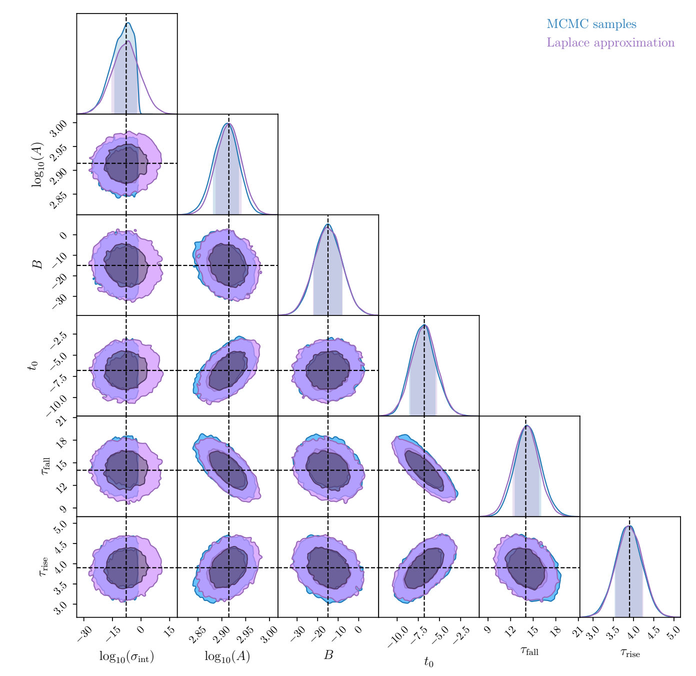
$$D_{spt} = A \frac{e^{-(t-t_0)/\tau_{\text{fall}}}}{1 + e^{-(t-t_0)/\tau_{\text{rise}}}} + B + A\epsilon_{\text{int}}(t) + \epsilon_{D,spt}$$

Bayesian model

$$P(\boldsymbol{\theta} \,|\, \boldsymbol{D}_{sp}, \boldsymbol{t}) \propto P(\boldsymbol{D}_{sp} \,|\, \boldsymbol{t}, \boldsymbol{\theta}) \; P(\boldsymbol{\theta})$$

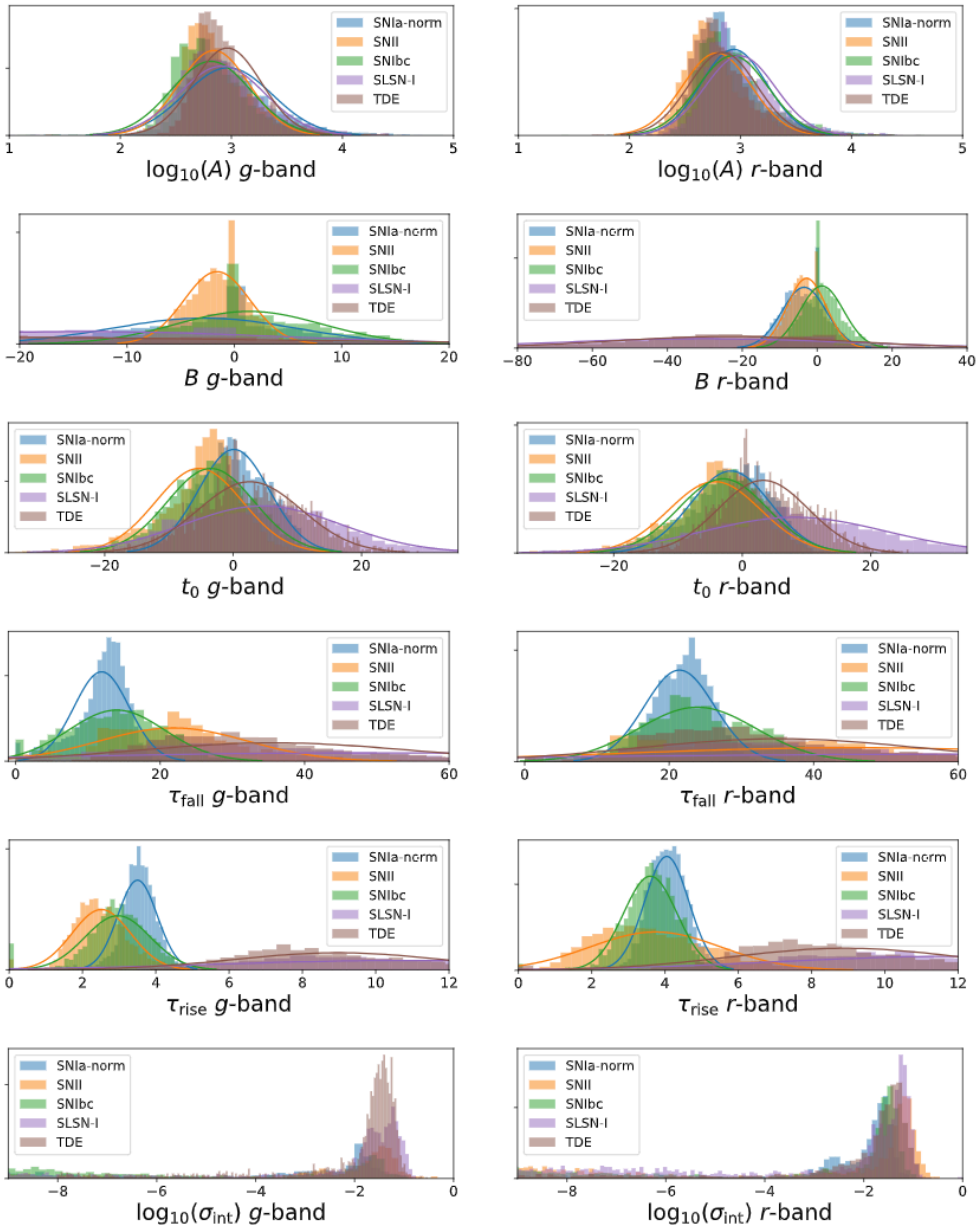
Posterior ∝ Likelihood x Prior

$$P(\boldsymbol{D}_{sp} | \boldsymbol{t}, \boldsymbol{\theta}) = \prod_{t=-70}^{T} N\left(D_{spt} \mid F_{spt}(\boldsymbol{\theta}), A^2\sigma_{\text{int}}^2 + \sigma_{D,spt}^2\right)$$

