The quest to undercover Open Clusters parameters

using QuadTrees and Artifical Neural Networks

L. Cavallo*, L. Spina, G. Carraro, L. Magrini, E. Poggio, T. Cantat-Gaudin, M. Pasquato, S. Lucatello, S. Ortolani, J. Ulloa-Schiappacasse

*Lorenzo Cavallo

Phd student

Università degli Studi di Padova

Firenze - November 2023

The Milky Way is a spiral galaxy like thousands ones in the Universe

Credit: European Space Agency & NASA

Credit: European Space Agency & NASA

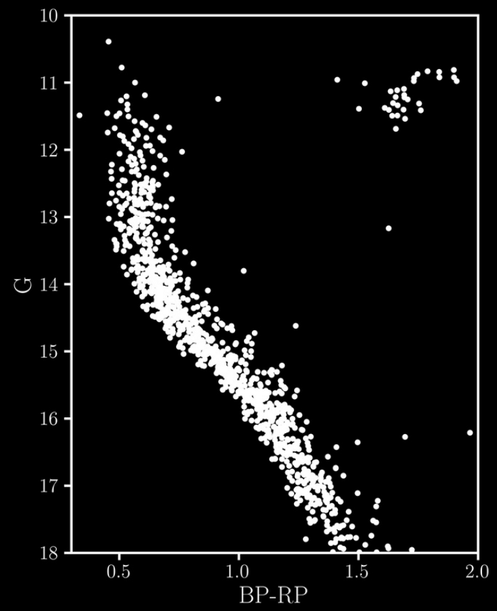
We still have a series of open questions:

How was the Galactic disk formed? Does it have a structure? Is it evolving with time? Do insulated stars form alone or are the relics of disrupting clusters? etc...

OPEN CLUSTERS

Groups of stars that have formed together and thus they **share the same age, and chemical composition** and are bounded by gravity.

Credit: ESO/G. Beccari

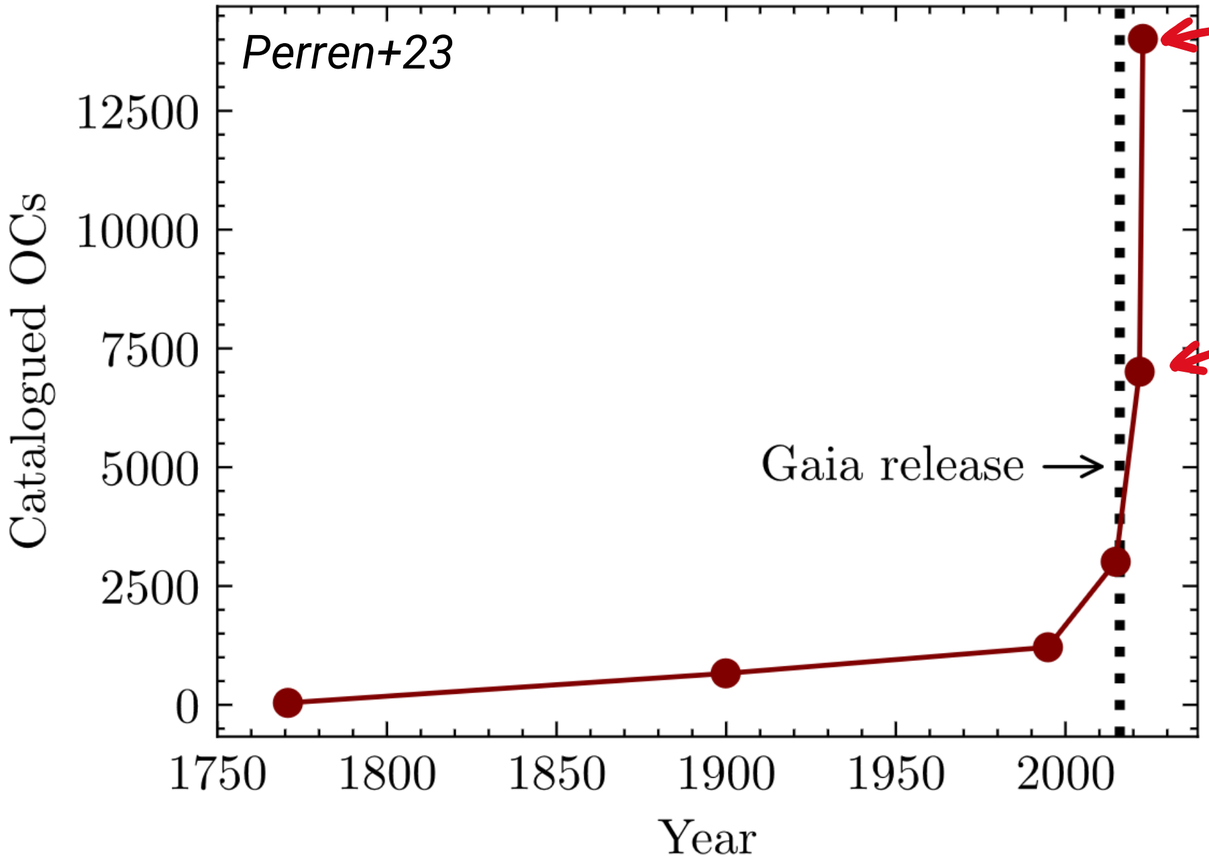


Members from Hunt & Reffert 2023

Their age, metallicity, and distance can be precisely (at least **more than in the case of individual stars**) recovered using photometry

WHAT IS HAPPENING NOW?

long story short: a lot of new clusters!



Perren et al. (2023) —— ~14000 ——

Hunt & Reffert (2023) —— ~7500 ——

!!! TOO MANY CLUSTERS !!!

automated tools for the characterisation of OCs are needed

They should be **<u>fast and reliable</u>**. At the same time, **<u>efficiently extract informations</u>** from the availble data and future ones (e.g. **<u>Gaia DR4 and/or LSST</u>**)

PREVIOUS SOLUTIONS

Cantat-Gaudin et al. (2020)

Used an **Artificial Neural Network** trained with a set of well-studied clusters available in literature

Hunt & Reffert (2023)

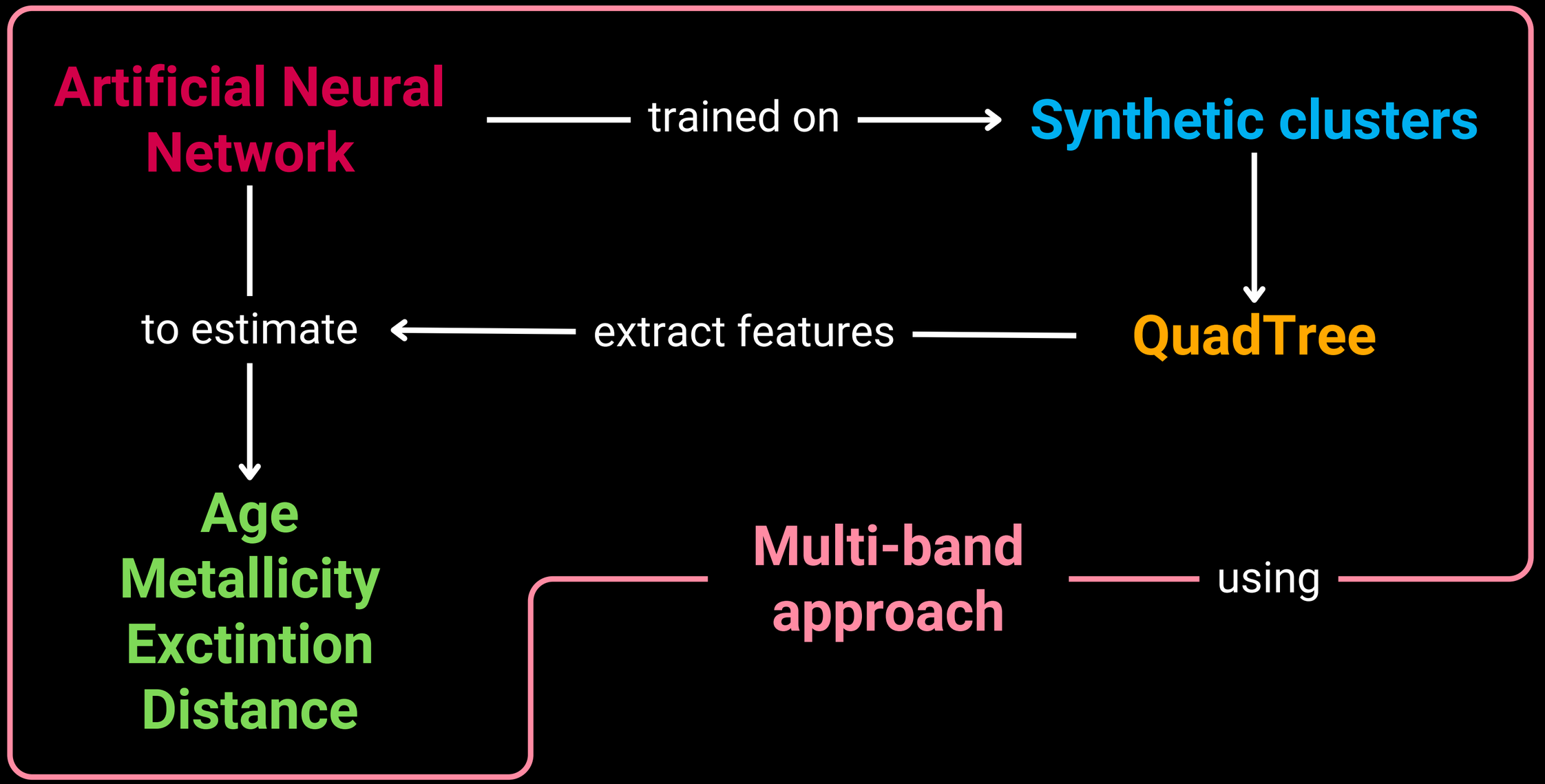
Used an **Artificial Neural Network** trained with synthetic clusters

Dias et al. (2021)

Minimize likelihood function imposing strong priors (e.g. metallicity gradient and extinction map) to reduce the parameters space

