A&A 674, A29 (2023)
https://doi.org/10.1051/0004-6361/202243750
© The Authors 2023

Astronomy & Astrophysics
Special issue

Gaia Data Release 3

Gaia Data Release 3

Analysis of RVS spectra using the General Stellar Parametriser from spectroscopy

A. Recio-Blanco[1,*], P. de Laverny[1], P. A. Palicio[1], G. Kordopatis[1], M. A. Álvarez[2], M. Schultheis[1], G. Contursi[1], H. Zhao[1], G. Torralba Elipe[2], C. Ordenovic[1], M. Manteiga[3], C. Dafonte[2], I. Oreshina-Slezak[1], A. Bijaoui[1], Y. Frémat[4], G. Seabroke[5], F. Pailler[6], E. Spitoni[1], E. Poggio[1], O. L. Creevey[1], A. Abreu Aramburu[7], S. Accart[6], R. Andrae[8], C. A. L. Bailer-Jones[8], I. Bellas-Velidis[9], N. Brouillet[10], E. Brugaletta[11], A. Burlacu[12], R. Carballo[13], L. Casamiquela[10,14], A. Chiavassa[1], W. J. Cooper[15,16], A. Dapergolas[9], L. Delchambre[17], T. E. Dharmawardena[8], R. Drimmel[16], B. Edvardsson[18], M. Fouesneau[8], D. Garabato[2], P. García-Lario[19], M. García-Torres[20], A. Gavel[18], A. Gomez[2], I. González-Santamaría[2], D. Hatzidimitriou[21,9], U. Heiter[18], A. Jean-Antoine Piccolo[6], M. Kontizas[21], A. J. Korn[18], A. C. Lanzafame[11,22], Y. Lebreton[23,24], Y. Le Fustec[12], E. L. Licata[16], H. E. P. Lindstrøm[16,25,26], E. Livanou[21], A. Lobel[4], A. Lorca[27], A. Magdaleno Romeo[12], F. Marocco[28], D. J. Marshall[29], N. Mary[30], C. Nicolas[6], L. Pallas-Quintela[2], C. Panem[6], B. Pichon[1], F. Riclet[6], C. Robin[30], J. Rybizki[8], R. Santoveña[2], A. Silvelo[2], R. L. Smart[16], L. M. Sarro[31], R. Sordo[32], C. Soubiran[10], M. Süveges[33], A. Ulla[34], A. Vallenari[32], J. Zorec[35], E. Utrilla[27], and J. Bakker[7]

(Affiliations can be found after the references)

Received 10 April 2022 / Accepted 17 May 2022

ABSTRACT

Context. The chemo-physical parametrisation of stellar spectra is essential for understanding the nature and evolution of stars and of Galactic stellar populations. A worldwide observational effort from the ground has provided, in one century, an extremely heterogeneous collection of chemical abundances for about two million stars in total, with fragmentary sky coverage.

Aims. This situation is revolutionised by the *Gaia* third data release (DR3), which contains the parametrisation of Radial Velocity Spectrometer (RVS) data performed by the General Stellar Parametriser-spectroscopy, GSP-Spec, module. Here we describe the parametrisation of the first 34 months of *Gaia* RVS observations.

Methods. GSP-Spec estimates the chemo-physical parameters from combined RVS spectra of single stars, without additional inputs from astrometric, photometric, or spectro-photometric BP/RP data. The main analysis workflow described here, MatisseGauguin, is based on projection and optimisation methods and provides the stellar atmospheric parameters; the individual chemical abundances of N, Mg, Si, S, Ca, Ti, Cr, Fe I, Fe II, Ni, Zr, Ce and Nd; the differential equivalent width of a cyanogen line; and the parameters of a diffuse interstellar band (DIB) feature. Another workflow, based on an artificial neural network (ANN) and referred to with the same acronym, provides a second set of atmospheric parameters that are useful for classification control. For both workflows, we implement a detailed quality flag chain considering different error sources.

Results. With about 5.6 million stars, the *Gaia* DR3 GSP-Spec all-sky catalogue is the largest compilation of stellar chemo-physical parameters ever published and the first one from space data. Internal and external biases have been studied taking into account the implemented flags. In some cases, simple calibrations with low degree polynomials are suggested. The homogeneity and quality of the estimated parameters enables chemo-dynamical studies of Galactic stellar populations, interstellar extinction studies from individual spectra, and clear constraints on stellar evolution models. We highly recommend that users adopt the provided quality flags for scientific exploitation.

Conclusions. The *Gaia* DR3 GSP-Spec catalogue is a major step in the scientific exploration of Milky Way stellar populations. It will be followed by increasingly large and higher quality catalogues in future data releases, confirming the *Gaia* promise of a new Galactic vision.

Key words. stars: fundamental parameters – stars: abundances – methods: data analysis – Galaxy: stellar content

1. Introduction

The chemo-physical characterisation of stars is at the core of stellar physics and Galactic studies, but also, through the analysis of unresolved stellar populations, of extragalactic physics. Stellar spectra encode a wealth of information that we have now learned to decrypt. The light emitted by a star is absorbed by the atoms and molecules present in its own atmosphere. This creates spectral absorption lines whose profiles depend on the physical properties of the star and the abundances of the different absorbing chemical species. Our understanding of stellar spectra, used to decode the enclosed information on the nature of stars, relies on a complex and extensive theoretical framework, including (among others) nuclear, atomic, and molecular physics, stellar atmosphere physics, element nucleosynthesis, and radiative transfer theory.

Before development of the necessary background theoretical knowledge, stellar spectra motivated the definition of stellar types and luminosity classes. These were the fruit of a classification effort categorising stars based on the identification and strength of their spectral features. Therefore, the chemo-physical parametrisation of stellar spectra has its roots in the large observational campaigns of the beginning of the 20th century (cf. Cannon & Pickering 1918) and the seminal works leading to the Morgan-Keenan classification (cf. Morgan et al. 1943).

* Corresponding author: A. Recio-Blanco,
e-mail: Alejandra.Recio-Blanco@oca.eu

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

The development of CCD detectors and, more recently, of multiobject facilities has resulted in the ability of even small telescopes to acquire large numbers of stellar spectra. Pioneering projects such as the Geneva Copenhaguen Survey (Nordström et al. 2004), RAVE (Steinmetz et al. 2006), and SEGUE (Yanny et al. 2009), followed by archival parametrisation projects like AMBRE (de Laverny et al. 2012) and a worldwide observational effort illustrated by the *Gaia*-ESO Survey (Gilmore et al. 2012), LAMOST (Zhao et al. 2012), APOGEE (Majewski et al. 2017), and GALAH (Martell et al. 2017) characterise era of Galactic spectroscopic surveys. In parallel to the above-mentioned ground-based efforts, the design and preparation of the *Gaia* space mission, including the Radial Velocity Spectrometer (RVS; for a historical overview see Katz et al. 2004; Cropper et al. 2018, and references therein), opened new horizons in the observation of Milky Way stellar populations, and delivered on the promise of an unprecedentedly extensive spectroscopic survey (Wilkinson et al. 2005).

This rapid evolution of observational capabilities brought to the fore the need for automated parametrisation tools, enabling fast and homogeneous processing of extensive data sets. Once again, pioneering efforts followed the trail of the first spectroscopic surveys (e.g. Allende Prieto et al. 2006, among others) and the *Gaia* space project (Recio-Blanco et al. 2006). A variety of mathematical approaches have been developed and applied since then. These include different optimisation, projection, and classification methods used as part of model-driven or data-driven approaches (see e.g. Recio-Blanco 2014; Allende Prieto 2016; Jofré et al. 2019, and references therein).

Gaia observations started on 25 July 2014. The wavelength range covered by the RVS is [846−870] nm, and its medium resolving power is $R = \lambda/\Delta\lambda \sim 11\,500$ (Cropper et al. 2018). The present work describes the parametrisation of the first 34 months of *Gaia* RVS observations by the General Stellar Parametriser from spectroscopy (GSP-Spec) module of the Astrophysical parameters inference system (Apsis, Creevey et al. 2023). Apsis is the heritage of the previously described scientific pathway and the outcome of a long-term effort: from the development of innovative methodologies (Recio-Blanco et al. 2006; Bijaoui et al. 2010, 2012; Kordopatis et al. 2011) to their integration into the *Gaia* Data Processing and Analysis Consortium (DPAC) framework (Bailer-Jones et al. 2013), their tailoring to *Gaia*/RVS prelaunch characteristics (Recio-Blanco et al. 2016; Dafonte et al. 2016) and their first publication as part of the *Gaia* Data Release 3 (DR3; Gaia Collaboration 2023c). This effort results in the largest catalogue of stellar chemo-physical parameters ever published, which is simultaneously the first of its kind from a space spectroscopic survey and with all-sky coverage.

Section 2 presents the GSP-Spec goals and output parameters. This is followed by a description of the input *Gaia* RVS data (Sect. 3) and reference synthetic spectra grids (Sect. 4). The spectral line selection used for individual abundance analysis is explained in Sect. 5. The two GSP-Spec analysis workflows, MatisseGauguin and artificial neural network (ANN), are described in detail in Sects. 6 and 7, respectively. Section 8 presents the performed validation of GSP-Spec outputs as part of *Gaia* DR3 operations, and defines the implemented quality flags. Section 9 is devoted to the comparison of GSP-Spec results to literature data and suggested calibrations. Finally, in Sects. 10 and 11 we present the GSP-Spec catalogue and our conclusions.

2. Goals and outputs of GSP-Spec

The GSP-Spec module implements a purely spectroscopic treatment. It estimates the stellar chemo-physical parameters from combined RVS spectra of single stars. No additional information is considered from astrometric, photometric, or spectro-photometric BP/RP data[1].

In particular, GSP-Spec estimates (i) the stellar effective temperature $T_{\rm eff}$, reported as `teff_gspspec`; (ii) the stellar surface gravity expressed in logarithm $\log(g)$[2], reported as `logg_gspspec`; (iii) the stellar mean metallicity [M/H][3] defined as the solar-scaled abundances of all elements heavier than He and reported as `mh_gspspec`; (iv) the enrichment of α-elements[4] with respect to iron ([α/Fe]), reported as `alphafe_gspspec`; (v) the individual abundances of 13 chemical species ([N/Fe], [Mg/Fe], [Si/Fe], [S/Fe], [Ca/Fe], [Ti/Fe], [Cr/Fe], [Fe I/M], [Fe II/M][5], [Ni/Fe], [Zr/Fe], [Ce/Fe], [Nd/Fe], (reported as `Xfe_gspspec` or `Xm_gspspec`, with X being the chemical species), including the number of used spectral lines (`Xfe_gspspec_nlines`) and the line-to-line scatter (`Xfe_gspspec_linescatter`); (vi) a CN differential abundance proxy reported as `cnew_gspspec`; (vii) the equivalent width (EW) and fitting parameters of the diffuse interstellar band (DIB) at 862 nm, reported as `dibew_gspspec` `dibp1_gspspec` and `dibp2_gspspec`; (viii) a goodness-of-fit (gof) over the entire spectral range reported as `logchisq_gspspec`; and (ix) a quality flag chain (the use of which is highly recommended) considering different error sources affecting the output parameters and reported as `flags_gspspec`.

Two different procedures, MatisseGauguin and ANN, described in Recio-Blanco et al. (2016), are applied to estimate the stellar atmospheric parameters ($T_{\rm eff}$, $\log(g)$, [M/H], and [α/Fe]). Individual chemical abundances and DIB parameters are estimated by the GAUGUIN algorithm (Recio-Blanco et al. 2016) and Gaussian fitting methods (Zhao et al. 2021), respectively, and are only produced by the MatisseGauguin analysis workflow[6]. The goodness-of-fit and the quality flag chain are provided for both the MatisseGauguin and ANN parametrisation. It is worth noting that, for each star, parameter uncertainties are estimated from 50 Monte-Carlo realisations[7] of its RVS spectrum, considering flux uncertainties. For each realisation, a new complete analysis is implemented, including atmospheric parameters, individual chemical abundances, and CN and DIB parameters. From this analysis, upper and lower confidence values are respectively provided from the 84th and 16th quantiles of the resulting parameter and abundance distributions and reported with the suffix `_upper` and `_lower`, respectively (cf. Sect. 6.7).

The DR3 GSP-Spec analysis is available through two archive tables: the MatisseGauguin workflow provides 101 fields for 5 594 205 stars in the *AstrophysicalParameters* table, and the

[1] A separate Apsis module, the GSP from photometry is in charge of the stellar parametrisation from BP/RP data, using constraints from astrometric and stellar isochrones (Andrae et al. 2023).
[2] g being in cm s^{-2}.
[3] In the following, we adopt the standard abundance notation for a given element X: $[{\rm X/H}] = \log({\rm X/H})_\star - \log({\rm X/H})_\odot$, where (X/H) is the abundance by number, and $\log\epsilon(X) \equiv \log({\rm X/H}) + 12$.
[4] O, Ne, Mg, Si, S, Ar, Ca, and Ti are considered as α-elements and vary in lockstep.
[5] Fe I and Fe II abundance enhancements with respect to the mean metallicity are estimated and respectively called *fem_gspspec* and *feIIm_gspspec* in the *AstrophysicalParameters* table.
[6] It is worth mentioning that MatisseGauguin algorithms have been conceived assuming a white Gaussian noise framework.
[7] This number of realisations has been optimised through simulations to ensure a good sampling of the associated parameter distributions, and taking into account the computation time allocated to GSP-Spec. We note that this Monte-Carlo procedure does not take into account uncertainties in the radial velocity correction, which have been considered through analysis flags (cf. Sect. 8.2).

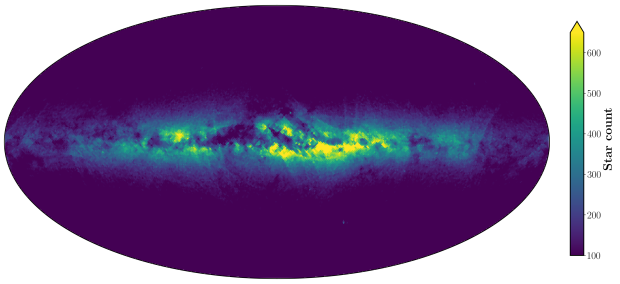


Fig. 1. Global all-sky spatial density distribution of all the GSP-Spec parametrised stars. This HEALPix map (Górski et al. 2005) in Galactic coordinates has a spatial resolution of 0.46° and at least 100 stars are contained in each resolution element.

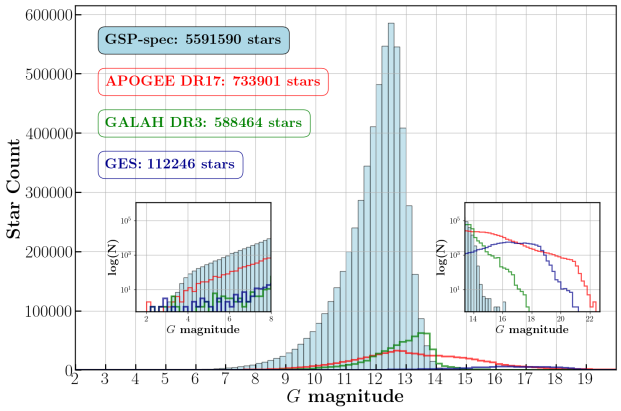


Fig. 2. *Gaia*-magnitude distribution of all the GSP-Spec parametrised stars. The APOGEE, GALAH, and GES magnitude distributions are shown for comparison in red, green, and blue, respectively.

ANN workflow provides 13 fields for 5 524 387 stars in the *AstrophysicalParametersSupp* table with an added `_ann` suffix in the parameter names. Figure 1 shows the spatial distribution in (l, b) Galactic coordinates of all the GSP-Spec parametrised stars. One can see that most stars are located close to the Galactic plane, as expected, although larger latitudes are still very well sampled with at least 100 stars per resolution element. The small-scale structures close to the Galactic plane are caused by interstellar absorption. Figure 2 illustrates the magnitude distribution of all the GSP-Spec parametrised stars in the G-band. The parametrised stars can be seen to cover a large range of magnitudes, starting from the brightest objects (about 4000 of them have $G < 6$, i.e. about two-thirds of the sky visible to the naked eye) to the faintest ones up to $G \sim 16$ (more than half a million and ∼100 000 have $G > 13$ and $G > 13.5$, respectively). This very high number statistics can also be appreciated for the magnitude bins with the highest number of stars. For instance, the bin $12.4 \leq G$-mag<12.6 contains as many stars as published by the large ground-based spectroscopic survey GALAH. For comparison, Fig. 2 also shows the magnitude distributions of the largest ground-based spectroscopic surveys whose spectral resolution is larger than the RVS one: APOGEE-DR17 (Abdurro'uf et al. 2022), GALAH-DR3 (Buder et al. 2021), and the *Gaia*-ESO Survey (GES) (Gilmore et al. 2022; Randich et al. 2022). The highest number statistics of the *Gaia* GSP-Spec catalogue is achieved for $G < 13.6$ mag. For magnitudes fainter than about $G \sim 14.0$, APOGEE dominates with about 100 000 stars. GES also complements *Gaia* DR3 data at such fainter magnitudes with several tens of thousands, while GALAH has only a few thousand stars fainter than this data release of GSP-Spec. We note that the number of stars parametrised by GSP-Spec will strongly increase with the next *Gaia* data releases, being about a factor ten larger in DR4 as a result of the spectra signal-to-noise ratio (S/N) increase with repeated observations (and hence with observing time).

