`StarHorse`: a Bayesian tool for determining stellar masses, ages, distances, and extinctions for field stars

A. B. A. Queiroz,[1,2]★ F. Anders,[3,2] B. X. Santiago,[1,2] C. Chiappini,[2,3] M. Steinmetz,[3]
M. Dal Ponte,[1,2] K. G. Stassun,[4] L. N. da Costa,[2,5] M. A. G. Maia,[2,5] J. Crestani,[1,2]
T. C. Beers,[6] J. G. Fernández-Trincado,[7,8] D. A. García-Hernández,[9,10]
A. Roman-Lopes[11] and O. Zamora[9,10]

[1]*Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, Porto Alegre, RS - 91501-970, Brazil*
[2]*Laboratório Interinstitucional de e-Astronomia, LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[3]*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[4]*Vanderbilt University, Department of Physics & Astronomy, VU Station B 1807, Nashville, TN 37235, USA*
[5]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[6]*Department of Physics and JINA Center for the Evolution of the Elements, University of Notre Dame, Notre Dame, IN 46556, USA*
[7]*Departamento de Astronomía, Universidad de Concepción, Casilla 160-C, Concepción, Chile*
[8]*Institut Utinam, CNRS UMR6213, Univ. Bourgogne Franche-Comté, OSU THETA, Observatoire de Besançon, BP 1615, F-25010 Besançon Cedex, France*
[9]*Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[10]*Universidad de La Laguna (ULL), Departamento de Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[11]*Departamento de Física, Facultad de Ciencias, Universidad de La Serena, Cisternas 1200, La Serena, Chile*

ABSTRACT

Understanding the formation and evolution of our Galaxy requires accurate distances, ages, and chemistry for large populations of field stars. Here, we present several updates to our spectrophotometric distance code, which can now also be used to estimate ages, masses, and extinctions for individual stars. Given a set of measured spectrophotometric parameters, we calculate the posterior probability distribution over a given grid of stellar evolutionary models, using flexible Galactic stellar-population priors. The code (called `StarHorse`) can accommodate different observational data sets, prior options, partially missing data, and the inclusion of parallax information into the estimated probabilities. We validate the code using a variety of simulated stars as well as real stars with parameters determined from asteroseismology, eclipsing binaries, and isochrone fits to star clusters. Our main goal in this validation process is to test the applicability of the code to field stars with known *Gaia*-like parallaxes. The typical internal precisions (obtained from realistic simulations of an APOGEE+*Gaia*-like sample) are $\simeq 8$ per cent in distance, $\simeq 20$ per cent in age, $\simeq 6$ per cent in mass, and $\simeq 0.04$ mag in A_V. The median external precision (derived from comparisons with earlier work for real stars) varies with the sample used, but lies in the range of $\simeq [0, 2]$ per cent for distances, $\simeq [12, 31]$ per cent for ages, $\simeq [4, 12]$ per cent for masses, and $\simeq 0.07$ mag for A_V. We provide `StarHorse` distances and extinctions for the APOGEE DR14, RAVE DR5, GES DR3, and GALAH DR1 catalogues.

Key words: stars: distances – stars: fundamental parameters – stars: statistics – Galaxy: stellar content.

1 INTRODUCTION

Improving the accuracy and precision of stellar distances and ages, as well as individual interstellar extinction measurements, is one of the major tasks of stellar astrophysics in the *Gaia* era. Although the parallaxes from the first data release of the *Gaia* mission (Gaia Collaboration et al. 2016)[1] provide a major improvement for stars in the solar vicinity ($d \lesssim 200$ pc), they do not yet reach the precision of spectrophotometric methods for the much larger distances probed by spectroscopic stellar surveys. Even after the final *Gaia* data release, foreseen for 2022, spectrophotometry will provide more precise distances for stars beyond 10 kpc.

★ E-mail: anna.queiroz@ufrgs.br

[1] http://sci.esa.int/gaia/

© 2018 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

A large amount of spectroscopic data for individual stars has become available in recent years from dedicated surveys such as the Sloan Extension for Galactic Understanding and Exploration (SEGUE; Yanny et al. 2009), the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al. 2017), the RAdial Velocity Experiment (RAVE; Steinmetz et al. 2006), the Galactic Archaeology with HERMES survey (GALAH; Martell et al. 2017), the LAMOST Experiment for Galactic Understanding and Exploration (LEGUE; Deng et al. 2012), and the *Gaia* ESO Survey (GES; Gilmore 2012). The combination of such data sets with broad-band photometric data and the astrometric solutions from *Gaia* allow for a much more detailed modelling of the chemodynamical history of the Milky Way. On the one hand, *Gaia*'s proper motions and parallaxes, complemented with radial-velocity measurements, enable us to measure stellar phase-space distribution functions with unprecedented precision over a Galactic volume of $\sim$8000 kpc^3. *Gaia*'s parallaxes (in combination with spectroscopy) also help to estimate stellar masses and ages. And for more distant populations, more accurate spectrophotometric distances can be achieved by improved calibrations in the *Gaia* volume. Such distance estimates are indispensable for mapping the chemical and kinematical properties of Galactic populations using large stellar surveys (e.g. Boeche et al. 2013, 2014; Anders et al. 2014; Mikolaitis et al. 2014; Nidever et al. 2014; Recio-Blanco et al. 2014; Carlin et al. 2015; Hayden et al. 2015; Bovy et al. 2016). If stellar ages are available as well, they can be used to probe the chemodynamical evolution of different Galactic components much more directly (e.g. Zoccali et al. 2003; Haywood et al. 2013; Mitschang et al. 2014; Anders et al. 2017; Mackereth et al. 2017).

In Santiago et al. (2016), we presented a Bayesian inference code to determine spectrophotometric distances for large survey samples, both in the optical and near-infrared (NIR). Since then, our group has extended that algorithm in several ways, improving the code's flexibility for different input data, updating priors and likelihood functions, and adding extinction, ages, and masses as parameters to be inferred by the method. In this paper, we demonstrate these new capabilities. The paper is structured as follows. In Section 2, we describe our method to estimate stellar parameters, distances, and extinctions. Section 3 presents the recent updates to our code. We provide an analysis of the performance of our code in terms of internal accuracy and precision in Section 4, using simulated stars, focusing especially on the new parameters mass, age, and extinction. We also discuss how biased spectroscopic parameters influence the estimated quantities. In Section 5, we compare our distances to several previous mass, age, and distance determinations that can be used as a reference. In Section 6, we describe a `StarHorse` application to a few spectroscopic surveys, with the purpose of delivering public releases of distances and extinction, We refrain from releasing ages and masses for the time being since their accuracy is still dependent on availability of *Gaia*-DR1 parallaxes and additional improvements. *Gaia*-DR2 will certainly improve their application to large volumes and data sets. We conclude the paper with a summary and future plans in Section 7.

2 THE METHOD

Our method uses a set of spectroscopically measured stellar parameters (typically effective temperature, T_{eff}, surface gravity, $\log g$, and overall metallicity [M/H]), photometric magnitudes, m_λ, and parallax, π, to estimate the mass, m_*, age, τ, distance, d, and extinction (in *V* band, A_V) for individual stars. The measured quantities are compared to predictions from stellar evolutionary models,

following a statistical approach that is similar to previous works (e.g. Burnett & Binney 2010; Burnett et al. 2011; Binney et al. 2014) and that generalizes the method presented in Santiago et al. (2016).

The calculations rely on three important assumptions. Most importantly, we assume that the stellar models are correct, which might not be true for metal-poor stars as well as other limitations in the current stellar models, i.e. that the object of interest follows a canonical single-star evolutionary track. We caution that this assumption, even if a star is apparently single, can be violated to various degrees in practice, leading first and foremost to incorrect stellar mass and/or age estimates (e.g. Brogaard et al. 2016; Yong et al. 2016; Fuhrmann & Chini 2017; Lagarde et al. 2017). The second assumption is that the observational uncertainties of the measured parameters follow a normal distribution. The third assumption is that the observed measurements are independent, a condition that can also be violated in practice.

We can then calculate the probability that a set of independent measured parameters $\vec{x} = \{x_1, ..., x_n\}$ with associated Gaussian uncertainties $\vec{\sigma}_x$ is drawn from a set of theoretical values $\vec{x}_0$, predicted by some model $\mathcal{M}$, by writing:

$$P(\vec{x}, \vec{\sigma}_x | \vec{x}_0) = \prod_i \frac{1}{\sqrt{2\pi}\,\sigma_{x_i}} \exp\left[-\frac{(x_i - x_{0_i})^2}{2\sigma_{x_i}^2}\right]. \qquad (1)$$

The above expression is called the likelihood of measuring the set $\{\vec{x}, \vec{\sigma}_x\}$, given a model $\mathcal{M}(\vec{x}_0[\vec{\theta}])$. Using Bayes's theorem, we now compute the posterior probability distribution (the probability of the model, given the data) as

$$P(\vec{x}_0 | \vec{x}, \vec{\sigma}_x) = \frac{P(\vec{x}, \vec{\sigma}_x | \vec{x}_0) P(\vec{x}_0)}{P(\vec{x}, \vec{\sigma}_x)}. \qquad (2)$$

The numerator contains the likelihood and the model priors $P(\vec{x}_0)$, and the denominator is the marginalized likelihood. It depends only on the measured parameters, being a constant through all the models that can be normalized out.

In our case, the model family $\mathcal{M}(\vec{\theta})$ consists of a grid of stellar models computed for different ages, metallicities, and initial masses, convolved with a grid of distances and extinctions (modifying the apparent magnitudes of each stellar model). To evaluate the probability of some specific model quantity, ϑ (usually one that cannot be measured directly), we now compute the marginal posterior probability distribution function (PDF) for this quantity, by integrating over all variables of equation (2), except ϑ:

$$p(\vartheta) := P(\vartheta | \vec{x}, \vec{\sigma}_x) = \int dx_{0,0} \ldots dx_{0,n}\, P(\vec{x}_0[\vec{\theta}] | \vec{x}, \vec{\sigma}_x). \qquad (3)$$

As mentioned earlier, a typical set of measured parameters includes $\vec{x} = \{[\text{M/H}], T_{\text{eff}}, \log g, m_\lambda, \pi\}$, or any subset of these. The model parameters ϑ we compute are mass, m_*, age, τ, distance, d, and *V*-band extinction, A_V. Our code delivers various statistics for the desired quantities. As in Rodrigues et al. (2014) and Santiago et al. (2016), for each quantity we compute the median of the marginalized posterior probability distribution, $p(\vartheta)$ (equation 3), along with its 5 per cent, 16 per cent, 84 per cent, and 95 per cent percentiles.

3 CODE UPDATES

In this section, we explain the technical details of our code in more detail (for an overview, see the flow diagram in Fig. 1). We
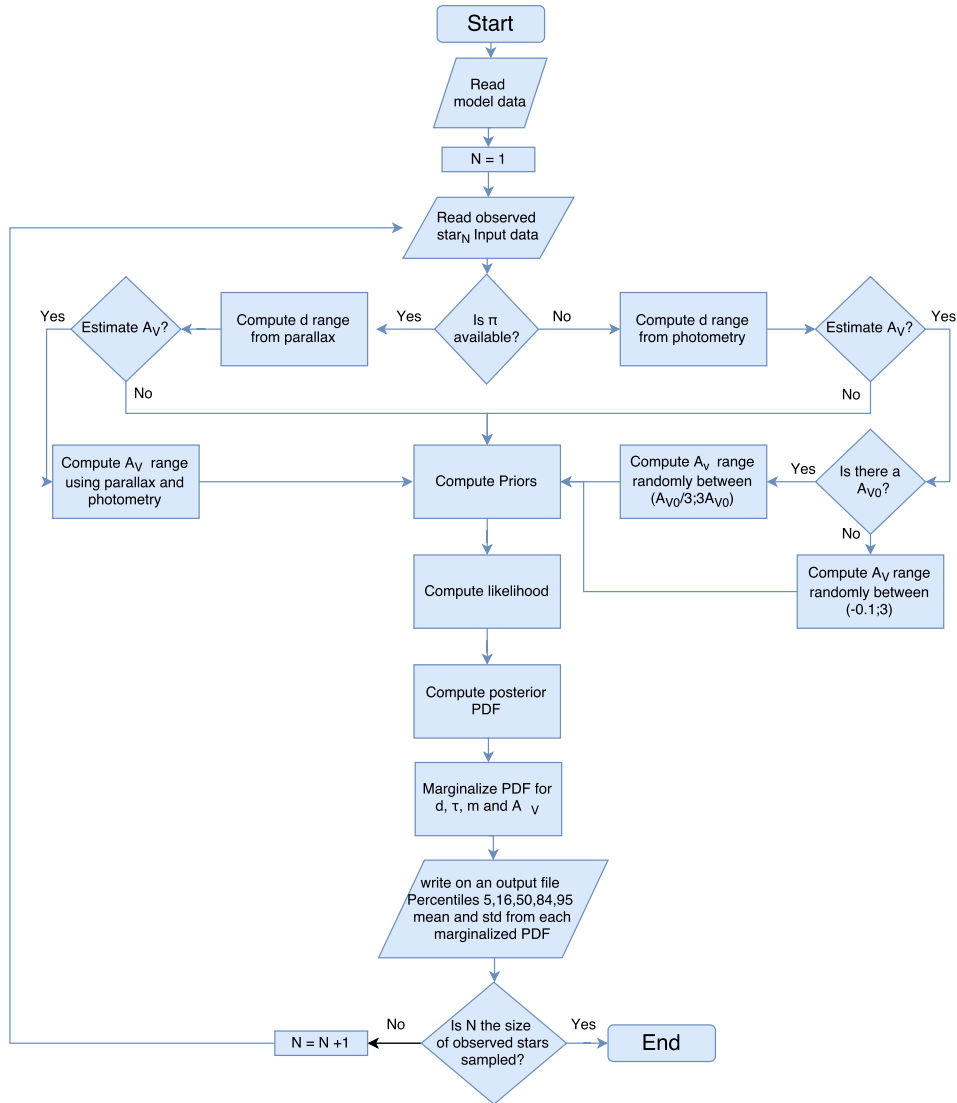
Figure 1. `StarHorse` flux diagram. See Sections 2 and 3 for details.

encourage the reader to contact the developers[2,3] for any questions or further details about the code. Via a parameter file, the user can choose the set of stellar models to be used, the available photometric and spectroscopic data, the treatment of extinction (whether to correct photometry for reddening or whether to include extinction as a parameter to be estimated), and the set of priors, amongst other options. Once the evolutionary models and the data are read in, the code operates according to the options chosen in the parameter file. These options, along with the other updates since Santiago et al. (2016), are detailed in the next subsections. Readers interested only in the overall performance of the code may skip these.

3.1 Including parallax as a measured parameter

To adapt our method to the new era of astrometric surveys like *Gaia*, JASMINE (Gouda 2012), VLBA (Melis et al. 2014), and

SKA (Imai et al. 2016), we introduced parallax as an optional measured input parameter for our code. As explained in the previous section, the likelihood can be extended for a generic group of measured parameters, so the method presents no difficulties to introduce the parallax in the likelihood, and it allows for much more precise estimates of stellar masses, ages, and extinction. When the user decides to use parallaxes as the primary input, we fix the range of distances for all models to be consistent with that measurement within 3σ (see Section 3.2.1). If this is not specified, the possible range of distances to be probed for each stellar model is derived by matching an observed apparent magnitude, m_λ (within $\pm 3\sigma_{m\lambda}$), in some filter to the corresponding model absolute magnitude.

3.2 Stellar parameters posterior

Currently our code can determine distances, ages, masses, and extinctions, given a set of measured parameters by marginalizing the joint posterior PDF. Below we explain in more detail how we build the values of distance, extinction, ages, and masses covered by the PDF.

[2] Anna Barbara Queiroz, Email: anna.queiroz@ufrgs.br.
[3] Friedrich Anders, Email: fanders@aip.de.

3.2.1 Distance

If no reliable parallaxes are available, or if they are only available for a subset of stars, the range of distances to be probed comes from the available measured apparent magnitudes. We choose a master filter, λ, and create an array of length N_d that ranges from $m_{\lambda 0} \pm 3\sigma_{m_{\lambda 0}}$, where $m_{\lambda 0}$ stands for intrinsic measured apparent magnitudes in the master filter. For each value of this array, we compute the distance modulus, $(m_{\lambda 0} - M_\lambda)$, for the absolute magnitudes in the model grid; these values are then finally transformed into an array of possible distances, d. When A_V is not being estimated together with distance, the code assumes that the given magnitudes are previously corrected by the known extinction. See Section 3.2.2 for details of the estimation of A_V. As explained in Section 3.1, if the user decides to use parallax measurements as the primary input, we build the distance array by inverting the array of allowed parallaxes. We then transform d to intrinsic distance moduli that do not depend on colour or extinction. The d and $(m - M)$ values are then used in the priors and likelihood to build the posterior PDF.

3.2.2 Extinction

When multiband photometry over a sufficient wavelength range is available, one can use the measured colours of a star to estimate interstellar dust extinction. When the intrinsic magnitudes are constrained by spectroscopic measurements, this extinction measurement can become very precise (e.g. Rodrigues et al. 2014). Our code can now also be used to determine extinction towards stars, by adding another free dimension to the model space.

