The Astrophysical Journal, 900:28 (20pp), 2020 September 1 https://doi.org/10.3847/1538-4357/aba43b
© 2020. The American Astronomical Society. All rights reserved.

MINESweeper: Spectrophotometric Modeling of Stars in the Gaia Era

Phillip A. Cargile[1], Charlie Conroy[1], Benjamin D. Johnson[1], Yuan-Sen Ting (丁源森)[2,3,4,5,6], Ana Bonaca[1], Aaron Dotter[1], and Joshua S. Speagle[1]
[1] Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA 02138, USA; pcargile@cfa.harvard.edu
[2] Institute for Advanced Study, Princeton, NJ 08540, USA
[3] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[4] Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[5] Research School of Astronomy & Astrophysics, Australian National University, Cotter Road, Weston, ACT 2611, Australia
Received 2019 August 29; revised 2020 May 15; accepted 2020 July 6; published 2020 August 27

Abstract

We present `MINESweeper`, a tool to measure stellar parameters by jointly fitting observed spectra and broadband photometry to model isochrones and spectral libraries. This approach enables the measurement of spectrophotometric distances, in addition to stellar parameters such as $T_{\rm eff}$, $\log g$, [Fe/H], [α/Fe], and radial velocity. `MINESweeper` employs a Bayesian framework and can easily incorporate a variety of priors, including Gaia parallaxes. Mock data are fit in order to demonstrate how the precision of derived parameters depends on evolutionary phase and signal-to-noise ratio. We then fit a selection of data in order to validate the model outputs. Fits to a variety of benchmark stars including Procyon, Arcturus, and the Sun result in derived stellar parameters that are in good agreement with the literature. We then fit combined spectra and photometry of stars in the open and globular clusters M92, M13, M3, M107, M71, and M67. Derived distances, [Fe/H], [α/Fe], and $\log g - T_{\rm eff}$ relations are in overall good agreement with literature values, although there are trends between metallicity and $\log g$ within clusters that point to systematic uncertainties at the $\approx$0.1 dex level. Finally, we fit a large sample of stars from the H3 Spectroscopic Survey in which high-quality Gaia parallaxes are also available. These stars are fit *without* the Gaia parallaxes so that the geometric parallaxes can serve as an independent test of the spectrophotometric distances. Comparison between the two reveals good agreement within their formal uncertainties after accounting for the Gaia zero-point uncertainties.

Unified Astronomy Thesaurus concepts: Stellar physics (1621); Stellar properties (1624); Stellar spectral lines (1630); Stellar effective temperatures (1597); Stellar colors (1590); Stellar ages (1581); Stellar luminosities (1609); Stellar masses (1614); Stellar radii (1626); Stellar abundances (1577); Stellar distance (1595); Stellar evolutionary models (2046)

1. Introduction

The field of Galactic astronomy is undergoing a dramatic change driven by multiple large-scale ground-based photometric and spectroscopic surveys and the space-based Gaia mission (Gaia Collaboration et al. 2016). Together, these observational efforts promise to provide new, much sharper maps of the stellar components of the Galaxy. In turn, these maps should deliver new insight into the formation and assembly history of the Galaxy, its dynamical state, and possibly the nature of dark matter.

In order to make progress, the raw observations must be converted into physical quantities. These include stellar properties such as $T_{\rm eff}$, $\log g$, [Fe/H], [α/Fe], mass, and age, as well as 3D positions and velocities. In previous work the measurement of these quantities has been performed in separate steps. For example, different methods for analysis of stellar spectra can deliver, with varying degrees of reliability, estimates of $T_{\rm eff}$, $\log g$, radial velocities, and abundances (e.g., Smiljanic et al. 2014; Jofré et al. 2019). In order to estimate masses, ages, and distances, one often combines these derived stellar parameters with model stellar isochrones (e.g., Pont & Eyer 2004; Jørgensen & Lindegren 2005; Takeda et al. 2007; Jurić et al. 2008; Breddels et al. 2010; Burnett & Binney 2010; Binney et al. 2014; Xue et al. 2014; Wang et al. 2016; Mints & Hekker 2017, 2018; Anders et al. 2019). In essence, this approach translates stellar properties that can be measured independent of distance and age, such as $T_{\rm eff}$, $\log g$, and [Fe/H], into a predicted age and absolute luminosity. Comparison to the observed flux enables a measurement of distance (an estimate of the extinction must also be available).

For relatively bright and nearby sources, Gaia DR2 now provides very precise parallaxes. For fainter and more distant sources the Gaia parallax becomes quite uncertain, and hence distances are also uncertain and depend on various subtleties such as the method of parallax inversion and assumed priors (e.g., Bailer-Jones et al. 2018; Luri et al. 2018). Therefore, even in the Gaia era, there is a need for complementary distance constraints provided by stellar isochrones (e.g., Mints & Hekker 2018; Sanders & Das 2018) or empirically calibrated data-driven models (e.g., Leistedt & Hogg 2017; Anderson et al. 2018). Furthermore, even with precise parallaxes, model isochrones are still necessary for estimating masses and ages.

A limitation of existing isochrone-based methods is the reliance on stellar parameters that have been derived elsewhere (e.g., from a separate spectroscopic pipeline). This presents several challenges. It is difficult to accurately propagate the uncertainties in the stellar parameters to the isochrone-based parameters (unless the full covariance matrices are preserved). Moreover, there is no guarantee that the derived stellar parameters will coincide with any isochrone. In fact, owing to various uncertainties associated with deriving stellar parameters, there are many cases in which unphysical $T_{\rm eff}$—$\log g$

[6] Hubble Fellow.

relations along the main sequence (MS) have been derived from fitting spectroscopic data (e.g., Valenti & Fischer 2005; Adibekyan et al. 2012; Holtzman et al. 2015). Consequently, deriving isochrone-based properties in such cases would incur large systematic uncertainties.

Schönrich & Bergemann (2014) proposed a unified framework in which isochrone-based stellar parameters are derived directly from the available data, including spectra, photometry, and parallaxes, as well as prior information on other quantities where available. They argue that this approach is optimal in that all the available data are brought to bear on the problem, resulting in the most robust available parameter constraints. In this paper we present MINESweeper, a new tool that follows the same general philosophy of Schönrich & Bergemann (2014). Specifically, the program incorporates a variety of observational constraints, including Gaia parallaxes, optical spectra, and broadband photometry, in order to estimate stellar parameters and distances.

2. Model

MINESweeper is a program to jointly model observed stellar spectra and photometry using stellar isochrones. Brief descriptions of early versions of this code can be found in several previous studies where we have determined stellar parameters (e.g., Dotter et al. 2017; Rodriguez et al. 2017). In this section we provide a detailed overview of the basic ingredients and our approach to the computation of the likelihood and priors.

2.1. Generative Model for Stellar Spectra and Photometry

Our model consists of two main parts, (1) a stellar evolution model that predicts the physical parameters ($T_{\rm eff}$, $\log g$, [Fe/H], etc.) of a star given its mass, evolutionary state, and initial composition, and (2) a stellar spectral model that takes these parameters and, along with a model for interstellar reddening, predicts the observable spectrum and photometry from this star. These two components are described in the sections below.

2.1.1. MIST Stellar Evolution Models

We use the MIST stellar evolution models, full details of which can be found in Choi et al. (2016) and Dotter et al. (in preparation). MIST uses the stellar evolution code MESA (Paxton et al. 2011) to compute evolutionary sequences in 1D for stars in the mass range of 0.1–300 $M_{\odot}$. Stars are evolved from pre-MS contraction either through the end of carbon burning or until the white dwarf cooling stage. A variety of free parameters, including those controlling mass-loss rates and convection, are tuned to match observational constraints. MIST models distinguish between the initial composition of the star and its current surface composition. These two can differ owing to diffusion and mixing near the surface layers (e.g., Dotter et al. 2017). We refer to the initial composition with an i subscript, e.g., $[{\rm Fe/H}]_i$.

In this work we use MIST v2.0 rotating models (Dotter et al., in preparation). A key improvement in v2.0 compared to previous models is the inclusion of nonsolar abundance ratios (i.e., variable $[\alpha/{\rm Fe}]_i$). In addition, the treatment of the surface boundary conditions is significantly improved, which results in better agreement with the red giant branch (RGB) temperature scale.

We incorporate the MIST v2.0 models into MINESweeper using an N-dimensional linear interpolator in order to predict the effective temperature, surface gravity, surface composition, and bolometric luminosity of stars. Adopting the recommendation given in Dotter (2016), we have chosen to interpolate the MIST models on equivalent evolutionary points (EEP), initial mass, and initial composition. EEP is monotonically related to the age, but it is defined in such a way that evolutionary phases in which there are rapid changes in either surface or interior stellar properties are adequately captured in the isochrone tables.

The program samples in initial composition ($[{\rm Fe/H}]_i$ and $[\alpha/{\rm Fe}]_i$), initial mass M_i, and evolutionary stage (EEP). The specification of these parameters makes a unique prediction for the spectrum and photometry, which we describe next.

2.1.2. Model Spectra and Photometry

A given EEP, initial mass, and composition fully specify the stellar surface properties, including $T_{\rm eff}$, $\log g$, [Fe/H], and $[\alpha/{\rm Fe}]$. Model spectra and photometry are assigned to these parameters using our custom grid of stellar spectral models. The synthetic spectra are calculated using the 1D LTE plane-parallel atmosphere and radiative transfer codes ATLAS12 and SYNTHE maintained by R. Kurucz (Kurucz 1970; Kurucz & Avrett 1981; Kurucz 1993). The line list used in the radiative transfer calculation was provided by R. Kurucz and has been empirically tuned to the observed, ultra−high-resolution spectra of the Sun and Arcturus. We have adopted a constant microturbulence of $v_{\rm micro} = 1\,{\rm km\,s^{-1}}$. We have computed models with variable α-abundance by scaling the α-elements (O, Ne, Mg, Si, S, Ar, Ca, and Ti) in lockstep. The grid is computed with $\Delta\log g = 0.5$ over the range $[-1, 5]$, $\Delta[{\rm Fe/H}] = 0.25$ over the range $[-4, +0.5]$, $\Delta[\alpha/{\rm Fe}] = 0.2$ over the range $[-0.2, +0.6]$, and an irregular (approximately log-spaced) $T_{\rm eff}$ grid with 40 points between 3500 and 15,000K. The synthetic spectra in our custom grid we computed at a resolving power of $R = 1{,}000{,}000$, and they were subsequently smoothed to a resolution of $R = 42{,}000$ and rebinned to critical sampling before building the generative model. We note that the MINESweeper code is capable of including different wavelength ranges and resolutions; investigating the effect of these changes on inferred stellar parameters is left to future work.

Photometry is computed from each synthetic spectrum for a large set of photometric systems. Details of the specific filter curves and zero-points can be found in Choi et al. (2016). Additionally, we have updated the filter curves and zero-points for Gaia G, BP, and RP with the latest DR2 photometric calibration (Evans et al. 2018; Maíz Apellániz & Weiler 2018). We also include extinction in our model, using the extinction curve of Cardelli et al. (1989), characterized by $A_{\rm V}$ with $R_{\rm V} = 3.1$.

2.1.3. Neural Networks for Rapid Prediction of Stellar Spectra and Photometry

In order to determine the spectrum and photometry for any set of interpolated MIST stellar labels, we need to be able to quickly evaluate the spectral models at arbitrary $T_{\rm eff}$, $\log g$, and [Fe/H]. As explained in Ting et al. (2019), the use of artificial neural networks to predict synthetic spectra (and in our case also photometry) has advantages over other interpolation methods like quadratic functions or Gaussian processes.

Namely, neural networks are very flexible and can easily handle nonlinear behavior of spectra, and they can be evaluated extremely rapidly, which is important in Monte Carlo−like sampling techniques.

Here, we take a similar approach to The Payne as described in Ting et al. (2019), to which we direct the interested reader for more details. The main difference is that instead of emulating the flux variation of individual pixels, here we emulate the model spectrum as a whole. Emulating the whole spectrum allows one to extract correlations between adjacent pixels, which generally leads to better emulations. This change is also reflected in the latest version of The Payne[7]. We train a multilayered feed-forward artificial neural network on the grid of normalized synthetic spectra. We adopt a training set of 18,000 model spectra and hold out 2000 spectra as the validation set. The network we use in MINESweeper has four layers (two hidden layers) with 300 neurons per layer and rectified linear unit (ReLU) activation functions. We employ an adaptive cross-validation procedure to optimize the weights and biases (i.e., the coefficients for the individual ReLU activation functions) of the neural network, where we investigate the training at regular epochs, and through standard validation techniques (e.g., holdout method, leave-one-out method, etc.). We adopt a batch size of 512 and train for $\approx 10^6$ epochs. In a similar way, we train a multilayered feed-forward artificial neural network on synthetic photometry derived from our same synthetic spectral grid. A three-layer (one hidden layer) network with 128 neurons per layer is trained on 10,000 model photometric data sets for all of the bands of interest (e.g., Pan-STARRS *ugriz*, Gaia DR2, Two Micron All Sky Survey [2MASS], etc). We also use a multiepoch training schema and cross-validation techniques to validate the accuracy of these networks.

Our final validation of the MINESweeper neural network results in predicted spectral and photometric fluxes that agreed with a "held-out" test model data set to a median residual of $\leqslant$0.01–0.1% over all wavelength pixels and photometric bands, and over $T_{\rm eff} = 3500$–15000 K, $\log g = -1$ to 5, $[{\rm Fe/H}] = -4$ to $+0.5$, $[\alpha/{\rm Fe}] = -0.2$ to $+0.6$, and $A_{\rm V} = 0$–5. The code is written in Python using the machine-learning package PyTorch (Paszke et al. 2019) and takes ∼24 hr to train the full network on modern CPUs with GPU acceleration.

Once the neural network is trained, it is used within MINESweeper to provide spectra and photometry at arbitrary locations in the grid.

2.1.4. Continuum Normalization

Accurate flux calibration of observed spectra is very challenging. Since we are simultaneously able to fit the broadband photometry in MINESweeper, which captures the overall shape of the stellar spectral energy distribution (SED), we can forgo the information content contained in the broadband shape of the spectrum. In particular, we work with continuum-normalized model spectra, and therefore we must also continuum-normalize the observed spectra. In the example cases presented below we have performed an initial preparation of the data by using twilight observations in order to approximately continuum-normalize the observations (this procedure takes into account wavelength-dependent throughput variations). In addition, during the fitting process we multiply the observed spectrum by a fourth-order Chebyshev polynomial whose coefficients, T_n, are free parameters.

[7] https://github.com/tingyuansen/The_Payne

2.2. Fitting Spectrophotometric Data

In MINESweeper we use a Bayesian framework to infer the physical properties of a star given a set of input data. Using Bayes's theorem, we sample the posterior distribution by evaluating the likelihood of observing the stellar spectrum and/or photometric data given a set of model parameters (known as the likelihood probability distribution, or just likelihood), the probability of our model (the prior), and normalizing these probabilities over all likelihood space (the evidence). Calculating the evidence is necessary when performing model comparison, as well as formally weighting individual posterior distributions in hierarchical modeling. In the following sections, we describe the calculation of the combined spectroscopic and photometric likelihood probability, a general description of typical priors, and the method we use to estimate the posterior probability and infer stellar properties.