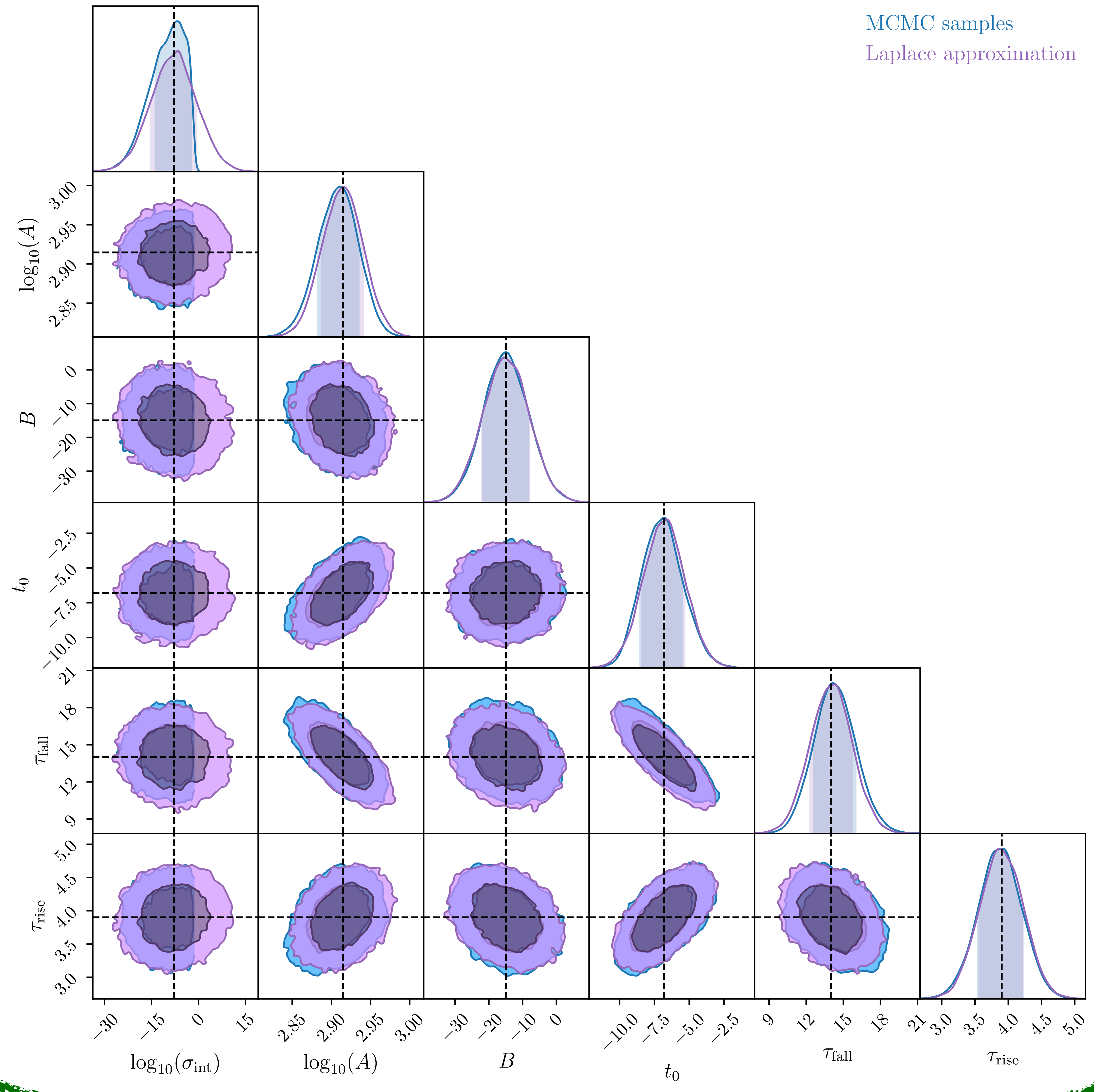


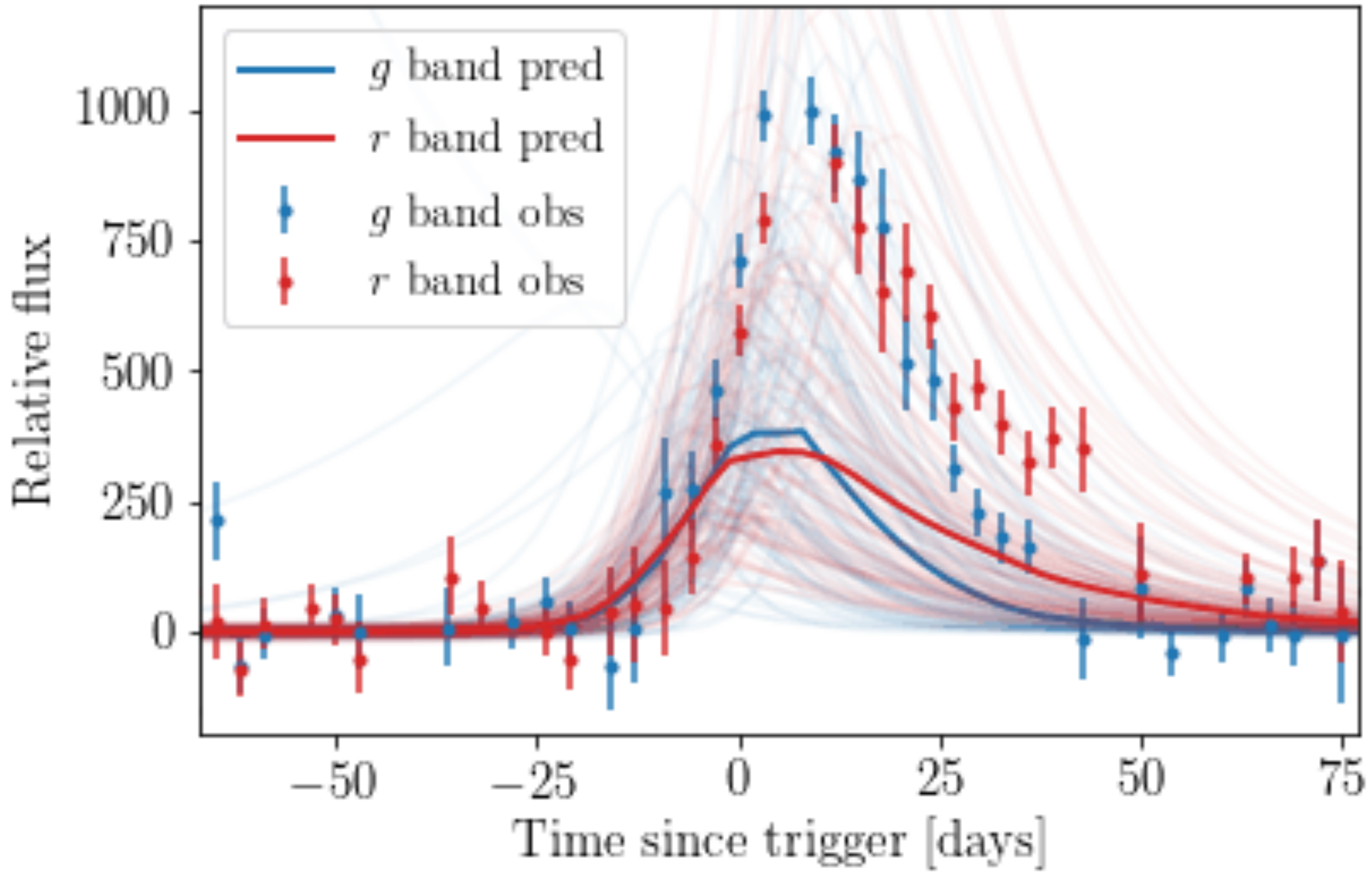
Bazin model

$$D_{spt} = A\frac{e^{-(t-t_0)/\tau_{\text{fall}}}}{1+e^{-(t-t_0)/\tau_{\text{rise}}}} + B + A\epsilon_{\text{int}}(t) + \epsilon_{D,spt}$$

