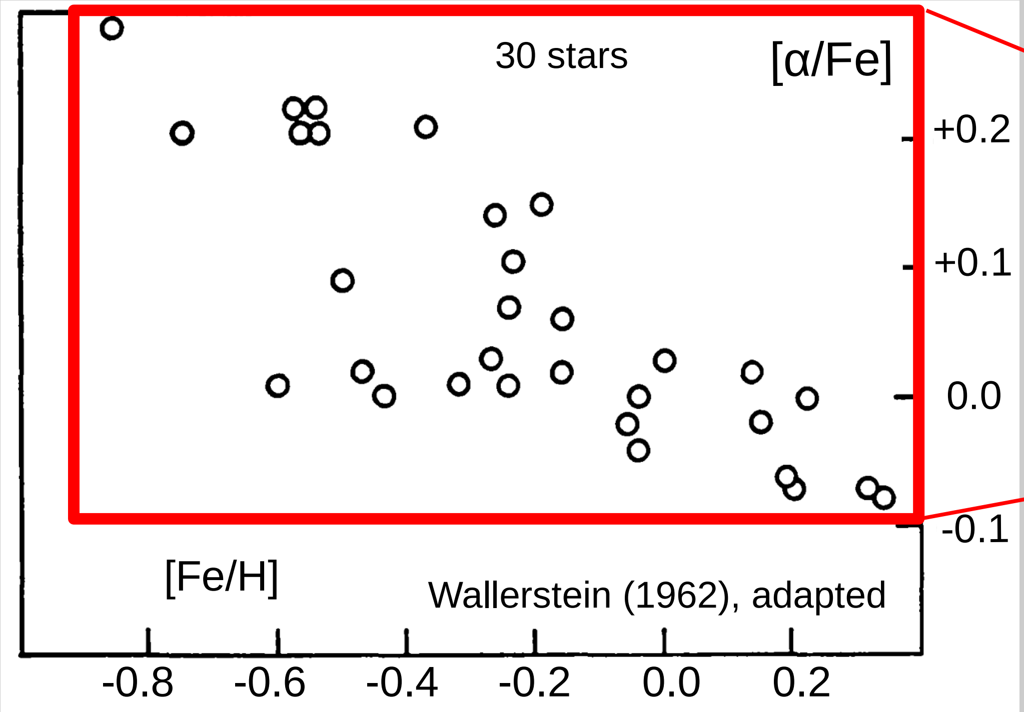
Machine-learning methods for stellar spectroscopy in the era of large scale surveys

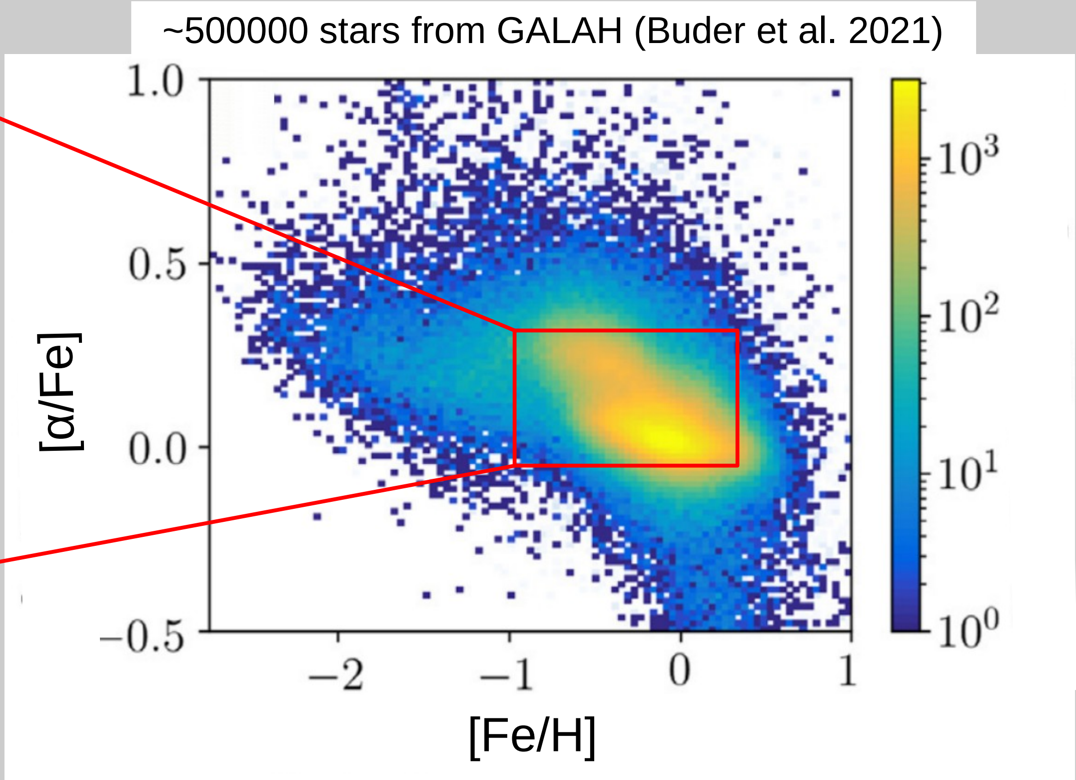
Guillaume Guiglion (GG)

with S. Nepal, C. Chiappini, M. Ambrosch, M. Steinmetz, M. Valentini, S., G. Matijevič and R. de Jong

The Milky Way Assembly Tale – 29/05/2024

Stellar abundances for Galactic Archaeology





Stellar kinematics for Galactic Archaeology → eg. Paul, Teresa, Danny's talks

Stellar ages for Galactic Archaeology → Today's session on ages :))

The need for large spectroscopic surveys

RAVE
RADIAL VELOCITY EXPERIMENT

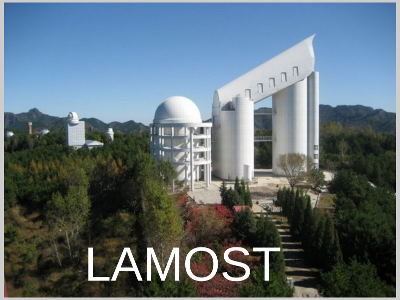
5×10^5

GaiaESO

10^5

ESO
AMBRE

$>10^4$



10^6

GALAH

5×10^5

APOGEE
SDSSIII

$>5 \times 10^5$

gaia

10^6

The need for large spectroscopic surveys

RAVE — Radial Velocity Experiment

5×10^5

GaiaESO

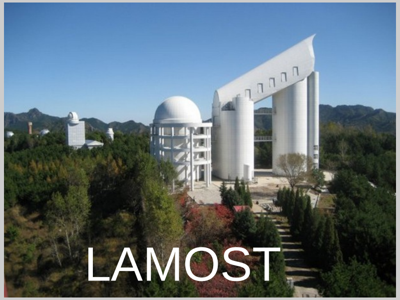
10^5

ESO Ambre

$>10^4$

WEAVE

$>10^6$



10^6

GALAH

10^6

4MOST

$>10^7$

MOONS — Multi Object Optical and Near-infrared Spectrograph

$>10^6$

APOGEE — SDSSIII

10^6

gaia

10^7

WST

$>10^7$

For stellar parametrization:

→ One can use standard spectroscopy

e.g.:

- **SME** (Valenti & Piskunov; @GALAH)
- **FERRE** (Allende-Prieto et al. 2006; @PRISTINE)
- **MATISSE** (Recio-blanco et al. 2006; @RAVE, @*Gaia*-RVS)
- **GAUGUIN** (GG et al. 2016, 2018; @GES, @RAVE)

For stellar parametrization:

→ One can use standard spectroscopy

e.g.:

- **SME** (Valenti & Piskunov; @GALAH)
- **FERRE** (Allende-Prieto et al. 2006; @PRISTINE)
- **MATISSE** (Recio-blanco et al. 2006; @RAVE, @*Gaia*-RVS)
- **GAUGUIN** (GG et al. 2016, 2018; @GES, @RAVE)

→ One can use machine-learning

e.g.:

- **Cannon** (Ness et al. 2015; @RAVE, @GALAH)
- **Payne** (Ting et al. 2019; @LAMOST, @APOGEE)
- **CNN**
  - → Leung & Bovy 2019 (@APOGEE)
  - → Bialek et al. 2019 (@GES UVES)
  - → GG et al. 2020 (@RAVE)
  - → Ambrosch, GG et al. 2023 (@GES HR10&21)
  - → Nepal, GG et al. 2023 (@GES HR15)
  - → GG et al. 2024 (@*Gaia* RVS)

For stellar parametrization:

→ One can use standard spectroscopy

e.g.:

- **SME** (Valenti & Piskunov; @GALAH)
- **FERRE** (Allende-Prieto et al. 2006; @PRISTINE)
- **MATISSE** (Recio-blanco et al. 2006; @RAVE, @*Gaia*-RVS)
- **GAUGUIN** (GG et al. 2016, 2018; @GES, @RAVE)

→ One can use machine-learning

e.g.:

- **Cannon** (Ness et al. 2015; @RAVE, @GALAH)
- **Payne** (Ting et al. 2019; @LAMOST, @APOGEE)
- **CNN**
  - → Leung & Bovy 2019 (@APOGEE)
  - → Bialek et al. 2019 (@GES UVES)
  - → GG et al. 2020 (@RAVE)
  - → Ambrosch, GG et al. 2023 (@GES HR10&21)
  - → **Nepal, GG et al. 2023 (@GES HR15)**
  - → **GG et al. 2024 (@*Gaia* RVS)**

Why do we want to use CNNs ?

- Adapted for large data sets, fast prediction power (*Gaia*, 4MOST ...)

- Very versatile: allow to combine different types of data (spectra, magnitudes, distances ...)

- Provide robust measurements from noisy data

→ Some literature:
LeCun et al. 1989
LeCun & Bengio 1995
Ciresan et al. 2011

CNN: basic concept

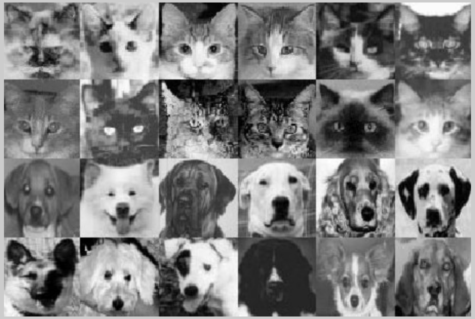
?

CNN: basic concept

?

1/ Build a **training sample**

Data | **Labels**



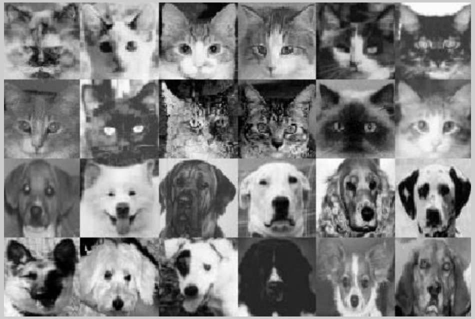
Cat
Dog
...
Dog
Cat

CNN: basic concept

→ ?

1/ Build a **training sample**

Data



Labels

Cat
Dog
...
Dog
Cat

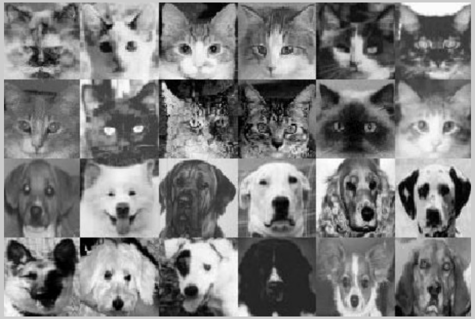
2/ Train a model between data and labels

Data → CNN: High-D non-linear function → **Labels**

CNN: basic concept

→ ?

1/ Build a **training sample**

Data	**Labels**
	Cat Dog ... Dog Cat

2/ Train a model between data and labels

Data → CNN:
High-D
non-linear function → **Labels**

3/ Predict the type of animal

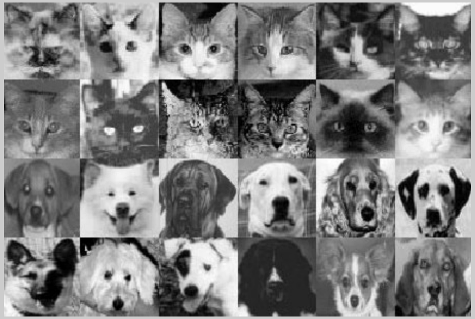
 → CNN model → **Dog**

CNN: basic concept

→ ?

1/ Build a **training sample**

Data | **Labels**



Cat
Dog
...
Dog
Cat

2/ Train a model between data and labels

Data → CNN: High-D non-linear function → **Labels**

3/ Predict the type of animal

→ CNN model → **Dog**

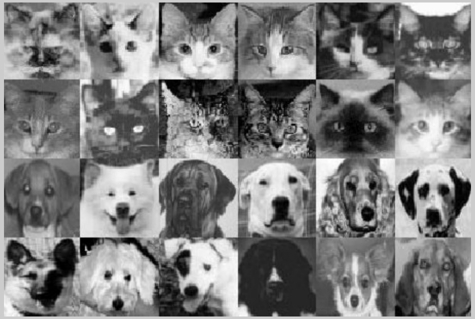
→

CNN: basic concept

→ ?

1/ Build a **training sample**

Data **Labels**



Cat
Dog
...
Dog
Cat

2/ Train a model between data and labels

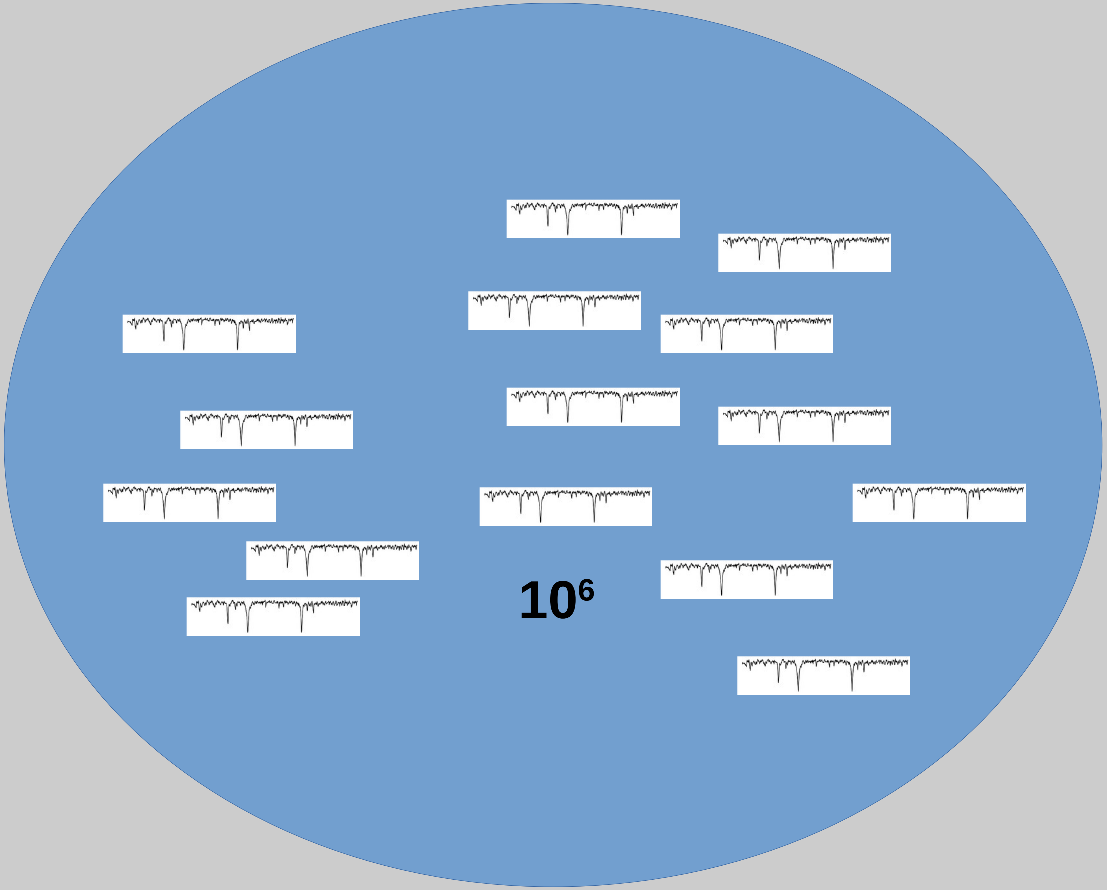
Data → CNN: High-D non-linear function → **Labels**

3/ Predict the type of animal

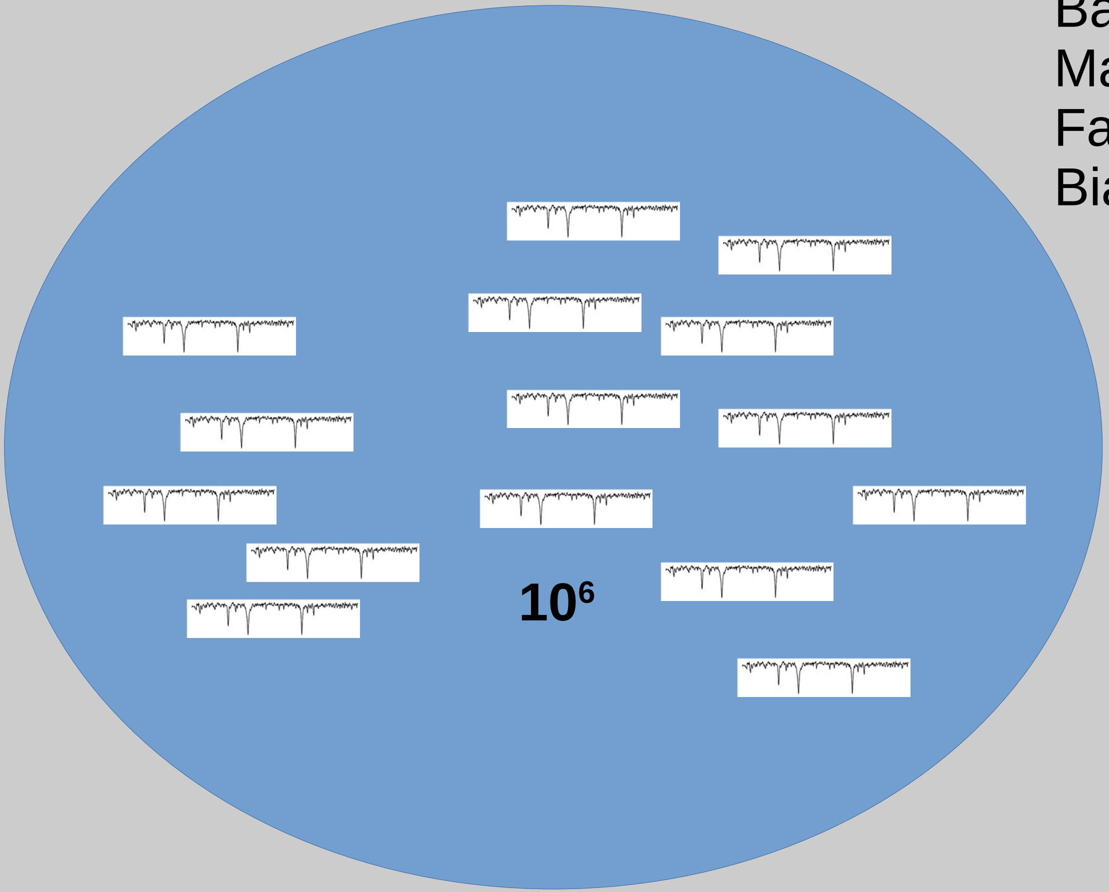
→ CNN model → **Dog**

→ **%$#5**

E.g. measuring T_{eff}, log(g), [Fe/H] in:



