



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

Lorenzo Cavallo^{1,2} , Lorenzo Spina^{1,2,3} , Giovanni Carraro¹ , Laura Magrini³ , Eloisa Poggio^{4,5} ,
Tristan Cantat-Gaudin^{6,7} , Mario Pasquato^{1,8,9,10} , Sara Lucatello¹¹ , Sergio Ortolani^{1,2} , and Jose Schiappacasse-Ulloa^{1,2}

¹ Dipartimento di Fisica e Astronomia, Università di Padova, Vicolo dell'Osservatorio 3, I-35122 Padova, Italy; giovanni.carraro@aas.org

² INAF—Padova Observatory, Vicolo dell'Osservatorio 5, I-35122 Padova, Italy

³ INAF, Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy

⁴ Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, F-06304 Nice Cedex 4, France

⁵ INAF—Osservatorio Astrofisico di Torino, Via Osservatorio 20, I-10025 Pino Torinese, TO, Italy

⁶ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, E-08028 Barcelona, Spain

⁷ Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

⁸ Département de Physique, Université de Montréal, Montréal, Québec H3T 1J4, Canada

⁹ Mila—Quebec Artificial Intelligence Institute, Montreal, Quebec, Canada

¹⁰ Ciela, Computation and Astrophysical Data Analysis Institute, Montreal, Quebec, Canada

¹¹ Institute for Advanced Studies, Technische Universität München, Lichtenbergstraße 2 a, D-85748, Garching bei München, Germany

Received 2023 August 3; revised 2023 October 27; accepted 2023 October 28; published 2023 December 7

Abstract

With the unprecedented increase in the number of known star clusters, quick and modern tools are needed for their analysis. In this work, we develop an artificial neural network (ANN) trained on synthetic clusters to estimate the age, metallicity, extinction, and distance of Gaia open clusters. We implement a novel technique to extract features from the color–magnitude diagram of clusters by means of the QuadTree tool, and we adopt a multiband approach. We obtain reliable parameters for ~ 5400 clusters. We demonstrate the effectiveness of our methodology in accurately determining crucial parameters of Gaia open clusters by performing a comprehensive scientific validation. In particular, with our analysis we have been able to reproduce the Galactic metallicity gradient as it is observed by high-resolution spectroscopic surveys. This demonstrates that our method reliably extracts information on metallicity from color–magnitude diagrams (CMDs) of stellar clusters. For the sample of clusters studied, we find an intriguing systematic older age compared to previous analyses present in the literature. This work introduces a novel approach to feature extraction using a QuadTree algorithm, effectively tracing sequences in CMDs despite photometric errors and outliers. The adoption of ANNs, rather than convolutional neural networks, maintains the full positional information and improves performance, while also demonstrating the potential for deriving cluster parameters from simultaneous analysis of multiple photometric bands, beneficial for upcoming telescopes like the Vera Rubin Observatory. The implementation of ANN tools with robust isochrone fit techniques could provide further improvements in the quest for open cluster parameters.

Unified Astronomy Thesaurus concepts: [Open star clusters \(1160\)](#); [Milky Way disk \(1050\)](#); [Fundamental parameters of stars \(555\)](#)

1. Introduction

Open stellar clusters are groups of stars formed together from the same molecular clouds. They are simple stellar populations whose members are born approximately at the same time, and share the same location in the sky, distances, kinematics, and initial chemical composition (e.g., De Silva et al. 2006, 2007; Bland-Hawthorn et al. 2010; Armillotta et al. 2018). Given all these properties, open clusters have long been considered benchmarks in the context of the determination of stellar properties such as ages, distances, and chemical composition. In fact, these quantities can be derived for cluster members with extremely high precision compared to what is normally achievable for individual field stars.

For that reason, open clusters are the ideal laboratories for a large variety of studies concerning stellar evolution and the Galaxy as a whole (Adams & Myers 2001; Portegies Zwart et al. 2001; Krumholz & McKee 2019; Wright 2020; Cantat-

Gaudin 2022). Numerous stellar evolution processes can be investigated with open cluster members: star formation (Lada & Lada 2003), progression in the Hertzsprung–Russell diagram (Vandenberg 1983; Gaia Collaboration et al. 2018), evolution of rotational properties (Meibom et al. 2011; Gruner & Barnes 2020; Godoy-Rivera et al. 2021), mixing episodes (Dumont et al. 2021; Magrini et al. 2021a), and diffusion processes (Bertelli Motta et al. 2018; Souto et al. 2018). Clusters have also long been used as tracers of the structure and evolution of the Milky Way (Moffat & Vogt 1973; Janes & Adler 1982; Dias & Lépine 2005; Moitinho et al. 2006; Carraro & Costa 2009; Carraro et al. 2010; Xu et al. 2018; Reid et al. 2019; Cantat-Gaudin et al. 2020; Poggio et al. 2021) and the distribution of element across the Galactic disk (Janes 1979; Frie et al. 2002; Donor et al. 2018, 2020) and have offered perspectives on how this distribution has evolved with time (Carraro et al. 1998; Myers et al. 2022; Spina et al. 2022; Magrini et al. 2023). Thanks to their ages and precisely determined chemical patterns, open clusters are excellent calibrators for chemical clocks: powerful, alternative, chemically based diagnostics of ages for field stars, such as the carbon-to-nitrogen abundance ratio (Casali et al. 2019; Spoo et al. 2022).

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

The entire life cycle of open clusters is deeply connected to several aspects of Galaxy evolution. For instance, their birth and death can be strongly influenced by large-scale nonaxisymmetric features in the Milky Way, such as spiral arms, as well as giant molecular clouds on a small scale (Piskunov et al. 2006; Morales et al. 2013; Piskunov et al. 2018; Wright 2020; Anders et al. 2021). Therefore, it has been proven to be important to understand the dynamical evolution and dissipation of open clusters (Bravi et al. 2018; Carrera et al. 2019; Tang et al. 2019; Yeh et al. 2019; Meingast et al. 2021; Pang et al. 2021; Casamiquela et al. 2022) and also compare their demography to that of field stars (Soubiran et al. 2018; Spina et al. 2020).

Open clusters are also critical in testing the possibility of using the chemical composition of stars to identify the environment where they formed (Freeman & Bland-Hawthorn 2002; Mitschang et al. 2013; Blanco-Cuaresma et al. 2015; Garcia-Dias et al. 2019; Spina et al. 2022c). Beyond their value in understanding the Galaxy and stellar evolution, open clusters are frequently used as reference objects to assess the quality of the large astronomical data sets (Babusiaux et al. 2023; Gaia Collaboration et al. 2023; Randich et al. 2022) and calibrators for stellar yields models (Maiorca et al. 2012; Magrini et al. 2021b).

All this broad variety of studies is grounded on reliable determinations of open cluster properties such as distance, age, and extinction. The advent of Gaia astrometric and photometric data (Gaia Collaboration et al. 2016) resulted in substantial advancement in our knowledge of the open cluster population. Current astrometric data allow a much more complete census of stellar associations and their stellar members than pre-Gaia catalogs (e.g., Dias et al. 2002; Kharchenko et al. 2013). All that, in conjunction with the precise Gaia photometry, allows astronomers to appreciate tight and clear sequences in color-magnitude diagrams, outlined by the members of such stellar populations, fundamental pieces of information from which we can extract key physical parameters of open clusters. As a consequence of this progress, automated tools for the characterization of open clusters based on advanced statistical methods have become increasingly common (e.g., Cantat-Gaudin et al. 2020; Dias et al. 2021; Hunt & Reffert 2023).

In 2022 June, Gaia Collaboration published the third data release (DR3; Gaia Collaboration et al. 2023b) that contains the photometry (G , G_{BP} , and G_{RP}) of more than 1.5 billion objects. From these data, clustering algorithms are discovering more and more star clusters (e.g., Hunt & Reffert 2023; Perren et al. 2023) for which a proper parameter estimation is required. To do that, one method often used is the so-called isochrone fitting, which provides good results, but is usually performed case by case and so is not efficient when used in samples that contain hundreds or even thousands of clusters.

Recently, Dias et al. (2021; hereafter D21) have determined the parameters, via isochrone fitting procedure, of 1743 open clusters contained in Gaia DR2. This kind of procedure also incurs significant computational overhead. To address this challenge and overcome the high computational cost, D21 made use of strong priors in cluster extinction and metallicity, which can guide the procedure toward more efficient and focused computations, resulting in improved performance. However, the introduction of such priors can introduce biases into the analysis. Furthermore, the entire procedure is

susceptible to the presence of faint and red nonmember stars, which can significantly impact its performance and reliability.

In recent years, due to the ever-growing number of open clusters and the considerable computational demands of the isochrone fitting procedure, several researchers have explored the utilization of machine learning algorithms to estimate the parameters of open clusters. Cantat-Gaudin et al. (2020; hereafter CG20) have used an artificial neural network (ANN; Rumelhart et al. 1986) to estimate the age, extinction, and distance of 1867 Gaia open clusters. Their neural network has been trained using well-known clusters that possess precisely defined parameters. Considering the utilization of an observational sample of open clusters which are typically younger than a few 100 Myr, there may exist potential biases in their analysis.

Hunt & Reffert (2023; hereafter HR23) employed the Gaia DR3 data set to conduct a comprehensive search for open clusters across the entire sky. They utilized the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm (Campello et al. 2013) to identify potential cluster candidates. These candidates were then subjected to validation using a quasi-Bayesian convolutional neural network (CNN) designed to distinguish genuine clusters from spurious ones. In fact, variable extinction—which is a frequent condition near the Galactic midplane—and random density fluctuations can make unbound groups of stars appear as relative overdensities in sky coordinates, parallaxes, and proper motions (e.g., Kounke et al. 2020). Furthermore, a similar neural network was trained to estimate the ages, extinctions, and distances of the identified clusters. By employing these techniques, HR23 generated a unified catalog of star clusters consisting of 7167 entries. Their training set was assembled with synthetic clusters, which suggests a potential absence of biases compared to the case of CG20.

In this work, we employ a group of ANNs trained on synthetic OCs to estimate the age, metallicity, extinction, and distance modulus of a sample of clusters obtained from HR23. Our method of analysis has several differences from the techniques used in the past. The most significant element of novelty is that we use multiple color-magnitude diagrams (CMDs) constructed using Gaia and Two Micron All-Sky Survey (2MASS) photometry. By incorporating additional information, our approach provides a more comprehensive analysis of the cluster properties. Moreover, developing and testing an algorithm that conjointly analyzes multiple photometric bands is strategic in view of next-generation telescopes, such as the Vera Rubin Observatory which will provide photometry from six filters for $\sim 20 \times 10^9$ stars (LSST Science Collaboration et al. 2017).

Instead of directly feeding the neural networks with the CMD images, with our approach, we first extract the relevant features of the distributions of stars within these diagrams with the QuadTree algorithm (Schiappacasse-Ulloa et al. 2023) and then we feed the neural net with this information only. By doing so we significantly reduce the complexity of the network that will be trained compared to typical networks that need to process full images. Besides all that, our approach is different in many other ways from past techniques of analysis. It is by exploring and developing new techniques that we enhance the accuracy and comprehensiveness of parameter estimation for the population of Galactic open stellar clusters.

This paper is structured as follows. In Section 2 we describe the data set employed in our study. In Section 3 we give details on the model we use for the parameter estimation and how it has been trained. The results and comparison with previous studies are discussed in Section 4. In Section 5 we present a scientific validation of our methodology to derive cluster parameters. Finally, conclusions are discussed in Section 6.

2. Data

Recently, HR23 presented a homogeneous sample of open clusters (see Section 2.1), providing estimates for age, extinction, and distance. Building upon their work, we extend their analysis by employing a distinct method to estimate age, extinction, distance, and metallicity for the same sample of clusters. To accomplish this, we utilize a group of ANNs trained with a set of synthetic open clusters generated using a numerical Monte Carlo code developed specifically for this purpose (see Section 2.2).

In this section, we describe the catalog of open clusters for which we estimate the parameters and the training set we used for this purpose.

2.1. Catalog of Open Clusters

HR23 conducted an all-sky blind search for open clusters in the Gaia DR3 data set using the HDBSCAN algorithm. Candidates identified by HDBSCAN have been validated with a quasi-Bayesian CNN trained to discriminate between real and fake clusters from pixelated images of their CMDs. A similar neural network is then used to infer the age, extinction, and distance of detected clusters. They then produced a homogeneous star cluster catalog composed of 7167 clusters (~ 2400 of them were not present in the previous catalogs). From the initial catalog, we select only star members with $G \leq 18$. It is important to note that the full catalog contains open clusters, globular clusters, and moving groups and that, as it results from HR23's analysis, some of these 7167 stellar associations may not be real.

From HR23 we select the stellar associations with ≥ 10 members ending up with 6413 clusters. In addition to the Gaia photometry, in our study, we also use the 2MASS J and H magnitudes for each member of these clusters, when available. More specifically, we use only the 2MASS magnitudes that have signal-to-noise ratio greater than 7 (2MASS photometric quality flag equal to A or B) and photometric errors ≤ 0.3 mag.

2.2. Synthetic Color–Magnitude Diagrams

An isochrone is the locus of points in the CMD of stars sharing the same age and chemical composition. The isochrone alone is not informative on the number of stars that are in a certain evolutionary stage (i.e., that populate different zones of the CMD). This information is given by the initial mass function (IMF) and the initial mass of the cluster. Moreover, in order to develop a synthetic CMD of an open cluster the effects on the CMD of extinction, distance, and photometric errors have to be considered.

In this work, we develop a numerical Monte Carlo code that simulates a CMD of an open cluster (i.e., simple stellar population) of age t , metallicity $[\text{Fe}/\text{H}]$, affected by an extinction A_V , situated at a distance d , and composed by a number of members N_{star} . For each star in our synthetic

Table 1

Distribution of Parameters of Our Synthetic Clusters Used to Train Our ANN

Parameter	Distribution
$\log \text{Age}$	$\mathcal{U}(6.7, 10.1)$
$[\text{Fe}/\text{H}]$	$\mathcal{U}(-0.8, 0.6)$
A_V	$\mathcal{U}(0, 5)$
dMod	$\mathcal{U}(0, 15)$
$\log N_{\text{stars}}$	$\mathcal{U}(1.0, 3.5)$

clusters, we compute the apparent magnitude in the following bands: G , G_{BP} , G_{RP} , J , and H .

