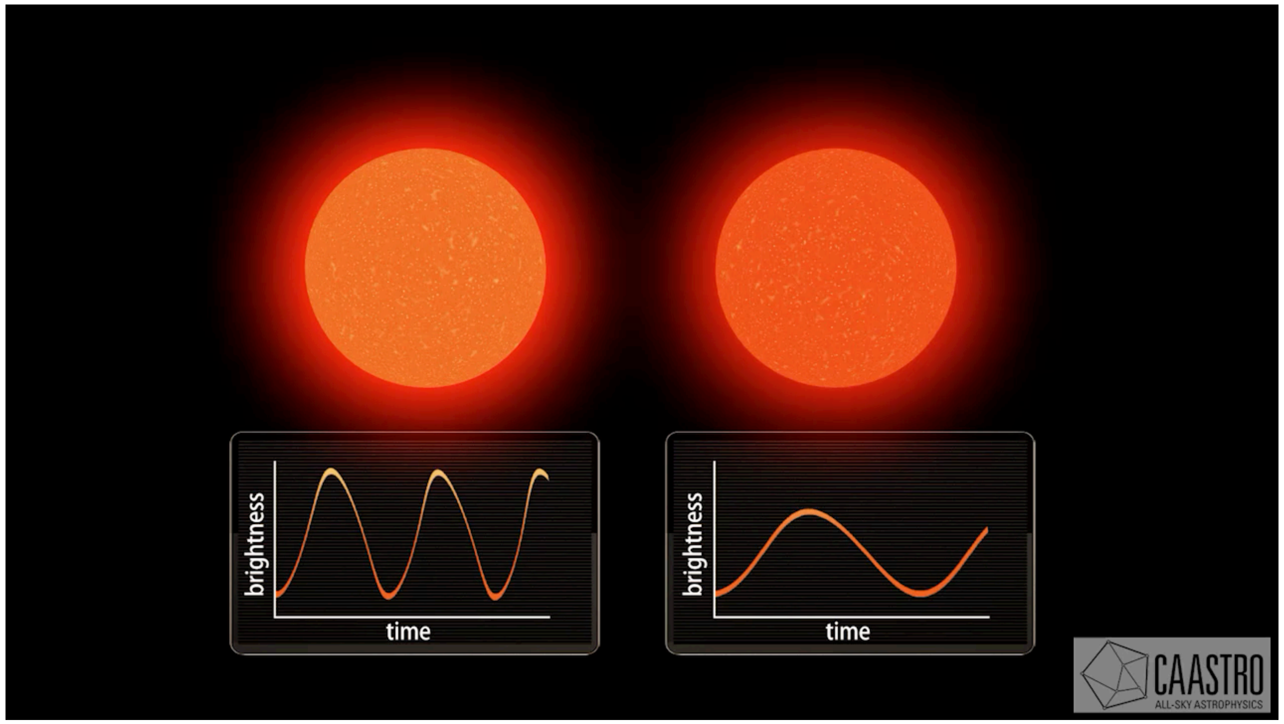
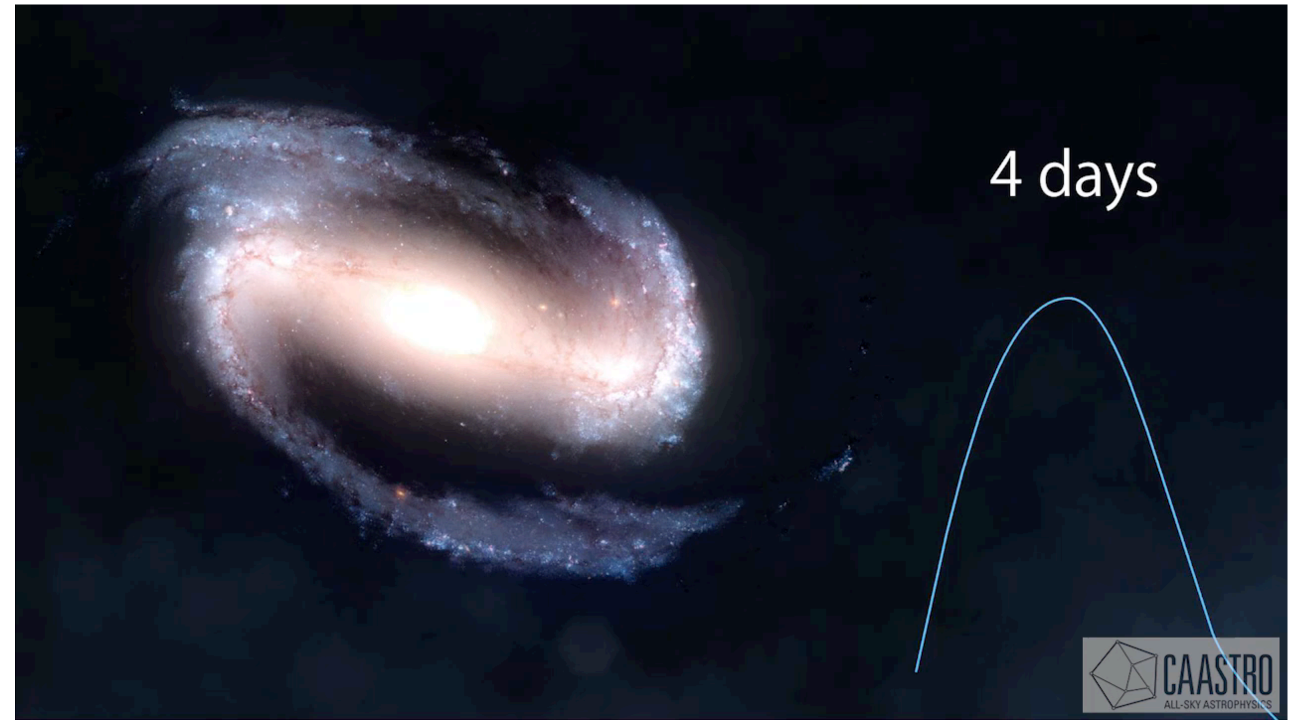
Bayesian Neural Networks and supernova classification

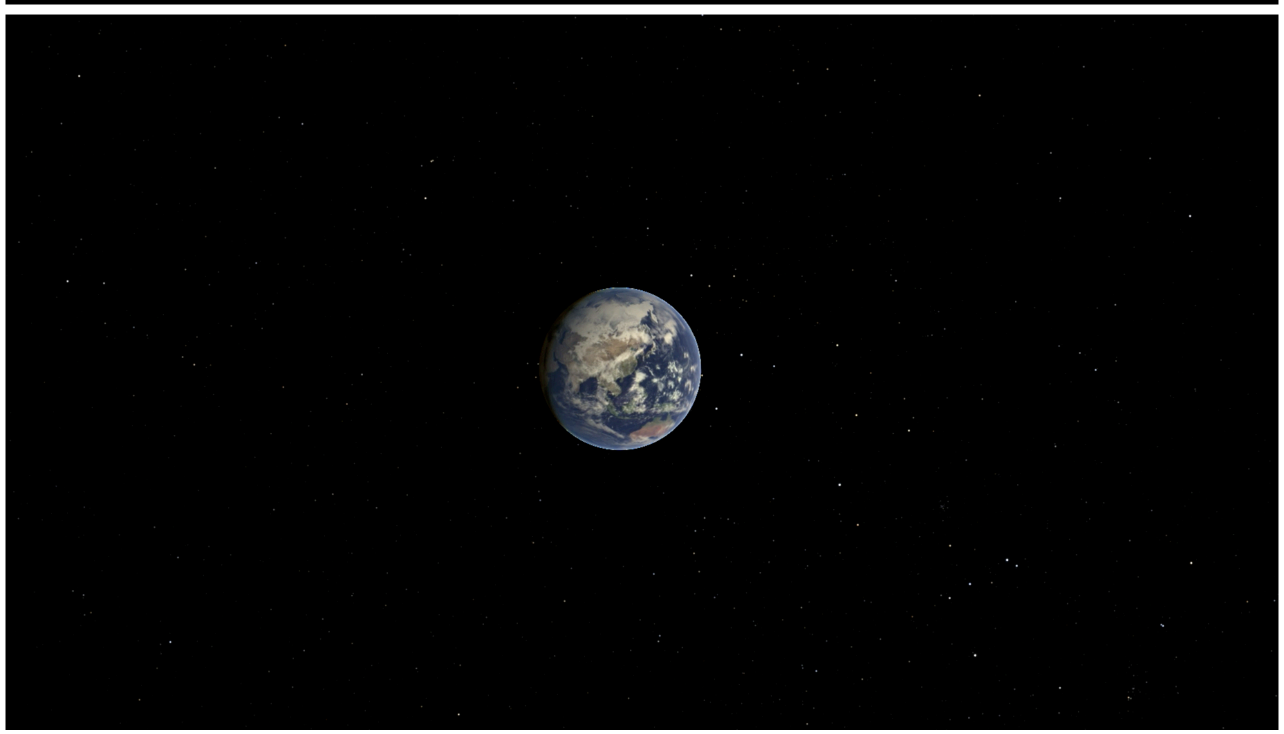
Anais Möller
ARC DECRA Fellow & Lecturer
co-PI FINK
SWINBURNE UNIVERSITY OF TECHNOLOGY
Australian Government
Australian Research Council

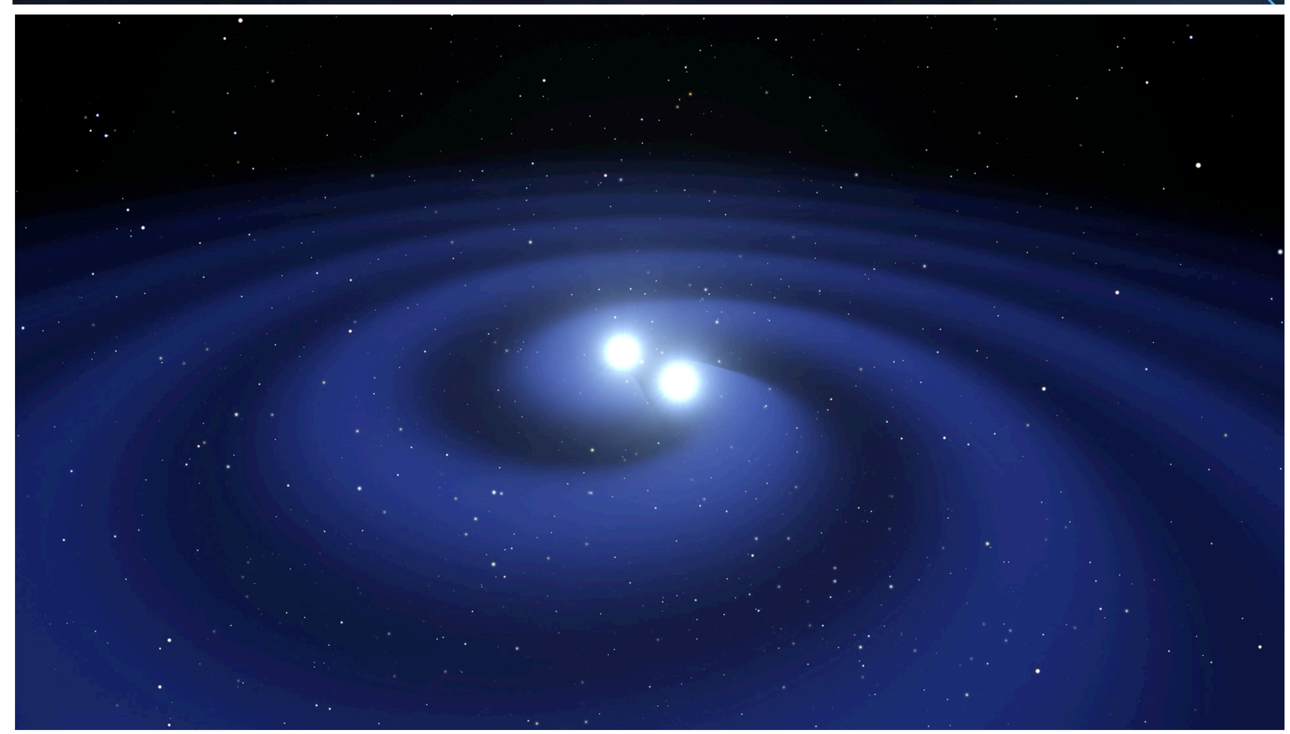
May 19, 2025
Bayesian Deep Learning for Cosmology and Time Domain Astrophysics