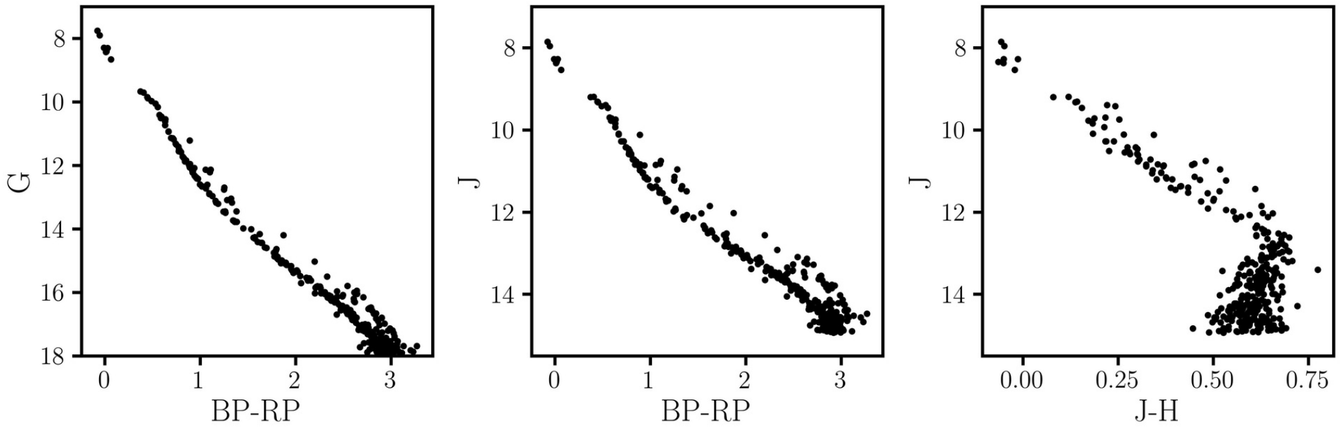
OUR SOLUTION

arXiv:2311.03009



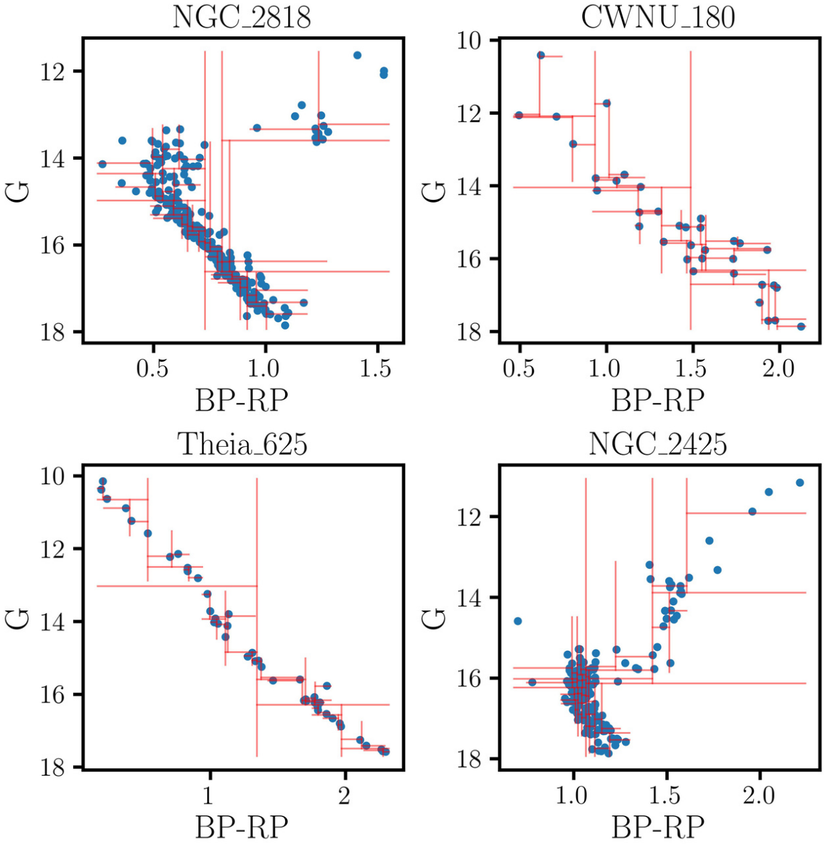
MULTI-BAND APPROACH

Use both optical (Gaia) and near-infrared (2MASS) photometry



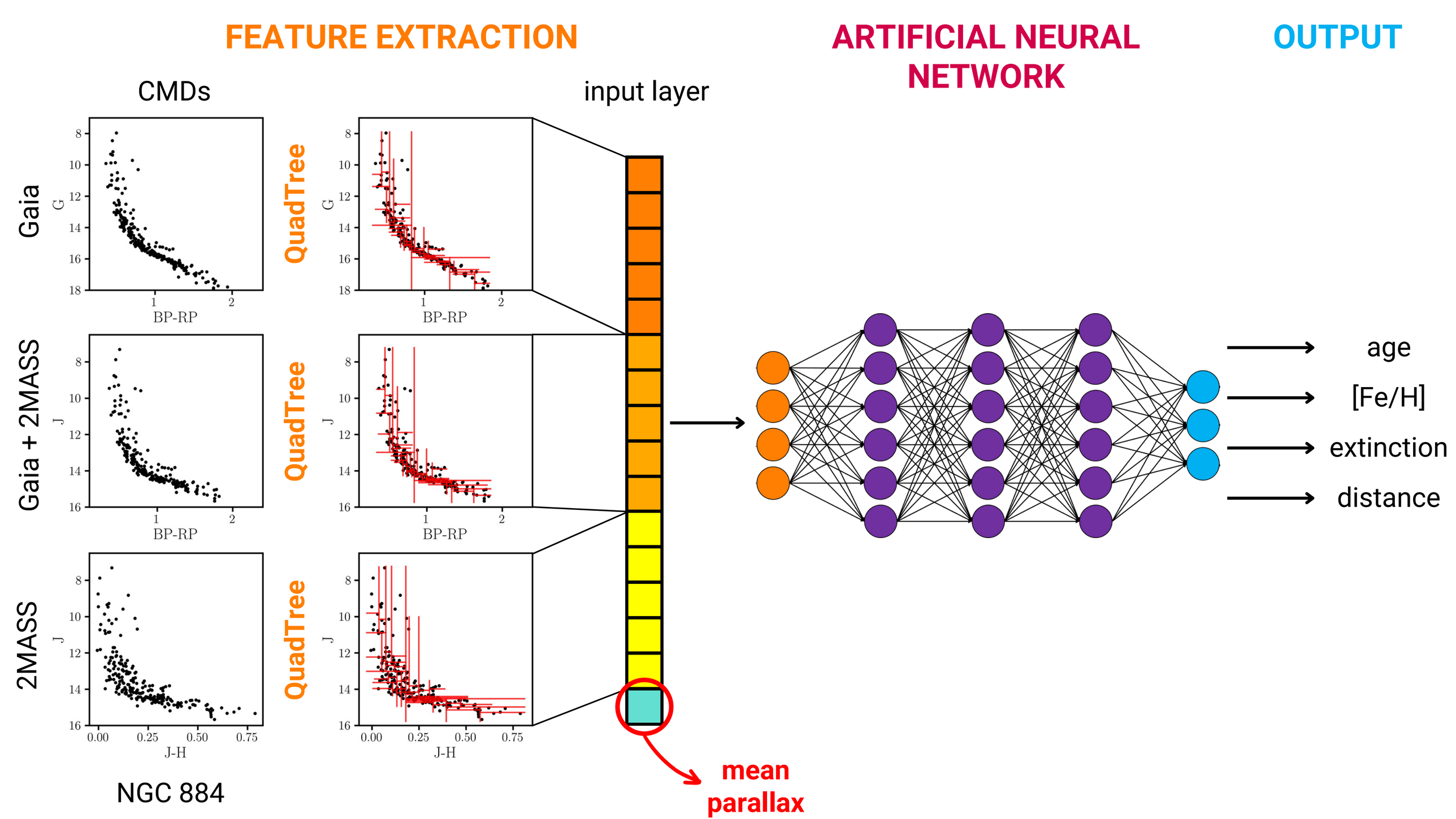
QUADTREE

Tree data structure which is used to represent object distribution in 2-dimensional space



METALLICITY ESTIMATION

Metallicity determination solely based on photometry is an extremely difficult task. However, with the adoption of the **multi-band approach** and the use of the **QuadTree** **is possible to extract this information from CMDs**


FEATURE EXTRACTION
ARTIFICIAL NEURAL NETWORK
OUTPUT
CMDs
input layer
Gaia
Gaia + 2MASS
2MASS
QuadTree
QuadTree
QuadTree
BP-RP
J-H
G
J
age
[Fe/H]
extinction
distance
mean parallax
NGC 884

FEATURE EXTRACTION

ARTIFICIAL NEURAL NETWORK

OUTPUT

CMDs

input layer

Gaia

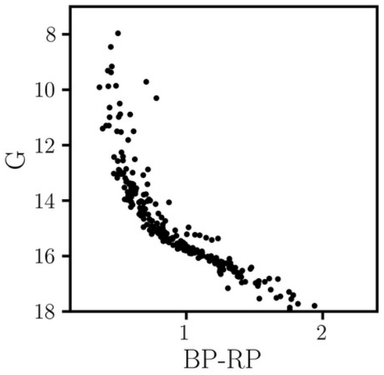
Gaia + 2MASS

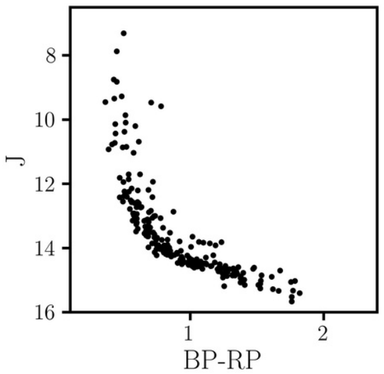
2MASS

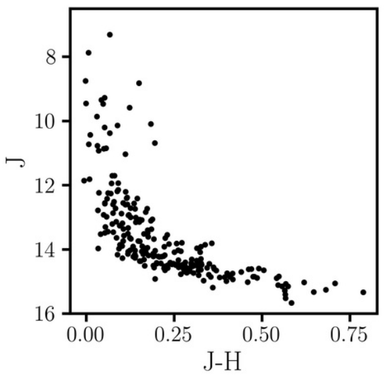
QuadTree

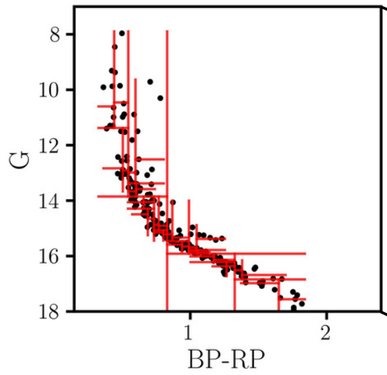
QuadTree

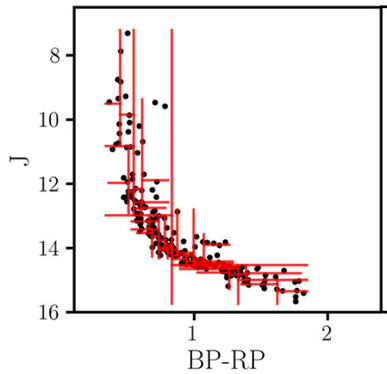
QuadTree

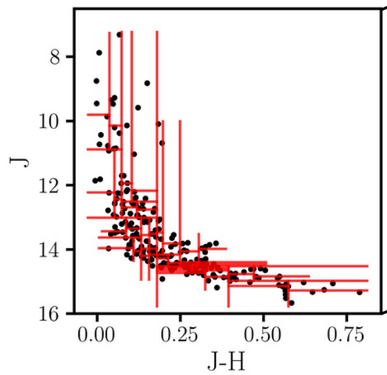


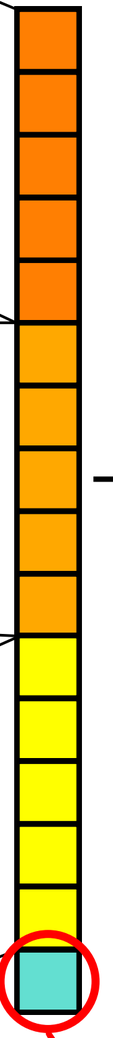


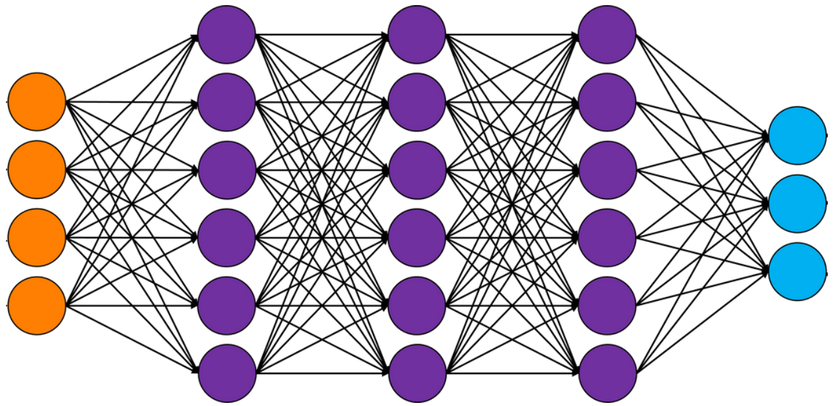












age

[Fe/H]