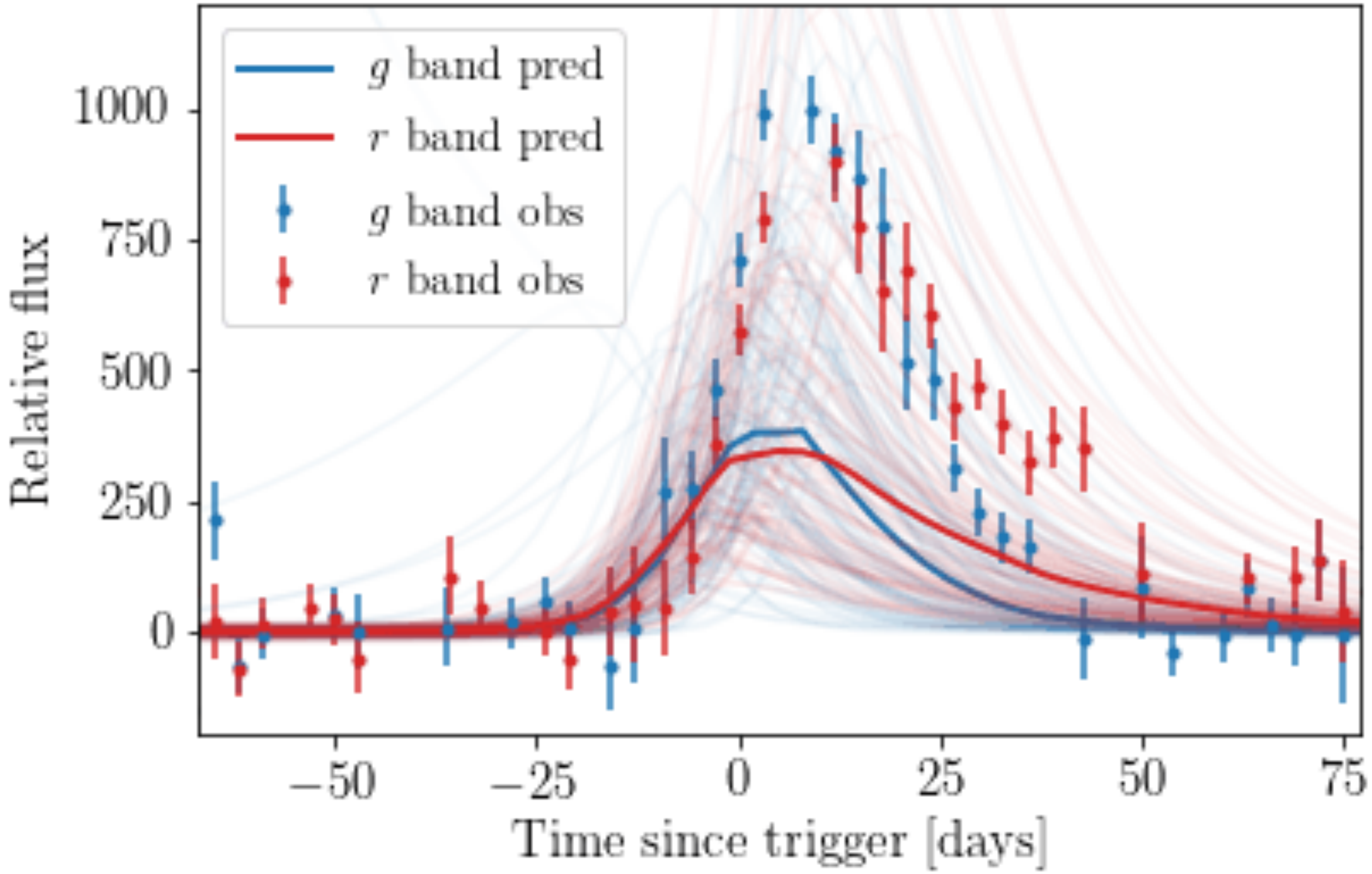


Bayesian model



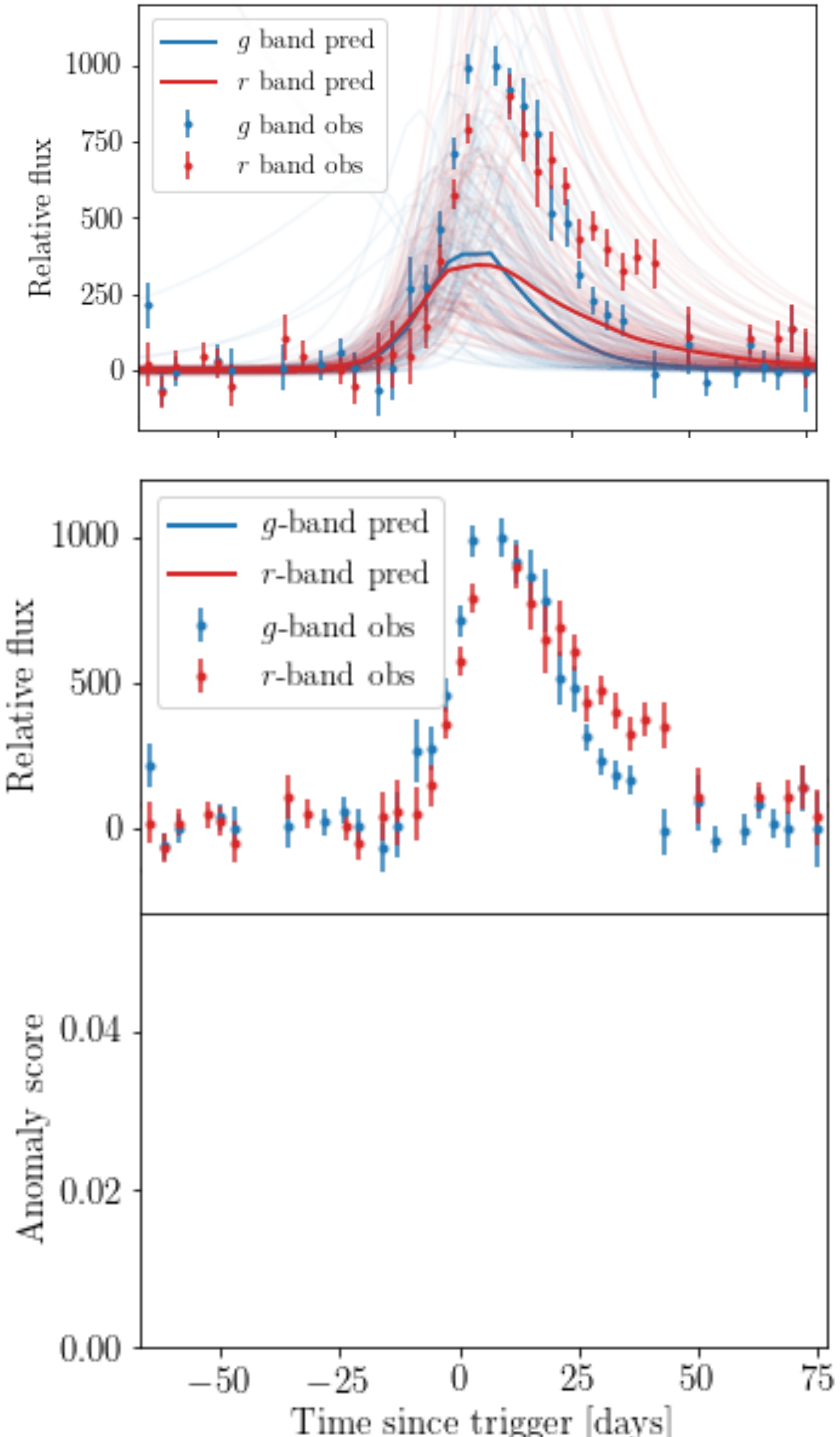


Build regressive model of common transients

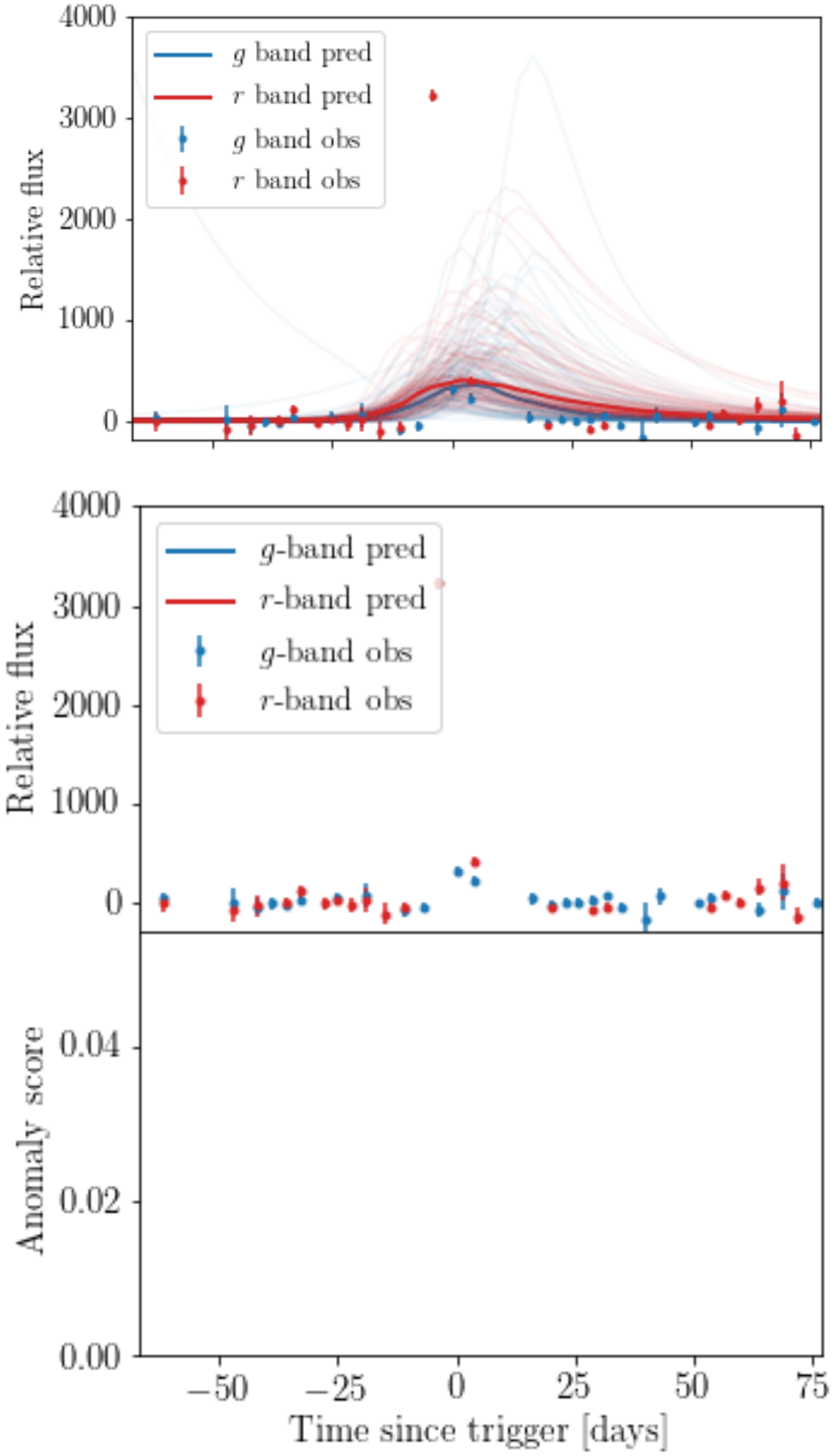


$$\text{Anomaly score} = \frac{1}{n_p}\sum_{p=1}^{n_p}\frac{(\text{pred}_p - \text{obs}_p)^2}{\sigma^2_{\text{pred},p} + \sigma^2_{\text{obs},p}}$$