When parallaxes are not available, we build a distance moduli that comes from the apparent magnitude $(m_\lambda - M_\lambda)$, as explained in Section 3.2.1, but now the measured magnitudes are assumed not to be intrinsic. The distance moduli must then be corrected by an a priori unknown extinction: $(m_{\lambda 0} - M_\lambda) = (m_\lambda - M_\lambda) - A_\lambda$. For each stellar model and each possible distance modulus, $(m_\lambda - M_\lambda)$, we thus create N_{A_V} random A_V values from a previously defined range of possible A_V. If there is no initial guess of the A_V for the given star, this range of A_V values is kept fixed as $[-0.1, 3.0]$. If some expectation for A_V is available (an A_V prior, $A_{V,0}$), we probe extinction in the range $[A_{V,0}/3, 3 \cdot A_{V,0}]$. We then transform the A_V to A_λ values using a chosen extinction curve (Schlafly et al. 2016 by default; see Section 3.4 for a discussion, and subtract it from $(m_\lambda - M_\lambda)$]. Since the model space is large, we usually use $N_{A_V} = 3$ to lower the computational cost. As long as the spectroscopic measurements do not confine the solution to a very small volume in model space, the marginalized PDFs over extinction and distance remain well sampled.

When parallax information is available, the dereddened distance modulus array, $(m - M)_0$, is determined directly from the parallax. To determine the extinction, we then use the reddened distance modulus arrays built from the apparent magnitudes, $(m_\lambda - M_\lambda)$, and the difference between those two naturally delivers A_λ.

3.2.3 Masses and ages

Because masses and ages are quantities provided by the grid of evolutionary models, they are simply repeated over the additional dimensions of distance and extinction. Therefore, once we have a PDF from equation (3), we can directly estimate these parameters by marginalizing over the distance and extinction dimensions.
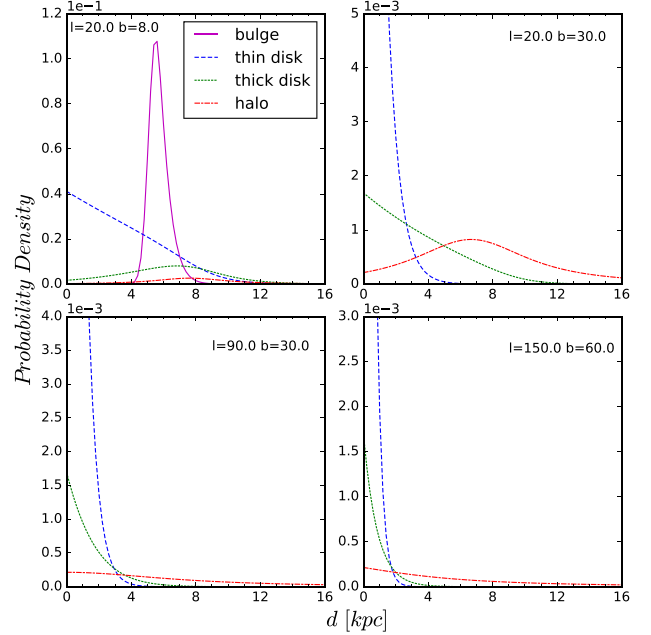
Figure 2. Spatial profile for the Galactic components used in the `StarHorse` priors. Each panel shows the behaviour of the prior probability density with distance for a given direction in Galactic coordinates (l, b).

3.3 New spatial, chemical, and age priors

Our code uses several priors that summarize our prior knowledge about the initial mass function (IMF) and the stellar population structure of the Milky Way, including the thin and the thick disc, and the stellar halo. For each Galactic component, these priors include the spatial density distribution, a metallicity distribution function (MDF) and an age distribution. In Santiago et al. (2016), we adopted the same priors as Binney et al. (2014). In the current version, we updated our default structural parameters (solar position, scale lengths, scale heights, normalisations) from Binney et al. (2014) to Bland-Hawthorn & Gerhard (2016).

Since APOGEE and other surveys are now probing also the inner regions of the Milky Way, we have also added simple spatial priors for the bulge/bar. The simplest choice of bulge spatial prior is a spherical exponential model with an exponential e-folding length of 0.5–1.0 kpc. We also added the oblate model described by Dehnen & Binney (1998). Finally, as our default model, we included a bar-like bulge model from Robin et al. (2012). Their models assume the bulge to be a triaxial ellipsoid, either boxy or discy (or yet a combination of both depending on the plane of projection), and with density laws that can be a $sech^2$, an exponential, or a Gaussian. Our code has been tested with one- and two-component model priors. Our default model is the S ellipsoid (bar component) taken from the 'S+E' case listed in table 2 of Robin et al. (2012). This 'S+E' model was the minimum likelihood one amongst the models presented by the authors. The E component was removed based on the revision of the thick-disc structure and its extrapolation towards the inner Galactic regions made by Robin et al. (2014), which effectively rendered the classical ellipsoid bulge unnecessary. We refer the interested reader to Robin et al. (2012) and Robin et al. (2014) for more details.

Fig. 2 shows the contributions of each Galactic component spatial distribution as a function of distance for four representative directions. In the upper left-hand panel, we show a direction

Component	Mean age (Gyr)	σ age (Gyr)	Mean [M/H]	σ[M/H] (dex)
Thin disc	5.0	4.0	-0.2	0.3
Thick disc	10.5	2.0	-0.6	0.5
Halo	12.5	1.0	-1.6	0.5
Bulge	10.0	3.0	0.0	0.5

towards the inner Galaxy $(l, b) = (20, 8)$; in this case, the bar/bulge component is the dominant population at heliocentric distances of ∼6 kpc, its density decreases rapidly towards greater distances. The upper right-hand panel shows a direction towards the same Galactic longitude, but at higher latitude. In this case, we miss the bulge and notice that each of the other components dominates at a certain distance range. The lower panels show directions away from the Galactic Centre. In the lower left-hand panel $(l, b) = (90, 30)$, we see that the contribution for the discs dominates out to 3.5 kpc; in the lower right-hand panel $(l, b) = (150, 60)$, the halo dominates already for $d > 2$ kpc.

Our new age and metallicity priors for the four Galactic components are all assumed to be Gaussians. The corresponding mean and standard deviation values for each case are provided in Table 1. The motivation for this change is twofold: (1) simplicity: they are simple functions, easily computed, which makes them ideal for the computationally intensive parameter estimate process used here; (2) they are made broad enough to accommodate most or all of the recent age and metallicity distributions found in the literature, which are not only diverse, but also often conflicting. We assume that the impact of this choice is minor, though we are aware that the age and metallicity distribution are not necessarily Gaussians; by taking this approach, we avoid making our priors too specific, but do not completely overlook the knowledge accumulated about the different Galactic components.

We note that the previous age priors from Santiago et al. (2016) assigned zero probability to disc stars older than 10 Gyr and to thick disc or halo stars younger than this value. Recent results found in the literature pose a challenge to such simple age step functions. One example is the discovery of young α-enhanced stars, likely thick-disc members (Chiappini et al. 2015). The previous metallicity priors were also narrower, specially for the thin disc, essentially ruling out any thin disc star more metal-poor than [Fe/H] $\simeq -0.6$. The changes made to these components also make them more in sync with our current understanding of the bulge populations, for which there is also recent evidence for a larger fraction of stars younger than $\simeq 5$ Gyr (Bensby et al. 2013; Valle et al. 2015).

3.4 A new default extinction curve

The adopted extinction curve, i.e. the dependence of the absolute extinction on wavelength, has of course an influence on the distances and extinctions provided by our code.

To quantify this effect, we recently implemented the one-parameter extinction curve presented by Schlafly et al. (2016), derived from APOGEE spectroscopy in combination with Pan-STARRS1, 2MASS, and WISE data, as an alternative to the Cardelli, Clayton & Mathis (1989) curve that was determined based on very few stars. For the mostly low-latitude APOGEE DR14 sample, where the sensitivity of our results to the chosen extinction law is arguably highest, we tested the effect of changing the extinction curve from Cardelli et al. (1989) to Schlafly et al. (2016).

Our tests showed that, if both optical (APASS) and near-infrared magnitudes are available (i.e. for the brighter stars), the systematic effect of the adopted extinction curve is almost negligible: median differences in the inferred parameters range in the 1 per cent regime, although there are some weak systematic trends with age. The main differences in the inferred parameters arise when only 2MASS magnitudes are available: in this case, the results using Schlafly et al. (2016) yield ∼5 per cent lower A_V, ∼9 per cent greater ages, ∼0.6 per cent greater distances, and ∼11 per cent higher A_V. We now use Schlafly et al. (2016) as StarHorse's default extinction curve, since this extinction curve is much more realistic, and its adoption also slightly improves the convergence of the code for high-extinction stars (i.e. distances to ∼10 000 more APOGEE DR14 stars could be found).

4 TESTS WITH SIMULATED STARS

To validate our code, we carried out several tests using simulated stars. Since one of our main goals is to apply StarHorse to field stars with reliable parallaxes to infer masses and ages, we include parallax in the set of observed quantities of our simulated stars and use them to constrain our distance range.

The first sample of simulated stars we used is identical with the set used in Santiago et al. (2016), except for the distance range and the assumed spectroscopic uncertainties. It consists of 5000 randomly drawn PARSEC models (Bressan et al. 2012), convolved with Gaussian errors in the spectrophotometric parameters and parallaxes (see details in Table 3). The error values for the 'high-resolution' version of this simulation were inspired by the spectroscopic uncertainties from the APOGEE DR14 results, while the low-resolution case was based on typical uncertainties from the SEGUE and RAVE surveys. In both cases, the samples contain stars with distances between 0.05 and 1 kpc (typical *Gaia*-TGAS distances), random Galactic positions, and extinctions from Schlegel, Finkbeiner & Davis (1998).

The PARSEC simulated stars are useful to map the internal accuracy and precision of our estimated distances, ages, masses, and extinctions over a wide input parameter range. The caveat is that the PARSEC sample sets are not representative of a real magnitude-limited sample of stars in our Galaxy. By randomly picking PARSEC models, we tend to oversample young stars, regardless of metallicity, relative to most survey data in the nearby Galaxy. This also means that our prior knowledge about Galactic stellar populations does not apply to these simulations, so that in this case we set all priors to unity.

To test the code in more realistic scenario, we also use a TRILEGAL (Girardi et al. 2012) population-synthesis simulation of an APOGEE-TGAS-like sample of giant stars. The details of this simulation are given in Table 2; we describe the main features briefly here. The underlying stellar models of TRILEGAL are from Marigo et al. (2008), which are similar but not identical to our default PARSEC 1.2S models. We used a Chabrier (2003) lognormal IMF for all Galactic components (thin disc, thick disc, bulge, and stellar halo), and the default spatial distribution, density normalization, star formation rate (SFR), and age–metallicity relation (AMR) for all components (see Table 2). Extinction was assumed to result from an exponential dust disc with calibration at infinity of $A_V = 0.0378$ mag for the Galactic poles, and the photometry is in the *UBVRIJHK* system (Maíz Apellániz 2006). The Solar position and Solar height above the disc were assumed to be $R_\odot = 8.7$ kpc and $Z_\odot = 24.2$ pc, respectively (deviant from the values of our Galactic priors, $R_\odot = 8.2$ kpc and $Z_\odot = 11.1$ pc; see Section 3.3).

Table 2. Adopted geometry, local or central calibration and SFR+AMR for the Galactic components simulated with TRILEGAL.

Component	Spatial distribution	Local/central calibration	SFR+AMR
Thin disc	Squared hyperbolic secant Girardi et al. (2005)	Local 55.40 $M_\odot$ pc^{-2}	2-Step age + Fuhrmann (1998) + α-enh Girardi et al. (2005)
Thick disc	Squared hyperbolic secant Girardi et al. (2005)	Local 0.001 $M_\odot$ pc^{-3}	11–12 Gyr const. + $Z = 0.008$ with δ[M/H] = 0.1 dex
Halo	Power law de Jong, Radburn-Smith & Sick (2009)	Local 0.0001 $M_\odot$ pc^{-3}	12–13 Gyr + Ryan, Norris & Beers (1996) [M/H] distribution
Bulge	Triaxial bulge Vanhollebeke, Groenewegen & Girardi (2009)	Central 406 $M_\odot$ pc^{-3}	10 Gyr + Zoccali et al. (2003) [M/H] + 0.3 dex

Table 3. Summary of the reference data: parameter ranges, uncertainties, and provenance.

Sample	$\sigma(\pi)$ (mas)	d range (kpc)	$\sigma(T_{\rm eff})$ (K)	$T_{\rm eff}$ range (K)	$\sigma \log g$	$\log g$ range	σ [M/H]	[M/H] range	σ mag	mag range (V mag)	filters
PARSEC high resolution	0.3	0.05–1	70.0	3000–7000	0.08	1–5	0.03	−2.5–0.5	0.025	−2–24	$BVgriJHK_s$
PARSEC low resolution	0.3	0.05–1	95.0	3000–7000	0.24	1–5	0.12	−2.5–0.5	0.025	−2–24	$BVgriJHK_s$
TRILEGAL	0.3	0.05–1	70.0	3000–7000	0.08	1–4.1	0.03	−2.5–0.5	0.025	4–13	$BVRIJHK_s$
Detached eclipsing binaries	0.04	0.01–65	80	4320–5730	0.02	1–3.6	0.12	−1.1–0.1	0.03	0.7–18	$BVRI$
Other eclipsing binaries	0.36	0.03–1	256.0	3880–30300	0.02	2.9–4.5	0.1	0.03–0.2	0.06	5–12	$BVRIJHK_s$
CoRoGEE stars	0.007	0.8–10	90.0	4000–5500	0.05	1.4–3.0	0.03	−2.5–0.5	0.025	11–16	$BVgriJHK_s$
OCCASO clusters stars	0.06	0.05–6	60.0	4300–5300	0.1	1.7–3.2	0.2	0.0–0.4	0.03	8–15	JHK_s

To simulate APOGEE-TGAS-like observations, we convolved the TRILEGAL stellar parameters with Gaussian errors, as in the PARSEC (*high-resolution*) sample, and introduced cuts in $\log g$ and distance. The uncertainty values and stellar parameter ranges for TRILEGAL sample are again listed in Table 3.

The main aim of using this TRILEGAL simulation was to test the impact of the different Galactic priors in our parameter estimates. Therefore, we ran StarHorse on the TRILEGAL sample for three prior configurations: (i) no spatial, age, and metallicity priors, only the IMF; (ii) IMF and spatial priors only; and (iii) all priors.

4.1 Internal accuracy and precision

Fig. 3 shows the results of our PARSEC simulated-star tests (high-resolution case). The first two rows show the relative errors in distance, $(d_{\rm SH} − d_{\rm True})/d_{\rm True}$, where $d_{\rm SH}$ are the distances estimated by our code (SH standing for StarHorse). Each panel shows these same errors as a function of a different parameter. The last panels on the right show the relative distance errors (top row) and uncertainties (second row) mapped on to the $\log g$ versus $T_{\rm eff}$ diagram. The first line of Table 4 shows the relative distance error values that correspond to the 5 per cent -ile, 16 per cent -ile, 50 per cent -ile, 84 per cent -ile, and 95 per cent -ile positions of the relative error distribution, when their signs are omitted. We also list the median value of the full error distribution to quantify the presence of systematic trends. For example, we see that 50 per cent of the simulated stars have distance errors of less than 6 per cent and 84 per cent have distance errors below 16 per cent. There is no strong systematic trend with any of the parameters, apart from an increase in the errors for larger distances and for lower $\log g$ values (giants). We also note that the discreteness of the PARSEC model grid used is visible in most of the panels.

The remaining rows of Fig. 3 show the same type of plots, but now with the relative errors in age, mass, and A_V. As before, the percentiles of the relative error distributions are listed in Table 4. As in the case of distances, the mass estimates (rows 5 and 6 in Fig. 3) do not suffer from any clear systematics with the main

parameters, apart from the trend of increasing errors with distance. There is a subset of mostly subgiant and dwarf stars ($\log g > 3.5$, and $m_{\rm ast} < 0.8 m_\odot$) with very well determined masses. From Table 4, we see that 50 per cent (84 per cent) of the estimated masses agree with the true values within 7 (22) per cent, and that the outliers are predominantly young (massive) evolved stars, which are rare in the Milky Way field.

As for ages, shown in the third and fourth rows of Fig. 3, catastrophic errors, of 100 per cent or more, occur for about 15 per cent of the stars. These stars are not restricted to a small subset of parameter space. But age is the main parameter leading to these catastrophic errors, which are more frequent for $\tau < 1$ Gyr. There is also some dependence on distance (i.e. low parallaxes, for which the parallax error is relatively large) and mass (or $\log g$). Since dwarfs change their positions in the spectroscopic Hertzsprung–Russell diagram only slightly on long time-scales, ages for dwarfs are more uncertain; the age PDF will tend to be very flat. The A_V errors shown in the last two rows are within 0.05 (0.1) mag for 50 per cent (84 per cent) of the stars, with no clear systematic effects.

Fig. 4 shows the results of our TRILEGAL test sample, in the same style as Fig. 3, for the case where all priors (IMF, spatial, age, and metallicity) are used. Although the stellar-parameter range of the TRILEGAL sample is similar to the PARSEC simulations, the resulting error distributions vary considerably when compared to the PARSEC results. There is a relative scarcity of low-mass and young stars in the TRILEGAL simulations when compared to PARSEC. Most dwarfs are removed by the APOGEE colour cut, and most red M dwarfs ($m_* < 0.8 m_\odot$, $\log g > 4$) are too faint to be seen. Likewise, most stars, even in the thin-disc component, are older than $\tau \simeq 500$ Myr. Relatively fast evolutionary stages, like post-HB phases, are also less frequent in the TRILEGAL simulations.