Time-domain

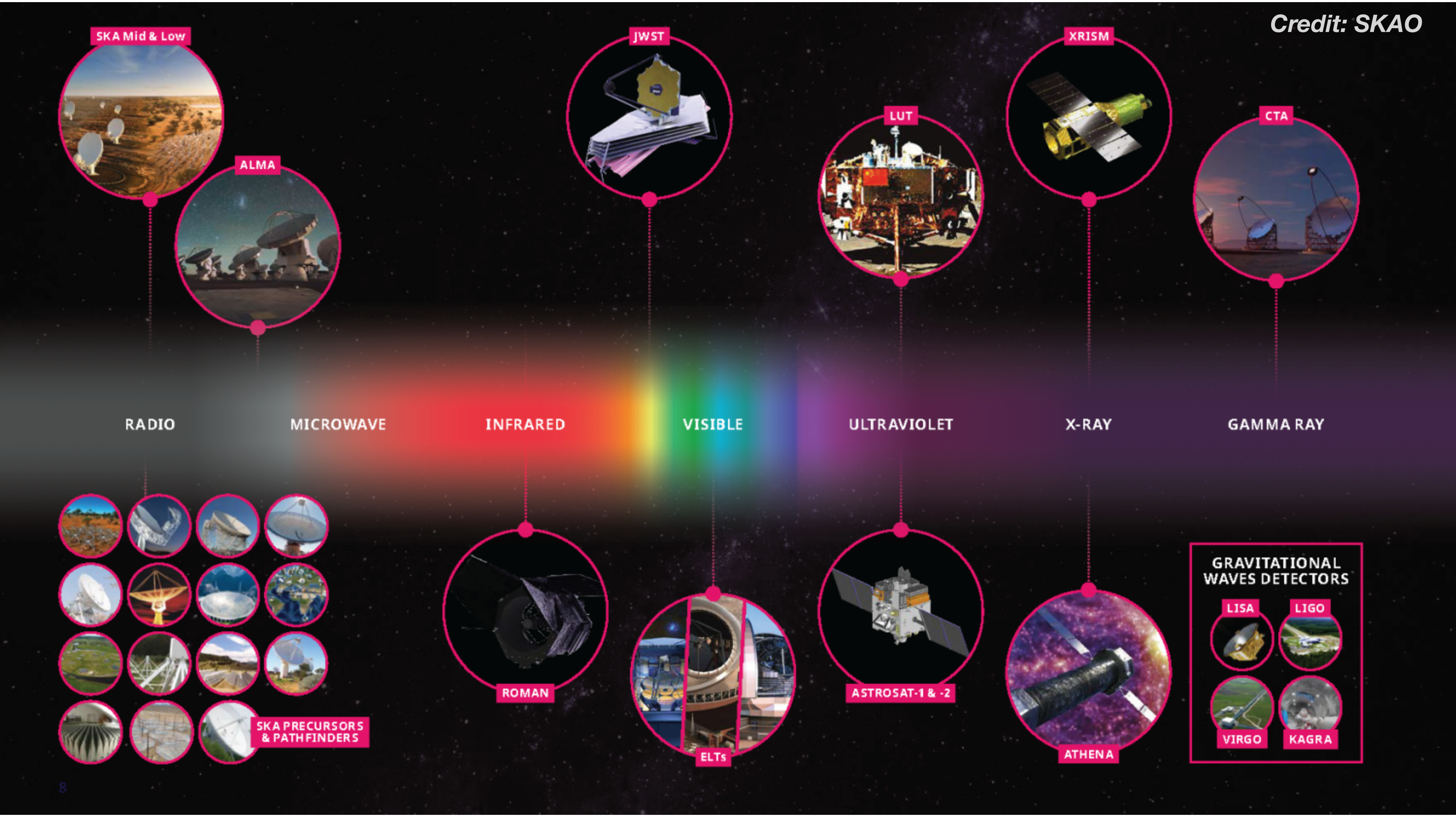




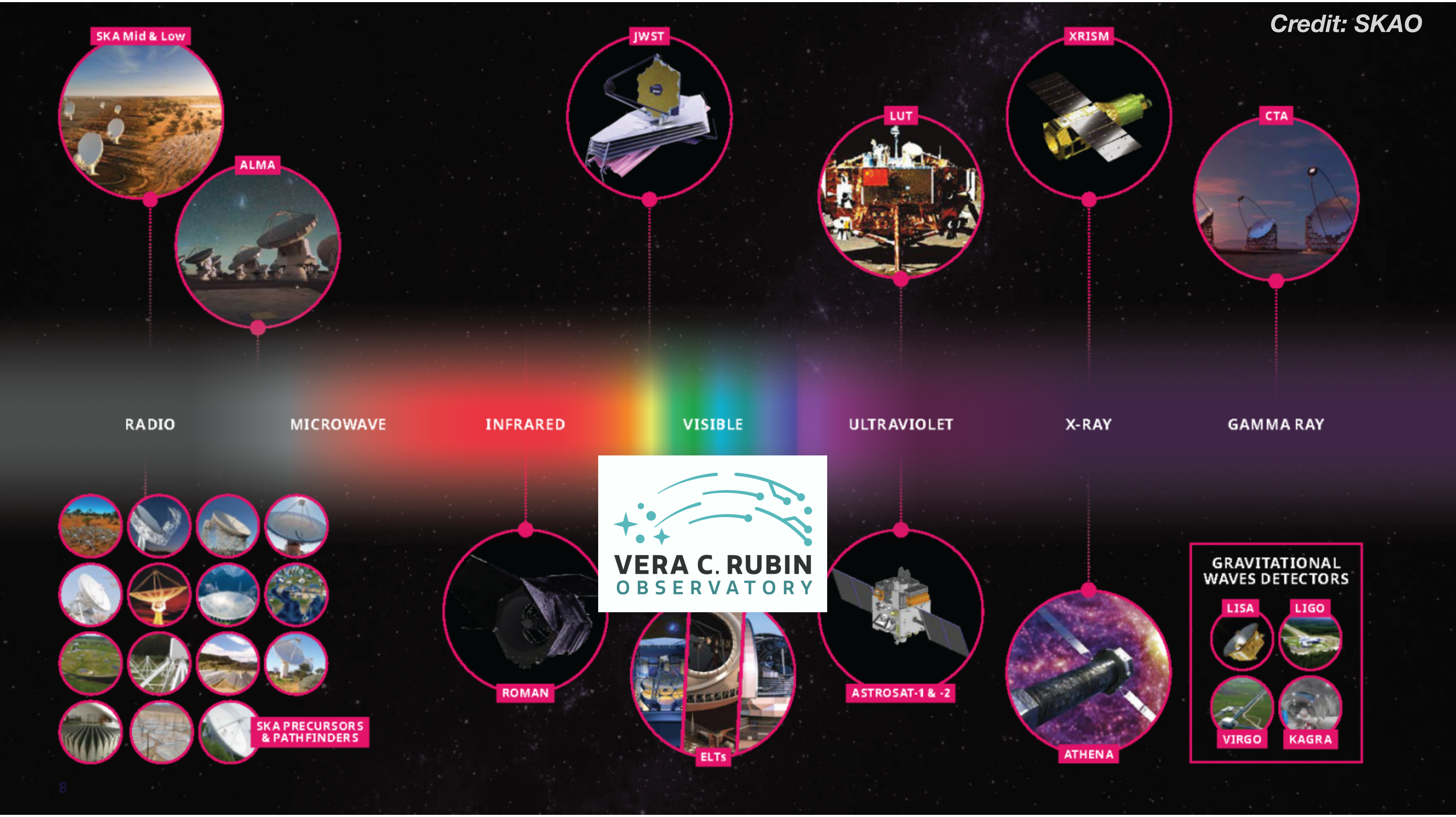




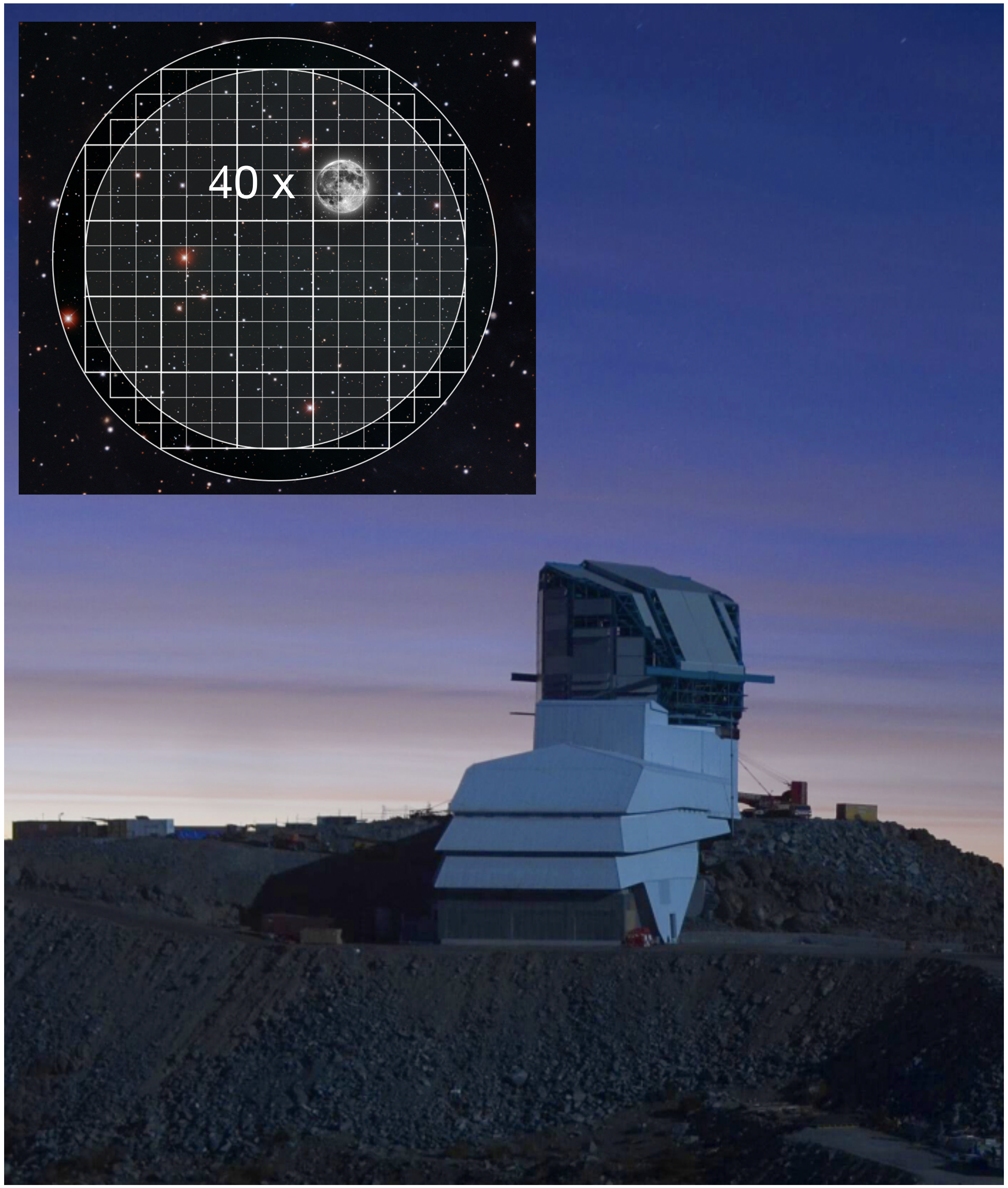
Generational leap



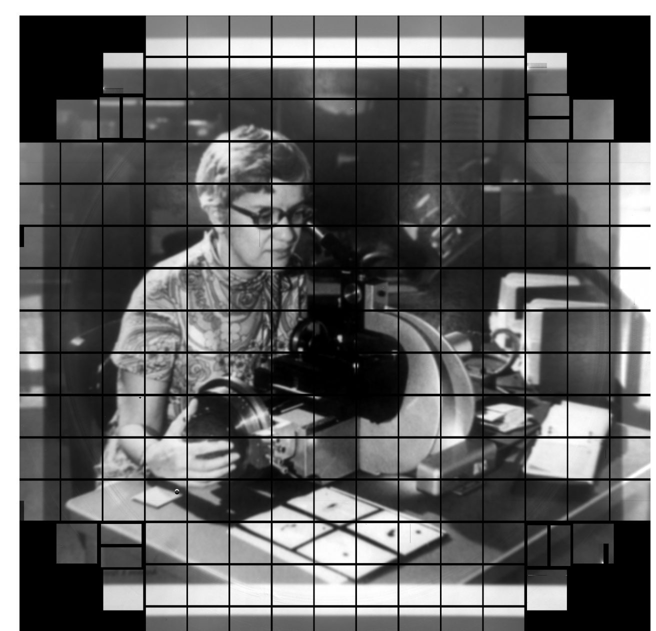
Generational leap



Vera C. Rubin Observatory

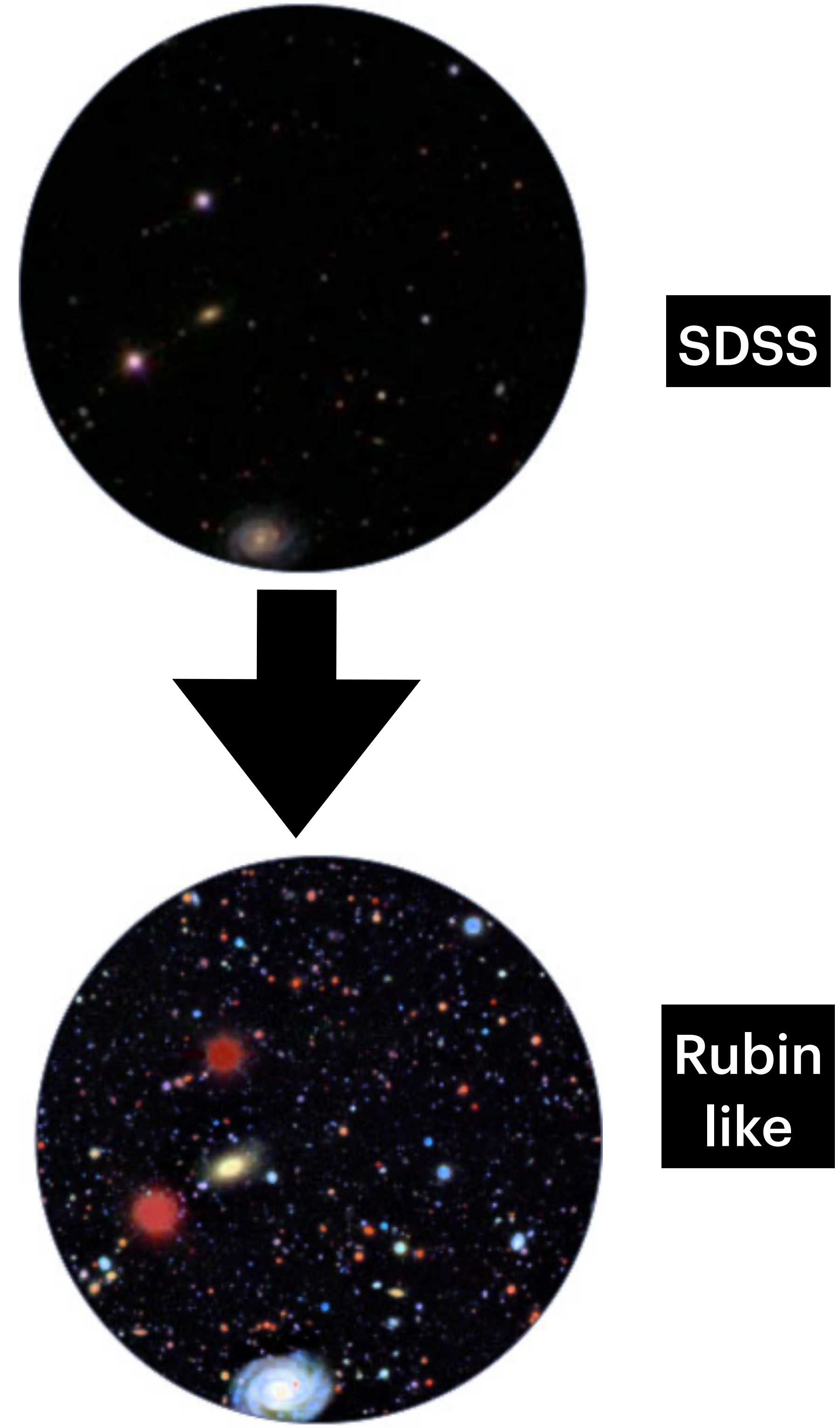


- Cerro Pachón, Chile
- Telescope: 8.4m diameter mirror
- World's largest CCD camera
  - 3.2 Gpixel camera
  - 0.2"/pixel
- Filters *ugrizy* (320-1050nm)
- Field of View: 9.6 deg^2



Filter	Single exposure	10-year
u	**23.9**	26.1
g	**24.5**	27.4
r	**24.7**	27.5
i	**24.0**	26.8
z	**23.3**	26.1
y	**22.1**	24.9

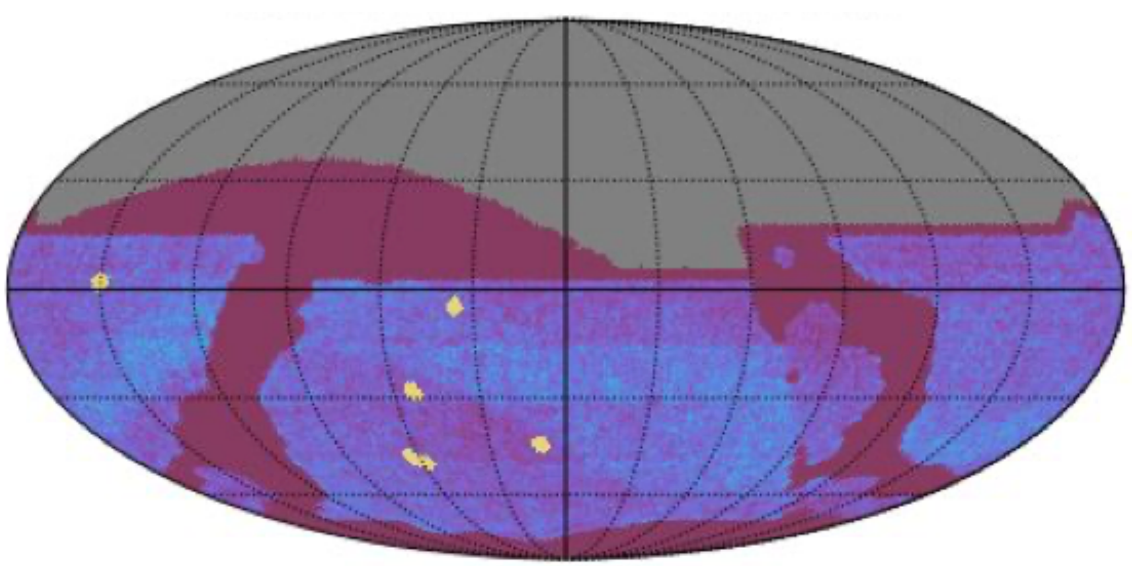
5-sigma point source depth



LSST Corporation

LSST

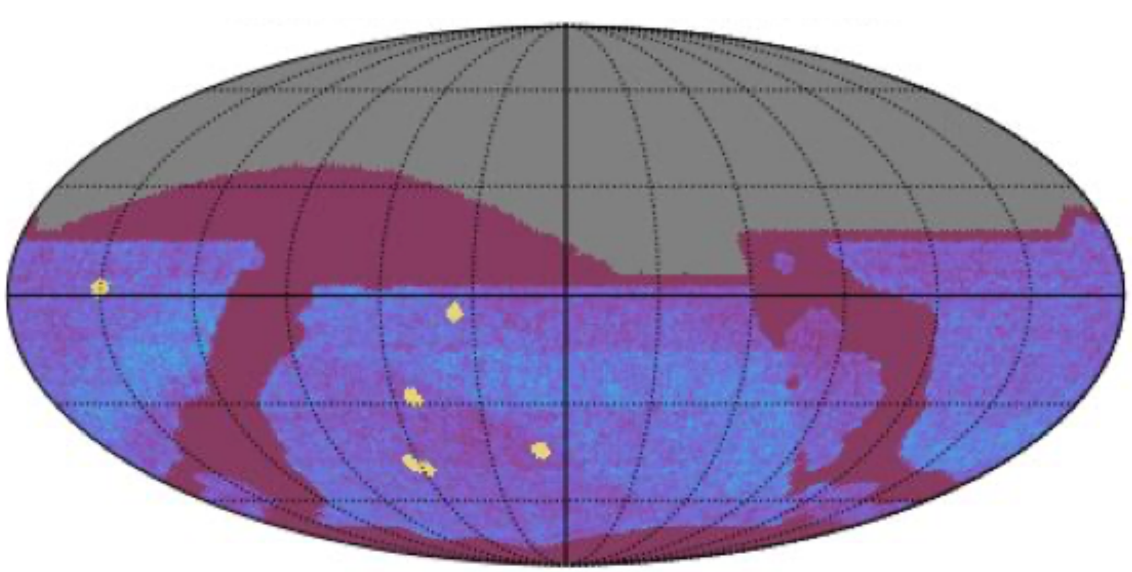
- 10 years
- WFD + 5 Deep drilling fields
- $m_r \approx 24.7$ mag



A. Möller | amoller@swin.edu.au

LSST

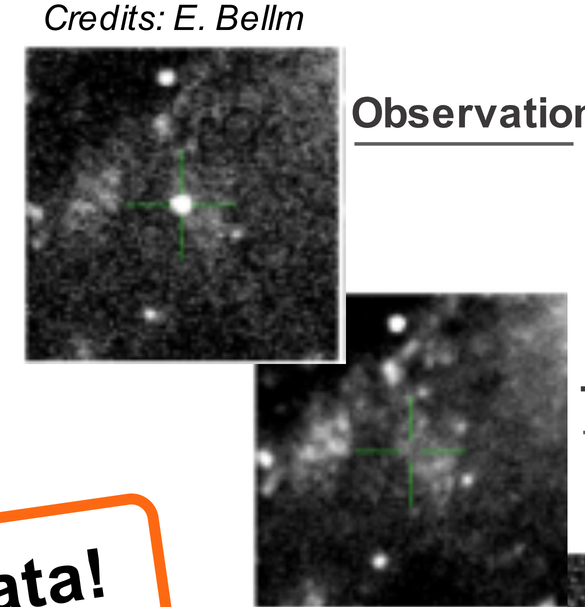
- 10 years
- WFD + 5 Deep drilling fields
- $m_r \approx 24.7$ mag

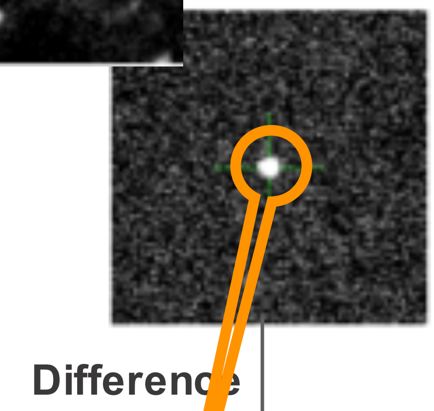


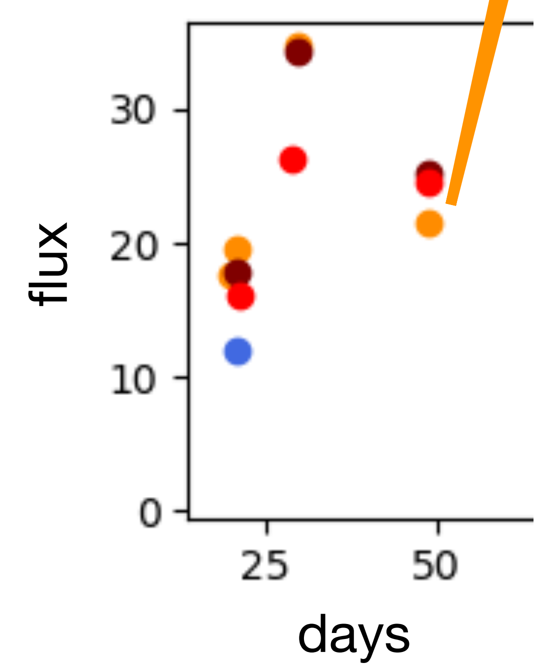
Public data!

Transients and variables

10 million/night







Ingested, enriched, filtered by

fink-portal.org

See Julien Peloton’s talk Tuesday!

