

Impact of selection criteria on the structural parameters of the Galactic thin and thick discs

S. Alinder¹, T. Bensby¹ and P. J. McMillan²

¹ Lund Observatory, Division of Astrophysics, Department of Physics, Lund University, Box 118, SE-221 00 Lund, Sweden

² School of Physics & Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK.
e-mail: simon.alinder@fysik.lu.se

Received XX November 2025; accepted XX August 202X

ABSTRACT

Context. The dual nature of the Milky Way disc is well-established. It contains a thick and a thin disc that differ in chemical, kinematic, structural, and spatial properties. Simultaneously, there is significant overlap in the distributions of these properties, and especially so at higher metallicities. Accurately distinguishing between these major structural components is crucial for understanding the formation and evolution of the Milky Way. Yet multiple selection methods exist to classify stars as thin or thick disc stars, each with its own advantages and limitations.

Aims. We aim to investigate how different classification methods for categorising stars into the thick and thin disc populations influence the determination of structural properties of the two discs.

Methods. We apply five different selection methods to define samples of stars that are likely to be members of either the thick or the thin disc. Two methods use cuts in the $[\alpha/\text{Fe}]$ - $[\text{Fe}/\text{H}]$ and $[\text{Mg}/\text{Mn}]$ - $[\text{Al}/\text{Fe}]$ planes; one method uses a dynamical separation in J_ϕ - J_z space; one uses an age-based cut; and the last method uses a kinematic likelihood method. For each method, we derive relative density profiles of each disc component as functions of height above the Galactic plane and Galactocentric radius, and fit these to a simple two-exponential disc model. For our analysis we use red giant stars from APOGEE DR17 and stellar ages from astroNN.

Results. Methods based on abundance or age data produce the cleanest separations, while kinematic and dynamical methods suffer higher contamination due to difficulties in separating well-mixed populations. The thin disc scale heights show a clear flaring as they increase with radius, while the thick disc stays approximately constant at around 1 kpc over most radii for all methods. All methods find the thin disc to have a longer scale length than the thick disc, with the difference being greatest for the chemical selection methods. A scale length of the thick disc of 2.0 kpc leads to one of between 2.3 and 3.0 kpc for the thin disc.

Conclusions. The method chosen to distinguish between Galactic components can significantly impact the view of the Galaxy that is produced. Abundance-based approaches offer a clear separation but require spectroscopy, which is not as widely available, while kinematic and dynamical methods work on larger samples but suffer from less clear separations. Every method recovers a longer scale length for the thin disc than the thick disc, unlike the geometric method, used with luminosity profiles in external galaxies.

Key words. Galaxy: disc – Galaxy: structure – Galaxy: kinematics and dynamics

1. Introduction

The two-component model of the Milky Way disc, comprising a thick and a thin disc, has been a cornerstone of Galactic structure studies for decades (see e.g. [Jurić et al. 2008](#); [Chang et al. 2011](#); [Hayden et al. 2017](#); [Kawata et al. 2026](#)). Such a model typically relies on fitting two exponential profiles to observed stellar density distributions, like [Gilmore & Reid \(1983\)](#), who observed stars in the direction of the Galactic south pole and found they needed two exponentials to fit their data, one with a scale height of 300 pc and one with a scale height of 1350 pc. These are now known as the thin and thick discs, respectively.

The thick and the thin discs each have distinct structural, kinematic, and chemical characteristics. The thick disc reaches higher above the Galactic plane and is less radially extended ([Bensby et al. 2011](#); [Bovy et al. 2012b](#)). Its stars are also older, have lower metallicities, and are enhanced in α -elements compared to the thin disc stars ([Fuhrmann 1998](#); [Bensby et al. 2003](#); [Reddy et al. 2003](#); [Katz et al. 2021](#)). Kinematically selected samples of thick-disc stars have a higher velocity dispersion and a greater asymmetric drift than thin-disc stars, which has been used to preferentially select stars from either disc ([Fuhrmann](#)

[1998](#); [Gilmore & Reid 1983](#); [Bensby et al. 2003](#); [Binney & Tremaine 2008](#)). In contrast, the thin disc’s kinematics are colder with a lower velocity dispersion, and a rotational velocity close to the circular speed of the Milky Way ([Casetti-Dinescu et al. 2011](#)).

Analogous dual-disc structures have been identified in external galaxies as breaks in their luminosity profiles when seen edge-on (e.g. [Comerón et al. 2012](#)). In contrast to the Milky Way, the thick discs in these external galaxies always have scale-lengths longer than the thin disc ones. [Minchev et al. \(2017\)](#) argue for a different interpretation of these observations. By using chemodynamical modelling, they claim that the observed thick discs are many smaller mono-age populations that are stacked and flared. They argue that discs forming inside-out produce nested flared populations with the same age that, when summed, mimic the geometry of a thick disc without requiring a separate disc component. In this view, there should be a gradient in age in the *geometrically* defined thick disc as the older populations have flared more, leaving a sequence of ages in the radial direction (see also [Martig et al. 2016](#)). The data available for external galaxies is far more limited, so fitting one or two luminosity profiles to the observed height above mid-plane is often the only vi-

able means of distinguishing disc components beyond the Milky Way.

How the Milky Way came to possess two discs has been discussed for many years (see, for example, the review by [Rix & Bovy 2013](#)). The two major theories for the formation mechanism of the thick disc are in-situ formation or ex-situ formation. In an in-situ scenario, high turbulence naturally leads to large scale heights ([Bournaud et al. 2009](#)) or radial migration causes stars to reach greater Z -heights ([Loebman et al. 2011](#)). In an ex-situ scenario, dynamical heating or gas-rich mergers, such as the Gaia-Enceladus-Sausage event, heat a pre-existing thin- or proto-disc and trigger centrally concentrated starbursts ([Helmi et al. 2018](#); [Belokurov et al. 2018](#)). Current consensus suggests both processes played roles: an early, rapid in-situ build-up supplemented by later merger-driven heating (e.g. [Pinna et al. 2024](#)). Following the formation of the thick disc, the Milky Way entered a more quiescent phase characterised by slower, inside-out growth of the thin disc. Semi-analytic and chemical evolution models, especially two-infall scenarios ([Chiappini et al. 1997](#)), propose a gap in the star formation of ~ 1 Gyr, followed by a phase of gradual accretion of enriched gas and slower star formation that builds an extended, kinematically colder thin disc ([Haywood et al. 2016](#); [Katz et al. 2021](#)). However, there are chemical evolution models that do not include this gap and are still able to produce thin and thick discs (e.g. [Schönrich & Binney 2009](#); [Schönrich & McMillan 2017](#)). In either scenario, accreted gas is mixed with the enriched gas already present, which, together with dynamical processes like radial migration of stars, results in the metallicity gradients and age-radius correlations that are observed today ([Cerqui et al. 2025](#)).

We currently have access to a wealth of high-quality data of stars in the Milky Way thanks to the European Space Agency’s astrometric *Gaia* mission ([Gaia Collaboration et al. 2016a,b, 2018, 2021, 2023](#)), and successful spectroscopic surveys like APOGEE ([Majewski et al. 2017](#); [Abdurro’uf et al. 2022](#)), the *Gaia*-ESO survey ([Gilmore et al. 2022](#); [Randich et al. 2022](#)), GALAH ([Buder et al. 2025](#)), and LAMOST ([Cui et al. 2012](#)). With these surveys, the disc of the Milky Way can be analysed in great depth.

However, categorising any star as belonging to the thick or thin disc is far from straightforward. It is possible to categorise stars based on geometric, kinematic, dynamical, chemical, or chronological criteria, each using different aspects of the stellar populations ([Martig et al. 2016](#)). The choice of selection depends on the scientific question being considered, but the selections are not always equivalent and can lead to differing interpretations. For example, the Galaxy has stars at high Z -altitude in the outer part of the disc. Under a geometric interpretation, these are thick disc stars, and the Galaxy therefore has a thick disc that extends far. Under a chemical interpretation, however, those stars would be classified as thin disc stars, and the Galaxy has a short thick disc and a flared thin disc ([Bensby et al. 2011](#)).

It is important to investigate whether different definitions of these components lead to different selections of groups of stars with different properties or not, because using different methods may yield inconsistent classifications, raising questions about how we interpret the structure of the disc and leading to possible confusion when structures defined in different ways are called the same thing. In this paper, we seek to systematically compare two chemical, a kinematic, a dynamical, and a chronological selection method by using them to fit a simple model of the vertical density profile of the Milky Way, and the density fraction of each disc component in the plane of the disc. The different methods for categorising stars we will examine are:

- Abundance-based selections, which look at stars in chemical spaces and try to find distinct populations. Several different elements can and have been used; [Buckley et al. \(2024\)](#) explore which elemental abundances are most effective for distinguishing components of the disc from each other. α elements like magnesium are frequently used, as there are two populations discernible in an α -metallicity plot ([Fuhrmann 1998](#); [Bensby et al. 2003](#); [Reddy et al. 2003](#)).
- A kinematic selection, in contrast, relies on stellar velocities in three dimensions. Thick disc stars have higher velocity dispersions and distinct orbital characteristics compared to thin disc stars (e.g. [Bensby et al. 2003](#)).
- Dynamical methods go a step further by incorporating both positions and velocities, evaluated within the context of a model of the Galactic gravitational potential. This allows for the computation of quantities such as orbital actions or energy, offering a more physically grounded classification framework ([Robin et al. 2022](#)).
- An age-based selection is based on stellar ages and relies on the assumption that thick disc stars formed earlier in the history of the Galaxy ([Haywood et al. 2013](#); [Bensby et al. 2014](#)). Stellar ages are notoriously difficult to determine with high precision for large samples, making this method challenging to apply in practice, at least until accurate age estimates become available for a larger number of stars.

We also compute the surface density fractions of the two discs and derive possible scale lengths for them. We test methods that are used frequently, that are used infrequently, and that could potentially be used in the future, with the goal of examining how each method influences the inferred properties of the thick and thin disc populations, thereby clarifying to what extent methodological choices affect scientific conclusions about the structure of the Galaxy.

In Sect. 2 we present the data and how the stellar sample was selected, and in Sect. 3 we describe the methods we used to categorise the thin and thick disc samples. The model we use to compute scale parameters of the discs is presented in Sect. 4, followed by the results in Sect. 5. Lastly, the results are discussed in Sect. 6, and in Sect. 7 we provide a summary and conclusion.

2. Data

We use the “Main Red Star Sample” data from APOGEE DR17 ([Abdurro’uf et al. 2022](#)), which was constructed with the intention of having a selection function that is easy to understand. That it only contains red giants also means that it is a high-luminosity population that probes a large part of the Galaxy, allowing us to probe the thin and thick discs well beyond the solar neighbourhood. Furthermore, the APOGEE data is a large dataset with many different elemental abundances, and as such, it is frequently used by the community, which makes our results easier to compare against others. We apply several quality cuts to ensure that our sample of red giant stars is of high quality (see Appendix A for details on how we processed the APOGEE data). It should be noted that the inconsistencies in some abundances in the APOGEE data, notably $[Al/Fe]$, which is important for one of our selection methods, are only present for dwarf stars, and not the red giants ([Jönsson et al. 2020](#)).

To evaluate the five proposed selection methods, we will fit a simple axisymmetric model of the Galactic disc to our data. We restrict our data to only include stars with Galactocentric radial distances (R_{gal}) between 4 kpc and 14 kpc, to ensure the model is valid. The inner edge is restricted by the presence of

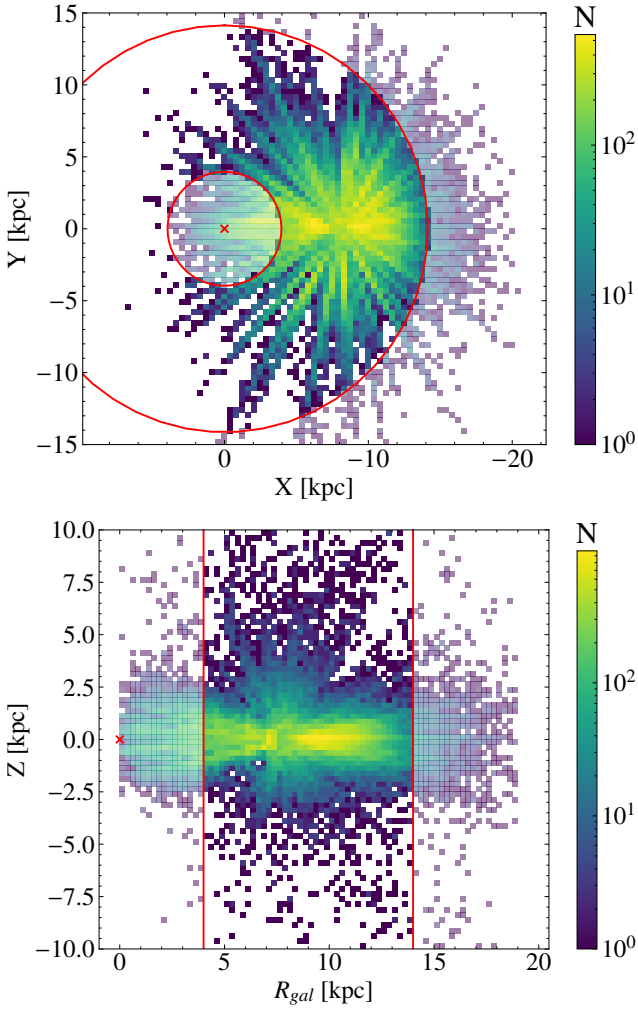


Fig. 1. Selection of stars for this study. The parts of the sample that fall outside our radial limits, shown with red lines at $R_{\text{gal}} = 4$ kpc and $R_{\text{gal}} = 14$ kpc, are shown with transparency. The Galactic centre is marked with a red cross. Top: X-Y view of the distribution of selected stars. Bottom: R_{gal} -Z view of the distribution of selected stars.

the bar, which is not an axisymmetric feature (Wegg et al. 2015), and the outer edge is restricted by the limit where our model breaks down due to low star counts. A top-down X-Y, and R_{gal} -Z visualisation of the distribution of our data in the Galaxy is shown in Fig. 1.

Even though the selected sample is large and contains stars from across the Galaxy, it is not necessarily representative of the true population of the Galaxy. We need to consider the selection function to obtain a representative sample. The selection function expresses the probability that a star is included in a survey given the survey’s magnitude limits and the location of the star. To find the selection function for our sample, we build on the work done in the `gaiaunlimited` package¹ (Cantat-Gaudin et al. 2024) by using a modified version of the `apogee_sf` method that is provided for this purpose. Our modifications allowed us to compute a selection function for 85% of our sample instead of 63% as in the original implementation. See Appendix B for details on this process.

Some of the selection methods we will use require the [Mg/Fe], [Mn/Fe], and [Al/Fe] abundance ratios. The sample is

Table 1. Number of stars selected for each component for each method.

