Foundation Models From Outside Astronomy, and Their Utility for Blazar Variability

Sasha (Alexander) Plavin

Black Hole Initiative Fellow, Harvard → Jansky Fellow, NRAO

Are foundational time-series models here already?

Sasha (Alexander) Plavin

Black Hole Initiative Fellow, Harvard → Jansky Fellow, NRAO

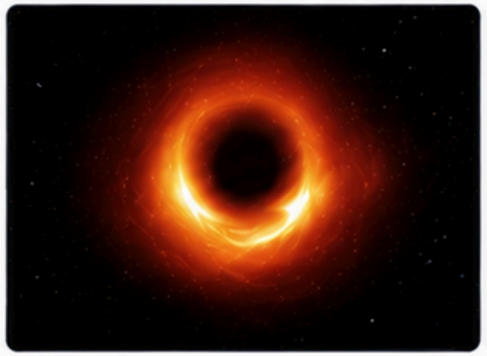
Inputs

Text (question)

Do we see a full ring morphology here?

Images

Natural image



Astronomical image

Submit input + question to VLM/LLM (ChatGPT)

VLM/LLM

Generate answer (structured output)

Answer (JSON)

```
{
    "full_ring": true,
    "confidence": 0.92,
    "reason": "Bright ring
     surrounding central
     dark region."
}
```

Submit input to VLM/LLM

VLM/LLM

Get embedding

Regress / Classify

Prediction

e.g., score = 0.93

or

e.g., full ring

how about time series?

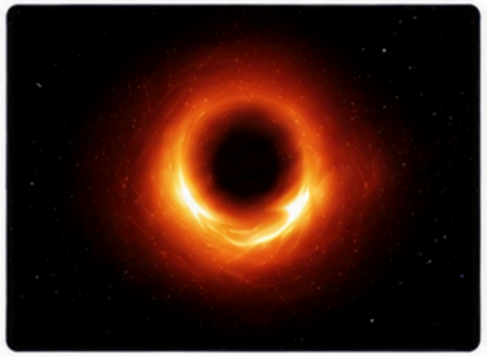
Inputs

Text (question)

Do we see a full ring morphology here?

Images

Natural image



Astronomical image

Submit input + question to VLM/LLM (ChatGPT)

VLM/LLM

Generate answer (structured output)

Answer (JSON)

```
{
    "full_ring": true,
    "confidence": 0.92,
    "reason": "Bright ring
     surrounding central
     dark region."
}
```

Submit input to VLM/LLM

VLM/LLM

Get embedding

Regress / Classify

Prediction

e.g., score = 0.93

or

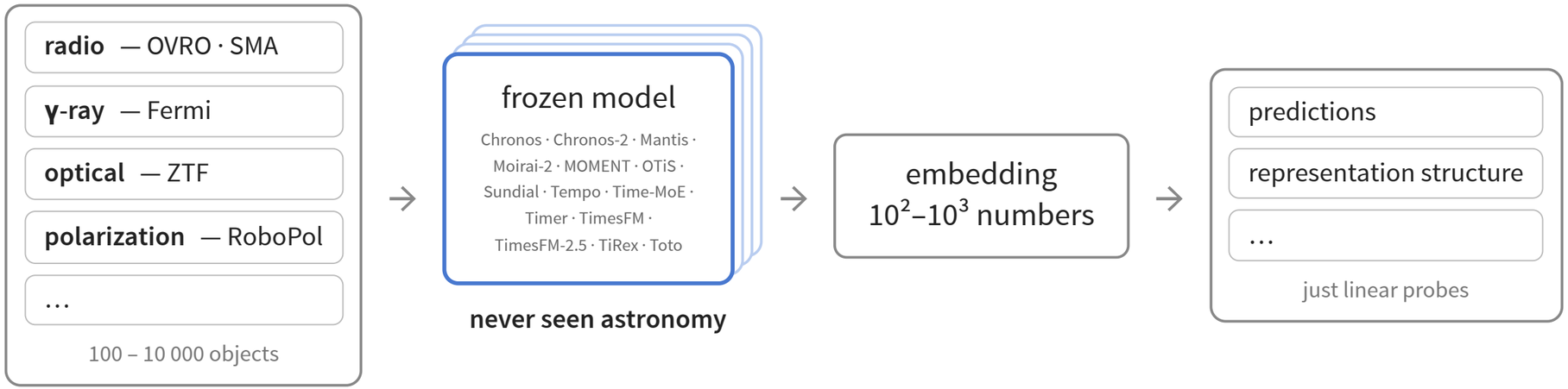
e.g., full ring

Astronomy time series
in small diverse data regime:

- light curves, polarization, geometry, ...
- bespoke analysis & connections
- infeasible to train dedicated models