VERA C. RUBIN OBSERVATORY

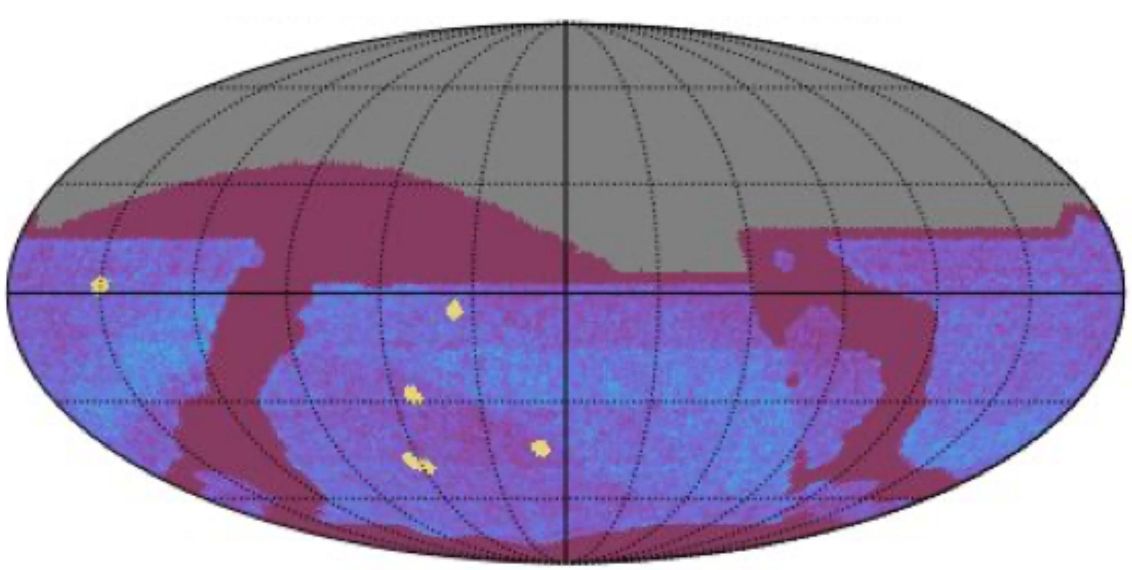
DARK ENERGY SURVEY

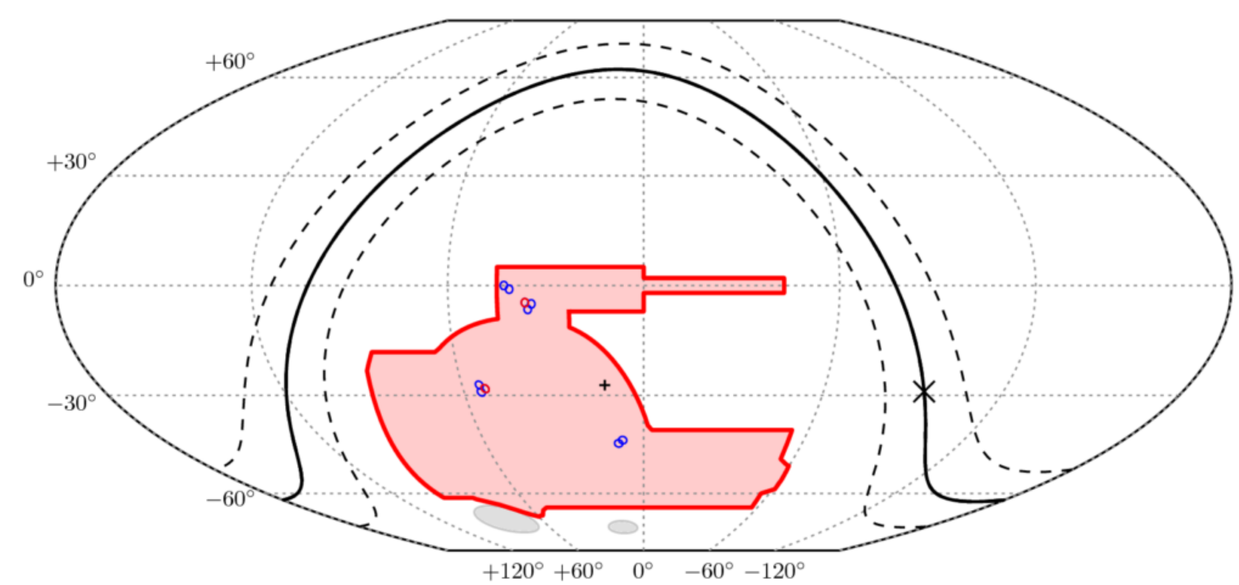
LSST

- 10 years
- WFD + 5 Deep drilling fields
- $m_r \approx 24.7$ mag

Supernova survey

- 5 years
- 10 Deep drilling fields
    - $m \approx 23.5\text{-}24.5$ mag





LSST

>1 million SN Ia

billions of detected objects

30.000 live transients (TiDES)

Frohmaier et al. 2025

Supernova survey

>2.000 high-quality SN Ia

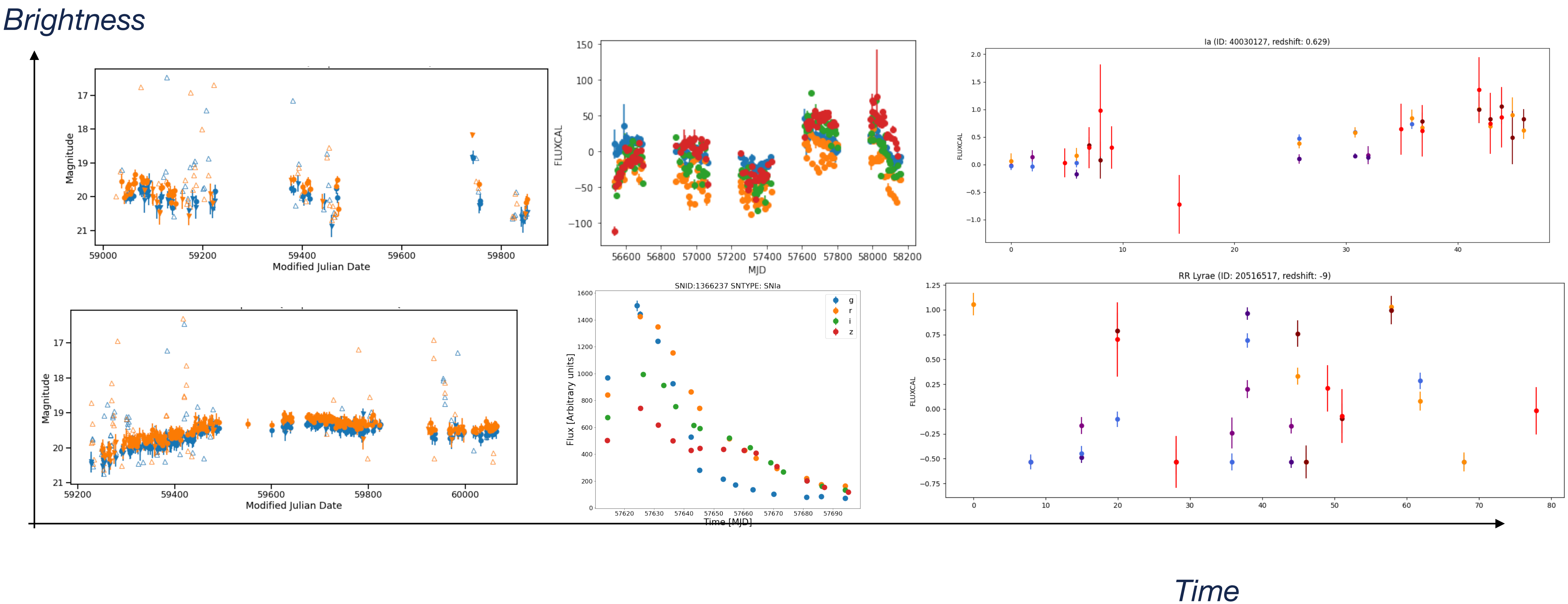
30.000 SN candidates

425 spectroscopically SN Ia

Smith, D'Andrea, Sullivan, AM et al. 2018

Photometric classification

Time-series in astrophysics



Light-curves

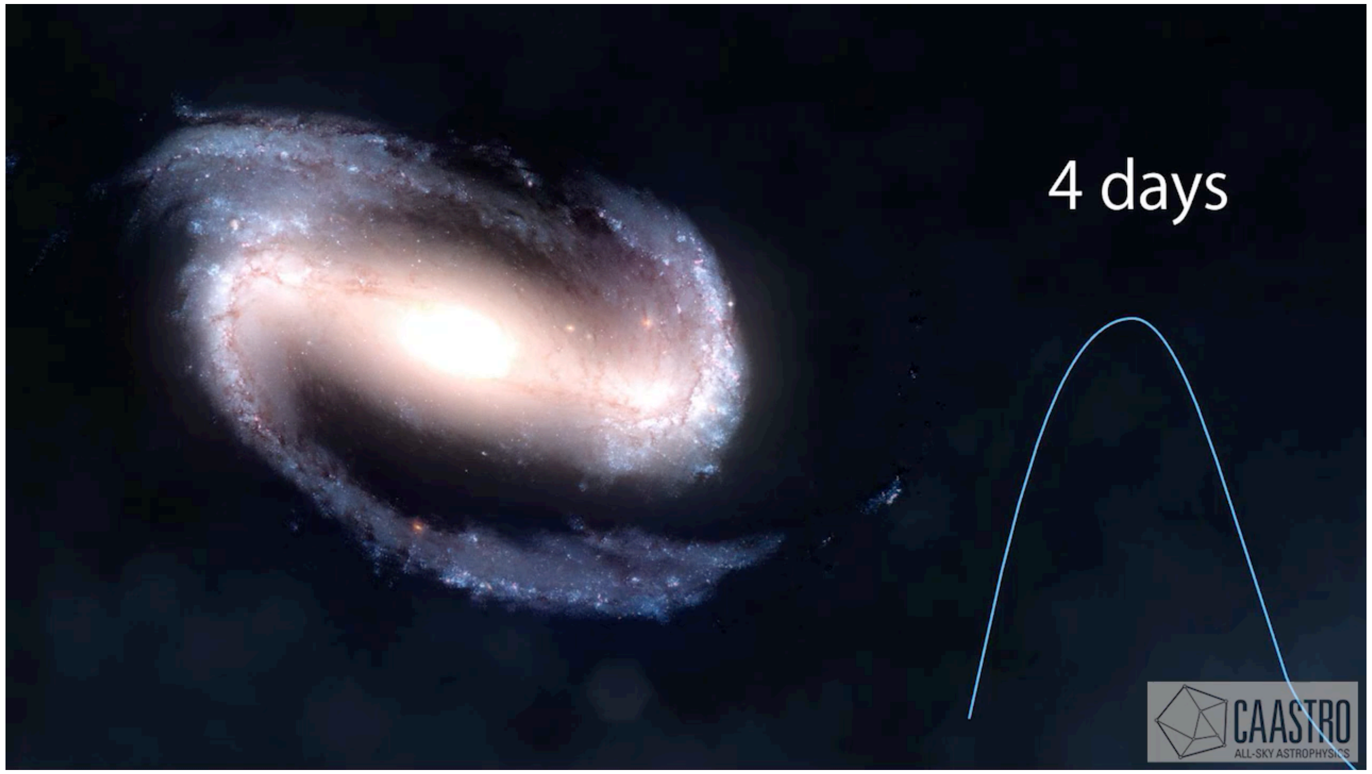
- Irregularly sampled
- Subject to weather conditions
- Partial information (e.g. photometry vs spectra)
- Only a small subset of events have been properly characterised…

Supernovae

- used for cosmology for precision statistical analyses

- extragalactic transient

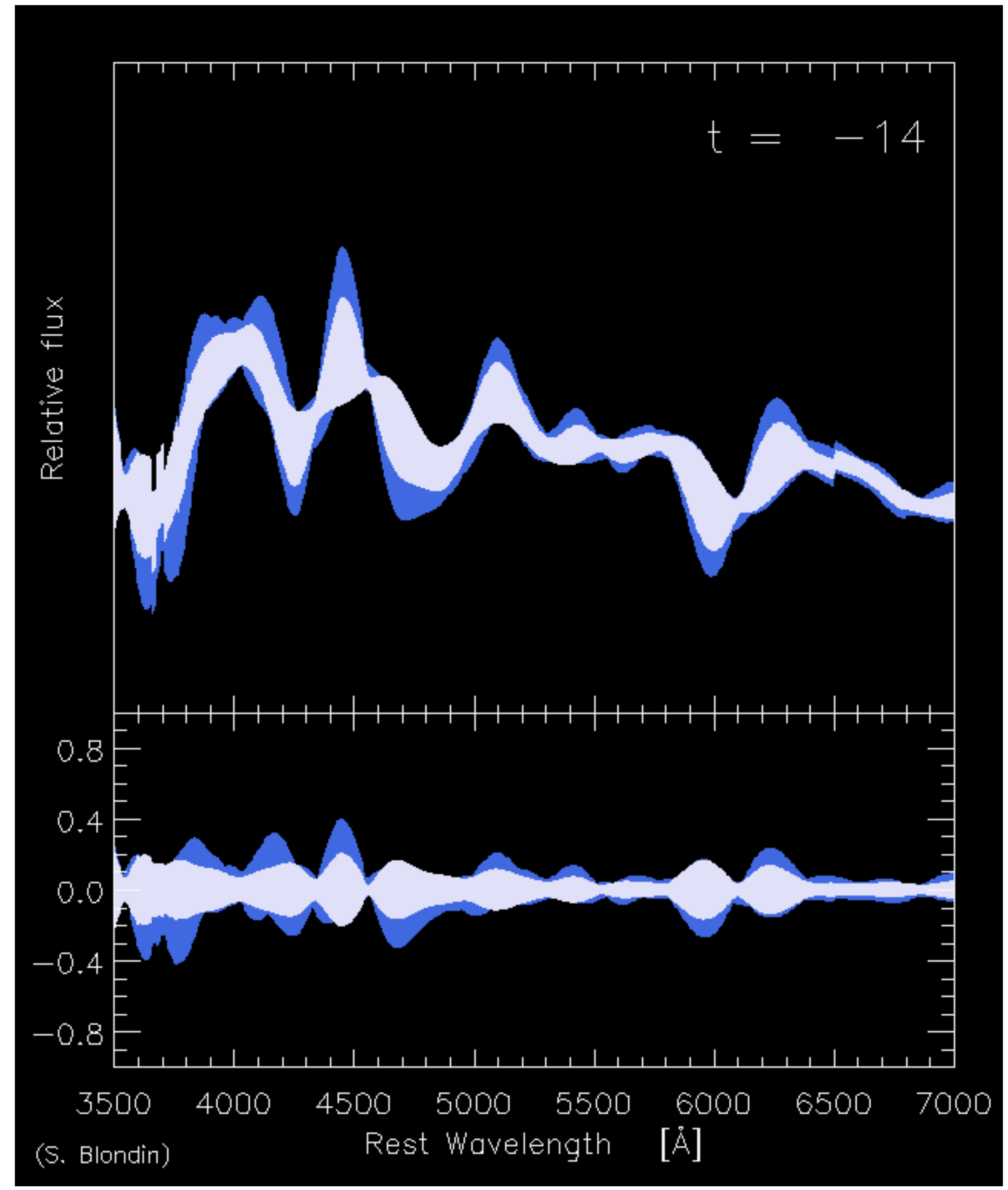
- different types: Ia, core-collapse, peculiar…



“Traditional classification”

Classification challenges:

- Spectroscopic resources are scarce
- Spectra time constraints
- Selection of biased samples

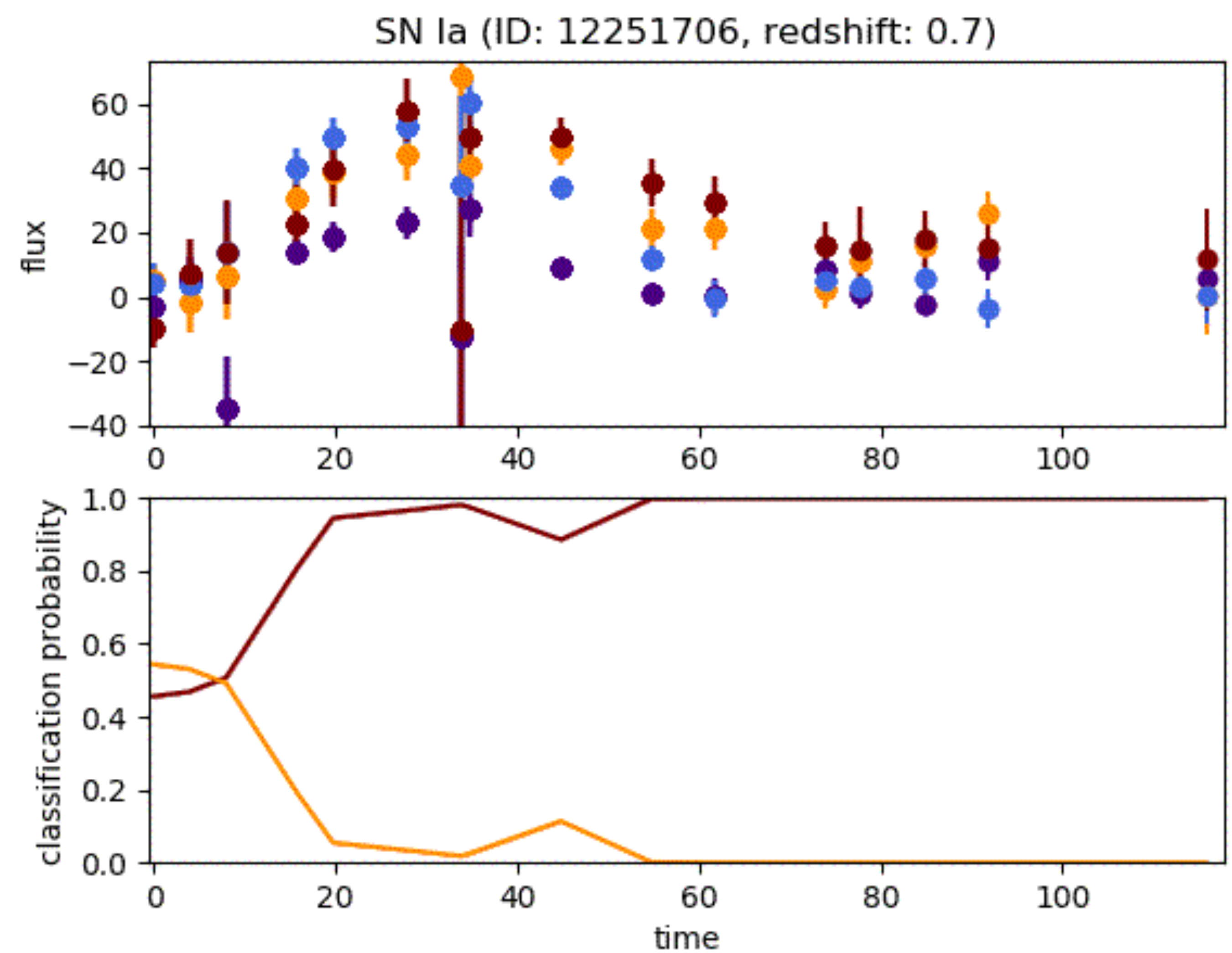


SuperNNova
open source photometric classification

Möller et al. 2020

GitHub: *supernnova/SuperNNova*

- Recurrent Neural Networks (RNNs):
    - LSTM
    - GRU
- Bayesian RNNs
    - MC dropout (Gal+2016)
    - Bayes by Backprop (Fortunato+2017)
    - SWA-G (Maddox+2019)
- Convolutional NN

