The Pristine survey

Nicolas F. Martin, Else Starkenburg, Zhen Yuan, Morgan Fouesneau, Anke Ardern-Arentsen, Francesca de Angeli, Felipe Gran, Martin Montelius, Samuel Rusterucci, René Andrae, et al.

A&A, 692, A115 (2024)
https://doi.org/10.1051/0004-6361/202347633

Astronomy & Astrophysics

The Pristine survey

XXIII. Data Release 1 and an all-sky metallicity catalogue based on *Gaia* DR3 BP/RP spectro-photometry

Nicolas F. Martin[1,2,⋆], Else Starkenburg[3,⋆], Zhen Yuan (袁珍)[1], Morgan Fouesneau[2], Anke Ardern-Arentsen[4], Francesca De Angeli[4], Felipe Gran[5], Martin Montelius[3], Samuel Rusterucci[1,3], René Andrae[2], Michele Bellazzini[6], Paolo Montegriffo[6], Anna F. Esselink[3], Hanyuan Zhang[4], Kim A. Venn[7], Akshara Viswanathan[3], David S. Aguado[8,9], Giuseppina Battaglia[8,9], Manuel Bayer[3], Piercarlo Bonifacio[10], Elisabetta Caffau[10], Patrick Côté[11], Raymond Carlberg[12], Sébastien Fabbro[11], Emma Fernández-Alvar[8,9], Jonay I. González Hernández[8,9], Isaure González Rivera de La Vernhe[5], Vanessa Hill[5], Rodrigo A. Ibata[1], Pascale Jablonka[13,10], Georges Kordopatis[5], Carmela Lardo[14], Alan W. McConnachie[11], Camila Navarrete[5], Julio Navarro[7], Alejandra Recio-Blanco[5], Rubén Sánchez Janssen[15], Federico Sestito[7], Guillaume F. Thomas[8,9], Sara Vitali[16], and Kristopher Youakim[17]

[1] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
[2] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[3] Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands
[4] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[5] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Nice, France
[6] INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Piero Gobetti 93/3, 40129 Bologna, Italy
[7] Department of Physics and Astronomy, University of Victoria, PO Box 3055, STN CSC, Victoria, BC V8W 3P6, Canada
[8] Instituto de Astrofísica de Canarias, 38205 La Laguna, Tenerife, Spain
[9] Universidad de La Laguna, Dept. Astrofísica, 38206 La Laguna, Tenerife, Spain
[10] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92195 Meudon, France
[11] NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada
[12] Department of Astronomy & Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada
[13] Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
[14] Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
[15] UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[16] Núcleo de Astronomía, Facultad de Ingeniería y Ciencias Universidad Diego Portales, Ejército 441, Santiago, Chile
[17] Department of Astronomy, Stockholm University, AlbaNova University Centre, 106 91 Stockholm, Sweden

Received 2 August 2023 / Accepted 27 August 2024

ABSTRACT

We used the spectro-photometric information of ~219 million stars from *Gaia*'s Data Release 3 (DR3) to calculate synthetic, narrow-band, metallicity-sensitive *CaHK* magnitudes that mimic the observations of the Pristine survey, a survey of photometric metallicities of Milky Way stars that has been mapping more than 6500 deg^2 of the northern sky with the Canada–France–Hawaii Telescope since 2015. These synthetic magnitudes were used for an absolute recalibration of the deeper Pristine photometry and, combined with broadband *Gaia* information, synthetic and Pristine *CaHK* magnitudes were used to estimate photometric metallicities over the whole sky. The resulting metallicity catalogue is accurate down to [Fe/H] ~ −3.5 and is particularly suited for the exploration of the metal-poor Milky Way ([Fe/H] < −1.0). We make available here the catalogue of synthetic $CaHK_{\rm syn}$ magnitudes for all stars with BP/RP information in *Gaia* DR3, as well as an associated catalogue of more than ~30 million photometric metallicities for high signal-to-noise FGK stars. This paper further provides the first public data release of the Pristine catalogue in the form of higher quality recalibrated Pristine *CaHK* magnitudes and photometric metallicities for all stars in common with the BP/RP spectro-photometric information in *Gaia* DR3. We demonstrate that, when available, the much deeper Pristine data greatly enhance the quality of the derived metallicities, in particular at the faint end of the catalogue ($G_{\rm BP} \gtrsim 16$). Combined, both photometric metallicity catalogues include more than two million metal-poor star candidates ($[{\rm Fe/H}]_{\rm phot} < -1.0$) as well as more than 200 000 and ~8000 very and extremely metal-poor candidates ($[{\rm Fe/H}]_{\rm phot} < -2.0$ and < -3.0, respectively). Finally, we show that these metallicity catalogues can be used efficiently, among other applications, for Galactic archaeology, to hunt for the most metal-poor stars, and to study how the structure of the Milky Way varies with metallicity, from the flat distribution of disk stars to the spheroid-shaped metal-poor halo.

Key words. catalogs – surveys – stars: abundances – Galaxy: abundances

⋆ Corresponding authors; nicolas.martin@astro.unistra.fr; estarkenburg@astro.rug.nl. Both authors contributed equally.

1. Introduction

In any given closed environment, the lowest metallicity stars are also the oldest ones. As such, they are often thought of as the ultimate targets that allow us to look back in time to the infancy of our Galaxy through the methods of Galactic archaeology (see, e.g., Tinsley 1980; Beers & Christlieb 2005; Belokurov 2013; Frebel & Norris 2015; Helmi 2020, for detailed reviews). It is likely that most, if not all, of the first (metal-free) stars are inaccessible to us, because they are very massive and, therefore, short-lived (for a recent review, see Klessen & Glover 2023, and references therein). As a consequence, the most metal-poor stars we can still observe today not only contain information on the earliest buildup of the proto-galaxy that would later become the Milky Way (MW), but they also hold invaluable clues about the properties of the first stars and their mass function. They give us unique information on times long gone and complement studies that focus on the high-redshift Universe through observations. For example, recent results from the *James Webb* Space Telescope include observations of damped Lyman-alpha systems (e.g., Welsh et al. 2023), high-redshift, magnified stars (Welch et al. 2022a,b), or the possible direct observation of the impact of the first stars on the photoionization of gas (Maiolino et al. 2024).

However, these important tracers of the early Galaxy are very rare among the more metal-rich and younger populations of the MW. In the Solar neighborhood, looking out of the Galactic plane through the dense foreground of predominantly young and metal-rich disk stars, it has been estimated that stars with less than one-thousandth of the solar metallicity represent, at most, one star in a thousand. Below this metallicity value, the metallicity distribution function (MDF) may decline even more steeply than at higher metallicities (e.g., Youakim et al. 2020; Bonifacio et al. 2021; Yong et al. 2021), leading to a very challenging search for "extremely metal-poor" stars (EMP stars; [Fe/H] < -3.0) or "ultra-metal-poor" stars (UMP stars, following the terminology from Beers & Christlieb 2005; [Fe/H] < -4.0). At the moment, databases collecting detailed chemical abundance studies of such stars (see, e.g., the SAGA database, Suda et al. 2008, and online updates[1]) contain a few hundred EMP stars and significantly fewer UMP stars, with only ∼40 such stars currently known (e.g., Sestito et al. 2019).

Historically, one of the most successful avenues to isolate the low-metallicity end of the MDF, starting with the "very metal-poor" regime (VMP stars; [Fe/H] < -2.0), has been to employ objective-prism spectroscopy in the region of the metallicity-sensitive Ca H& K lines (Bond 1970; Bidelman & MacConnell 1973; Bessell 1977). The two most recent such endeavors, the HK survey (Beers et al. 1992) and the Hamburg-ESO survey (Christlieb et al. 2008), led to the buildup of the first significant samples of EMP stars via the spectroscopic follow-up of apparently metal-deficient stars identified in the surveys. In parallel, they helped discover the first confirmed UMP stars (Christlieb et al. 2002; Frebel et al. 2005). Large, low-resolution spectroscopic surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Large Sky Area Multi-Object Fibre Spectroscopic Telescope (LAMOST; Zhao et al. 2006) have also proven to be a major source of stars at the low-metallicity end of the MDF. Although the survey strategy was not designed to specifically target EMP and UMP candidates, samples of millions of stellar spectra naturally included a sizable number of such stars. These were discovered through the application of specifically developed pipelines and analyses (e.g., the Turn-Off Primordial Stars survey – TOPoS – Caffau et al. 2013; and see also Aguado et al. 2016 and Li et al. 2018).

The community has now mainly shifted toward using narrow-band photometric surveys to identify these exceptional stars, thanks to both the preponderance of wide-field CCD imagers and advances in filter technology that allow for the construction of well-behaved medium- and narrow-band filters (i.e., near top-hat and providing stable photometry over a wide focal plane). Two such surveys have been driving the field over the last decade: the SkyMapper panoptic survey (Wolf et al. 2018; Onken et al. 2019; Chiti et al. 2021), in the south, that includes a Strömgren v filter that covers the metallicity-sensitive Ca H & K lines in addition to the more classical *ugriz* broadband filters; and the Pristine survey (Starkenburg et al. 2017a), in the north, that relies on a specifically tailored narrow-band filter centered on the Ca H & K lines and is mounted on the MegaCam imager at the Canada-France-Hawaii Telescope (CFHT). The combination of the metallicity-sensitive medium- or narrow-band photometry with broadband photometry has led to large samples of VMP and EMP candidates. Dedicated spectroscopic follow-up surveys have confirmed that both surveys produce samples of VMP stars with a high purity and that they are efficient at isolating true EMP stars, with success rates of ∼20% when quality flags are carefully considered (Youakim et al. 2017; Aguado et al. 2019; Da Costa et al. 2019; Marino et al. 2019). Both surveys have also led to the discovery of a handful of new UMP stars (Starkenburg et al. 2018; Kielty et al. 2021; Nordlander et al. 2019; Lardo et al. 2021), with the most iron-deficient star known to date being among them (Keller et al. 2014). Very recently, additional new surveys that follow similar strategies, such as J-PLUS (Cenarro et al. 2019), S-PLUS (Almeida-Fernandes et al. 2022), and SAGES (Fan et al. 2023), have also started to produce results (Galarza et al. 2022; Placco et al. 2022; Yang et al. 2022; Huang et al. 2023).

The data provided by the recent *Gaia* Data Release 3 (DR3; Gaia Collaboration 2023b) promise to bridge the two techniques – very low-resolution spectroscopy and narrow-band photometry – with the release of the spectro-photometric observations conducted with the Blue Prism (BP) and the Red Prism (RP) on board the spacecraft (Carrasco et al. 2021; De Angeli et al. 2023; Montegriffo et al. 2023). Specifically, the BP prism includes the region of the Ca H & K lines so the *Gaia* spectro-photometry is expected to perform well at constraining the metallicity of a star, even in the EMP regime (Witten et al. 2022; Xylakis-Dornbusch et al. 2022). For more generic MW stars, the BP/RP information is already proving invaluable to derive accurate stellar parameters, including [Fe/H] (Andrae et al. 2023a,b; Bellazzini et al. 2023; Zhang et al. 2023; Xylakis-Dornbusch et al. 2024), despite being currently limited to the fairly bright sky in DR3 ($G \lesssim 17.65$).

The release of the DR3 BP/RP information is particularly exciting in the context of the Pristine survey. As shown by Gaia Collaboration (2023a), the BP/RP information – distributed by the consortium as coefficients from the projection of the observed low resolution spectra on a set of basis functions – can be used to reproduce synthetic photometry for any filter that overlaps the large wavelength coverage of the observations (330–1050 nm). In particular, they show that synthetic Pristine $CaHK_{\rm syn}$ magnitudes compare favorably with the actual Pristine observations (see their Fig. 32). This brings a unique opportunity: (1) to supersede the previous, relative calibration of the Pristine survey and calibrate all of the narrow-band $CaHK$ photometry on an absolute scale provided by the $CaHK_{\rm syn}$

[1] http://sagadatabase.jp

Table 1. Description of the columns of the photometry catalogue that includes both the synthetic $CaHK_{\rm syn}$ and the Pristine $CaHK$ photometric information.

Column	Description	Unit	Type
RA	*Gaia* DR3 right ascension	ICRS (J2016)	float
Dec	*Gaia* DR3 declination	ICRS (J2016)	float
source_id	*Gaia* DR3 source_id	–	longint
CaHK_flux	CaHK mean flux	e^-/s	float
d_CaHK_flux	uncertainty on the CaHK mean flux	e^-/s	float
CaHK_syn	$CaHK_{\rm syn}$, synthetic Pristine-like $CaHK$ magnitude	mag	float
d_CaHK_syn	$\delta CaHK_{\rm syn}$, uncertainty on $CaHK_{\rm syn}$	mag	float
Pvar	Probability for the source to be variable, as defined in Section 5	–	float
RA_Pr	Pristine right ascension	ICRS (J2015.5)	float
Dec_Pr	Pristine declination	ICRS (J2015.5)	float
CaHK_Pr	$CaHK$ Pristine magnitude	mag	float
d_CaHK_Pr	Uncertainty on the $CaHK$ Pristine magnitude	mag	float
merged_CASU_flag	Morphology and quality flag[2]	–	int

magnitudes[3], and (2) to build a catalogue of *Gaia*-based synthetic $CaHK_{\rm syn}$ Pristine-like photometry over the whole sky. The latter can be pushed through an updated version of the Pristine photometric metallicity model to yield photometric metallicities that are accurate down to the EMP regime.

This paper describes both these efforts. In addition to information based on the shallower $CaHK_{\rm syn}$ magnitudes, we assemble and distribute the first public Data Release (DR1) of the Pristine survey that now includes the newly recalibrated photometry and metallicities from the updated photometric metallicity model. This data release includes all sources observed in the Pristine survey by the end of March 2024 that have BP/RP information in *Gaia* DR3.

The paper is organized as follows. Section 2 presents the catalogue of *Gaia*-based synthetic $CaHK_{\rm syn}$ photometry based on the BP/RP information. Section 3 provides a detailed update to the data reduction of the now ∼11,500 images of the Pristine survey and, in particular, explains the calibration of the survey photometry onto the absolute scale provided by the *Gaia* synthetic magnitudes. In Section 4, we describe how we handle the complex issue of correcting observed magnitudes from extinction, while Section 5 presents a probabilistic model to isolate likely variable stars that can have erroneous photometric metallicities. Section 6 describes the updated Pristine $(CaHK, G, G_{\rm BP}, G_{\rm RP}) \rightarrow [{\rm Fe/H}]_{\rm phot}$ model that is now based on *Gaia* information alone to complement the Pristine photometry. The catalogues of photometric metallicities generated from pushing the *Gaia*-based $CaHK_{\rm syn}$ and the Pristine $CaHK$ magnitudes through the model are detailed in Section 7, in which we also compare the resulting data sets with spectroscopic catalogues of metal-poor stars and other metallicity catalogues based on *Gaia* DR3. In the same section, we also provide advice on how to best use the photometric metallicity catalogues. In Section 8, we show two possible applications of the catalogues, one to study Galactic globular clusters and another to map the MW as a function of metallicity. Finally, Section 9 summarizes the contents of the paper.

Accompanying this paper, we provide a number of large catalogues that are accessible through the CDS.

The $CaHK_{\rm syn}$ photometric catalogue that contains the synthetic, Pristine-like $CaHK_{\rm syn}$ magnitudes for all stars with BP/RP coefficients information in *Gaia* DR3. It is described in Section 2. The Pristine DR1 photometric catalogue contains the Pristine $CaHK$ magnitudes for all stars in the Pristine footprint as of the end of March 2024 that also have BP/RP information in *Gaia* DR3. It is described in Section 3. Both photometric catalogues are merged into a single file[4], whose content is described in Table 1.

The Pristine-*Gaia* synthetic metallicity catalogue contains the photometric metallicities created from pushing relevant stars from the $CaHK_{\rm syn}$ catalogue through the Pristine $(CaHK, G_{\rm BP}, G, G_{\rm RP}) \rightarrow [{\rm Fe/H}]_{\rm phot}$ model. It contains $[{\rm Fe/H}]_{\rm CaHKsyn}$ photometric metallicities over the whole sky and is most accurate in the bright regime ($G_{\rm BP} \lesssim 16.0$). The catalogue is available online[5] and described in Section 7 and Table 2.

The Pristine DR1 metallicity catalogue contains the photometric metallicities created from pushing the relevant stars from the Pristine $CaHK$ DR1 catalogue through the Pristine $(CaHK, G, G_{\rm BP}, G_{\rm RP}) \rightarrow [{\rm Fe/H}]_{\rm phot}$ model. It retains high S/N photometry for the full magnitude range of the *Gaia* BP/RP data, and therefore provides more accurate $[{\rm Fe/H}]_{\rm Pristine}$ photometric metallicities but over a smaller footprint of ∼6500 deg^2. The catalogue is available online[6] and described in Section 7 and Table 3.

2. The *Gaia* synthetic *CaHK* catalogue

Gaia-based synthetic $CaHK$ magnitudes, $CaHK_{\rm syn}$, are calculated using GaiaXPy[7] for all 219.2 million objects for which the

[2] Merged flag based on the CASU_flag of individual detections, as mentioned in Section 3.3.1. The flag discriminates between saturated stars (-9), likely point sources (-2), very likely point sources (-1), extended sources (+1), discrepant flags from individual detections (+2), merged source from individual detections with inconsistent photometry (+3).

[3] A similar effort was also successfully undertaken by the J-PLUS collaboration to calibrate their multi-narrowband photometry, including photometry from their own CaHK filter (López-Sanjuan et al. 2024).

[4] 219.2 million rows, 6.4 million of which also include Pristine DR1 photometry.

[5] 52.3 million rows.

[6] 6.4 million rows.

[7] https://pypi.org/project/GaiaXPy/

Table 2. Description of the columns of the Pristine-*Gaia* synthetic metallicity catalogue.

Column	Description	Unit	Type
source_id	*Gaia* DR3 source_id	–	longint
RA	*Gaia* DR3 right ascension	ICRS (J2016)	float
Dec	*Gaia* DR3 declination	ICRS (J2016)	float
E(B-V)	Schlegel et al. (1998) extinction value	mag	float
G_0	De-reddened *Gaia G* magnitude	mag	float
d_G	δG, uncertainty on the *Gaia G* magnitude	mag	float
BP_0	De-reddened *Gaia* G_{BP} magnitude	mag	float
d_BP	δG_{BP}, uncertainty on the *Gaia* G_{BP} magnitude	mag	float
RP_0	De-reddened *Gaia* G_{RP} magnitude	mag	float
d_RP	δG_{RP}, uncertainty on the *Gaia* G_{RP} magnitude	mag	float
CaHK_0	De-reddened Pristine-like $CaHK_{\rm syn}$ magnitude	mag	float
d_CaHK_syn	$\delta CaHK_{\rm syn}$, uncertainty on the Pristine-like $CaHK_{\rm syn}$ magnitude	mag	float
FeH_CaHKsyn	Photometric metallicity $[{\rm Fe/H}]_{\rm phot}$ based on the $CaHK_{\rm syn}$ magnitude and using the (favored) giant model	dex	float
FeH_CaHKsyn_16th	Lower bound of the 68% confidence interval on FeH_CaHKsyn based on the Monte Carlo sampling of the magnitude uncertainties	dex	float
FeH_CaHKsyn_50th	Median value of FeH_CaHKsyn based on the Monte Carlo sampling of the magnitude uncertainties	dex	float
FeH_CaHKsyn_84th	Upper bound of the 68% confidence interval on FeH_CaHKsyn based on the Monte Carlo sampling of the magnitude uncertainties	dex	float
mcfrac_CaHKsyn	Fraction of the Monte Carlo samples of the magnitude uncertainties that fall outside the model color–color space	–	float
G_0_dw	Same as G_0 but for the dwarf model	mag	float
BP_0_dw	Same as BP_0 but for the dwarf model	mag	float
RP_0_dw	Same as RP_0 but for the dwarf model	mag	float
CaHK_0_dw	Same as CaHK_0 but for the dwarf model	mag	float
FeH_CaHKsyn_dw	Same as FeH_CaHKsyn_dw but for the dwarf model	dex	float
FeH_CaHKsyn_dw_16th	Same as FeH_CaHKsyn_dw_16th but for the dwarf model	dex	float
FeH_CaHKsyn_dw_50th	Same as FeH_CaHKsyn_dw_50th but for the dwarf model	dex	float
FeH_CaHKsyn_dw_84th	Same as FeH_CaHKsyn_dw_84th but for the dwarf model	dex	float
mcfrac_CaHKsyn_dw	Same as mcfrac_CaHKsyn but for the dwarf model	–	float
Pvar	Probability for the source to be variable, as defined in Section 5	–	float
RUWE	*Gaia* DR3 Renormalised Unit Weight Error	–	float
Cstar	*Gaia* DR3 corrected flux excess, C^*, as defined in equation (6) of Riello et al. (2021)	mag	float
Cstar_1sigma	Normalized standard deviation of C^* for the G magnitude of this source, as defined in equation (18) of Riello et al. (2021)	–	float

Gaia DR3 includes BP/RP coefficients. We follow the procedure presented in Gaia Collaboration (2023a), integrating the BP/RP spectra under the curve of the MegaCam CaHK[8] filter, available on the CFHT website[9]. This narrow-band filter has the significant advantage of presenting a near-top-hat transmission curve and of being centered on the Calcium H & K lines (near 395 nm) that are very sensitive to the metallicity of a MW FGK star (see Figure 1 of Starkenburg et al. 2017a). By construction, this catalogue covers the fairly bright sky as the *Gaia* catalogue only includes the BP/RP information for objects detected with sufficiently high signal-to-noise (S/N). The added difficulty of working toward the blue end of the BP spectrum, which has lower S/N than most of the BP/RP wavelength coverage, means that uncertainties on the $CaHK_{\rm syn}$ magnitudes can be very large. We nevertheless include them in the catalogue, but the reader should be aware that quality cuts and a careful treatment of uncertainties are necessary and essential to use this catalogue. For instance, in the main case presented in this paper, that of calculating photometric metallicities based on the Pristine $(CaHK, G, G_{\rm BP}, G_{\rm RP}) \rightarrow [{\rm Fe/H}]_{\rm phot}$ model, we only consider stars with $CaHK_{\rm syn}$ uncertainties $\delta CaHK_{\rm syn} < 0.1$. The left-hand panel of Figure 1 shows that this limit is reached for $G_{BP} \sim$15–17. This large range stems from the broad distribution of colors of the catalogue stars: since the CaHK filter is very blue the resulting S/N of a blue or very red star observed under the same conditions by *Gaia* leads to much larger $CaHK_{\rm syn}$ uncertainties for the latter.

From Figure 1, it is evident that the *Gaia*-based $CaHK_{\rm syn}$ is most useful at the bright end and that, when available, the Pristine data usually has much higher quality (right-hand panel). The new catalogue is nevertheless extremely valuable for two main reasons: it gives us a ground truth against which we can finally calibrate the Pristine survey to a very high level of photometric accuracy, and it provides an all-sky (bright) equivalent to Pristine on which we can apply the Pristine model to build an all-sky

[8] In what follows, CaHK refers to the filter and *CaHK* refers to the magnitudes obtained using this filter.
[9] https://www.cfht.hawaii.edu/Instruments/Filters/megaprimenew.html

Table 3. Description of the columns of the Pristine DR1 metallicity catalogue.