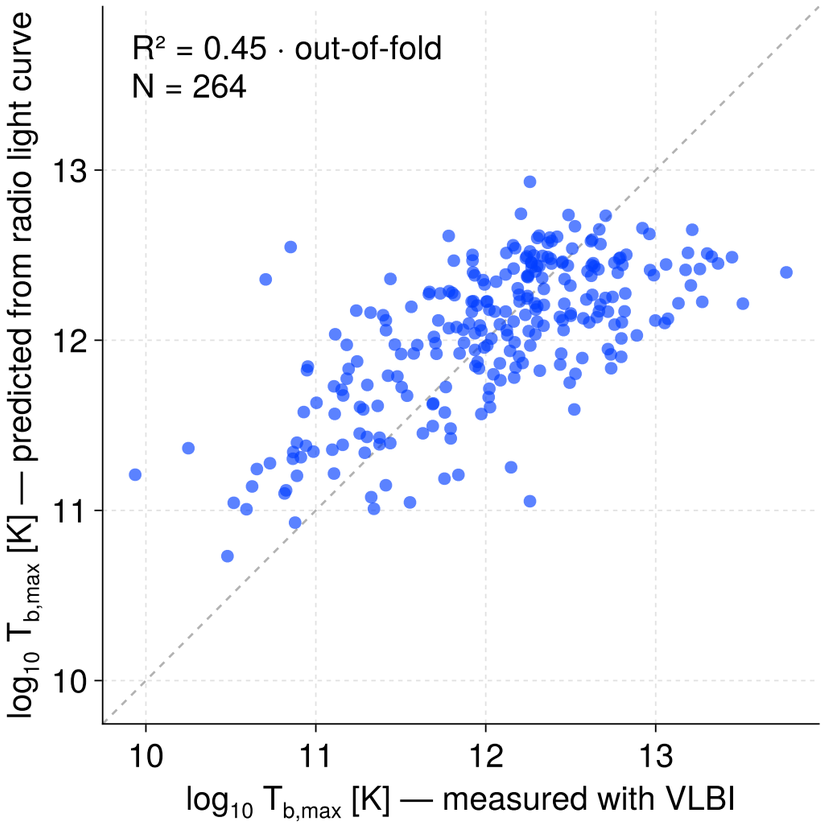
Examples: Li+25 for StarEmbed, Auer+25 for non-astro FM-as-extractors

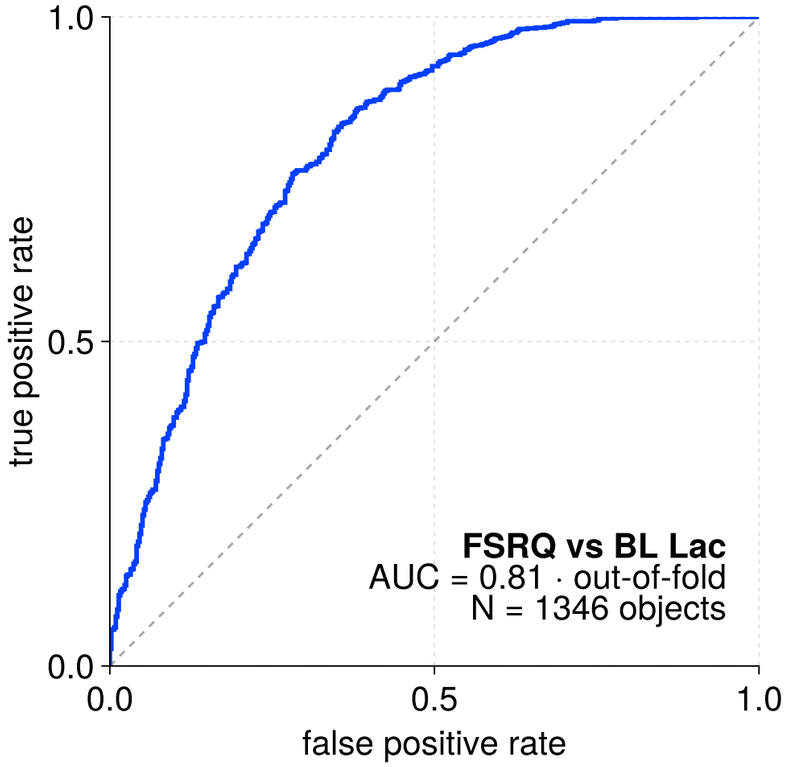
what do we do?


radio — OVRO · SMA
γ-ray — Fermi
optical — ZTF
polarization — RoboPol
…
100 – 10 000 objects
frozen model
Chronos · Chronos-2 · Mantis · Moirai-2 · MOMENT · OTiS · Sundial · Tempo · Time-MoE · Timer · TimesFM · TimesFM-2.5 · TiRex · Toto
never seen astronomy
embedding
10^2–10^3 numbers
predictions
representation structure
…
just linear probes