The first two rows of Fig. 4 show that there is a small trend towards underestimating distances overall, especially for hot dwarfs and subgiants (high $T_{\rm eff}$ and $\log g$ clouds in the upper row). Errors in all four parameters also tend to systematically increase with distance. Slight general trends are also observed in the sense of overestimating ages, and underestimating masses and extinction (these latter two

Parsec Simulated Stars
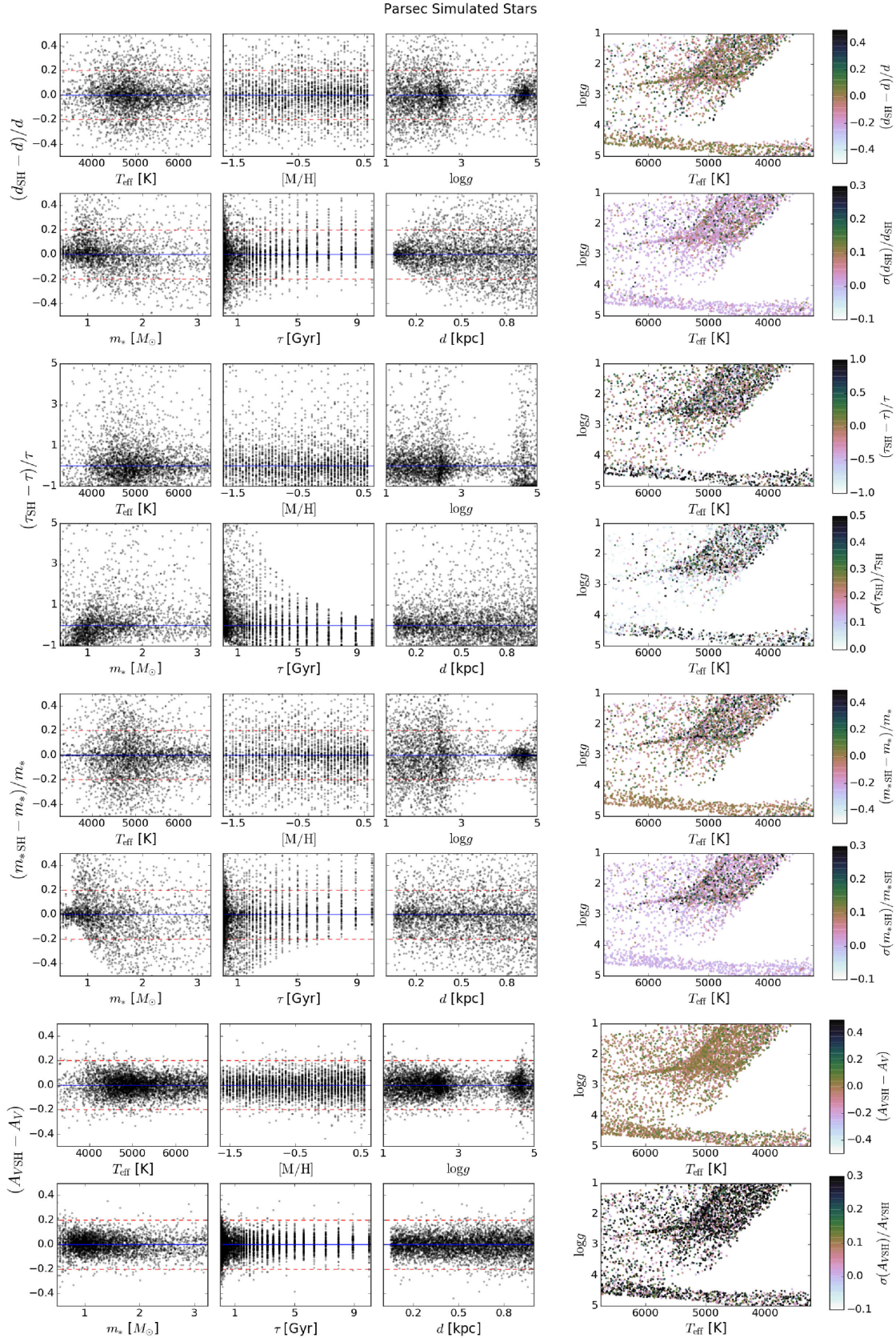
Figure 3. `StarHorse` relative errors for the PARSEC high-resolution simulations. For each parameter, the six panels on the left show the relative errors, d (top rows), τ (third and fourth rows), m_* (fifth and sixth rows), A_V (last rows) as a function of the true parameters. The solid blue line is the identity line, and the dashed red lines correspond to ± 20 per cent errors (except for ages: 40 per cent). The panels on the right show the relative errors (top panel) and uncertainties (bottom panel) in the $\log g$ versus $T_{\rm eff}$ plane.

Table 4. Results of the external validation tests. We report the percentiles of the modulus of the relative deviation distribution and the median of the relative deviation distribution for each parameter estimated by StarHorse. Data are shown for all reference samples used.

Sample	Parameter	P_5	P_{16}	P_{50}	P_{84}	P_{95}	Median
PARSEC	d	0.005	0.015	0.063	0.163	0.258	0.0
High resolution	τ	0.002	0.109	0.350	0.972	4.203	0.0
	m_*	0.001	0.012	0.074	0.222	0.391	0.0
	Av	0.005	0.014	0.048	0.107	0.162	0.0
PARSEC	d	0.005	0.016	0.060	0.172	0.289	-0.001
Low resolution	τ	0.014	0.110	0.398	1.120	4.550	0.002
	m_*	0.003	0.017	0.083	0.247	0.429	-0.001
	Av	0.016	0.049	0.169	0.354	0.499	0.004
TRILEGAL	d	0.007	0.023	0.079	0.171	0.240	-0.04
All priors	τ	0.017	0.054	0.188	0.516	0.937	0.11
	m_*	0.006	0.019	0.063	0.132	0.206	-0.04
	Av	0.003	0.012	0.043	0.099	0.162	0.007
Detached	d	0.001	0.002	0.017	0.041	0.047	-0.015
EBs	τ	0.009	0.022	0.246	0.764	1.94	0.22
	m_*	0.008	0.033	0.124	0.165	0.194	-0.12
	Av	0.017	0.030	0.072	0.186	0.354	-0.07
Other	d	0.032	0.071	0.178	0.298	0.419	-0.178
EBs	m_*	0.005	0.021	0.099	0.168	0.23	0.048
	Av	0.021	0.044	0.130	0.269	0.338	0.075
CoRoGEE	d	0.000	0.001	0.003	0.008	0.014	0.0
	τ	0.013	0.047	0.120	0.650	1.253	0.16
	m_*	0.003	0.011	0.043	0.097	0.164	0.0
	Av	0.009	0.024	0.070	0.158	0.270	-0.04
OCCASO	d	0.000	0.001	0.0244	0.130	0.310	0.0
Cluster members	τ	0.028	0.090	0.310	0.626	0.968	-0.07

parameters seem to be more affected for low $T_{\rm eff}$ stars). Still, typical errors are of order $\simeq 8$ per cent, $\simeq 19$ per cent, $\simeq 6$ per cent, and $\simeq 0.04$ mag, respectively, for distances, ages, masses, and A_V (Table 4). We also notice that the absence of a large number of young main-sequence stars significantly reduces the occurrence of catastrophic age errors in TRILEGAL.

To test the influence of our prior knowledge, the TRILEGAL results shown in Fig. 4 (all priors included) can be compared to Fig. A2 in Appendix A. That figure shows the corresponding results for the case when no priors were adopted. In general, the results in the *all priors* and in the *no priors* cases are quite similar, indicating that the accuracy of our code is robust to prior assumptions. We also tested the case where only the spatial priors were used, but not the age and MDF ones. As expected for this intermediate case, the panels are again very similar to those shown in Figs 4 and A2.

This is further corroborated by Fig. 5, where we show the relative error distributions of age, distance, mass, and A_V for the three combinations of priors, and separated for subgiants/hot dwarfs and giants. The histograms confirm the trends seen in the scatter plots, and they also show that the parameter estimates are not strongly dependent on the priors adopted, at least out to our maximum distance of 1 kpc (TGAS volume). We note that the spatial density profiles SFH and MDF used by TRILEGAL are not the same as those in StarHorse, which shows the importance of adding basic and non-restrictive priors in the parameter inference.

### 4.2 Effect of systematic stellar parameter errors

As in Santiago et al. (2016), we also tested the effect of systematic offsets in the observed quantities on our results, using the PARSEC

samples. The results are shown in Fig. 6. Again we split the simulated samples into giants and dwarfs. Each column corresponds to a given parameter for which shifts were applied, keeping the other parameters at their observed (but systematics-free) value. Each panel shows the mean, the median, and the dispersion around the mean values of the relative error (over all stars), as a function of the shift parameter.

In almost all panels, the difference between the mean and median relative errors is very small, attesting to the existence of relatively few outliers. The dispersion around the mean is rarely larger than 20 per cent (or 0.2 mag in the case of extinction) in most cases studied. The exception is the relative age error, for which the mean is much farther from zero than the median, and the dispersion is of order $\simeq 100$ per cent in the case of giants, and even larger for dwarfs (see also the discussion above). On average, however, the effect of systematic errors on our estimated parameters are typically less than ± 10 per cent. In the following, we discuss some more conspicuous effects.

Systematic errors in $T_{\rm eff}$ affect almost all inferred parameters for both dwarfs and giants. The effect of (under)overestimating temperatures on dwarfs is perhaps simpler to interpret, as it leads to best matching models of (lower)higher masses, therefore (less)more luminous. The apparent distance modulus is correspondingly biased too (low)high, yielding either (smaller)larger inferred distances or A_V, or a mixture of both. For the giants, evolutionary time-scales are shorter, making age a central parameter. In their case, an (under)overestimated $T_{\rm eff}$ requires a (older)younger [and therefore (less)more massive and luminous] progenitor, and leads to concordance models of higher apparent distance moduli.
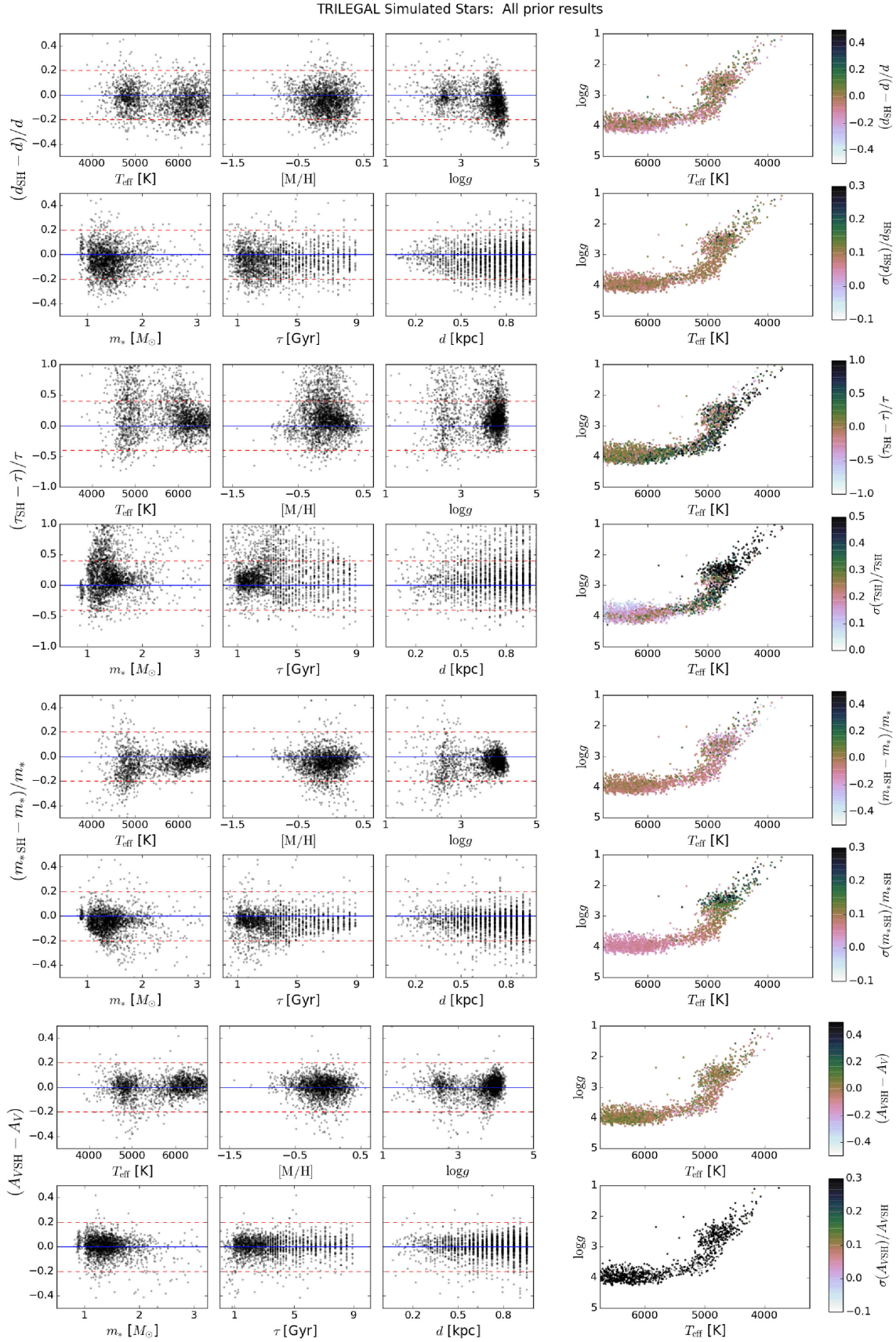
Figure 4. Relative distance, age, mass, and *V*-band extinction errors for the TRILEGAL simulations in the same style as Fig. 3.
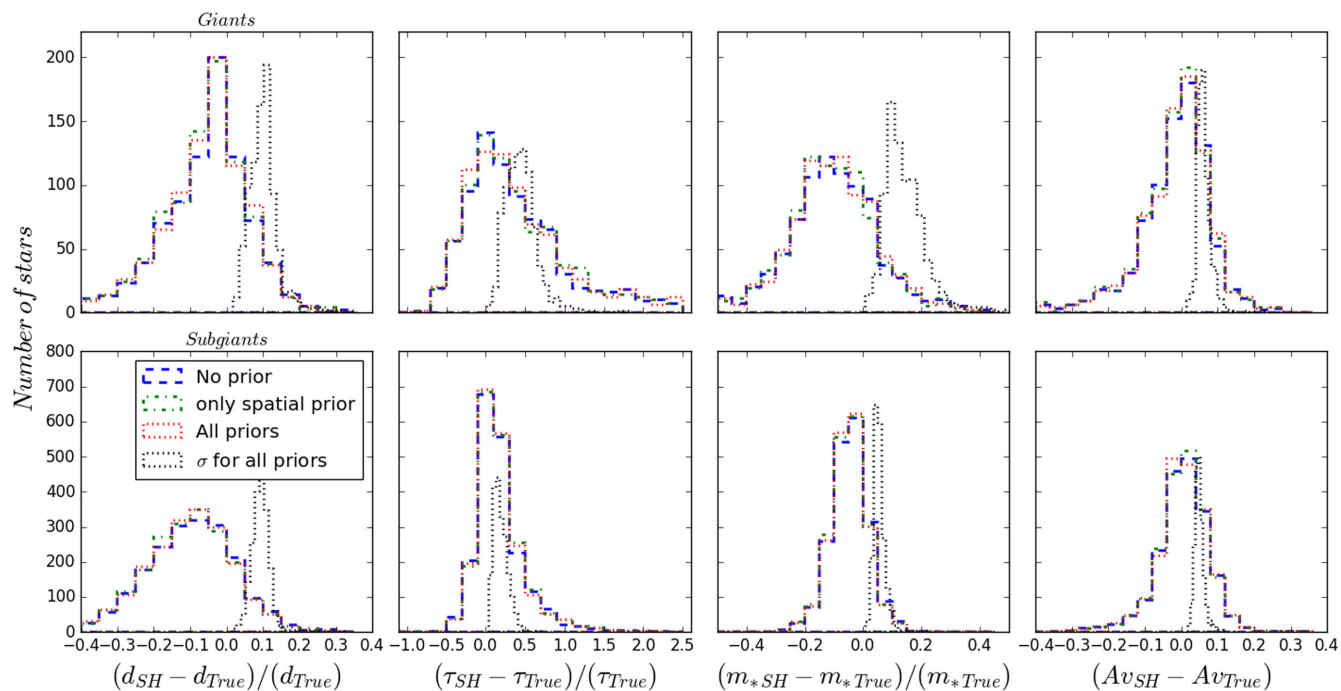
Figure 5. Testing priors with TRILEGAL simulations. The figure shows the distributions of relative errors and uncertainties in distance (leftmost column), age (second column from left), mass (third column from left), and A_V (rightmost column). The top panels show the results for giant stars ($\log g < 3.5$) and the bottom ones are for subgiant stars ($4.1 > \log g > 3.5$). The different lines in each panel correspond to the three different prior assumptions used. The blue line represents the results when no spatial, age, and metallicity priors are adopted, the green line is based on spatial priors only, and the red line shows the results for all priors. The black lines represent the distribution of StarHorse uncertainties normalized by the parameter, only for the all priors case.

The systematic effects of metallicity biases on the inferred parameters are of lower amplitude as compared to $T_{\rm eff}$. This is consistent with $T_{\rm eff}$ having a tighter correlation with the photometric parameters, and hence more strongly affecting the likelihood functions. Estimated masses are affected when the metallicity is biased. An (over)underestimated metallicity leads to a better match of a given star to models of (higher)lower metallicity, which will be of (higher)lower mass for a fixed luminosity (i.e. fixed apparent magnitudes and distance). For giants, the higher(lower) masses will again require younger(older) progenitors.