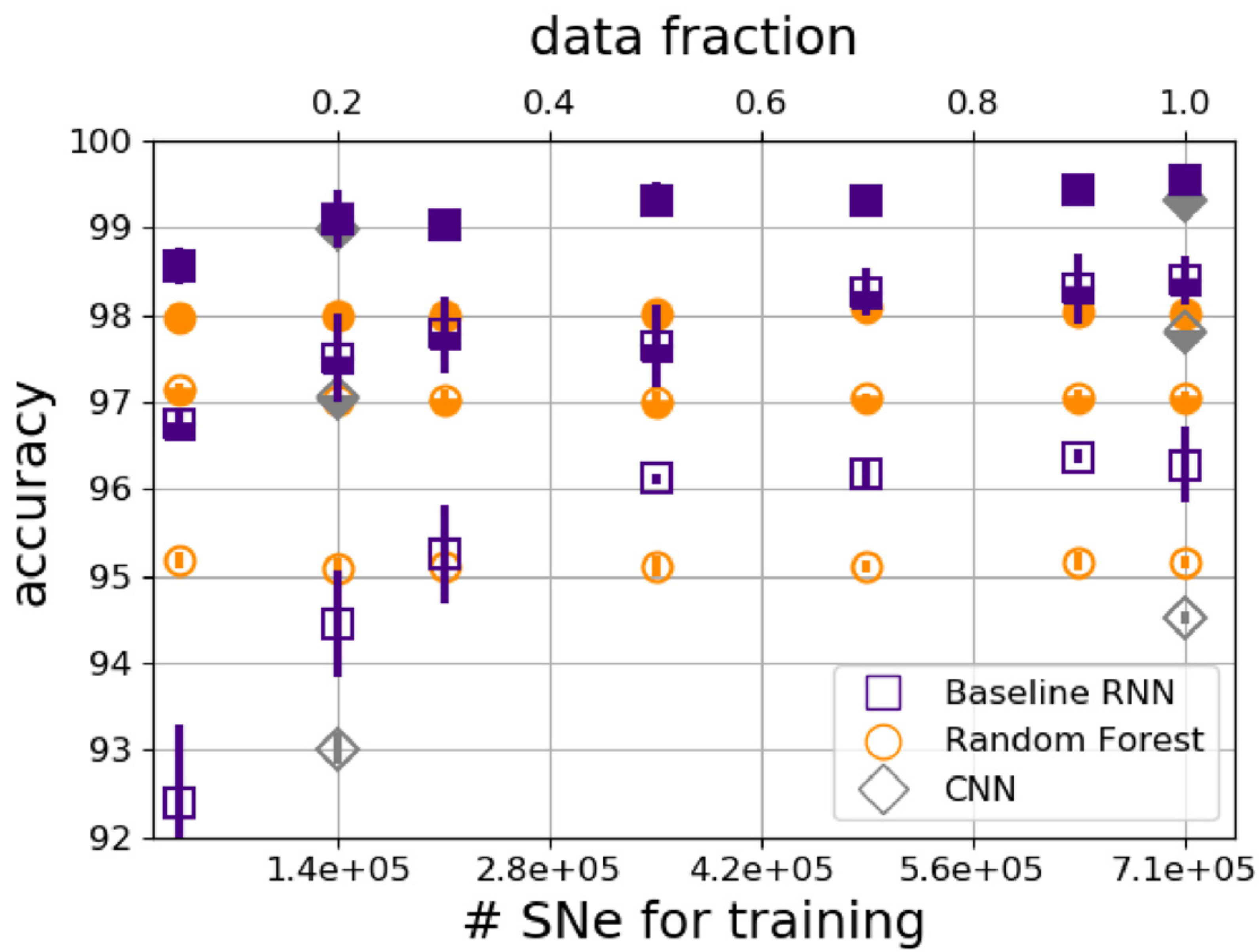


Performance depends on training

High accuracy >98%

- with or without redshift information
- Complete or partial light-curves

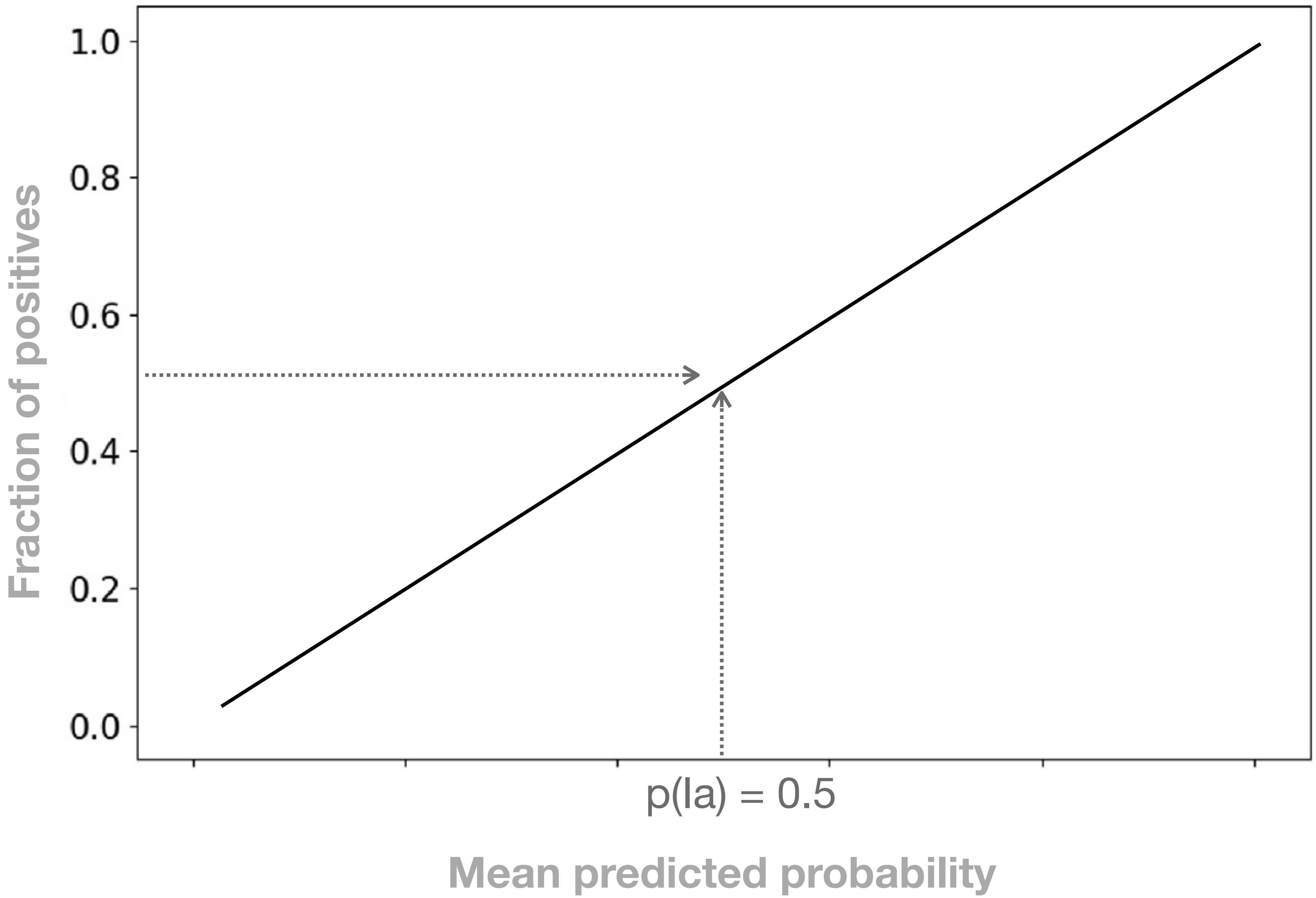
- Needs large simulations for training



NNs robust?

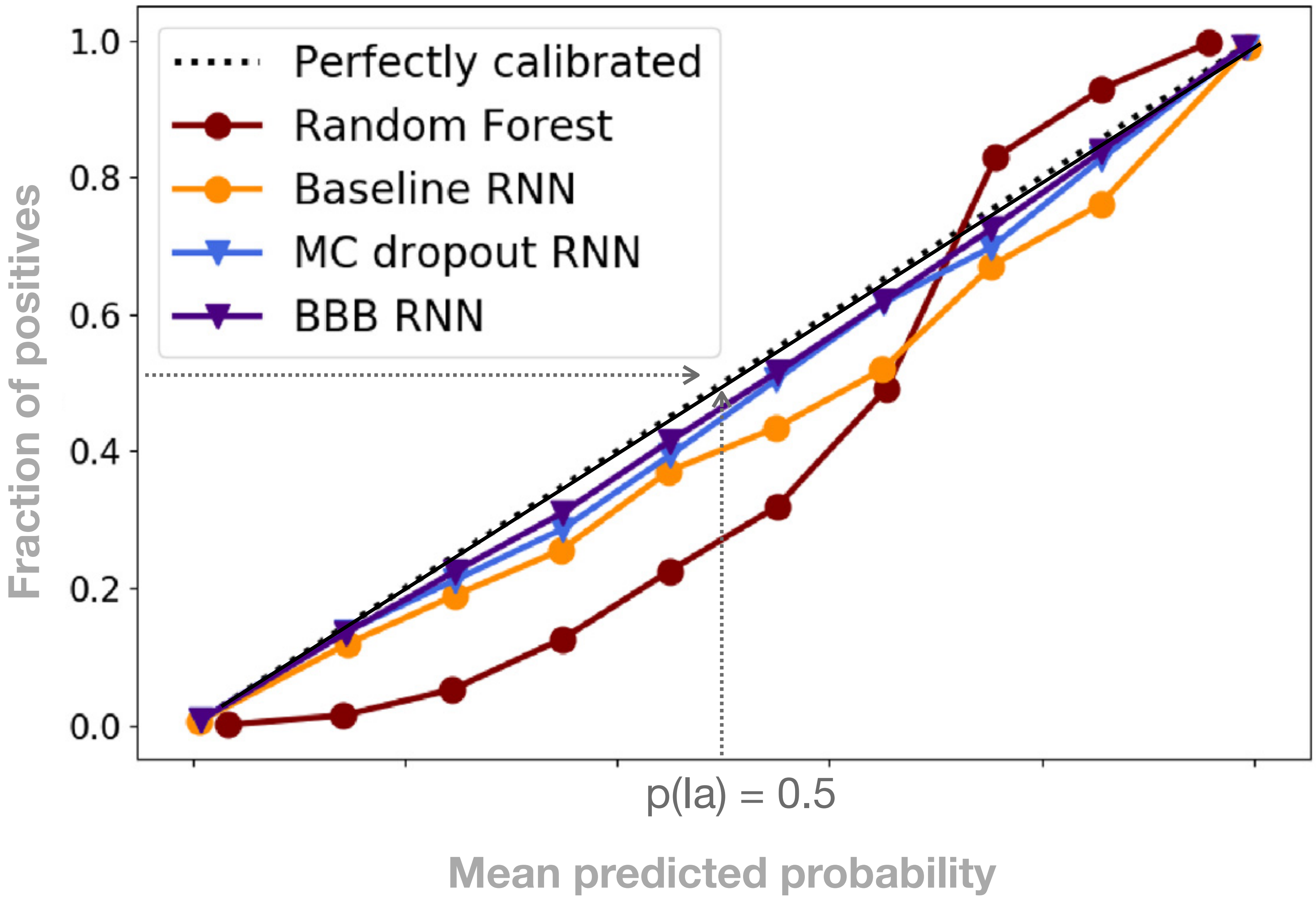
Calibration = Reliability diagrams

De Groot+ 1983, Niculezcu-Mizil+ 2005, Guo+ 2017

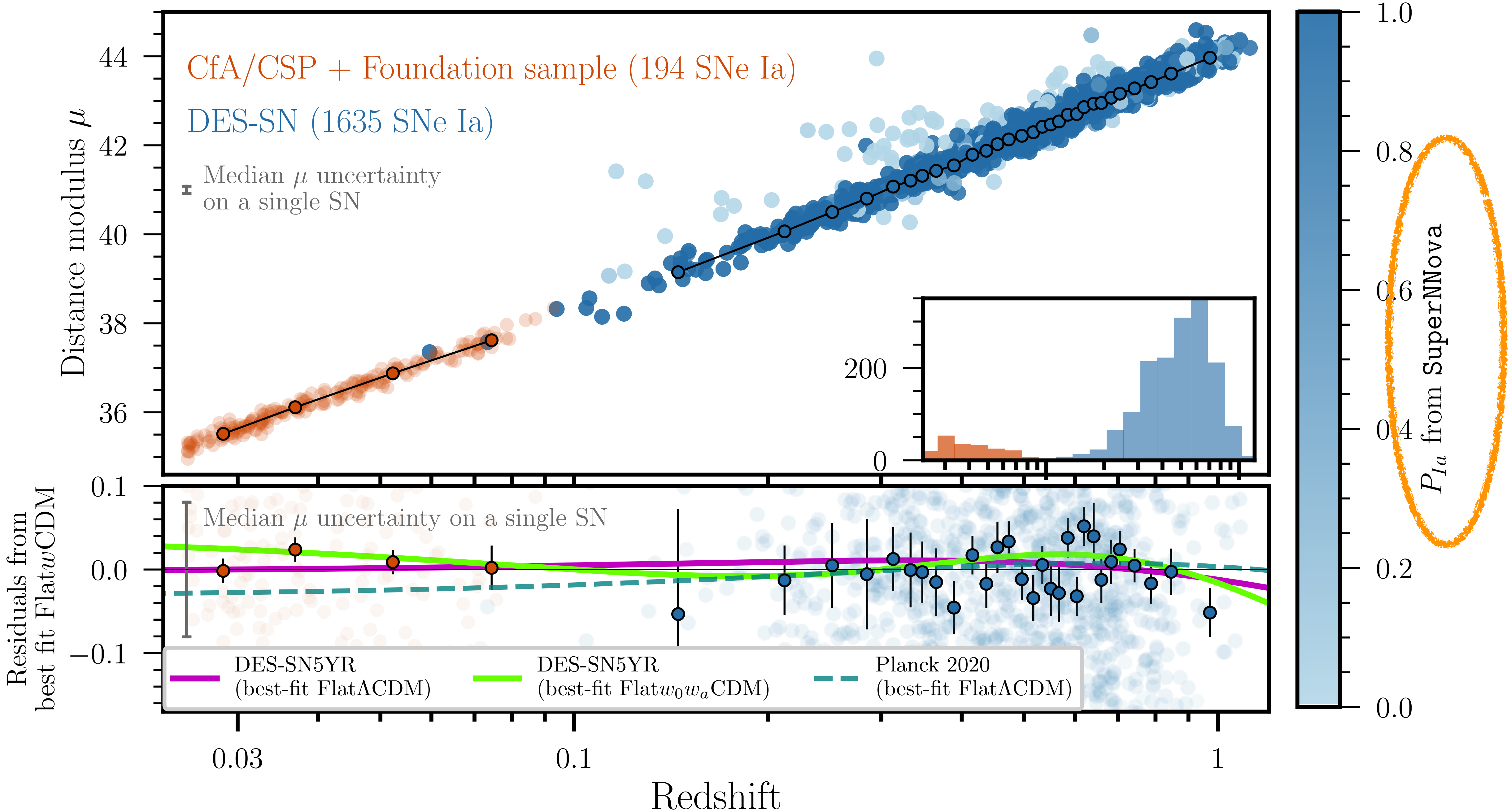


Calibration = Reliability diagrams

De Groot+ 1983, Niculezcu-Mizil+ 2005, Guo+ 2017



DES 5-year SNIa



Largest high-z SN Ia sample from a single survey for cosmology

NNs interpretable?

Uncertainties

Lack of knowledge about the truth…

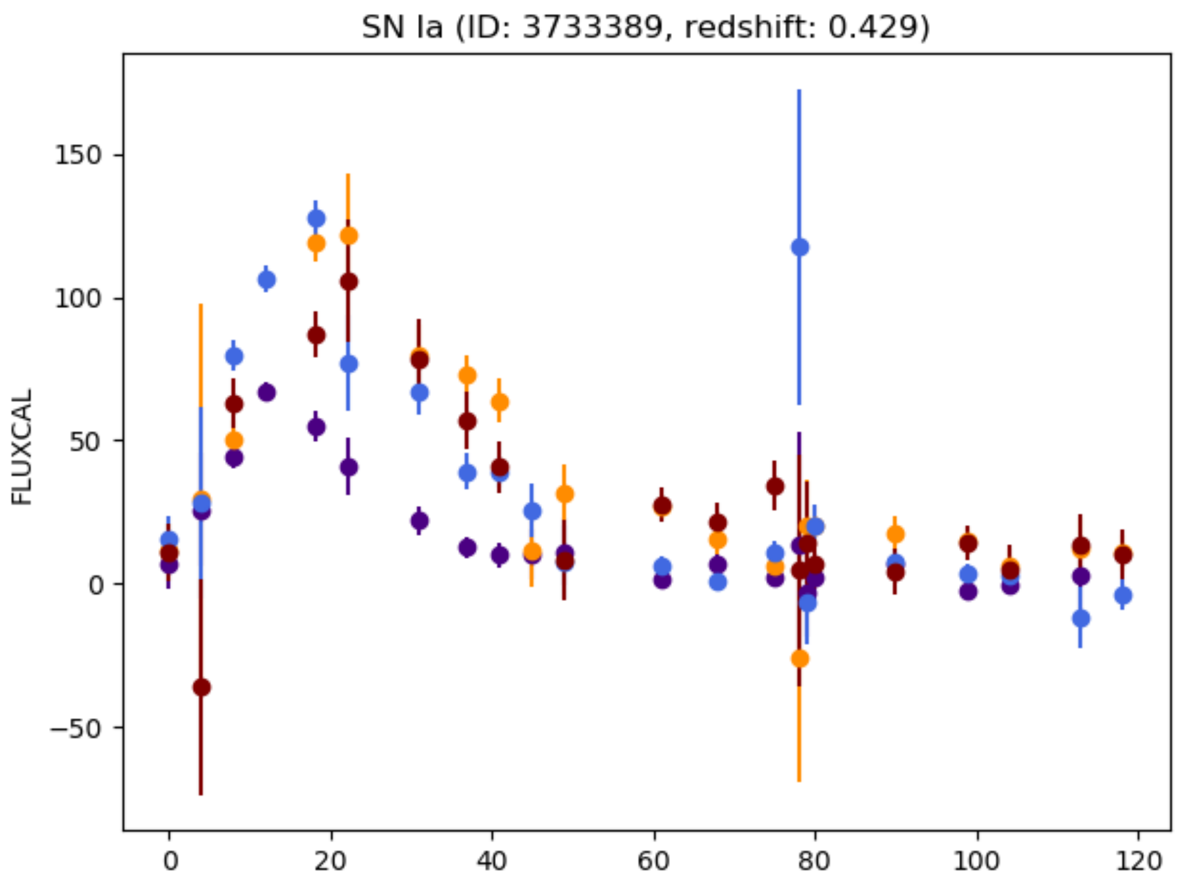
Uncertainties

- Aleatoric

"due to the random nature of the way the observed objects are created and the way we make observations."