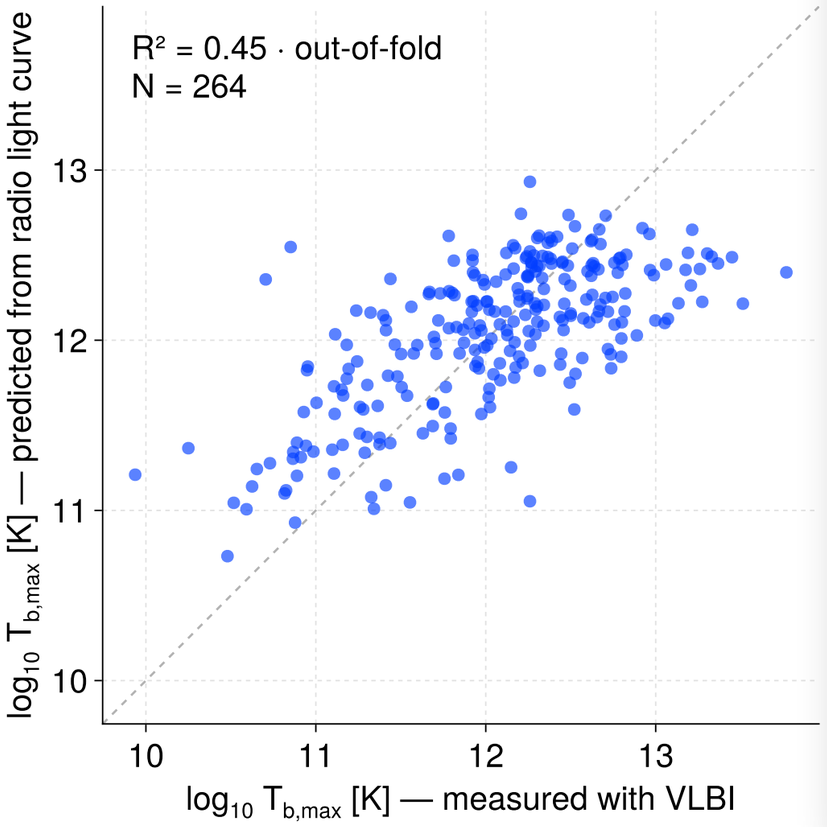
what do we get?

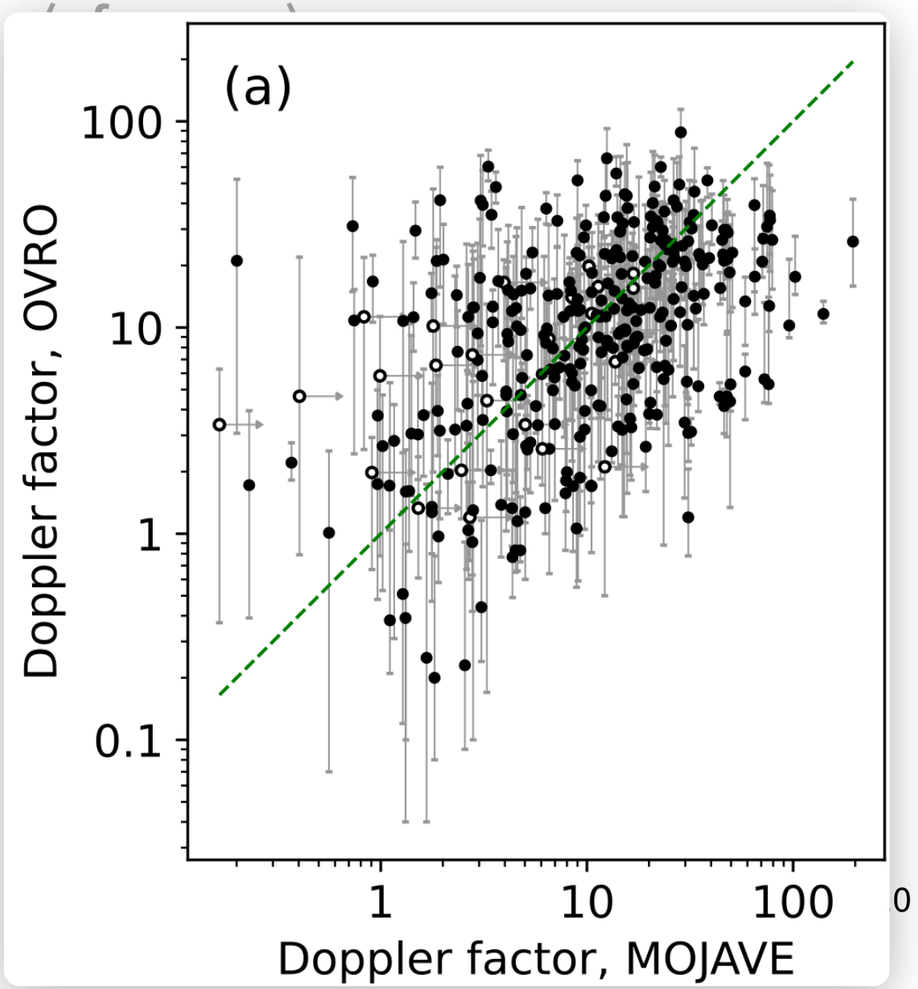
predictions (of course)