In the first step, the code searches among a grid of theoretical isochrones the one of a stellar population of age t and metallicity $[\text{Fe}/\text{H}]$. For the theoretical grid of models, we use version 1.2S of PARSEC isochrones (Bressan et al. 2012; Chen et al. 2014; Tang et al. 2014; Chen et al. 2015). The selected isochrone gives the relation between the absolute magnitude M_λ (where λ indicates a generic band) and the mass of the star. At this point, we transform the absolute magnitudes of the isochrone into apparent ones (m_λ) as:

$$m_\lambda = M_\lambda + \text{dMod} + A_V \times c_\lambda \quad (1)$$

where M_λ is the absolute magnitude in the generic band λ , dMod is the distance modulus ($\text{dMod} = 5 \log(d) - 5$), A_V is the extinction in the V band, and c_λ is the ratio of the extinction coefficients in the λ and V band. The coefficients c_λ depend both on the A_V of the cluster and the effective temperature (T_{eff}) of stars and thus it needs to be computed star-to-star. To do that, we used the Python package `pysvo`¹² (Rodrigo et al. 2012), which allows you to handle data coming from the Theoretical Model Services maintained by the Spanish Virtual Observatory. For this computation, we assume an extinction curve with $R_V = 3.1$ (Cardelli et al. 1989).

Then, we generate N_{stars} synthetic stars with a mass randomly chosen accordingly with the Salpeter (1955) IMF. With the isochrone computed in the first step, we obtain the apparent magnitude of each star.

To train our neural network we use a sample that consists of 50,000 synthetic open clusters generated with the CMD generator described above. The synthetic population of OCs has been generated with parameters distributed as reported in Table 1. For each synthetic OC, we also derive its expected parallax as the inverse of the distance.

During the development of our tool, we considered well-known astronomical phenomena that have the potential to sculpt the CMD and, consequently, influence our ANN results. Phenomena such as stellar rotation, binaries (see Cordoni et al. 2023; Donada et al. 2023), extended main-sequence turnoff, and differential extinction, among others, undeniably play important roles in shaping the CMD and can introduce complexities and potential sources of error when not accounted for.

Our decision to employ a simplified, yet plausible, model to test the effectiveness of the QuadTree algorithm in extracting CMD features, and the overall capability of the ANN to efficiently extract parameters like $\log \text{Age}$, A_V , dMod, and notably, the metallicity, was driven by a strategic choice to establish a robust baseline.

¹² <https://github.com/fjaellet/pysvo>

It is important to acknowledge that the aforementioned astronomical processes have a tangible impact on the CMD and omitting them could indeed be a source of error. However, it is crucial to highlight a pivotal aspect of our ANN: it does not “see” the position of individual cluster members. Instead, it processes a series of coordinates that delineate zones of point overdensities, thereby lacking the capability to discriminate between a CMD that presents a sequence of binary stars and one where the main sequence is more sparse than predicted by theoretical models.

2.3. *t*-Distributed Stochastic Neighbor Embedding

Before we use our synthetic sample of OCs to train the ANN, we need to check whether or not the training set is a good generalization of all the possible observations. As we anticipated in Section 1, the ANN that we use to determine the OCs’ parameters receives features extracted from cluster CMDs. More specifically, it is fed with 63 input features that are extracted from the set of three CMDs built for each of the clusters on the basis of Gaia and 2MASS photometry (see Section 3.2 for details), plus an additional input feature which is the parallax. The procedure of feature extraction from the CMDs is described in Section 3.1. Here, we aim to demonstrate that distributions of all the features we extract from the observed CMDs overlap with those of the same features extracted from the synthetic CMDs.

In the fields of data analysis and machine learning, understanding and visualizing complex high-dimensional data sets is a critical task. Traditional visualization techniques often struggle to effectively capture the intricate relationships and structures hidden within such data. However, *t*-Distributed Stochastic Neighbor Embedding (*t*-SNE; van der Maaten & Hinton 2008) maps offer a powerful approach to address this challenge.

t-SNE maps provide a means of visualizing high-dimensional data in a reduced dimensional space while attempting to preserve the inherent relationships and structures among the data points, especially at small scales. This is accomplished by minimizing a measure of dissimilarity between the distance distribution in the high-dimensional space and the low-dimensional projection; an accessible explanation of the algorithm is provided by Wattenberg et al. (2016). The resulting *t*-SNE map provides a visually intuitive representation of the data, where clusters, patterns, and similarities are discernible. Points that are similar in the original space tend to be grouped together in the *t*-SNE map, enabling analysts to identify distinct clusters or regions of interest.

In Figure 1 we plot the superimposed *t*-SNE maps obtained by the Gaia CMDs of synthetic (color coded according to their distance modulus) and observed (black dots) clusters. In that plot, we notice that the clusters of HR23 are perfectly embedded in the synthetic sample generated by us and that we use to train the neural network. It is also clear that the observed clusters cover only a portion of the area filled by the synthetic clusters. As observing an individual *t*-SNE embedding does not per se guarantee that these relations are also reflected in our original feature space, we have examined various plots similar to Figure 1 obtained for different values of the perplexity hyperparameter, which sets the effective scale at which structure is rendered faithfully. The results appear robust to such changes.

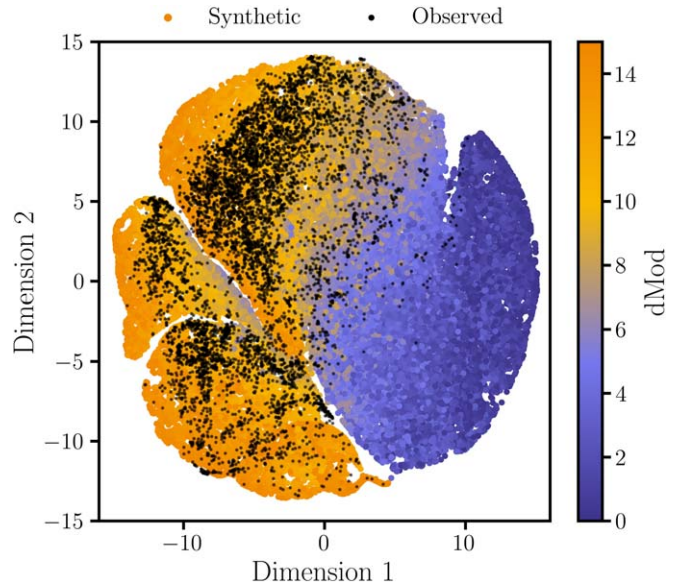


Figure 1. Superimposed *t*-SNE maps of synthetic and observed OCs. Synthetic clusters are color coded based on their distance modulus, while observed clusters are plotted with black dots.

3. Artificial Neural Network

An ANN is an algorithm that maps input data into output parameters through a network of nodes and weights. In our case, we feed the ANN with parallaxes of OCs and information extracted from their CMDs (input) and obtain their age, metallicity, extinction, and distance modulus (output). Between the input and output layers, hidden layers of neurons are placed. These layers increase the complexity of the network and its predictive potential. However, as the network becomes more complex, it develops a susceptibility to overfitting. In fully connected ANNs (as the one developed in this work), each neuron in a layer receives from (all) the nodes of the previous layer an input, that is transformed with a nonlinear function and transmitted to the neurons of the next layer.

In this work, we train an ANN with synthetic OCs and then use it to estimate the parameters of the sample of clusters from HR23. Below we describe the method used to build the sample of synthetic OCs used to train our ANN (Section 2.2), its input layer (Section 3.1), architecture (Section 3.2), and training and performance (Section 3.3).

3.1. Features Extraction

To build the input layer of our ANN we need to pre-process the CMD of each cluster and transform it into a one-dimensional array. To do that we compute the QuadTree of CMDs.

A QuadTree is a tree data structure that is used to represent object distribution in two-dimensional space. It takes a distribution of points on a xy -plane, in the first level of the tree the algorithm divides the xy -plane into four different areas that contain the same number of points. It thus saves the coordinates of the edges of these areas and so it obtains one x -coordinate (x_1) and two y -coordinates (y_1, y_2). In the second level, it repeats the procedure in each of the four areas of the first level. It then obtains again one x -coordinate and two y -coordinates per area. At the end of the second level, it obtains two arrays of coordinates: ($x_1^{1st}, x_2^{2nd}, \dots, x_5^{2nd}$) and ($y_1^{1st}, y_2^{1st}, y_3^{2nd}, \dots, y_{10}^{2nd}$); where superscripts 1st and 2nd label

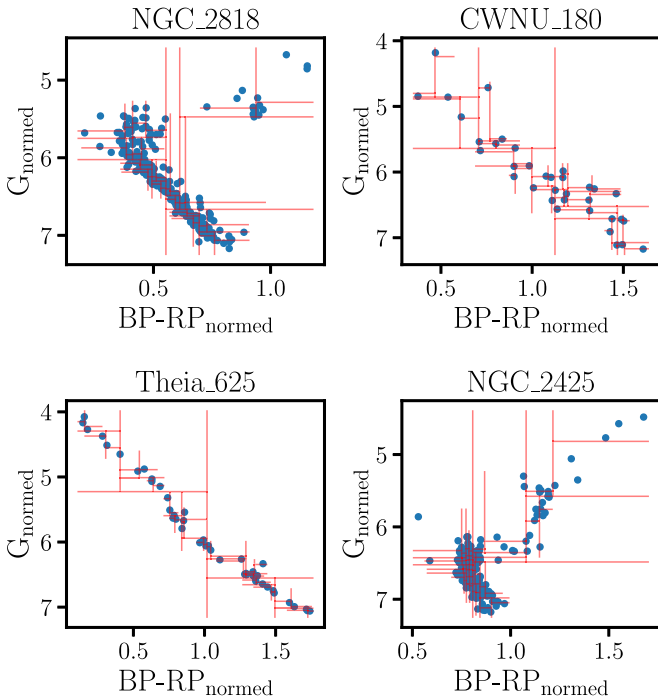


Figure 2. QuadTrees of Gaia CMD of four clusters contained in HR23. Note that magnitudes and colors are standardized as described in Section 3.1.

coordinates that come from the first and second levels, respectively. It repeats the same procedure also for the third and last level, and it obtains two arrays: $(x_1^{1st}, \dots, x_{21}^{3rd})$ and $(y_1^{1st}, \dots, y_{42}^{3rd})$. Then we concatenate them into a unique array with 63 entries.

In a typical CMD, the absolute star-to-star difference in magnitude is larger than the one in color. In this situation, the ANN may weigh more the differences in magnitude and less the one in color. To avoid this situation we standardize the CMD diagrams by dividing each quantity by its standard deviation computed on the whole sample of synthetic CMDs. We then compute the QuadTrees of these standardized CMDs.

In Figure 2 we plot the QuadTrees of four selected clusters. It is possible to notice that the algorithm can trace the sequence of OCs with a fair number of stars, as well as in the case of OCs with few tens of members.

3.2. The Neural Network Architecture

In this work, we tested many ANNs with various architectures, input layers, and characteristics. In the following lines, we are going to describe the combination of characteristics that leads to the best performance.

Our ANN is made of fully connected layers and is implemented in TensorFlow (Abadi et al. 2016). The input layer receives the 63×3 features produced by the QuadTree (see Section 3.1) extracted from the three different CMDs ($(G_{BP} - G_{RP})$ versus G , $((G_{BP} - G_{RP}))$ versus J , and $(J - H)$ versus J) of the cluster that we want to study plus the mean parallax of its members. In particular, the input layer has 190 neurons divided in 3×63 entries from the flattened QuadTrees of the three CMDs and one with the mean parallax of the cluster. As described in Section 3.1, the QuadTrees are computed on standardized CMDs. In this layer, we add a Dropout function that randomly switches off the 10% of neurons of the layer. We verified that this helps the ANN deal

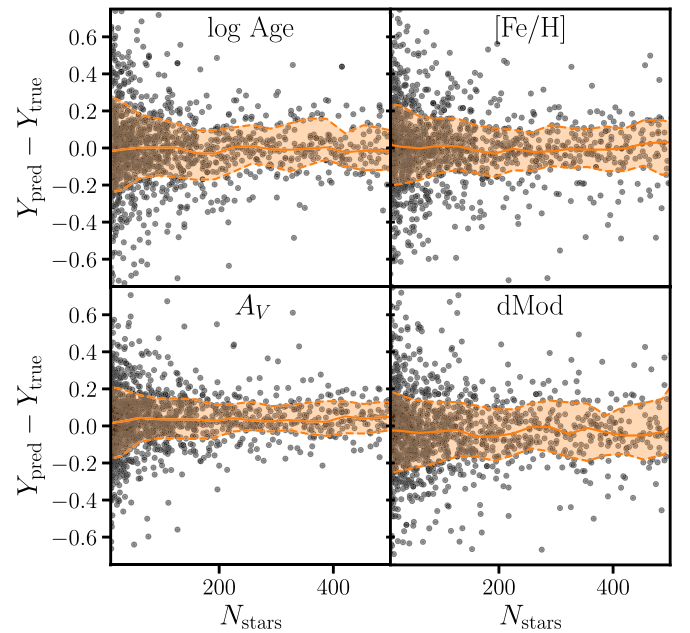


Figure 3. Difference between the parameters estimated by our ANN and the true ones as a function N_{stars} for ~ 2000 synthetic clusters contained in the test sample (black dots). In orange, we plot the mean and standard deviations of $Y_{pred} - Y_{true}$.

with contaminants present in CMDs of real OCs that produce out-layers in the QuadTree.

After the input layer, we place five hidden layers, each one composed of 500 neurons. For the activation function of these hidden layers, we use the Leaky version of the Rectified Linear Unit (LeakyReLU) with $\alpha = 0.3$. As output, the ANN gives the logarithm of the age ($\log t$), metallicity ($[Fe/H]$), visual extinction (A_V), and distance modulus (dMod) of a cluster given its CMDs and the mean parallax.

3.3. Training and Performance

In general, an ANN is deterministic, i.e., at a certain input will produce a certain output without a metric that represents how well determined that output is. In order to estimate that metric we trained, 300 ANNs on different subsamples of 10,000 clusters randomly selected among the total synthetic sample (50,000).

To train each ANN we divide the 10,000 synthetic clusters into train ($\sim 80\%$) and test ($\sim 20\%$) samples. The train is performed with the Adam solver (Kingma & Ba 2014) that optimizes the mean squared error (MSE) loss function.

In order to evaluate the reliability of the ANN to obtain the parameter of a cluster given its CMDs and parallax we analyze its performance on ~ 2000 synthetic clusters that are part of the test sample. In Figure 3 we plot the difference between the parameters estimated by our ANN and the real ones. As expected we notice that the parameters of clusters with fewer members are more uncertain. On average, our ANN estimates $\log Age$, $[Fe/H]$, A_V , and distance modulus with an uncertainty of $\lesssim 0.2$. These uncertainties are a combination of the errors made by the ANN itself and the ambiguities generated by photometric errors (embedded in the CMDs of our synthetic clusters, see Section 2.2). By aggregating the predictions from these 300 ANNs we enhance the reliability of the estimates and gain a more comprehensive understanding of the uncertainties introduced by our ANNs.