2.2.1. Likelihood

The likelihood probability function within MINESweeper is computed by combining the individual likelihoods of an observed spectrum and/or measured photometric magnitudes given our model. This procedure starts with a given set of EEP, $[{\rm Fe/H}]_i$, $[\alpha/{\rm Fe}]_i$, and M_i. Atmospheric parameters ($T_{\rm eff}$, $\log g$, $[{\rm Fe/H}]$, $[\alpha/{\rm Fe}]$) are then predicted for the appropriate MIST model. Spectral and photometric models are computed for these parameters using our trained neural network. The synthetic spectrum is then convolved with a broadening function that is composed of two kernels: an instrument line profile characterized by a resolving power (R), and broadening from the star itself ($v_{\rm stellar}$), which includes the stellar projected rotation velocity and large-scale macroturbulent velocity. We then apply the continuum normalization to the model and calculate the likelihood for observed spectrum and photometry.

The likelihood for a spectrum with flux (f_o) over a wavelength range (λ) given a predicted flux (f_m) from our model is

$$\mathcal{L}_{\rm spec} = \prod_{\lambda} \frac{1}{\sqrt{2\pi}\,\sigma_\lambda} \exp - \frac{(f_{\lambda,o} - f_{\lambda,m})^2}{2\sigma_\lambda^2}, \quad (1)$$

where σ_λ is the measured flux error in the observed spectrum. In our procedure the model is always interpolated onto the observed wavelength grid.

Similarly, the likelihood of a set of observed photometry given predicted synthetic photometry is

$$\mathcal{L}_{\rm phot} = \prod_{i} \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp - \frac{(m_{i,o} - m_{i,m})^2}{2\sigma_i^2}, \quad (2)$$

where the product is over all filters i, with observed and model magnitudes m_o and m_m, respectively, and measured photometric errors σ_i. We could also have calculated this likelihood probability using observed fluxes instead of magnitudes. However, we choose to work in magnitudes, as the dominant source of error for the majority of the bright targets we will fit in later sections are the photometric zero-points, which are approximately Gaussian in magnitude and not in flux.

Multiplying these two likelihood probabilities together gives us the total likelihood probability of observing a spectrum and

Table 1
Summary of Parameters and Priors Used in MINESweeper

Parameter	Prior Functions		
	Mocks	Clusters	H3
Sampled Parameters			
Mass ($M_\odot$)	Kroupa IMF	Kroupa IMF	Kroupa IMF
EEP	U(200,808)	U(200,808)	U(200,808)
$[Fe/H]_i$	U(−4.0, +0.5)	U(−4.0, +0.5)	U(−4.0, +0.5)
$[\alpha/Fe]_i$	U(−0.2, +0.6)	U(−0.2, +0.6)	U(−0.2, +0.6)
Distance (kpc)	U(1, 50)	U(0, 50)	Galactic model[a]
v_{rad}[b] (km s^{-1})	U($v_{rad,init}$−25, $v_{rad,init}$+25)	U($v_{rad,init}$ −25, $v_{rad,init}$ +25)	U($v_{rad,init}$ −25, $v_{rad,init}$ +25)
$v_{stellar}$ (km s^{-1})	U(0, 15) × S($v_{stellar}$)[c]	U(0, 15) × S($v_{stellar}$)	U(0, 15) × S($v_{stellar}$)
T_n	…[d]	N(1,[0.5, 0.1, 0.05, 0.05])	N(1,[0.5, 0.1, 0.05, 0.05])
A_V	N(0.1,0.015)	U(0,1.5)	N($A_{V,SFD}$,0.15 × $A_{V,SFD}$)
Derived Parameters			
Age (Gyr)	U(4, 14)	M67: U(0.5, 8), others: U(6, 14)	Galactic model[a]
π (mas)	N(π,σ_π)	N($\pi_{GaiaDR2}$, $\sigma_{\pi_{Gaia\ DR2}}$)	N($\pi_{Gaia\ DR2}$,$\sigma_{\pi_{Gaia\ DR2}}$)

Notes. U(x,y) indicates a uniform prior from x to y, while N(μ,σ) indicates a normal prior.
[a] See text for a description of our Galactic model.
[b] $v_{rad,init}$ is an initial guess on the radial velocity of the star, typically taken from a preliminary optimization (see text for details).
[c] S($v_{stellar}$) is a sigmoid prior on the stellar broadening with an exponential drop-off at 7 km s^{-1} for giants and 10 km s^{-1} for dwarfs and subgiants.
[d] Mock spectra are synthesized as continuum-normalized models; therefore, no Chebyshev polynomial is included when fitting them.

photometry given a predicted model:

$$\mathcal{L}_{total} = \mathcal{L}_{spec} \times \mathcal{L}_{phot}. \quad (3)$$

The likelihood accounts for individual measurement errors in the data (assuming that they are normally distributed) and can also incorporate systematic uncertainties in the data (e.g., known systematic biases in the photometric calibration of surveys; see, e.g., Portillo et al. 2020).

We also identify and handle outliers in the photometry by performing an initial χ^2 optimization to the SED in order to find photometric data that are clearly inconsistent with this preliminary model. We classify the photometric data as outliers if they are $>5\sigma$ away from this preliminary SED fit. Once identified, these outlier data are reweighted by increasing the uncertainty such that they are 1σ away from the preliminary best-fit SED model. This procedure effectively identifies obviously erroneous photometry. One could also consider more complex noise models for both the photometry and spectroscopy (e.g., Hogg et al. 2010). This will be the subject of future work.

2.2.2. Priors

MINESweeper allows for a wide range of prior probability distributions of spectroscopic and photometric parameters. There are two types of parameters in MINESweeper: those that are explicitly sampled, and extra parameters that are specified by the sampled parameters but not explicitly fit. These can have priors assigned to them (e.g., a uniform age prior). In our setup the sampled parameters are M, EEP, $[Fe/H]_i$, $[\alpha/Fe]_i$, distance, v_{rad}, $v_{stellar}$, A_V, and four Chebyshev coefficients (T_n). Examples of extra parameters include the age and parallax. These parameters and the associated priors are listed in Table 1. In addition, the MINESweeper code also outputs posterior predictions for any other "derivative" stellar parameters predicted by the MIST models (stellar radius, bolometric luminosity, CNO abundance predictions, etc.)

For the stellar mass, we adopt a prior given by the Kroupa (2001) initial mass function (IMF). The parameters $[Fe/H]_i$ and $[\alpha/Fe]_i$ are uniformly sampled over the available grid range, namely, $-4.0 < [Fe/H]_i < +0.5$ and $-0.2 < [\alpha/Fe]_i < +0.6$. For the EEPs we adopt a uniform prior between the zero-age MS (EEP = 200) and the end of the early-AGB phase (EEP = 808). Because the relation between EEP and age is complex and nonlinear, drawing uniformly in EEP will imprint a complicated age bias (generally in favor of short evolutionary phases). We remove this bias by applying a prior ($\propto \frac{dAge}{dEEP}$) at any given stellar mass and metallicity that results in the probability of drawing any age equally likely. By removing this age bias in this way, we then are able to apply more general age priors (e.g., uniform over the age of the universe) without preferentially being biased toward these ages of rapid stellar evolution.

When fitting mock data and the star cluster data, we adopt a uniform prior on the distance from 0 to 50 kpc. For the H3 data, we incorporate a more physical distance prior based on a simple Galactic number density model similar to other studies (Schönrich & Bergemann 2014; Bailer-Jones et al. 2018; Anders et al. 2019). We use a 3D Galactic model that includes three components: thin-disk, thick-disk, and halo populations. Each of these components has specific shapes and relative fractional contributions based on Bland-Hawthorn & Gerhard (2016) for the thin and thick disks and Xue et al. (2015) for the Galactic halo. In a similar way, we place a prior on stellar age on the H3 data based on this Galactic density model, with the thin disk having a uniform prior from 4 to 14 Gyr, a Gaussian prior for the thick disk centered at 10 Gyr with a σ of 2 Gyr, and a Gaussian prior for the halo centered at 12 Gyr with a σ of 2 Gyr. For the mocks and cluster modeling, we use the uniform priors given in Table 1. We use the measured Gaia DR2 parallaxes as a prior for the cluster and H3 data. We assume

that the Gaia parallax measurements are normally distributed about the true parallax (Luri et al. 2018).

We place a strong prior on the radial velocity, $v_{\rm rad}$, by first performing a least-squares optimization to find a best-fit velocity offset between the observed spectrum and a model spectrum for a star with a $T_{\rm eff}$ and [Fe/H] estimated from an initial photometry-only least-squares fit to the stellar SED. We then adopt a uniform prior that is centered at this initial velocity $\pm 25\,{\rm km\,s^{-1}}$.

The stellar broadening kernel is determined by two parameters: $v_{\rm stellar}$, and the instrumental resolution with resolving power R. For $v_{\rm stellar}$ we adopt a sigmoid prior function with an exponential probability decay at 7 and 10 $\rm km\,s^{-1}$ for giants and dwarfs, respectively. We typically choose to fix the instrumental resolving power to a value determined through fitting a line-spread function to Th-Ar or solar-twilight calibration spectra. In this way, we reduce the dimensionality of the model by one free, highly degenerate parameter.

The Chebyshev coefficients are given Gaussian priors with unit mean and widths of 1.0, 0.25, 0.1, and 0.1. These were determined based on experimentation from fitting the various data sets described below.

And finally, the prior on the total extinction value $A_{\rm V}$ when we fit the mocks is a Gaussian distribution centered at the input "truth" value of 0.1 mag. For the cluster data, we use a less informative, uniform prior of $A_{\rm V} = 0$–1.5 mag. For the H3 data we adopt a Gaussian prior with mean determined by the Schlegel et al. (1998) dust maps with a width of 15% (we have adopted the 14% lower normalization of the dust maps provided by Schlafly & Finkbeiner 2011).

2.2.3. Inference of Stellar Parameters

We approximate the posterior distribution function in `MINESweeper` using the dynamic nested sampling package `dynesty` (Speagle 2020). Nested sampling is particularly well suited for fitting data in `MINESweeper`, as it is more efficient in sampling complex, multimodal, and/or highly covariant posterior distributions compared to other popular Markov Chain Monte Carlo samplers. Nested sampling also directly computes the Bayesian evidence and has a well-defined stopping criterion based on evidence estimation. In this study, we use a nested sampler with 750 live points and a stopping criterion of $\Delta \log Z = 0.01$, where Z is the Bayesian evidence. We have tested these choices to ensure that our results are not sensitive to their exact values. With these `dynesty` settings, for a single star we can sample a full posterior distribution inferred from a medium-resolution spectrum with several thousand pixels and 5−10 bands of photometry in 3−4 CPU hours.

3. Model Validation

In this section we present a variety of tests of the `MINESweeper`-derived parameters. We begin in Section 3.1 with a suite of mock tests that demonstrate the measurement precision of derived parameters as a function of signal-to-noise ratio (S/N), metallicity, and stellar evolutionary phase. Turning to data, in Section 3.2 we fit high-S/N spectra of well-studied benchmark stars that have independently determined stellar parameters. In Section 3.3 we present fits to six star clusters spanning a wide range in metallicity, and in Section 3.4 we fit data from the H3 survey and compare `MINESweeper`-derived distances to parallax-based distances.

Throughout our validation tests, we consider photometry from Pan-STARRS *griz*; Gaia G, BP, and RP; 2MASS JHK_s; and WISE W1 and W2, which together provide coverage across the bulk of the stellar SED for the stellar types considered here. For the mock tests we synthesize photometry in these systems. Where relevant, we include a minimum error floor of $\sigma_{\rm Pan\text{-}STARRS} = 0.02$, $\sigma_{\rm 2MASS} = 0.05$, and $\sigma_{\rm WISE} = 0.04$ mag. In the case of the cluster and H3 data we include Sloan Digital Sky Survey (SDSS) photometry where available, with an error floor of $\sigma_{\rm SDSS} = 0.02$.

For spectra, we focus on a spectral wavelength range and resolution delivered by the MMT Hectochelle instrument (Szentgyorgyi et al. 2011) with the RV31 order-blocking filter, which provides a spectral resolution of $R = 32{,}000$ across the wavelength range 5150−5350 Å. This wavelength range contains the Mg I triplet, commonly used to determine the surface gravity of MS and giant stars. The H3 and cluster data were acquired with this setup. For simplicity and consistency, the mock spectra are also generated with this wavelength range and resolution, and the high-S/N, high-resolution benchmark spectra are smoothed and downsampled to mimic this configuration.

3.1. Mock Data

We first test `MINESweeper` using simulated data that are drawn from the same models used to fit the data. These mock data allow us to determine the precision of derived parameters as a function of S/N, in the limit where the models are a perfect representation of the data. We create mock stars residing on a 9 Gyr isochrone at three metallicities: $[{\rm Fe/H}]_i = -2.0$, -1.0, and 0.0. The stars are placed at a distance of 10 kpc. For most of the analysis in this section we focus on three archetypal stellar types: a warm MS star, a turnoff (TO) star, and a first ascent RGB star. These three types allow us to explore the parameter recovery for three very different regimes within old stellar populations.

Figure 1 shows an example set of posterior distributions resulting from `MINESweeper` modeling of a 9 Gyr solar-metallicity RGB star. The spectrum for this mock has errors corresponding to an S/N ($\rm S/N_{spec}$) of 10 $\rm pixel^{-1}$. Note that we are not adding noise to the actual spectrum. The parallax prior for this mock assumes a parallax S/N ($\rm S/N_{\pi} = \pi/\sigma_{\pi}$) of 1.0.

The posteriors show that we recover all of the input stellar parameters. Figure 1 also shows significant correlations between parameters, resulting in nonsymmetric posteriors, as well as nonlinear covariances. These include "trivial" covariances, e.g., between $\log R$, distance, and $\log L$, that are a consequence of their definitions. Other correlations are more interesting, such as between age and distance, and between [Fe/H], $\log g$, and $T_{\rm eff}$. The latter set is correlated in such a way as to preserve the line depths in the spectrum (higher temperatures go with higher metallicities and higher $\log g$).