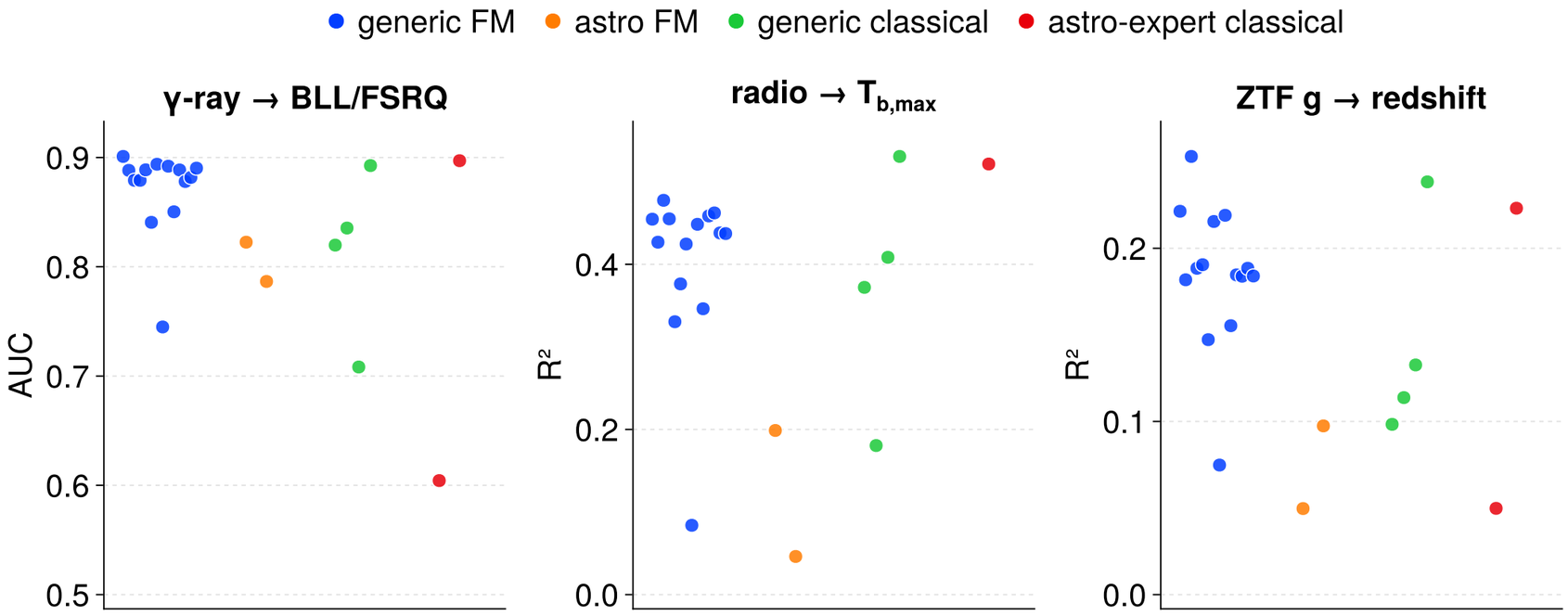
prediction


R² = 0.45 · out-of-fold
N = 264
log10 Tb,max [K] — predicted from radio light curve
log10 Tb,max [K] — measured with VLBI
10
11
12
13


(a)
Doppler factor, OVRO
Doppler factor, MOJAVE
0.1
1
10
100

predictions: quality

the audit: **31** physical targets × **6** bands × **20+** extractors — cross-validated, out-of-fold