extinction

distance

An Artificial Neural Network is deterministic

mean parallax

NGC 884

FEATURE EXTRACTION

ARTIFICIAL NEURAL NETWORK

CMDs

input layer

Gaia

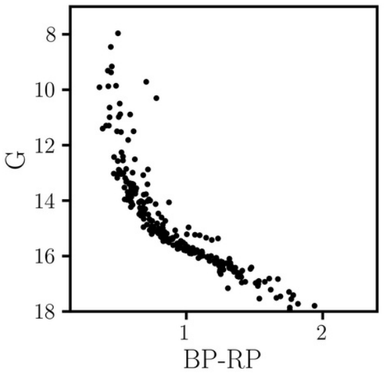
Gaia + 2MASS

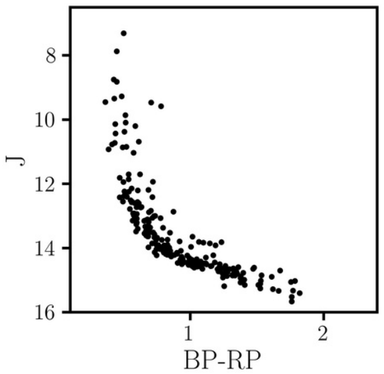
2MASS

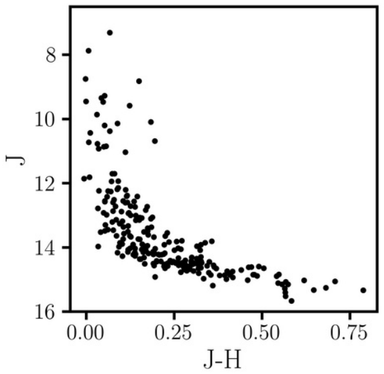
QuadTree

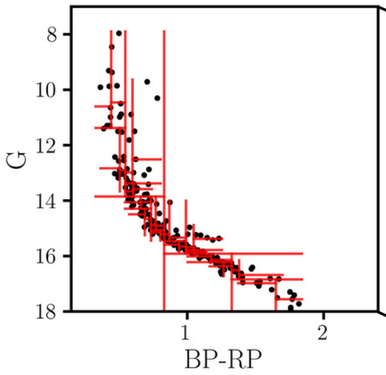
QuadTree

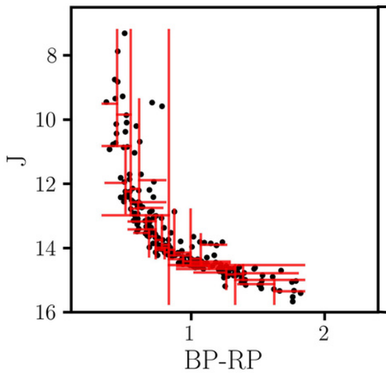
QuadTree

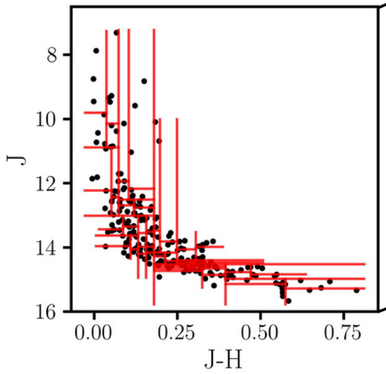


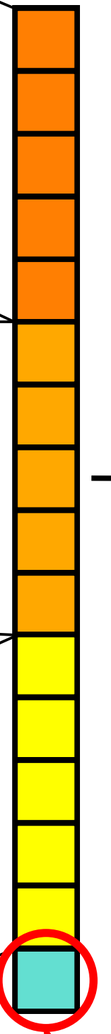


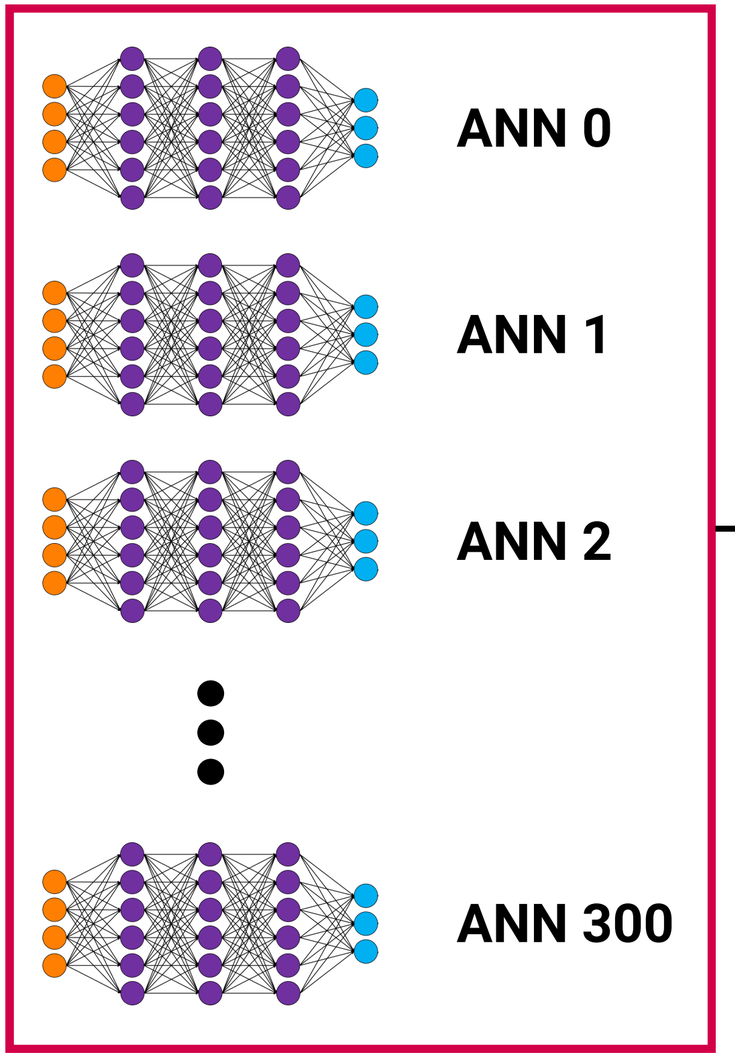










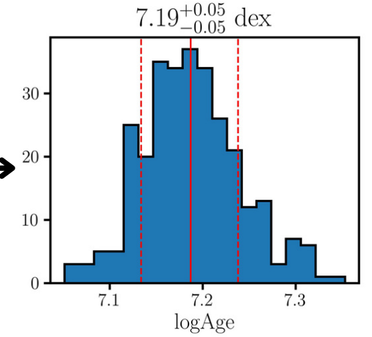


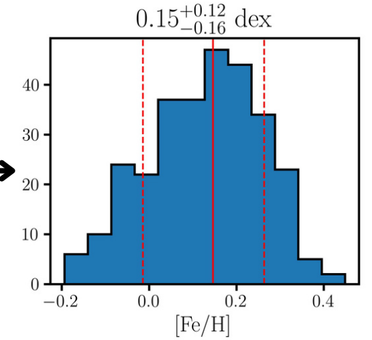
ANN 0

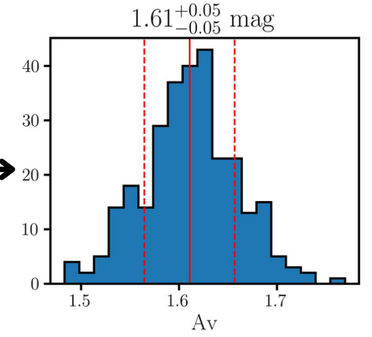
ANN 1

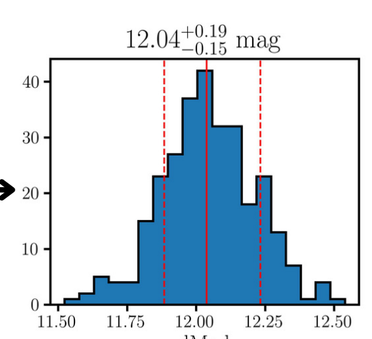
ANN 2

ANN 300







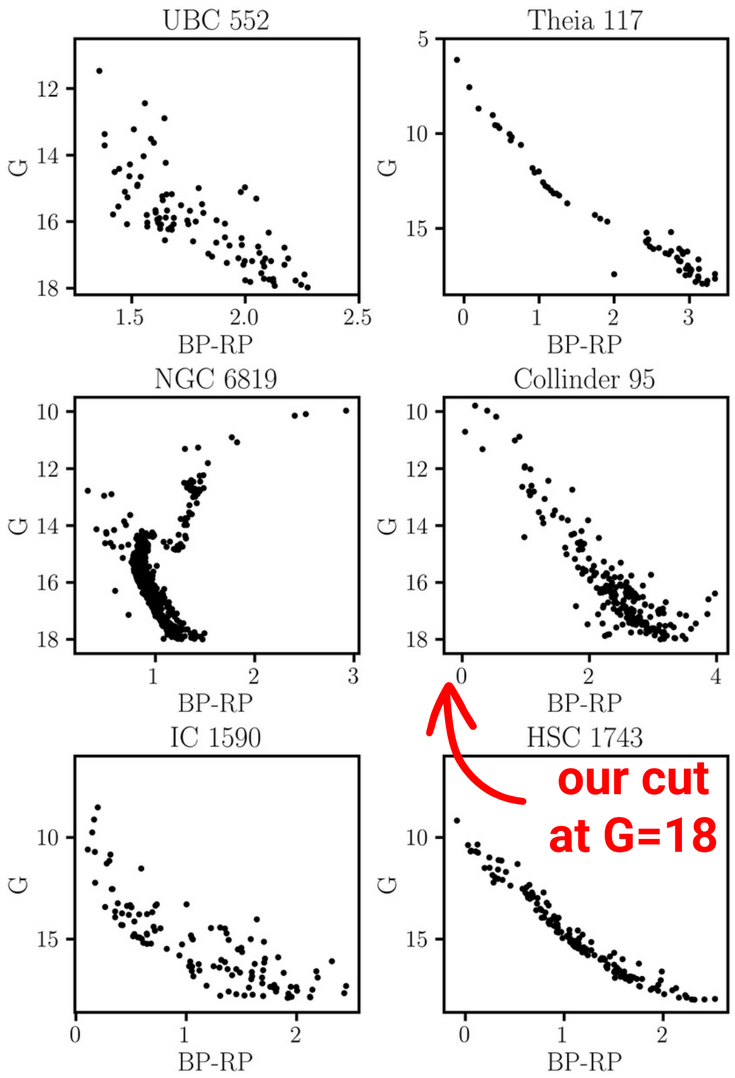


mean parallax

NGC 884

SAMPLE OF CLUSTER

Hunt & Reffert (2023)