The GSP-Spec analysis module is coded in Java following DPAC requirements, and is executed at the Data Processing Centre C hosted by the Centre National d'Etudes Spatiales (CNES) in Toulouse, France. During DR3 operations, about 6.9 million spectra were processed by the module in ∼110 000 h, spread over ∼2100 cores (execution time of around 130 h, all the cores not being fully dedicated to GSP-Spec). The necessary RAM to run GSP-Spec is 25–30 GB. Therefore, the total execution time to derive the two sets (MatisseGauguin and ANN) of four atmospheric parameters, the 13 individual chemical abundances, the CN differencial abundance proxy, the DIB fitting parameters, and all the associated uncertainties and goodness of fit is about one second per spectrum for one Monte-Carlo realisation of the noise.

An illustration of the GSP-Spec parameterisation was published as a *Gaia* Image of the Week[8]. GSP-Spec parameters are also used in the *Gaia* DR3 chemical cartography analysis (Gaia Collaboration 2023a,b). GSP-Spec is the main spectroscopic parametriser module of the *Gaia* Apsis pipeline, independent of other modules, and feeds some of them executed afterwards in the module chain. The GSP-Spec methodology was largely tested on ground-based spectroscopic observations resulting from different projects, such as RAVE (Steinmetz et al. 2006), GES (Recio-Blanco et al. 2014), and AMBRE (de Laverny et al. 2013), among others.

3. Input *Gaia* RVS data

As input, GSP-Spec uses combined RVS spectra (averaged over multiple transits) and their flux uncertainties per wavelength pixel (wlp) over the 846–870 nm spectral domain. Prior to the GSP-Spec module operations, the stellar radial velocity ($V_{\rm Rad}$, Katz et al. 2023) is used to Doppler shift RVS CCD spectra to the rest frame before combining them into a mean RVS spectrum (Seabroke et al., in prep.). The actual RVS wavelength range extends into the filter wings (845–872 nm, see Cropper et al. 2018, Fig. 16), and the cut to 846–870 nm minimises border effects. In addition, the spectra are normalised at the local pseudo-continuum and are resampled to a wavelength bin width of 0.01 nm (2400 wavelength points, *wlp* hereafter) by the DPAC/Coordination Unit6 (CU6) pipelines.

It is important to note that GSP-Spec reassesses the continuum placement during the parameterisation procedure (see Sect. 6.3). Moreover, the spectra are rebinned from 2400 to 800 *wlp*, sampled every 0.03 nm (without reducing the spectral resolution thanks to the RVS oversampling), which increases their S/N. The RVS spectra analysed by GSP-Spec during DR3 operations were selected to have $S/N > 20$ before resampling. The considered S/N corresponds to the `rv_expected_sig_to_noise` value provided by the CU6 analysis (Seabroke et al., in prep.).

It is worth mentioning that, although the mean RVS spectra serve as an input to GSP-Spec, subsequent filtering of $V_{\rm Rad}$ was not propagated to GSP-Spec outputs for DR3. This means that there are a very small number of stars with GSP-Spec parameters, but not $V_{\rm Rad}$ (Appendix A). A subset of RVS mean spectra (999 995, all having $V_{\rm Rad}$) are published for the first time in

[8] https://www.cosmos.esa.int/web/gaia/iow_20210709

DR3 (Seabroke et al., in prep.). These articles detail the overlap of the published mean RVS spectra with GSP-Spec parameters (and other *Gaia* parameters).

4. Input and training synthetic spectra grids

GSP-Spec performs a model-driven parametrization for which stellar flux dependencies on atmospheric parameters and surface chemical abundances are interpreted through the comparison of the observed spectra with theoretical (*synthetic*) ones. For this purpose, we have computed large grids of synthetic RVS spectra with different combinations of stellar atmospheric parameters ($T_{\rm eff}$, log(g), [M/H] and [α/Fe]) and individual chemical abundances ([X/Fe], with X being the considered element, with the exception of Fe I and Fe II for which [X/M] is used). They span the entire parameter space of Galactic stellar populations with a detailed coverage that allows to reach the required parametrization precision. The use of these grids is three-fold: (i) training the GSP-Spec MATISSE (cf. Sect. 6.1) and ANN (cf. Sect. 7) algorithms before their application; (ii) acting as reference models for the algorithm performing on-the-fly regressions (GAUGUIN), and (iii) anchoring the normalization and DIB analysis procedures to reference flux values.

As a consequence, a 4-dimensional grid of spectra in $T_{\rm eff}$, log(g), [M/H] and [α/Fe] (cf. Sect. 4.2) and 5-dimensional grids for twelve chemical elements with the fifth dimension being [X/Fe] (cf. Sect. 4.3) are provided as input for GSP-Spec together with the learning functions of the parametrisation algorithms. These synthetic spectra are calculated through a procedure previously implemented for the AMBRE Project (de Laverny et al. 2012). We refer to a detailed description of the AMBRE grid to de Laverny et al. (2013). In the following, we particularly focus on several improvements considered for the GSP-Spec module.

4.1. Set of MARCS atmosphere models

The reference spectra are computed using MARCS atmosphere models (Gustafsson et al. 2008). We first selected 13 848 models that covered the following parameter space: 2600 to 8000 K for $T_{\rm eff}$ in steps of 200 or 250 K (below or above 4000 K, respectively), −0.5 to 5.5 for log(g) (step of 0.5 dex), and −5.0 to 1.0 dex for the mean metallicity (step of 0.25 dex for [M/H] > –2.0 dex and 0.5 dex for lower [M/H] values). For each metallicity, all the available [α/Fe]-enrichments were considered. In practice, this corresponds to models with [α/Fe]-values varying between at most −0.4 dex and +0.8 dex, around the classical relation observed for Galactic populations: [α/Fe] = 0.0 dex for [M/H]≥ 0.0 dex, [α/Fe] = +0.4 dex for [M/H]≤ −1.0 dex and [α/Fe] = −0.4 × [M/H] for −1.0 ≤ [M/H]≤ 0.0 dex. We point out, however, that not all values of [α/Fe] were always available for a given set of $T_{\rm eff}$, log(g), and [M/H]. Moreover, we only selected models for dwarfs (defined as log(g) > 3.5) with plane-parallel geometry and a microturbulent-velocity parameter of 1.0 km s^{-1} whereas spherical geometry with a mass of 1 $M_{\odot}$ and $V_{\rm micro}$ = 2 km s^{-1} were considered for giants (log(g) ≤ 3.5). Then, in order to improve the covering of the parameter space (particularly in the [α/Fe] dimension for which we adopted a step of 0.1 dex), we filled this first selection of MARCS models by models interpolated linearly, using the tool developed by T. Masseron and available on the MARCS website[9]. The resulting grid of MARCS atmosphere models adopted in the present work contains 35 803 models.

4.2. The 4-D spectra grid in $T_{\rm eff}$, log(g), [M/H] and [α/Fe]

For each adopted MARCS atmosphere model, a synthetic spectrum has been computed with the TURBOSPECTRUM code (version 19.1.2, Plez 2012) between 842.0 nm and 874.0 nm (i.e. a wider spectral domain than the one covered by the RVS spectra, in order not to be affected by border effects when simulating the RVS-like spectra) and adopting an initial wavelength step of 0.001 nm (i.e. corresponding to a spectral resolution larger than ∼300 000). We considered the Solar abundances of Grevesse et al. (2007), and specific atomic and molecular line lists. These line lists contain millions of lines and have been checked (and, when necessary, some atomic lines were calibrated) with observed spectra of benchmark reference stars (see Contursi et al. 2021, for more details). For dwarfs (defined as above for the MARCS models by log(g) > 3.5), the spectra were computed assuming one-dimensional plane-parallel atmospheric model while for giants (log(g) ≤ 3.5) a spherical geometry is considered. Both cases assume hydrostatic and local thermodynamic equilibria. Similar stellar masses as in the MARCS models were adopted for the computation. Moreover, consistent [α/Fe]-enrichments were considered in the model atmosphere and the synthetic spectrum calculation. Finally, we used an empirical law for the microturbulence parameter. This parametrized relation is a function of $T_{\rm eff}$, log(g) and [M/H] and has been derived from $V_{\rm micro}$ literature values for the *Gaia*-ESO Survey (Bergemann et al., in prep.). The spectra were computed in the air and then converted into vacuum wavelengths thanks to the relation of Birch & Downs (1994). It is worth noting that no stellar rotation or macro-turbulence broadening were included in these spectra. The impact of this assumption in the derived stellar parameters has been estimated from simulations and accounted through quality flags (Appendix C.1). These flags are a function of the *vbroad* parameter value of each star (available in the *gaia_source* table) but also of $T_{\rm eff}$, log(g) and [M/H].

The high-resolution spectra were then convolved and resampled in order to mimic real observed RVS spectra. For that purpose, we adopted a broadening instrumental profile corresponding to the RVS spectral resolution, keeping only the 846–870 nm domain and adopting the sampling of 0.03 nm chosen for the parametrisation within GSP-Spec (800 *wlp*, see Sect. 3). In practice, this convolution was performed thanks to tools developed for the DR3 version of the CU6 pipeline (Sartoretti et al. 2018). It assumes a Gaussian ALong-scan line spread function and adopts the median resolving power value known at the beginning of CU8's DR3 processing phase (R = 11 500, Cropper et al. 2018).

Finally, for the stellar atmospheric parameters estimation (see Sects. 6 and 7), this original grid of RVS-like synthetic spectra has been filled adopting a cubic Catmull-Rom (Catmull et al. 1974), a quadratic or linear 1D interpolation, depending on the number of neighbour models available. The final 4D grid contains 51,373 spectra with a constant step of 250 K, 0.5, 0.25 dex and 0.1 dex in $T_{\rm eff}$, log(g), [M/H] and [α/Fe], respectively. Figure 3 illustrates the covered parameter space.

4.3. 5D spectra grids for individual chemical abundance estimations

For the derivation of individual chemical abundances with the GAUGUIN method (Sect. 6.4), we have computed sets of 5D grids for which the first four dimensions are the ones of the 4D

[9] https://marcs.astro.uu.se/

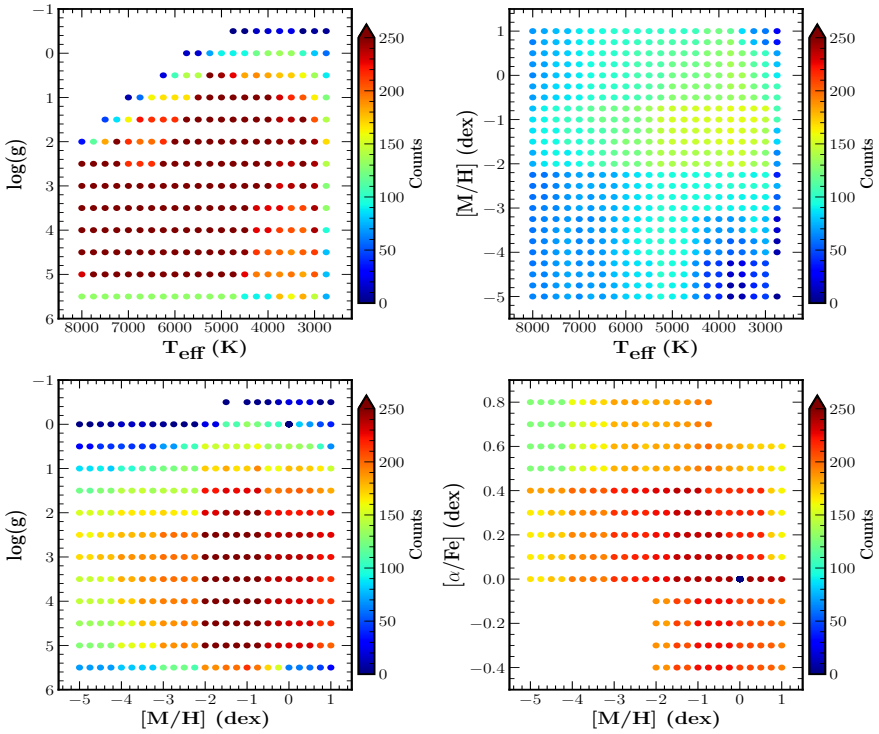


Fig. 3. Distribution in the 4D parameter space of the GSP-Spec reference grid, that contains the 51 373 synthetic spectra adopted for the stellar parametrisation. The colour-code refers to the number of available spectra in each 2D projection. For the derivation of the chemical abundance of a given chemical element X with the GAUGUIN method, 21 spectra are computed for most combinations of the four atmospheric parameters by varying the individual abundance of X (12 different species were considered: N, Mg, Si, S, Ca, Ti, Cr, Fe, Ni, Zr, Ce, Nd).

grid described above while the fifth dimension corresponds to the abundance values of a specific chemical species [X/Fe] (with the exception of Fe I and Fe II for which [X/M] is used). The considered chemical elements, X, are N, Mg, Si, S, Ca, Ti, Cr, Fe, Ni, Zr, Ce, Nd. These species have been chosen due to the availability of at least one of their atomic lines in the RVS spectral domain, following a careful line quality selection (see Sect. 5).

For these 5D grids, we considered a subsample of the MARCS models selected in Sect. 4.1: $T_{\rm eff} > 3500$ K, [M/H] > -3.0 dex and any values of $\log(g)$ and [α/Fe], except for Ca, Fe and Ti. Some atomic lines of these three atoms can indeed be detected at the very metal-poor regime and we therefore computed their 5D grids for any [M/H] values, i.e. down to [M/H] $= -5.0$ dex. The adopted variations in the chemical element dimension are from -2.0 to $+2.0$ dex around $\epsilon(\mathrm{X}) = \epsilon(\mathrm{X})_{\odot}$+[M/H]+$K_{\alpha}$, with a step of 0.2 dex (i.e. 21 different abundance values). K_{α} is assumed to be equal to zero for all elements except the α-species for which it follows a similar variation with the metallicity as [α/Fe]: $K_{\alpha} = 0.0$ for [M/H]≥ 0.0 dex, $K_{\alpha} = +0.4$ for [M/H]≤ -1.0 dex and $K_{\alpha} = -0.4 \times$ [M/H] for $-1.0 \leq$ [M/H]≤ 0.0 dex.

In total, we have computed twelve 5D grids of ~478 400 spectra each, except for Ca, Fe and Ti whose grids contain ~590 750 spectra since they cover the entire metallicity regime of the atmosphere model grids.

5. Line and wavelength interval selection for individual abundance analysis

As mentioned above, the reference synthetic spectra grids contain all the atomic and molecular lines collected by Contursi et al. (2021). Most of these lines are too weak and/or blended and can therefore not easily be used to derive reliable chemical diagnostics. To choose the adequate spectral intervals for individual abundance estimation, we implemented a careful line selection procedure and a thorough definition of the wavelength intervals for abundance estimation and local normalisation described below.

5.1. Selection of unblended lines

First, we looked for unblended lines through visual inspection of synthetic spectra at high-resolution (R ~100 000) and at the resolution of RVS (R ~11 500). The atmospheric parameters of four well-known reference stars were adopted: two cool giants (Arcturus and μ Leo), one cool dwarf (the Sun), and one hot dwarf (Procyon)[10]. In particular, we looked at (i) the flux contribution of each chemical species (including the 12 atomic elements and the most abundant molecules) by computing specific spectra with highly enhanced abundances, and (ii) the existing blends assuming super-solar metallicities and high enhancements in α-elements. This led to an initial selection of about 130 isolated atomic lines belonging to a dozen different atoms and five CN lines[11] that could be useful for chemical diagnostics. In particular, we identified interesting lines of some heavy elements (Zr, Ce, and Nd) and one line of singly ionised iron at λ = 858.794 nm, as suggested by Contursi et al. (2021), to complement iron abundance based on Fe I lines (see Sect. 8.7.2). The correct simulation of these lines was verified through the comparison of synthetic spectra to high-resolution observed spectra for the four mentioned benchmarks.

Second, the previous selection was confirmed by examining the observed RVS spectra of a few stars with atmospheric parameters close to those of the reference ones. By visual inspection, we kept only the lines showing the highest sensitivity to abundance variations in at least one of the inspected spectra,

[10] The adopted parameters for these stars can be found in Contursi et al. (2021).

[11] In our tests, CN was the sole identified molecule with rather unblended lines but this work has to be extended towards cooler stars ($T_{\rm eff} < 4000$ K) for future *Gaia* releases.

excluding those for which blends were still suspected from lines of different chemical elements within ~0.3 nm.