Selection method	Thin disc	Thick disc	Halo
[Mg/Mn] - [Al/Fe]	8.4×10^4	1.9×10^4	1.9×10^3
[Mg/Fe] - [Fe/H]	7.7×10^4	2.7×10^4	1.9×10^3
Dynamical	7.9×10^4	2.6×10^4	1.0×10^3
Age	7.3×10^4	1.7×10^4	1.1×10^3
Kinematic	1.4×10^4	5.7×10^3	1.0×10^3

Notes. The chronological method uses a smaller sample of stars with age data, and the kinematic method is applied to a smaller space and therefore include fewer stars.

therefore limited to stars that have these values. With this requirement, we are left with a sample of 106 163 stars as the basic sample for all methods.

Some of the methods we use require additional data not included in APOGEE DR17. For instance, the dynamical method requires information on the orbits of the stars. Here we compute the actions for each star using AGAMA (Action-based Galaxy Modelling Architecture, Vasiliev 2019) with the best potential from McMillan (2017)². Then the age-based method requires stellar ages, which we get estimates of from Mackereth et al. (2019) via the `astroNN` catalogue for APOGEE DR17. This is a catalogue containing the results from running the `astroNN` deep-learning code (Leung & Bovy 2019) on APOGEE spectra to estimate stellar ages, and other properties like elemental abundances and distances. We select stars with a relative age uncertainty not exceeding 50%, which limits the sample size for this method to 90 387 stars.

3. Selection methods

All methods we use to categorise stars by labelling them as either thin disc, thick disc, or halo. The halo category is intended to contain all non-disc stars in the sample. This often means this category includes halo stars, accreted stars, or peculiar disc stars with properties that separate them from other disc stars. Table 1 contains a summary of the total number of stars categorised by each method as thin disc, thick disc, or halo, respectively.

3.1. Selection using the [Mg/Mn]-[Al/Fe] plane

One of the most promising methods for separating stars into Galactic components works by grouping stars based on position in the [Mg/Mn]-[Al/Fe] plane. This relies on the fact that the interstellar medium is enriched with magnesium primarily via type II core-collapse supernovae, while manganese is an iron-peak element primarily produced by thermonuclear type Ia supernovae. Because core-collapse supernovae are caused by massive stars with short lifetimes, manganese abundance will lag magnesium abundance, allowing [Mg/Mn] to act as an indicator of time of formation (Nomoto et al. 2013). Aluminium enhances the effect by being mostly produced in massive stars, with yields that rise

² The galactocentric coordinate system used in this study places the Sun on the negative X-axis ($\phi = 180^\circ$) at a distance of 8.122 kpc and a height of 20.8 pc, with the Y-axis in the same direction as $l = 90^\circ$, and the Z-axis is in the same direction as $b = 90^\circ$. Galactic azimuth, ϕ , decreases in the direction of Galactic rotation, and the Sun has velocity components $V_{R,\odot} = -12.9$, $V_{\phi,\odot} = -245.6$, $V_{Z,\odot} = -7.78$ km s⁻¹. (Reid & Brunthaler 2004; Drimmel & Poggio 2018; GRAVITY Collaboration et al. 2018; Bennett & Bovy 2019). For the computations and definitions of coordinates, we use `Astropy` v7.0 (Astropy Collaboration et al. 2022).

¹ <https://github.com/gaia-unlimited/gaiaunlimited>

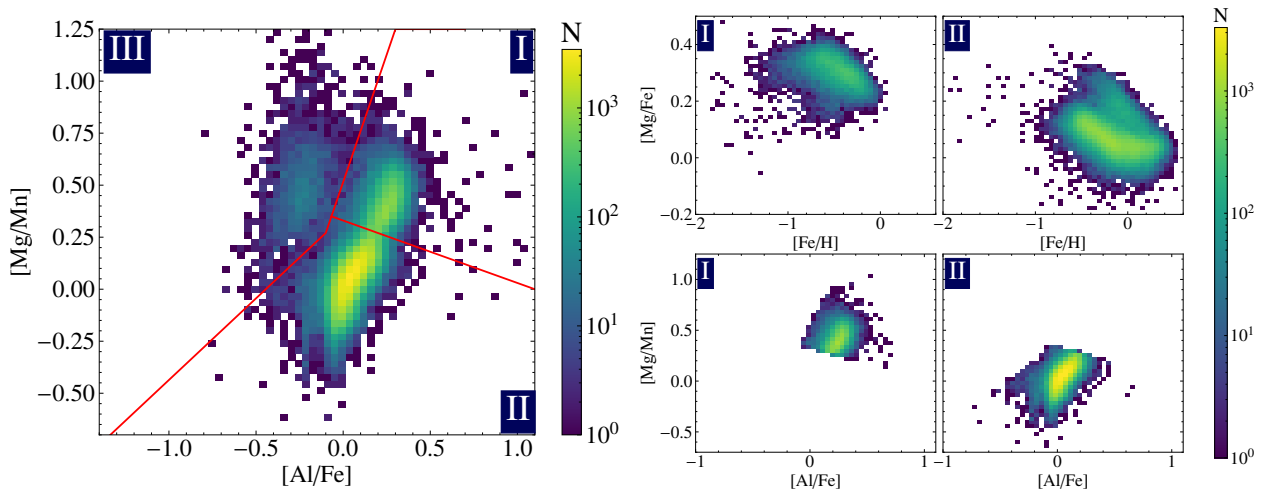


Fig. 2. How the accreted population was split from the thin- and thick discs using the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane. Left: How the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane was divided into different regions, marked with red lines. Region I is the thick disc, region II is the thin disc, and region III is the halo or accreted stars. Right: Distribution of stars from the selection. The panels show $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ in the top row and $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ in the bottom row. Left columns: Region I, the thick disc; Right columns: Region II, the thin disc.

with metallicity and therefore correlate with intense, sustained star formation (Nordlander & Lind 2017, and the references therein). Environments with slow, prolonged chemical evolution, such as accreted dwarf galaxies, remain at lower metallicity for longer and therefore exhibit systematically low $[\text{Al}/\text{Fe}]$ (Nissen & Schuster 2010; Hawkins et al. 2015; Das et al. 2020).

This method builds on the work of Hawkins et al. (2015) and Das et al. (2020), who identified the separation between accreted stars and disc stars in this plane. The low- $[\text{Al}/\text{Fe}]$ part of this plane was found by Das et al. (2020) to be mostly pure accreted stars in the high $[\text{Mg}/\text{Mn}]$ area and Al-poor thin disc stars in the low $[\text{Mg}/\text{Mn}]$ area. We select three regions in this plane to use as the thick disc, thin disc, and halo. The left panel in Fig. 2 shows the lines we draw in the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane to separate the accreted stars (Region III), leaving us with the disc stars in the high- $[\text{Al}/\text{Fe}]$ region (Regions I & II). The line separating the in-situ stars and the accreted stars is bent to include the population of old low-Al thin disc stars found by Das et al. (2020) and Feuillet et al. (2022). The disc population shows two overdensities, which Hawkins et al. (2015) identified with the thick disc (Region I), centred around $([\text{Mg}/\text{Mn}] = 0.5, [\text{Al}/\text{Fe}] = 0.3)$, and the thin disc (Region II), centred around $([\text{Mg}/\text{Mn}] = 0.0, [\text{Al}/\text{Fe}] = 0.0)$.

The results of this selection are shown in the smaller panels in Fig. 2, with the distributions in the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ and $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ planes for the thick and thin disc selections in the left and right columns, respectively. We can see that the thick disc, Region I, forms a short continuous sequence in $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$, and reaches $[\text{Fe}/\text{H}] \approx 0$. There is also a small group of stars below the sequence at $([\text{Mg}/\text{Fe}] \approx 0.18, [\text{Fe}/\text{H}] \approx -0.6)$, which would traditionally be classified as belonging to the low- α thin disc. The thin disc, Region II, here shows two sequences, one at lower $[\text{Mg}/\text{Fe}]$ and a smaller one at higher $[\text{Mg}/\text{Fe}]$, which would traditionally be classified as belonging to the high- α thick disc. The distributions in the X-Y plane, the $R_{\text{gal}}-Z$ plane, the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane, and the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane for each of the three regions, can be found in Appendix D in Fig. D.1.

3.2. Selection using the $[\alpha/\text{Fe}]-[\text{Fe}/\text{H}]$ plane

The most common method of separating the thin and thick discs is by using the $[\alpha/\text{Fe}]-[\text{Fe}/\text{H}]$ plane, because it shows a relatively clear separation between the two populations (Fuhrmann 1998; Bensby et al. 2003). Similarly to the method based on $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$, this one is based on the theory that different types of supernova occur at different time scales during the evolution of the Galaxy and produce different amounts of different elements. The thick disc stars have higher values of $[\alpha/\text{Fe}]$ than thin disc stars because the thick disc stars are relatively more enriched by type II supernovae, which are caused by massive stars with shorter lifetimes and yield more α -elements than iron, than type Ia supernovae, which are caused by white dwarfs in evolved binaries and produce relatively more iron.

We use Mg to represent α since it gives a clear distinction between the two discs when using the APOGEE DR17 data. In this method, a dividing line is manually chosen in the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane³. All the stars below the line are assumed to be thin disc (low- α) stars, and those over the line are assumed to be thick disc (high- α) stars. To remove accreted and halo stars, we use a simple cut-off at $[\text{Fe}/\text{H}] = -1.0$. Stars with metallicities below this limit are assumed to be accreted or halo stars, and those above are assumed to be disc stars.

The results of this selection are shown in Fig. 3 with the distributions in the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ and $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ planes shown for the thick (Region I) and thin (Region II) discs. We can see that the distributions for both Regions are similar to those found in the previous Sect. 3.1. Stars from Region I are densest around $([\text{Mg}/\text{Mn}] \approx 0.4, [\text{Al}/\text{Fe}] \approx 0.15)$ but reach lower values of $[\text{Mg}/\text{Mn}]$, down to approximately -0.1 . Region II has the densest part of the selection at $([\text{Mg}/\text{Mn}] \approx 0.0, [\text{Al}/\text{Fe}] \approx 0.0)$, and also includes a group of low-Al stars previously identified as halo, located around $([\text{Mg}/\text{Mn}] \approx 0.3, [\text{Al}/\text{Fe}] \approx -0.2)$. Further results of this selection, including the distributions in the X-Y plane, the $R_{\text{gal}}-Z$ plane, the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane, and the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane for each of the three regions, can be found in Appendix D in Fig. D.2.

³ We investigate another method for placing the line in this space, based on using a copula in the $[\alpha/\text{Fe}]-[\text{Fe}/\text{H}]$ plane (Patil & Bovy 2024), in Appendix C.

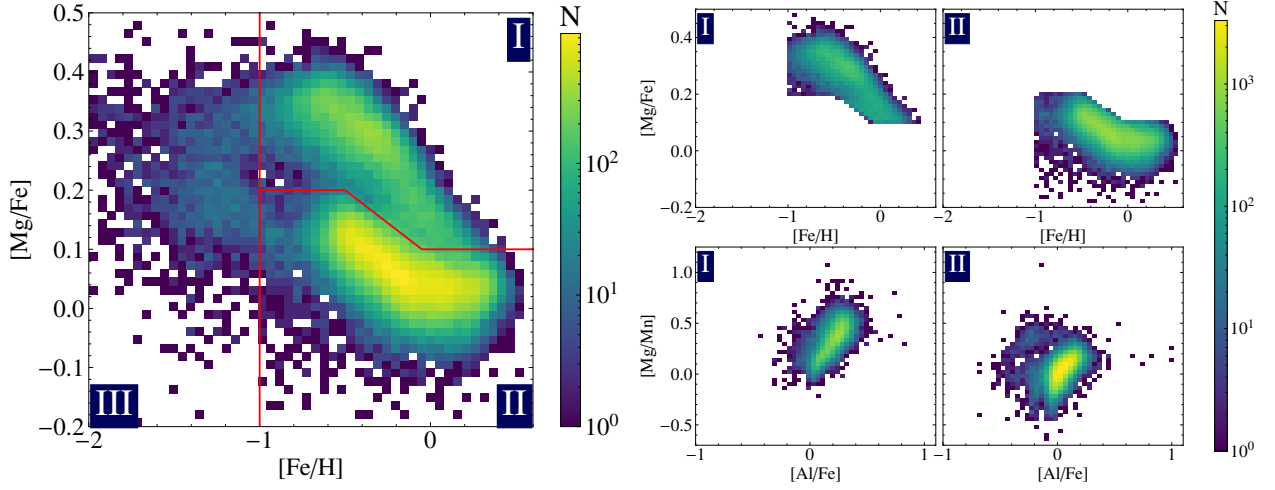


Fig. 3. How the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane was split into the thin disc, thick disc, and halo. Left: Region I is the thick disc, Region II is the thin disc, and Region III is accreted and halo stars, separated by red lines. Right: Distribution of stars from the selection. The panels show $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ in the top row and $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ in the bottom row. Left columns: Region I, the thick disc; Right columns: Region II, the thin disc.

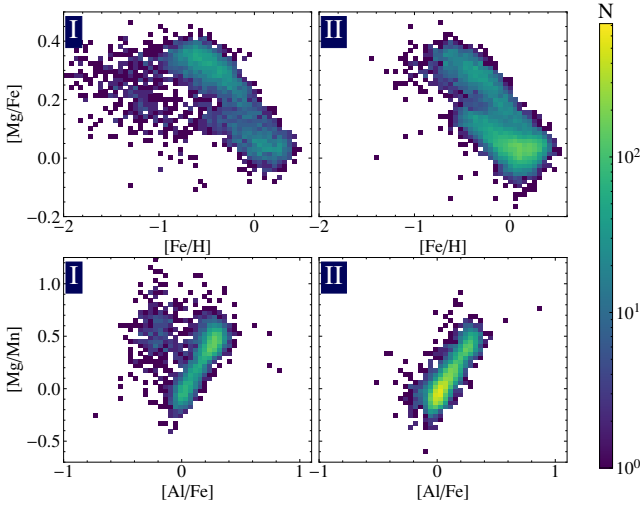


Fig. 4. Distribution of stars from the kinematic selection. Top row: $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$. Bottom row: $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$. Left: Region I, the thick disc; Right: Region II, the thin disc.

3.3. Kinematic selection

The kinematics of stars offer another way of distinguishing between the thin and thick discs. This method uses the fact that thin and thick disc populations in the Milky Way retain the kinematic imprints of their distinct formation histories. Stars in the thick disc tend to move on hotter, more vertically extended orbits and lag behind the local standard of rest in rotational velocity, whereas thin disc stars follow colder, more circular paths. By quantifying these differences statistically, it is possible to assign each star a likelihood of belonging to one component or another, independent of chemical composition.

