The Equilibrium Model of Disk Galaxy Metallicity Gradients

JAMES W. JOHNSON[1, 2, 3] AND COLLABORATORS

[1]*The Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, CA 91101, USA*
[2]*Department of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH 43210, USA*
[3]*Center for Cosmology & Astroparticle Physics (CCAPP), The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA*

ABSTRACT

We quantify the evolution of the Milky Way disk radial metallicity gradient as traced by APOGEE red giants with age estimates from machine learning algorithms. We measure gradients of $\nabla[\text{O/H}] = -0.062 \pm 0.001$ kpc^{-1} and $\nabla[\text{Fe/H}] = -0.070 \pm 0.003$ kpc^{-1} in our full sample. Splitting into bins of stellar age, we find only minimal evolution in terms of both slope and normalization up to ages of ~9 Gyr. Galactic chemical evolution models reproduce this result when the metallicity gradient tracks a decline in the equilibrium abundance with radius (i.e. $\dot{Z} \approx 0$ because $Z \approx Z_{\text{eq}}$, and $Z_{\text{eq}} \propto e^{-R}$). Wind driven outflows with a mass loading factor $\eta \equiv \dot{\Sigma}_{\text{out}}/\dot{\Sigma}_{\star}$ that increases exponentially with radius are one possible origin of this equilibrium gradient, which our measurements indicate was reached early in the disk lifetime. If extended to other spirals, the equilibrium model offers a qualitative explanation to the surprisingly high abundances observed in high redshift galaxies with JWST. We find evidence of the Sgr dSph's first infall, which presents as a perturbation from the gradient's equilibrium state caused by an accretion induced starburst ~6 Gyr ago. We postulate that the relationship between the radial metallicity gradient in the disk and its inside-out formation is a correlation *without* causation. In our models, its origin lies with a decline in the amount of star formation per unit accretion, which tracks the increase in η with radius.

Keywords: methods: numerical — galaxies: abundances — galaxies: evolution — galaxies: star formation — galaxies: stellar content

1. INTRODUCTION

Radial metallicity gradients are ubiquitous in spiral galaxies. First quantified through nebular emission lines in Galactic HII regions (Aller 1942; Searle 1971; Shaver et al. 1983), the metal mass fraction Z declines roughly exponentially with radius (e.g. Wyse & Silk 1989; Zaritsky 1992). Due to the successes of spectroscopic surveys, measurements are now available for thousands of external galaxies out to increasingly high redshift (e.g. Maiolino & Mannucci 2019; Sánchez 2020), placing these gradients among the most fundamental observable properties of spiral galaxies. The presence and ubiquity of metallicity gradients have generally been interpreted as a consequence of "inside-out" growth, in which the inner regions of spiral galaxies assemble first with the outskirts following suit on longer timescales (e.g. Matteucci & Francois 1989; White & Frenk 1991; Kauffmann 1996; Bird et al. 2013).

Within the Milky Way (MW), stellar population metallicities tend to decline with slope of order ~−0.1 dex/kpc toward large radii (e.g. Cheng et al. 2012; Frinchaboy et al. 2013; Myers et al. 2022). Though the shape is the same, Boardman

et al.'s (2022?) population study of spiral galaxy metallicity gradients with MaNGA[1] (Bundy et al. 2015) indicates that this slope is interestingly steep compared to the background population. The deviation is not to the extent that the MW is an abnormal disk, at least not to the level of precision important for this paper. Nonetheless, its position as an outlier could be an indication that the MW is interestingly unique in at least one regard.

With tracers of underlying stellar populations at different ages, we can constrain the evolution in the disk metallicity gradient empirically. Such direct measurements are available only for the MW since large sample of external MW mass galaxies are only accessible with integrated light. Despite the sample sizes, the detailed evolution in the disk metallicity gradient is only weakly constrained due to the difficulties associated with stellar age inferences (e.g. Soderblom 2010; Chaplin & Miglio 2013). Most investigations thus far have compared samples of various spectral types that trace different ages of underlying populations, such as OB stars (e.g. Daflon

---

[1] MaNGA: Mapping Nearby Galaxies at Apache Point Observatory.

& Cunha 2004), Cepheid variables (e.g. Andrievsky et al. 2004; Luck et al. 2006; Yong et al. 2006), open clusters (e.g. Friel 1995; Chen et al. 2003; Magrini et al. 2009a), and planetary nebulae (e.g. Maciel et al. 2003; Henry et al. 2010; Stanghellini & Haywood 2010; citations mined from intro of Magrini et al. 2016).

Advancements in precision stellar age estimates have recently improved the landscape by allowing measurements with large samples. Willett et al. (2023) measured the evolution in the disk metallicity gradient using asteroseismic age estimates. Lu et al. (2022) used a sample of LAMOST subgiants with precise ages inferences from the main sequence turnoff. In this paper, we leverage the newfound power of age measurements from machine learning algorithms. Imig et al. (2023) recently made these measurements using APOGEE red giants in the DistMass age catalog (Stone-Martinez et al. 2023). We also use APOGEE red giants in this paper and extend the measurements to two additional age catalogs constructed with neural networks.

In good qualitative agreement with Imig et al.'s (2023) findings, we find that both the slope and normalization of the metallicity gradient evolve only minimally out to surprisingly high ages (~9 Gyr). We highlight the age independence of the *normalization* of the metallicity gradient in particuar. This result is contrary to expectations from "classic" GCE models (see e.g. the reviews by Tinsley 1980 and Matteucci 2021), which might postulate that ongoing star formation drives ongoing metal production. Stellar metallicities would tend to decline toward high ages in this scenario. Our empirical results, which have now been seen with three different machine learning age catalogs, dispute this expectation.

To explain the origin of this result, we dedicate this paper to presenting the *equilibrium model* (EM) of metallicity gradients in disk galaxies. We invoke the EM first and foremost as an interpretation of this empirical result with Galactic chemical evolution (GCE) models. Our discussion here casts the EM as a class of models leading to similar predictions rather than one specific parameterization. We qualitatively define three criteria that distinguish EM predictions from other GCE models in section 3.1. Their defining feature is that ISM abundances are predicted to decline with radius but variations with time are small after the first ~few Gyr of disk formation.

Discussion of gradients in the literature tends to focus on the slope of the metallicity-radius relation, but the EM is first and foremost an argument about its normalization rooted in its namesake: equilibrium abundances. The notion of chemical equilibrium dates back to Larson (1972), who showed that ISM abundances evolve toward a value Z_{eq} where metal production by stars is balanced by the H gained through accretion. Weinberg et al. (2017) demonstrated that Z_{eq} falls when feedback driven outflows are included. For a given choice of star formation history (SFH), models with stronger

outflows experience more vigorous accretion in order to fulfill the mass budget. The net effect is that high mtallicity material is removed from the disk and replaced with lower metallicity material due.

This connection between outflows and equilibrium abundances has been central to discussion of the mass-metallicity relation (MZR) in galaxies over the past several years. Higher mass galaxies tend to be higher metallicity in terms of both gas (e.g. Tremonti et al. 2004; Andrews & Martini 2013; Zahid et al. 2012) and stellar populations (e.g. Gallazzi et al. 2005; Kirby et al. 2013; Simon 2019). This trend has been attributed to low mass galaxies ejecting gas in a wind more readily than high mass galaxies due to their weaker gravitational fields (Finlator & Davé 2008; Peeples & Shankar 2011; Lilly et al. 2013; Chartab et al. 2023; Sanders et al. 2021). Under this interpretation, the mass loading factor $\eta \equiv \dot{M}_{out}/\dot{M}_\star$ relating the rate of mass ejection to the SFR should scale with stellar mass approximately as (e.g. Muratov et al. 2015).

Also motivated by chemical equilibrium arguments, our GCE models in Johnson et al. (2021, hereafter J21) assumed that η increases approximately exponentially with radius (see equation 7 therein). While the MZR literature invokes variations in chemical equilibrium *between* galaxies, this model invokes variation *within* the Galactic disk. This choice enforced the correct gradient to arise by placing Z_{eq} as a function of radius on top of the observed gradient, making the gradient more of an assumption than a prediction. We simply asserted this origin of the metallicity gradient in J21 to focus on a different science goal (understanding the influence of stellar migration on the disk abundance structure). Testing these hypotheses, both $\eta \propto e^R$ and the EM in general, is our central interest in this paper.

In section 2 below, we present our empirical measurements of the evolution in the disk metallicity gradient. We describe our GCE models in section 3 below. We define the criteria distinguishing the EM from other GCE models in section 3.1 and discuss the parameter choices we focus on in this paper. Our comparison cases are described in section 3.2. We present empirical evidence supporting the equilibrium model in section 4 and show enlightening model-to-model comparisons. We discuss the EM's broad range of implications in section 5 and summarize our conclusions in section **??**.

2. EMPIRICAL AGE AND METALLICITY GRADIENTS

There are many spectroscopic surveys with publicly available data to choose from to characterize the age and abundance structure of the MW, such as LAMOST (Luo et al. 2015), GALAH (De Silva et al. 2015; Martell et al. 2017), *Gaia*-ESO (Gilmore et al. 2012), and APOGEE (Majewski et al. 2017). APOGEE provides excellent coverage of the Galactic disk by targeting luminous evolved stars accessible at large distances, making it particularly well-suited for
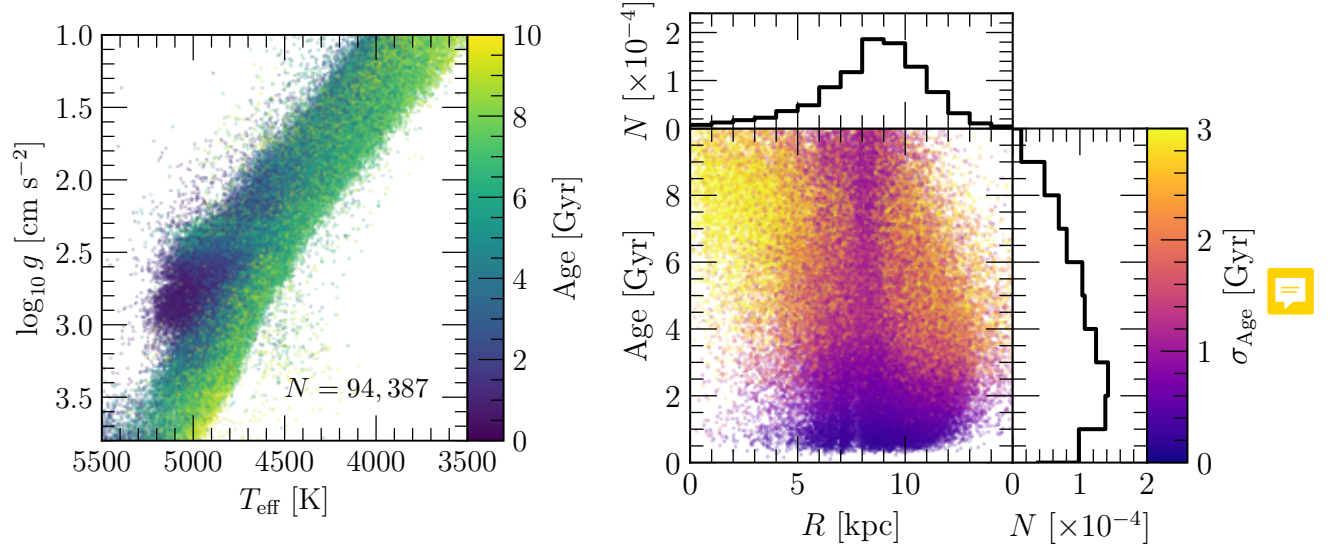
Figure 1. Our sample, the AstroNN value added catalog (Mackereth et al. 2019, see discussion in section 2.1). **Left**: The Kiel diagram, colour coded by stellar age according to the colorbar. **Right**: The Galactocentric radii and ages of each star, color coded by the reported uncertainties in the age estimates according to the colorbar. Top and right panels show the 1-D distributions in Galactocentric radius and age. **Summary**: By drawing stars from across the red giant branch, we construct a large sample with excellent coverage of the Galacti disk, with ~80,000 stars between $R = 5$ and 12 kpc. If the reported uncertainties are reliable, then stellar ages are precisely determined ($\langle\sigma_{\mathrm{Age}}\rangle \approx 1$ Gyr) across much of the Galactic disk up to ages of ~4 Gyr and up to ~10 Gyr in the solar annulus ($R = 7 - 9$ kpc).

this task. It is also less susceptible to dust obscuration since its spectra are at near-IR wavelengths ($\lambda = 1.51 - 1.70$ μm; Wilson et al. 2019).

2.1. *The Sample*

In this paper, we use the AstroNN value added catalog[2] for APOGEE's seventeenth data release (DR17; Abdurro'uf et al. 2022). In constructing the original catalog, Mackereth et al. (2019) used AstroNN (Leung & Bovy 2019a) to train a Bayesian convolutional neural network on APOGEE DR14 spectra and asteroseismic data from APOKASC-2 (Pinsonneault et al. 2018). Retrained on APOGEE DR17 spectra, the value added catalog improves the performance at low metallicity by incorporating asteroseismic ages from Montalbán et al. (2021) and additionally provides individual stellar abundances (Leung & Bovy 2019a) and distances (Leung & Bovy 2019b) retrained with *Gaia*-EDR3 (Gaia Collaboration et al. 2021). We filter the sample based on the following selection criteria:

- STAR_BAD == 0
- EXTRATARG == 0
- S/N $\geq$ 80
- log g = 1 − 3.8
- T_{eff} = 3500 − 5500 K

To avoid potential contamination by the main sequence, we additionally exclude stars that have both log $g > 3$ and $T_{\mathrm{eff}} < 4000$ K.

Uncertainties become substantial in the AstroNN catalog above ages of $\tau \gtrsim 8 - 10$ Gyr (Leung et al. 2023). We therefore focus on $\tau \leq 10$ Gyr stars, and impose an upper limit on Galactocentric radius of $R \leq 15$ kpc. We also focus on populations within $|z| \leq 0.5$ kpc of the disk midplane throughout this paper, because our sample most closely traces thin disk stellar populations. These cuts yield a final sample of $N = 94,387$ red giant and red clump stars, which we visualize in Fig. 1.

The left panel shows the Kiel diagram color coded by stellar age. Young populations are preferentially located in the red clump, whereas the oldest stars distribute themselves much more evenly along the red giant branch. This effect is primarily driven by actual changes in the distribution of stars along the giant branch as a function of population age (citation). However, systematic uncertainties in APOGEE tend to present as spurious correlations with T_{eff} and log g (e.g. Jönsson et al. 2018; Eilers et al. 2022). Leung et al.'s (2023) comparison of their age estimates with the AstroNN catalog indicates that systematic uncertainties are much more substantial for $\tau \gtrsim 8 - 10$ Gyr old stars. For younger populations, they may play a role in shaping the diversity in measurements of the gradient traced by stellar populations of different ages.