The case of $\log g$ is such that it affects distances more strongly. This has been investigated before by Santiago et al. (2016), with similar results. An (under)overestimate in $\log g$ leads to an (over)underestimate in the distances, since the data for a star become more consistent with models of stars (more)less luminous than it actually is.

Parallaxes (last column of Fig. 6) also predominantly affect distances, in the expected sense. A slight effect on masses and ages of giants can also be seen, since an (under)overestimated parallax will require (more)less luminous giants, therefore shifting the models towards (younger)older ages with (higher)lower mass progenitors.

5 EXTERNAL VALIDATION

Up to this point, we have shown StarHorse results for simulated stars, for which we previously know all their stellar parameters. Although TRILEGAL delivers realistic star-counts simulations for the Galaxy, real data can present different behaviour from the assumptions we made on the simulations. In this section, we test StarHorse results for observed data. We choose samples of stars from eclipsing binaries (EBs), asteroseismology, and open clusters.

Another important difference with respect to the validation carried out in the previous section is that here we will compare our estimated distances, ages, masses, and extinction values out to distances larger than 1 kpc.

5.1 Eclipsing binaries

In this section, we show StarHorse validations with samples of EBs. EBs can give precise stellar masses and radii, which in turn yield precise surface gravities. If temperature is also available, these stars can provide a good benchmark for estimating distances, ages, masses, and extinctions. Here, we show tests on two samples: one is made up of detached binary systems with individual values of stellar parameters, the other has EBs that are photometrically unresolved.

5.1.1 Detached EBs

We use the sample of EBs from Ghezzi & Johnson (2015) as a first comparison sample to our StarHorse estimates. Those authors carried out a literature search for detached binary systems with at least one evolved star to be used as a benchmark for the determination of masses. This sample contains a total of 26 binaries with photometry, atmospheric parameters, parallaxes, ages, masses, and extinctions. The majority of the sample is composed of stars from the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC), but some stars are near the Solar vicinity ($d < 1$ kpc). As we see in Table 3, this sample contains mainly giant and subgiant stars ($\log g < 3.6$) with low metallicity. To run StarHorse we used as input the photometry $T_{\rm eff}$, [Fe/H], and π given in table 1 of Ghezzi & Johnson (2015) and $\log g$ given in their table 2.
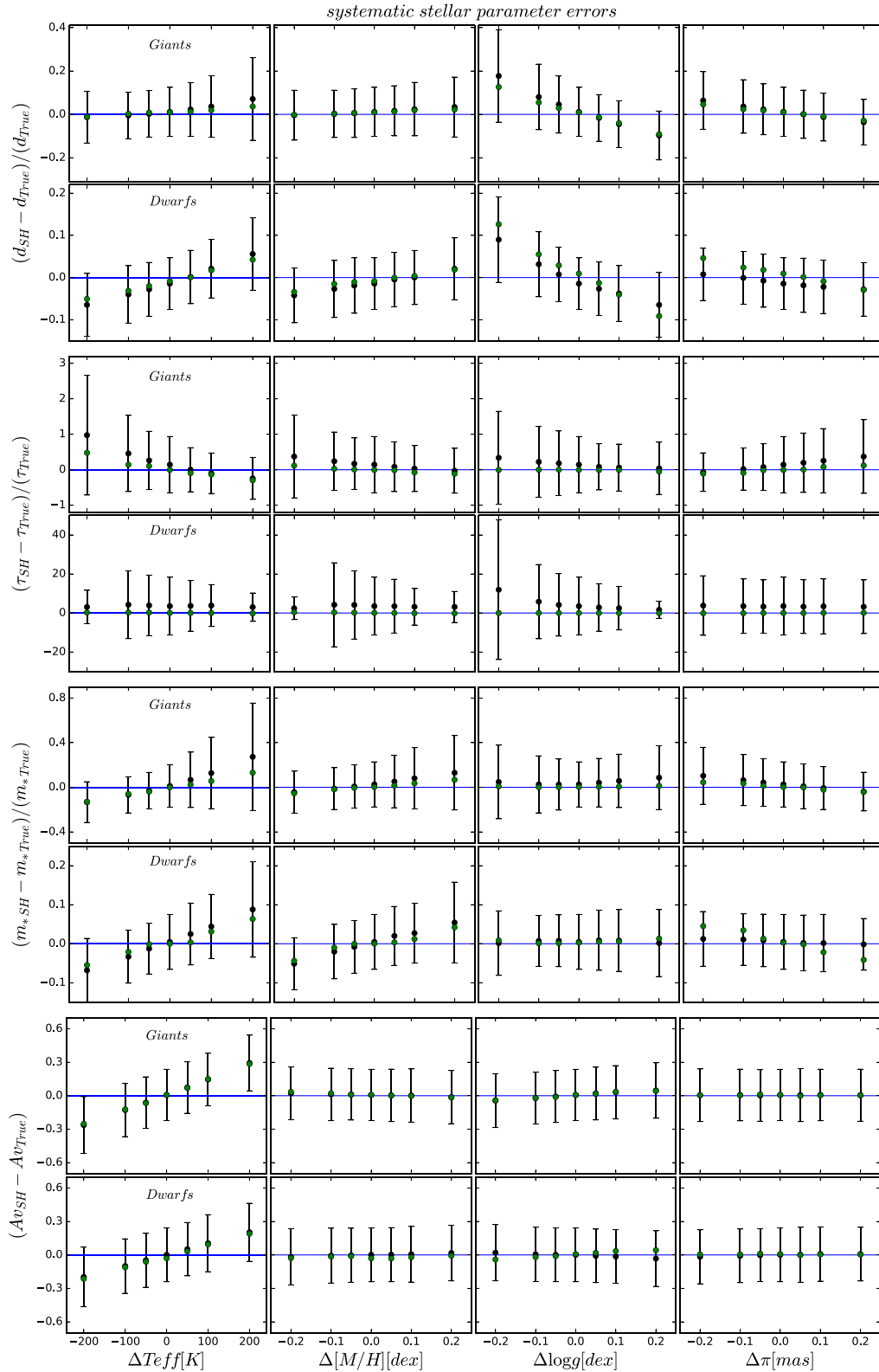
Figure 6. Relative errors per parameter (first two rows: d; third and fourth rows: τ; fifth and sixth rows: $m_{\rm ast}$; seventh and eighth rows: A_V) obtained when different systematic offsets are introduced to the stellar parameter at a time. For this test, we use the simulated PARSEC sample with high-resolution errors, but only 500 stars. The panels show the results for giant stars with $\log g < 4$, and dwarf stars with $\log g > 4$ separately. From the left, the first column shows the results for offsets to $T_{\rm eff}$, the second column shows the offsets to [M/H], the third column shows the offsets to $\log g$, and the last column shows the offsets to parallax π. In each panel, the black (green) dots represent the mean (median) of the distribution of relative errors in each parameter, whereas the standard deviation from the mean is shown as the error bars.
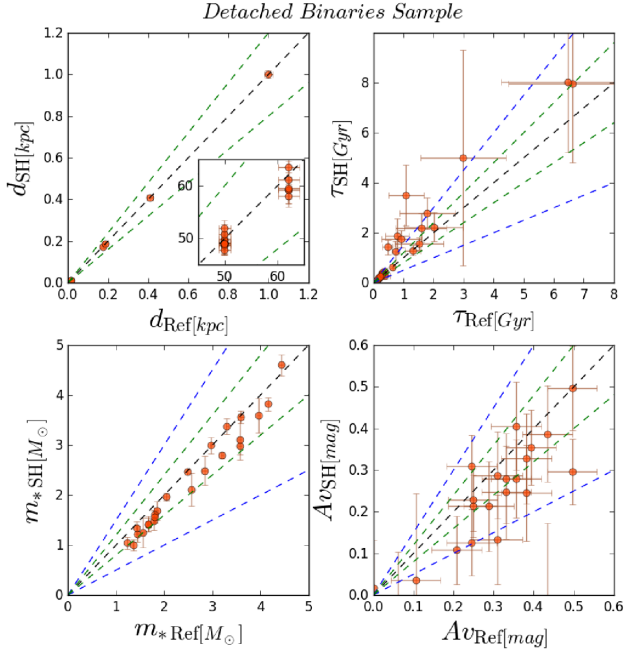
Figure 7. Comparison between our distance (upper left-hand panel), age (upper right-hand panel), mass (lower left), and extinction (lower right) results with those from asteroseismology and the   code for the sample of Ghezzi & Johnson (2015) binaries. The black dashed line is the identity line. The green dashed lines correspond to ±20 per cent deviates, whereas the blue dashed lines (in the last three panels only) correspond to ±50 per cent. The insert on the bottom right shows the same comparison for the LMC/SMC stars.

`StarHorse` recovered distances, ages, masses, and extinctions for 24 out of 26 stars. Fig. 7 compares these results to the more fundamental determinations of Ghezzi & Johnson (2015). For the distances, shown in the upper left-hand panel of Fig. 7, the reference distance is the inverse of the *Hipparcos* input parallaxes used by those authors. Because the sample includes stars either in the Solar Neighbourhood, ($d < 1$ kpc), as well as stars at LMC/SMC distances, we show these latter as a separate inset in the figure. The agreement is excellent, with only a small degradation for the larger distances to LMC and SMC stars, whose parallax uncertainties are larger (see the panel inset). The mean uncertainty on the parallax of this sample is lower than the simulated samples (see Table 3). This is probably the reason for such a good agreement in distances. In the same figure, we show the comparison between `StarHorse` and   da Silva et al. (2006) ages (which incorporate results from asteroseismology using a method similar to `StarHorse`; upper right-hand panel), and the comparison between `StarHorse` masses and those taken from asteroseismology scaling relations (lower left-hand panel; for the references on the masses, see table 2 in Ghezzi & Johnson 2015). We also show the comparison between `StarHorse` A_V and $3.1E(B - V)$ (lower right-hand panel), where the references for $E(B - V)$ are in table 1 of Ghezzi & Johnson (2015).

From Fig. 7, we see that `StarHorse` yields ages that are systematically larger than those from   . The median age offset is of 22 per cent. Still, most of the `StarHorse` ages appear to agree with   ages to within 50 per cent. The two age determinations are also consistent with each other for most stars, considering both error bars. Table 4 shows that 50 per cent (84 per cent) of the estimated ages have errors below 25 per cent (76 per cent). As for the

masses, the agreement between our estimates and those from asteroseismology are somewhat better. `StarHorse` masses tend to be smaller by 12 per cent. The error bars and discrepancies relative to   values are relatively smaller as well, with most of stars having errors of 17 per cent or smaller. The extinction estimates also agree well with the ones from the literature, despite the large error bars. Most of the stars have an A_V error below 0.18 mag, with a moderate systematic effect (-0.07 mag). We note that the systematics between `StarHorse` and the reference sample values cannot be due to systematics in the input atmospheric parameters, since both methods used the same data as input.

5.1.2 Other EBs

As a second comparison with EBs, we use the sample from Stassun & Torres (2016), which contains 156 systems. Their sample is composed of stars with precise stellar radii and effective temperatures. Most of the stars have also available masses, *Gaia* parallaxes, metallicities, and magnitudes in at least one the following filter systems: Tycho (Høg et al. 2000), APASS (Henden & Munari 2014), Str"omgren (Casagrande et al. 2014), and 2MASS (Cutri et al. 2003). Their distances and extinctions were estimated by performing fits to the broad-band photometric spectral energy distributions of the binary systems. The range of the parameters is described in Table 3. Although most of the parameters are individual for each star, the magnitudes are systemic; these magnitudes describe the binary system, not each star individually. This can include a bias in our likelihood, and as most of the sample is made by systems with similar masses, this will affect both primary and secondary stars, though the effect would be greater on the secondary stars. It is very important to proceed with this test, since approximately 50 per cent of Solar-type field stars are binary or multiple systems (Raghavan et al. 2010; Moe & Di Stefano 2017), and most of them are photometrically unresolved.

To run `StarHorse` we only select stars with available masses and limit the distance to 1 kpc. Fig. 8 compares the `StarHorse` results for these stars with the ones of Stassun & Torres (2016). The `StarHorse` distances are, in general, underestimated in relation to the ones estimated by Stassun & Torres (2016). Most of our distance estimates are smaller by 20 per cent (see Table 4), the effect being larger for the secondary members. This is probably a direct result of the systemic magnitudes we are using, since each star is assigned a brighter magnitude than it actually is, the amplitude of the effect being larger for secondaries. Therefore, our code will tend to match it with models at nearer distances. The masses estimated by `StarHorse` are in reasonable agreement with the reference sample, as shown in Table 4: 50 per cent of the stars present relative errors smaller than 10 per cent. We should keep in mind that the surface gravities for this EBs sample are based on the quoted radii and masses. The estimated extinction values do not present systematic deviation with the extinction given by Stassun & Torres (2016), but as the extinction strongly depends on the distance models, they also should be affected by the systemic magnitudes.

5.2 Asteroseismology: CoRoGEE

We also use the CoRoT-APOGEE sample (CoRoGEE; Anders et al. 2017) to evaluate the accuracy of our `StarHorse` results. CoRoGEE contains seismic measurements (from CoRoT) combined with high-resolution spectra (from APOGEE) for more than 600 stars, this sample has also estimates of distance, age, mass, and A_V from
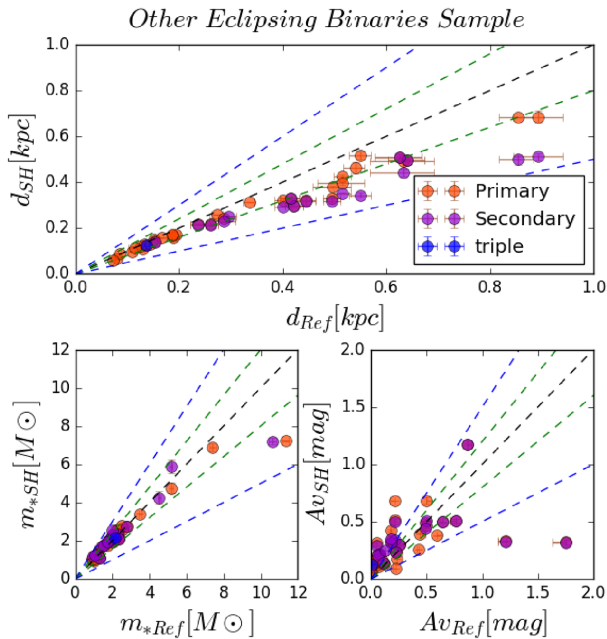
Figure 8. Comparison between our distance (upper panel), mass (lower left-hand panel), and extinction (lower right-hand panel) results with those from the EB sample of Stassun & Torres (2016). The dashed lines correspond to the same deviate as in Fig. 7. The orange dots represent the primary star, the purple dots represent the secondary star, and the blue dots represent the detected triple systems.

the   code. They cover a wide range of Galactocentric distances, metallicities, and ages (see Table 3), but are all red giants stars. To run `StarHorse`, we use as input the atmospheric parameters and total metallicity given by APOGEE, and the parallaxes as the inverse of the distance given by   . We then compare our estimates with the ones from   . The comparison is shown in Fig. 9. The estimated distances are in excellent agreement (upper left-hand panel), since the well-constrained input parallaxes are used by `StarHorse` when building the marginalized distance PDF for each star.

The upper right-hand panel of Fig. 9 shows the comparison between `StarHorse` and   ages. The discreteness of our age grid is visible in the plot. The scatter is large, but most of the stars have ages that agree within ±50 per cent of each other. The median and 84 per cent-ile positions in the distribution of age discrepancies are 12 per cent and 65 per cent, respectively (Table 4). As in the case of the EBs, `StarHorse` tends to yield larger ages than those based on asteroseimology, but this time with a smaller median systematic (16 per cent). In general, the results are quite similar to those from the previous section. This is also true for the extinction estimates, shown in the bottom right-hand panel of Fig. 9. The mass estimates tend to show a better agreement in the case of CoRoGEE, with no systematic trend and smaller errors (median value of 4 per cent) when compared to the EBs.

5.3 The OCCASO clusters

As a third comparison sample, we use the data from the Open Clusters Chemical Abundances from Spanish Observatories (OCCASO) sample (Casamiquela et al. 2016, 2017). This sample contains a total of 128 stars from 18 clusters, covering Galactocentric distances out to 6 kpc, and a small range in metallicity. OCCASO contains
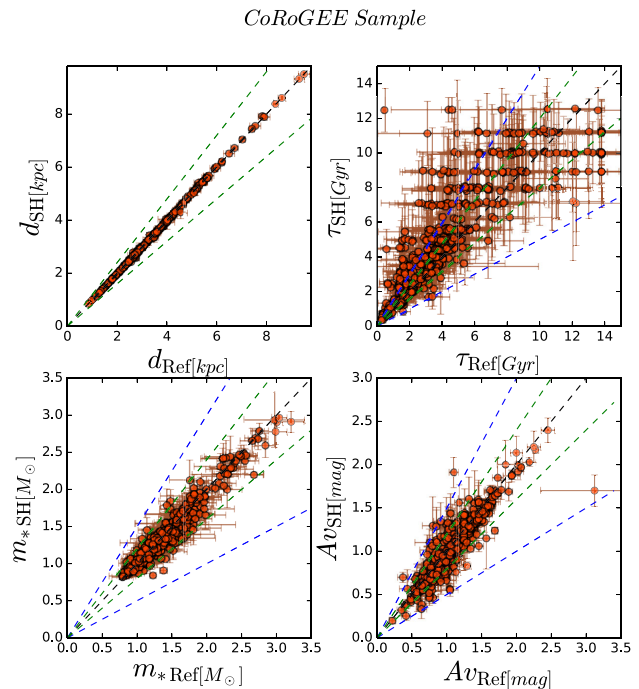
Figure 9. Same panels as in Fig. 7, but now showing the comparison to the CoRoGEE sample.
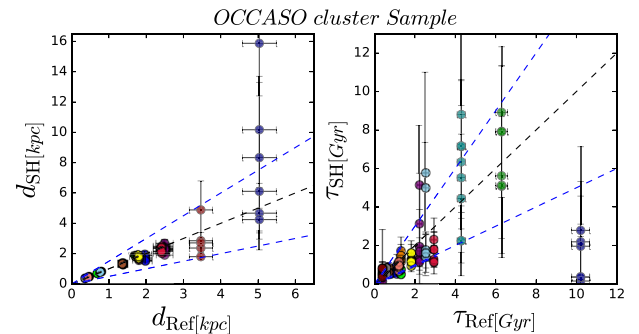
Figure 10. Same as the upper panels in Figs 7 and 9, but now showing the comparison of determined distances and ages to those for the star cluster sample from OCCASO, described in Section 5.3. The stars in the figure are coloured as an identification of their cluster host. In this figure, we show only the identity and ±50 per cent deviating lines.

only red clump stars, for a better spectroscopic resolution in spectroscopy. The age and distance estimates for these clusters are based on isochrone fitting, with a typical uncertainty of 0.2 mag in distance modulus and a mean age uncertainty of 0.2 Gyr. The input parameters used for `StarHorse` were parallaxes, converted from the isochrone distances, metallicities, and atmospherical parameters for each star from high-resolution spectroscopy from OCCASO survey; the mean uncertainty on these parameters can be seen in Table 3.