Column	Description	Unit	Type
source_id	*Gaia* DR3 source_id	–	longint
RA	*Gaia* DR3 right ascension	ICRS (J2016)	float
Dec	*Gaia* DR3 declination	ICRS (J2016)	float
E(B-V)	Schlegel et al. (1998) extinction value	mag	float
merged_CASU_flag	Morphology and quality flag[10]	–	int
G_0	De-reddened *Gaia G* magnitude	mag	float
d_G	δG, uncertainty on the *Gaia G* magnitude	mag	float
BP_0	De-reddened *Gaia* G_{BP} magnitude	mag	float
d_BP	δG_{BP}, uncertainty on the *Gaia* G_{BP} magnitude	mag	float
RP_0	De-reddened *Gaia* G_{RP} magnitude	mag	float
d_RP	δG_{RP}, uncertainty on the *Gaia* G_{RP} magnitude	mag	float
CaHK_0	De-reddened Pristine *CaHK* magnitude	mag	float
d_CaHK	$\delta CaHK$, uncertainty on the Pristine *CaHK* magnitude	mag	float
FeH_Pristine	Photometric metallicity $[Fe/H]_{phot}$ based on the Pristine *CaHK* magnitudes and using the (favored) giant model	dex	float
FeH_Pristine_16th	Lower bound of the 68% confidence interval on FeH_Pristine based on the Monte Carlo sampling of the magnitude uncertainties	dex	float
FeH_Pristine_50th	Median value of FeH_Pristine based on the Monte Carlo sampling of the magnitude uncertainties	dex	float
FeH_Pristine_84th	Upper bound of the 68% confidence interval on FeH_Pristine based on the Monte Carlo sampling of the magnitude uncertainties	dex	float
mcfrac_Pristine	Fraction of the Monte Carlo samples of the magnitude uncertainties that fall outside the model color–color space	–	float
G_0_dw	Same as G_0 but for the dwarf model	mag	float
BP_0_dw	Same as BP_0 but for the dwarf model	mag	float
RP_0_dw	Same as RP_0 but for the dwarf model	mag	float
CaHK_0_dw	Same as CaHK_0 but for the dwarf model	mag	float
FeH_Pristine_dw	Same as FeH_Pristine_dw but for the dwarf model	dex	float
FeH_Pristine_dw_16th	Same as FeH_Pristine_dw_16th but for the dwarf model	dex	float
FeH_Pristine_dw_50th	Same as FeH_Pristine_dw_50th but for the dwarf model	dex	float
FeH_Pristine_dw_84th	Same as FeH_Pristine_dw_84th but for the dwarf model	dex	float
mcfrac_Pristine_dw	Same as mcfrac_Pristine but for the dwarf model	–	float
Pvar	Probability for the source to be variable, as defined in Section 5	–	float
RUWE	*Gaia* DR3 Renormalised Unit Weight Error	–	float
Cstar	*Gaia* DR3 corrected flux excess, C^*, as defined in equation (6) of Riello et al. (2021)	mag	float
Cstar_1sigma	Normalized standard deviation of C^* for the G magnitude of this source, as defined in equation (18) of Riello et al. (2021)	–	float

catalogue of MW metal-poor stars. It is however important to keep in mind that this catalogue is impacted by the *Gaia* scanning law (e.g., Cantat-Gaudin et al. 2023). In particular, the S/N and minimum number of visits required for the publication of the BP/RP coefficients in the *Gaia* DR3 archive translates to a nonuniform inclusion of stars in the catalogue. More generically, the varying number of visits at different locations of the sky that are imparted by the *Gaia* scanning law translates into a nonuniform S/N over the sky, even for stars with the same properties. This is visible in the left-hand panel of Figure 2, which presents the density of stars in the $CaHK_{syn}$ catalogue. Patches of the sky with significantly lower densities of stars are clearly visible and result in a complicated footprint and inhomogeneous densities of stars in the resulting $CaHK_{syn}$ catalogue.

The full catalogue that contains $CaHK_{syn}$ values and their uncertainties are made available online at CDS for the 219.2 million sources with BP/RP coefficients from *Gaia* DR3 and its content is described in Table 1.

3. The Pristine *CaHK* survey

3.1. The Pristine photometric data

The design of the Pristine survey, along with its core science goals, are presented in detail in Starkenburg et al. (2017a). In a nutshell, the Pristine survey is a narrow-band photometric survey that relies on the CaHK narrow-band filter that was procured in 2014 by the CFHT for its wide-field imager MegaCam (Boulade et al. 2003). Thanks to the large $\sim 1 \times 1\,deg^2$ field of view of the MegaCam camera and the dedication of the CFHT engineering and service observing crew, we have now gathered ~11 500 images with the CaHK filter. These cover more than $6500\,deg^2$ of the north and south Galactic caps. The current coverage of the Pristine survey is shown in the right-hand panel of Figure 2. The observational set-up and the data reduction of Pristine remain in general the same as those described in Starkenburg et al. (2017a) for the first two semesters of observations. In what follows, we mainly focus on the updates to the data reduction pipeline.

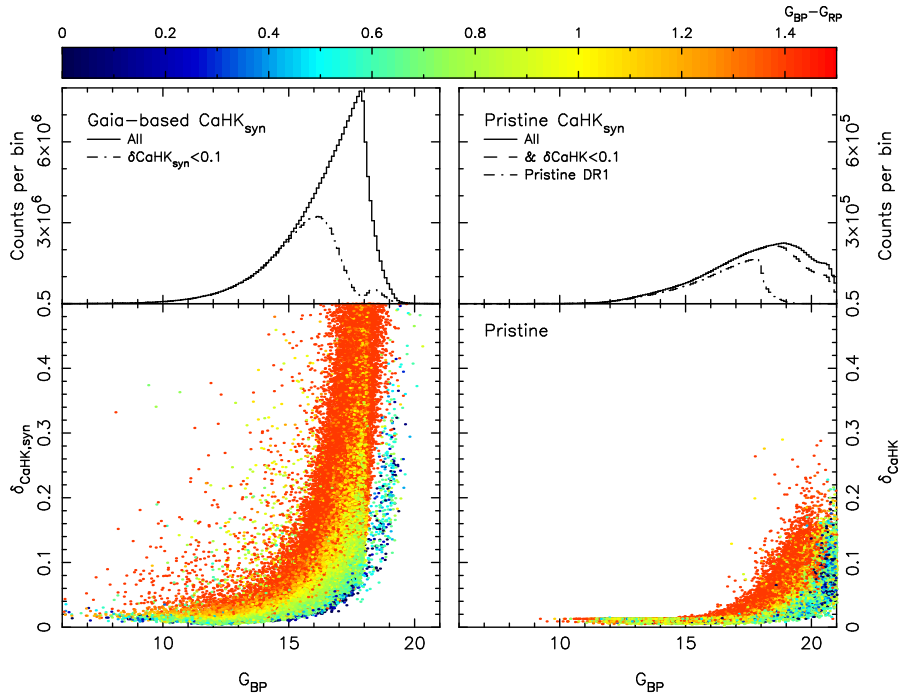

Fig. 1. Quality of the *CaHK* magnitudes. Bottom panels: *CaHK* magnitude uncertainties for stars in a representative sample of the *Gaia*-based $CaHK_{\rm syn}$ (left) and Pristine (right) catalogues. The colors code the $G_{\rm BP} - G_{\rm RP}$ color of a given star and show the clear impact of a star's color on the *CaHK* signal-to-noise in both samples. The sharp faint edge in the *Gaia*-based synthetic catalogue is driven by the DR3 $G < 17.65$ cut and the small number of fainter sources correspond to specific fainter targets included in the catalogue by the *Gaia* collaboration (De Angeli et al. 2023). The Pristine survey, on the other hand, has significantly higher signal-to-noise and only reaches photometric uncertainties $\delta CaHK = 0.1$ at $19.0 \lesssim G \lesssim 21.0$. Top panels: histograms of star counts as a function of magnitude, with the line style tracking different samples, as labeled in the panels.

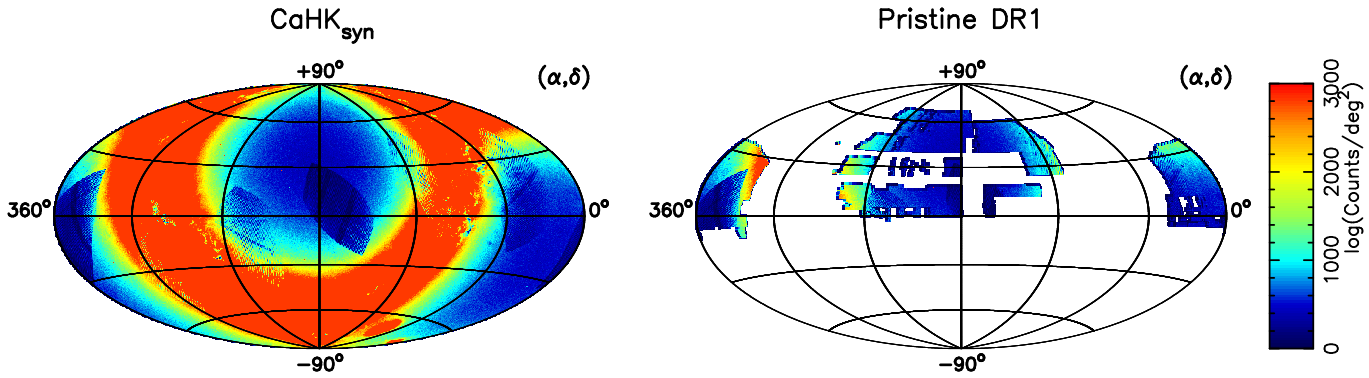

Fig. 2. Maps of the source density in the $CaHK_{\rm syn}$ catalogue (left) and in its cross-match with the Pristine catalogue (right). The localized artifacts in the density values, particularly visible in the map of $CaHK_{\rm syn}$ sources are a consequence of the *Gaia* scanning law.

Since semester 2016B, Pristine is set up as a CFHT snapshot (i.e., poor-weather) program and we enforce no restriction on observing conditions. While this means that, over time, some of the fields had to be reobserved because the initial observing conditions proved too challenging, it also ensures that a sizable number of images are regularly observed. This strategy also leads to a footprint that is not contiguous and is instead driven by poor weather or gaps in the observing queue of normal CFHT programs. While we have been striving, over the years, to fill in the holes in the coverage, some of these remain, especially in the region of the northern Galactic cap (right-hand panel of Figure 2). Our agreement with the *William Herschel* Telescope Enhanced Area Velocity Explorer (WEAVE) Galactic Archaeology spectroscopic survey for the follow-up of EMP star candidates (Jin et al. 2024) leads us to favor covering the expected footprint of the "WEAVE Galactic Archaeology Low Resolution high latitude" (WEAVE-GA LR-highlat) subsurvey that focuses on halo regions ($|b| \gtrsim 25°$ and $\delta > 0°$).

The initial survey strategy, when Pristine was observed as a normal program at CFHT and that we presented in Starkenburg et al. (2017a), included exposures of 1×100 s or 2×100 s, depending on the rankings of our programs. Since 2016B, we have settled on a single 200 s exposure for all fields. After the images are observed, they are preprocessed by the CFHT staff with the `elixir` software to remove most of the instrumental signatures (de-biasing, flat-fielding; Magnier & Cuillandre 2004). All individual preprocessed exposures are retrieved from the Canadian Astronomy Data Center (CADC) archive at this

stage and we use the pipeline from the Cambridge Astronomy Survey Unit (CASU; Irwin & Lewis 2001; Irwin et al. 2004) for the next steps of data reduction. Initially, in Starkenburg et al. (2017a), only the central 36 CCDs of the images were used, but we now treat all 40 CCDs of the imager in our analysis, including the "ears" that are formed by the 4 CCDs to the east and west of the central square-degree of the field-of-view[10].

We refine the astrometry of the images downloaded from the archive and we now use *Gaia* DR2 (Lindegren et al. 2018)[11] as an astrometric reference instead of 2MASS. The small exposure times, combined with the sometimes low density of stars toward the MW halo, the limited number of photons that pass through the very blue and narrow filter, and the sometimes challenging observing conditions mean that a small number of images fail at this astrometry stage. The corresponding fields are then sent back in the queue to be reobserved later.

After the astrometry is refined, we also use the CASU pipeline to perform aperture photometry on the images. This step can be made difficult by the low S/N of some observations and the absence of many high S/N, bright stars from which the pipeline can build a reliable point-spread function (PSF) and, then, a reliable aperture correction. This step, which is conducted independently for each CCD of each image, sometimes fails and generates obviously wrong aperture corrections. In such cases, the average aperture correction of successful CCDs on a given image is propagated to the CCDs of the same image that failed the photometry step. In total, this is necessary for a few percent of all the CCDs. We show below that, after calibration, the resulting photometry is very stable despite these necessary corrections.

The CASU pipeline automatically assigns to each source a flag that tracks its morphology: whether it is point-source-like (`CASU_flag = -1` for very likely stars, `-2` for likely stars, or `-9` for saturated stars with extrapolated photometry), appears extended (`+1` for a source that is likely extended), or resembles noise (`0`). However, the difficulty to sometimes determine a good PSF from a limited number of high S/N stars on poorly populated CCDs means that this flag is not always reliable. From experience and numerous tests, we find that this step is still reliable to discriminate between spurious and real astrophysical objects but does not always work efficiently to separate point sources from extended sources. This is not dramatic since Pristine is a fairly bright survey tailored to the *Gaia* depth and all of the envisaged Pristine science requires the contribution of broad-band photometry (SDSS before, now typically *Gaia*), whose star/galaxy discriminator can be used to assign a morphological class to a given source. Consequently, we keep in the catalogue sources that are not defined as noise by the CASU pipeline (`CASU_flag` $\neq$ `0`).

3.2. Data calibration

Previous versions of the Pristine photometry were calibrated relatively to each other using the stellar locus of red dwarf stars (using SDSS photometry, $(g-i)_0 > 1.2$; Starkenburg et al. 2017a) because, at the time, there was no absolute photometric scale to calibrate against. With the availability of the all-sky, beautifully and absolutely calibrated *Gaia*-based $CaHK_{\rm syn}$ catalogue, we now rework our procedure to calibrate against this catalogue. In Starkenburg et al. (2017a), we showed that two different effects needed to be taken into account for the calibration: an offset in the calibration of a given field that depends on the observing conditions (clouds, dust in the atmosphere, reflectivity of the mirror) and a variation of the zero point of the $CaHK$ magnitudes as a function of the location on the field of view (FoV). This latter effect is well known and not subtle, and can produce magnitude differences larger than 0.1 for the same star depending on whether it is observed at the center of the field or its outskirts (see Starkenburg et al. 2017a and, e.g., Figure 3 of Ibata et al. 2017a; see also Regnault et al. 2009).

Until now, the contribution of both effects to the calibration was determined separately by comparing the median location of the red part of the stellar locus, first dealing with the zero-point offset before building a smooth model of magnitude offsets as a function of the location on the FoV by combining all the observations of a given semester, once again using the red part of the stellar locus. There are, however, reasons to believe that the shape and location of these offsets subtly changes each time the camera is placed on the telescope, that is for every MegaCam "run" that happens once or twice a month (Ibata et al. 2017a). A quick comparison between the catalogue generated using the previous calibration and the $CaHK_{\rm syn}$ catalogues shows that, in general, the previous zero-point offsets are good, with only a small fraction of fields showing significant offsets. Similarly, the FoV correction was adequate, but nevertheless shows some artifacts when compared to the $CaHK_{\rm syn}$ catalogue. We also find that the systematic floor on the $CaHK$ uncertainties, which was previously determined to be 18 mmag from a mosaic of overlapping fields observed under good conditions during the same semester, was significantly underestimated. This number was in fact closer to 40 mmag once taking into account all survey fields observed under very different conditions over a period of 8 years.

3.2.1. Updated calibration model

Taking the $CaHK_{\rm syn}$ magnitudes as the truth to calibrate against, we now perform the two corrections simultaneously for all images of a given run and allow for more flexible and detailed, run-specific FoV models. To achieve this goal, we build a simple neural network model, `PhotCalib`[12], that we apply to the independent photometric catalogues of all ∼11 500 Pristine images observed during 85 MegaCam runs between semesters 2015A and 2023B (until run 23Bm03 in October 2023; no additional data where gather after this run until April 2024).

In practice, `PhotCalib` is fed with all the photometric catalogues of Pristine images observed in run j, to which it assigns a zero-point offset, $zp(i)$, as a free parameter for each image i, and models the FoV correction of this run, $FOV_j(X, Y)$, as a smooth function that depends on the (X, Y) position in the field. For a star of uncalibrated magnitude $CaHK_{\rm uncalib}$, the calibrated magnitude, $CaHK_{\rm calib}$, is therefore

$$CaHK_{\rm calib} = CaHK_{\rm uncalib} + zp(i) + FOV_j(X, Y). \quad (1)$$

Specifically, we use a free n-D parameter tensor to describe the $zp(i)$ offsets and the FOV_j model uses three fully connected neural layers, with each layer having 200 neurons, followed by an activation and a normalization layer. The input data for calibration are the positions of the stars (X, Y) in the FoV, the image number i, and both the $CaHK_{\rm uncalib}$ and $CaHK_{\rm syn}$ catalogues

[10] https://www.cfht.hawaii.edu/Instruments/Imaging/Megacam/specsinformation.html

[11] Switching to *Gaia* (E)DR3 leads to no significant improvement of the astrometry compared to *Gaia* DR2. To avoid having to reprocess the thousands of images observed before December 2020, we continue using *Gaia* DR2 as a reference. The resulting rms on the astrometry of stars is consistently better than 0.1″.

[12] https://github.com/zyuan-astro/PhotCalib

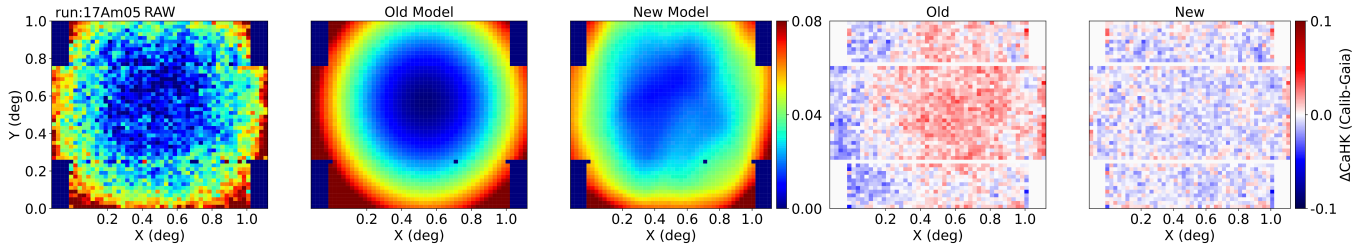

Fig. 3. Field-of-view corrections of run 17Am05. From left to right: the un-calibrated data after the field zero-point offset corrections ($CaHK_{\rm uncalib} + zp(i) - CaHK_{\rm syn}$); the old analytic model; the new FoV model from `PhotCalib`; the median residual over the field of view between the calibrated and the synthetic *Gaia* values ($\Delta CaHK = CaHK_{\rm calib} - CaHK_{\rm syn}$) from the old and the new model.

for a given run. The loss function is simply the uncertainty-weighted chi-square between the $CaHK_{\rm calib}$ values, as defined in the equation above, and the corresponding $CaHK_{\rm syn}$ values. For fast convergence, we use the stochastic gradient descent optimizer, which remains fast even for the largest run of almost 1000 images and 404 144 stars.

3.2.2. Setup and application

To ensure that we determine the calibration models on reliable stars, we apply the following quality cuts to the Pristine data set:

- `CASU_flag = -1`;
- $\delta CaHK_{\rm syn} < 0.1$;
- no star within radius $r_{\rm max}$ from the center of globular clusters, with $r_{\rm max}$ the rough estimate of the size of a cluster defined by Vasiliev & Baumgardt (2021);
- no star with $|CaHK_{\rm uncalib} - CaHK_{\rm syn,med}| > 0.2$, with $CaHK_{\rm syn,med}$ the median $CaHK_{\rm syn}$ value of the considered field (these stars are likely variable stars[13] or catastrophic failures).

Figure 3 visualizes the results of the FoV model for a randomly chosen run, 17Am05 (89 994 stars with $\delta CaHK_{\rm syn} < 0.1$, a number that is typical of most runs, although see below for the special handling a particularly small runs). In the leftmost panel, we show the median magnitude offsets $\Delta CaHK = CaHK_{\rm uncalib} + zp(i) - CaHK_{\rm syn}$, after correcting the photometry of all images by the corresponding zero-point offset. It makes the need for the FoV correction particularly evident. The old FoV model is shown in the second panel of the figure and, while it reproduced the global shape of the offsets, it was clearly not perfect. The FoV model determined by `PhotCalib` is shown in the third panel. Its increased flexibility more subtly captures the irregularities in the FoV offsets. This is confirmed in the fourth and fifth panels of the figure that present the residuals after applying the two FoV models. While median magnitude offsets with the $CaHK_{\rm syn}$ magnitudes remain small with the previous model (usually within ±0.03) and do not invalidate the past Pristine catalogues, the application of `PhotCalib` yields significantly flatter photometry.

We achieve similar results to those presented in Figure 3 for most runs. Five of the runs contain fewer than 3000 good quality stars each and yield visually poorer results for their FoV models. For those, we combine two small neighboring runs (21Bm03 and 21Bm04) with 1073 and 2889 stars, respectively, to obtain a single model for both runs. For the other three small runs (16Bm02, 19Am05, and 22Am04), the number of stars is always less than 10% of that of the neighboring runs. For those, we chose to use the FoV model from the two neighboring runs and only let `PhotCalib` determine the zero-point offsets of each image. From these two models, we compare the distribution of the residuals after the calibration and pick the one with the smallest intrinsic dispersion as the best model.

Overall, contiguous runs show subtle variations in their FoV models, with a range of ±0.02 mag. It justifies the choice of building a model for each run when possible, but these variations are small enough that using the FoV model of adjoining runs for those with a small number of stars is still a reasonable choice and much better than not applying a FoV correction.

3.2.3. Dealing with a sharp feature in runs 19Am05, 19Am06 and 19Bm01

Three of the runs – 19Am05, 19Am06, and 19Bm01 – show a problematic sharp feature in the map of residuals, as shown in the left-hand panel of Figure 4. The CFHT staff tracked this feature to the presence, during the summer of 2019, of a hair on one of the MegaCam lenses. They removed it after run 19Bm01 but the direct consequence is that all the photometric catalogues observed during a period of about 3 months suffer from this additional source of localized absorption. It causes the $CaHK$ values to get fainter by up to ∼0.15 mag. `PhotCalib` under-corrects the magnitudes at the location of the feature and over-corrects them around it as the jump in the data are locally so sharp that the gradient descent method fails when reaching the boundaries of the feature. Fortunately, two of these three runs (19Am06, 19Bm01) are the largest runs in the survey and contain hundreds of images (300 000–400 000 good quality stars for each run). We are therefore in the fortunate position to be able to build a refined version of the calibration for this specific situation.

We first use the FoV model from the closest large run (19Am03) to determine temporary zero-point offsets for all the images in run 19Am06 (Step 0, left-hand panel of Figure 4). We then determine the boundaries of the region of the feature (region A) by grouping together neighboring pixels that show median $\Delta CaHK$ deviations larger than the 97.5th percentile of the entire distribution. The pixels outside region A define region B. We then apply `PhotCalib` only to stars in region B, which yields the final zero-point offsets as well as a reliable $FOV_{\rm 19Am06}$ model in regions not affected by the feature (Step 1; central panel of Figure 4). After Step 1, the $zp(i)$ values are frozen for all the stars in this run and, to learn the behavior of the feature, we set a different neural network containing one convolutional and one linear layer, with each accompanied by an activation and a normalization layer like for the regular model. We apply this new neural network only to region B with

[13] Applying a more specific variability cut that relies on the $P_{\rm var}$ quantity defined later, in Section 5, only mildly impacts the calibration, with differences of less than 3 mmag on the $zp(i)$ values and significantly less than the systematic floor of 13 mmag that results from the new calibration process (see Section 3.3).

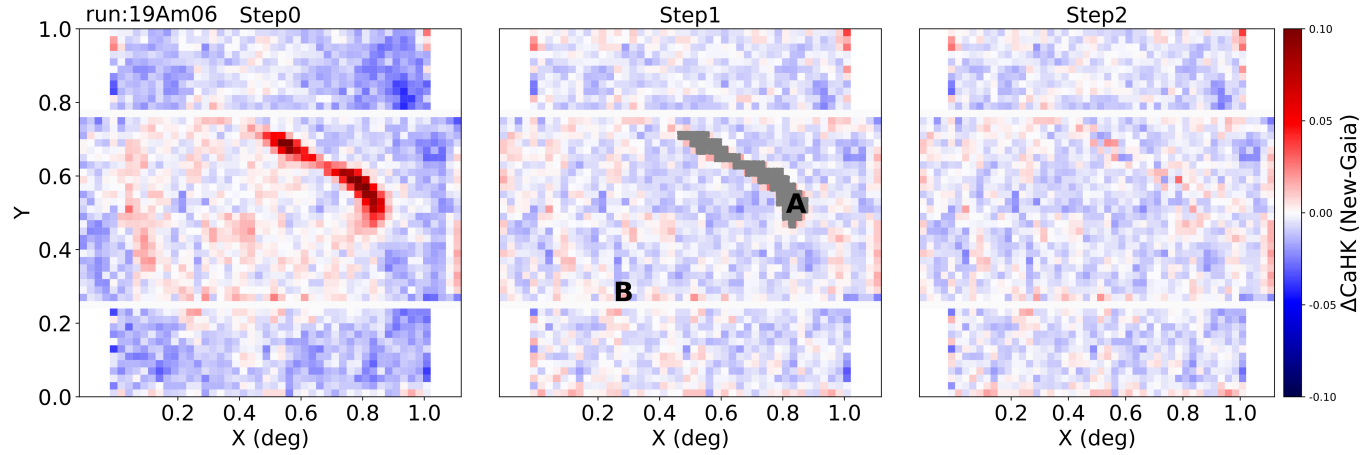

Fig. 4. Median residuals per pixel over the field of view between the calibrated and the synthetic *Gaia* values ($CaHK_{\rm calib} - CaHK_{\rm syn}$) for run 19Am06 that is affected by the sharp feature. Step 0: results after applying only the regular model. Step 1: the masked region of the feature is defined as region A (gray) and the rest is region B, which has been corrected by applying the regular model only on region B. Step 2: after applying the modified model on region A, the sharp feature is removed. The corrections in region B are the same as in Step 1.