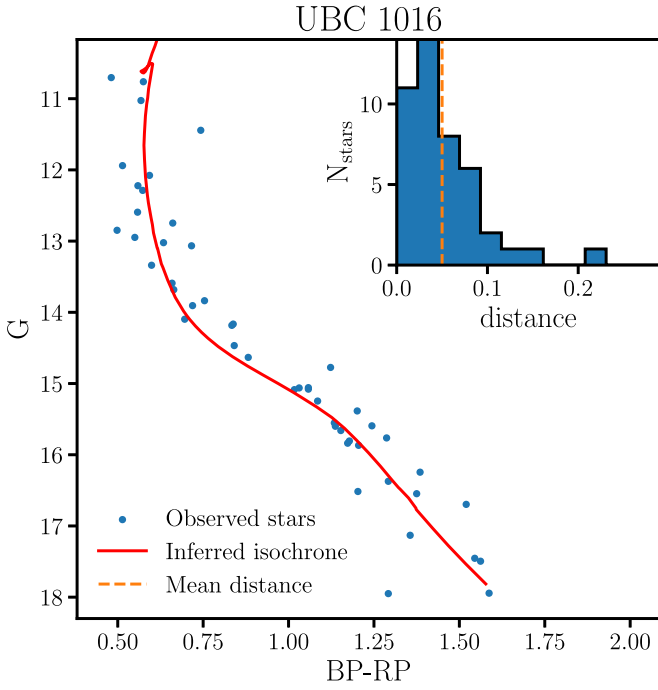


Figure 4. Example of the computation of the quality metric developed in this work. With blue dots, the members of the open cluster UBC 1016 on which we superimpose the isochrone inferred by the ANN (plotted with a solid red line). In the top right corner, we plot the histogram of the distances between the inferred isochrone and the members of the cluster. On the histogram, we plot with a vertical dotted orange line the mean value of the distribution of distances.

3.4. Parameter Estimation

Once the 300 ANNs are trained, we proceed with the parameter estimation of the real clusters. To do so we proceed as follows: (i) the three CMDs are standardized as described in Section 3.1; (ii) we compute the QuadTree of each CMD, merge them, and add the mean parallax of the cluster members obtaining the input layer; (iii) the input layer is fed to the 300 ANNs and each of them gives an estimate for $\log\text{Age}$, $[\text{Fe}/\text{H}]$, A_V , and $d\text{Mod}$. Thus, for each parameter, we obtain a distribution from which we extract the mean (our estimated value) and 1σ confidence interval, i.e., its uncertainty.

Given a large number of objects, it is impossible to perform a visual inspection of the CMDs of each cluster with theoretical isochrones computed with the estimated parameters to check if there is a good match. We thus need a metric that indicates the quality of the match between the estimated isochrone and the observed CMD. To do that, we developed a metric based on the distance of stars in the Gaia CMD from the isochrone predicted by our neural networks. Given the Gaia CMD of a cluster, for each cluster member, we normalize the colors and magnitudes (see Section 3.1), we compute the distance of the star from the closest point of the inferred isochrone, and we average them obtaining the mean distance of the CMD from the isochrone (μ_{dist}). In Figure 4 we plot an example of the computation just described, for the cluster UBC 1016.

We divide the sample of clusters analyzed in this work into four different categories based on the value of the metric (μ_{dist}) described above: gold, silver, bronze, and wood. Given the distribution of μ_{dist} , plotted in the right panel of Figure 5, we compute its standard deviation $\sigma_{\mu_{\text{dist}}}$ (hereafter $\sigma_{\mu_{\text{dist}}} \equiv \sigma_{\mu}$) and subsequently, we label our parameter estimate of a cluster as:

(i) gold, if μ_{dist} is within σ_{μ} from the mode of the distribution; (ii) silver, μ_{dist} between 1σ and 2σ ; (iii) bronze, μ_{dist} between 2σ and 3σ ; (iv) wood, μ_{dist} beyond 3σ . As plotted in the right panel of Figure 5 ~ 5300 ($>80\%$) clusters belong to the gold sample and so their inferred isochrones have a good match with the observational data.

About 200 clusters ($\sim 3\%$) belong to the wood sample, where ANN struggles to extract parameters that are in agreement with the observed CMDs. Typically the members of this wood group present CMDs that are poorly populated (low number of stars) have high star-to-star scatter in both magnitudes and colors or contain strong contaminants.

3.5. The Multiband Approach

As discussed in the previous sections, our ANN uses multiple CMDs to estimate the parameters of Gaia open clusters. In particular, it uses a combination of magnitudes and colors from Gaia and 2MASS: $(BP - RP)$ versus G , $(BP - RP)$ versus J , and $(J - H)$ versus J . The use of these three different CMDs comes in handy to partially disentangle the degeneracies between age, metallicity, extinction, and distance taking advantage of the fact that different bands have different sensibilities on these parameters.

To test the improvements given by our multiband approach, we trained another group of 300 ANNs with an input layer of 64 neurons composed by the 63-entry QuadTree array of the Gaia CMD ($(G_{BP} - G_{RP})$ versus G) and the parallax of the clusters. From a visual inspection of a few hundred clusters is clear that the performance of the ANNs that use only the Gaia CMDs is worse than in the case of the multiband ones. In particular, the additional two CMDs seem to help the ANN to partially disentangle the degeneracy between the four parameters. In Figure 6 we show an insightful example of this. We plot in the three CMDs the isochrones as predicted by HR23 (purple solid lines), our single-band ANNs (red dotted lines), and our multiband ANNs (red solid line) for UBC 1516. We also report the estimated parameters of the corresponding plotted isochrones. HR23 and our single-band ANN predict similar parameters for the cluster. On the other side, multiband ANN benefits from infrared 2MASS data, producing isochrones consistent across all CMDs. Note, that the difference between $\log\text{Age}$ inferred by single-band and multiband ANNs is ~ 1 dex.

4. Results And Comparison With The Literature

In this section, we compare our estimated parameters to the ones contained in three recent works present in the literature that used different approaches to the parameter estimation problem from CMDs: CG23, D21, and HR23. Our aim is to highlight the techniques used by these authors to ease the comprehension of the differences and similarities in the results. In Figure 7 we plot the comparison between parameters derived by CG20, D21, HR23, and our ANN for open clusters with a CMD class >0.5 . The CMD class is the probability that a given cluster is a real one as derived by the neural network developed by HR23. We observe that there is a generally good agreement between our results and those from previous works. A satisfactory global agreement is also present in Figure 8, which shows the open cluster age distribution obtained by us, and the other works from the literature. In the figure, we use open clusters with a CMD class >0.5 that belong to our gold sample. However, a careful inspection of Figures 7 and 8

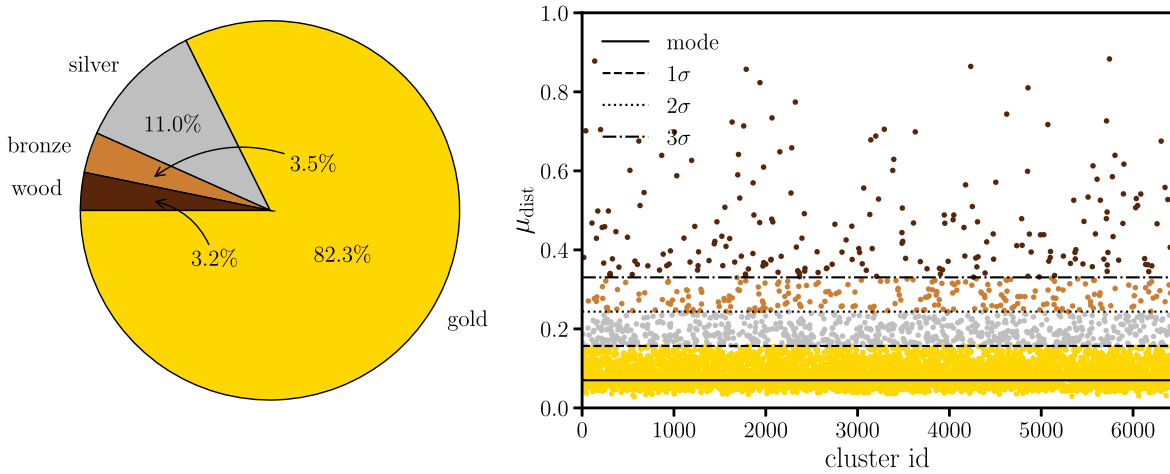


Figure 5. Left: fraction of clusters contained in gold (82.3%), silver (11.0%), bronze (3.5%), and wood (3.2%) samples. Right: distribution of μ_{dist} for the 6413 clusters analyzed in this work. Clusters are colored according to their “quality” sample (gold, silver, bronze, or wood).

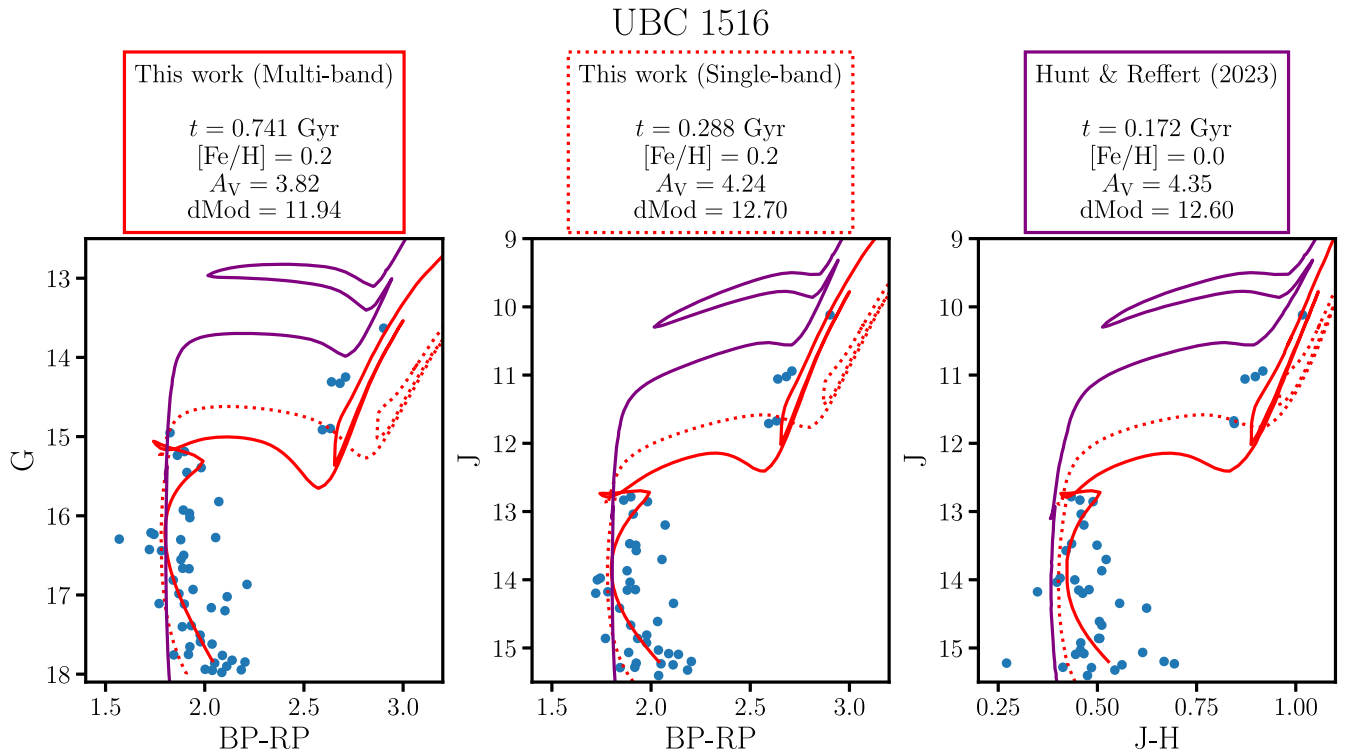


Figure 6. Isochrones and parameters for UBC 1516 inferred by [HR23](#) (solid purple line), our single-band ANN (dotted red line), and multiband ANN (solid red line). The predicted isochrones are superimposed on the corresponding CMDs of the cluster. Note that the isochrone inferred by [HR23](#) is at solar metallicity. This value is fixed and not actually derived by their ANN.

shows interesting details and small discrepancies between these studies that here we investigate in more detail in order to get insights into the different methods of analysis.

In [Figure 9](#) we directly compare the isochrones predicted with these different methods to understand differences and their impact on the overall differences between the studies.

4.1. Comparison with [CG20](#)

[CG20](#) developed a fully connected neural network that estimates the age, extinction, and distance modulus of OC using Gaia photometry. They have trained their neural network using a sample of real clusters whose parameters were

previously determined by [Bossini et al. \(2019\)](#) via isochrone fitting. Their training set was mostly composed of a subsample of well-known clusters whose features are then used to infer the parameters of a much larger sample of open clusters. They analyzed a sample of 2017 clusters contained in previous works ([Castro-Ginard et al. 2018, 2019](#); [Liu & Pang 2019](#); [Cantat-Gaudin & Anders 2020](#); [Castro-Ginard et al. 2020](#)) and they estimated the parameters for 1867 of them using Gaia DR2 photometry.