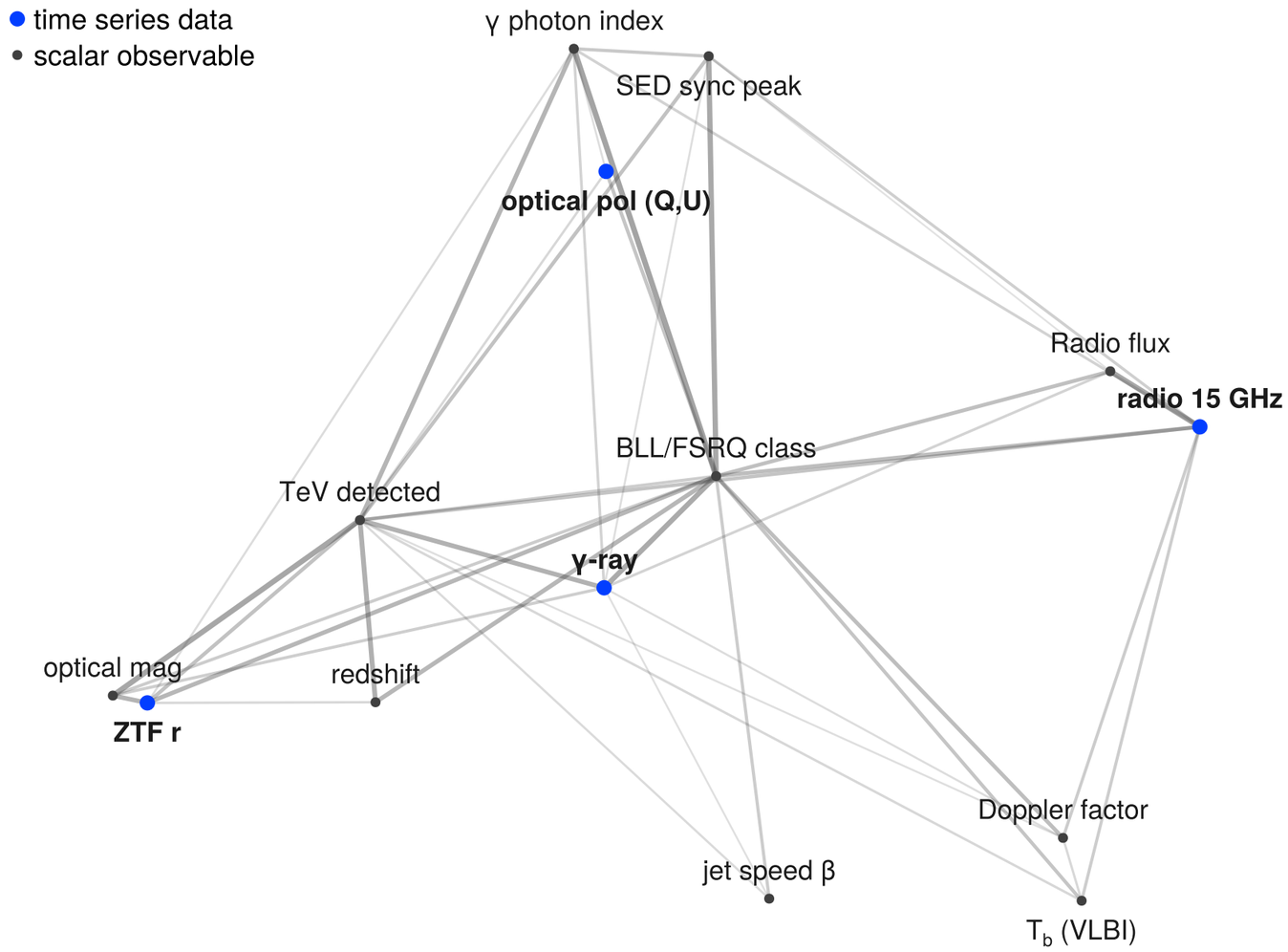


my personal pick: chronos from amazon

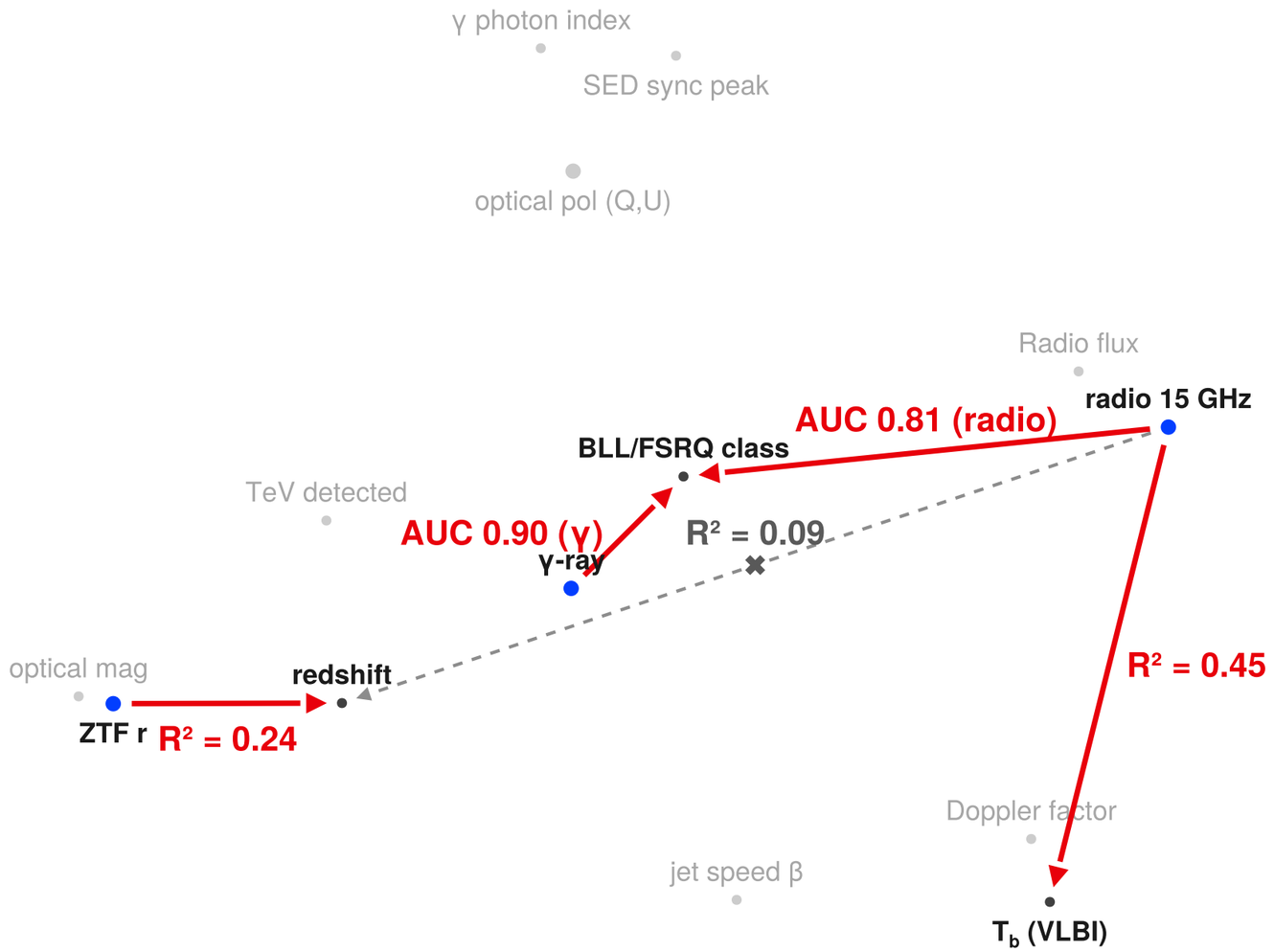
even better than astro-FM (Astromer, StarCLR) – at zero cost for us