Fig. 10 shows the comparison to `StarHorse` distances and ages. Each point is a star, for which the cluster's distance and age are attributed. The distances agree with no systematics and median (84 per cent-ile) discrepancies of 2 per cent (13 per cent). Only a few of stars from the most distant cluster resulted with `StarHorse` distances significantly above the one from isochrone fitting. The ages exhibit larger relative discrepancies (31 per cent and 63 per cent,

respectively, for the 50 per cent-ile and 84 per cent-ile positions). We notice from the figure that the stars with a high discrepancy on distance also results in a high discrepancy in age (dark blue points), these points belong to the NGC 6791 cluster – one of oldest open clusters in the Milky Way. Still, for most of the stars, the age estimates are consistent within the relatively large error bars.

6 RELEASED DATA PRODUCTS

We applied StarHorse to a few spectroscopic surveys together with astrometric and photometric measurements to deliver public distances and extinction catalogues. For the moment we are not releasing ages and masses because their estimates are still subject to considerable improvement. Amongst the several spectroscopic surveys available to the community, we chose amongst those that have been more widely used and which have high-to-medium resolution.

All catalogues are available to the community via the LIneA web page.[4] As our method is subject to further improvements (e.g. incorporating new *Gaia* information, new stellar tracks with denser model grids, future releases of the adopted data bases, and new extinction priors), this will lead to new versions of the released data; therefore, we suggest the reader to check for next releases at the LIneA web page.

The following subsections explain the procedure used to determine distances and extinctions for each spectroscopic survey. We applied StarHorse in these samples with all priors described in Section 3, and in all cases, the following corrections were made to each survey catalogue:

(i) *Polish photometry*
In some few cases in which a 2MASS (Cutri et al. 2003) magnitude of a star exists, but the associated uncertainty is −9999.99, we substitute this value by 0.2 mag. Similarly, if the quoted APASS magnitude uncertainty is 0, we set it to be 0.15 mag, and introduce an error floor of 0.02 mag.

(ii) *Correct metallicities for [α/Fe] enhancement*
Since the PARSEC 1.2S stellar models do not yet include models with non-Solar [α/Fe] ratio, we correct for this effect in the data, following Salaris, Chieffi & Straniero (1993), by defining a total metallicity [M/H] as

$$[M/H] = [Fe/H] + \log(0.638 \cdot [\alpha/Fe]^{10} + 0.362)$$

$$\sigma_{[M/H]} = \sqrt{\sigma_{[Fe/H]}^2 + \sigma_{[\alpha/Fe]}^2}.$$

6.1 APOGEE catalogues

The APOGEE-2 survey (Majewski et al. 2017) is a programme from the Sloan Digital Sky Survey (SDSS-IV; Blanton et al. 2017). It is a spectroscopic survey conducted in the near-infrared, with high resolution ($R \sim 20\,500$) and high signal-to-noise ratio (S/N > 100). It maps the Galaxy through all populations and it has targeted, as of its latest release (DR14; Abolfathi et al. 2017), about 270 000 stellar spectra. As a near-infrared survey of high resolution, APOGEE has the advantage to study with more detail the stars located in the dusty regions of our Galaxy, such as the bulge and disc. We applied StarHorse to the latest APOGEE release DR14.

6.1.1 APOGEE DR14 ASPCAP

We used the APOGEE DR14 allStar summary catalogue (a Star ts[5]). The file contains photometry in the 2MASS *JHK_s* passbands (Cutri et al. 2003), as well as chemical abundances and atmospheric parameters derived by the APOGEE Stellar Parameters and Chemical Abundances (ASPCAP; García Pérez et al. 2016). Starting from this file, we took the following pre-processing steps before running StarHorse:

(i) *Cross-match with photometry and astrometry*
Using TOPCAT (Taylor 2005), we did a positional cross-match so as to add information from APASS DR9 (Henden & Munari 2014) (207 604 matches) and *Gaia* DR1/TGAS (Gaia Collaboration et al. 2016) (46 033 matches). Our cross-match used a maximum separation of 5 arcsec.

(ii) *Selecting reliable results*
We discarded sources with ASPCAP flags containing the words 'BAD' or 'NO_ASPCAP'. In addition, we selected only sources with valid 2MASS *H* magnitudes. This resulted in a total of 226 323 stars.

(iii) *Using uncalibrated ASPCAP results*
When no calibrated ASPCAP results for one or more spectroscopic parameters are available (stars outside the calibration ranges; e.g. log *g* values for dwarfs), we use uncalibrated ASPCAP results. For those objects we inflated the uncertainties. Our ad hoc conservative uncertainty estimates for these cases amount to 150 K in T_{eff}, 0.3 dex in log *g*, 0.15 dex in [M/H], and 0.1 dex in [α/M].

(iv) A_V *prior*
As priors for the *V*-band extinction, we used the APOGEE targeting extinction values (Zasowski et al. 2013), derived by the RJCE method (Majewski, Zasowski & Nidever 2011), by setting $A_{V, prior} = A(K_s)_{Targ}/0.12$. When RJCE estimates were not available, we used $E(B − V)$ estimates from (Schlegel et al. 1998) to estimate the prior A_V. As explained in Section 3.2.2, the posterior A_V values are allowed to lie in a very broad interval around the prior values.

We then applied StarHorse to the resulting catalogue. Our results are available through an SDSS-IV value-added catalogue (VAC) of APOGEE stellar distances, as part of SDSS DR14[6] and on the LIneA web page. The output format of this catalogue is described in Table 5. Fig. 11 summarizes the results for APOGEE DR14. Fig. 12 shows the associated uncertainty distributions.

6.1.2 APOGEE DR14 Cannon

As the APOGEE DR14 also contains stellar parameters and chemical abundances derived by the data-driven method called Cannon (Ness et al. 2015; Casey et al. 2016), we also applied StarHorse with this input. We used the APOGEE DR14[7] allStar summary catalogue (a Star a o ts). Which contains the atmospheric stellar parameters, chemical abundances, and its uncertainties. Starting from this file we did the following steps to prepare a input to StarHorse:

(i) *Cross-match with photometry and astrometry*

---

[4] http://www.linea.gov.br/020-data-center/acesso-a-dados-3/spectrophotometric-distances-starhorse-code/

[5] Available at https://data.sdss.org/sas/dr14/apogee/spectro/redux/r8/allStar-l31c.2.fits.

[6] http://www.sdss.org/dr14/irspec/spectro_data/

[7] http://www.sdss.org/dr14/irspec/spectro_data/

Table 5. General description of the release distances and extinctions.

Column	Description	units
OBJECT ID	Survey's object ID name	
glon	Galactic longitude	degrees
glat	Galactic latitude	degrees
dist05	5th percentile of the stars's distance PDF	kpc
dist16	16th percentile of the stars's distance PDF	kpc
dist50	50th percentile of the stars's distance PDF	kpc
dist84	84th percentile of the stars's distance PDF	kpc
dist95	95th percentile of the stars's distance PDF	kpc
meandist	Mean of the stars's distance PDF	kpc
diststd	standard deviation of the stars's distance PDF	kpc
AV05	5th percentile of the stars's extinction PDF	mag
AV16	16th percentile of the stars's extinction PDF	mag
AV50	50th percentile of the stars's extinction PDF	mag
AV84	84th percentile of the stars's extinction PDF	mag
AV95	95th percentile of the stars's extinction PDF	mag
meanAV	Mean of the star's extinction PDF	mag
stdAV	Standard deviation of the stars's extinction PDF	mag
SH_INPUTFLAGS	`StarHorse` flags regarding the input data	
SH_OUTPUTFLAGS	`StarHorse` flags regarding the output data	

We carried out a positional cross-match, using TOPCAT (Taylor 2005), with photometry from APASS (Henden & Munari 2014) and from 2MASS (Cutri et al. 2003).

(ii) *Av prior*

As priors for the *V*-band extinction we used the APOGEE targeting extinction values (Zasowski et al. 2013), derived by the RJCE method (Majewski et al. 2011), by setting $A_{V,\,\mathrm{prior}} = A(K_s)_\mathrm{Targ}/0.12$.

The catalog for this sample is available at LIneA web page, the output format of this catalogue is described in Table 5, Fig. 12 shows the associated uncertainty distributions, and in Appendix B we summarize the results for APOGEE Cannon on Fig. B1.

6.2 The *Gaia*-ESO sample

The GES (Gilmore et al. 2012) is a large public spectroscopic survey with high resolution that covers all Milky Way components and open star clusters of all ages and masses. The final GES release is expected to include about 10^5 stars.

We downloaded the *Gaia*-ESO data release 3 (DR3) from the ESO catalogue facility.[8] This catalogue contains a total of 25 533 stars, including the Milky Way field, open clusters, and calibration stars. The catalogue contains 100 columns, from which we selected the necessary information to run `StarHorse`: atmospheric parameters, elemental abundances, and 2MASS *JHK_s* passbands (Cutri et al. 2003). From this catalogue, we proceeded with the following steps, before applying `StarHorse`:

(i) *Select only field stars and reliable sources*

From this catalogue, we selected only Milky Way field stars, which are the main targets intended for `StarHorse`. This leaves us with 7870 stars. We then adopted the following quality criteria: relative errors in T_eff less than 5 per cent; errors in $\log g$ lower than 0.4 dex; and errors in metallicity lower than 0.2 dex, with no cuts in the abundances. From the 7870 field stars, 6316 of them meet these criteria:

(ii) *Calculate an overall [α/Fe] abundance*

The PARSEC models do not list individual elemental abundances, only the total metallicity value, [M/H]. The models do not include non-Solar abundances in α elements either. Therefore, we converted the GES abundances in the total metallicity [M/H] before running `StarHorse`. For that purpose, we calculated the overall [α/Fe] abundance as follows:

$$[\alpha/\mathrm{Fe}] = \frac{1}{n}\sum_{i=1}^{n}[\mathrm{X_i/H}] - [\mathrm{Fe/H}],$$

where X_i refers to the elements O, S, Ti, Ca, and Mg, and the [X_i/H] abundances were calculated using the solar values from Asplund et al. (2009). The error in [α/Fe] was propagated in quadrature from the error on each elemental abundance and the error in [Fe/H].

(iii) *Cross-match with photometry*

To obtain more precise and reliable extinction estimates we decided to include APASS (Henden & Munari 2014) magnitudes for this sample. For this, we carried out a positional cross-match with APASS DR9 (Henden & Munari 2014) using TOPCAT (Taylor 2005), with a maximum separation of 1 arcsec. Of the 6316 stars, 5719 stars have APASS magnitudes.

(iv) A_V *prior*

No A_V prior was applied to this sample; therefore, the posterior probability function for A_V always ranges from 0 to 3 (mag).

After carrying out these steps, we used this final catalogue with 6316 stars as input to `StarHorse`. The code delivered an output catalogue with 6011 stars with available distances and extinctions. The columns of the released distances and extinction catalogue are shown in Table 5, the uncertainty distribution are shown in Fig. 12, and in the Appendix B Fig. B2 we summarize the `StarHorse` results for GES.

6.3 GALAH sample

The GALAH (Martell et al. 2017) is a spectroscopic survey that will target about 1 million stars with the high-resolution ($R \sim 28\,000$) instrument High Efficiency and Resolution Multi-Element Spectrograph (HERMES; Sheinis et al. 2015) at the Anglo-Australian Telescope (AAT). The main goal of the project is to provide a detailed star-formation history for the thick and thin discs. Therefore,
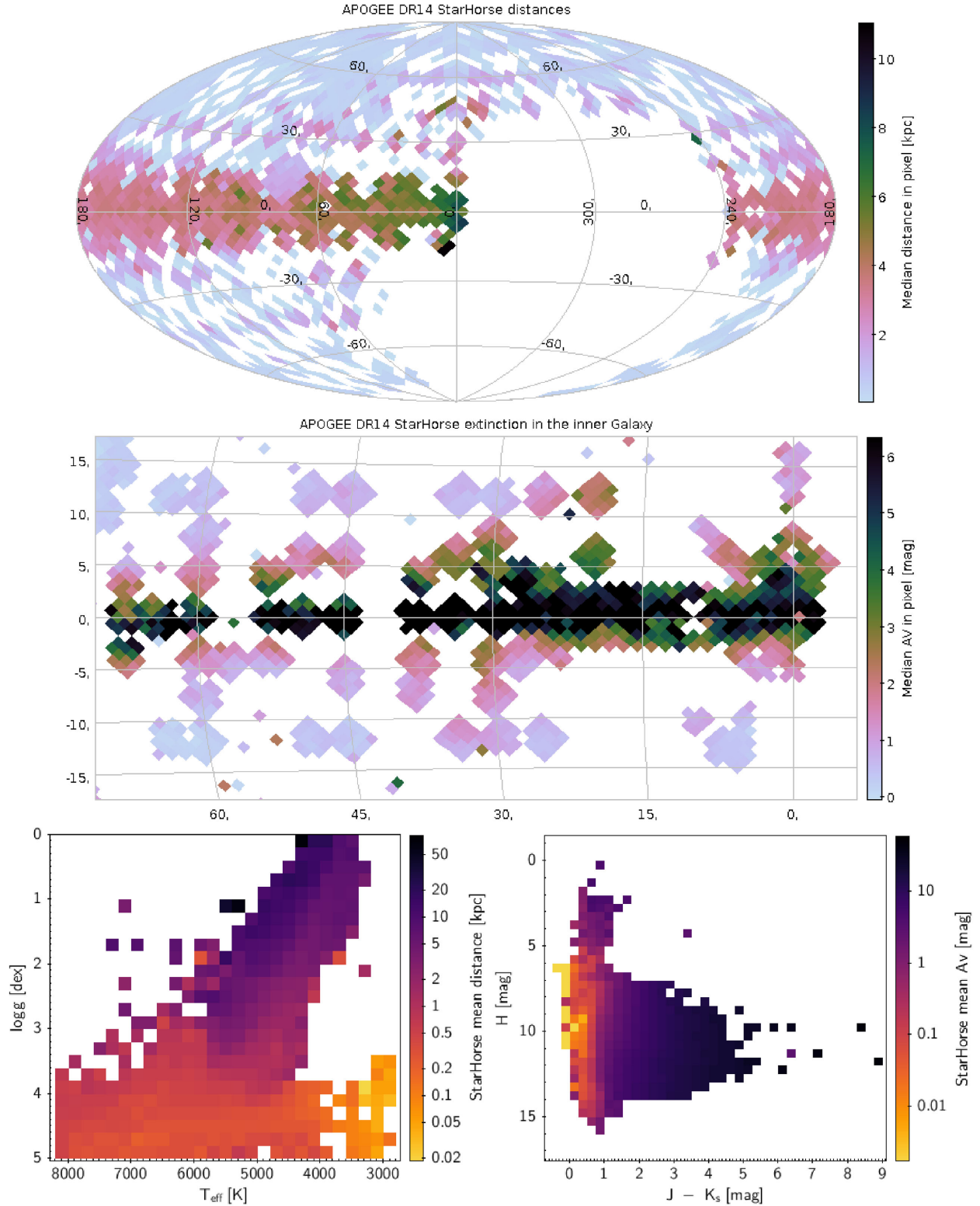
Figure 11. Illustration of the APOGEE DR14 distance and extinction results from `StarHorse`. Top panel: Aitoff projection of median APOGEE distances per HealPix cell in Galactic coordinates. Middle panel: Resulting median A_V per HealPix cell in the inner Galaxy. Bottom left-hand panel: Spectroscopic Hertzsprung–Russell diagram, colour-coded by median distance in each pixel. Bottom right-hand panel: 2MASS colour–magnitude diagram, colour-coded by median extinction in each pixel.