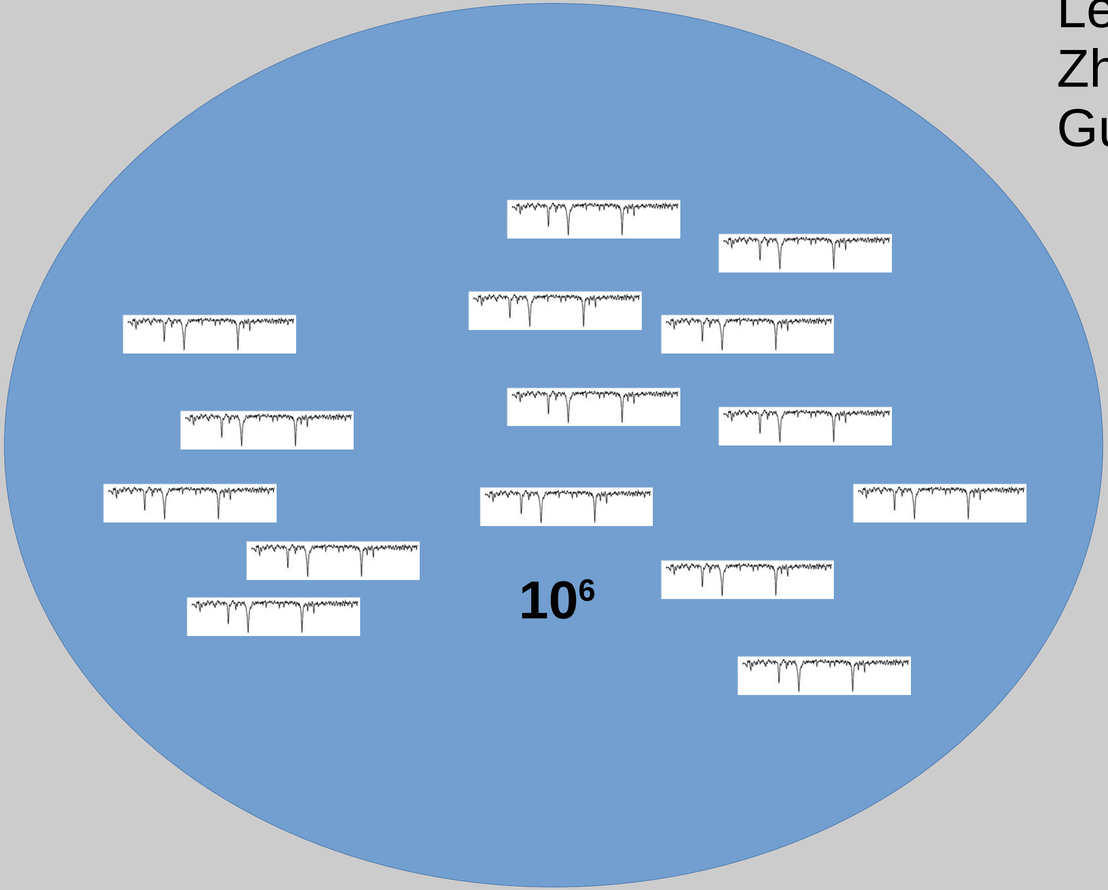
E.g. measuring T_{eff}, log(g), [Fe/H] in:



Model driven:

Bailer-Jones et al. 1997
Manteiga et al. 2010
Fabbro et al. 2018
Bialek et al. 2020

E.g. measuring T_{eff}, log(g), [Fe/H] in:

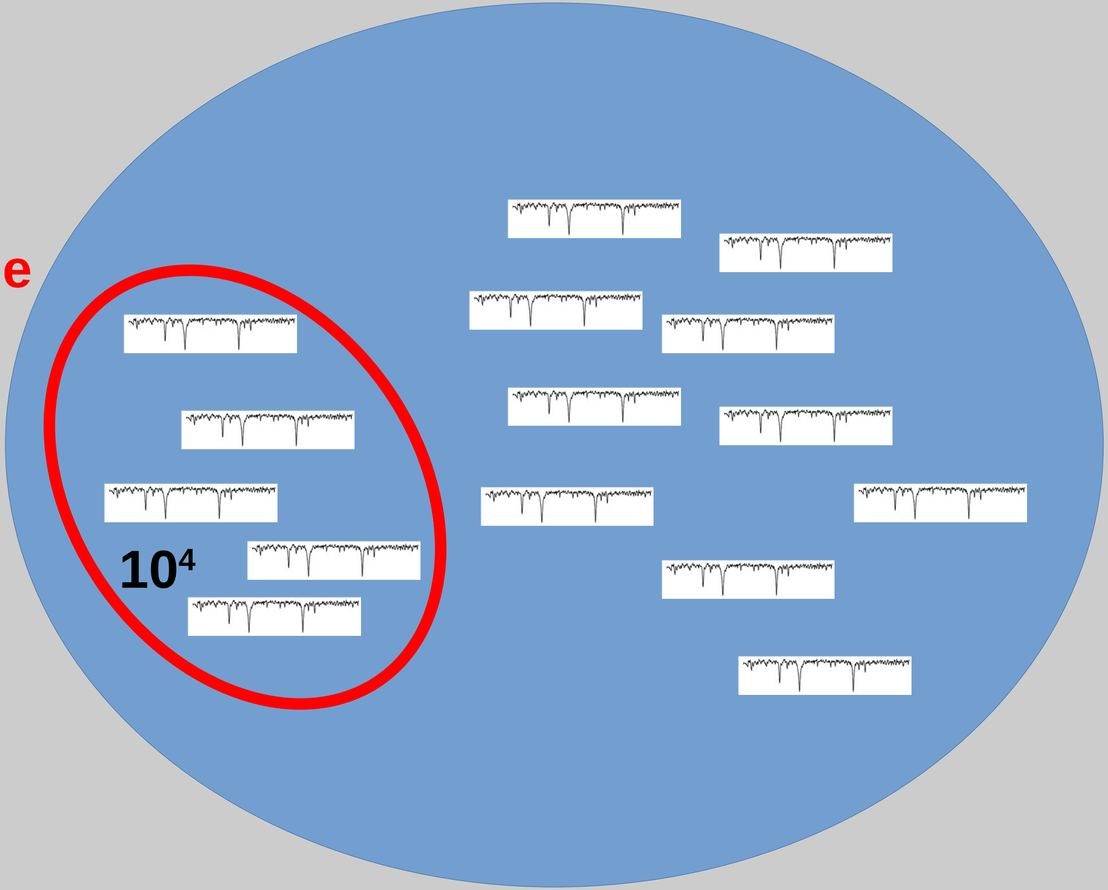


Data driven:
Leung & Bovy 2019
Zhang et al. 2019
Guiglion et al. 2020

E.g. measuring T_{eff}, log(g), [Fe/H] in:

T_{eff}, log(g), [M/H] from standard spectroscopy

→ **Training sample**



E.g. measuring T_{eff}, log(g), [Fe/H] in:

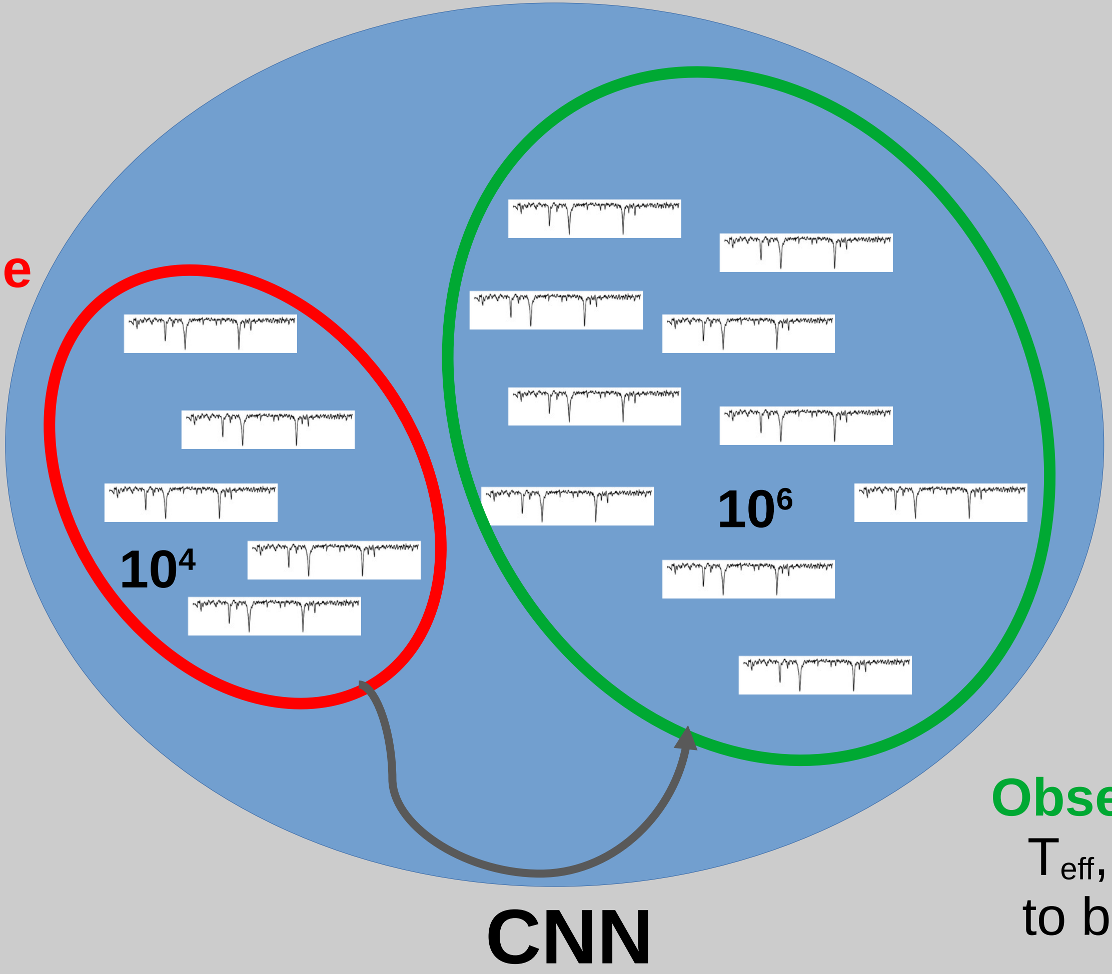
T_{eff}, log(g), [M/H]
from standard
spectroscopy
→ **Training sample**

Observed sample:
T_{eff}, log(g), [M/H]
to be determined

E.g. measuring T_{eff}, log(g), [Fe/H] in:

T_{eff}, log(g), [M/H] from standard spectroscopy

→ **Training sample**



Observed sample:

T_{eff}, log(g), [M/H] to be determined

How is CNN learning?
CNN for *Gaia*-ESO

A&A 671, A61 (2023)
https://doi.org/10.1051/0004-6361/202244765

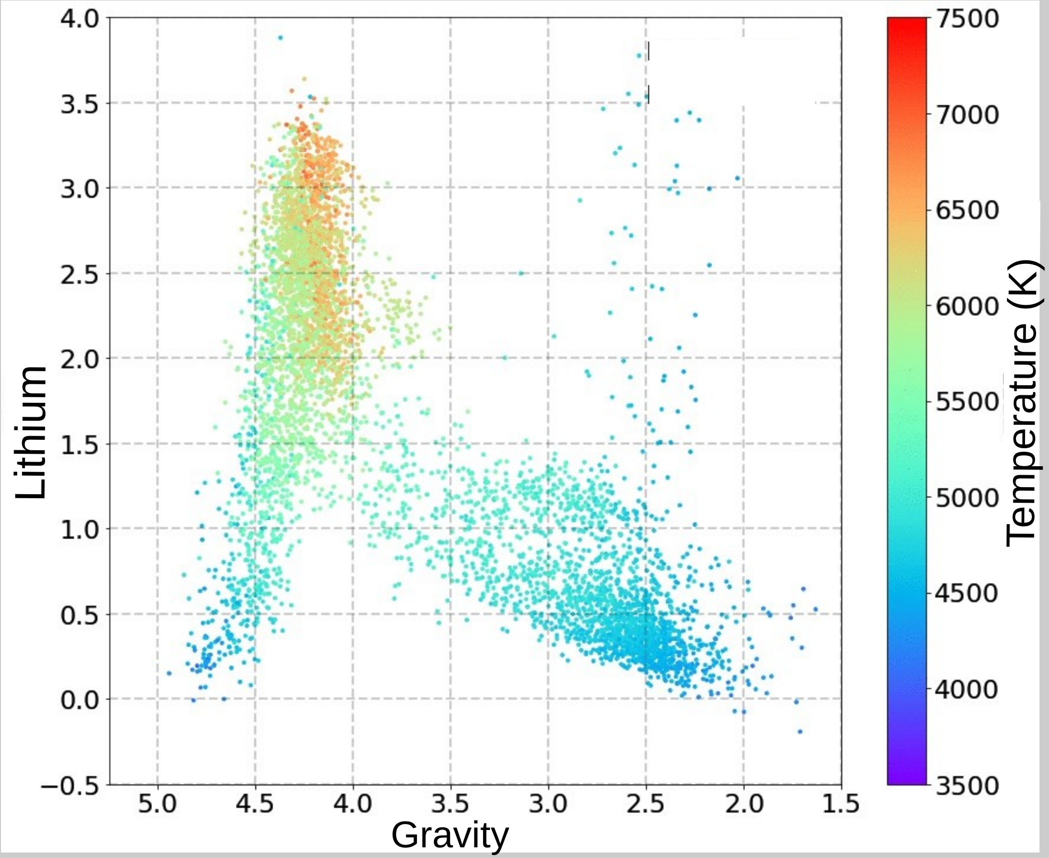
Astronomy & Astrophysics

The *Gaia*-ESO Survey: Preparing the ground for 4MOST and WEAVE galactic surveys

Chemical evolution of lithium with machine learning★,★★,★★★

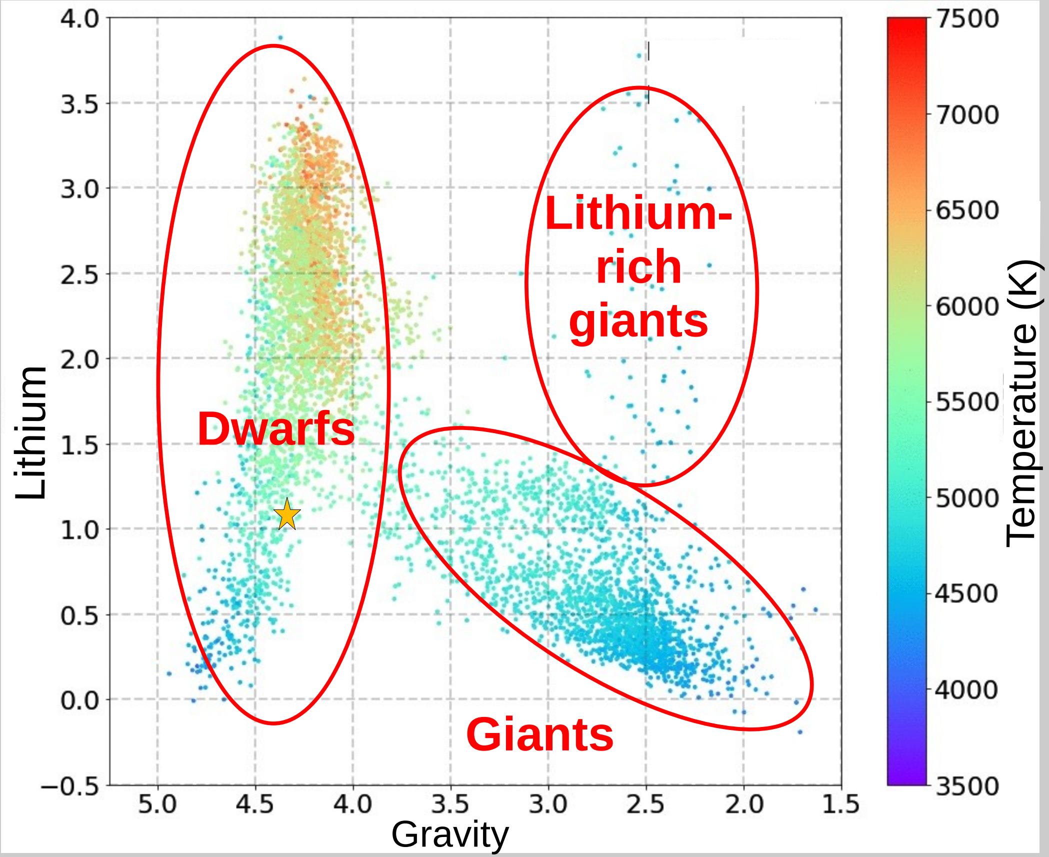
S. Nepal[1,2], G. Guiglion[3,1], R. S. de Jong[1], M. Valentini[1], C. Chiappini[1], M. Steinmetz[1], M. Ambrosch[4], E. Pancino[5], R. D. Jeffries[6], T. Bensby[7], D. Romano[8], R. Smiljanic[9], M. L. L. Dantas[9], G. Gilmore[10], S. Randich[5], A. Bayo[11], M. Bergemann[12,3], E. Franciosini[5], F. Jiménez-Esteban[13], P. Jofré[14], L. Morbidelli[5], G. G. Sacco[5], G. Tautvaišienė[4], and S. Zaggia[15]

Measurements of lithium in Milky Way stars with CNN & GES



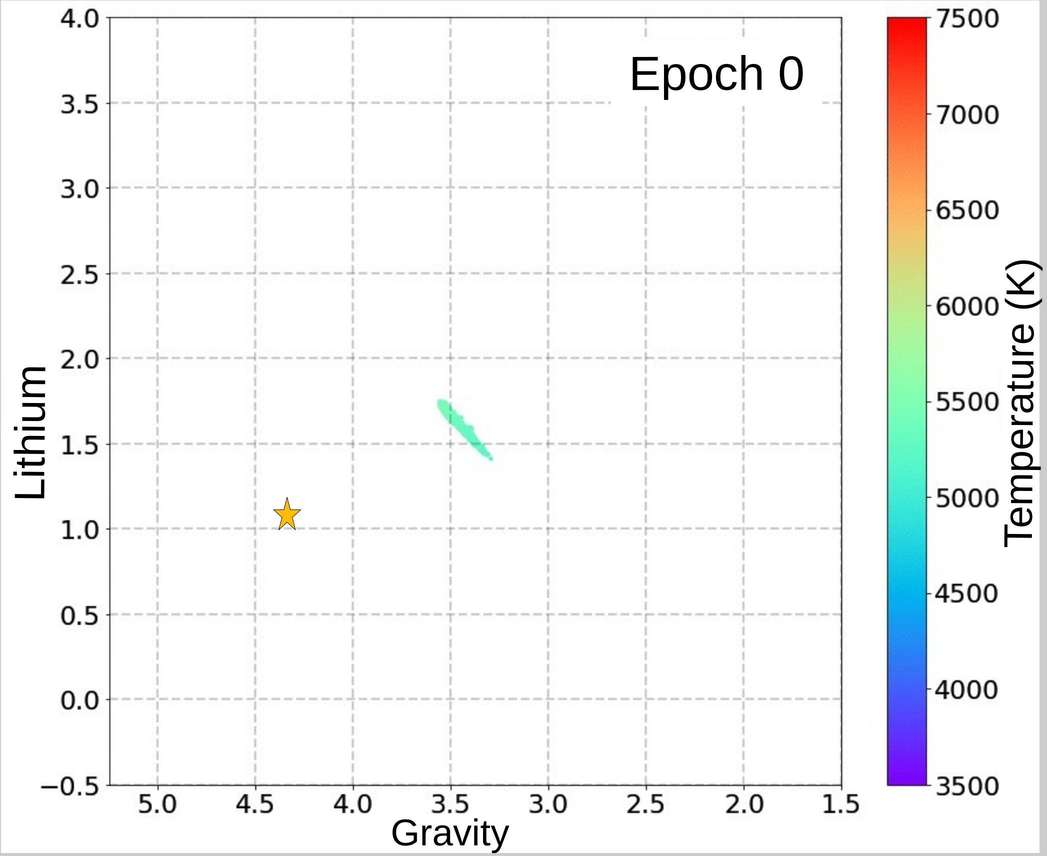
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