"cannot be reduced through greater understanding"

uncertainties in the input data, e.g. noise or other effects of data acquisition



Uncertainties

- Epistemic

“lack of knowledge about ***underlying model of the data*** as well as the ***form of the neural network***, the ***way it is trained***, the choice of ***cost function*** used to characterise how well the network performs, etc”

ignorance about the model that generated the classification

Uncertainties

- non-representative training sets
- finding/rejecting unknowns (anomalies, OOD)
- Rigorous ML: bias-free, meaningful uncertainties

- Epistemic

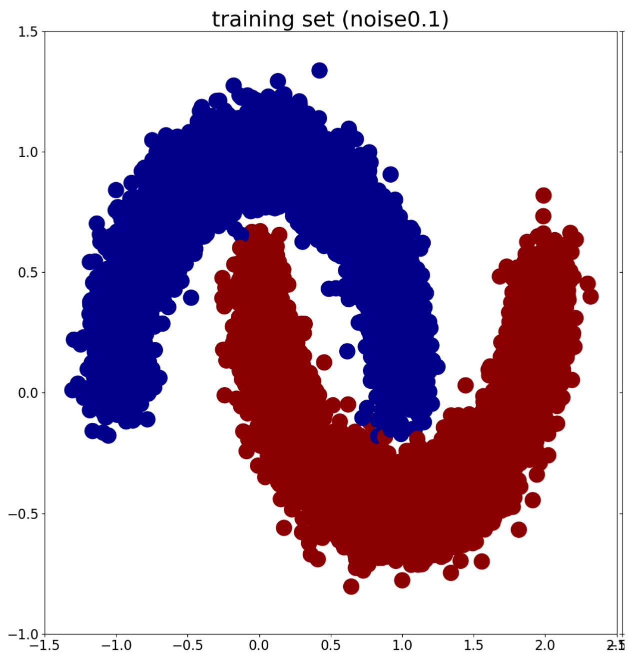
"lack of knowledge about ***underlying model of the data*** as well as the ***form of the neural network***, the ***way it is trained***, the choice of ***cost function*** used to characterise how well the network performs, etc"

ignorance about the model that generated the classification

Meaningful uncertainties

training

classification probability

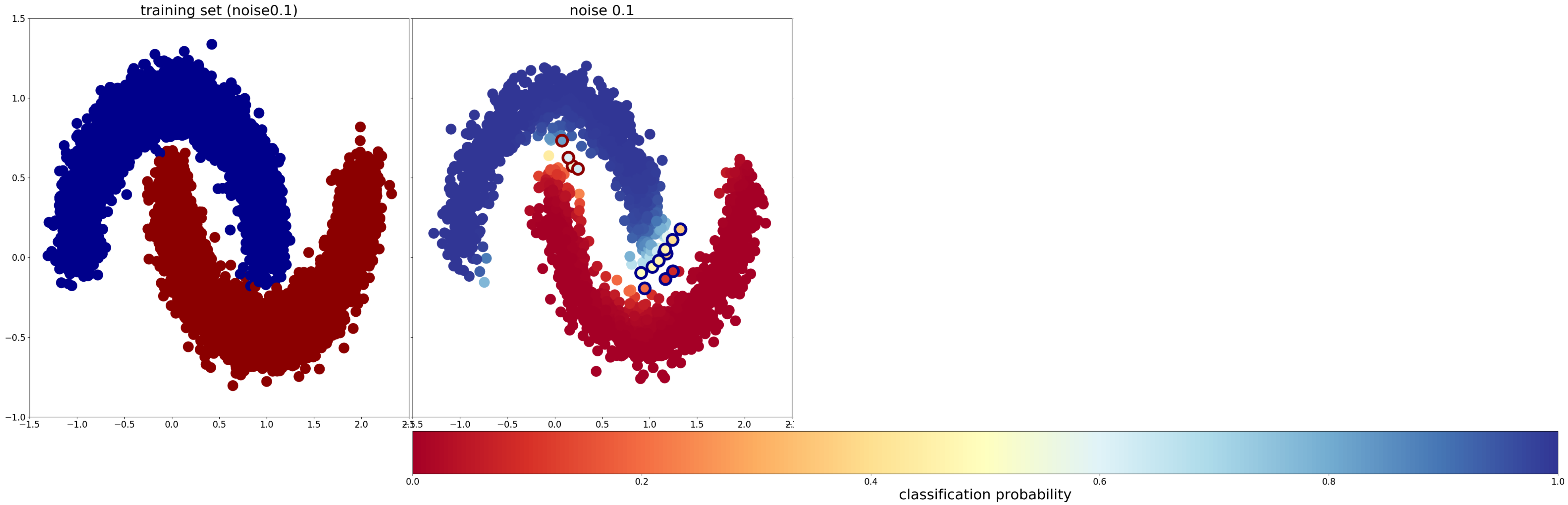


A. Möller

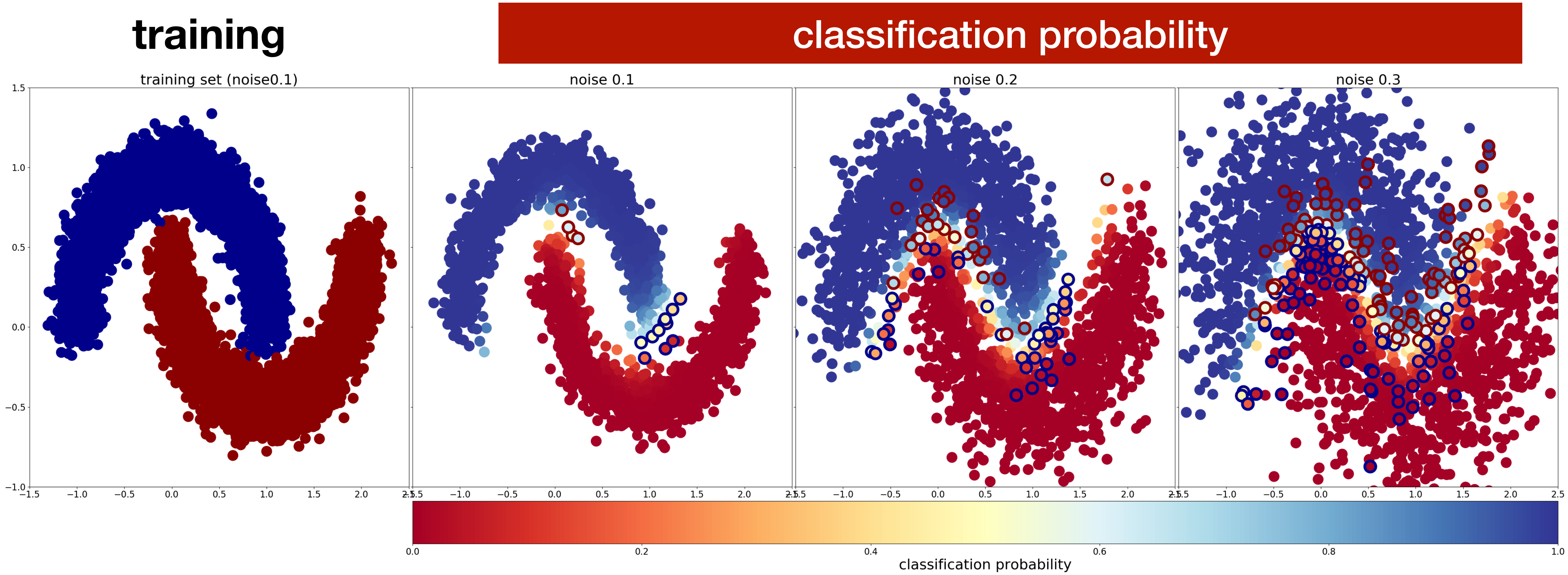
Meaningful uncertainties

training

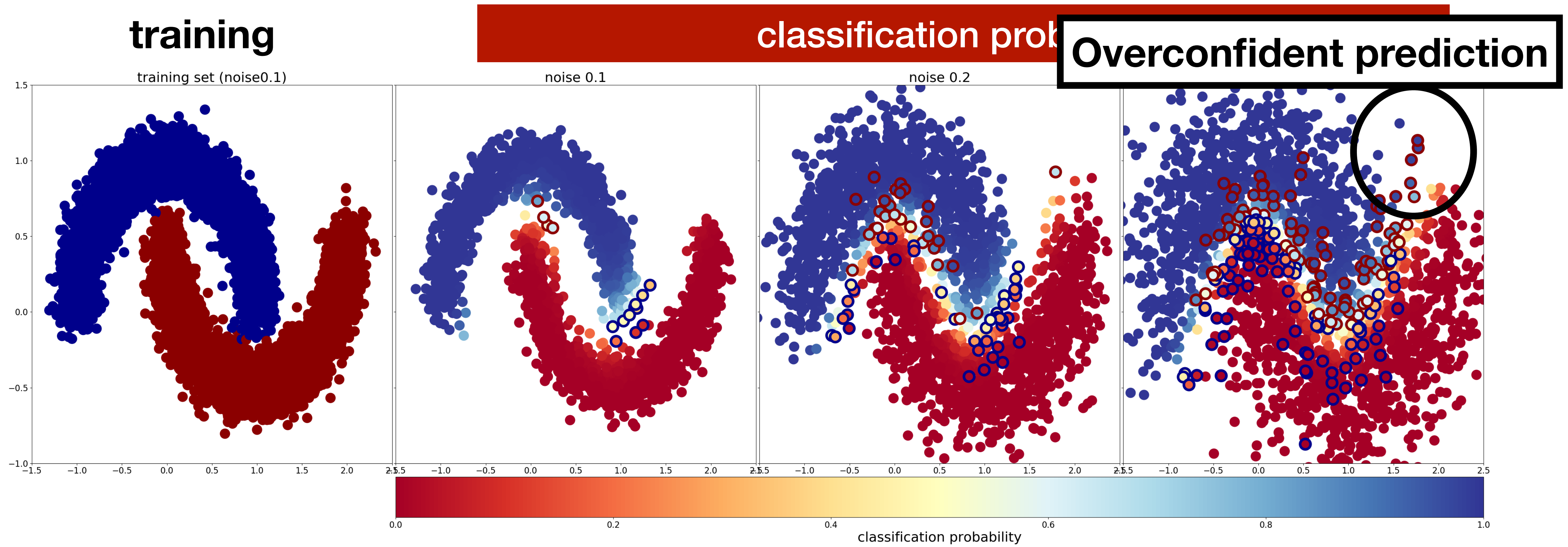
classification probability



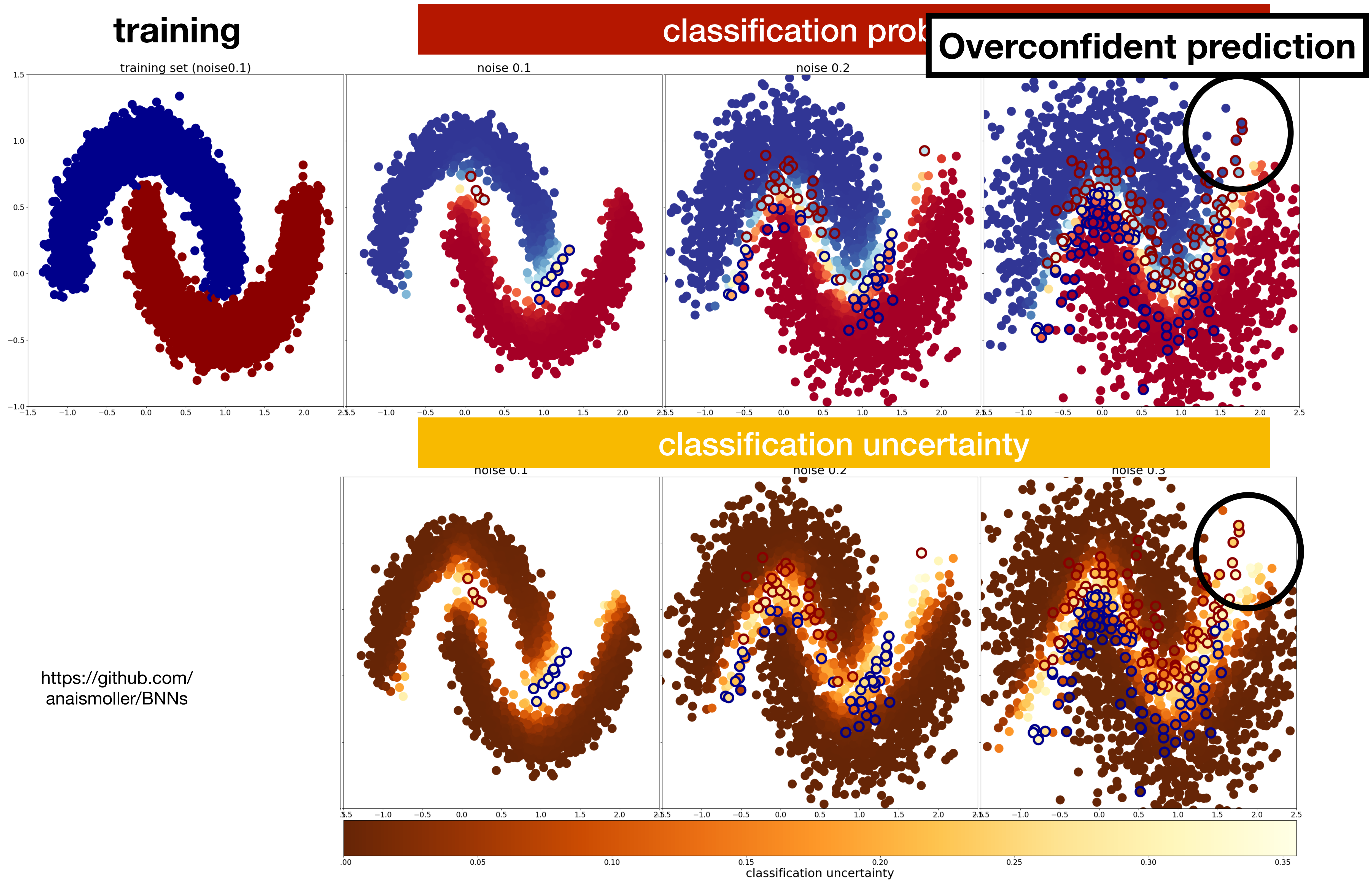
Meaningful uncertainties



Meaningful uncertainties



Meaningful uncertainties



https://github.com/
anaismoller/BNNs

Physicists:

Epistemic uncertainties:
express our ignorance about the model that generated the data.