5.2. Selection of abundance and local normalization windows

To further optimise the line selection and the chemical analysis procedure, we carefully defined two wavelength windows around the selected lines used in the abundance estimation and the local normalisation, respectively. These were defined after visual inspection of the Arcturus, solar, and Procyon spectra at the RVS resolution, maximising the wavelength domains (and therefore the information on the abundance and the continuum placement) and avoiding nearby lines. To ensure the reliability of the finally selected windows, chemical abundances were derived using GAUGUIN for a set of about 10 000 RVS spectra and slight variations of the window interval. This allowed us to exclude window definitions producing very discrepant results ($\gtrsim$0.5 dex) with respect to other lines of the same element and their average value.

For line doublets and triplets, the merger of each line within a single abundance determination and normalisation window was adopted whenever possible. These are referred to as *merged multiplets* hereafter. In the particular case of the Ca II IR triplet, to minimise NLTE effects, two abundance windows were defined at the line wings, avoiding the cores (i.e. up to six independent abundance estimates can be provided for the three Ca II lines).

5.3. Line selection based on line-to-line abundance scatter

Finally, from the above set of unblended lines, we performed an additional selection to optimise the line-to-line scatter. For some species (N, Si, S, Ca, Ti, Cr, Fe I), the final element abundance was computed by combining the results of the different single line (or merged multiplets) abundances of the same element (cf. Sect. 6.8)[12]. For that purpose, we computed 50 Monte-Carlo flux realisations of the above-mentioned set of RVS spectra, considering their corresponding flux covariances. For each spectral line, the median and the inter-quartile range (IQR) were estimated from the derived abundance distribution. We then explored all the possible line combinations to evaluate the contribution of each line to the mean abundance, as well as its effect on the total number of estimates. A mean abundance is derived for each line combination, weighted by the inverse of the individual line abundance IQR. These weights were set to zero if they had IQR > 0.5 dex to avoid low-quality estimations. For each chemical element, the combination of lines minimising the line-to-line scatter and maximising the correlation[13] of the mean abundance with [M/H] (for iron-peak elements) or [α/Fe] (for α-elements) was selected.

The final list of 33 lines (some being merged multiplets) selected for the individual abundance analysis is provided in Appendix B and Table B.1, together with their associated windows for chemical analysis and normalisation. We refer also to the two figures of Sect. 6 for some examples of observed and model spectra that help to identify most of these lines. We note that Zr, Ce, and Nd lines passed all the above tests and are therefore adopted for abundance estimation. Similarly, the Fe II line is also conserved. As a consequence, the GSP-Spec module is able to estimate abundances of neutral and singly ionised iron (see Sect. 8.7.2).

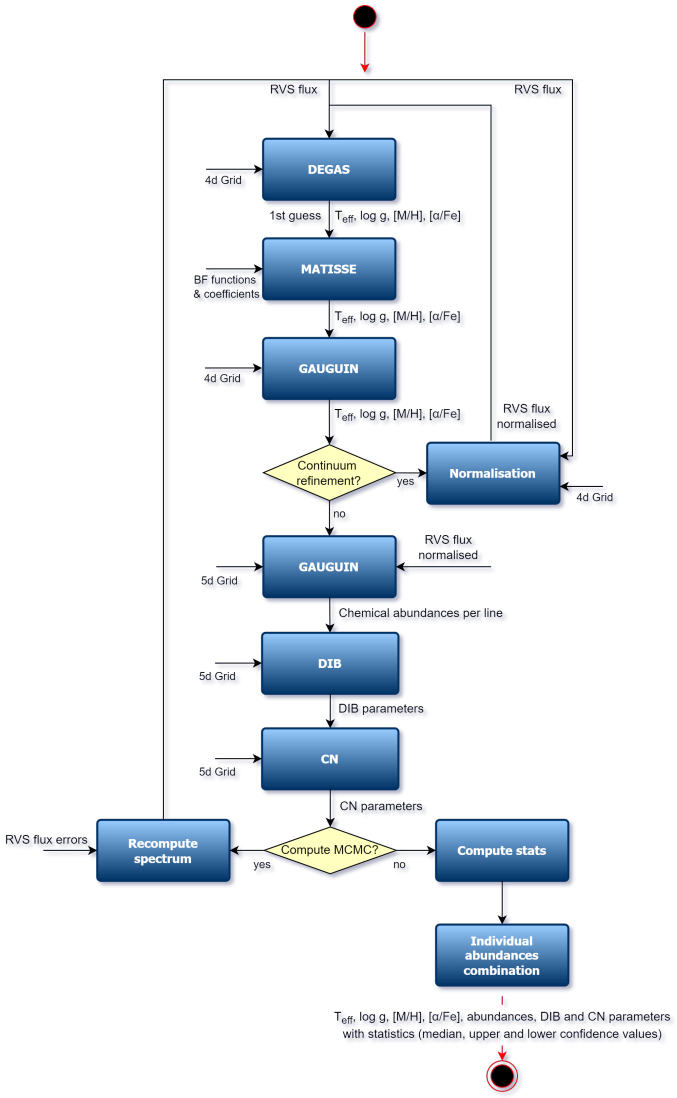


Fig. 4. Complete MatisseGauguin workflow that estimated stellar atmospheric parameters ($T_{\rm eff}$, log(g), [M/H], and [α/Fe]), individual chemical abundances of 12 species, CN, and DIB parameters (see Sect. 6 for detailed description).

6. GSP-Spec MatisseGauguin analysis workflow

This section describes the MatisseGauguin analysis workflow in sequential order. We reiterate that MatisseGauguin produces the GSP-Spec fields published in the *AstrophysicalParameters* table, including stellar atmospheric parameters, individual chemical abundances, and DIB parameters.

The complete workflow of MatisseGauguin is summarised in Fig. 4. In addition, to illustrate the MatisseGauguin parametrisation performance, the challenging automated fit of two observed high-S/N spectra[14] is presented in Figs. 5 and 6. The presented synthetic spectra are computed using the atmospheric parameters and chemical abundances estimated by GSP-Spec for those

[12] Other derived abundances (Mg, Ni, Fe II, Zr, Ce, Nd) rely on only a single line or a single abundance determination in the case of merged multiplets.

[13] We first performed a linear fit and then obtained the slope, the intercept, the median, and the MAD of the distance |data-fit|.

[14] These two spectra are not part of the set of RVS spectra published in the *Gaia* DR3.

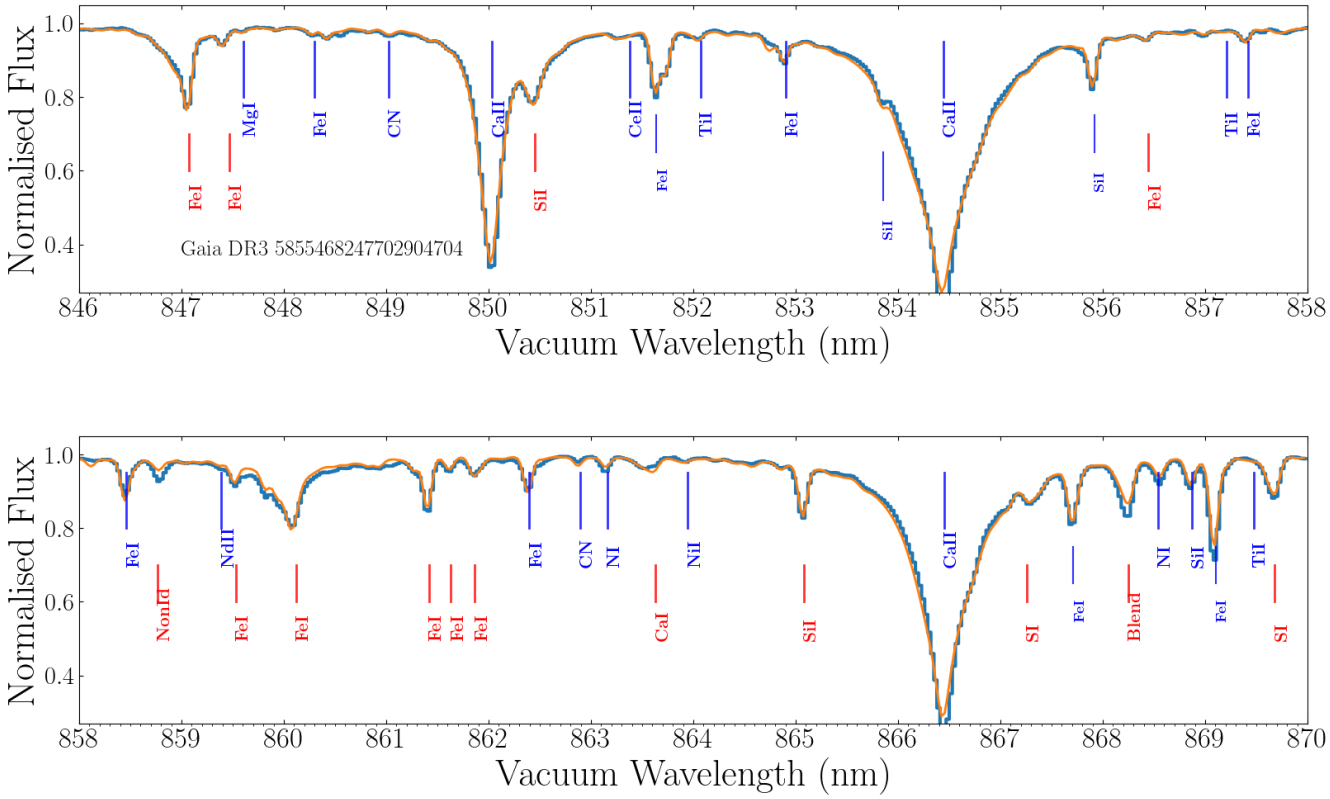


Fig. 5. Observed (blue histogram) and synthetic (orange line) spectra of the Cepheid variable star *Gaia* DR3 5855468247702904704. The observed spectrum has a very high S/N (equal to 884) and its histogram bin size corresponds to the wavelength sampling adopted for the analysis (0.03 nm, 800 *wlp*). The synthetic spectrum was computed from the GSP-Spec MatisseGauguin atmospheric parameters ($T_{\rm eff} = 5477$ K, $\log(g) = 1.44$, [M/H] = 0.07 dex, [α/Fe] = 0.11 dex) and individual chemical abundances, was then convolved by a rotational profile to reproduce the CU6 estimated broadening velocity (15.6 km s^{-1}) and, finally, was degraded to the RVS spectral resolution and sampling. The atomic lines identified in blue belong to the chemical species whose abundances were derived by the GAUGUIN method (the local normalisation performed for the chemical analysis of these selected lines was not considered in the figure for clarity reasons). The lines in red were not analysed in the shown spectrum because of suspected blends in the present case. The feature around 868.3 nm is a blend of SI+FeI+SiI plus probably other potential unidentified lines. The NonId feature at ~858.8 nm is a blend of the Fe II line described in Sect. 8.7.2 (seen in orange) and of unidentified lines that cannot be reproduced with the present line list.

two stars. The identification of several spectral lines is included in the figures. It is worth noting that the combination of the automated MatisseGauguin parametrisation with our reference spectra is able to find an excellent match with the observations that confirms the quality and the high precision of the observed RVS spectra and the input reference spectra grids.

6.1. MATISSE stellar atmospheric parameters

To initialise the whole MatisseGauguin procedure, a first guess of $T_{\rm eff}$, $\log(g)$, [M/H], and [α/Fe] is derived using the DEGAS decision-tree method (Bijaoui et al. 2010), which considers the entire parameter space of the 4D grid (see also Kordopatis et al. 2011, 2013, for first applications to observed spectra).

Subsequently, the MATISSE algorithm (Recio-Blanco et al. 2006) is applied following an iterative procedure in the parameter estimation. This allows the user to overcome problems caused by a non-linear variation of the spectral flux with the stellar parameters. MATISSE is a local multi-linear regression method, resulting from the projection of the full input spectrum onto a set of vectors (called BF functions in Fig. 4). These vectors (and the associated coefficients) account for the sensitivity, at each wavelength, of the stellar flux to variations of a given parameter ($\Delta T_{\rm eff}$, $\Delta\log(g)$, Δ[M/H] or Δ[α/Fe]); they are derived during a training phase based on the noise-free 4D reference grids, and correspond to regions of the entire parameter space, spanning ±500 K in $T_{\rm eff}$, ±0.5 dex in $\log(g)$, ±0.25 dex in [M/H], and ±0.20 dex in [α/Fe]. The noise optimisation is taken into account by employing a Landweber algorithm during the covariance matrix inversion and which is adapted to each scientific application (see Recio-Blanco et al. 2006, for more details). The MATISSE projection is first applied at the DEGAS solution in a local environment of ±500 K in $T_{\rm eff}$, ±0.5 dex in $\log(g)$, ±0.25 dex in [M/H], and ±0.20 dex in [α/Fe] (corresponding to the parameter space region of each training function). This produces a second solution around which MATISSE is applied again. This iterative procedure is repeated until convergence (i.e. the solution stays within the local environment), within a maximum of ten iterations.

6.2. GAUGUIN refinement of the atmospheric parameters

The GAUGUIN algorithm is then applied around the final MATISSE solution of the previous step, considering a local environment of ±250 K in $T_{\rm eff}$, ±0.5 dex in $\log(g)$, ±0.25 dex in [M/H], and ±0.20 dex in [α/Fe]. GAUGUIN (Bijaoui et al. 2012; Recio-Blanco et al. 2016) is a classical, local optimisation method implementing a Gauss-Newton algorithm. It is based on

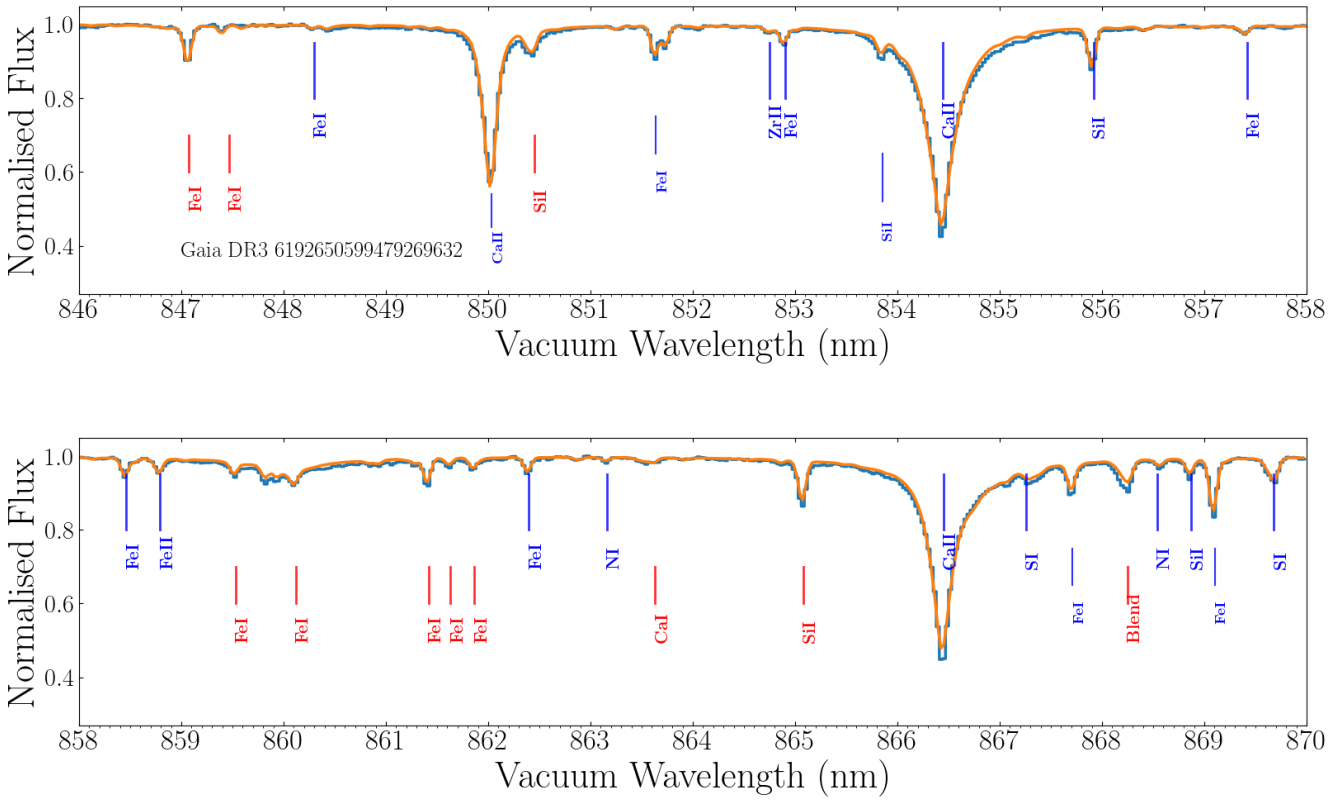


Fig. 6. Same as Fig. 5 but for the hot dwarf *Gaia* DR3 6192650599479269632 whose MatisseGauguin atmospheric parameters are $T_{\rm eff} = 6754$ K, $\log(g) = 4.38$, [M/H] = −0.03 dex, and [α/Fe] = 0.15 dex (S/N = 408). No rotational profile was applied as no broadening velocity was estimated (suspected low-rotating star).

a local linearisation around a given set of parameters that are associated with a reference synthetic spectrum (via linear interpolation of the derivatives). It is designed to find the direction in the parameter space that has the highest negative gradient as a function of distance (defined as the flux difference between the observed and synthetic spectra). Once this direction is found, the method proceeds in an iterative way, by modifying the initial guess of the studied parameter and re-calculating the gradient again, until convergence of the parameter solution. A few iterations are carried out through linearisation around the new solution until the algorithm converges towards the minimum distance. In practice, and to avoid trapping into secondary minima, we recall that GAUGUIN is initialised by parameters independently determined by MATISSE. At the end of this process, the final MatisseGauguin solution in $T_{\rm eff}$, $\log(g)$, [M/H], and [α/Fe] is provided as input to the spectrum normalisation procedure.