tie with classical – networks learn expert-crafted features

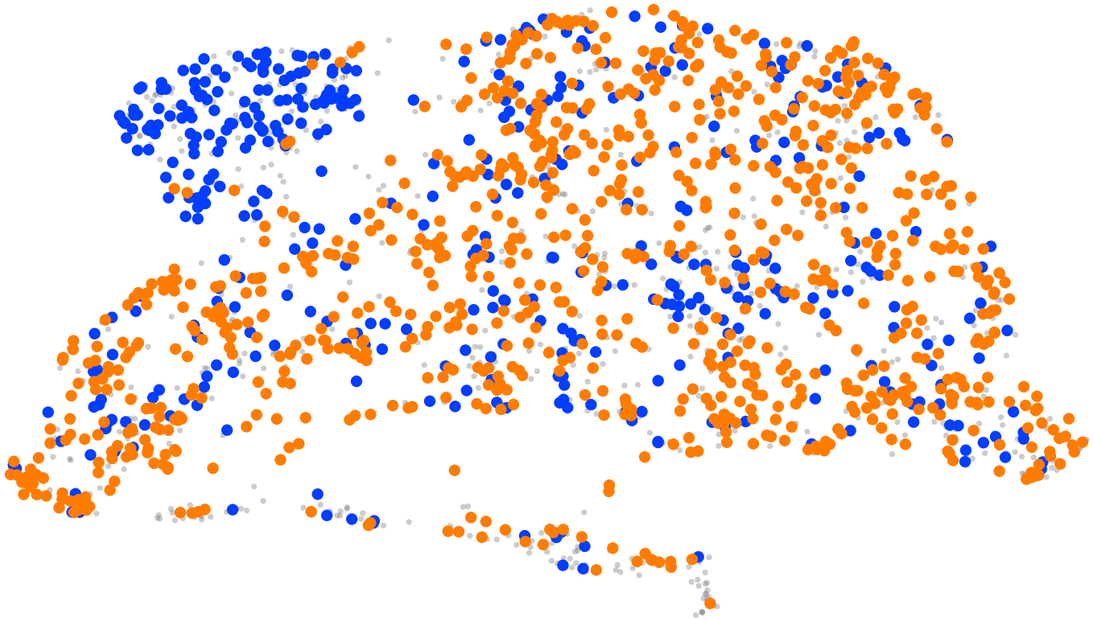
predictability navigation chart



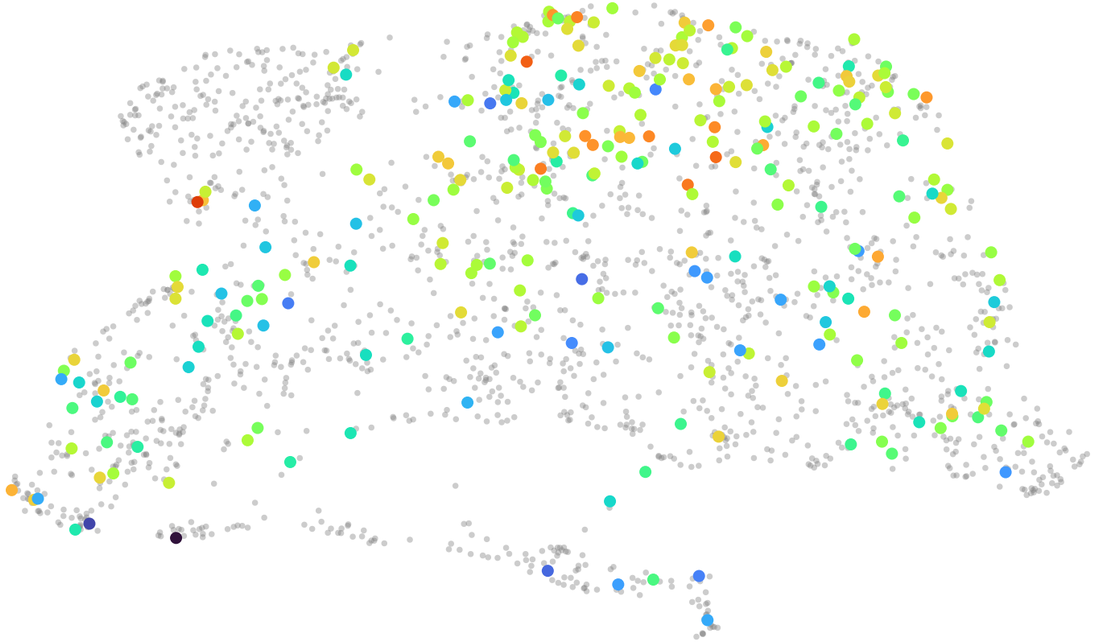
predictability navigation chart



the representation space