Data Products Uncertainty Distributions
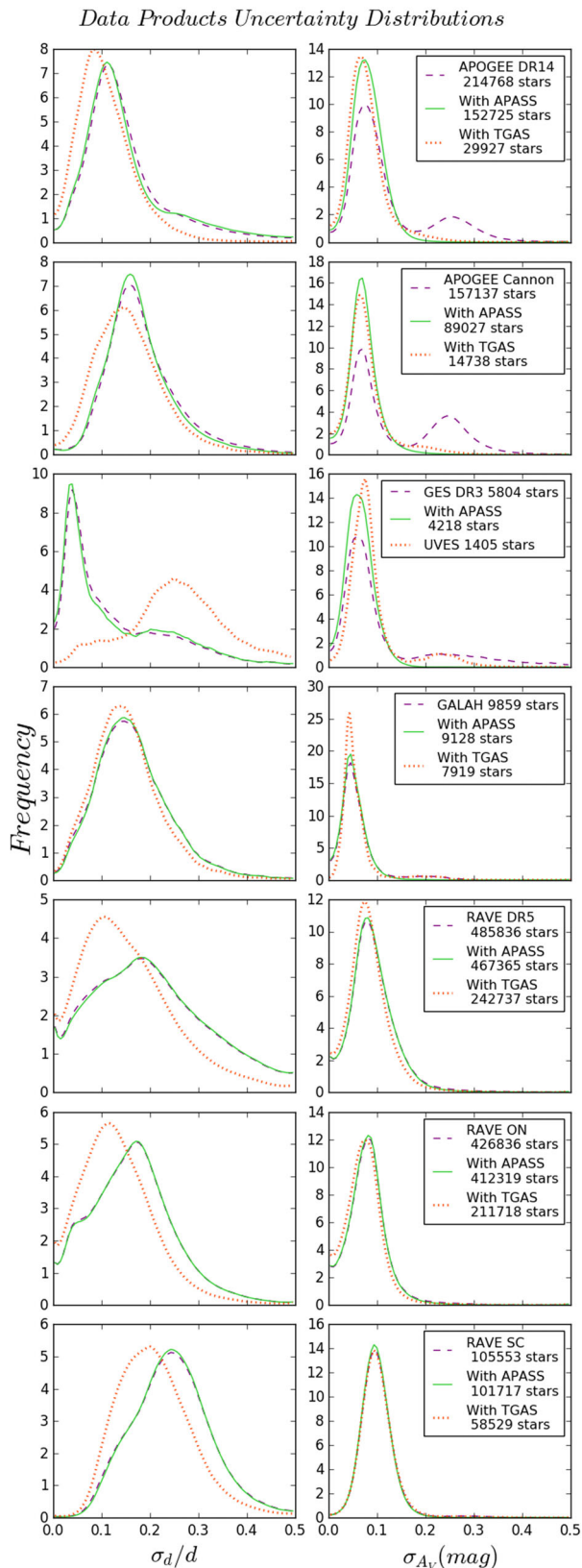
Figure 12. `StarHorse` relative distance (left-hand panels) and absolute extinction (right-hand panels) uncertainty distributions for all release data products. The legend specifies the stars with available APASS photometry, parallaxes from TGAS or UVES spectroscopy (GES sample).

the survey covers mainly the disc, but it also has some fields towards the bulge and halo. A first public data release of GALAH is already available. We then applied `StarHorse` to estimate distances and extinction for GALAH DR1 stars. The catalogue[9] contains 10 680 stars with [Fe/H], [α/Fe], log$g$, and $T_{\rm eff}$ measurements derived by the Cannon method (Ness et al. 2015). We proceeded with the following steps to have an input catalogue ready for `StarHorse`:

(i) *Stellar parameter uncertainties*

The GALAH DR1 catalogue does not provide individual uncertainties for the stars. We therefore used the values recommended by Martell et al. (2017): σ([Fe/H]) = 0.056 dex, σ(log g) = 0.17 dex, and $\sigma(T_{\rm eff})$ = 51 K. As there is no mention of the uncertainty in [α/Fe], so we assumed the error to be same as for σ[Fe/H].

(ii) *Cross-match with photometry and astrometry*

Since the GALAH DR1 catalogue has stars in common with the *Gaia* DR1, we carried out a positional cross-match with *Gaia* using TOPCAT (Taylor 2005), with a maximum separation of 1 arcsec. From the 10 680 DR1 stars, 7919 stars have parallax available. We obtained photometry for the sample by cross-matching with the 2MASS (Cutri et al. 2003) (10 680 matches) and APASS (Henden & Munari 2014) (9263 matches) catalogues.

(iii) *Av prior*

We use the reddening given by the GALAH DR1 catalogue, which is derived by comparison of absolute magnitudes with the apparent magnitude *V* from APASS (Henden & Munari 2014) and *J* magnitude from 2MASS(Cutri et al. 2003). We assumed then that $A_V = 3.1E(B - V)$.

After these steps were completed, we used this final file as input to `StarHorse` and derived distances and extinctions. The code returned 10 623 distances and extinctions. The columns for the released catalogue are described in Table 5, Fig. 12 shows the associated uncertainty distributions and Fig. B3 in Appendix B shows the main results for this sample.

6.4 The RAVE catalogues

The RAVE (Steinmetz et al. 2006) is one of the largest spectroscopic surveys of the Milky Way. RAVE has already delivered spectra for almost 500 K stars that were randomly targeted in an area of 20 K square degrees of the Galactic Southern hemisphere. In addition, RAVE is currently the survey that contains the largest number of stars in common with TGAS ($\sim$200 K stars). The survey works with a multiobject spectrograph deployed on 1.2-m UK Schmidt Telescope of the Australian Astronomical Observatory (AAO). The spectra have a medium resolution of ($R \sim 7500$) and cover the Ca-triplet region (8410–8795 Å). Given that RAVE has a medium resolution is pioneering amongst the large spectroscopic surveys, covers a large area, and has the largest overlap with the *Gaia* sample, we decided to apply `StarHorse` to the entire survey, and to make the estimated distances and extinctions available to the community. The following subsections explain how we proceed with the RAVE input catalogues to execute `StarHorse` and the description of the released distances-extinction catalogues. All RAVE catalogues were downloaded from the RAVE website.[10]

---

[9] https://cloudstor.aarnet.edu.au/plus/index.php/s/OMc9QWGG1koAK2D
[10] https://www.rave-survey.org/project/

6.4.1 The RAVE DR5 catalogue

The data release 5 (DR5; Kunder et al. 2017) is the latest RAVE data release. It contains spectra for 483 330 stars. We downloaded the publicly released catalogue called RAVE_DR5, which contains spectral parameters and radial velocities derived by the SPARV pipeline (Zwitter et al. 2008; Siebert et al. 2011). The catalogue also contains astrometry from *Gaia*-DR1 (215 590), and photometry from 2MASS and APASS. We note that very recently, McMillan et al. (2017) updated the DR5 catalogue parameters and derived distances using feedback from the *Gaia* DR1 parallaxes, but our VAC presented here is based on the public DR5 data. From this catalogue, we proceeded with the following steps before running `StarHorse`:

(i) *Spectral parameters*

We use the calibrated atmospheric parameters, which are named in the catalogue as $T_{\mathrm{eff}NK}$, log g_{NK}, and Met$_{NK}$. For the uncertainties, if the error spectral analysis is available, we use the maximum between the two values: $\sigma T_{\mathrm{eff}K}$ and StdDev$T_{\mathrm{eff}K}$; otherwise, we use the maximum between 70 K and $\sigma T_{\mathrm{eff}}f_K$. We worked analogously with the other parameters.

(ii) A_V *prior*

As explained in Section 3.2.2, we can use a prior value of extinction to build the A_V posterior probability function. We use as extinction prior in *V* band the maps of $E(B - v)$ from Schlegel et al. (1998) for this catalogue.

6.4.2 RAVE-SC catalogue

The RAVE-SC catalogue has stars from DR5 with gravity from seismic calibrations (Valentini et al. 2017); therefore, this sample is only composed by giants. We downloaded the catalogue named as RAVE_Gravity_SC. Step 1 from the previous Section 6.4.1 was also applied to this sample. We use the overall [α/Fe] abundance and the [Fe/H], given by the catalogue to calculate a total metallicity as defined by Salaris et al. (1993), with a fixed uncertainty of 0.2 dex. The atmospheric parameters were used as they were given by the catalogue; we use the following temperature and surface gravity columns: Teff_IR and logg_SC. For the A_V prior we use the Schlegel et al. (1998) $E(B - V)$ maps.

6.4.3 RAVE-on catalogue

The RAVE-on catalogue (Casey et al. 2017) has stars from DR5 with parameters derived by the Cannon method (Casey et al. 2016). We downloaded the catalogue named as RAVE-ON. The atmospheric parameters and [Fe/H] were used directly from this catalogue. The following steps were applied before applying `StarHorse`:

(i) *Calculate an overall [α/Fe] abundance*

The Cannon provides the individual abundances for the stars. We then calculated [α/Fe] as the simple average between the individual abundances when they are available, exactly as described in Section 6.2 with X_i as O, Mg, Ca, and Si.

(ii) A_V *prior*

No $A_V$prior was applied to this sample; therefore, the posterior probability function for A_V always ranges from 0 to 3 mag.

All RAVE catalogs are available in the LIneA web page, the output format follows the columns described in Table 5, in Fig. 12 we show the uncertainty distributions and in Figs B3–B6 in the Appendix B we summarize the results for the RAVE catalogs.

6.5 `StarHorse` FLAGS

All released data product catalogues have two columns that describe the `StarHorse` input data, SH_INPUTFLAGS, and the `StarHorse` output data, SH_OUTPUTFLAGS, as shown in Table 5. The input flags specify which parameters were used in the likelihood calculation to estimate the distances and extinctions given. For example, if the temperature was available for that star a flag will appear as "TEFF", and if temperature was not available in the calculation a flag will appear as "uncalTEFF". The other parameters are specified as follows in the input flag: LOGG (surface gravity), PARALLAX (parallax), MH (metallicity), *JHK*s (2Mass filters), and BVgri (APASS filters). If the input flag contains ALPHAM, it means that the alpha elements were available in the calculation of the total metallicity of the star. The input flags also indicate you if we use a A_V prior as the AVprior flag. The output flags tell us if the number of models that have converged in the likelihood calculation is too small. If less than 10 models are consistent with the star a NUMMODELS_BAD flag will appear, while if the number of models is between 10 and 30, a NUMMODELS_WARM flag will appear. The output flags also indicate if the estimated extinction is negative (NEGATIVE_EXTINCTION_WARN), if it is too high (HIGH_EXTINCTION_WARN), or if the estimated extinction has a bright 2MASS source (EXTINCTION_BAD_BRIGHT2MASS).

7 SUMMARY AND FUTURE PERSPECTIVES

We have presented a code that computes distances, ages, masses, and extinctions for field stars with photometric, spectroscopic, and astrometric data. It is based on Bayesian inference, computing the marginal posterior distributions for the data given a set of stellar models. The code represents a significant improvement over the one presented by Santiago et al. (2016) in several aspects. The most important one is the ability to estimate ages, masses, and extinction, in addition to the spectrophotometric distances presented by those authors. The updated code, which we call `StarHorse`, is also capable of incorporating the parallax as an additional observational quantity in the statistical analysis (Fig. 1). Updated spatial, metallicity, and age priors for the Galactic components, now including the bulge as well, are presented (Table 1). In addition, `StarHorse` is now more flexible in terms of the input data and the choice of observational quantities to be used within them.

The new code was validated using simulated and real stars. These latter are samples with reliable parallax (or distance) data, including field giants with asteroseismic data, EBs used as benchmarks for stellar evolutionary codes, or cluster stars with well-known distances and ages, usually from isochrone fitting, often in combination with spectroscopic data. For EBs that are not detached, the distances present an offset in relation to the reference ones, our distances being usually smaller by 20 per cent in this specific case. The discrepancy is larger for secondary stars than for the primary ones, which is what one expects from using systemic photometric measurements. In all cases, age is the single most difficult parameter to infer, yielding median errors that range from 12 per cent to 35 per cent for quality spectrophotometric data, depending on the sample (Table 4). Errors larger than 100 per cent in age may result for $\simeq$15 per cent of the stellar models, most of them younger than $\tau \simeq 1$ Gyr. In a realistic flux-limited sample, as simulated by the TRILEGAL code or for real stars, the fraction of such catastrophic age errors is reduced to $\simeq$5 per cent of the stars. Our results for stellar ages, either based on simulated or real stars, also indicate a systematic trend of `StarHorse` overestimating ages by 10–20 per cent.

As for spectrophotometric masses, we obtain consistent results over all validation samples used, in the sense that errors < 20 per cent are observed for most (84 per cent) of the stars in any sample. The median error varies depending on the quality of the parallax used as a constraint. For typical *Gaia*-TGAS errors of 0.3 mas, the typical distance errors are around 15 per cent. For real stars used as reference, median `StarHorse` A_V errors are of 0.07 mag, with the 84 per cent *-ile* error at 0.15–0.20 mag. For TRILEGAL and PARSEC synthetic stars, the relative errors are a bit smaller, around 0.04 mag.

We note that the error estimates based on comparison to real samples may be overestimated, considering that the some of the discrepancy may be attributed to the methods used to obtain the reference quantities for comparison. In fact, Rodrigues et al. (2017) report that ages and masses from asteroseismology are typically obtained with a precision of 19 per cent and 5 per cent, respectively, which are comparable to the errors we quote in this analysis.

`StarHorse` has already been used to infer distances and extinction values for stars from APOGEE DR14. These parameters, in turn, may be used in connection to APOGEE abundances and radial velocities, to study the properties of the main Galactic populations, and their spatial variations, as was previously done by Anders et al. (2014), and Fernández-Alvar et al. (2016) using distances from Santiago et al. (2016). For more local samples, such as *Gaia*-TGAS and RAVE, reliable parallax information can be included in the Bayesian method to yield masses and ages, as validated in this paper, allowing for a more detailed modelling of the chemodynamical history of our Galaxy (Anders et al. 2017).

Finally, we have run `StarHorse` on different public catalogues from the RAVE collaboration, as well as on GES, and GALAH public data releases. These are available for download at the LIneA web site.[11]

ACKNOWLEDGEMENTS

REFERENCES

Abolfathi B. et al., 2017, ApJS, preprint (arXiv:1707.09322)
Anders F. et al., 2014, A&A, 564, A115
Anders F. et al., 2017, A&A, 597, A30
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Astropy Collaboration et al., 2013, A&A, 558, A33
Bayo A., Rodrigo C., Barrado Y Navascués D., Solano E., Gutiérrez R., Morales-Calderón M., Allard F., 2008, A&A, 492, 277
Bensby T. et al., 2013, A&A, 549, A147
Binney J. et al., 2014, MNRAS, 437, 351
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Blanton M. R. et al., 2017, AJ, 154, 28
Boeche C. et al., 2013, A&A, 559, A59
Boeche C. et al., 2014, A&A, 568, A71
Bovy J., Rix H.-W., Schlafly E. F., Nidever D. L., Holtzman J. A., Shetrone M., Beers T. C., 2016, ApJ, 823, 30
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127

Brogaard K. et al., 2016, Astron. Nachr., 337, 793
Burnett B., Binney J., 2010, MNRAS, 407, 339
Burnett B. et al., 2011, A&A, 532, A113
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Carlin J. L. et al., 2015, AJ, 150, 4
Casagrande L. et al., 2014, MNRAS, 439, 2060
Casamiquela L. et al., 2016, MNRAS, 458, 3150
Casamiquela L. et al., 2017, MNRAS, 470, 4363
Casey A. R., Hogg D. W., Ness M., Rix H.-W., Ho A. Q., Gilmore G., 2016, ApJ, preprint (arXiv:1603.03040)
Casey A. R. et al., 2017, ApJ, 840, 59
Chabrier G., 2003, PASP, 115, 763
Chiappini C. et al., 2015, A&A, 576, L12
Cutri R. M. et al., 2003, VizieR On-line Data Catalog: II/246
da Silva L. et al., 2006, A&A, 458, 609
de Jong R. S., Radburn-Smith D. J., Sick J. N., 2009, in Andersen J., Nordström B., Bland-Hawthorn J., eds, Proc. IAU Symp. 254, The Galaxy Disk in Cosmological Context. Cambridge Univ. Press, Cambridge, p. 19
Dehnen W., Binney J., 1998, MNRAS, 294, 429
Deng L.-C. et al., 2012, Res. Astron. Astrophys., 12, 735
Fernández-Alvar E., Allende Prieto C., Beers T. C., Lee Y. S., Masseron T., Schneider D. P., 2016, A&A, 593, A28
Fuhrmann K., 1998, A&A, 338, 161
Fuhrmann K., Chini R., 2017, ApJ, 834, 114
Gaia Collaboration et al., 2016, A&A, 595, A2
García Pérez A. E. et al., 2016, AJ, 151, 144
Ghezzi L., Johnson J. A., 2015, ApJ, 812, 96
Gilmore G., 2012, in Aoki W., Ishigaki M., Suda T., Tsujimoto T., Arimoto N., eds, ASP Conf. Ser. Vol. 458, Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way. Astron. Soc. Pac., San Francisco, p. 147
Gilmore G. et al., 2012, Messenger, 147, 25
Girardi L., Groenewegen M. A. T., Hatziminaoglou E., da Costa L., 2005, A&A, 436, 895
Girardi L. et al., 2012, in Miglio A., Montalbán J., Noels A., eds, TRI-LEGAL, a TRIdimensional modeL of thE GALaxy: Status and Future. Springer-Verlag, Berlin, p. 165
Gouda N., 2012, in Aoki W., Ishigaki M., Suda T., Tsujimoto T., Arimoto N., eds, ASP Conf. Ser. Vol. 458, Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way. Astron. Soc. Pac., San Francisco, Springer-Verlag, Berlin, p. 417
Hayden M. R. et al., 2015, ApJ, 808, 132
Haywood M., Di Matteo P., Lehnert M. D., Katz D., Gómez A., 2013, A&A, 560, A109
Henden A., Munari U., 2014, Contrib. Astron. Obs. Skalnate Pleso, 43, 518
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Høg E. et al., 2000, A&A, 357, 367
Imai H., Burns R. A., Yamada Y., Goda N., Yano T., Orosz G., Niinuma K., Bekki K., 2016, preprint (arXiv:1603.02042)
Kunder A. et al., 2017, AJ, 153, 75
Lagarde N., Robin A. C., Reylé C., Nasello G., 2017, A&A, 601, A27
Mackereth J. T. et al., 2017, MNRAS, 471, 3057
Maíz Apellániz J., 2006, AJ, 131, 1184
Majewski S. R., Zasowski G., Nidever D. L., 2011, ApJ, 739, 25
Majewski S. R. et al., 2017, AJ, 154, 94
Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883
Martell S. L. et al., 2017, MNRAS, 465, 3203