6.3. Spectra re-normalisation and iterations on atmospheric parameters

The parameter solution of the previous step is used to re-estimate the continuum placement. This step is particularly important in the case of cool stars, which have pseudo-continuum flux values that can be much lower than one. The continuum placement and normalisation procedure is described in detail in Santos-Peral et al. (2020). In this step, the spectrum flux is normalised over the entire RVS wavelength domain. For this purpose, the observed spectrum (O) is compared to an interpolated synthetic one from the 4D reference grid (S) with the same atmospheric parameters. First, the most appropriate wavelength points of the residuals (Res = S/O) are selected using an iterative procedure implementing a linear fit to Res followed by a σ-clipping. The residual trend is then fitted with a third-degree polynomial. Finally, the refined normalised spectrum is obtained after dividing the observed spectrum by a linear function resulting from the fit of the residuals.

This renormalised spectrum is then fed back to the first step described in Sect. 6.1 to re-estimate the atmospheric parameters using the new spectra normalisation. This loop is performed five times (a sufficient number to reach convergence), iterating on the parameters and the continuum placement.

The parameters of the converged solution in $T_{\rm eff}$, $\log(g)$, [M/H], and [α/Fe] is then saved, as well as the final normalised spectrum. A goodness-of-fit (gof) between the observed and a synthetic spectrum interpolated to the atmospheric parameters is computed. The logarithm of this gof is reported in the *AstrophysicalParameters* table under the *logchisq_gspspec* field. The provided gof value reports the goodness of fit with respect to the observed spectrum, not including Monte-Carlo variations of the flux (see Sect. 6.7).

6.4. GAUGUIN chemical abundances per spectral line

Considering the final atmospheric parameters solution and normalised spectrum, each of the 33 selected atomic lines (see Table B.1) is then analysed with GAUGUIN to estimate the chemical abundance of the related chemical element causing the line absorption.

First, for each line l associated with the chemical element X, a specific 1D grid in the $[X/Fe]$ abundance space is generated. To this purpose, the corresponding 5D grid presented in Sect. 4.3 is interpolated at the stellar $T_{\rm eff}$, $\log(g)$, [M/H], and [α/Fe] values

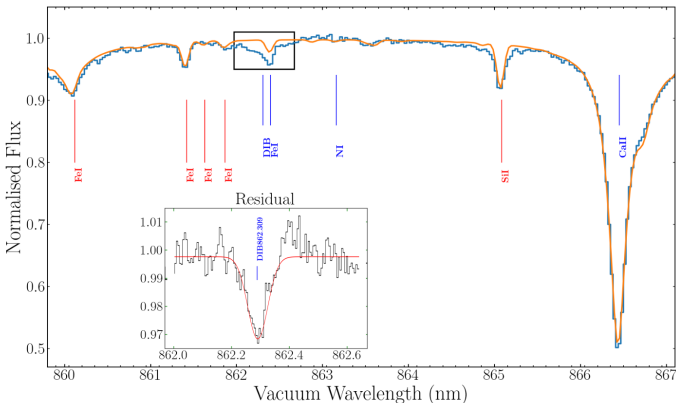


Fig. 7. Similar to Fig. 5 but for the metal-poor hot subgiant *Gaia* DR3 4378933739135936000 around its DIB feature. The insert is a zoom onto the flux residual between observed and model spectra around the DIB. It has been renormalised and the DIB characteristics are measured thanks to the Gaussian fit shown in red (EW = 0.0244 nm and central wavelength p_1 = 862.309 nm). The MatisseGauguin atmospheric parameters of this star are $T_{\rm eff}$ = 6414 K, $\log(g)$ = 3.75, [M/H] = −0.61 dex, and [α/Fe] = +0.42 dex (S/N = 293 and CU6 broadening velocity equal to 17.1 km s^{-1}).

of the adopted MatisseGauguin solution (cf. Sect. 6.3). This 1D reference spectra grid covers the entire normalisation wavelength range. It includes a large range of abundance variations in $\epsilon(X)$. Second, a local normalisation around the line is performed (Santos-Peral et al. 2020). A minimum quadratic distance is then calculated between the reference grid and the observed spectrum, providing a first guess of the abundance estimate $[X/Fe]^l_0$. This initial guess is then optimised using the GAUGUIN algorithm, which iterates through linearisation around the successive new solutions. The algorithm stops when the relative difference between two consecutive iterations is less than a given value (one-hundredth of the grid abundance step) and provides the final abundance estimation of each line ($[X/Fe]^l$).

6.5. Diffuse interstellar band parameters

Once the atmospheric parameters and the individual abundances have been derived, the next step of the MatisseGauguin workflow is to evaluate the presence of any DIB signature around ∼862 nm. For each RVS spectrum, we first perform a local renormalisation on the spectrum around the DIB feature (over 35 Å around 862 nm). We then pass a preliminary detection of the DIB profile by fitting a Gaussian profile to produce initial guesses for the fitting and eliminate cases where noise is at the same level as or exceeds the depth of the possible detection of the DIB. Only detections above the 3 σ-level are considered as true detections. In order to perform the main fitting process of the DIB, we then separate our sample into cool ($3500 < T_{\rm eff} \leq 7000$ K) and hot star samples ($T_{\rm eff} \geq 7000$ K). For cool stars, the observed spectrum is divided by a synthetic spectrum whose atmospheric parameters are provided by MatisseGauguin. The residual, assumed to correspond to the DIB profile, is then renormalised and fitted by a Gaussian function (see Fig. 7). For hot stars for which no lines are found close to the DIB feature, a Gaussian process similar to Kos (2017) is applied where the DIB profile is fitted by a Gaussian process regression (Gershman & Blei 2012).

For each detected DIB feature, we determine its equivalent width (EW), the central wavelength of the fitted Gaussian (p_1), its depth (p_0), the width of the Gaussian profile (p_2), and their uncertainties which are estimated based on Monte-Carlo Markov chain realisations (see Sect. 6.7). We remind the reader that, for a Gaussian, $p_2 = FWHM/(2*\mathrm{sqrt}(2*\ln 2))$ with $FWHM$ being the full width at half maximum. The EW is computed assuming a Gaussian profile: EW = $\sqrt{2\pi} \times |p_0| \times p_2/C$ where C is the continuum level.

Finally, two quality flags (DIB_q and QF) for the DIB parameters were implemented in order to allow a selection of the best determinations, depending on the science application (see e.g. Gaia Collaboration 2023b). The first quality flag DIB_q is included in the GSP-Spec quality flag string chain and its value varies from zero (highest quality) to five (lowest quality) and is equal to nine when no DIBs are measured; we refer to Sect. 8.9 for its definition. The second flag QF is defined during the preliminary detection of the DIB profile and provides the reason why DIB_q has been fixed to nine for a given spectrum.

If the depth of the fitted profile is smaller than 3-σ the noise level, we do not consider this case to be a true detection and assign it $QF = -1$. Finally, stars with effective temperatures cooler than 3500 K are automatically disregarded because their spectrum is crowded by molecular lines, leading to undetectable DIB ($QF = -2$).

6.6. Cyanogen differential abundance proxy

In the spectra of cool stars, a couple of cyanogen lines can be seen (their wavelength identification can be found in Fig. 5, although the lines are weaker in the illustrated spectrum with respect to cooler stars). Five interesting CN lines were initially identified when building the line list. The tests performed in the line-selection process presented in Sect. 5 selected one of these CN lines as a reliable CN over- or underabundance proxy in the spectra of cool stars.

This CN line is centred at 862.884 nm and a window of 0.15 nm has been selected around it for its analysis. As for the DIB feature of cool stars, the observed spectrum is divided by the corresponding synthetic spectrum, interpolated to the atmospheric parameters of the star derived by MatisseGauguin. This synthetic spectrum assumes the solar-scaled values of carbon and nitrogen abundances [C/Fe] = [N/Fe] = 0.0 dex. We then estimated the EW of the residual by adopting the same Gaussian fitting procedure as for the DIB parameters of cool stars. This CN proxy is therefore an indicator of the strength of the line with respect to the standard value, and reveals a CN underabundance or overabundance (positive or negative EW, respectively). In addition, the central wavelength and the width of the residual feature are also derived from the above-mentioned Gaussian fit, as already implemented for the DIB.

6.7. Propagation of flux uncertainties

The estimation of a star's atmospheric parameters, chemical abundances from individual lines, DIB, and CN-index parameters described above is performed from the input RVS spectrum, without considering the associated flux uncertainties per wlp. To estimate parameter uncertainties induced by the spectral noise, the complete MatisseGauguin workflow is rerun 50 times to analyse the same number of different Monte-Carlo realisations of the stellar spectrum. Upon each realisation, the input stellar flux per wlp F_i is modified according to the corresponding flux uncertainty of that wlp, σ_{Fi}. In particular, each realisation is computed by adding or subtracting a ΔF_i at each wlp i, randomly

sampling a Gaussian distribution of standard deviation equal to σ_{Fi} and centred at zero.

The complete Monte-Carlo implementation produces a total set of 50 values for each estimated parameter ($T_{\rm eff}$, log(g), [M/H], [α/Fe], individual line abundances [X/Fe]^l, DIB, and CN indexes). For each of the corresponding parameter distributions, we compute the median and the lower and upper confidence values, from the 50th, 16th, and 84th quantiles, respectively. The median value of each parameter is saved as the adopted parameter estimation in the GSP-Spec catalogue. Both the lower and upper confidence levels are also published. In summary, this procedure allows parameter uncertainties to be properly estimated for each star, and for them to be tailored to the quality of the associated spectrum, but also to its stellar type and chemical abundance pattern. It is important to note that, in this way, the uncertainties on individual line abundances [X/Fe]^l propagate the atmospheric parameters ones, as new [X/Fe]^l values are computed upon each realisation for the new $T_{\rm eff}$, log(g), [M/H], and [α/Fe] estimations. In addition, asymmetric uncertainties around the finally considered median value are provided thanks to the lower and upper confidence levels. This Monte Carlo treatment is made possible thanks to the extremely fast application of the GSP-Spec analysis (cf. Sect. 2).

6.8. Individual element chemical abundances

As explained in Sect. 6.4, GAUGUIN provides chemical abundances for each of the 33 atomic lines of Table B.1, called [X/Fe]^l. The final chemical abundances per element [X/Fe] are derived by combining the independent abundance estimates [X/Fe]^l of all the available lines of the same species. To this purpose, a mean abundance per element is calculated, weighted by the inverse of the [X/Fe]^l uncertainty of each line (defined as the upper minus lower confidence values of the [X/Fe]^l abundance distribution provided by the 50 Monte-Carlo realisations). The published abundances are [N/Fe], [Mg/Fe], [Si/Fe], [S/Fe], [Ca/Fe], [Ti/Fe], [Cr/Fe], [Fe I/M], [Fe II/M][15], [Ni/Fe], [Zr/Fe], [Ce/Fe], and [Nd/Fe]. Their associated lower and upper confidence values are also published and were calculated as the weighted mean of the [X/Fe]^l ones.

7. The GSP-Spec ANN workflow

The ANN algorithm is based on supervised learning and provides a different parameterisation of the RVS spectra, independent from the MatisseGauguin workflow. ANN projects the RVS spectra onto the label space of the astrophysical parameters. We trained the network on the same grid of reference synthetic spectra as MatisseGauguin (see Sect. 4), in this case adding noise according to the different S/N scales in the observed spectra (Manteiga et al. 2010).

The ANN architecture is feed-forward with three fully connected neuron layers. The input layer has as many neurons as *wlp* in the spectrum (800) whereas the output layer has four neurons corresponding to the number of estimated parameters. The number of neurons in the hidden layer was empirically determined between 50 and 100 for nets trained with low- to high-S/N spectra, respectively. In the same way, we determined the learning rate in the range [0.001, 0.2]. The activation function

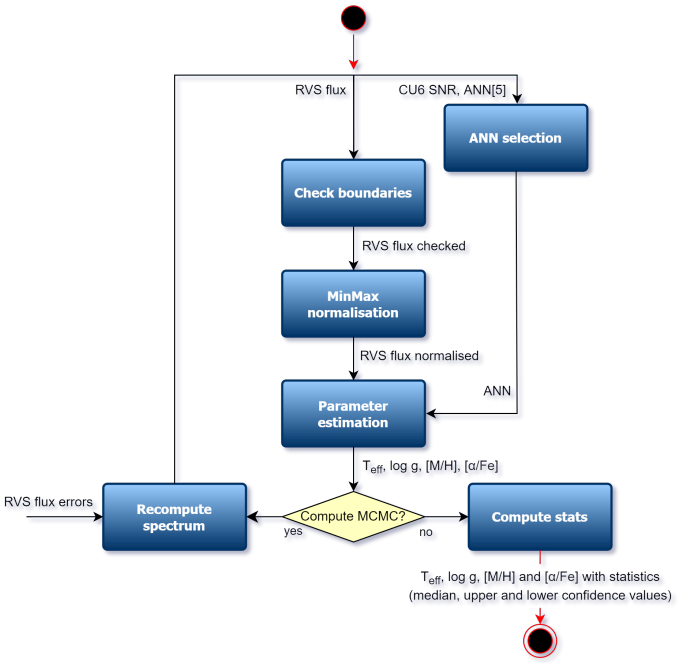


Fig. 8. ANN workflow that provides the second set of the main stellar atmospheric parameters ($T_{\rm eff}$, log(g), [M/H] and [α/Fe]).

Table 1. Equivalent S/Ns between ANN networks and RVS spectra.

ANN	25	30	35	40	50
RVS	[20, 24]	(24, 40]	(40, 68]	(68, 108]	(108, ∞]

selected for input and output layers is linear, whereas the logistic function was selected for the hidden layer.

The training procedure is performed with the backpropagation function, which can be interpreted as a problem of minimisation of the error existing between the obtained and desired outputs. In order to avoid overtraining, and to select the ANN that leads to the best generalisation, the early stopping procedure was used, finalising the training process when the performance starts to degrade, and obtaining the net that minimises the error.

The effectiveness of the ANN depends on the input ordering. For that reason, we perform ten trainings with different ordering, selecting the one with minimum error. For each train, weights initialise in the range [−0.2,0.2], and we established a limit of 1000 iterations because we observed that beyond that number, the training process does not improve but the computational cost increases.

The ANN parameterisation procedure that estimates the second set of GSP-Spec atmospheric parameters ($T_{\rm eff}$, log(g), [M/H] and [α/Fe]) is published in the *AstrophysicalParametersSupp* table and is summarised in Fig. 8. Specifically, the present ANN version included in GSP-Spec proceeds as follows:

ANN selection: ANN behaves well in the presence of noise (Manteiga et al. 2010), confirming that it is a robust method when estimating astrophysical parameters for relatively low-S/N spectra. As there is no noise model for the *Gaia* RVS spectra, we empirically determined the relation between the noise given by CU6 and the Gaussian noise that we need to use to train the nets. The corresponding values are shown in Table 1. For each RVS input spectrum, we then used its S/N value,

[15] We provide iron abundances with respect to the mean metallicity following the implementation of the reference grids. The classical [Fe/H] can be easily obtained by adding [M/H] to [Fe I/M] or [Fe II/M].

provided by CU6, to select which net performs the parameter estimation.

Check boundaries: Some RVS spectra have zero flux values at the beginning or at the end of their spectral range. These are often caused by radial velocity corrections and could lead to large flux variations in the borders and cause ANN malfunctions. To avoid this behaviour, we truncated these zero flux values and adopted the mean of the flux spectrum for these *wlp*.

Normalisation: A minimum–maximum scaling procedure is applied to the RVS spectra, equalising it to avoid geometric biases during the training stage in order to guarantee that all the inputs are in a comparable range.

Parameter estimation: Once the net has been selected, it is fed with the normalised spectrum to estimate $T_{\rm eff}$, log(g), [M/H] and [α/Fe]. The net returns these estimations normalised, and so a denormalisation procedure is applied to return the values in the expected range.

Monte-Carlo iterations using flux uncertainties: The same procedure as for MatisseGauguin (see Sect. 6.7) is also applied for ANN to estimate the parameter uncertainties caused by flux errors. We therefore obtain the median and the lower and upper confidence values of each AP again.