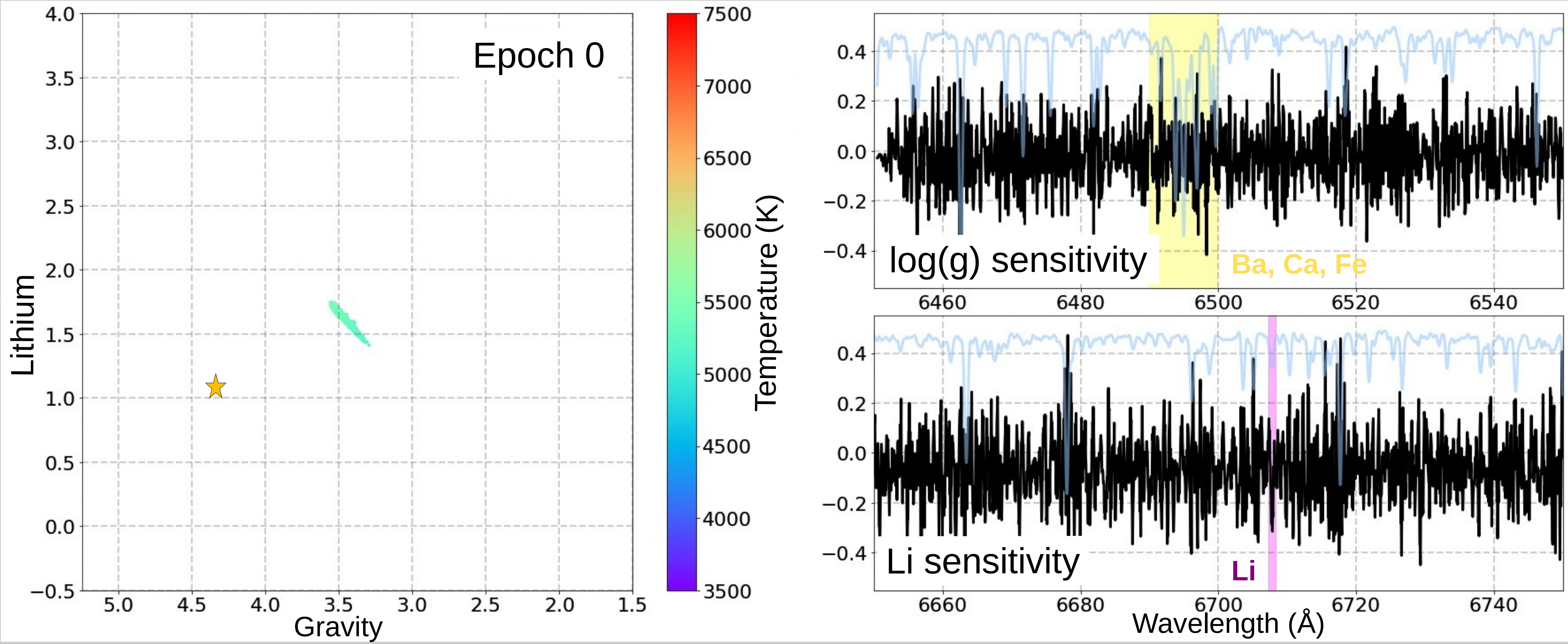
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



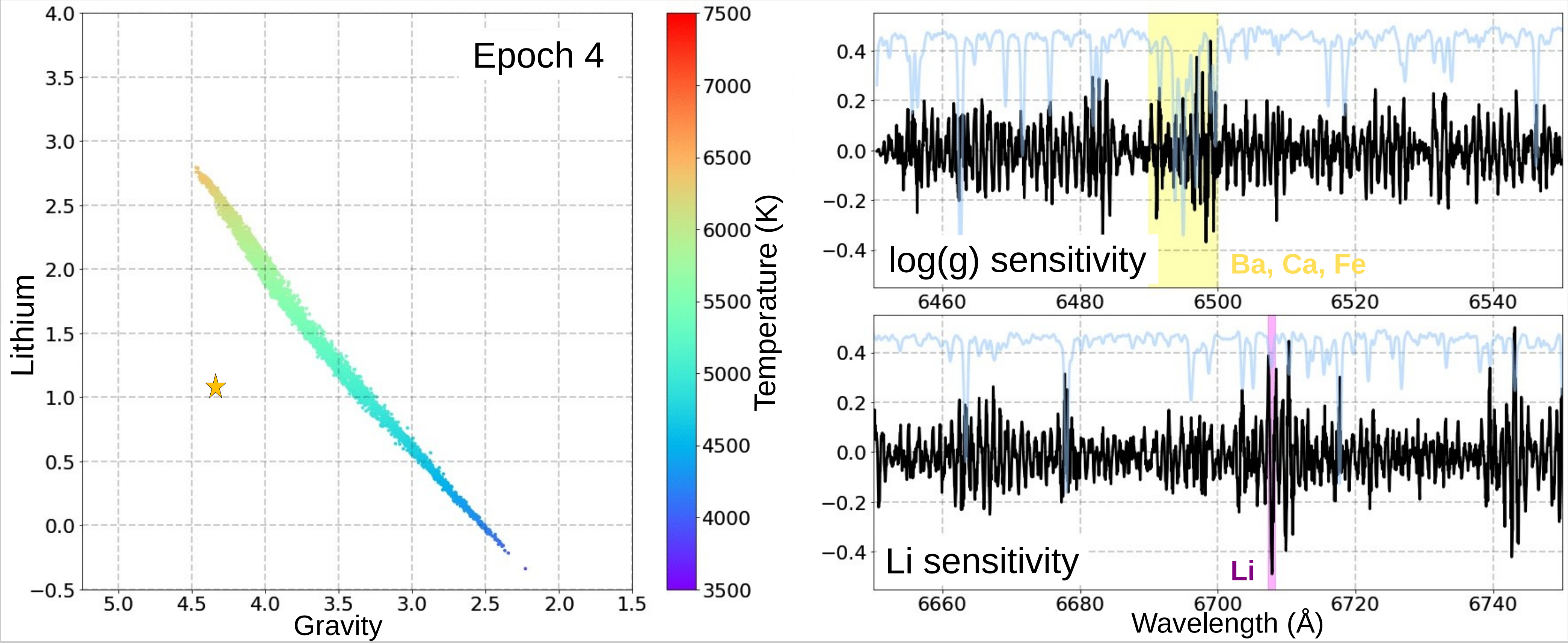
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



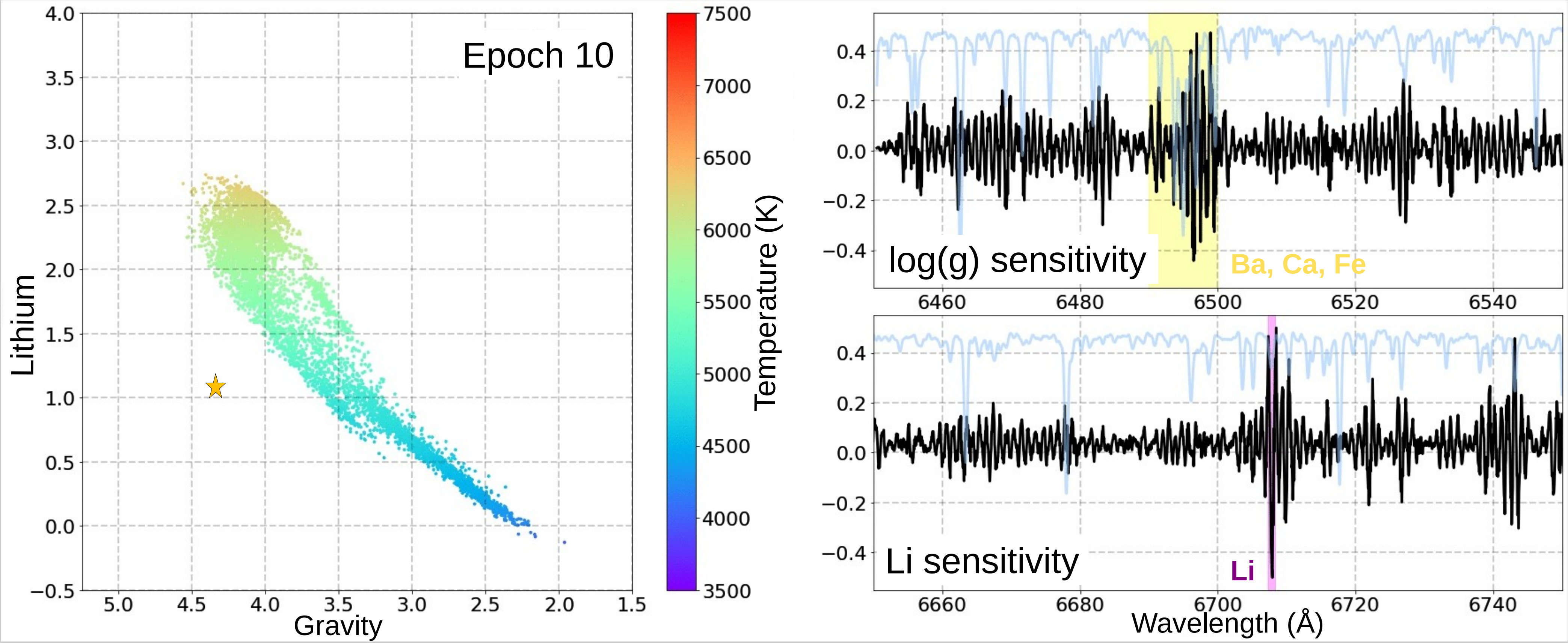
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



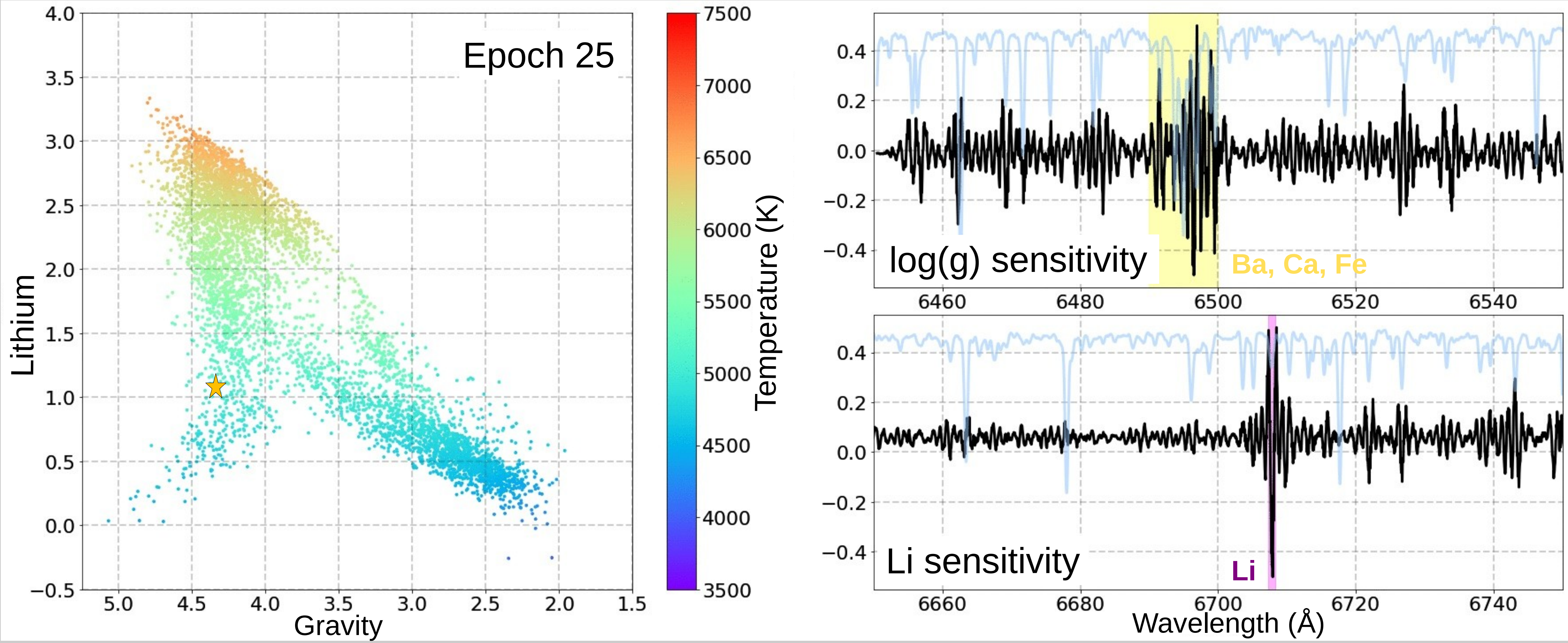
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



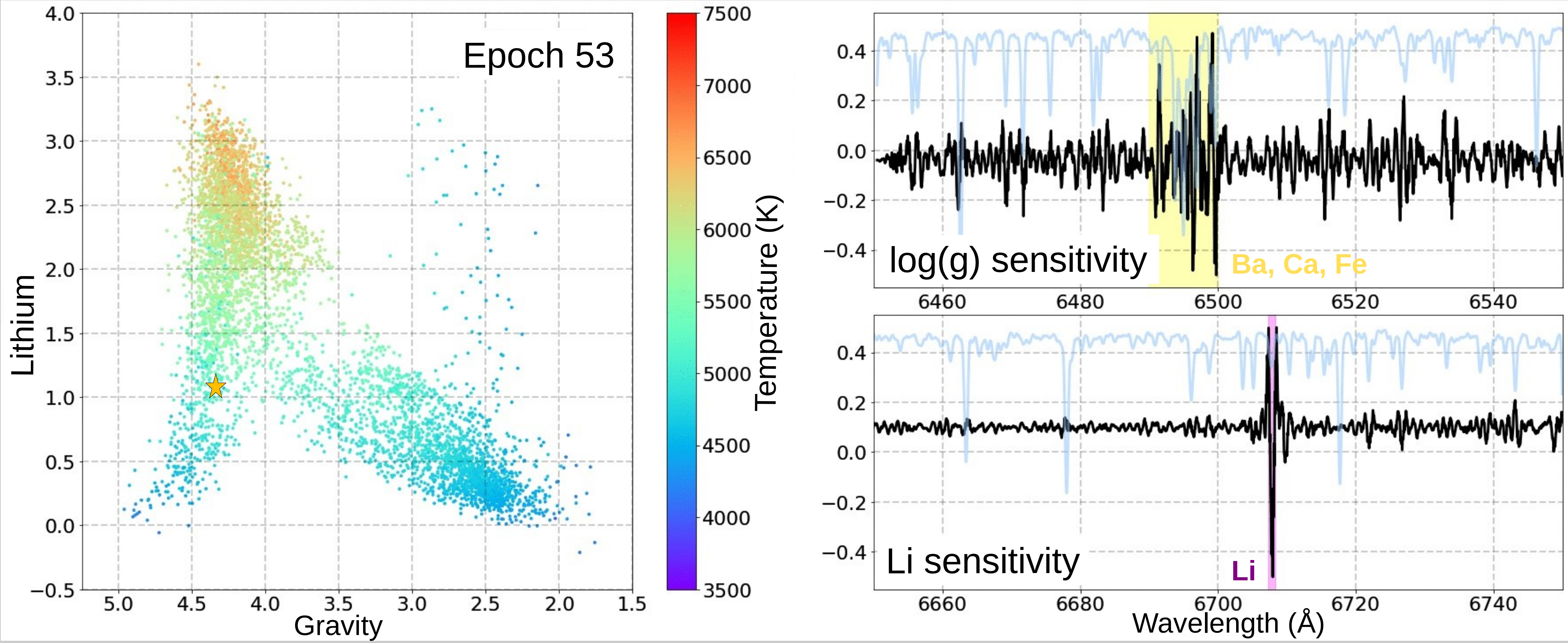
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



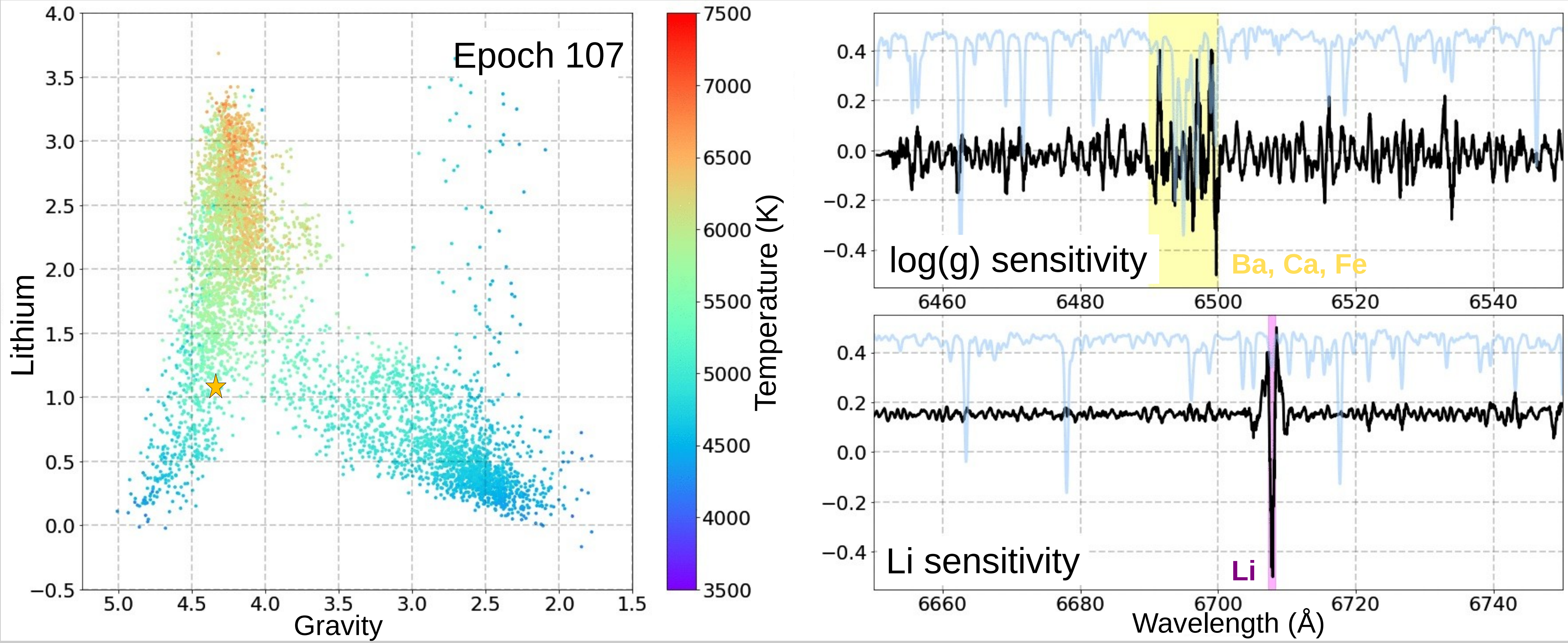
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