McMillan P. J. et al., 2017, MNRAS, preprint (arXiv:1707.04554)
Melis C., Reid M. J., Mioduszewski A. J., Stauffer J. R., Bower G. C., 2014, Science, 345, 1029
Mikolaitis Š. et al., 2014, A&A, 572, A33
Mitschang A. W., De Silva G., Zucker D. B., Anguiano B., Bensby T., Feltzing S., 2014, MNRAS, 438, 2753
Moe M., Di Stefano R., 2017, ApJS, 230, 15
Ness M., Hogg D. W., Rix H.-W., Ho A. Y. Q., Zasowski G., 2015, ApJ, 808, 16
Nidever D. L. et al., 2014, ApJ, 796, 38
Oliphant T. E., 2007, Comput. Sci. Eng., 9, 90
Raghavan D. et al., 2010, ApJS, 190, 1
Recio-Blanco A. et al., 2014, A&A, 567, A5
Robin A. C., Marshall D. J., Schultheis M., Reylé C., 2012, A&A, 538, A106
Robin A. C., Reylé C., Fliri J., Czekaj M., Robert C. P., Martins A. M. M., 2014, A&A, 569, A13
Rodrigues T. S. et al., 2014, MNRAS, 445, 2758
Rodrigues T. S. et al., 2017, MNRAS, 467, 1433
Ryan S. G., Norris J. E., Beers T. C., 1996, ApJ, 471, 254
Salaris M., Chieffi A., Straniero O., 1993, ApJ, 414, 580
Santiago B. X. et al., 2016, A&A, 585, A42
Schlafly E. F. et al., 2016, ApJ, 821, 78
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Sheinis A. et al., 2015, J. Astron. Telesc. Instrum. Syst., 1, 035002
Siebert A. et al., 2011, AJ, 141, 187
Stassun K. G., Torres G., 2016, AJ, 152, 180
Steinmetz M. et al., 2006, AJ, 132, 1645
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
Valentini M. et al., 2017, A&A, 600, A66
Valle G., Dell'Omodarme M., Prada Moroni P. G., Degl'Innocenti S., 2015, A&A, 577, A72
Vanhollebeke E., Groenewegen M. A. T., Girardi L., 2009, A&A, 498, 95
Yanny B. et al., 2009, AJ, 137, 4377
Yong D. et al., 2016, MNRAS, 459, 487
Zasowski G. et al., 2013, AJ, 146, 81
Zoccali M. et al., 2003, A&A, 399, 931
Zwitter T. et al., 2008, AJ, 136, 421

APPENDIX A: ADDITIONAL SIMULATION RESULTS

Here, we show the results of PARSEC simulations when parallaxes are not used to constrain distances or the likelihood. Fig. A1 can then be compared to the case shown in the main body of the paper (Fig. 3). All estimated parameters are subject to larger errors, especially A_V. The systematic distance and mass error dependences on true stellar age and mass become very pronounced when parallaxes are not used.

We also show the results of the TRILEGAL simulations for the case where no priors are adopted to be compared to those shown in Fig. 4, for the *All priors* case. Fig. A2 shows the results for TRILEGAL simulations with no spatial, MDF, or ADF (see Section 3.3).
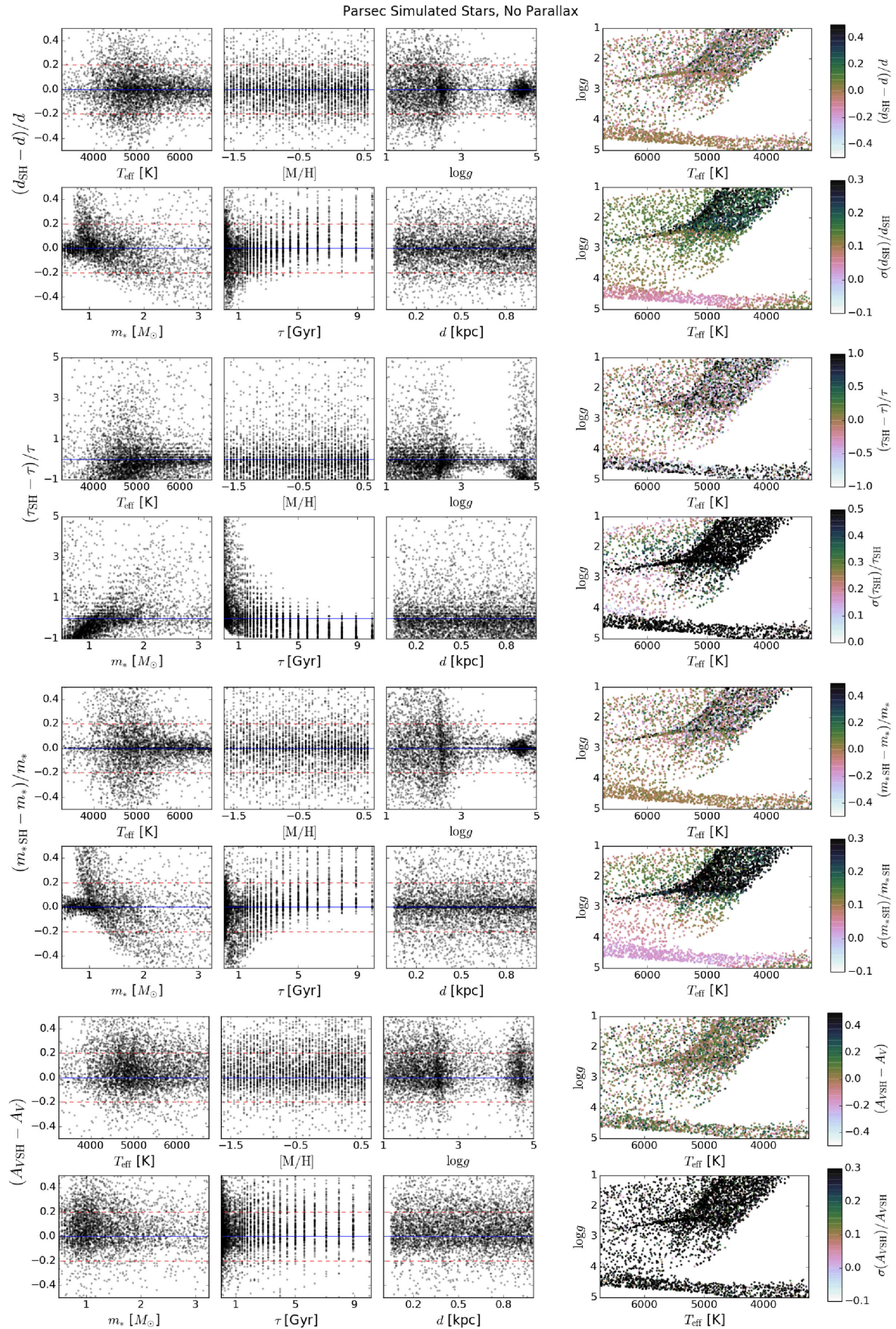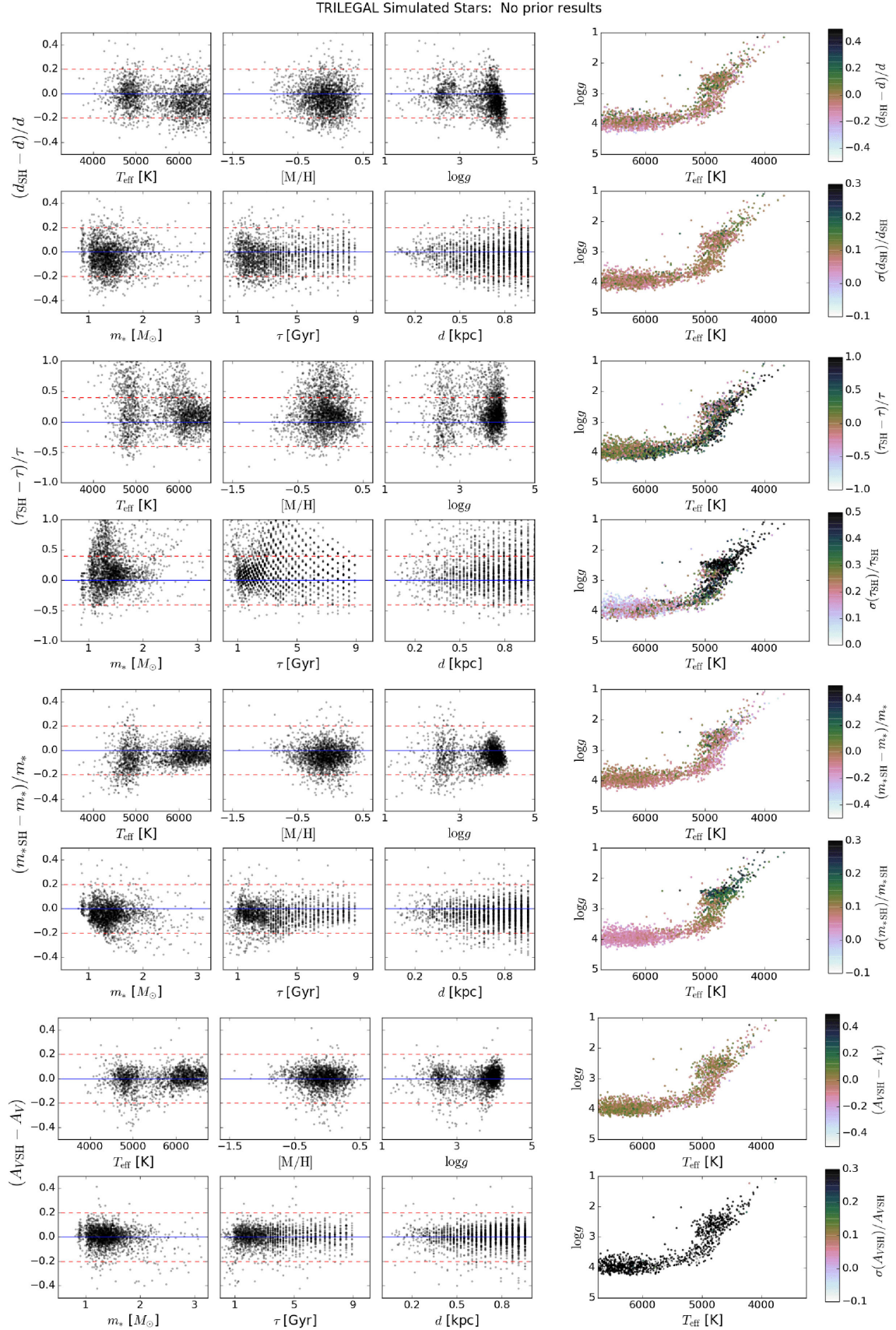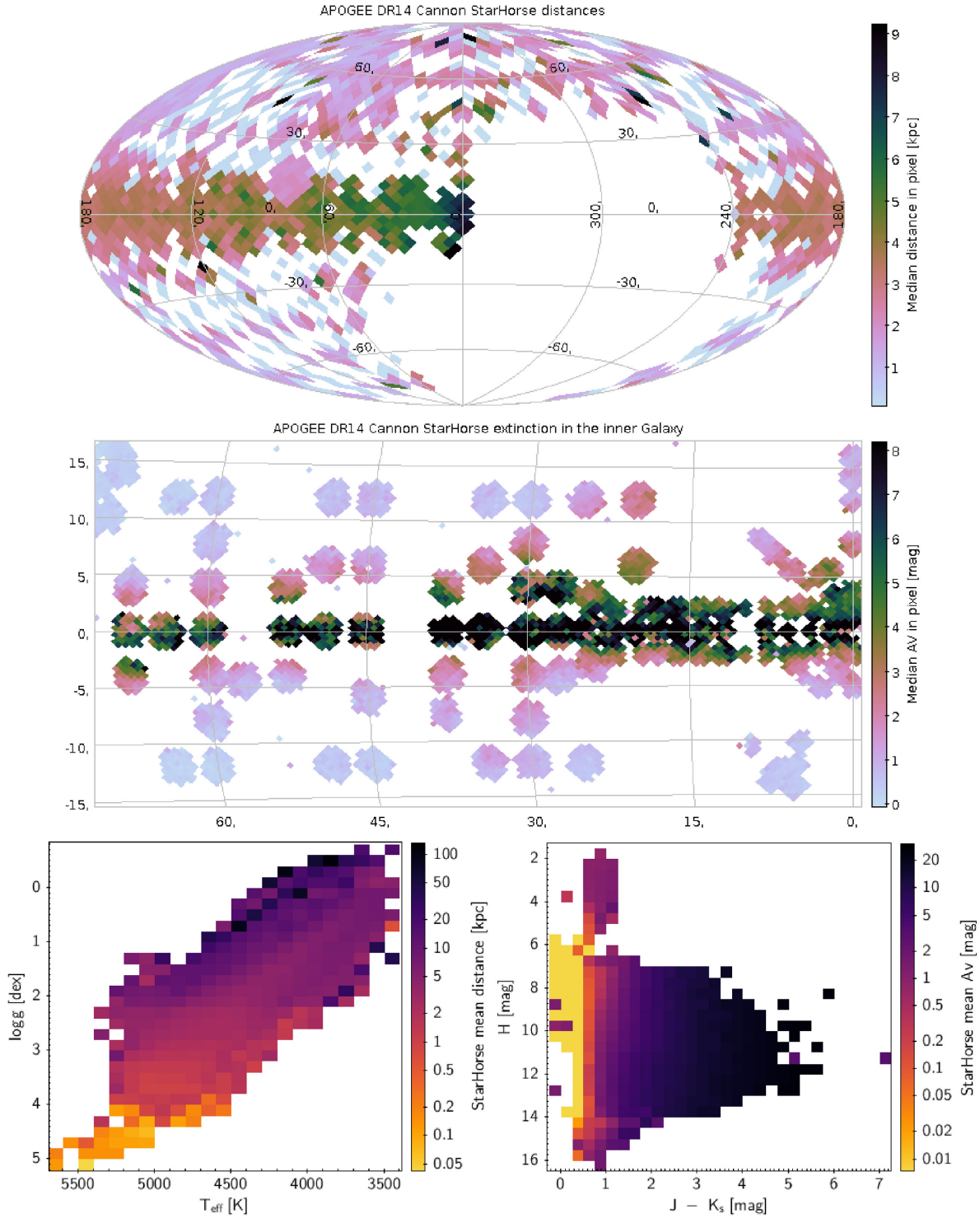
Parsec Simulated Stars, No Parallax

Figure A1. Same panels as in Fig. 3, but now showing the results from `StarHorse` when the constraint provided by the parallax is not used.

Figure A2. Same panels as in Fig. 4, but now showing the results from StarHorse when no priors in metallicity, age, and spatial distribution are adopted.