Following Bensby et al. (2003), we ascribe a relative probability for each star to belong to the thick disc relative to the thin disc, or the thick disc relative to the halo, based on their kinematics. To compute the probabilities, we assume each Galactic component has a Gaussian velocity distribution function, then compute the relative likelihood for each star to belong to these distributions, weighted by the relative abundance of that component. Bensby et al. (2003) explicitly states that the values used

are only valid for the solar neighbourhood at the Galactic mid-plane. This is also how this method is used in practice, for example by Fernández-Alvar et al. (2025), who work within a cylindrical volume 250 pc in radius and 1 kpc in height, centred on the Sun. For this reason, we restrict this model to only be applied to the radial bin that includes the solar neighbourhood with an additional restriction on azimuthal angle, meaning it covers an annular section of $7 \text{ kpc} < R_{\text{gal}} < 9 \text{ kpc}$, covering 60° , centred on the Sun, with no restriction in height. See Bensby et al. (2003) for a detailed description of the method.

The results of the selection are as shown in Fig. 4 with the distributions in the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ and $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ planes for the thick and thin discs. Compared to the chemical selection methods in Sect. 3.1 and 3.2, we see that the thick disc, Region I, mostly contains stars from the high- α sequence but also stars from the low- α sequence, and stars with lower metallicity ($[\text{Fe}/\text{H}] > -1$) or low $[\text{Al}/\text{Fe}]$ stars that were earlier classified as halo. The thin disc, Region II, mostly contains stars from the low- α sequence, but a clear signal of the high- α sequence can also be seen. Further results, including the distributions in the X-Y plane, the $R_{\text{gal}}-Z$ plane, the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane, and the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane for each of the three regions, can be found in Appendix D in Fig. D.3.

3.4. Dynamical selection

With the wealth of information provided by the data from the *Gaia* mission, one can hope to categorise stars based on their dynamical properties, removing the need to rely on abundance data, which is difficult to attain in similar quantities.

Our attempt to categorise stars using dynamical properties is done in two steps. First, we need to find a suitable action plane for categorising stars into Galactic components, and for this, we turn to azimuthal action (J_ϕ) and vertical action (J_z). Orbital actions like J_ϕ and J_z are adiabatic invariants that quantify the oscillatory motions of a star. Thin disc stars are typically on near-circular orbits in the Galactic plane, with J_ϕ values near that of the local standard of rest, whereas thick disc stars tend to have more eccentric orbits and therefore greater asymmetric drift. This means they have a smaller V_ϕ and thus J_ϕ for a given radius (Binney & Tremaine 2008). J_z measures the amplitude of vertical oscillations away from the Galactic plane. It is related

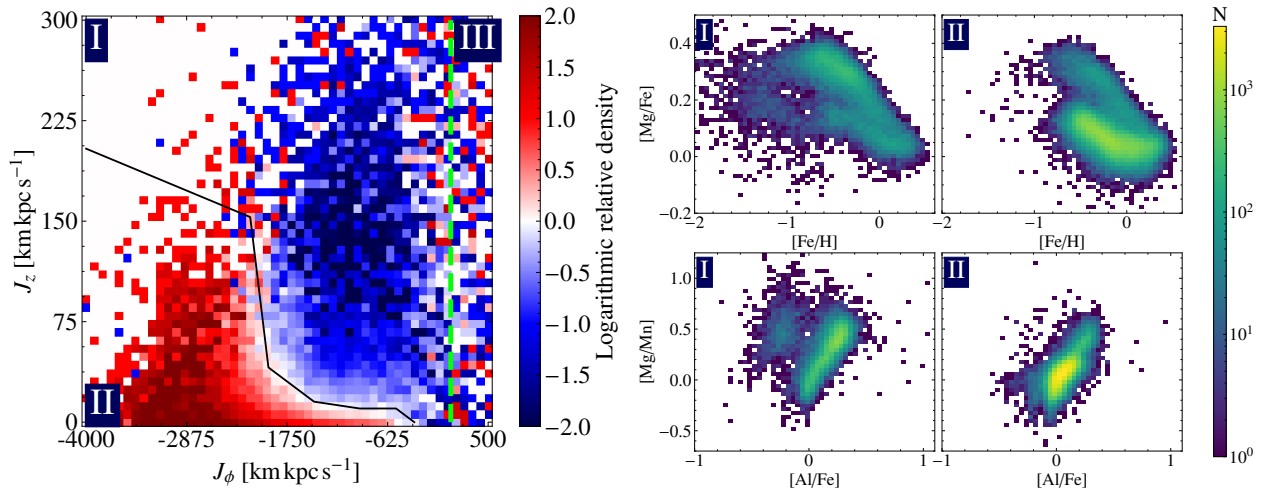


Fig. 5. How the J_ϕ - J_Z plane was split into the thin disc, thick disc, and halo. Left: Relative density of the chemically defined high- and low- α populations in the J_ϕ - J_Z plane. The colour of each pixel is logarithmic relative density as defined by $\log(N_{thin}) - \log(N_{thick})$, where red means thin disc dominated and blue means thick disc dominated. The disc components are separated by a black line following the area where $\log(N_{thin}) - \log(N_{thick}) = 0$. Region I is the thick disc, Region II is the thin disc, and the dashed line at $J_\phi = 0$ delineates Region III, which is the halo and accreted stars. Right: Distribution of stars from the selection. The panels show [Mg/Fe]-[Fe/H] in the top row and [Mg/Mn]-[Al/Fe] in the bottom row. Left columns: Region I, the thick disc; Right columns: Region II, the thin disc.

to the vertical energy and increases with the maximum height reached by a star's orbit. Thin disc stars, which formed in a cold, rotationally supported layer, have small vertical excursions and thus low J_Z . Thick disc stars, which are kinematically hotter, whether due to early turbulent formation, heating by mergers, or radial migration, have larger J_Z due to their higher vertical velocity dispersion. Because J_ϕ and J_Z are integrals of motion in an axisymmetric potential, they remain relatively stable over many orbital periods and are less sensitive to the instantaneous phase of a star's motion than its velocity. The stability of the J_ϕ - J_Z plane makes it an effective diagnostic for separating populations with different kinematic temperatures, in theory. In this plane, the thin disc occupies the high- J_ϕ , low- J_Z area, the thick disc appears at lower J_ϕ and higher J_Z , and halo stars cluster at low J_ϕ with a wide range of J_Z (Binney & Tremaine 2008).

This leaves the second step, we need to find a way to divide this plane into areas corresponding to membership in the different Galactic components. In order to identify the exact locations of the regions for the thin and thick discs, we sorted stars into the high- and low- α components using the abundance-based method described in section 3.2, and plotted their relative densities in the J_ϕ - J_Z plane in the left panel in Fig. 5. The white area with $\log(N_{thin}) - \log(N_{thick}) = 0$ is taken as the boundary between the thin and thick disc regions. This leaves the thin disc assigned to the area which roughly covers $J_\phi \lesssim -2000$ km kpc s⁻¹ along with the area around $J_Z = 0$ km kpc s⁻¹ and the thick disc roughly in the area spanning $J_Z > 10$ km kpc s⁻¹ and $J_\phi \gtrsim -2000$ km kpc s⁻¹. To capture non-disc stars, we select stars with positive J_ϕ , marked by a dashed green line. $J_\phi > 0$ means the stars are on a polar or retrograde orbit and can not reasonably be said to be a member of the disc.

The results of this selection are as shown in the right part of Fig. 5 with the distributions in the [Mg/Fe]-[Fe/H] and [Mg/Mn]-[Al/Fe] planes for the thick and thin disc selections in the left and right columns, respectively. The results are similar to those of the kinematic selection in Sect. 3.3, except with more stars due to the larger selection and a slightly better separation on the discs. Region I contains stars from both the high- α sequence and the high-[Mg/Fe] part of the low- α sequence. It also contains stars with

low [Fe/H] or low [Al/Fe] stars that were earlier classed as accreted. Region II shows two sequences, a strong low- α sequence and a faint high- α sequence that does not extend to [Fe/H] lower than -1 . Further results of this selection, including the distributions in the X-Y plane, the R_{gal} -Z plane, the [Mg/Fe]-[Fe/H] plane, and the [Mg/Mn]-[Al/Fe] plane for each of the three regions, can be found in Appendix D in Fig. D.4.

3.5. Selection using age

Several of the methods presented above are attempts at using proxies for the age of the stars to classify them into different Galactic components. In principle, if the age of the stars can be estimated directly, that should be a very powerful method for categorisation since it is believed that the majority of thick disc stars formed earlier than the thin disc in the history of the Milky Way (Freeman & Bland-Hawthorn 2002). Most stars, unfortunately, do not carry any direct age markers. Instead, age estimates must be inferred from other observable properties, such as chemical abundances, luminosities, or oscillation frequencies. With the growing availability of high-quality data, it has become possible to construct large-scale catalogues containing stellar age estimates. These estimates provide an opportunity to attempt classification based directly on stellar formation time.

We use the astroNN value-added catalogue from the Sloan Digital Sky Survey to get age estimates for as large a fraction of our APOGEE sample as possible, ending up with 90 387 stars after quality cuts, 85% of the size of the APOGEE sample we use. The simplest way to make a selection based on age data is to pick a threshold age and divide the sample into two groups. For example, Thulasidharan et al. (2024) adopted a fixed age boundary in this manner. We select 8 Gyr as the cut-off between thick disc and thin disc stars. 8 Gyr is chosen because Jurić et al. (2008); Bensby et al. (2014) and Martig et al. (2016), among others, indicate that this is approximately between the times when the thick disc and the thin disc formed, so 8 Gyr should capture a majority of the stars in the thick disc. In the left panel of Fig. 6, we have plotted [Mg/Fe] against age as a way of visualising the relationship between age and chemical enrichment. We see a

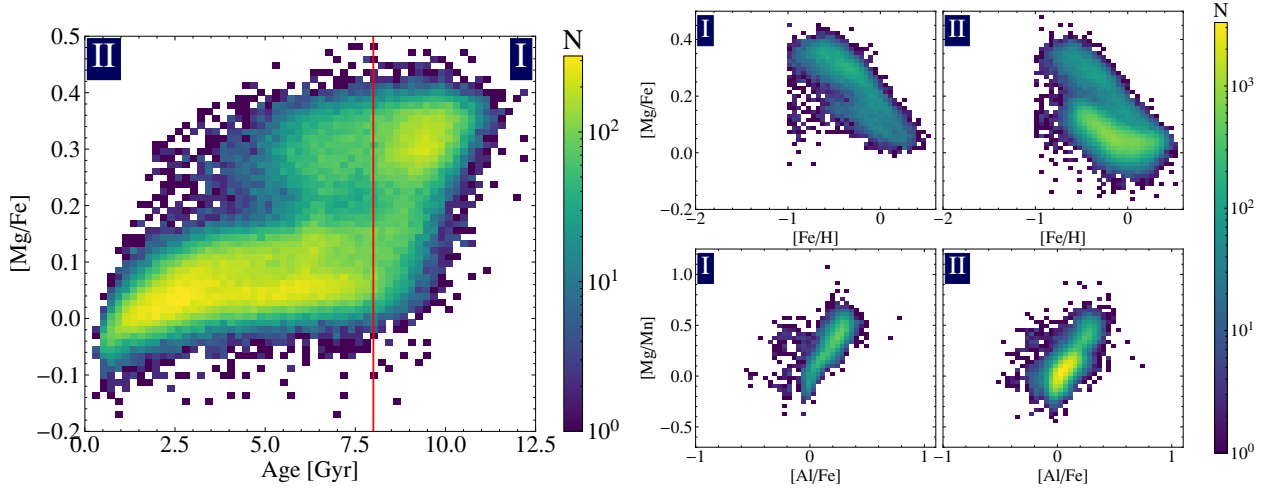


Fig. 6. The age method to select thin and thick disc stars. Left: The number density in the $[\text{Mg}/\text{Fe}]$ - Age plane. Region I is the thick disc, Region II is the thin disc, separated at 8 Gyr (vertical red line). Right: Distribution of stars from the selection by age. The panels show $[\text{Mg}/\text{Fe}]$ - $[\text{Fe}/\text{H}]$ in the top row and $[\text{Mg}/\text{Mn}]$ - $[\text{Al}/\text{Fe}]$ in the bottom row. Left columns: Region I, the thick disc; Right columns: Region II, the thin disc.

sequence of low- $[\text{Mg}/\text{Fe}]$ stars that are young, with ages going from almost 0 to 9 Gyrs and with the densest parts younger than 7 Gyrs, and a sequence of high- $[\text{Mg}/\text{Fe}]$ stars that are mostly older than 4 Gyrs and up to 11 Gyrs old, with the densest part older than 8 Gyrs. When attempting to separate these sequences as well as possible using age, 8 Gyrs presents the best choice.

Extracting the non-disc sample is not as simple as one might expect. In Fig. 6, we see no visible “tail” of old, non-disc stars on the high- $[\text{Mg}/\text{Fe}]$ sequence. More careful studies of the age structure of the Milky Way reveal that, while such a tail should exist, it would not account for all of the accreted stars as they occupy the region between 8 and 12 Gyr (Das et al. 2020), which makes using a cut in age to extract halo and accreted stars is unviable. We therefore use a cut in metallicity to select the accreted population, the same as was used in section 3.2, with $[\text{Fe}/\text{H}] < -1.0$ taken to mean a star is not a disc star. The results of this selection are shown to the right in Fig. 6 with the distributions in the $[\text{Mg}/\text{Fe}]$ - $[\text{Fe}/\text{H}]$ and $[\text{Mg}/\text{Mn}]$ - $[\text{Al}/\text{Fe}]$ planes for the thick and thin discs in the left and right columns, respectively. In the top row, we see the metallicity cut at $[\text{Fe}/\text{H}] = -1$, which removed the accreted stars. Region I includes the high- α sequence and some of the most adjacent parts of the low- α sequence, while Region II clearly shows the low- α sequence and a low-density part of the nearby high- α sequence. Further results of this selection, including the distributions in the X-Y plane, the R_{gal} -Z plane, the $[\text{Mg}/\text{Fe}]$ - $[\text{Fe}/\text{H}]$ plane, and the $[\text{Mg}/\text{Mn}]$ - $[\text{Al}/\text{Fe}]$ plane for each of the three regions, can be found in Appendix D in Fig. D.5.

4. Scale height model

We investigate how stars are distributed between the thin disc, thick disc, and halo components at different R_{gal} . To do this, we divide our data into radial bins and then further into vertical $|Z|$ bins, and distribute the stars into the three components for each bin. The vertical bins are dynamic, meaning their positions and widths are adapted such that each contains as close to the same amount of stars as possible. This leads to densely packed bins close to the plane where stellar number density is high, and only a few further from it. We chose to use a total of 25 vertical bins to provide sufficient resolution to sample the furthest stars. This

increases the importance of the dense region close to the disc compared to the diffuse region far from it when fitting the model.

We seek to estimate the likelihood that the observed distribution of stars in each bin comes from a given model, which predicts a certain proportion of stars in each of the three components. This is a classic statistical problem of categorising a fixed number of objects into a fixed number of categories, according to some underlying probabilities. The appropriate statistical tool for describing this kind of situation is the multinomial distribution, which allows us to calculate the likelihood of seeing the observed numbers of stars in each category, given the expected proportions. That is, the model predicts that a certain fraction of stars should belong to the thin disc, another to the thick disc, and a final fraction to the halo, with all fractions adding up to one. The multinomial distribution tells us the likelihood of observing the actual star counts we see in a bin, given these predicted fractions. This likelihood function forms the basis for evaluating how well different models explain the observed data.