Figure 2 shows recovered stellar parameters in Kiel diagrams ($\log g - T_{\rm eff}$) for eight evolutionary phases ranging from lower MS through the TO, RGB, horizontal branch (HB), and asymptotic giant branch (AGB). Rows show results for three metallicities: $[{\rm Fe/H}]_i = -2.0$, -1.0, 0.0. Columns show results for $\rm S/N_{spec} = 2$, 5, 10. In all cases we have assumed a parallax S/N ($\rm S/N_{\pi}$) of 1.0, and broadband photometries from Pan-STARRS, Gaia, 2MASS, and WISE are included in

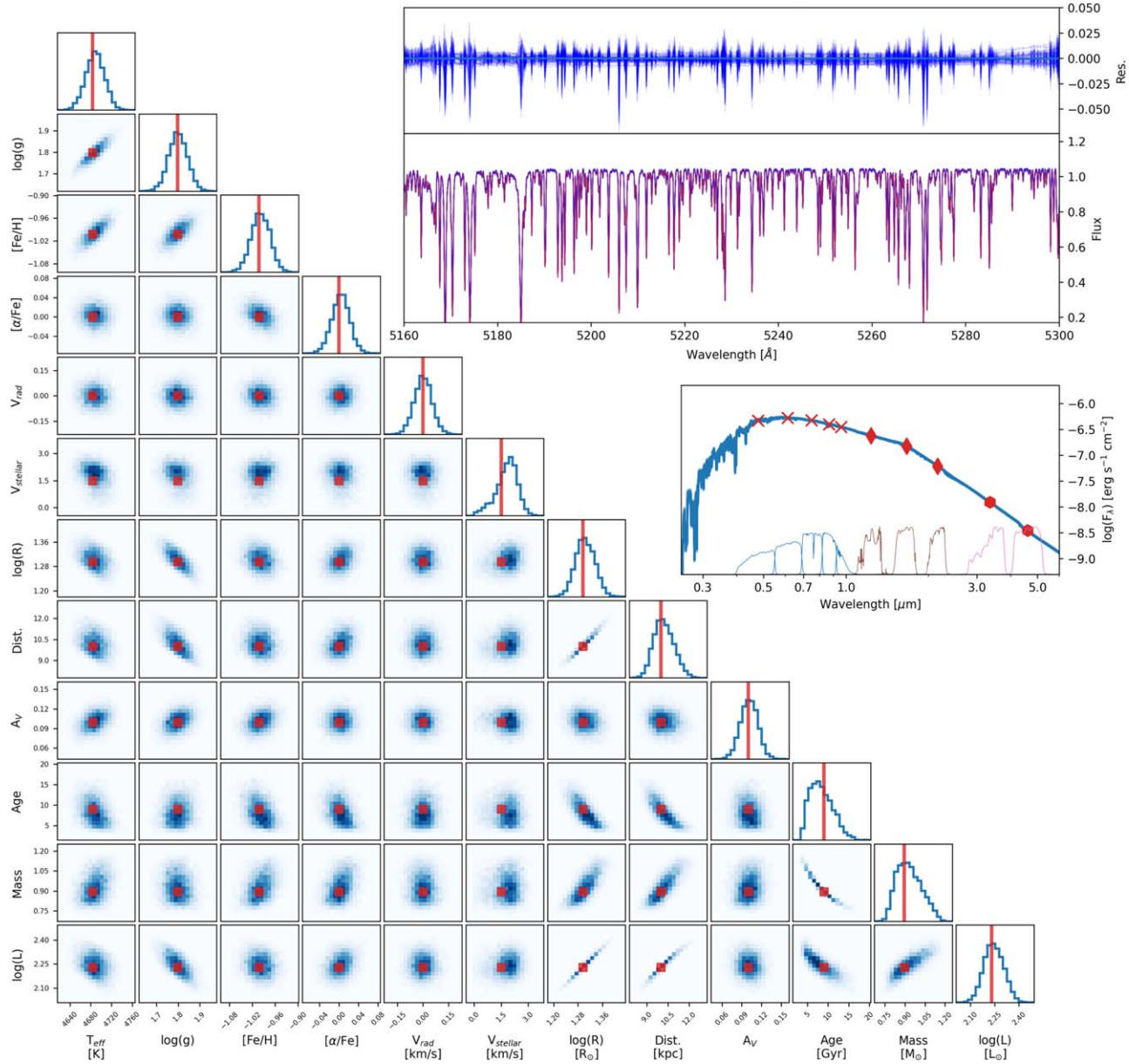


Figure 1. Example posterior distributions for a `MINESweeper` fit to a solar-metallicity RGB star with spectrum $S/N_{spec} = 10$ pixel^{-1} and parallax $S/N_{\pi} = 10$. For clarity, only a subset of free parameters are shown here. In each panel, the stellar parameters used to create the mock data are indicated by red squares or lines. Plotted on the right are the mock spectrum (top) and SED (bottom), along with 1000 models randomly sampled from the posteriors (blue lines). Above the mock spectrum, we also plot the relative residual for the 1000 posterior draws (blue lines).

the fitting. Also shown is the 9 Gyr isochrone from which the mock data were drawn.

In general, the posteriors agree very well with the true values (shown as star symbols). It is noteworthy that this remains true even at $S/N_{spec} = 2$ at low metallicity. `MINESweeper` struggles in the case of the TO star, especially at low S/N_{spec} and low metallicity. The primary reason for this is the age prior. The TO is a relatively short-lived phase, and so the prior strongly prefers the higher-$\log g$ solution at the same temperature. High S/N and/or high metallicity are required for the spectrum to favor the lower-$\log g$ TO solution.

The parameter space explored here was tailored to the stellar populations and data quality relevant for the H3 survey (see Section 3.4 below), and so one should be careful not to generalize these results too broadly. In particular, the combined use of photometry and spectra is essential here, as the former strongly constrains $T_{\rm eff}$, which in turn allows the spectrum to place stronger constraints on $\log g$ and [Fe/H] even at low S/N_{spec}.

We now focus on the precision with which `MINESweeper` recovers distances. Figure 3 shows the fractional distance uncertainties for the MS, TO, and RGB stars as a function of metallicity, S/N_{π}, and S/N_{spec}. The parallax prior is also shown as a gray line in each panel.

The top panels show the fractional distance uncertainties as a function of parallax S/N. From these plots it is clear that the parallax prior dominates the distance uncertainties for $S/N_{\pi} \gtrsim 10$. At lower S/N_{π} the uncertainties are governed by the spectrum (and photometry, which is also included in the fit). At low S/N_{spec} (left panel) the distance precision depends strongly on stellar type but not on metallicity, while at high S/N_{spec} (right panel) there is a strong metallicity dependence

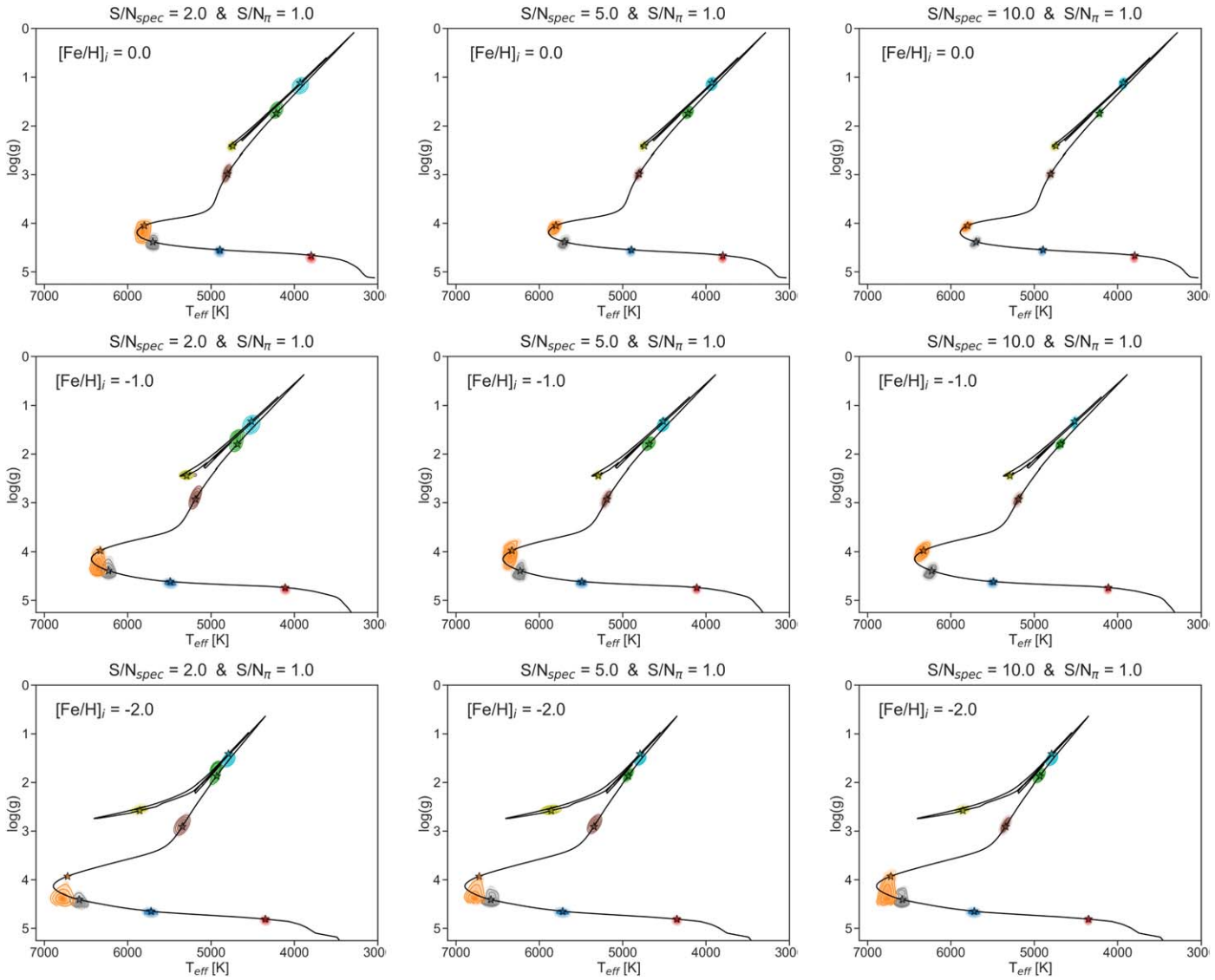


Figure 2. Kiel diagrams showing the posterior distributions resulting from modeling mock spectra and photometry with parallax $S/N_{\pi} = 1$, for $[Fe/H]_i = 0.0$ (top row), $[Fe/H]_i = -1.0$ (middle row), and $[Fe/H]_i = 2.0$ (bottom row). Mocks were fit with photometry and a spectrum with spectral $S/N_{spec} = 2$ pixel^{-1} (left column), spectral $S/N_{spec} = 5$ pixel^{-1} (middle column), and $S/N_{spec} = 10$ pixel^{-1} (right column). Colors indicate the evolutionary state of the mock star ranging from lower MS through the TO, first ascent giant branch, horizontal branch, and asymptotic giant branch. Contours indicate 1σ, 2σ, and 3σ credible intervals. True isochrone positions of these mock stars are given by the star symbols.

for the TO star (see also the right panels of Figure 2). The distances are well constrained for the MS star in all cases owing to the nature of stellar isochrones: if the star can be determined to be a dwarf, not near the TO, with a well-measured $T_{\rm eff}$ from photometry, then its location in the H-R diagram can be determined quite precisely.

The bottom panels of Figure 3 show the fractional distance uncertainties as a function of S/N_{spec}. It is noteworthy that the distance precision for the RGB star does not begin to significantly improve with increasing S/N_{spec} until $S/N_{spec} \gtrsim 10$. For the TO star the distance precision starts improving for $S/N_{spec} \gtrsim 2$, except for the lowest metallicity, which shows no discernible improvement with S/N_{spec}. Finally, the MS star shows a more continuous increase in distance precision with S/N_{spec} for all metallicities.

In Figure 4, we show for several stellar properties the output posterior distributions from our `MINESweeper` modeling of the mock data with $S/N_{spec} = 2$, 5, and 10 pixel^{-1} and $S/N_{\pi} = 1.0$. Here we focus on the MS, TO, and RGB stars with $[Fe/H]_i = -1.0$.

The posteriors for the three stellar types have different correlations and systematic offsets as a result of the information content of the spectrum and photometry, as well as the shape of the isochrone at the three different evolutionary phases. In many cases, increasing S/N_{spec} has the expected effect of reducing the posterior regions. Interesting exceptions to this rule include the distance constraint for the TO star (as noted earlier) and the age constraint for the MS and RGB stars (right panels). For the TO star, while the distance uncertainty is not reduced, the systematic bias is removed at higher S/N. In the case of the MS and RGB stars, they have weak age constraints for the well-known reason that the isochrones at those locations are nearly independent of age for the range of ages considered here. We also note that in all cases [Fe/H] is recovered with no bias and small uncertainties, even at $S/N_{spec} = 2.0$.

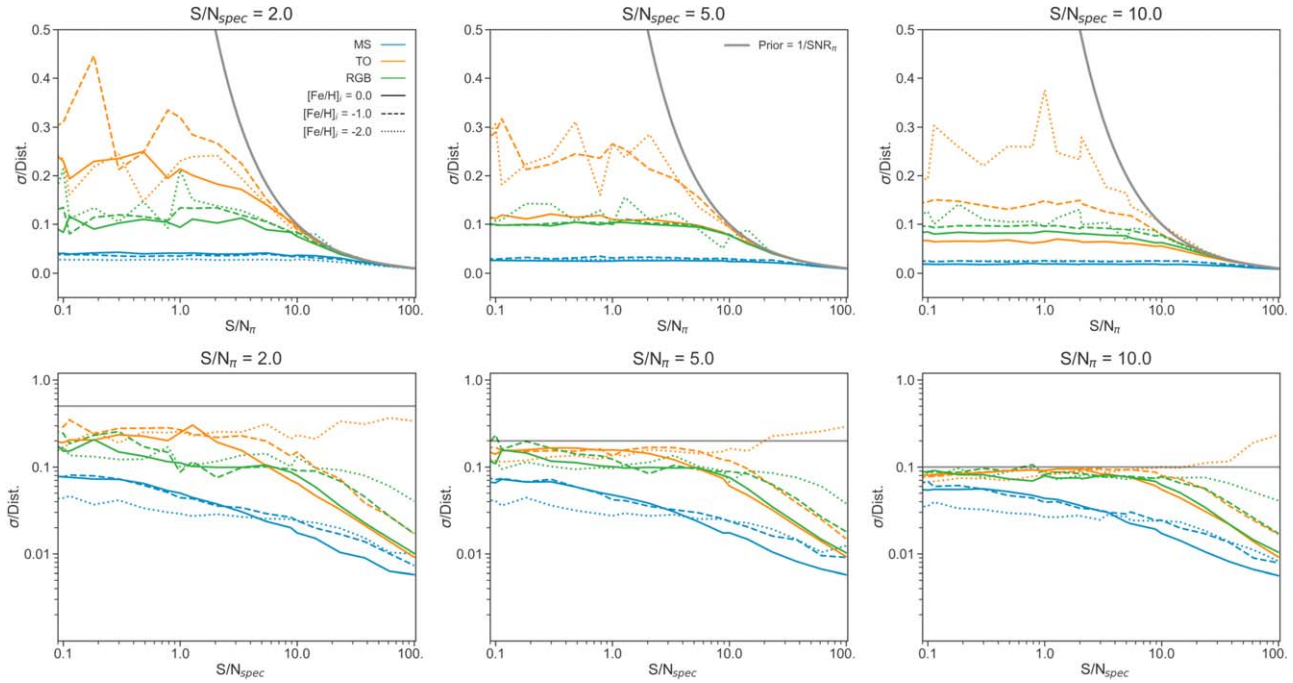


Figure 3. Fractional distance uncertainty as a function of parallax S/N (top panels) and spectrum S/N (bottom panels) for three evolutionary phases: MS, TO, and RGB, each at three metallicities: $[Fe/H]_i = -2.0$, -1.0, and 0.0. The parallax S/N prior is shown as a gray line in each panel. Distances and uncertainties are determined for each mock based on the median and 1σ credible interval, respectively.