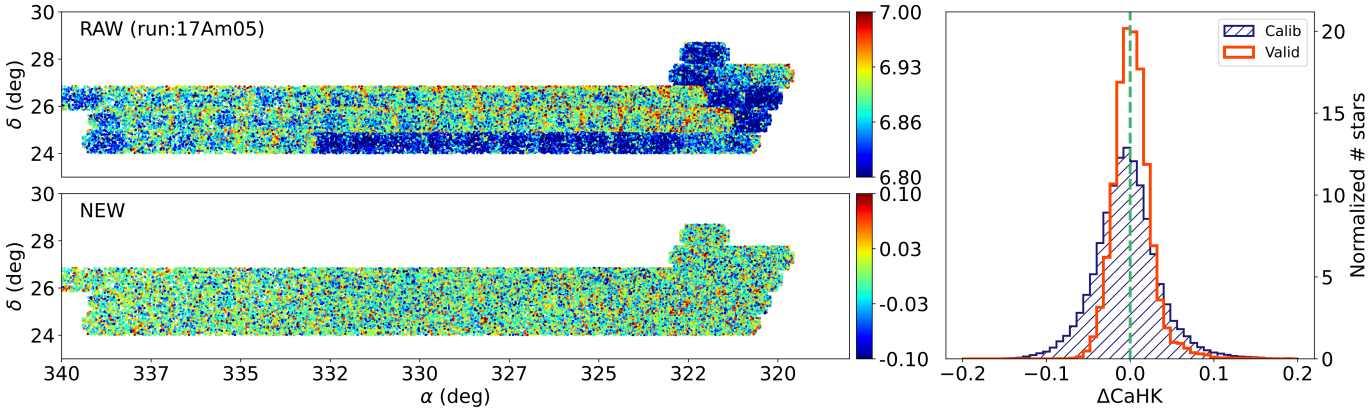

Fig. 5. Effect of the calibration procedure on the *CaHK* magnitudes. *Left*: maps of the *CaHK* offsets between the uncalibrated (top) and calibrated (bottom) Pristine catalogues and the *Gaia* $CaHK_{\rm syn}$ magnitudes (the color codes $CaHK_{\rm uncalib} - CaHK_{\rm syn}$ and $CaHK_{\rm calib} - CaHK_{\rm syn}$, respectively) for run 17Am05. *Right*: histogram of a random sampling of the residual ($\Delta CaHK = CaHK_{\rm calib} - CaHK_{\rm syn}$) for 1% of stars in the calibration ($\delta CaHK < 0.05$; blue) and validation ($\delta CaHK < 0.015$; orange) samples.

fixed $zp(i)$ values. We then simply tile together the models from the two regions in Step 2 (right-hand panel of Figure 4), which successfully removes the feature while, at the same time, leaves its surroundings un-affected. We follow a similar approach to calibrate 19Bm01, with the only difference being that the FoV model used in Step 0 is the model from region B of 19Am06. The third run affected by this feature is 19Am05, which contains only 5 images and 1142 good-quality stars. For this run, we apply the recipe we designed above to deal with small runs and use the FoV model from its large neighboring run, 19Am06.

3.2.4. Final results

The left-hand panels of Figure 5 give a visual representation of the success of the calibration by showing a map of the differences between the Pristine *CaHK* magnitudes and the *Gaia* synthetic $CaHK_{\rm syn}$ magnitudes before and after calibration for stars of good-quality synthetic data ($\delta CaHK_{\rm syn} < 0.05$). The chosen 17Am05 run combines data that are affected by different amounts of absorption from the observing conditions, which produce the obvious field-to-field offsets in the top-left panel. The magnitude variations with the location on the FoV are also visible. After calibration, both of these effects are handled and the data are very flat. This is summarized for the full data set in the right-hand panel: the blue histogram represents a random sampling of 1% of the calibration sample with $\delta CaHK_{\rm syn} < 0.05$ over the full data set, and the orange histogram shows a random sample of 1% of the validation sample ($\delta CaHK_{\rm syn} < 0.015$), still over the full data set. After calibration, the mean difference is 0.006 mag and the distribution, which looks Gaussian-like, has a dispersion that is barely larger than the *Gaia* photometric-uncertainty cut used to build the sample (∼18 mmag).

3.3. The Pristine CaHK DR1

3.3.1. Merging of repeat individual detections

While the Pristine survey was designed as a single-exposure survey, there is nevertheless a significant fraction of survey stars that are observed multiple time, either from the repeat observations of low-quality fields that could be salvaged, from the 2′ overlap between the central, square portion of neighboring fields, or from the 4 side CCDs (the "ears") that were not assumed to be available when the survey tiling was constructed. In total

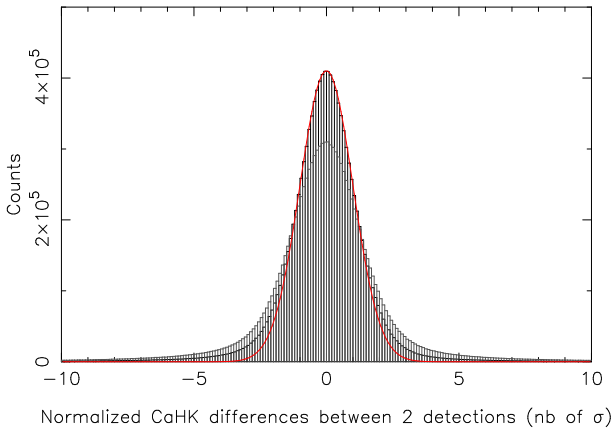

Fig. 6. Distribution of normalized magnitude differences between repeat observations in Pristine. The gray data corresponds to the reduced, calibrated data, while the black histogram also includes a systematics uncertainty floor of 13 mmag, added in quadrature to the photometric uncertainties coming out of the pipeline. The black histogram is a good approximation of the expected distribution (red line) but shows a small fraction of objects in the tails of the distribution produced by incompatible repeat observations (variable stars, artifacts, etc.). These represent ∼7.5% of repeats.

∼20% of sources in the detection catalogue correspond to repeat observations. To increase the signal-to-noise of those objects, we merge their individual detections using a 0.5″ search radius. The magnitudes of the individual detections are combined into the magnitude and corresponding uncertainty of the merged detection using an arithmetic mean weighted by the uncertainties of the individual measurements.

Figure 6 shows the resulting magnitude differences between repeat measurements, normalized by their uncertainties (gray histogram). Perfect data would result in a Gaussian centered on zero, with width unity (red line). It is not the case here, which shows that there remains some level of systematic uncertainties in the photometric calibration, as is expected in any data set. The black histogram, whose central distribution is a good match to the red model has a systematic uncertainty of 0.013 mag added in quadrature to the photometric uncertainties of individual detections. This value is entirely compatible with the width of the orange histogram in Figure 5 and allows us to conclude that the Pristine magnitudes are calibrated at the 13 mmag level. The tails of the black histogram in Figure 6 correspond to variable stars or stars with problematic photometry. In the catalogue of merged detections, we flag with `merged_CASU_flag = +2` merged detections whose `CASU_flag` is not consistent between the different individual detections (∼20% of merges). The `merged_CASU_flag` otherwise duplicates the consistent value of `CASU_flag` from the different individual detections. Finally, we flag with `merged_CASU_flag = +3` merged objects that have two individual detections that are incompatible at more than the 5σ level (2.9% of repeats for the full Pristine survey, including faint sources).

3.3.2. Catalogue

With this paper, we make public the part of the Pristine data that overlaps with the $CaHK_{\rm syn}$ catalogue; that is, we provide the Pristine information on the merged detections for all 6 311 676 stars with *Gaia* DR3 BP/RP coefficients that overlap the Pristine footprint. The density of stars in this catalogue is shown in the right-hand panel of Figure 2. When available, the Pristine photometric information for these sources is added to the photometric catalogue described earlier in Table 1.

4. Extinction correction

The *Gaia* broad-band filters that we use in our photometric metallicity model (see Section 6) are so broad that the extinction correction in a given filter depends on the underlying properties of the star ($T_{\rm eff}$, $\log g$, [Fe/H]), as well as the extinction toward that star. The *Gaia* team provides a relation to derive the extinction coefficients as a function of $T_{\rm eff}$ (or $G_{\rm BP} - G_{\rm RP}$) and A_0 (the monochromatic extinction at 541.4 nm)[14]. We follow a similar methodology as described there to re-derive this relation including a dependence on [Fe/H], using the `dustapprox` package (Fouesneau et al. 2022) with Kurucz synthetic spectra (Castelli & Kurucz 2003), adopting the Fitzpatrick (1999) extinction law with $R_V = 3.1$, the Riello et al. (2021) *Gaia* passbands and the MegaCam CaHK passband.

The synthetic grid covers $4000 \leq T_{\rm eff} \leq 10\,000$ K with step of 250 K, $0.0 \leq \log g \leq 5.0$ with steps of 0.5, and $-2.5 \leq$ [Fe/H] $\leq +0.5$ with steps of 0.5 dex, with the addition of spectra at [Fe/H] = +0.2 and [Fe/H] = −4.0. We adopt [α/Fe] = +0.4 for [Fe/H] ≤ -1.0 and [α/Fe] = 0.0 for [Fe/H] ≥ -0.5. We fit separate relations for giants (all [Fe/H], $4000 \leq T_{\rm eff} \leq 7500$ K, $\log g = 4.0$ for $T_{\rm eff} > 5250$ K and $\log g = -8.3000 + 0.0023\,T_{\rm eff}$ for cooler stars) and dwarfs (all $T_{\rm eff}$, $\log g = 4.5$ and all [Fe/H] for $T_{\rm eff} \leq 7500$ K but only [Fe/H] ≥ -0.5 for hotter stars).

We compute synthetic photometry for $0.01 \leq A_0 \leq 2.99$ with steps of 0.02 and we then fit polynomial relations of the following form to the synthetic photometry from `dustapprox`:

$$\begin{aligned} k_f = {} & a_0 + a_1 T + a_2 A_0 + a_3 [\mathrm{Fe/H}] + a_4 T^2 + a_5 A_0^2 + a_6 [\mathrm{Fe/H}]^2 \\ & + a_7 T^3 + a_8 A_0^3 + a_9 [\mathrm{Fe/H}]^3 + a_{10} T A_0 + a_{11} T [\mathrm{Fe/H}] + a_{12} A_0 [\mathrm{Fe/H}] \\ & + a_{13} A_0 T^2 + a_{14} T A_0^2 + a_{15} [\mathrm{Fe/H}] T^2 + a_{16} T [\mathrm{Fe/H}]^2 + a_{17} [\mathrm{Fe/H}] A_0^2 \\ & + a_{18} A_0 [\mathrm{Fe/H}]^2 + a_{19} T A_0 [\mathrm{Fe/H}], \end{aligned} \quad (2)$$

with $T = T_{\rm eff}/5040$ K and k_f the extinction coefficients in a given filter. The best fit polynomial coefficients are listed in Table A.1 (giants) and A.2 (dwarfs) in the Appendices. The giant solution as a function of $T_{\rm eff}$ for different A_0 and [Fe/H] is presented in Figure 7 for the three *Gaia* filters. The sensitivity to the stellar parameters for the narrow CaHK filter is negligible (<0.3%) and not shown in the Figure.

In this work we use the Schlegel et al. (1998, hereafter SFD) extinction map, which provides $E(B-V)$ instead of A_0. We derive A_0/A_V as a function of the stellar parameters following the same methodology as above[15], as shown in the bottom-right panel of Figure 7. Finally, the extinction in filter f is given by $A_f = R_f E(B-V)$ with $R_f = k_f \frac{A_0}{A_V} 3.1$, which is multiplied by an additional factor of 0.86 when using $E(B-V)$ from the SFD map (from the recalibration of the SFD map by Schlafly & Finkbeiner 2011).

We iteratively derive the best extinction coefficients for each individual star from the photometry, together with $T_{\rm eff}$ and [Fe/H]. For $T_{\rm eff}$, we adopt the Casagrande et al. (2021) *Gaia* EDR3 photometric $T_{\rm eff}$ relation that depends on $(G_{BP} - G_{RP})_0$, [Fe/H], and $\log g$ (we fix the latter to 3.0 as it only has a minor

[14] https://www.cosmos.esa.int/web/gaia/edr3-extinction-law

[15] This ratio is not equal to one because A_V is for a broad filter and depends on the stellar parameters, whereas A_0 does not.

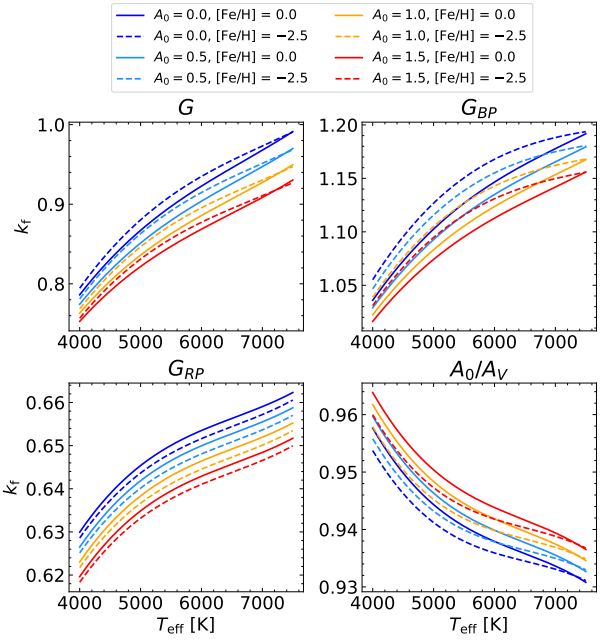

Fig. 7. Extinction coefficients for the three *Gaia* filters and the ratio A_0/A_V as a function of the effective temperature $T_{\rm eff}$, for different values of A_0 and [Fe/H] (see legend).

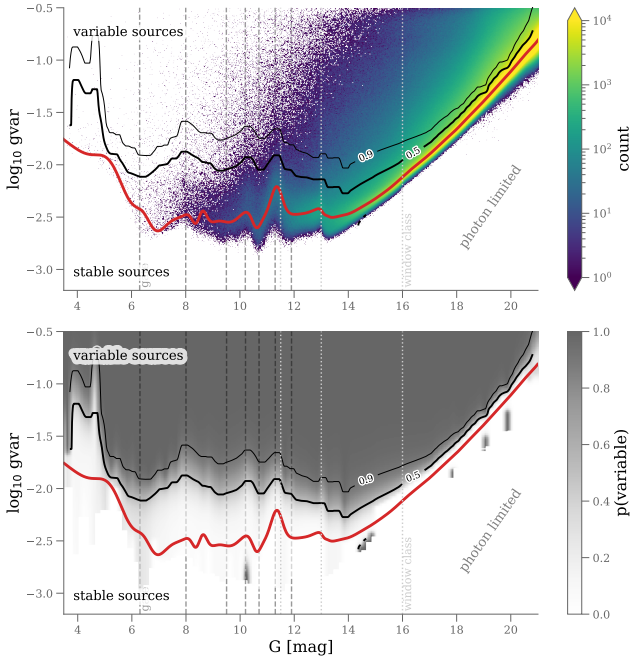

Fig. 8. Variability distribution as a function of the G magnitude for 1 806 185 593 sources from *Gaia* DR3. The red line indicates our B-spline model that tracks the mean location of non-variable sources. We note that for $G < 13$, the photometry is affected by the gating and windowing system onboard *Gaia*. We indicate the approximate G magnitudes of the eleven gates with the dashed gray lines and the three window class changes resulting in overall offsets in the photometric calibration (details in Riello et al. 2021). Between these positions, the configuration affects the photometry in various manners, generating apparent wiggles and discontinuities. In DR3, the reported uncertainties include an additional calibration error on a single measurement of 2 mmag corresponding to the apparent floor ($\log_{10} g_{\rm var} \sim -2.7$) at the bright end.

effect). An estimate of [Fe/H] is then derived using our photometric metallicity model (see Section 6). If a star falls outside the metallicity grid, we assign it [Fe/H] = −4.0 if the star is above the edge of the metallicity grid and [Fe/H] = 0.0 otherwise. If it falls outside the validity range of the Casagrande et al. (2021) relation (3500–9000 K), $T_{\rm eff}$ is set to the closest valid $T_{\rm eff}$. Initial guesses for $A_{(BP-RP)}$, A_G and A_{CaHK} are $E(B-V)_{SFD} \times 1, 2$ and 4, respectively, and the initial guess for $A_0/A_V = 1$. We find that the extinction coefficients converge quickly and that five iterations are sufficient.

For a typical halo giant star ($T_{\rm eff}$ = 5000 K, [Fe/H] = −1.0, $E(B-V) = 0.05$), we find $R_{CaHK} = 3.918$ for the SFD map, which is very similar to the $CaHK$ extinction coefficient of 3.924 that our team has been using previously (Starkenburg et al. 2017a).

5. Probabilistic variability model

Variable stars are one of the main sources of contamination in low-metallicity catalogues based on photometry (Starkenburg et al. 2017a; Lombardo et al. 2023). Previously, we relied on the repeat observations of sources in Pan-STARRS 1 to quantify the variability of a source (Hernitschek et al. 2016). As we strive to rely solely on Pristine and *Gaia* information to characterize observed sources over the full sky coverage, we have moved toward using a single *Gaia*-based diagnostic of variability (Fernández-Alvar et al. 2021; Lombardo et al. 2023), which we update here with a probabilistic approach.

Gaia observes repeatedly 1.8 billion sources quasi-simultaneously in G, G_{BP}, and G_{RP} integrated photometric bands and the *Gaia* DR3 contains the mean photometry for these sources based on 34 months of multi-epoch observations. The repetition of the measurements provides precise mean photometric values but, additionally, the associated uncertainties can tell us about the photometric variations of a source (see also Belokurov et al. 2017). Assuming Poisson variation of the measurements, we can define a variability metric as the fractional flux uncertainty scaled by the square root of the number of observations. For the G band, we defined this variability metric by

$$\mathrm{gvar} = \sqrt{N_{obs,G}} \cdot \frac{\sigma_G}{f_G}, \tag{3}$$

where $N_{obs,G}$ is the number of observations contributing to the photometric value, and $\frac{\sigma_G}{f_G}$ is the fractional flux uncertainty. This quantity corresponds to `sqrt(phot_g_n_obs)/phot_g_mean_flux_over_error` in the `gaia_source` *Gaia* tables[16]. The top panel of Fig. 8 shows the complex distribution of this metric as a function of apparent G-magnitude. The majority of sources at any given magnitude resides close to the photon-noise limit of the instrument, i.e., small values of `gvar`. Complexity arises at the bright end, where the gating and windowing system on board *Gaia* affects the photometry in various manners generating apparent wiggles and discontinuities in that distribution.

Defining which *Gaia* sources are intrinsically variable ones requires finding the photon-limited sequence[17] and modeling the dispersion around it. We adapt the B-Spline approach of Riello et al. (2021) to derive a mean uncertainty model (see the

[16] https://gea.esac.esa.int/archive/

[17] Technically, systematics and instrumental effects dominate the noise budget of sources at the bright end.

Gaia science performance pages[18]). We set the spline knots such that we capture the main features (gates) and the sharp transitions due to window class changes (e.g., $G = 13.0$). We assume that the spline location identifies the "stable" sources and that they follow an almost Gaussian distribution around it (law of large numbers). In contrast, we expect the "variable" sources (strictly, the non-stable sources) to be rare and thus distribute more closely to a Poisson distribution toward larger values of `gvar`. In practice, we model the distributions at a given G-magnitude by the mixture of two t-distributions using the spline as a prior on the location of the peaks. A t-distribution is a continuous equivalent to a Poisson distribution, which encompasses a Gaussian distribution. The mixture ratio, $P_{\rm var}$, defines the probability of a source to be variable. The lower panel of Figure 8 shows our model predictions as a function of `gvar` and G-magnitude. Because we assume the distributions at different G magnitudes are independent, our model can show artifacts at the lower bound of the data. However, it captures the complexity of the distribution, especially for magnitudes brighter than $G = 15.0$.

We report $P_{\rm var}$ values in the $CaHK_{\rm syn}$ and Pristine DR1 photometric catalogues of Table 1. An aggressive cut on the probability of being variable, $P_{\rm var} < 0.3$, can be used to produce a sample of stars whose *Gaia* magnitudes are unlikely to be affected by variability. This cut flags out 20% of the 219.2 million stars with BP/RP information.

6. The Pristine ($CaHK$,G,$G_{\rm BP}$,$G_{\rm RP}$) → [Fe/H]$_{\rm phot}$ model

We use both the Pristine $CaHK$ magnitudes as well as the *Gaia*-based synthetic $CaHK_{\rm syn}$ magnitudes to determine photometric metallicity estimates, [Fe/H]$_{\rm phot}$. The method used here is in essence very similar to the method described in Starkenburg et al. (2017a), but has been updated and adapted to use *Gaia* broadband photometry (G, $G_{\rm BP}$, $G_{\rm RP}$) rather than SDSS broadband photometry.

6.1. Spectroscopic training sample

Our method is limited to FGK stars as, for hotter stars, the strength of the Ca H & K absorption lines is too weak to be a reliable estimator of metallicity. At the other end of the temperature range, very cool M stars contain very prominent molecular bands that significantly drop the level of the pseudo-continuum in the relevant wavelength regime, making the measurement of the Ca H & K absorption features challenging. We therefore restrain our analysis to the range $0.5 < (G_{\rm BP} - G_{\rm RP})_0 < 1.5$, which approximately covers evolutionary stages between the upper main sequence, the turn-off, and the tip of the red giant branch for an old, (very) metal-poor stellar population. This color interval corresponds to a temperature range of approximately $3900 < T_{\rm eff} < 7000$ K.

Following Starkenburg et al. (2017a), we use a training sample to provide a mapping from the de-reddened ($CaHK$, G, $G_{\rm BP}$, $G_{\rm RP}$) color space onto [Fe/H]$_{\rm phot}$. A first, major, constituent of this sample are the SDSS/SEGUE stars (Smee et al. 2013; Yanny et al. 2009) that are in the Pristine footprint. We restrict ourselves to those spectra with an average signal-to-noise ratio per pixel larger than 25 over the wavelength range 400–800 nm. We further require that the SDSS pipeline provides a value for $\log g$, an adopted $T_{\rm eff} < 7000$ K, a radial velocity with an uncertainty $<10\,{\rm km\,s^{-1}}$, and an adopted spectroscopic metallicity, [Fe/H]$_{\rm adop}$, with an uncertainty <0.2 dex. We also limit the sample to stars with nominal `n` flags, except for stars that also show the `g'` or `G` flag indicating a G-band feature. For outer halo science cases we are particularly interested in red giant stars and because these are less numerous in the sample, we complement our SDSS sample with APOGEE giants, stars that have APOGEE derived $\log g < 3.9$ and $T_{\rm eff} < 5800$ K (Majewski et al. 2017; Wilson et al. 2019; Abdurro'uf et al. 2022). The added APOGEE stars are furthermore restricted to spectra with a signal-to-noise ratio larger than 50, `FE_H_FLAG = 0`, and no `STARFLAG` raised. The APOGEE [Fe/H] values are shifted down by 0.1 dex to match the SDSS/SEGUE scale we have adopted. On the Pristine side, we restrict the sample to high-quality data: we use `CASU_flag = -1` and $CaHK$ photometric uncertainties $\delta CaHK < 0.05$.