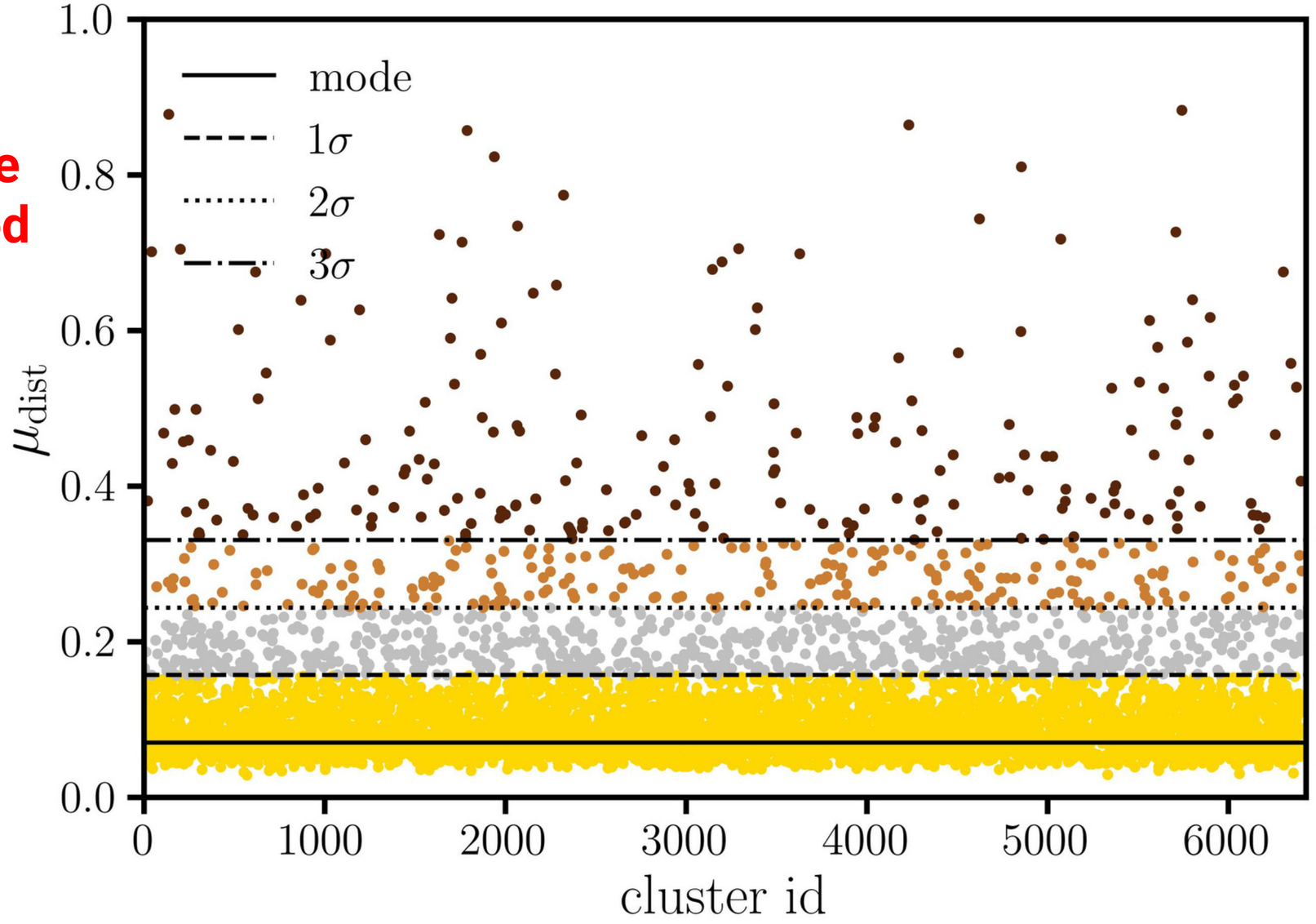
- **<u>Blind all-sky search</u>** using 729 million sources from *Gaia* DR3 down to G~20

- Using the HDBSCAN algorithm they recover **<u>7167 clusters</u>** **(<u>2387 new candidates</u>)**

We select the **<u>stellar associations with more then 10 members</u>** (in all the CMDs) ending up with <u>6413 clusters</u>

HOW GOOD ARE OUR ESTIMATES?

“mean distance from the inferred isochrone”



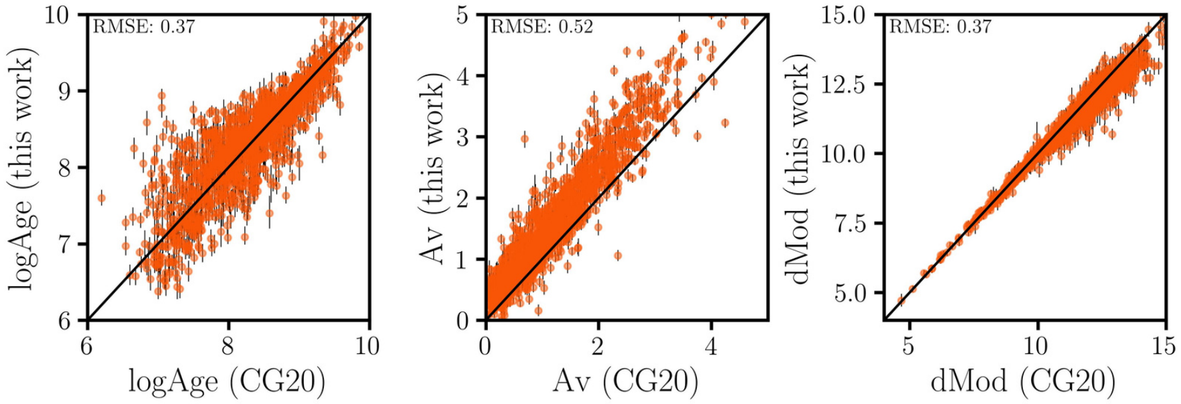
WOOD
3%

BRONZE
4%

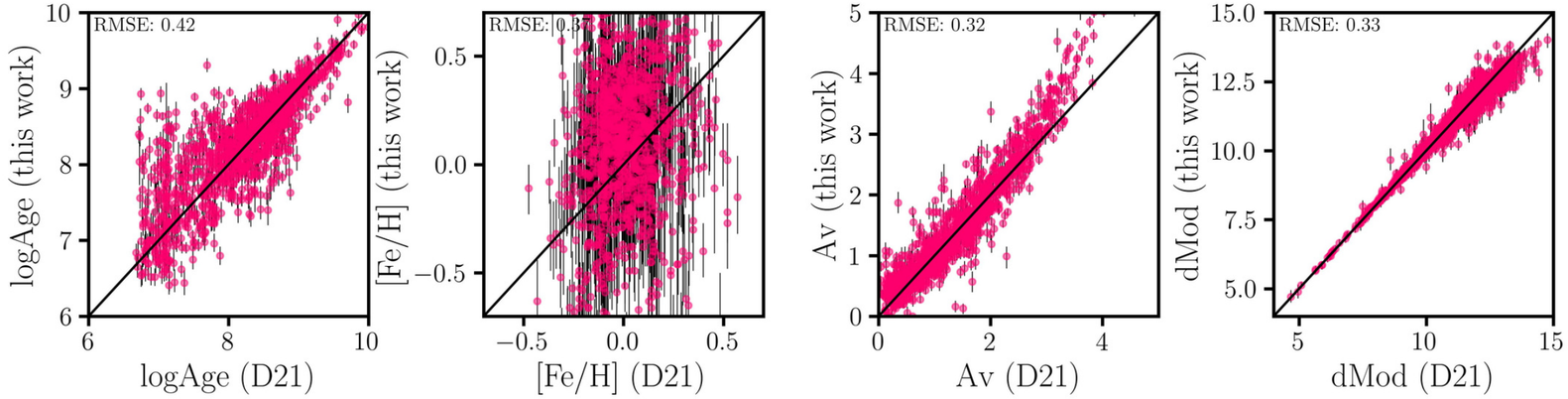
SILVER
11%

GOLD
82%

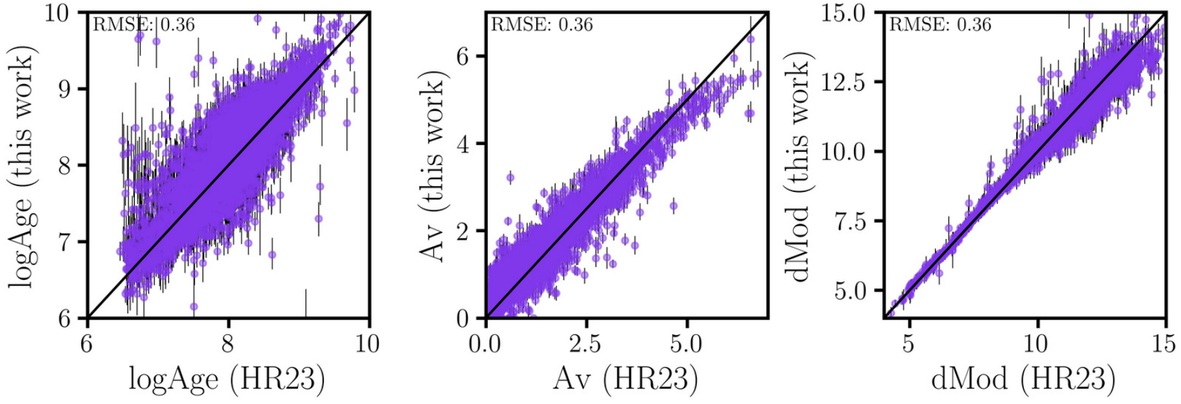
Cantat-Gaudin et al. (2020)



Dias et al. (2021)



Hunt & Reffert (2023)