SNIa

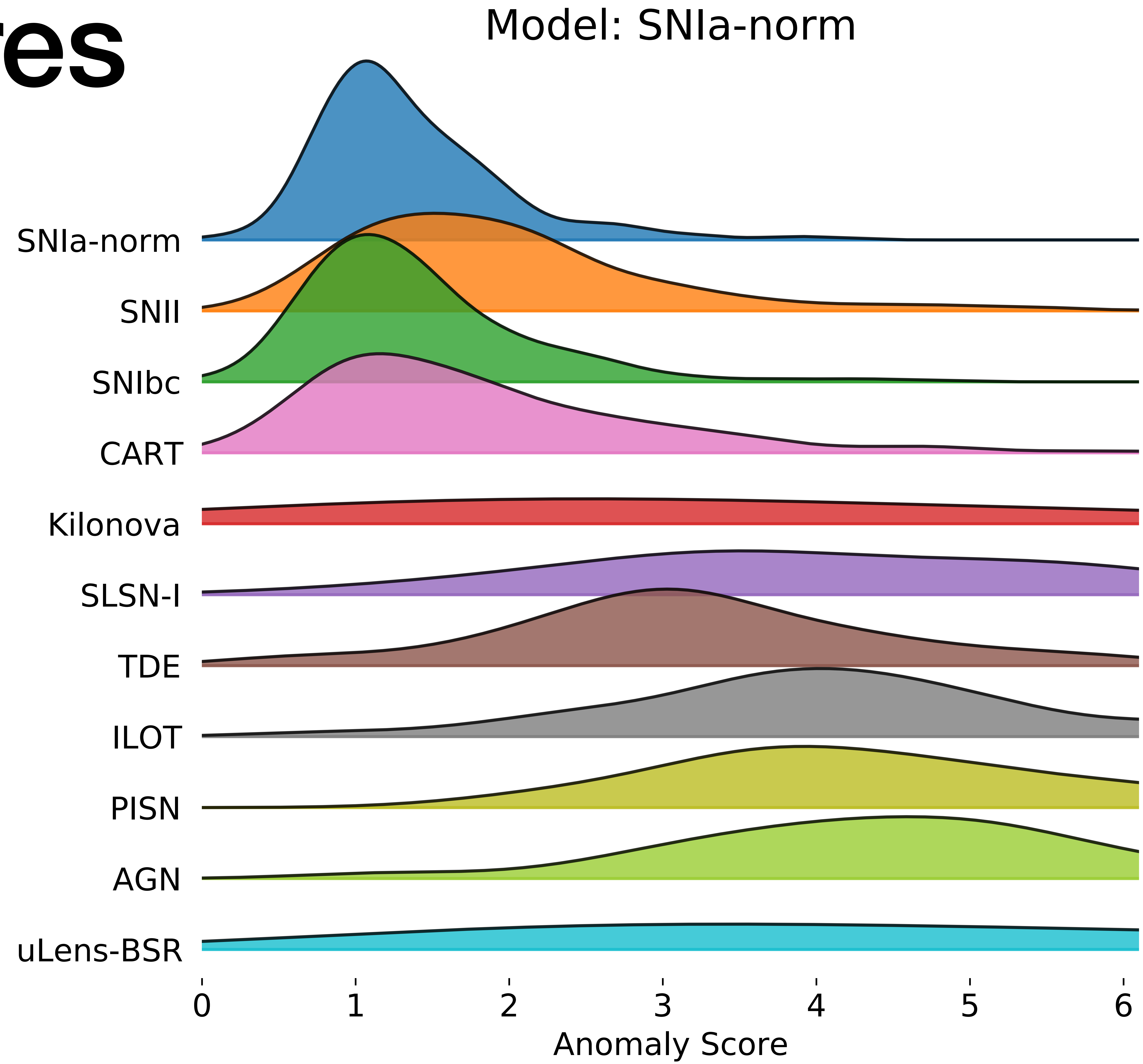


Kilonova



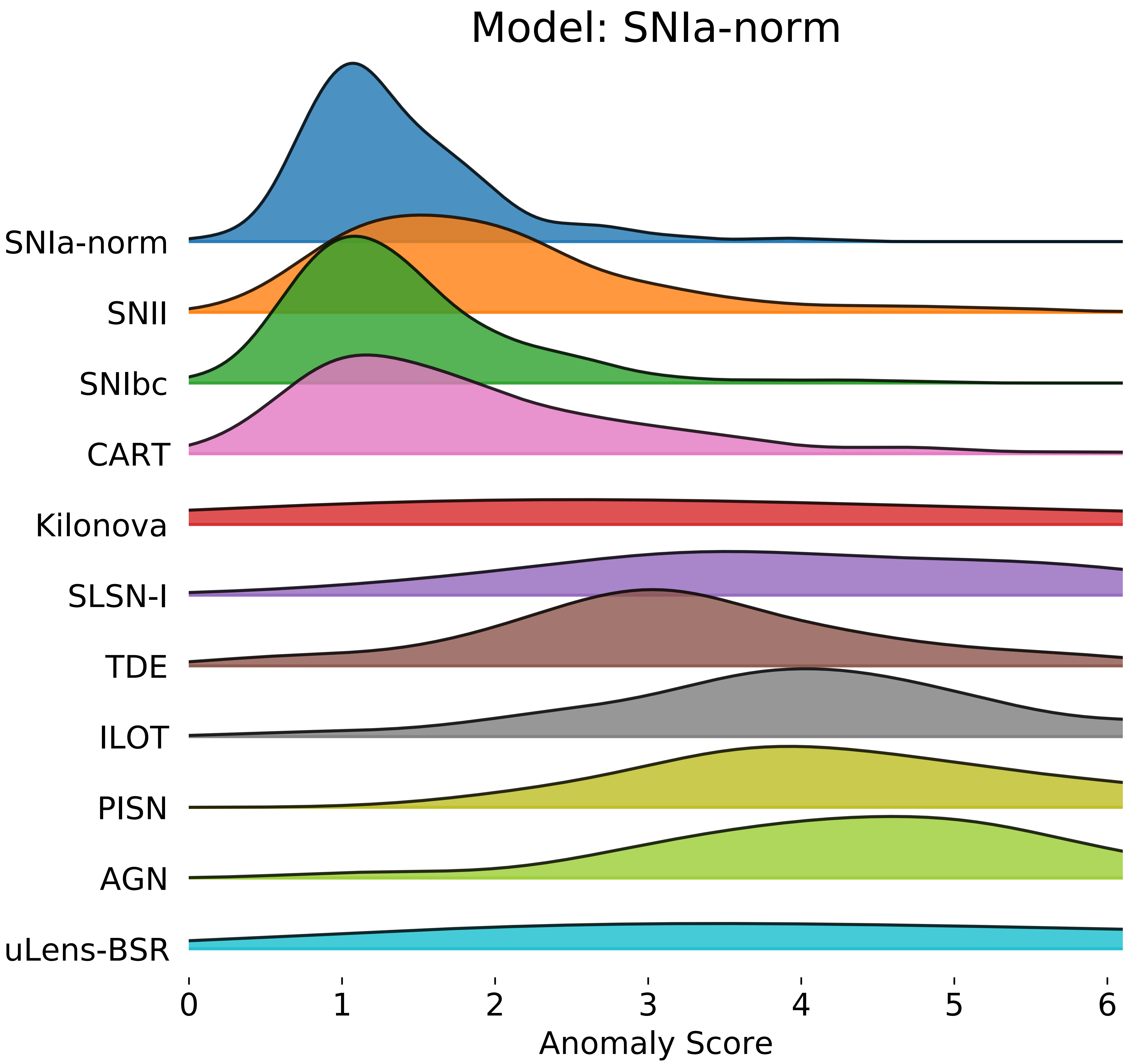
Anomaly scores

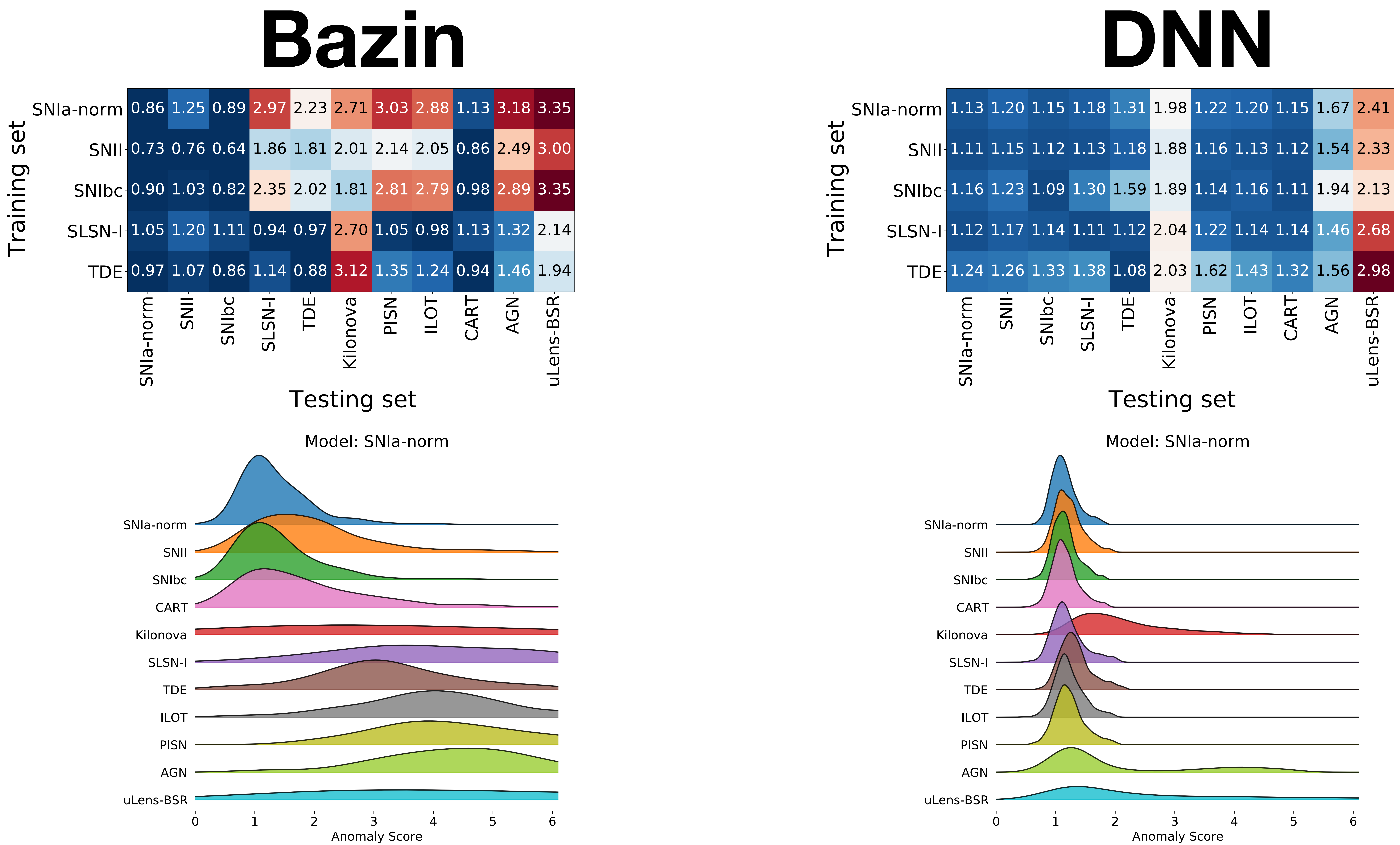
- Common supernovae have similar scores when using the Bazin SNIa model
- Anomalous classes have higher anomaly scores