8. Validation and *flags_gspspec* quality flag chain

The GSP-Spec output after operations has been carefully checked and validated, considering different potential error sources. Following this validation procedure, a quality flag chain (*flags_gspspec*) is implemented (cf. Table 2)[16]. In this chain, a value of 0 is the best, and 9 is the worst, generally implying the parameter masking. This allows the user to publish all kinds of quality results, satisfying the more or less restrictive needs of different science applications. Nevertheless, this implies that considering these quality flags is mandatory for correct use of the GSP-Spec parameters and abundances. If not applied, results of low quality for a given application could be unconsciously included in the analysis, severely affecting its conclusions.

The following subsections review the different reasons for failure, potential bias, and the uncertainty sources considered in the GSP-Spec validation, and following the characters ordering in the quality flag chain. Several associated figures and tables can be found in Appendix C.

8.1. Parameterisation biases induced by rotational and macroturbulence line broadening (vbroad flags)

GSP-Spec is trained with reference spectra assuming no rotation (see Sect. 4). At the RVS spectral resolution, the parameterisation tolerance to broadened spectra through rotational (V sin i) and/or macroturbulence broadening has to be flagged according to tests with synthetic data.

Potential biases in $T_{\rm eff}$, log(g), and [M/H] induced by stellar rotation were therefore modelled using a dedicated set of synthetic RVS spectra, which were broadened with different V sin i values from 0 to 70 km s^{-1}. For simplicity, we assumed in what follows that the line broadening factor produced by CU6 (*vbroad*) is well reproduced by only mimicking a stellar rotation. The estimated biases ($\Delta T_{\rm eff}$, Δlog(g), Δ[M/H]) induced by rotational broadening are a function of $T_{\rm eff}$ and log(g). Metallicity dependencies are also observed, with metal-poor objects being more affected than metal-rich ones (a consequence of their smaller number of lines that can be used for the parametrisation).

First, using this data set, we identified the limiting V sin i values inducing a bias larger than $\Delta T_{\rm eff} = 2000$ K. We then modelled the parameter dependence of the V sin i values leading to that maximum admitted bias by fitting a third-order polynomial with variable $T_{\rm eff}$, log(g), and [M/H], as shown in Fig. C.1. To avoid extrapolation issues, upper and lower limits were imposed on the third-order interpolation polynomial during the post-processing. This function was finally adopted during post-processing to mask the corresponding GSP-Spec results (Flag *vbroad* $T = 9$ in Table 2). Similarly, we applied this procedure to define three other values for this quality flag, depending on the amplitude of the predicted induced bias in $T_{\rm eff}$: Flag *vbroad* $T = 0$, 1, or 2 for stars with a possible bias $\Delta T_{\rm eff} \leq 250$ K, $250 < \Delta T_{\rm eff} \leq 500$ K, and $500 < \Delta T_{\rm eff} < 2000$ K, respectively.

Exactly the same procedure was adopted for defining the flags associated with a bias in log(g) and [M/H] induced by the rotational and macroturbulence line broadening. Their detailed definition is given in Table C.1.

8.2. Parameterisation biases induced by radial velocity uncertainty (vrad flags)

In a very similar way, we investigated the possible bias induced by radial velocity uncertainties, because the GSP-Spec parametrisation is performed whilst assuming that the observed spectra are perfectly at rest-frame. The examination of GSP-Spec unfiltered results reveals that large $V_{\rm Rad}$ errors (provided by CU6) are preferentially found in specific regions of the output atmospheric parameter space (combinations of $T_{\rm eff}$, log(g), and [M/H] where no stars are expected, or at extremely high or low [α/Fe]). This is an important illustration of the expected parametrisation sensitivity to $V_{\rm Rad}$ uncertainties.

We therefore investigated the amplitude of possible biases in $T_{\rm eff}$, log(g), and [M/H] caused by $V_{\rm Rad}$ errors varying between 0 and 10 km s^{-1} using specific synthetic spectra. Again, metal-poor stars (with a lower number of lines available for the parametrisation) were found to be more affected than metal-rich ones. As described above for the *vbroad* flags, specific third-order polynomials with variable $T_{\rm eff}$, log(g), and [M/H] were then fitted to define the values associated with three *vrad* flags. Their precise definition is given in Table C.2.

8.3. Parameter uncertainties due to flux noise (fluxNoise flag)

The parametrisation is affected by uncertainties in the observed fluxes, that is, the noise at each wavelength leading to a mean S/N over the entire wavelength domain. To quantify this effect and as already explained in Sect. 6.7, flux uncertainties are taken into account by GSP-Spec through 50 Monte-Carlo realisations of the spectral flux for each star. The GSP-Spec parameterisation is then performed for those 50 spectrum realisations and parameter uncertainties (noted σ, hereafter) are defined from the 16th and 84th quantiles of the obtained distributions. To enable a rapid selection of results in the GSP-Spec catalogue from the estimated parameter uncertainties, we defined a specific quality flag ($fluxNoise$). This flag simultaneously considers uncertainties in $T_{\rm eff}$, log(g), [M/H], and [α/Fe], labelling results of progressively higher precision from $fluxNoise = 5$ to $fluxNoise = 0$. The exact conditions imposed during the post-processing for the noise uncertainty quality flags are indicated in Tables C.3 and C.4 for MatisseGauguin and ANN, respectively. It is worth

[16] We note that the flags associated with the ANN results correspond to the first 12 flags of this table.

Table 2. Definition of each character in the GSP-Spec quality flag string chain (*flags_gspspec*), including the possible values (Col. 3) and the related subsection and tables providing further information (Col. 4).

Chain character number - name	Considered quality aspect	Possible adopted values	Related subsection and table
1 **vbroadT**	*vbroad* induced bias in $T_{\rm eff}$	0,1,2,9	8.1 & C.1
2 **vbroadG**	*vbroad* induced bias in $\log(g)$	0,1,2,9	8.1 & C.1
3 **vbroadM**	*vbroad* induced bias in [M/H]	0,1,2,9	8.1 & C.1
4 **vradT**	$V_{\rm Rad}$ induced bias in $T_{\rm eff}$	0,1,2,9	8.2 & C.2
5 **vradG**	$V_{\rm Rad}$ induced bias in $\log(g)$	0,1,2,9	8.2 & C.2
6 **vradM**	$V_{\rm Rad}$ induced bias in [M/H]	0,1,2,9	8.2 & C.2
7 **fluxNoise**	Flux noise induced uncertainties	0,1,2,3,4,5,9	8.3 & C.3, C.4
8 **extrapol**	Extrapolation level of the parametrisation	0,1,2,3,4,9	8.4 & C.5, C.6
9 **negFlux**	Negative flux *wlp*	0,1,9	8.5 & C.7
10 **nanFlux**	NaN flux *wlp*	0,9	8.5 & C.7
11 **emission**	Emission line detected by CU6	0,9	8.5 & C.7
12 **nullFluxErr**	Null uncertainties *wlp*	0,9	8.5 & C.7
13 **KMgiantPar**	KM-type giant stars	0,1,2	8.6 & C.8
14 **NUpLim**	Nitrogen abundance upper limit	0,1,2,9	8.7 & C.9
15 **NUncer**	Nitrogen abundance uncertainty quality	0,1,2,9	8.7 & C.10
16 **MgUpLim**	Magnesium abundance upper limit	0,1,2,9	8.7 & C.9
17 **MgUncer**	Magnesium abundance uncertainty quality	0,1,2,9	8.7 & C.10
18 **SiUpLim**	Silicon abundance upper limit	0,1,2,9	8.7 & C.9
19 **SiUncer**	Silicon abundance uncertainty quality	0,1,2,9	8.7 & C.10
20 **SUpLim**	Sulphur abundance upper limit	0,1,2,9	8.7 & C.9
21 **SUncer**	Sulphur abundance uncertainty quality	0,1,2,9	8.7 & C.10
22 **CaUpLim**	Calcium abundance upper limit	0,1,2,9	8.7 & C.9
23 **CaUncer**	Calcium abundance uncertainty quality	0,1,2,9	8.7 & C.10
24 **TiUpLim**	Titanium abundance upper limit	0,1,2,9	8.7 & C.9
25 **TiUncer**	Titanium abundance uncertainty quality	0,1,2,9	8.7 & C.10
26 **CrUpLim**	Chromium abundance upper limit	0,1,2,9	8.7 & C.9
27 **CrUncer**	Chromium abundance uncertainty quality	0,1,2,9	8.7 & C.10
28 **FeUpLim**	Neutral iron abundance upper limit	0,1,2,9	8.7 & C.9
29 **FeUncer**	Neutral iron abundance uncertainty quality	0,1,2,9	8.7 & C.10
30 **FeIIUpLim**	Ionised iron abundance upper limit	0,1,2,9	8.7 & C.9
31 **FeIIUncer**	Ionised iron abundance uncertainty quality	0,1,2,9	8.7 & C.10
32 **NiUpLim**	Nickel abundance upper limit	0,1,2,9	8.7 & C.9
33 **NiUncer**	Nickel abundance uncertainty quality	0,1,2,9	8.7 & C.10
34 **ZrUpLim**	Zirconium abundance upper limit	0,1,2,9	8.7 & C.9
35 **ZrUncer**	Zirconium abundance uncertainty quality	0,1,2,9	8.7 & C.10
36 **CeUpLim**	Cerium abundance upper limit	0,1,2,9	8.7 & C.9
37 **CeUncer**	Cerium abundance uncertainty quality	0,1,2,9	8.7 & C.10
38 **NdUpLim**	Neodymium abundance upper limit	0,1,2,9	8.7 & C.9
39 **NdUncer**	Neodymium abundance uncertainty quality	0,1,2,9	8.7 & C.10
40 **DeltaCNq**	Cyanogen differential equivalent width quality	0,9	8.8 & C.12
41 **DIBq**	DIB quality flag	0,1,2,3,4,5,9	8.9 & C.13

Notes. Flag names are split into three categories: Parameter flags (green), individual abundance flags (blue), and EW flags (orange). All flags concern the MatisseGauguin parameters, while only the parameter flags except KMgiantPar are applied to ANN results.

noting that stars with extremely poor quality parameters, such as, for instance, those without any distinction between giants and dwarfs ($\sigma\log(g) > 2$ dex) or between F, G, and K stellar types ($\sigma T_{\rm eff} > 2000$ K) are filtered out during the post-processing ($flux Noise = 9$) and do not appear in the finally published catalogue.

8.4. Extrapolation level (extrapol flag)

Due to extrapolation, the GSP-Spec parameter solution could be located outside the parameter space of the training grid (cf. Fig. 3) for either one or several parameters. In addition, censored training occurs near the grid borders. In order to flag those extrapolated results for which the parametrisation is less reliable, we have implemented a specific flag (*extrapol*) that is indicative of the extrapolation level.

The definition of this flag is reported in Tables C.5 and C.6 for MatisseGauguin and ANN, respectively, depending on the availability (or not) of a gof and the distance between the parameter solutions and the grid borders. The flag value depends on the level of extrapolation: from results near the grid limits ($extrapol = 4$) to no extrapolation at all (and therefore a more

reliable solution, $extrapol = 0$). Again, sources without a gof and with $T_{\rm eff}$ values outside the 2500 to 9000 K interval or $\log(g)$ values outside the -1 to 6 dex range were filtered out ($extrapol = 9$) during the post-processing and do not appear in the final catalogue.

8.5. RVS flux issues or emission line flags

MATISSE and GAUGUIN being model-driven methods that essentially aim to maximise the goodness of fit between an observation and a set of templates, any significant and/or systematic difference between the RVS spectra and the reference grid can introduce biases in the results. These differences can be associated with the RVS spectra processing, or be inherent to the stellar physics assumptions adopted when computing the reference grid (stellar activity being one example). When such issues randomly affect a *wlp*, then it can be very difficult (if not impossible) to properly take them into account during the analysis. We have implemented four specific flags to identify such cases. Their definition is presented below and is summarised in Table C.7.

RVS spectral anomalies can manifest as *wlp* that have a negative flux (flag *negFlux*), or a flux (or associated variance) that is not a number (*nanFlux* and *nullFluxErr* flags, respectively). Whereas such caveats do not necessarily alter the RV determination, they can hamper parameterisation estimates relying specifically on the affected *wlp*. For instance, some tens of stars have a couple of *wlp* with negative flux. They are predominantly found in the cores of the strongest Ca II lines and result from an oversubtraction of the straylight during the spectrum production. This leads to a modified line profile and could indeed affect the parametrisation. Similarly, NaN flux values can appear in the spectra. As explained in Seabroke et al. (in prep.), *wlp* are masked in the CCD sample. When these are averaged, a chance alignment of these masks when there are few CCD spectra pixels contributing to a particular wavelength bin in the combined spectrum could lead to a NaN flux value, which happens more often near the edges. The GSP-Spec treatment partly overcomes this problem thanks to the rebinning (from 2400 to 800 *wlp*) of the oversampled input spectra. For this rebinning, a median flux is computed every three *wlp*, excluding NaN values. As a consequence, NaN flux values in the rebinned spectra only remain if the three averaged *wlp* are equal to NaN. To filter out those rare cases, we have implemented the specific *nanFlux* flag. Finally, if no flux variance is associated with a *wlp*, then the derived parameter uncertainty is unreliable or impossible to estimate. This is reported by the *nullFluxErr* flag.

As presented in Table C.5, while the *nanFlux* and the *nullFluxErr* flags lead to a systematic exclusion of the source from the final catalogue (only values equal to 9 have been implemented), the *negFlux* flag can also be equal to 1 (one or two *wlp* with negative flux values) or 0 (no negative *wlp* at all). However, for the reasons described above, we recommend preferentially selecting stars with $negFlux$=0.

On the other hand, emission lines due to stellar activity are inherent to the stellar properties and carry important information about the observed star. However, the physical conditions that lead to the emission lines are not considered in our grid of synthetic spectra. Therefore, if a star shows signs of activity, its GSP-Spec parameters should also be discarded and considered unreliable. We used the $CU6_is_emission$ flag provided by the CU6 to detect such stars, and forced them to have a GSP-Spec flag $emission = 9$ to reject them.

8.6. Parametrisation quality of K and M type giants

The parametrisation of cool stars with effective temperatures below 4000 K is known to be complex due to their crowded spectra, which results from the increasing presence of atomic and, especially, molecular lines. This aggravates normalisation issues and parameter degeneracies, in particular for metal-rich stars. During the GSP-Spec validation process, a correlation was found between the minimum flux value ($F_{\rm min}$) of the spectra of giant stars with $T_{\rm eff} \lesssim 4000$ K and their estimated $\log(g)$. In particular, in this cool temperature regime, objects with higher $\log(g)$ values present larger $F_{\rm min}$ values than expected when compared to those of slightly hotter giants with similar $\log(g)$ and S/N values. This reveals a parametrisation problem, as the pseudo-continuum should present lower values for cooler stars for which the line-crowding increases, and not vice versa. We have therefore implemented a specific flag (*KM-typestars*) that takes this issue into account. This flag depends on the $F_{\rm min}$ value and the gof in order to take account of the influence of the S/N on $F_{\rm min}$. As reported in Table C.8, stars with *KM-typestars* equal to 1 and 2 have corrected $T_{\rm eff}$ and $\log(g)$ with uncertainties reflecting the GSP-Spec parameterisation problems encountered for these stars: $T_{\rm eff} = 4250 \pm 500$ K and $\log(g) = 1.5 \pm 1$.

8.7. Quality of individual chemical abundances

We checked the reliability of all the abundance estimates, including their uncertainties across the Kiel diagram ($\log(g)$ vs. $T_{\rm eff}$ plot) and taking into account the S/N. As a result of this process, we defined two flags for each individual abundance. Their definitions are given in Table C.9 and C.10 (with associated coefficients in Table C.11).

On one hand, as expected, the estimation quality depends on the strength of the spectral lines of the studied element, which varies with $T_{\rm eff}$, $\log(g)$, [M/H], and the abundance of the element itself. To help the user to deal with this effect, we implemented the individual abundance upper limit flag ($XUpLim$), which is an indicator of the line depth with respect to the noise level. This flag is based on an estimate of the detectability limit (*upper-limit*) that depends on the line atomic data, the stellar parameters, the line broadening, and the S/N. We note that, for the definition of this $XUpLim$ flag, we adopted a GSP-Spec internal estimate of the S/N that could slightly differ from the published `rv_expected_sig_to_noise`. The closer is the derived abundance to this upper limit, the higher the flag value and the abundances should therefore be used more cautiously.

On the other hand, for low-S/N spectra (with a limiting S/N depending on the analysed lines), the reliability of the associated abundance uncertainties can be underestimated. This is due to the fact that the maximum allowed abundance value in the reference grids is [X/Fe] = 2.0 dex, preventing higher values in the abundance distribution associated with the flux noise Monte-Carlo realisations. This effect depends on the line detectability and the S/N. As a consequence, we defined a second individual abundance flag ($XUncer$) labelling the reliability of the associated abundance uncertainty taking into account its dependence on the stellar type ($T_{\rm eff}$, $\log(g)$, and [M/H]), the S/N estimate, and/or the gof. Moreover, it is worth noting that the distance between the [X/Fe] upper confidence level and the grid upper border is also a good indicator of the estimate reliability.