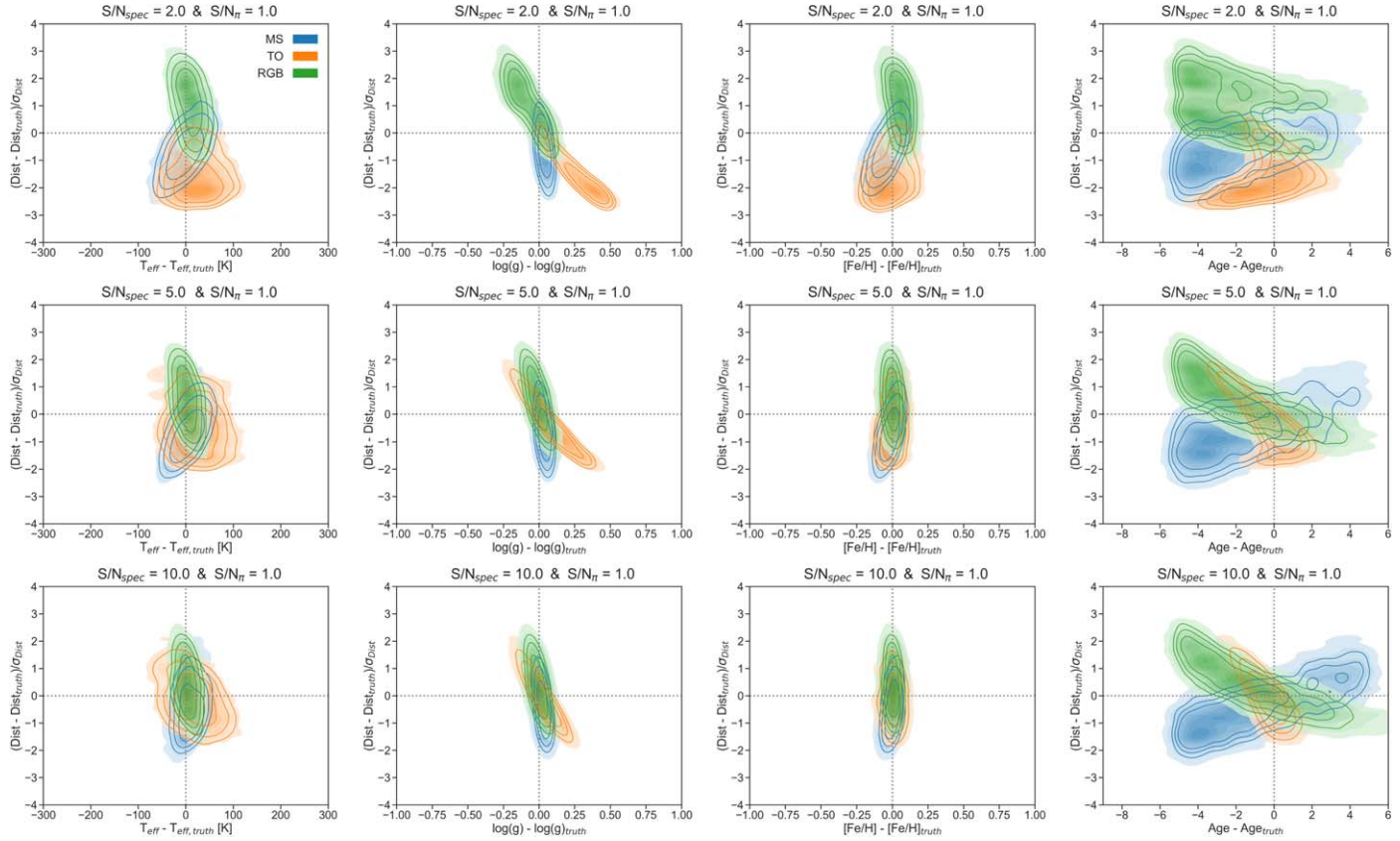


Figure 4. Constraints on stellar parameters relative to the true input values for three evolutionary phases (MS, TO, RGB) at $[Fe/H]_i = -1.0$ and parallax $S/N_\pi = 1.0$. Rows show results for $S/N_{spec} = 2$ (top), 5 (middle), and 10 (bottom).

Figure 5 explores the accuracy of the recovered distances and uncertainties. In this figure we show the measured distance minus the true distance in units of the reported uncertainty. This quantity should lie within ± 1 if the measured distances and uncertainties are accurate. We show this quantity for the MS, TO, and RGB stars for metallicities

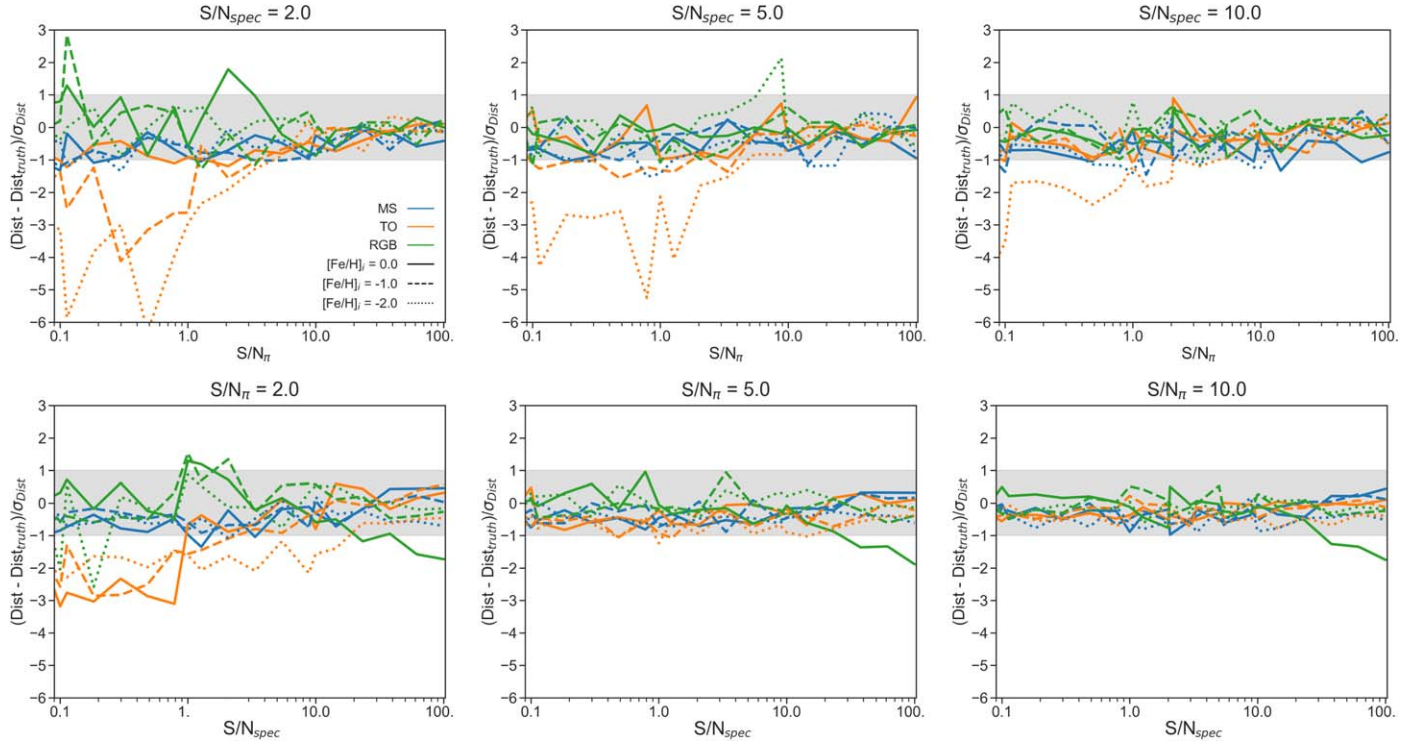


Figure 5. Accuracy of recovered distances and reported uncertainties of mock data. The *y*-axis shows the best-fit distance minus the true distance in units of the measurement uncertainty. Each panel shows three evolutionary phases (MS, TO, RGB) at three metallicities: $[Fe/H]_i = -2.0$, -1.0, and 0.0. The top panels show results as a function of parallax S/N for three S/N_{spec}, while the bottom panels show results as a function of S/N_{spec} for three choices of parallax S/N. In general, the measured distances agree with the true distances to within $\approx 1\sigma$. Significant deviations are limited to low-metallicity TO stars at low S/N.

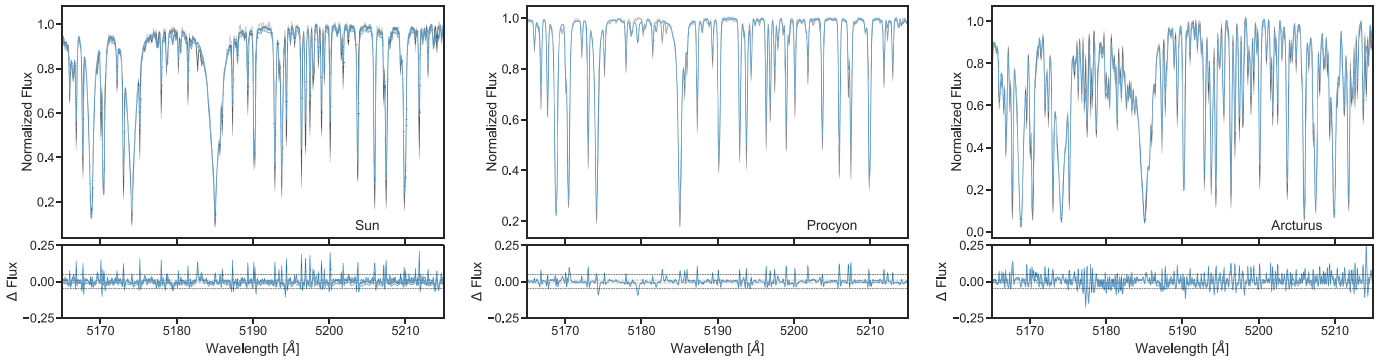


Figure 6. Gaia FGK benchmark spectra and best-fit models (top panels) with residuals (bottom panels) for the Sun (left), Procyon (middle), and Arcturus (right). Spectra were smoothed to a common resolution of $R = 32{,}000$ before fitting. Parameters inferred from `MINESweeper` modeling of these spectra are given in Table 2. Also displayed with the residuals are lines indicating 1% (dotted) and 5% (dashed) residuals.

$[Fe/H]_i = -2.0$, -1.0, and 0.0 and as a function of S/N_{spec} and S/N_π.

In most cases the reported distances and uncertainties are within the expected range (demarcated by the gray band), indicating that these quantities are well measured by `MINESweeper`. There are two outlier cases. First are the TO stars at low S/N, especially at $[Fe/H] = -2.0$. The distances deviate by several σ compared to the true values, as noted earlier in this section. The reason for this is that at low S/N the prior plays an increasingly important role, and the TO star is a less likely phase to find a star compared to the higher-$\log g$ solution at the same T_{eff}. The distances in this case are therefore systematically underestimated. This bias is removed for $S/N_{spec} > 2$ except for $[Fe/H]_i = -2.0$, where the bias is lifted for $S/N_{spec} > 10$.

The second outlier case is at very high S/N_{spec} for the solar-metallicity RGB star. In this case the issue is related to the posterior solutions becoming poorly sampled, as the uncertainties are dropping to below 1%. We have added a noise floor of 1% in the distance uncertainties to partially alleviate this issue. In any event, this issue is not relevant in the analysis of real data, as systematic uncertainties dominate the error budget at the percent level.

3.2. FGK Benchmark Stars

We now turn to validation of `MINESweeper` with real data. In this section we model seven well-studied "benchmark" stars: the Sun, ϵ Eri, Procyon A (hereafter Procyon), HD 84937, HD 140283, Arcturus, and HD 122563. These stars were chosen because they contain multiple spectra in the Gaia Benchmark Stars online database (Blanco-Cuaresma et al. 2014) and because they span a range in [Fe/H], T_{eff}, and $\log g$. They are

Table 2
MINESweeper Parameters for Benchmark Stars

Star	$T_{\rm eff}$ (K)	$\log g$	[Fe/H]	[α/Fe]	M ($M_\odot$)	L ($L_\odot$)	Age (Gyr)	Dist. (pc)	$N_{\rm spec}$[a] Phot.[b]
Sun									
This work	5829 ± 25	4.45 ± 0.06	0.08 ± 0.02	0.03 ± 0.02	1.04 ± 0.01	1.06 ± 0.10	2.88 ± 2.50	... ± ...	9
This work[c]	5812 ± 22	4.42 ± 0.02	0.07 ± 0.04	0.03 ± 0.02	1.03 ± 0.03	1.11 ± 0.09	4.56 ± 0.04	... ± ...	9
Lit.	5772 ± 1	4.44 ± ...	0.00 ± ...	0.00 ± ...	1.00 ± ...	1.00 ± ...	4.57 ± 0.01	... ± ...	⋯
ϵ Eri									
This work	5026 ± 37	4.59 ± 0.04	−0.10 ± 0.03	0.02 ± 0.04	0.81 ± 0.04	0.37 ± 0.01	6.07 ± 7.05	3.25 ± 0.25	4
Lit.	5076 ± 30	4.60 ± 0.03	−0.09 ± 0.06	0.01 ± 0.12	0.80 ± 0.06	0.32 ± 0.01	0.4–0.9	3.20 ± 0.01	G,2M
Procyon									
This work	6598 ± 24	3.93 ± 0.05	−0.01 ± 0.02	0.04 ± 0.01	1.52 ± 0.04	7.86 ± 1.10	1.89 ± 0.07	3.66 ± 0.26	6
Lit.	6554 ± 84	4.00 ± 0.02	−0.04 ± 0.08	0.08 ± 0.11	1.50 ± 0.07	6.88 ± 0.36	1.5–2.5	3.51 ± 0.01	2M
HD 84937									
This work	6380 ± 84	3.89 ± 0.09	−2.30 ± 0.06	0.52 ± 0.03	0.81 ± 0.02	3.98 ± 0.79	10.31 ± 0.95	95.17 ± 11.18	5
Lit.	6356 ± 97	4.06 ± 0.04	−2.09 ± 0.08	0.33 ± 0.14	0.75 ± 0.04	2.09 ± 0.03	11.0–13.5	72.52 ± 4.64	2M
HD 140283									
This work	5860 ± 287	3.59 ± 0.19	−2.59 ± 0.16	0.44 ± 0.05	0.77 ± 0.03	5.57 ± 1.51	12.97 ± 1.90	64.98 ± 12.40	5
Lit.	5522 ± 105	3.58 ± 0.11	−2.43 ± 0.10	0.10 ± 0.12	0.68 ± 0.17	4.65 ± 0.09	$\geqslant$12.0	62.06 ± 0.28	G,2M
Arcturus									
This work	4297 ± 9	1.66 ± 0.03	−0.64 ± 0.02	0.41 ± 0.00	1.28 ± 0.08	229.92 ± 21.33	3.65 ± 0.63	12.24 ± 0.49	5
Lit.	4286 ± 35	1.64 ± 0.09	−0.53 ± 0.08	0.37 ± 0.14	1.03 ± 0.21	198.05 ± 5.06	4.0–10.0	11.28 ± 0.05	2M
HD 122563									
This work	4517 ± 61	1.15 ± 0.32	−2.91 ± 0.07	0.50 ± 0.12	1.17 ± 0.40	907.52 ± 349.47	3.19 ± 9.18	493.53 ± 87.28	5
Lit.	4587 ± 60	1.61 ± 0.07	−2.74 ± 0.22	0.39 ± 0.24	0.86 ± 0.03	343.66 ± 14	⋯	290.36 ± 5.22	G,2M,W

Notes. Literature parameters are adopted from Jofré et al. (2014, 2015), Heiter et al. (2015), and Sahlholdt et al. (2019), except for distances, which are derived directly from parallaxes given in Gaia DR2. Luminosities have also been updated to reflect the DR2 parallaxes. The values given for MINESweeper are the weighted mean and standard deviation measured from the set of fitted spectra.
[a] Number of spectra analyzed.
[b] Photometric data used in the fitting: G: Gaia DR2; 2M: 2MASS; W: WISE.
[c] Parameters include a strong age prior based on meteorite data: age = 4.567 ± 0.01 Gyr (Connelly et al. 2012).

nearby and bright, and hence have precisely determined parallax-based distances and very high S/N data.