All the categorisation methods are tested by fitting a simple model for the density fraction of each Galactic component as a function of Z at a given R_{gal} . The base of the model is an exponential density distribution as a function of height,

$$N(Z) = \exp\left(-\frac{|Z|}{H}\right) \quad (1)$$

where H is the scale height.

Two instances of this model (N_1 and N_2) are then combined with a constant term for the halo to form the model of the distributions of stars in Z ,

$$N_{\text{total}}(Z) = f_d N_1(Z) + (1 - f_d) N_2(Z) + C_{\text{halo}} \quad (2)$$

where f_d is the thick disc fraction in the plane. These equations do not give the true density of these components, but are proportional to it. We are dealing with relatively small volumes of the Galaxy, and it is therefore acceptable to approximate the halo with a constant term. This also allows us to not need to fit a halo model in addition to the disc models, simplifying the process and removing a potential source of uncertainty.

Our data consists of star counts in the thin and thick discs and in the halo. The data is split into several radial bins, and for each, the stars are categorised using one of the methods described in

Sect. 3. We then seek the model of the Galaxy that best fits this categorisation. Formulated mathematically, this becomes a question of categorising n objects into k ($= 3$) categories with fixed probabilities. The solution to this problem is the multinomial distribution, and we can express the probability for our model as

$$P(n_m, n_{tk}, n_h | p_m, p_{tk}) = \frac{n!}{n_m! n_{tk}! n_h!} p_m^{n_m} p_{tk}^{n_{tk}} p_h^{n_h}, \quad (3)$$

where n is the total number of stars, n_m , n_{tk} , and n_h are the numbers of stars categorised as belonging to the thin disc, thick disc, and halo, respectively, and p_m , p_{tk} , and p_h are the fractions of stars in the thin disc, thick disc, and halo, respectively. To avoid having to compute very large factorials, we recast this equation using Gamma functions,

$$P(n_m, n_{tk}, n_h | p_m, p_{tk}) = \frac{\Gamma(n+1)}{\Gamma(n_m+1)\Gamma(n_{tk}+1)\Gamma(n_h+1)} p_m^{n_m} p_{tk}^{n_{tk}} p_h^{n_h}. \quad (4)$$

We seek to maximise the product of P for all bins. A convenient way of doing this is to take the logarithm of the probability first, then maximise the results, making the function

$$\log P(n_m, n_{tk}, n_h | p_m, p_{tk}) = \sum_i \log P_i(n_m, n_{tk}, n_h | p_m, p_{tk}) \quad (5)$$

where the sum is over the vertical bins. We use the `emcee` (Foreman-Mackey et al. 2013) Markov Chain Monte Carlo (MCMC) package, which applies the Goodman-Weare algorithm (Goodman & Weare 2010), to fit the model parameters using Eq. 5 to find the most probable model for each radial bin.

We use our model of the scale heights as functions of R_{gal} to derive estimates of the scale lengths of the disc components. By integrating Eq. 1, over Z , we find the ratio of the surface densities of the thick and thin disc components, Σ_{tk} and Σ_m , to be

$$\frac{\Sigma_m}{\Sigma_{tk}}(R_{\text{gal}}) = \frac{(1-f_d)H_m(R_{\text{gal}})}{f_d H_{tk}(R_{\text{gal}})}. \quad (6)$$

If we assume both discs are decreasing exponentially with R_{gal} , we can state the ratio of surface densities of those discs as

$$\frac{\Sigma_m}{\Sigma_{tk}}(R_{\text{gal}}) \propto e^{R_{\text{gal}} \left(\frac{R_m - R_{tk}}{R_m R_{tk}} \right)}, \quad (7)$$

where R_m and R_{tk} are the scale lengths of the thin and thick discs, respectively.

We use bootstrapping in two ways. First, to account for the survey selection function. We treat the inverse of the selection function as a probability density function within the bootstrap procedure. This assigns each star a probability of being drawn in the resampled dataset, which is then used for the rest of the procedure. Second, we use bootstrapping to estimate the uncertainties in our model fits. For each categorisation method, we repeat the categorisation multiple times with a newly bootstrapped sample, taking the selection function into account each time. The scatter in the resulting density fractions of each component is taken to be their standard deviation, which we use as the fractional uncertainty for each component in the fitting procedure.

5. Results

5.1. Scale heights

Figure 7 shows how each method fits the thin disc, thick disc, and halo across a range of R_{gal} . The model parameters are also

presented in Table 5.1. This table and figure show the results for selected bins in R_{gal} . The radial bins are 2 kpc wide and spaced 1 kpc apart, meaning they overlap. We note that all methods show the same general trends, that the thin disc becomes more dominant at larger R_{gal} , while the thick disc becomes weaker and is almost completely missing at large R_{gal} .

The [Mg/Mn]-[Al/Fe] method and the [Mg/Fe]-[Fe/H] method show similar results. The [Mg/Mn]-[Al/Fe] method finds a slightly higher fraction of thin disc stars at $R_{\text{gal}} < 9$ kpc, meaning a correspondingly slightly smaller thick disc fraction in the same regions. This is due to this method including some stars that would usually be included in the high- α sequence in the selection for the thin disc. At higher R_{gal} , the results are similar.

The dynamical method gives selections that seem to prefer fitting a higher thick disc fraction. Unlike other methods, this approach, because of how the selection is made, always categorises every star with low vertical action ($J_z \leq 10$ kpc km s⁻¹) as a thin-disc member, except when $J_\phi \approx 0$ kpc km s⁻¹, which leads to a selection where the mid-plane of the disc is found to be purely thin disc for all radial bins except the inner-most one, but the thin disc density falls off faster with $|Z|$ than other methods, particularly interior of $R_{\text{gal}} = 10$ kpc. That this method does not identify genuine thick-disc stars that happen to lie in the mid-plane is a notable weakness.

The age method finds a thin disc fraction that decreases more slowly with $|Z|$ interior of $R_{\text{gal}} = 9$ kpc than other methods, despite the thin disc fraction in the plane being lower, and decreases faster than others exterior of $R_{\text{gal}} = 10$ kpc. We also find a larger fraction of halo or accreted stars at $R_{\text{gal}} > 10$ kpc. The reason this method does not have an identical halo fraction to the [Mg/Fe]-[Fe/H] method, despite sharing the definition for halo stars, is that the samples are not identical. The age method only includes stars with reliable age estimates, which is 85% of the full sample.

The kinematic method is only applied in the $7 \text{ kpc} < R_{\text{gal}} < 9$ kpc bin. It finds a lower fraction of thin disc stars in the plane compared to the other methods, and the vertical profiles of both the thin and thick discs are most similar to the age-based method. It also finds a large fraction of halo or accreted stars, mostly at the expense of the thick disc, meaning the thick disc fraction is smaller compared to the other methods at high $|Z|$. Finally, it finds that the thin disc fraction does not drop to zero by $|Z| = 4$ kpc as for the chemical and dynamical methods, nor stays as large as for the age-based method, but declines almost linearly to just a few per cent at $|Z| = 4$ kpc.

5.2. Disc parameters

Figure 8 shows an overview of how each parameter of our model varies with R_{gal} for each categorisation method. The thick disc scale height has relatively large uncertainties, but the different methods agree on a value of between 1.0 kpc and 1.25 kpc in the middle of the R_{gal} range, around the solar neighbourhood. We see no trends that are consistent for all methods, and the scale heights stay between 0.5 and 2 kpc. The thin disc scale height increases with R_{gal} , showing an exponential increase from between 100 and 400 pc at the inner R_{gal} edge to between 1 kpc and 2 kpc at the outer R_{gal} edge. The age-based method is an exception to this. It shows a thin disc scale height that stays at approximately 400 to 500 pc throughout the R_{gal} range, though it also shows an increase in the outermost bins, similar to the other methods. The fraction of thick disc stars in the plane (the Disc ratio) decreases approximately exponentially with increasing R_{gal} , starting with values between 4% and 14%, and decreasing to less than 2% in the outer disc at $R_{\text{gal}} \geq 12$ kpc.

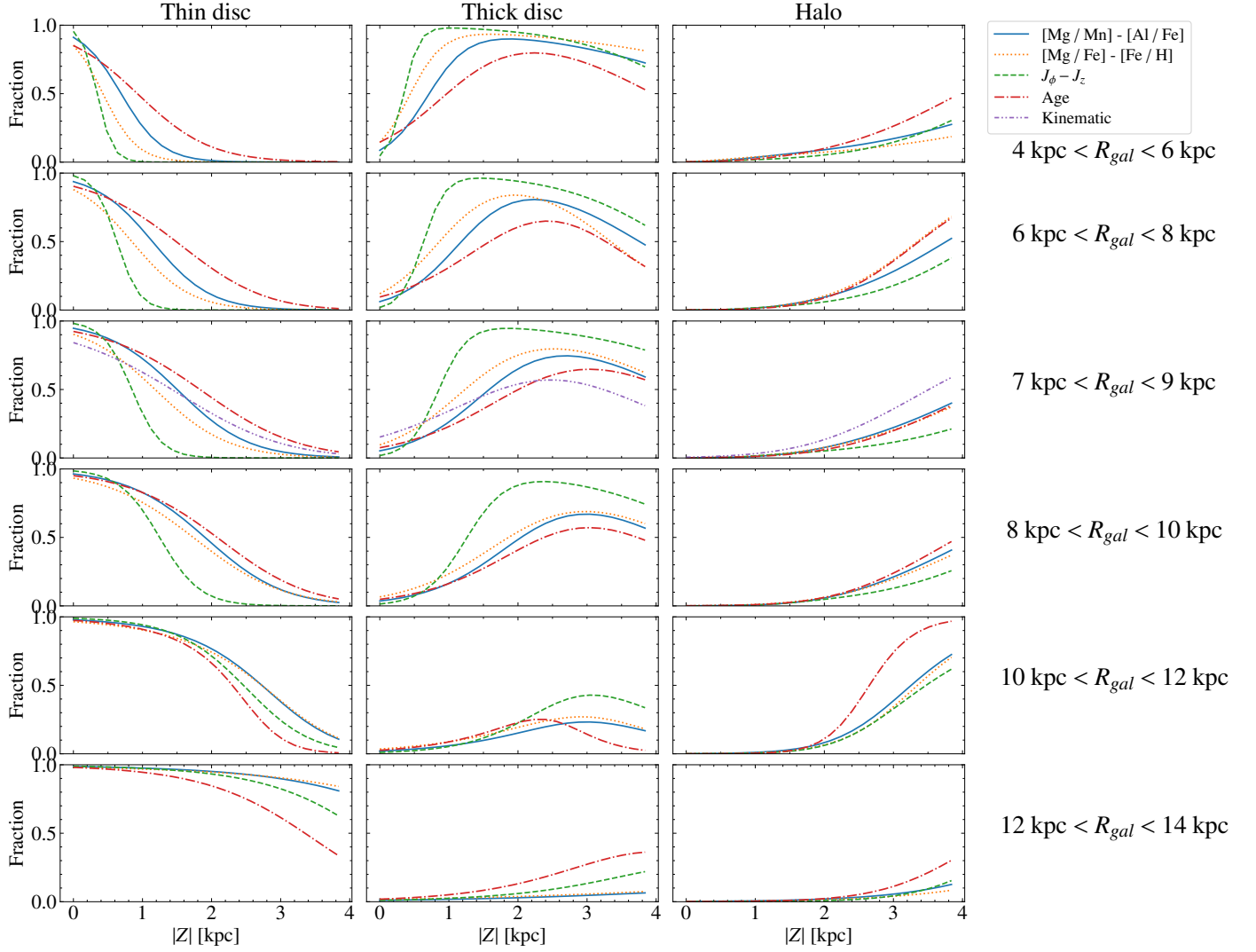


Fig. 7. The relative fraction of the thin disc, thick disc, and halo in selected R_{gal} bins, based on the methods presented in the Sect. 3. From top to bottom, bins of R_{gal} are shown, starting at $4 < \text{kpc } R_{\text{gal}} < 6 \text{ kpc}$ and ending at $12 < \text{kpc } R_{\text{gal}} < 14 \text{ kpc}$. The methods are shown as lines of different colours and styles: the selection using a line in the $[\text{Mg}/\text{Mn}]$ - $[\text{Al}/\text{Fe}]$ plane is a solid blue line, the selection using a line in the $[\text{Mg}/\text{Fe}]$ - $[\text{Fe}/\text{H}]$ plane is a dotted orange line, the Dynamical selection is a dashed green line, the selection using Age is a dash double-dotted red line, and the Kinematic selection is a dash double-dotted purple line that is only present on the $7 < \text{kpc } R_{\text{gal}} < 9 \text{ kpc}$ row.

In Fig. 9, we show the surface density ratios derived for each method in the top panel. We see that all methods produce rising approximately exponential curves, indicating that the scale length of the thin disc is larger than the scale length of the thick disc, as defined by each method. The lines are at different heights and mostly parallel to each other, rarely crossing, meaning that the computed surface density ratios grow at approximately the same pace for all methods. The marker for the kinematic method is slightly below the lines of the other methods, meaning that this method estimates a higher thick disc surface density, which is also seen as a higher disc ratio in Fig. 8. In the bottom panel in Fig. 9, we plot the combinations of thin and thick disc scale lengths in Eq. 7 that produce the best fits to the profiles seen in the top panel. In all cases, the thin disc scale length is larger than the thick disc scale length.