Anomaly scores

Training set \ Testing set	SNIa-norm	SNII	SNIbc	SLSN-I	TDE	Kilonova	PISN	ILOT	CART	AGN	uLens-BSR
SNIa-norm	0.86	1.25	0.89	2.97	2.23	2.71	3.03	2.88	1.13	3.18	3.35
SNII	0.73	0.76	0.64	1.86	1.81	2.01	2.14	2.05	0.86	2.49	3.00
SNIbc	0.90	1.03	0.82	2.35	2.02	1.81	2.81	2.79	0.98	2.89	3.35
SLSN-I	1.05	1.20	1.11	0.94	0.97	2.70	1.05	0.98	1.13	1.32	2.14
TDE	0.97	1.07	0.86	1.14	0.88	3.12	1.35	1.24	0.94	1.46	1.94

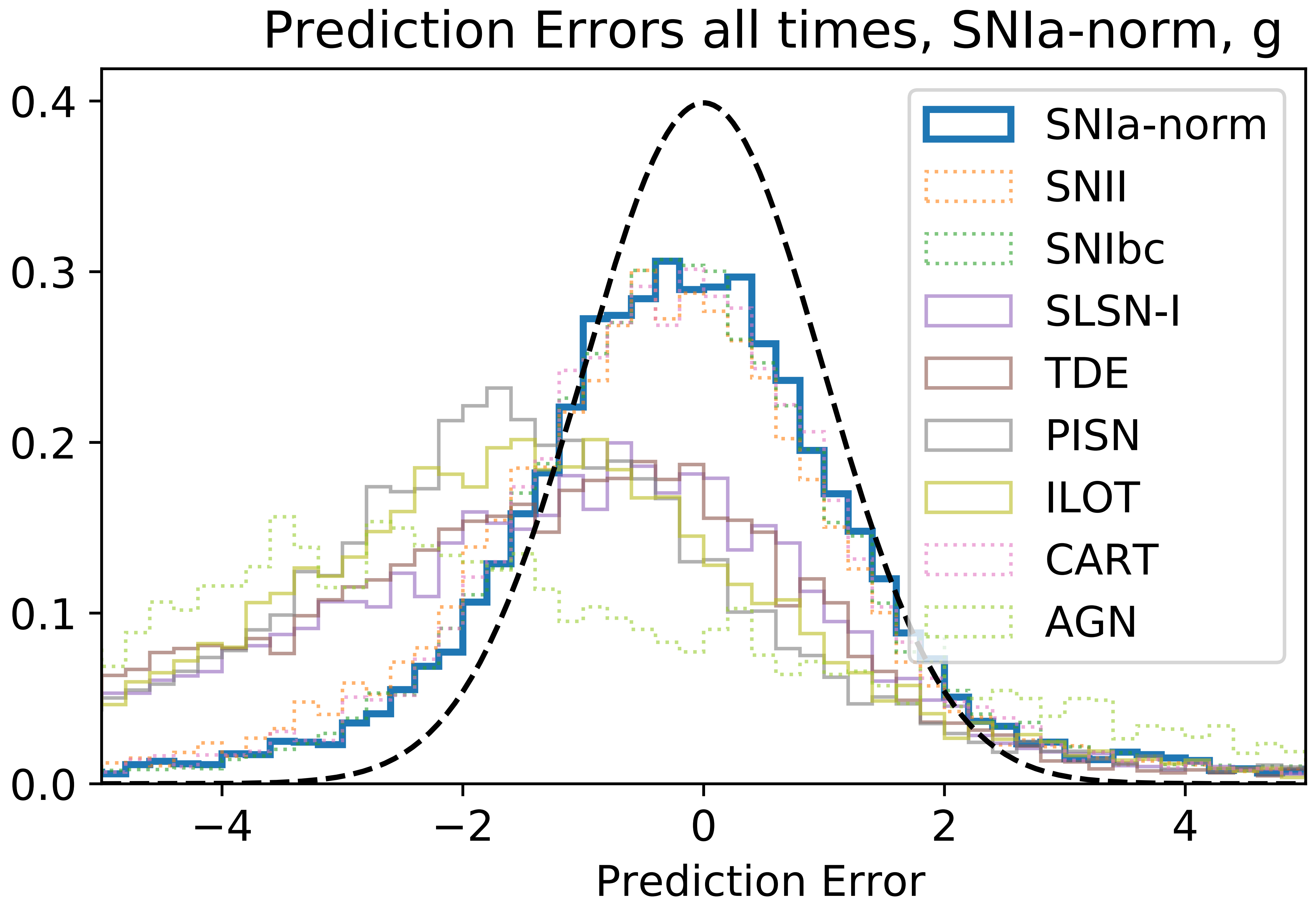




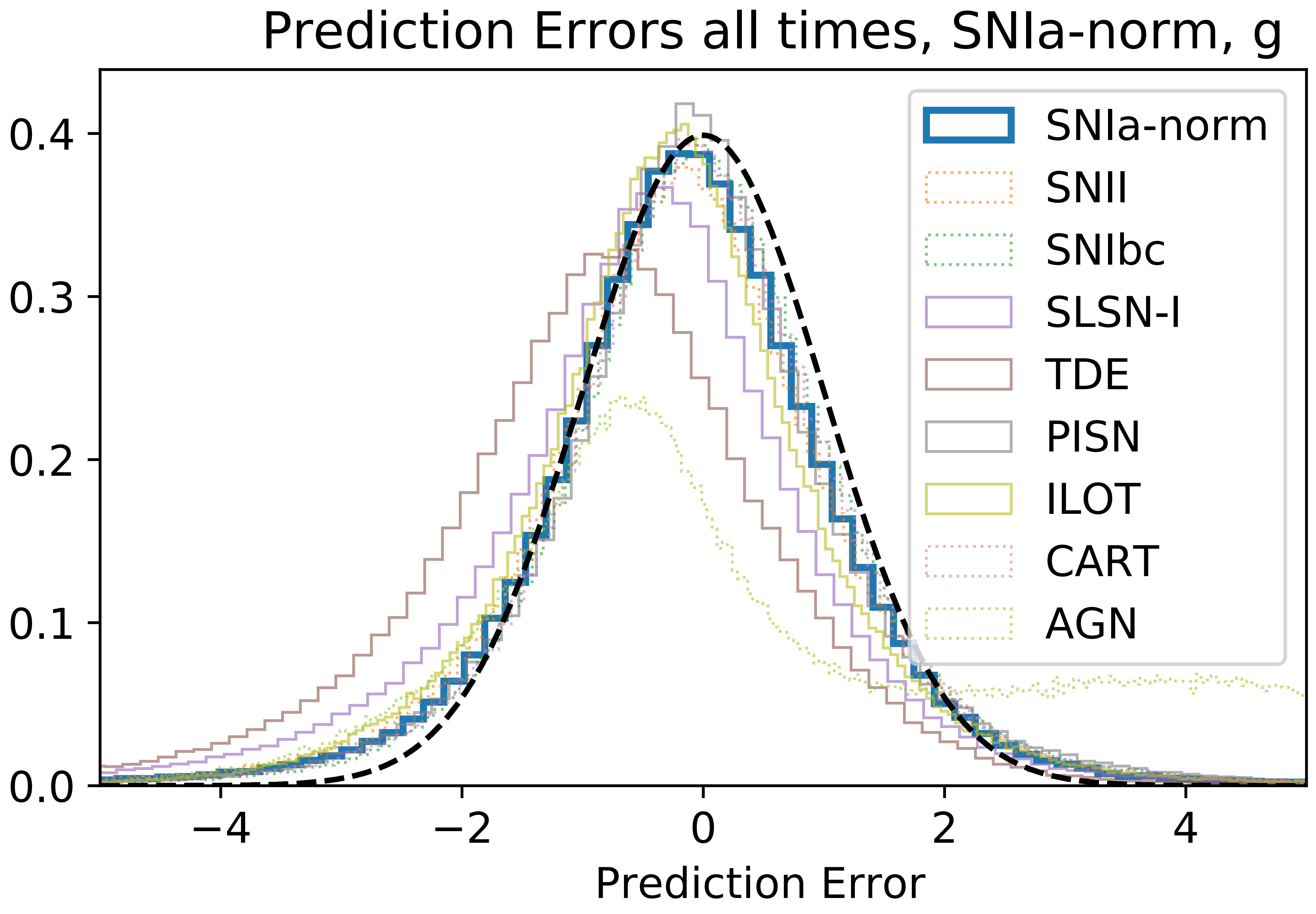
Bazin
DNN
Training set
Testing set
SNIa-norm 0.86 1.25 0.89 2.97 2.23 2.71 3.03 2.88 1.13 3.18 3.35
SNII 0.73 0.76 0.64 1.86 1.81 2.01 2.14 2.05 0.86 2.49 3.00
SNIbc 0.90 1.03 0.82 2.35 2.02 1.81 2.81 2.79 0.98 2.89 3.35
SLSN-I 1.05 1.20 1.11 0.94 0.97 2.70 1.05 0.98 1.13 1.32 2.14
TDE 0.97 1.07 0.86 1.14 0.88 3.12 1.35 1.24 0.94 1.46 1.94
SNIa-norm SNII SNIbc SLSN-I TDE Kilonova PISN ILOT CART AGN uLens-BSR
SNIa-norm 1.13 1.20 1.15 1.18 1.31 1.98 1.22 1.20 1.15 1.67 2.41
SNII 1.11 1.15 1.12 1.13 1.18 1.88 1.16 1.13 1.12 1.54 2.33
SNIbc 1.16 1.23 1.09 1.30 1.59 1.89 1.14 1.16 1.11 1.94 2.13
SLSN-I 1.12 1.17 1.14 1.11 1.12 2.04 1.22 1.14 1.14 1.46 2.68
TDE 1.24 1.26 1.33 1.38 1.08 2.03 1.62 1.43 1.32 1.56 2.98
Model: SNIa-norm
SNIa-norm
SNII
SNIbc
CART
Kilonova
SLSN-I
TDE
ILOT
PISN
AGN
uLens-BSR
0 1 2 3 4 5 6
Anomaly Score