Spectra for these stars were obtained from the Gaia Benchmark Stars online database, an archive of high-resolution and high-S/N spectra of FGK stars (Blanco-Cuaresma et al. 2014). We use all available spectra in the archive. These spectra uniformly are of high resolution ($R \approx 80{,}000$) and S/N ($S/N_{\rm spec} \geqslant 200$). We limit the wavelength range to $\lambda =$ 5150–5350 Å and convolve the spectra with a Gaussian instrument profile to a resolving power of $R = 32{,}000$. We choose this setup because it is similar to the MMT/Hectochelle data analyzed in later sections. As is the case with many "naked-eye" stars, these Gaia Benchmark Stars lack high-quality photometry in many modern surveys. Therefore, we include photometry only if available from three well-calibrated photometric surveys: Gaia, 2MASS, and/or WISE. Each available spectrum plus photometry is fit with MINESweeper. We adopt a flat distance prior of 0–1 kpc and a flat prior in log (age) from 0.1 to 14 Gyr. Independent parallax information is not used as a prior in the fitting. The Sun was fit without photometry, and so no distance is reported.

A summary of best-fit parameters for these stars is given in Table 2; best-fit models to the spectra of the Sun, Arcturus, and Procyon are shown in Figure 6; and a comparison between the MINESweeper and values from the literature for $T_{\rm eff}$, log g, [Fe/H], and distances is presented in Figure 7. The quoted uncertainties from this work represent the scatter among the fits to the independent spectra available for each star. The formal 1σ uncertainties reported by MINESweeper are always much smaller than this value owing to the very high $S/N_{\rm spec}$. The literature sources are adopted directly from the Gaia benchmark online table,[8] and they are based on spectroscopic analysis from Jofré et al. (2014, 2015) and Heiter et al. (2015) and isochrone fitting from Sahlholdt et al. (2019). In MINESweeper all of the $\alpha-$elements are varied in lockstep. However, the available wavelength range is dominated by the Mg I triplet, so we expect that [α/Fe] is most sensitive to [Mg/Fe]. For the literature values, we therefore adopt [Mg/Fe] as a tracer of [α/Fe].

Overall, the MINESweeper parameters are in good agreement with the parameters published in the literature for these stars. In nearly all cases parameters agree to within $\lesssim$0.1 dex or the quoted uncertainties, whichever is larger. It is particularly noteworthy that we are able to recover the literature parameters given the wide range of values: metallicities range from −2.7 to +0.0, log g values range from 4.6 to 1.6, and temperatures range from 4286 to 6550 K. The most significant outlier is HD 122563. Previous work has also found this star to lie well off of standard stellar evolution models given its measured metallicity, $T_{\rm eff}$, and bolometric luminosity (Creevey et al. 2012, 2019), and so the fact that this star is an outlier is not surprising.

Our analysis shows how sensitive ages for MS stars like the Sun are to small changes in measured parameters (see Table 2). For the Sun, a change of $T_{\rm eff} \approx 10$ K results in an age difference of ≈2 Gyr. We have therefore refit the Sun with a

[8] www.blancocuaresma.com/s/benchmarkstars

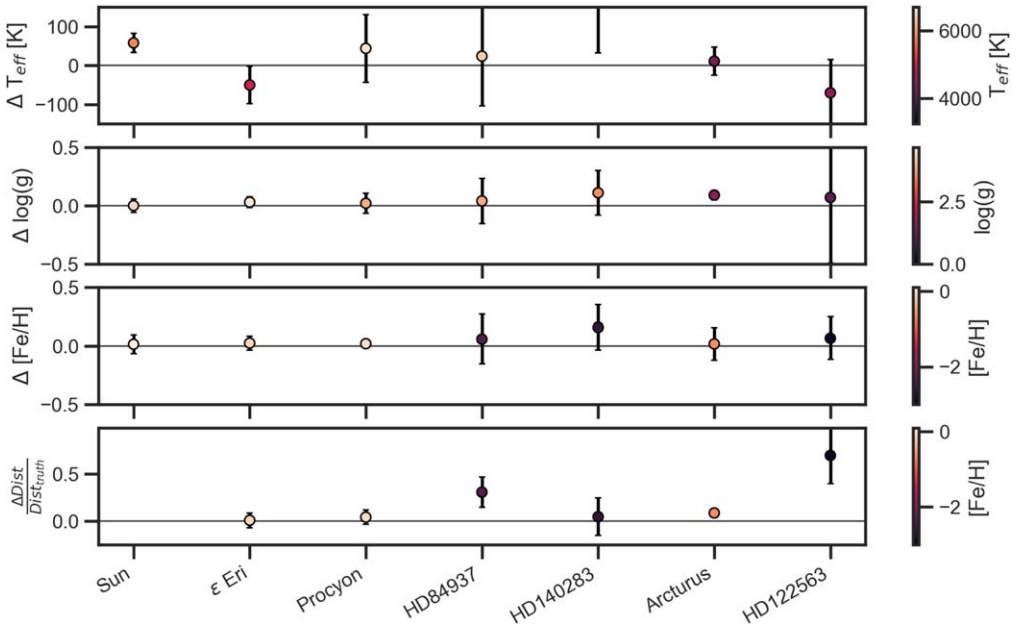


Figure 7. Comparison of MINESweeper parameters and literature values for a set of benchmark stars. The differences are MINESweeper parameters minus the literature value. The error bars represent the MINESweeper uncertainties added in quadrature with the literature uncertainties, and in some cases they are smaller than the symbol size. See Table 2 for literature references. For HD 140283 the difference in $T_{\rm eff}$ is 338 K and is therefore beyond the range of the figure.

Table 3
Cluster Parameters

Cluster	Number of Stars	Lit.				MINESweeper			
		Age (Gyr)	Dist. (kpc)	[Fe/H][a]	[α/Fe][b]	Age (Gyr)	Dist. (kpc)	[Fe/H]	[α/Fe]
M92	38	13.25	17.7	−2.23, −2.31	0.14–0.47	⋯	9.26 ± 2.46	−2.13 ± 0.15	0.29 ± 0.12
M13	80	12.9	12.5	−1.50, −1.53	0.13–0.26	⋯	8.33 ± 1.33	−1.44 ± 0.06	0.19 ± 0.05
M3	82	12.6	10.5	−1.40, −1.50	0.15–0.25	⋯	10.66 ± 1.22	−1.34 ± 0.08	0.17 ± 0.04
M107	40	12.75	6.75	−1.01, −1.02	0.24–0.51	⋯	5.80 ± 1.07	−0.97 ± 0.09	0.35 ± 0.10
M71	124	12.5	4.44	−0.68, −0.78	0.21–0.49	9.92 ± 1.37	4.27 ± 0.65	−0.72 ± 0.13	0.35 ± 0.09
M67	178	3.58	0.88	−0.01	−0.05–0.05	3.92 ± 0.37	0.80 ± 0.08	0.04 ± 0.06	0.06 ± 0.05

Notes. Uncertainties from MINESweeper modeling are based on the 68th percentile scatter weighted by the maximum posterior probability ($\propto$maximum likelihood) for the individual clusters. They therefore include both the measurement uncertainty and any systematic bias. The formal errors on the mean values are much smaller and so are not reported here.

[a] Literature values of globular clusters for [Fe/H] are listed twice, from Mészáros et al. (2015) and Carretta et al. (2009a).

[b] Literature values for [α/Fe] are quoted as a range encompassing the results of Mészáros et al. (2015), Carretta et al. (2009a), and the literature compilation of Pritzl et al. (2005).

strong prior on the age derived from meteorite data (Connelly et al. 2012), and we show the resulting derived parameters for this case also in Table 2. The parameters are nearly unchanged with respect to the less constrained fit.

3.3. Star Clusters

Another important validation test for MINESweeper is to fit stars belonging to well-studied star clusters. These populations offer important tests for several reasons: First, a subset of open and globular clusters have been extensively studied for decades, and so their spectroscopic parameters and distances are relatively well known. Second, stars within a cluster are all at approximately the same distance. And third, all stars within a cluster were born with the same abundances, so inspection of $[{\rm Fe/H}]_i$ and $[\alpha/{\rm Fe}]_i$ as a function of $T_{\rm eff}$ and $\log g$ offers a very strong test of the precision of the reported abundances. This third point is complicated by two effects: First, element diffusion and mixing of the surface layers of a star will result in differences between photospheric abundances and the initial abundances of a star (see, e.g., Dotter et al. 2017). MIST includes both of these effects, so our reported initial abundances should be independent of evolutionary phase. Second, it is now well known that globular clusters harbor substantial light-element abundance variations *within* individual clusters (e.g., Gratton et al. 2004). This includes variation in helium and oxygen, which can significantly alter a star's location in the H-R diagram. These caveats must be kept in mind when using globular clusters as calibrators.

3.3.1. Data

Here we fit data collected for members of clusters spanning a wide range of metallicities: globular clusters M92, M13, M3, M107, and M71, as well as the solar-metallicity open cluster M67. Literature parameters for each individual cluster are

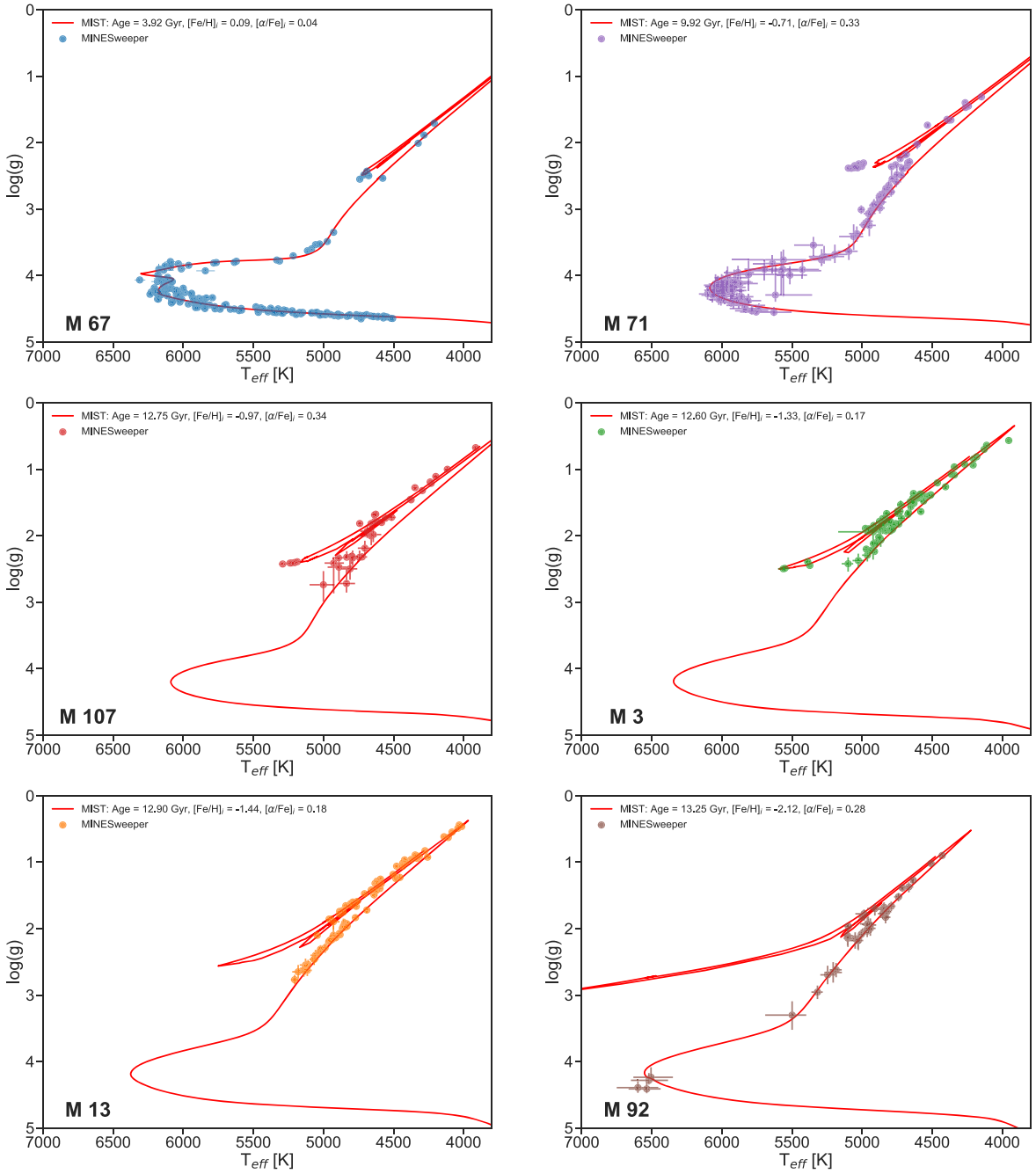


Figure 8. Kiel diagrams for clusters M67 (top left), M71 (top right), M107 (middle left), M3 (middle right), M13 (bottom left), and M92 (bottom right). For each cluster, we plot the `MINESweeper`-measured $\log g$ and $T_{\rm eff}$ for stars with Gaia DR2 parallax priors included in the modeling. Errors on the points give the 68th percentile credible interval for these parameters. Overplotted are the `MIST` isochrone predictions (red lines) for estimated average cluster ages and metallicities. In the cases of M107, M3, M13, and M92, the isochrone ages are based on literature values.

given in Table 3. For the globular clusters, the [Fe/H] values are taken from Mészáros et al. (2015) by default, with a nominal uncertainty of 0.05 dex. We also quote [Fe/H] values from Carretta et al. (2009a) in some cases to highlight the literature uncertainty associated with the metallicity scale. Ages for the globular clusters are adopted from Dotter et al. (2010). For the open cluster M67, literature [Fe/H] and [α/Fe] are adopted from Liu et al. (2016), while a literature age is adopted from Stello et al. (2016), who estimate the age from astroseismology of M67 giants.

For [α/Fe], we adopt a range for each cluster, determined by the literature compilation of Pritzl et al. (2005) and the results presented in Carretta et al. (2009a) and Mészáros et al. (2015). The latter two works measure [Mg/Fe], which we adopt as a proxy for [α/Fe]. We note that it is difficult to report a single value of "[α/Fe]" for a cluster for two reasons. First, the α-elements do not vary in lockstep, and so formally "[α/Fe]" should be some appropriate mass-weighted average. Second, the light α-elements (e.g., O, Mg) show star-to-star variation within clusters, in some cases with $\approx$1 dex variation (e.g., Carretta et al. 2009b). So a large sample of stars must be studied in order to report a reliable average for a cluster.