The top row of [Figure 7](#) shows the predictions from our ANN in comparison to those obtained by [CG20](#). Focusing on the age estimates we note that some clusters being younger than 100 Myr for [CG20](#) have been found to be older by our

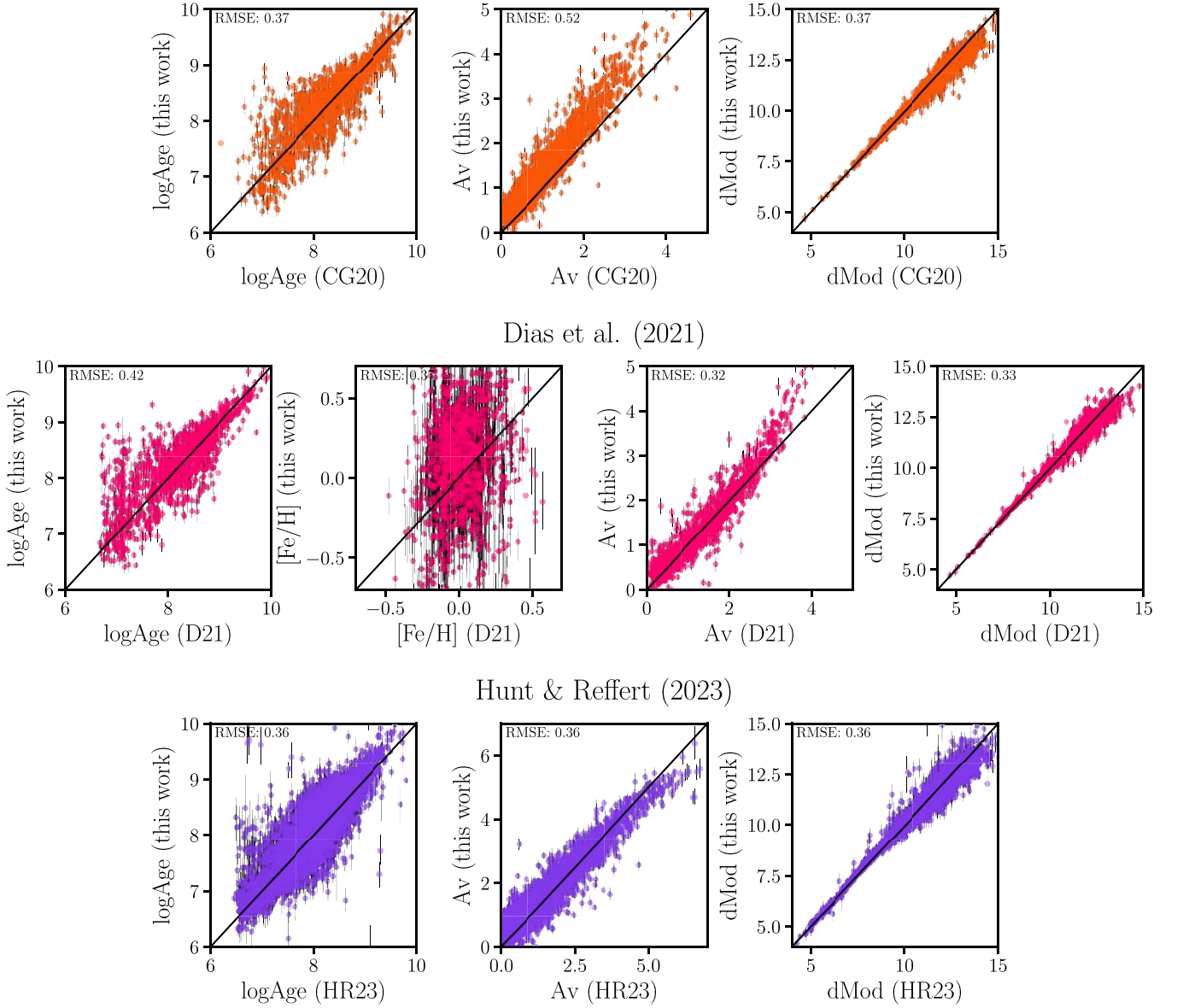


Figure 7. Comparison between parameters derived in this work with the ones present in literature. In the top left corners is reported the value of the root main square error. Top row: comparison of age, extinction, and distance for clusters that are in common with [CG20](#). Middle row: comparison of age, $[\text{Fe}/\text{H}]$, extinction, and distance for clusters in common with [D21](#). Bottom row: comparison of age, extinction, and distance for clusters contained in [HR23](#).

ANN. In some cases, the difference is of ~ 1 order of magnitude in age. The cause of this discrepancy is not uniquely clear, but it is probably related to two possible pitfalls carried out by the training set that has been used.

As described above, the CNN developed by [CG20](#) has been trained on real open clusters. This strategy has the clear advantage of training the CNN with features that are identical to those that are found in nature instead of using models and synthetic CMDs. However, as open clusters are typically younger than a few 100 Myr (see [Figure 8](#)), their training set was not well balanced across the entire age range. That might have introduced a bias in their model favoring an age underestimation for some open clusters, in particular those where a marked turnoff is still not present (e.g., age $\lesssim 500$ Myr; see [Figure 8](#)). As reported in [CG20](#), the authors initially trained the ANN with synthetic CMDs. Interestingly, while the ANN was efficient at retrieving the input parameters

from the synthetic CMDs, they did not perform as well when faced with actual, observed Gaia CMDs.

Another possible problem is related to the method used to estimate parameters for the training set: the isochrone fitting technique ([Bossini et al. 2019](#)). Although they focus on nearby populated clusters with very sharp CMDs. As we discuss in [Section 4.2](#), isochrone fitting is also prone to age underestimation when the cluster CMD is affected by faint contaminants.

The extinction values estimated by [CG20](#) are systematically lower than the ones retrieved by our ANN. On average, ours are ~ 0.4 mag larger than the one obtained by [CG20](#). We do not find a clear explanation for this trend. However, the extinctions determined by [CG20](#) are in similar disagreement also with [D21](#) and [HR23](#). This is caused by the fact that in the case of strong differential extinction and broadened CMDs, they tend to fit the blue edge of the sequence, ending up with lower extinctions.

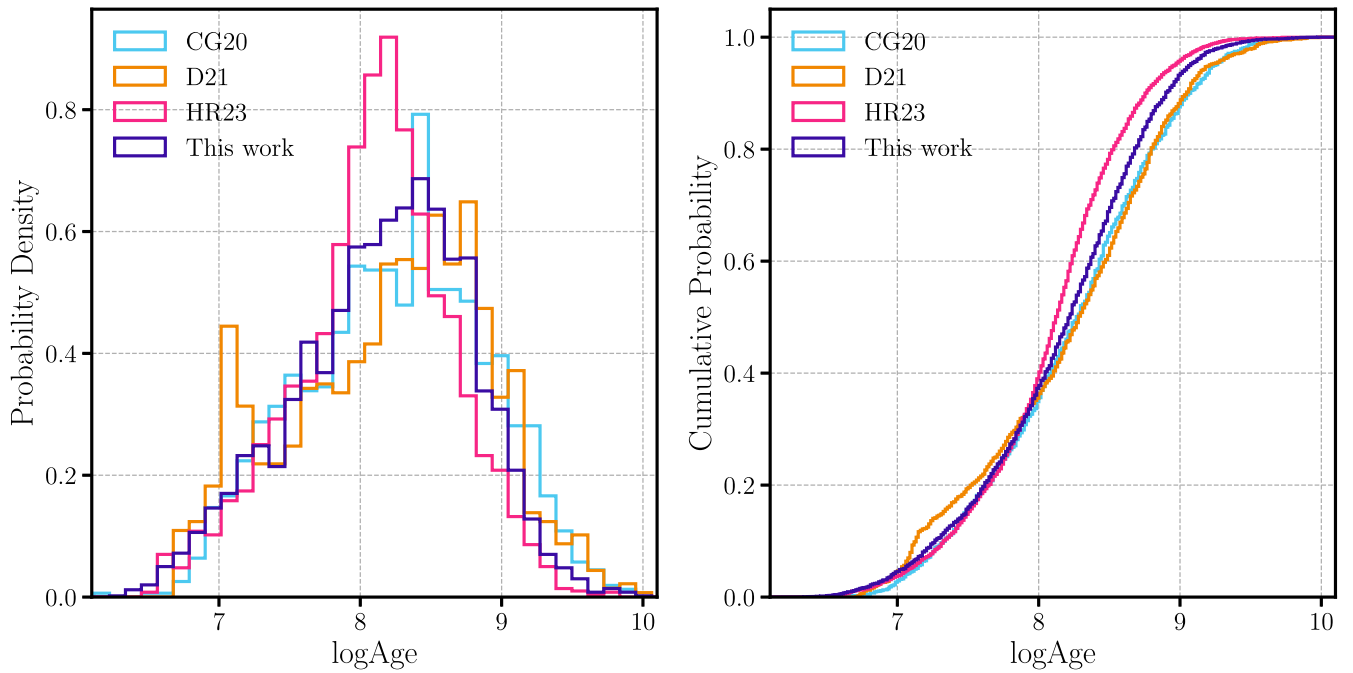


Figure 8. Age distributions of Gaia open clusters as predicted by [CG20](#), [D21](#), [HR23](#), and in this work. Left panel: differential age distribution. Right panel: cumulative age distribution.

Looking at the last plot of the top row of [Figure 7](#) is seen that our ANN tends to indicate lower distances compared to [CG20](#). This can be caused by the deviation present in the extinction coefficient. The general overestimate in A_V (redder and fainter) is partially balanced by an underestimation of the distance (brighter).

4.2. Comparison with [D21](#)

[D21](#) used Gaia DR2 data to estimate age, $[\text{Fe}/\text{H}]$, extinction, and distance modulus for 1743 open clusters retrieved via an isochrone fitting procedure using PARSEC isochrones. They analyzed a sample of 2237 clusters (for 1743 of them they successfully determine parameters) composed of clusters previously discovered by [Cantat-Gaudin et al. \(2018\)](#), [Castro-Ginard et al. \(2019\)](#), [Liu & Pang \(2019\)](#), [Sim et al. \(2019\)](#), [Cantat-Gaudin & Anders \(2020\)](#), [Castro-Ginard et al. \(2020\)](#), [Ferreira et al. \(2020\)](#), and [Monteiro et al. \(2020\)](#). As priors of their analysis, they used the Galactic metallicity gradient presented in [Donor et al. \(2020\)](#) for $[\text{Fe}/\text{H}]$ and the 3D extinction map developed by [Capitanio et al. \(2017\)](#) for A_V .

The comparison with the parameters derived by [D21](#) is plotted in the middle row of [Figure 7](#). We report a mild disagreement between our ages and those of [D21](#). More specifically, our ANN tends to assign larger ages to some clusters that have been labeled as young (i.e., $\lesssim 100$ Myr) by [D21](#). This behavior is extremely similar to that observed in the comparison with [CG20](#). The effects are visible also in [Figure 8](#) where both [D21](#) and [CG20](#) find more open clusters with ages ranging between 10 and a few tens of Myr than us. In particular, the analysis by [D21](#) results in a spike at ~ 10 Myr in the cluster age distribution.

During our investigation of the clusters with discrepant age determinations, we discovered that for many of these problematic cases, the membership from [D21](#) (and sometimes also that from [CG20](#)) contains a considerable number of faint and red contaminants which can greatly impact the performance

and reliability of the isochrone fitting procedure. The isochrone fitting method is particularly susceptible to the distribution of stars in the region of the CMD where most stars are located: the low-luminosity end of the sequence. That happens because the automatic fitting procedure wants to minimize the distance between the stars and the isochrone, regardless of where the star is located in the CMD. This strategy is somewhat different from an isochrone fitting carried out by eye, for which one gives more importance to very few stars at the turnoff or in the red clump instead of the multitudes in the lower main sequence. Therefore, if the cluster sample includes a significant number of red contaminants, the automatic fitting procedure is forced to capture most of them with a prominent pre-main sequence. On the other side, the spread of faint stars at the bottom of the CMDs may be attributed not only to contamination but also to systematic effects arising from photometric errors and blends at low signal-to-noise ratios ([Mateo 1988](#); [Vallenari & Ortolani 1993](#)). When working with a magnitude-limited (or S/N-limited) sample, overestimated luminosities are more likely to occur at the faint end. Moreover, due to the tilted nature of the main sequence with redder stars at fainter magnitudes, this effect can lead to an overestimation of pre-main-sequence (red) stars, as discussed above.

An example of the bias reported above is shown in [Figure 9](#) (left panel), where we plot the CMD of UBC 276. Although the stellar members found by [HR23](#) align in a very neat sequence (blue points), those used by [D21](#) for the fitting procedure (empty circles) include a large number of red contaminants. As a result, isochrone fitting by [D21](#) sees these latter as part of the pre-main sequence of a 6 Myr old open cluster. Instead, all the other methods result in much larger ages. Another interesting case is UBC 562 (central panel of [Figure 9](#)), where both the results by [D21](#) and [CG20](#) are off from the upper cluster sequence. However, they both succeed in fitting the lower sequence which is rich in stars, most of which are contaminants according to [HR23](#). It is difficult to fully understand this

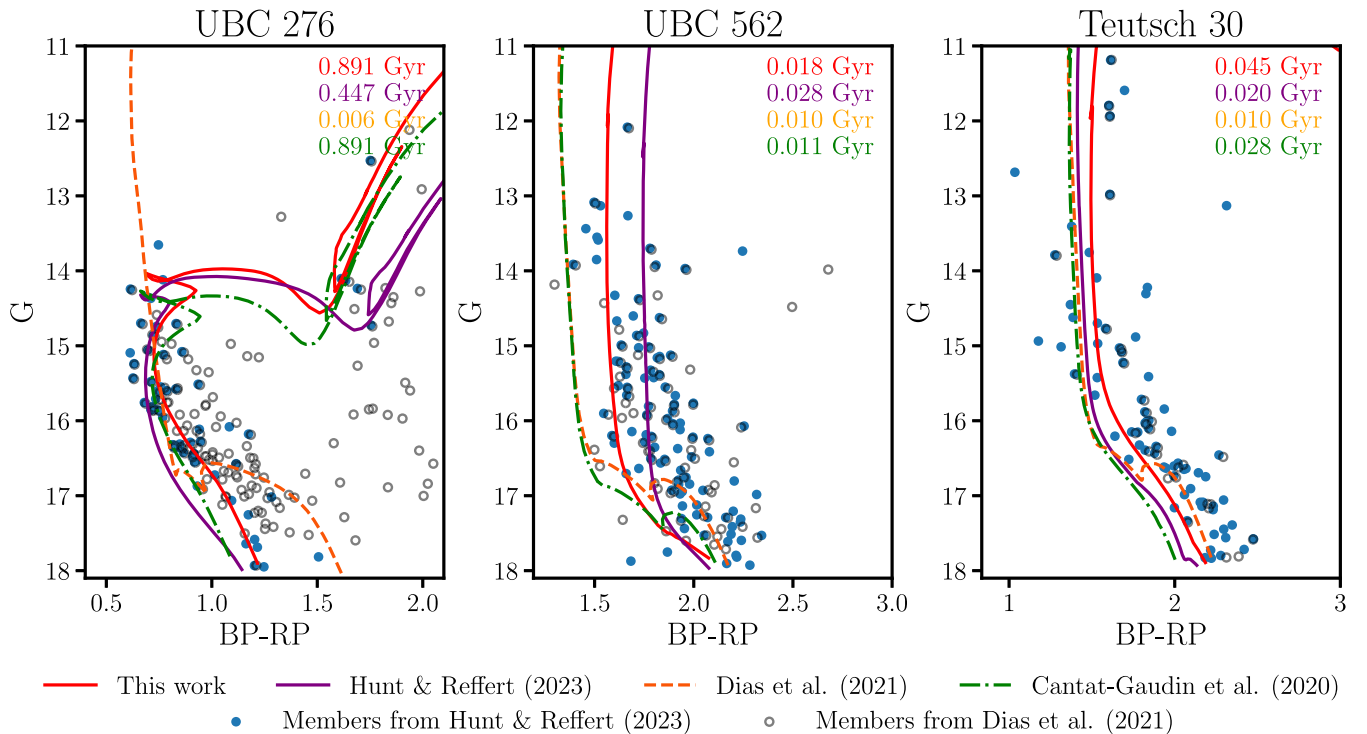


Figure 9. Predicted isochrones for three open clusters UBC 276 (left panel), UBC 562 (middle panel), and Teutsch 30 (right panel) from *CG20* (dotted–dashed green line), *D21* (dashed orange line), *HR23* (solid purple line), and this work (solid red line). In the top right, we annotate the ages that correspond to the plotted isochrones (colored accordingly). With blue and full dots we plot the members of the clusters retrieved by *HR23*, and with empty black dots we plot the ones of *D21*.

strange behavior; however, it is possible that a large contamination of nonmembers in the low-luminosity end of the sequence can produce deep potential wheels in the isochrone fitting loss function and force the entire procedure to fit these low-luminosity stars with pre-main sequences.