The right hand plot shows the radii and ages of each star, color-coded by their age uncertainties, along with the 1-D distributions in both Galactocentric radius and age. Our sample
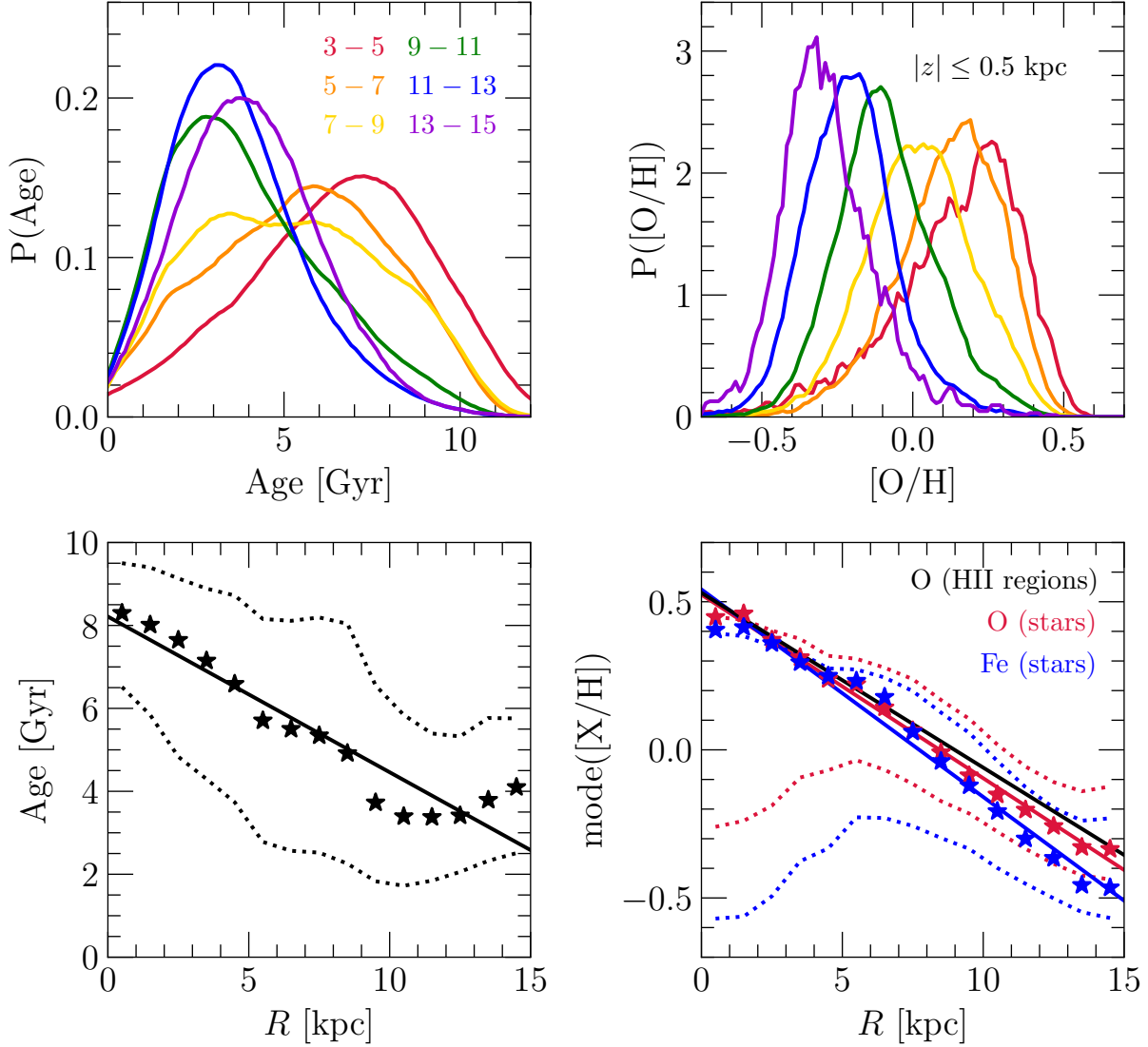
Figure 2. Radial age (left) and metallicity gradients (right) for stars within $|z| \leq 0.5$ kpc of the disk midplane in our sample. **Top**: Distributions in stellar age and [O/H] in 2-kpc wide bins of Galactocentric radius, color-coded according to the legend in the top-left panel. Each distribution is box-car smoothed with a window width equivalent to the median measurement uncertainty in each radial bin. **Bottom**: Median age and the peak of the metallicity distribution in 1-kpc bins of radius (stars). Dotted lines denote the 16th and 84th percentiles of the distribution in each radial bin, while solid lines mark the line of best fit to the corresponding summary statistics (parameters of which are given in Table 1). We additionally plot the gradient in [O/H] measured from Galactic HII regions by Méndez-Delgado et al. (2022, black). **Summary**: In agreement with previous work, stellar populations tend to be younger and more metal-poor at large radii.

achieves excellent coverage of the Galactic disk, particularly in the $R \approx 5 - 12$ kpc range. Through this paper, we generally focus on bins in radius and age that 1 kpc and 1 Gyr wide. The statistical uncertainties reach this level of precision or better across much of the Galactic disk up to ages of ~4 Gyr. Near the solar radius (assumed to be $R_\odot = 8$ kpc), this level of precision remains up to the age limit of our sample. We discuss sources of uncertainty further in section 2.4 below. (grab this later, it's a good callback but it's redundant here). Taken at face value, the reported uncertainties indicate that

our measurements most tightly constrain the enrichment history of the disk over the last few Gyr across all radii and over most of the disk lifetime near the sun.

2.2. Radial Gradients

Fig. 2 shows age and abundance distributions of our sample in bins of Galactocentric radius. We have not made any corrections for the survey selection function, so these are distributions of observed APOGEE stars satisfying our selection cuts as opposed to mass-weighted distributions. Stars tend to be young in the outer Galaxy and old in the inner

Galaxy, with skewed tails toward old and young populations, respectively. This result is a natural consequence of inside-out Galaxy formation, whereby the inner regions of the disk assemble at high redshift and the outer regions follow suit on longer timescales (e.g., Bird et al. 2013; White & Frenk 1991). Abundances follow a similar pattern, shifting from a metal-rich mode to a metal-poor mode with increasing radius. In practice, these radial gradients are quantified in a "compressed" manner – the information contained in the bulk shift in stellar population properties is collapsed into a summary statistic that varies with radius.

To that end, we first sort stars into 1-kpc wide bins in Galactocentric radius. We then compute the median age $\tau_{1/2}$ and the 16th and 84th percentiles of the distribution in each radial bin. The top left panel of Fig. 2 shows the resulting age gradient. A linear regression indicates a median trend of $\nabla\tau_{1/2} = -0.375 \pm 0.036$ Gyr/kpc with an intercept of 8.21 ± 0.31 Gyr. The median age closely follows this line of best fit at $R \lesssim 9$ kpc, but flattens off in the outer disk and reverses sign slightly.

Though we quantify the age gradient in terms of the median trend, the mode is our summary statistic of choice in quantifying metallicity gradients. We discuss our motivation behind this choice in section 5.2 below. In short, our GCE models suggest that the mode is resilient against modification by stellar migration. To mitigate noise in our inferences introduced by poisson fluctuations, we first fit a skew normal distribution to the MDF in each radial bins and determine the mode through numerical optimization.

The top right panel of Fig. 2 shows the resultant gradients in [O/H] and [Fe/H].[3] As expected, stars tend to decline in metallicity with increasing radius. Linear regressions indicate slopes of ∇[O/H] $= -0.062 \pm 0.001$ kpc^{-1} and ∇[Fe/H] $= -0.070 \pm 0.003$ kpc^{-1}, in reasonable agreement with previous measurements from APOGEE (e.g. Frinchaboy et al. 2013; Myers et al. 2022). For comparison, we additionally plot Méndez-Delgado et al.'s (2022) fit to the gas-phase O gradient traced by Galactic HII regions including their estimated correction for temperature inhomogeneities. The two O gradients are consistent within their 1σ uncertainties.

2.3. *Evolution of the Abundance Gradient*

Because our central interest in this section is to quantify the evolution of the disk metallicity gradient over time, we now repeat our measurements in section 2.2 above in 1 Gyr bins of stellar age. After sorting based on both age and radius, we fit for the mode of the MDF only in bins containing at least 200

---

[3] We follow conventional notation where [X/Y] $\equiv \log_{10}(N_x/N_y) - \log_{10}(N_x/N_y)_\odot$.

Table 1. A summary of the linear regressions applied to radial metallicity and age gradients in this paper (see discussion in sections 2 and 5.3). We use a pivot point at $R_{\rm gal} = 8$ kpc in all regressions (i.e. $y = m(R - R_\odot) + b$).

Age Range	Slope	Value at 8 kpc
[O/H]		
All Stars	-0.062 ± 0.001 kpc^{-1}	0.028 ± 0.006
≤ 9 Gyr	-0.062 ± 0.002 kpc^{-1}	0.028 ± 0.007
$0 - 1$ Gyr	-0.055 ± 0.011 kpc^{-1}	0.018 ± 0.022
$1 - 2$ Gyr	-0.056 ± 0.005 kpc^{-1}	0.028 ± 0.013
$2 - 3$ Gyr	-0.059 ± 0.005 kpc^{-1}	0.026 ± 0.016
$3 - 4$ Gyr	-0.066 ± 0.004 kpc^{-1}	0.014 ± 0.012
$4 - 5$ Gyr	-0.079 ± 0.004 kpc^{-1}	0.030 ± 0.012
$5 - 6$ Gyr	-0.080 ± 0.007 kpc^{-1}	0.051 ± 0.021
$6 - 7$ Gyr	-0.055 ± 0.008 kpc^{-1}	0.073 ± 0.025
$7 - 8$ Gyr	-0.049 ± 0.002 kpc^{-1}	0.125 ± 0.007
$8 - 9$ Gyr	-0.049 ± 0.007 kpc^{-1}	0.103 ± 0.030
$9 - 10$ Gyr	-0.001 ± 0.005 kpc^{-1}	0.029 ± 0.018
[Fe/H]		
All Stars	-0.070 ± 0.003 kpc^{-1}	-0.019 ± 0.013
≤ 9 Gyr	-0.068 ± 0.003 kpc^{-1}	-0.023 ± 0.014
$0 - 1$ Gyr	-0.068 ± 0.010 kpc^{-1}	0.045 ± 0.020
$1 - 2$ Gyr	-0.072 ± 0.005 kpc^{-1}	0.028 ± 0.014
$2 - 3$ Gyr	-0.077 ± 0.006 kpc^{-1}	-0.006 ± 0.017
$3 - 4$ Gyr	-0.083 ± 0.005 kpc^{-1}	-0.044 ± 0.015
$4 - 5$ Gyr	-0.096 ± 0.007 kpc^{-1}	-0.032 ± 0.020
$5 - 6$ Gyr	-0.097 ± 0.012 kpc^{-1}	-0.028 ± 0.035
$6 - 7$ Gyr	-0.066 ± 0.012 kpc^{-1}	0.012 ± 0.036
$7 - 8$ Gyr	-0.051 ± 0.004 kpc^{-1}	0.082 ± 0.014
$8 - 9$ Gyr	-0.061 ± 0.008 kpc^{-1}	0.013 ± 0.031
$9 - 10$ Gyr	0.038 ± 0.006 kpc^{-1}	-0.207 ± 0.022
Age		
All Stars	-0.375 ± 0.036 Gyr kpc^{-1}	5.21 ± 0.16 Gyr

stars. We estimate uncertainties in the mode in via jackknife resampling.

The top panels of Fig. 3 show the resultant measurements of the gradient's evolution in [O/H] and [Fe/H]. There is only minimal change in the gradient across most of the age range probed by our sample. Both the slope and normalization were established early in the disk lifetime, nearly 10 Gyr ago. To demonstrate this point further, the lower left panel of Fig. 3 shows the MDF in the solar annulus ($R = 7 - 9$ kpc) in the same 1 Gyr age bins. The distribution shifts only slightly toward supersolar [O/H] with increasing age.

We also apply linear regressions to the [O/H] $- R$ and [Fe/H] $- R$ trends. The lower middle and lower right panels
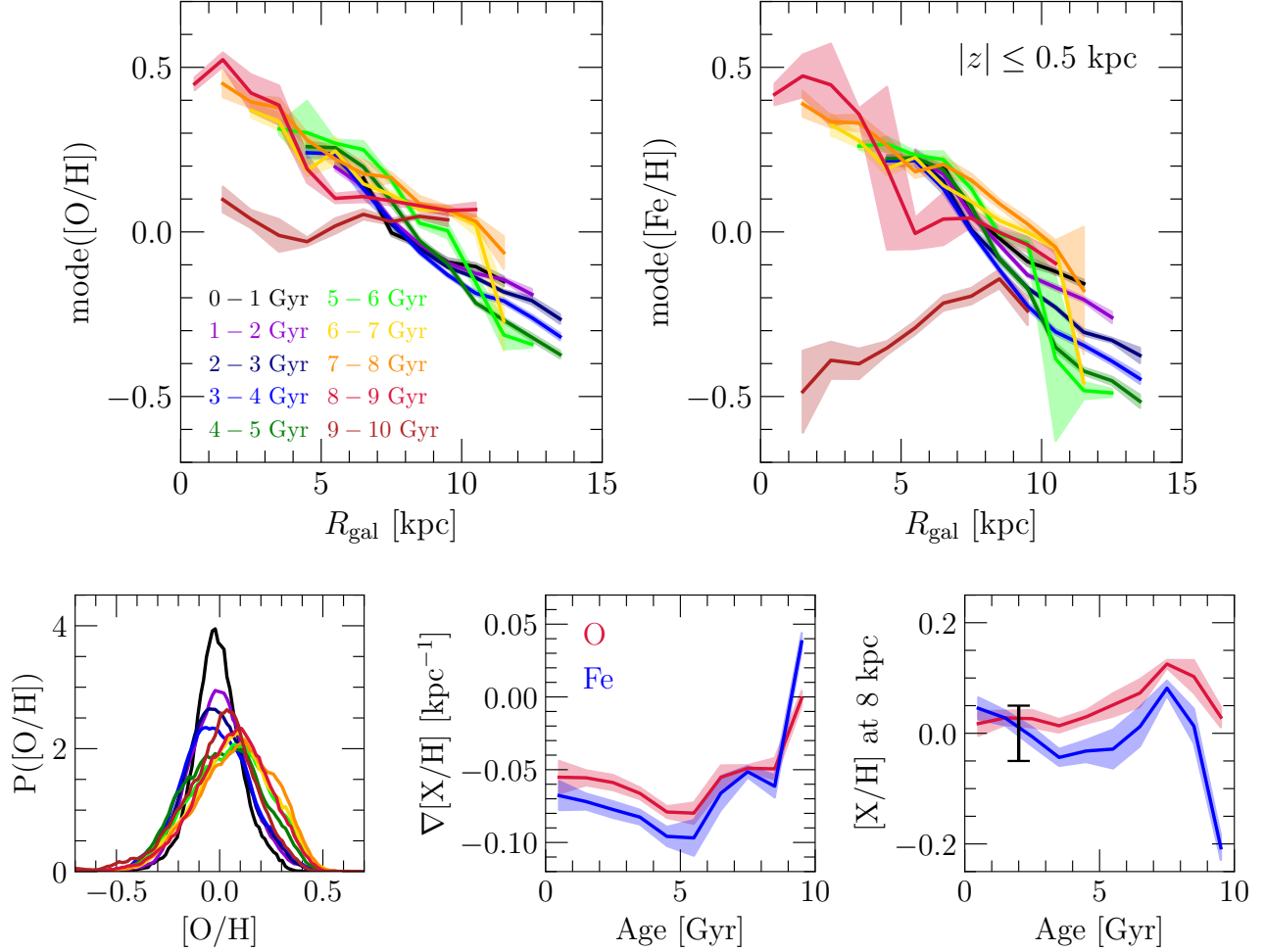
Figure 3. Radial metallicity gradients as a function of stellar age. **Top**: The peak of the [O/H] (left) and [Fe/H] (right) distributions in 1-kpc bins of radius, color-coded by stellar age according to the legend in the top-left panel. Shaded regions denote the statistical uncertainty in the mode of the MDF estimated through jackknife resampling (see discussion in section 2.3). **Bottom**: The distribution in [O/H] in the solar annulus ($R = 7 - 9$ kpc; left) in the same age bins as the top panels. Middle and right-hand panels show the slope and value at $R = 8$ kpc, respectively, inferred from linear regressions to the metallicity gradients in each age bin, parameters of which are reported in Table 1. **Summary**: The radial metallicity gradient of stars in the Galactic disk is largely age-independent up to ∼9 Gyr in terms of both slope and normalization.

show the slopes and intercepts as functions of the age bin, which are also reported in Table 1. Our measurements indicate that to first order, the gradient was established ∼8 − 9 Gyr ago. The trend held steady at ∇[O/H] ≈ ∇[Fe/H] ≈ −0.05 kpc^{-1} for ∼3 Gyr, then steepened to ∇[O/H] ≈ −0.07 kpc^{-1} and ∇[Fe/H] = −0.09 kpc^{-1} before gradually returning to its origin value up until the present day. Section 4 is dedicated to demonstrating that the scale of these variations with time is better described as perturbations of some equilibrium state as opposed to holistic changes in the abundances on Gyr timescales. In particular, the steepening of the gradient seen ∼6 Gyr ago is consistent with an accretion induced burst of star formation caused by the Sgr dSph's first infall (citation).

2.4. *Caveats*

Since our sample was produced by training AstroNN's deep learning capabilities on APOGEE spectra, the inferred ages may be biased by known correlations between ages and spectroscopically inferred abundances (e.g., the age-[O/Fe] relation; Feuillet et al. 2019). To assess the impact of this potential systematic uncertainty, we have replicated our measurements in this section with the Leung et al. (2023) age catalog, which we present in Appendix **??**. They carefully ensured that these estimates are independent of any abundance information in the spectra. We find similar results with both age catalogs, indicating that these biases in AstroNN are not at a level that would change our conclusions.

Because Leung et al.'s (2023) ages are reasonably unbiased to metal abundances, measurements with this catalog are perhaps more reliable than with AstroNN. However, their ages

are limited to surface gravities of $\log g = 2.5 - 3.6$ due to the training sample. As a result, the AstroNN value added catalog is a factor of ~2.5 times larger, yielding more precise summary statistics. The addition of the most luminous giants also improves the radial range of the Galactic disk with sufficiently many stars to compute the mode of the MDF. We therefore elect to focus on our measurements with the AstroNN catalog in this paper.

Particularly with regard to our measurement of the age gradient, the APOGEE selection function is a significant source of uncertainty. APOGEE primarily targets stars with 2MASS (Skrutskie et al. 2006) magnitudes $7 < H < 13.8$ on a grid of sightlines at Galactic latitudes of $b = 0, \pm4°$, and $\pm8°$ (targeting is described in detail in Zasowski et al. 2013, 2017, Beaton et al. 2021, and Santana et al. 2021). Quantifying the impact of this sampling bias in APOGEE first requires relating a star's apparent magnitude and color to a selection fraction based on its location on the 2-D sky (Bovy et al. 2016; Lian et al. 2022; Mackereth et al. 2017). One must then combine this mapping with a 3-D dust map to account for reddening and some model of the intrinsic number density of potential targets based on stellar properties from isochrones (e.g. Bressan et al. 2012; Dotter 2016; Choi et al. 2016; Haemmerlé et al. 2019) and an initial mass fuction (IMF; e.g., Kroupa 2001; Chabrier 2003), which then results in the selection fraction as a function of Galactic position, metallicity, and age (see also Mackereth & Bovy 2020).

