Machine-learning for stellar spectroscopy: past, present and future

Guillaume Guiglion (GG; ZAH/LSW, MPIA, guiglion@mpia.de)

Department of Physics and Astronomy, Uppsala University, Thu. 12 Oct 2023

C. Chiappini

M. Valentini

M. Steinmetz

G. Matijevič

Š. Mikolaitis

I. Minchev

S. Nepal

A. Queiroz

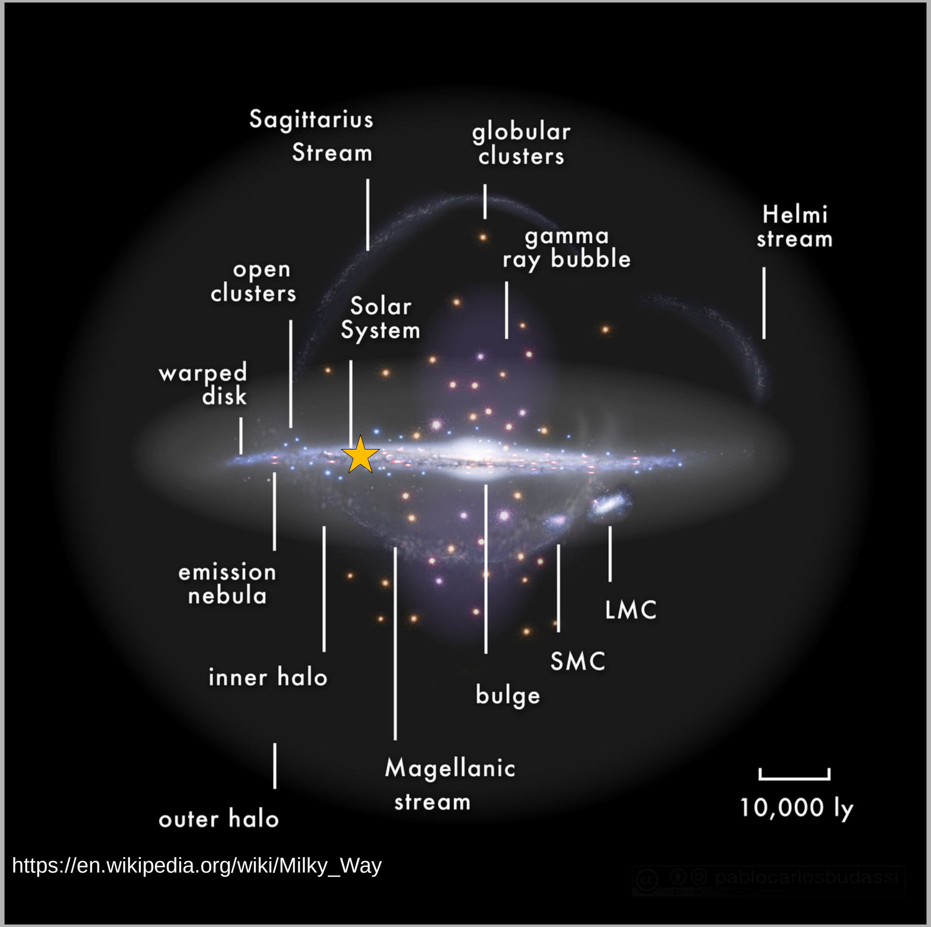
M. Ambrosch

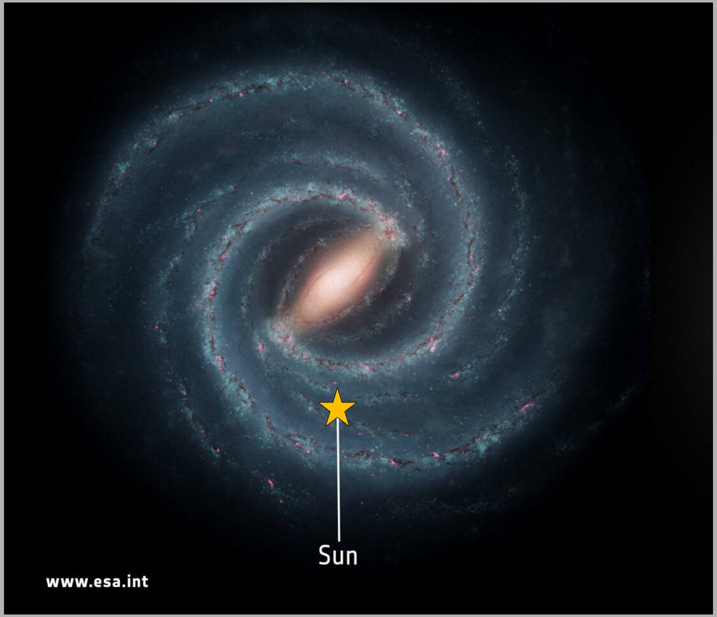
K. Youakim

M. Bergemann

Context: Galactic Archaeology

→ studying the formation and evolution of the Milky Way and it’s local volume

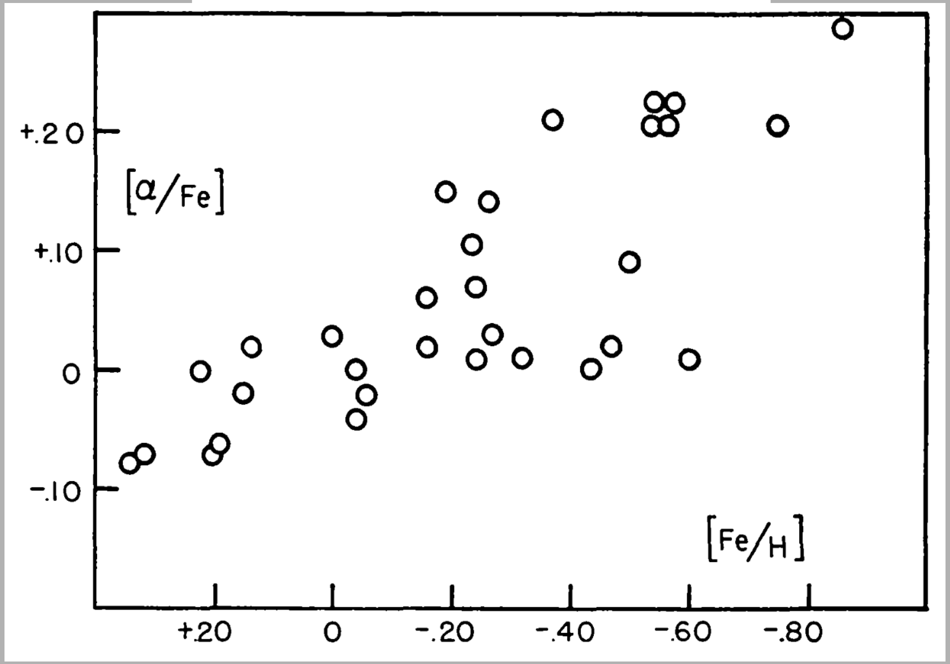




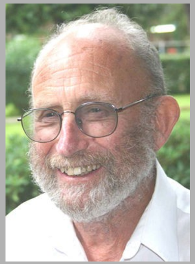
→ **Stellar chemistry in essential for Galactic Archaeology**

→ **The Tinsley-Wallerstein diagram**

30 stars from Wallerstein (1962)



Beatrice Tinsley

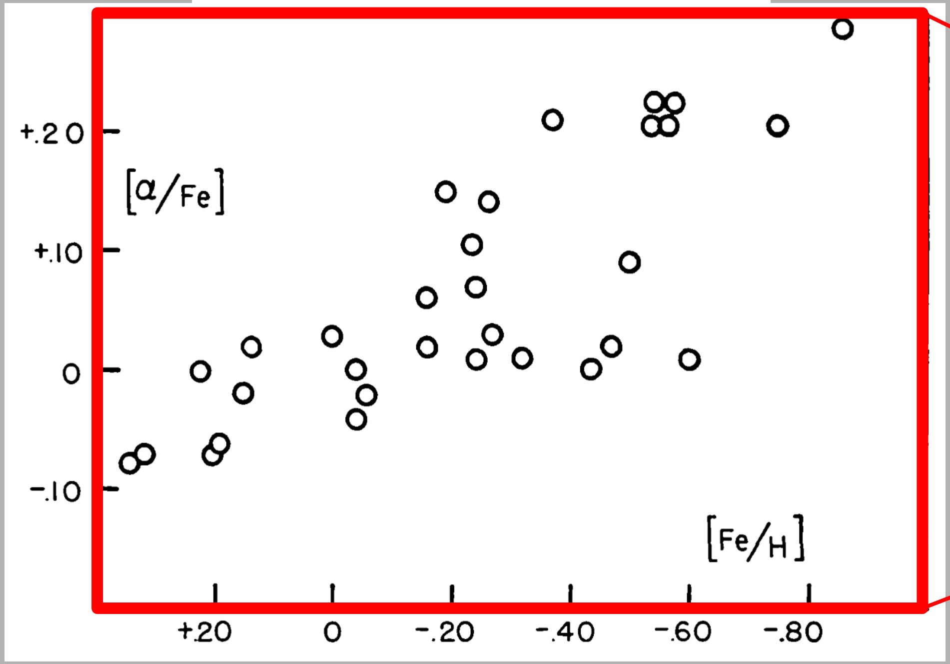


George Wallerstein

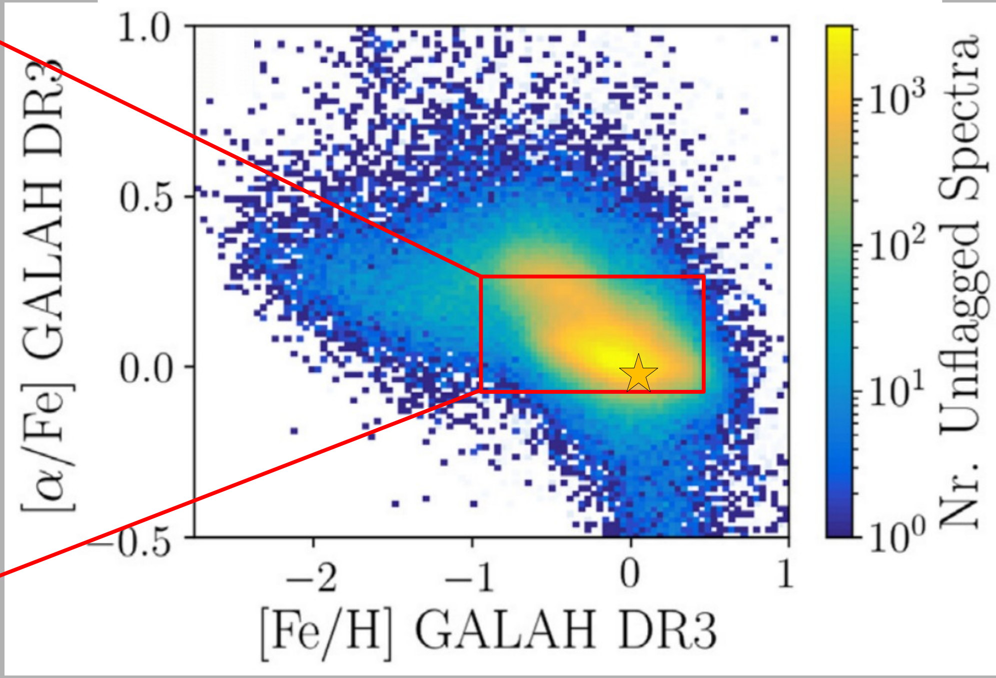
→ **Stellar chemistry in essential for Galactic Archaeology**

→ **The Tinsley-Wallerstein diagram**

30 stars from Wallerstein (1962)



~500000 stars from GALAH (Buder et al. 2021)



Other types of elements:

→ Fe-peak: Z, Cu, Ni, Co, Fe, Mn, Cr, V

→ neutron-capture: Sr, Y, Zr, Ba, La, Ce, Pr, Sm, Eu, Gd, Dy, … (Battistini & Bensby 2016)

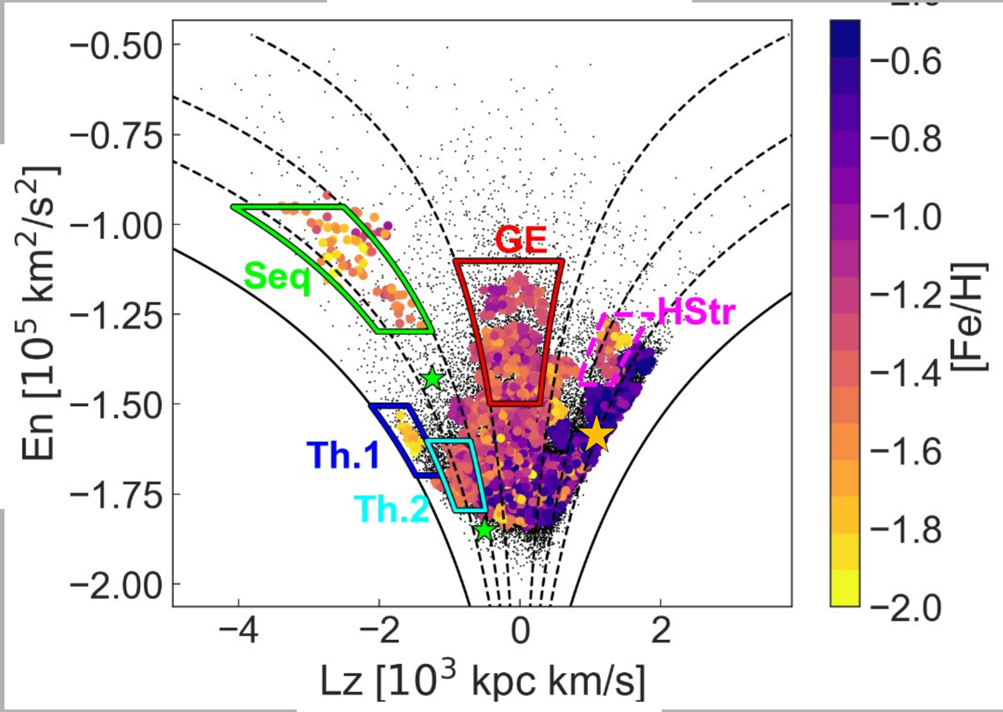
→ Light: Li, B, Be, C, N (Randich & Magrini 2021)

→ Odd-Z: Sc, K, Al, Na

→ Stellar chemistry in essential for Galactic Archaeology

→ But stellar kinematics and stellar ages too :)

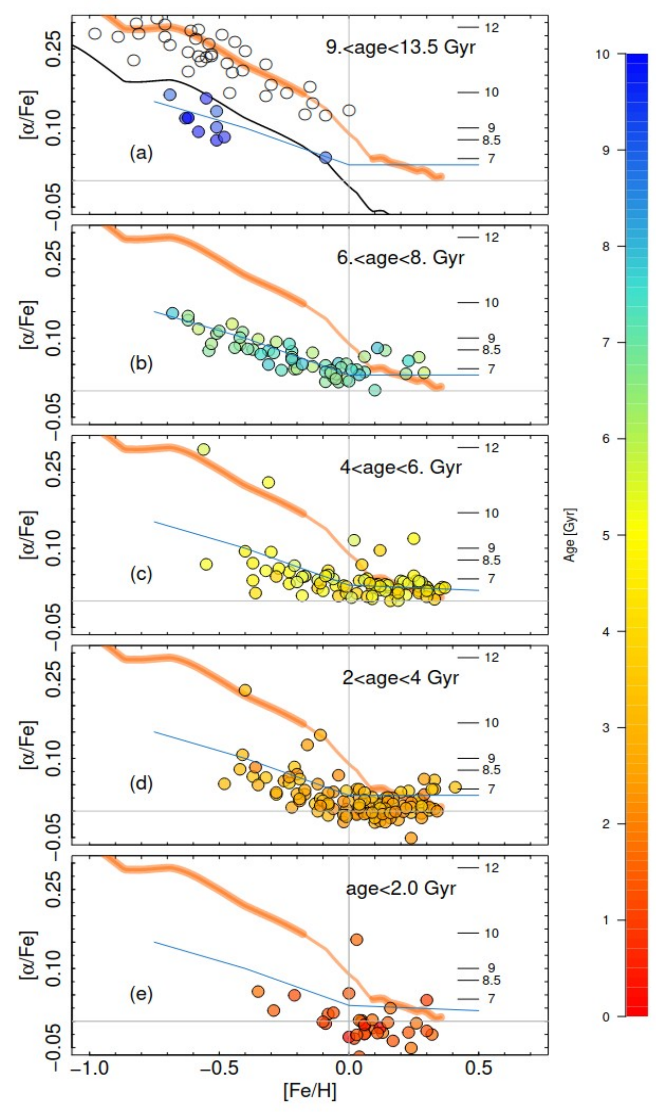
Koppelmann et al. 2019



(Energy)