Another valuable example is Teutsch 30, plotted in the right panel of Figure 9. In this case, *D21* fit the low-mass end of the sequence with a pre-sequence phase (widely discussed above). On the other side, both *CG20* and *HR23* seem to follow the blue part of the CMD losing the group of bright stars at $G < 12$ (members of the cluster for both *D21* and *HR23*). We find several other cases where the isochrone fitting procedure is affected by red and faint contaminants in a very similar way. This problem—which is the cause of the spike at 10 Myr in Figure 8—is especially important for those cases where the number of nonmembers is comparable to the one of “real” members. Furthermore, only young open clusters (e.g., age $\lesssim 100$ Myr) can have their ages underestimated by isochrone fitting. On the contrary, the isochrone fit method has been proven to be highly reliable for analyzing old stellar populations, as it is facilitated by the presence of strong and distinct age constraints such as the turnoff, red clump, and prominent giant branches. In this sense, it is reassuring to observe a general agreement between our age estimates for older open clusters ($\log \text{Age} > 8.5$) and those from *CG20* and *D21*.

From the clusters plotted in Figure 9 is possible to notice that *HR23* tends to predict isochrones slightly bluer than the ones of our ANN. This tendency will be discussed in details in Section 4.3.

Metallicity determination solely based on CMDs is an extremely difficult task. For that reason *D21* have fitted their isochrones imposing two strong priors on $[\text{Fe}/\text{H}]$, and also on

A_v . Instead, our strategy is to let the ANN decide which is the best $[\text{Fe}/\text{H}]$ without any bias or prior. As a result, we observe a large scatter between our $[\text{Fe}/\text{H}]$ determinations and those of *D21*, with a total root mean square error (RMSE) that reaches ~ 0.35 dex. However, with the values of $[\text{Fe}/\text{H}]$ and distance modulus predicted by our ANN we are able to obtain the metallicity gradient observed in the Galaxy (see Section 5.2).

Just as for $[\text{Fe}/\text{H}]$, *D21* have imposed priors also for the A_v parameter. However, we notice a general agreement between us and *D21*, with only a clear trend at high extinction factors. Namely, our ANN tends to overestimate A_v of high extinction clusters (i.e., with $A_v > 3$ mag). We do not find a clear cause of this trend.

The distance estimations of our ANN are in general good agreement with the ones obtained by *D21* with the exception of a slight underestimation trend for the most distant clusters.

4.3. Comparison with *HR23*

HR23 used a quasi-Bayesian CNN trained with synthetic clusters to estimate age, extinction, and distance of clusters contained in their catalog. In the bottom row of Figure 7 we only compare the parameters obtained by our ANN of cluster labeled as “real” by *HR23* (i.e., median CMD class > 0.5 , see *HR23*). The clusters used for this comparison are the same plotted for *HR23* in the cluster age distribution of Figure 8.

The major discrepancy between our results and those by *HR23* is particularly evident in Figure 8 as a quasi-systematic shift in the age distributions. More specifically, the age distribution found by *HR23* is mostly shifted toward lower ages and also seems to be narrower than ours. Additionally, *HR23* have a lower fraction of clusters older than $\log \text{Age} > 8.5$ dex compared to all the other works discussed in the present paper, which are mutually consistent.

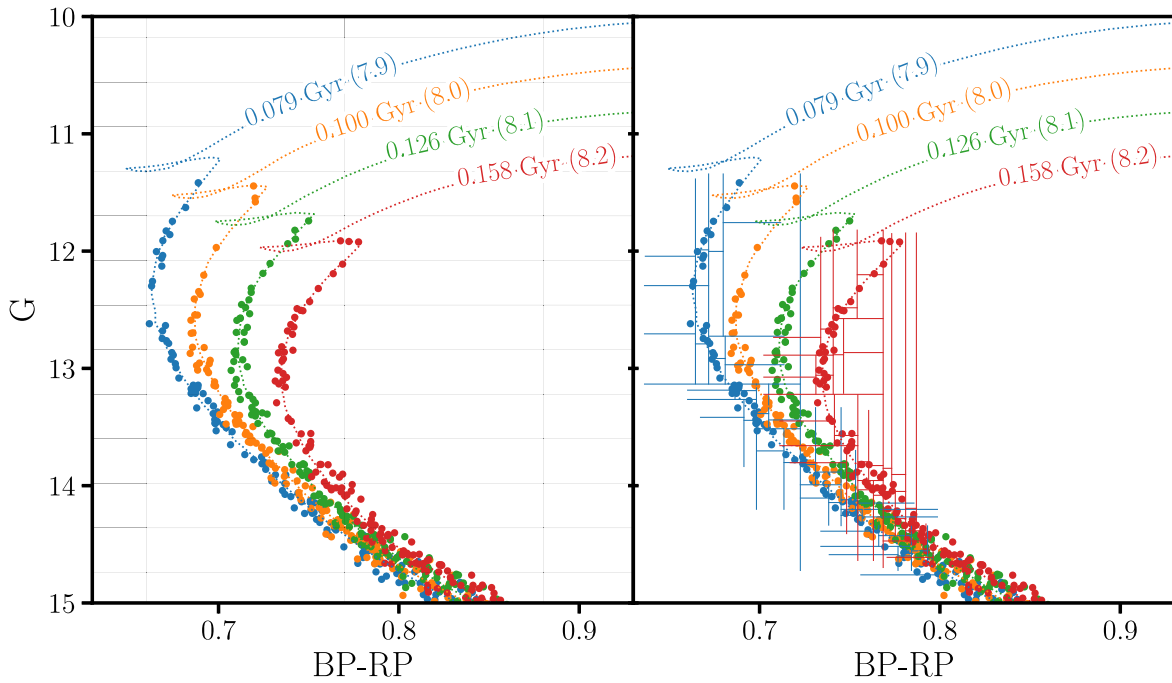


Figure 10. Synthetic CMDs generated by the code described in Section 2.2. The mock clusters are generated with $N_{\text{star}} = 400$, $[\text{Fe}/\text{H}] = 0$, $A_v = 2$, $d\text{Mod} = 12$ mag, and $\log\text{Age}$ between 7.9 and 8.2. Single members of simulated clusters are plotted with dots colored according to their age. With dotted lines, we plot the theoretical isochrones used to generate the synthetic CMDs. Left panel: synthetic CMDs superimposed to the grid used by HR23. Right panel: synthetic CMDs superimposed to the QuadTrees of the simulated clusters with $\log\text{Age} = 7.9$ (blue) and $\log\text{Age} = 8.2$.

This peculiarity may suggest that the CNNs of HR23 tend to underestimate the age of clusters older than 500 Myr considering them in the age range $8 \lesssim \log\text{Age} \lesssim 8.4$. Hence, their distribution peaks at $\log\text{Age} \sim 8.1$, instead of $\log\text{Age} \sim 8.5$ such as it is consistently found by the other works compared in Figure 8. This difference between us and HR23 is particularly surprising because both the studies are based on the same sample of stellar members, thus the different outputs must entirely be the products of the different approaches of analysis. Given that we both use neural networks to estimate the parameters, the difference must be due to the ways HR23 and us extract information from the CMDs.

More specifically, HR23 feed their CNN with 2D histograms of the cluster CMD. These images are formed by 32×32 pixels and each pixel brings information on the density of stars falling within that small portion of the grid. The pixels have sizes of 0.38 mag in G and 0.11 mag in $(G_{\text{BP}} - G_{\text{RP}})$. Instead, our strategy is to extract the relevant information using a QuadTrees algorithm (see Section 3.1) whose results are then directly fed to our ANN. In Figure 10 we show how these different methods can affect the final results. Namely, we plot four synthetic clusters with ages in the range $7.9 \leq \log\text{Age} \leq 8.2$. In the left panel, we plot the pixelation grid used by HR23, on the right panel we show the feature extraction carried out by the QuadTrees. From the left panel of Figure 10 is clear that all the synthetic CMDs fall into the same color bins due to the pixelation and—even though they have different ages—they become indistinguishable to the neural network.

When a neural network has to indiscriminately choose between a range of outputs, it will likely choose the most probable one. Typically there are more young isochrones than old ones crossing a specific color bin. Therefore, when the model has to choose between all the isochrones fitting a given pixel it is biased toward younger ages which are considered the

most probable by the network. By contrast, the QuadTree technique produces coordinates that are always different between clusters of different ages (see right panel of Figure 10). In conclusion, pixelation is not as efficient as QuadTree in exploiting at maximum the full potential of precise magnitudes produced by Gaia. This loss of resolution is probably further magnified by the differential reddening introduced by HR23 in their synthetic CMDs and that has the effect of making the cluster sequences tickier.

A consequence of this loss of resolution due to pixelation is also seen in Figure 5 of HR23 (top row), which plots the performance of the ANN to retrieve the parameters of the test sample of clusters. In their third plot, the one on differential reddening, it is possible to notice that the model is not able to correctly recognize clusters with no extinction and therefore assigns to them a nonzero value. It is likely that their model sees the clusters with no differential reddening and small differential reddening as indistinguishable due to the loss of resolution caused by pixelation. Therefore, it assigns at least a small amount of differential reddening to all clusters simply because that is the most *probable* option as it has been learned by the CNN.

On top of pixelation, there is another key difference between the methods used by HR23 and our analysis which is related to the type of neural network. More specifically, HR23 adopt convolutional neural networks (CNNs) fed with images of CMDs. The convolution operation in combination with max pooling makes CNNs approximately invariant to translation. This type of inductive bias has significantly contributed to the success of CNNs for classification tasks such as image recognition, where the algorithm has to identify specific objects or patterns within an image: typically the ability to detect patterns independently from their location within the image can greatly improve their generalization ability. On the other hand,

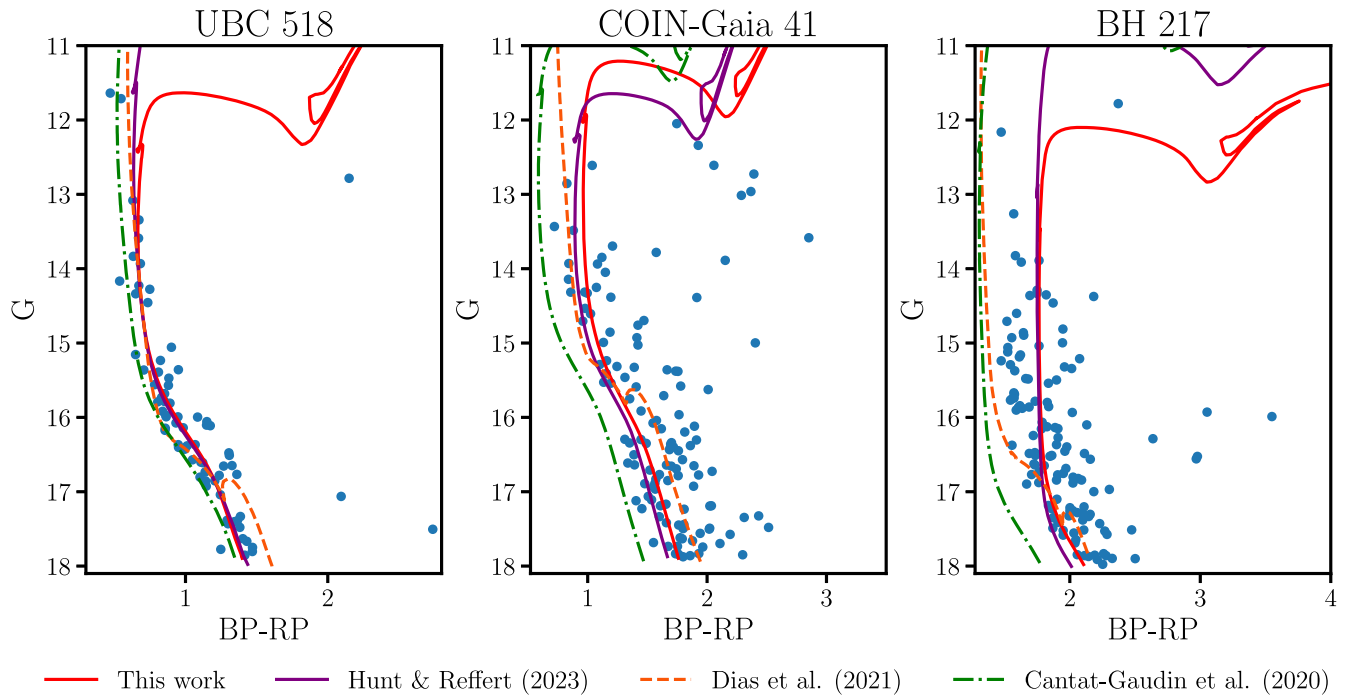


Figure 11. Predicted isochrones for three open clusters UBC 518 (left panel), COIN-Gaia 41 (middle panel), and BH 217 (right panel) from **CG20** (dotted–dashed green line), **D21** (dashed orange line), **HR23** (solid purple line), and this work (solid red line). With blue dots, we plot the members of the clusters retrieved by **HR23**.

this translation invariance can hurt the performance of CNNs in cases where the position of the objects in the image matters. This is exactly the case of analyzing the sequences of stars in a CMD, where positional information is critical and translation always results in different cluster parameters. Instead, our analysis is based on ANNs fed with the arrays of numerical values generated by QuadTrees. Notably, ANNs excel in preserving most of the information passed to the input neurons. Although that can limit the generalization ability of the algorithm, it is certainly a valuable trait for the analysis of CMDs.

Besides the systematic shifts discussed above, in the bottom row of Figure 7, it is possible to notice also a group of outliers, reported to be extremely young ($\log \text{Age} < 7$) by **HR23**, that our ANN classifies as intermediate/old age ($\log \text{Age} > 7.5$). These clusters are generally poorly populated and in some cases very sparse. Part of the discrepancies arise from clusters that are contained in our bronze/wood sample. In other cases, the discrepancies are not caused by the ANN approach used by us and **HR23** but rise from an intrinsic limitation of using CMDs to estimate the age of clusters. In particular, in some cases, there is the absence of clear features (e.g., subgiant branch, turnoff) useful to remove the degeneration in age. It is clear that the limitation is exponentially higher in clusters with few members (as in these cases). As expected, part of the outliers in $\log \text{Age}$ are also out-layers in A_v and $d\text{Mod}$.