Medium-resolution spectra were collected for stars in these clusters as part of the calibration phase of the H3 survey using

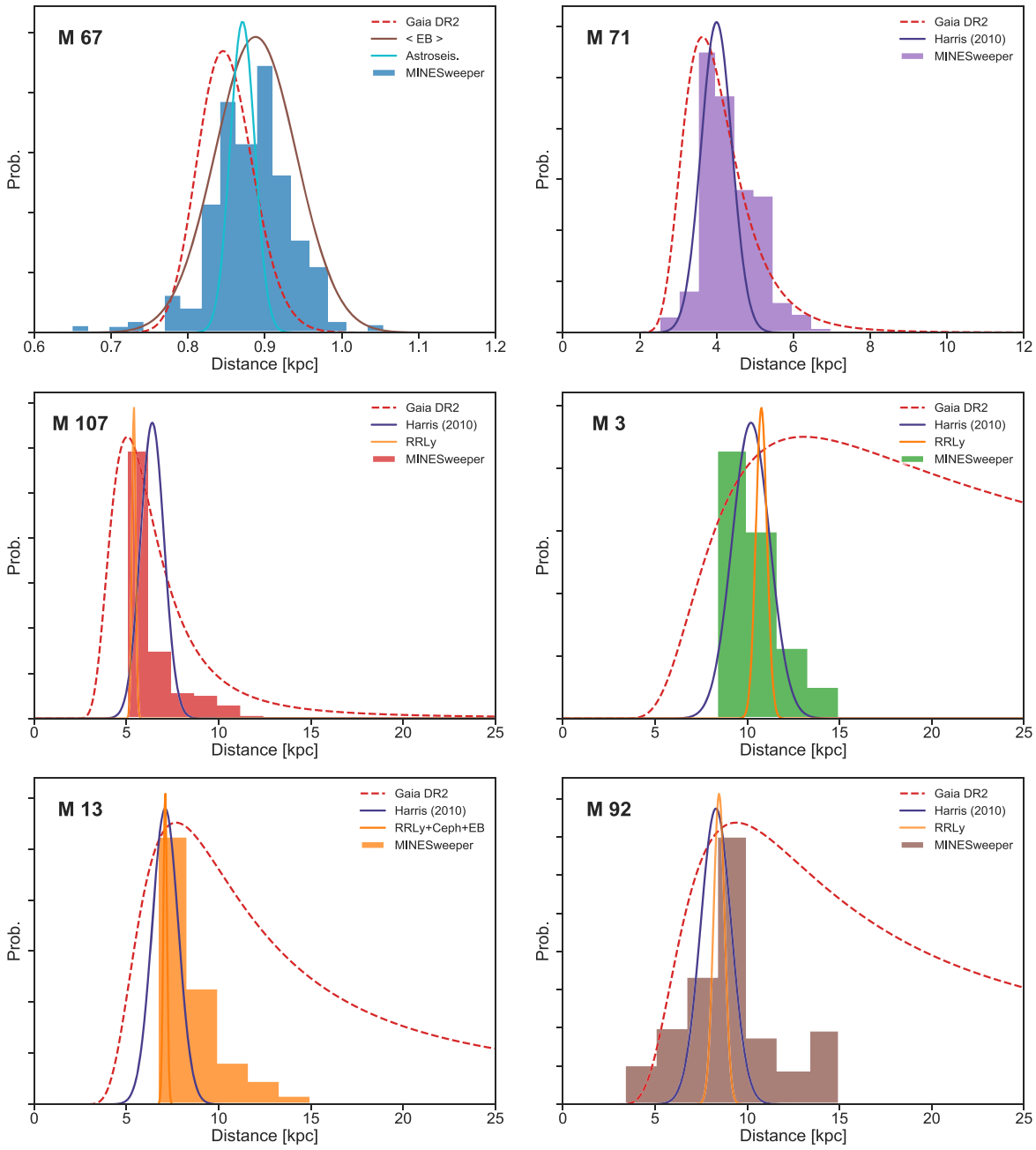


Figure 9. Derived distances for stars in six star clusters. `MINESweeper`-derived distances include the Gaia DR2 priors. Individual stellar distances have been weighted by their `MINESweeper`-derived posterior probabilities. The red dashed line shows the distance constraints based on Gaia DR2 parallaxes alone (Gaia Collaboration et al. 2018a). The other lines are for cluster distances measured using alternative techniques (see text for references). The `MINESweeper`-derived distances are in good agreement with the existing independent constraints.

the Hectochelle instrument on the MMT (see Section 3.4). Owing to the varying distances, in some clusters spectra were obtained for stars over a wide range of $\log g$, while in others spectra were collected only for the giants. Targets were identified by a spatial selection and the requirement that they reside near an isochrone appropriate for the cluster parameters.

We fit both the spectroscopy and, where available, broad-band photometry with `MINESweeper`. We include the Gaia DR2 parallax as a Gaussian prior. The DR2 parallaxes have known systematic zero-point issues. Following Lindegren et al. (2018), Schönrich et al. (2019), and Leung & Bovy (2019), we therefore add an offset of $+0.05$ to the Gaia parallaxes and also add this value in quadrature to the reported uncertainties.

For each cluster, likely members were identified through a combination of coordinates, radial velocities, and Gaia proper motions and parallaxes. Stars with $S/N_{spec} < 3$ are removed from further analysis. The median S/Ns of the cluster members are 9, 15, 20, 25, 31, and 35 for M71, M107, M92, M67, M3, and M13, respectively. We tabulate the results of our `MINESweeper` modeling for likely members in Table 3, where we provide the weighted median values for the cluster members and the reported errors are the 68th percentile range for each parameter.

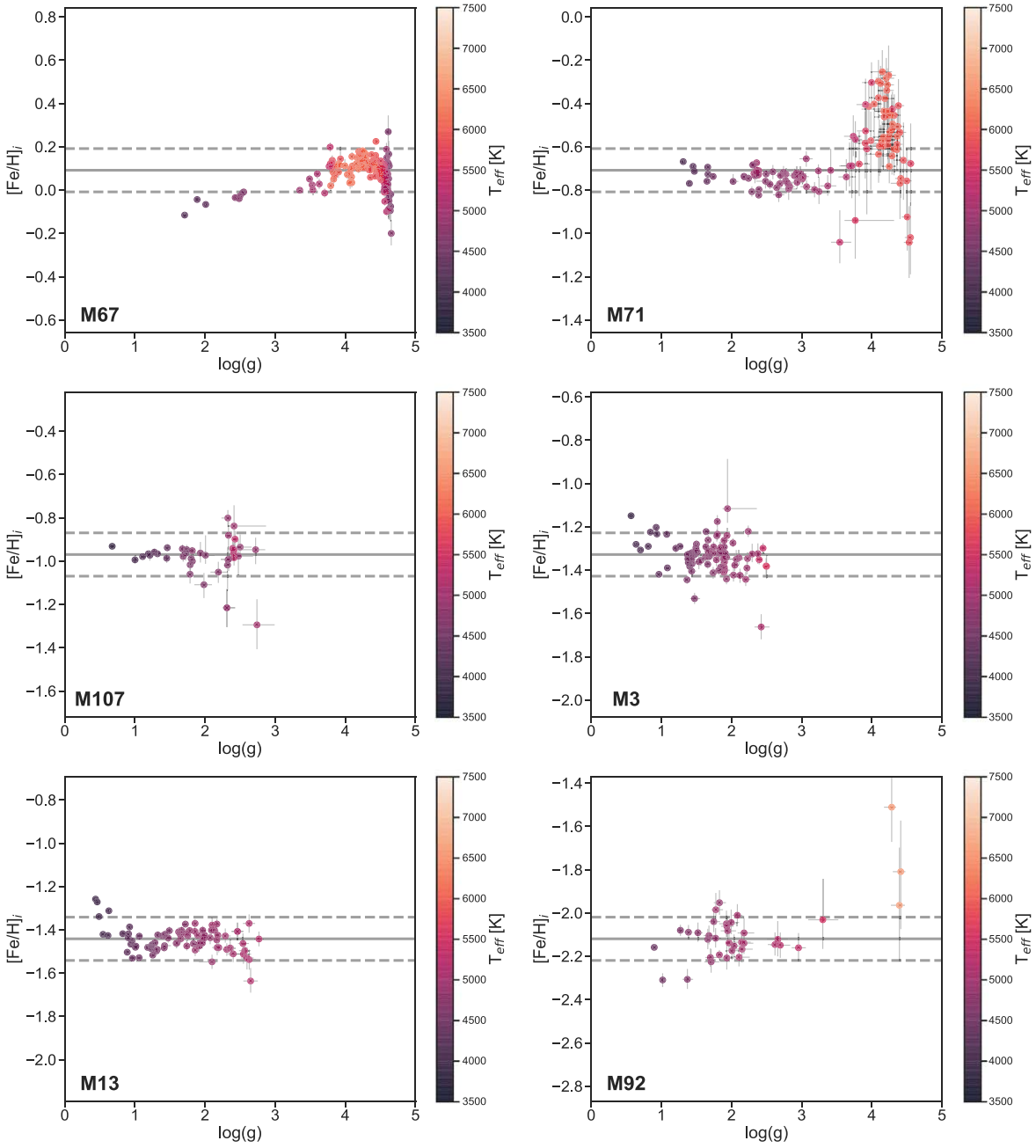


Figure 10. Derived metallicities as a function of $\log g$ for six star clusters. Plotted on the y-axis is the initial stellar [Fe/H] as opposed to the more commonly reported surface abundance. The color scale of the points indicates the $T_{\rm eff}$ of the stars. The dashed black lines indicate the weighted median $[{\rm Fe/H}]_i$ for each cluster, and the gray dashed lines mark ± 0.1 around the median value and are intended to guide the eye.

3.3.2. Results

Figure 8 shows the resulting Kiel diagram for each cluster based on this analysis, along with `MIST` isochrones at the derived median cluster metallicity and age. In the case of the clusters with no or only a few TO stars (M107, M3, M13, and M92) we use the literature age for the displayed isochrone. Most points within a cluster lie along a single isochrone, which is a nontrivial success of the method because, while an individual star must reside on an isochrone, there was no guarantee that all stars would reside on the *same* isochrone. M71 presents an interesting case, as the core He-burning sequence is unusually extended in $T_{\rm eff}$ for such a metal-rich cluster. This extension might reflect the observed spread in light elements for this cluster (Cordero et al. 2015). This effect is not included in our isochrone tables, so the tension there is not surprising.

In Figure 9 we compare the `MINESweeper` distances for cluster members to several literature sources. We show the Gaia DR2 cluster parallax estimates from Gaia Collaboration et al. (2018b), in which we include the +0.05 zero-point correction and also add that as an additional term in the error budget. For the nearby clusters (in particular M67) the Gaia parallaxes offer a useful constraint, while for the most distant clusters they provide very weak constraints. We also include the Harris (2010) tabulation of distances for globular clusters (with a 10% distance uncertainty for reference) and RR Lyrae standard candles where available (Benedict et al. 2011;

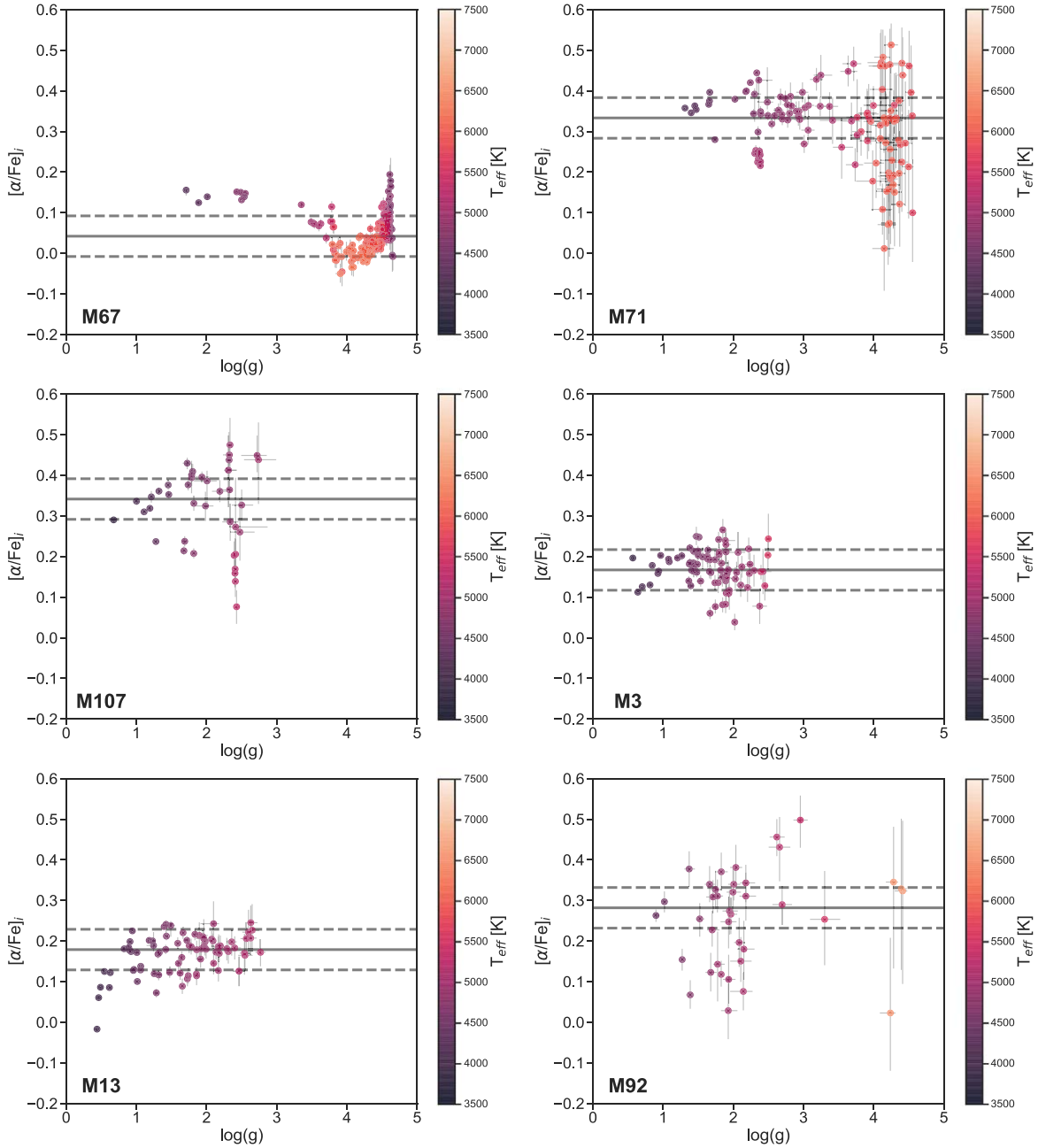


Figure 11. Same as in Figure 10, but now for $[\alpha/\mathrm{Fe}]$ as a function of $\log g$. Gray dashed lines mark ± 0.05 around the median value for each the cluster and are intended to guide the eye.

Deras et al. 2018, 2019; Hernitschek et al. 2019). Finally, for M67 we include distances measured from eclipsing binaries (Yakut et al. 2009; Sandquist et al. 2018) and from astroseismology of giants in the cluster (Stello et al. 2016). Overall the `MINESweeper`-derived distances are in good agreement with the literature values.