Neural networks are too good at regression for anomaly detection

Bazin

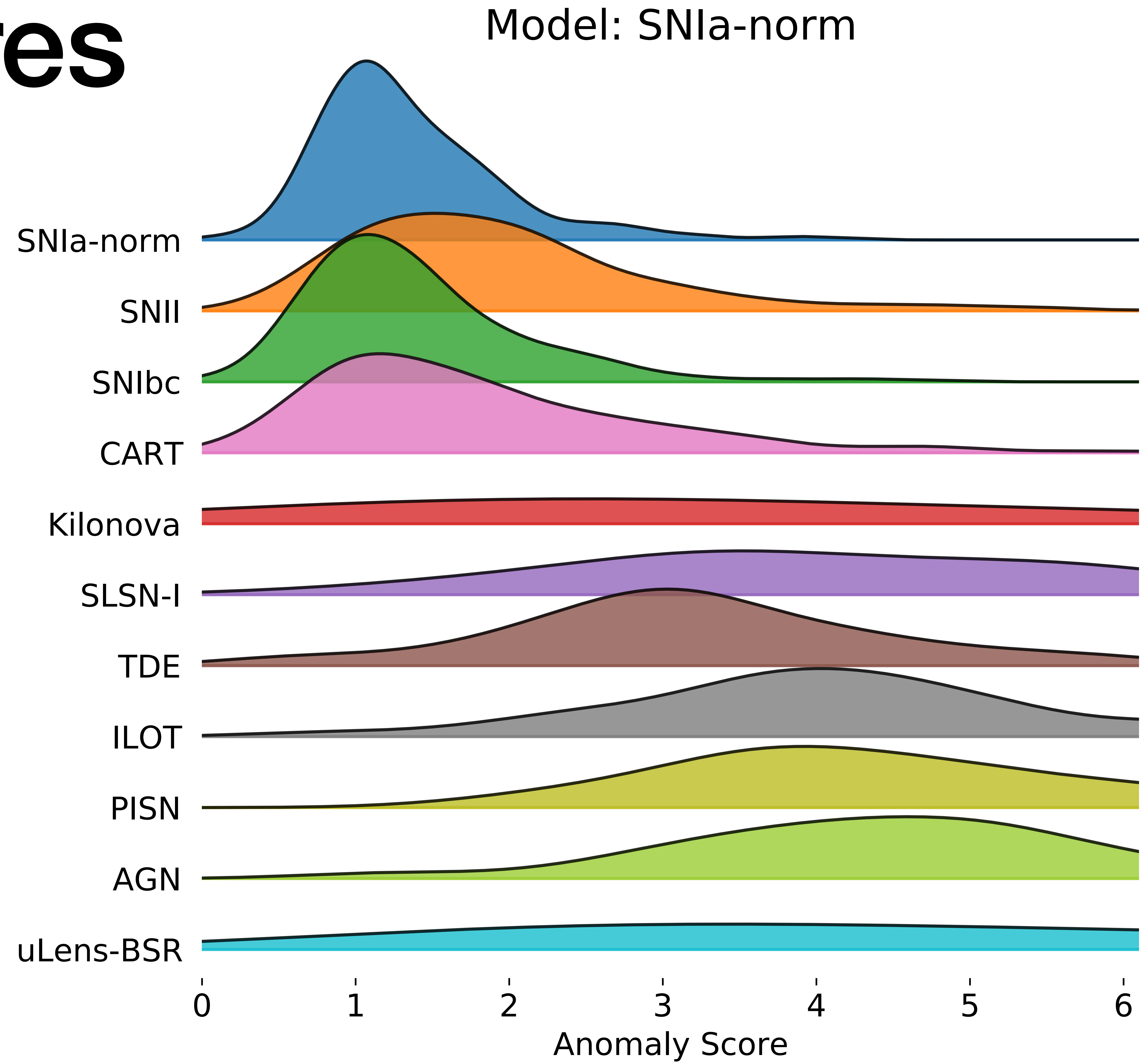


DNN

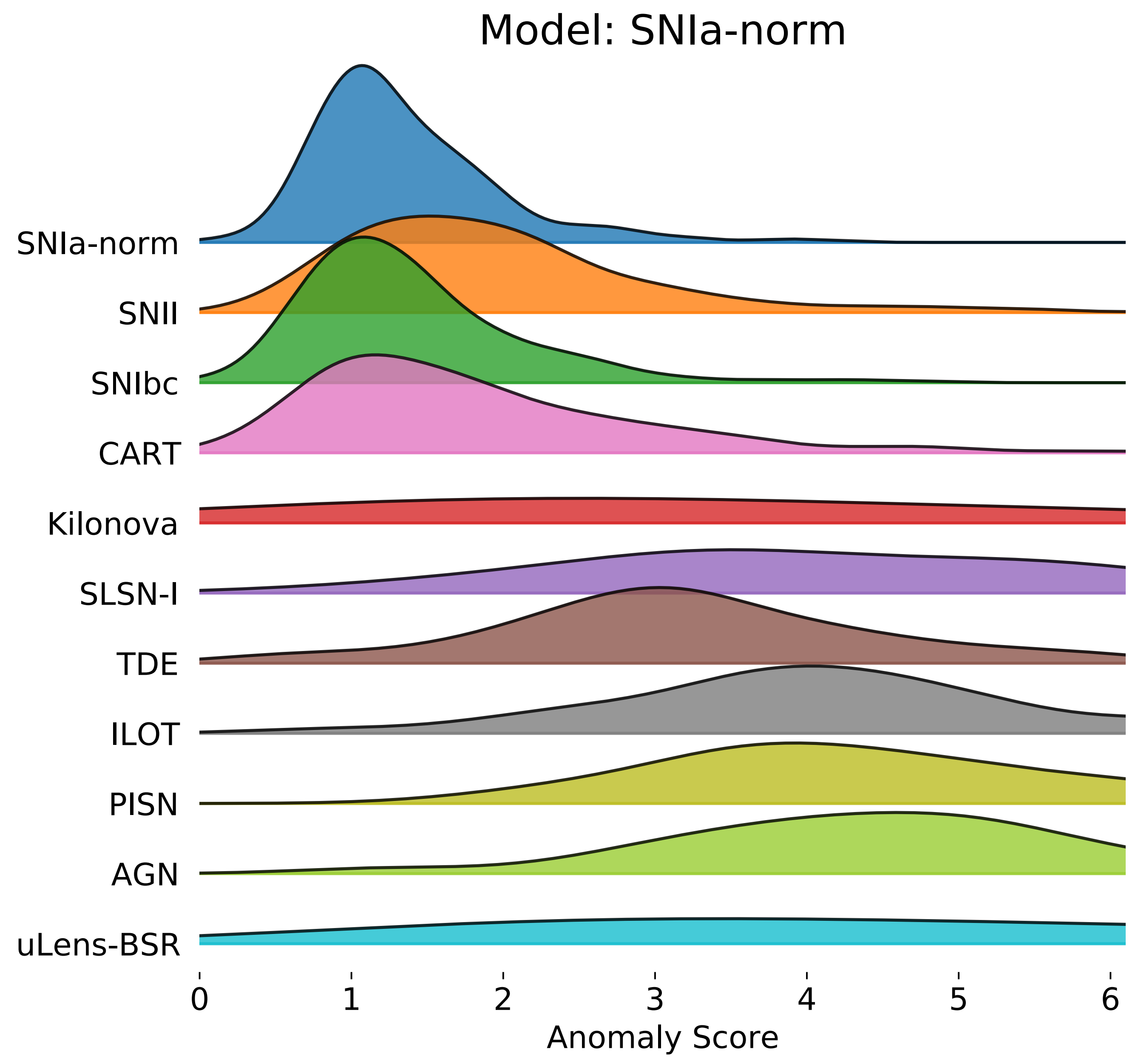


$$\text{Prediction Error} = \frac{y_{pred} - y_{obs}}{\sigma_{obs}}$$

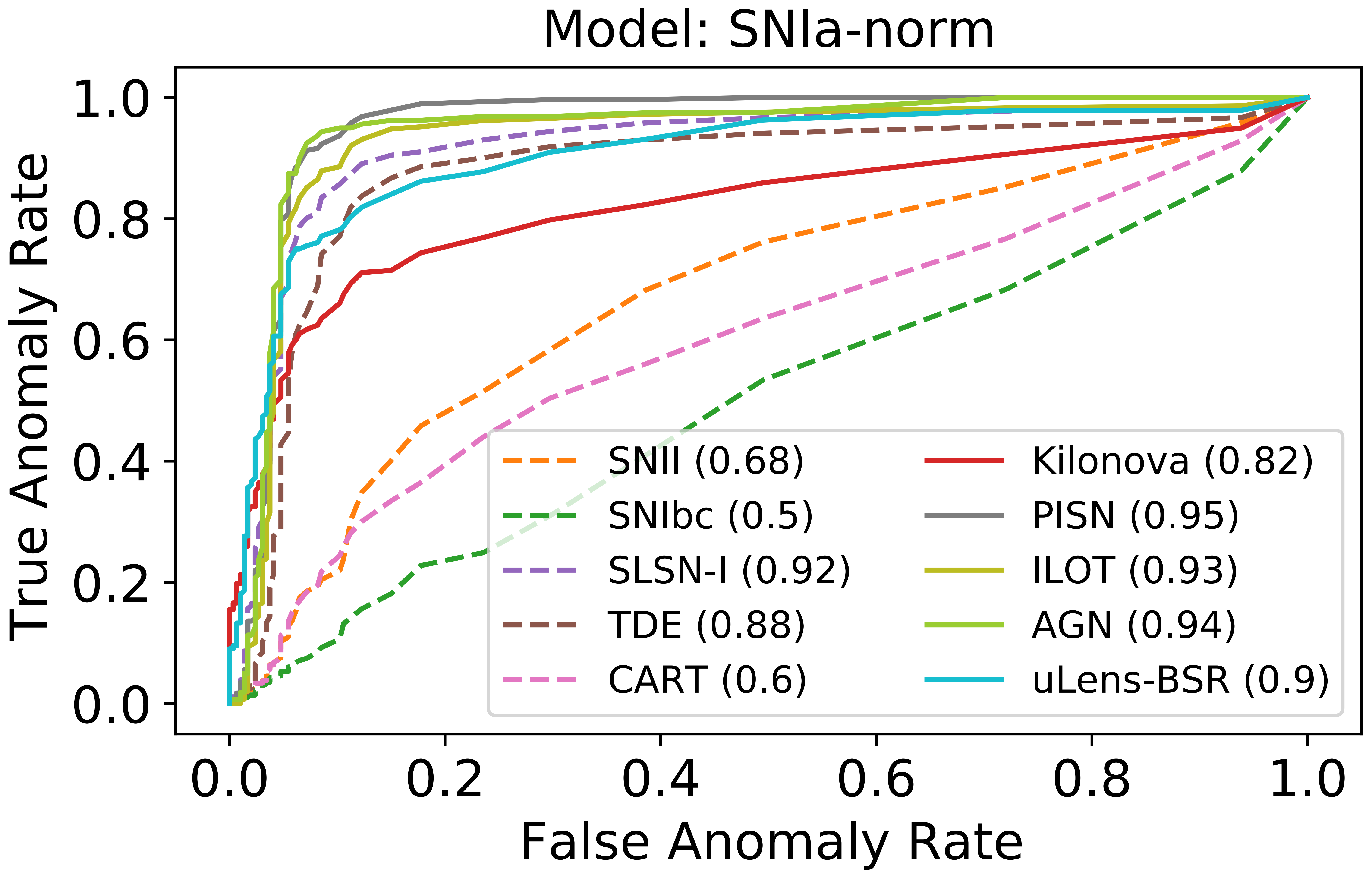
Anomaly scores



ROC Curves



Model: SNIa-norm
SNIa-norm