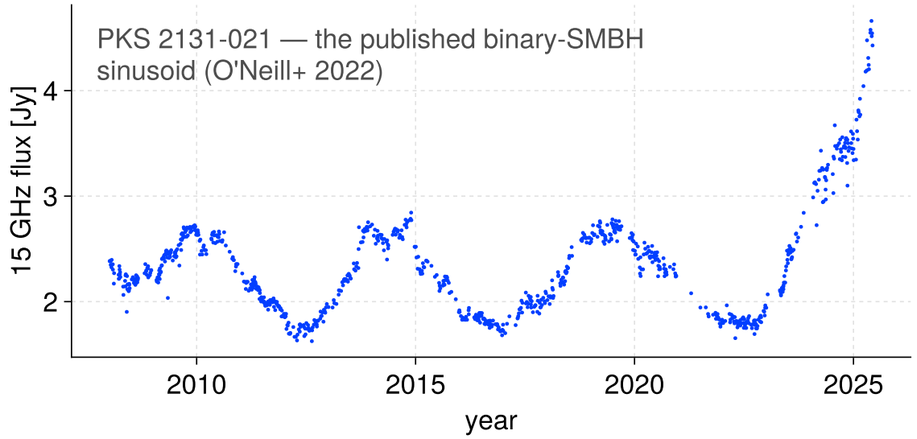


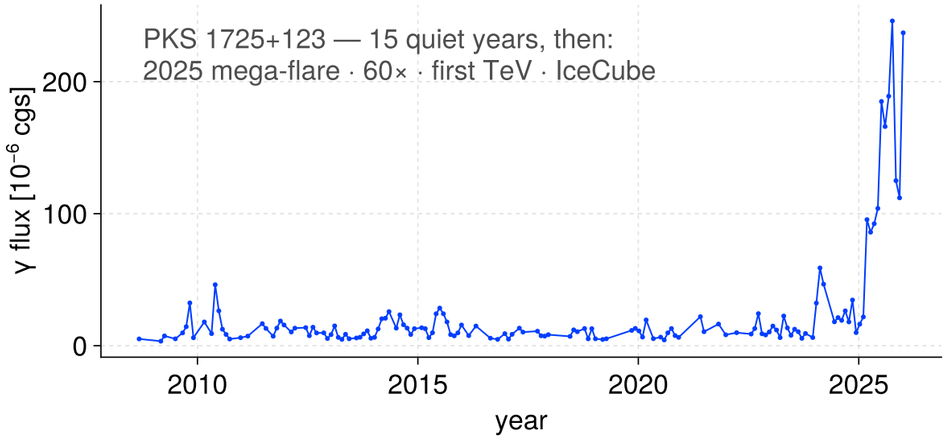
by quasar type: strong structure

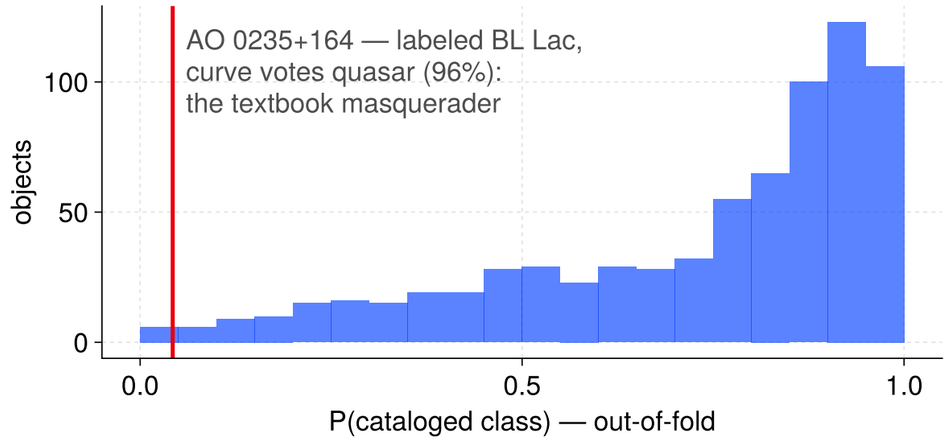


by VLBI core brightness: subdominant despite successful prediction

anomaly detection (manually inspected top-12)