(Angular momentum)

Haywood et al. 2019



The need for large spectroscopic surveys

RAVE

RADIAL VELOCITY EXPERIMENT

$5x10^5$ stars

GaiaESO

10^5 stars

ESO

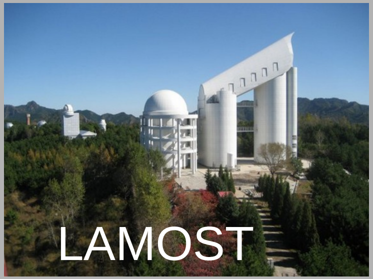
AMBRE

$>10^4$ stars

gaia

$>5x10^6$ spec

SDSS $>10^6$ stars



10^6 stars

GALAH

$>5x10^5$ stars

MOONS

Multi Object Optical and Near-infrared Spectrograph

$>10^6$ stars

Maunakea Spectroscopic Explorer

$>10^6$ stars

APOGEE

SDSSIII

$>5x10^5$ stars

4MOST

$30x10^6$ stars

WEAVE

$>10^6$ stars

Gaia**:** Gaia Collaboration, Vallenari et al. 2022
AMBRE/ESO: Guiglion et al. 2016
LAMOST: Zhang et al. 2021
Gaia**-ESO:** Romano et al. 2021
GALAH: Gao et al. 2020
APOGEE: Abdurro'uf et al. 2022
4MOST: de Jong et al. 2019
WEAVE: Jin et al. 2022
MSE: Bergemann et al. 2019
MOONS: Cirasuolo et al. 2020

What we measure from a star:

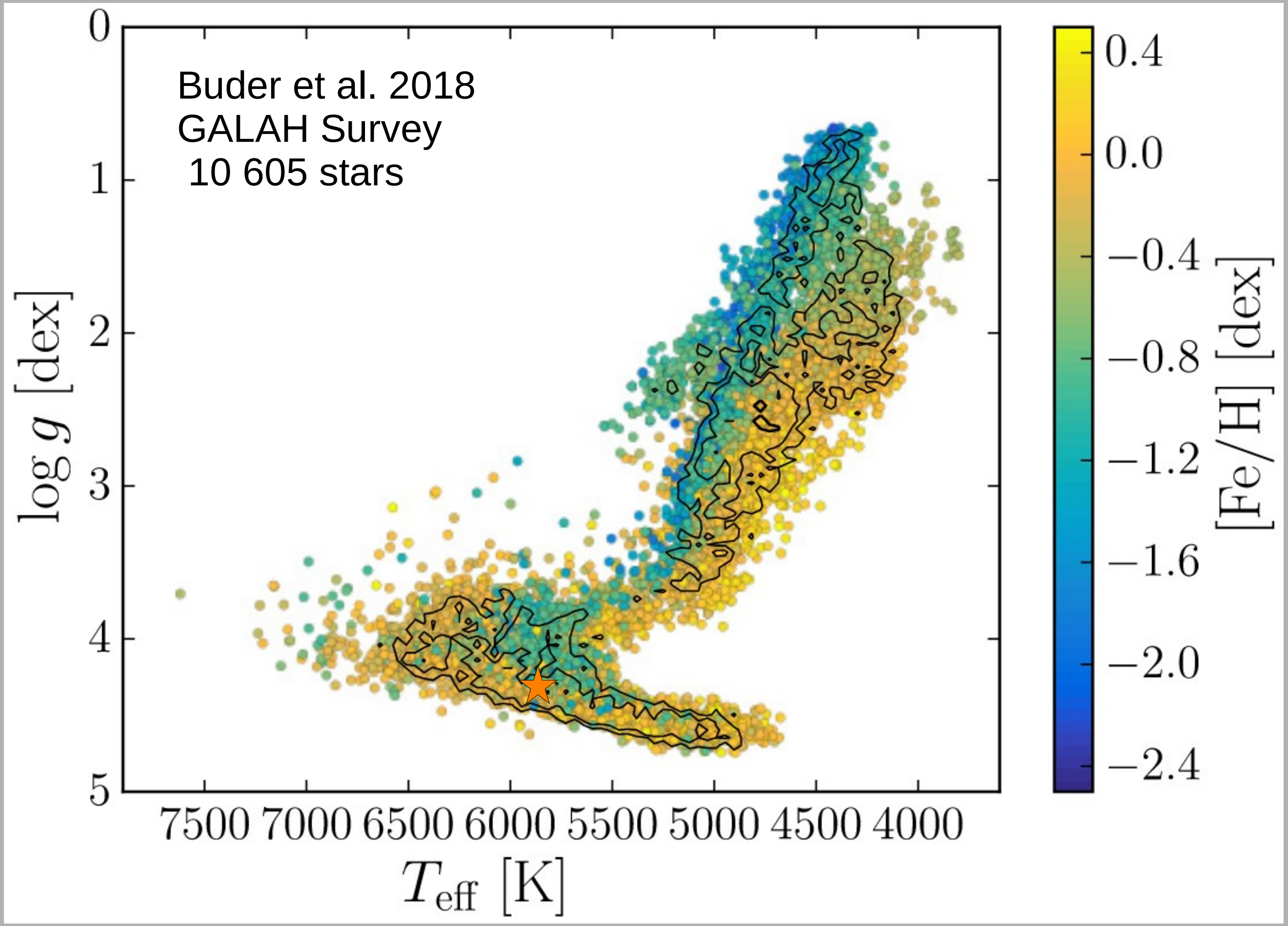
→ **Atmospheric parameters:**
  - → Effective temperature **T_{eff}**
  - → Surface gravity **log(g)**
  - → Overall metallicity **[M/H]**

→ **Average abundance ratios**
  - → For instance **[α/M]** with α goes for α-elements (Mg, Si, Ca, O, Ti, Ne, S)

→ **Individual chemical abundances**
  - → **[X/Fe] with X = {Mg, Si, Ti, Ni, Fe, Ba, Eu, ….}**

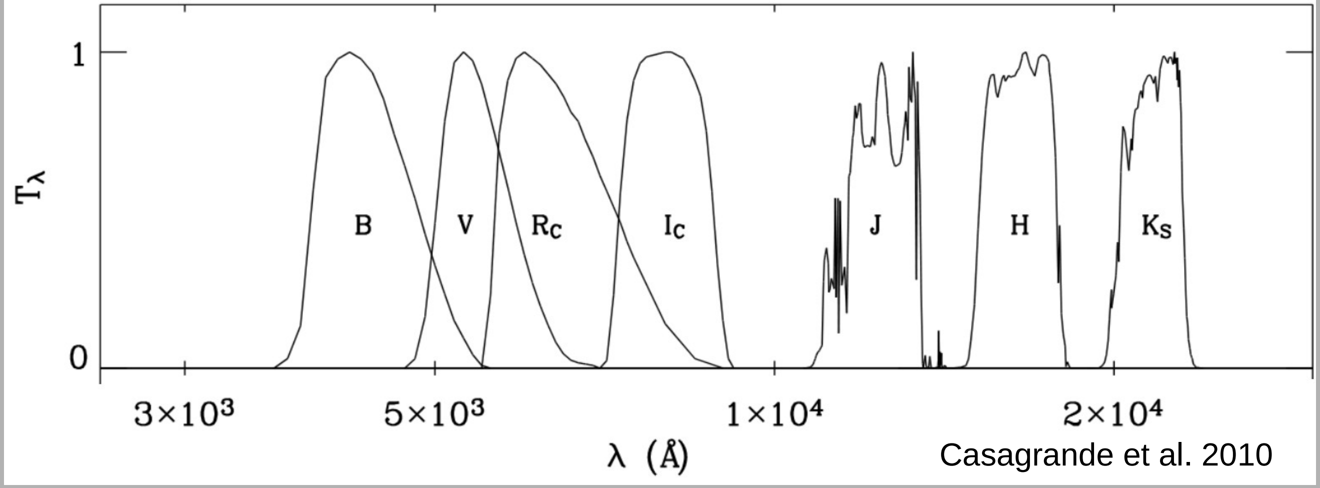
Many other parameters, such as rotation, activity, mass, age ...

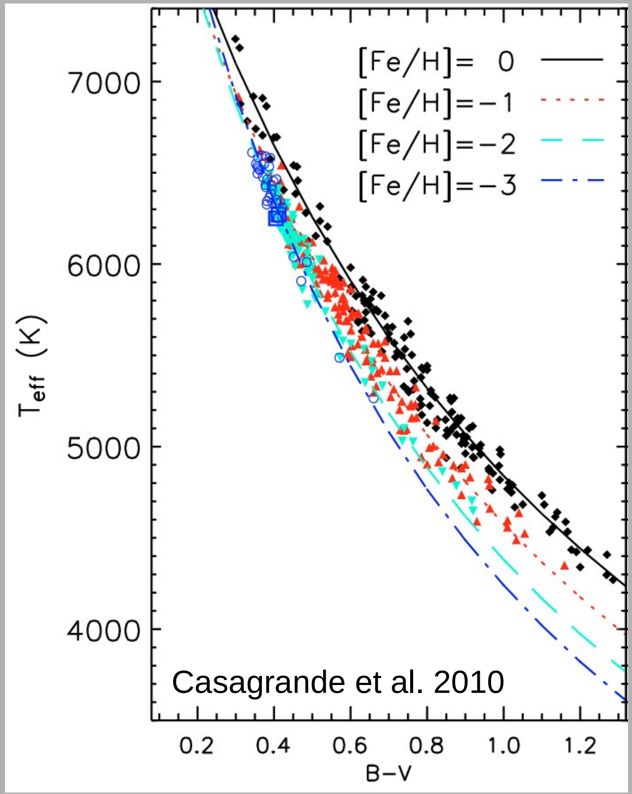
What type of stars are we interested in ?



How do we measure astrophysical quantities?

→ Using photometry (stellar magnitudes)





→ See also Casagrande et al. 2021 (+ references there-in)

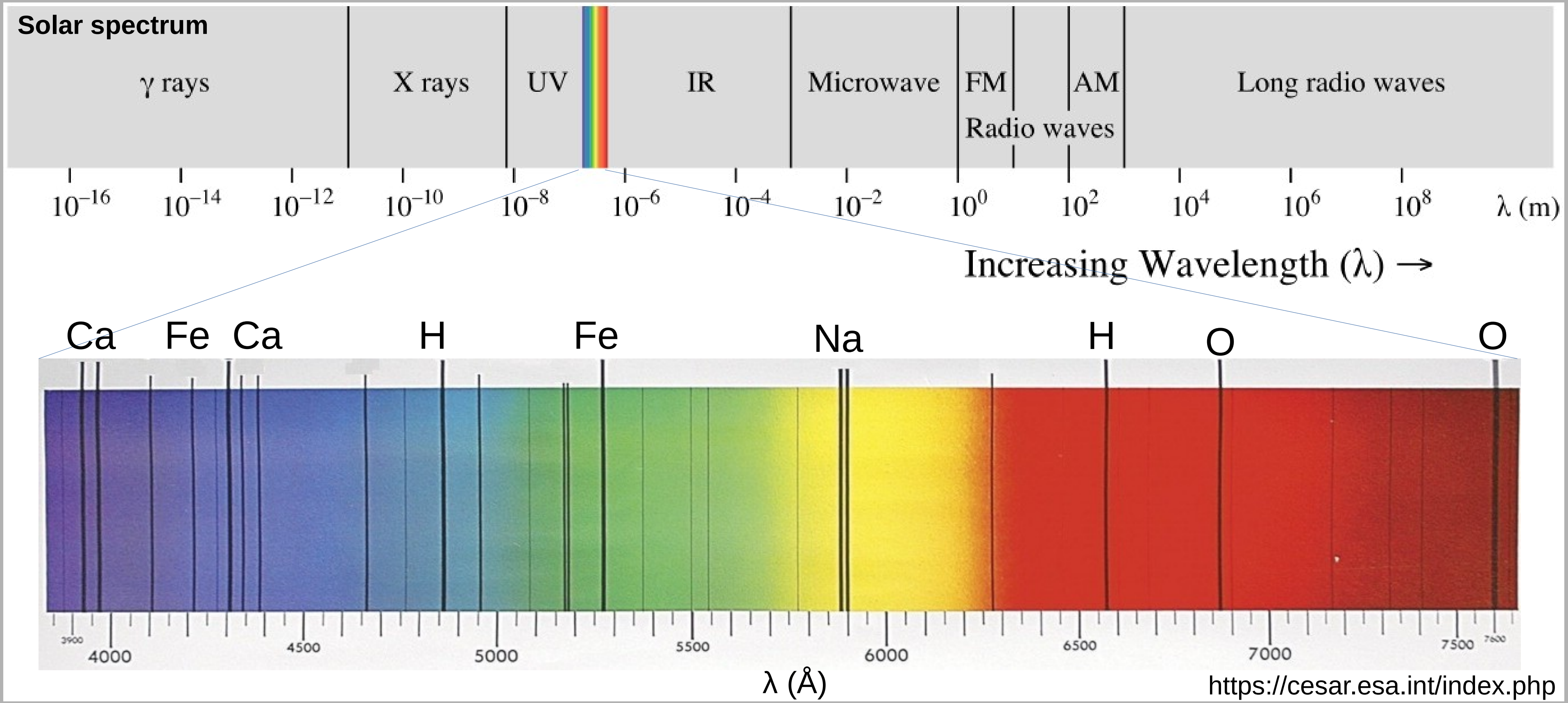
→ Using stellar evolutionary models + magnitudes + parallaxes (distances)

- → Can measure T_{eff}, log(g), [M/H]
- → StarHorse code: Queiroz et al (2018, 2020, 2023), Anders et al. (2019, 2022)

Standard analysis of stellar spectra

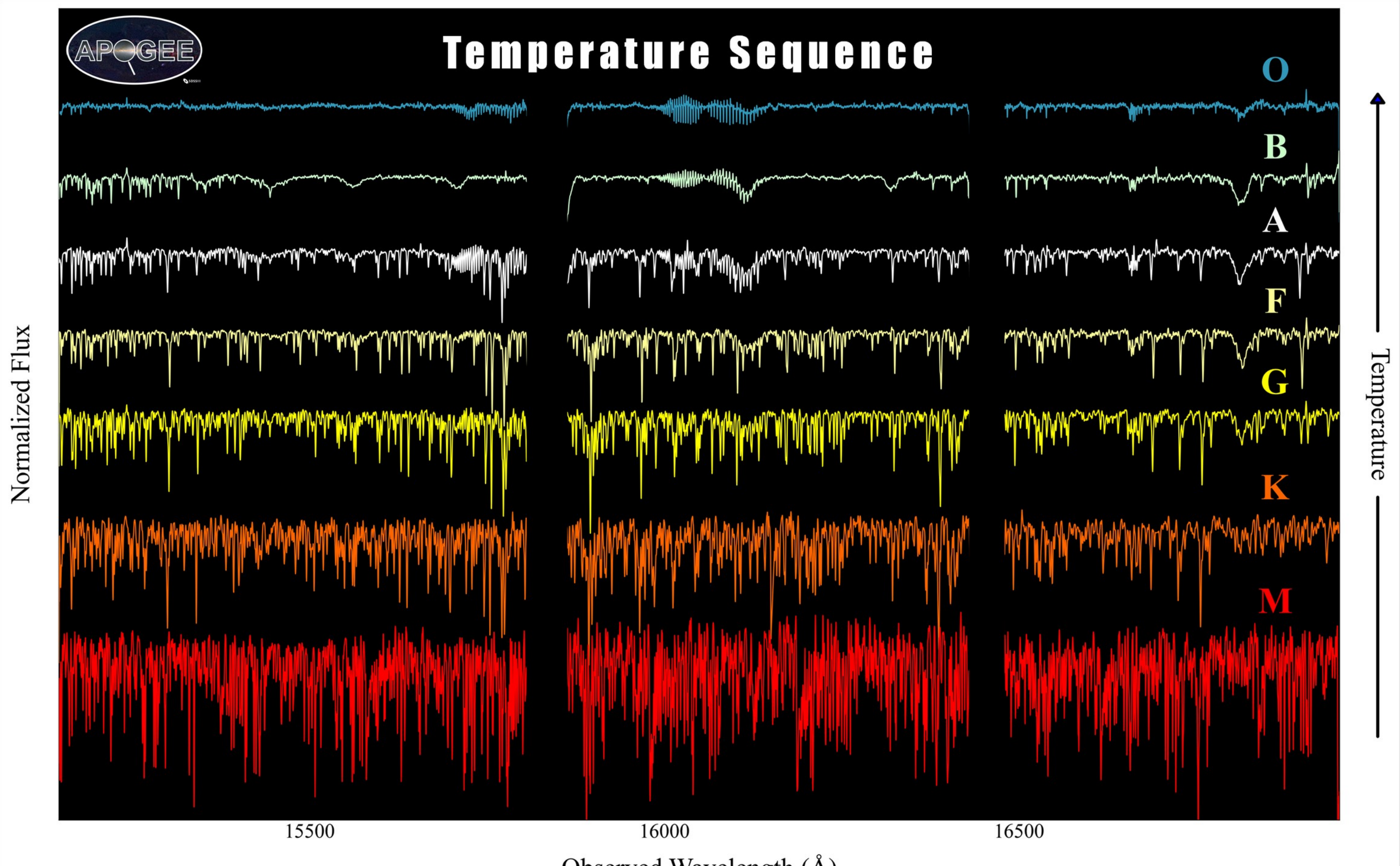
How do we measure astrophysical quantities?