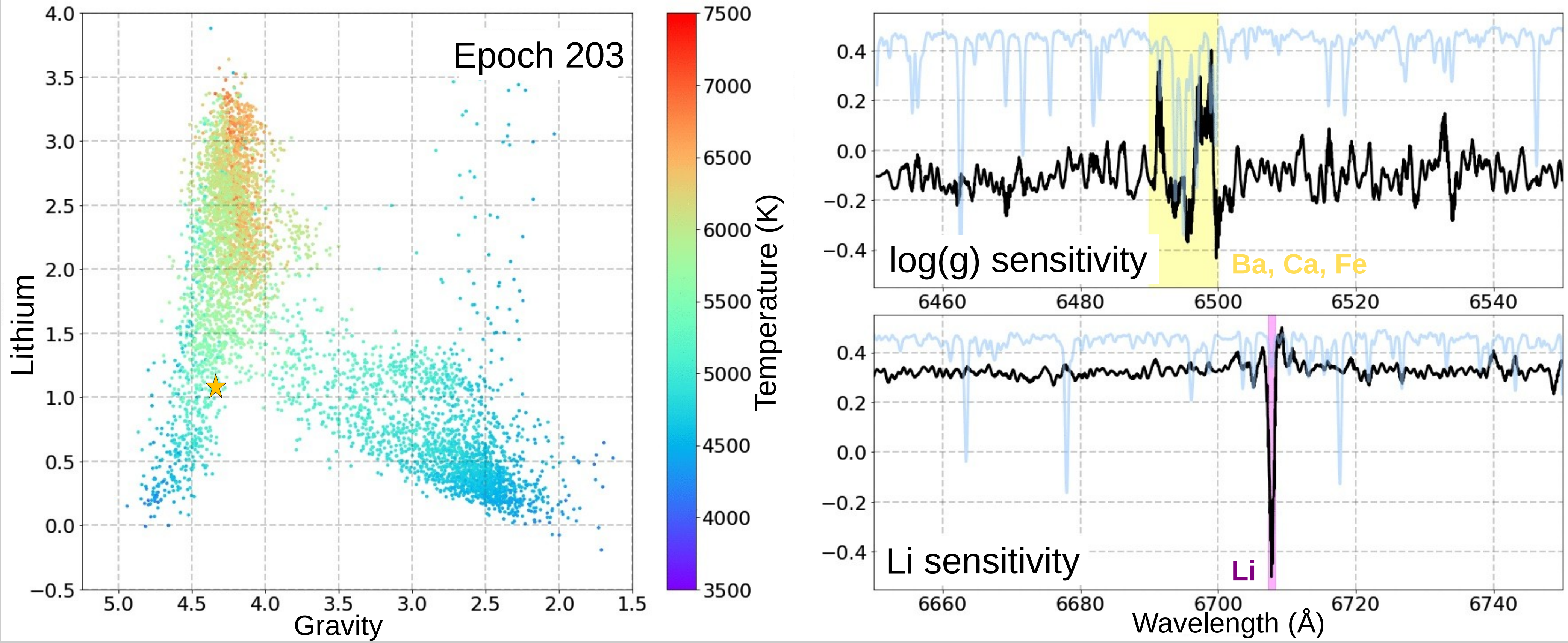
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



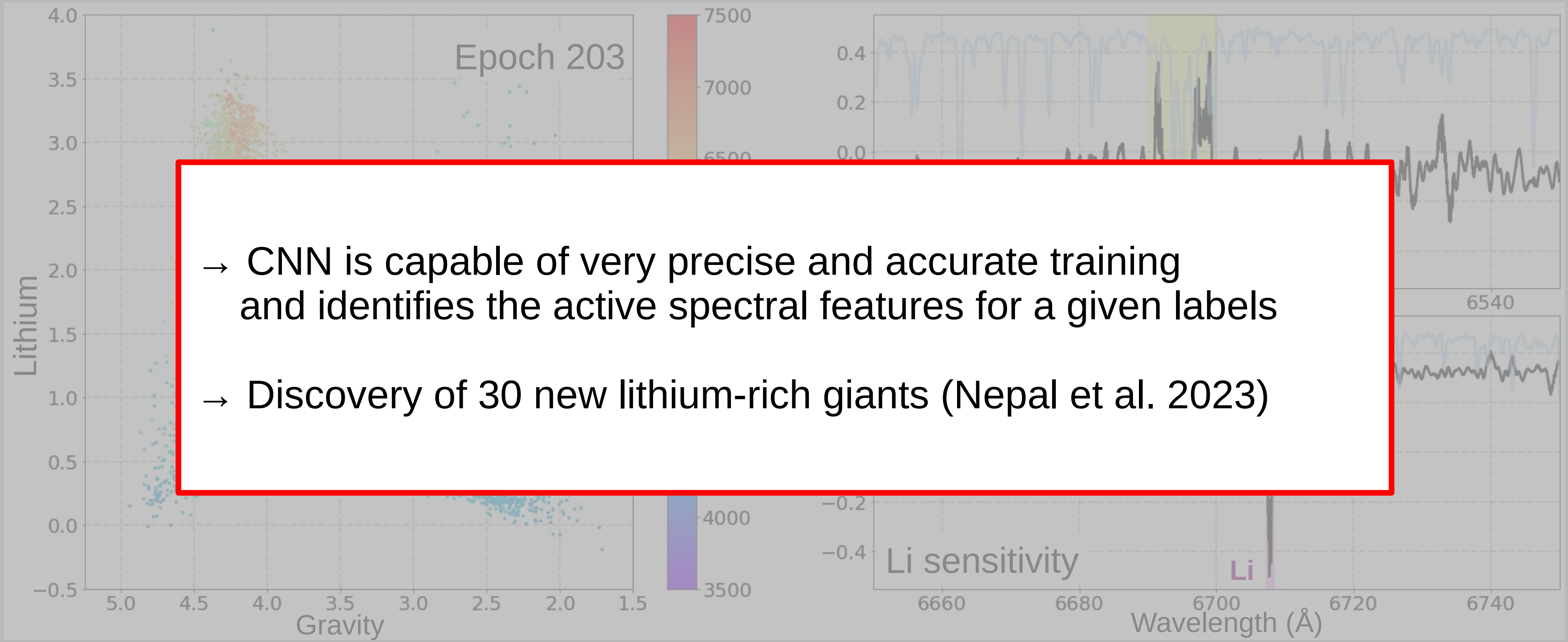
Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



Nepal et al. 2023

Measurements of lithium in Milky Way stars with CNN & GES



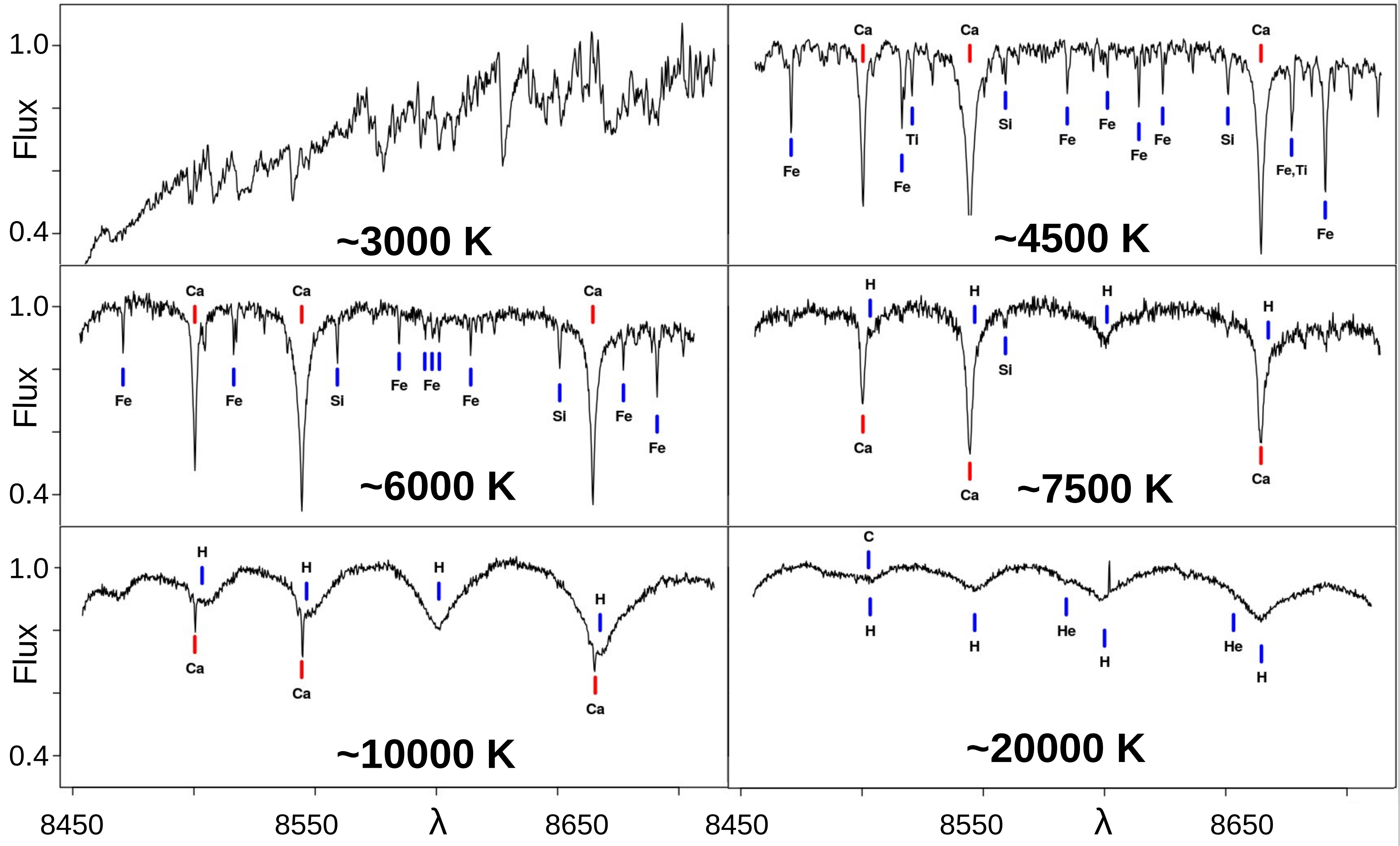
→ CNN is capable of very precise and accurate training
and identifies the active spectral features for a given labels

→ Discovery of 30 new lithium-rich giants (Nepal et al. 2023)

Nepal et al. 2023

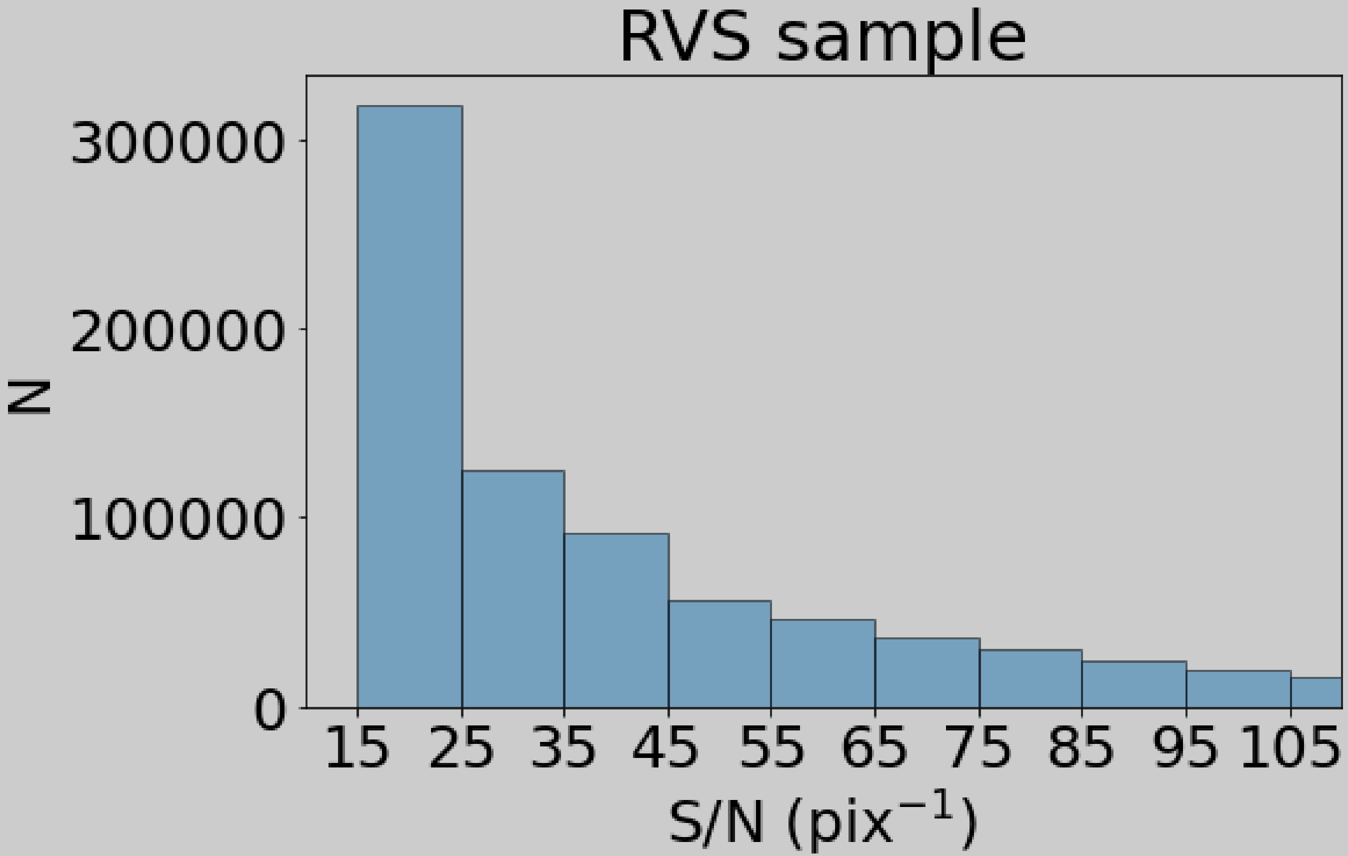
gaia RVS + CNN

Gaia DR3 June 2022: 10^6 RVS spectra, R~11500 (Katz et al. 2022)