The APOGEE selection function should not significantly impact our central conclusion. Selection alters the distribution of ages at a given location and metallicity, but red giant luminosities and lifetimes have little metallicity dependence near solar abundances (citation). Therefore, a representation such as Fig. 3 is relatively robust. Selection effects do, however, cast doubt on the significance of the steepening of the gradient seen around ages of ~6 Gyr. In the near future, it may be particularly constraining to replicate the measurements in this section with subgiants from SDSS-V's Milky Way Mapper program (Kollmeier et al. 2017).

3. GALACTIC CHEMICAL EVOLUTION MODELS

We use our previous GCE models from J21, which we integrate using the publicly available VERSATILE INTEGRATOR FOR CHEMICAL EVOLUTION (VICE; Johnson & Weinberg 2020). Following previous models with similar motivations (e.g. Schönrich & Binney 2009; Minchev et al. 2013, 2014), these models discretize the Galactic disk into 200 $\delta R = 100$ pc annuli spanning $R = 0$ to 20 kpc. Each ring is coupled to its neighbors via the exchange of stellar populations, which can carry their nucleosynthetic yields with them, but is otherwise described by a conventional one-zone GCE model (see e.g. the reviews by Tinsley 1980 and Matteucci 2021). The detailed description of this model can be found in section 2

of J21, and a short summary can be found in section 3 of Johnson et al. (2023). In this section, we provide an overview of the model components relevant to this paper.

Perhaps one more short paragraph on higher level details, such as disk lifetime and stellar migration, how many stellar populations we have in the model, the notion that we're tracking populations instead of individual stars so need to weight by mass, etc.

We focus on alpha and iron-peak elements in this paper, taking O and Fe as representative cases thereof. Production of these metals is dominated by core collapse supernovae (CCSNe) and Type Ia supernovae (SNe Ia; Johnson 2019). In the interest of flexibility, VICE accepts mathematical parameterizations of population-averaged stellar yields directly. They are defined as the total metal production per unit star formation, integrated over all events and encompassing all effects arising from, e.g., binary interactions and stellar explodability. We discuss our yields in the context of these processes in section X below.

We adopt the following choices in three out of our four GCE models:

- $y_O^{CC} = 0.0057$
- $y_O^{Ia} = 0$
- $y_{Fe}^{CC} = 4.5 \times 10^{-4}$
- $y_{Fe}^{Ia} = 8.4 \times 10^{-4}$,

where the superscripts CC and Ia denote the particular SN type. 35% (65%) of Fe arises from CCSNe (SNe Ia), very nearly the same breakdown as J21. We approximate CCSN production to be instantaneous. For example, a hypothetical 1000 $M_\odot$ stellar population would produce 5.7 $M_\odot$ of O through CCSNe. SN Ia events are known to instead follow a considerably extended delay-time distribution (DTD). In Johnson et al. (2021), we took a single power-law DTD τ^{-1} with a minimum delay time of $t_D = 150$ Myr based on the observed volumetric SN Ia rate as a function of redshift (Maoz & Mannucci 2012). Dubay et al. (2024) recently presented a suite of GCE models with different choices of the DTD from the literature. Based on their recommendation, we revise our choice in the shape and use an exponential DTD $R_{Ia} \propto e^{-\tau/\tau_{Ia}}$ with an e-folding timescale of $\tau_{Ia} = 1.5$ Gyr.

These yields are on a scale where the sum from both types of SNe is exactly the solar abundance[4] of the element in question. Weinberg et al. (2023) advocated for such a scale based on Rodríguez et al.'s (2022) measurements of the mean ^{56}Ni yield from Type II SNe based on the radioactive tails of their lightcurves. Our original yields in J21 were nearly a

---

[4] We use Asplund et al.'s (2009) measurements of O and Fe abundances in the solar photosphere.

Table 2. A summary of the GCE parameters we use in this paper (top) and their input parameters (bottom).

Model Name	$\eta_\odot$	R_η	$y/Z_\odot$	Star formation History	Calibrated to Reproduce
ηExp-$y3$	2.4	7 kpc	3	$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t)$ (Eq. X)	Radial stellar age gradient
ηExp-$y1$	0.4	7 kpc	1	$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t)$	Radial stellar age gradient
η0.4-$y1$	0.4	∞ (η = constant)	1	$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t)$	∇[O/H] observed in HII regions
η0-$y1$	0	∞ (η = constant)	1	$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t)$	∇[O/H] observed in HII regions
ηevolExp-$y3$	2.4	7 kpc ($t \le 8$ Gyr) 21.7 kpc ($t > 8$ Gyr)	3	$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t)$	Radial stellar age gradient
ηExp-$y3$+burst	2.4	7 kpc	3	$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t)\,(1 + A_b\Phi(t|t_b,\sigma_b,\alpha_b))$	Radial stellar age gradient

Parameter	Section	Description	Value
$\dot{\Sigma}_\star$	N/A	The surface density of star formation	Model-specific (see above)
Σ_g	N/A	The surface density of the ISM	Model-specific (see above)
$\dot{\Sigma}_{\text{in}}$	N/A	The surface density of accretion	Model-specific (see above)
$\tau_\star$	3	The star formation efficiency timescale $\tau_\star \equiv \Sigma_g/\dot{\Sigma}_\star$	Eq. X
y_O^{CC}	3	The population-averaged yield of O from CCSNe.	$0.0057(y/Z_\odot)$
$y_{\text{Fe}}^{\text{CC}}$	3	The population-averaged yield of Fe from CCSNe.	$(4.5 \times 10^{-4})(y/Z_\odot)$
y_O^{Ia}	3	The population-averaged yield of O from SNe Ia.	0
$y_{\text{Fe}}^{\text{Ia}}$	3	The population-averaged yield of Fe from SNe Ia.	$(8.4 \times 10^{-4})(y/Z_\odot)$
τ_{disk}	3	The duration of star formation in the disk.	13.2 Gyr
η	3.1	The outflow mass loading factor $\eta \equiv \dot{\Sigma}_{\text{out}}/\dot{\Sigma}_\star$	Model-specific (see above)
$\eta_\odot$	3.1	The value of η at $R = R_\odot$, the Galactocentric radius of the sun.	Model-specific (see above)
R_η	3.1	The scale radius of the increase in η.	Model-specific (see above)

factor of three larger. We will demonstrate in section 4 that the origin of our results in section 2 is sensitive to the overall normalization of stellar yields. To this end, we also present a model in which the yields are amplified by a factor of three, which places them only slightly larger than the original yields from J21.

Dubay et al. (2024) also updated the prescription for stellar migration in the J21 models. Originally, we "tagged" stellar populations in the GCE model with "analogs" from the h277 hydrodynamic simulation, adopted their change in orbital radii $\Delta R = R - R_{\text{birth}}$ at face value, and moved the stellar populations with some assumed time dependence ($\sqrt{t}$ in the fiducial model). The vast majority of stellar populations are matched with a star particle that formed nearby in terms of both radius and time, but the algorithm runs into sampling issues due to a low number of star particles at particularly large radii and early times. Dubay et al. (2024) describe a procedure that closely approximates the h277 predictions with sampling from normal distributions to determine radial

displacements (further details can be found in their Appendix X). We retain this updated formalism here.

We parameterize the shape of the SFH using J21's "inside-out" model, which we refer to in this paper as the "rise-fall" model to differentiate it from the inside-out formation of the disk. At fixed radius, the SFR evolves with time according to

$$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t) = \left(1 - e^{-t/\tau_{\text{rise}}}\right) e^{-t/\tau_{\text{sfh}}}, \qquad (1)$$

where τ_{rise} and τ_{sfh} control the timescales on which the SFR rises at early times and falls at late times, respectively. The normalization is set at each radius such that the predicted surface density gradient at late times is consistent with Bland-Hawthorn & Gerhard's (2016) findings. We calibrate τ_{rise} and τ_{sfh} slightly differently for each GCE model in our primary set, which we present in Appendix B. In short, each model reasonably reproduces the observed radial gradients in stellar age and O abundances in HII regions. In section 3.3 below, we describe a model in which we impose an accretion induced burst of star formation on top of this background SFH.

We have also updated the star formation law used in these GCE models. In J21, we used a three component power-law relationship between the surface densities of gas and star formation, $\dot{\Sigma}_\star \propto \Sigma_g^N$ based on aggregate observations by Bigiel et al. (2010) and Leroy et al. (2013, see Fig. 2 in Krumholz et al. 2018). In practice, the three component power-law complicates our model calibration in Appendix B. We therefore use the more classical single power-law prescription with $\dot{\Sigma}_\star \propto \Sigma_g^N$ with $N = 1.5$ based on Kennicutt (1998). This relation has been seen to vary from galaxy to galaxy, with considerable differences in slopes and zero-points (citation), making it a nuisance parameter in GCE models. Much of the uncertainty can be traced to the ongoing debate about the CO-to-H_2 conversion factor (citation). Nonetheless, de los Reyes & Kennicutt (2019) and Kennicutt & de los Reyes (2021) argue that it remains a viable prescription for galaxy formation models.

With the rest of this section, we present the GCE model parameter choices that we focus on in this paper. We define the equilibrium model in section 3.1 below, along with the models that fall under its umbrella. In section **??**, we describe our comparison cases for testing the equilibrium model, which we dub "sub-equilibrium" models. Section 3.3 presents two additional comparisons in which the model-to-model comparisons yield useful insights in our discussion of previous results in the literature in section 5.4.

We provide Table 2 as a reference. In its top half, we summarize the GCE models that receive the bulk of this paper's attention. In its bottom half, we provide section numbers, descriptions, and values for the most important input parameters in the models. We highlight our "primary set" of GCE models, which refers to the ηExp-y3, ηExp-y1, η0.4-y1, and η0-y1 models listed at the top of Table 2. These are the models whose predictions we compare against the data, while the remaining are invoked only for the sake of model-to-model comparisons.

3.1. *The Equilibrium Model*

We invoke the equilibrium model first and foremost as a qualitative interpretation of our measurements in section 2. We therefore define it as a class of models with similar predictions as opposed to a specific parameterization. In this subsection, we describe the generalized definition of the equilibrium model and discuss our specific choice of parameters that satisfies the requirements we describe.

At its core, the equilibrium model postulates that metallicity gradients arise in spiral galaxies for two reasons: the equilibrium abundance itself declines with radius, and the ISM tends to reach equilibrium early in the disk lifetime. More specifically, a GCE model falls under the umbrella of the equilibrium model if its predictions satisfy each of the following criteria:
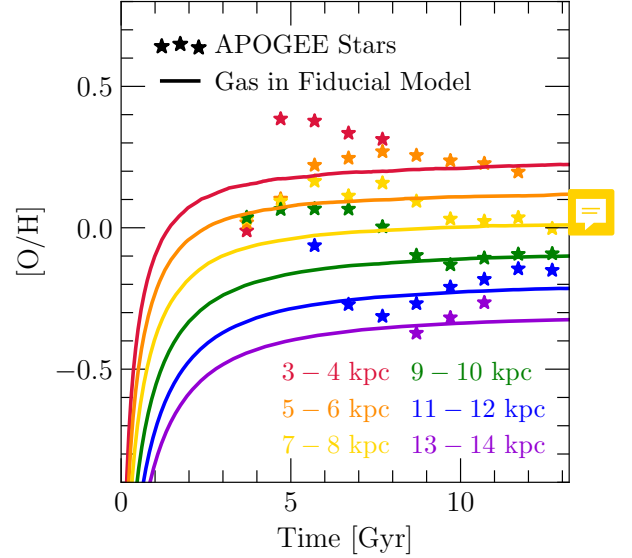
Figure 4. A visualization of the equilibrium model's main qualitative argument. Stars mark the mode of the MDF measured in bins of stellar age and Galactocentric radius (see Fig. 3 and discussion in section 2), color-coded according to the legend and shown as a function of time rather than age (assuming a disk lifetime of $\tau_{disk} = 13.2$ Gyr. Lines show the ISM abundance at the center of each radial bin as predicted by our ηExp-y3 GCE model (see discussion in section 3). **Summary**: Stellar abundances at fixed radius are nearly age independent across much of the disk up to ages of $\sim 4-5$ Gyr. This result is an indication that ISM abundances have varied only minimally for the past several Gyr.

Criterion 1. At all radii within the Galactic disk, the metal abundance in the ISM evolves toward a value at which metal production by stars is balanced by the accretion of fresh hydrogen. This *local equilibrium abundance* Z_{eq} is at a value typically observed in the Galactic disk.

Criterion 2. The equilibrium abundance itself declines exponentially with radius. This *equilibrium gradient*,

$$\nabla_{eq} \equiv \frac{\partial}{\partial R} \log_{10}\left(\frac{Z_{eq}}{Z_\odot}\right) \propto -R, \qquad (2)$$

is an intrinsic property of the Galactic disk set by variations in GCE parameters with radius.

Criterion 3. The evolution toward equilibrium is fast ($\lesssim 1$ Gyr timescales). Consequently, the equilibrium gradient is reached early in the disk lifetime.

The combination of all three criteria leads to the defining feature of the equilibrium model: that ISM abundances decline with radius, but do not evolve significantly with time. In section 5 below, we discuss predictions that generally follow as a consequence of this sort of enrichment history.

Mathematically, the equilibrium model can be written as

$$[\alpha/\text{H}]_\odot \approx [\alpha/\text{H}]_{\odot,\text{eq}}$$
$$\nabla[\alpha/\text{H}] \approx \nabla[\alpha/\text{H}]_{\text{eq}}, \qquad (3)$$

where $[\alpha/\text{H}]_\odot$ is the alpha element abundance at the solar radius $R_\odot$, $[\alpha/\text{H}]_{\odot,\text{eq}}$ is the local equilibrium abundance at $R_\odot$, and $\nabla[\alpha/\text{H}]_{\text{eq}}$ is the equilibrium gradient in alpha elements. We tend to focus our discussion in this paper on O abundances rather than Fe, because alpha elements approach equilibrium significantly faster than iron-peak elements in GCE models (Weinberg et al. 2017). This difference arises due to the short timescales (of order $\lesssim 50$ Myr) associated with massive star enrichment and the extended DTD of SN Ia events (characteristic delay times of order ~ 1 Gyr). In order to fully utilize the constraining power of our sample though, we compare our predicted abundances of both O and Fe with our data in section 4 below.

In one-zone GCE models where the late-time SFH is exponential, the equilibrium alpha element abundance is closely approximated by

$$Z_{\alpha,\text{eq}} = \frac{y_\alpha^{\text{CC}}}{1 + \eta - r - \tau_\star/\tau_{\text{sfh}}}, \qquad (4)$$

where $\tau_\star \equiv \Sigma_{\text{g}}/\dot{\Sigma}_\star$ is the star formation efficiency (SFE) timescale. Often referred to as the "depletion time" in the observational literature on galaxies, this parameter is more an efficiency descriptor in GCE models, because it is also the *fractional* rate of star formation. $r \approx +0.4$ is the recycling correction, approximately accounting for the return of stellar envelopes back to the ISM (Weinberg et al. 2017; this value is appropriate for a Kroupa 2001 IMF).

Each of the models that we present in this paper that reasonably satisfy the criteria defining the equilibrium model invoke an exponential increase in the mass loading factor $\eta \propto e^R$ according to

$$\eta = \eta_\odot e^{(R-R_\odot)/R_\eta}, \qquad (5)$$

where $\eta_\odot$ is the value of η at the solar radius, and R_η is its scale radius. In these models, the value of R_η approximately sets the slope of the equilibrium gradient. The reason for this close relationship can be understood by considering the limit that η is large. Under that approximation, $Z_{\alpha,\text{eq}} \approx 1/\eta \propto e^{-R}$, which ultimately satisfies Criterion 2.

The yield-outflow degeneracy also becomes most apparent in this limit (see discussion in section 5.6). However, the timescale to reach the equilibrium abundance turns out to be sensitive to their overall scale (see discussion in section X). We therefore present two models with different overall scales of stellar yields but with the same yield ratios. In the first, we take the fiducial scale described at the beginning of this section, in which population-averaged stellar yields

are exactly the solar abundance, $y/Z_\odot = 1$. In the second, we take $y/Z_\odot = 3$. This higher yield model is the one that ultimately performs the best in its comparison with our data. We therefore take it as a fiducial case and construct additional variations of it in sections 3.3 below.

We set the values of τ_{rise} and τ_{sfh} as functions of radius, which control the shape of the SFH (see equation 1), such that the radial gradient in stellar ages is approximately reproduced. Appendix B describes the procedure, which amounts to associating the 50th percentile of the SFH with a median age that we can forecast without having to run the model. The age gradient that we measure is likely affected by bias introduced by the APOGEE selection function, so the age gradients that these models predict at likely flawed. However, this bias is not a concern for this paper. Ours is a qualitative comparison, so are only concern is that the models predict the age gradient to have an accurate shape. Though our calibration procedure is not perfect (see discussion in Appendix B), the models that fail to reproduce our results are unsuccessful for reasons unrelated to our calibration procedure.

3.2. *Comparison Cases*

Paragraph: Comparison cases have either constant or zero η, because this takes away the equilibrium gradient based on the above discussion.