As mentioned earlier, the Pristine footprint has grown enormously since the publication of Starkenburg et al. (2017a), from about $1000\,{\rm deg^2}$ to $\sim 6500\,{\rm deg^2}$ and, naturally, this commensurately enlarges the training sample, with a total of now over 66 000 stars that also have a *Gaia* DR3 counterpart. Nevertheless, the range of EMP and UMP stars is not well sampled as these stars are exceedingly rare (see, e.g., Youakim et al. 2017; Bonifacio et al. 2021; Yong et al. 2021). We attempt to mitigate the sparsity of this important region of parameter space in two ways. First, we complement the SDSS sample with known VMP, EMP, and UMP stars. We add overlapping high-resolution observations of stars in the Boötes I dwarf galaxy (Feltzing et al. 2009; Gilmore et al. 2013; Ishigaki et al. 2014; Frebel et al. 2016), as well as VMP stars from the third data release of the LAMOST survey, taken from Li et al. (2018). For this latter sample, we correct for spurious stars in the low effective temperature range, as described in Sestito et al. (2020), but then use the most recent LAMOST DR8 catalogue. We add stars from the PASTEL sample (Soubiran et al. 2016), as also used by Huang et al. (2022). Finally, we add targets followed up from the Pristine survey and confirmed to be VMP, EMP, or UMP stars (Youakim et al. 2017; Aguado et al. 2019; Venn et al. 2020; Kielty et al. 2021; Lardo et al. 2021; Lucchesi et al. 2022). We note that we use the same training sample, with $CaHK$ magnitudes derived from the Pristine survey rather than the synthetic $CaHK_{\rm syn}$, for the training of both catalogues presented in this paper. The distribution of these stars in the Kiel diagram is shown in the left-hand panel of Figure 9.

6.2. The Pristine color–color space

To further sample the extremely low-metallicity regime, we calculate a grid of stellar spectra with [Fe/H] = −3.0 and [α/Fe] = +0.4 using MARCS (Model Atmospheres in Radiative and Convective Scheme) stellar atmospheres and the Turbospectrum code (Alvarez & Plez 1998; Gustafsson et al. 2008; Plez 2008). To derive magnitudes, these spectra are integrated under the *Gaia* EDR3 filter curves and that of the CaHK MegaCam filter. We also add a synthetic spectral grid calculated with the lowest metallicity atmospheres publicly available from MARCS and no metals at all in the step to create the spectra using Turbospectrum.

The color–color space used to determine [Fe/H]$_{\rm phot}$ combines *Gaia* $(G_{\rm BP} - G_{\rm RP})_0$ information as a temperature proxy and a combination of the *Gaia* broad-band filters and the Pristine narrow-band $CaHK$, as illustrated in the two right-most panels of Figure 9. From left to right, these include dwarf ($\log g > 3.9$ or $T_{\rm eff} > 5800$ K) and giant stars ($\log g < 3.9$), respectively, and the

[18] https://www.cosmos.esa.int/web/gaia/science-performance#photometric%20performance

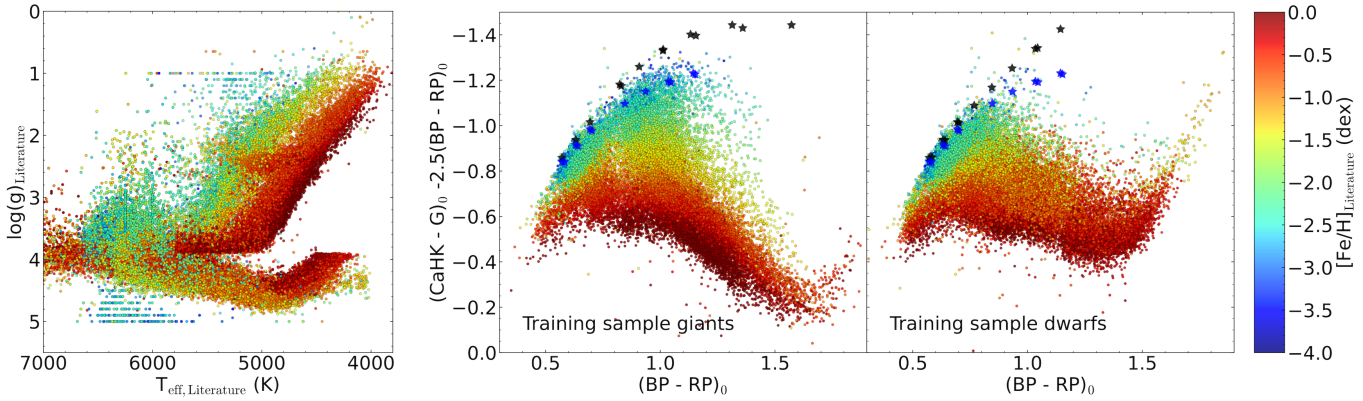

Fig. 9. Temperatures and surface gravities from our training sample (left) and their behavior in the Pristine color–color space for giants ($\log g < 3.9$; middle panel) and dwarf stars ($\log g > 3.9$ or $T_{\rm eff} > 5,800$ K; right-hand panel). The synthetic color resulting from synthetic spectra for stars with [Fe/H] = −3.0 (blue) and (hypothetical) stars with no heavy metals (black) are overplotted with larger star symbols. Stars are plotted from the most metal-rich to the most metal-poor, so as to emphasize the metal-poor region. The sharp edge at $\log g \sim 3.9$ in the left panel is due to the APOGEE sample that is limited to giants only (see text). The very low-metallicity stars from Aguado et al. (2019) have less physical $\log g$ determinations, showing up as horizontal lines in this panel. However, this does not significantly affect their metallicity, or location in the color–color space (see Aguado et al. 2019, for a detailed discussion). The magnitudes shown here are de-reddened following the procedure described in Section 4.

two panels show both the training sample of observed stars (dots) and the synthetic results mentioned above (star symbols). Here, the magnitudes of stars are extinction corrected as described in Section 4 by assuming the spectroscopic parameters of the stars. As expected, the stars arrange themselves in an orderly fashion so that, at fixed color, more metal-poor stars are higher up along the y axis of the figure as the *CaHK* magnitude becomes brighter with the lower absorption of the smaller Ca H & K spectral features of a more metal-poor star.

6.3. The algorithm

The algorithm that builds the model linking a location in color–color space to a metallicity value is described in detail in Starkenburg et al. (2017a) and Fernández-Alvar et al. (2021) and is available online[19]. In short, the color–color space is divided into a grid and, for each cell of the grid, the mean metallicity of the training sample is determined. The main difficulty in this determination is that the smaller number of metal-poor stars compared to more metal-rich stars means that any scattering from measurement uncertainties always scatters more metal-rich stars into the metal-poor regime rather than the other way around, biasing the metallicity upwards. We aim to mitigate this effect by sigma-clipping outliers beyond 2σ from the mean metallicity of a given pixel and, additionally, by imposing that, at a given $(G_{\rm BP} - G_{\rm RP})_0$ (the x-axis in Figure 9), the metallicity decreases monotonically with decreasing $[(CaHK - G)_0 - 2.5(G_{\rm BP} - G_{\rm RP})_0]$ (the y axis in Figure 9) for [Fe/H] < −1.0. In other words, stars that show less *CaHK* absorption for similar broadband colors are only allowed to get a more metal-poor value. The synthetic colors calculated from the modeled spectra of EMP stars are used to anchor the model in the lowest metallicity regime. Grid points closer to the "no-metals" line than to the synthetic [Fe/H] = −3.0 models are set to $[{\rm Fe/H}]_{\rm phot} = -4.0$. In addition, we account for possible systematics in the color of these spectra by setting cells to [Fe/H] = −4.0 down to 0.2 magnitudes below the "no-metals" line (i.e., 0.2 magnitudes above the black symbols in Figure 9)[20]. Finally, the resulting grid is smoothed with a kernel of 2-pixel dispersion to remove outlying grid points caused by a few metal-rich stars scattered into a metal-poor region. We assign no metallicity value to stars outside of the $0.5 < (G_{\rm BP} - G_{\rm RP})_0 < 1.5$ color range, more than 0.2 magnitudes above the "no-metals" line, or more than 0.1 magnitudes below a polynomial fit to the solar metallicity stars (the last two restrictions are near the top and bottom of the color–color plots of Figure 9, respectively).

As mentioned above, and since the extinction correction of the magnitudes used by the model depends on the stellar parameters of a star, including its metallicity, we cannot simply de-redden the magnitudes and push them through the photometric metallicity model as we did before when we relied on SDSS broad-band magnitudes. We therefore determine both the extinction coefficients and the metallicity of the stars iteratively, as mentioned at the end of Section 4.

This framework is used to build both a model for dwarfs ($\log g > 3.9$ or $T_{\rm eff} > 5800$ K) and for giants ($\log g < 3.9$). As the Pristine survey does not provide distance information for the observed stars, one does not a priori know whether a given object is a dwarf or a giant. However, since the Pristine survey and model are mainly focussed on metal-poor stars that often happen to be giants in the MW halo, we favor the giant metallicities as our generic output, $[{\rm Fe/H}]_{\rm phot}$. We nevertheless provide the photometric metallicities derived using the dwarf model, $[{\rm Fe/H}]_{\rm phot,dw}$, which should be preferred for stars that are specifically known to be dwarf stars. When the evolutionary stage of a star is known (for instance via parallax-based distances built from *Gaia*; e.g., Bailer-Jones et al. 2021), one should of course use the most appropriate photometric metallicity values.

For a given star of de-reddened magnitudes ($CaHK_0$, G_0, $G_{\rm BP,0}$, $G_{\rm RP,0}$) and associated Gaussian uncertainties (not including uncertainties in the de-reddening), the probability distribution function (PDF) on any photometric metallicity estimate is determined through 100 Monte Carlo samplings of the magnitude uncertainties. $[{\rm Fe/H}]_{\rm phot}$ and $[{\rm Fe/H}]_{\rm phot,dw}$ are redetermined

[19] https://github.com/ankearentsen/Pristine_code

[20] A clear consequence of this choice is that the model cannot return a metallicity $[{\rm Fe/H}]_{\rm phot} < -4.0$, but this regime is so poorly sampled that any model would only be a wild extrapolation beyond this limit.

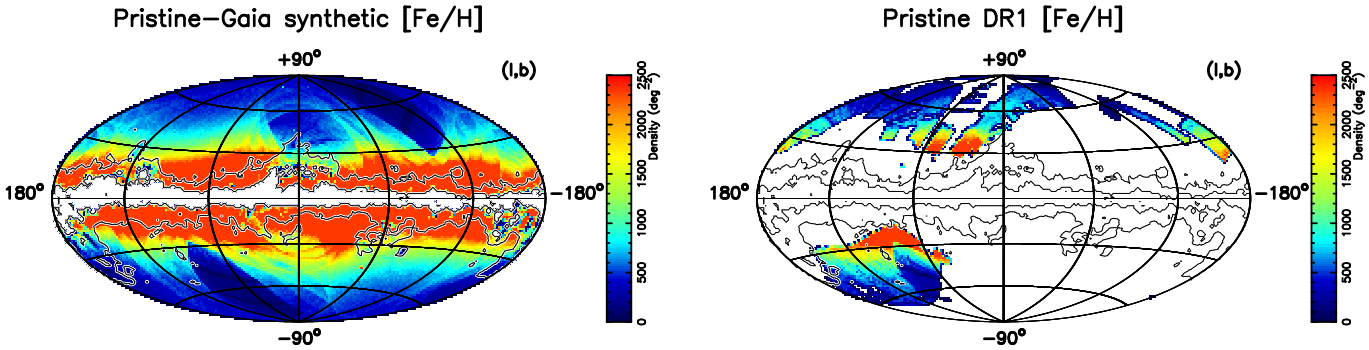

Fig. 10. Density of sources with photometric metallicities in the Pristine-*Gaia* synthetic catalogue (left-hand panel) and the DR1 catalogue (right-hand panel). The black contours are lines with $E(B-V) = 0.15$ and 0.5. By construction, regions with extinction $E(B-V) > 0.5$ are removed from the photometric metallicity catalogues and produce the white regions in the left-hand map. The impact of the *Gaia* scanning law on the S/N of the BP/RP information and, consequently, on that of the $CaHK_{\rm syn}$ magnitudes, is clearly visible; it is responsible for most of the irregular features in this map. The map of Pristine metallicities is more limited on the sky but also denser, owing to the higher S/N Pristine data: some stars with BP/RP information do not have high enough S/N to make it through the enforced $\delta CaHK_{\rm syn} < 0.1$ cut but have a high-enough S/N in Pristine to generate a Pristine metallicity.

for each sampling of the uncertainties. Even though the uncertainties on the magnitudes are assumed to be symmetrical, the resulting PDFs on the photometric metallicities can be very asymmetrical as the narrow-band color does not vary linearly with decreasing metallicity. This can be appreciated in Figure 9, in which two stars with the same metallicity difference are much closer together if they are metal-poor. For this reason, we do not quantify the PDF with a simple dispersion but instead provide the values of the 16th, 50th, and 84th percentiles of a given Monte Carlo PDF. While we usually use the metallicities calculated directly from the favored values of the magnitudes, the median values (50th percentile) are also systematically reported in the catalogues. Additionally, we report `mcfrac`, the fraction of Monte Carlo samples for each star that falls within the limits of the metallicity grid.

6.4. Applicability of the model

Given the main science goals of the Pristine survey, which focus on metal-poor stars, most of the effort on the model is placed in the region $[{\rm Fe/H}]_{\rm phot} < -1.0$ to avoid being driven and dominated by the much more numerous stars above this metallicity cut at the expense of the metal-poor region. A direct consequence is that, while we do provide the photometric metallicities of stars with $[{\rm Fe/H}]_{\rm phot} > -1.0$ and the Pristine model provides reasonable results in this range (see the next section), these metallicities may be biased in unexpected ways (e.g., from the increased impact of the gravity on the depth of the Ca H & K lines, changes in the mean $[\alpha/{\rm Fe}]$ abundance, or non-LTE and 3D effects on the line strengths). Alternative [Fe/H] estimates based on the *Gaia* BP/RP information and trained, in particular, on large spectroscopic surveys (APOGEE, LAMOST) that are vastly dominated by stars with $[{\rm Fe/H}] > -1.0$ may be more appropriate to explore this regime (e.g., Andrae et al. 2023b; Zhang et al. 2023). We stress that our results are in particular optimized to study the low-metallicity regime ($[{\rm Fe/H}] < -1.0$) of FGK stars and that it performs best for stars with $4000 \lesssim T_{\rm eff} \lesssim 6000$ K.

7. The photometric metallicity catalogues

Both photometric catalogues presented in Sections 2 and 3 are pushed through the model described in Section 6 to generate the Pristine-*Gaia* synthetic and the Pristine DR1 photometric metallicity catalogues, respectively. The content of the resulting catalogues are presented in Tables 2 and 3. The sky density of sources in both photometric metallicity catalogues is displayed in Figure 10.

7.1. Building the catalogues

Both datasets are processed in the exact same way and, with Pristine now calibrated onto the $CaHK_{\rm syn}$ photometric system, are expected to yield complementary catalogues of $[{\rm Fe/H}]_{\rm phot}$ that are on the same scale. In addition to the selection in the color–color space of Figure 9 that was discussed in the previous section[21], we further implement two cuts to the photometric catalogues before pushing them through the model. A cut on the $CaHK$ photometric uncertainties, $\delta CaHK_{\rm syn} < 0.1$ or $\delta CaHK < 0.1$ and a cut on the extinction, $E(B-V) < 0.5$, remove noisy photometric measurements and stars that may be plagued by systematics that stem from large extinction corrections. We note that the uncertainties on the photometric metallicities are often asymmetric and larger on the metal-poor side. The catalogues list the photometric metallicities calculated for the favored magnitude values (which we favor), along with the median metallicity obtained from sampling the magnitude uncertainties and the limits of the central 68% confidence interval, $[{\rm Fe/H}]_{\rm phot,16th}$ and $[{\rm Fe/H}]_{\rm phot,84th}$. In what follows, we approximate the photometric metallicity uncertainty, $\delta[{\rm Fe/H}]_{\rm phot} = 0.5([{\rm Fe/H}]_{\rm phot,84th} - [{\rm Fe/H}]_{\rm phot,16th})$, as half the range of metallicity spanned by the central 68% confidence interval.

Figure 11 shows the typical uncertainties attached to the photometric metallicities in the two catalogues. As expected, the metallicities based on the Pristine $CaHK$ data have significantly smaller uncertainties than those determined from the *Gaia*-based $CaHK_{\rm syn}$ magnitudes. The photometric metallicity uncertainties for the latter can be quite significant, especially in the lowest metallicity regimes and/or for bluer stars, or for magnitudes fainter than $G_{\rm BP} \sim 16.0$. The bottom-right panel of the figure, which shows the Pristine photometric metallicity uncertainties as a function of the Pristine photometric metallicities, is quite

[21] These are: $0.5 < (G_{\rm BP} - G_{\rm RP})_0 < 1.5$, $[(CaHK - G)_0 - 2.5(G_{\rm BP} - G_{\rm RP})]_0 < -0.6$, and, at most, 0.2 mag above the "no-metal" line in the figure.

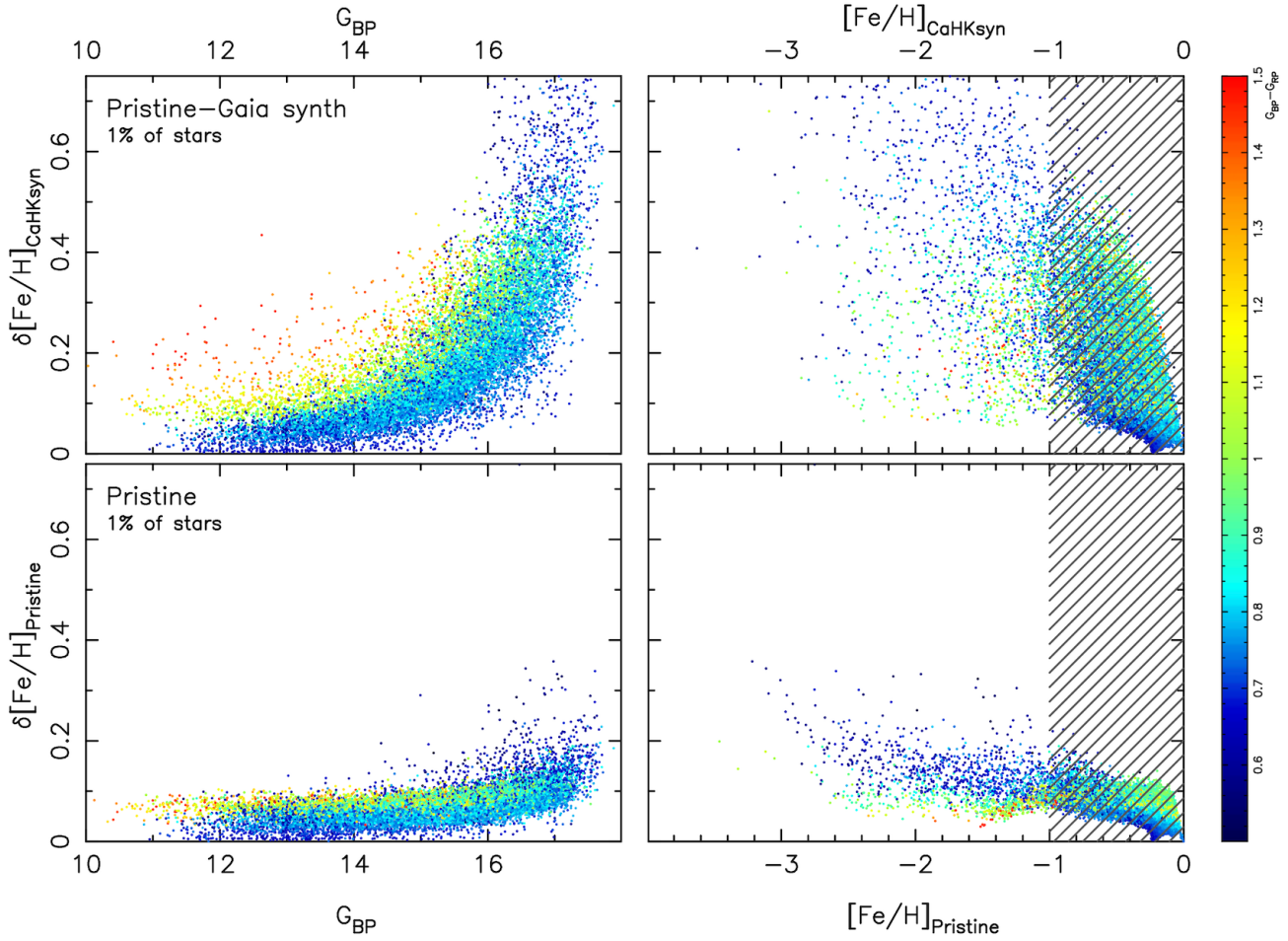

Fig. 11. Mean metallicity uncertainties as a function of the magnitude of a star or its photometric metallicities for a random 1% of stars in the Pristine-*Gaia* synthetic catalogue (top) and the Pristine DR1 catalogue (bottom). Dots are color-coded by the $G_{BP} - G_{RP}$ color of a star. The dashed regions corresponds to the regime for which metallicities might be hampered by systematics due to a stronger impact of the gravity of a star on its derived metallicity (see Section 6). The uncertainties are a complex function of the magnitude of a star, its color, and its metallicity because of the nontrivial mapping of metallicities on the color–color space shown in Figure 9. The metallicities based on Pristine magnitudes show much lower uncertainties owing to their significantly higher S/N.

structured and much more so than its equivalent for the Pristine-*Gaia* synthetic photometric metallicities. This effect is due to the complex translation of magnitude uncertainties into photometric metallicity uncertainties. In the case of the Pristine-*Gaia* synthetic metallicities, the much larger uncertainties on $\delta CaHK_{\rm syn}$ means that these dominate the model grid effects.

In addition to the complex impact of color on the derived photometric metallicities and associated uncertainties (at fixed magnitude, redder stars have a lower S/N in the Ca H & K region but, at the same time, a fixed difference in photometric metallicities corresponds to a larger difference in $CaHK$), these uncertainties also scale with the photometric metallicity since iso-metallicity lines are closer together in the color–color space of Figure 9 for decreasing metallicities. Any sharp quality cut on the photometric metallicity uncertainties, therefore, is more detrimental to more metal-poor stars, at the risk of removing the most metal-poor stars. On the other hand, not including any quality cut on the photometric metallicity uncertainties would include a large number of the much more numerous metal-rich stars, scattered into the realm of the rare stars of lowest metallicities. This effect is of course worse for the lower quality Pristine-*Gaia* synthetic catalogue. Keeping these complications in mind and depending on the science goal, we alternatively use quality cuts of $\delta[{\rm Fe/H}]_{\rm phot} < 0.3$ dex and $\delta[{\rm Fe/H}]_{\rm phot} < 0.5$ dex in what follows.

7.2. A comparison of the two Pristine catalogues

To better understand the limitations of the more noisy Pristine-*Gaia* synthetic catalogue, we can use stars that have both Pristine-*Gaia* synthetic and Pristine metallicities. Figure 12 presents this comparison for a randomly selected subsample of stars in common between the two photometric metallicity catalogues. As expected because the Pristine $CaHK$ magnitudes are now calibrated using the $CaHK_{\rm syn}$ magnitudes, the metallicities calculated from the two sets of magnitudes are consistent with each other and follow the one-to-one line in the left-hand panel of the figure. The uncertainties on the Pristine-*Gaia* synthetic metallicities, even though large, do not show any obvious bias when restricting the sample to stars with $\delta[{\rm Fe/H}]_{\rm phot} < 0.3$ dex (large dots) and $\delta[{\rm Fe/H}]_{\rm phot} < 0.5$ dex (small dots). In this panel, we have applied different quality cuts and we show in the other three panels the stars that were removed by those.