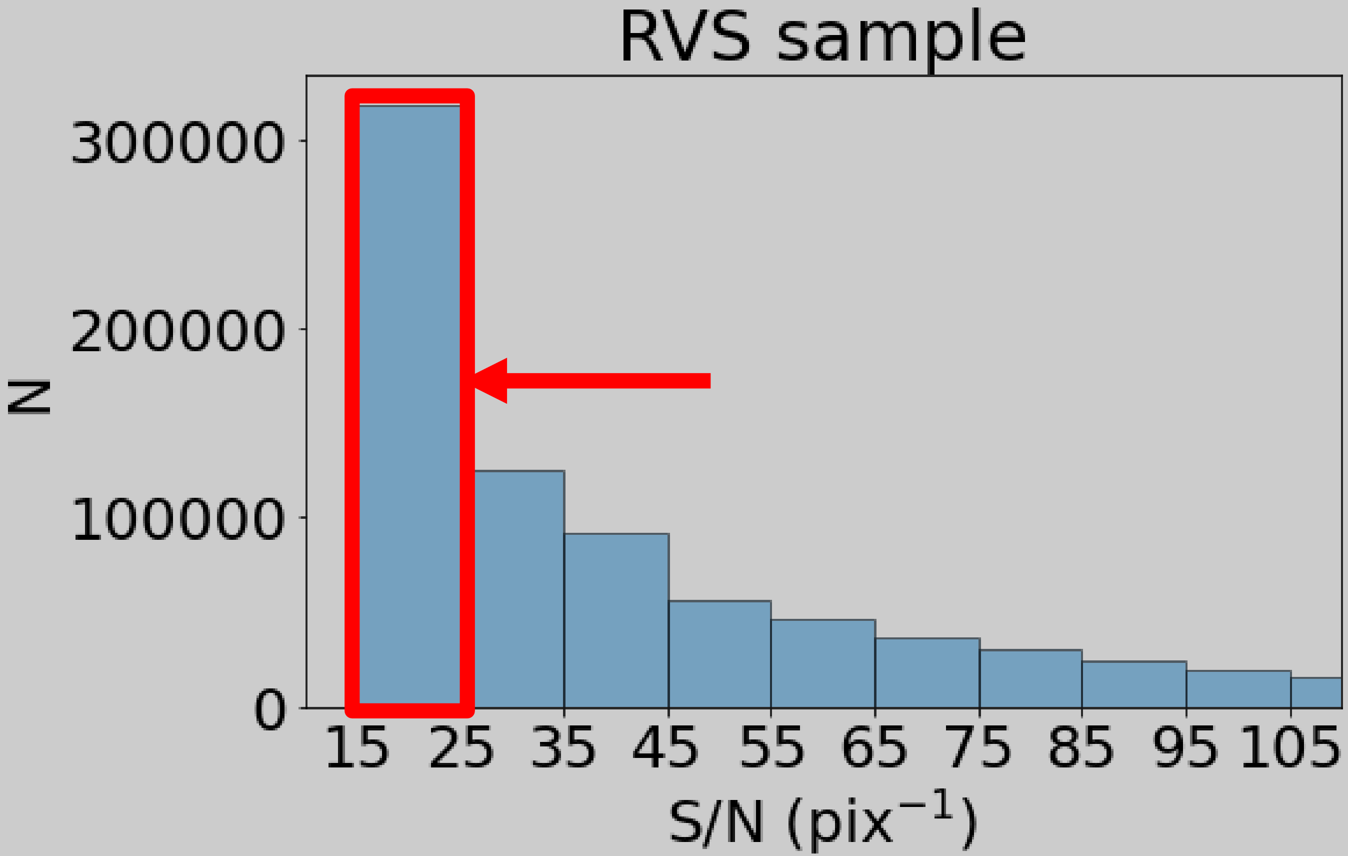


ESA/Gaia/DPAC/Observatoire de Paris-Meudon/Olivier Marchal & David Katz

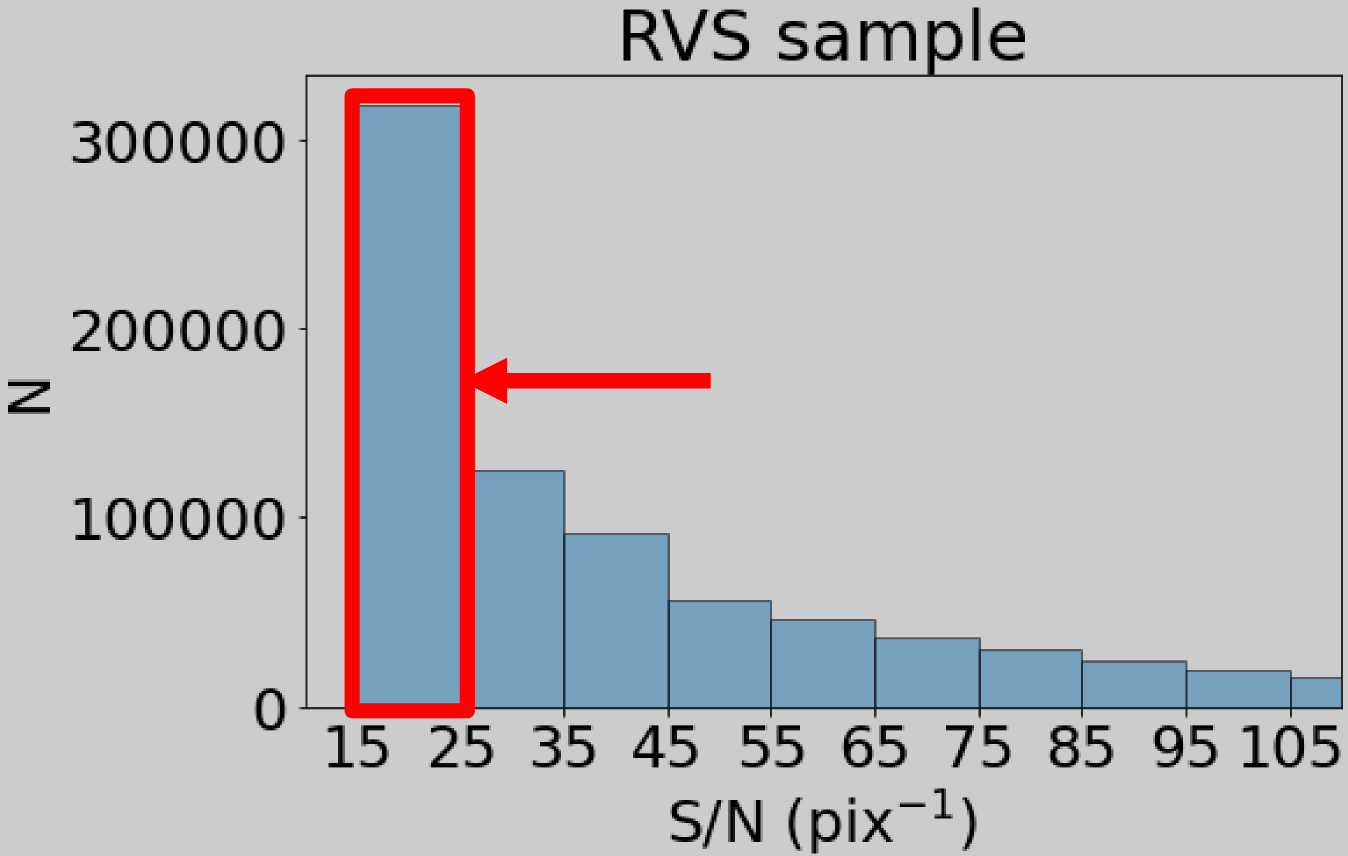
Motivations



Motivations

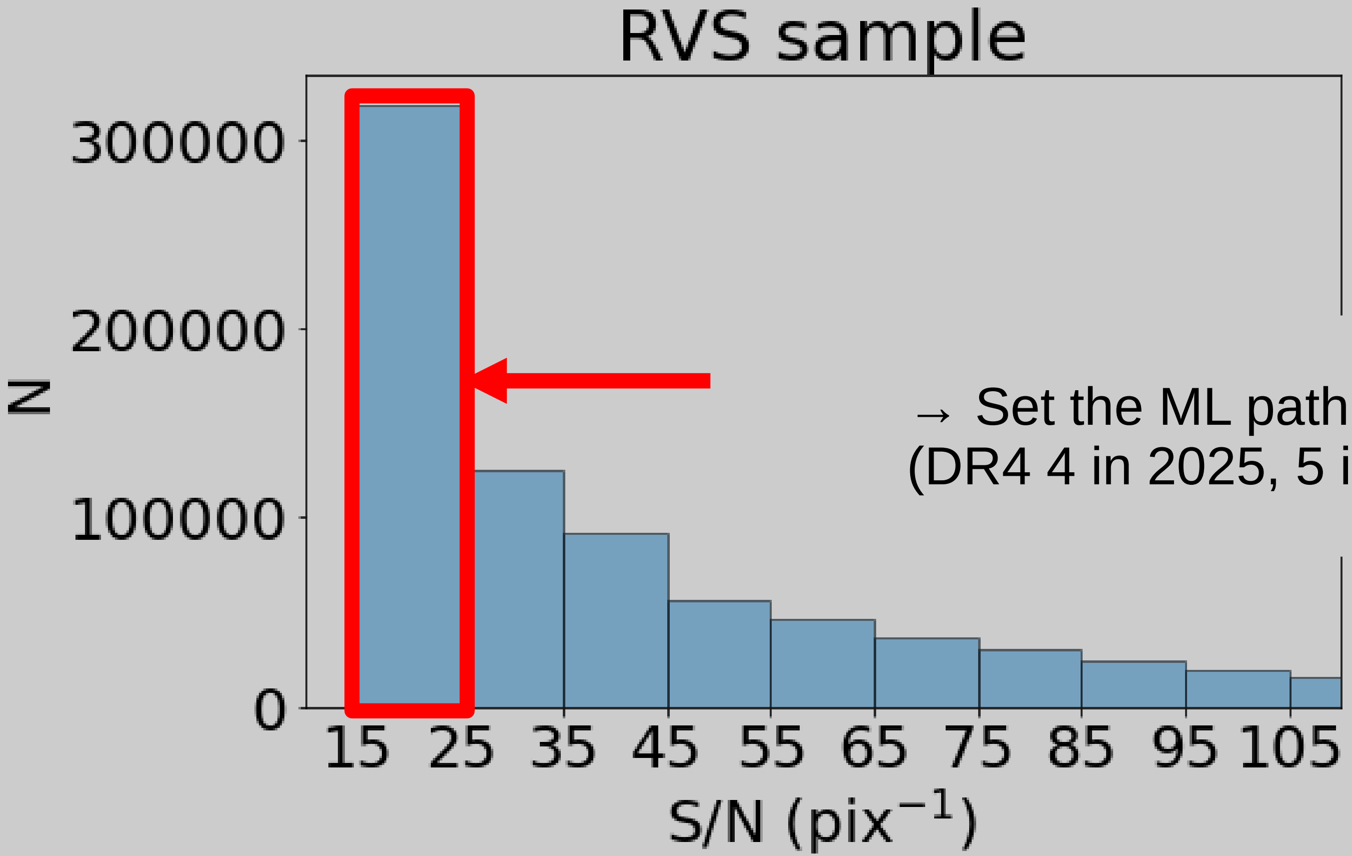


Motivations



No GSP-Spec labels with 13 “good” flags within 15<S/N<25
→ **Leverage the low-S/N RVS sample**

Motivations



→ Set the ML path for *Gaia* RVS analysis (DR4 4 in 2025, 5 in 2027)

No GSP-Spec labels with 13 “good” flags within 15<S/N<25

→ **Leverage the low-S/N RVS sample**

A&A, 682, A9 (2024)
https://doi.org/10.1051/0004-6361/202347122

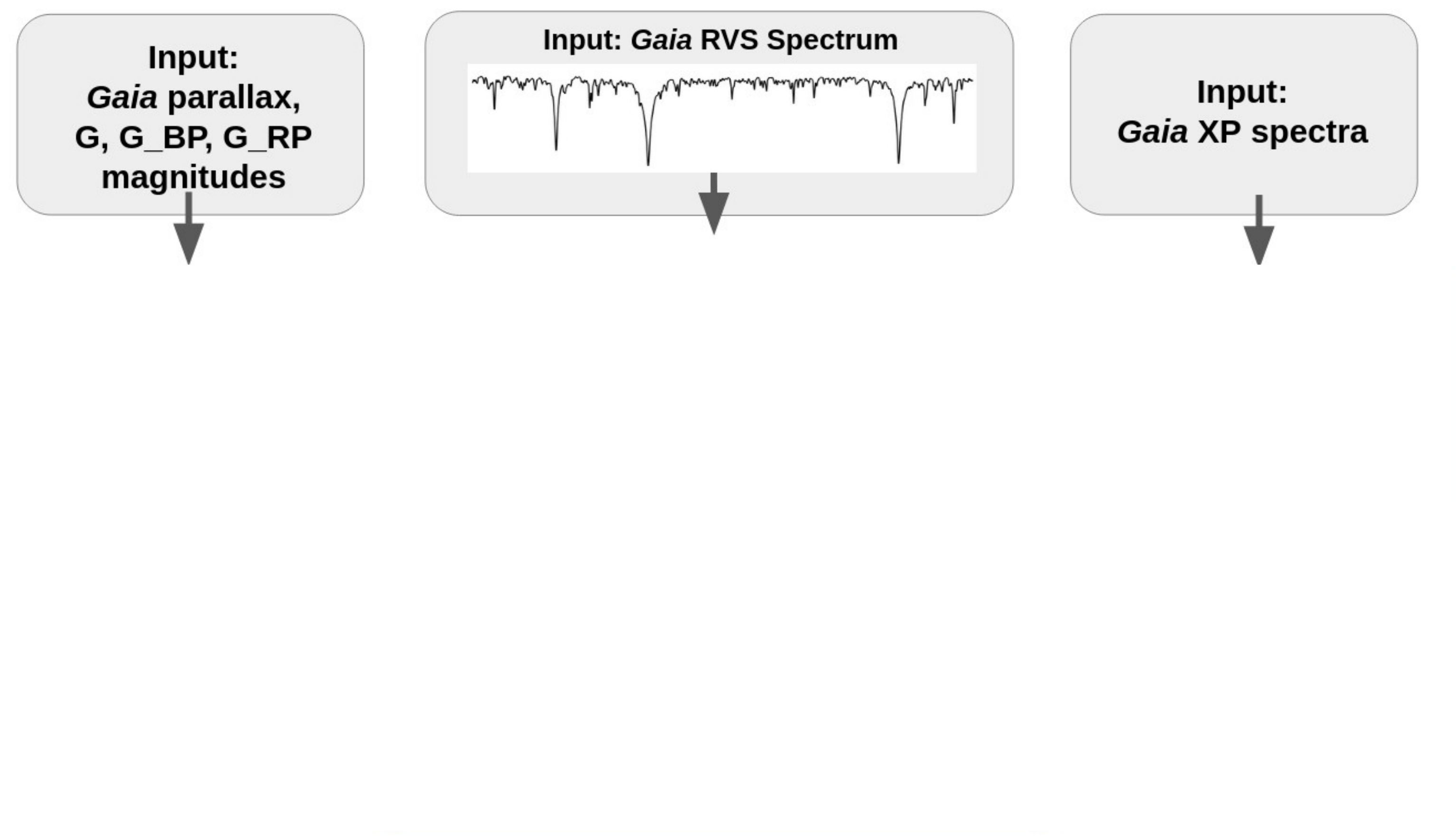
gaia

Astronomy & Astrophysics

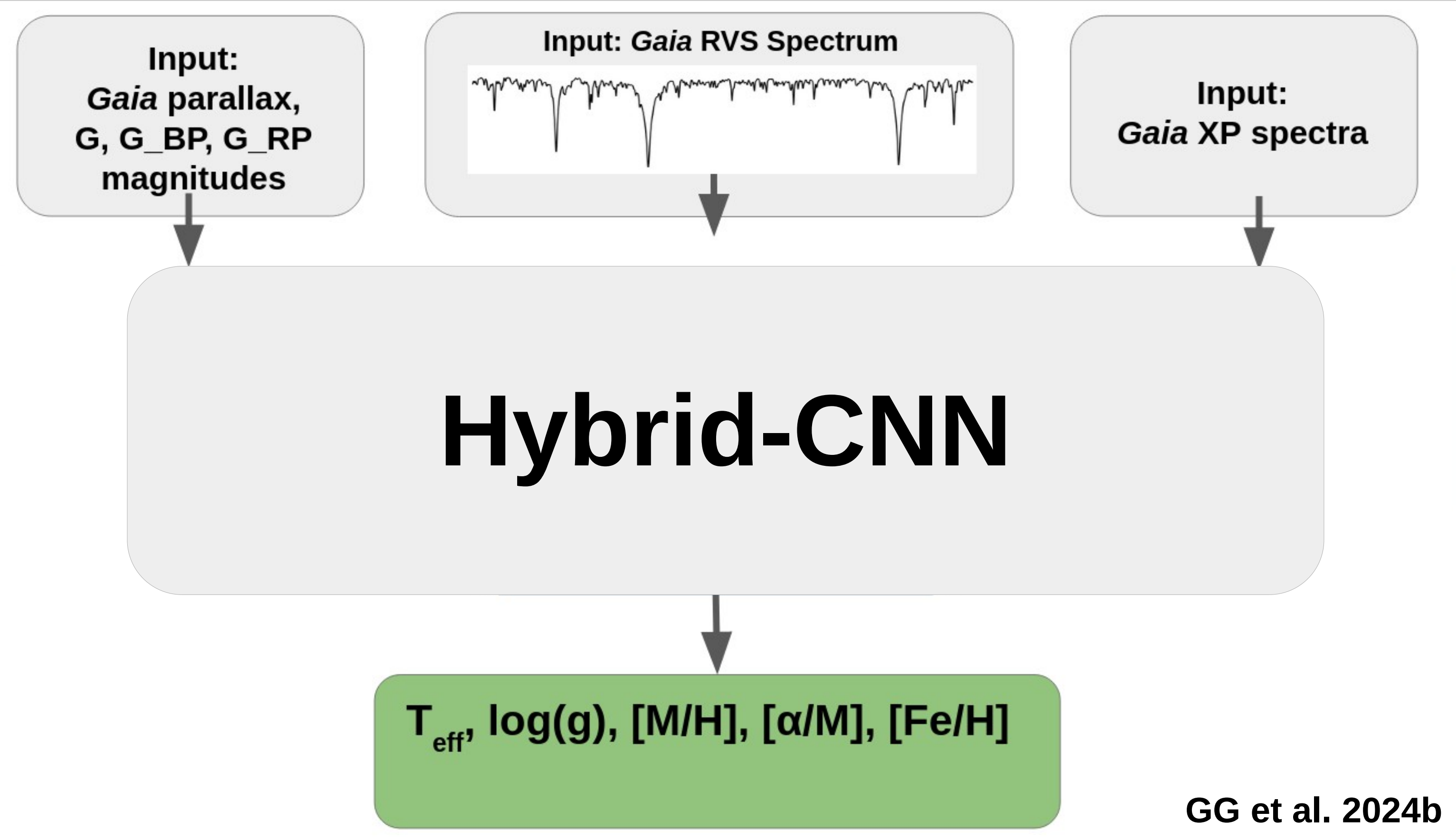
Beyond *Gaia* DR3: Tracing the [α/M] – [M/H] bimodality from the inner to the outer Milky Way disc with *Gaia*-RVS and convolutional neural networks*

G. Guiglion[1,2,3], S. Nepal[3,4], C. Chiappini[3], S. Khoperskov[3], G. Traven[5], A. B. A. Queiroz[3], M. Steinmetz[3], M. Valentini[3], Y. Fournier[3], A. Vallenari[6], K. Youakim[7], M. Bergemann[2], S. Mészáros[8,9], S. Lucatello[10,11], R. Sordo[6], S. Fabbro[12], I. Minchev[3], G. Tautvaišienė[13], Š. Mikolaitis[13], and J. Montalbán[14]

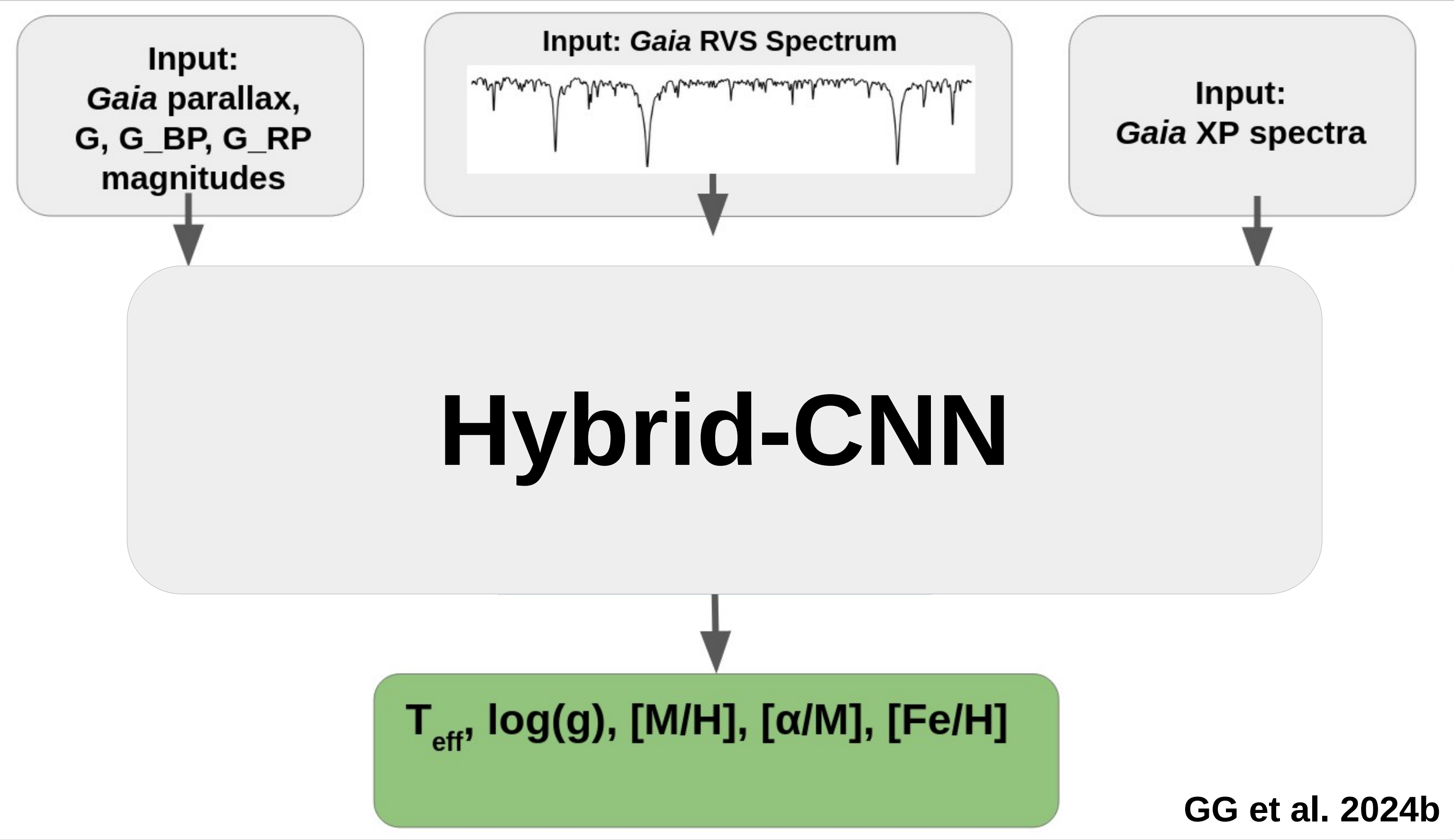
A hybrid Convolutional Neural-Network for *Gaia*-RVS analysis