Paragraph: $\eta = 0$ is a common choice for modeling the MW, we just do $\eta = 0.4$ as an intermediate case. Normalization is picked so that it has $\eta_\odot$ from the ηExp-$y1$ model.

Paragraph: These models need lower yields because reasons.

Paragraph: Calibrate these models to the present day gasphase metallicity gradient. Age gradient comes out broadly consistent anyway, which is great considering there were calibration challenges.

3.3. *Perturbative Models*

Paragraph: We construct two variations of our fiducial model that include perturbations from the equilibrium state. We highlight them in this paper as their differences between the fiducial model and how they compare to the data leads to useful insight.

Paragraph: The first is a case in which we build in a burst of star formation. Show the equation and say that it's a modification of our outerburst model from J21. Parameters are tuned to reproduce this supposed dilution event seen in Fig. 3.

Paragraph: The next is a model in which R_η suddenly changes, leading to a sudden change in the slope of the equilibrium gradient.

4. RESULTS

We begin the presentation of our results with a handful of comparisons between the models themselves. We highlight
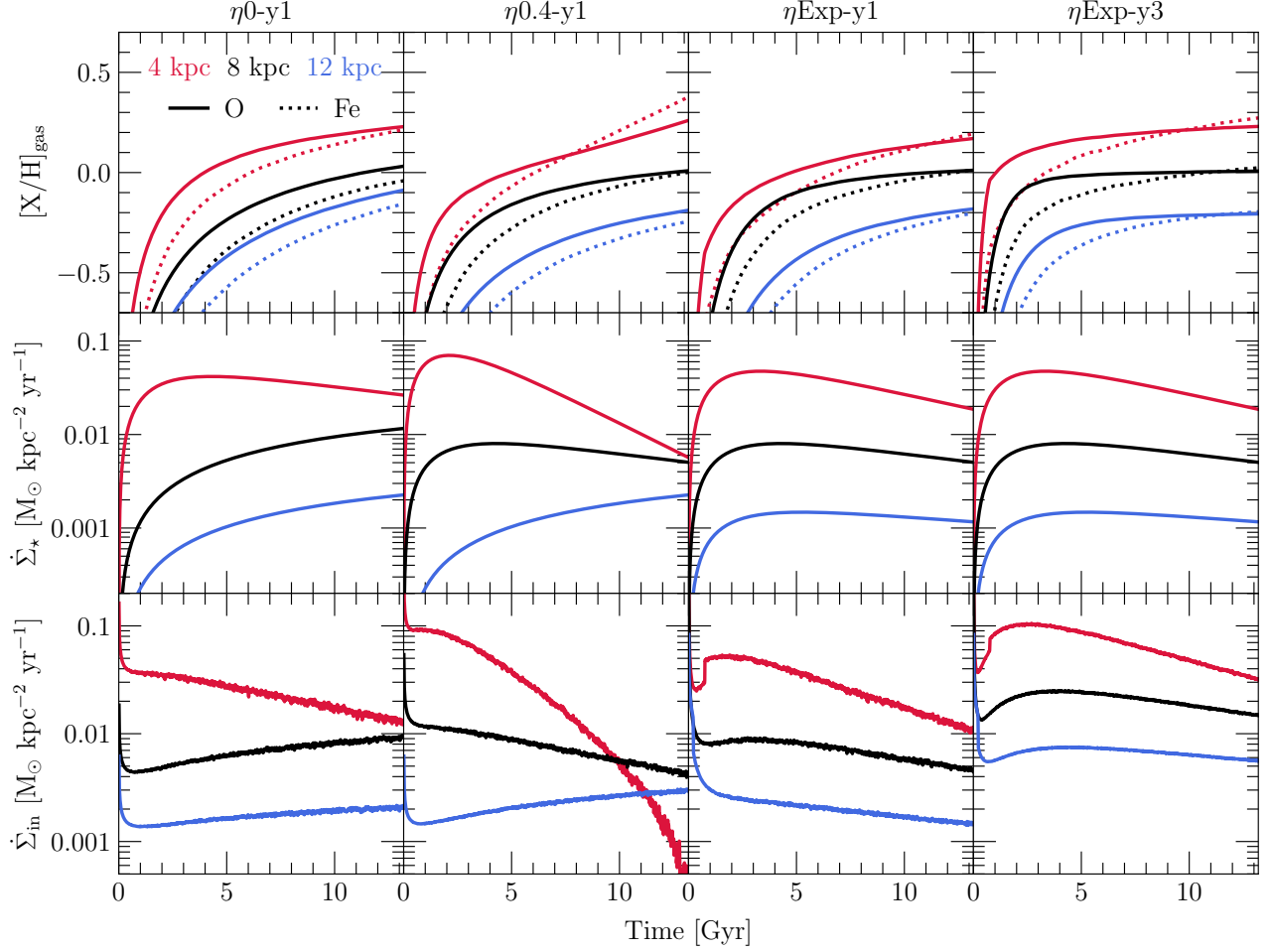
Figure 5. Enrichment (top), star formation (middle), and infall histories (bottom) for our primary set of GCE models, labeled at the top of each column of panels (see Table 2 and discussion in section 3). We visualize the evolution at Galactocentric radii of $R = 4$ (red), 8 (black), and 12 kpc (blue) in all panels. O and Fe abundances in the top row are distinguished with solid and dotted lines, respetively. **Summary**: Each model approximately reproduces the radial metallicity gradient observed in HII regions (Méndez-Delgado et al. 2022; see Fig. **??** in Appendix B), but evolves to that point differently due to differences in GCE parameters.

the underlying behavior that makes the models physically different, which provides useful insight in comparisons with the data later in this section.

Fig. 5 shows the enrichment and evolutionary histories predicted by our primary set of models (see Table 2 and discussion in section 3). All models reach similar ISM abundances at the present day, because we have calibrated their input parameter to do so (see Appendix B). However, each model reaches the present-day abundances differently.

The SFHs in the outer disk in the sub-equilibrium models are quite extended, and in some cases monotonically increasing until the present day. This arises as a necessary component of their calibration. The steadily increasing SFH requires a steadily increasing accretion rate to fulfill the mass budget (see discussion in section 4.1 below). The continual buildup of fresh hydrogen helps keep the metallicities low.

The $\eta0.4$-$y1$ model has the strongest changes in the shape of the SFH with radius. The inclusion of outflows at this level makes it challenging to reach super-solar abundances. As a consequence, our calibration finds the "gas-starved ISM" effect noted by Weinberg et al. (2017) as a manner in which to reach the required gas-phase abundances in the inner disk. This term refers to the increase in the equilibrium abundance in the limit of low τ_{sfh} (see equation 4). In this scenario, metals are ejected to a gas reservoir that is rapidly losing its H because accretion is not keeping up with the current rate of star formation. Our calibration has dialed in a particular $\tau_{\mathrm{sfh}} - R$ relation in the inner disk that leads to a close approximation of the present day O abundance gradient in the ISM.

Our ηExp-$y1$ and ηExp-$y3$ models have similarly shaped SFHs. This similarity arises because τ_{rise} and τ_{sfh} are tuned to reproduce the observed median stellar age as a function of
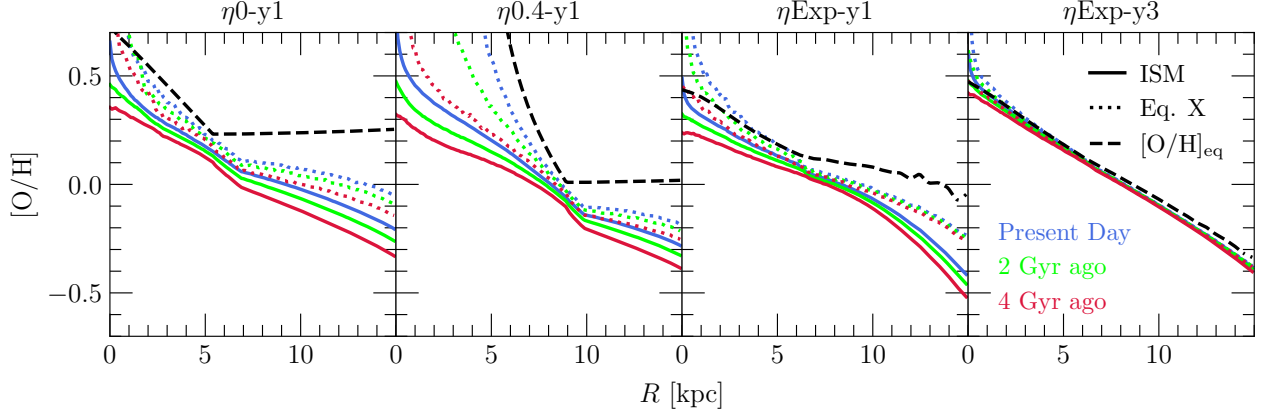
Figure 6. Evolution in gas-phase O abundance gradients in comparison to their equilibrium abundances. Solid lines show the abundances at the present day and at two recent snapshots predicted by each of the four GCE models in our primary set (see Table 2 and discussion in section 3). The black dashed line marks the equilibrium abundance as a function of radius R (see equation **??**). Dotted lines mark approximations thereof from equation X evaluated at the same three snapshots. **Summary**: Close to the equilibrium abundance, the ratio of star formation per unit infall becomes close to time independent. The equilibrium gradient tracks changes in this ratio with radius.

radius (see Appendix B). A detail that turns out to be particularly noteworthy is that the ηExp-$y3$ model has an infall history that is a factor of ~3 higher at all times and radii. We also call attention to the fact that the ηExp-$y3$ model reaches the equilibrium across at most radii significantly earlier than the ηExp-$y1$ model. We argue in section 4.1 below that this relationship makes the equilibrium model favor nucleosynthesis models in which the scale of stellar yields is large ($y/Z_\odot \gtrsim 2$).

4.1. *The Equilibrium Gradient*

In this subsection, we demonstrate that the disk metallicity gradient tracks a decline in the ratio of star formation per unit infall when in an equilibrium state. The relation we find, which is most accurate for alpha elements, is given by

$$\nabla_{eq} = \frac{1}{\ln 10} \left(\frac{\partial \ln \dot{\Sigma}_\star}{\partial R} - \frac{\partial \ln \dot{\Sigma}_{in}}{\partial R} \right). \qquad (6)$$

We begin by first showing this result predicted by our GCE models. Compute scale radius r_{in} based on the observed gradient and disk scale radius, but not here (somewhere in discussion section).

Fig. 6 shows [O/H] abundances in the ISM at the present day and at two recent snapshots for each model in our primary set. We also plot the equilibrium gradient for each model, evaluated according to equation 4 for all radii (we use the value of $\tau_\star$ at the final snapshot). The ηExp-$y1$ and ηExp-$y3$ models are reasonably close to their equilibrium gradients (within ~0.1 dex across most radii), perhaps with the exception of the ηExp-$y1$ model at the very outer edge of the disk. As a result, the evolution in gas phase abundances between the recent snapshots and the present day is minimal. In the $\eta0$-$y1$ and

$\eta0.4$-$y1$ models, there is significant evolution over the most recent ~few Gyr. As expected, the equilibrium abundances are much higher in these models.

For comparison, we plot equation **??** below, which is a proxy for the equilibrium gradient based on the ratio of star formation per unit infall. In models where an equilibrium gradient arises, the ratio $\dot{\Sigma}_\star/\dot{\Sigma}_{in}$ reaches some time independent value that depends on radius. The value that it settles to determines the local equilibrium abundance.

In the $\eta0$-$y1$ and $\eta0.4$-$y1$ models, neither the ratio $\dot{\Sigma}_\star/\dot{\Sigma}_{in}$ nor the gas phase metallicity gradients reach an equilibrium state. There is evolution all the way up to the present day. An equilibrium abundance exists at all radii in the $\eta = 0$ case, but only outside of $R \approx 5 - 6$ kpc in the $\eta = 0.4$ case. However, equilibrium is not reached anywhere in these models, even at the radii where it has a physical value.

Based on this behavior in the predicted equilibrium abundances, the $\eta0.4$-$y1$ scenario is an interesting theoretical case. Its SFH in the inner disk (see Fig. 5) decline so quickly (low τ_{sfh}) that the denominator of $Z_{\alpha,eq}$ inverts sign, and the model predicts $Z_{\alpha,eq} < 0$. These cases are not unphysical; rather, they simply predict the abundances to rise with time forever, because they can never possibly reach their equilibrium state (see discussion in Appendix B).

Fig. 7 illustrates the physical origin of these differences in predictions in our primary set. The left panel shows the [O/H] abundance in the ISM as a function of time at $R = 8$ kpc, our assumed Galactocentric radius of the Sun. The models reach solar metallicity on vastly different timescales. While the ηExp-$y3$ model spends ~10 Gyr near solar metallicity, the $\eta0$-$y1$ model reaches it only ~3 Gyr ago.
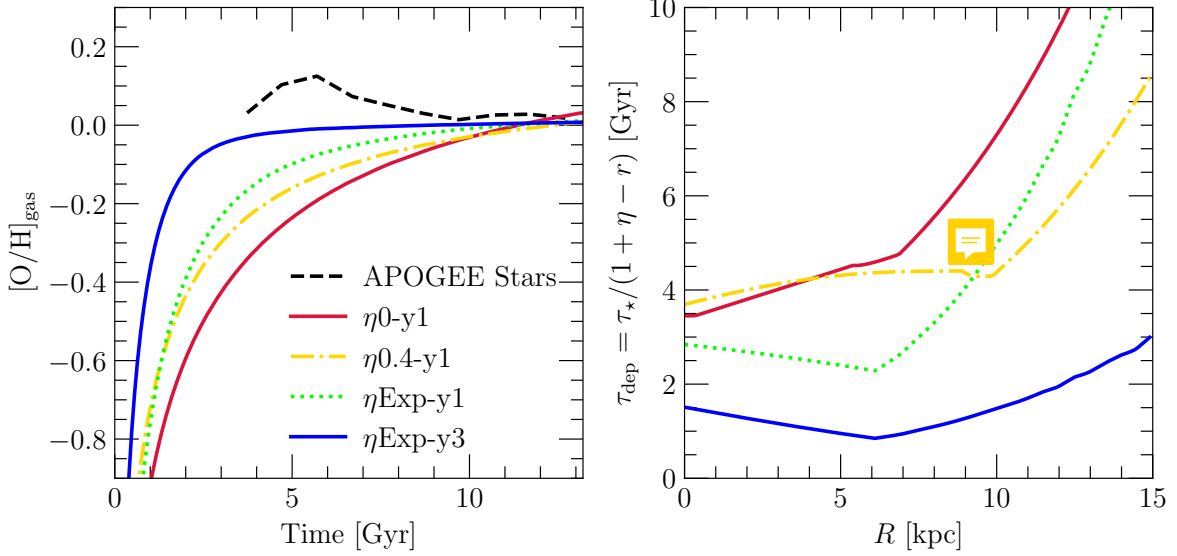
Figure 7. The key difference between our GCE models. **Left**: the O abundance in the $R = 8$ kpc ring as a function of time as predicted by our primary set of GCE models (see Table 2 and discussion in section 3), labeled according to the legend. For comparison, we plot the values inferred from our linear regressions reported in Table 1 (identical to the red line in the lower right panel of Fig. 3, but shown as a function of time assuming a disk age of $\tau = 13.2$ Gyr; see discussion in section 3). **Right**: The timescale of the baryon cycle (see equation X) as a function of radius at the present day predicted the same GCE models in the left panel. **Summary**: Reaching equilibrium early in the disk lifetime is a natural consequence of a fast baryon cycle, meaning that a typical fluid element does not remain in the ISM for long before being incorporated into new stars or ejected in an outflow.

In the right panel, we plot the *baryon cycle timescale*[5] as a function of radius, defined as

$$\tau_{\text{bary}} = \frac{\tau_\star}{1 + \eta - r}, \qquad (7)$$

for each model at the present day. This timescale describes the average amount of time a fluid element remains present in the ISM before being incorporated into new stars or lost in an outflow. Weinberg et al. (2017) demonstrate analytically that one-zone GCE models with low τ_{bary} reach equilibrium more quickly than high τ_{bary} models for a given choice of SFH (see discussion in Appendix B). Fig. 7 indicates that our GCE models make similar predictions in a radially resolved manner. (grab this later, it's a good callback, but redundant here) The Galaxy being in chemical equilibrium for a substantial fraction of the disk lifetime arises in GCE models when the baryon cycle is fast (i.e. inflows and outflows are vigorous as opposed to slow).