→ **Using stellar spectra**



How do stellar spectra correlate with astrophysical parameters ?

→ **Example: effective temperature**



https://www.sdss3.org/dr10/irspec/spectro_basics.php

T_{eff} (K)

> 20000

[10000-20000]

[7500-1000]

[6000-7500]

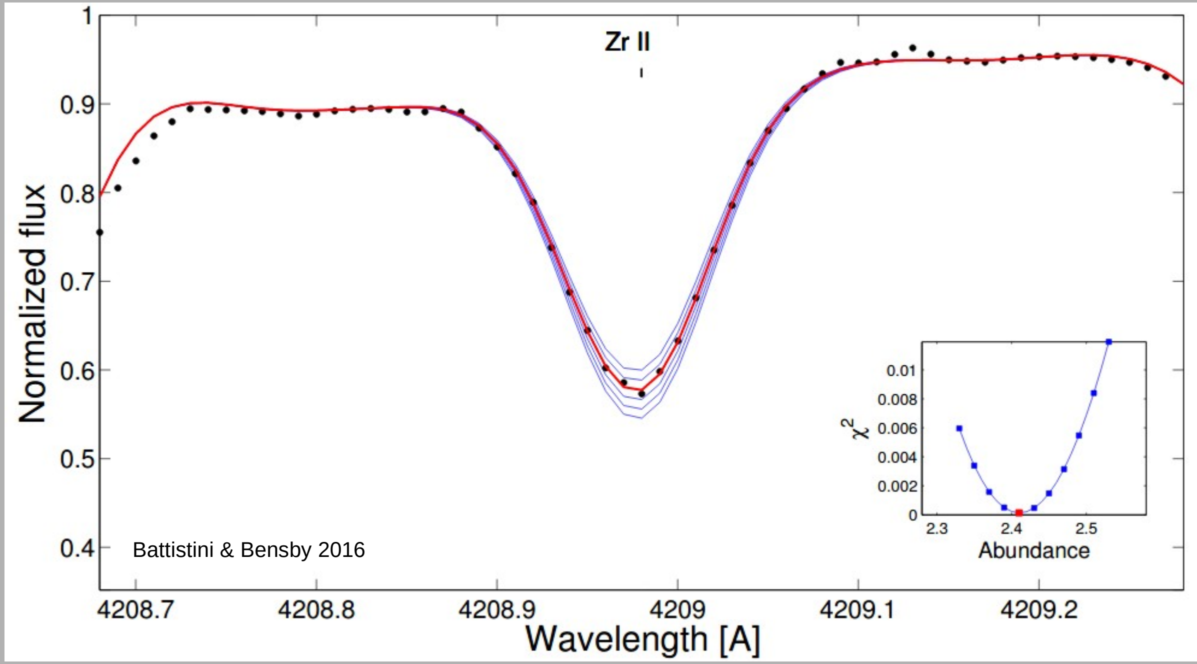
[4500-6000]

[3000-4500]

< 3000

How do we measure chemical abundances ?

→ One popular method: spectral fitting



To create model spectra, we need:

Model atmosphere
(e.g MARCS, Gustafsson et al. 2008)

+

Radiative transfer code
(e.g. Turbospectrum)

+

$\mathbf{T_{eff}}$
log(g)
[M/H]

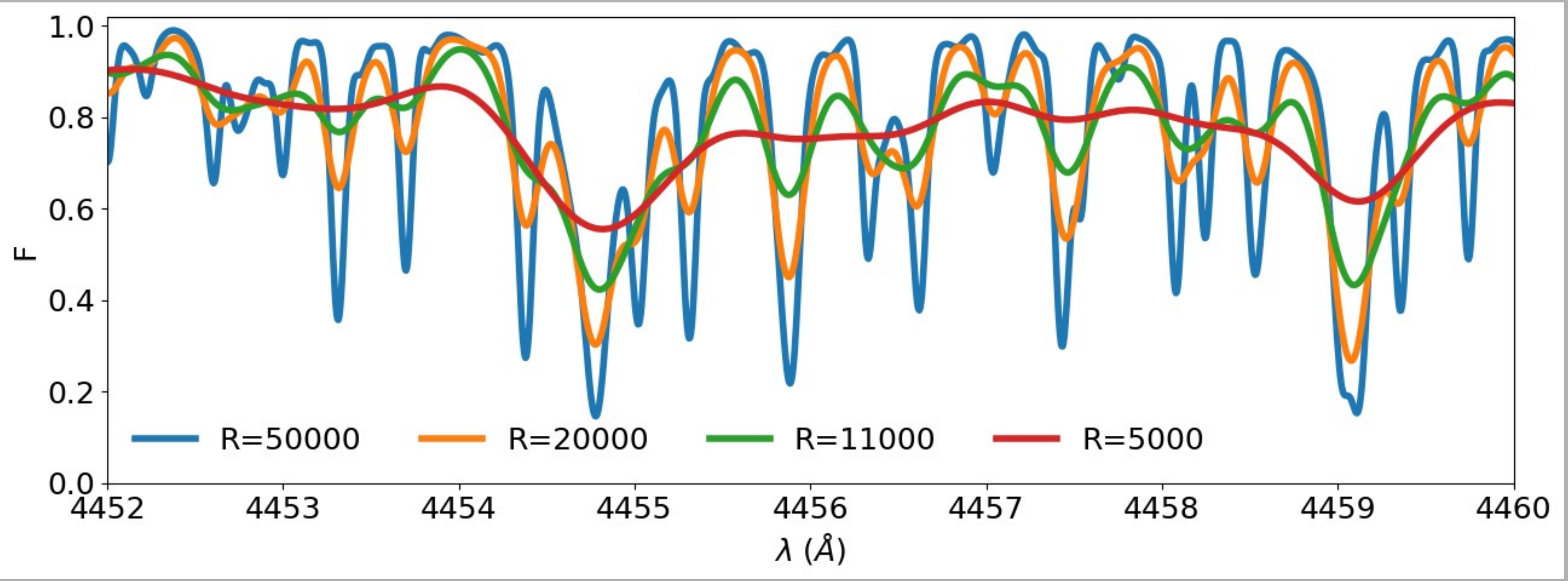
+

Atomic data
(VALD, Piskunov et al. 1995,
GES, Heiter et al. 2021)

Challenges to face:
1D vs 3D
LTE vs. NLTE
(Lind et al. 2009, Wang et al. 2021, Amarsi et al. 2020)
Low-res vs. High-res
Blends
Rotation, Turbulence

More details on chemical abundance derivation:
→ Jofré, Heiter, and Soubiran (2019)

The impact of spectral resolution on abundance determination



→ **Lower spectral resolution:**

  → **Less clean spectral features to rely on**

  → **Less elements to be measured**

  → **Lower precision**

Machine-learning analysis of stellar spectra

The world of machine-learning

Supervised Learning			Unsupervised Learning	
Regression	Classification	Neural networks	Clustering	Dimensionality Reduction
Linear Regression	Logistic Regression	Convolutional Neural Networks (CNN)	K-means	t-Distributed Stochastic Neighbor Embedding (t-SNE)
K-Nearest Neighbors Regression (KNN)	K-Nearest Neighbors Classification (KNN)	Generative Adversarial Networks (GAN)	Gaussian Mixture Models (GMM)	Locally Linear Embedding (LLE)
Random Forest Regression	Random Forest Classification	Long Short Term Memory Networks (LSTM)	Hierarchical Agglomerative Clustering (HAC)	Uniform Manifold Approximation and Projection (UMAP)
Decision Tree Regression (CART)	Decision Tree classification (CART)	Gated Recurrent Units (GRU)	Density-Based spatial Clustering of Applications with Noise (DBSCAN)	Multidimensional Scaling (MDS)
Support Vector Regression (SVR)	Support Vector Machines (SVM)	Feedforward Neural Networks (FFNN)		Principal Component Analysis (PCA)
Multivariate Adaptive regression Splines (MARS)	Extreme Gradient Boosting (XGBoost)			Isomap Embedding
	Gradient Boosted Trees			
	Naive Bayes			

Webb & Good 2023

→ In the current talk, I will discuss only on Convolutional Neural-Networks

Basic concepts of Convolutional Neural-Networks (CNN)
→ Practical example: Cat and dog classification

→ ?

→ Step 1: build a training sample

Data

Labels

Cat
Dog
...
Dog
Cat

→ Step 2: generate (train) a model between data and labels

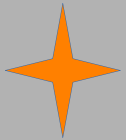
Data → CNN: High-D non-linear function composed of neurons, characterized by weights and biases → **Labels**

→ Step 3: predict the type of animal on a picture

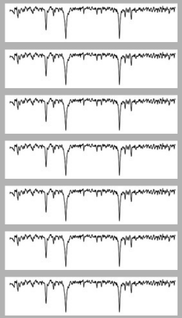
→ CNN model → **Dog**

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

CNNs for stellar spectroscopy

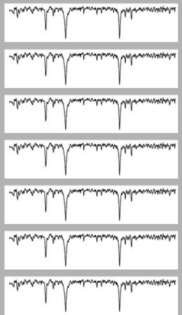
→ **Example: Measuring temperature of the star**  **which spectrum is** 

→ **Step 1: build a training set**
(with labels determined from standard spectroscopic methods)

Data = spectra	**Labels = T_{eff}**
	5234 4834 2730 7598 … … …

→ **Step 2: train a CNN**

Input: Spectra



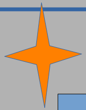
→

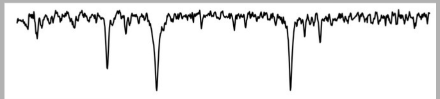
CNN:
High-D non-linear function composed of neurons, characterized by weights and biases

→

Output: Labels

5234
4834
2730
7598
…
…
…

→ **Step 3: predict temperature of** 



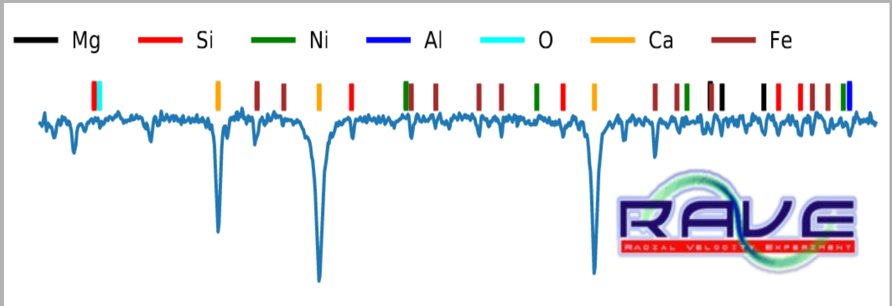
→ Trained CNN →

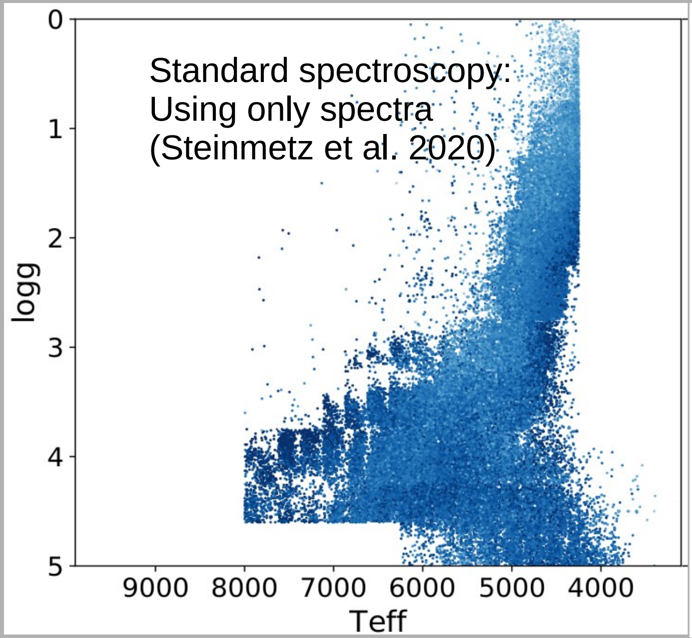
T_{eff} = 5812 K

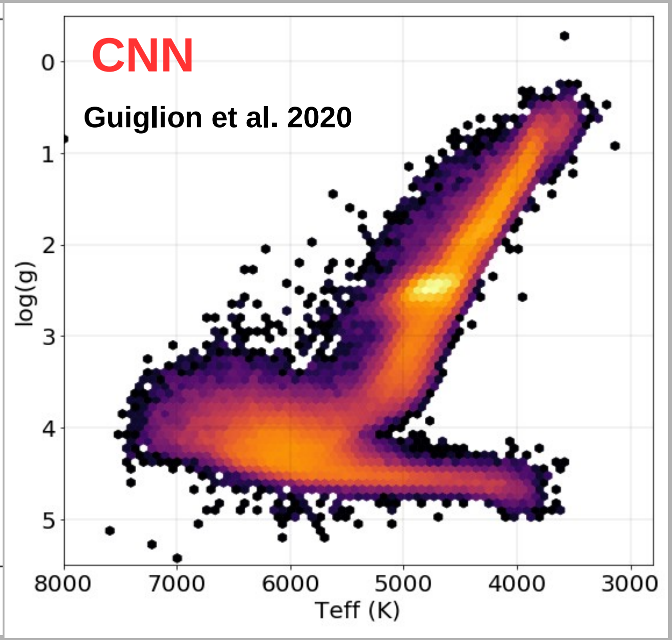
Past applications of CNNs for stellar spectroscopy

Our experience with CNNs and RAVE spectra

→ 1st application of CNNs combining RAVE spectra, *Gaia* magnitudes, and parallaxes
→ Training set: 4000* with labels from APOGEE DR16 (R~22000)
→ Transfer high-quality labels to low-resolution RAVE spectra (R~7500)







See also:
Bailer-Jones et al. 1997
Leung & Bovy 2019
Fabbro et al. 2018
Zhang et al. 2019
Bialek et al. 2020

→ Such particular combination of data allows to break the spectral degeneracies inherent to RAVE spectra (and likely to be present in *Gaia* RVS spectra)