Bayesian deep learning

$$\mathscr{P}\left(\mathbf{y}\,|\,\mathbf{x},\,\mathbf{w}\right)$$

classification data weights

$$NLL = \min_{\mathbf{w}} \sum_{i=1}^{N} -\log \mathscr{P}(\mathbf{y}_i\,|\,\mathbf{x}_i,\,\mathbf{w})$$

$$\mathscr{P}(\mathbf{y}\,|\,\mathbf{x}) = \int \mathscr{P}(\mathbf{y}\,|\,\mathbf{x}, \mathbf{w})\mathscr{P}\left(\mathbf{w}\,|\,\mathscr{D}\right) d\mathbf{w}$$

Distribution of weights

$$\mathscr{P}(\mathbf{y}\,|\,\mathbf{x}) = \int \mathscr{P}(\mathbf{y}\,|\,\mathbf{x}, \mathbf{w})\mathscr{P}\left(\mathbf{w}\,|\,\mathscr{D}\right) d\mathbf{w}$$

Posterior is intractable for deep NNs

$$\mathscr{P}(\mathbf{y}\,|\,\mathbf{x}) = \int \mathscr{P}(\mathbf{y}\,|\,\mathbf{x},\mathbf{w})\mathscr{P}\left(\mathbf{w}\,|\,\mathscr{D}\right) d\mathbf{w}$$

Posterior is intractable for deep NNs

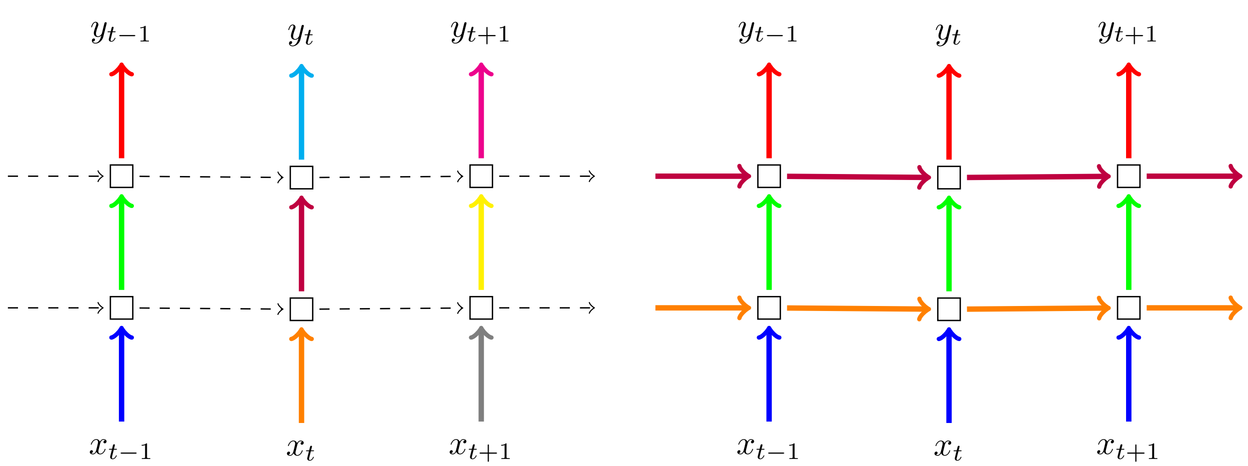
$$\mathscr{P}(\mathbf{w}\,|\,\mathscr{D}) \approx q(\mathbf{w}\,|\,\theta)$$ variational distribution

Training minimisation $$\hat{\theta} = \min_{\theta} \mathbf{KL}\left(q(\mathbf{w}\,|\,\theta)\,||\,\mathscr{P}(\mathbf{w}\,|\,\mathscr{D})\right)$$

BNNs Variational Inference

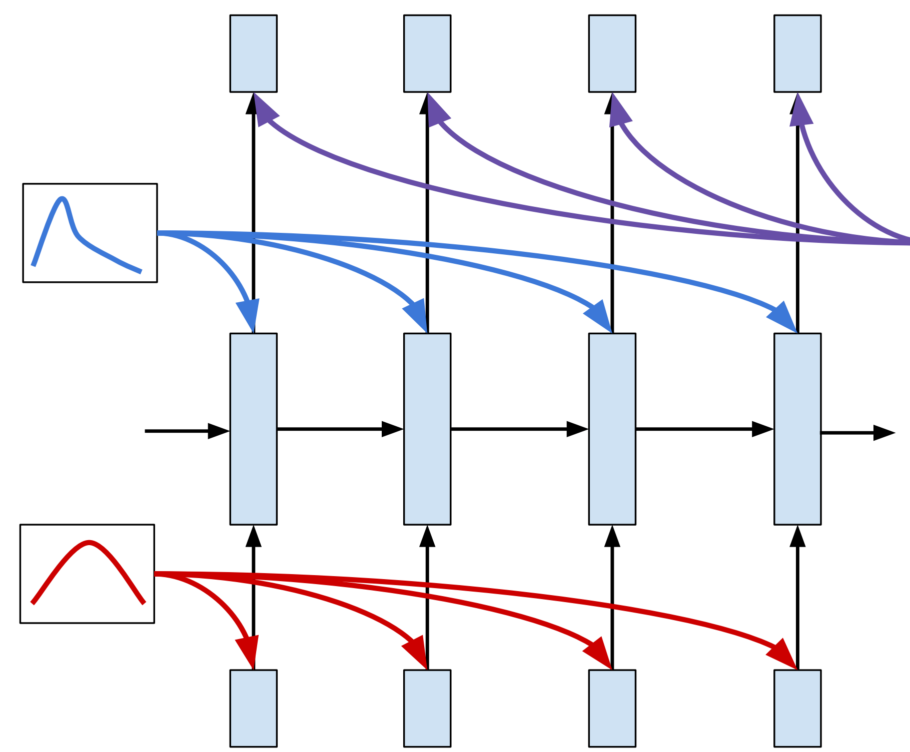
Approximating the variational distribution $q(\mathbf{w}|\theta)$

1. MC dropout
Gal & Ghahramani 2016



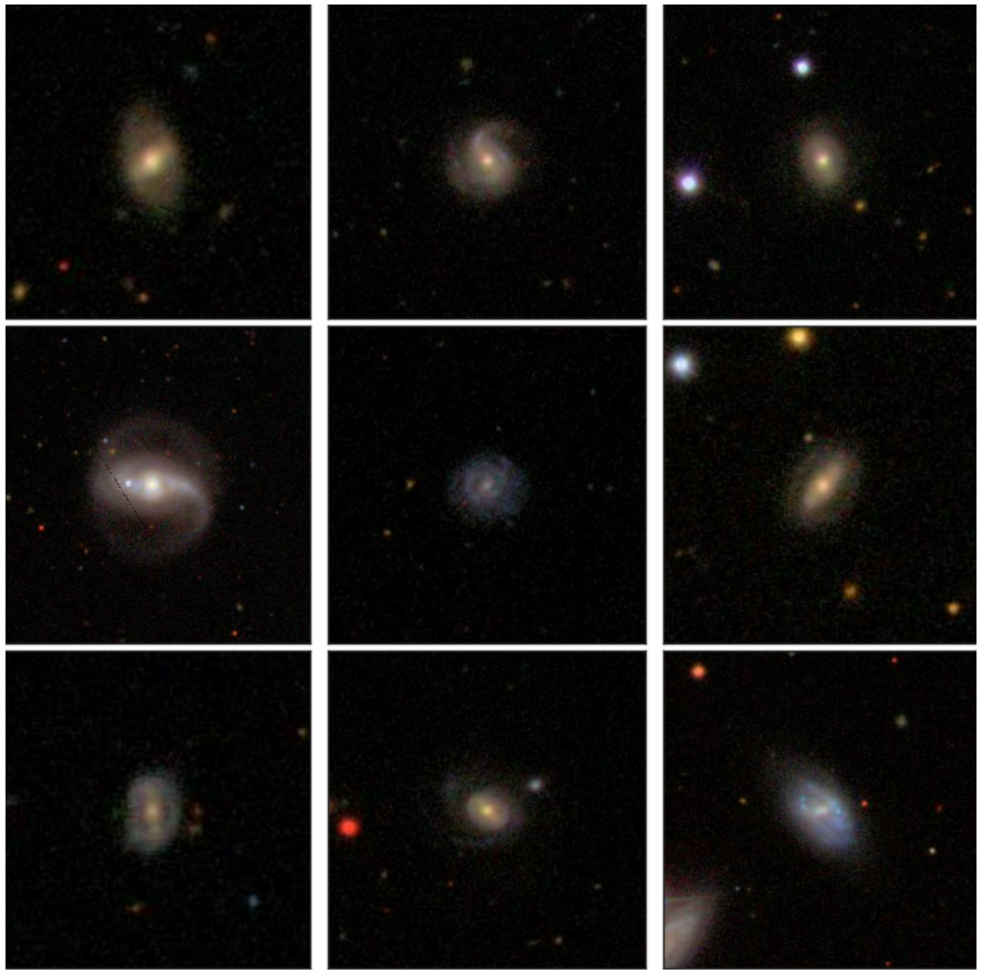
(a) Naive dropout RNN (b) Variational RNN

2. Bayes by Backprop
Fortunato+ 2017



Other methods in literature e.g. SWA-G (Maddox+2019), Flipout (Wen+2018)

Some examples in literature



Walmsley+ 2020

- **Bayes by Backprop** Fortunato+ 2017
  - SN *classification*: Möller+ 2020

- **SWA-G** Maddox+2019
  - Planetary *regression*: RNN Cranmer+2021

- **MC Dropout** Gal & Ghahramani 2016
  - Regression:
    - Strong lensing: Perreault Levasseur+ 2017
    - 21cm regression: Hortúa+2020
    - Stellar ages: Weaver+2024
    - CC SN GW: Nunes+2024
    - AGN properties; Tien+2025
    - Galaxies properties: Ginolfi+2024
  - *Classification*
    - Galaxy: AL+CNN Walmsley+ 2020
    - SN Möller+ 2020

BNNs for classification

DARK ENERGY SURVEY

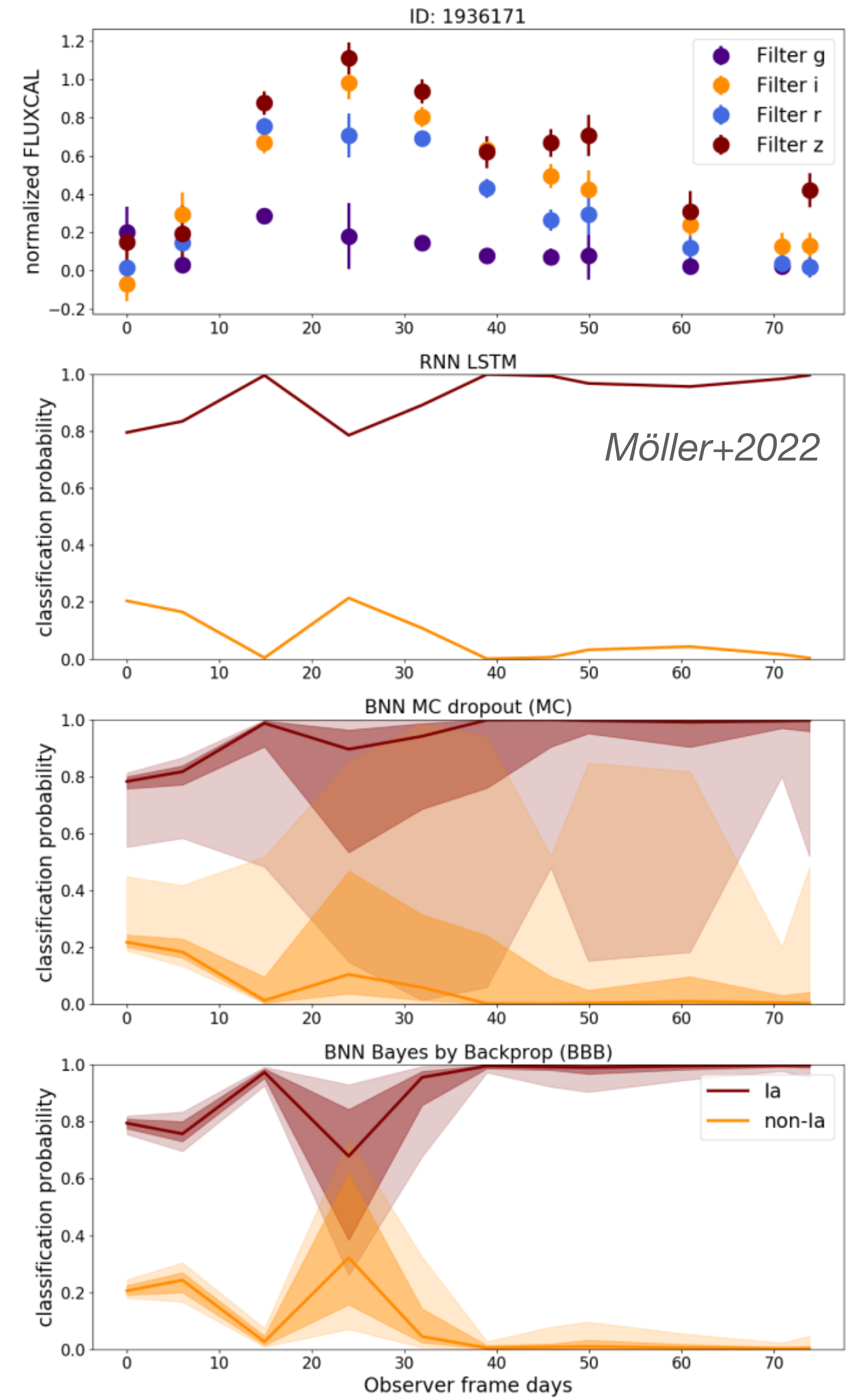
LSTM

MC dropout

Gal+ 2016

Bayes by Backprop

Fortunato+ 2017



“meaningful classification uncertainties”