SNII
SNIbc
CART
Kilonova
SLSN-I
TDE
ILOT
PISN
AGN
uLens-BSR
0 1 2 3 4 5 6
Anomaly Score

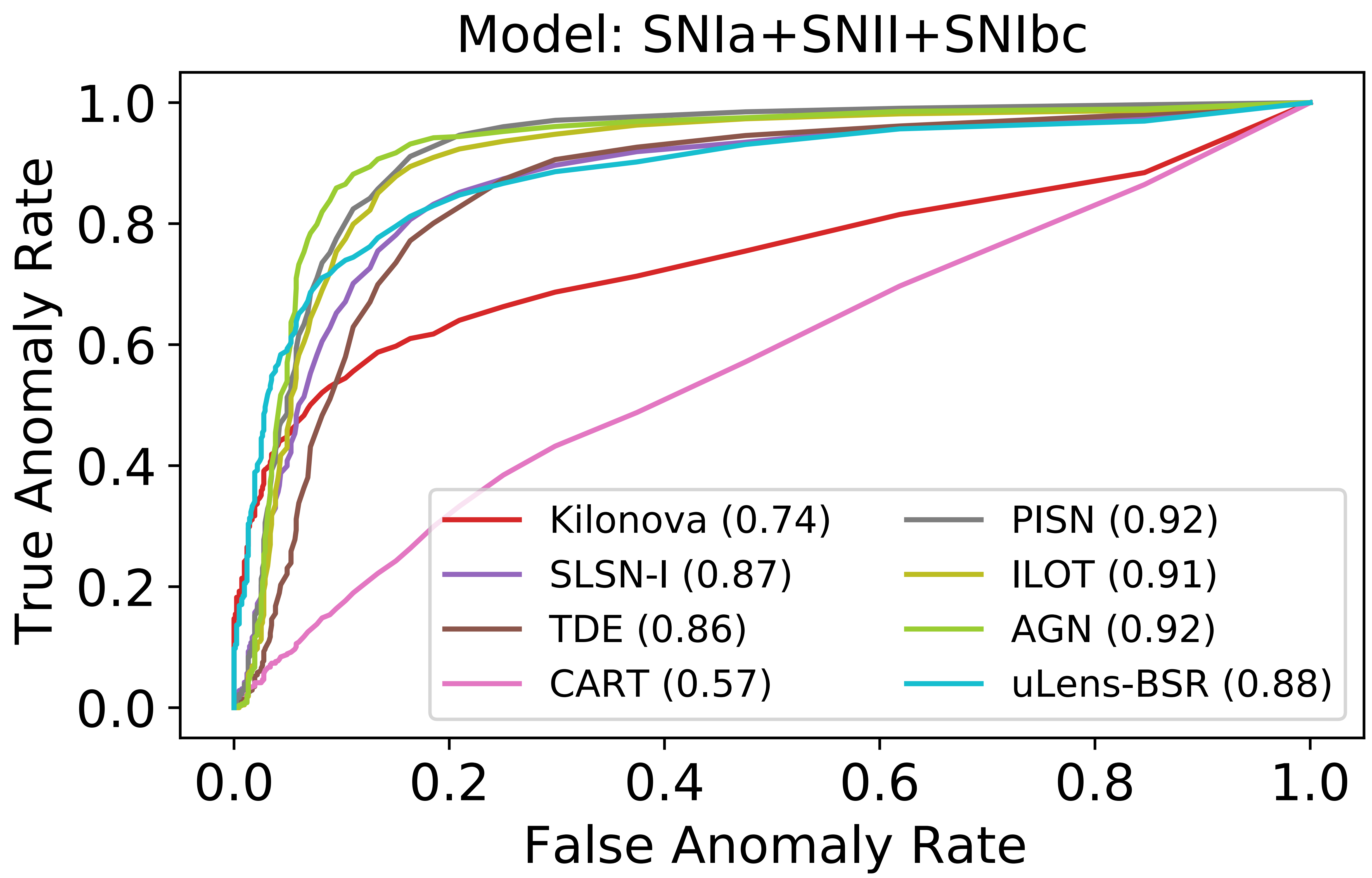


Model: SNIa-norm
True Anomaly Rate
False Anomaly Rate
0.0 0.2 0.4 0.6 0.8 1.0
SNII (0.68)
SNIbc (0.5)
SLSN-I (0.92)
TDE (0.88)
CART (0.6)
Kilonova (0.82)
PISN (0.95)
ILOT (0.93)
AGN (0.94)
uLens-BSR (0.9)