Chemical evolution of lithium with CNN from stellar spectra

→ **Why is lithium important ?**

→ **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**

S. Nepal (AIP)

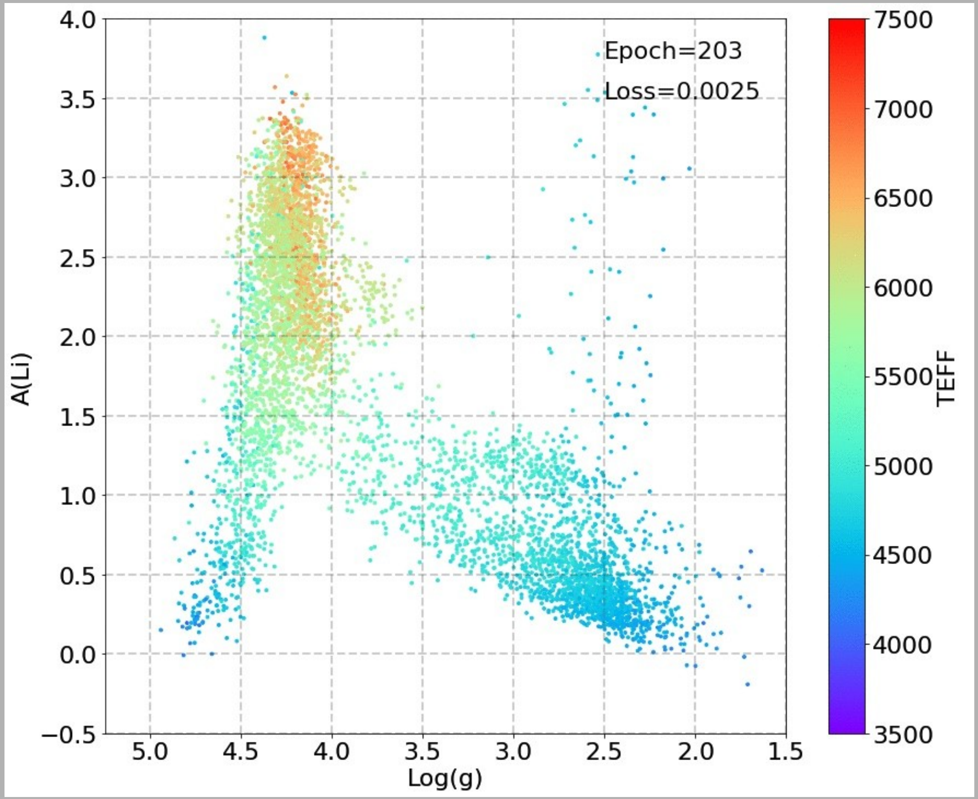
Chemical evolution of lithium with CNN from stellar spectra

S. Nepal (AIP)

→ **Why is lithium important ?**

→ **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**



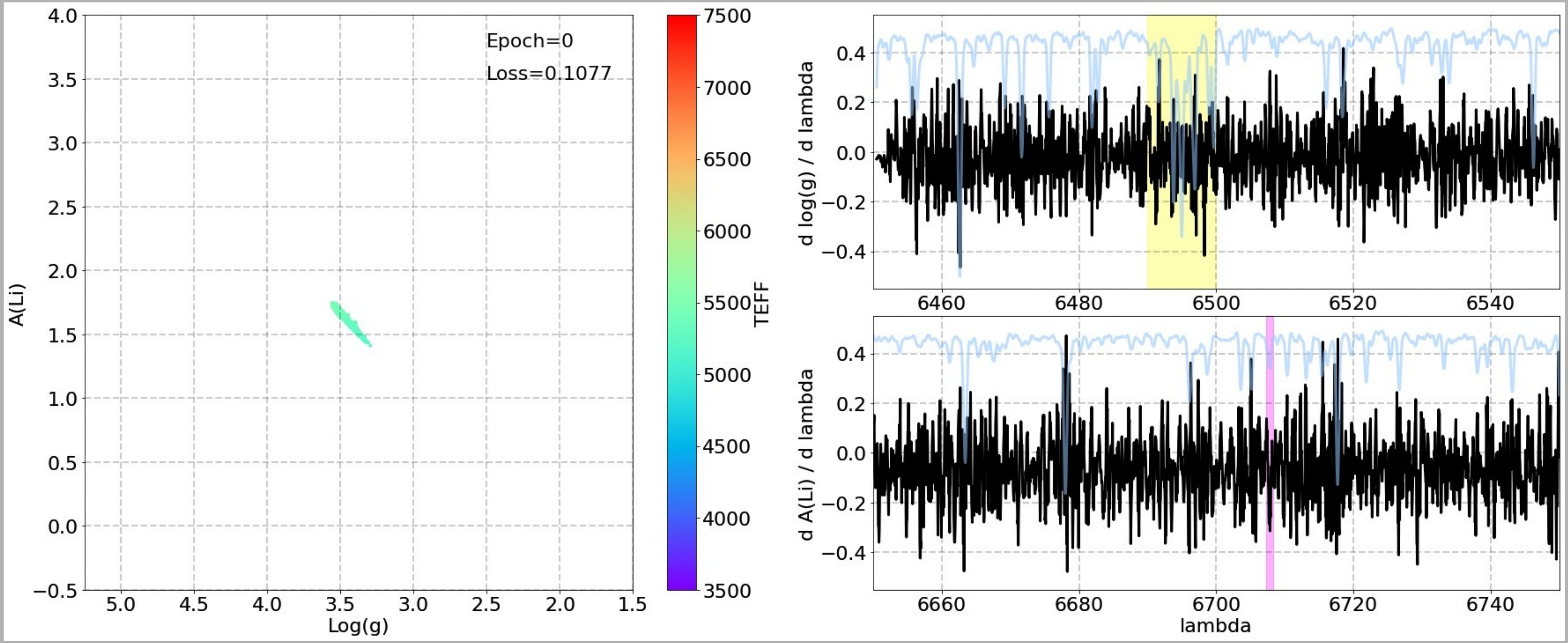
Chemical evolution of lithium with CNN from stellar spectra

S. Nepal (AIP)

→ **Why is lithium important ?**

→ **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**

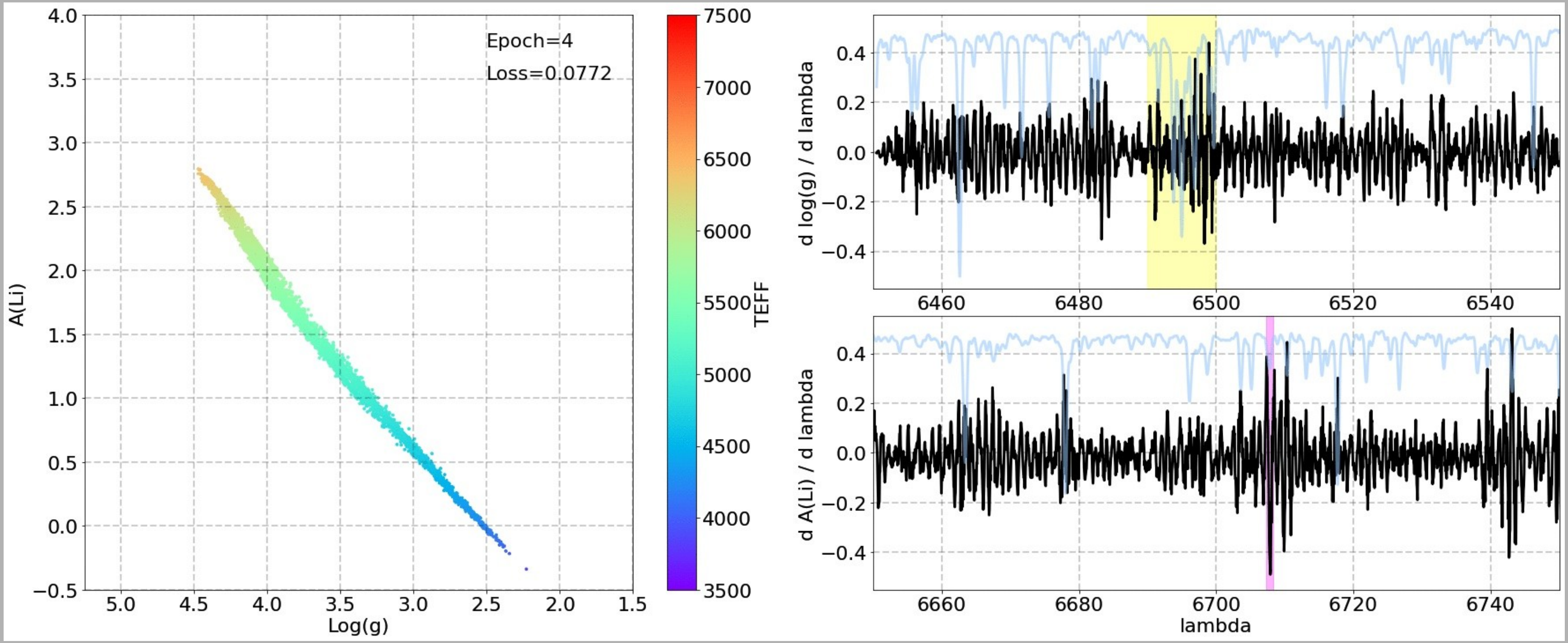


Chemical evolution of lithium with CNN from stellar spectra

S. Nepal (AIP)

→ **Why is lithium important ?**
  → **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**

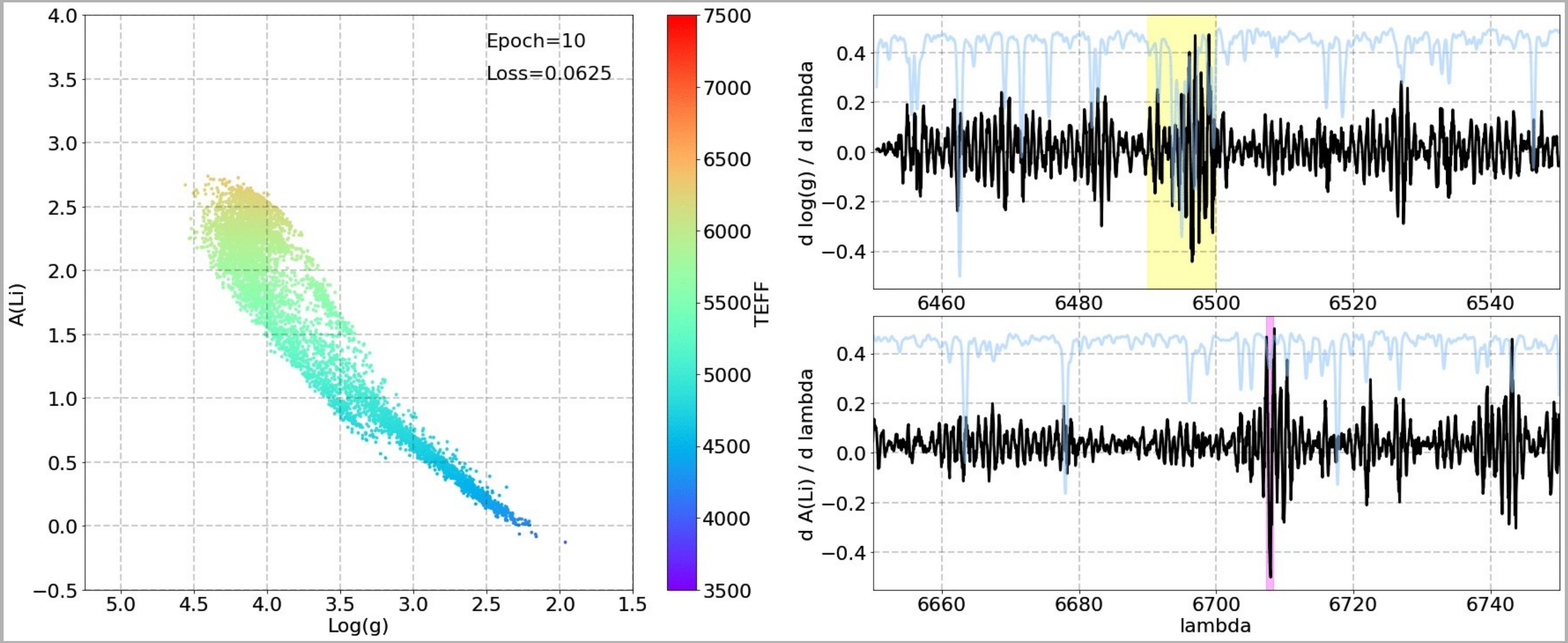


Chemical evolution of lithium with CNN from stellar spectra

→ **Why is lithium important ?**
  → **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**


S. Nepal (AIP)

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**



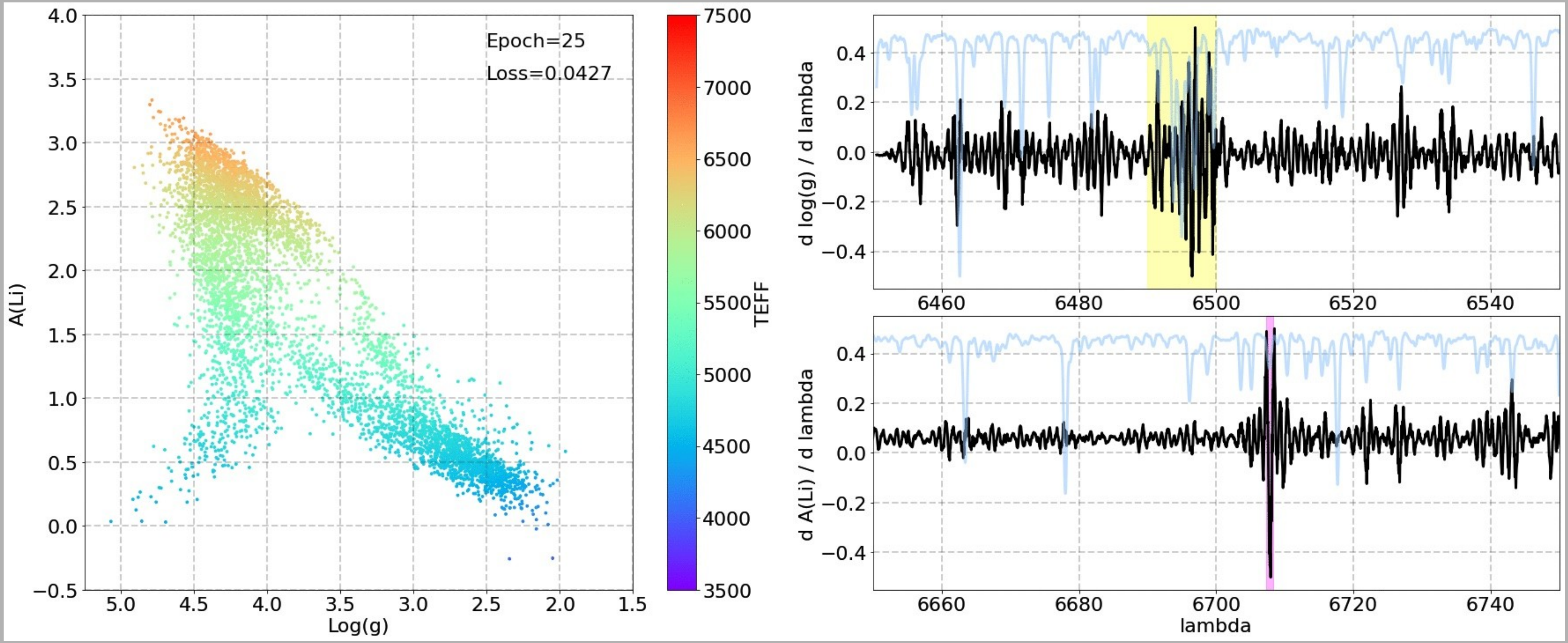
Chemical evolution of lithium with CNN from stellar spectra