With the remainder of this subsection, we mathematically justify this relationship between the equilibrium gradient ∇_{eq} and the ratio of star formation per unit infall $\dot{\Sigma}_\star / \dot{\Sigma}_{\text{in}}$. In

a given ring embedded within our multi-zone GCE models (see discussion in section 3), the time derivative of the gas surface density[6] follows a summation over source and sink terms:

$$\dot{\Sigma}_g = \dot{\Sigma}_{\text{in}} - \dot{\Sigma}_\star (1 + \eta - r), \qquad (8)$$

where $\eta \dot{\Sigma}_\star \equiv \dot{\Sigma}_{\text{out}}$ is the surface density of the outflow and $r \dot{\Sigma}_\star \equiv \dot{\Sigma}_r$ is the surface density of stellar envelopes being returned back to the ISM (see discussion in section 3.1). This expression enforces the conservation of mass within GCE models. From here, it is trivial to isolate $\dot{\Sigma}_\star / \dot{\Sigma}_{\text{in}}$:

$$\frac{\dot{\Sigma}_\star}{\dot{\Sigma}_{\text{in}}} = \left[1 + \eta - r + \tau_\star \frac{\dot{\Sigma}_g}{\Sigma_g} \right]^{-1}, \qquad (9)$$

where we have also substituted in the SFE timescale $\tau_\star \equiv \Sigma_g / \dot{\Sigma}_\star$ (see discussion in section

We now write the production rate of alpha elements, whose enrichment is approximated to happen instantaneously following star formation in our GCE models. We can use equation 7 from Weinberg et al. (2017) directly:

$$\dot{\Sigma}_\alpha = y_\alpha^{\text{CC}} \dot{\Sigma}_\star - Z_\alpha \dot{\Sigma}_\star (1 + \eta - r). \qquad (10)$$

The rate of change in the alpha metallicity $Z_\alpha = \Sigma_\alpha / \Sigma_g$ then follows from quotient rule and substituting in equations 8, 9,

---

[5] In J21, we refer to τ_{bary} as the "depletion time." We change nomenclature in this paper, because this name has multiple meanings in the literature. We suspect that additional components of the baryon cycle, such as radial gas flows, might add additional terms to the denominator (see discussion in section 5.6). We use this term as it refers generically to all processes adding and removing matter from the star forming reservoir relative to the SFR.

[6] Because our rings are fixed in radius, their area is constant, so the time-derivatives of surface density are ultimately time derivatives of mass.

and [10]:

$$\dot{Z}_\alpha = \frac{\Sigma_g \dot{\Sigma}_\alpha - \Sigma_\alpha \dot{\Sigma}_g}{\Sigma_g^2}$$

$$= \frac{y_\alpha^{CC}}{\tau_\star} - \frac{Z_\alpha}{\tau_\star}\left(1 + \eta - r + \tau_\star \frac{\dot{\Sigma}_g}{\Sigma_g}\right) \quad (11)$$

$$= \frac{y_\alpha^{CC}}{\tau_\star} - \frac{Z_\alpha}{\tau_\star}\left(\frac{\dot{\Sigma}_{in}}{\dot{\Sigma}_\star}\right).$$

The local equilibrium abundance then follows by letting $\dot{Z}_\alpha = 0$. In detail, $Z_{\alpha,eq}$ is best written in terms of our GCE parameters according to equation [4]. However, the equilibrium model by definition predicts the ISM to be close to the local equilibrium abundance across much of the disk. We then expect the following relation to hold across much of the Galactic disk

$$Z_\alpha \approx Z_{\alpha,eq} = y_\alpha^{CC} \frac{\dot{\Sigma}_\star}{\dot{\Sigma}_{in}}. \quad (12)$$

Equation [6], defined at the beginning of this subsection, then follows by converting the metallicities from Z_α into [α/H] and differentating with respect to radius. (Move this to a discussion subsection somewhere). This is a particularly interesting implication of the equilibrium model, because it could be used to infer accretion rates from observables (modulo the scale of yields, see discussion in section X). We caution based on Fig. [7] that the accuracy of the relation breaks down if future empirical tests find that the equilibrium model is inaccurate.

4.2. *Metallicity Distributions in the Solar Annulus*

Paragraph: Fig. [8] shows the MDF in the solar neighborhood for a couple age bins predicted by the model. The $\eta0$-$y1$ model struggles the most significantly, overpredicting the frequency of subsolar metallicity stars at old ages.

Paragraph: This arises due to the behavior shown in Fig. [7] and discussed in section [4.1]. If the approach to equilibrium isn't fast enough, the model predicts sub-solar metallicity populations at ages where the data indicate they should not exist.

Short paragraph: Comparison against Fe shows the same results but to a greater extent. The restitution delay of Fe makes it harder to reach equilibrium on short timescales. This difference is at the heart of subsequent comparisons in this section.

4.3. *Disk Metallicity Gradient*

Paragraph: We start with the total gradient shown in Fig. [X]. Oh my god the sub-equilibrium models have the stars below the gas, which isn't observed.

Paragraph: Now we split into bins of age in Fig. [10], and we see the same effects but now spread out along the time axis. The ηExp-$y3$ is the only model that appears to escape the effects of an evolving normalization in the $\sim 5 - 10$ kpc range of radius.

Short paragraph: The results of both comparisons arise because the abundances are continually evolving across much of the disk lifetime in the $\eta0$-$y1$ and $\eta0.4$-$y1$ models.

4.4. *The Age-Metallicity Relation*

Paragraph: Fig. [11] presents shows the comparison as Fig. [10], but against the age axis at fixed radius as opposed to vice versa. In the $\eta = 0$ and $\eta = 0.4$ models, the abundances at high ages, especially in the inner disk, tend to underpredict the observations by as much as $\sim 0.3 - 0.4$ dex in some cases.

Short paragraph: We demonstrate in this subsection that this result is closely related to the realization in recent years that the AMR is C-shaped.

Paragraph: We construct a quick and dirty measurement of the AMR in our sample without correcting for selection effects. We are mostly interested in its qualitative shape, which is likely not rewritten by selection effects. Use methodology akin to Feuillet et al. (2018, 2019) where we in abundances instead of ages.

Short paragraph: Results are shown in Fig. [12]. It has a C-shape, which has been noted previously, and was a point of discussion in J21 (see their section X).

Paragarph: We apply the same measurements to our model predictions, weighting by the mass accordingly. The ηExp-$y1$ and ηExp-$y3$ models perform the best at reproducing the shape, in particular the turn toward higher ages at $\sim$solar metallicity. Other models predict characteristic ages of super solar metallicity stars to continue to decline, in tension with the data.

For discussion at some point: The C-shaped AMR supports the EM. The presence of old, super-solar metallicity stars near the solar neighborhood is difficult to reproduce without it.

4.5. *Perturbations and Sudden Events*

Paragraph (or 2): Now we inspect the perturbative models described in section [3.3]. Outerburst model shown in Fig. [13]. The gas-phase gradient deviates from its equilibrium state near the time of maximum accretion. The result is an accretion induced burst in star formation that dilutes the ISM, re-enriches, and relaxes on $\sim$Gyr timescales. The timing of the burst seems to line up with the first infall of the Sgr dSph, a point we discuss further in section X below.

Paragraph: Fig. [14] shows the model in which R_η suddenly increases at $t = 8$ Gyr. We see a shallowing of the gradient over ~ 2 Gyr following the parameter change. We do not see such behavior in the data, but Imig et al. (2023) have argued in this direction (see their Fig. 17). We return to this point once more in section [5] below in the context of our comparison against Lu et al. (2022).

5. DISCUSSION

In section [4], we measured the stellar metallicity gradient in the Galactic disk to be interestingly close to age invariant
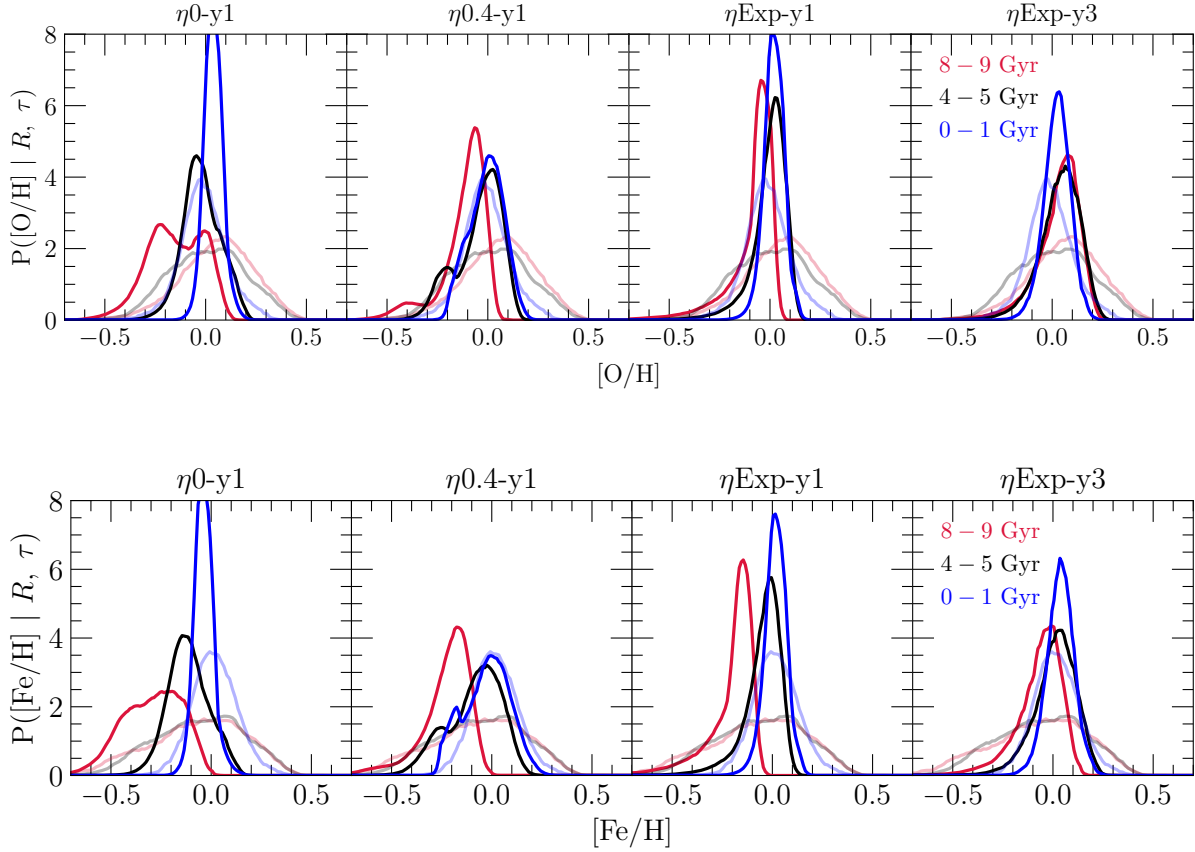
Figure 8. MDFs for different mono-age populations in the solar annulus (Galactocentric radius $R = 7 - 9$ kpc and midplane distances $|z| \leq 0.5$ kpc), denoted according to the legend in the right hand panels. Transparent lines show the observed distribution in our APOGEE sample (i.e. those in the top panels are identical to the corresponding curves in the lower left panel of Fig. 3; see discussion in section 2). Solid lines show the predictions of the GCE model marked at the top of each panel (see Table 2 and discussion in section 3). **Summary**: Only the ηExp-$y3$ model successfully reproduces the age independence of both the O and Fe distributions up to ages of ~9 Gyr.

in terms of both slope and normalization. We have made this inference using APOGEE (Majewski et al. 2017) red giants with ages inferred from machine learning algorithms (Mackereth et al. 2019). Perhaps most importantly, we find consistent results using the Leung et al. (2023) age catalog (see Appendix A). The authors carefully demonstrated that these estimates are not biased by the chemical abundance information in the APOGEE spectra.

We highlight the age invariance of the *normalization* of the metallicity-radius trend in particular (see Fig. 3.

In terms of both normalization and slope, we measure the stellar metallicity gradient in the Galactic disk to be interestingly close to age invariant. We have made this inference from APOGEE red giants with ages inferred from machine learning algorithms (Mackereth et al. 2019). In particular, we find the same result using the Leung et al. (2023) catalog, in which the authors carefully demonstrate that the inferred ages are most sensitive to spectral features other than the emission lines from which chemical abundances are inferred (par-break

here). We highlight the age invariance of the *normalization* of the metallicity gradient in particular. This realization is contrary to expectations from "classical" chemical evolution arguments, whereby might expect abundances to generally evolve upward with time due to ongoing star formation.

Potentially start the discussion section with implications for the cosmological context? Extremely worth front-loading the notion that the equilibrium model offers a great explanation for the high abundances at high redshift seen by JWST. Maybe save it for later since I've mentioned it in the draft abstract. Start with SFR/IFR demonstration instead? Best to move that math to this subsection.