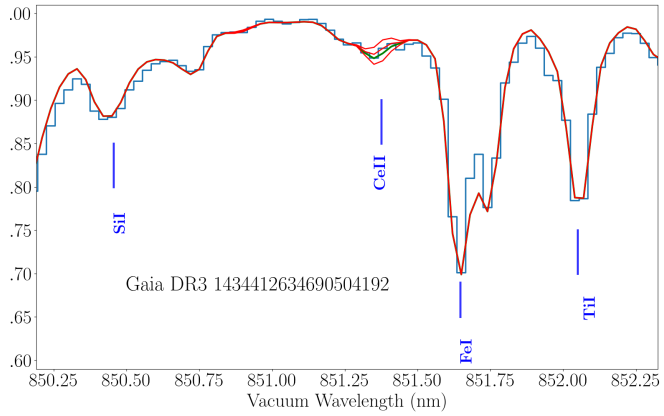


Fig. 9. Fit of the RVS spectrum (blue histogram) of the RGB star *Gaia* DR3 1434412634690504192 around its cerium line. The model in green corresponds to the GAUGUIN solution [Ce/Fe] = 0.26 dex (in excellent agreement with the literature value) whereas those in orange have [Ce/Fe] = −2.0 dex (almost no cerium) and ±0.2 dex around the GAUGUIN abundance, respectively. The S/N is 907 and the broadening velocity is equal to 10.4 km s^{-1}. See text for more details.

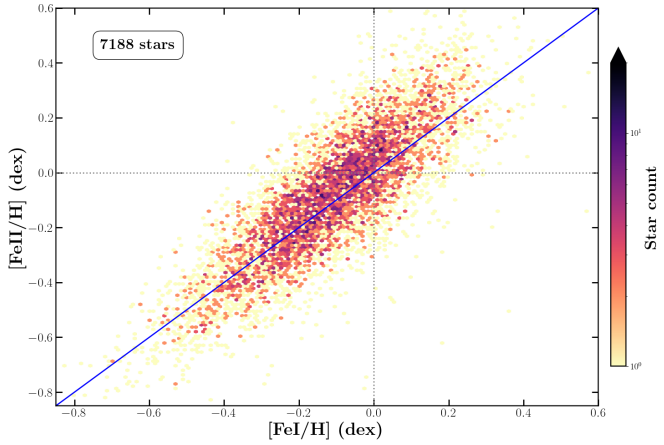


Fig. 10. Comparison between iron abundances measured from the proposed Fe II line at 858.79 nm and from all the other Fe I lines. The Spearman correlation coefficient is equal to 0.82. See text for more details.

8.7.1. Validation of heavy element abundances

An important illustration of the quality of the GSP-Spec abundance analysis with GAUGUIN is provided by the derivation of heavy element abundances, the estimation of which seemed too challenging for the RVS resolution. As an example, Fig. 9 shows the RVS spectrum of a red giant branch (RGB) star around its cerium line. The Ce abundance ([Ce/Fe] = 0.26 dex with lower and upper confidence levels being 0.18 and 0.38 dex, respectively) was derived from the MatisseGauguin parameters: $T_{\rm eff}$ = 4157 K, log(g) = 1.09, [M/H] = −0.4 dex, and [α/Fe] = 0.12 dex. The GSP-Spec abundance flags are $CeUpLim = CeUncer = 0$. This star was previously analysed by Forsberg et al. (2019) who derived a very consistent [Ce/Fe] = 0.22 dex, adopting very similar atmospheric parameters. It is important to note that GSP-Spec cerium abundances are on the same scale as those found by Forsberg et al. (2019) with a null median difference for the overlapping sample. This confirms the high quality of the GSP-Spec chemical analysis and of the *Gaia*/RVS spectra.

8.7.2. Validation of the singly ionised iron abundance

The specific case of Fe II abundances merits discussion. When building the GSP-Spec line list, Contursi et al. (2021) identified an unknown line at 858.79 nm (in the vacuum) and proposed that it is actually an Fe II feature. Because of its unblended nature in the RVS spectra of hot stars (see Fig. 6), it has been included in the line list used by GAUGUIN for the individual abundance analysis (Table B.1).

In Fig. 10, Fe II abundances are compared to Fe I ones in the atmospheric parameters regime where both estimates are possible at the same time. Both iron abundances were calibrated as suggested in Sect. 9. We selected stars with all 13 atmospheric parameter flags equal to zero together with a rather strict quality selection using the two abundance flags: $XUpLim \leq 1$ and $XUncer = 0$. We also selected only stars in which the Fe II line is easily detected ($6000 < T_{\rm eff} < 7200$ K). The agreement between both iron abundances is excellent. The Spearman correlation coefficient is equal to 0.82 and increases up to 0.89 when selecting the ∼2500 stars with $S/N > 300$.

We can therefore safely conclude that this 858.79 nm line is indeed a very good metallicity proxy and probably corresponds to an absorption produced by an iron-peak element, Fe II being the best candidate as suggested by Contursi et al. (2021).

8.8. Quality of cyanogen differential equivalent width (DeltaCNq)

To validate CN parameters, literature data were used to identify cool RGB and AGB stars for which CN lines are expected to be present. The flag associated with the EW of this CN abundance proxy (*DeltaCNq*) is defined in Table C.12; it depends on the three line-broadening flags (*vbroad*), the S/N, the gof, and the measured line position (p_1).

8.9. DIB quality flag (DIBq)

To quantify the quality of the DIB analysis, we defined a specific flag, ranging from $DIBq = 0$ (highest quality) to 5 (lowest quality). When no DIB is measured ($DIBq = 9$), another flag QF, not included in the flag chain, details the reasons as to why no measurements were performed (see Sect. 6.5). Its definition depends on the p_0 and p_2 parameters but also on the global noise level (*Ra*) defined by the standard deviation of the (*data* − *model*) residual between 860.5 and 864 nm as well as on the local noise level *Rb*, that is, the (*data* − *model*) residual within the DIB profile. Table C.13 explains the definition of the $DIBq$ flag and Fig. C.2 shows its flow chart. As discussed in Gaia Collaboration (2023b), we recommend the adoption of the most reliable DIB parameters ($DIBq = 0,1,2$) as well as (i) a good central wavelength measurement $862.0 < p1 < 862.6$ nm, (ii) a rather small uncertainty on the EW measurement (err(EW)/EW < 0.35), and (iii) a good stellar parametrisation (first 13 GSP-Spec flag being smaller than 2).

9. Known parameter and abundance biases

After the previous evaluation of the parameter quality through a flagging system, internal and external biases were studied, taking into account the implemented flags. The result of this analysis is presented in this section for MatisseGauguin atmospheric parameters and abundances (Sect. 9.1 includes a summary of the proposed solutions at the end) and for ANN atmospheric parameters (Sect. 9.2). Several figures and tables associated with this

section can be found in Appendix E. In some cases, simple calibrations with low-degree polynomials are suggested. It is worth noting that published DR3 GSP-Spec data are deliberately uncalibrated, and so users are able to (i) use the raw data that come from the GSP-Spec processing, (ii) apply, whenever suggested, the calibrations presented in this paper, and (iii) perform a new calibration tailored to their scientific analysis.

On one hand, although specific work has been done on the optimisation of the reference synthetic spectra grids, the observed biases can be partially due to mismatches between observations and reference synthetic spectra if some physical aspects not considered in the modelling (e.g. stellar rotation, macroturbulence, departures from local thermodynamic and hydrostatic equilibria) become non-negligible for some parameters of certain types of stars. This has been partially taken into account with parameter flags (e.g. *vbroadT*, *vbroadG*, *vbroadM* flags). We recall that the parametrisation of cool stars is often challenging (see e.g. Sect. 8.6 and Soubiran et al. 2022), and even higher resolution surveys in the literature can exhibit biases.

On the other hand, it is worth noting that the observed biases with respect to the literature can also have their origin in methodological and theoretical assumption differences with respect to those adopted in this work, such as different atmosphere models, atomic data, or reference solar abundances. In addition, several ground-based spectroscopic surveys have applied *adhoc* offset corrections as a result of their calibration procedures. Finally, the presented global biases with respect to the literature depend on the relative proportion of stars in the various reference catalogues as a function of the S/N and the analysed parameter space.

Finally, it is important to mention that reference catalogues have their own biases. Although literature references are generally calibrated (while *Gaia* archive data are not), this does not remove all the existent trends, as shown by some recent works (e.g. Soubiran et al. 2022). As a consequence, it cannot be excluded that the observed trends in the comparison with external catalogues are partly due to biases that are still present in the literature data.

As a consequence of all the above mentioned points, the results of the bias analysis presented in the following have to be cautiously and thoroughly considered. We recommend that the user adapt any bias correction to the targeted scientific goal and selected sample.

9.1. GSP-Spec MatisseGauguin biases

The MatisseGauguin workflow produces both atmospheric parameters and individual chemical abundances. Estimation biases have been evaluated for each case and are presented in the two following subsections. We have chosen to present the [α/Fe] biases together with those of individual abundances, as the underlying spectral indicators are dominated by the Ca II IR triplet lines and, as a consequence, the behaviour of [α/Fe] is very similar to that of the [Ca/Fe] abundance. Section 9.1.3 summarises the observed biases and the proposed solutions.

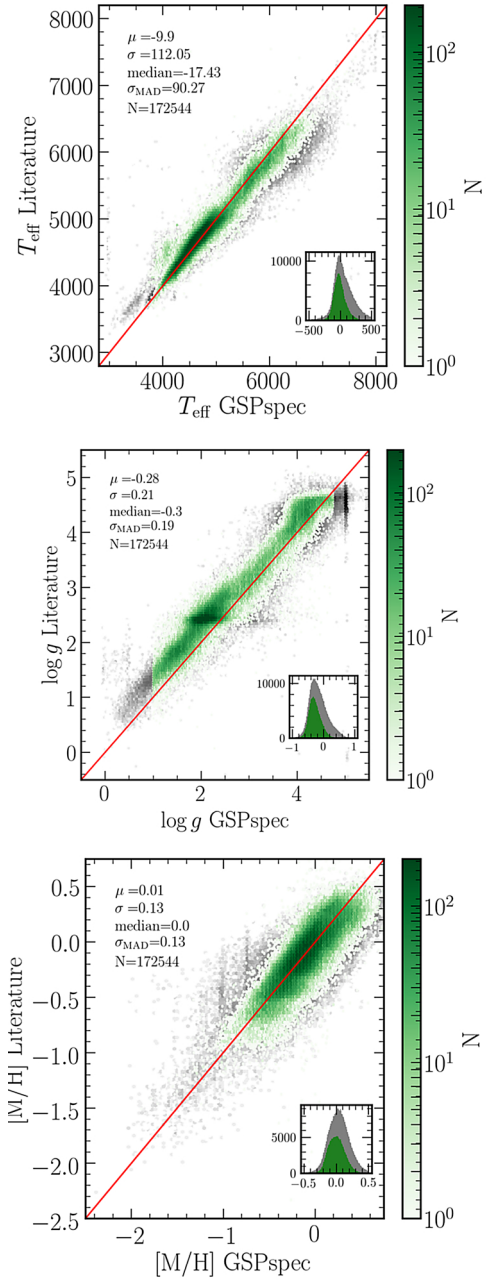


Fig. 11. Density plots comparing GSP-Spec MatisseGauguin parameters with literature data (APOGEE-DR17, GALAH-DR3, RAVE-DR6). Green and grey show the best- and medium-quality subsamples, respectively (see text for details about these samples). The histograms inside each plot show the difference between the literature and the GSP-Spec parameters. Mean (μ), standard deviation (σ), median, robust standard deviation (derived from the MAD), and the number of stars (N) of the offsets for the best-quality subset are annotated inside each box.

9.1.1. Analysis of $T_{\rm eff}$, log(g), and [M/H]

In this section, we compare the GSP-Spec MatisseGauguin $T_{\rm eff}$, log(g), and [M/H] with the latest data releases of three major ground-based spectroscopic surveys, namely APOGEE-DR17 (Abdurro'uf et al. 2022), GALAH-DR3 (Buder et al. 2021), and RAVE-DR6 (Steinmetz et al. 2020). We filtered the literature samples based on both the associated uncertainties of the published parameters (≤500 K, 0.5, 0.3 dex, for $T_{\rm eff}$, log(g) and metallicity/iron abundance, respectively) and the reliability flags (following the suggestions of each of the respective surveys). In total, a sample of ~8×10^5 stars (among which ~7.5×10^5 unique targets) were selected in such a way. The three panels in Fig. 11 show how the main atmospheric parameters compare when all of the first 13 GSP-Spec flags are equal to zero (best quality sample, ~1.7×10^5 stars plotted in green) and when we allow them to be smaller than or equal to one, except for the *KMgiantPar*, which we insist must be equal to zero and the *fluxnoise* flag

Table 3. Polynomial coefficients for the calibration of the Matisse-Gauguin gravities and metallicities.

Parameter	p_0	p_1	p_2	p_3	p_4
$\log(g)$	0.4496	−0.0036	−0.0224		
[M/H]	0.274	−0.1373	−0.0050	0.0048	
$[M/H]_{OC}$	−0.7541	1.8108	−1.1779	0.2809	−0.0222

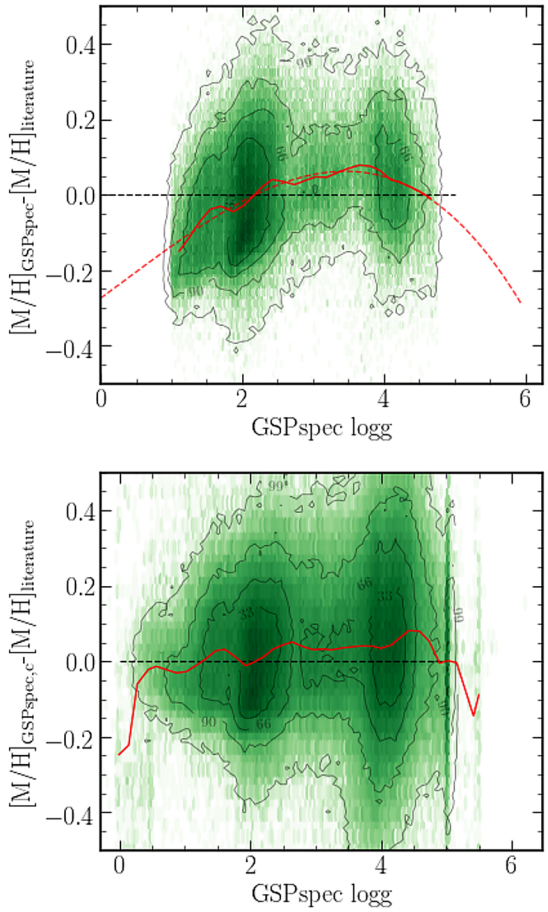


Fig. 12. Comparison of GSP-Spec and literature metallicities. *Top*: 2D histogram of the differences between the GSP-Spec metallicities and the literature values as a function of uncalibrated $\log(g)$ for our best-quality sample. The red full line is the running mean of the difference, and the dashed line is the fit to the running mean, defining the correction to apply. *Bottom*: medium-quality sample showing the differences between the calibrated metallicities and the literature values.

that we relax to smaller than or equal to three (medium quality sample, plotted in grey, $\sim 3.7 \times 10^5$ stars).

For our best quality sample, we find a median offset for $T_{\rm eff}$, $\log(g)$, [M/H] of −17 K, −0.3 dex and 0.0 dex, respectively, and a robust standard deviation (i.e. ∼1.48 times the median absolute deviation) of 90 K, 0.19 dex, and 0.13 dex. These trends are globally similar when taking into account each reference catalogue separately (see Appendix D).

Whereas $T_{\rm eff}$ and [M/H] are globally well recovered (however, see next paragraph), $\log(g)$ determination is slightly biased. GSP-Spec MatisseGauguin finds consistently lower gravities, the offset being larger for giants than for dwarfs. Based on these findings, we suggest the following calibration for $\log(g)$:

$$\log(g)_{\rm calibrated} = \log(g) + \sum_{i=0}^{2} p_i \cdot \log(g)^i. \quad (1)$$

The p_i coefficients were obtained by fitting the trends with respect to the above-mentioned literature compilation, and are reported in the first row of Table 3.