→ **Why is lithium important ?**

→ **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

S. Nepal (AIP)

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**

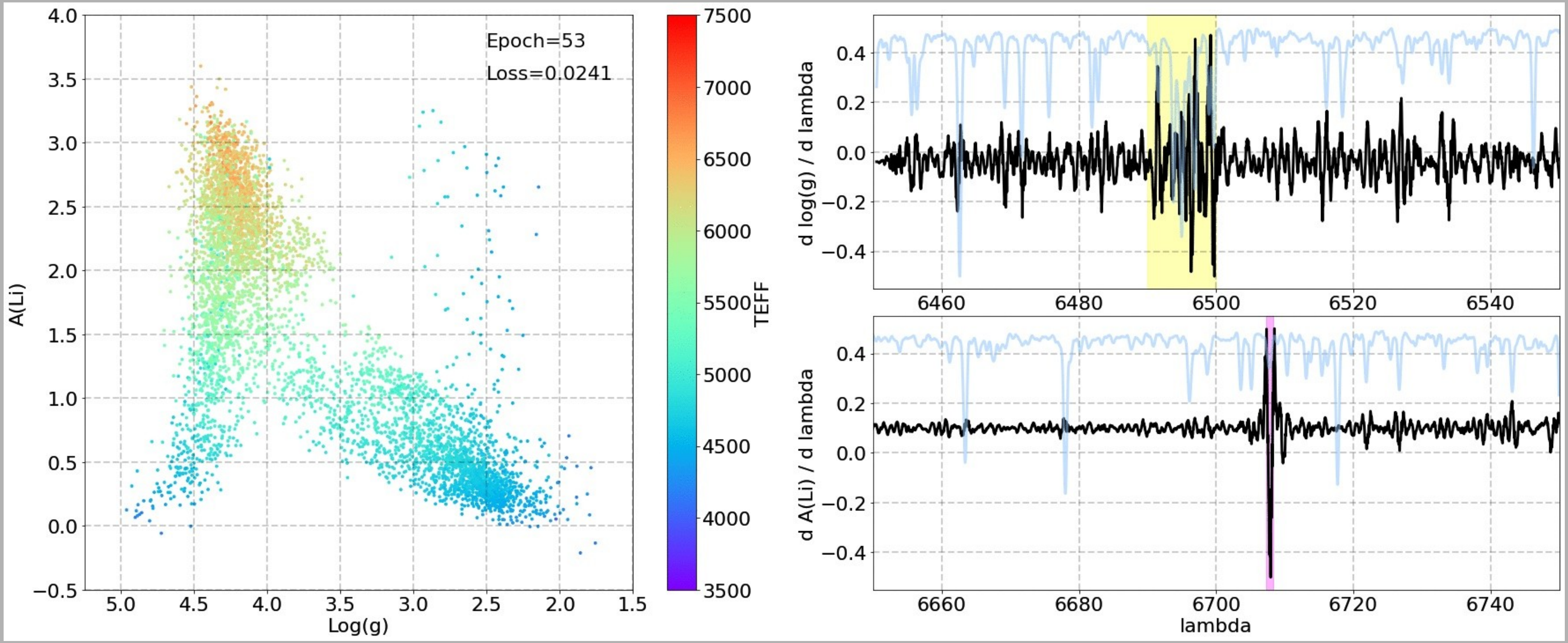


Chemical evolution of lithium with CNN from stellar spectra

→ **Why is lithium important ?**
  → **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

S. Nepal (AIP)

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**



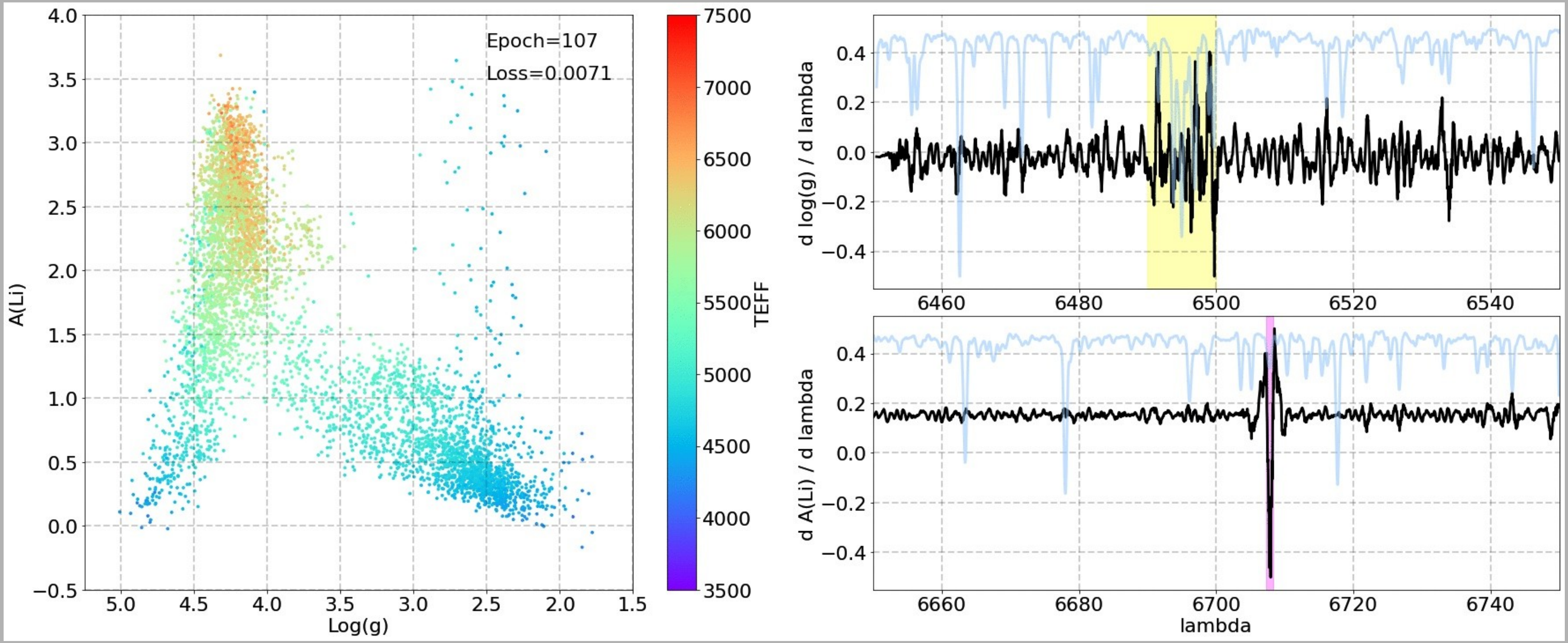
Chemical evolution of lithium with CNN from stellar spectra

S. Nepal (AIP)

→ **Why is lithium important ?**

→ **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**



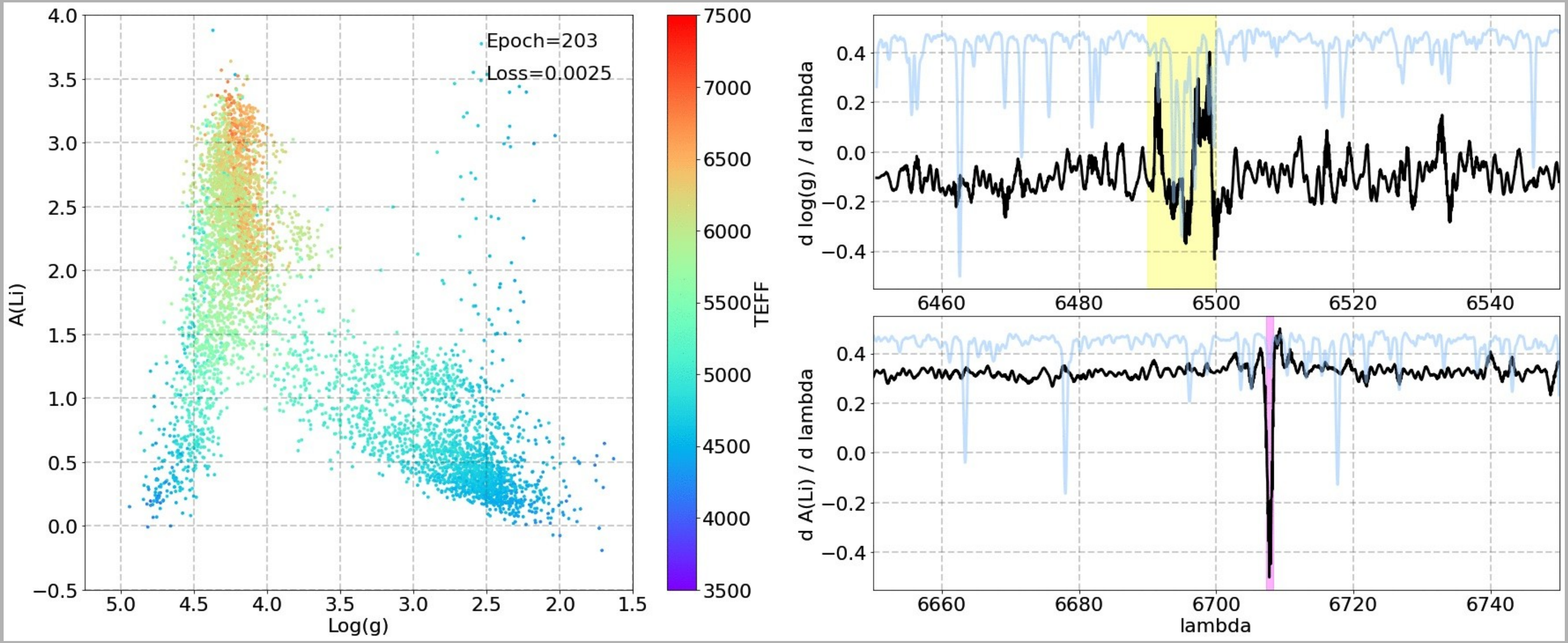
Chemical evolution of lithium with CNN from stellar spectra

S. Nepal (AIP)

→ **Why is lithium important ?**

→ **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**



Nepal, GG et al. (2023)
https://github.com/SamirNepal/Li_CNN_2022

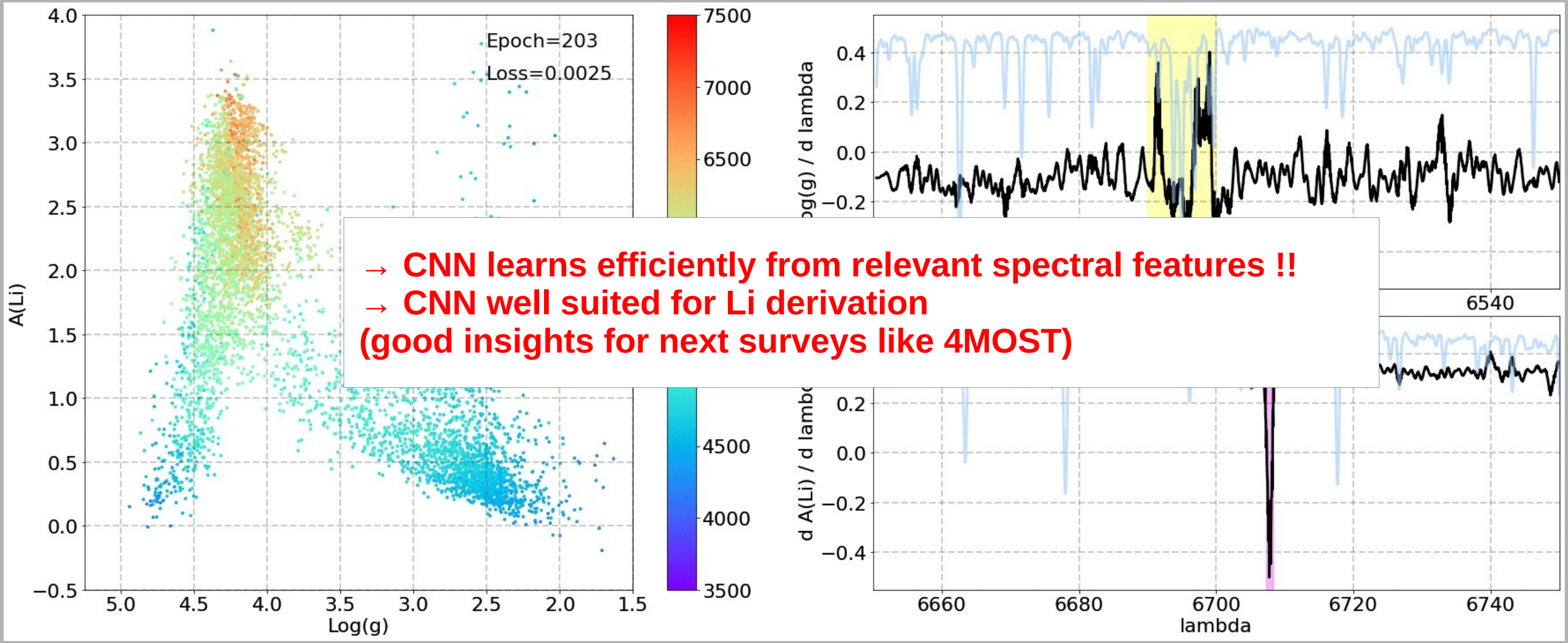
Chemical evolution of lithium with CNN from stellar spectra

S. Nepal (AIP)

→ **Why is lithium important ?**

→ **Chemical evolution of Li in the Milky Way still unclear (e.g. Guiglion et al. 2019)**

→ **Training set: 7000 stars with *Gaia*-ESO spectra, to derive T_{eff}, log(g), [Fe/H], A(Li)**

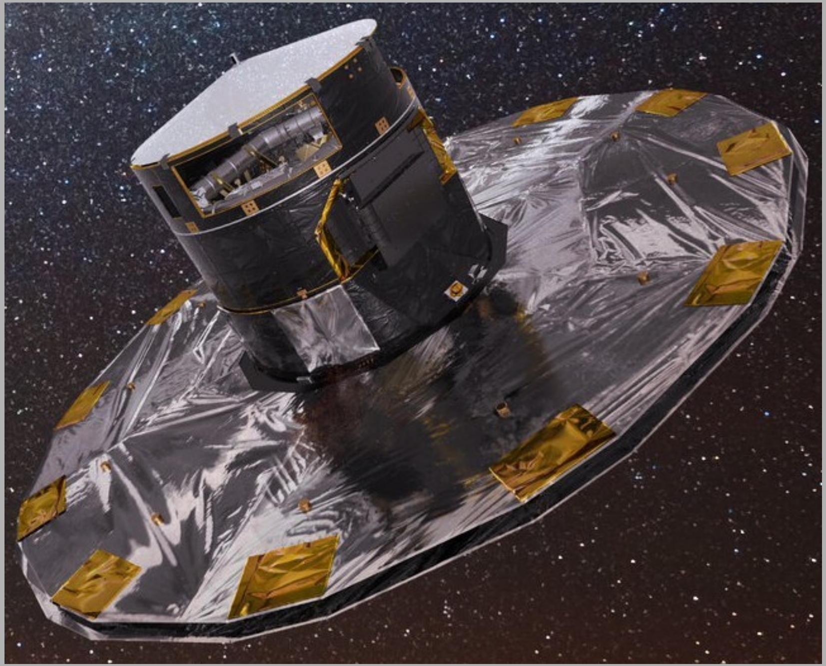


→ CNN learns efficiently from relevant spectral features !!
→ CNN well suited for Li derivation
(good insights for next surveys like 4MOST)

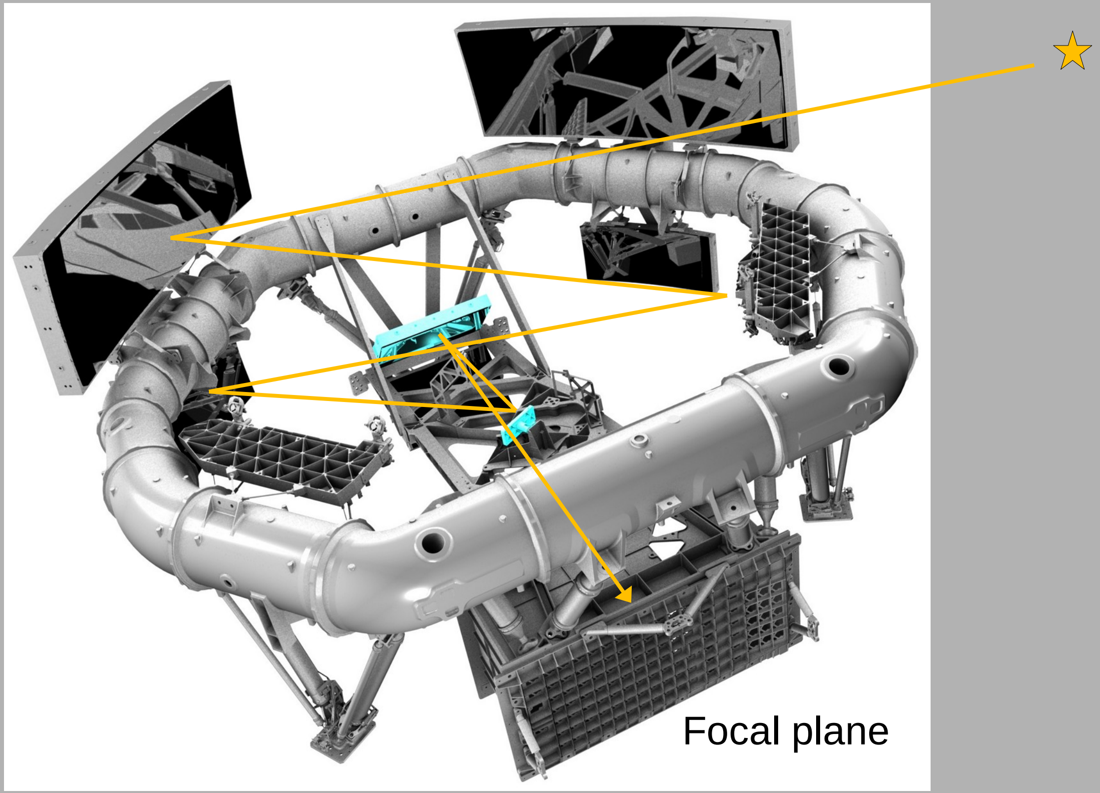
Nepal, GG et al. (2023)
https://github.com/SamirNepal/Li_CNN_2022