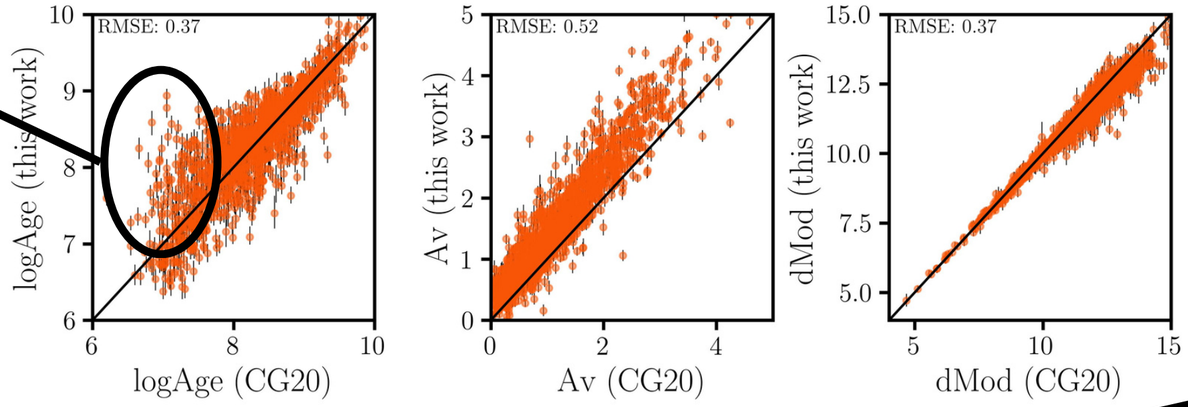


Young clusters (for CG20 and D21) are older for our ANN

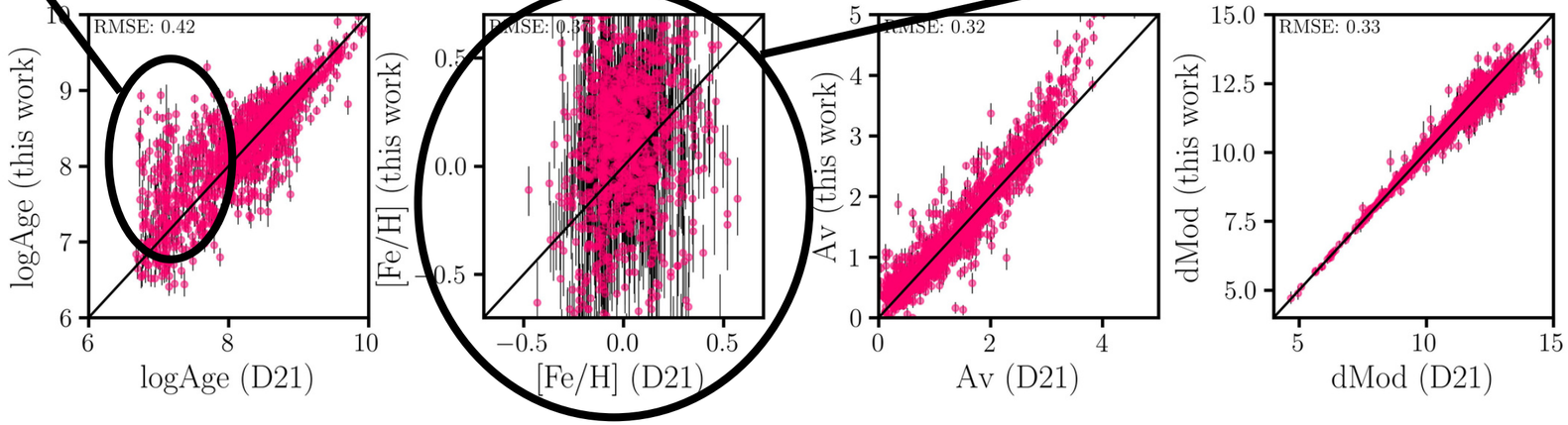
Random values of [Fe/H] (?)

Large scatter at intermediate ages logAge ~ 8

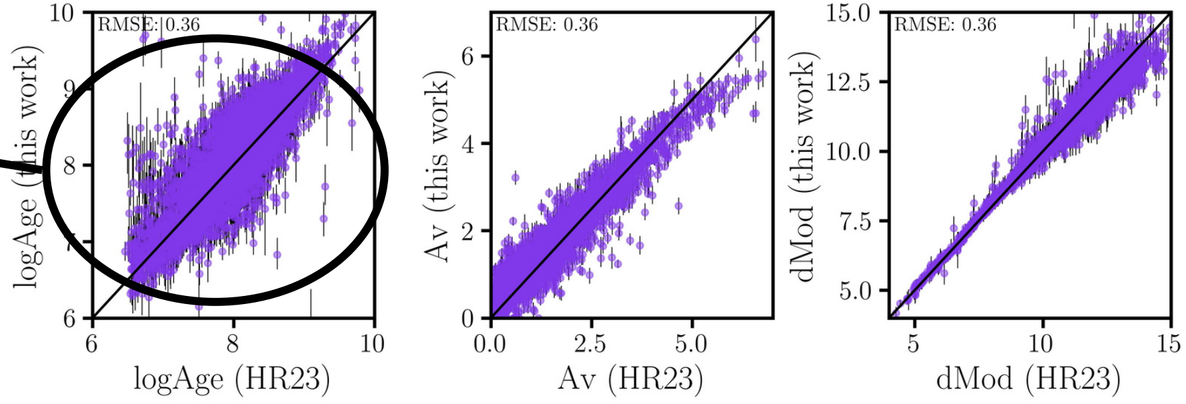
Cantat-Gaudin et al. (2020)



Dias et al. (2021)

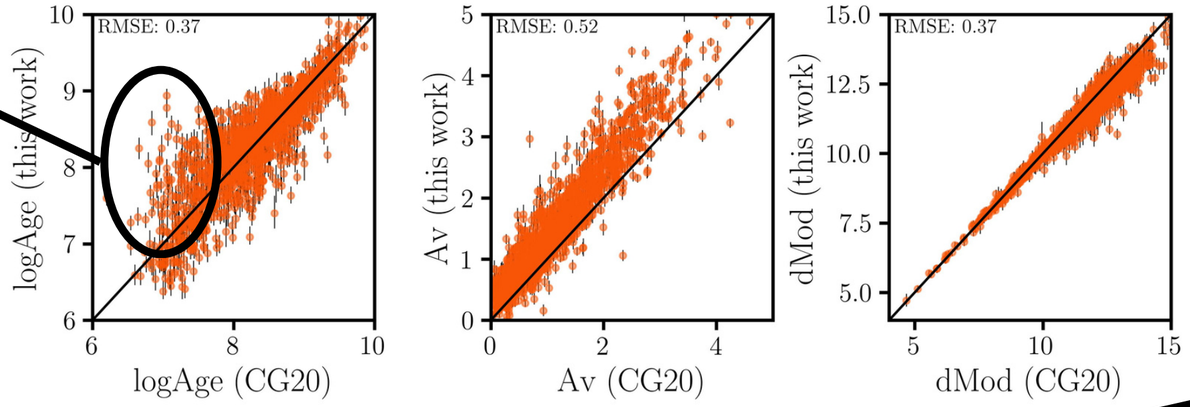


Hunt & Reffert (2023)

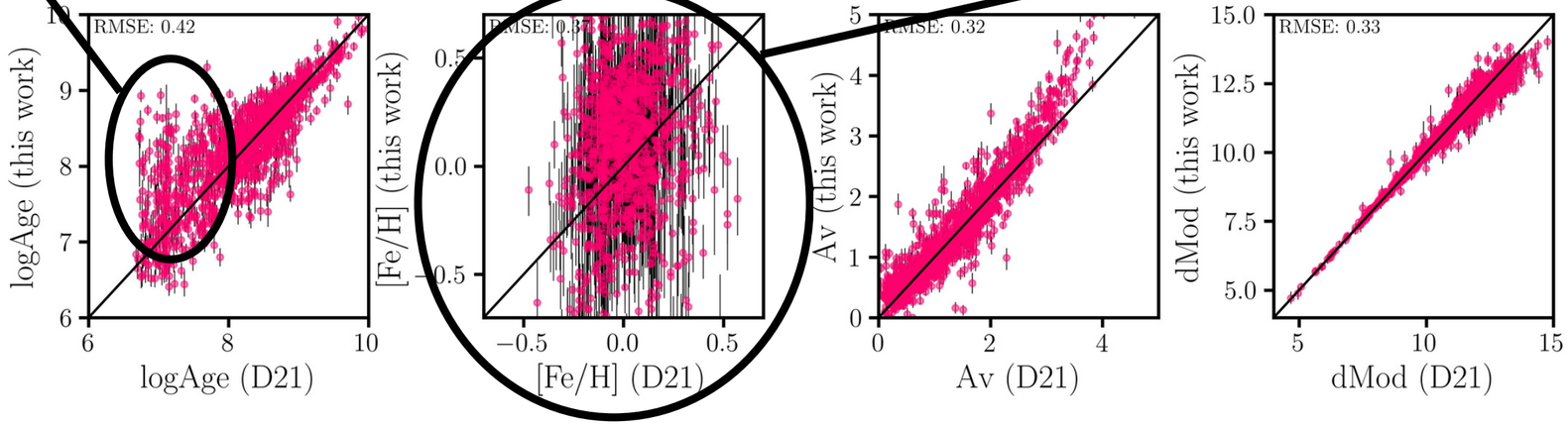


Young clusters (for CG20 and D21) are older for our ANN

Cantat-Gaudin et al. (2020)



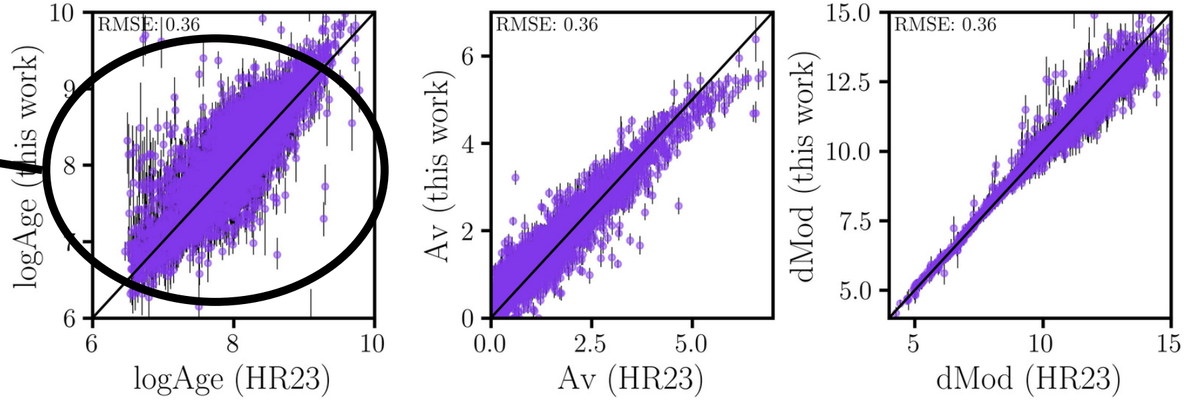
Dias et al. (2021)