Figures 10 and 11 show measured $[\mathrm{Fe/H}]_i$ and $[\alpha/\mathrm{Fe}]_i$ values as a function of $\log g$ for each cluster. These plots are valuable for diagnosing systematic trends in the derived abundances, especially since the reported abundances are the *initial*, rather than the surface, values (the former should not vary with evolutionary state of the star). In each panel the points are color-coded by $T_{\rm eff}$. The weighted medians are shows as black dashed lines, and gray lines indicate ± 0.1 for $[\mathrm{Fe/H}]_i$ and ± 0.05 for $[\alpha/\mathrm{Fe}]_i$ and are meant to guide the eye.

We begin by discussing M67. This cluster has the most members in our sample at 181 spanning the widest range in $\log g$. In addition, the Gaia parallaxes are high S/N for this cluster ($\langle \mathrm{S/N}_\pi \rangle = 28$), and this cluster has the highest metallicity in the sample, so the data place very demanding constraints on the model. There is a clear systematic trend in $[\mathrm{Fe/H}]_i$ with $\log g$. In fact, $[\mathrm{Fe/H}]_i$ is monotonically related to the derived $T_{\rm eff}$, with a mean offset between the cool and warm stars of ≈ 0.15 dex. Similar systematic behavior is seen in $[\alpha/\mathrm{Fe}]_i$. There are several possible reasons for these offsets including mismatches between the `MIST` isochrones, limitations in the atomic line lists, non-LTE effects, or our choice to fix the microturbulence parameter (see Section 3.3.3 below for additional discussion). Trends of this magnitude as a function of $\log g$ or $T_{\rm eff}$ are common in the literature (e.g., Valenti & Fischer 2005; Adibekyan et al. 2012;

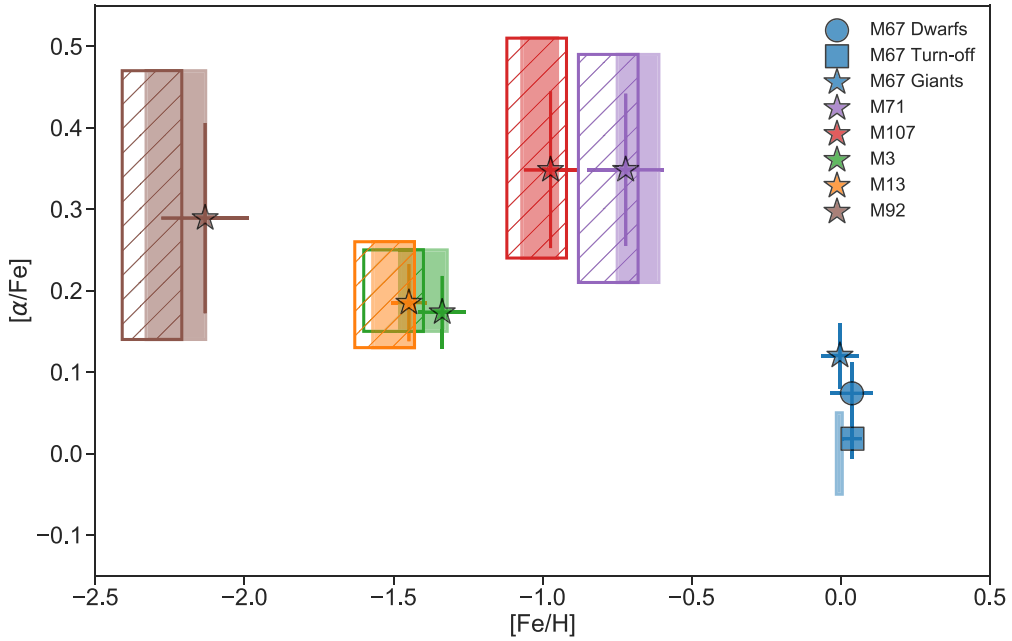


Figure 12. Distributions of surface metallicity and α-abundance ([α/Fe]) for the open cluster M67 (blue) and globular clusters M71 (purple), M107 (red), M3 (green), M13 (orange), and M92 (brown). The star symbols and associated error bars are the weighted median and 1σ values from the `MINESweeper` analysis. The ranges of literature values for each cluster are given by the shaded and hatched regions (for the metallicity scales of Carretta et al. 2009a and Mészáros et al. 2015, respectively). The M67 results are presented separately for dwarfs (circle), TO stars (square), and giants (star). Our [Fe/H] values are in good agreement with the literature, given the ≈0.1 dex systematic uncertainty in the overall metallicity scale. It is difficult to draw meaningful conclusions from the comparison in [α/Fe] given the large range in literature estimates and the inconsistent definitions of "α."

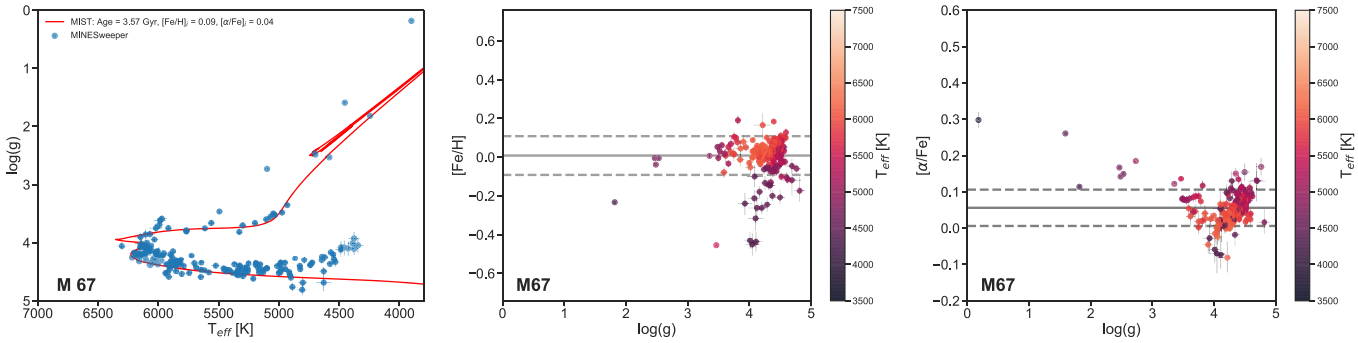


Figure 13. Stellar parameters for M67 cluster members derived without isochrone constraints. The Kiel diagram (left panel) shows reasonable agreement with the expected isochrone locus (red line), though there are several misplaced giants and an upturn in $\log g$ among the coolest dwarfs (compare with Figure 8). Surface [Fe/H] (middle panel) and surface [α/Fe] (right panel) are shown as a function of $\log g$ (compare with Figures 10 and 11). The star with $\log g$<1 has [Fe/H] < -1, which is below the bottom range of the middle panel. It is noteworthy that the trend between [α/Fe] and $\log g$ is very similar to the case with isochrone constraints (Figure 11), suggesting that that this behavior is due to a limitation in our spectral models.

Holtzman et al. 2015), and in particular for studies focusing on M67 (Gao et al. 2018; Souto et al. 2019).

M71 shows broadly similar behavior to M67 for [Fe/H]$_i$, in which the hotter stars prefer higher metallicities compared to the cooler giants. Among the giants, there is no noticeable systematic trend between [Fe/H]$_i$ and $\log g$. Moreover, Figure 11 shows no trend between [α/Fe]$_i$ and either $\log g$ or $T_{\rm eff}$. However, the uncertainties are larger when compared to the other clusters, as the S/Ns of the M71 spectra are the lowest among our sample (M71 also has high reddening of $A_{\rm V} \approx 0.7$).

The remaining four clusters show broadly similar behavior. Aside from a small number of TO stars in M92, the cluster members are confined to $0 < \log g < 3$. Among the giants there is no mean trend between [Fe/H]$_i$ and $\log g$ in any of M107, M3, M13, and M92. Moreover, the scatter among the measured [Fe/H]$_i$ is <0.1 dex for all four clusters. Similar conclusions hold for [α/Fe]$_i$, although the scatter is slightly larger for M107 and M92.

Figure 12 shows our measured [Fe/H] and [α/Fe] values for the six clusters along with 68th percentile ranges. Here we show the average surface abundances in order to compare to literature parameters. For M67 we divided the sample into dwarfs ($\log g \geqslant 4.4$), TO stars ($4.0 < \log g < 4.4$), and giants ($\log g \leqslant 4.0$). For M67 the literature abundances are adopted from Liu et al. (2016), while for other clusters we show the results from both Mészáros et al. (2015) and Carretta et al. (2009a). These two sources have a ≈0.1 dex offset in their overall metallicity scale.

Based on these star cluster tests, we conclude that `MINESweeper` is able to recover the known abundances of [Fe/H] and [α/Fe] for clusters that span a wide range in metallicity. We do not detect any systematic behavior between

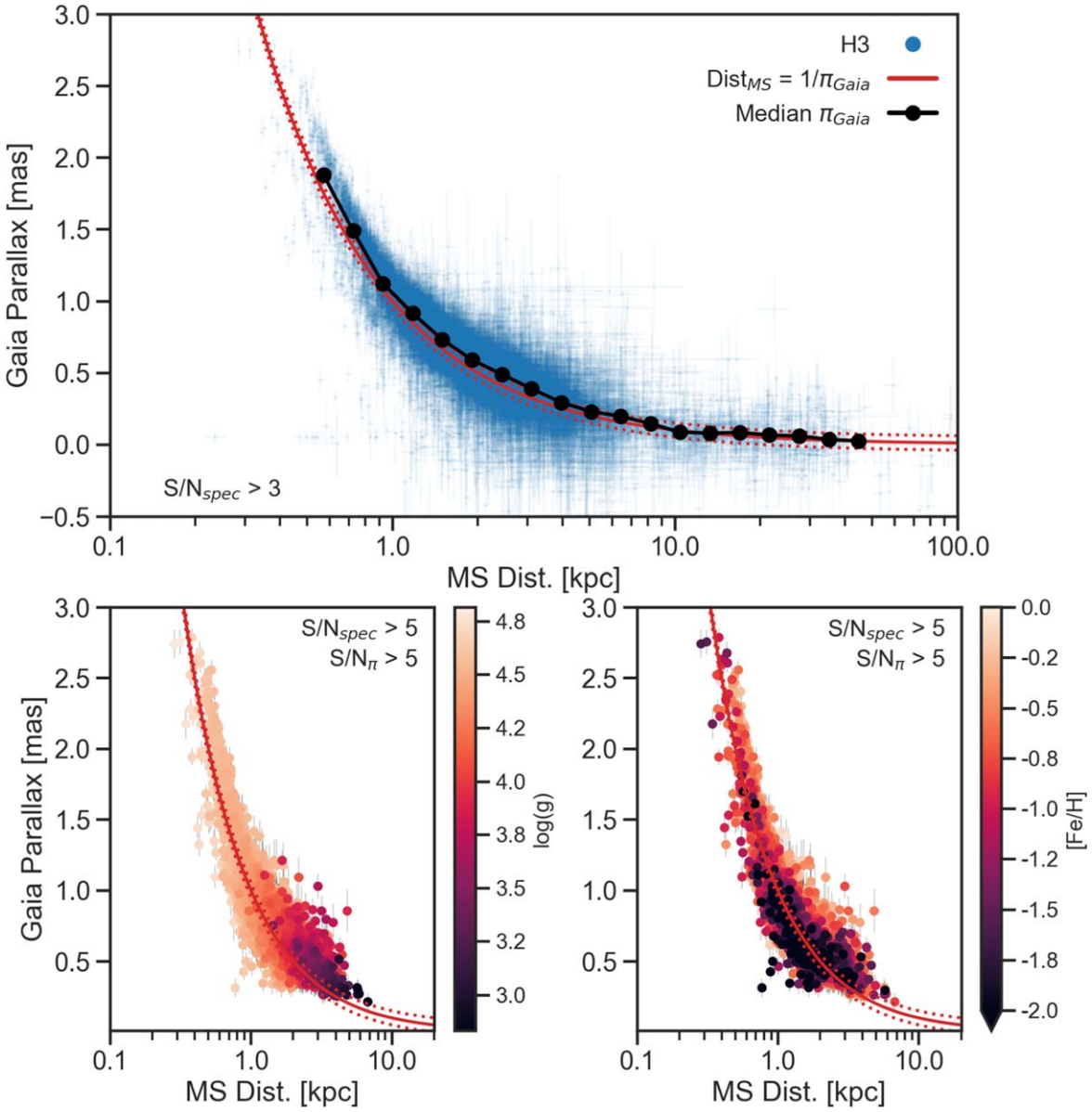


Figure 14. Comparison of Gaia parallaxes to measured `MINESweeper` distances for a set of stars observed as part of the H3 survey. The top panel shows the comparison for the full sample. The red line shows the one-to-one line, and dotted red lines indicate the Gaia DR2 zero-point uncertainty of 0.05 mas. Black symbols show the binned median Gaia parallax in `MINESweeper` distance bins. The errors on the binned medians are smaller than the data points. Bottom panels show the same comparison for a subset of the data with high-quality parallaxes and H3 spectra. The color of the points shows the individual stellar surface gravity (left) and metallicity (right). `MINESweeper` distances in this figure were computed *without* including Gaia parallax priors.

measured abundances and other stellar parameters ($\log g$ and $T_{\rm eff}$) among the giant populations of these clusters ($\log g < 3$). In the two clusters with dwarf stars (M67 and M71) we identify systematics at the 0.1–0.2 dex level. Resolving this issue will be the subject of future work.

3.3.3. Testing the Effect of the Isochrone Prior with M67

Figures 10 and 11 showed systematic offsets in the derived abundances as a function of $\log g$ within individual clusters where both giant and dwarf stars were available. In this section we test whether those offsets can be attributed to the requirement that stars lie on model isochrones. In particular, we have refit all of the M67 data using `MINESweeper` in a mode where the star is not required to lie on a model isochrone. In practice, this means that only the surface parameters ($T_{\rm eff}$, $\log g$, [Fe/H] [α/Fe], $v_{\rm rad}$, $v_{\rm stellar}$, along with $A_{\rm V}$) were fit to the data.

Figure 13 shows the result of this test. The Kiel diagram (left panel) shows overall good agreement with the expected isochrone locus. A few giants are displaced to substantially higher $\log g$ than allowed by the isochrones. More significantly, the cool dwarfs ($T_{\rm eff} < 5000$ K) display an upturn in the Kiel diagram that is in strong disagreement with the expected isochrone behavior. This behavior has been seen in many previous analyses (e.g., Adibekyan et al. 2012; Bensby et al. 2014; Holtzman et al. 2015), although the physical origin of the trend is uncertain.

The middle and right panels of Figure 13 show the surface [Fe/H] and [α/Fe] abundances versus $\log g$ for the fits where stars are not required to lie on a model isochrone. Compared with the default fits in Figure 10, the isochrone-free results

show increased scatter for the dwarfs and a few giants at very low metallicity ([Fe/H] < -1.0, which is off the plot in Figure 13). Perhaps more interesting is the behavior of [α/Fe] with $\log g$. Here the result is very similar to the default fits shown in Figure 11, although with slightly larger scatter. The fact that the U-shaped behavior is clear in both fitting modes strongly suggests that the cause of the behavior is due to the underlying spectral models. Possible reasons include microturbulence and non-LTE effects.