Furthermore, we note that despite finding, overall, a zero offset in metallicity, a further investigation of the trends compared to the literature shows that giants ($\log(g) \lesssim 1.5$) have slightly underestimated metallicities, whereas dwarfs ($\log(g) \gtrsim 4$) have slightly overestimated values (see top plot of Fig. 12). These trends can be corrected by fitting a low-order polynomial to the residuals as a function of *uncalibrated* $\log(g)$, and correcting the raw metallicities by this polynomial. The correction takes the form of:

$$[M/H]_{\rm calibrated} = [M/H] + \sum_{i=0}^{\rm deg} p_i \cdot \log(g)^i. \quad (2)$$

The p_i coefficients are provided in Table 3. Two different corrections are proposed. The first one, a third-order polynomial, was obtained by fitting the trends with respect to the above-mentioned literature compilation. The result of this calibration is illustrated in the bottom plot of Fig. 12. The second proposed correction, a fourth-order polynomial, is based on a set of open cluster stars with known metallicity from the literature and high membership probability (Cantat-Gaudin et al. 2020; Castro-Ginard et al. 2022; Tarricq et al. 2021). The advantage of open cluster data is that they ensure a constant metallicity at all $\log(g)$ values for the same object. However, as open clusters are thin-disc objects, the considered [M/H] range is restricted to the metal-rich regime. This alternative correction is illustrated in Fig. 13 and reported in the last row of Table 3.

9.1.2. Analysis of [α/Fe] and individual chemical abundances

To evaluate, calibrate, and remove possible gravity dependencies on the measured [α/Fe], [Fe I/H], [Fe II/H], and [X/Fe] abundance values (with X being an arbitrary element), we follow the strategy described below. It assumes that the abundance distribution (expressed relative to the solar values) should be close to zero in the solar neighbourhood for stars with metallicities close to solar and velocities close to the Local Standard of Rest (to avoid stars with large eccentricities). This strategy furthermore has the advantage of avoiding any calibration based on external catalogues. The procedure that we carry out is the following.

We first select only stars that have their first 13 quality flags (see Table 2) less or equal to one, except for their *KMgiantPar* flag and *extrapol* flag which we set to be equal to zero. In addition, we also impose that the abundance flag associated with the upper limit (*XUpLim*) is equal to zero, whereas the one associated with the uncertainties (*XUncer*) is set to less than or equal to one. Finally, we set an upper limit for their uncertainty (defined as the difference between the upper value and the lower value divided by two) and the line scatter to be less than 0.2 dex for both.

Amongst the selected stars, we further select the ones that are located within 0.25 kpc of the Sun; have a global metallicity [M/H]= 0.0 ± 0.25 dex (to avoid possible effects due to metallicty zero-point offsets); and have an azimuthal velocity V_ϕ close

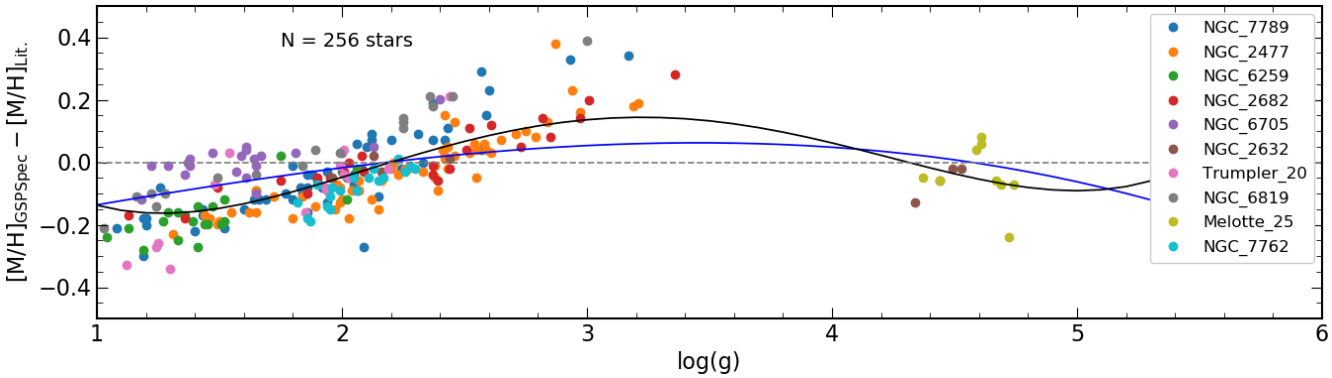


Fig. 13. Metallicity bias with respect to the literature as a function of log(g) for the open cluster stars, excluding dwarfs with S/N lower than 50. The colour code used for each cluster is indicated in the legend. The solid blue line corresponds to the general metallicity correction while the black line refers to that specifically obtained from the open clusters.

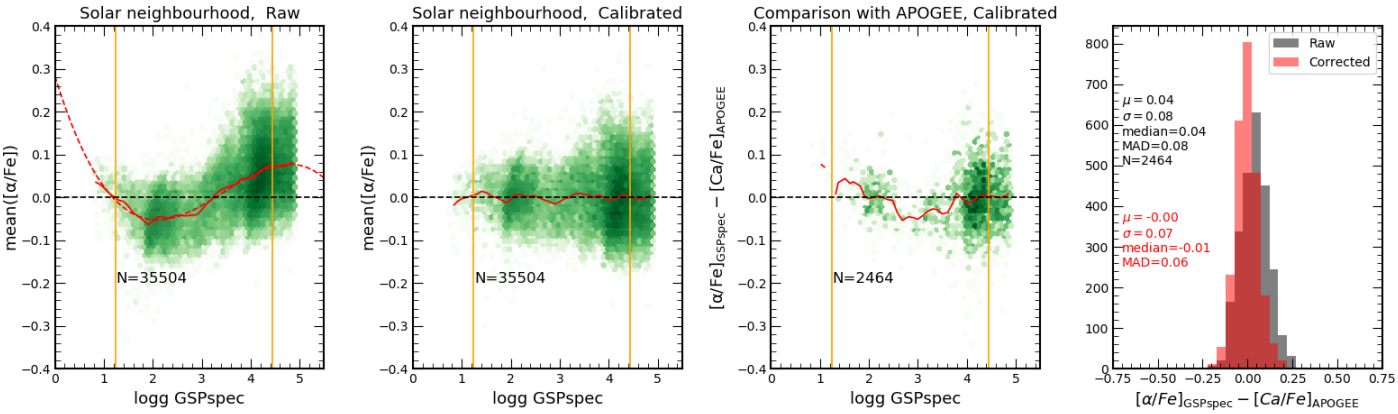


Fig. 14. Correction of [α/Fe] trends as a function of log(g). *Left panel*: the 2D histogram of the stars with 3750 K $\leq T_{\rm eff} <$ 5750 K, log(g)$<$ 4.9 in green, with all of their quality flags equal to zero, located at the solar neighbourhood, with velocities close to the LSR and metallicities close to solar values in the raw (i.e. uncalibrated) [α/Fe]-log(g) space, colour-coded by log(N). The running mean is plotted as a full red line, and its fit is the red dashed line. The dashed black line is included as a visual reference for the y-axis. Vertical orange lines indicate the log(g) range over which the calibration is assumed to be reliable (differences between the fit and the running mean smaller than 0.05 dex). The second panel is similar to the left one, but the calibration has now been applied. *Third panel*: the difference between the calibrated [α/Fe] and the calcium values from APOGEE DR17 as a function of MatisseGauguin log(g), where we have relaxed the *extrapol* flag to be less than or equal to one. Finally, the right panel shows the histograms of the differences compared to the literature data before (in grey) and after (in red) the calibration. Quantifications of the mean, median, standard deviation, and robust standard deviation (1.4826· MAD) are shown in the *top left corner* for the uncalibrated values (in grey) and in the *bottom right corner* for the calibrated values (in red).

to the Local Standard of Rest ($V_{\rm LSR} \pm 25\,{\rm km\,s^{-1}}$)[17]. By choosing such a sample, we ensure that we select stars with a high probability of having, on average, similar chemical properties to the Sun. Therefore, their $[X_1/X_2]$ abundance distributions (with X_1 and X_2 associated with two different elements or families of elements) are expected to be centred on zero.

We then compute the running mean of $[X_1/X_2]$ as a function of log(g), in bins of δlog(g)= 0.2 dex (red full line on the first row of plots in Figs. 14, E.1, and E.2). This trend, for an unbiased abundance estimation, should be centred on zero, regardless of the dispersion of the underlying distribution (which is a manifestation of either a true Galactic dispersion, or of the precision of GSP-Spec, or both).

Finally, we fit the trend defined by the running mean with a third- or fourth-order polynomial (choosing the correct compromise, depending on the data behaviour, to avoid overfittings), where each point has a weight inversely proportional to the dispersion of $[X_1/X_2]$ within the considered log(g)-bin. This fit defines the correction that could be applied to our data (red dashed line on the leftmost panels in Figs. 14, E.1, and E.2). The correction takes the form of:

$$[X_1/X_2]_{\rm calibrated} = [X_1/X_2] + \sum_{i=0}^{\rm deg} p_i \cdot \log(g)^i, \quad (3)$$

where deg = 3 or 4 and X_2 is either Fe or H, depending on the chemical species (see Table 4).

We also define the log(g) range over which the calibration is expected to be valid (vertical orange lines in the figures). The latter is evaluated by estimating the difference between the running mean and the fit, $\Delta_{\rm fit}$, and excluding the points at log(g) $\pm$ 0.4 from the boundaries for which $\Delta_{\rm fit}$ is larger than 0.05 dex (chosen arbitrarily). We note that the application of the calibration outside this log(g) confidence range should be used with caution, if not avoided. To increase the validity log(g) range for the abundances with high number statistics ([α/Fe] and [Ca/Fe]), we propose another calibration by relaxing the GSP-Spec quality flag associated to the extrapolation ($\leq$1). This leads to an alternative fourth-order polynomial fitting (second and third last rows of Table 4) and allows a qualitative view of how the correction

[17] Velocities are computed as in Gaia Collaboration (2023a).

Table 4. Polynomial coefficients, recommended parameter intervals, and *extrapol* flag values for Matisse-Gauguin [α/Fe] and individual abundance calibrations (Eq. (3)).

Element	p_0	p_1	p_2	p_3	p_4	Recommended interval		*extrapol* flag
	As a function of log(g)					Min log(g)	Max log(g)	
[α/Fe]	−0.5809	0.7018	−0.2402	0.0239	0.0000	1.01	4.85	0
[Ca/Fe]	−0.6250	0.7558	−0.2581	0.0256	0.0000	1.01	4.85	0
[Mg/Fe]	−0.7244	0.3779	−0.0421	−0.0038	0.0000	1.30	4.38	0
[S/Fe]	−17.6080	12.3239	−2.8595	0.2192	0.0000	3.38	4.81	0
[Si/Fe]	−0.3491	0.3757	−0.1051	0.0092	0.0000	1.28	4.85	0
[Ti/Fe]	−0.2656	0.4551	−0.1901	0.0209	0.0000	1.01	4.39	0
[Cr/Fe]	−0.0769	−0.1299	0.1009	−0.0200	0.0000	1.01	4.45	0
[Fe I/H]	0.3699	−0.0680	0.0028	−0.0004	0.0000	1.01	4.85	0
[Fe II/H]	35.5994	−27.9179	7.1822	−0.6086	0.0000	3.53	4.82	0
[Ni/Fe]	−0.2902	0.4066	−0.1313	0.0105	0.0000	1.41	4.81	0
[N/Fe]	0.0975	−0.0293	0.0238	−0.0071	0.0000	1.21	4.79	0
[α/Fe]	−0.2838	0.3713	−0.1236	0.0106	0.0002	0.84	4.44	≤1
[Ca/Fe]	−0.3128	0.3587	−0.0816	−0.0066	0.0020	0.84	4.98	≤1
	As a function of $t = T_{\rm eff}/5750$					Min $T_{\rm eff}$	Max $T_{\rm eff}$	
[α/Fe]	−6.6960	20.8770	−21.0976	6.8313	0.0000	4000	6830	≤1
[Ca/Fe]	−7.4577	23.2759	−23.6621	7.7657	0.0000	4000	6830	≤1
[S/Fe]	0.1930	−0.2234	0.0000	0.0000	0.0000	5700	6800	≤1

Notes. The uncertainties associated with these coefficients are provided in Table E.1.

behaves outside the log(g)-confidence range of the third-order polynomial.

We then verify on the same sample that the correction improves the trends (second column of plots of Figs. 14, E.1, and E.2).

Finally, we use literature data, which contain a wider variety of metallicities, to verify that the calibration is indeed improving the offsets (third and fourth columns of Figs. 14, E.1, and E.2). The literature data we use in this case are composed of APOGEE-DR17 and GALAH-DR3 for all of the elements except sulphur, and AMBRE for this latter abundance (Perdigon et al. 2021). In the case of [α/Fe], the comparison is made with respect to literature [Ca/Fe] values, as in the RVS domain the [α/Fe] indicators are dominated by the CaT lines. We note that, for these abundance comparisons, no agreement was required between GSP-Spec and the literature in the related stellar atmospheric parameters or the assumed solar abundances.

It can be seen from Fig. 14, E.1, and E.2 that the provided calibrations for the [α/Fe] and individual chemical abundance offsets significantly decrease their gravity dependence (and even remove it completely for several species), and that they set them close to the solar values. Moreover, the comparison with literature data is also improved, reducing the offset and/or the dispersion. The values of the polynomial coefficients of Eq. (3) (p_i) together with their domain of validity in log(g), to avoid extrapolations, are listed in Table 4. We also provide the uncertainties on the polynomial coefficients (derived from the fit) in Table E.1, as a possible criterion to evaluate the robustness of the calibration. We note that, for some elements, a lower order polynomial might be sufficient to fit the data, but the verification made on the datasets suggests that we nevertheless correct without overfitting with a third-order polynomial.

Interestingly, the methodology described above does not allow calibration of the Zr and Nd abundances, as as insufficient number of stars are selected with the criteria previously described. Furthermore, we note that the distribution of the literature [Nd/Fe] values (GALAH-DR3, Buder et al. 2021) found for our solar neighborhood sample does not peak at 0 dex; therefore, an offset correction of [Nd/Fe] would be meaningless. For the same reason, we do not apply any correction to the GSP-Spec [Ce/Fe] abundances since our cross-match with literature (Hinkel et al. 2014; Abdurro'uf et al. 2022; Forsberg et al. 2019; Buder et al. 2021) reveals a similar offset in the cerium distribution with respect to the solar value. It is also important to note that the log(g) domain covered by the abundances of these three heavy elements is not very large, minimising gravity trends in the results.

Finally, it could be convenient for some scientific purposes to calibrate the abundance trends as a function of the effective temperature instead of the gravity. This is particularly the case when hot dwarf stars are included in the used sample. For this reason, we provide an example of this alternative calibration applied to [α/Fe] and illustrated in Fig. 15. The derived third-order polynomial is provided at the end of Table 4. In those cases, the log(g) variable in Eq. (3) should be replaced by the effective temperature. The use of similar calibrations as a function of $T_{\rm eff}$ for other chemical abundances or to correct gravity trends has to be evaluated by the user, depending on the target sample and scientific goals. Such calibrations are not provided here for clarity.

9.1.3. Summary of GSP-Spec MatisseGauguin biases and proposed solutions

In this section, we summarise the observed MatisseGauguin parameter and abundance biases, as well as the recommended solutions.

Effective temperature: No significant biases are observed in $T_{\rm eff}$. For hot stars, rotational mismatches could nevertheless affect the results.

Proposed solution: The *vbroadT* flag (or the vbroad value) has to be checked and/or used to clean the samples.

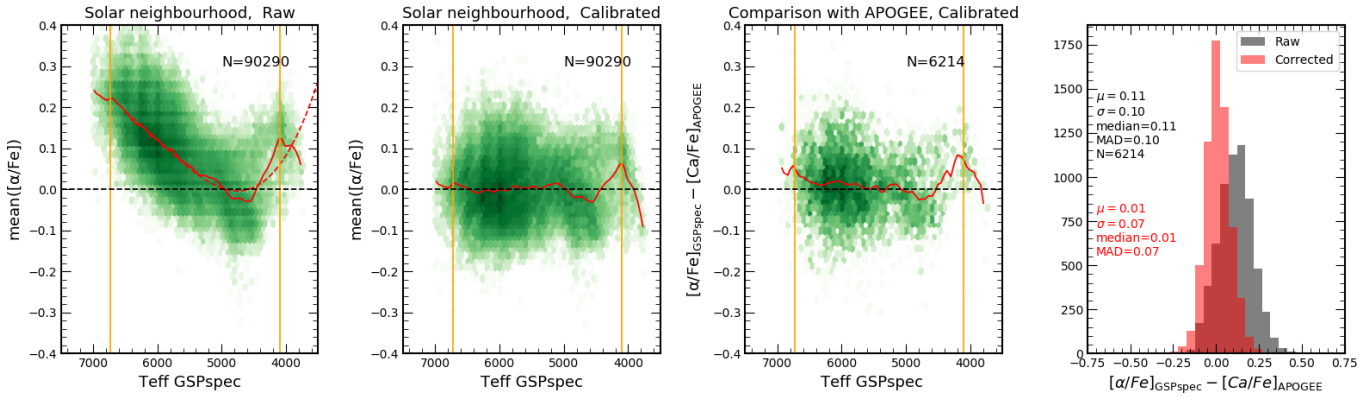


Fig. 15. Same as Fig. 14 but using the effective temperature as a reference parameter, instead of log(g). The associated polynomial coefficients and applicability intervals are provided in Table 3.