APPENDIX B: ADDITIONAL DATA RELEASED ANALYSIS

Figure B1. Illustration of the APOGEE DR14 Cannon distance and extinction results from `StarHorse`. The panels and conventions are the same as in Fig. 11.
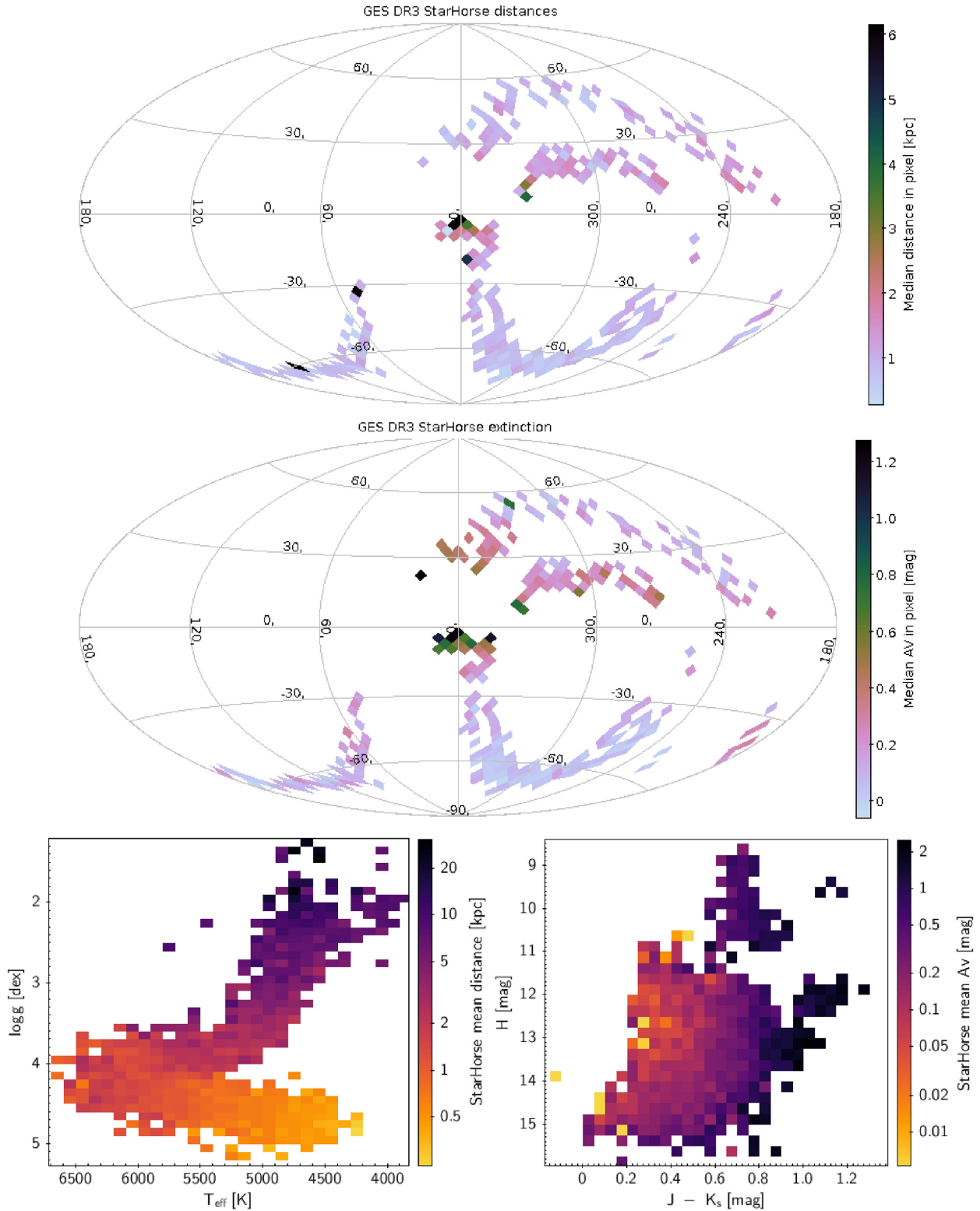
Figure B2. Illustration of the GES DR3 distance and extinction results from `StarHorse`. The panels and conventions are the same as in Fig. 11.
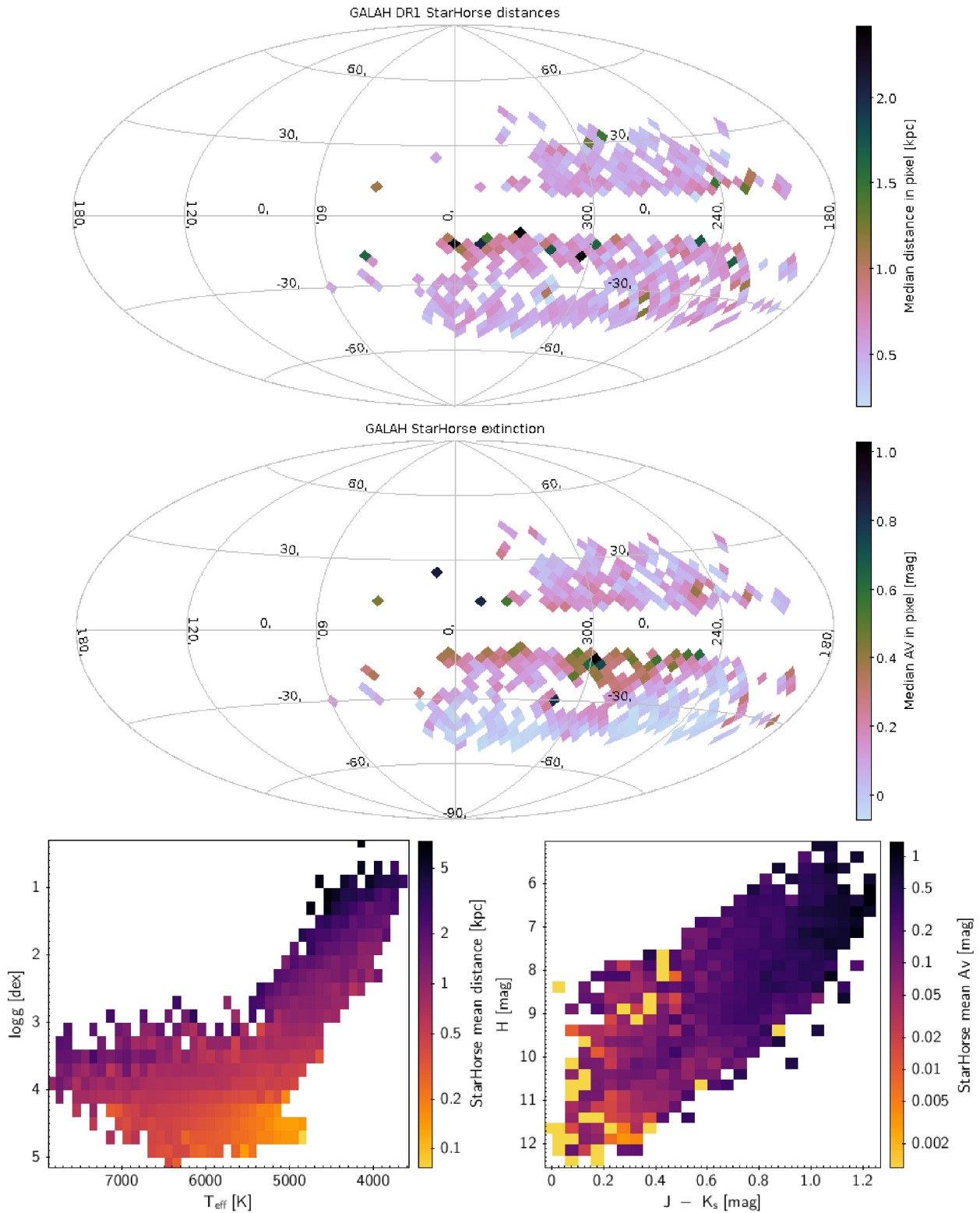
Figure B3. Illustration of the GALAH DR1 distance and extinction results from `StarHorse`. The panels and conventions are the same as in Fig. 11.
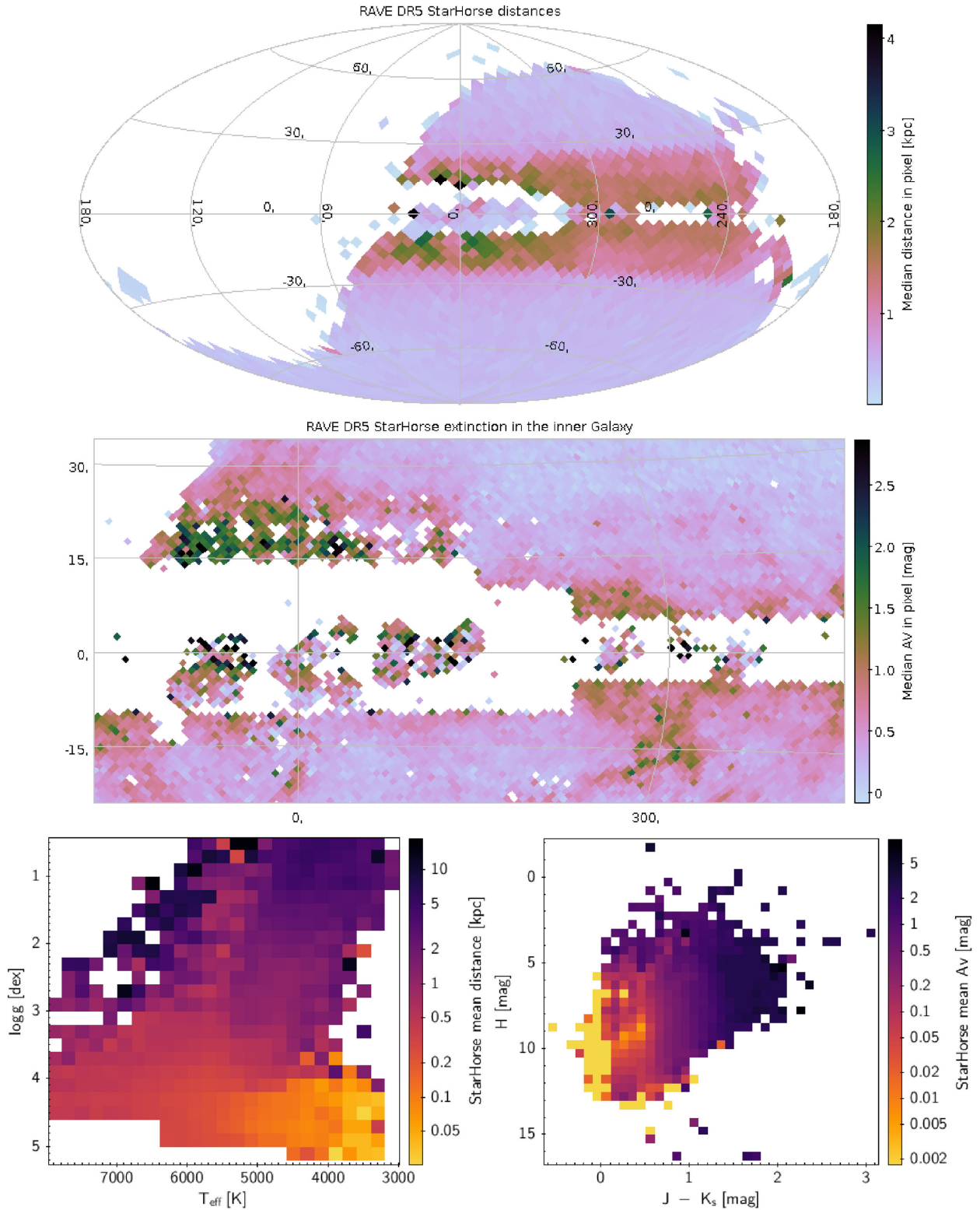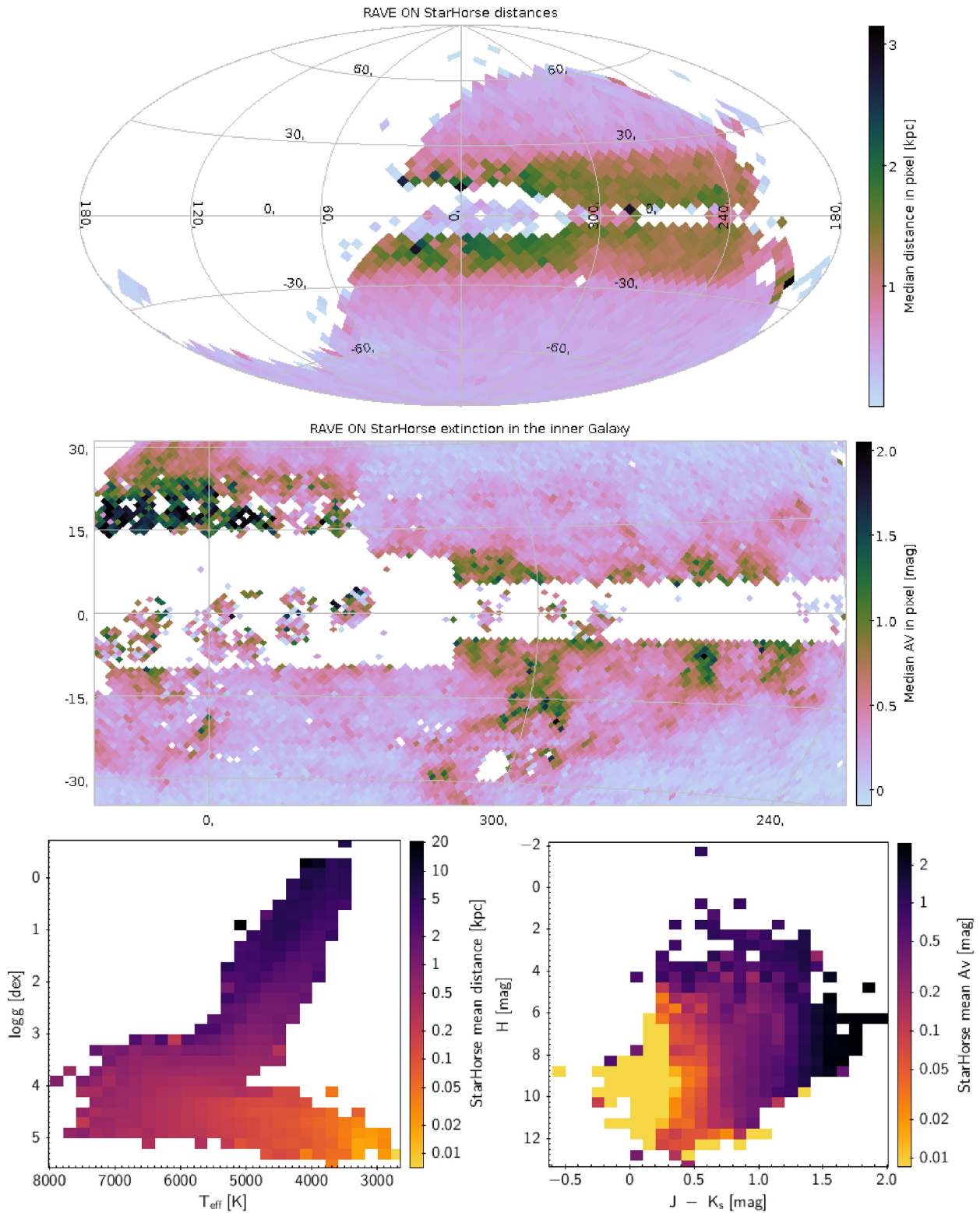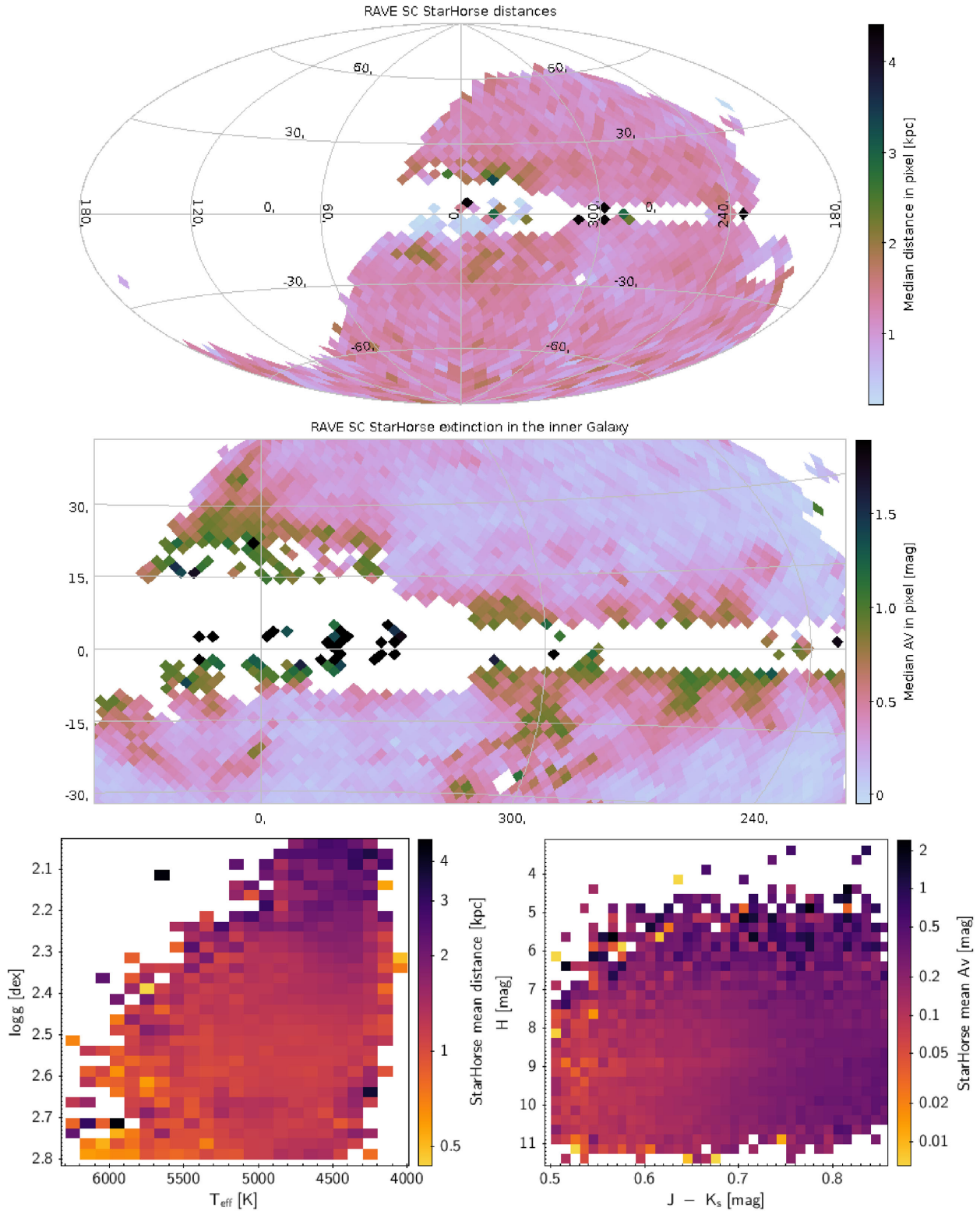
Figure B4. Illustration of the RAVE DR5 distance and extinction results from `StarHorse`. The panels and conventions are the same as in Fig. 11.

Figure B5. Illustration of the RAVE ON distance and extinction results from `StarHorse`. The panels and conventions are the same as in Fig. 11.

Figure B6. Illustration of the RAVE SC distance and extinction results from `StarHorse`. The panels and conventions are the same as in Fig. 11.

This paper has been typeset from a TEX/LATEX file prepared by the author.