A hybrid Convolutional Neural-Network for *Gaia*-RVS analysis



A hybrid Convolutional Neural-Network for *Gaia*-RVS analysis



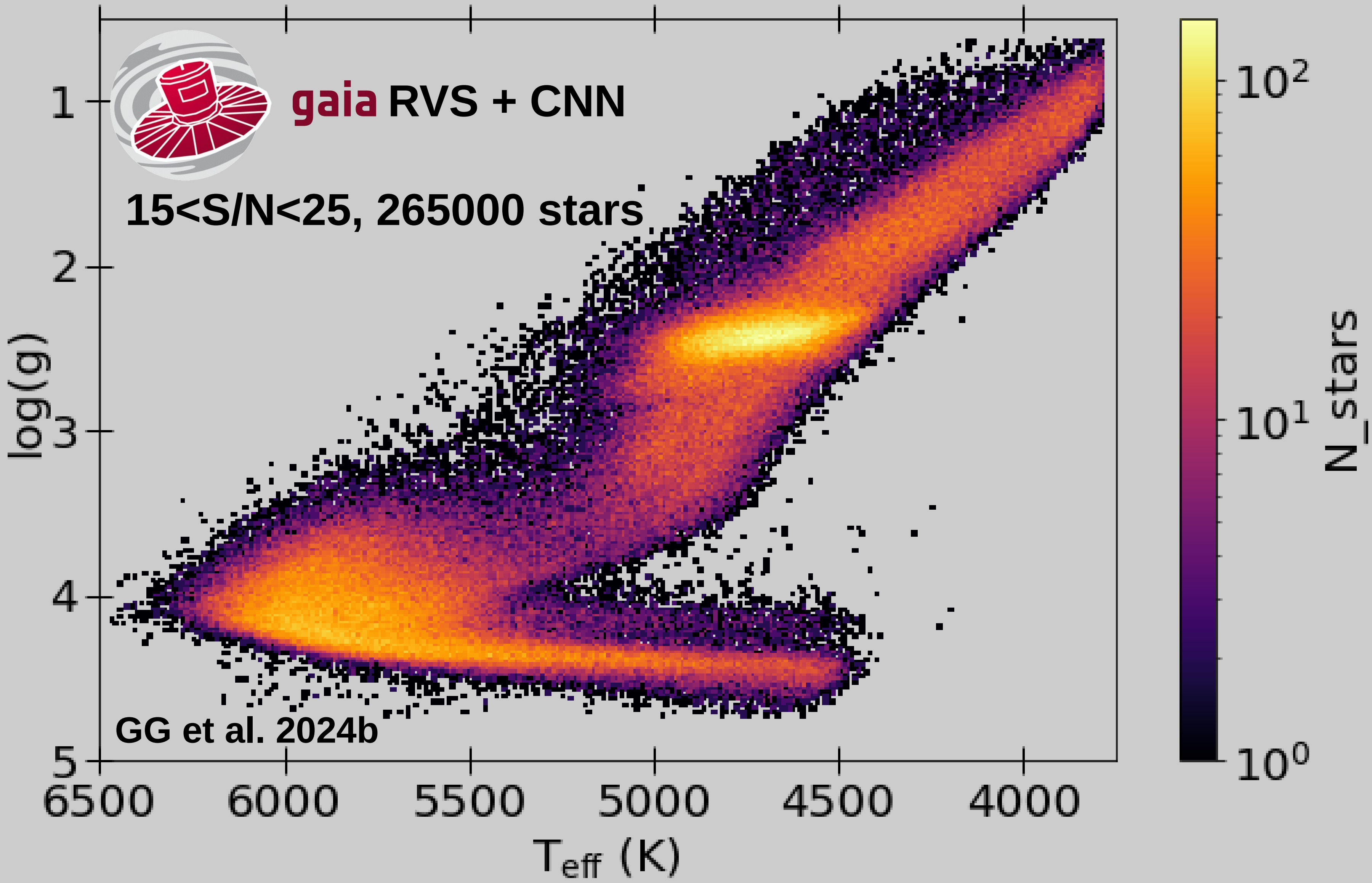
Training sample labels

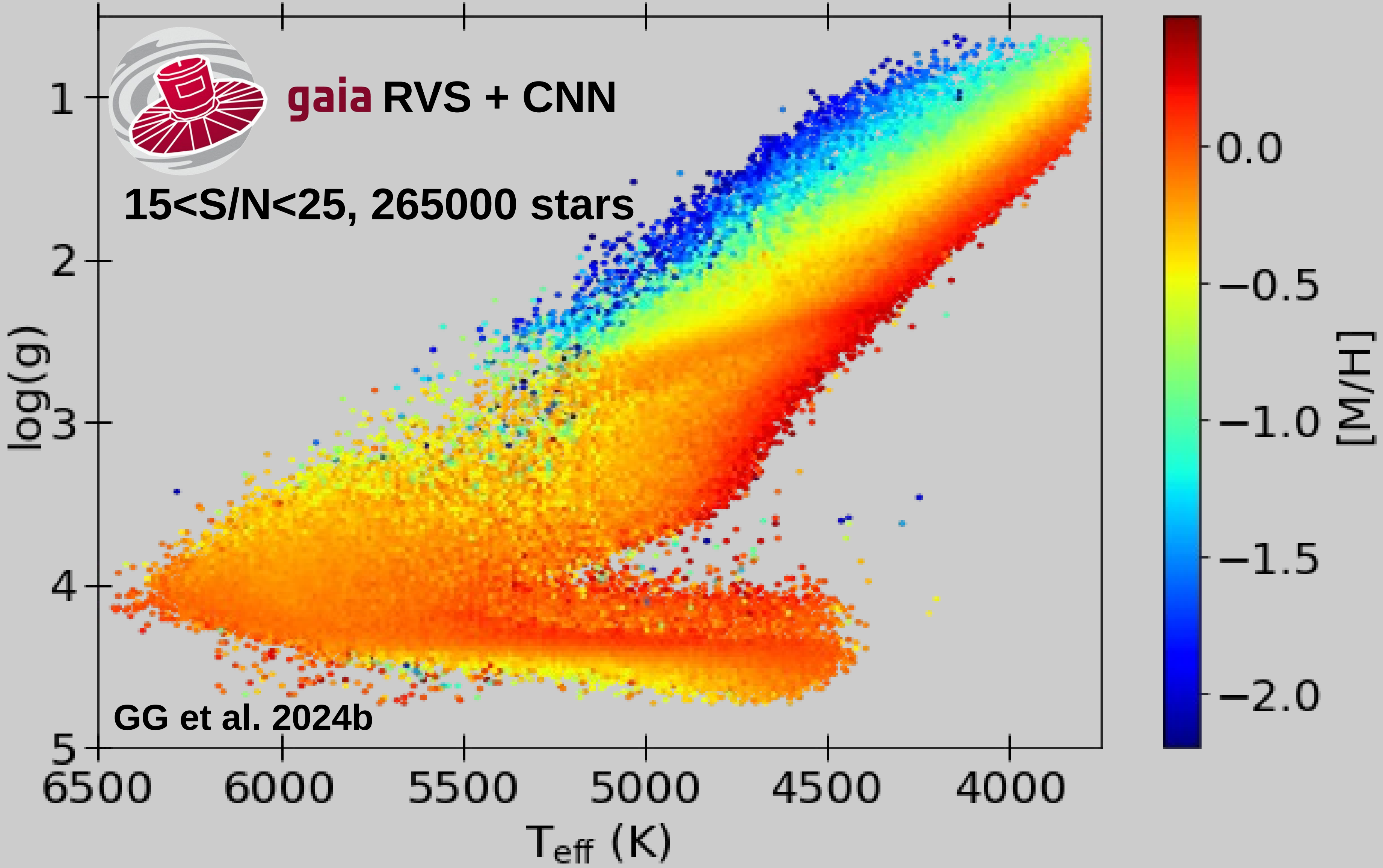
APOGEE

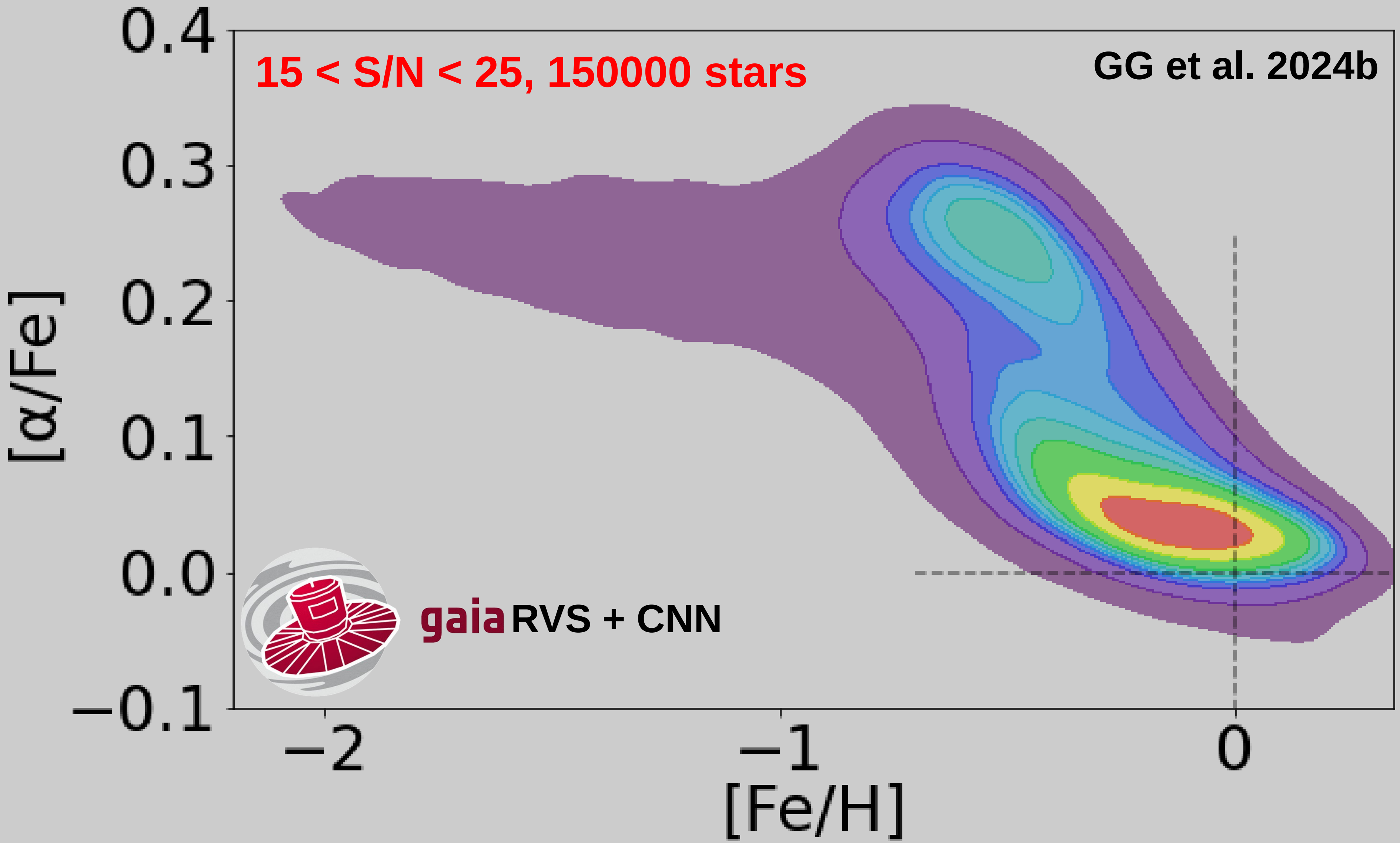
R~22000

→ Observed sample: 841300 RVS stars

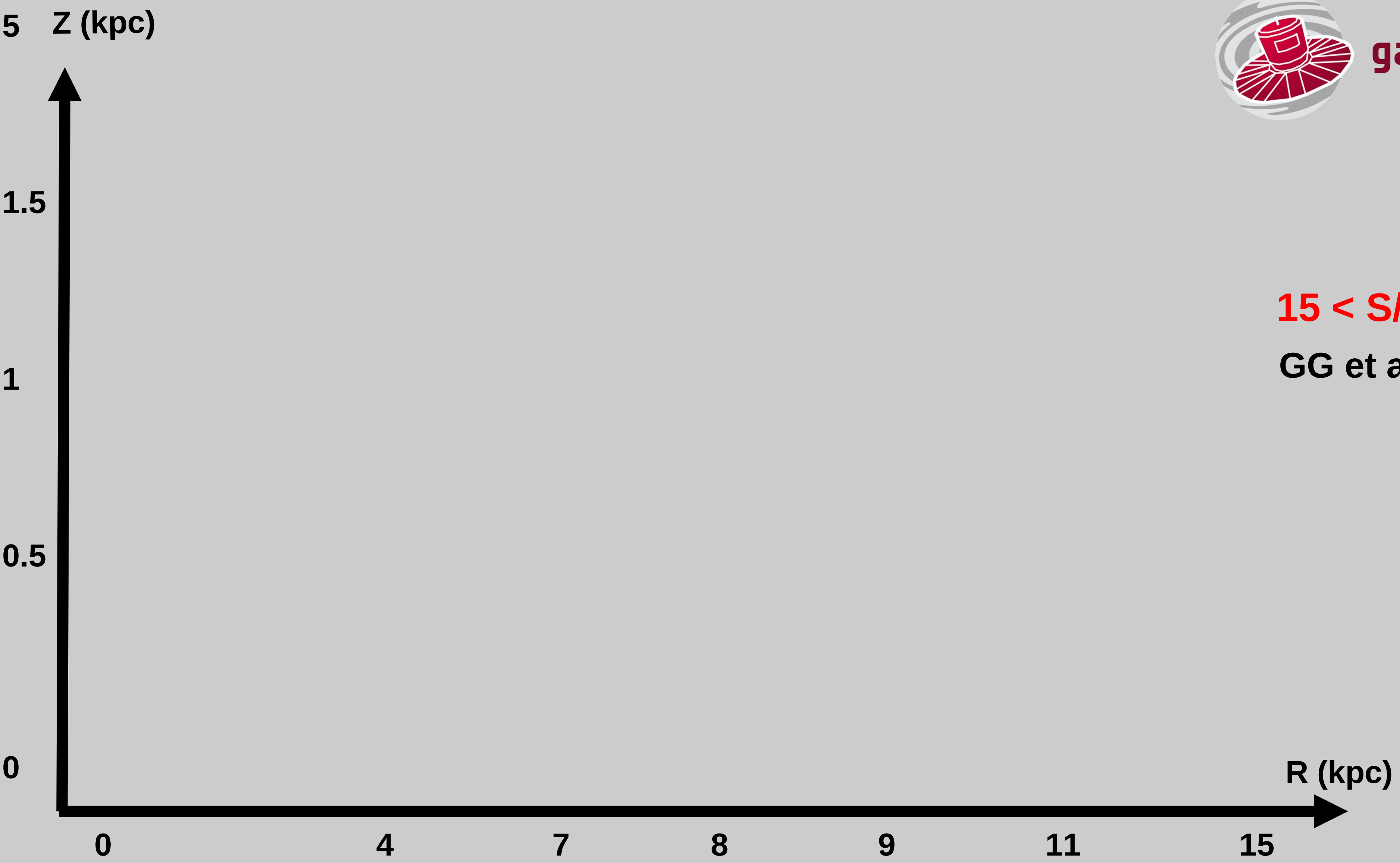
→ Prediction time 3300 stars / second


gaia RVS + CNN
15<S/N<25, 265000 stars
GG et al. 2024b
log(g)
1
2
3
4
5
6500
6000
5500
5000
4500
4000
T_eff (K)
N_stars
10^0
10^1
10^2


gaia RVS + CNN
15<S/N<25, 265000 stars
GG et al. 2024b
log(g)
1
2
3
4
5
6500
6000
5500
5000
4500
4000
T_{eff} (K)
[M/H]
0.0
−0.5
−1.0
−1.5
−2.0

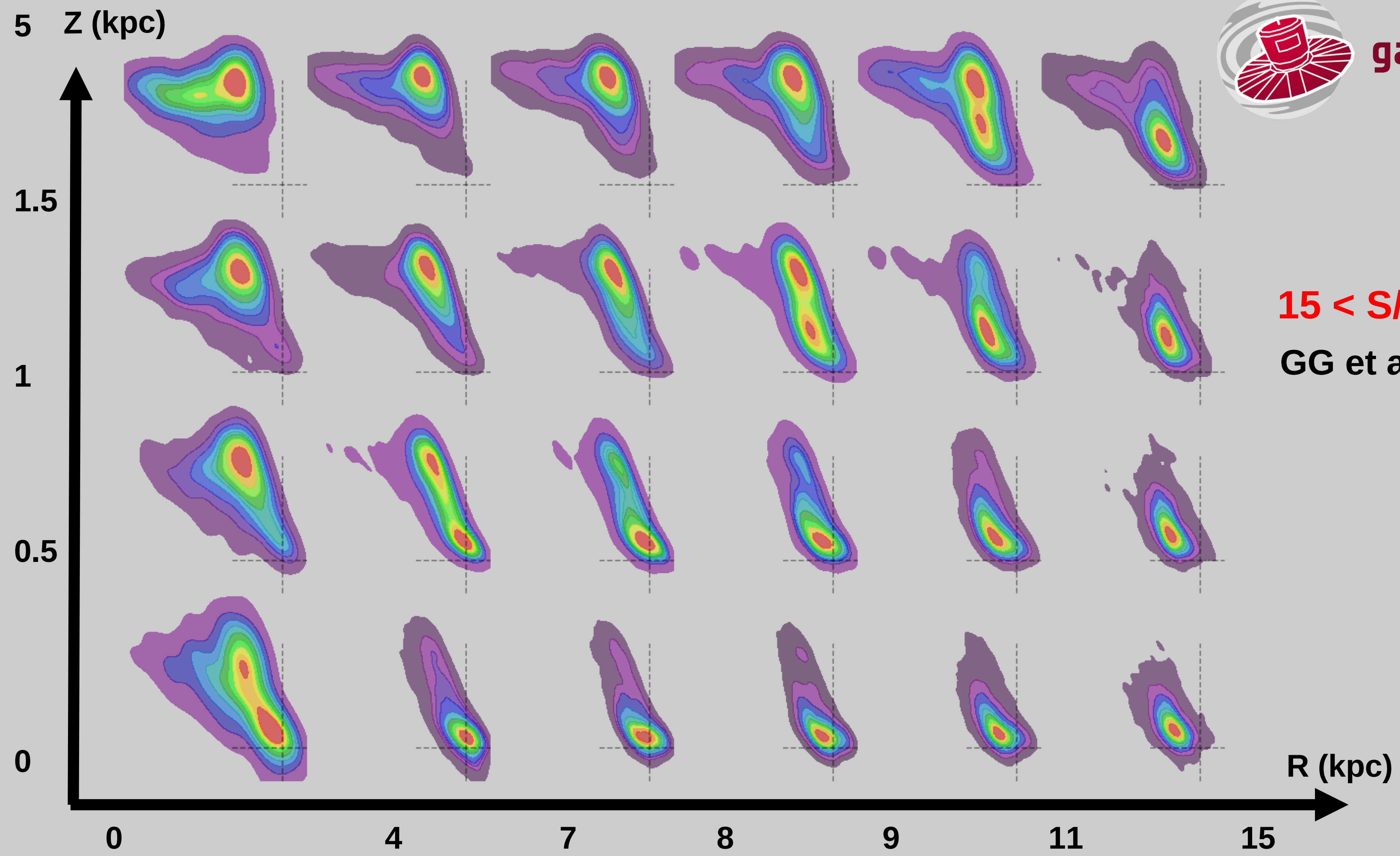

15 < S/N < 25, 150000 stars
GG et al. 2024b
gaia RVS + CNN
[α/Fe]
[Fe/H]
0.4
0.3
0.2
0.1
0.0
−0.1
−2
−1
0