A potentially useful measure is the height at which the disc goes from being dominated by thin disc stars to being dominated by thick disc stars. We show an estimate of this in Fig. 10, where the height at which the two disc components of our model are of equal density is plotted for the different selection methods across

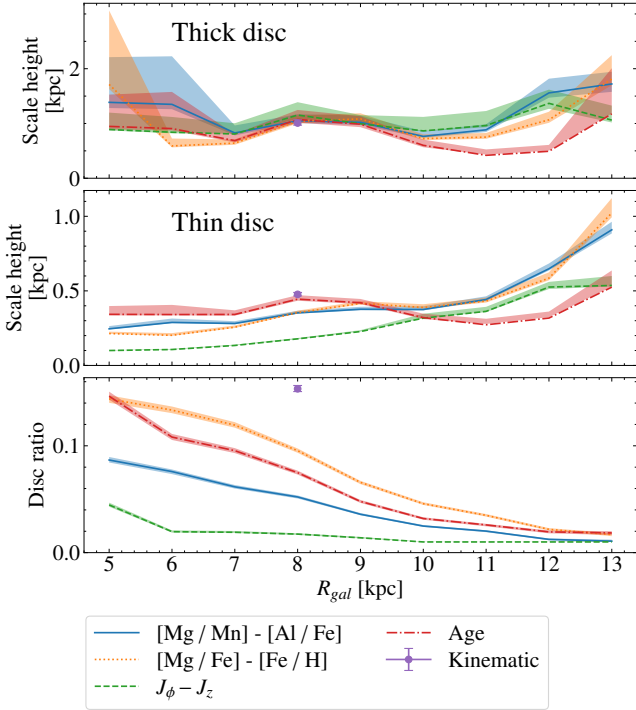
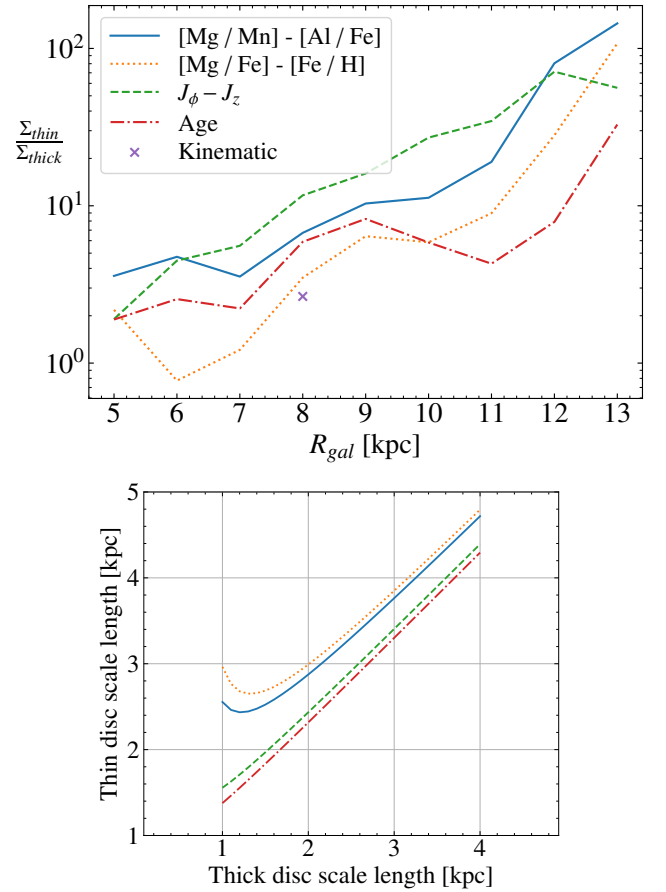
R_{gal} . In the figure, all methods show the same general trend of exponential increase, though with some spread in initial values and slopes. The exception is the age-based method, again. It has the highest initial value and the lowest final value, and shows only a linear increase with R_{gal} .

6. Discussion

Every categorisation method used here is designed to enable comparison between methods, not only to study the properties of the disc components. To achieve this, every star is assigned to a component, and each component is measured as a fraction of the total stellar density. Leaving some stars unclassified would lower the measured fractions and prevent them from summing to 1, meaning we would no longer measure the components as parts of the total population, just the parts of the population that the method can sort into one of the components. This requirement sacrifices some sample purity: a method that allows an “uncertainty zone” would produce cleaner samples but would not classify all stars. Unlike how methods like these would be

Table 2. Model parameters for each method for bins centred on given distance in kpc.

Parameter	Method	$R = 5$ kpc	$R = 7$ kpc	$R = 8$ kpc	$R = 9$ kpc	$R = 11$ kpc	$R = 13$ kpc
Scale height thick disc	[Mg / Mn] - [Al / Fe]	1.4	0.83	1.0	1.0	0.88	1.7
–	[Mg / Fe] - [Fe / H]	1.7	0.64	1.0	1.1	0.75	1.8
–	$J_\phi - J_z$	0.89	0.81	1.2	0.99	0.96	1.1
–	Ages	0.95	0.69	1.1	0.99	0.42	1.2
–	Kinematics	–	–	1.0	–	–	–
Scale height thin disc	[Mg / Mn] - [Al / Fe]	0.25	0.28	0.35	0.38	0.44	0.91
–	[Mg / Fe] - [Fe / H]	0.21	0.26	0.35	0.42	0.43	1.0
–	$J_\phi - J_z$	0.1	0.13	0.18	0.23	0.36	0.54
–	Ages	0.34	0.34	0.44	0.42	0.27	0.52
–	Kinematics	–	–	0.48	–	–	–
Disc ratio	[Mg / Mn] - [Al / Fe]	0.087	0.062	0.052	0.036	0.02	0.011
–	[Mg / Fe] - [Fe / H]	0.14	0.12	0.095	0.066	0.035	0.017
–	$J_\phi - J_z$	0.045	0.019	0.017	0.014	0.01	0.01
–	Ages	0.15	0.095	0.075	0.048	0.026	0.018
–	Kinematics	–	–	0.15	–	–	–

Notes. Units for scale heights are kpc and the Disc Ratio is the density fraction of the thick disc in the plane. The distance R is R_{gal} for the centre of the 2 kpc wide radial bin.**Fig. 8.** Values of the model parameters for different selection methods over R_{gal} . The shaded regions are the 1σ uncertainties of the model fit. The x-axis values are the centres of each radial bin. The bins are 2 kpc wide. Top panel: Scale height of the thick disc. Middle panel: Scale height of the thin disc. Bottom panel: Thick disc fraction in the plane.**Fig. 9.** Top: Surface density ratios over R_{gal} for the different methods. Bottom: Best-fit values of the disc scale-lengths for each method.

used, ours lack such zones, which means we expect our results to be slightly different, except for the abundance-based methods, which typically do not include “uncertainty zones”.

The five methods we tested are not the only possible ways of differentiating the components of the disc from each other. Buckley et al. (2024) explore the possibility of using different elemental abundances for identifying different groups of stars with different origins. They find that [Mg/Fe] offers the greatest separation between the thick and thin discs but is poor at differ-

entiating disc stars and accreted stars, which is in line with our results. They also find evidence for multiple chemically distinct populations in the thick disc while noting the limited utility of [Mg/Fe] for distinguishing between them as well, instead using combinations of Cu, Zn, Al, Ba, Sc, and Y.

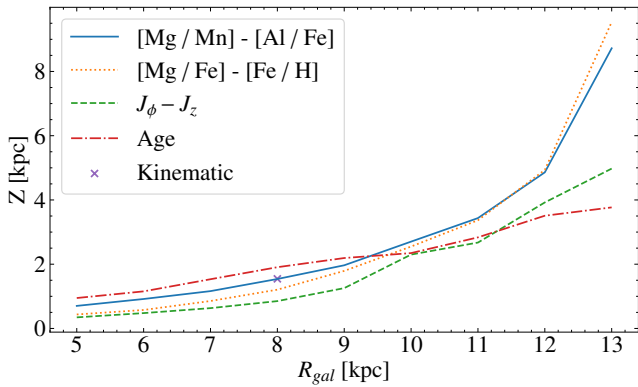


Fig. 10. The Z -height where the models for the thick and thin discs have equal density for different selection methods across R_{gal} .

Of the methods we test, the kinematic method is the most difficult to evaluate due to its being used in only one radial bin. Schönrich & Binney (2009) discuss a comparison between a chemical selection and a kinematic one, finding that the kinematic selection is likely to miscategorise stars because the distribution of kinematic properties of thick disc stars is Gaussian and overlaps with the distribution of kinematic properties of thin disc stars. Bovy et al. (2012a) further claim that the Milky Way possesses a single monotonic distribution in the disc and that imposing a boundary between the “thick and thin discs” is not motivated by the continuous distribution. These results agree with ours, insofar as there is no sharp distinction between the kinematic properties of the populations of the thick and thin discs.

The dynamical method represents a hope that an increased understanding of the stellar contents of the Galaxy, chiefly provided by the *Gaia* mission, will allow a categorisation of stars based on dynamics, without the need for spectroscopy. It is therefore noteworthy that our method produces results that stand out from the others in the inner disc, as seen in the top four rows in Fig. 7. This is the result of difficulties in separating well-mixed populations. These mixed populations are also seen by Binney & Vasiliev (2024) as they explore several different areas in chemodynamical spaces for the purpose of constructing a chemodynamical model of the Milky Way. Here, we consider their results through the lens of finding groups analogous to the thin and thick discs in these spaces. For example, in their Fig. 6, they show several plots of the $[\text{Mg}/\text{Fe}]$ - $[\text{Fe}/\text{H}]$ plane coloured by mean action (J_ϕ , J_z , or J_R), where each plot contains stars in a range of J_z and J_ϕ . Many of the plots show a smooth and gradual variation in the measured action across the sample in chemical space. This would make a complete categorisation based purely on dynamics challenging, which is in line with what we find.

In both the thin disc in Fig. 8 and in Fig. 10, the age-based method produces a line which is approximately linear when the other methods are approximately exponential. This is consistent with an age gradient in the discs, which would make the inner parts have a larger fraction of younger stars compared to the outer disc or the other methods. This is consistent with the age gradients discussed by Martig et al. (2016). The $[\text{Mg}/\text{Fe}]$ -age structure in Fig. 6 shows several interesting properties which are beyond the scope of this paper to investigate, but still worth mentioning. Most obvious is that the high- α sequence is older than the low- α sequence. The low- α sequence also seems to be split in two, with an upper part plateauing at $[\text{Mg}/\text{Fe}] \approx 0.12$ at > 3.5 Gyr, while the lower one plateaus at $[\text{Mg}/\text{Fe}] \approx 0.5$ at the same age. The high- α sequence shows two “clumps” of higher density,

one at approximately 6.5 Gyr and the other at approximately 9.5 Gyr, matching with the red clump and red giants, respectively, and a lower density region with ages $\lesssim 5$ Gyr, which match the young α -rich stragglers noted by Haywood et al. (2013) and further discussed by Cerqui et al. (2023). For this study, we conclude that there is no one vertical line that can be drawn in the plot to separate the high- and low- α sequences with a high degree of purity. It does seem to be possible to get a relatively clean sample of thin disc stars this way, by cutting at 5 Gyr, but that will lead to a significantly incomplete sample that is still contaminated with a small population of stragglers.

Imig et al. (2025) conduct an investigation similar to ours. They use red giants from APOGEE and a method similar to ours to account for the selection function to get a representative sample. They combine measurements in age, metallicity, and α abundance to fit a model of the Galaxy and derive the scale lengths and scale heights for the thin and thick discs, as well as the shape of their profiles. They recover a short scale-length, large scale-height, high- α component and a larger scale-length, shorter scale-height, low- α one. The scale heights they derive increase with R_{gal} , metallicity, and age, and the scale lengths depend on whether the measurement is weighted by mass or light. They find a scale length for the low- α component of either 2.74 or 3.32 kpc, depending on weighing, and a scale length of 1.58 or 1.59 kpc for the high- α component, shown in their Fig. 13. The largest scale heights are in the most metal-rich populations in the inner disc, and the shortest scale heights are in the older, metal-poor populations in the outer disc, as shown in their figures 10 and 11. The scale heights for the low- α component range from about 100 pc to about 1 kpc, and for the high- α component, the range goes from about 400 pc to greater than 1.5 kpc, which the high- α component showing a stronger flaring than the low- α component. This seems to be in tension with our results, which show no noticeable flaring in the thick disc, as well as those of, for example, Bovy et al. (2016) or Mackereth et al. (2017), but is in line with the model proposed by Minchev et al. (2015).

7. Summary and conclusions

Using red giant stars from APOGEE DR17 and astroNN, we have compared five different methods for categorising stars as belonging to the thick or thin disc. These methods categorise stars by:

1. separating two populations in the $[\text{Mg}/\text{Mn}]$ - $[\text{Al}/\text{Fe}]$ plane,
2. separating two populations in the $[\alpha/\text{Fe}]$ - $[\text{Fe}/\text{H}]$ plane,
3. selecting areas in the J_ϕ - J_z plane,
4. making a cut in estimated ages and,
5. making a statistical measure based on velocity dispersions.

Each method produces the relative densities of each component as a function of absolute height above the Galactic mid-plane $|Z|$ at a given Galactocentric radius R_{gal} . By fitting these results to a simple 1D model of the Milky Way, we can compare these methods. All methods broadly recovered the structure of the Galactic disc expected from literature: the thick disc dominates closer to the Galactic centre, while the thin disc becomes more dominant with R_{gal} and shows significant flaring. The fraction of thick disc stars in the mid-plane decreases with Galactocentric radius, reaching a minimum of a few per cent at $R_{\text{gal}} \geq 12$ kpc.

All methods find that the thin disc has a longer scale length than the thick disc. Based on the relationship between the discs, a thick disc scale length of 2.0 kpc, as found by Bensby et al. (2011) and Bovy et al. (2012b), corresponds to a thin disc scale

length of between 2.3 and 3.0 kpc, depending on the method. A longer thick disc scale length of 3.6 kpc, as by [Jurić et al. \(2008\)](#), would give a thin disc scale length between 3.9 and 4.4 kpc.

All methods provide samples with some degree of mixing of the Galactic components. This is because the physical populations in the Galaxy are fairly well mixed, particularly in velocity and action space, making studies like this inherently difficult. The method using the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane is able to cleanly separate the disc populations from the accreted population, but not the discs from each other, while the method using the $[\alpha/\text{Fe}]-[\text{Fe}/\text{H}]$ plane can separate the high- α population from the low- α population well, but is less able to separate out the accreted population. Our dynamical method is faced with well-mixed populations and categorises a large number of stars as belonging to the thick disc. The age-based method encounters the issue of overlapping ages for the oldest parts of the thin disc and the youngest part of the thick disc, which presents a formidable obstacle for selecting a clean sample by age alone, so our attempt categorises more stars than the other methods as thin disc stars, particularly in the inner disc. We only apply the kinematic method in the radial bin that includes the solar neighbourhood because this is the only region where this method is valid, and there it encounters the same mixed populations as the dynamical method.

Our results show that the choice of selection method can meaningfully influence the derived properties of the Milky Way's discs. Abundance-based selections provide higher purity but require spectroscopic data, while kinematic or dynamical methods are more widely applicable but less precise. Some superior selection could be achieved by combining multiple methods, as is currently underway by, for example, [Neitzel et al. \(2025\)](#). The opportunities for more high-quality research will be greatly expanded in the near future when the precise astrometric data from *Gaia* is combined with the spectroscopic data provided by 4MOST ([de Jong et al. 2019](#)), in particular the high- and low-resolution surveys of the Milky Way bulge and disc ([Bensby et al. 2019](#); [Chiappini et al. 2019](#)), and WEAVE ([Jin et al. 2024](#)).

Acknowledgements.

PM gratefully acknowledges support from a project grant from the Swedish Research Council (Vetenskapsrådet, Reg: 2021-04153). TB and SA acknowledge support from project grant No. 2018-04857 from the Swedish Research Council. Some of the computations in this project were completed on computing equipment bought with a grant from The Royal Physiographic Society in Lund. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. This research has made use of NASA's Astrophysics Data System. This work made use of the following software packages for Python, *AstroPy* ([Astropy Collaboration et al. 2022](#)), *emcee* ([Foreman-Mackey et al. \(2013\)](#)), *Numpy* ([Harris et al. 2020](#)), *Matplotlib* ([Hunter 2007](#)), *SciPy* ([Virtanen et al. 2020](#)).