Present application of CNNs for stellar spectroscopy

Gaia: ESA's billion star surveyor

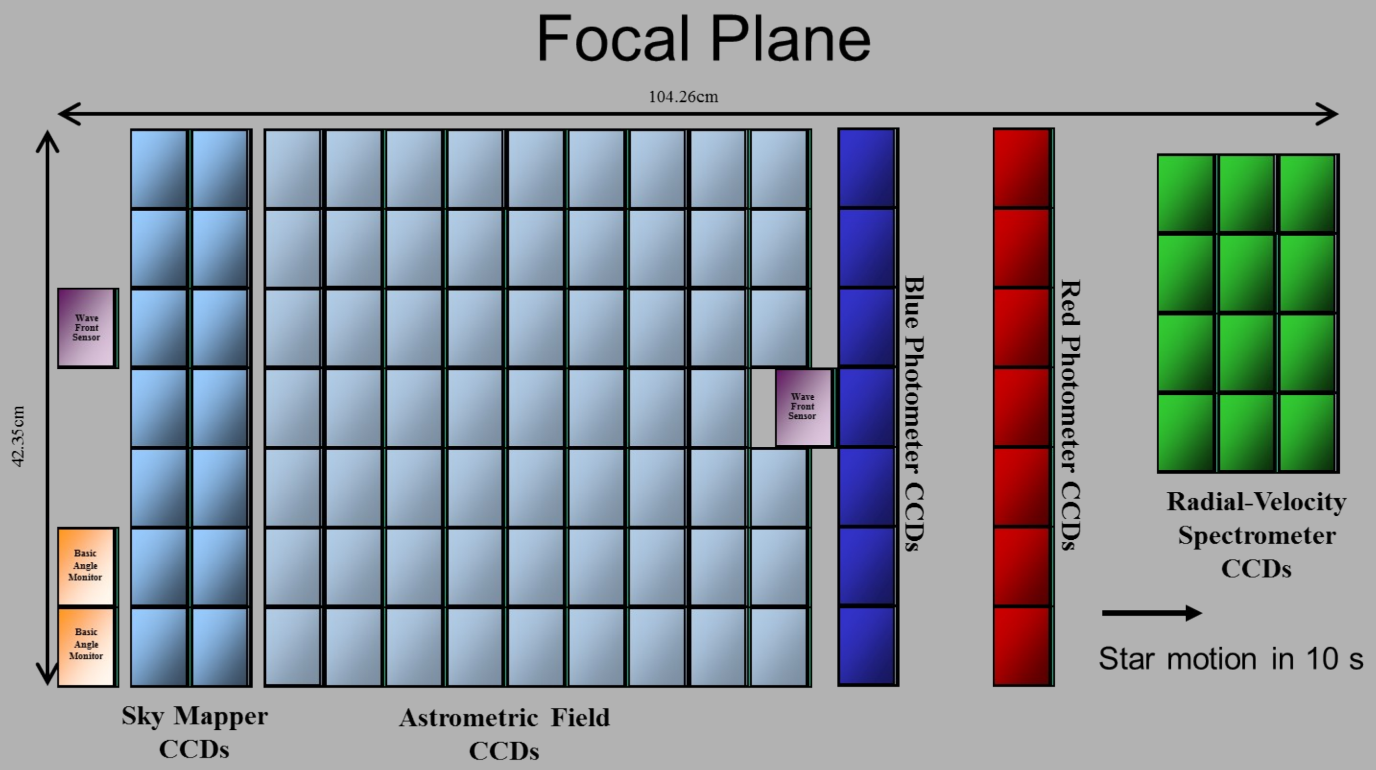


https://www.esa.int/Enabling_Support/Operations/Gaia_s_biggest_operation_since_launch



https://www.cosmos.esa.int/web/gaia/instruments

Gaia: ESA's billion star surveyor



Position & Brightness

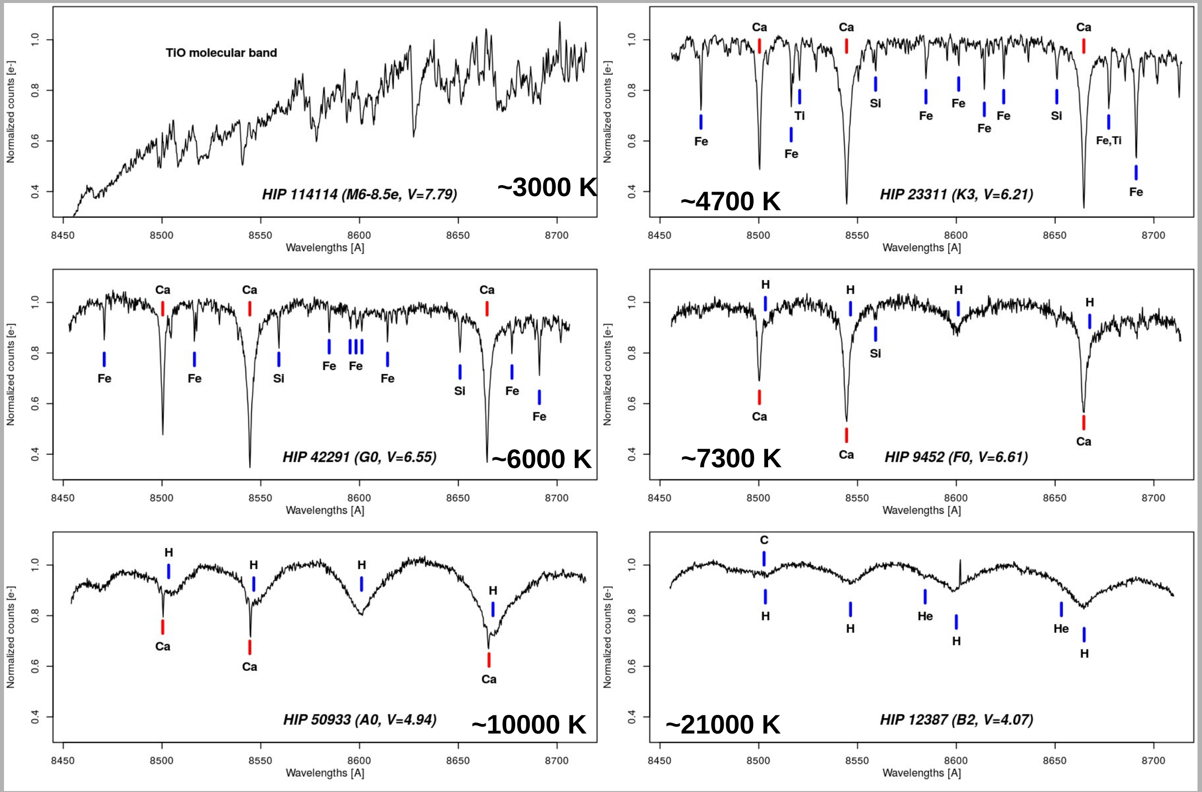
Low-resolution Spectra (Blue & Red)

Intermediate-resolution spectra (abundances :))

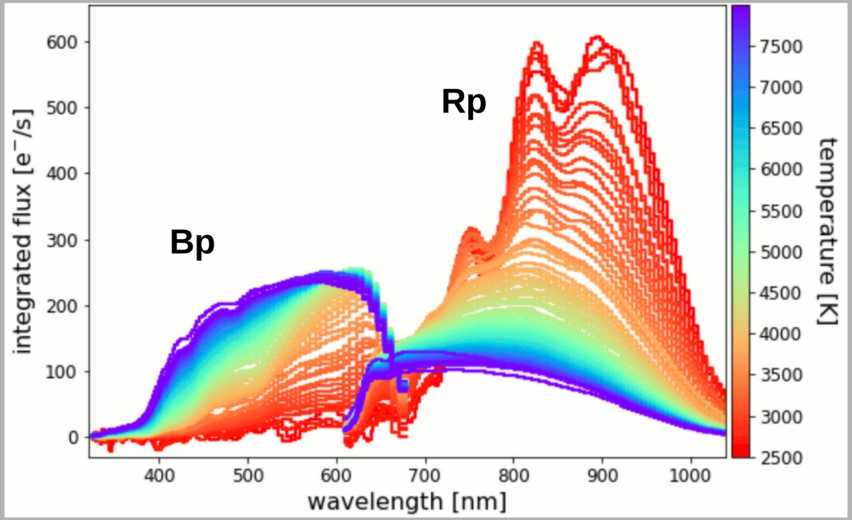
https://www.cosmos.esa.int/web/gaia/payload-module

What Gaia DR3 gave us:

→ 10^6 RVS spectra, R~11500 (Katz et al. 2022)



→ 220 millions BP&RP spectra R~30-100 (De Angeli et al. 2022)



Credit: ESA/Gaia/DPAC- CC BY-SA 3.0 IGO, R. Andrae

→ 1.5×10^9 parallaxes (Lindegren et al. 2021)

→ 1.8×10^9 G mags

→ 1.5×10^9 BP & RB mags

→ See Recio-Blanco et al. 2023 for standard spectroscopic analysis of RVS spectra

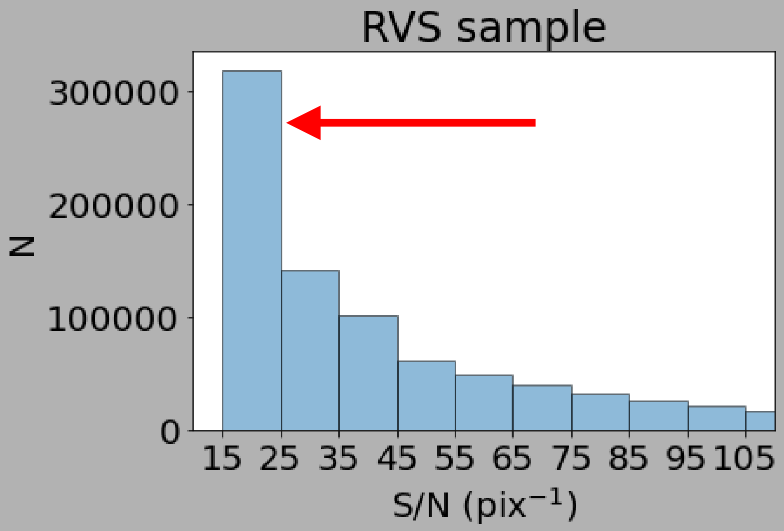
Can we exploit in a homogeneous way *Gaia* spectra (RVS + BP/RP) magnitudes (G, Bp, Rp) and parallaxes for supercharged stellar parametrization ?

Analysis of the 1 million *Gaia* RVS-spectra with CNNs

Beyond *Gaia* DR3:
tracing the $[\alpha/M] - [M/H]$ bimodality
from the Inner to the outer Milky Way disc
with *Gaia* RVS and Convolutional Neural-Networks

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

→ **Resubmitted :)**



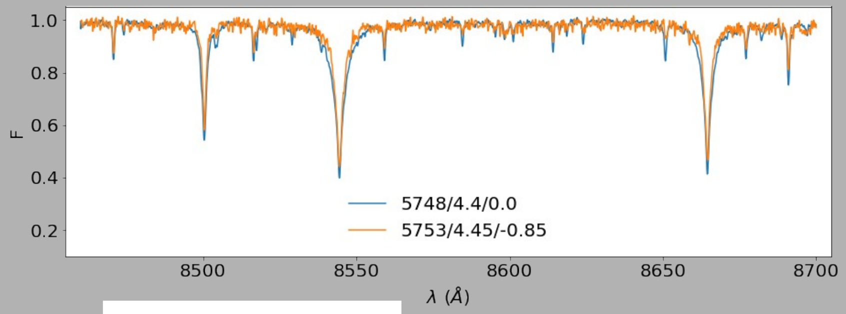
Motivations and goals:

→ **Use homogeneously the full *Gaia* data product**

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

→ **Set the machine-learning path for *Gaia* data analysis (DR4 in 2025, DR5 in 2027)**

Analysis of the 1 million *Gaia* RVS-spectra with CNNs



Training sample

R~22000

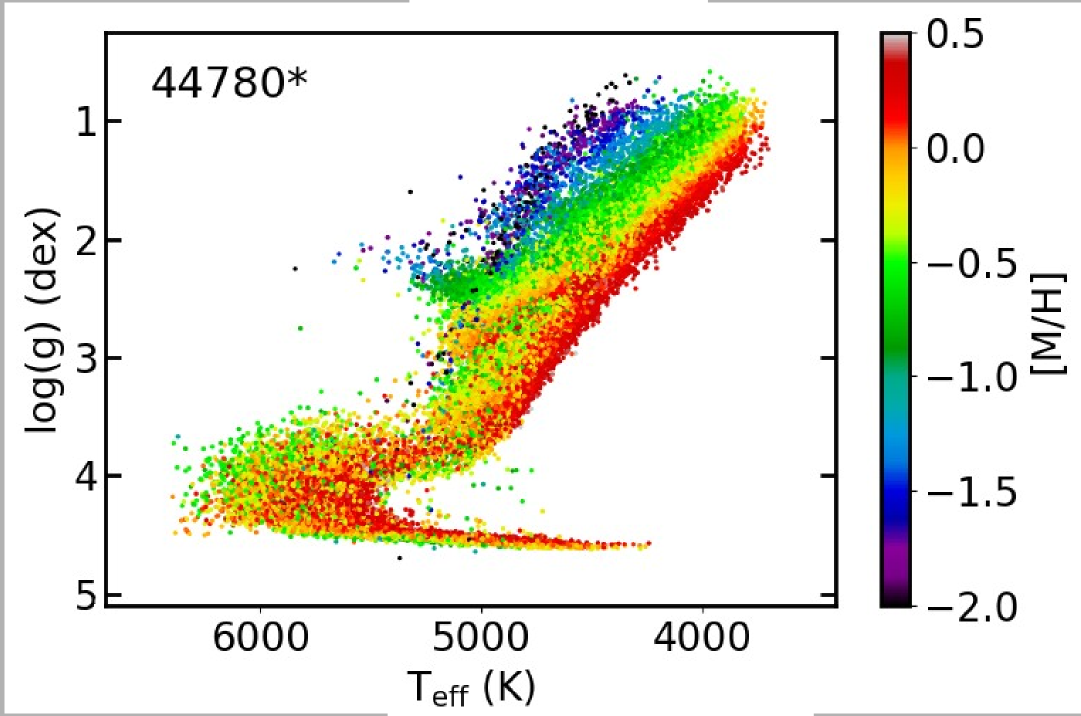
Knowledge transfer from high-quality high-res APOGEE labels