gaia RVS
+
CNN



15 < S/N < 25

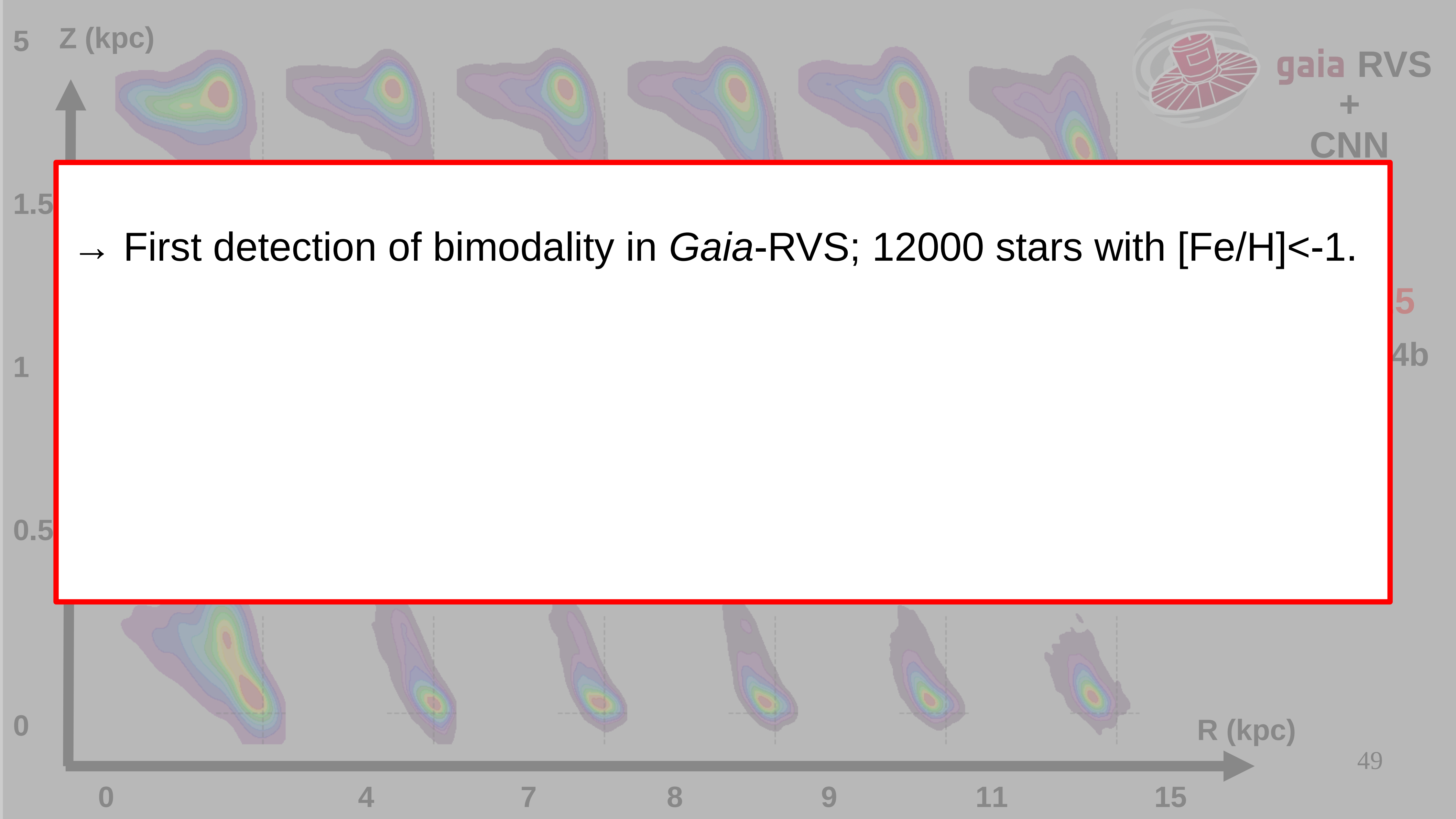
GG et al. 2024b



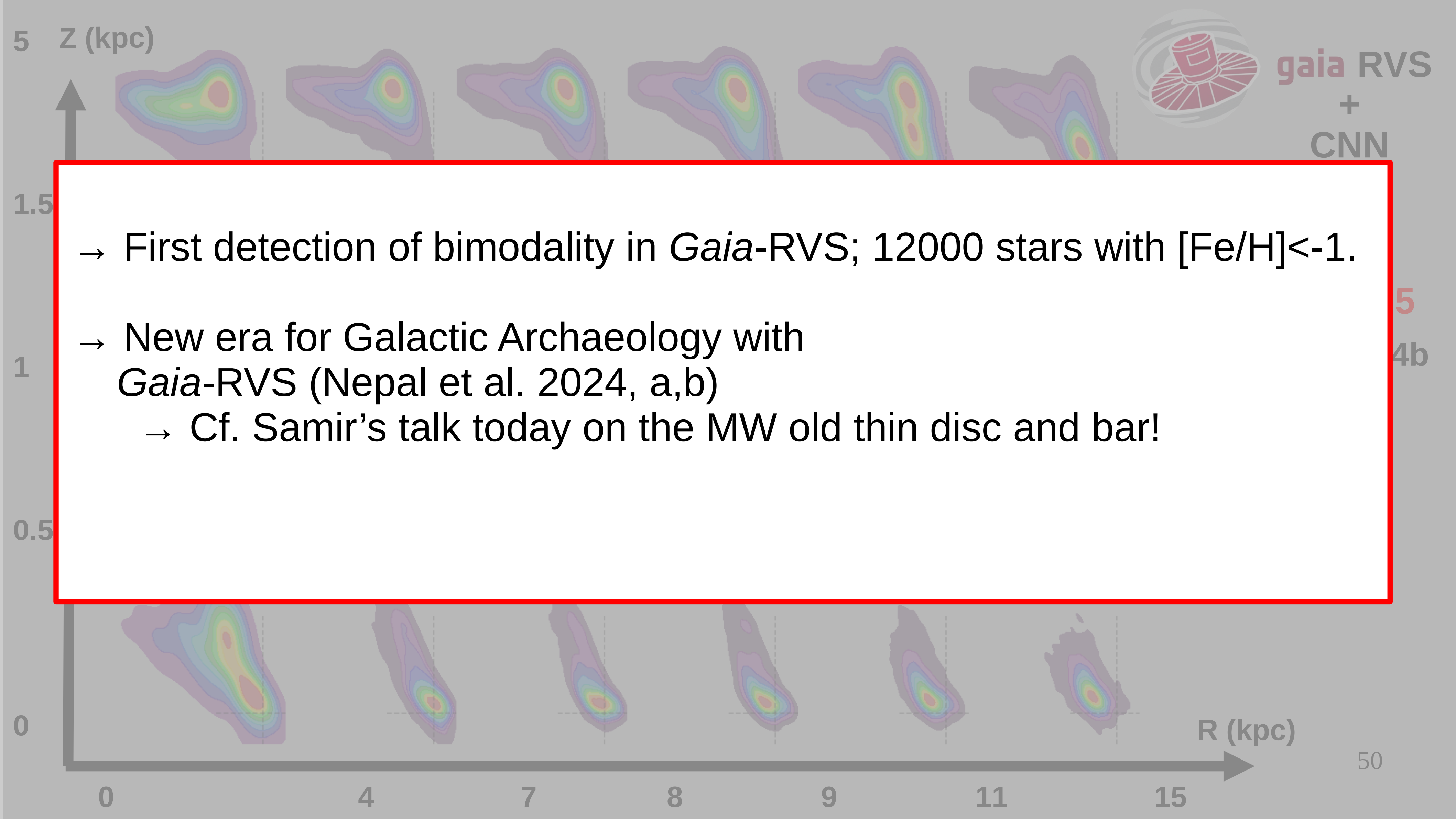
gaia RVS + CNN

15 < S/N < 25

GG et al. 2024b



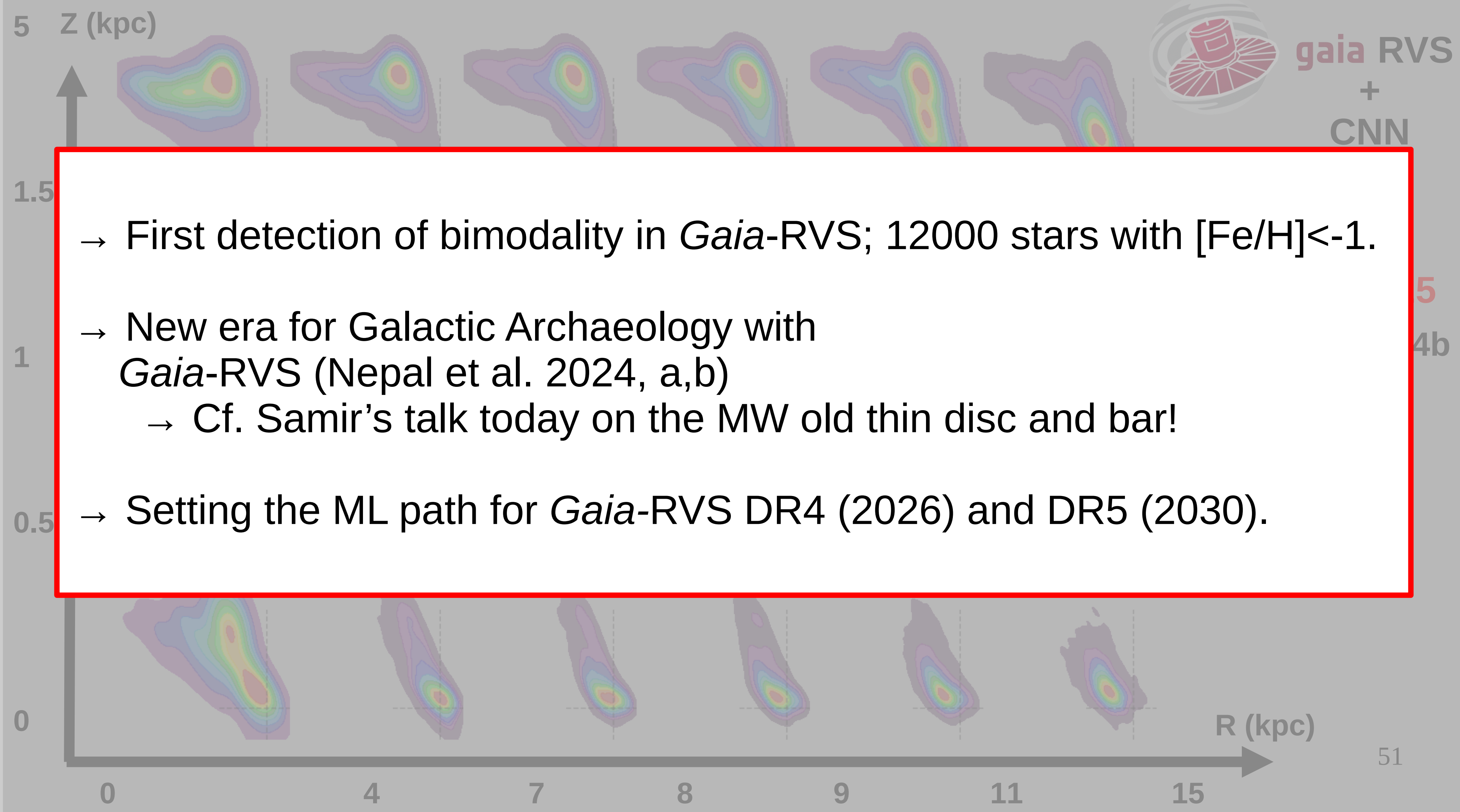
→ First detection of bimodality in *Gaia*-RVS; 12000 stars with [Fe/H]<-1.



→ First detection of bimodality in *Gaia*-RVS; 12000 stars with [Fe/H]<-1.

→ New era for Galactic Archaeology with *Gaia*-RVS (Nepal et al. 2024, a,b)

→ Cf. Samir’s talk today on the MW old thin disc and bar!



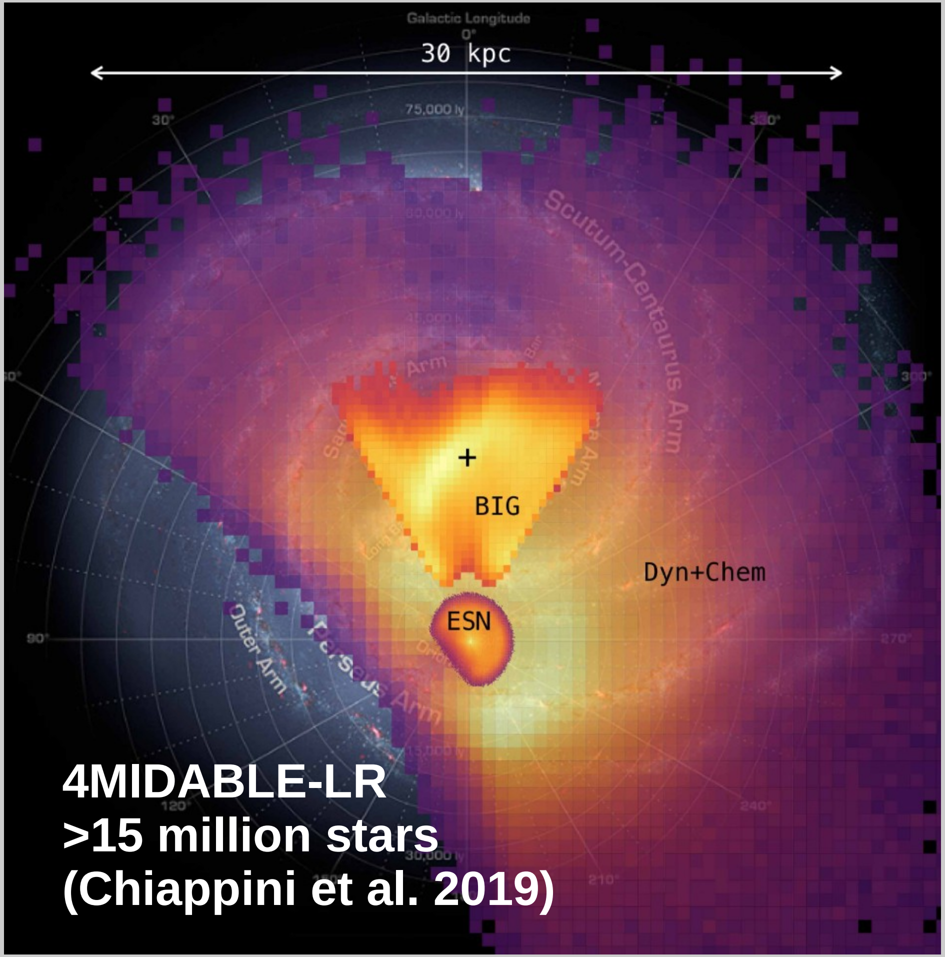
→ First detection of bimodality in *Gaia*-RVS; 12000 stars with [Fe/H]<-1.

→ New era for Galactic Archaeology with *Gaia*-RVS (Nepal et al. 2024, a,b)

  → Cf. Samir's talk today on the MW old thin disc and bar!

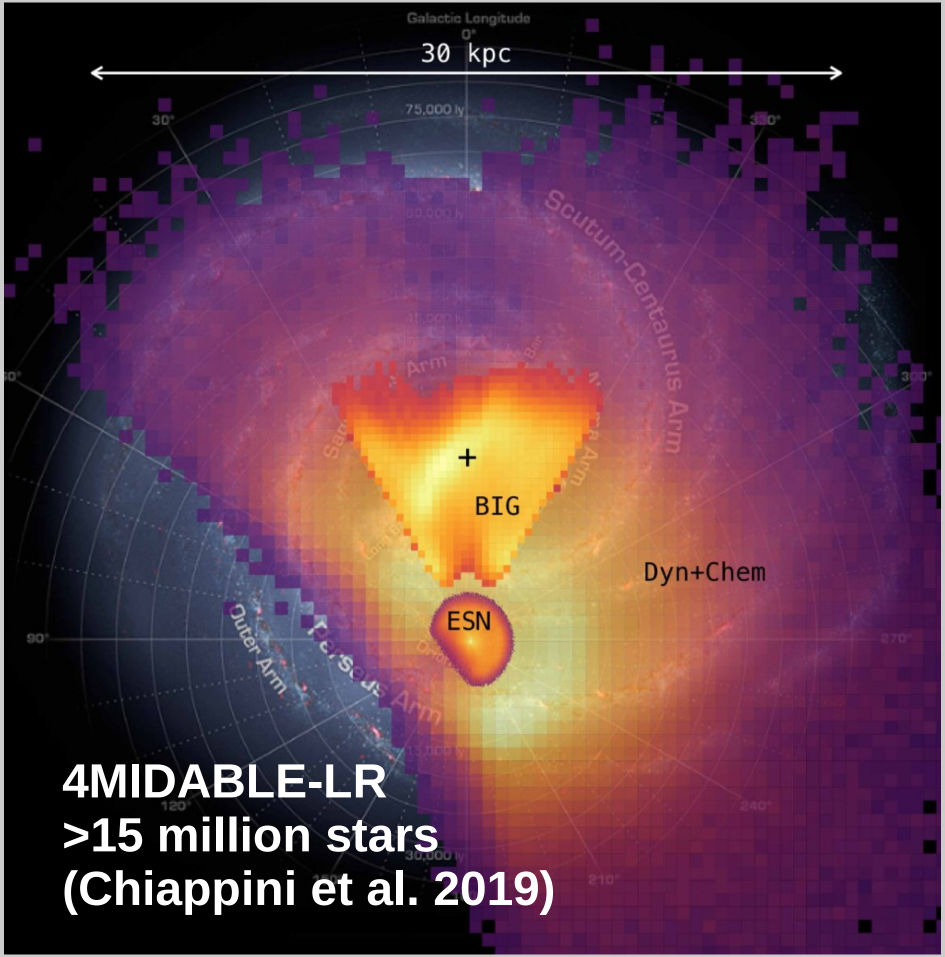
→ Setting the ML path for *Gaia*-RVS DR4 (2026) and DR5 (2030).

Why using CNN on low-res spectra ?



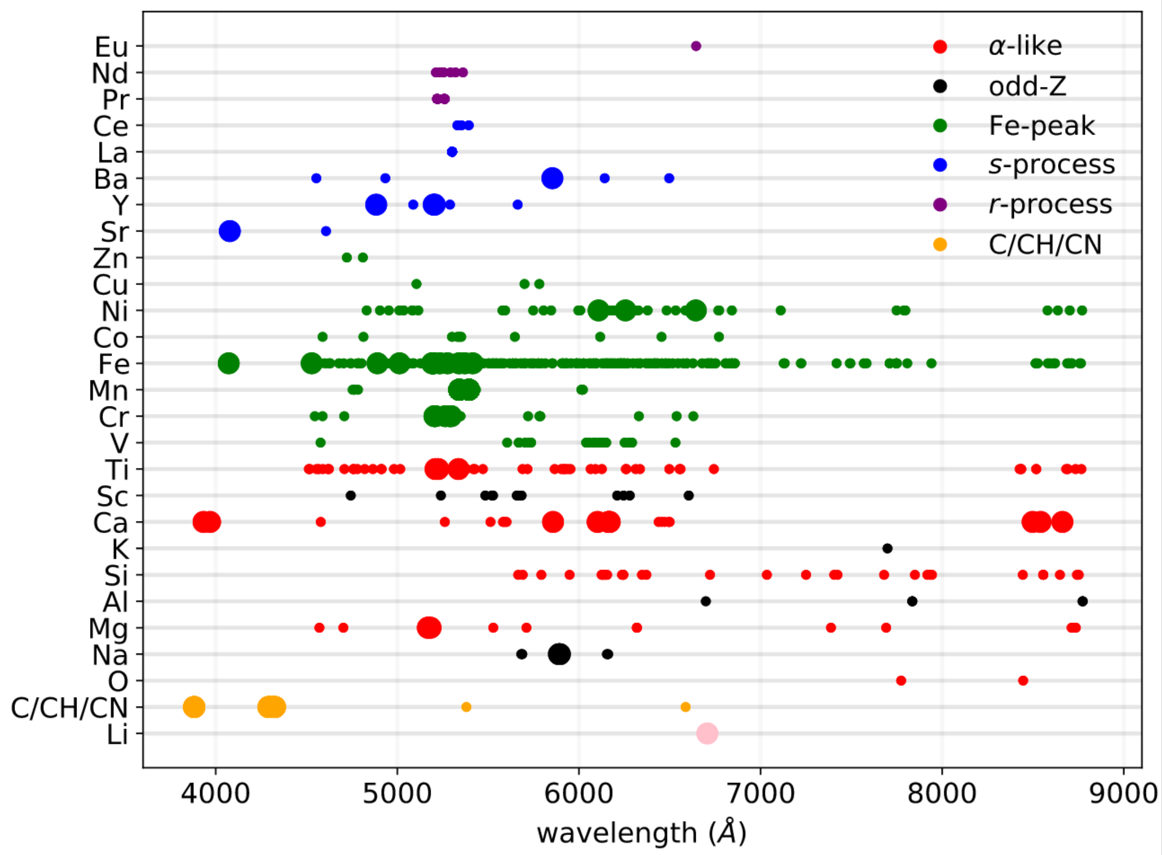
4MIDABLE-LR ESO proposal 2020

Why using CNN on low-res spectra ?