The second panel of the figure shows stars that appear variable according to our model (Section 5; $P_{\rm var} > 0.3$). As expected, variability negatively affects both metallicity catalogues, with a

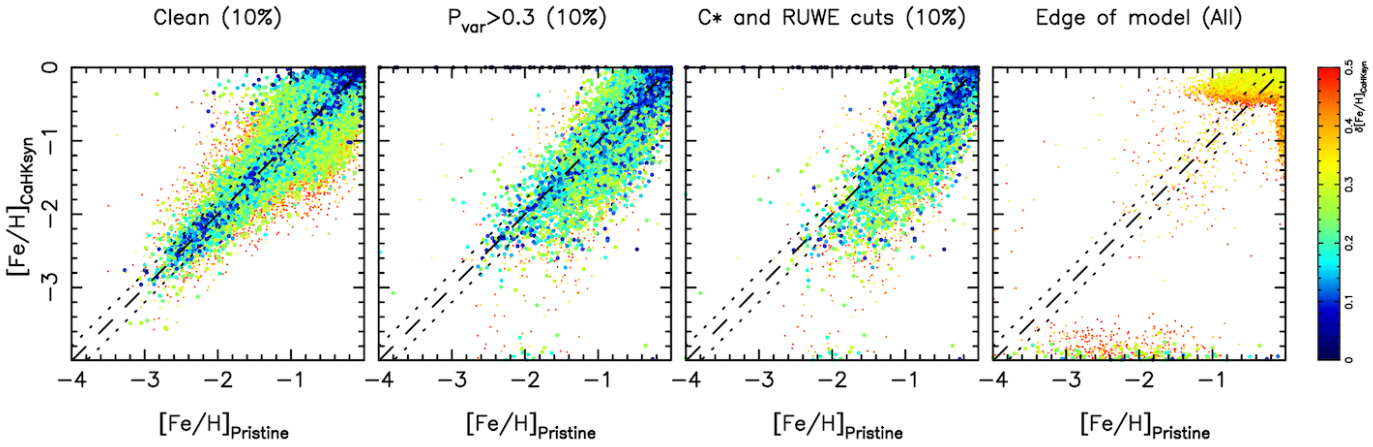

Fig. 12. Comparison of the photometric metallicities in the Pristine-*Gaia* synthetic and the Pristine DR1 catalogues for a random 10% of the overlap between the two catalogues. Points are colored by the uncertainties on the photometric metallicities from the Pristine-*Gaia* synthetic catalogue and small and large dots correspond to stars with a loose and tight quality cut on the metallicity uncertainties, respectively (δ[Fe/H]$_{\rm phot} < 0.5$ and 0.3). The dashed lines correspond to the one-to-one lines and the dotted lines show offsets of 0.2 dex from them. As expected, the metallicities derived from the two sets of magnitudes are consistent since they are on the same photometric system. The next three panels show the impact of quality cuts to clean the sample by removing stars with a probability of being variable ($P_{\rm var} > 0.3$; second panel), flagged as potentially problematic in *Gaia* (RUWE> 1.4 and $|C^*| > \sigma_{C^*}$; third panel), or at the edge of the grid (fourth panel; here we shown all the catalogue stars that follow this criterion). The different quality cuts remove stars that do not behave as expected.

larger distribution than in the first panel. The impact of variability is present in both catalogues, but the effect is likely worse for the Pristine catalogue than for the Pristine-*Gaia* synthetic catalogue because the Pristine *CaHK* photometry of a star was taken at a different time from the broad-band *Gaia* magnitudes. In the past, this has led to contamination of the Pristine EMP-candidate sample when variability information was not available (Lombardo et al. 2023). We strongly encourage applying a similar cut on $P_{\rm var}$ when using the catalogues.

The third panel of Figure 12 shows the impact of cleaning the data with two cuts based on the *Gaia* diagnostics: RUWE > 1.4 and $|C^*| > 1 \times \sigma_{C^*}$, as defined by Riello et al. (2021, equations (6) and (18)). There is a significant overlap between stars removed with these cuts and the variability cut.

Finally, the fourth panel of the Figure shows the impact of removing stars that are at the edge of the metallicity model: [Fe/H]$_{\rm phot} = -4.0$ or 0.0, [Fe/H]$_{\rm phot,16th} = -4.0$ or 0.0, or a large fraction of Monte Carlo samplings are at the edge of the grid (`mcfrac` < 0.8). In particular at the most metal-poor edge of the grid, this removes a small number of objects from the Pristine-*Gaia* synthetic catalogue that have random metallicities compared to the higher S/N Pristine DR1 metallicities. Of course, this latter cut may have the consequence of removing true UMP stars that fall very close to the no-metals line in Figure 9. In the right-most panel of Figure 12, there are clearly a small number of objects removed by this cut that are EMP candidates in the higher quality Pristine catalogue. It should therefore be used with caution when searching for or working on such extreme stars.

7.3. Best practices when using the photometric metallicity catalogues

Given the checks shown in the previous subsection and, in particular, in Figure 12, we recommend a list of quality cuts when using the two catalogues. Except if otherwise specified, these cuts are implemented throughout the rest of the paper:

1. [Fe/H]$_{\rm phot} < 0.0$ and [Fe/H]$_{\rm phot} > -4.0$, as well as a fraction of the Monte Carlo samplings of the magnitude uncertainties that are within the color range of the metallicity model more than 80% of the time (`mcfrac` > 0.8). These cuts remove stars located at the edge of the metallicity grid[22].
2. Photometric uncertainties δ[Fe/H]$_{\rm phot} < 0.5$ dex or, depending on the science case, δ[Fe/H]$_{\rm phot} < 0.3$ dex[23].
3. $P_{\rm var} < 0.3$ to remove stars that are potentially variable[24].
4. *Gaia* quality cuts RUWE < 1.4[25] and $|C^*| < 1 \times \sigma_{C^*}$, with C^* the corrected flux excess and σ_{C^*} the width of its distribution at a given magnitude, as defined in equations (6) and (18) of Riello et al. (2021)[26].
5. `|CASU_flag|=1` or `|CASU_flag|=2` for stars observed by Pristine to ensure that potentially different individual detections of the source did not have conflicting `CASU_flags` when they were merged (see Sub-section 3.3.1).

Finally, we reiterate that most of the effort in the metallicity model was placed in the low metallicity regime ([Fe/H]$_{\rm phot} < -1.0$). While we show below that the photometric metallicities above this limit still behave well, studies specifically focussing on [Fe/H]$_{\rm phot} > -1.0$ may wish to use other catalogues that are more specifically designed for this regime.

[22] However, bear in mind that this restriction (as well as the [Fe/H]$_{\rm phot} > -4.0$ restriction above) may in fact remove some true UMP stars that would be located near the "no-metals" line in the color–color space, or scattered upward of it by noise in the photometry.
[23] We note, however, that the confidence intervals are in fact asymmetrical and that, in some cases, this cut may be too crude.
[24] This cut could be loosened at the bright end, for which the S/N of the *Gaia* magnitudes is so high that even tiny variations of the magnitude of a star would lead to a high $P_{\rm var}$ while not significantly affecting [Fe/H]$_{\rm phot}$.
[25] Yuan et al. (2015) has shown that unresolved binaries can lead to stars in binary systems being assigned a more metal-poor photometric metallicity than their true metallicity. Xu et al. (2022) recommends using the stricter RUWE < 1.1 for analyses where this is expected to be an issue.
[26] This C^* restriction as stated here is quite strict and likely removes some good stars. For a more inclusive selection, one might opt for 2 or $3\sigma_{C^*}$ rather than $1\sigma_{C^*}$.

Table 4. Numbers of candidate stars below a given metallicity limit in the two catalogues.

Quality cut	Number of stars	$[Fe/H]_{phot} < -1.0$	$[Fe/H]_{phot} < -2.0$	$[Fe/H]_{phot} < -3.0$
Pristine-*Gaia* synthetic catalogue				
$E(B-V) < 0.5$ & $\delta[Fe/H]_{phot} < 0.5$	31 605 960	1 927 650	177 488	6331
$E(B-V) < 0.3$ & $\delta[Fe/H]_{phot} < 0.5$	27 882 616	1 774 800	166 143	5770
$E(B-V) < 0.3$ & $\delta[Fe/H]_{phot} < 0.3$	23 254 626	829 320	84 384	1585
Pristine DR1 catalogue				
$E(B-V) < 0.5$ & $\delta[Fe/H]_{phot} < 0.5$	3 812 595	480 602	57 786	1801
$E(B-V) < 0.3$ & $\delta[Fe/H]_{phot} < 0.5$	3 810 843	480 263	57 752	1798
$E(B-V) < 0.3$ & $\delta[Fe/H]_{phot} < 0.3$	3 783 243	469 066	53 636	795

Notes. All quality cuts listed in Section 7.3 are applied to the catalogues to yield these numbers. Additional cuts on the $E(B-V)$ extinction values and the photometric uncertainties $\delta[Fe/H]_{phot}$ are listed in the first column.

Applying these various quality cuts results in 31 605 960 stars with reliable photometric metallicities in the Pristine-*Gaia* synthetic catalogue and 1 927 650 of these correspond to metal-poor star candidates ($[Fe/H]_{phot} < -1.0$). Adopting various cuts on extinction ($E(B-V) < 0.5$ or, more strictly, <0.3) and $\delta[Fe/H]_{phot}$ (<0.5 or, more robustly, <0.3) yields the numbers listed in Table 4. In particular, restricting to a high-quality sample, the catalogue lists 84 384 VMP star candidates and 1585 EMP star candidates. From our own follow-up of Pristine selected EMP star candidates, we expect that most of those stars truly have $[Fe/H] < -2.5$ and that >20% of EMP star candidates indeed have $[Fe/H] < -3.0$ when followed up with spectroscopy (Youakim et al. 2017; Aguado et al. 2019 and see also Caffau et al. 2017, 2020; Venn et al. 2020; Kielty et al. 2021; Lardo et al. 2021; Lucchesi et al. 2022; Sestito et al. 2023), although it is important to remove the contamination from potentially variable stars (Bonifacio et al. 2019; Lombardo et al. 2023). The Pristine DR1 catalogue is overall smaller, with 3 783 243 good-quality photometric metallicities, among which 469 066 metal-poor star candidates, 53 636 VMP star candidates, and 795 EMP star candidates, but contains many fainter targets.

7.4. Internal and external comparisons with metallicity catalogues

7.4.1. Internal comparisons

To test the quality of the resulting photometric metallicities, Figure 13 presents a comparison of the calculated photometric metallicities with the spectroscopic metallicities of stars from the training set that was used to build the $(CaHK, G, G_{BP}, G_{RP}) \rightarrow [Fe/H]_{phot}$ model. Beyond the cleaning of the training sample already mentioned in Section 6, we apply the additional quality criteria of Section 7.3 with the exception that we place not cut on the $\delta[Fe/H]_{phot}$ uncertainty. We restrict this analysis to giant stars (with $\log g < 3.9$) that are cooler than 6000 K. We also leave out the sample from Li et al. (2018) from this specific comparison, because that particular data set shows significantly more scatter. By construction, there should be good agreement between the photometric and spectroscopic metallicities for this sample and it is indeed the case. The left-hand panels of the figure show that, at the bright end ($G_{BP} < 16.0$, where the uncertainties on the $CaHK_{syn}$ uncertainties and photometric metallicity uncertainties remain reasonably small; cf. Figures 1 and 11), both input $CaHK$ catalogues yield very similar results. The histograms in the right-hand panels show the distribution of the differences between the photometric and spectroscopic metallicities. Their confidence intervals, bounded by their 16th and 84th percentiles (vertical lines) correspond to average uncertainties of 0.11 and 0.15 dex, even when including the regime with $[Fe/H]_{phot} > -1.0$, for which the model is not specifically designed. For fainter magnitudes, however, the *Gaia*-Pristine synthetic metallicities become significantly less accurate than their Pristine counterpart, with uncertainties that reach 0.33 dex (vs. 0.14 dex), as expected from the higher S/N of the Pristine photometry. The medians of the distributions of metallicity differences inform us on the photometric metallicity bias, which remains small in all cases (<0.05 dex).

Since the start of the Pristine survey 8 years ago, the team has been actively pursuing spectroscopic follow-up of promising (very) metal-poor targets (Youakim et al. 2017; Caffau et al. 2017; Aguado et al. 2019; Bonifacio et al. 2019; Caffau et al. 2020; Venn et al. 2020; Kielty et al. 2021; Lardo et al. 2021; Lucchesi et al. 2022; Martin et al. 2022; Yuan et al. 2022; Sestito et al. 2023; Lombardo et al. 2023). Figure 14 focusses specifically on a comparison with these follow-up efforts in high-resolution as presented in Caffau et al. (2017); Bonifacio et al. (2019); Venn et al. (2020); Kielty et al. (2021); Lardo et al. (2021); Lucchesi et al. (2022); Martin et al. (2022); Yuan et al. (2022); Lombardo et al. (2023); Sestito et al. (2023). Since many of these stars also feature in our training sample, this Figure does not present an independent validation of our photometric metallicities, but it is insightful to see how the photometric metallicities compare to these high-resolution results. While some of the samples followed up suffered from contamination of higher metallicity sources due to subtle saturation effects in the SDSS photometry used at the time, or variable stars (see, in particular, Caffau et al. 2017; Bonifacio et al. 2019; Venn et al. 2020; Lucchesi et al. 2022; Lombardo et al. 2023 for details), it is clear from this comparison that the current catalogues and recommended quality cuts would result in a very successful sample selection, as the more metal-rich stars are now correctly identified as such. We do, however, still see a clear trend with temperature and, as expected, the photometric metallicities are less accurate for stars with higher temperatures above ~6 000 K, a regime for which the lines of equal metallicity are closer together in our color–color space.

7.4.2. External comparisons

A more interesting and powerful test of the photometric metallicities in the metal-poor regime comes from their comparison with a set of spectroscopic metallicities of metal-poor stars obtained with high-resolution spectroscopy that were not included in the

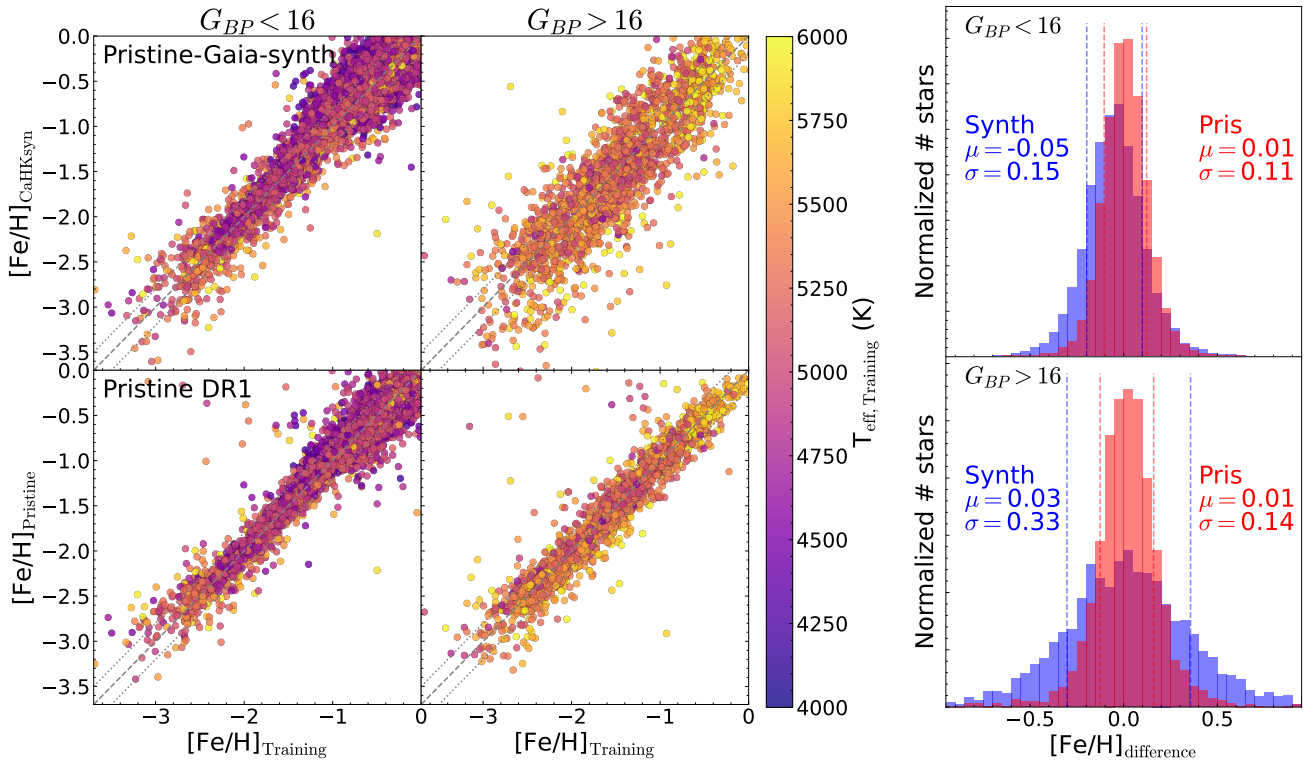

Fig. 13. Comparisons of the Pristine-*Gaia* synthetic (top row of the left-hand panels) and Pristine DR1 (bottom row of the left-hand panels) photometric metallicities with the spectroscopic metallicities of the cleaned giant star training set (see text for details). The right-hand panels show the histograms of metallicity differences for the two catalogues, with the Pristine-*Gaia* synthetic catalogue in blue and Pristine DR1 in red. The catalogues are split as a function of $G_{\rm BP}$ magnitude, into the bright end ($G_{\rm BP}$ < 16; top) and the fainter end ($G_{\rm BP}$ < 16; bottom). At brighter magnitudes than this limit, both catalogues are similarly competitive, as can be appreciated from the confidence interval of the metallicity differences in the right-hand panels (16th to 84th percentiles; vertical dashed lines) and the average of these percentile offsets quoted as σ in the two panels. At fainter magnitudes, the Pristine DR1 catalogue continues to remain very competitive while the Pristine-*Gaia* synthetic catalogue suffers from the lower signal-to-noise of the $CaHK_{\rm syn}$ magnitudes, leading to a wider distribution of metallicity differences.

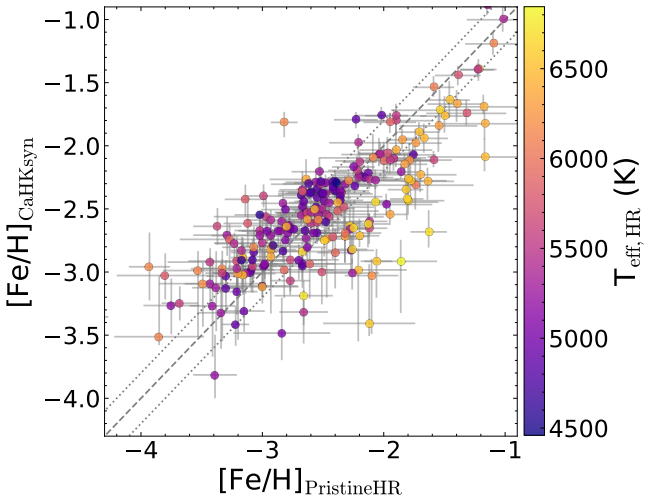

Fig. 14. Comparison of the metallicities from the Pristine-*Gaia* synthetic catalogue with the high-resolution spectroscopic metallicities accumulated by the Pristine collaboration. The photometric metallicities are in very good agreement with their spectroscopic counterparts throughout the metallicity scale, which means that the issues (variable stars, subtly saturated SDSS stars, ...) that led us to erroneously target stars with [Fe/H]$_{\rm PristineHR}$ > −2.0 as VMP candidates are now resolved. There remains an offset between the photometric and spectroscopic metallicities with the temperature of a star (the color of the dots). Stars with $T_{\rm eff}$ > 6000 K do tend to have lower photometric metallicities than their spectroscopic metallicities.

training sample. These include four sets of high-resolution spectroscopic observations (Hansen et al. 2018; Sakari et al. 2018; Yong et al. 2021; Li et al. 2022; R > 25 000). All these stars are mainly located in the southern sky (where the coverage of the Pristine survey is extremely limited) or were only recently released, which explains why they were not taken into account when building the training set. They consequently provide a completely independent test of the quality of the photometric metallicities. To test the consistency of the method over a range of metallicities and/or temperatures, we deliberately chose to include in this comparison large samples of homogeneously analyzed stars, rather than a compilation of individual follow-up results from the literature. This comparison is presented in Figure 15 and shows an excellent agreement between the photometric metallicities we calculate and the high-quality spectroscopic metallicities. Despite increased uncertainties in the EMP regime ([Fe/H] < −3.0), the Pristine metallicities are able to discriminate between stars of metallicities [Fe/H] ∼ −2.0, [Fe/H] ∼ −3.0 and even [Fe/H] ∼ −3.5. The right-hand panel of the Figure however illustrates that this capability decreases somewhat for hotter stars. Nevertheless, for most FGK stars, the catalogue of metallicities we publish here should be particularly suited to hunt for the low end of the MDF over the whole sky, just as the Pristine metallicities have been used to track such stars over the Pristine footprint (Starkenburg et al. 2018; Youakim et al. 2017; Aguado et al. 2019; Arentsen et al. 2020a; Caffau et al. 2020; Venn et al. 2020; Kielty et al. 2021; Lucchesi et al. 2022; Sestito et al. 2023).

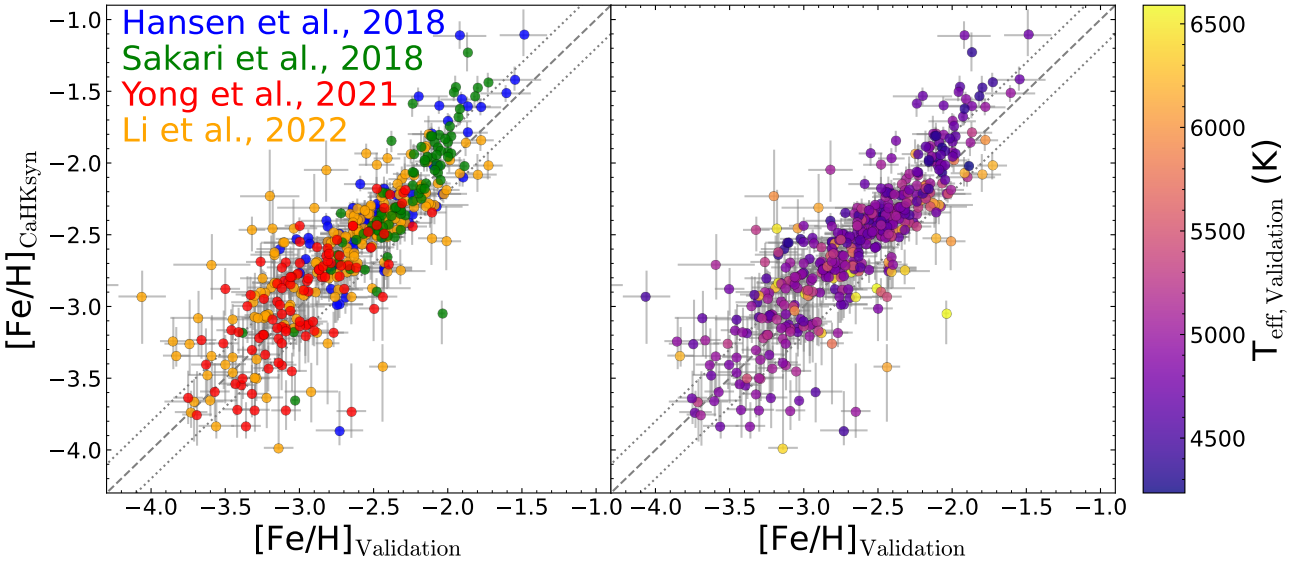

Fig. 15. Comparison of the Pristine-*Gaia* synthetic photometric metallicities with metallicities from high-resolution spectroscopy samples of metal-poor stars (Hansen et al. 2018; Sakari et al. 2018; Yong et al. 2021; Li et al. 2022; $R > 25\,000$). In the left-hand panel, color traces the origin of the spectroscopic metallicities, as listed in the label. The photometric metallicities are of high enough quality that they track the spectroscopic metallicities down to the lowest metallicities available. In the right-hand panel, the color codes the temperatures of the same stars, showing that the agreement is generally better for cooler rather than hotter stars. In both panels, the dashed line corresponds to the 1-to-1 line and the dotted lines represent offsets of 0.2 dex.

7.4.3. Comparisons with other *Gaia*-based metallicity catalogues

Finally, Figure 16 shows a succinct comparison of current *Gaia*-based metallicity catalogues with the giant stars in our training set (left-hand set of panels) and the validation sample presented in the previous sub-section (right-hand set of panels). In particular, we show a comparison of the spectroscopic metallicities of those two sets with the metallicities of catalogues from:

- the Pristine-*Gaia* synthetic and Pristine DR1 photometric metallicity catalogues, as already shown above (top rows);
- Recio-Blanco et al. (2023), based on the *Gaia* Radial Velocity Spectrometer (RVS) spectra around the Calcium triplet region, with metallicities derived using the General Stellar Parametriser-spectroscopy (GSP-Spec), with the recommended set of flags for low metallicity stars[27] (middle left-hand panel);
- Andrae et al. (2023a), with metallicities derived using the General Stellar Parameterizer from Photometry (GSP-Phot) applied to the *Gaia* astrometry, photometry, and BP/RP spectrophotometry (middle right-hand panel);
- Andrae et al. (2023b, v2.1 catalogue), based on a data-driven algorithm, XGBoost, applied to the BP/RP coefficients (excluding the region of the Ca H&K region because of low S/N) and AllWISE magnitudes, trained on a sample of APOGEE data complemented by a small number of VMP stars (bottom left-hand panel);
- Zhang et al. (2023), with a catalogue of stellar parameters (including metallicities) obtained by forward modeling the stellar type, extinction, and parallax of stars onto the space of XP spectra and near-infrared 2MASS and WISE photometry, using the information of the LAMOST survey for training (bottom right-hand panel).