$T_{eff,}$ log(g), [M/H], [α/M], [Fe/H]

to intermediate-res RVS

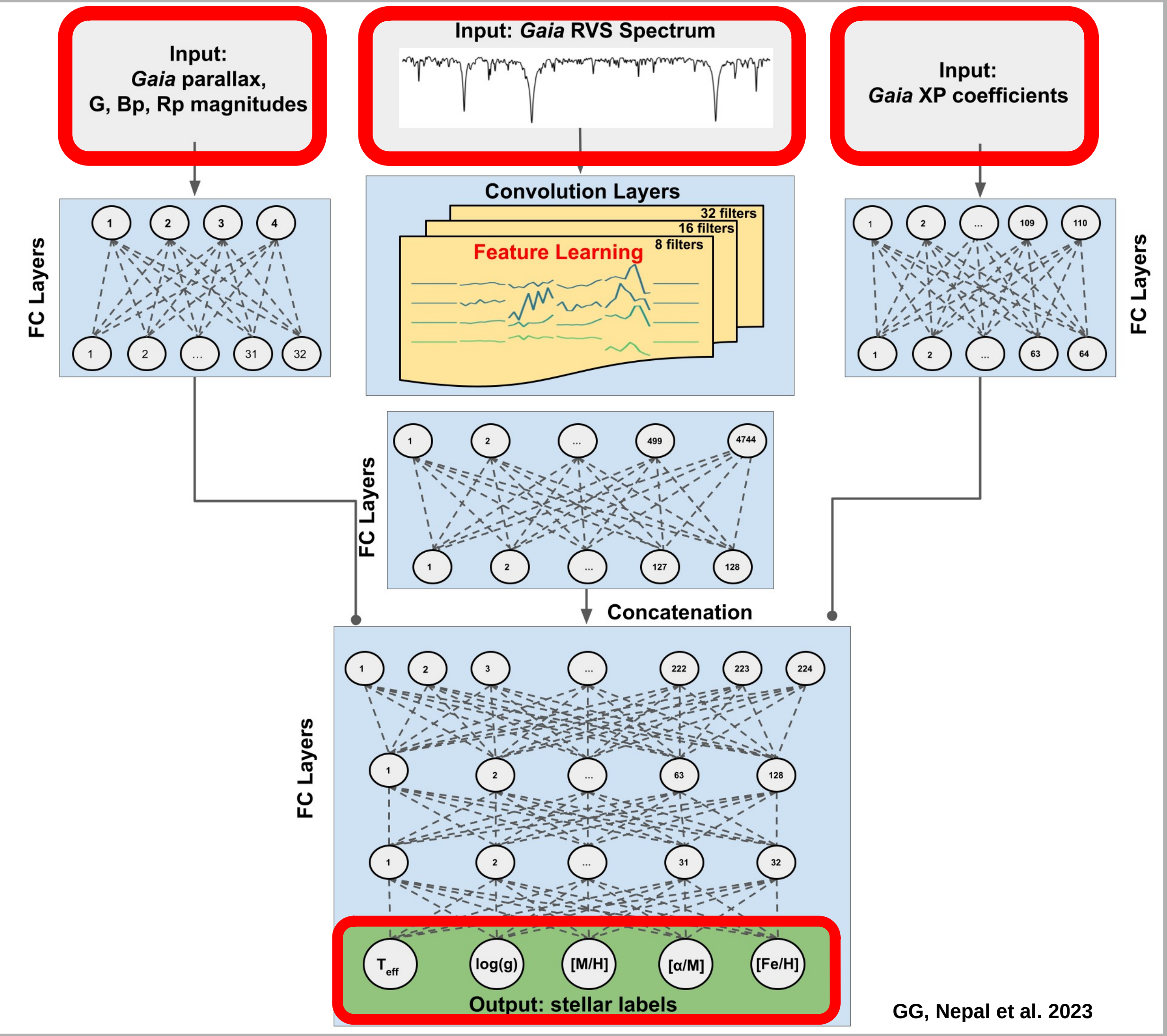
R~11400

Training sample



GG, Nepal et al. 2023

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



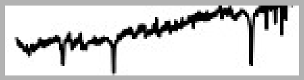
→ **Prediction time**

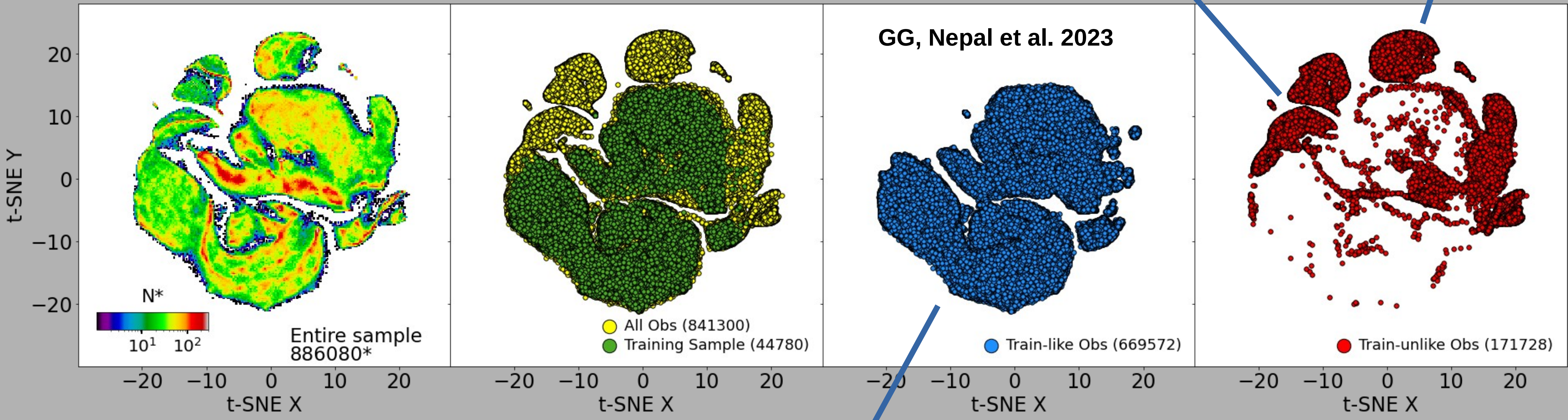
4 labels in 3300 stars / second

How to ensure that a label falls within the training sample limits ?

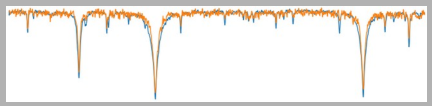
→ **Labels within T_{eff}, log(g), [M/H], [α/M], [Fe/H], G, and parallax limits of training sample.**

→ **t-SNE classification of RVS spectra**



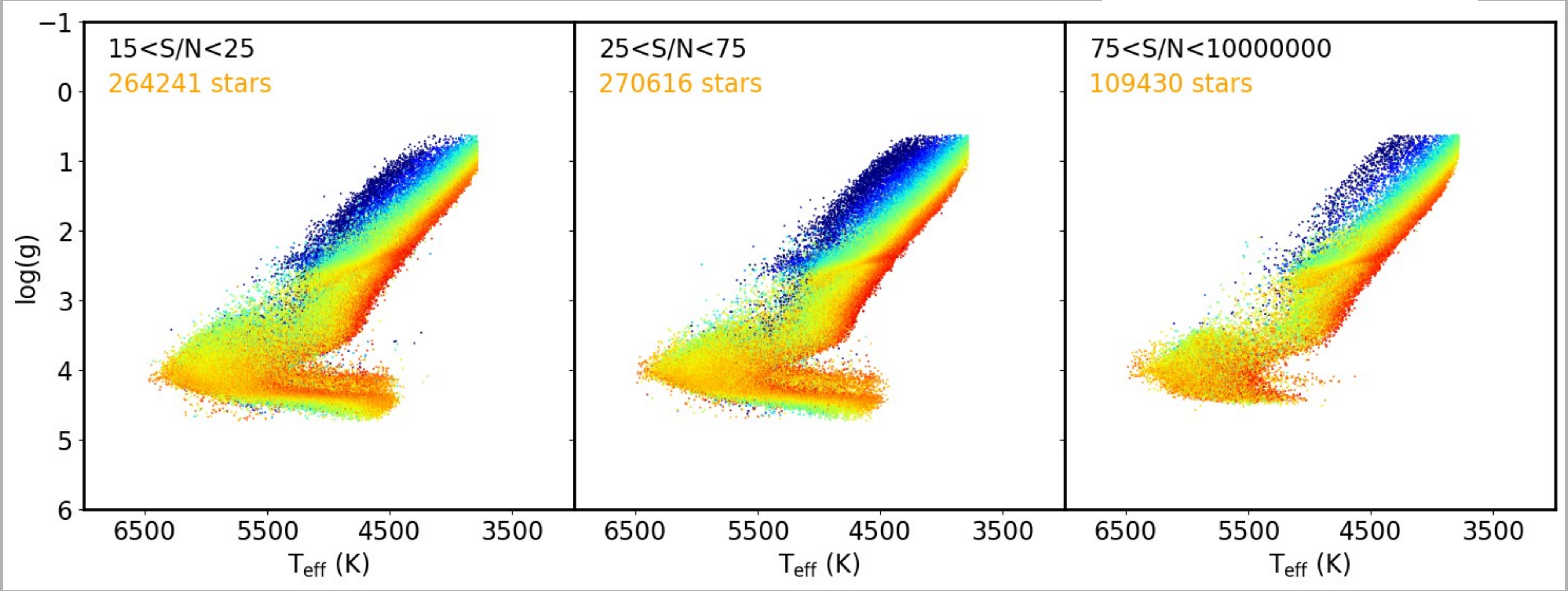


→ **644287 RVS stars within TS**



Robust estimates of T_{eff}, log(g), [M/H] for 690000 *Gaia* stars

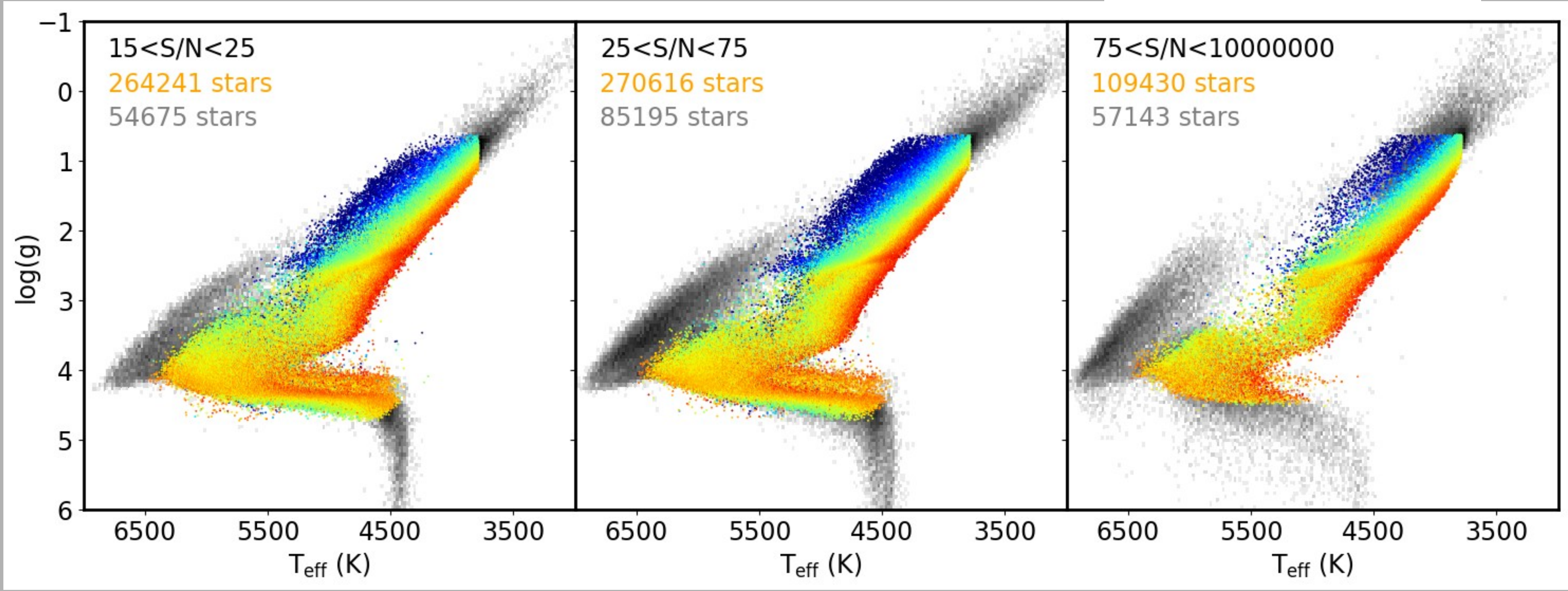
GG, Nepal et al. 2023



→ **By adding magnitudes, parallaxes and XP data, CNN is able to break spectral degeneracies in *Gaia* RVS spectra.**

Robust estimates of T_{eff}, log(g), [M/H] for 690000 *Gaia* stars

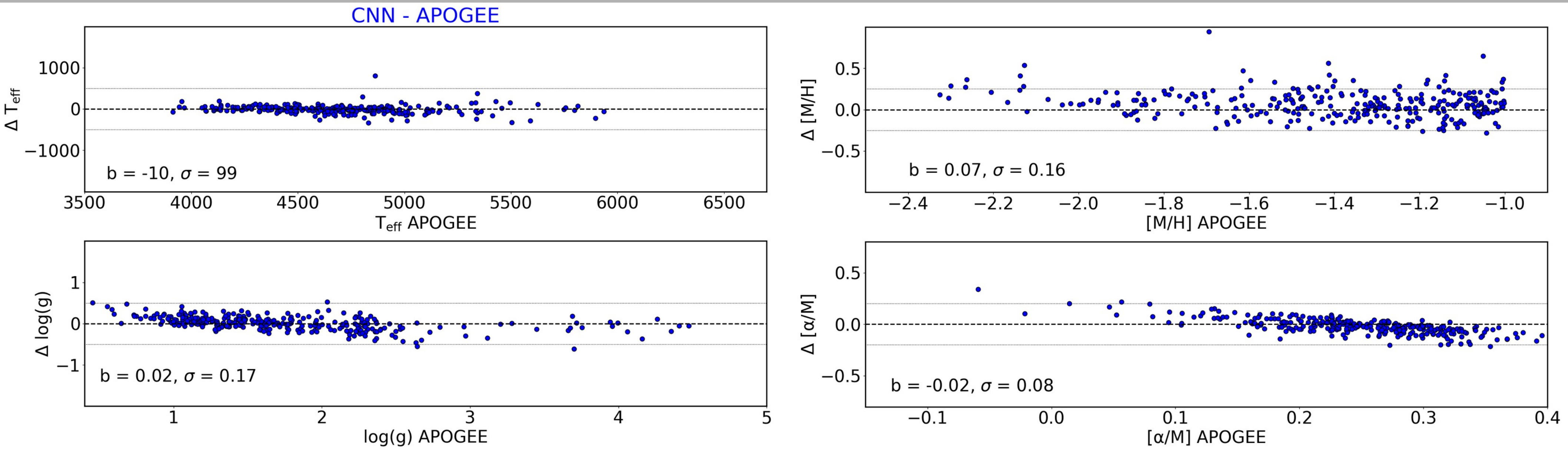
GG, Nepal et al. 2023



→ **By adding magnitudes, parallaxes and XP data, CNN is able to break spectral degeneracies in *Gaia* RVS spectra.**