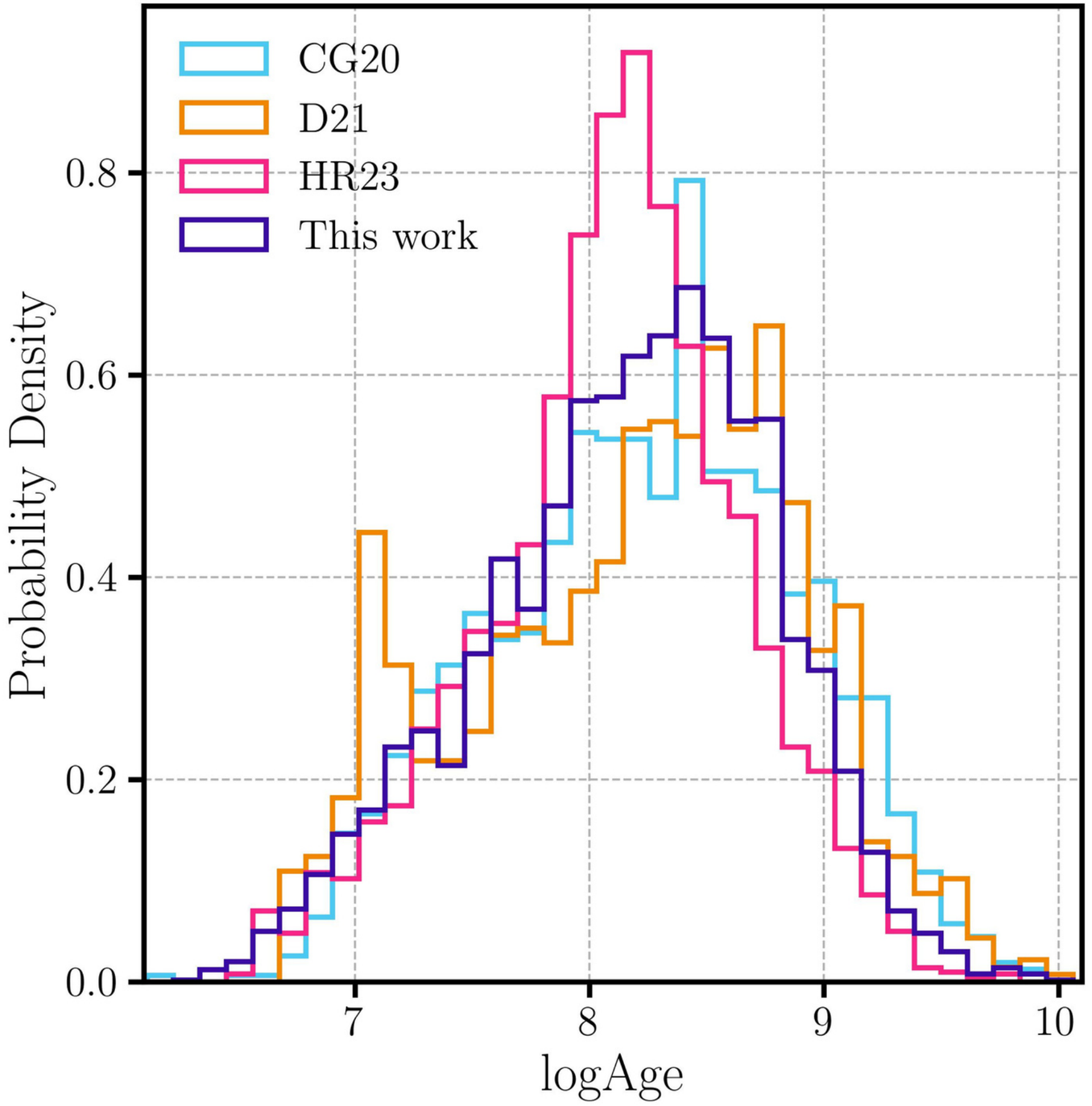
Random values of [Fe/H] (?)

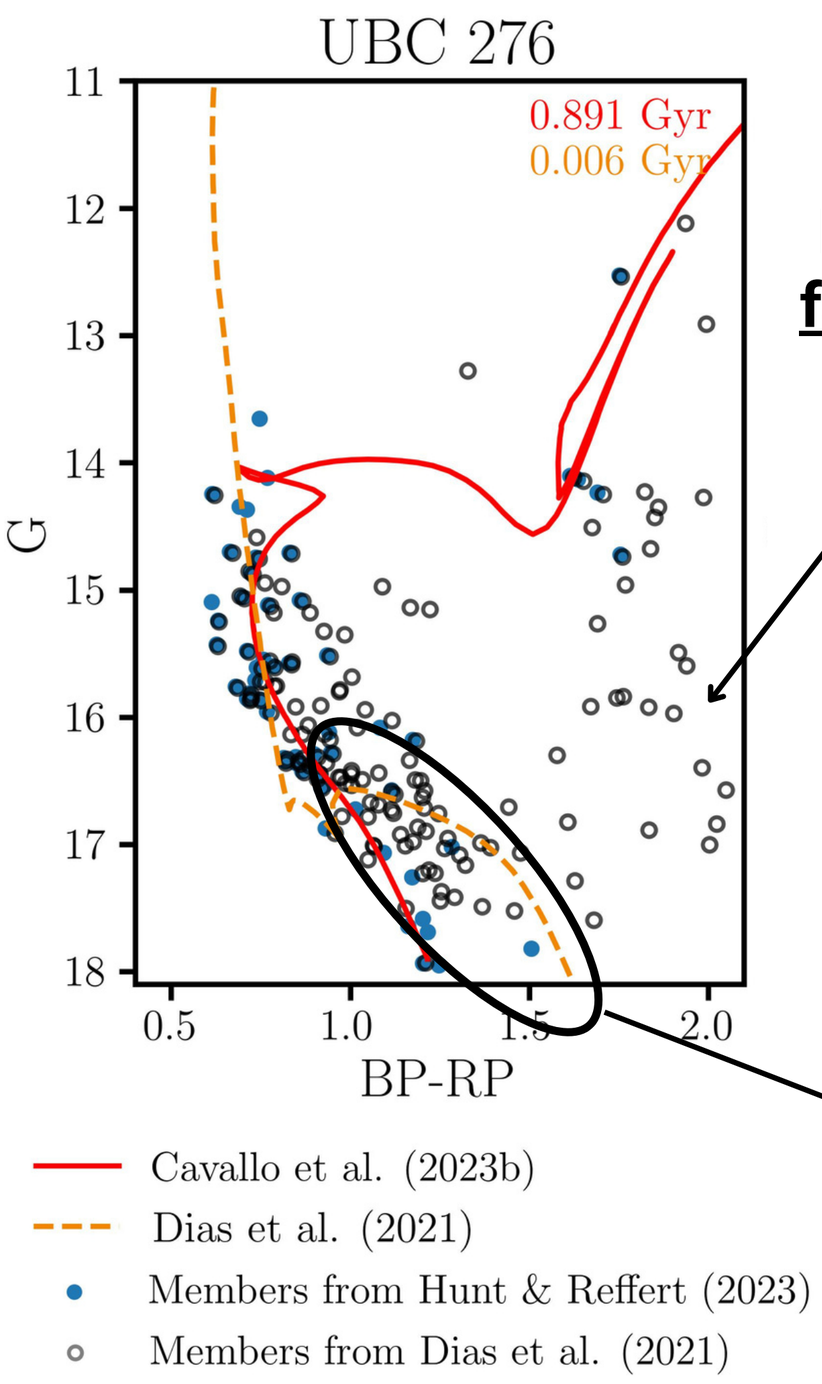
SPOILER: NO!

Hunt & Reffert (2023)



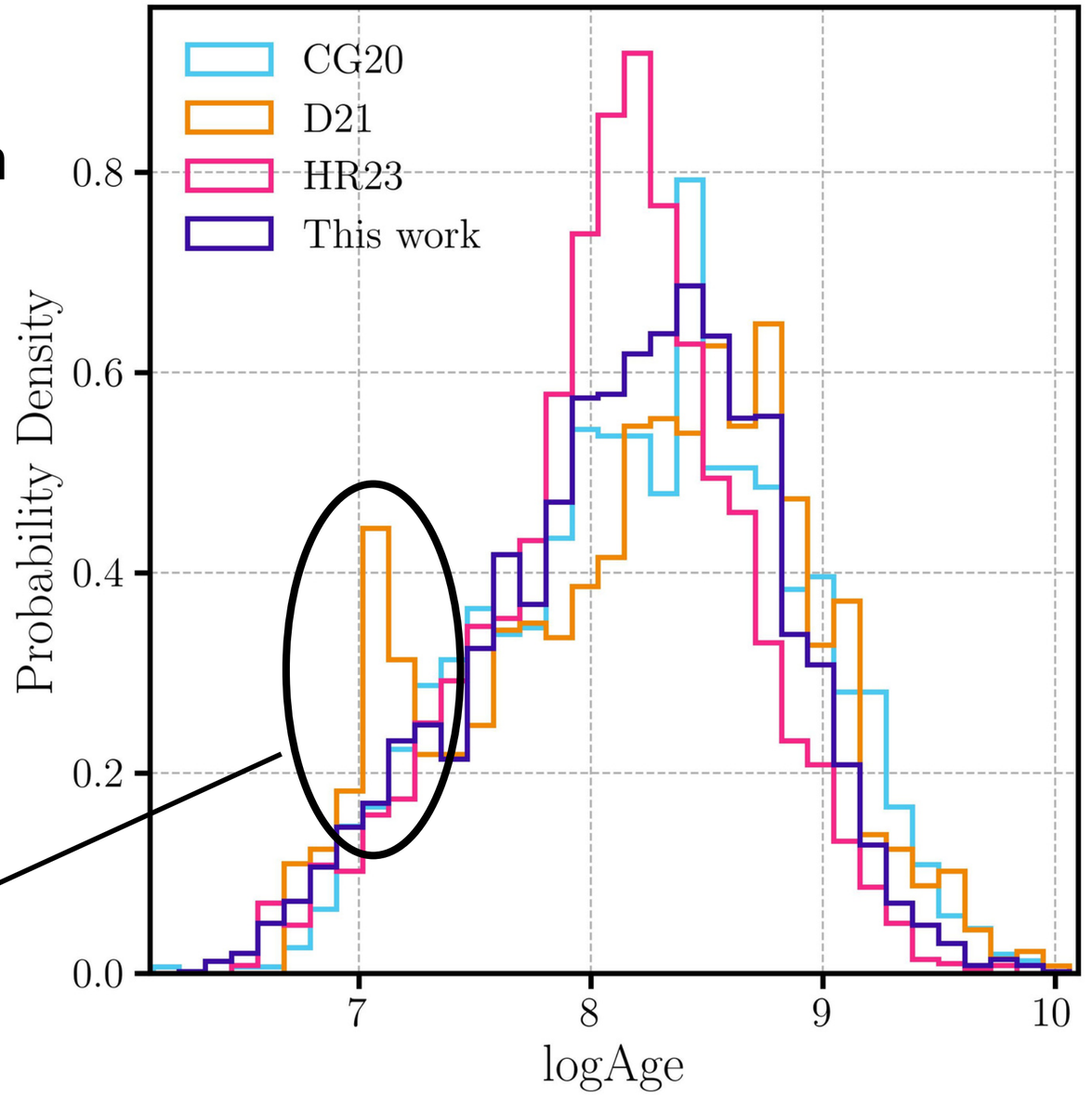
Large scatter at intermediate ages logAge ~ 8