5.1. *Mass Loaded Outflows Versus Radial Gas Flows*

5.2. *Radial Migration*

5.3. *Birth Radii Inferences*

5.4. *Comparison to Literature Results*
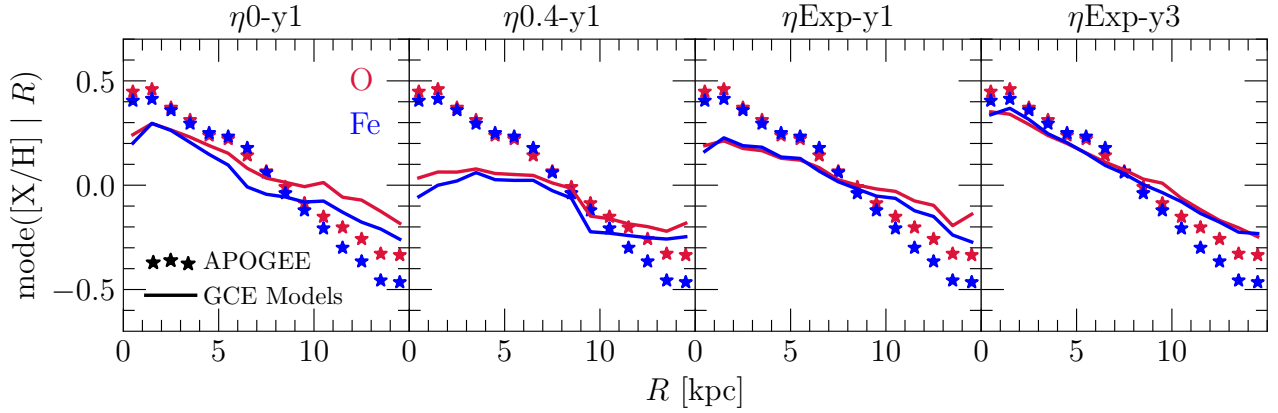
Figure 9. Comparing predicted stellar metallicity gradients with our empirical measurements. Stars denote our measurements from APOGEE in O (red) and Fe (blue) and are identical to the points in Fig. 2 (see discussion in section 2.3). Lines denote the predictions from our primary set of GCE models, marked at the top of each panel (see Table 2 and discussion in section 3). **Summary**: Since the models are calibrated to reproduce the observed gradient in the ISM, it is not guaranteed that they will also reproduce the observed stellar gradient. The ηExp-$y3$ model offers the best explanation of the consistency between gradients in gas and stars seen in Fig. 2.
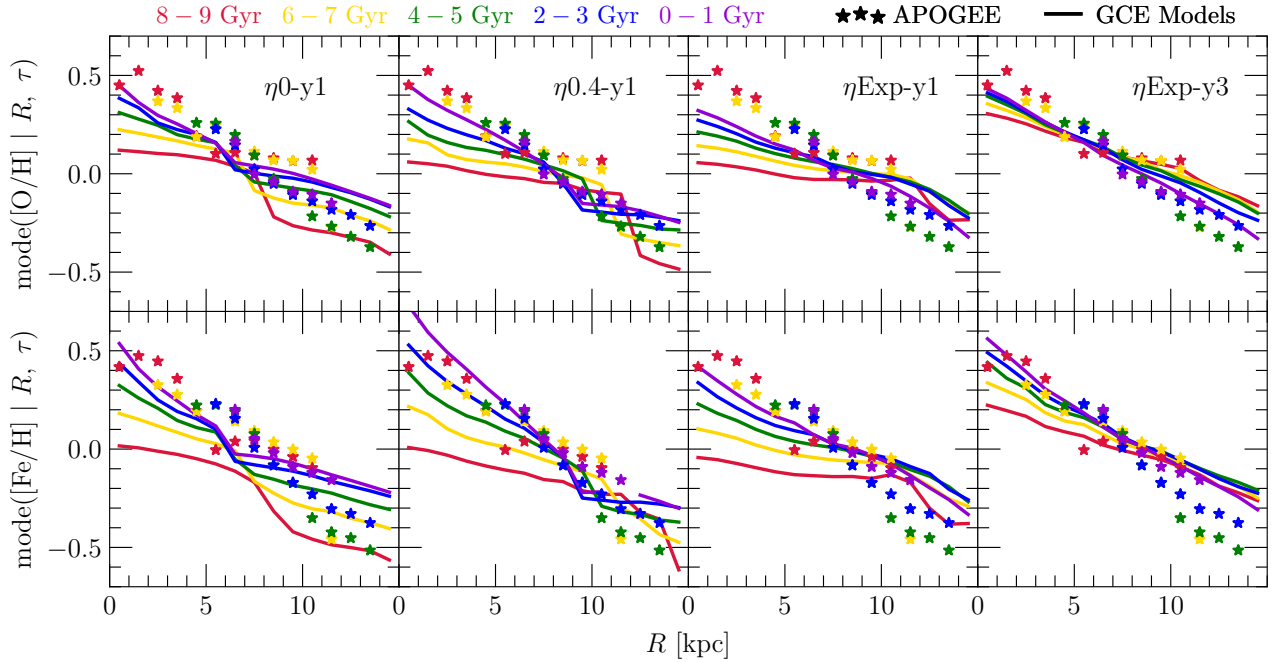
Figure 10. Comparing metallicity gradients in O (top) and Fe (bottom) for different ages of stellar populations (color coded according to the legend at the top left) between our measurements and our models. Stars mark our measurements with the AstroNN value added catalog (see discussion in section 2) and are identical to the corresponding points in Fig. 3. Lines mark the corresponding measurements from our primary set of GCE models, denoted by the text in the top row of panels (see Table 2 and discussion in section 3). **Summary**: The ηExp-$y3$ model is the best explanation of the measurements, because it reproduces the lack of evolution in the gradient with age. Other models see evolution in the gradient normalization across a broad range of disk radii.
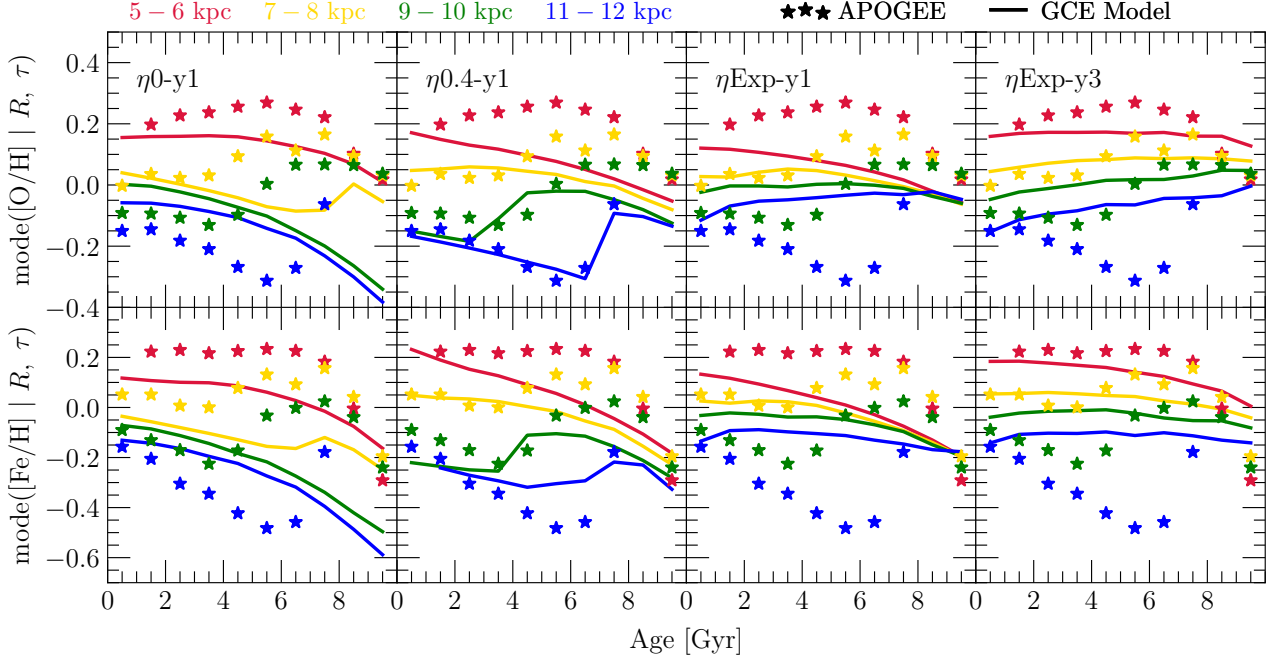
Figure 11. The same comparison as Fig. 10, but shown as a function of stellar age at fixed radius as opposed to the inverse (color coded according to the legend in the top left). **Summary**: The models that fail to reproduce our primary result, the age independence of the stellar disk metallicity gradient (see Fig. **??**), underpredict abundances (by ∼0.3 − 0.4 dex in some cases) for populations older than ∼6 Gyr.
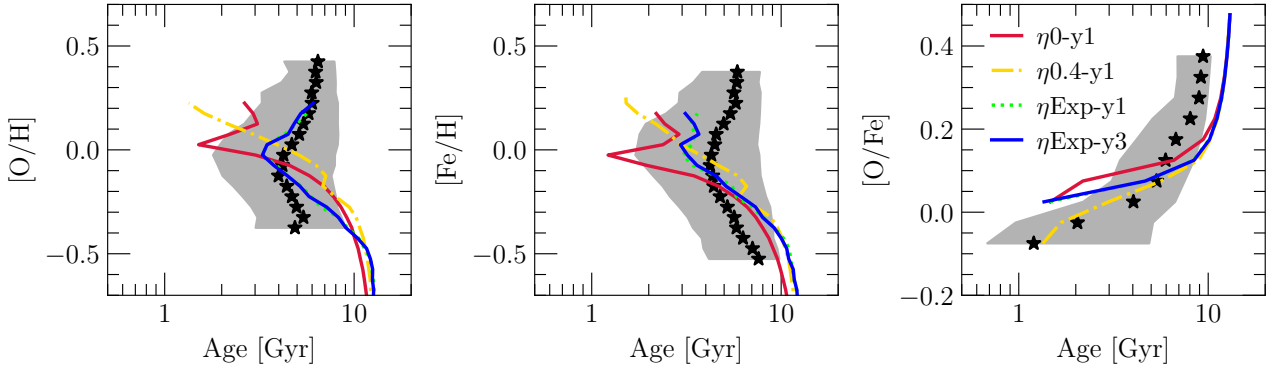
Figure 12. A comparison of the age metallicity relation for all stars in the solar annulus ($R = 7 − 9$ kpc; $|z| ≤ 0.5$ kpc) between our models and our data. Black stars denote the median age in a narrow bin of [O/H] (left), [Fe/H] (middle), or [O/Fe] (right), with the 16th and 84th percentiles of the age distribution marked by the shaded region. Lines of different colors and styles mark the GCE models in our primary set according to the legend in the right panel (see Table 2 and discussion in section 3). **Summary**: The ηExp-$y1$ and ηExp-$y3$ models better reproduce the shape of the AMR than the $\eta0$-$y1$ and $\eta0.4$-$y1$ models. This difference originates in the former predicting super solar abundances for old stellar populations where the latter fails.
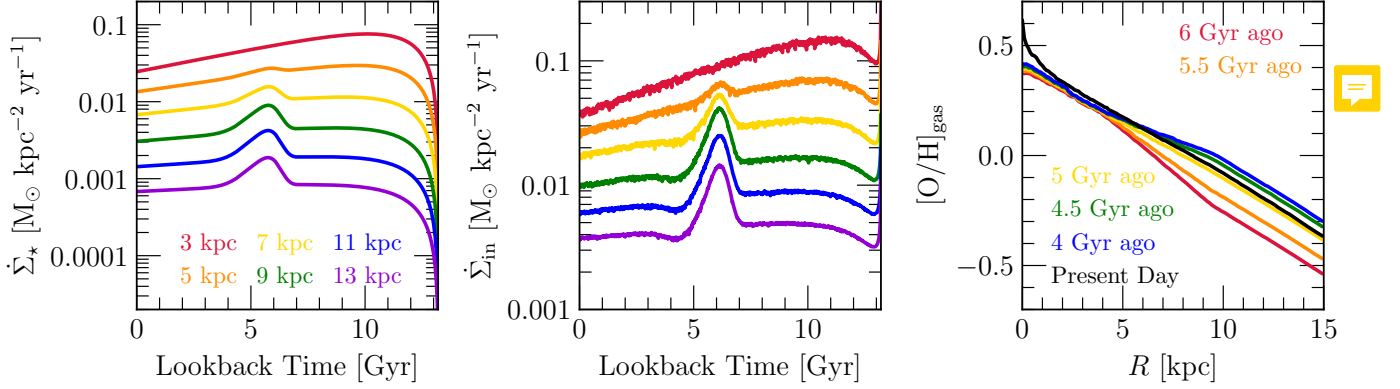
Figure 13. A variation of our ηExp-$y3$ model with an accretion induced burst of star formation in the outer disk ~6 Gyr ago. Left and middle panels show the input SFH (defined according to equation X) and the predicted infall history at six Galactocentric radii, color coded according to the legend in the left panel. The right panel shows the [O/H] gradient in the ISM at the present day (black) and at a handful of snapshots during and after the starburst (color coded according to the legend in this panel). **Summary**: The level of variation in the stellar metallicity gradient seen in ~$4-6$ Gyr old populations (see Fig. 3 and discussion in section 2.3) is broadly consistent with an accretion induced burst of star formation caused by the first passage of the Sgr dSph. This major merger is predicted to present as a perturbation of the metallicity gradient's equilibrium state that relaxes on ~Gyr timescales.

Turn this section into a generic discussion in comparison to literature results. Pull in the discussion of the equilibrium abundance from MaNGA (see Nick Boardman's SDSS MWAG talk slides). Should touch on how the realization that the AMR is flat fits into this picture and was an indication in this direction. Worth pointing out the Willett et al. (2023) paper, in which the authors claimed to have found evolution in both the slope and normalization of the gradient, but if you look closely at their Fig. 7, their actual measurements are consistent with the gradient slope and intercept being constant with age. Also worth pointing out Imig et al.'s (2023) measurements of the gradient as a function of age, which do show some flattening (see their Fig. 17?).

5.5. *Empirical Constraints on Stellar Yields*

5.6. *The Yield-Outflow Degeneracy*

5.7. *Metallicity Gradients in the Cosmological Context*

6. CONCLUSIONS

APPENDIX

A. THE LEUNG ET AL. (2023) AGE CATALOG

In this appendix, we replicate our results in Fig. 3 with the Leung et al. (2023) catalog of stellar ages. We have focused our discussion in this paper on the AsTroNN catalog (Mackereth et al. 2019), but it is possible that the neural network simply learned the correlations between chemical abundances and ages. In this case, our key empirical result in Fig. 3 would be subject to considerable systematic uncertainties. Leung et al. (2023) demonstrated that their estimates do not contain any significant amount of information on stellar abundances (see discussion in section 2.4), making their catalog an excellent comparison case to validate our measurements.

They mitigate this potential issue by compressing the spectra into lower dimensional representations of themselves (i.e., a *latent space*) using a variation encoder-decoder algorithm (e.g., LeCun et al. 2015). They then train a modified random forest algorithm to predict similarly compressed lightcurves trained on *Kepler* photometry Borucki et al. (2010). They demonstrate that this latent space contains little if any information on chemical abundances and stellar parameters, as intended, after which they are able to estimate ages by augmenting the latent space with T_{eff} and $\log g$ and decompressing the lightcurves.

The Leung et al. (2023) training sample spanned a much narrower range in surface gravity ($\log_{10} g = 2.5 - 3.6$), which lowers the sample size by a factor of ~2.5. We therefore use 2 Gyr as opposed to 1 Gyr bins in age and compute the mode of the MDFs for bins with at least 100 as opposed to 200 stars. Fig. 22 shows the results, with the lines of best fit to the metallicity gradients in each age bin reported in Table 3. As in Fig. 3, we find that both the slope and normalization of the disk metallicity gradient are independent of age up to ~10 Gyr, the full range of ages that we probe here. We can then safely conclude that our central results are not affected by any learned correlations between chemical abundances and stellar ages in the AsTroNN catalog.
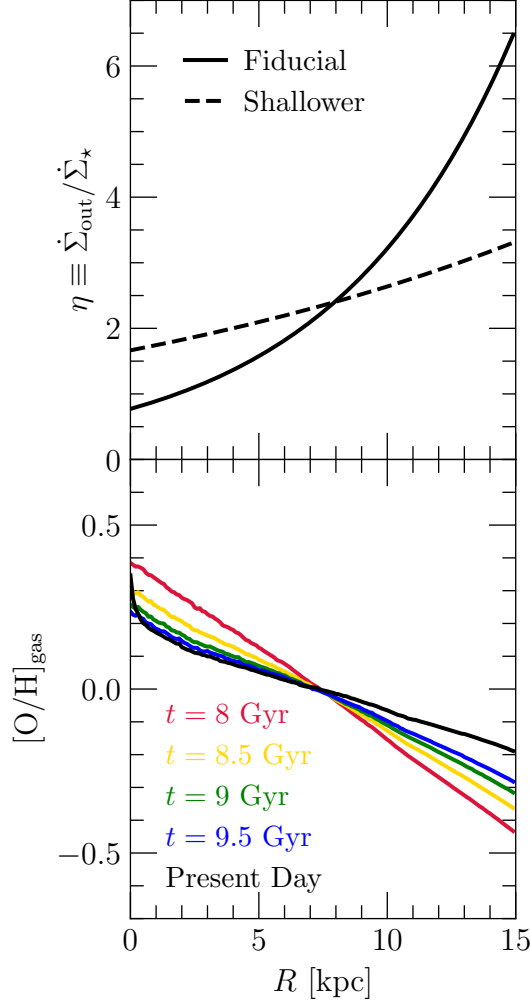
Figure 14. The impact of building a sudden change in R_η at $t = 8$ Gyr into our ηExp-$y3$ model (see Table 2 and discussion in section 3). **Top**: The mass loading factor η as a function of radius before (solid) and after (dashed) the perturbation. **Bottom**: The [O/H] abundance in the ISM at the present day (black) and at a handful of snapshots immediately following the parameter change (lines color coded according to the legend). **Summary**: Parameter changes lead to changes in the equilibrium abundance with time, which can lead to shifts in the slope of the metallicity gradient. Though we do not find any obvious signs of such events in our sample, the literature has seen some recent arguments in favor of evolution in its slope (see discussion in section 5.4).

B. GCE MODEL CALIBRATIONS

REFERENCES

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35,
    doi: 10.3847/1538-4365/ac4414

Aller, L. H. 1942, ApJ, 95, 52, doi: 10.1086/144372

Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140,
    doi: 10.1088/0004-637X/765/2/140

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009,
    ARA&A, 47, 481, doi: 10.1146/annurev.astro.46.060407.145222

Table 3. A summary of the linear regressions applied to radial age and metallicity gradients in this paper (see discussion in §X).

Age Range	Slope	Value at 8 kpc
[O/H]		
0 − 2 Gyr	$-0.051 \pm 0.007 \text{ kpc}^{-1}$	0.035 ± 0.017
2 − 4 Gyr	$-0.054 \pm 0.004 \text{ kpc}^{-1}$	0.034 ± 0.012
4 − 6 Gyr	$-0.054 \pm 0.006 \text{ kpc}^{-1}$	0.038 ± 0.014
6 − 8 Gyr	$-0.079 \pm 0.009 \text{ kpc}^{-1}$	0.119 ± 0.019
8 − 10 Gyr	$-0.034 \pm 0.010 \text{ kpc}^{-1}$	0.122 ± 0.018
[Fe/H]		
0 − 2 Gyr	$-0.067 \pm 0.004 \text{ kpc}^{-1}$	0.050 ± 0.011
2 − 4 Gyr	$-0.083 \pm 0.005 \text{ kpc}^{-1}$	-0.002 ± 0.012
4 − 6 Gyr	$-0.072 \pm 0.010 \text{ kpc}^{-1}$	0.006 ± 0.022
6 − 8 Gyr	$-0.105 \pm 0.020 \text{ kpc}^{-1}$	0.090 ± 0.042
8 − 10 Gyr	$-0.073 \pm 0.023 \text{ kpc}^{-1}$	0.114 ± 0.039
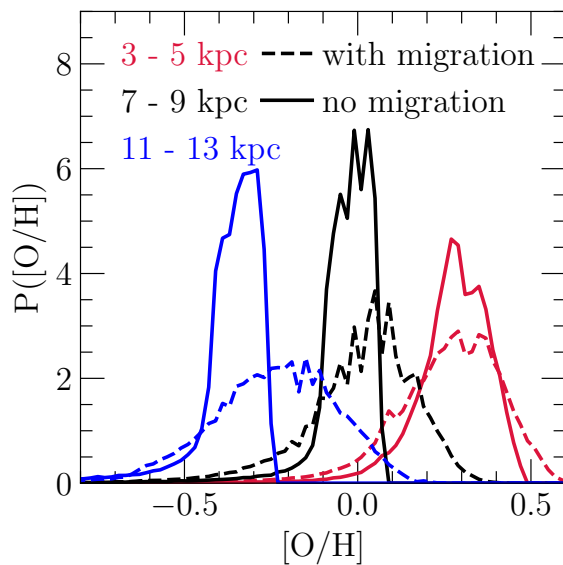
Figure 15. Distributions in [O/H] for stellar populations with (dashed) and without (solid) radial migration from our ηExp-$y3$ model (see discussion in section 5.2). Lines are color coded by their 2-kpc bin in radius according to the legend in the upper left. **Summary**: In the radial equilibrium model, stellar migration enhances the tails of the metallicity distribution but has only a minimal impact on the position of the peak, consistent with Grand et al.'s (2014) chemodynamic simulation. Y-axis label $P([\text{O/H}]) \rightarrow P([\text{O/H}] \,|\, R)$.