- Out-of-distribution events (anomalies)
- Training set representativity

With simulations

Out of Distribution

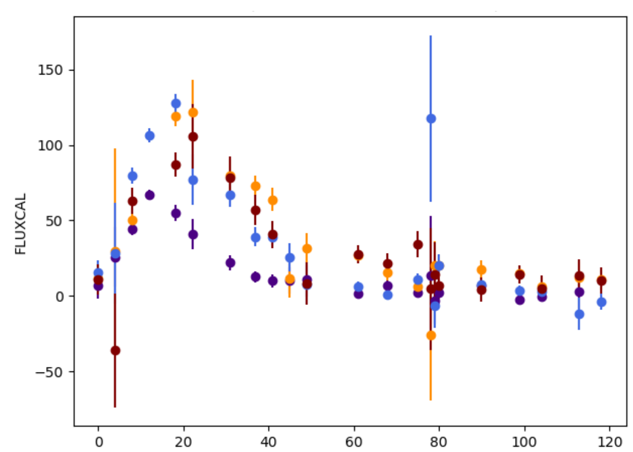
training set

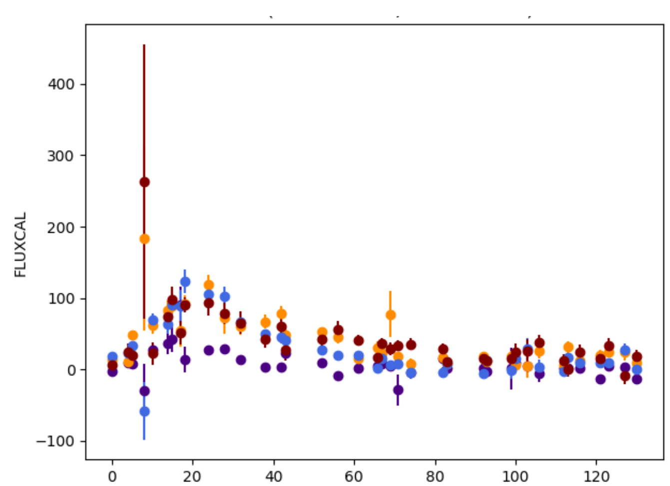
to classify

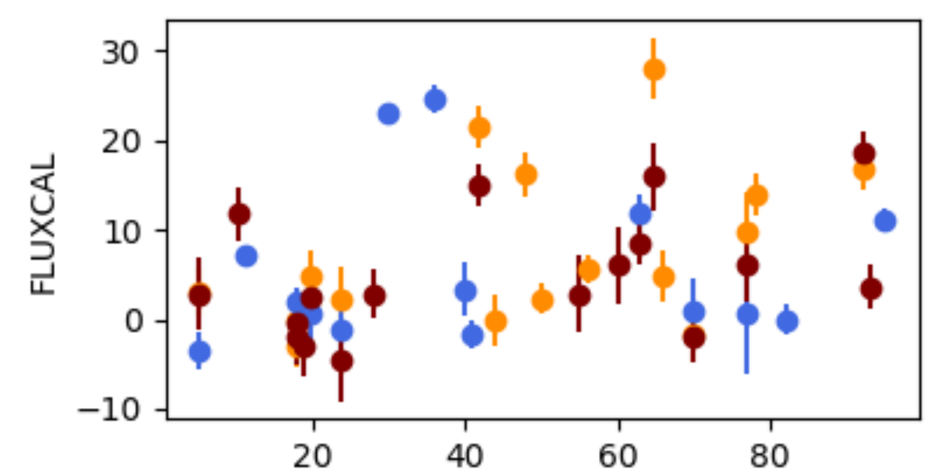
Out of Distribution

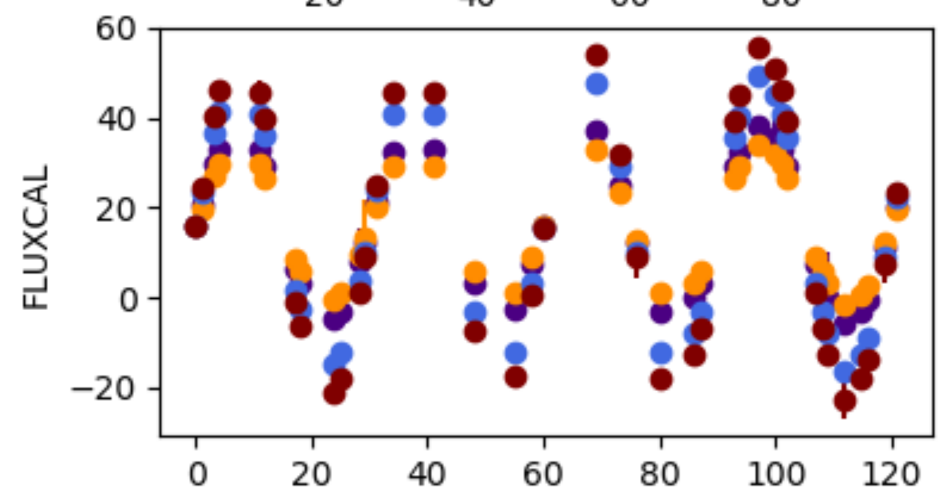
training set

to classify

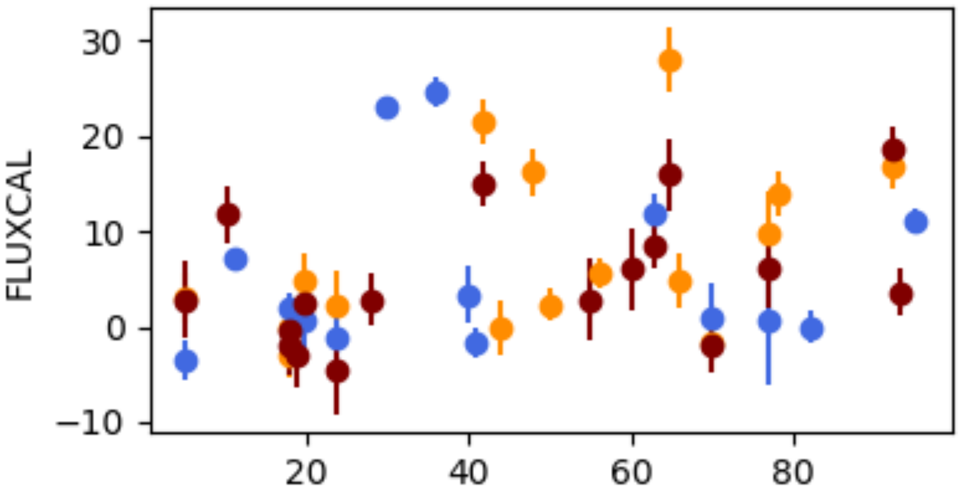








Out of Distribution

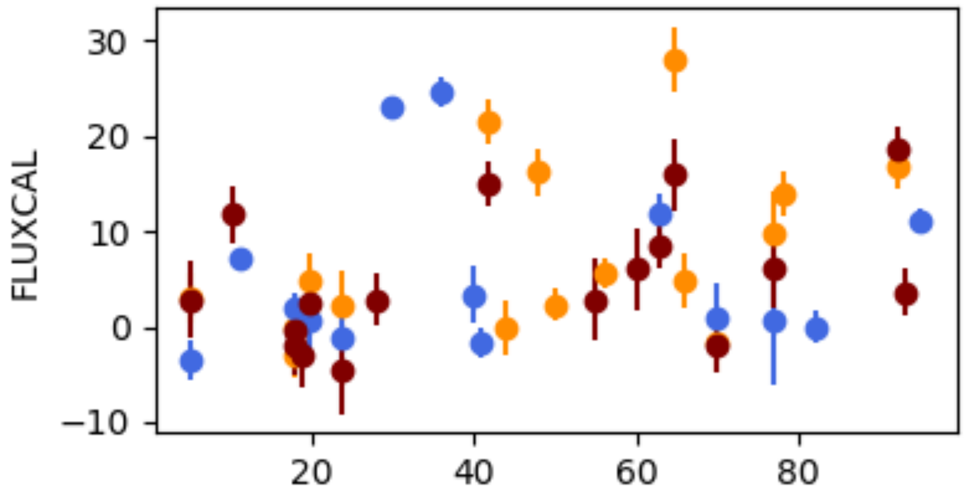


RNN

low probability for any class

0.2

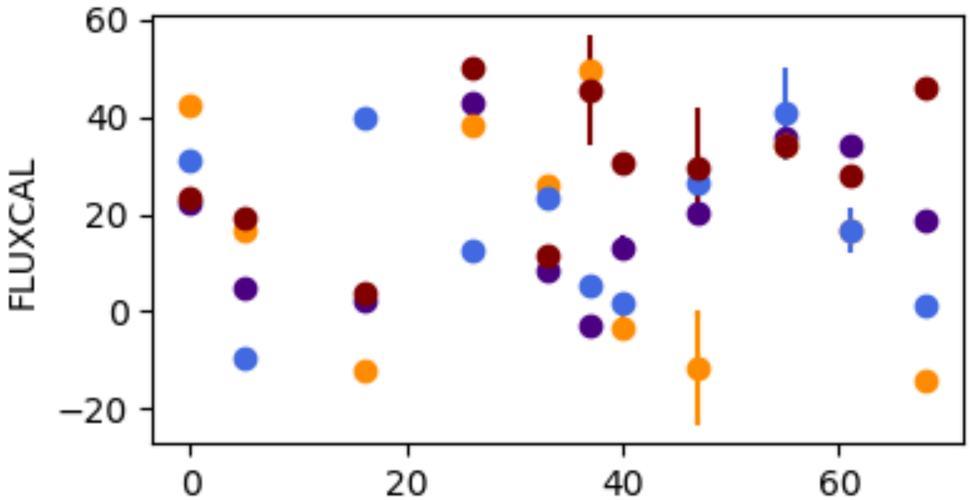
Out of Distribution



RNN

low probability for any class

0.2

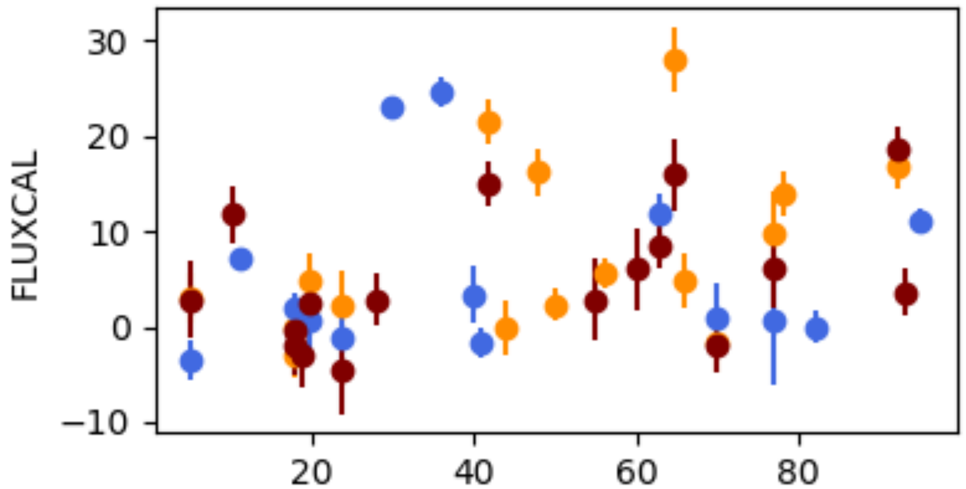


overconfident prediction for the difficult class

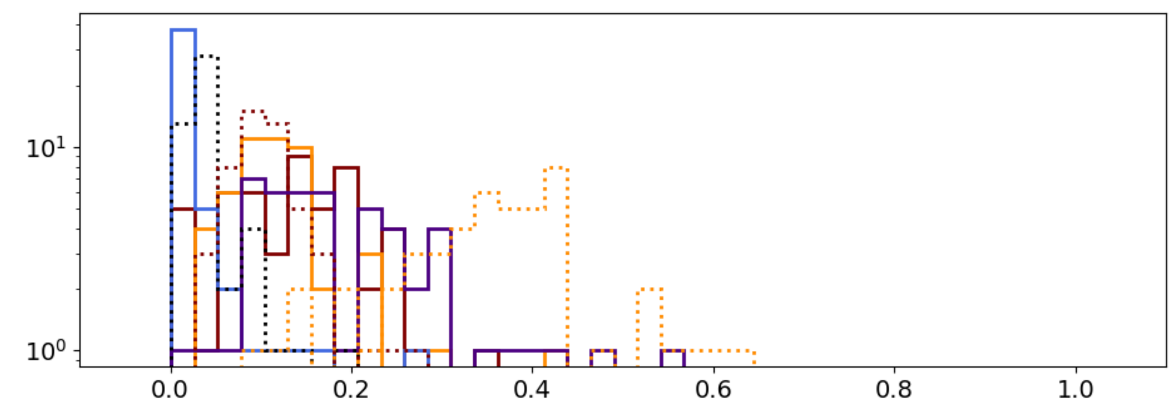
0.8!

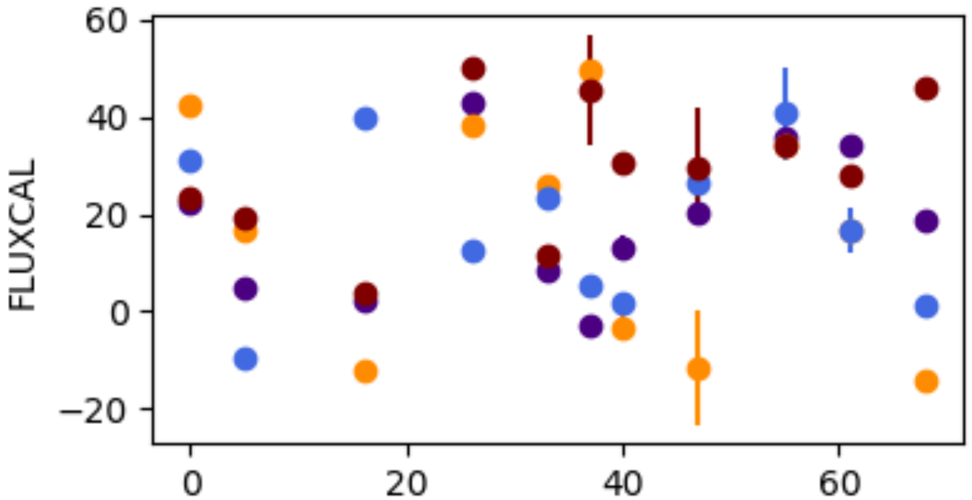
Out of Distribution

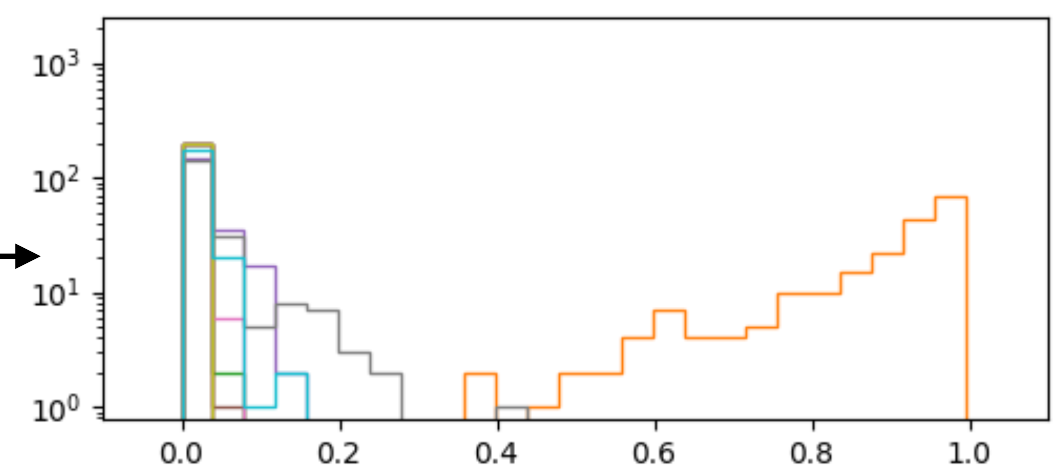
Distribution of predictions



BRNN







"an increase in average classification uncertainties for these anomalies"

For both MC dropout and BBB (although different amplitudes)

Metrics

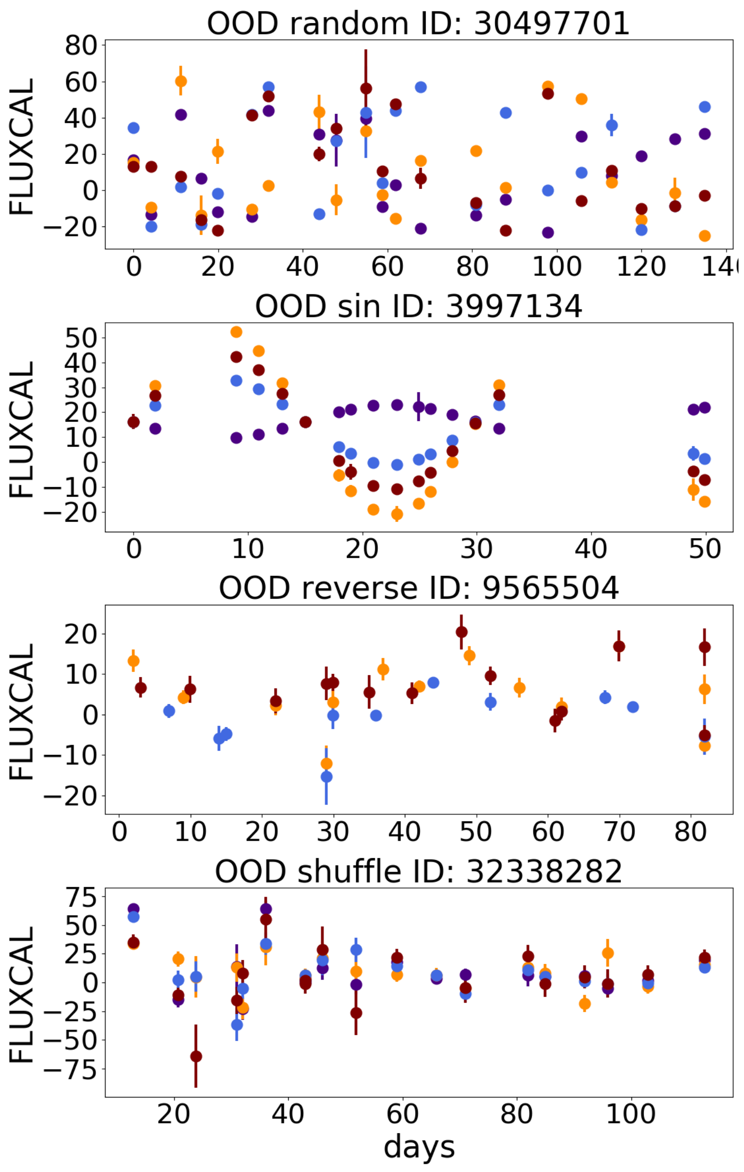
Entropy (Fortunato+ 2017)

$$H_m[\mathcal{D}_t] = \sum_{i=1}^{N} p_m(\mathbf{y}_i|\mathbf{x}_i) log\left(\frac{1}{p_m(\mathbf{y}_i|\mathbf{x}_i)}\right).$$

$$\Delta H = \quad \overline{H_{m_1}}[\mathcal{D}_t] - \overline{H_{m_2}}[\mathcal{D}_t] \quad or = \overline{H_m}[\mathcal{D}_{t_1}] - \overline{H_m}[\mathcal{D}_{t_2}].$$

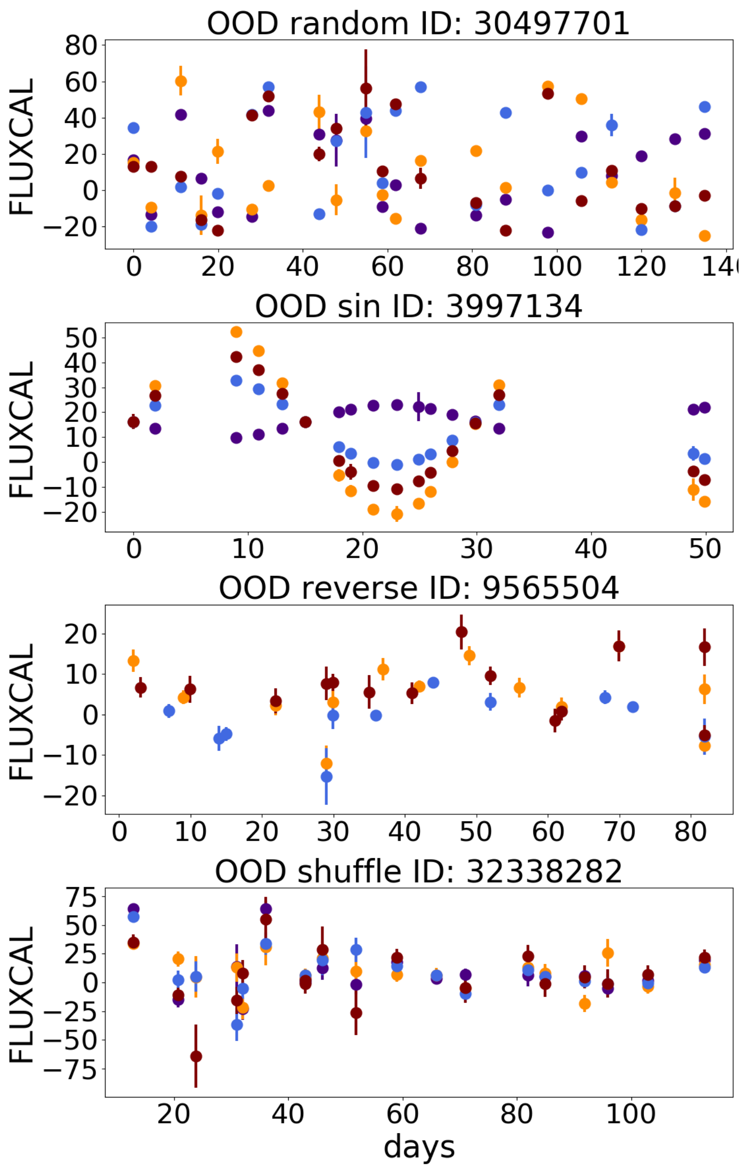
OOD should show high entropy compared to non-OOD

Out of Distribution



OOD high entropy
BBB largest entropy gap

Out of Distribution



OOD negative entropy for MC dropout

¯_(ツ)_/¯

???????

Representativity

- Tagged datasets are small ~thousands of supernovae and are biased
- Simulations are usually used for training and benchmarking

simulation

peak brightness *i*

→ train classification algorithm

Representativity

- Tagged datasets are small ~thousands of supernovae and are biased
- Simulations are usually used for training and benchmarking

simulation

peak brightness *i*

train classification algorithm

data

peak brightness *i*

classify the data

accuracy decreases ! *Lochner+ 2016, Charnock+2017*

Representativity

Model 1:
trained with
simplistic
simulation

peak brightness *i*

SuperNNova
open source photometric
classification

Model 2:
trained with
representative
simulation

peak brightness *i*

classify

representative
simulation

Representativity

SuperNNova
open source photometric classification

Model 1:
trained with
simplistic
simulation

peak brightness *i*

classify

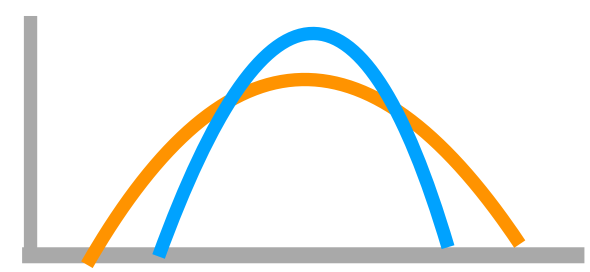
Model 2:
trained with
representative
simulation

peak brightness *i*

representative simulation

accuracy changes slightly (<prob> are not the most indicative)

non-representative models give larger uncertainties!