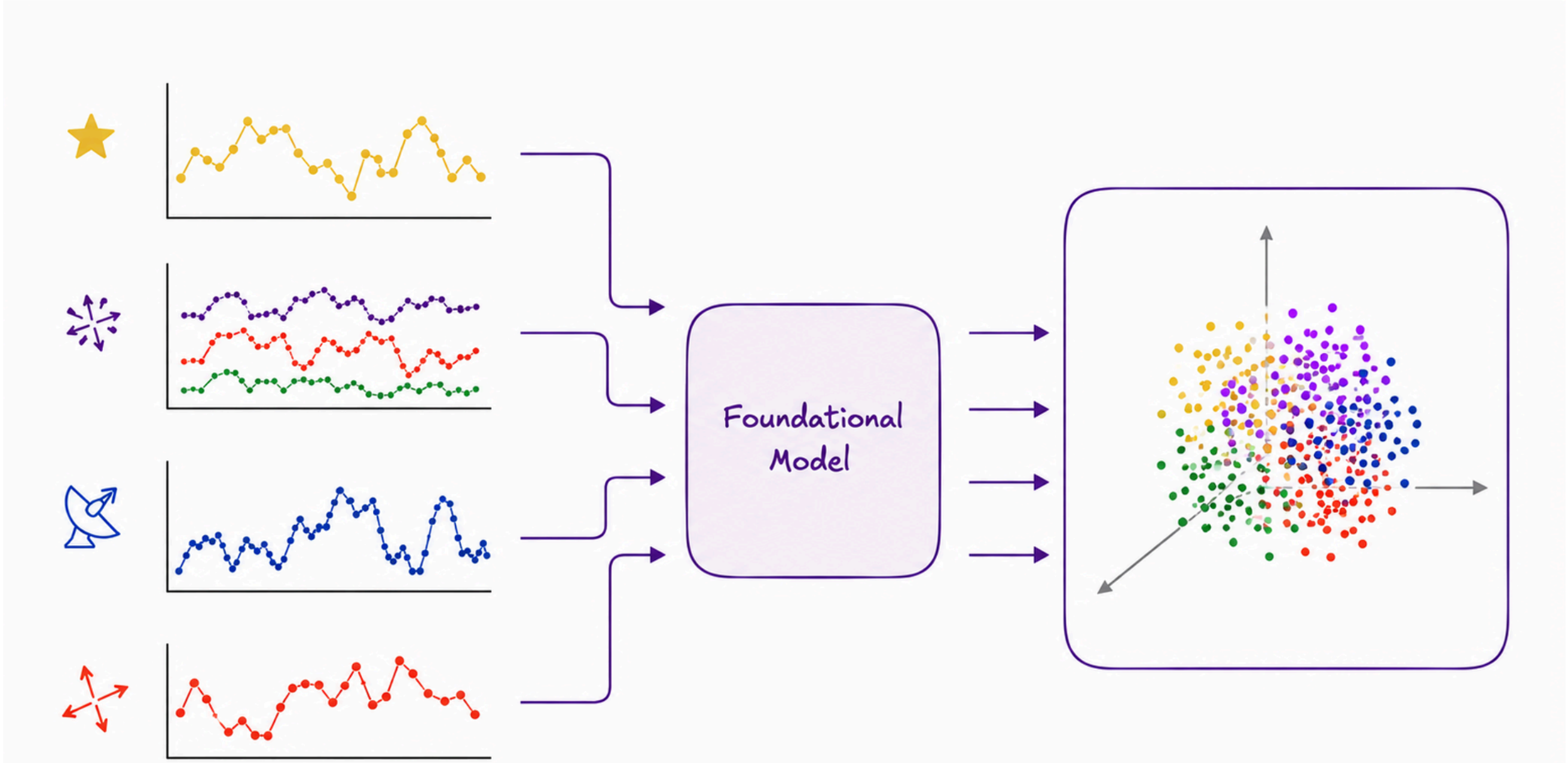




… also: an outdated redshift caught,
and more, known and unknown — found automatically

What's still hard

- extremely irregular time coverage
- joint multi-timeseries embeddings: delays, time correlations, ...
- shared embedding spaces: not directly yet; maybe learn simple linear transforms?



Are foundational time-series models here already?

tie hand-crafted features and bespoke analysis – for free

- frozen weights + simple linear probe
- from prediction to unfamiliar field exploration
- ultimately, any data shape: image, timeseries, ...
  - now – the worst these models will ever be

every feature extractor

generic foundation models (14)

Chronos · Chronos-2 · Mantis · Moirai-2 · MOMENT · OTiS · Sundial · Tempo · Time-MoE · Timer · TimesFM · TimesFM-2.5 · TiRex · Toto — from Amazon, Google, Datadog, Salesforce, Tsinghua, ...

astro foundation models (2)

Astromer · StarCLR — pretrained on astronomical light curves, fed their native irregular cadence

generic classical features (4)

catch22 · tsfresh · signatures · stats — hand-crafted time-series feature banks + simple summary statistics

astro-expert classical features (2)

gp-bank · lcf — GP/DRW variability-model fits and astronomy light-curve feature sets

all compared at the same settings: linear probe on top, cross-validated, out-of-fold