As mentioned above, the top row of panels in the left-hand set is a beneficial comparison for the Pristine metallicities since it shows a comparison with the sample that was used to build the Pristine model. While this comparison is only including stars on the giant branch (with $\log g < 3.9$) and $T_{\rm eff} < 6000$ K, a comparison with the turn-off, or main-sequence stars is shown in the appendix (see Figures B.1 and B.2). We note that, generally, the performance is better for giants and moreover decreases toward hotter stars. The top row of panels in the right-end set, however, provides an external comparison with the validation sample that was not used to build any of the models, except the Andrae et al. (2023b) model that was partially trained on the Li et al. (2022) sample that constitutes about a quarter of the validation set. In all panels, the color of the data points encodes the magnitude of a star (left-hand set of panels) or its $T_{\rm eff}$ (right-hand set of panels).

These comparisons are very informative and show the strengths and weaknesses of the different metallicity data sets that can be quite complementary and should be used preferentially for different science goals. Stars with GSP-Spec metallicities are all bright since they require RVS spectra[28]. Because of that, they do not contain a lot of metal-poor stars. Applying the strict flags listed in Recio-Blanco et al. (2023) is absolutely essential to clean up the data in the VMP regime and to prevent being dominated by spurious metallicity values (small points in the corresponding panel). Despite this clean-up, there remains stars with GSP-Spec metallicities that deviate from those of the high-resolution validation sample[29]. The strength of these spectra however lies in the additional information available for them, including, of course, the radial velocities that provide full 6-dimensional phase-space information, from the rich multi-dimensional abundance information, and from the high-quality $\log g$, $T_{\rm eff}$, and further information that they can provide (e.g., Kordopatis et al. 2023).

[27] We use the set of flags described in their Sub-section 10.5 and in the caption of their Figure 26.

[28] Even though *Gaia* RVS spectra are observed down to $G \sim 15$, not all of these have high signal-to-noise. GSP-Spec analyzed spectra with signal-to-noise ratios > 20 yield atmospheric parameters and the corresponding stars typically have G <12–13.

[29] At least in part because GSP-Spec metallicities have large uncertainties in the VMP regime.

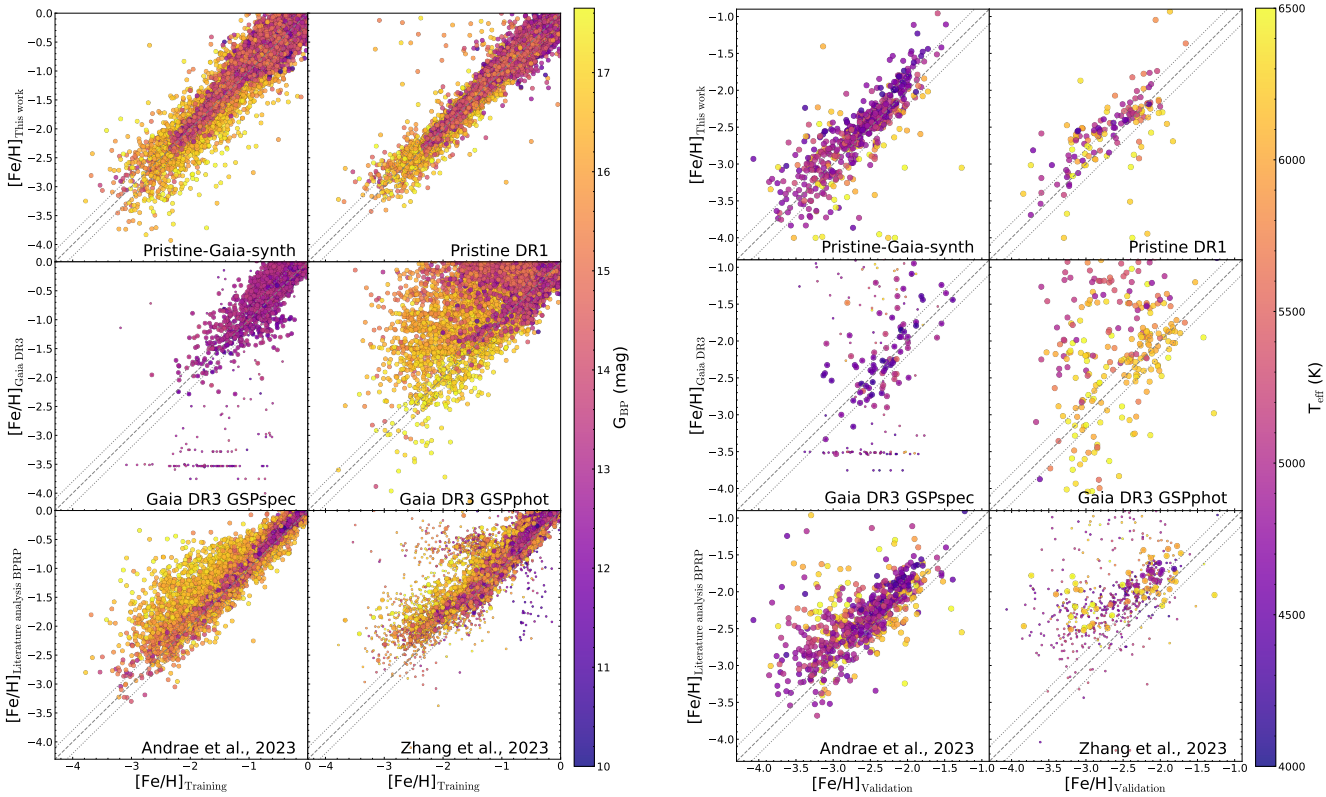

Fig. 16. Comparison of various metallicity catalogues based on *Gaia* DR3 with the giant stars in the training (left, see text for the construction of this sample) and with the validation sample at lower metallicities (right, see text for details) and their spectroscopic metallicities. After the comparison with the Pristine-*Gaia* synthetic catalogue, and the Pristine DR1 catalogue, the panels show the comparisons for the *Gaia* GSP-Spec metallicities based on the *Gaia* RVS spectra (Recio-Blanco et al. 2023, smaller circles correspond to stars flagged as unreliable in this catalogue, following the same criteria discussed in their Section 10.5 and represented in their Figure 26), the *Gaia* GSP-Phot catalogue based on the BP/RP information (Andrae et al. 2023a), the XGBoost algorithm (Andrae et al. 2023b) and the Zhang et al. (2023, smaller circles correspond to stars flagged to have unreliable metallicities) catalogue that use the BP/RP information and some additional non-*Gaia* photometry. The color encodes the magnitude of a star in the left-hand set of panels and their temperature in the right-hand set of panels. The dashed line corresponds to the 1-to-1 line and the dotted lines represent offsets of 0.2 dex.

GSP-Phot, XGBoost, and the Zhang et al. (2023) metallicities go significantly deeper as they were derived using the BP/RP information and are more directly comparable to the two Pristine metallicity catalogues. The GSP-Phot metallicities are definitely challenged and include a sizable number of stars with erroneously high metallicities. This is particularly problematic when comparing with the validation data set in the VMP regime (corresponding bottom panel), where strong GSP-Phot systematics lead to metallicities that cover the full range.

The XGBoost (Andrae et al. 2023b) and Zhang et al. (2023) catalogues, which are both based on the BP/RP coefficients but also supplemented by information external to *Gaia*, provide metallicities that are much better behaved over the full metallicity range, especially at the bright end (magenta points in the right-hand set of panels). However, in the VMP regime, even though the metallicities of both catalogues are significantly better behaved than GSP-Phot values, a non-negligible fraction of stars are systematically assigned more metal-rich photometric metallicities compared to their spectroscopic counterpart (as reported by both Andrae et al. 2023b and Zhang et al. 2023). These stars are still often assigned metallicities below $[Fe/H]_{phot} < -2.0$, but this would limit the ability to build a pure and complete sample of stars with $[Fe/H] < -2.5$ or < -3.0. This behavior is somewhat worse for the Zhang et al. (2023) catalogues, possibly because Andrae et al. (2023b) took the extra step of complementing their APOGEE training sample with a specific set of VMP stars. A similar overestimation of the metallicities is also visible for $[Fe/H]_{spec} < -1.0$ for fainter stars when comparing with the training set and could indicate that this bias is related to the S/N of the spectral features, either because a star is VMP, or because it is faint. Both catalogues do work well for bright and/or fairly metal-rich stars (despite a small number of catastrophic failures in the Zhang et al. 2023 catalogue), the latter of which constitute most of the MW stars. Both samples also have the advantage of providing metallicities in all regions of the MW: using near-infrared photometry in conjunction with the BP/RP coefficients and training on APOGEE or LAMOST stars that cover a wide range of extinction values helps XGBoost and the Zhang et al. (2023) algorithm to separate the impact of extinction on the input features.

None of these four external catalogues were specifically built to treat VMP stars. We may therefore expect that the Pristine-*Gaia* synthetic and the Pristine DR1 catalogues handle the very low-metallicity range better and this is indeed what we can see in the top rows of panels in Figure 16. As already discussed above, the Pristine metallicities are well aligned with the spectroscopic metallicities, despite a small tendency for them to be underestimated in the lowest metallicity regime. In the $[Fe/H] > -1.5$ regime, we note that the XGBoost metallicities show less scatter than the Pristine metallicities at the bright end and, at the metal-rich end, the Pristine metallicities are clearly suboptimal compared to the XGBoost or most of the Zhang et al. (2023)

values. A comparison of the plots for the two Pristine catalogues emphasizes what we already mentioned above: the Pristine DR1 catalogue has higher S/N than the Pristine-*Gaia* synthetic catalogue (left-hand set of panels) but the latter covers a significantly larger footprint and allows for a more thorough comparison, in particular with the high-resolution validation set (right-hand set of panels).

7.5. Impact of quality cuts on the search for UMP stars

The search for the most metal-poor stars and, in particular, UMP stars is one of the key uses for the metallicity catalogues. However, the dearth of good calibrators and models at those metallicities led us to fix the lowest metallicities output by the model described in Section 6 to $[Fe/H]_{phot} = -4.0$. In addition, this metallicity limit is assigned to all stars within 0.2 mag of the $[Fe/H] = -4.0$ model in the color–color space of Figure 9 and we consequently recommend not using stars with $[Fe/H]_{phot} = -4.0$ for a generic use of the catalogue. One may therefore naturally wonder whether the proposed quality cuts negatively impact the search for the most metal-poor stars by simply flagging them out.

To explore this question, we cross-match the list of 42 UMP stars provided by Sestito et al. (2019) with the Pristine-*Gaia* synthetic metallicity catalogue. There are 23 stars in common between the two catalogues, with the other stars being too faint to have a metallicity in the Pristine-*Gaia* synthetic catalogue. Figure 17 shows the distribution of those 23 stars in the Pristine color–color space. Reassuringly, almost all stars are within the region $-4.0 < [Fe/H]_{phot} < -3.0$. Within this extremely low-metallicity region, they are not specifically located at the lower metallicity edge of the model (the upper edge in the figure). This could be due to the large amount of carbon present in these stars that could increase the absorption in the CaHK filter leading to photometric metallicities that are biased high (see Sub-section 7.7 below). In particular, this is the case for the most problematic star, SDSS J081554.26+472947.5, that appears in red in the figure, with $[Fe/H]_{phot} = -1.84^{+0.39}_{-0.57}$ in the Pristine-*Gaia* synthetic catalogue. This star is also present in the Pristine DR1 catalogue, with a much more consistent $[Fe/H] = -3.64^{+0.46}_{-0.36}$, which means that its erroneous metallicity in the Pristine-*Gaia* synthetic catalogue is driven by its large $CaHK_{syn}$ uncertainties, $\delta CaHK_{syn} = 0.06$, and its blue color, for which the iso-metallicity lines are closer in the Pristine color–color space.

While there is a handful of true UMP stars that are located above the $[Fe/H] = -4.0$ theoretical model (black line), most of these are below the $[Fe/H]_{phot} = -4.0$ threshold, as defined by the model given the pixelization of the color–color space (gray line; see sub-Section). Most of these stars would not be removed by the $[Fe/H]_{phot} > -4.0$ recommended above. There remains two genuine UMP stars, represented by square symbols in the figure and that both have spectroscopic $[Fe/H] < -5.0$, that are located above this line and would be rejected by the recommended cut. These two stars also have large $CaHK$ uncertainties, as can be seen from their large vertical error bars and they are being pushed in the $[Fe/H]_{phot} = -4.0$ region may simply be a consequence of these large uncertainties. Nevertheless, a catalogue user that aims at isolating candidate UMP stars may wish to loosen the recommended cut to $[Fe/H]_{phot,84th} > -4.0$. This change would prevent the removal of the two egregious UMPs.

7.6. Impact of extinction and latitude

Even though the largest extinction regions, with $E(B-V) > 0.5$, are excluded from this study, we stress that, as we adopt an integrated extinction map that has its limitations (Schlegel et al. 1998), much care should be taken for stars in high extinction regions that are listed in the catalogue. To test possible systematics with $E(B-V)$, we compare our Pristine-*Gaia* synthetic photometric metallicities with spectroscopic metallicities from APOGEE DR17[30] (Abdurro'uf et al. 2022) and LAMOST DR8[31] (Zhao et al. 2012), which both have many stars across a range of extinction. We select giant stars with spectroscopic $[Fe/H] < -1.0$ passing basic spectroscopic quality criteria, remove globular cluster stars from the sample, and only keep stars with $\delta[Fe/H]_{phot} < 0.3$ dex. To investigate spatial trends, we limit APOGEE to the inner Galaxy ($|l| < 45°$) and split the LAMOST sample in two: $l < 90°$ (inner Galaxy) and $l > 90°$ (anticenter). The result is shown in Fig. 18. Up until $E(B-V) = 0.3$, the comparison between photometric and spectroscopic metallicities is excellent, to within ~0.05 dex. Between $0.3 < E(B-V) < 0.5$, systematic offsets begin to appear, but, on average, they remain <0.2 dex. There is an opposite trend for the inner Galaxy and the anticenter, which might be related to systematics in the reddening map and/or to different distance distributions for stars in the sample (nearby stars might be over-corrected and distant stars might be under-corrected for extinction when using an integrated extinction map). In summary, our photometric metallicities appear to be good for $E(B-V) < 0.3$, while for $0.3 < E(B-V) < 0.5$ some biases appear to be creeping in[32].

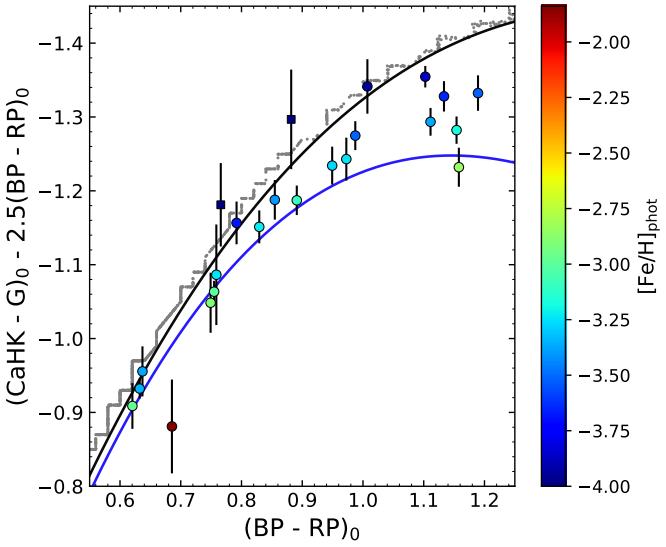

Fig. 17. Distribution in the Pristine color–color space of the 23 known UMPs in common with the Pristine-*Gaia* synthetic catalogue. Stars are color-coded by their photometric metallicities $[Fe/H]_{phot}$. The blue and black lines are polynomial fits to the expected location of stars with $[Fe/H] = -3.0$ (blue) and $[Fe/H] = -4.0$ (black) based on the spectral synthesis described in Section 6.2 and also represented by star symbols in Figure 9. The gray points correspond to stars with $-4.0 < [Fe/H]_{phot} < -3.99$ and highlight the location of the region with $[Fe/H]_{phot} = -4.0$, as assigned by the Pristine model. Only two stars, represented by squares, are above this region and technically rejected by the $[Fe/H]_{phot} > -4.0$ cut recommended in Sub-section 7.3. Because these stars have large uncertainties, requesting only $[Fe/H]_{phot,84th} > -4.0$ would select them back in the sample of high-quality stars.

[30] https://www.sdss4.org/dr17/
[31] https://www.lamost.org/dr8/
[32] In their study to calibrate the J-PLUS photometry from the *Gaia* XP data, López-Sanjuan et al. (2024) find a potential trend of their calibration with extinction, which they venture could be the sign of a degradation of the calibration of the XP data in regions of high extinction. It could explain some of the systematic deviations we measure when the extinction increases.

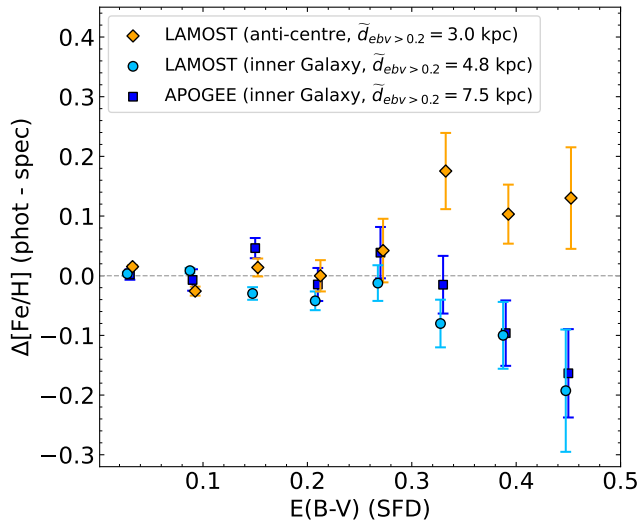

Fig. 18. Median difference between Pristine-*Gaia* synthetic photometric metallicities and APOGEE/LAMOST spectroscopic metallicities as a function of the extinction values $E(B-V)$ from the map of Schlegel et al. (1998, i.e., SFD) used to deredden the magnitudes. The samples are split in Galactic longitude to investigate spatial trends, with APOGEE having mostly stars in the inner Galaxy and the LAMOST sample covering two spatial regions: toward the inner Galaxy or the anti-center. The median distance of high extinction stars with $E(B-V) > 0.2$ in each sample are given in the legend. Error bars correspond to $\sigma/\sqrt{N}$, with σ the standard deviation between $[Fe/H]_{phot}$ and $[Fe/H]_{spec}$ in a given bin after clipping 3-sigma outliers, and N the number of stars in this bin.

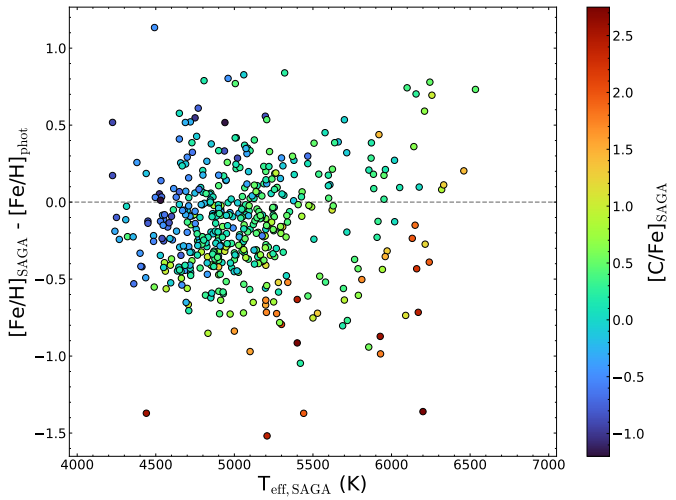

Fig. 19. Impact of the carbon abundance of a star on the derived Pristine photometric metallicities. The difference between the photometric metallicities from the Pristine-*Gaia* synthetic catalogue and the spectroscopic metallicities from the SAGA database is plotted against the spectroscopic effective temperature and color-coded by the carbon abundance, [C/Fe]. Stars that are strongly carbon-enhanced show a clear bias and have overestimated Pristine metallicities, especially for lower T_{eff}.

Additionally, comparisons of our results with the spectroscopic samples reveal a larger fraction of metal-rich contaminants at low latitudes, close to the Galactic plane, even at lower extinction values in the Schlegel et al. (1998) maps. We recommend additional caution in these regions. As they are overwhelmingly dominated by metal-rich stars, any source of contamination from those in the metal-poor regime, even small, can become problematic.

7.7. Impact of carbon

The presence of strong carbon bands in the spectrum of a star can significantly bias the estimated photometric metallicity as carbon absorption in the wavelength range of the CaHK filter are interpreted as stronger Ca H & K lines. This is of particular importance in the VMP regime because a sizable fraction of VMP and EMP stars are expected to be carbon-enhanced ([C/Fe] > 0.7), even though different data sets yield different conclusions on the exact value of this fraction (e.g., Arentsen et al. 2022). From the Pristine spectroscopic follow-up presented in Aguado et al. (2019), Arentsen et al. (2022) showed that selecting EMP candidates from the Pristine photometric metallicities leads to a dearth of cool (T_{eff} < 5500 K) Carbon Enhanced Metal-Poor (CEMP) stars (see also Caffau et al. 2020). To investigate the effect of this bias, we compare the metallicities from the Pristine-*Gaia* synthetic catalogue with the spectroscopic metallicities from the SAGA database (Suda et al. 2008). As SAGA is a compilation of different sources the selection function may well be biased with regards to carbon, but for comparative purposes it suits our needs well. We use the recommended version of the data set from April 7, 2021 and limit the comparison to stars with $-4.0 < [Fe/H]_{SAGA} < -1.0$ to match the recommended range for the photometric metallicities. The catalogues are crossmatched with a search radius of 1″ and [M/H] values are used for stars without [Fe/H] values in SAGA.

Figure 19 shows the difference in metallicity between Pristine and SAGA, color-coded by their carbon abundance, [C/Fe], as listed in the SAGA database. It is clear that stars with large carbon abundances have photometric metallicities that are significantly overestimated. In the cool regime, for stars with T_{eff} < 5500 K, most photometric metallicities are overestimated for stars with [C/Fe]> 0.7, with a mean discrepancy of 0.70 dex. For the hotter stars with T_{eff} > 5500 K the mean deviation is noticeably smaller at 0.33 dex. The temperature dependance of these discrepancies are in good agreement with the bias found by Arentsen et al. (2022). Because of this, we caution against drawing conclusions about the fraction of stars with significant carbon enhancement in samples derived from the Pristine catalogues. A more comprehensive analysis of the causes of this bias is under study (Montelius et al., in prep.).