Beaton, R. L., Oelkers, R. J., Hayes, C. R., et al. 2021, AJ, 162, 302, doi: 10.3847/1538-3881/ac260c

Bigiel, F., Leroy, A., Walter, F., et al. 2010, AJ, 140, 1194, doi: 10.1088/0004-6256/140/5/1194

Bird, J. C., Kazantzidis, S., Weinberg, D. H., et al. 2013, ApJ, 773, 43, doi: 10.1088/0004-637X/773/1/43

Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529, doi: 10.1146/annurev-astro-081915-023441

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977, doi: 10.1126/science.1185402

Bovy, J., Rix, H.-W., Green, G. M., Schlafly, E. F., & Finkbeiner, D. P. 2016, ApJ, 818, 130, doi: 10.3847/0004-637X/818/2/130

Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127, doi: 10.1111/j.1365-2966.2012.21948.x

Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7, doi: 10.1088/0004-637X/798/1/7

Chabrier, G. 2003, PASP, 115, 763, doi: 10.1086/376392

Chaplin, W. J., & Miglio, A. 2013, ARA&A, 51, 353, doi: 10.1146/annurev-astro-082812-140938

Chartab, N., Newman, A. B., Rudie, G. C., Blanc, G. A., & Kelson, D. D. 2023, arXiv e-prints, arXiv:2310.12200, doi: 10.48550/arXiv.2310.12200

Cheng, J. Y., Rockosi, C. M., Morrison, H. L., et al. 2012, ApJ, 746, 149, doi: 10.1088/0004-637X/746/2/149

Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102, doi: 10.3847/0004-637X/823/2/102

de los Reyes, M. A. C., & Kennicutt, Robert C., J. 2019, ApJ, 872, 16, doi: 10.3847/1538-4357/aafa82

De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604, doi: 10.1093/mnras/stv327

Dotter, A. 2016, ApJS, 222, 8, doi: 10.3847/0067-0049/222/1/8

Eilers, A.-C., Hogg, D. W., Rix, H.-W., et al. 2022, ApJ, 928, 23, doi: 10.3847/1538-4357/ac54ad

Feuillet, D. K., Frankel, N., Lind, K., et al. 2019, MNRAS, 489, 1742, doi: 10.1093/mnras/stz2221

Feuillet, D. K., Bovy, J., Holtzman, J., et al. 2018, MNRAS, 477, 2326, doi: 10.1093/mnras/sty779

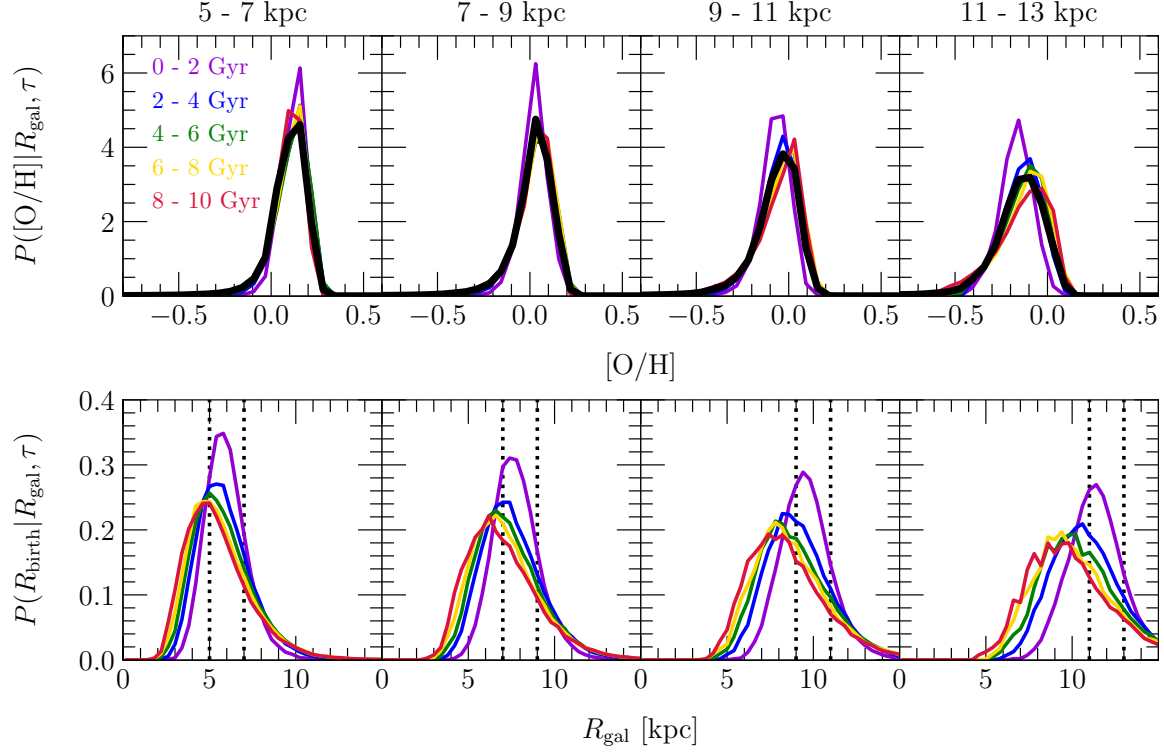Finlator, K., & Davé, R. 2008, MNRAS, 385, 2181, doi: 10.1111/j.1365-2966.2008.12991.x

Figure 16. Metallicity (top) and birth radius (bottom) distributions predicted by our ηExp-$y3$ model for populations within $|z| \leq 0.5$ kpc of the disk midplane. **Top**: Distributions in [O/H] in 2-kpc bins of Galactocentric radius, marked at the top of each panel. Lines denote mono-age populations at those radii, color coded according to the legend in the top left panel. The black lines the distribution associated with the full population. **Bottom**: Distributions in Galactic birth radius in the same bins of present day radius and age as the top panels. **Summary**: Despite old stellar populations migrating several kpc of their lifetimes, the metallicity gradient in the Galactic disk is sufficiently shallow that radial migration does not substantially shift the peak of the MDF.

Frinchaboy, P. M., Thompson, B., Jackson, K. M., et al. 2013, ApJL, 777, L1, doi: 10.1088/2041-8205/777/1/L1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1, doi: 10.1051/0004-6361/202039657

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41, doi: 10.1111/j.1365-2966.2005.09321.x

Gilmore, G., Randich, S., Asplund, M., et al. 2012, The Messenger, 147, 25

Grand, R. J. J., Kawata, D., & Cropper, M. 2014, MNRAS, 439, 623, doi: 10.1093/mnras/stt2483

Haemmerlé, L., Eggenberger, P., Ekström, S., et al. 2019, A&A, 624, A137, doi: 10.1051/0004-6361/201935051

Imig, J., Price, C., Holtzman, J. A., et al. 2023, ApJ, 954, 124, doi: 10.3847/1538-4357/ace9b8

Johnson, J. A. 2019, Science, 363, 474, doi: 10.1126/science.aau9540

Johnson, J. W., & Weinberg, D. H. 2020, MNRAS, 498, 1364, doi: 10.1093/mnras/staa2431

Johnson, J. W., Weinberg, D. H., Vincenzo, F., Bird, J. C., & Griffith, E. J. 2023, MNRAS, 520, 782, doi: 10.1093/mnras/stad057

Johnson, J. W., Weinberg, D. H., Vincenzo, F., et al. 2021, MNRAS, 508, 4484, doi: 10.1093/mnras/stab2718

Jönsson, H., Allende Prieto, C., Holtzman, J. A., et al. 2018, AJ, 156, 126, doi: 10.3847/1538-3881/aad4f5

Kauffmann, G. 1996, MNRAS, 281, 475, doi: 10.1093/mnras/281.2.475

Kennicutt, Robert C., J. 1998, ApJ, 498, 541, doi: 10.1086/305588

Kennicutt, Robert C., J., & de los Reyes, M. A. C. 2021, ApJ, 908, 61, doi: 10.3847/1538-4357/abd3a2

Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102, doi: 10.1088/0004-637X/779/2/102

Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv e-prints, arXiv:1711.03234, doi: 10.48550/arXiv.1711.03234

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

Krumholz, M. R., Burkhart, B., Forbes, J. C., & Crocker, R. M. 2018, MNRAS, 477, 2716, doi: 10.1093/mnras/sty852

Larson, R. B. 1972, Nature Physical Science, 236, 7, doi: 10.1038/physci236007a0

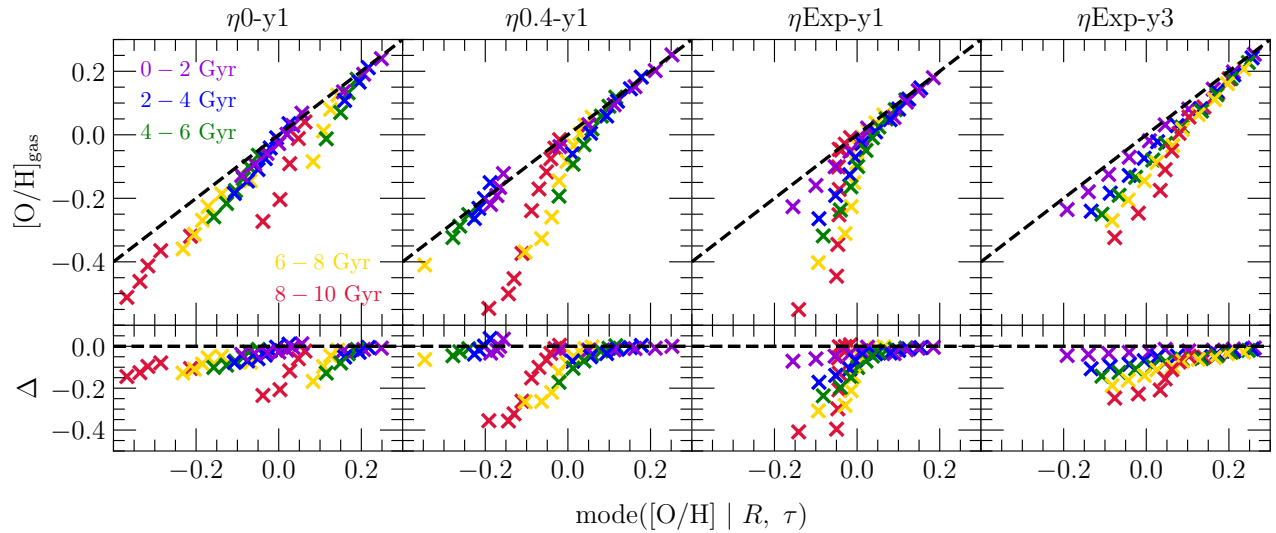LeCun, Y., Bengio, Y., & Hinton, G. 2015, Nature, 521, 436, doi: 10.1038/nature14539

Figure 17. Extending the argument of Fig. 16 to the other GCE models in our primary set (marked at the top of each pair of panels; see Table 2). Each panel compares the [O/H] abundance in the ISM with the peak of the MDF predicted in bins of Galactocentric radius and age 1 kpc and 2 Gyr wide (color coded by age according to the legend in the left panel). The ISM abundance is taken from the center of each age and radius bin. Black dashed lines denote the one-to-one relation. Bottom panels show the difference $[O/H]_{gas} -$ mode$([O/H] \mid R, \tau)$. **Summary**: The ISM abundance at a particular radius and lookback time is typically overestimated by the observed mode of the MDF of the local stellar population at the corresponding age, but only by $\sim 0.2 - 0.3$ dex (less for younger and/or more metal-rich populations).

Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, AJ, 146, 19, doi: 10.1088/0004-6256/146/2/19

Leung, H. W., & Bovy, J. 2019a, MNRAS, 483, 3255, doi: 10.1093/mnras/sty3217

—. 2019b, MNRAS, 489, 2079, doi: 10.1093/mnras/stz2245

Leung, H. W., Bovy, J., Mackereth, J. T., & Miglio, A. 2023, MNRAS, doi: 10.1093/mnras/stad1272

Lian, J., Zasowski, G., Mackereth, T., et al. 2022, MNRAS, 513, 4130, doi: 10.1093/mnras/stac1151

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119, doi: 10.1088/0004-637X/772/2/119

Lu, Y., Minchev, I., Buck, T., et al. 2022, arXiv e-prints, arXiv:2212.04515, doi: 10.48550/arXiv.2212.04515

Luo, A. L., Zhao, Y.-H., Zhao, G., et al. 2015, Research in Astronomy and Astrophysics, 15, 1095, doi: 10.1088/1674-4527/15/8/002

Mackereth, J. T., & Bovy, J. 2020, MNRAS, 492, 3631, doi: 10.1093/mnras/staa047

Mackereth, J. T., Bovy, J., Schiavon, R. P., et al. 2017, MNRAS, 471, 3057, doi: 10.1093/mnras/stx1774

Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, MNRAS, 489, 176, doi: 10.1093/mnras/stz1521

Maiolino, R., & Mannucci, F. 2019, A&A Rv, 27, 3, doi: 10.1007/s00159-018-0112-2

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94, doi: 10.3847/1538-3881/aa784d

Maoz, D., & Mannucci, F. 2012, PASA, 29, 447, doi: 10.1071/AS11052

Martell, S. L., Sharma, S., Buder, S., et al. 2017, MNRAS, 465, 3203, doi: 10.1093/mnras/stw2835

Matteucci, F. 2021, A&A Rv, 29, 5, doi: 10.1007/s00159-021-00133-8

Matteucci, F., & Francois, P. 1989, MNRAS, 239, 885, doi: 10.1093/mnras/239.3.885

Méndez-Delgado, J. E., Amayo, A., Arellano-Córdova, K. Z., et al. 2022, MNRAS, 510, 4436, doi: 10.1093/mnras/stab3782

Minchev, I., Chiappini, C., & Martig, M. 2013, A&A, 558, A9, doi: 10.1051/0004-6361/201220189

—. 2014, A&A, 572, A92, doi: 10.1051/0004-6361/201423487

Montalbán, J., Mackereth, J. T., Miglio, A., et al. 2021, Nature Astronomy, 5, 640, doi: 10.1038/s41550-021-01347-7

Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, MNRAS, 454, 2691, doi: 10.1093/mnras/stv2126

Myers, N., Donor, J., Spoo, T., et al. 2022, AJ, 164, 85, doi: 10.3847/1538-3881/ac7ce5

Peeples, M. S., & Shankar, F. 2011, MNRAS, 417, 2962, doi: 10.1111/j.1365-2966.2011.19456.x

Pinsonneault, M. H., Elsworth, Y. P., Tayar, J., et al. 2018, ApJS, 239, 32, doi: 10.3847/1538-4365/aaebfd

Rodríguez, Ó., Maoz, D., & Nakar, E. 2022, arXiv e-prints, arXiv:2209.05552, doi: 10.48550/arXiv.2209.05552

Sánchez, S. F. 2020, ARA&A, 58, 99, doi: 10.1146/annurev-astro-012120-013326

Sanders, R. L., Shapley, A. E., Jones, T., et al. 2021, ApJ, 914, 19, doi: 10.3847/1538-4357/abf4c1

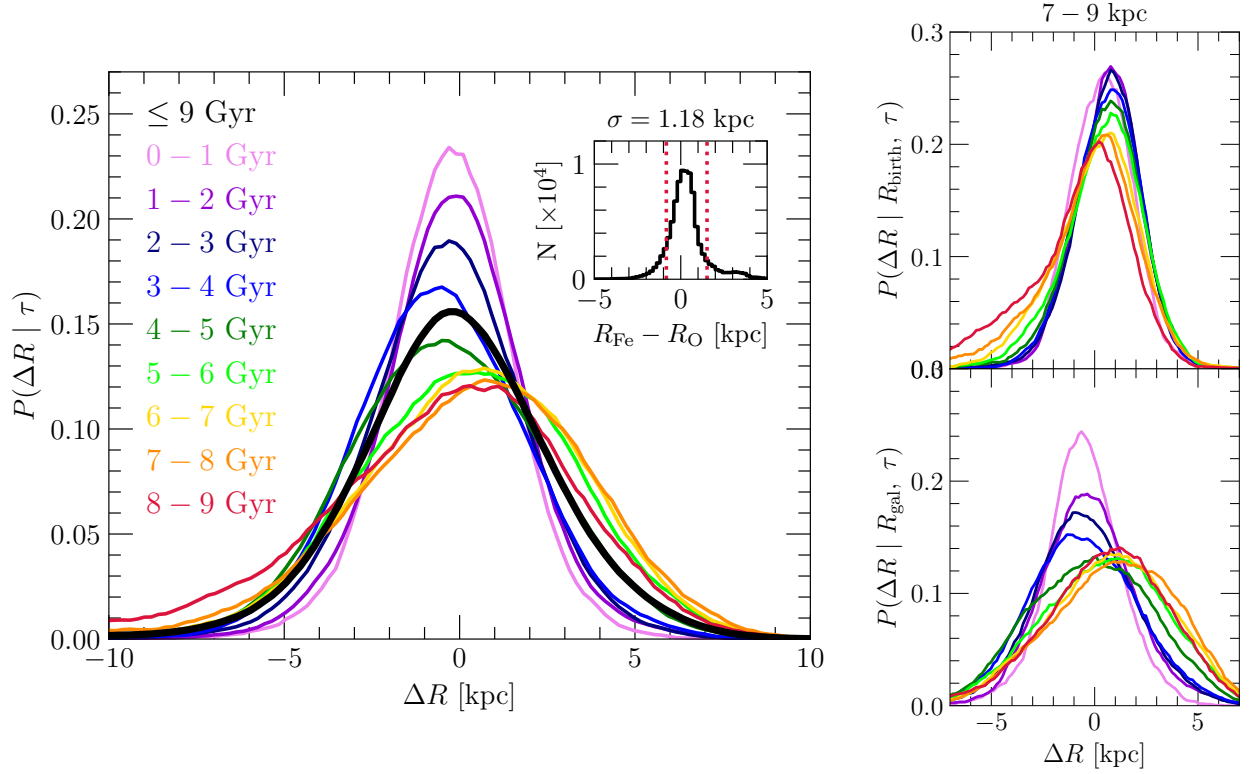**Figure 18.** Distributions in the radial displacements $\Delta R \equiv R_{gal} - R_{birth}$ of stars inferred by assuming the equilibrium model and inverting the metallicity gradient (see discussion in section 5.3). All distributions are box-car smoothed with a window of width ± 1 kpc. **Left**: The distribution in ΔR for stars at all radii, color coded by age according to the legend. The inset shows the distribution of differences in birth radii inferred by switching between [O/H] and [Fe/H] gradients (we use [O/H] here). **Top-right**: The same as the left panel, but for stars inferred to have been born in the solar annulus. **Bottom-right**: The same as the left panel, but for stars in the solar annulus at the present day. **Summary**: When inferring birth radii from abundances alone, we recover known qualitative trends in radial displacement and age, consistent with the equilibrium model.