References

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, *ApJS*, 259, 35
 Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, 935, 167
 Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, 478, 611
 Bennett, M. & Bovy, J. 2019, *MNRAS*, 482, 1417
 Bensby, T., Alves-Brito, A., Oey, M. S., Yong, D., & Meléndez, J. 2011, *ApJ*, 735, L46
 Bensby, T., Bergemann, M., Rybizki, J., et al. 2019, *The Messenger*, 175, 35
 Bensby, T., Feltzing, S., & Lundström, I. 2003, *A&A*, 410, 527
 Bensby, T., Feltzing, S., Lundström, I., & Ilyin, I. 2005, *A&A*, 433, 185

Bensby, T., Feltzing, S., & Oey, M. S. 2014, *A&A*, 562, A71
 Binney, J. & Tremaine, S. 2008, *Galactic Dynamics: Second Edition*
 Binney, J. & Vasiliev, E. 2024, *MNRAS*, 527, 1915
 Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, *ApJ*, 707, L1
 Bovy, J., Rix, H.-W., & Hogg, D. W. 2012a, *ApJ*, 751, 131
 Bovy, J., Rix, H.-W., Liu, C., et al. 2012b, *ApJ*, 753, 148
 Bovy, J., Rix, H.-W., Schlafly, E. F., et al. 2016, *ApJ*, 823, 30
 Buckley, N., Das, P., Jofré, P., Yates, R. M., & Hawkins, K. 2024, *MNRAS*, 534, 1985
 Buder, S., Kos, J., Wang, X. E., et al. 2025, *PASA*, 42, e051
 Cantat-Gaudin, T., Fouesneau, M., Rix, H.-W., et al. 2024, *A&A*, 683, A128
 Casetti-Dinescu, D. I., Girard, T. M., Korchagin, V. I., & van Altena, W. F. 2011, *ApJ*, 728, 7
 Cerqui, V., Haywood, M., Di Matteo, P., Katz, D., & Royer, F. 2023, *A&A*, 676, A108
 Cerqui, V., Haywood, M., Snaith, O., Di Matteo, P., & Casamiquela, L. 2025, *A&A*, 699, A277
 Chang, C.-K., Ko, C.-M., & Peng, T.-H. 2011, *ApJ*, 740, 34
 Chiappini, C., Matteucci, F., & Gratton, R. 1997, *ApJ*, 477, 765
 Chiappini, C., Minchev, I., Starenburg, E., et al. 2019, *The Messenger*, 175, 30
 Comerón, S., Elmegreen, B. G., Salo, H., et al. 2012, *ApJ*, 759, 98
 Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, *Research in Astronomy and Astrophysics*, 12, 1197
 Das, P., Hawkins, K., & Jofré, P. 2020, *MNRAS*, 493, 5195
 de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, *The Messenger*, 175, 3
 Drimmel, R. & Poggio, E. 2018, *Research Notes of the American Astronomical Society*, 2, 210
 Fernández-Alvar, E., Ruiz-Lara, T., Gallart, C., et al. 2025, *arXiv e-prints*, arXiv:2503.19536
 Feuillet, D. K., Feltzing, S., Sahlholdt, C., & Bensby, T. 2022, *ApJ*, 934, 21
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
 Freeman, K. & Bland-Hawthorn, J. 2002, *ARA&A*, 40, 487
 Fuhrmann, K. 1998, *A&A*, 338, 161
 Gaia Collaboration, Antoja, T., McMillan, P. J., et al. 2021, *A&A*, 649, A8
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016a, *A&A*, 595, A2
 Gaia Collaboration, David, P., Mignard, F., et al. 2023, *A&A*, 680, A37
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016b, *A&A*, 595, A1
 Gilmore, G., Randich, S., Worley, C. C., et al. 2022, *A&A*, 666, A120
 Gilmore, G. & Reid, N. 1983, *MNRAS*, 202, 1025
 Goodman, J. & Weare, J. 2010, *Communications in Applied Mathematics and Computational Science*, 5, 65
 GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2018, *A&A*, 615, L15
 Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357
 Hawkins, K., Jofré, P., Masseron, T., & Gilmore, G. 2015, *MNRAS*, 453, 758
 Hayden, M. R., Recio-Blanco, A., de Laverny, P., Mikolaitis, S., & Worley, C. C. 2017, *A&A*, 608, L1
 Haywood, M., Di Matteo, P., Lehnert, M. D., Katz, D., & Gómez, A. 2013, *A&A*, 560, A109
 Haywood, M., Lehnert, M. D., Di Matteo, P., et al. 2016, *A&A*, 589, A66
 Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Nature*, 563, 85
 Holtzman, J. A., Hogg, D. W., Shetrone, M., et al. 2018, *AJ*, 156, 125
 Hunter, J. D. 2007, *Computing in Science & Engineering*, 9, 90
 Inig, J., Holtzman, J. A., Zasowski, G., et al. 2025, *ApJ*, 990, 203
 Jin, S., Trager, S. C., Dalton, G. B., et al. 2024, *MNRAS*, 530, 2688
 Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, *AJ*, 160, 120
 Jurić, M., Ivezić, Z., Brooks, A., et al. 2008, *ApJ*, 673, 864
 Katz, D., Gómez, A., Haywood, M., Snaith, O., & Di Matteo, P. 2021, *A&A*, 655, A111
 Kawata, D., Grand, R. J. J., Hunt, J. A. S., & Ciucă, I. 2026, in *Encyclopedia of Astrophysics*, Vol. 4, 38–60
 Leung, H. W. & Bovy, J. 2019, *MNRAS*, 483, 3255
 Loebman, S. R., Roškar, R., Debattista, V. P., et al. 2011, *ApJ*, 737, 8
 Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, *MNRAS*, 489, 176
 Mackereth, J. T., Bovy, J., Schiavon, R. P., et al. 2017, *MNRAS*, 471, 3057
 Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, 154, 94
 Martig, M., Minchev, I., Ness, M., Fouesneau, M., & Rix, H.-W. 2016, *ApJ*, 831, 139
 McMillan, P. J. 2017, *MNRAS*, 465, 76
 Minchev, I., Martig, M., Streich, D., et al. 2015, *ApJ*, 804, L9
 Minchev, I., Steinmetz, M., Chiappini, C., et al. 2017, *ApJ*, 834, 27
 Neitzel, A. W., Campante, T. L., Bossini, D., & Miglio, A. 2025, *A&A*, 695, A243
 Nissen, P. E. & Schuster, W. J. 2010, *A&A*, 511, L10
 Nomoto, K., Kobayashi, C., & Tominaga, N. 2013, *ARA&A*, 51, 457
 Nordlander, T. & Lind, K. 2017, *A&A*, 607, A75
 Patil, A. & Bovy, J. 2024, in *American Astronomical Society Meeting Abstracts*, Vol. 243, American Astronomical Society Meeting Abstracts #243, 129.03
 Pinna, F., Walo-Martín, D., Grand, R. J. J., et al. 2024, *A&A*, 683, A236
 Randich, S., Gilmore, G., Magrini, L., et al. 2022, *A&A*, 666, A121
 Reddy, B. E., Tomkin, J., Lambert, D. L., & Allende Prieto, C. 2003, *MNRAS*, 340, 304
 Reid, M. J. & Brunthaler, A. 2004, *ApJ*, 616, 872
 Rix, H.-W. & Bovy, J. 2013, *A&A Rev.*, 21, 61
 Robin, A. C., Bienaymé, O., Salomon, J. B., et al. 2022, *A&A*, 667, A98
 Schönrich, R. & Binney, J. 2009, *MNRAS*, 399, 1145
 Schönrich, R. & McMillan, P. J. 2017, *MNRAS*, 467, 1154
 Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, 131, 1163
 Thulasidharan, L., D'Onghia, E., Benjamin, R., et al. 2024, *arXiv e-prints*, arXiv:2412.12304
 Vasiliev, E. 2019, *MNRAS*, 482, 1525
 Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Medicine*, 17, 261
 Wegg, C., Gerhard, O., & Portail, M. 2015, *MNRAS*, 450, 4050

Appendix A: Data selection

To get a sample that is representative of the general stellar population to the greatest extent possible while also containing stars with high-quality spectral data, we use the quality flags provided by APOGEE. We use both quality flags for each star and element flags for each element.

We require that the flag `EXTRATARG == 0`, which indicates that the star was randomly selected, see Appendix B. We require that the flag `MEMBERFLAG == 0`, to remove any stars that are members of clusters or dwarf galaxies.

Further, we require that the following APOGEE quality flags are not set,

- BAD_PIXELS
- VERY_BRIGHT_NEIGHBOR
- STAR_BAD
- CHI2_BAD
- VSINI_WARN
- DUPLICATE
- PERSIST_HIGH
- SUSPECT_BROAD_LINES
- SUSPECT_ROTATION
- M_H_BAD
- CHI2_WARN

We make use of APOGEE’s element flags for each abundance considered in our study, $[\text{Fe}/\text{H}]$, $[\text{Mg}/\text{Fe}]$, $[\text{Mn}/\text{Fe}]$, and $[\text{Al}/\text{Fe}]$. We require that the following flags for each element are not set,

- GRIDEDGE_BAD
- CALRANGE_BAD
- OTHER_BAD
- PARAM_MISMATCH_BAD
- TEFF_CUT

We also require the flag `FE_H_FLAG == 0`, which removes entries with poor $[\text{Fe}/\text{H}]$ values.

Finally, we use these additional cuts to select only red giants,

- $\text{teff} < 7000 \text{ K}$
- $\log g < 2.8$
- $\text{SNRev} > 80$

SNRev is the revised S/N estimate, which is used instead of SNR because it avoids persistence issues present in some of the detectors. It is recommended to be used in place of SNR by Holtzman et al. (2018).

Appendix B: Selection function

We use a modified version of the `apogee_sf` method from the `gaiaunlimited` package⁴ to compute the selection function for the stars in our sample. The method for obtaining the APOGEE selection function in `gaiaunlimited` is described in Cantat-Gaudin et al. (2024). This method is based on comparing the number of stars in APOGEE to the number of stars in 2MASS (Skrutskie et al. 2006) for each field and colour-magnitude bin. 2MASS is assumed to be a representative intermediary sample that can be used for this purpose. Using this package, we find that 37% of the stars in our sample were assigned a selection function of exactly zero, meaning that a selection function can not be described and that the star can not be used.

⁴ <https://github.com/gaia-unlimited/gaiaunlimited>

We modify `apogee_sf` to expand the selection and use stars in fields (named areas on the sky) with mixed designs (selections of stars), as well as fields that were not included in the `gaiaunlimited` data by using a more relaxed definition of the Main Red Star Sample compared to Cantat-Gaudin et al. (2024). In their work, a star is part of the Main Red Star Sample if it has the `EXTRATARG == 0`, a `PROGRAMNAME` of either `bulge`, `disk`, `disk1`, `disk2` or `apogee`, and a `FIELD` name that follows the pattern of `LLL±BB` where L and B are Galactic longitudes and latitudes. We relax these conditions by only using the `EXTRATARG == 0` requirement.

The value of the selection function is computed in two parts: the selection fraction and the extinction correction. The selection fraction is computed by dividing the number of stars in APOGEE by the number of stars in 2MASS for each region of the colour-magnitude diagram for each field. The extinction correction is the fraction of 2MASS stars within the magnitude range, ignoring the colour limits, that have extinction values from 2MASS. The final selection function is computed by multiplying the selection fraction by the extinction correction. This updated version of `apogee_sf` returns a selection function of exactly zero for 15% of our sample instead of 37%.

Appendix C: Chemical selection with Copula split

We investigate the method of separating the high- and low- α sequences in the $[\text{Mg}/\text{Fe}]$ direction following Patil & Bovy (2024), who present a method using copulas to select the lowest density regions between the two. A copula is, in this context, a distribution function for which the marginal probability distribution of each variable is uniform on the interval $[0, 1]$. Patil & Bovy (2024) contains the original description of the algorithm implemented here, as well as the mathematical background for this method⁵. In it, the probability distribution functions (PDFs) of the sample along the axes of the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane are transformed into flat distributions, see Fig. C.1. An automated procedure then identifies an area in the transformed plane suitable for splitting the distribution and places a number of points in it. This procedure considers the probability density space as a scalar field. The region separating the upper and lower sequences is identified as an area of high divergence in the gradient of this scalar field, see Fig. C.2. Within this region, the algorithm places points at the extreme left or right points of the probability density contour lines, depending on their orientation, as shown in Fig. C.3.

A weakness of this implementation of the point placing algorithm is that it is not normally able to find points to place at the edges of the space. For this reason, a couple of additional points are placed manually in the area where the algorithm can not identify optimal placements, as a way to guide the fitted line to the edges. These points are placed according to the same criteria as the algorithm uses and are marked in red in Fig. C.4. A third-degree polynomial is fitted to the points. We use a third-degree polynomial because it is the simplest polynomial that adequately captures the shape of the separation between the upper and lower regions in probability density. The polynomial curve is then transformed back to the chemical plane and used to separate the sequences, shown in Fig. C.5, together with the manual line we used in our method described in Sect. 3.2. It shows that

⁵ The original paper Patil & Bovy (2024) only contains the description of an algorithm to select the points of a copula, but no implementation. For this reason, we developed an implementation of the algorithm as a part of this project.

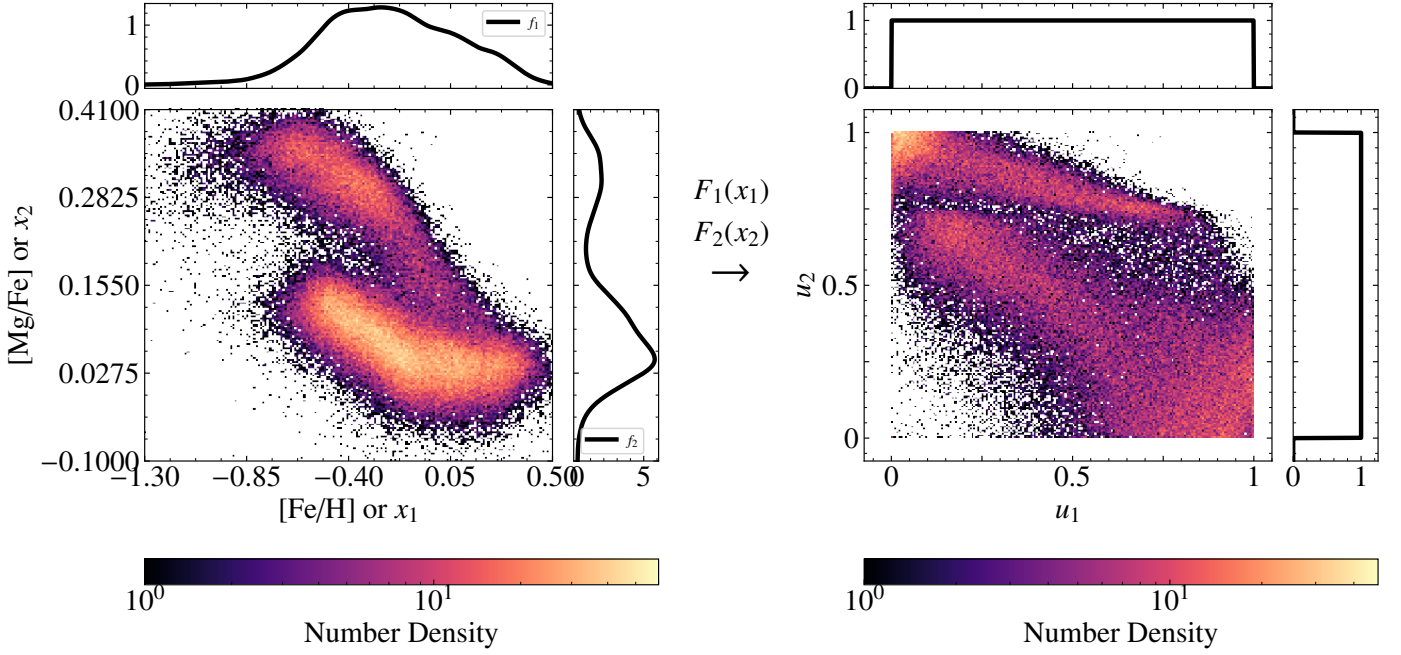


Fig. C.1. Transformation of the sample in the [Mg/Fe]-[Fe/H] plane into copula space with flat probability distributions. Reproduction of Fig. 5 from Patil & Bovy (2024).