→ **CNN results are as good as the training set can be.**

CNN performances for halo stars
→ 15<S/N<25



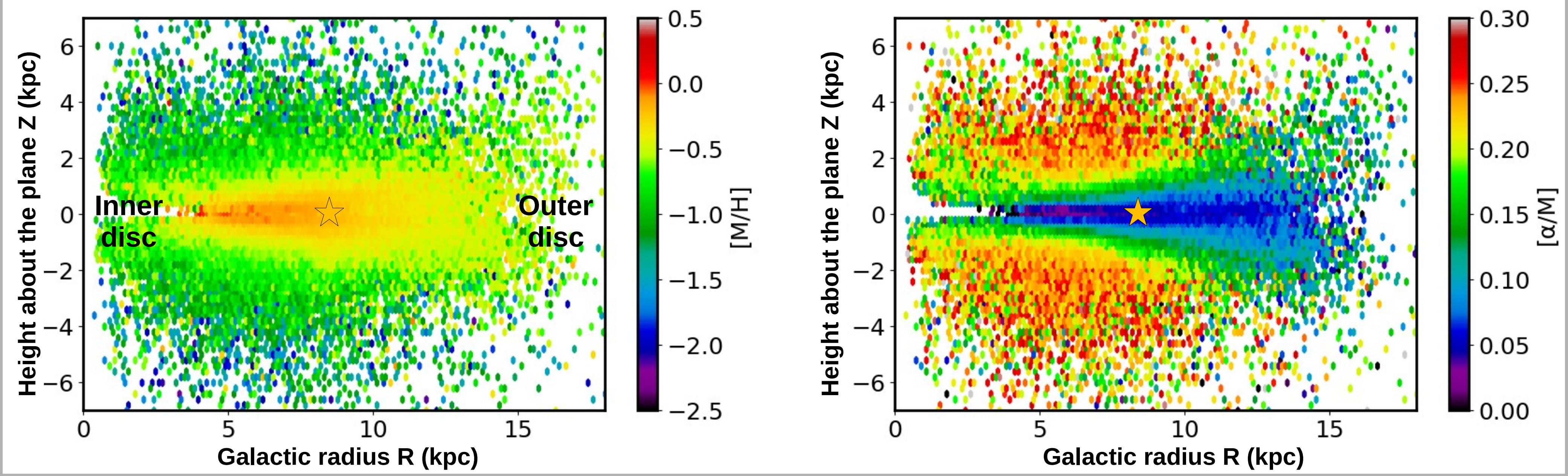
GG et al. 2023

→ CNN provides precise and accurate labels down to [M/H]=-2.4 dex

→ More external validation with GALAH, OCs, and GSP-Phot

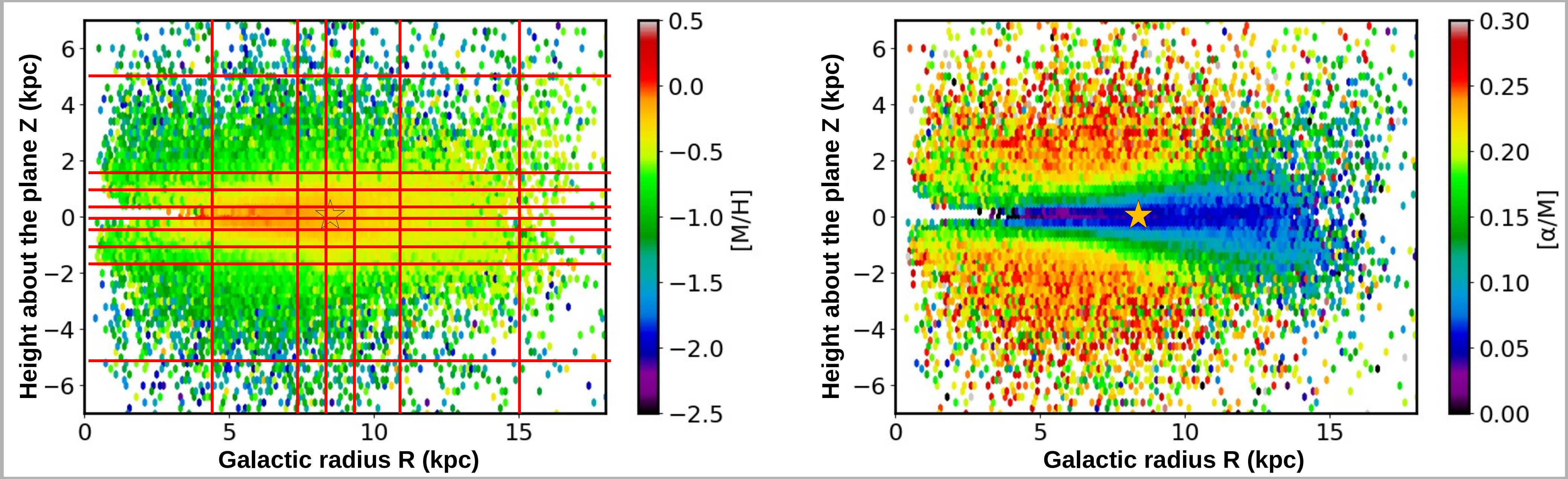
Chemical cartography of the Milky Way, for Inner to Outer regions with *Gaia* and CNN

→ We selected giants, to probe large distances, and limit possible systematics → 147416 stars
→ Galactic radius and Height adopted from Nepal et al. In prep. (using StarHorse distances).



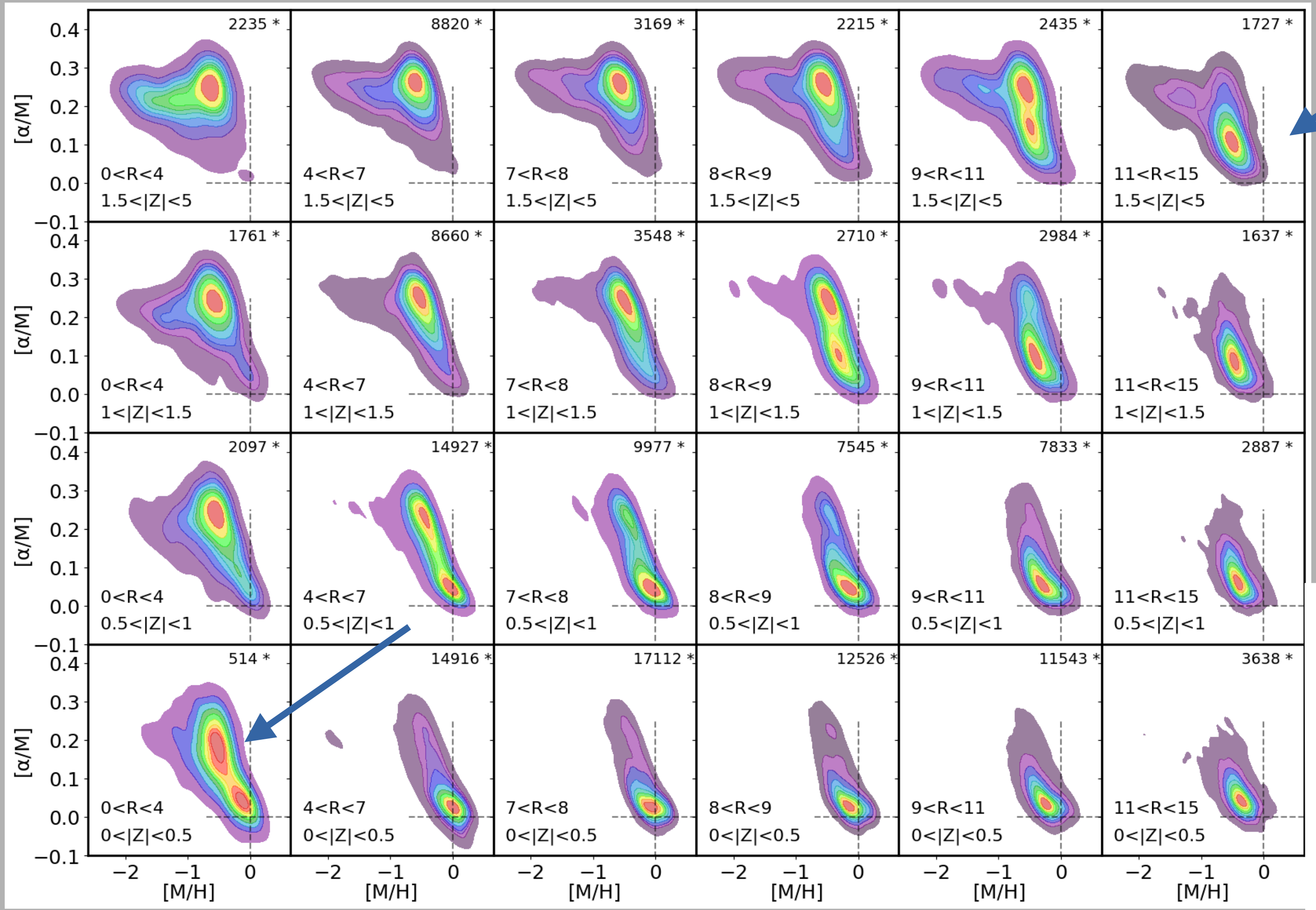
Chemical cartography of the Milky Way, for Inner to Outer regions with *Gaia* and CNN

→ We selected giants, to probe large distances, and limit possible systematics → 147416 stars
→ Galactic radius and Height adopted from Nepal et al. In prep. (using StarHorse distances).



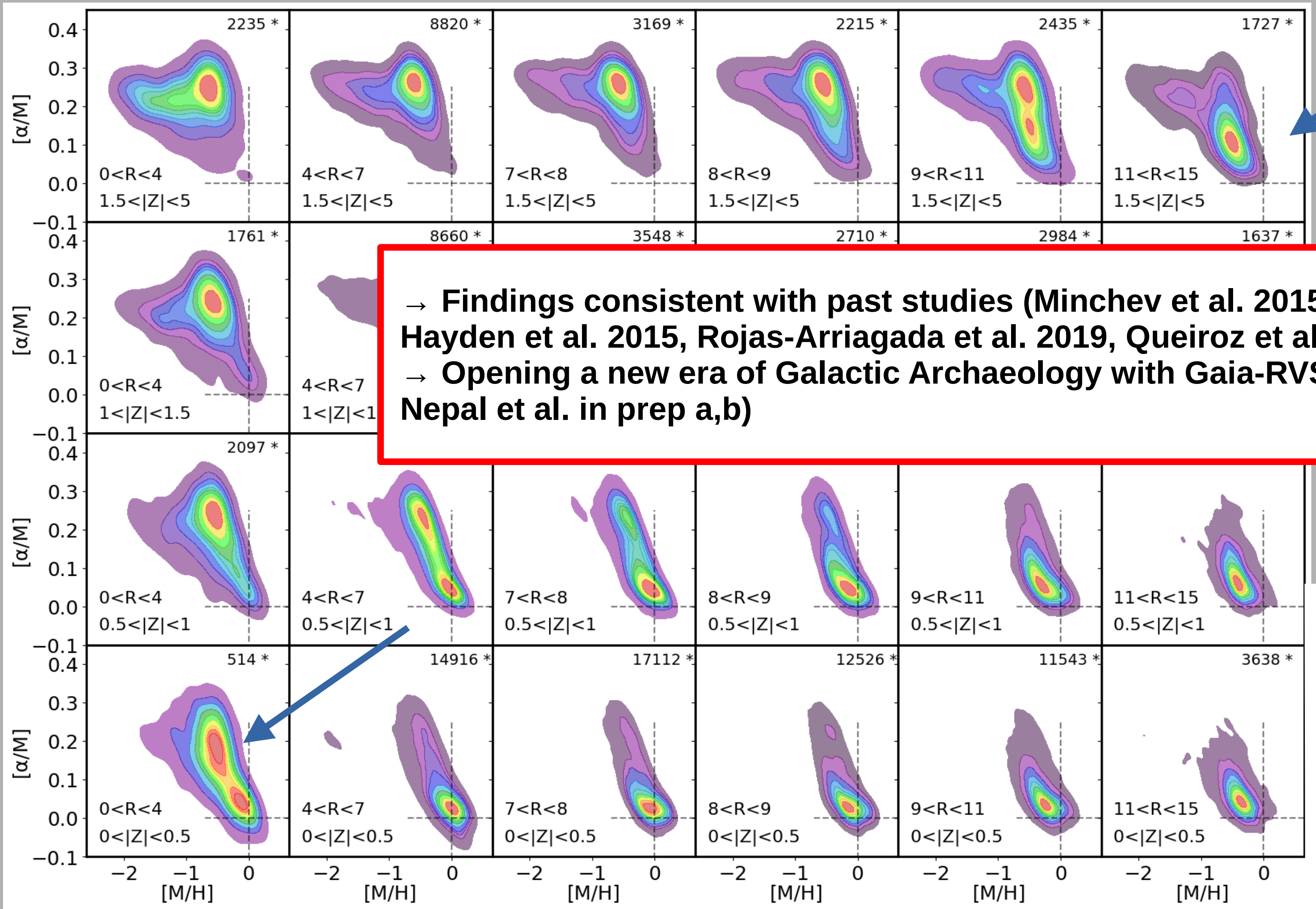
→ Studying the chemical abundance pattern [α/M] vs. [M/H] as function of R and Z

Chemical cartography of the Milky Way, for Inner to Outer regions with *Gaia* and CNN



GG, Nepal et al. 2023

Chemical cartography of the Milky Way, for Inner to Outer regions with *Gaia* and CNN

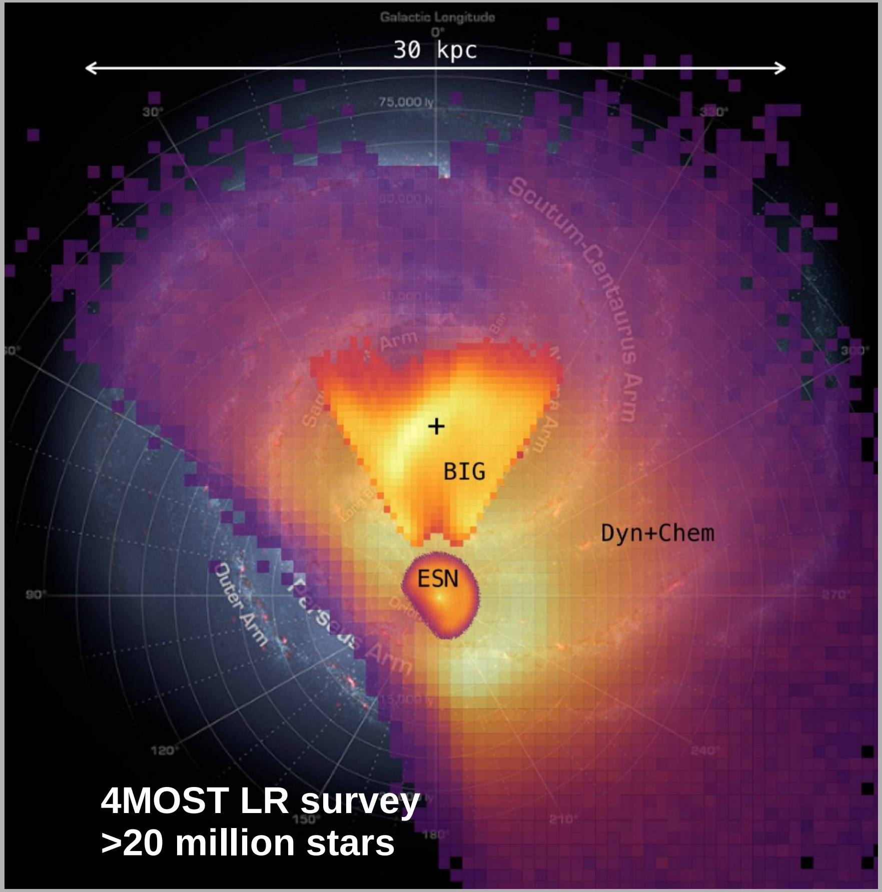


→ Findings consistent with past studies (Minchev et al. 2015, Anders et al. 2014, Hayden et al. 2015, Rojas-Arriagada et al. 2019, Queiroz et al. 2020, 2021).
→ Opening a new era of Galactic Archaeology with Gaia-RVS (Guiglion et al. In prep, Nepal et al. in prep a,b)

GG, Nepal et al. 2023

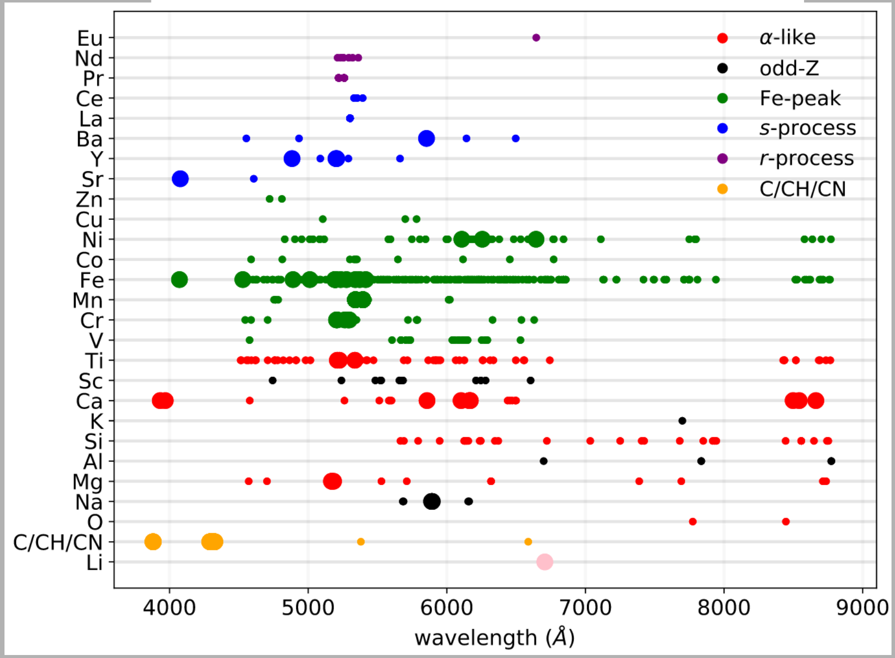
Future application of CNNs for stellar spectroscopy

4MOST (de Jong et al. 2019)
→ 4MIDABLE-LR Disc and Bulge surveys (Chiappini et al. 2019)



4MIDABLE-LR ESO proposal 2020

>20 elements to be measured at R=5000



4MIDABLE-LR ESO proposal 2020

→ **Developing CNN for 4MIDABLE-LR D1(>) spectral analysis.**

Summary:

- CNN is an optimal method for combining full *Gaia* data product
  → Leveraging the large set of low S/N RVS spectra

- CNN parametrization is fast and robust (several 10^3 stars per second)

Insights:

- Standard spec. and ML methods complement each other !

- Future spectroscopic surveys will strongly benefit from CNNs

- CNN parametrization mainly reliable within the training sample limits
  → The training sample should be built in a pro-active way
  E.g.: IWG3 for 4MOST

guiglion@mpia.de