8. Science tests

8.1. Science test 1: The metallicity of globular clusters

An illustration of both the quality and the scientific usefulness of the photometric metallicities we provide is presented in Figure 20. The figure shows the photometric metallicities for all stars in the vicinity of four different MW globular clusters (GCs) that have been chosen to span a wide range in metallicity. Here, we use stars with $\delta[Fe/H]_{phot} < 0.5$ dex. Close to the center of the clusters, crowding becomes an issue and the holes in the data are produces by the imposed quality cuts (Subsection 7.3), especially the *Gaia* quality cuts. Apart from these central regions without data, the member stars of metal-poor GCs are very easy to isolate from the foreground contamination of MW stars and the resulting photometric metallicities are all in excellent agreement with their literature metallicities, despite spanning almost 2 dex (the red lines in the two first columns of the figure; 47 Tuc, [Fe/H] = −0.76 ± 0.02; M 5, [Fe/H] = −1.33 ± 0.02; M 15, [Fe/H] = −2.33 ± 0.02; M 92, [Fe/H] = −2.35 ± 0.05;

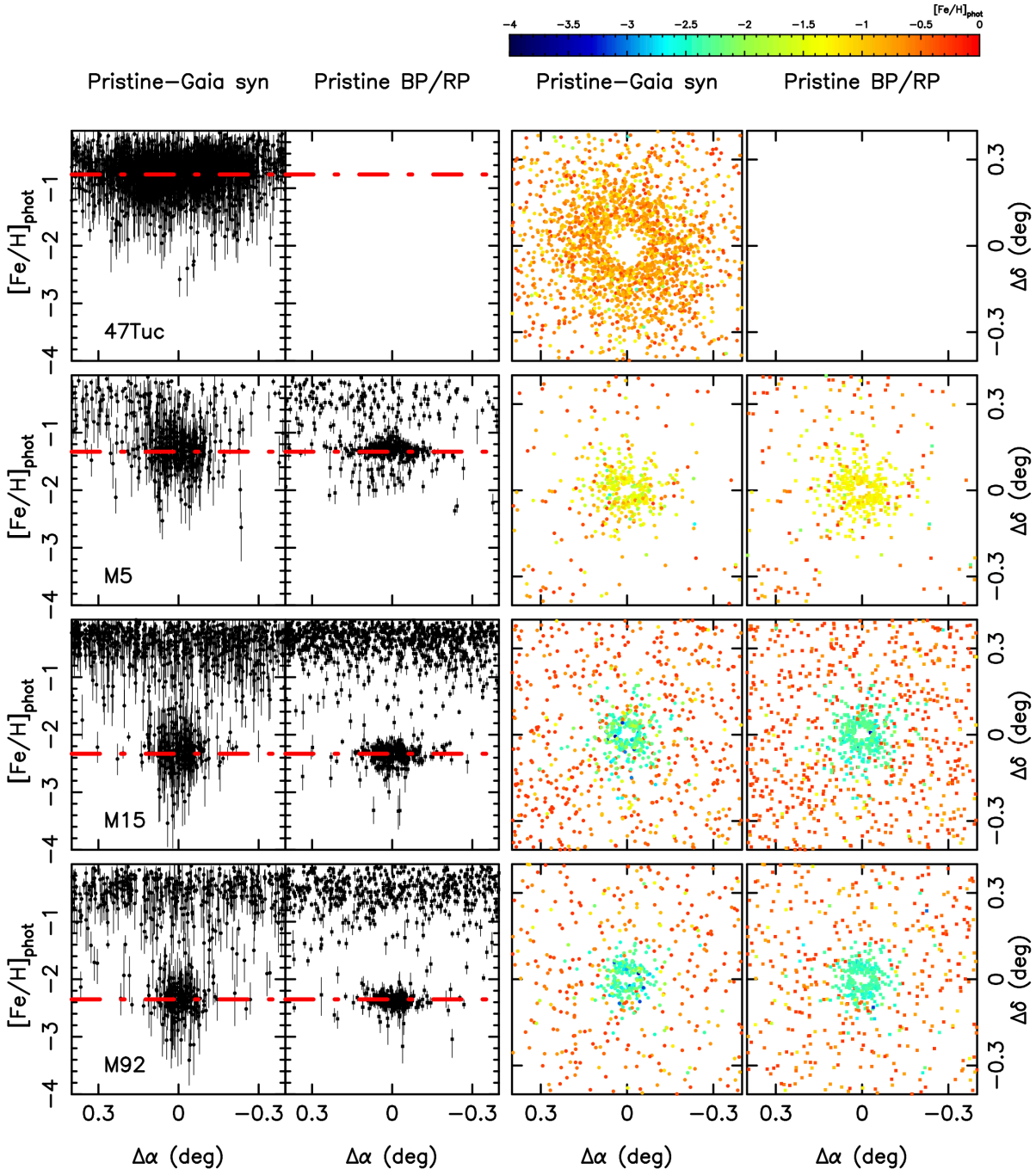

Fig. 20. Metallicity of stars in both the Pristine-*Gaia* synthetic and Pristine DR1 catalogues around four globular clusters that span a wide metallicity range. From top to bottom, 47 Tuc, M 5, M 15, and M 92. The two left-most columns show the distribution of metallicities of stars in both catalogues as a function of right ascension. They are in good agreement with the known metallicity of the clusters, indicated by the red dot-dashed lines (Carretta et al. 2009). The two right-most columns of panels show the distribution of stars around the clusters, color-coded by metallicity. With the metallicity information, likely members of the clusters are very easy to pick out among the field contamination.

Carretta et al. 2009). In the case of the two most metal-poor clusters, M 15 and M 92, it becomes trivial to separate metal-rich, MW contaminants from member stars and opens up an exciting discovery space for "extra-tidal" cluster member stars, or cluster stellar stream stars for those that are disrupting (e.g., M 92; Thomas et al. 2020). When a cluster appears in both metallicity catalogues, the higher signal-to-noise of the Pristine photometry, combined with the $\delta CaHK < 0.1$ quality cut to calculate $[\mathrm{Fe/H}]_{\mathrm{phot}}$, translate to significantly more stars with Pristine DR1 metallicities.

To further explore the accuracy of our photometric metallicities, we systematically examine the whole population of MW GCs. Using catalogues of known GCs (Harris 1996; Baumgardt & Vasiliev 2021), we select stars around GCs that also have *Gaia* color-magnitude and proper motion information compatible with the properties of the clusters. We then extract photometric metallicities from the Pristine-*Gaia* synthetic and Pristine DR1 catalogues for those stars, once again using the quality cuts presented in Section 7.3, except that we use a stricter $\delta[\mathrm{Fe/H}]_{\mathrm{phot}} < 0.3$ dex. For those clusters with a well-populated red giant branch

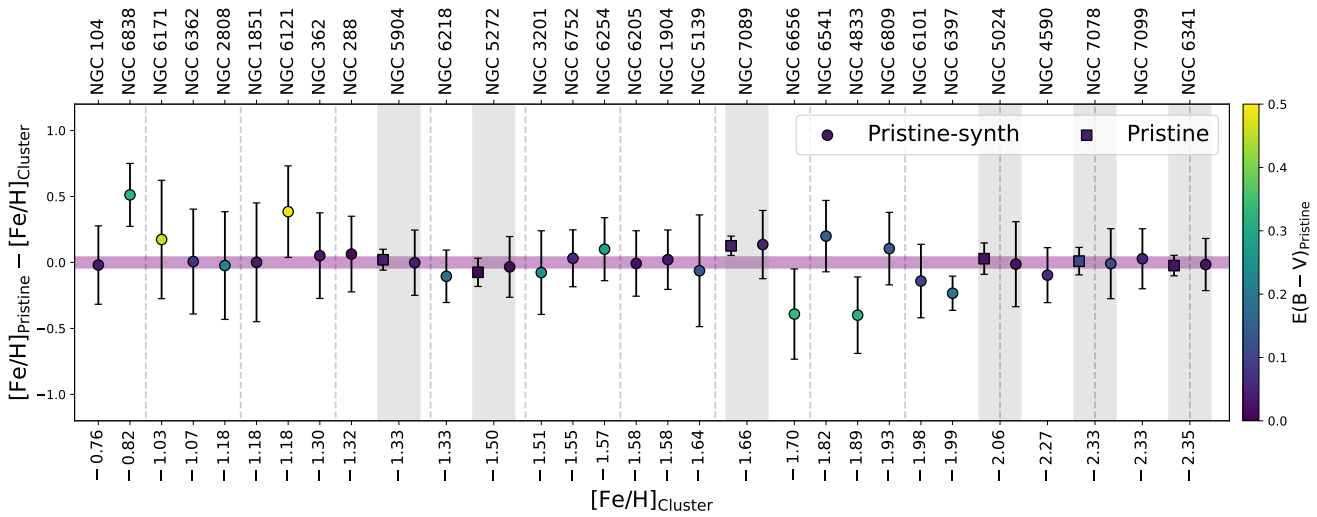

Fig. 21. Difference between the mean metallicity of 35 GCs with at least 100 stars in the Pristine-*Gaia* synthetic (circles) and Pristine DR1 (squares) catalogues, compared to the literature values (Carretta et al. 2009), over a broad range of metallicities. The gray bands highlight clusters that are present in both catalogues. Cluster names and metallicities are listed, and the symbols are color-codded by the $E(B-V)$ value toward each GC. The expected zero purple line and the vertical dashed lines are highlighted for visualization purposes.

(more than 100 stars above the sub-giant branch) we compute the mean difference between the Pristine metallicities and the literature value by fitting a Gaussian to that distribution.

In total, we can compare results from the photometric metallicity catalogues and literature metallicities (Carretta et al. 2009) for 35 clusters that cover the entire metallicity distribution of MW GCs. The results are shown in Figure 21 and highlight a very good agreement over the full metallicity range. There may be the hint of a mild overestimation of the Pristine metallicities at the metal-rich end. However, some of these differences also appear driven by the redenning (coded by the color of the symbols in the figure). More highly reddened clusters tend to yield mean Pristine DR1 and Pristine-*Gaia* synthetic metallicities that deviate more significantly from the literature values. Nevertheless, this comparison with MW GCs emphasizes the very good quality of the Pristine photometric metallicities that were assembled without any information on the metallicity of the GC clusters.

8.2. Science test 2: Slicing the MW by metallicity

A second illustration of the power of the all-sky catalogue of photometric metallicities we publish is presented in Figure 22, where we show density maps of MW stars for different metallicity slices. With the aim for these maps to be all-sky, we only use the Pristine-*Gaia* synthetic catalogue and include stars with $\delta[\mathrm{Fe/H}]_{\mathrm{phot}} < 0.3$ dex, along with the different quality cuts described in Subsection 7.3. We further add a restriction on the parallax to remove nearby stars ($\varpi + \delta\varpi < 1''$). Changes in the shape of the MW[33] with metallicity are directly evident, with the most metal-rich stars ([Fe/H] > -1.0, map I) exhibiting a clear domination of the disk that transitions to the rounder distribution of the halo as more and more metal-poor samples are explored with $-1.5 <$ [Fe/H] < -1.0 (map II), $-2.5 <$ [Fe/H] < -1.5 (map III), and finally $-4.0 <$ [Fe/H] < -2.5 (map IV).

Some of the features in those maps are however directly related to the *Gaia* scanning law (Cantat-Gaudin et al. 2023) and it can be difficult to visualize the exact shape of the MW stars intersected by the complex volume probed by the catalogue. We therefore also show maps of density ratios in the bottom four panels of the figure. These have the advantage of exhibiting fewer artifacts and more clearly highlight the differences in the morphology of the MW in the different metallicity slices.

With the mean metallicity of the halo being [Fe/H] ~ -1.5 (e.g., Ryan & Norris 1991; Carney et al. 1996; Ivezić et al. 2008; Schörck et al. 2009; Beers et al. 2012; Ibata et al. 2017b; Youakim et al. 2020), the ratio of the two most metal-rich maps, II and I, clearly highlights the transition from the disk-dominated metallicity regime to the rounder halo-dominated regime. Interestingly, the III-to-II density ratio map implies that, overall, the density of stars with $-1.5 <$ [Fe/H] < -1.0 and $-2.5 <$ [Fe/H] < -1.5 are fairly similar, except in the anticenter direction that is clearly more metal-rich. This is possibly due to the numerous stellar features in this region, globally referred to as the Galactic Anticenter Stellar Structure (GASS; e.g., Newberg et al. 2002; Ibata et al. 2003; Slater et al. 2014; Morganson et al. 2016) and now thought to be perturbed disk material sent to high latitudes (e.g., Bergemann et al. 2018; Laporte et al. 2020; Li et al. 2021).

Focussing on the low metallicity end of the halo with the IV/II map and, especially, the IV/III map (bottom row of the Figure) highlights that the relative density of the most metal-poor stars is higher toward the outer MW than toward the central Galactic regions (in the bottom-right panel of the figure, the directions toward the inner galaxy has lower density ratios that other direction). This is reminiscent of the findings of Starkenburg et al. (2017b) and El-Badry et al. (2019) that lines of sight toward the inner galaxy are privileged directions to find the oldest stars but that they suffer from the domineering presence of in situ stars formed as or after the halo of the MW assembled. Additional information, such as kinematics and detailed chemical abundances, are essential to sift through these stars to isolate samples of accreted stars that trace the assembly

[33] The absence of the Magellanic Clouds from these maps stems from the relative brightness of the Pristine-*Gaia* synthetic catalogue. Being distant and very red, giant branch stars from the Clouds are effectively removed by the $\delta CaHK < 0.1$ cut. The Schlegel et al. (1998) dust map is also not very reliable toward the central regions of the two satellite galaxies and, in this region, the de-reddened magnitudes should be used with caution.

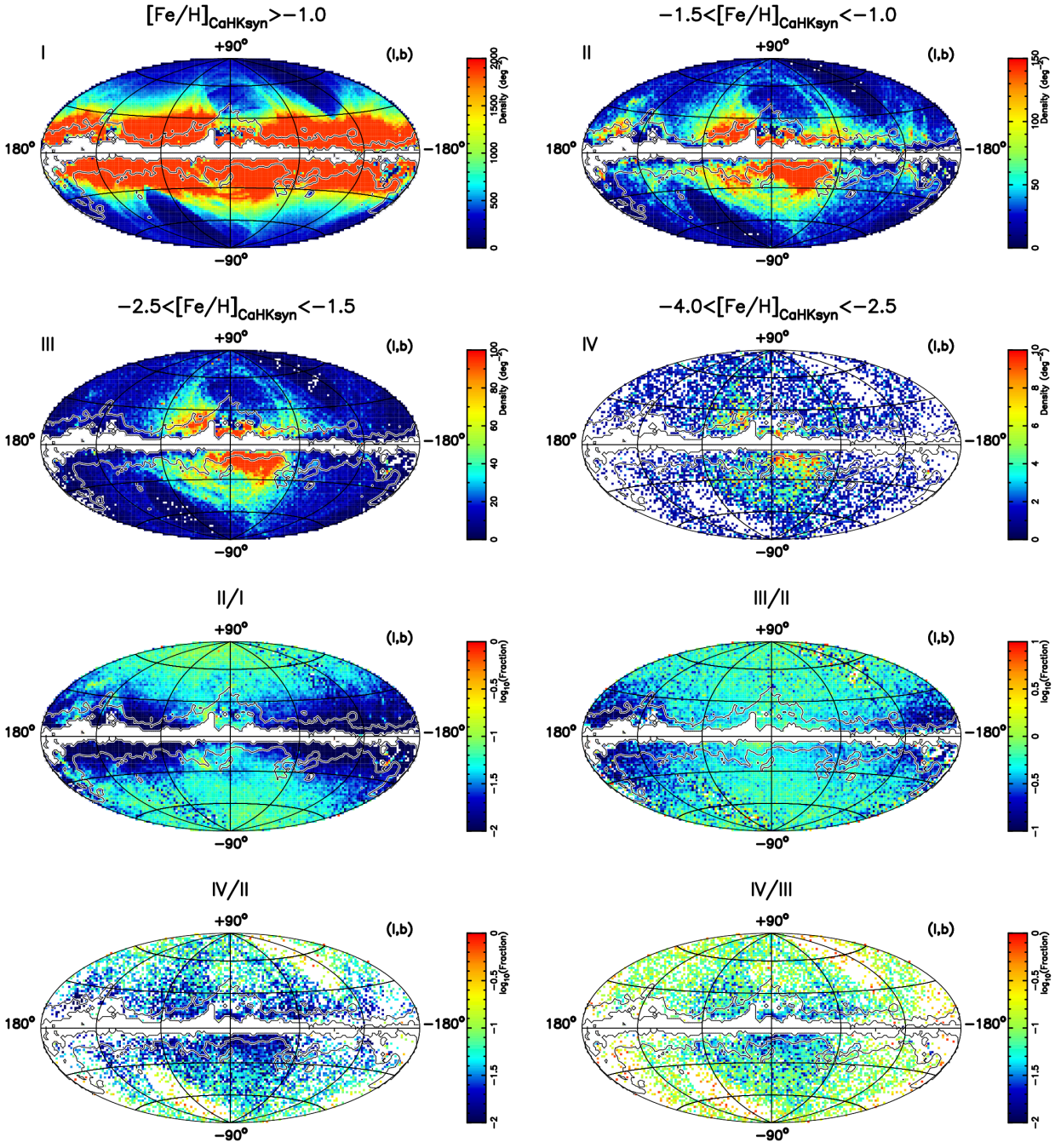

Fig. 22. Density maps for different slices of Pristine-*Gaia* synthetic metallicities, projected in Galactic coordinates. The contours correspond to $E(B-V) = 0.5$ and $E(B-V) = 0.15$. From top-left to bottom-right, the four top panels display density for metallicity ranges $[\mathrm{Fe/H}]_{\mathrm{CaHKsyn}} > -1.0$ (panel I), $-1.5 < [\mathrm{Fe/H}]_{\mathrm{CaHKsyn}} < -1.0$ (panel II), $-2.5 < [\mathrm{Fe/H}]_{\mathrm{CaHKsyn}} < -1.5$ (panel III), and $-4.0 < [\mathrm{Fe/H}]_{\mathrm{CaHKsyn}} < -2.5$ (panel IV). To overcome the inhomogeneous survey and highlight the transition in the shape of the MW, the bottom four panel shows ratios of density maps. From top-left to bottom-right, II/I, III/II, IV/II, and IV/III.

of the proto-MW (e.g., Arentsen et al. 2020b; Belokurov & Kravtsov 2022; Rix et al. 2022).

More detailed information on the "mono-abundance" relative shape of the MW is not directly measurable from these maps that probe an irregular volume of the Galaxy around the Sun's position. However, the regular shape of the maps in Figure 22 implies that it should be straightforward to parametrize the shape of the MW as a function of metallicity. Provided, of course, that the very complex selection function imparted onto the data by the *Gaia* scanning law, the choices made by the *Gaia* consortium as to which star would be included in the catalogue of BP/RP coefficients and, of course, our own color and S/N cuts. This will be the focus of future work.

9. Summary

In this paper, we have taken advantage of the recently released *Gaia* DR3 spectrophotometric BP/RP information to build a catalogue of synthetic $CaHK_{\mathrm{syn}}$ magnitudes mimicking the narrow-band photometry of the Pristine survey that, centered on the calcium H & K lines, is sensitive to the metallicity of a star. We seize this opportunity to reprocess the full ∼11 500 images

of the Pristine survey gathered since 2015 and present the current state of the survey that now covers more than 6500 deg^2. In particular, we use the *Gaia*-based synthetic observations to refine the calibration of the Pristine photometry with a neural-net based algorithm, `PhotCalib`. We show that the new photometric catalogue is significantly flatter than before, better accounting for systematics across the field of view, and accurate at the 13 mmag level (from a previously underestimated ∼40 mmag with the previous calibration). The Pristine model that derives the photometric metallicity of a star from *CaHK* magnitudes and broadband magnitudes is updated to rely solely on *Gaia* broadband information (G, G_{BP}, G_{RP}). As part of this new model, we also present a more reliable way to iteratively deal with the extinction correction that is challenging with the *Gaia* broadband magnitudes and now folds in the metallicity of a star. Finally, as photometrically variable sources are a significant source of spurious metallicities, we build a probabilistic variability model based on the photometric uncertainties of the *Gaia* sources.

Pushing both the Pristine *CaHK* magnitudes and the BP/RP-based synthetic $CaHK_{syn}$ magnitudes through the Pristine model, we build and distribute two catalogues of photometric metallicities for reliable stars ($E(B-V) < 0.5$ and $\delta CaHK < 0.1$ or $\delta CaHK_{syn} < 0.1$): the Pristine-*Gaia* synthetic catalogue and the Pristine DR1 catalogue for stars in common between Pristine and the BP/RP catalogue of *Gaia* DR3. The latter serves as the first data release of the Pristine survey. We show that both catalogues can be used to build reliable samples of metal-poor stars and are particularly well suited to track VMP and EMP stars. The two catalogues are also complementary: the Pristine-*Gaia* synthetic catalogue provides photometric metallicities over most of the sky but the Pristine DR1 catalogue (which is, by definition, limited to the Pristine footprint) contains significantly higher quality metallicities that go significantly fainter.

We conclude by showing test cases of how both catalogues of photometric metallicities can very efficiently determine the metallicity of MW GCs and make it easy to isolate extra-tidal cluster stars when the cluster's metallicity separates from the metallicity of MW field (more metal-rich) stars. Finally, we use the all-sky catalogue of Pristine-*Gaia* synthetic photometric metallicities to display the change of morphology of the MW with metallicity, from a flat, disk-like distribution at high metallicity to a spheroidal halo as the metallicity decreases.

The release of the BP/RP information in the *Gaia* DR3 catalogue truly opens a new age in the chemo-dynamical decomposition of the metal-poor MW, over the full MDF. And we can but look forward to the *Gaia* DR4 data that will provide BP/RP information at higher S/N and for fainter stars.

Data availability

Catalogues described in Tables 1, 2, and 3 are available at the CDS via anonymous ftp to `cdsarc.cds.unistra.fr` (`130.79.128.5`) or via `https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/692/A115`.

Acknowledgements. We are deeply grateful to the CFHT staff for performing the Pristine observations in queue mode, for their reactivity in adapting the schedule, and for answering our questions during the data-reduction process. We are also particularly grateful to the *Gaia* consortium for providing data of such wonderful quality and variety to the astronomy community. NFM, ZY, and RAI gratefully acknowledge support from the French National Research Agency (ANR) funded project "Pristine" (ANR-18-CE31-0017) along with funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No. 834148). ES, MB, and MM acknowledge funding through VIDI grant "Pushing Galactic Archaeology to its limits" (with project number VI.Vidi.193.093) which is funded by the Dutch Research Council (NWO). AA acknowledges support from the Herchel Smith Fellowship at the University of Cambridge and a Fitzwilliam College research fellowship supported by the Isaac Newton Trust. M.B. and P.M. acknowledge the support to activities related to the ESA/*Gaia* mission by the Italian Space Agency (ASI) through contract 2018-24-HH.0 and its addendum 2018-24-HH.1-2022 to the National Institute for Astrophysics (INAF). EFA acknowledge support from the Agencia Estatal de Investigación del Ministerio de Ciencia e Innovación (AEI-MCINN) under grant "At the forefront of Galactic Archaeology: evolution of the luminous and dark matter components of the Milky Way and Local Group dwarf galaxies in the *Gaia* era" with reference PID2020-118778GB-I00/10.13039/501100011033. EFA also acknowleges support from the 'María Zambrano' fellowship from the Universidad de La Laguna. JIGH and CAP acknowledge financial support from the Spanish Ministry of Science and Innovation (MICINN) project PID2020-117493GB-I00. We benefited from the International Space Science Institute (ISSI) in Bern, CH, thanks to the funding of the team "Pristine". This research was supported by the International Space Science Institute (ISSI) in Bern, through ISSI International Team project 540 (The Early Milky Way). DA also acknowledges financial support from the Spanish Ministry of Science and Innovation (MICINN) under the 2021 Ramón y Cajal program MICINN RYC2021-032609. G.B. acknowledges support from the Agencia Estatal de Investigación del Ministerio de Ciencia en Innovación (AEI-MICIN) and the European Regional Development Fund (ERDF) under grant number AYA2017-89076-P, the AEI under grant number CEX2019-000920-S and the AEI-MICIN under grant number PID2020-118778GB-I00/10.13039/501100011033. Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is `www.sdss.org`. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University. Guoshoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope LAMOST) is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences. Author contributions: NFM and ES are co-PIs of the Pristine survey, designed the project and led the writing of the paper. NFM handled the data reduction and ES built the metallicity model, the training set (with KV's help) and worked on the comparisons with external data sets. ZY took care of the calibration. MF coordinated the $CaHK_{syn}$ calculations with FDA. AA-A developed the extinction calculations with HZ's help. AA-A, AE, and AV updated the code of the metallicity model to its public version. FG, MM, and SR developed the comparisons with globular clusters, carbon-rich stars, and UMP stars. RA, MB, and PM provided their *Gaia* expertise. All other authors are active members of the Pristine collaboration and took part in the writing and editing of the paper.