CG20
D21
HR23
This work
Probability Density
0.0
0.2
0.4
0.6
0.8
7
8
9
10
logAge

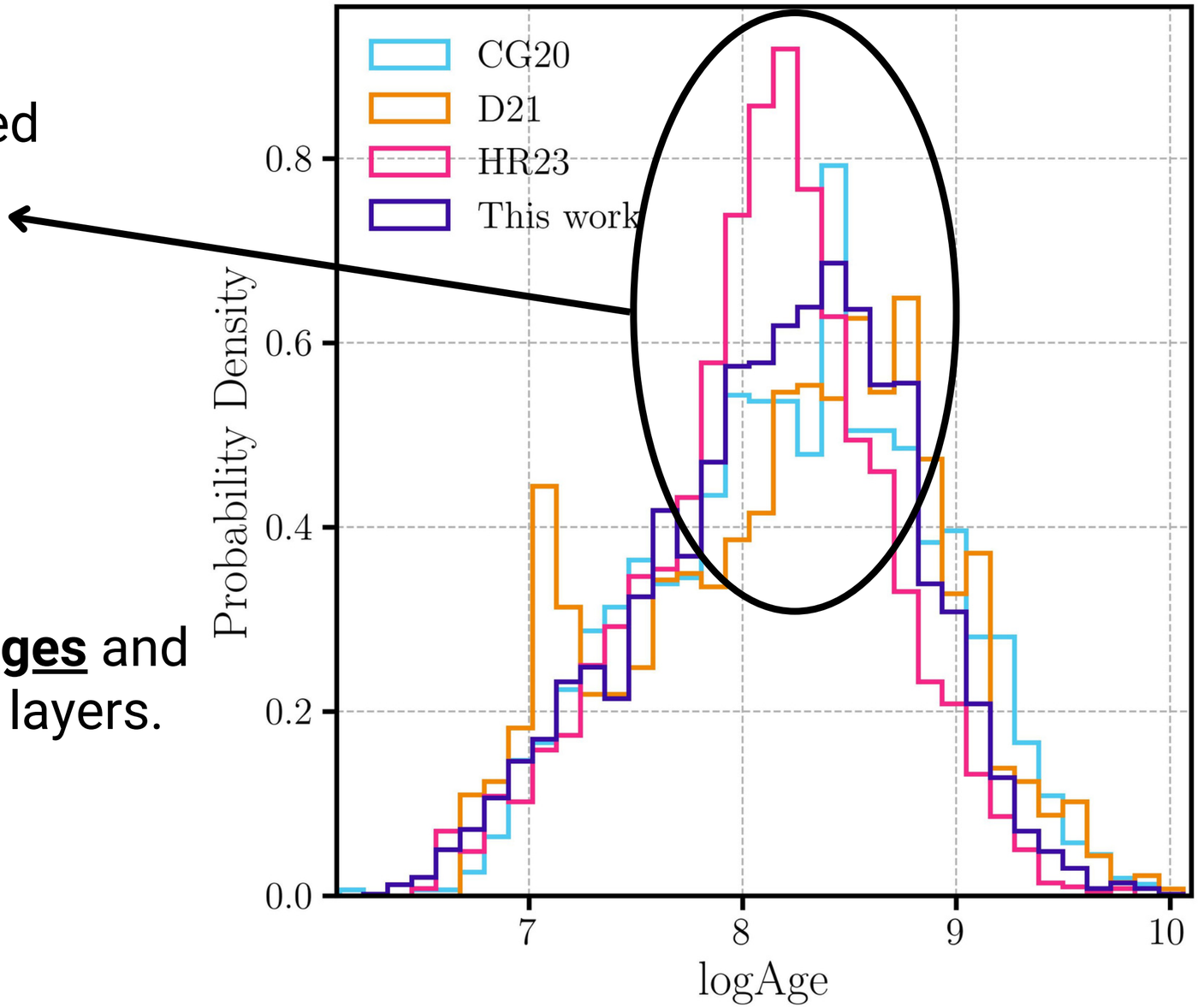


Presence of **red and faint contaminants** in the membership of Dias+21

Pre-sequence



Narrower age distribution compared with the others



FEATURE EXTRACTION

HR23 **pixelized the CMDs into 32x32 images** and then processed them with convolutional layers.
Similar to CG20

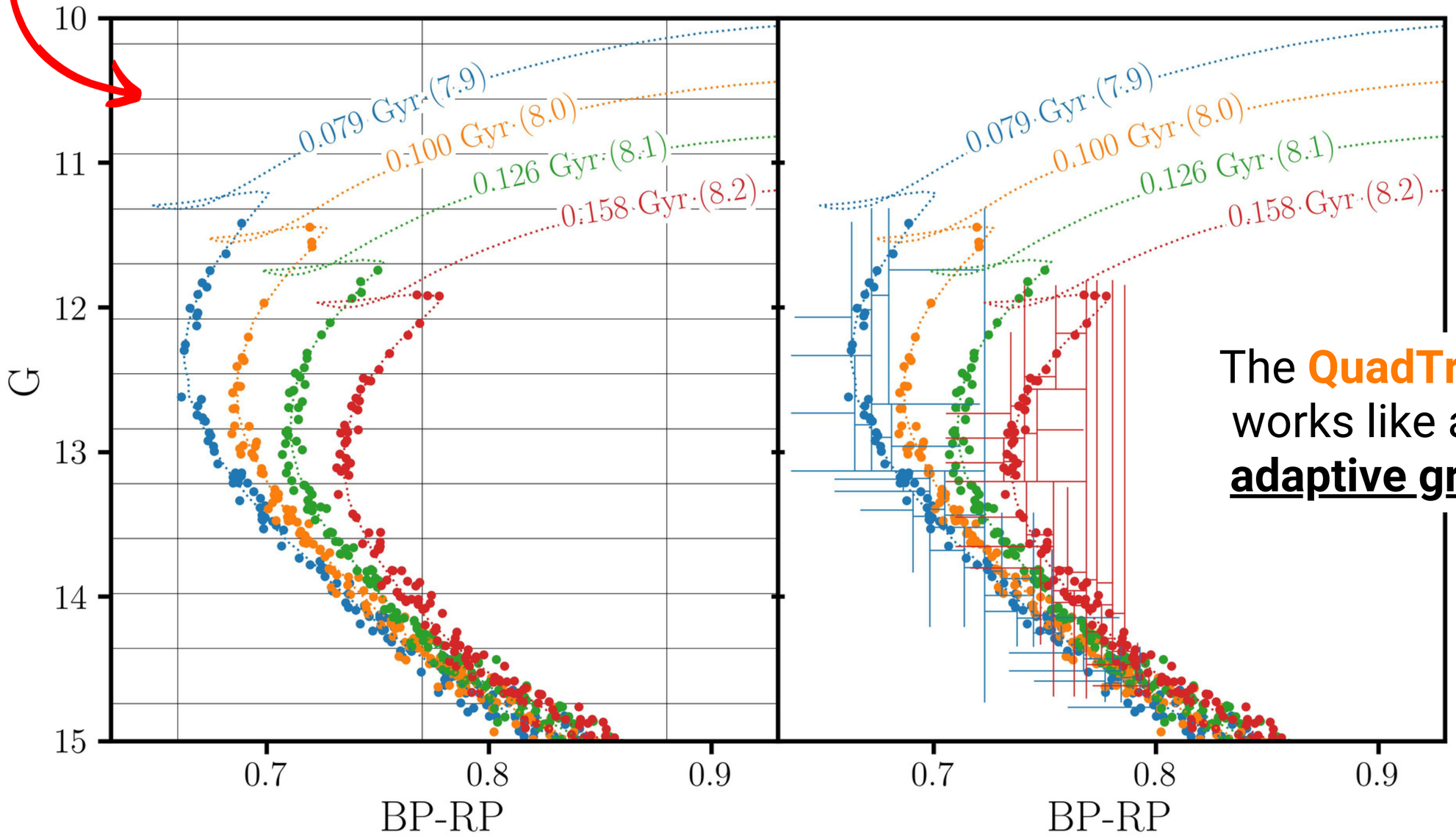
PIXEL SCALE

0.38 mag

0.11 mag

grid used by Hunt & Reffert (2023)

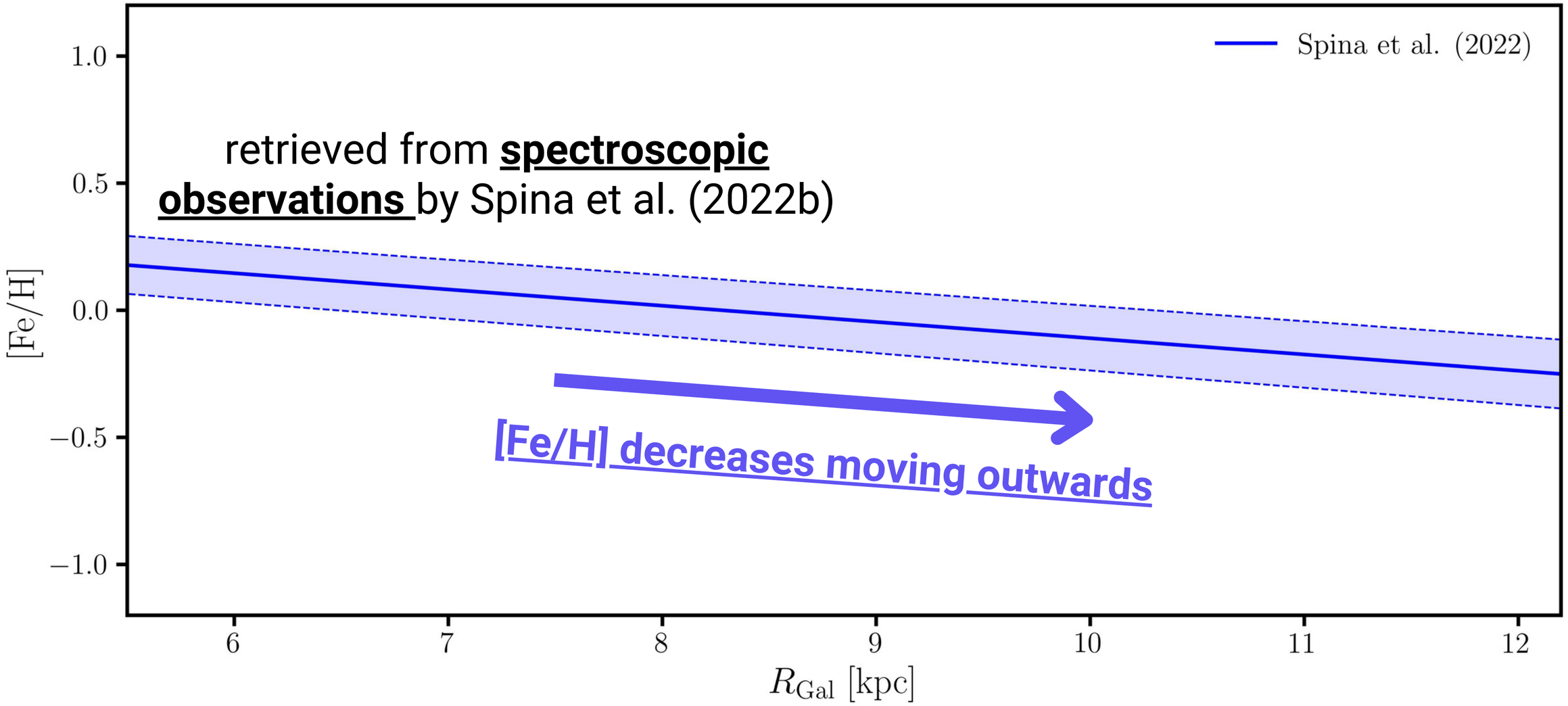
Lose available **information** provided by the **precise Gaia photometry**