ROC Curves

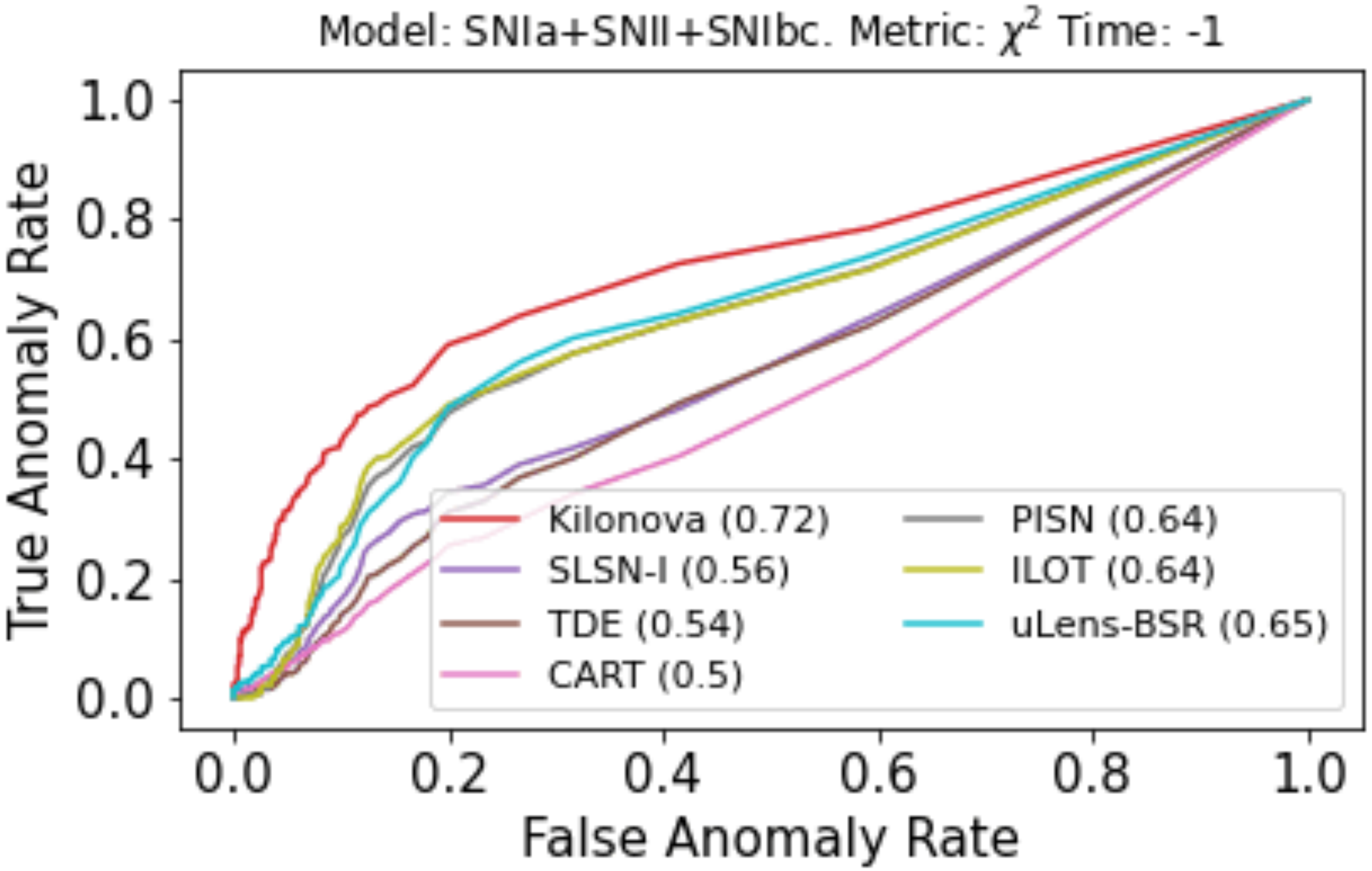
Common Supernovae

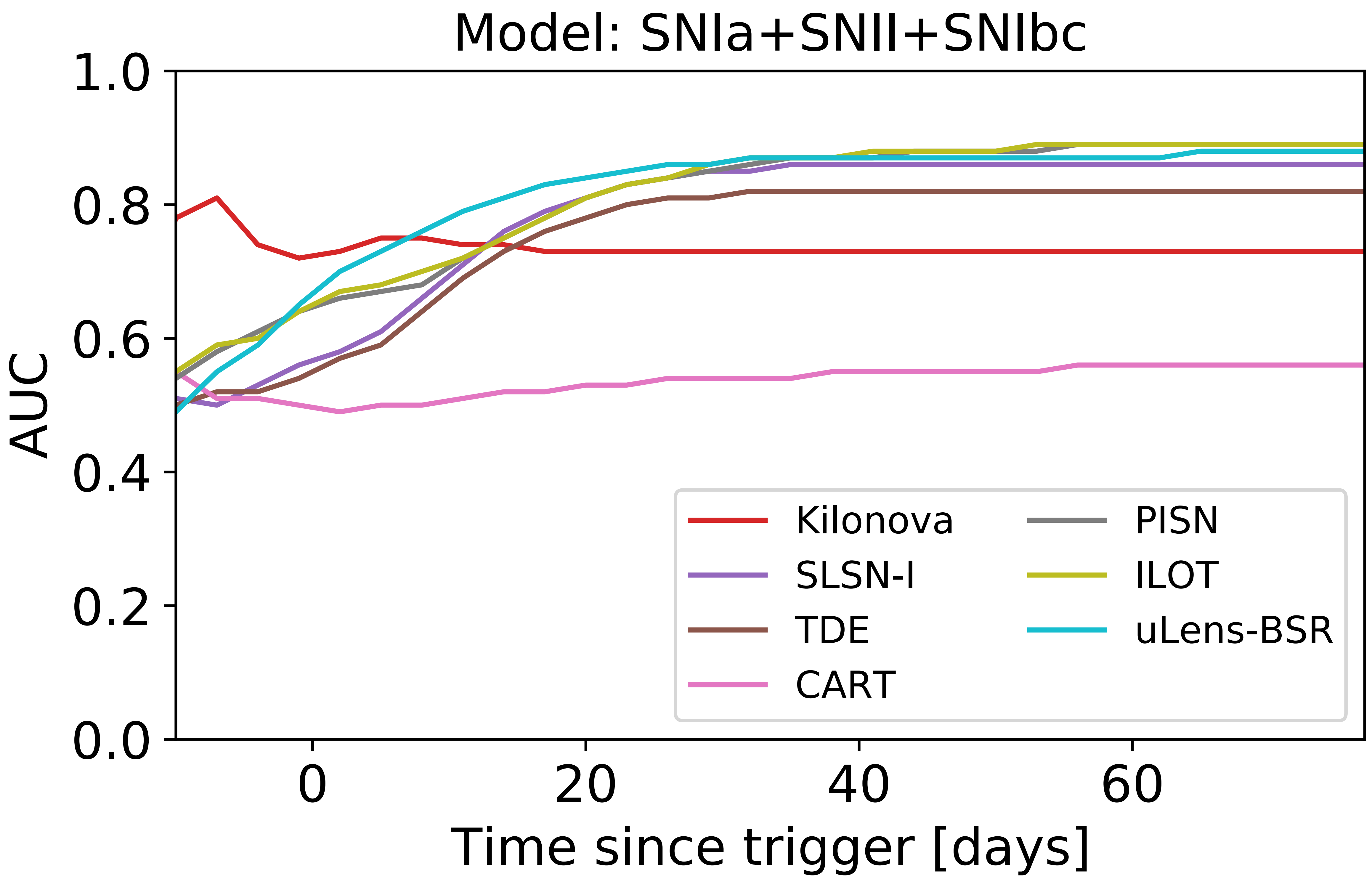
ROC Curves



Model: SNIa+SNII+SNIbc
True Anomaly Rate
False Anomaly Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Kilonova (0.74)
SLSN-I (0.87)
TDE (0.86)
CART (0.57)
PISN (0.92)
ILOT (0.91)
AGN (0.92)
uLens-BSR (0.88)

Real-time ROC curves





Conclusion

- Developed two frameworks to model common transient classes using
  - (1) a Deep Neural Network
  - (2) a Bayesian model based on the Bazin function
- Built models of the SNIa, SNII, SNIbc, SLSNe, TDE transient classes
- Can detect anomalies in real-time, useful for prioritised follow-up in new large scale transient surveys
- Fast and scaleable to model tens of thousands of events that will be discovered in LSST and ZTF within a few seconds