Santana, F. A., Beaton, R. L., Covey, K. R., et al. 2021, AJ, 162, 303, doi: 10.3847/1538-3881/ac2cbc

Schönrich, R., & Binney, J. 2009, MNRAS, 396, 203, doi: 10.1111/j.1365-2966.2009.14750.x

Searle, L. 1971, ApJ, 168, 327, doi: 10.1086/151090

Shaver, P. A., McGee, R. X., Newton, L. M., Danks, A. C., & Pottasch, S. R. 1983, MNRAS, 204, 53, doi: 10.1093/mnras/204.1.53

Simon, J. D. 2019, ARA&A, 57, 375, doi: 10.1146/annurev-astro-091918-104453

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708

Soderblom, D. R. 2010, ARA&A, 48, 581, doi: 10.1146/annurev-astro-081309-130806

Stone-Martinez, A., Holtzman, J. A., Imig, J., et al. 2023, arXiv e-prints, arXiv:2311.17887, doi: 10.48550/arXiv.2311.17887

Tinsley, B. M. 1980, FCPh, 5, 287, doi: 10.48550/arXiv.2203.02041

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898, doi: 10.1086/423264

Weinberg, D. H., Andrews, B. H., & Freudenburg, J. 2017, ApJ, 837, 183, doi: 10.3847/1538-4357/837/2/183

Weinberg, D. H., Griffith, E. J., Johnson, J. W., & Thompson, T. A. 2023, arXiv e-prints, arXiv:2309.05719, doi: 10.48550/arXiv.2309.05719

Wenger, T. V., Balser, D. S., Anderson, L. D., & Bania, T. M. 2019, ApJ, 887, 114, doi: 10.3847/1538-4357/ab53d3

White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52, doi: 10.1086/170483

Willett, E., Miglio, A., Mackereth, J. T., et al. 2023, MNRAS, 526, 2141, doi: 10.1093/mnras/stad2374

Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001, doi: 10.1088/1538-3873/ab0075

Wyse, R. F. G., & Silk, J. 1989, ApJ, 339, 700, doi: 10.1086/167329

Zahid, H. J., Bresolin, F., Kewley, L. J., Coil, A. L., & Davé, R. 2012, ApJ, 750, 120, doi: 10.1088/0004-637X/750/2/120

Zaritsky, D. 1992, ApJL, 390, L73, doi: 10.1086/186375

Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, AJ, 146, 81, doi: 10.1088/0004-6256/146/4/81
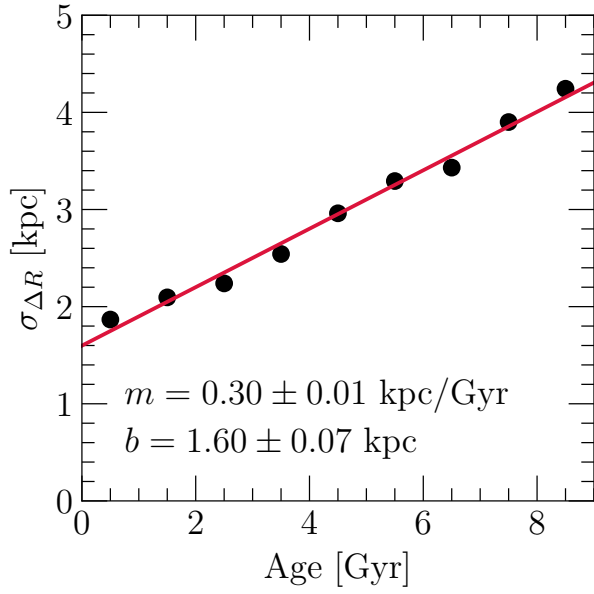
Figure 19. The inferred strength of stellar migration as a function of population age. Black points mark the standard deviation of the radial displacement distribution $P(\Delta R|\tau)$ in the top panel of Fig. 18 as a function of τ. The line of best fit $y = mx + b$ is marked in red. **Summery**: The non-zery y-intercept $b = 1.60 \pm 0.07$ kpc likely reflects the maximum achievable precision with this methodology due to systematic uncertainties, including non-axisymmetric variations in the ISM metallicity (Wenger et al. 2019) and simply the choice of element (see discussion in section 5.3).
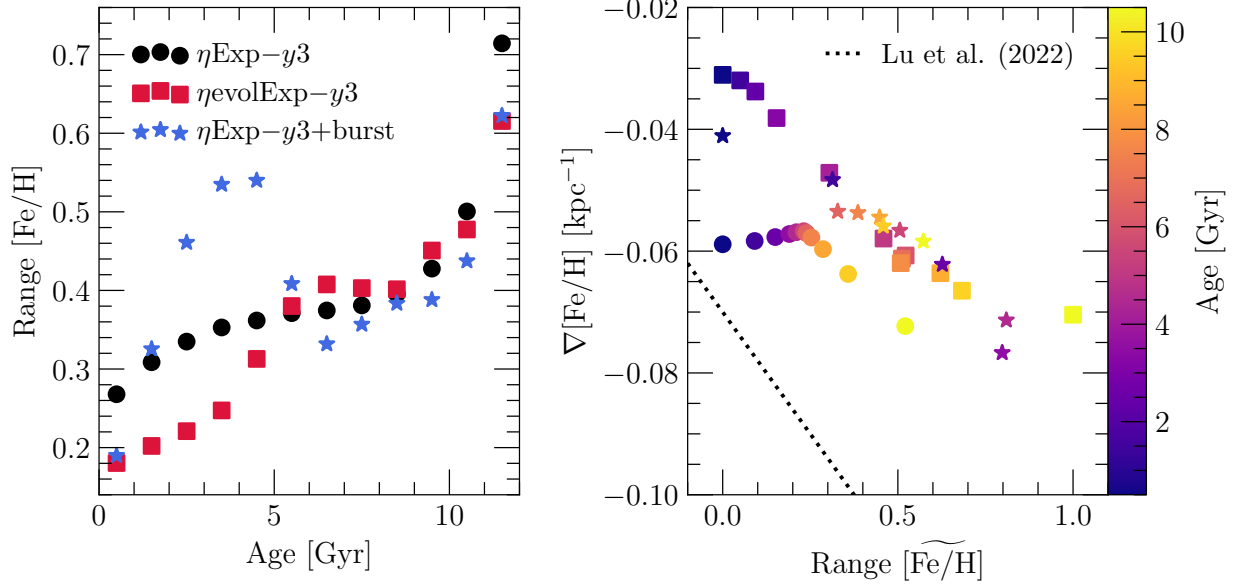
Figure 20. Comparing our GCE models to Lu et al.'s (2022) arguments by applying their methodology to the model predictions (see discussion in section 5.4 and 3). **Left**: The range of [Fe/H] as a function of stellar population age at $R = 7 - 10$ kpc. **Right**: The inferred slope of the abundance gradient ∇[Fe/H] as a function of the range of [Fe/H], normalized to the [0, 1] range. Motivated by the predictions of hydrodynamic simulations, they adopt the dotted black line to relate the normalized ranges to the gradient slope; the offset in normalization originates in the gradients they measure being steeper than what we find here. **Summary**: All three of the ηExp-$y3$, ηevolExp-$y3$, and ηExp-$y3$+burst models predict significantly different relationships between Range([Fe/H]) and Age. Interestingly, the ηExp-$y3$ does not predict a significantly sloped relation between ∇[Fe/H] and Range([Fe/H]) while the others do.
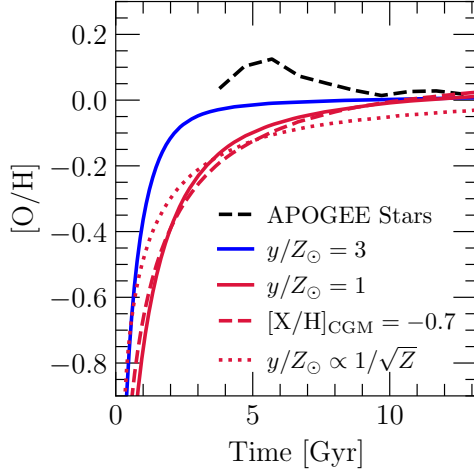
Figure 21. Variations in the ηExp-y1 model (all shown in red) intended to quicken its approach to equilibrium. Similar to the left panel of Fig. 7, all curves show the ISM O abundance at $R = 8$ kpc as a function of time. The dashed line marks a model in which the metallicity of accreting material is initially metal-free but reaches $[\text{O/H}]_{\text{CGM}} \approx [\text{Fe/H}]_{\text{CGM}} = -0.7$ on 1 Gyr timescales (see equation X). The dotted line shows a case in which we let all SN yields increase with a $1/\sqrt{Z}$ metallicity dependence, based on recent arguments about the Z-dependence of SN event rates (see discussion in section 5.6). The blue solid and black dashed line are the same as in Fig. 7.
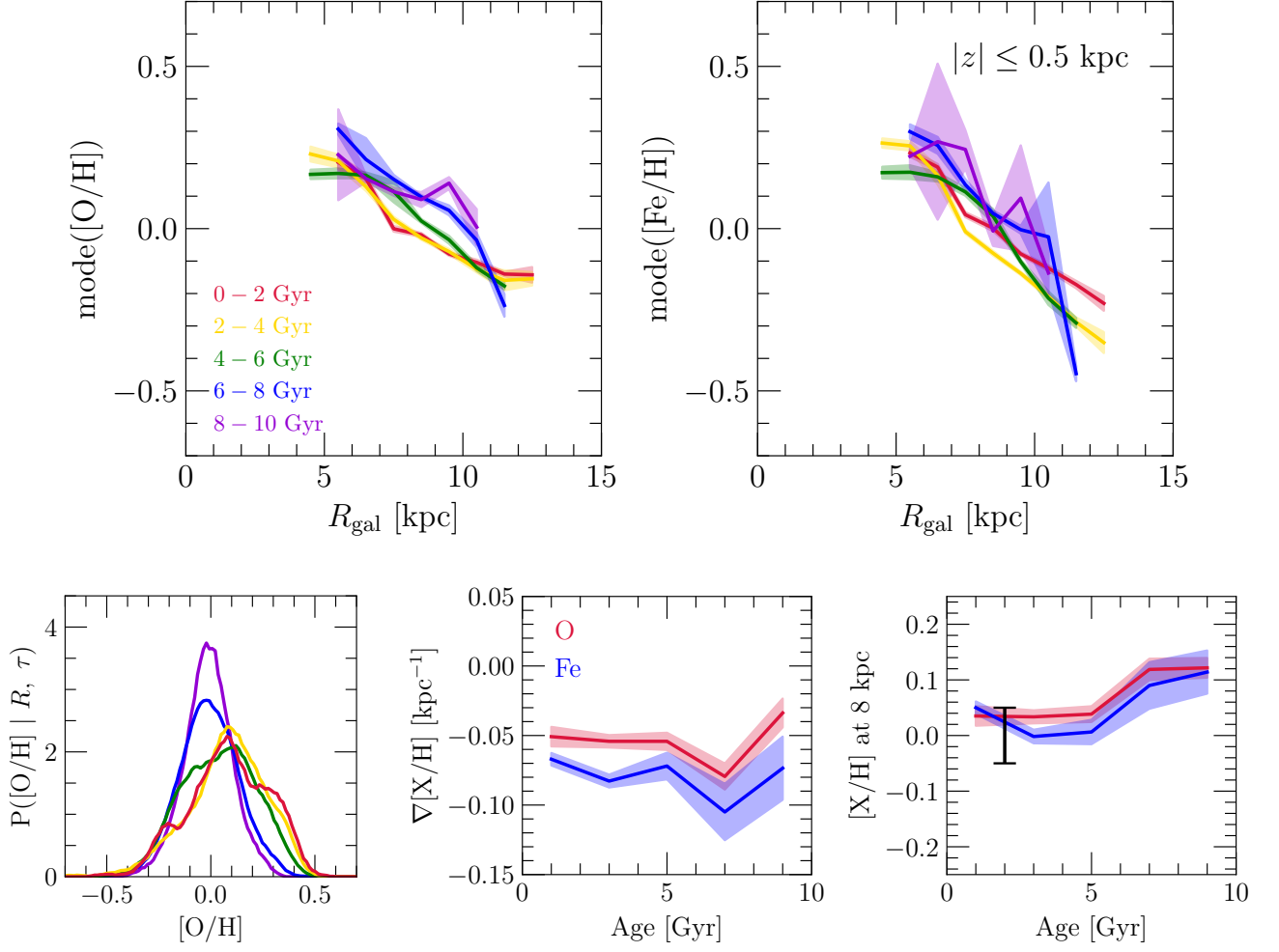
Figure 22. The same as Fig. 3, but using Leung et al.'s (2023) catalog of stellar ages as opposed to the AstroNN value added catalog. Because of the smaller sample, we use 2 Gyr bins in age as opposed to 1 Gyr according to the legend in the top left panel. **Summary**: We find the same results as in Fig. 3, so our central results should not be affected by any learned correlations between stellar abundances and ages in the AstroNN catalog.
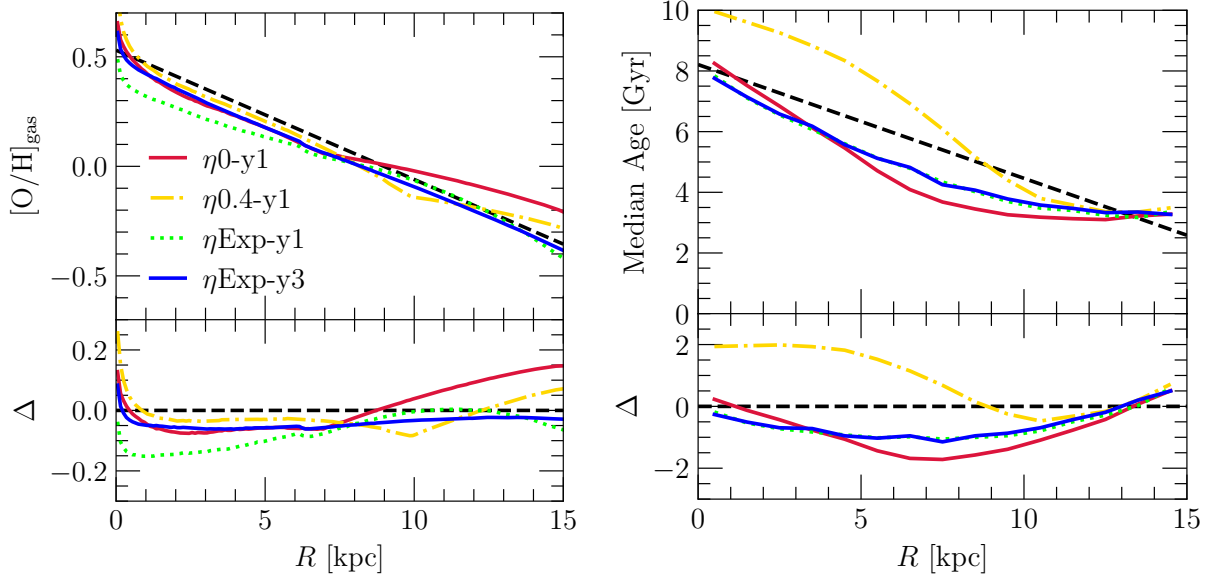
Figure 23. Calibration of our primary set of GCE models (see Table 2 and discussion in section 3 and Appendix B). **Left**: The predicted [O/H] gradients in the ISM at the present day (colored and styled lines marked by the legend). The black dashed line marks Méndez-Delgado et al.'s (2022) measurements in HII regions. The bottom panel shows the differences between the predicted and observed abundances as a function of radius. **Right**: The same as the left panel, but for the stellar age gradient. The black dashed line marks our linear regression in the bottom left panel of Fig. 2 (see fit parameters in Table 1). **Summary**: All models are calibrated to reproduce the observed ISM abundances at the present day, which they achieve sufficiently accurately for our purposes. Only the models invoking $\eta \propto e^R$ are calibrated to reproduce the observed age gradient, but those that are not still reasonably reproduce it.