4.4. Possible Pitfalls in Our Analysis

While testing our ANN we find that our model is very sensitive to the presence of stars in the red part of the CMD. This sensitivity is both a strength and a weakness, depending on the case.

We know that the red clump is a strong age indicator. Stars in the red clump have not completed the He burning and they are burning helium in the core, which is surrounded by an

H-burning shell. They have relatively similar masses and luminosities, resulting in a tight grouping in the CMD. Our ANN seems to have learned that the presence of stars in the red part of the CMD is a strong age indicator. This characteristic of our ANN is clearly visible in Figure 11 where we plot the isochrones inferred by **CG20**, **D21**, **HR23**, and our ANN for three different clusters. Focusing on UBC 518 (left panel of Figure 11) our ANN retrieves an isochrone (red solid line) that is significantly different from the ones obtained by the other works. In particular, the sensibility of our ANN leads to an intermediate-age isochrone that reaches the red and bright star at $G \sim 13$. If that star is a real member of the cluster we can conclude that our ANN has a better performance with respect to the other works. The cluster COIN-Gaia 41 (central panel of Figure 11) is another case where the sensibility of our ANN is functional. In fact, for this cluster, our ANN is able to correctly identify the stars that belong to the red clump (if real and are not contaminants). This trait of our ANN leads to an isochrone that seems more in agreement with the distribution of stars in the CMD compared to the ones derived by **CG20** and **D21**. We notice that the tool developed by **HR23** has produced an isochrone that is very similar to the one obtained by us.

However, this high sensibility to stars on the left side of the CMDs has the drawback of being susceptible to red contaminants. That feature is visible from the CMD of BH 217 (right panel of Figure 11). In that case, the isochrone inferred by our ANN is not compatible with the observed CMD. It is plausible that our ANN tends to indicate an intermediate-age isochrone with a prominent turnoff because of the presence of the group of faint and red stars (probably contaminants). It is worth noticing, that for this cluster **HR23** have obtained a more plausible isochrone (purple solid line) also compared to **CG20** and **D21**.

Therefore, it is clear that the ANN developed in this work is susceptible to red contaminants in the observed CMDs.

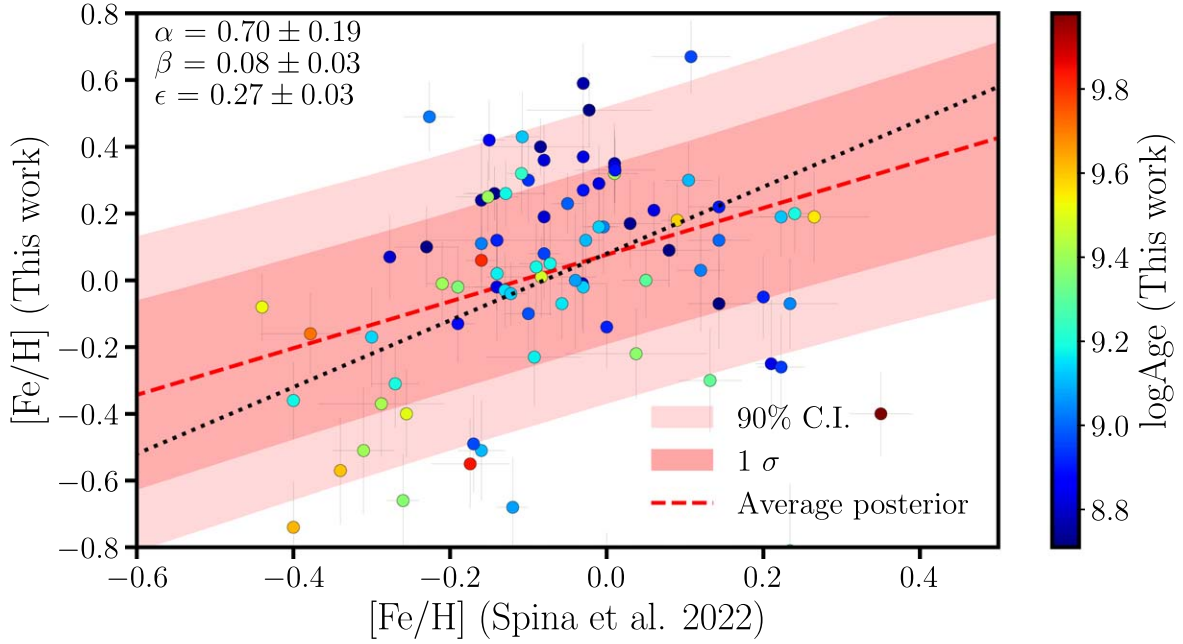


Figure 12. Comparison between $[\text{Fe}/\text{H}]$ derived in this work and the one obtained from spectroscopic data by Spina et al. (2022). Clusters are color coded based on $\log\text{Age}$ as determined by our ANN. The average posterior of the fitted relation is plotted by a dashed red line, while the red-shaded areas represent the 68% and 90% confidence intervals. In the top left corner, we provide the mean and standard deviations for the three free parameters. With a dotted black line, we plot the relation with $\alpha = 1$ and $\beta = 0.08$.

Currently, our ANN is sensible to red members regardless of their magnitude. Thus, even when only red and faint members are present in the CMD, the ANN tends to pass to a nonexistent red clump (see right panel of Figure 11). In order to mitigate this problem we use a dropout layer right after the input layer (see Section 3). However, the undesired effect is not completely removed. In future work, we plan to improve our ANN in order to avoid these misleading cases.

5. Scientific Validation

In the following, we present a comprehensive scientific validation of the proposed ANN methodology. We seek to demonstrate its effectiveness in accurately determining crucial parameters of Gaia OCs such as age, metallicity, extinction, and distance.

5.1. Comparison with Spectroscopic Data

In Section 4.2 we noticed that the $[\text{Fe}/\text{H}]$ predictions of our ANNs have a large scatter around the one of D21 (derived with an isochrone fitting procedure with the galactic gradient as a prior), with an RSME ~ 0.35 dex. To test and validate the metallicity predictions of our tool we compare them with a sample of OCs with spectroscopic data and thus $[\text{Fe}/\text{H}]$ estimates.

For this test, we use the catalog of 251 Galactic OCs produced by Spina et al. (2022). The sample has been built through the collective analysis of homogenized data, obtained from high-resolution spectroscopic surveys and programs, namely, APOGEE (Majewski et al. 2017), Gaia-ESO (Gilmore et al. 2012), GALAH (Buder et al. 2018), OCCASO (Casamiquela et al. 2016), and SPA (Origlia et al. 2019). From the original sample, we select clusters with precise $[\text{Fe}/\text{H}]$ estimates (i.e., with errors < 0.1 dex) that are present in our analyzed ones, ending up with 197 objects. We further selected those open clusters which: 9i) have ages greater than 500 Myr;

(ii) are contained in the gold sample; (iii) have at least a median CMD class > 0.5 (i.e., reliably “real”, see HR23). The final test sample is composed of 89 clusters.

To test the correlation between the metallicities estimated by our ANNs and the one obtained from spectroscopy we use a simple linear model $y_i = \alpha \times x_i + \beta$, where y_i and x_i are normal distributions defined as follows: $y_i = \mathcal{N}([\text{Fe}/\text{H}]_i^{\text{ANN}}, \sigma_{[\text{Fe}/\text{H},i]}^{\text{ANN}})$ and $x_i = \mathcal{N}([\text{Fe}/\text{H}]_i^{\text{S22}}, \sigma_{[\text{Fe}/\text{H},i]}^{\text{S22}})$. $\sigma_{[\text{Fe}/\text{H},i]}^{\text{ANN}}$ and $\sigma_{[\text{Fe}/\text{H},i]}^{\text{S22}}$ are the uncertainties associated with the predictions of our ANN on the i -th cluster and the ones of the estimates of Spina et al. (2022), respectively. Beyond α and β , we also introduce a third free parameter, ϵ , to account for the intrinsic scatter inherent in the data. This scatter is attributed to a blend of processes that are difficult to model and/or are neglected in our modeling. For α and β we chose priors that are $\mathcal{N}(1, 1)$ and $\mathcal{N}(0 \text{ dex}, 1 \text{ dex})$, respectively. For the prior of the intrinsic scatter ϵ we adopt a positive half-Cauchy distribution with $\gamma = 1$. Using the Python package PYMC3 (Salvatier et al. 2016) we compute 10,000 samples using a No-U-Turn Sampler (Hoffman & Gelman 2011); half of them have been discarded as part of the burn-in phase.

In Figure 12 we plot the selected open clusters as a function of their $[\text{Fe}/\text{H}]$ as predicted by Spina et al. (2022) and our ANNs, with relative errors, colored accordingly to their $\log\text{Age}$ (from our ANNs). With a red dashed line we plot the average relation, and with red-shaded areas the 68% and 90% confidence interval (C.I.) of the posteriors of the free parameters. In the top left corner of the figure, we report the mean values of α , β , and γ , with relative uncertainties. From this analysis, we found a mild correlation between the $[\text{Fe}/\text{H}]$ values, which proves that our ANN is able to extract information on the cluster metallicity from its CMDs. It is evident that the ANN does not merely produce random values within the $[\text{Fe}/\text{H}]$ training range. Instead, it can effectively differentiate between metal-poor and solar-metallicity clusters.

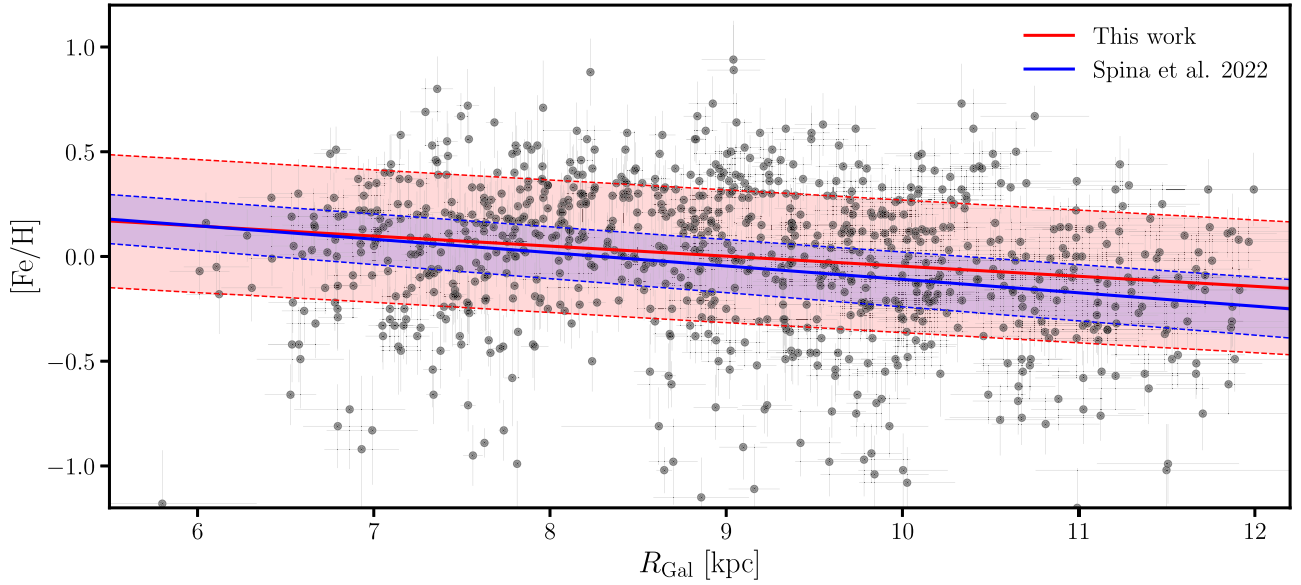


Figure 13. Values of $[\text{Fe}/\text{H}]$ as a function of the distance from the Galactic center (R_{Gal}). With black dots, we plot the clusters analyzed in this work and selected as reported in Section 5.2 and their respective uncertainties. With a red solid line, we plot the relation obtained by the mean values of α and β while a 68% confidence interval is plotted with a red-shaded area. We also superimpose the metallicity gradient obtained by Spina et al. (2022), mean with a solid blue line and 68% confidence interval with a blue-shaded area.

Table 2

Mean, σ , and 68% Confidence Interval of α , β , and ϵ Obtained from our Bayesian Analysis

Parameter	Mean	σ	68% C.I.
α [kpc^{-1}]	-0.048	0.009	[-0.057, -0.039]
β	0.43	0.08	[0.35, 0.51]
ϵ	0.313	0.009	[0.303, 0.322]

In particular, Figure 12 reveals a noticeable absence of points in the top left corner, while the region around $[\text{Fe}/\text{H}] = 0$ dex is densely populated.

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.

5.2. Metallicity Gradient

In our Galaxy, the metal content in disk stars decreases moving from the inner to the outer parts. This fact is known as the Galactic metallicity gradient. In particular, the metal content $[\text{Fe}/\text{H}]$ decreases as the distance to the Galactic center (R_{Gal}) increases. In the inner part ($R_{\text{Gal}} < 12$ kpc) of the Galaxy the $[\text{Fe}/\text{H}]$ - R_{Gal} relation is linear (see Spina et al. 2022).

As a validation test, we aim to verify that the $[\text{Fe}/\text{H}]$ and dMod parameters estimated by our ANN are consistent with the observed Galactic metallicity gradient. To do that, we first transform the distance modulus in R_{Gal} , assuming a distance of the Sun from the Galactic center of 8.122 kpc (GRAVITY Collaboration et al. 2018). For this test, we selected from the sample of open clusters those which: (i) have ages greater than 500 Myr; (ii) are contained in the gold sample; (iii) are located from 5 to 12 kpc from the Galactic center; (iv) have at least a median CMD class > 0.5 (i.e., reliably “real”, see HR23).

We model the gradient using a simple linear model $y_i = \alpha \times x_i + \beta$, where y_i and x_i are normal distributions defined as follows: $y_i = \mathcal{N}([\text{Fe}/\text{H}]_i, \sigma_{[\text{Fe}/\text{H}]_i})$ and $x_i = \mathcal{N}(R_{\text{Gal},i}, \sigma_{R_{\text{Gal},i}})$. $\sigma_{R_{\text{Gal},i}}$ and $\sigma_{[\text{Fe}/\text{H}]_i}$ are the uncertainties associated with the

predictions of our ANN on the i -th cluster. Note that in general the uncertainties on the parameters extracted by our ANN might be asymmetric. Given that the asymmetry is ~ 1 order of magnitude less than the uncertainty itself, we compute $\sigma_{R_{\text{Gal},i}}$ and $\sigma_{[\text{Fe}/\text{H}]_i}$ by averaging the uncertainties of our ANN. Besides α and β we also fit the intrinsic scatter of data with a third free parameter ϵ . The source of the scatter is a combination of processes that are difficult to model and predict, such as migration and systematics. For α and β we chose priors that are $\mathcal{N}(-0.068 \text{ dex kpc}^{-1}, 0.1 \text{ dex kpc}^{-1})$ and $\mathcal{N}(0.5 \text{ dex}, 1 \text{ dex})$, respectively. For the prior of the intrinsic scatter ϵ we adopt a positive half-Cauchy distribution with $\gamma = 1$. Using the Python package PYMC3 (Salvatier et al. 2016) we compute 10’000 samples using a No-U-Turn Sampler (Hoffman & Gelman 2011); half of them have been discarded as part of the burn-in phase.