References

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
Aguado, D. S., Allende Prieto, C., González Hernández, J. I., et al. 2016, A&A, 593, A10

Aguado, D. S., Youakim, K., González Hernández, J. I., et al. 2019, MNRAS, 490, 2241
Almeida-Fernandes, F., SamPedro, L., Herpich, F. R., et al. 2022, MNRAS, 511, 4590
Alvarez, R., & Plez, B. 1998, A&A, 330, 1109
Andrae, R., Fouesneau, M., Sordo, R., et al. 2023a, A&A, 674, A27
Andrae, R., Rix, H.-W., & Chandra, V. 2023b, ApJS, 267, 8
Arentsen, A., Starkenburg, E., Martin, N. F., et al. 2020a, MNRAS, 496, 4964
Arentsen, A., Starkenburg, E., Martin, N. F., et al. 2020b, MNRAS, 491, L11
Arentsen, A., Placco, V. M., Lee, Y. S., et al. 2022, MNRAS, 515, 4082
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Baumgardt, H., & Vasiliev, E. 2021, MNRAS, 505, 5957
Beers, T. C., & Christlieb, N. 2005, ARA&A, 43, 531
Beers, T. C., Preston, G. W., & Shectman, S. A. 1992, AJ, 103, 1987
Beers, T. C., Carollo, D., Ivezić, Ž., et al. 2012, ApJ, 746, 34
Bellazzini, M., Massari, D., De Angeli, F., et al. 2023, A&A, 674, A194
Belokurov, V. 2013, NewAR, 57, 100
Belokurov, V., & Kravtsov, A. 2022, MNRAS, 514, 689
Belokurov, V., Erkal, D., Deason, A. J., et al. 2017, MNRAS, 466, 4711
Bergemann, M., Sesar, B., Cohen, J. G., et al. 2018, Nature, 555, 334
Bessell, M. S. 1977, PASA, 3, 144
Bidelman, W. P., & MacConnell, D. J. 1973, AJ, 78, 687
Bond, H. E. 1970, ApJS, 22, 117
Bonifacio, P., Caffau, E., Sestito, F., et al. 2019, MNRAS, 487, 3797
Bonifacio, P., Monaco, L., Salvadori, S., et al. 2021, A&A, 651, A79
Boulade, O., Charlot, X., Abbon, P., et al. 2003, SPIE Conf. Ser., 4841, 72
Caffau, E., Bonifacio, P., Sbordone, L., et al. 2013, A&A, 560, A71
Caffau, E., Bonifacio, P., Starkenburg, E., et al. 2017, Astron. Nachr., 338, 686
Caffau, E., Bonifacio, P., Sbordone, L., et al. 2020, MNRAS, 493, 4677
Cantat-Gaudin, T., Fouesneau, M., Rix, H.-W., et al. 2023, A&A, 669, A55
Carney, B. W., Laird, J. B., Latham, D. W., & Aguilar, L. A. 1996, AJ, 112, 668
Carrasco, J. M., Weiler, M., Jordi, C., et al. 2021, A&A, 652, A86
Carretta, E., Bragaglia, A., Gratton, R., D'Orazi, V., & Lucatello, S. 2009, A&A, 508, 695
Casagrande, L., Lin, J., Rains, A. D., et al. 2021, MNRAS, 507, 2684
Castelli, F., & Kurucz, R. L. 2003, in Modelling of Stellar Atmospheres, 210, eds. N. Piskunov, W. W. Weiss, & D. F. Gray, A20
Cenarro, A. J., Moles, M., Cristóbal-Hornillos, D., et al. 2019, A&A, 622, A176
Chiti, A., Frebel, A., Mardini, M. K., et al. 2021, ApJS, 254, 31
Christlieb, N., Bessell, M. S., Beers, T. C., et al. 2002, Nature, 419, 904
Christlieb, N., Schörck, T., Frebel, A., et al. 2008, A&A, 484, 721
Da Costa, G. S., Bessell, M. S., Mackey, A. D., et al. 2019, MNRAS, 489, 5900
De Angeli, F., Weiler, M., Montegriffo, P., et al. 2023, A&A, 674, A2
El-Badry, K., Quataert, E., Weisz, D. R., Choksi, N., & Boylan-Kolchin, M. 2019, MNRAS, 482, 4528
Fan, Z., Zhao, G., Wang, W., et al. 2023, ApJS, 268, 9
Feltzing, S., Eriksson, K., Kleyna, J., & Wilkinson, M. I. 2009, A&A, 508, L1
Fernández-Alvar, E., Kordopatis, G., Hill, V., et al. 2021, MNRAS, 508, 1509
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fouesneau, M., Andrae, R., Sordo, R., & Dharmawardena, T. 2022, dustapprox https://github.com/mfouesneau/dustapprox
Frebel, A., & Norris, J. E. 2015, ARA&A, 53, 631
Frebel, A., Aoki, W., Christlieb, N., et al. 2005, Nature, 434, 871
Frebel, A., Norris, J. E., Gilmore, G., & Wyse, R. F. G. 2016, ApJ, 826, 110
Gaia Collaboration (Montegriffo, P., et al.) 2023a, A&A, 674, A33
Gaia Collaboration (Vallenari, A., et al.) 2023b, A&A, 674, A1
Galarza, C. A., Daflon, S., Placco, V. M., et al. 2022, A&A, 657, A35
Gavras, P., Rimoldini, L., Nienartowicz, K., et al. 2023, A&A, 674, A22
Gilmore, G., Norris, J. E., Monaco, L., et al. 2013, ApJ, 763, 61
Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, 951
Hansen, C. J., El-Souri, M., Monaco, L., et al. 2018, ApJ, 855, 83
Harris, W. E. 1996, AJ, 112, 1487
Helmi, A. 2020, ARA&A, 58, 205
Hernitschek, N., Schlafly, E. F., Sesar, B., et al. 2016, ApJ, 817, 73
Huang, Y., Beers, T. C., Wolf, C., et al. 2022, ApJ, 925, 164
Huang, Y., Beers, T. C., Yuan, H., et al. 2023, ApJ, 957, 65
Ibata, R. A., Irwin, M. J., Lewis, G. F., Ferguson, A. M. N., & Tanvir, N. 2003, MNRAS, 340, L21
Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017a, ApJ, 848, 128
Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017b, ApJ, 848, 129
Irwin, M., & Lewis, J. 2001, New Astron. Rev., 45, 105
Irwin, M. J., Lewis, J., Hodgkin, S., et al. 2004, SPIE Conf. Ser., 5493, 411
Ishigaki, M. N., Aoki, W., Arimoto, N., & Okamoto, S. 2014, A&A, 562, A146
Ivezić, Ž., Sesar, B., Jurić, M., et al. 2008, ApJ, 684, 287
Jin, S., Trager, S. C., Dalton, G. B., et al. 2024, MNRAS, 530, 2688
Keller, S. C., Bessell, M. S., Frebel, A., et al. 2014, Nature, 506, 463
Kielty, C. L., Venn, K. A., Sestito, F., et al. 2021, MNRAS, 506, 1438
Klessen, R. S., & Glover, S. C. O. 2023, ARA&A, 61, 65
Kordopatis, G., Schultheis, M., McMillan, P. J., et al. 2023, A&A, 669, A104
Laporte, C. F. P., Belokurov, V., Koposov, S. E., Smith, M. C., & Hill, V. 2020, MNRAS, 492, L61
Lardo, C., Mashonkina, L., Jablonka, P., et al. 2021, MNRAS, 508, 3068
Li, H., Tan, K., & Zhao, G. 2018, ApJS, 238, 16
Li, J., Xue, X.-X., Liu, C., et al. 2021, ApJ, 910, 46
Li, H., Aoki, W., Matsuno, T., et al. 2022, ApJ, 931, 147
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Lombardo, L., Bonifacio, P., Caffau, E., et al. 2023, MNRAS, 522, 4815
López-Sanjuan, C., Vázquez Ramió, H., Xiao, K., et al. 2024, A&A, 683, A29
Lucchesi, R., Lardo, C., Jablonka, P., et al. 2022, MNRAS, 511, 1004
Magnier, E. A., & Cuillandre, J. C. 2004, PASP, 116, 449
Maiolino, R., Übler, H., Perna, M., et al. 2024, A&A, 687, A67
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Marino, A. F., Da Costa, G. S., Casey, A. R., et al. 2019, MNRAS, 485, 5153
Martin, N. F., Venn, K. A., Aguado, D. S., et al. 2022, Nature, 601, 45
Montegriffo, P., De Angeli, F., Andrae, R., et al. 2023, A&A, 674, A3
Morganson, E., Conn, B., Rix, H.-W., et al. 2016, ApJ, 825, 140
Newberg, H. J., Yanny, B., Rockosi, C., et al. 2002, ApJ, 569, 245
Nordlander, T., Bessell, M. S., Da Costa, G. S., et al. 2019, MNRAS, 488, L109
Onken, C. A., Wolf, C., Bessell, M. S., et al. 2019, PASA, 36, e033
Placco, V. M., Almeida-Fernandes, F., Arentsen, A., et al. 2022, ApJS, 262, 8
Plez, B. 2008, Phys. Scr. Vol. T, 133, 014003
Recio-Blanco, A., de Laverny, P., Palicio, P. A., et al. 2023, A&A, 674, A29
Regnault, N., Conley, A., Guy, J., et al. 2009, A&A, 506, 999
Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3
Rix, H.-W., Chandra, V., Andrae, R., et al. 2022, ApJ, 941, 45
Ryan, S. G., & Norris, J. E. 1991, AJ, 101, 1865
Sakari, C. M., Placco, V. M., Farrell, E. M., et al. 2018, ApJ, 868, 110
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schörck, T., Christlieb, N., Cohen, J. G., et al. 2009, A&A, 507, 817
Sestito, F., Longeard, N., Martin, N. F., et al. 2019, MNRAS, 484, 2166
Sestito, F., Martin, N. F., Starkenburg, E., et al. 2020, MNRAS, 497, L7
Sestito, F., Venn, K. A., Arentsen, A., et al. 2023, MNRAS, 518, 4557
Slater, C. T., Bell, E. F., Schlafly, E. F., et al. 2014, ApJ, 791, 9
Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
Soubiran, C., Le Campion, J.-F., Brouillet, N., & Chemin, L. 2016, A&A, 591, A118
Starkenburg, E., Martin, N., Youakim, K., et al. 2017a, MNRAS, 471, 2587
Starkenburg, E., Oman, K. A., Navarro, J. F., et al. 2017b, MNRAS, 465, 2212
Starkenburg, E., Aguado, D. S., Bonifacio, P., et al. 2018, MNRAS, 481, 3838
Suda, T., Katsuta, Y., Yamada, S., et al. 2008, PASJ, 60, 1159
Thomas, G. F., Jensen, J., McConnachie, A., et al. 2020, ApJ, 902, 89
Tinsley, B. M. 1980, Fund. Cosmic Phys., 5, 287
Vasiliev, E., & Baumgardt, H. 2021, MNRAS, 505, 5978
Venn, K. A., Kielty, C. L., Sestito, F., et al. 2020, MNRAS, 492, 3241
Welch, B., Coe, D., Diego, J. M., et al. 2022a, Nature, 603, 815
Welch, B., Coe, D., Zackrisson, E., et al. 2022b, ApJ, 940, L1
Welsh, L., Cooke, R., Fumagalli, M., & Pettini, M. 2023, MNRAS, 525, 527
Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001
Witten, C. E. C., Aguado, D. S., Sanders, J. L., et al. 2022, MNRAS, 516, 3254
Wolf, C., Onken, C. A., Luvaul, L. C., et al. 2018, PASA, 35, e010
Xu, S., Yuan, H., Zhang, R., et al. 2022, ApJS, 263, 29
Xylakis-Dornbusch, T., Christlieb, N., Lind, K., & Nordlander, T. 2022, A&A, 666, A58
Xylakis-Dornbusch, T., Christlieb, N., Hansen, T. T., et al. 2024, A&A, 687, A177
Yang, L., Yuan, H., Xiang, M., et al. 2022, A&A, 659, A181
Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377
Yong, D., Da Costa, G. S., Bessell, M. S., et al. 2021, MNRAS, 507, 4102
York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579
Youakim, K., Starkenburg, E., Aguado, D. S., et al. 2017, MNRAS, 472, 2963
Youakim, K., Starkenburg, E., Martin, N. F., et al. 2020, MNRAS, 492, 4986
Yuan, H., Liu, X., Xiang, M., Huang, Y., & Chen, B. 2015, ApJ, 803, 13
Yuan, Z., Martin, N. F., Ibata, R. A., et al. 2022, MNRAS, 514, 1664
Zhang, X., Green, G. M., & Rix, H.-W. 2023, MNRAS, 524, 1855
Zhao, G., Chen, Y.-Q., Shi, J.-R., et al. 2006, Chinese J. Astron. Astrophys., 6, 265
Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, Res. Astron. Astrophys., 12, 723

Appendix A: Extinction coefficients

Tables A.1 and A.2 list the coefficients of the polynomial relations used to fit the synthetic photometry from `dustapprox`, as described in section 4 and equation 2.

Table A.1. Polynomial coefficients for k_f (Equation 2) for giants. Only coefficients larger than 0.001 are listed.

	factor	G	G_{BP}	G_{RP}	$CaHK$	A_0/A_V
a_0	1	−0.0897	0.3441	0.4009	1.5409	1.1753
a_1	T	1.9108	1.4676	0.5364	0.0377	−0.5199
a_2	A_0	0.0307	0.0302	−0.0059		0.0045
a_3	[Fe/H]	0.0113	0.0128	−0.0027	−0.0023	0.0059
a_4	T^2	−1.2696	−0.9191	−0.3920	−0.0252	0.3903
a_6	[Fe/H]^2	−0.0077	−0.0087			0.0020
a_7	T A_0	−0.0921	−0.0778	−0.0019		
a_8	T [Fe/H]	−0.0547	−0.0728	0.0042	0.0028	−0.0012
a_9	A_0 [Fe/H]	0.0023	0.0025			
a_{10}	T^3	0.3183	0.2122	0.1005	0.0060	−0.1017
a_{13}	T^2 A_0	0.0284	0.0274			
a_{15}	T^2 [Fe/H]	0.0300	0.0405	−0.0011		−0.0017
a_{16}	T [Fe/H]^2	0.0035	0.0038			−0.0011
σ		0.0017	0.0019	0.0005	0.0001	0.0005

Table A.2. Polynomial coefficients for k_f (Equation 2) for dwarfs. Only coefficients larger than 0.001 are listed.

	factor	G	G_{BP}	G_{RP}	$CaHK$	A_0/A_V
a_0	1	0.2554	0.6765	0.4620	1.5524	1.0985
a_1	T	0.9525	0.6622	0.3391	0.0091	−0.2832
a_2	A_0	0.0113	0.0062	−0.0061		0.0042
a_3	[Fe/H]	−0.0312	0.0107	−0.0214		0.0076
a_4	T^2	−0.4075	−0.2720	−0.1963	−0.0022	0.1637
a_5	$A_0{}^2$	0.0017	0.0011			
a_6	[Fe/H]^2	−0.0065	−0.0027	−0.0021		
a_7	T A_0	−0.0601	−0.0381	−0.0013		
a_8	T [Fe/H]	0.0195	−0.0475	0.0305		−0.0071
a_9	A_0 [Fe/H]	0.0024	0.0019			
a_{10}	T^3	0.0657	0.0393	0.0395		−0.0328
a_{13}	T^2 A_0	0.0158	0.0113			
a_{15}	T^2 [Fe/H]		0.0229	−0.0100		0.0016
a_{16}	T [Fe/H]^2	0.0026		0.0015		
σ		0.0022	0.0019	0.0012	0.0001	0.0008

Appendix B: Comparison of photometric metallicities with the training set for dwarf stars

Figures B.1 and B.2 are the same as Figure 16, but for the dwarf and main-sequence turnoff stars of the training sample, respectively, and color-coded by the temperature of a star. This comparison with the training sample shows somewhat poorer quality than for the giant stars but the degrading performance stems mainly from the main-sequence turnoff stars. This is understandable as these fairly hot stars have narrow spectral Ca H & K features that translate into a loss of metallicity resolution for a fixed photometric uncertainty. As shown in Figure 9, a fixed difference in $[Fe/H]_{phot}$ is significantly harder to measure as it corresponds to a smaller $CaHK$ difference for these blue stars than it is for a redder star. It is interesting to note on the other hand that the Andrae et al. (2023a) GSP-Phot values have less catastrophic outliers when compared to our dwarf training sample than for the giants, possibly because of the additional information brought by the parallax in the GSP-Phot analysis.

Appendix C: Comparison of the variability model with the *Gaia* variability classes

Figure C.1 presents the distribution of the different classes of variable sources in *Gaia* Gavras et al. (2023) in the variability space of Figure 8. The proposed quality cut that keeps stars with $P_{var} < 0.3$ does indeed remove most of these identified variable stars, irrespective of the variable class.

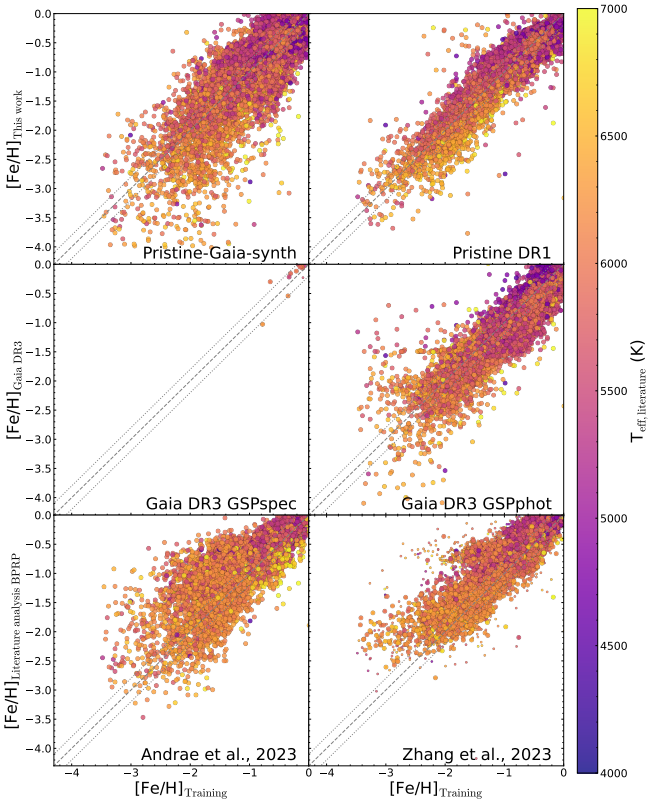

Fig. B.1. Comparison of various metallicity catalogues based on *Gaia* DR3 with the stars in the training sample with $\log g > 3.9$ corresponding to the main-sequence stars. After the comparison with the Pristine-*Gaia* synthetic catalogue, and the Pristine DR1 catalogue, the panels show the comparisons for the *Gaia* GSP-Spec metallicities based on the *Gaia* RVS spectra (Recio-Blanco et al. 2023, smaller circles correspond to stars flagged as unreliable in this catalogue, following the same criteria as in their Figure 26), the *Gaia* GSP-Phot catalogue based on the BP/RP information (Andrae et al. 2023a), the XGBoost algorithm that uses the BP/RP information and some additional non-*Gaia* photometry (Andrae et al. 2023b), the BP/RP-based metallicities from Zhang et al. (2023, smaller circles are flagged results). In all panels, the color encodes the temperature of the star from the training sample. The dashed line corresponds to the 1-to-1 line and the dotted lines represent offsets of 0.2 dex.

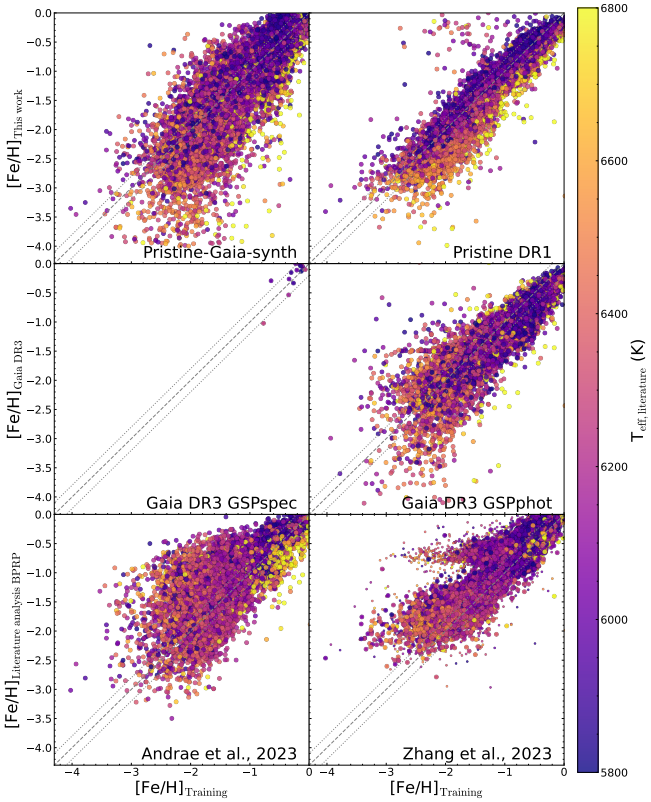

Fig. B.2. Similar to Figure B.1, but for training sample turn-off stars only (defined as $T_{\rm eff} > 5800$ K, regardless of their $\log g$). A clear trend with temperature can be seen in the Pristine DR1 catalogue, even in this relatively small temperature range. Our photometric metallicities estimates degrade for hotter stars.

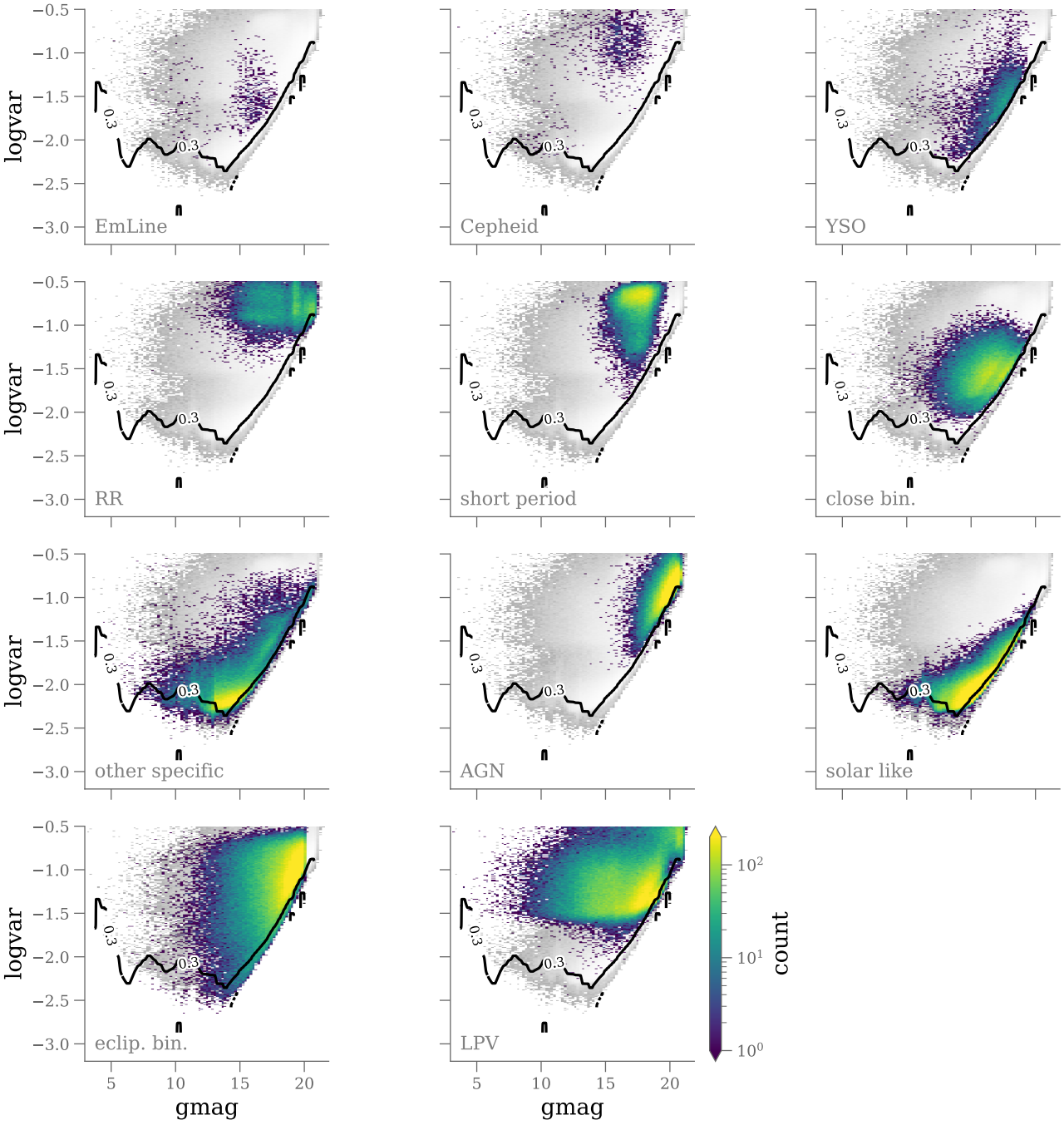

Fig. C.1. Distribution of the main *Gaia* DR3 variable classes, as defined by Gavras et al. (2023), in the variability space of Figure 8 used to define $P_{\rm var}$. All panels show the distribution of the variables identified in *Gaia* DR3 (`phot_variable_flag = 'VARIABLE'`) in gray and the density distribution of a given class of variables, as labeled in each panel. We only show classes with at least 5,000 identified sources and we overlay our $P_{\rm var} = 0.3$ threshold for comparison. The proposed cut effectively rejects the most obvious classes of variables.