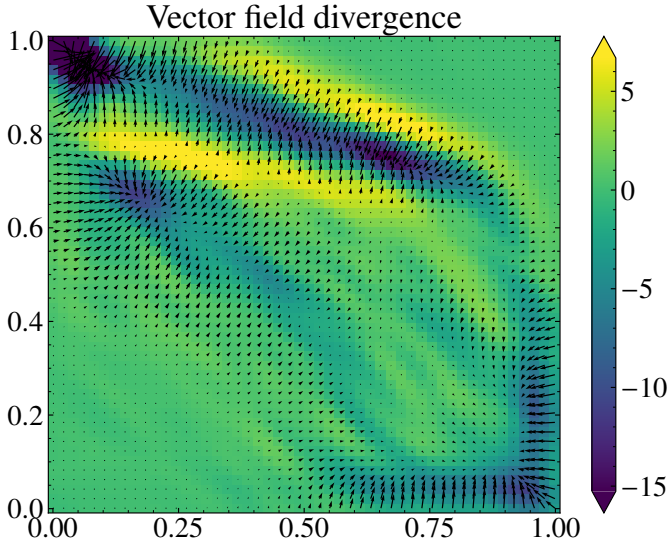


Fig. C.2. Divergence of the density scalar field of the probability distribution of copula space. The vectors show the gradient. The colour shows the divergence.

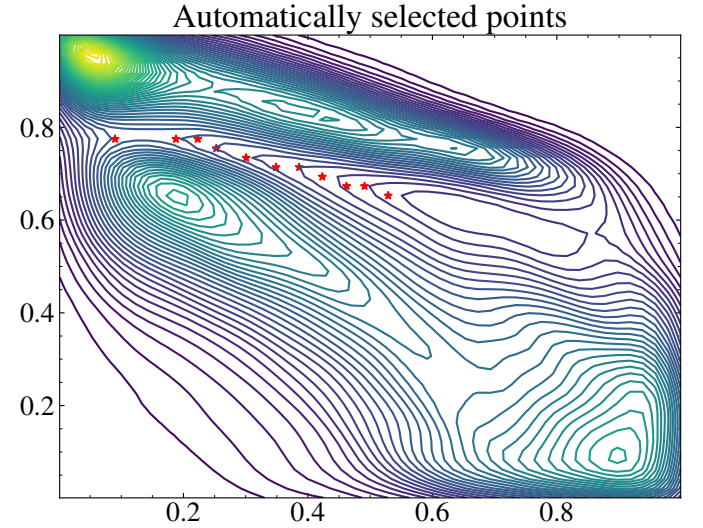


Fig. C.3. Algorithmically placed points on the contour lines of the gradient of the probability distribution.

Appendix D: Method categorisation figures

To give a clearer idea of where our different samples are located in physical and chemical space, we present here figures showing the results of each method presented in Sect. 3 with distributions in X-Y location, R_{gal} -Z location, the [Mg/Fe]-[Fe/H] plane, and the [Mg/Mn]-[Al/Fe] plane for the thin disc, thick disc, and halo categories. The distributions in the halo category (third row of each figure) are shown with a linear colour scale due to low numbers of stars, unlike the thick and thin discs, which use a logarithmic colour scale.

Fig. D.1 shows the results from the chemical selection using the [Mg/Mn]-[Al/Fe] plane in Sect. 3.1, in which we can see stars from Region III occupying the low-metallicity region, with $[\text{Fe}/\text{H}] > -0.5$ except for some scatter. Stars in Region II, the thin

it is possible to closely emulate the selection of a copula using a simple but well-informed line placed manually in the space.

The argument for using this method is that it allows one to make an unambiguous selection in the [Mg/Fe]-[Fe/H] plane, particularly regarding the placement of the line splitting the high- and low- α sequences in the [Mg/Fe] direction. We find this to be a useful application of this technique, but, as pointed out by Patil & Bovy (2024) themselves, the most efficient way of using it is to use pre-computed values rather than recreating the method, which is also how it is being used in practice by Imig et al. (2025).

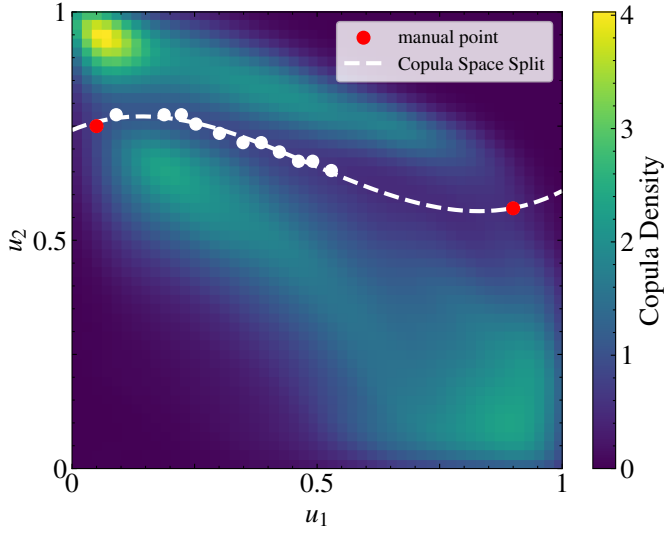


Fig. C.4. Illustration of our automated procedure for selecting the copula. White points are placed by the algorithm. Red points are placed manually. The white dashed line is a third-degree polynomial fitted to all the points.

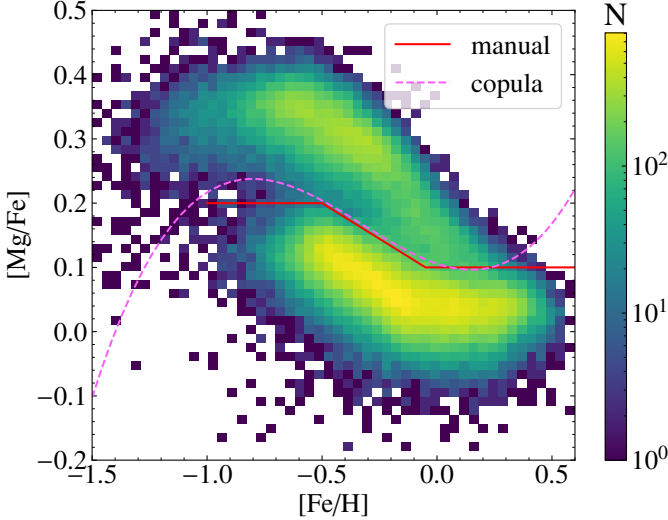


Fig. C.5. $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane with both a copula line and the manually drawn line from Sect. 3.2 shown.

disc, are mostly confined to within approximately 2 kpc of the mid-plane, while stars from Region I, the thick disc, are found approximately 5 kpc of the mid-plane, and stars from Region III are spread further than 10 kpc.

Fig. D.2 shows the results from the chemical selection using the $[\text{Mg}/\text{Fe}]-[\text{Fe}/\text{H}]$ plane in Sect. 3.2. We see stars from Region III concentrated into two regions of the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane. The biggest concentration is centred at approximately $([\text{Mg}/\text{Mn}] = 0.5, [\text{Al}/\text{Fe}] = -0.2)$, and a smaller one is centred at approximately $([\text{Mg}/\text{Mn}] = 0.6, [\text{Al}/\text{Fe}] = 0.2)$. Stars in Region II, the thin disc, are mostly confined to within about kpc of the mid-plane with some scatter, while stars from Region I, the thick disc, are found within about kpc of the mid-plane, and stars from Region III are spread evenly and further than 10 kpc from the mid-plane.

Fig. D.3 shows the results from the kinematic selection in Sect. 3.3. Stars in Region II, the thin disc, are mostly confined

to within approximately 2.5 kpc of the mid-plane, while stars from Region I, the thick disc, are dense out to about 4 kpc but found in lower numbers out to 10 kpc from the mid-plane, and stars from Region III are spread out to 10 kpc but with lower density throughout. Region III contains a small clump of stars from the high- α sequence, centred at approximately $([\text{Mg}/\text{Fe}] = 0.35, [\text{Fe}/\text{H}] = -0.6)$ and $([\text{Mg}/\text{Mn}] = 0.5, [\text{Al}/\text{Fe}] = -0.15)$.

Fig. D.4 shows the results from the dynamical selection in Sect. 3.4. We see that Region I, the thick disc, has a clear underdensity of stars running along the plane of the Galaxy for the length of the disc. Its highest density areas are two small triangular regions about 1 kpc from the place between the inner edge at $R_{\text{gal}} = 4$ kpc and $R_{\text{gal}} = 8$ kpc, and the density noticeably decreases beyond $R_{\text{gal}} = 10$ kpc. We also see that in the $R_{\text{gal}}-Z$ plot, Region II, the thin disc, is mostly confined to a wedge shape less than 1 kpc in height at the inner edge at $R_{\text{gal}} = 4$ kpc and extending to 4 kpc from the mid-plane at the outer edge at $R_{\text{gal}} = 16$ kpc. Region III occupies the regions categorised as accreted $([\text{Fe}/\text{H}] > -1)$ and a small but distinct group of high- α sequence stars located at approximately $([\text{Mg}/\text{Fe}] = 0.35, [\text{Fe}/\text{H}] = -0.6)$ and $([\text{Mg}/\text{Mn}] = 0.5, [\text{Al}/\text{Fe}] = 0.2)$.

Fig. D.5 shows the results from the age-based selection in Sect. 3.5. We see that Region III includes, apart from the accreted stars with low Al, a small part of the thick disc in the $[\text{Mg}/\text{Mn}]-[\text{Al}/\text{Fe}]$ plane with a group of stars centred at approximately $([\text{Mg}/\text{Mn}] = 0.6, [\text{Al}/\text{Fe}] = 0.2)$. We also see that stars in Region II, the thin disc, are mostly confined to within about 2 kpc of the mid-plane, while stars from Region I, the thick disc, are found approximately within 4 kpc of the mid-plane and decrease in density with radial distance, and stars from Region III are spread evenly and are present out to 10 kpc.

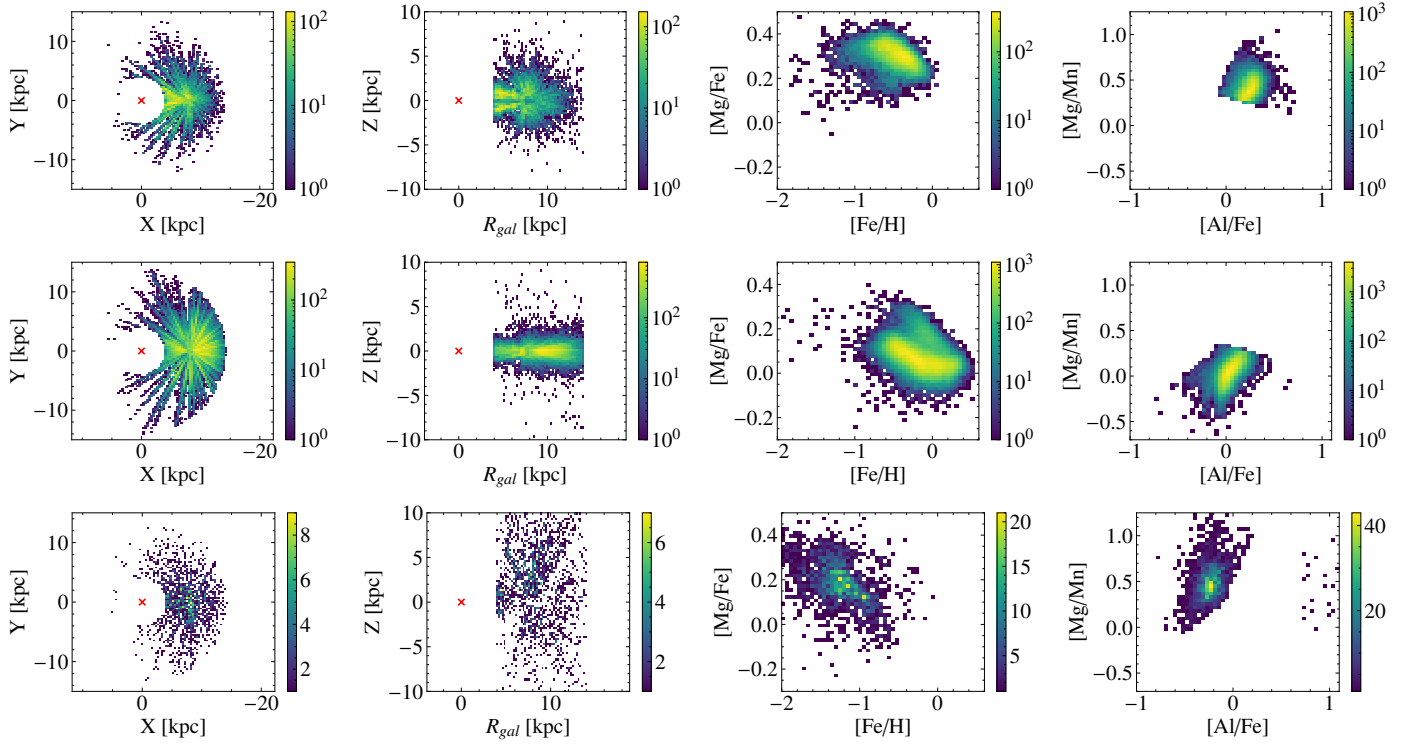


Fig. D.1. Distribution of stars from the chemical selection in Sect. 3.1. From left to right, the panels show X-Y location, R_{gal} -Z location, $[Mg/Fe]$ - $[Fe/H]$, and $[Mg/Mn]$ - $[Al/Fe]$. Top: Region I, the thick disc; Middle: Region II, the thin disc; Bottom: Region III, halo.