In Table 2 we report the mean deviation, standard deviation, and 68% confidence interval (C.I.) of the posteriors of α , β , and ϵ . In Figure 13 we plot with a red solid line the model obtained from the mean values of α and β obtained from our Bayesian analysis. The 68% confidence interval is plotted with a red-shaded area. We compare the gradient obtained in this work with the one retrieved from spectroscopic observations by Spina et al. (2022). We found that the gradient estimated from the predictions of our ANN is consistent with the one obtained by Spina et al. (2022). We also found that in our case the intrinsic scatter ϵ is ~ 3 times larger than the one found by Spina et al. (2022). The discrepancy likely arises from our estimation of metallicity using the CMD based on photometric data, which is less precise than determinations made through spectroscopy. Additionally, the uncertainties in the parameters might be underestimated, further contributing to the difference. Nevertheless, this result demonstrates that the advancement of our method (i.e., including $[\text{Fe}/\text{H}]$ as an output parameter) is real. The information on metallicity can be reliably extracted by the sequence of stars in the CMD built on Gaia and 2MASS photometry.

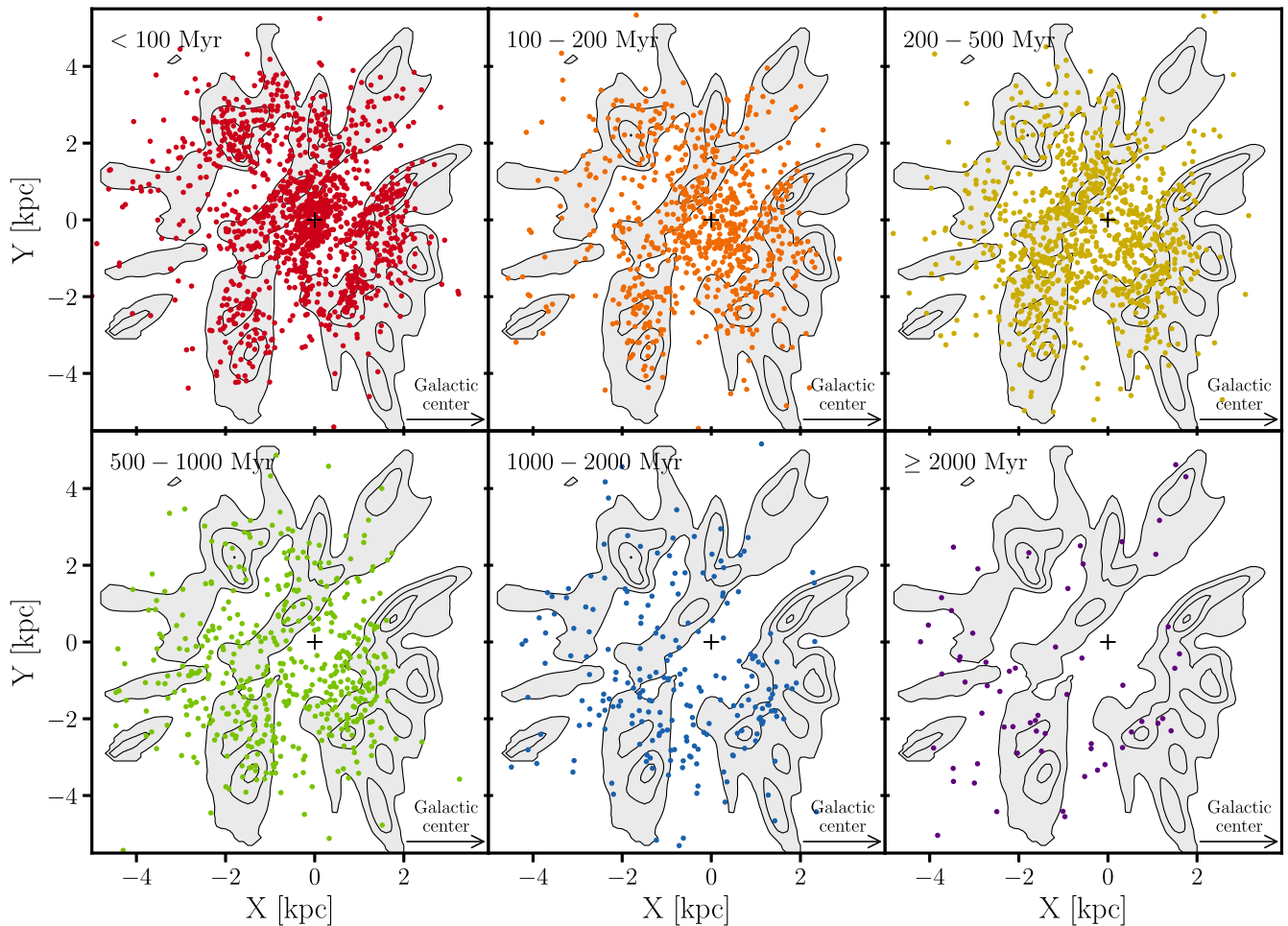


Figure 14. Spatial distribution of intrinsically bright open clusters divided into age bins superimposed on the overdensity map (in gray) measured by Poggio et al. (2021). The Sun is located in $(X, Y) = (0, 0)$, the Galactic Center is to the right, and Y is taken as positive in the direction of Galactic rotation.

5.3. The Distribution of Open Clusters within the Galactic Disk

The spatial distribution of open clusters within the Galactic disk is an important feature that we can inspect in order to validate our results. In Figure 14 we plot the distribution of open clusters as a function of the Galactic cartesian coordinates X and Y for six different age bins. Among the entire sample, here we only show the open clusters that (i) are included in the gold sample; (ii) have a $\text{CMDclass} > 0.5$; (iii) have at least 3 stellar members with absolute magnitude lower than 3 dex. This latter condition allows us to exclusively select the intrinsically bright open clusters, which is necessary to highlight the spatial features formed by open clusters in the disk (see, Poggio et al. 2021; Cantat-Gaudin 2022; Gaia Collaboration et al. 2023a; Drimme et al. 2023). In fact, a significant number of known stellar clusters are small and young associations located very close to the Sun. Some of them remain visible in the upper left panel even after the cut in magnitude. In Figure 14 we also show the position of the spiral arms from upper main-sequence stars in Poggio et al. (2021), shown as gray-shaded areas and contours. Starting from the upper left of these panels the contours show the Perseus Arm, the Local Arm, and finally the Sag-Car Arm in the lower right corner. From the comparison between the location of open clusters and spiral arms, we notice that younger clusters are more concentrated near the spiral arms. It is particularly evident the clump of clusters younger than 100 Myr is located in the

Perseus Arm at $(X, Y) \sim (-1.7, 2)$. Moreover, the higher density of clusters in the Local and Sag-Car arms is visible for ages up to 500 Myr. That is an expected behavior as spiral arms enhance the formation of giant molecular clouds, hence stellar associations. As young clusters had no time to migrate and disperse throughout the Galactic disk, they are still close to the spirals: the large-scale structures from which they were born. Instead, as we move toward the older clusters, these latter are more randomly distributed across the Galactic disk. It is therefore plausible that old clusters have significantly migrated from their original birthplaces. In Figure 15 we show the distributions of heights above the plane for open clusters in different age bins. Here we consider the same open clusters plotted in the left panel and with a distance from the Sun that is less than 2 kpc. Here we notice that the youngest clusters are spatially concentrated on the Galactic midplane. However, as we move to older clusters the distribution spreads at larger heights. The distribution of the oldest age bin is nearly flat. Similarly to what is observed in Figure 14 from the comparison with the spiral arms, this general behavior described above is widely expected. Open clusters typically form very close to the midplane where larger gas reservoirs are available for star formation. Once an open cluster is formed and gets older, it has time to migrate far above the Galactic midplane (e.g., Jílková et al. 2012; Tarricq et al. 2021; Viscasillas Vaézquez et al. 2023). On top of that, the open clusters living to close the disk

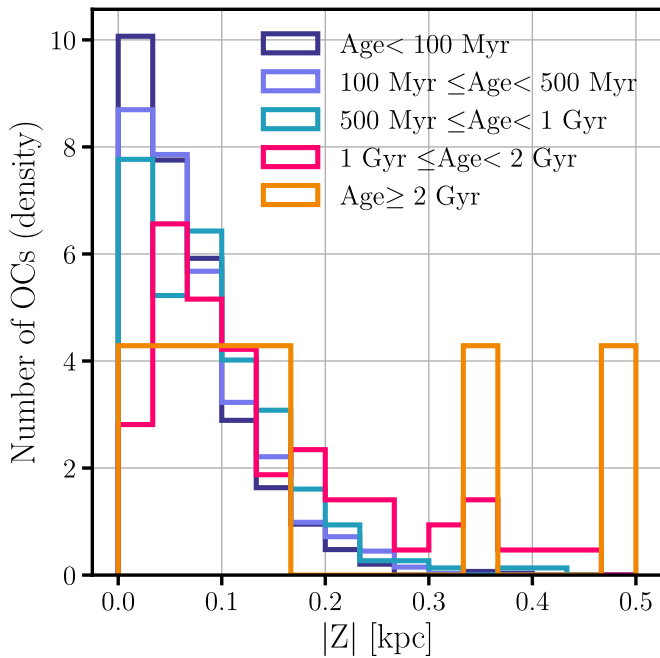


Figure 15. Distribution of heights above the Galactic midplane for intrinsically bright open clusters in different age bins.

have higher chances of being disrupted by perturbations of the Galactic potential, such as giant molecular clouds and spiral arms. All of that contributes to widening the distributions of older clusters.

6. Conclusions

In this work, we perform the parameter estimation of the stellar cluster (~ 6400) contained in the sample presented by HR23 using color–magnitude diagrams from both Gaia and 2MASS. To do that, we use a group of artificial neural networks trained on synthetic clusters based on PARSEC isochrones. To extract the features for our neural networks, we pre-process the color–magnitude diagrams using the QuadTree algorithm. Our conclusions can be summarized as follows:

1. We obtain credible estimates of age, metallicity, extinction, and distance for ~ 5400 clusters (i.e., the ones included in our gold sample). These parameters pass through a comprehensive scientific validation that demonstrates the reliability of our artificial neural network to determine the vital parameters of Gaia open clusters. We compared the metallicity estimates of our ANNs with ones obtained from spectroscopic studies obtaining promising results. Furthermore, with our parameters, we are able to reconstruct the Galactic metallicity gradient visible in the inner part of the disk. These results demonstrate that our method is able to reliably extract the information on metallicity from the color–magnitude diagram built on Gaia photometry. With the parameters obtained from our network, we derive the 3D distribution of clusters around the Sun. From that, we found that young clusters are the closest tracers of the spiral arms of the Galactic disk (see Poggio et al. 2021) and that there is a clear increase in the scale height of clusters with their age. Both of these things are widely

expected by current models and thus they represent a significant scientific validation.

2. We compare our results with the parameters obtained by previous works that used a wide variety of techniques: CG20, D21, and HR23. We thus investigate the possible sources of some discrepancies between this work and the ones cited above. The isochrone fitting technique used by D21 suffers from faint and red star contaminants and in these cases tends to fit a pre-sequence that is not present in the observed CMD. CG20 ANN has been trained with observed CMDs with known parameters (obtained via isochrone fitting) thus the trained neural network is potentially biased as the training sample. 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. With our analysis, we find systematically older age compared to the previous works presented here. This is an interesting result, with some possible relevant specific cases.
3. In this work, we make use of a novel approach to the feature extraction task by adopting a QuadTree algorithm. The QuadTree is capable of efficiently tracing the sequence even with photometric errors and outliers. The use of a QuadTree algorithm to extract numerical features from the CMDs paves the way to a second key element of novelty in the analysis, which is the use of ANNs. Previous works based on machine learning techniques have made use of CNNs (see, CG20 and HR23). However, these types of algorithms are approximately invariant to translation. That specific trait of CNNs is due to convolution and max pooling operations upon which these algorithms are designed, but also to the pixelation of the input CMDs images which necessarily results in a loss of spatial resolution. The translation invariance can hurt the analysis of sequences of stars in CMDs. Instead, ANNs fed with numerical arrays preserve the full positional information which is expected to improve the performances. Finally, with our analysis, we also demonstrate that cluster parameters can be derived from the simultaneous analysis of multiple photometric bands. This prospect will be particularly useful in view of next-generation telescopes, such as the Vera Rubin Observatory which will provide photometry from six filters for $\sim 20 \times 10^9$ stars.
4. In the last years clustering algorithms have found more and more new clusters so it become crucial to develop parameters and/or techniques to discriminate if a group of stars is a real cluster. We found that finding an isochrone that fits the CMD of a generic cluster is a necessary but not sufficient condition to classify a cluster as real. Further analysis and studies need to be done in the future to address this crucial problem.

In recent years, a series of works CG20, HR23, and this work have proved that artificial neural networks are powerful and reliable tools to analyze Gaia open clusters. They also proved performances similar (or potentially superior) to the isochrone fitting technique while having lower computational costs. On the other hand, isochrone fitting methods are costly in terms of computational time because of the large parameter space that they need to explore. For this reason, it is common to assume strong priors on (some) parameters (e.g., see D21) that can

introduce some biases. One possibility is to merge the two techniques taking advantage of the strong points of each one. In particular, is possible to perform a preliminary analysis by means of an ANN-based tool (e.g., similar to the one presented in this work) that can generate the distribution of the plausible parameters. The estimates of the ANN are then refined by taking them as priors of an isochrone fit procedure. This new approach will be developed and tested in future works. Future improvements to the model outlined in this work will incorporate more sophisticated and precise modeling of synthetic CMDs, taking into account factors like binaries, stellar rotation, and differential extinction.

Acknowledgments

We warmly thank the anonymous referee for their constructive and beneficial feedback, which has notably enhanced our manuscript.

L.C. thanks Dr. Laura Greggio for the insightful discussion on stellar evolution and stellar cluster analysis. This research

has made use of the [SVO Filter Profile Service](#) supported by the Spanish MINECO through grant AYA2017-84089.

M.P. acknowledges financial support from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 896248. S.L. acknowledges financial support by the PRIN INAF 2019 grant ObFu 1.05.01.85.14 (“Building up the halo: chemo-dynamical tagging in the age of large surveys”). J.S-U. was supported by the National Agency for Research and Development (ANID)/Programa de Becas de Doctorado en el extranjero/DOCTORADO BECASCHILE/2019-72200126.