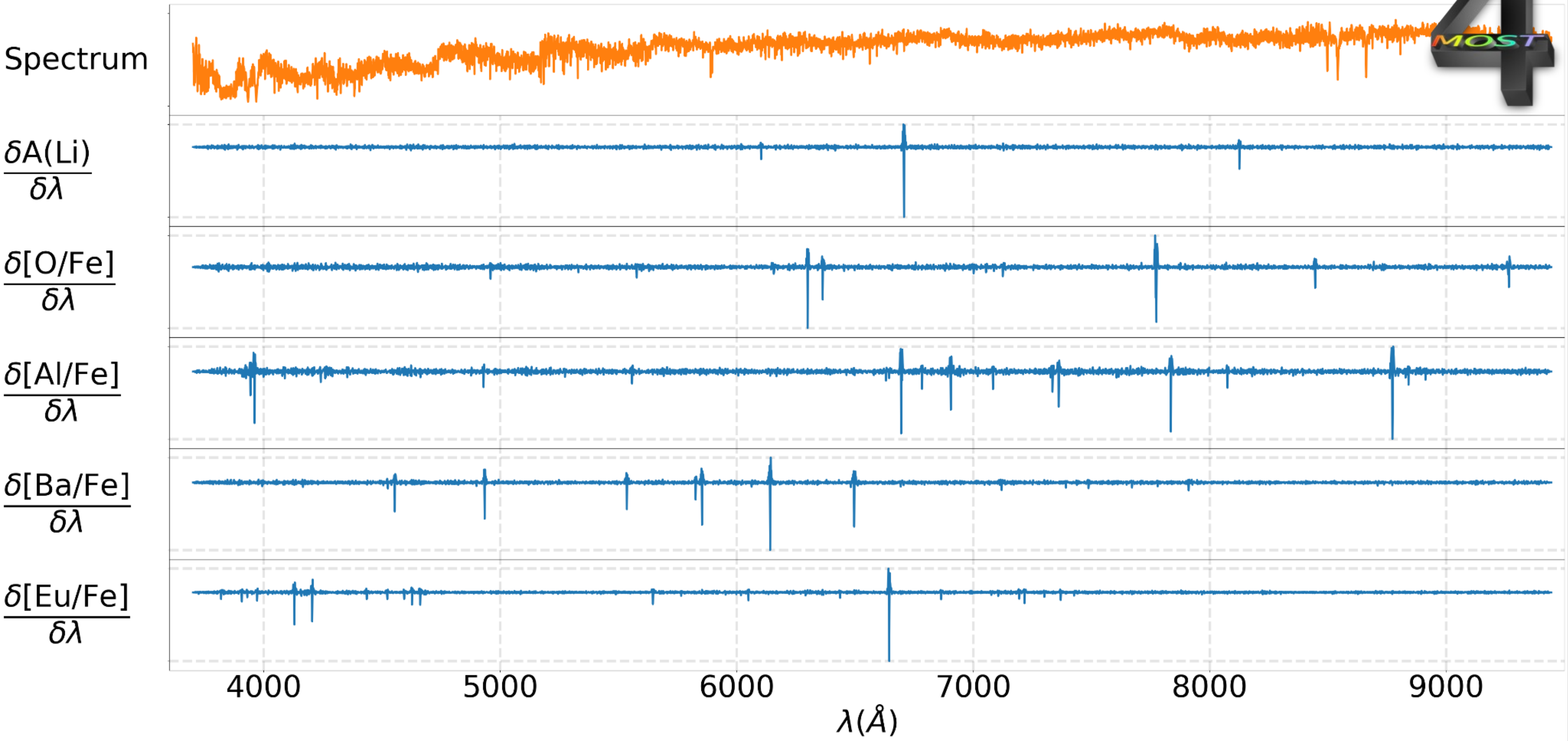
4MIDABLE-LR ESO proposal 2020

>20 elements to be measured at R=5000

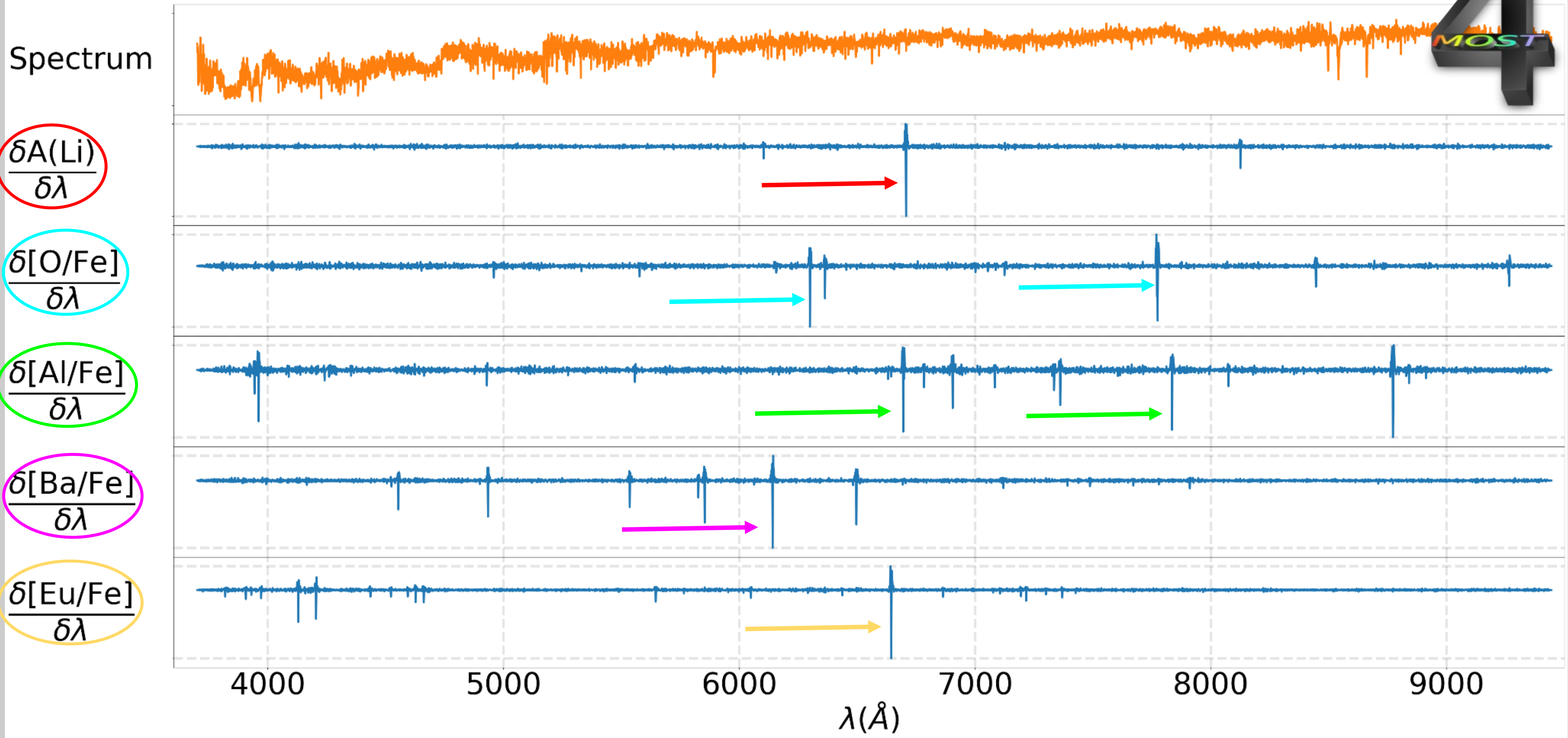


4MIDABLE-LR ESO proposal 2020

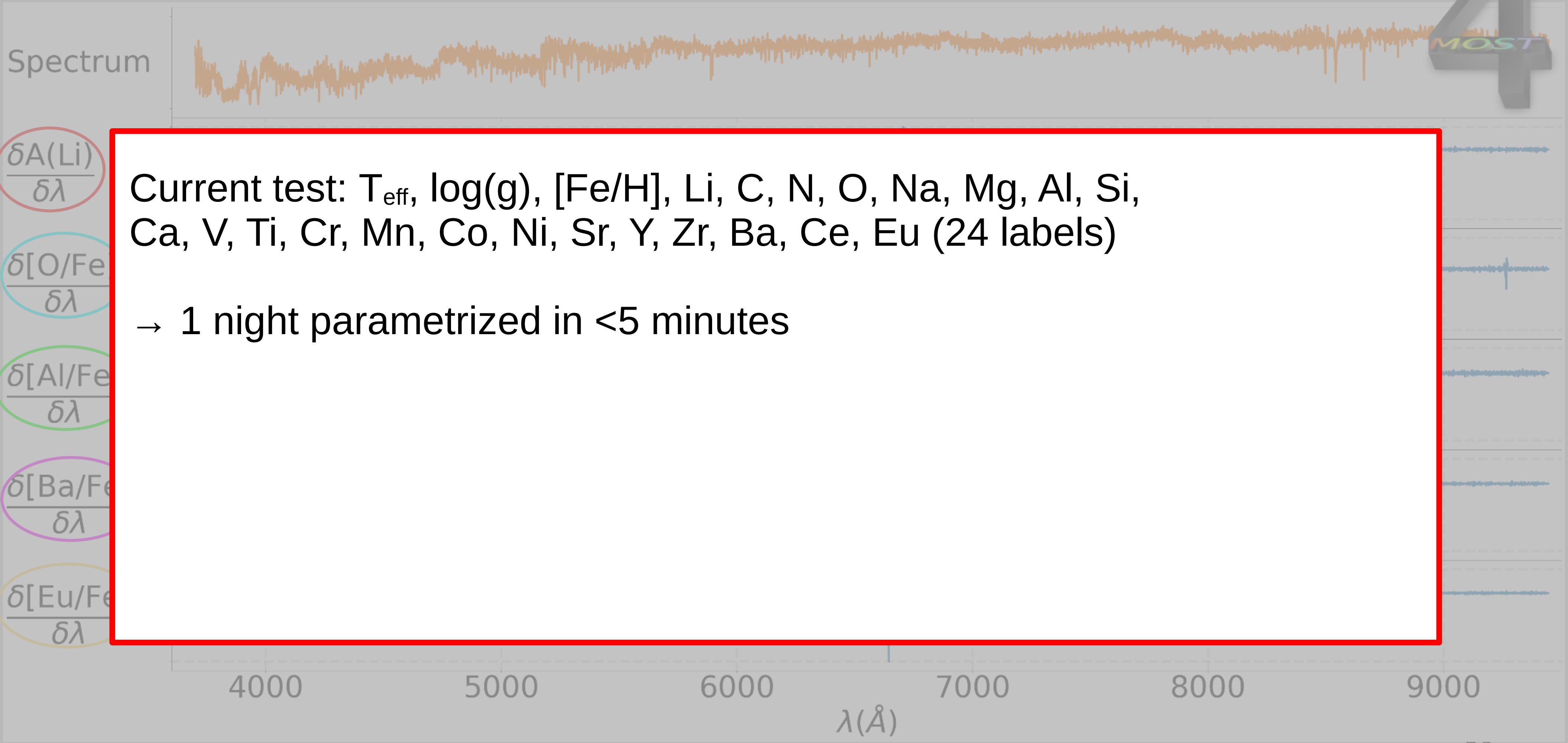
Developing CNN for 4MOST



Developing CNN for 4MOST



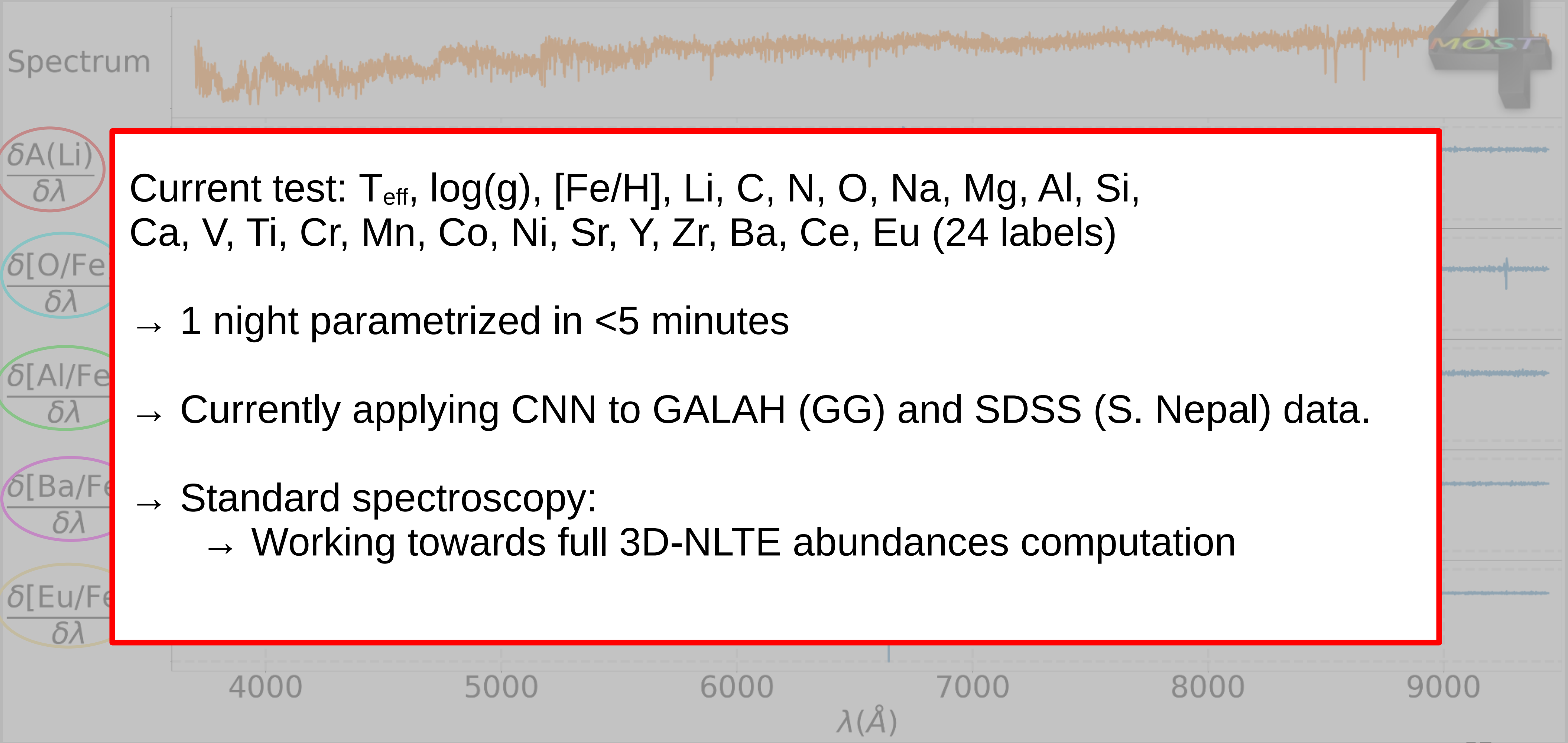
Developing CNN for 4MOST



Current test: T_{eff}, log(g), [Fe/H], Li, C, N, O, Na, Mg, Al, Si, Ca, V, Ti, Cr, Mn, Co, Ni, Sr, Y, Zr, Ba, Ce, Eu (24 labels)

→ 1 night parametrized in <5 minutes

Developing CNN for 4MOST



Current test: T_{eff}, log(g), [Fe/H], Li, C, N, O, Na, Mg, Al, Si, Ca, V, Ti, Cr, Mn, Co, Ni, Sr, Y, Zr, Ba, Ce, Eu (24 labels)

→ 1 night parametrized in <5 minutes

→ Currently applying CNN to GALAH (GG) and SDSS (S. Nepal) data.

→ Standard spectroscopy:
  → Working towards full 3D-NLTE abundances computation

Take-home messages

- Standard and ML spectroscopic methods complement each other !

Take-home messages

- Standard and ML spectroscopic methods complement each other !

- CNN performs extremely well for abundance measurements (RAVE, *Gaia*-ESO, *Gaia*-RVS, 4MOST)

Take-home messages

- Standard and ML spectroscopic methods complement each other !

- CNN performs extremely well for abundance measurements (RAVE, *Gaia*-ESO, *Gaia*-RVS, 4MOST)

- CNN in the context of the future large datasets !

(2024-2030) (2025-2030) (2026) (2030) (>2030)







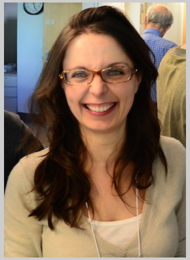








Merci !



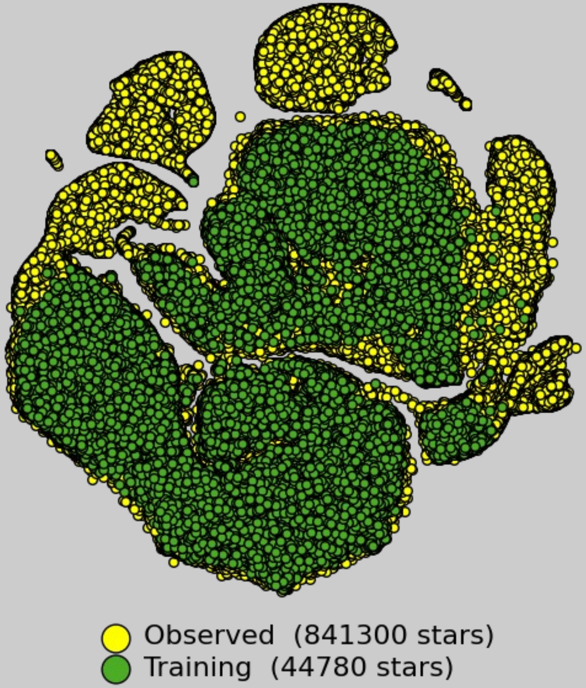


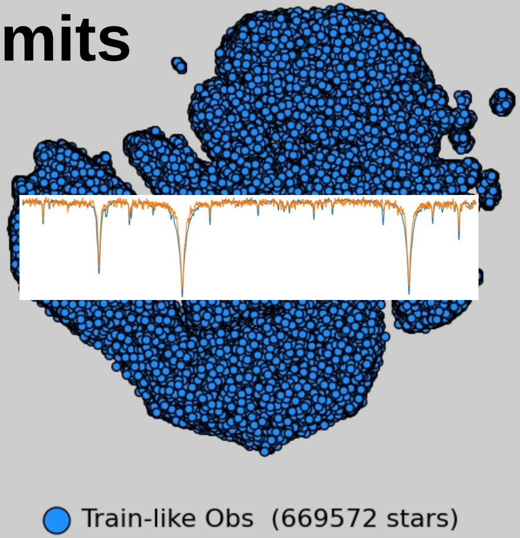

guiglion@mpia.de

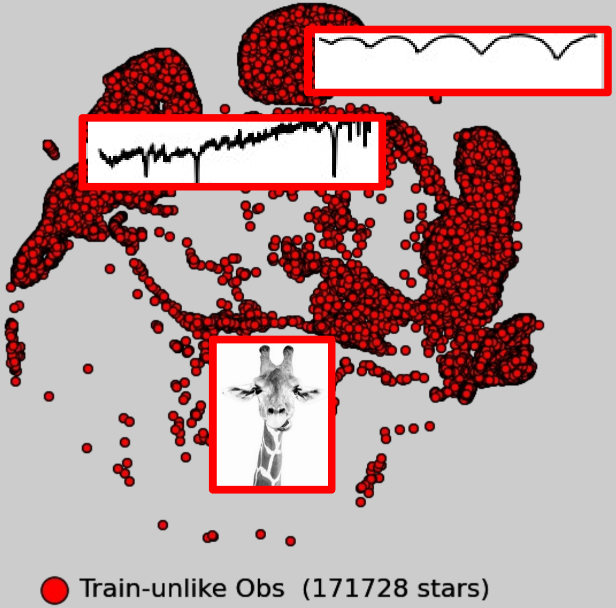
BONUS SLIDES

Selecting stars within the training sample limits

→ t-SNE classification of RVS spectra (adapted from Ambrosch, GG et al. 2023)







GG et al. 2024b

CNN uncertainties: deep ensemble approach