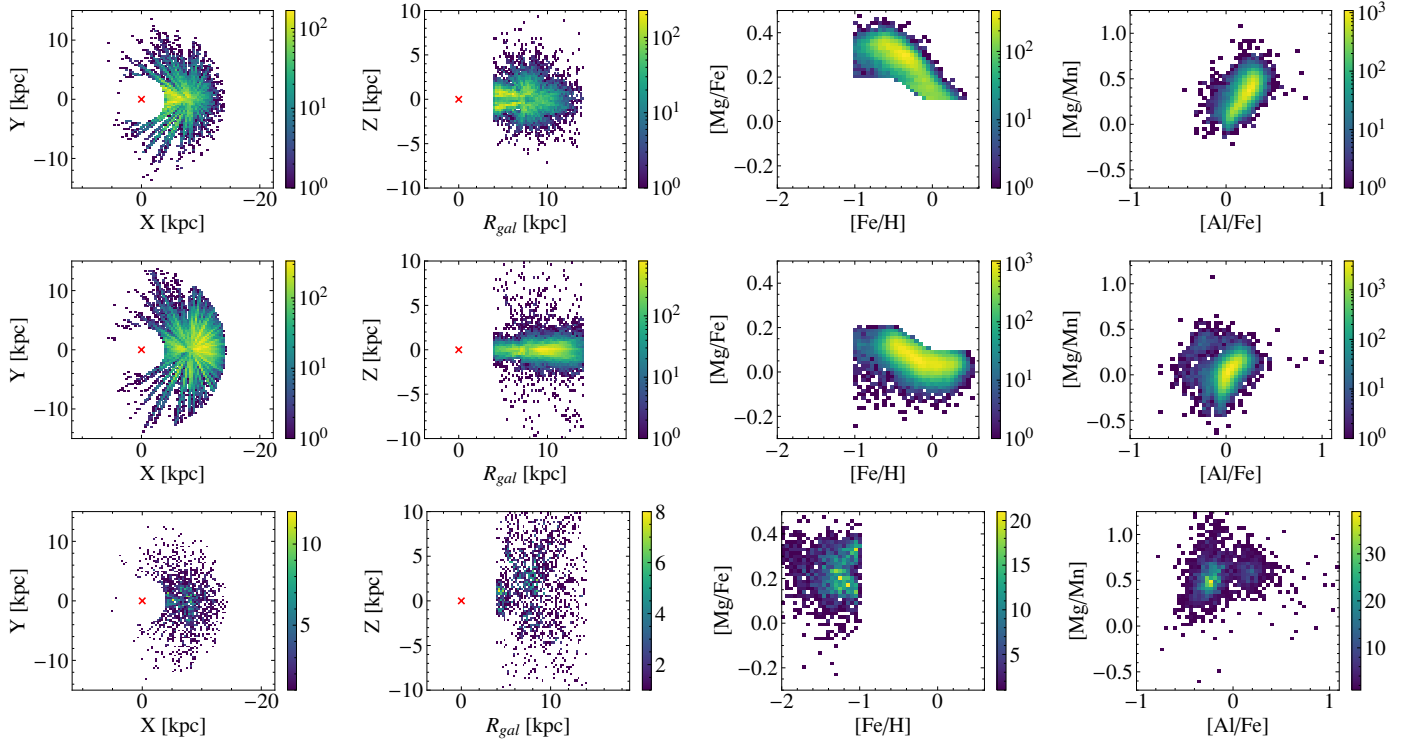


Fig. D.2. Distribution of stars from the chemical selection in Sect. 3.2. From left to right, the panels show X-Y location, R_{gal} -Z location, $[Mg/Fe]$ - $[Fe/H]$, and $[Mg/Mn]$ - $[Al/Fe]$. Top: Region I, the thick disc; Middle: Region II, the thin disc; Bottom: Region III, halo.

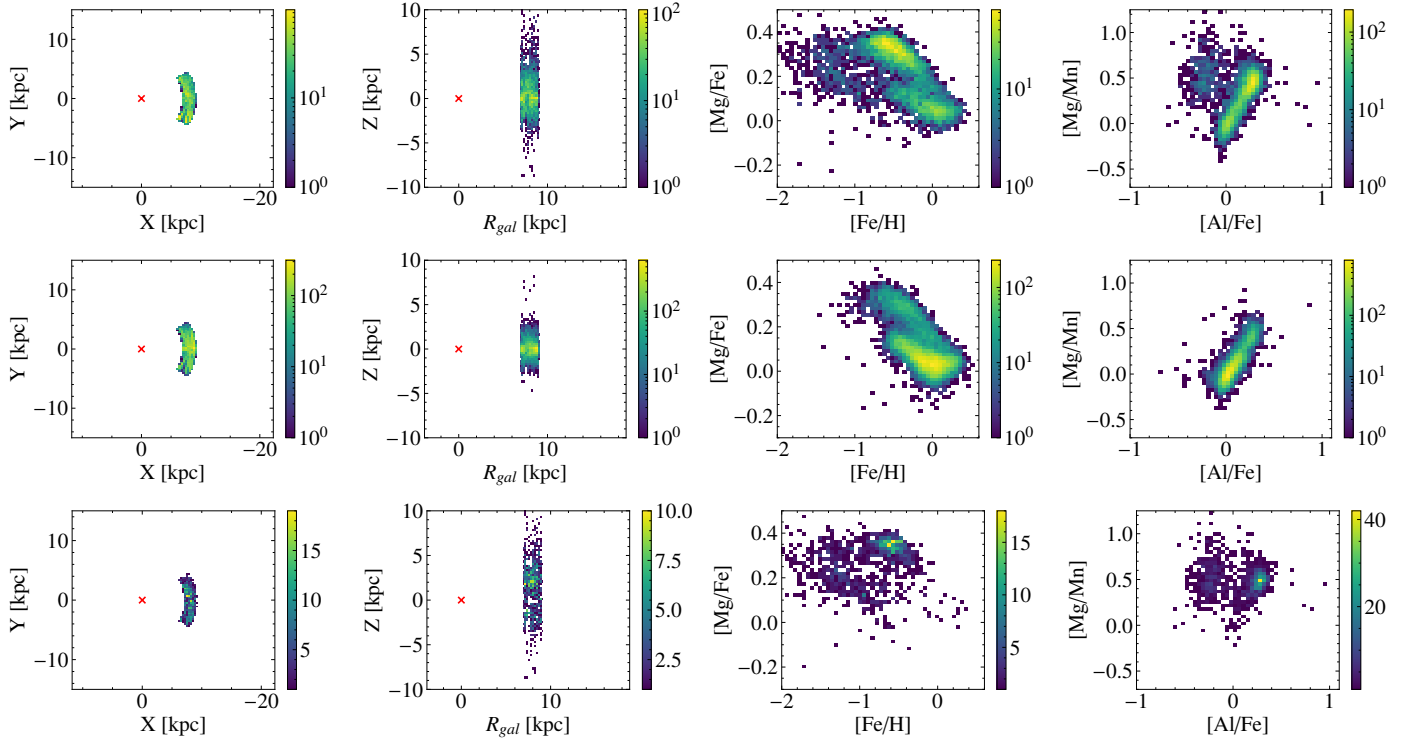


Fig. D.3. Distribution of stars from the kinematic selection in Sect. 3.3. From left to right, the panels show X-Y location, R_{gal} -Z location, $[Mg/Fe]$ - $[Fe/H]$, and $[Mg/Mn]$ - $[Al/Fe]$. Top: Region I, the thick disc; Middle: Region II, the thin disc; Bottom: Region III, halo.

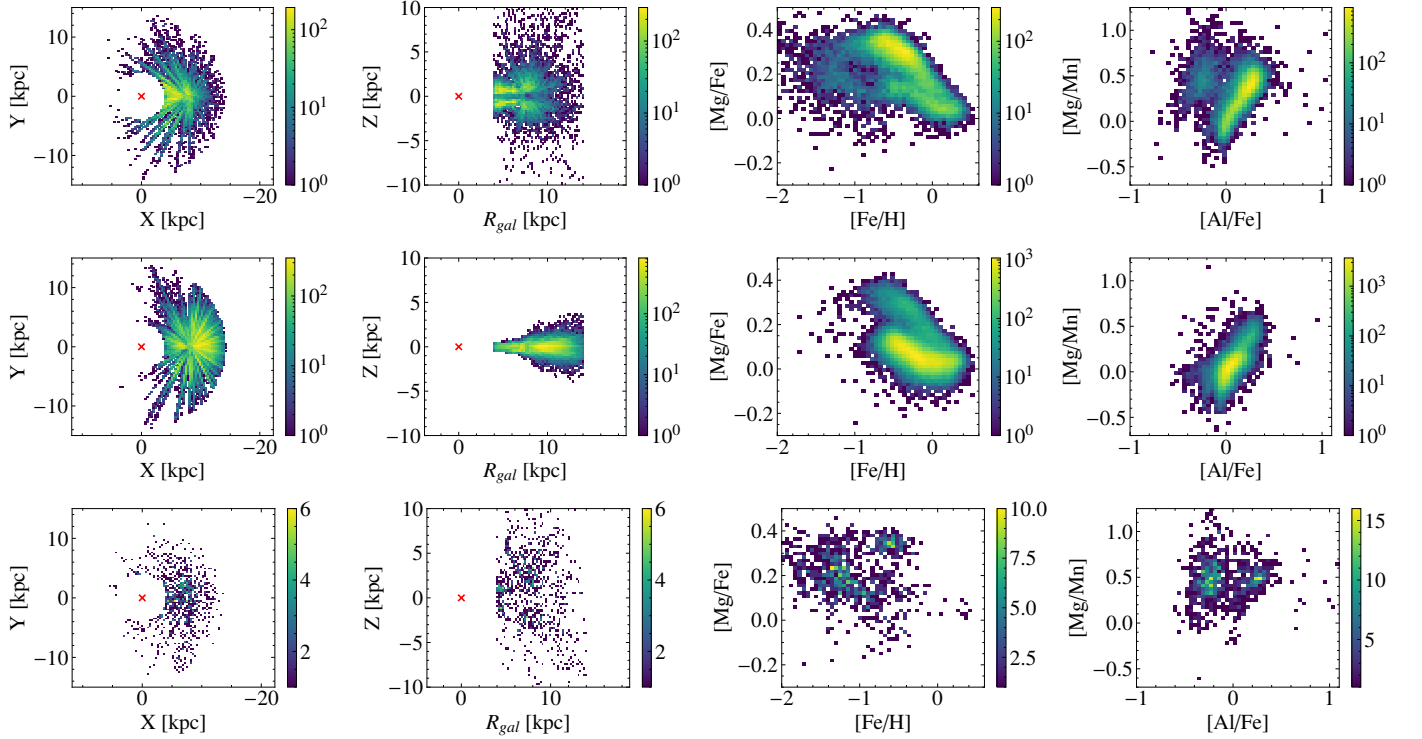


Fig. D.4. Distribution of stars from the dynamics selection in Sect. 3.4. From left to right, the panels show X-Y location, R_{gal} -Z location, $[Mg/Fe]$ - $[Fe/H]$, and $[Mg/Mn]$ - $[Al/Fe]$. Top: Region I, the thick disc; Middle: Region II, the thin disc; Bottom: Region III, halo.

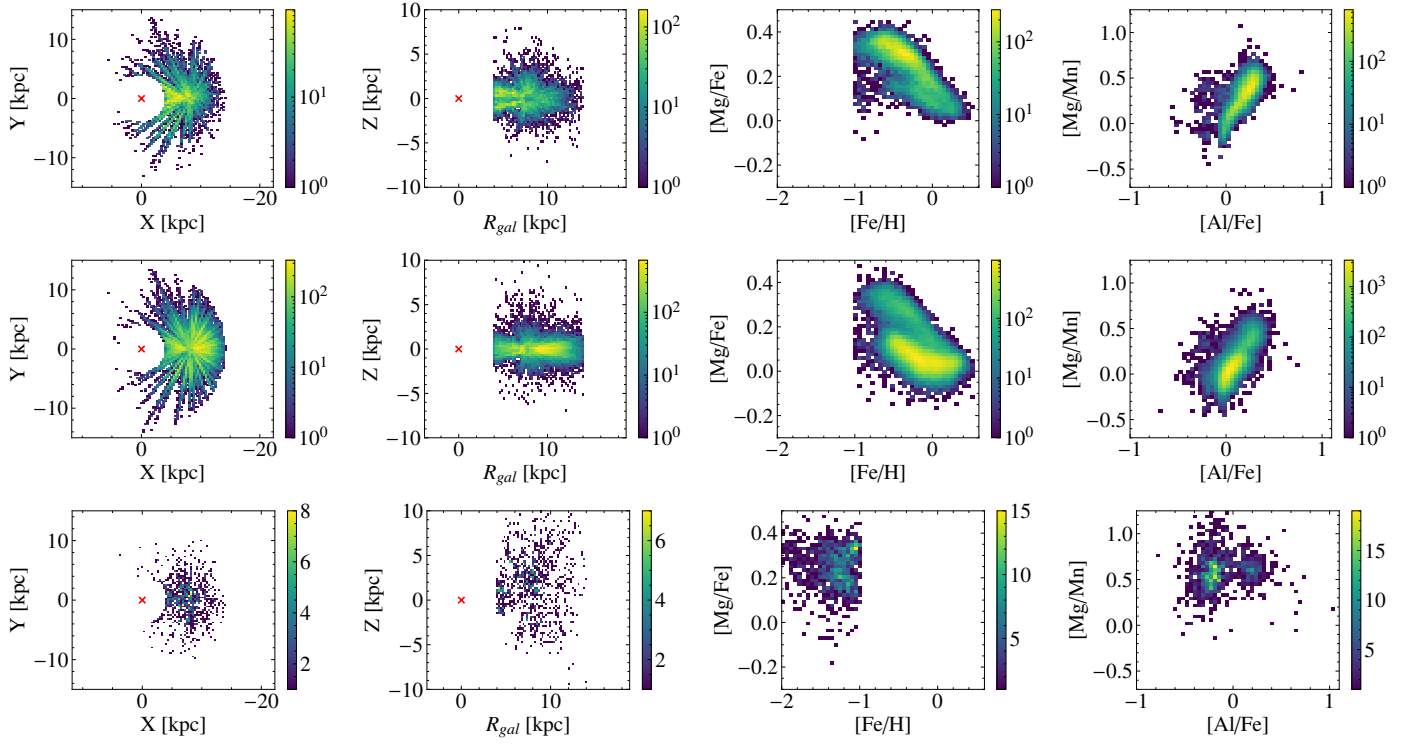


Fig. D.5. Distribution of stars from the age-based selection in Sect. 3.5. From left to right, the panels show X-Y location, R_{gal} -Z location, [Mg/Fe]-[Fe/H], and [Mg/Mn]-[Al/Fe]. Top: Region I, the thick disc; Middle: Region II, the thin disc; Bottom: Region III, halo.