Appendix Our Catalog of Open Clusters Parameters

Table 3 lists and describe all the entrances of the table available online that contains the catalog of 6413 clusters analyzed in this work.

Table 3
Description of the Columns in the Table Available Online

Column	Label	Unit	Description
1	Cluster	...	Cluster name
2	kind	...	[omg] Type of object (o: open cluster, m: moving group, g: globular cluster) from HR23
3	R.A.	[deg]	Mean R.A. of cluster members
4	Decl.	[deg]	Mean decl. of cluster members
5	plx	[mas]	Mean parallax of cluster members
6	e_plx	[mas]	Error on the parallax
7	CMDclass	...	[0/1] 50th percentile of CMD class from HR23
8	quality	...	[0/3] Quality of the predictions (0: gold sample 1: silver sample 2:bronze sample 3:wood sample)
9	logAge_16	[dex]	16th percentile of logAge predictions of the ANNs
10	logAge_50	[dex]	50th percentile of logAge predictions of the ANNs
11	logAge_84	[dex]	84th percentile of logAge predictions of the ANNs
12	FeH_16	[dex]	16th percentile of [Fe/H] predictions of the ANNs
13	FeH_50	[dex]	50th percentile of [Fe/H] predictions of the ANNs
14	FeH_84	[dex]	84th percentile of [Fe/H] predictions of the ANNs
15	Av_16	[mag]	16th percentile of V-band extinction predictions of the ANNs
16	Av_50	[mag]	50th percentile of V-band extinction predictions of the ANNs
17	Av_84	[mag]	84th percentile of V-band extinction predictions of the ANNs
18	dMod_16	[mag]	16th percentile of distance module predictions of the ANNs
19	dMod_50	[mag]	50th percentile of distance module predictions of the ANNs
20	dMod_84	[mag]	84th percentile of distance module predictions of the ANNs
21	X	[kpc]	X-coordinate in galactocentric galactic coordinates (as predicted by the ANNs)
22	Y	[kpc]	Y-coordinate in galactocentric galactic coordinates (as predicted by the ANNs)
23	Z	[kpc]	Z-coordinate in galactocentric galactic coordinates (as predicted by the ANNs)
24	Rgc	[kpc]	Distance from Galactic center (as predicted by the ANNs and assuming Sun Rgc = 8.122 kpc)

ORCID iDs

Lorenzo Cavallo  <https://orcid.org/0000-0001-5831-1889>
 Lorenzo Spina  <https://orcid.org/0000-0002-9760-6249>
 Giovanni Carraro  <https://orcid.org/0000-0002-0155-9434>
 Laura Magrini  <https://orcid.org/0000-0003-4486-6802>
 Eloisa Poggio  <https://orcid.org/0000-0003-3793-8505>
 Tristan Cantat-Gaudin  <https://orcid.org/0000-0001-8726-2588>
 Mario Pasquato  <https://orcid.org/0000-0003-3784-5245>
 Sara Lucatello  <https://orcid.org/0000-0001-8808-0073>
 Sergio Ortolani  <https://orcid.org/0000-0001-7939-5348>
 Jose Schiappacasse-Ulloa  <https://orcid.org/0000-0002-2179-9363>

References

- Abadi, M., Agarwa, A., Barham, P., et al. 2016, arXiv:1603.04467
 Adams, F. C., & Myers, P. C. 2001, *ApJ*, **553**, 744
 Anders, F., Cantat-Gaudin, T., Quadrino-Lodoso, I., et al. 2021, *A&A*, **645**, L2
 Armillotta, L., Krumholz, M. R., & Fujimoto, Y. 2018, *MNRAS*, **481**, 5000
 Babusiaux, C., Fabricius, C., Khanna, S., et al. 2023, *A&A*, **674**, A32
 Bertelli Motta, C., Pasquali, A., Richer, J., et al. 2018, *MNRAS*, **478**, 425
 Blanco-Cuaresma, S., Soubiran, C., Heiter, U., et al. 2015, *A&A*, **577**, A47
 Bland-Hawthorn, J., Krumholz, M. R., & Freeman, K. 2010, *ApJ*, **713**, 166
 Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019, *A&A*, **623**, A108
 Bravi, L., Zari, E., Sacco, G. G., et al. 2018, *A&A*, **615**, A37
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, **427**, 127
 Buder, S., Asplund, M., Duong, L., et al. 2018, *MNRAS*, **478**, 4513
 Campello, R. J. G. B., Moulavi, D., & Sander, J. 2013, in *Advances in Knowledge Discovery and Data Mining*, ed. J. Mining et al. (Berlin: Springer), 160
 Cantat-Gaudin, T. 2022, *Univ*, **8**, 111
 Cantat-Gaudin, T., & Anders, F. 2020, *A&A*, **633**, A99
 Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, *A&A*, **640**, A1
 Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, *A&A*, **618**, A93
 Capitano, L., Lallement, R., Vergely, J. L., Elyajouri, M., & Monreal-Ibero, A. 2017, *A&A*, **606**, A65
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
 Carraro, G., & Costa, E. 2009, *A&A*, **493**, 71
 Carraro, G., Ng, Y. K., & Portinari, L. 1998, *MNRAS*, **296**, 1045
 Carraro, G., Vázquez, R. A., Costa, E., Perren, G., & Moitinho, A. 2010, *ApJ*, **718**, 683
 Carrera, R., Bragaglia, A., Cantat-Gaudin, T., et al. 2019, *A&A*, **623**, A80
 Casali, G., Magrini, L., Tognelli, E., et al. 2019, *A&A*, **629**, A62
 Casamiquela, L., Carrera, R., Jordi, C., et al. 2016, *MNRAS*, **458**, 3150
 Casamiquela, L., Olivares, J., Tarricq, Y., et al. 2022, *A&A*, **664**, A31
 Castro-Ginard, A., Jordi, C., Luri, X., et al. 2018, *A&A*, **618**, A59
 Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, *A&A*, **635**, A45
 Castro-Ginard, A., Jordi, C., Luri, X., Cantat-Gaudin, T., & Balaguer-Núñez, L. 2019, *A&A*, **627**, A35
 Chen, Y., Bressan, A., Girardi, L., et al. 2015, *MNRAS*, **452**, 1068
 Chen, Y., Girardi, L., Bressan, A., et al. 2014, *MNRAS*, **444**, 2525
 Cordoni, G., Milone, A. P., Marino, A. F., et al. 2023, *A&A*, **672**, A29
 De Silva, G. M., Freeman, K. C., Asplund, M., et al. 2007, *AJ*, **133**, 1161
 De Silva, G. M., Sneden, C., Paulson, D. B., et al. 2006, *AJ*, **131**, 455
 Dias, W. S., Alessi, B. S., Moitinho, A., & Lépine, J. R. D. 2002, *A&A*, **389**, 871
 Dias, W. S., & Lépine, J. R. D. 2005, *ApJ*, **629**, 825
 Dias, W. S., Moitinho, H., Moitinho, A., et al. 2021, *MNRAS*, **504**, 356
 Donada, J., Anders, F., Jordi, C., et al. 2023, *A&A*, **675**, A89
 Donor, J., Frinchaoy, P. M., Cunha, K., et al. 2018, *AJ*, **156**, 142
 Donor, J., Frinchaoy, P. M., Cunha, K., et al. 2020, *AJ*, **159**, 199
 Drimme, R., Khanna, S., D'Ongnia, E., et al. 2023, *A&A*, **670**, A10
 Dumont, T., Charbonne, C., Palacios, A., & Borisov, S. 2021, *A&A*, **654**, A46
 Ferreira, F. A., Corradi, W. J. B., Maia, F. F. S., et al. 2020, *MNRAS*, **496**, 2021
 Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, **40**, 487
 Frie, E. D., Janes, K. A., Tavaréz, M., et al. 2002, *AJ*, **124**, 2693
 Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018, *A&A*, **616**, A10
 Gaia Collaboration, Drimme, R., Romero-Gómez, M., et al. 2023a, *A&A*, **674**, A37
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, **595**, A1
 Gaia Collaboration, Recio-Blanco, A., Kordopatis, C., et al. 2023, *A&A*, **674**, A38
 Gaia Collaboration, Vallenari, A., Brown, A., et al. 2023b, *A&A*, **674**, A1
 Garcia-Dias, R., Allende Prieto, C., Sánchez Almeida, J., & Alonso Palicio, P. 2019, *A&A*, **629**, A34
 Gilmore, G., Randich, S., Asplund, M., et al. 2012, *Msngr*, **147**, 25
 Godoy-Rivera, D., Pinsonneault, M. H., & Rebul, L. M. 2021, *ApJS*, **257**, 46
 GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2018, *A&A*, **615**, L15
 Gruner, D., & Barnes, S. A. 2020, *A&A*, **644**, A16
 Hoffman, M. D., & Gelman, A. 2011, arXiv:1111.4246
 Hunt, E. L., & Reffert, S. 2023, *A&A*, **673**, A114
 Janes, K., & Adler, D. 1982, *ApJS*, **49**, 425
 Janes, K. A. 1979, *ApJS*, **39**, 135
 Jílková, L., Carraro, G., Jungwiert, B., & Minchev, I. 2012, *A&A*, **541**, A64
 Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2013, *A&A*, **558**, A53
 Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980
 Kounke, M., Covey, K., & Stassun, K. G. 2020, *AJ*, **160**, 279
 Krumholz, M. R., & McKee, C. F. 2019, *ARA&A*, **57**, 227
 Lada, C. J., & Lada, E. A. 2003, *ARA&A*, **41**, 57
 Liu, L., & Pang, X. 2019, *ApJS*, **245**, 32
 LSST Science Collaboration, Marshall, P., Anguita, T., et al. 2017, arXiv:1708.04058
 Magrini, L., Smiljanic, R., Franciosini, E., et al. 2021a, *A&A*, **655**, A23
 Magrini, L., Vescovi, D., Casali, G., et al. 2021b, *A&A*, **646**, L2
 Magrini, L., Viscasillas Vázquez, C., Spina, L., et al. 2023, *A&A*, **669**, A119
 Maiorca, E., Magrini, L., Busso, M., et al. 2012, *ApJ*, **747**, 53
 Majewski, S. R., Schiavon, R. P., Frinchaoy, P. M., et al. 2017, *AJ*, **154**, 94
 Mateo, M. 1988, *ApJ*, **331**, 261
 Meibom, S., Barnes, S. A., Latham, D. W., et al. 2011, *ApJ*, **733**, L9
 Meingast, S., Alves, J., & Rottensteiner, A. 2021, *A&A*, **645**, A84
 Mitschang, A. W., De Silva, G., Sharma, S., & Zucker, D. B. 2013, *MNRAS*, **428**, 2321
 Moffat, A. F. J., & Vogt, N. 1973, *A&A*, **23**, 317
 Moitinho, A., Vázquez, R. A., Carraro, G., et al. 2006, *MNRAS*, **368**, L77
 Monteiro, H., Dias, W. S., Moitinho, A., et al. 2020, *MNRAS*, **499**, 1874
 Morales, E. F. E., Wyrowski, F., Schuller, F., & Menten, K. M. 2013, *A&A*, **560**, A76
 Myers, N., Donor, J., Spoo, T., et al. 2022, *AJ*, **164**, 85
 Origlia, L., Dalessandro, E., Sanna, N., et al. 2019, *A&A*, **629**, A117
 Pang, X., Li, Y., Yu, Z., et al. 2021, *ApJ*, **912**, 162
 Perren, G. I., Pera, M. S., Navone, H. D., et al. 2023, *MNRAS*, **526**, 4107
 Piskunov, A. E., Just, A., Kharchenko, N. V., et al. 2018, *A&A*, **614**, A22
 Piskunov, A. E., Kharchenko, N. V., Röser, S., Schilbach, E., & Scholz, R. D. 2006, *A&A*, **445**, 545
 Poggio, E., Drimme, R., Cantat-Gaudin, T., et al. 2021, *A&A*, **651**, A104
 Portegies Zwart, S. F., McMillan, S. L. W., Hut, P., & Makino, J. 2001, *MNRAS*, **321**, 199
 Randich, S., Gilmore, G., Magrini, L., et al. 2022, *A&A*, **666**, A121
 Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, *ApJ*, **885**, 131
 Rodrigo, C., Solano, E., & Bayo, A., 2012 *SVO Filter Profile Service Version 1.0, IVOA Working Draft 15 October 2012*
 Rumelhart, D., Hinton, G., & Williams, R. 1986, *Natur*, **323**, 533
 Salpeter, E. E. 1955, *ApJ*, **121**, 161
 Salvatier, J., Wiecki, T. V., & Fonnesbeck, C., 2016 PyMC3: Python probabilistic programming framework, Astrophysics Source Code Library, ascl:1610.016
 Schiappacasse-Ulloa, J., Pasquato, M., & Lucatello, S. 2023, arXiv:2306.15487
 Sim, G., Lee, S. H., Ann, H. B., & Kim, S. 2019, *JKAS*, **52**, 145
 Soubiran, C., Cantat-Gaudin, T., Romero-Gómez, M., et al. 2018, *A&A*, **619**, A155
 Souto, D., Cunha, K., Smith, V. V., et al. 2018, *ApJ*, **857**, 14
 Spina, L., Magrini, L., & Cunha, K. 2022, *Univ*, **8**, 87
 Spina, L., Magrini, L., Sacco, G. G., et al. 2022c, *A&A*, **668**, A16
 Spina, L., Nordlander, T., Casey, A. R., et al. 2020, *ApJ*, **895**, 52
 Spoo, T., Tayar, J., Frinchaoy, P. M., et al. 2022, *AJ*, **163**, 229
 Tang, J., Bressan, A., Rosenfield, P., et al. 2014, *MNRAS*, **445**, 4287
 Tang, S. Y., Pang, X., Yuan, Z., et al. 2019, *ApJ*, **877**, 12
 Tarricq, Y., Soubiran, C., Casamiquela, L., et al. 2021, *A&A*, **647**, A19
 Vallenari, A., & Ortolani, S. 1993, in *European Southern Observatory Conf. and Workshop Proc.*, Vol. 47 (München: European Southern Observatory), 125
 van der Maaten, L., & Hinton, G. 2008, *JMLR*, **9**, 2579
 Vandenberg, D. A. 1983, *ApJS*, **51**, 29
 Viscasillas Vázquez, C., Magrini, L., Spina, L., et al. 2023, arXiv:2309.17153
 Wattenberg, M., Viégas, F., & Johnson, I. 2016, *Distill*, **1**, 1
 Wright, N. J. 2020, *NewAR*, **90**, 101549
 Xu, Y., Hou, L. G., & Wu, Y. W. 2018, *RAA*, **18**, 146
 Yeh, F. C., Carraro, G., Montalto, M., & Seleznev, A. F. 2019, *AJ*, **157**, 115