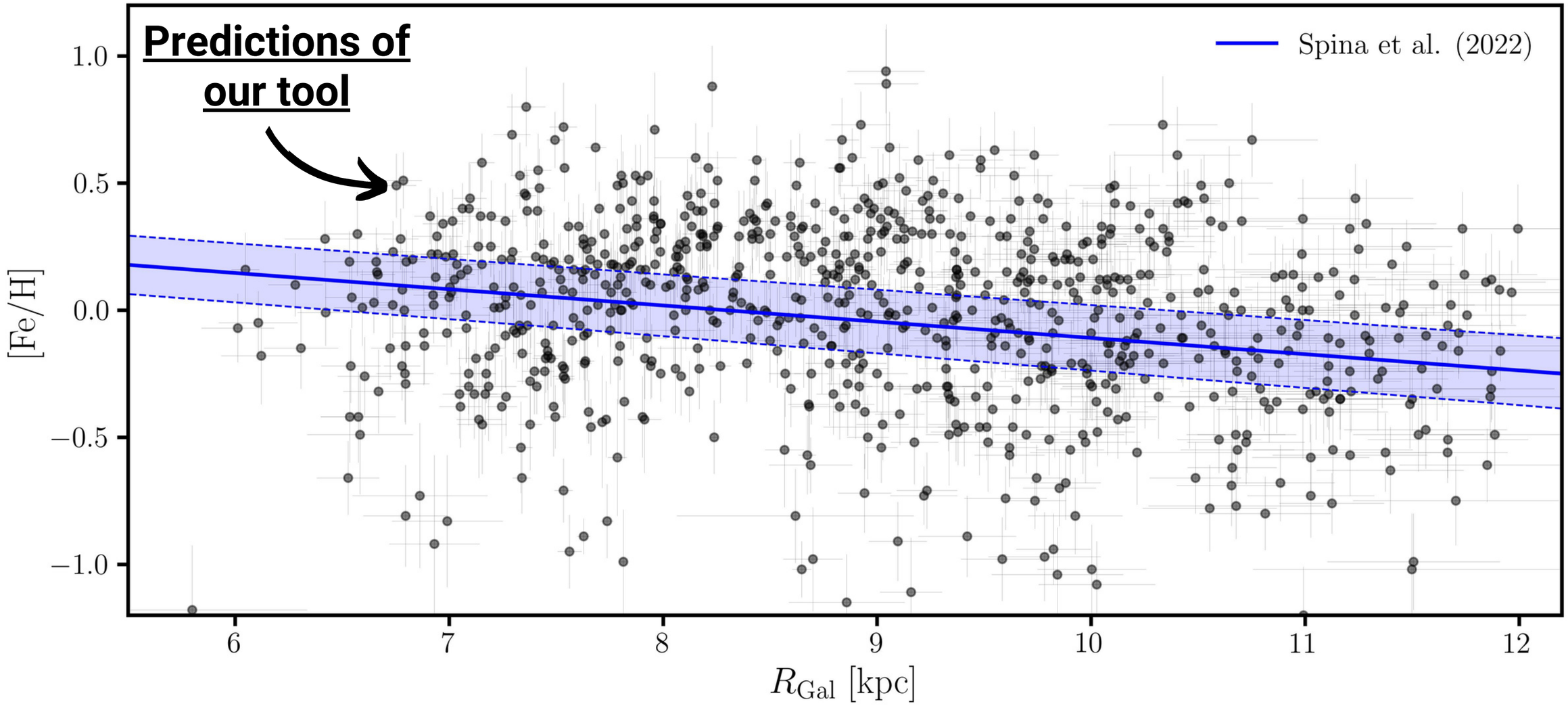
The **QuadTree** works like an **adaptive grid**

WHAT ABOUT THE METALLICITY?

In the inner part of the Galaxy the **metallicity gradient is linear**



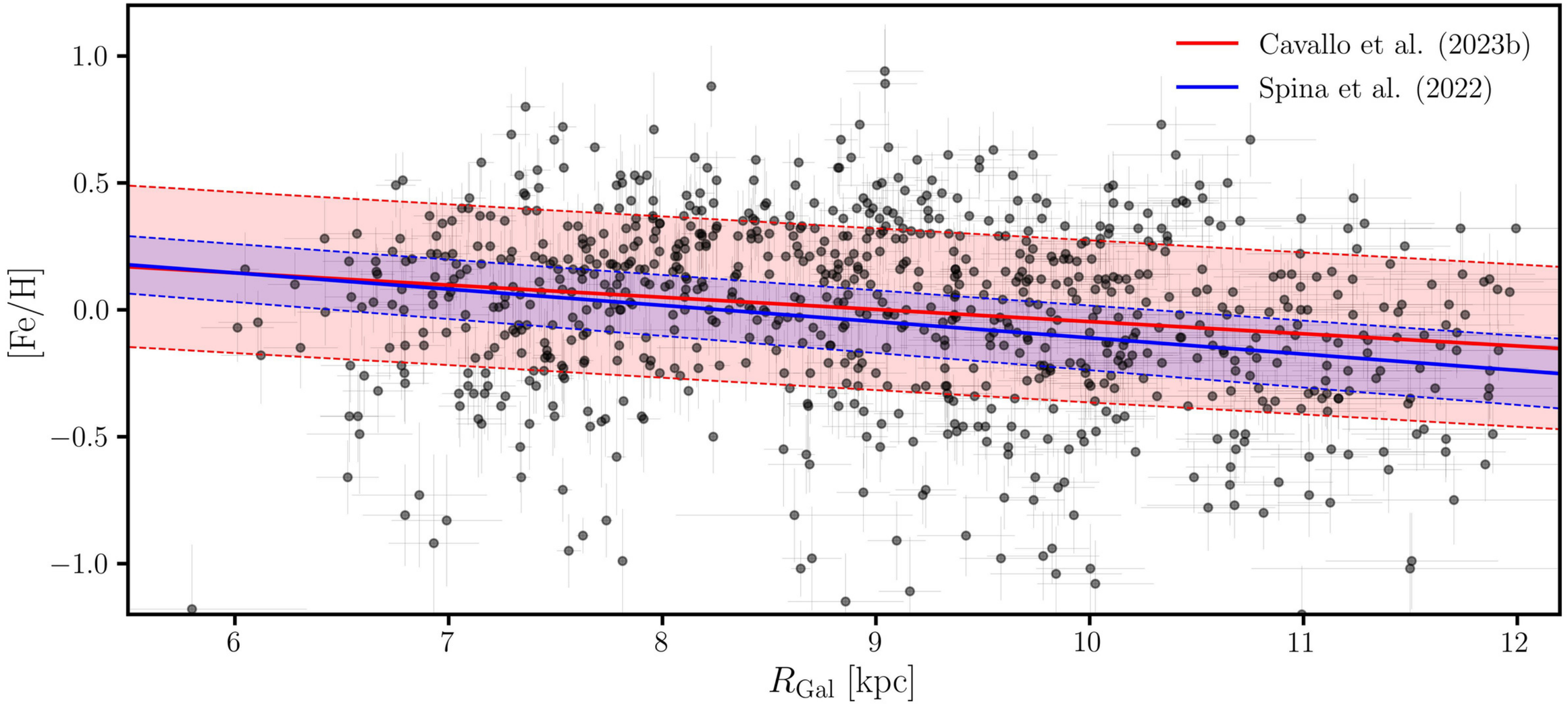
WHAT ABOUT THE METALLICITY?



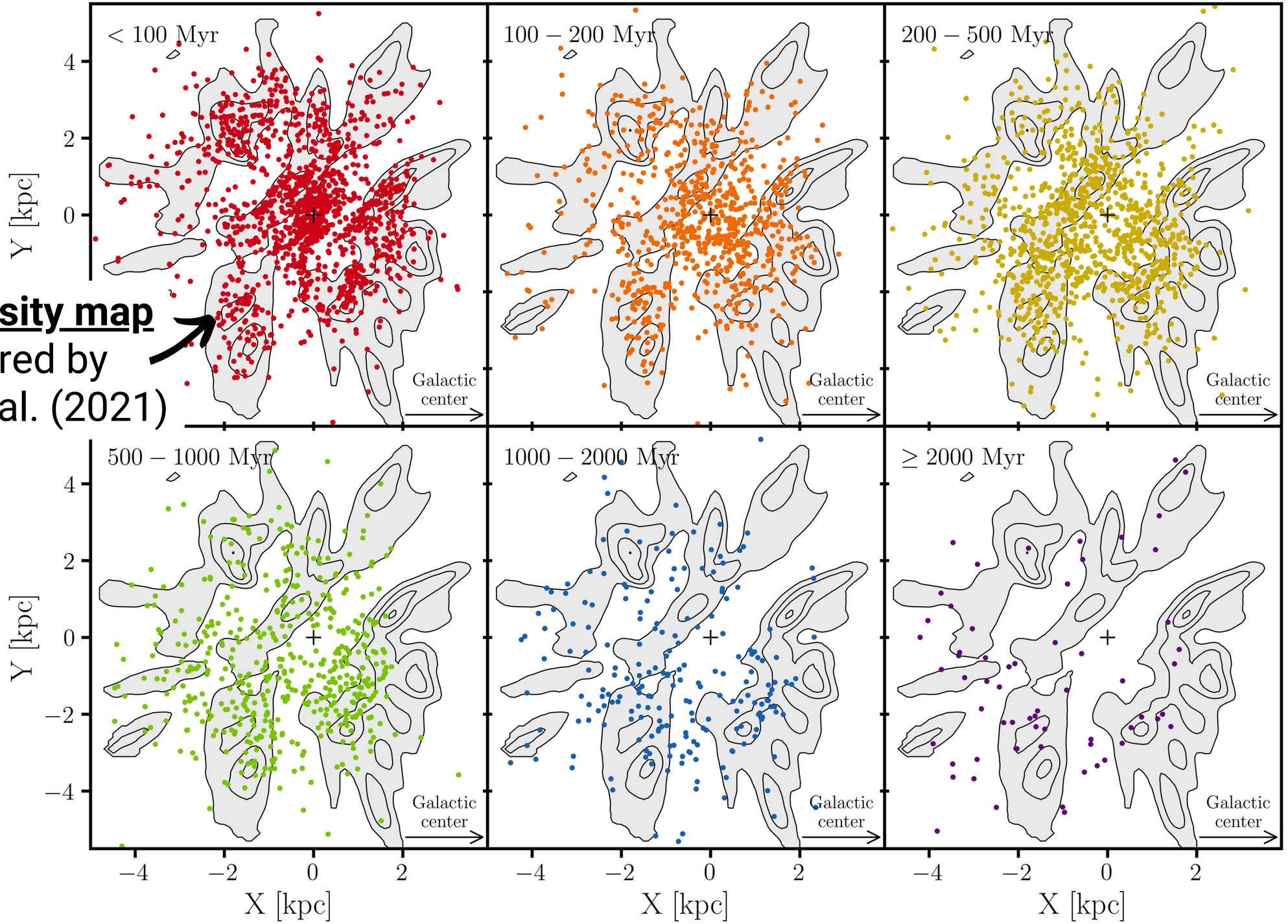
INTRINSIC SCATTER

0.313 ± 0.009

0.091 ± 0.006



Over-density map
measured by
Poggio et al. (2021)



CONCLUSIONS

Catalogue of ~5400 clusters from HR23 with **credible estimates of age, metallicity, extinction, and distance.**

Our tool is the **only one currently available** in the literature that is able to **reliably obtain the metallicity of a cluster** from its photometry and **without spectroscopic data.**

We find systematically older ages compared to the previous works. This is an interesting result, with some possible relevant specific cases.

LESSONS LEARNED 💡

#1

The **use of a grid in a CMD to extract its features** to feed an ANN, a procedure adopted by both CG20 and HR23, **seems to lose some important characteristics** of the sequences. On the other hand, the **QuadTree is capable of efficiently tracing the sequence** even with photometric errors and outliers.

#2

We demonstrate that **clusters' parameters can be derived from the simultaneous analysis of multiple photometric bands.** This prospect will be particularly useful in view of next-generation surveys (e.g. LSST and Gaia DR4)

Parameter Estimation for Open Clusters using an Artificial Neural Network with a QuadTree-based Feature Extractor

L. Cavallo, L. Spina, G. Carraro, L. Magrini, E. Poggio, T. Cantat-Gaudin,
M. Pasquato, S. Lucatello, S. Ortolani, J. Ulloa-Schiappacasse

Accepted in The Astronomical Journal (AJ)

arXiv:2311.03009