Prospects for Rubin

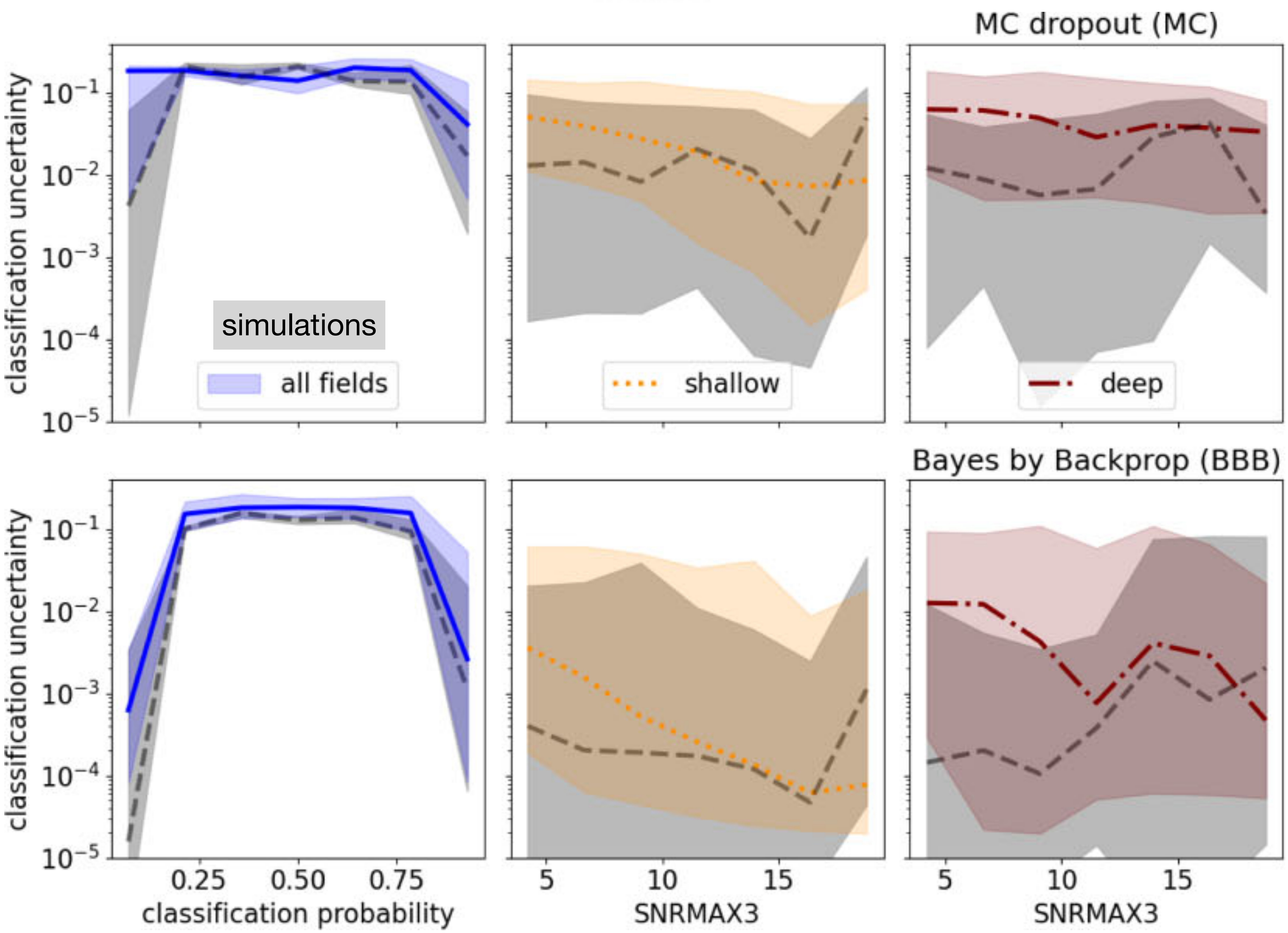
(i) can BNN uncertainties be used as an indicator of the **representativity of the training set** for a given data set?

(ii) can BNN **uncertainties replace selection cuts**?

With data

Aleatoric?

DARK ENERGY SURVEY



Uncertainty $\propto$ data SNR

Uncertainty $\not\propto$ simulations SNR

Representativity

DARK ENERGY SURVEY

trained with simulation

SuperNNova
open source photometric classification

DARK ENERGY SURVEY

SNIa Data

ZTF

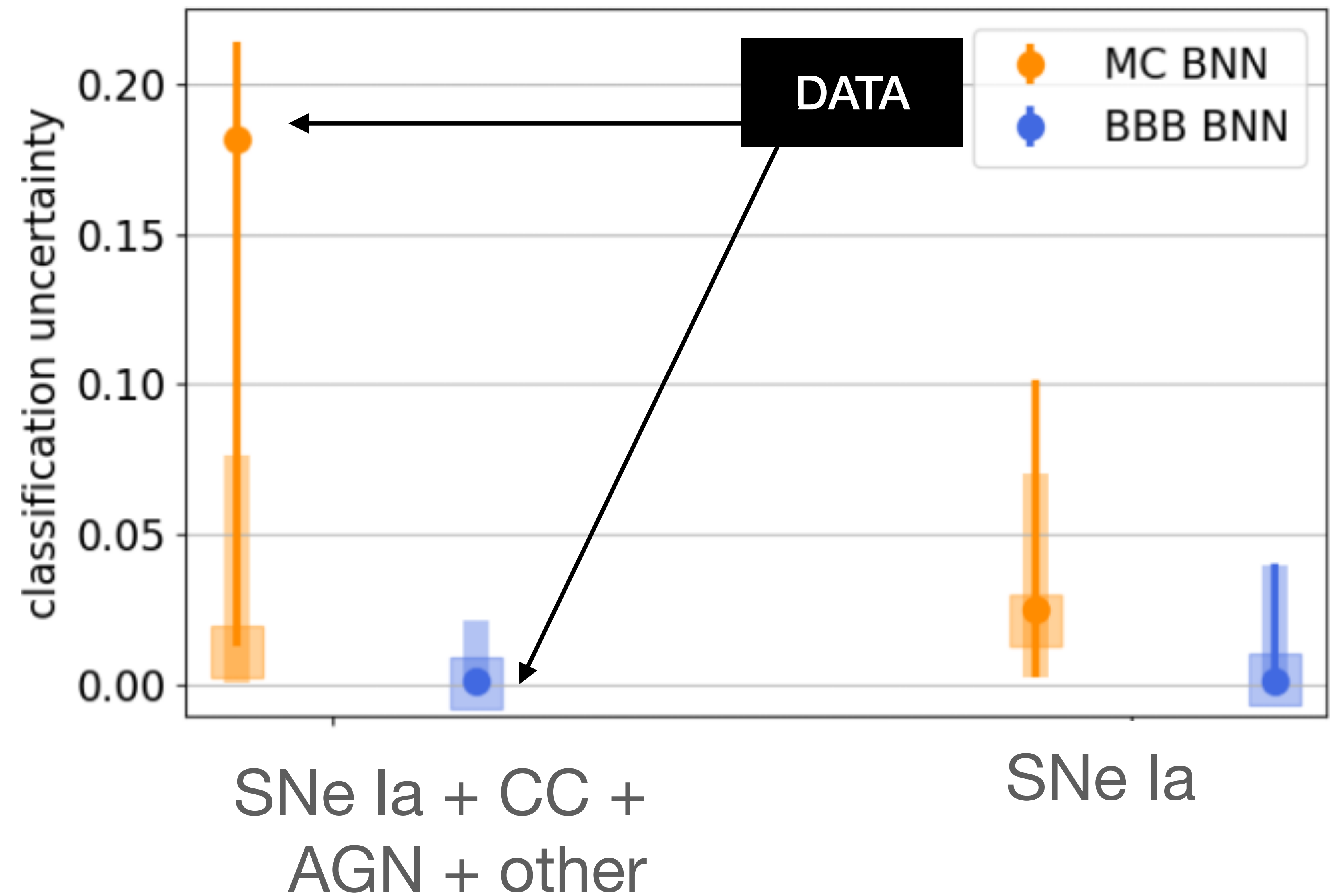
FINK

SNIa data

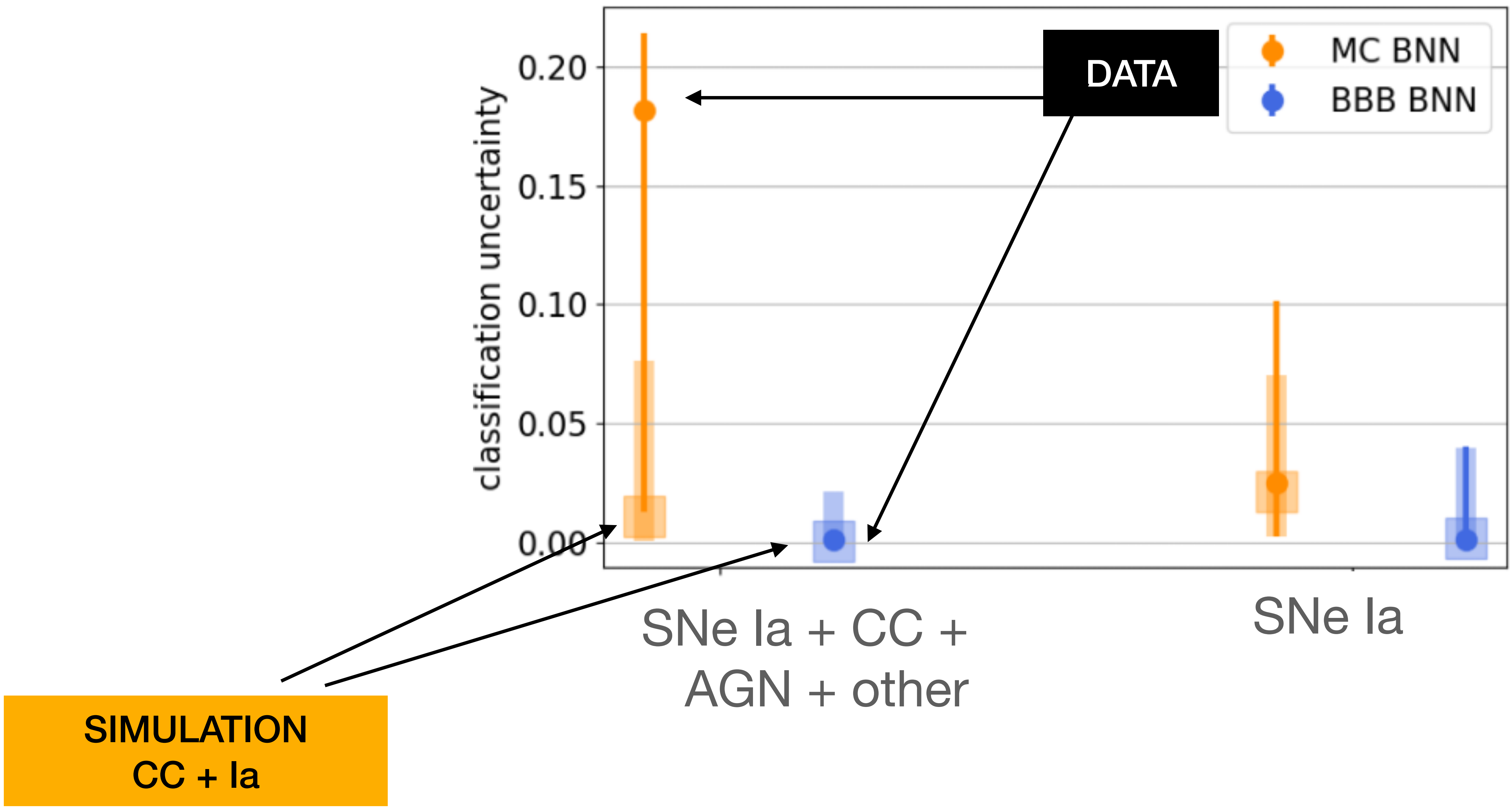
2 orders magnitude uncertainty increase!

Möller+ 2022b (ICML)

Representativity + OOD



Representativity + OOD

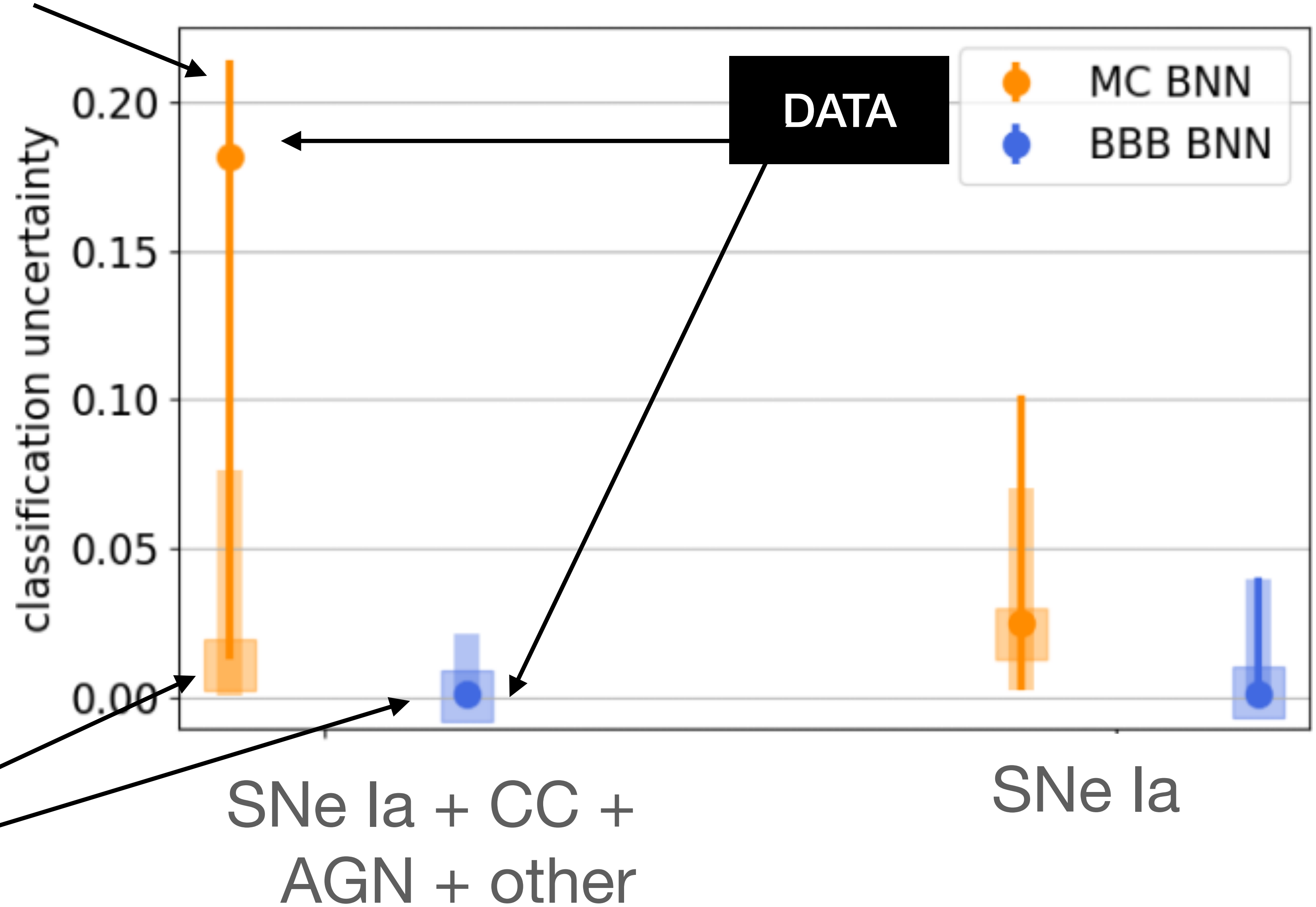


Möller+2022

Representativity + OOD

DARK ENERGY
SURVEY

OOD?



Uncertainty << with selection cuts than without for BBB
but not for MC

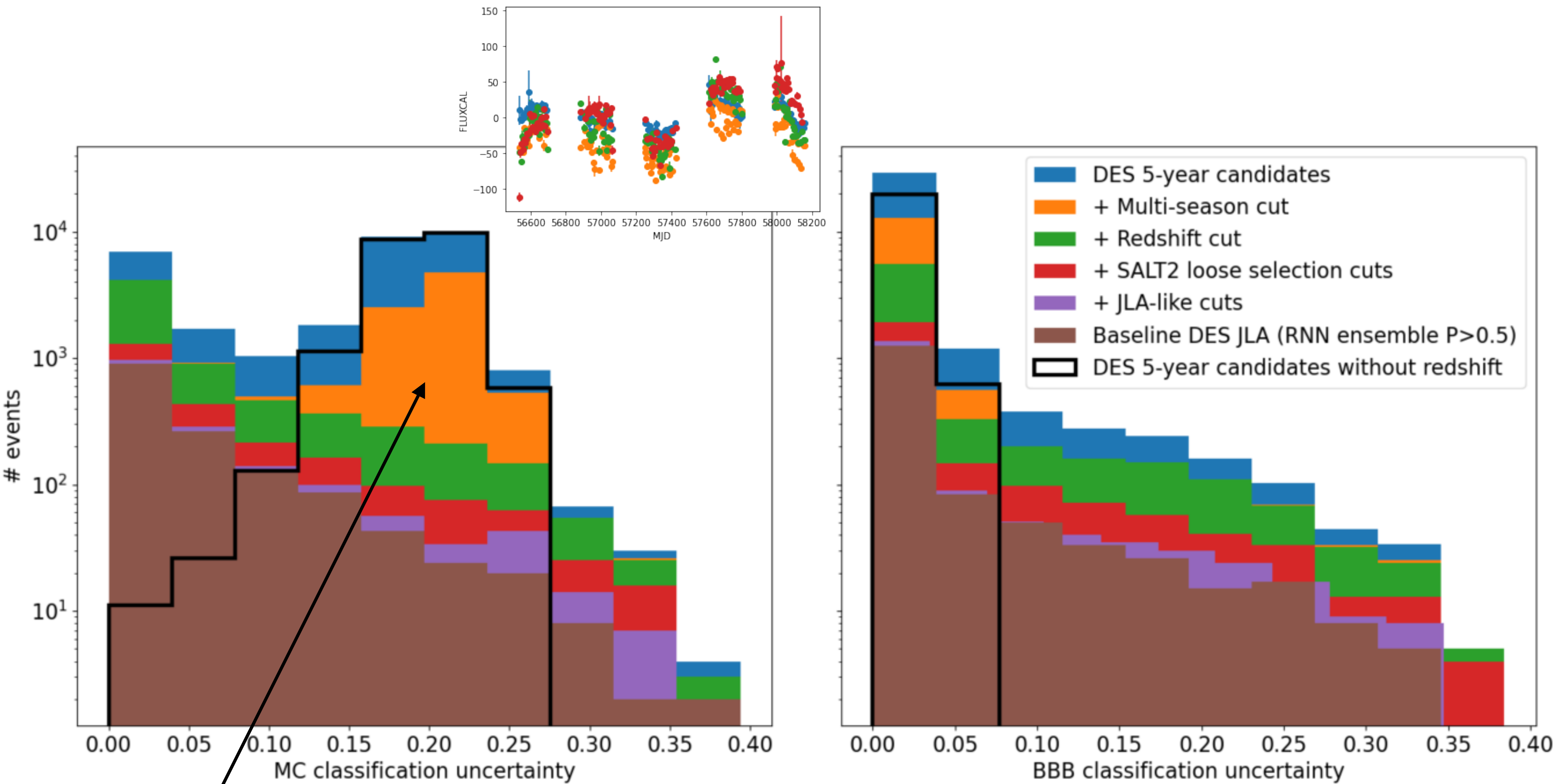
¯_(ツ)_/¯

???????

SIMULATION
CC + Ia

Möller+2022

Replace selection cuts?



MC dropout uncertainties
can identify OOD

(^_^)

Summary

BNNs **can provide classification uncertainties** for time-domain

Should we use these uncertainties for cosmology? Not yet…

- OOD and lack of training set representativity can be reflected
- Some behaviours are unexpected!
- It is hard to know if an uncertainty is “calibrated”

But the uncertainties can be insightful for large dataset analysis!