3.4. Application to the H3 Survey

In this section we present a final test of the MINESweeper-derived distances by comparing the spectrophotometric distances to Gaia parallaxes. For this test we utilize data collected through the H3 survey (Conroy et al. 2019). The survey is obtaining spectra of 200,000 stars in high-latitude fields ($|b| > 30°$) in order to map the stellar halo of the Galaxy. H3 employs the Hectochelle instrument on the MMT, which is delivering $R \approx 32{,}000$ spectra over the wavelength range 5150 −5300 Å. The targets are relatively bright ($15 < r < 18$) and hence have good SED coverage from Pan-STARRS, Gaia, 2MASS, and WISE (and SDSS where available). The main sample selection function is simple, consisting only of magnitude and Gaia parallax cuts.

MINESweeper is being used to derive stellar parameters for the H3 survey. In its default mode, the Gaia parallaxes are used as a prior in the fitting. However, in order to provide an independent test of the MINESweeper-based distances, we have refit a subset of the H3 data excluding the Gaia parallaxes.

The result of this test is shown in Figure 14. In the top panel we show all stars that have $S/N_{spec} > 3$, H3 quality flag $= 0$, and a stellar broadening $v_{stellar} < 5\,km\,s^{-1}$. The last cut was necessary, as there is a failure mode where MINESweeper occasionally attempts to fit a giant solution with a large broadening, when the star is in fact a dwarf (using the Gaia parallax as a prior tends to alleviate this failure mode). The one-to-one line is shown as a red line, and binned medians and error on the medians are shown as black circles and error bars (the latter are generally too small to be seen in the plot). We plot the Gaia parallaxes with a zero-point offset of 0.05 mas (Leung & Bovy 2019; Schönrich et al. 2019); in addition, red dotted lines show ±0.05 mas systematic uncertainties in the Gaia parallaxes (Lindegren et al. 2018; Schönrich et al. 2019). The bottom panels show results for a subset of data with higher spectroscopic S/N and with a Gaia parallax $S/N_{\pi} > 5$. In the bottom panels the points are color-coded by $\log g$ (left panel) and [Fe/H] (right panel).

Overall the agreement is good, especially at $d > 1$ kpc, which provides an independent check that the MINESweeper-based distances are reliable. At $d < 1$ kpc the MINESweeper-based distances tend to be overestimates compared to Gaia parallaxes. From inspection of the bottom panels it is clear that the discrepancy can be traced to metal-rich dwarf stars. Such stars are challenging to model not only because their cool temperatures result in strong atomic and molecular absorption feature but also because stellar isochrones have difficulty reproducing observations in this regime (e.g., Feiden & Chaboyer 2012; Choi et al. 2016).

We provide one additional test of the MINESweeper pipeline. The uniform H3 field placement resulted in serendipitous observations of the Sextans dwarf spheroidal galaxy. These stars were visually identified as a large overdensity of distant stars in one field. Closer inspection revealed 26 Sextans stars with $S/N > 2$ at the radial velocity associated with Sextans (224 km s^{-1}; Walker et al. 2009). The mean S/N of these stars is on the low end of the survey (3.3), and so they offer a good test of our ability to recover distances and metallicities at low S/N. Furthermore, being so distant, these stars all have low $\log g$ ($0.7 < \log g < 1.4$).

For the 26 Sextans members we find a mean heliocentric distance of 83 kpc with a 1σ scatter of 8 kpc. The mean reported 1σ error for this sample is 5 kpc, implying that the reported errors are underestimated by 60%. Lee et al. (2009) report a distance to Sextans of 86 kpc based on CMD fitting that included the horizontal branch and MS TO. We therefore conclude that our reported distances are accurate. Furthermore, the scatter among the individual measurements implies a distance precision of 10% for upper giant branch stars even at low S/N.

We turn now to the metallicities for Sextans. We find a mean metallicity of $\langle$[Fe/H]$\rangle = -1.92$ for our 26 Sextans members. Kirby et al. (2011) report a mean metallicity of −1.93 for Sextans (see also Walker et al. 2009). From this comparison we conclude that our metallicities are accurate, even for low-$\log g$ giants at low S/N.

4. Summary and Future Directions

In this paper we have presented MINESweeper, a new program to estimate stellar parameters and distances by combining broadband photometry, spectra, and parallaxes. We presented a suite of mock tests to gain intuition for the measurement precision as a function of stellar evolutionary phase, $[Fe/H]_i$, S/N_{spec}, and S/N_{π}. A key result is that at low parallax S/N (<1), distances can be determined to better than 20% even with low-S/N spectra, and as precise as 10% for RGB and 5% for MS stars.

We then presented several tests of the method using data. Moderate-resolution, high-S/N spectra of seven benchmark stars, including Arcturus, Procyon, and the Sun, were fit, and the resulting parameters agree well with literature values in most cases. Moderate-resolution spectra and broadband photometry of star clusters spanning a wide range of metallicities were also fit. The derived metallicities, [α/Fe], distances, and ages also agree well with literature estimates. We identified a modest level of systematic uncertainties in the $[Fe/H]_i$ and $[\alpha/Fe]_i$ values as a function of $\log g$ and T_{eff} within each cluster, at the ≈0.1 dex level. The systematic behavior is mostly confined to differences between dwarfs and giants (or low and high T_{eff}). The magnitude of the systematic behavior is similar to what has been reported in previous work (e.g., Cohen & Meléndez 2005; Holtzman et al. 2015; Mészáros et al. 2015). Systematic behavior is not apparent within the giant branch. Finally, we fit 22,000 stellar spectra from the H3 survey (Conroy et al. 2019). In these fits the Gaia parallaxes were not used, and the resulting measured distances were then compared to the independent Gaia parallaxes. The comparison showed good agreement as a function of distance, $\log g$, and [Fe/H]. We also showed that with H3 data MINESweeper recovers accurate distances and metallicities for the Sextans dwarf galaxy located at 86 kpc and with a mean metallicity of −1.9.

There are several directions for future improvement. On the spectral modeling side, we currently have fixed the microturbulence parameter. However, this parameter is known to vary with spectral type and evolutionary phase (e.g., Ramírez et al. 2013;

Jofré et al. 2014) and so should be included as an additional parameter in the fit. Stacked residuals have revealed several locations where the line list needs improvement.

On the fitting side, stellar binarity is not currently included in MINESweeper. However, it should be possible to identify binaries in cases where the Gaia parallax S/N is large and the star is on the MS. In such cases the "star" will sit above the MS owing to the combined light of the two stars (e.g., Widmark et al. 2018). A way forward is to model the stellar spectrum and SED with a combined model for two stars parameterized by a ratio of secondary to primary mass (for a similar binary modeling approach, see El-Badry et al. 2018). Finally, it would be straightforward to include prior information beyond parallax, e.g., interferometric and/or asteroseismic constraints (e.g., Mints & Hekker 2017).

MINESweeper was developed specifically to measure stellar parameters for the H3 survey, but it is a general tool that can be applied to spectrophotometric data of any kind. It is the ideal tool for extracting the maximum amount of information from ongoing and future stellar spectroscopic surveys in the Gaia era and beyond.

We thank Joseph Rodriguez for helping us with the name MINESweeper, and Robert Kurucz for his extensive and important work managing and updating his archive of atomic and molecular transition data and the ATLAS/SYNTHE suite of programs for the general astronomical community. Y.-S.T. is grateful to be supported by the NASA Hubble Fellowship grant *HST*-HF2-51425.001 awarded by the Space Telescope Science Institute.

Observations reported here were obtained at the MMT Observatory, a joint facility of the Smithsonian Institution and the University of Arizona. We thank the Hectochelle operators Chun Ly, ShiAnne Kattner, Perry Berlind, and Mike Calkins, as well as the support staff at the CfA and U. of Arizona. The computations in this paper were run on the Odyssey/Cannon CPU cluster supported by the FAS Division of Science, Research Computing Group at Harvard University. We thank the staff at Harvard FAS research computing, in particular Scott Yockel and Paul Edmon, for their continued support and troubleshooting.

This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

ORCID iDs

Phillip A. Cargile https://orcid.org/0000-0002-1617-8917
Charlie Conroy https://orcid.org/0000-0002-1590-8551
Yuan-Sen Ting (丁源森) https://orcid.org/0000-0001-5082-9536
Ana Bonaca https://orcid.org/0000-0002-7846-9787
Aaron Dotter https://orcid.org/0000-0002-4442-5700
Joshua S. Speagle https://orcid.org/0000-0003-2573-9832

References

Adibekyan, V. Z., Sousa, S. G., Santos, N. C., et al. 2012, A&A, 545, A32
Anders, F., Khalatyan, A., Chiappini, C., et al. 2019, A&A, 628, A94
Anderson, L., Hogg, D. W., Leistedt, B., Price-Whelan, A. M., & Bovy, J. 2018, AJ, 156, 145
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58
Benedict, G. F., McArthur, B. E., Feast, M. W., et al. 2011, AJ, 142, 187
Bensby, T., Feltzing, S., & Oey, M. S. 2014, A&A, 562, A71
Binney, J., Burnett, B., Kordopatis, G., et al. 2014, MNRAS, 437, 351
Blanco-Cuaresma, S., Soubiran, C., Jofré, P., & Heiter, U. 2014, A&A, 566, A98
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Breddels, M. A., Smith, M. C., Helmi, A., et al. 2010, A&A, 511, A90
Burnett, B., & Binney, J. 2010, MNRAS, 407, 339
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carretta, E., Bragaglia, A., Gratton, R., D'Orazi, V., & Lucatello, S. 2009a, A&A, 508, 695
Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009b, A&A, 505, 117
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Cohen, J. G., & Meléndez, J. 2005, AJ, 129, 303
Connelly, J. N., Bizzarro, M., Krot, A. N., et al. 2012, Sci, 338, 651
Conroy, C., Bonaca, A., Cargile, P., et al. 2019, ApJ, 883, 107
Cordero, M. J., Pilachowski, C. A., Johnson, C. I., & Vesperini, E. 2015, ApJ, 800, 3
Creevey, O., Grundahl, F., Thévenin, F., et al. 2019, A&A, 625, A33
Creevey, O. L., Thévenin, F., Boyajian, T. S., et al. 2012, A&A, 545, A17
Deras, D., Arellano Ferro, A., Lázaro, C., et al. 2019, MNRAS, 486, 2791
Deras, D., Arellano Ferro, A., Muneer, S., Giridhar, S., & Michel, R. 2018, AN, 339, 603
Dotter, A. 2016, ApJS, 222, 8
Dotter, A., Conroy, C., Cargile, P., & Asplund, M. 2017, ApJ, 840, 99
Dotter, A., Sarajedini, A., Anderson, J., et al. 2010, ApJ, 708, 698
El-Badry, K., Rix, H.-W., Ting, Y.-S., et al. 2018, MNRAS, 473, 5043
Evans, D. W., Riello, M., de Angeli, F., et al. 2018, A&A, 616, A4
Feiden, G. A., & Chaboyer, B. 2012, ApJ, 757, 42
Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, A&A, 616, A10
Gaia Collaboration, Helmi, A., van Leeuwen, F., et al. 2018b, A&A, 616, A12
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gao, X., Lind, K., Amarsi, A. M., et al. 2018, MNRAS, 481, 2666
Gratton, R., Sneden, C., & Carretta, E. 2004, ARA&A, 42, 385
Harris, W. E. 2010, arXiv:1012.3224
Heiter, U., Jofré, P., Gustafsson, B., et al. 2015, A&A, 582, A49
Hernitschek, N., Cohen, J. G., Rix, H.-W., et al. 2019, ApJ, 871, 49
Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv:1008.4686
Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, AJ, 150, 148
Jofré, P., Heiter, U., Soubiran, C., et al. 2014, A&A, 564, A133
Jofré, P., Heiter, U., Soubiran, C., et al. 2015, A&A, 582, A81
Jofré, P., Heiter, U., & Soubiran, C. 2019, ARA&A, 57, 571
Jørgensen, B. R., & Lindegren, L. 2005, A&A, 436, 127
Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864
Kirby, E. N., Lanfranchi, G. A., Simon, J. D., Cohen, J. G., & Guhathakurta, P. 2011, ApJ, 727, 78
Kroupa, P. 2001, MNRAS, 322, 231
Kurucz, R. L. 1970, SAO Special Rep., 309
Kurucz, R. L. 1993, SYNTHE Spectrum Synthesis Programs and Line Data (Cambridge, MA: Smithsonian Astrophysical Observatory)
Kurucz, R. L., & Avrett, E. H. 1981, SAO Special Rep., 391
Lee, M. G., Yuk, I.-S., Park, H. S., Harris, J., & Zaritsky, D. 2009, ApJ, 703, 692
Leistedt, B., & Hogg, D. W. 2017, AJ, 154, 222
Leung, H. W., & Bovy, J. 2019, MNRAS, 489, 2079
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Liu, F., Asplund, M., Yong, D., et al. 2016, MNRAS, 463, 696
Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9
Maíz Apellániz, J., & Weiler, M. 2018, A&A, 619, A180
Mészáros, S., Martell, S. L., Shetrone, M., et al. 2015, AJ, 149, 153
Mints, A., & Hekker, S. 2017, A&A, 604, A108
Mints, A., & Hekker, S. 2018, A&A, 618, A54
Paszke, A., Gross, S., Chintala, S., et al. 2019, arXiv:1912.01703
Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3
Pont, F., & Eyer, L. 2004, MNRAS, 351, 487
Portillo, S. K. N., Speagle, J. S., & Finkbeiner, D. P. 2020, AJ, 159, 165
Pritzl, B. J., Venn, K. A., & Irwin, M. 2005, AJ, 130, 2140
Ramírez, I., Allende Prieto, C., & Lambert, D. L. 2013, ApJ, 764, 78
Rodriguez, J. E., Zhou, G., Vanderburg, A., et al. 2017, AJ, 153, 256
Sahlholdt, C. L., Feltzing, S., Lindegren, L., & Church, R. P. 2019, MNRAS, 482, 895

Sanders, J. L., & Das, P. 2018, MNRAS, 481, 4093
Sandquist, E. L., Mathieu, R. D., Quinn, S. N., et al. 2018, AJ, 155, 152
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schönrich, R., & Bergemann, M. 2014, MNRAS, 443, 698
Schönrich, R., McMillan, P., & Eyer, L. 2019, MNRAS, 487, 3568
Smiljanic, R., Korn, A. J., Bergemann, M., et al. 2014, A&A, 570, A122
Souto, D., Allende Prieto, C., Cunha, K., et al. 2019, ApJ, 874, 97
Speagle, J. S. 2020, MNRAS, 493, 3132
Stello, D., Vanderburg, A., Casagrande, L., et al. 2016, ApJ, 832, 133
Szentgyorgyi, A., Furesz, G., Cheimets, P., et al. 2011, PASP, 123, 1188
Takeda, G., Ford, E. B., Sills, A., et al. 2007, ApJS, 168, 297
Ting, Y.-S., Conroy, C., Rix, H.-W., & Cargile, P. 2019, ApJ, 879, 69
Valenti, J. A., & Fischer, D. A. 2005, ApJS, 159, 141
Walker, M. G., Mateo, M., & Olszewski, E. W. 2009, AJ, 137, 3100
Wang, J., Shi, J., Zhao, Y., et al. 2016, MNRAS, 456, 672
Widmark, A., Leistedt, B., & Hogg, D. W. 2018, ApJ, 857, 114
Xue, X.-X., Ma, Z., Rix, H.-W., et al. 2014, ApJ, 784, 170
Xue, X.-X., Rix, H.-W., Ma, Z., et al. 2015, ApJ, 809, 144
Yakut, K., Zima, W., Kalomeni, B., et al. 2009, A&A, 503, 165