Surface gravity: A bias in log(g) is present. The median value is 0.3 dex on the entire parameter space. It shows a slight trend with $T_{\rm eff}$, getting worse as $T_{\rm eff}$ decreases. This could be related to the progressive dominance of the CaT lines as log(g) indicators (they become stronger along the giant branch as $T_{\rm eff}$ decreases), and to the absence of Paschen lines for $T_{\rm eff} \lesssim 5500$ K. No clear relation with line broadening mismatches seems to exist.

Proposed solution: A global correction is proposed based on literature data. For dwarf stars, a correction based on $T_{\rm eff}$ could offer more precise corrections, as the $T_{\rm eff}$ range is higher than the log(g) one. We generally advise to optimise the correction to the parameter space of the user.

Global metallicity: The observed log(g) bias seems to be associated with a slight [M/H] bias, presenting a similar behaviour with $T_{\rm eff}$ for $T_{\rm eff} \lesssim 5500$ K (not related to line broadening mismatches). For hotter stars, rotational mismatches could also cause a bias.

Proposed solution: Global corrections are proposed based on literature data of (i) field stars and (ii) open clusters. These corrections are only significant for giant stars in the low-log(g) regime. For hot stars, the *vbroadM* flag (or the vbroad value) has to be checked and/or used to clean the sample.

[α/Fe] and individual abundances: Atmospheric parameter biases are linked to abundance biases. In the GSP-Spec MatisseGauguin case, the main sources seem to be the log(g) bias and the rotational mismatch. On the contrary, the impact of the observed slight metallicity biases is probably reduced thanks to the fact that most abundances are derived with respect to iron. In the regime $T_{\rm eff} \lesssim 5500$ K, [α/Fe] and individual abundance biases are of small amplitude and show a very weak trend with $T_{\rm eff}$ and/or log(g). In the regime $T_{\rm eff} \gtrsim 5500$ K, abundance biases seem dominated by rotational missmatches.

Proposed solution: Global corrections are proposed as a function of log(g), based on a zero-point calibration to the Local Standard of Rest at solar metallicity. Alternatively, global corrections as a function of $T_{\rm eff}$ can be implemented (as in that proposed for [α/Fe]), with very similar results in the regime of $T_{\rm eff} \lesssim 5500$ K. For samples with a short $T_{\rm eff}$ and log(g) coverage, a constant shift to the solar value at [M/H] = 0 can be implemented if the user prefers to work with raw parameters (although the proposed calibrations are still valid). For dwarf stars, when including the hot temperature regime, a correction as a function of $T_{\rm eff}$ should be implemented. Alternatively, the sample can be cleaned using the *vbroadT*, *vbroadG,* and *vbroadM* (or the vbroad value). We generally advise optimisation of the correction to the parameter space of the user.

9.2. GSP-Spec ANN biases

The ANN workflow produces an alternative estimation of the stellar atmospheric parameters ($T_{\rm eff}$, log(g), [M/H], and [α/Fe]) that can be found in the *AstrophysicalParametersSupp* table. In the following, the observed ANN biases with respect to the literature are presented, proceeding in a similar way to that described for MatisseGauguin results (cf. Sect. 9.1.1).

Figure 16 shows the comparison with the literature for two ANN subsamples: the best quality, in green, and the medium quality, in grey. For our best-quality sample, we selected all the sources with the first eight flags equal to zero, excluding those with broadening and radial velocity issues, higher noise uncertainties, or extrapolations. The median offsets for the 274 592 stars of the best-quality sample are −114 K, −0.12 dex, and −0.24 dex for $T_{\rm eff}$, log(g), and [M/H], respectively, and the corresponding mean absolute deviations are 142 K, 0.28 dex, and 0.14 dex.

Compared to the literature, ANN results present a larger bias than MatisseGauguin in $T_{\rm eff}$ and [M/H], and a slightly lower bias in log(g). Nevertheless, these differences come partially from the fact that the ANN quality flags select a different reference subsample for the comparison statistics than the one used for GSP-Spec-MatisseGauguin. In particular, cooler giants are outside the ANN high-quality selection. Finally, the dispersion for the ANN parameterisation is also higher than for MatisseGauguin, particularly for $T_{\rm eff}$ and log(g).

We propose simple polynomial calibrations for $T_{\rm eff}$, log(g), and [M/H] based on the above comparison with the literature using the *best*-quality sample. It is important to note that the $T_{\rm eff}$ calibration of ANN is S/N dependent (cf. Appendix F) because the ANN algorithm was trained with synthetic spectra in five S/N levels (cf. Sect. 7).

We focus in the following on the high-S/N regime ($S/N_{\rm ANN} > 50$ corresponding to $S/N > 108$, cf. Table 1). For lower S/N values, we refer the reader to Appendix F, where the correct S/N optimisation of the algorithm is validated. We also highlight that, as the number of stars with $T_{\rm eff} > 6000$ K in the literature is small, the proposed corrections should not be applied beyond this temperature limit. The resulting calibrations take the form of:

$$X_{\rm calibrated} = X + \sum_{i=0}^{\rm deg} p_i \cdot X^i, \tag{4}$$

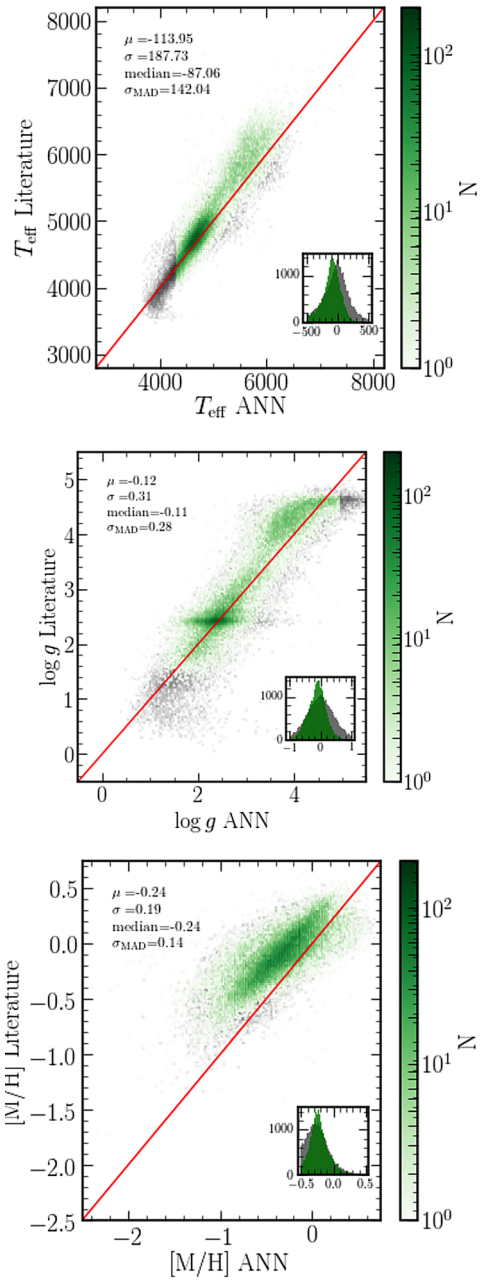


Fig. 16. Same as Fig. 11 but for GSP-Spec-ANN, published in the complementary table *AstrophysicalParametersSupp*. The reference high-quality subsample used for the comparison statistics is different from that shown in Fig. 11 (for GSP-Spec-MatisseGauguin), as imposed by the ANN quality flags.

where p_i coefficients for each parameter calibration can be found in Table 5. Moreover, similarly to MatisseGauguin results, we also suggest a calibration for [α/Fe], independent of literature data:

$$[\alpha/\mathrm{Fe}]_{\mathrm{calibrated}} = [\alpha/\mathrm{Fe}] + \sum_{i=0}^{3} p_i \cdot \log(g)^i. \quad (5)$$

In summary, although ANN parameters present slightly higher biases and uncertainties than MatisseGauguin ones (and they are therefore published in the complementary table AstrophysicalParametersSupp), their overall quality provides a methodologically different parametrisation, which could be useful, in particular, to test the MatisseGauguin classification in the low-S/N regime.

Table 5. Polynomial coefficients for the calibration of ANN parameters (at $S/N_{ANN} \sim 50$ for T_{eff}, see Appendix F for other S/N_{ANN} values).

Parameter	p_0	p_1	p_2	p_3
T_{eff}	12816	−8.1	1.65E−3	−1.07E−7
log(g)	−0.006	0.023		
[M/H]	0.092	−0.446	−0.07	
[α/Fe]	−0.038	0.099	−0.052	0.006

10. Illustration of GSP-Spec results

Illustrating all DR3 GSP-Spec results is obviously out of the scope of this paper. Two performance demonstration articles exclusively based on GSP-Spec MatisseGauguin parameters show their detailed application to Galactic chemo-dynamical studies of the disc and halo populations (Gaia Collaboration 2022a), and interstellar medium studies through the RVS diffuse interstellar band carrier (Gaia Collaboration 2022b). The homogeneous GSP-Spec treatment of the exhaustive all-sky RVS survey enables a chemo-physical parametrisation quality comparable to that of ground-based surveys of higher spectral resolution and wavelength coverage. Examples of this are the precision in the estimated individual chemical abundances (including heavy elements) allowing chemo-dynamical studies of Galactic stellar populations, DIB parameter estimation from individual spectra, and the precision in the atmospheric parameters providing clear constraints on stellar evolution models (see below).

In the following, we provide a few more examples of GSP-Spec results, focusing on (i) the number of parametrised stars in different quality regimes, (ii) the colour–effective temperature relation, (iii) an illustration of the T_{eff} spatial distribution, (iv) the atmospheric parameters of high-S/N spectra and associated constraints on stellar evolution models, and (v) the parametrisation of very metal-poor stars.

10.1. Number of parametrised stars in different quality regimes

As explained throughout this article, GSP-Spec has produced two sets of parameters (one from the MatisseGauguin workflow on the *AstrophysicalParameters* table, and another from the ANN workflow on the *AstrophysicalParametersSupp* table) for about 5.6 million stars from their RVS spectra. The total number of derived atmospheric parameters by both workflows and the number of GAUGUIN chemical abundances are illustrated in Fig. 17 (left panel for MatisseGauguin and right panel for ANN) and Fig. 18, respectively. In both figures, the total number of published parameters is shown together with the corresponding number for the best parametrised stars (from a high-quality selection, where all parameter flags are set to zero, including the abundance flags from MatisseGauguin). It is important to note that imposing that the full flag chain be equal to zero corresponds to very demanding requirements, including very low associated uncertainties. This selects about two million stars for the atmospheric parameters, whereas, for the chemical abundances (cf. Fig. 18), the number of estimates varies over several orders of magnitude from one element to another,

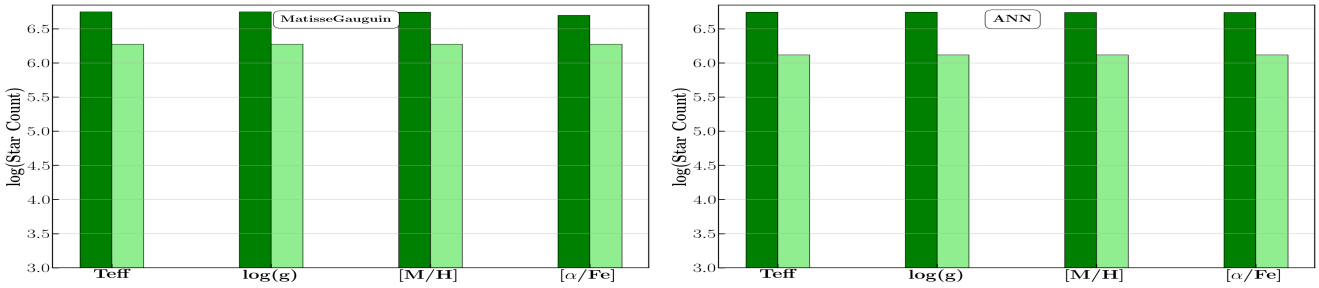


Fig. 17. Number of stars whose atmospheric parameters have been derived by MatisseGauguin and ANN (*left and right panels*, respectively). The dark green histograms refer to the whole sample whereas the light-green ones show only the very best parametrised stars with all their parameter quality flags equal to zero.

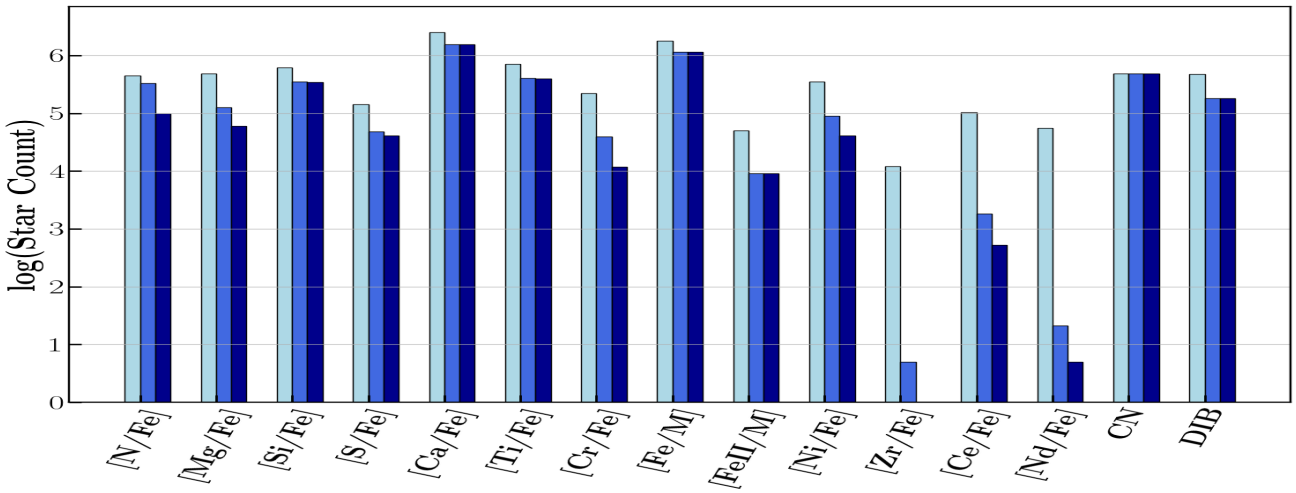


Fig. 18. Same as Fig. 17 but for the individual abundances derived by GAUGUIN plus the CN-abundance proxy and the DIB. The light-blue histogram (left bars) refers to the whole sample. The two other sets of bars (central and right bars) show only the very best stars with all their parameter flags and their abundance uncertainty quality equal to zero. The abundance upper limit flag is lower than or equal to one and equal to zero for the medium-blue and dark-blue bars, respectively.

as expected. In particular, calcium and iron (Fe I) are the most often derived species with estimates for around two millions stars, thanks to the Ca prominent lines and the numerous available iron lines. Abundances of heavy elements are derived for up to 10^4–10^5 stars, although these numbers strongly decrease when all the flags are used to filter (Ce being the heavy element with the highest number of estimates). However, we point out that this very strict quality filtering can be relaxed to increase number statistics, depending on the scientific goals of the user.

10.2. Colour–temperature relation

A classical way of validating effective temperature estimates is to verify their expected correlation with stellar colour. Figure 19 shows the trend between the (BP–RP) colour and the GSP-Spec $T_{\rm eff}$ estimates from MatisseGauguin. To consider the effect of extinction on the (BP–RP) colour, the points are colour coded according to the EW of the DIB derived for the same stars by GSP-Spec (only stars with the DIB flag equal to zero have been selected). First, it is observed that the lower envelope of the distribution corresponds to the lower DIB EW values, as expected from the correlation between DIB absorption and extinction. To quantify this observation, the median values of the distribution for the stars with a DIB EW lower than 0.05 Å (blue circles) can be compared to those whose DIB EW is equal to the median value of the distribution (0.07 Å for dwarf stars in the left panel, and 0.12 Å for giants in the right panel), plus a dispersion of ± 0.01 Å. Second, the observed relation is compared to a $T_{\rm eff}$ derived from the Casagrande et al. (2021) prescription (black dots) based on an implementation of *Gaia* and 2MASS photometry in the InfraRed Flux Method. No extinction has been considered in this case and the corresponding median values are shown as white circles. The Casagrande et al. (2021) predictions are in very good agreement with the low-extinction envelope of the GSP-Spec distribution (blue circles), validating the global behaviour of the estimated temperatures.

To complement this analysis, Fig. 20 presents the metallicty correlations of the colour–temperature relation for targets with all the parameter flags equal to zero. Again, the expected metallicty trend is observed in the low-extinction envelope. Interestingly, the higher extinction region above the lower envelope of the distribution is mainly occupied by metal-rich stars. This is expected from the fact that metal-rich stars are preferentially placed near the Galactic plane, where the interstellar extinction is higher.

10.3. Sky distribution of effective temperature estimates

Figure 21 presents the global all-sky spatial distribution in Galactic coordinates of the stars parametrised by GSP-Spec, colour-coded with their MatisseGauguin effective temperature

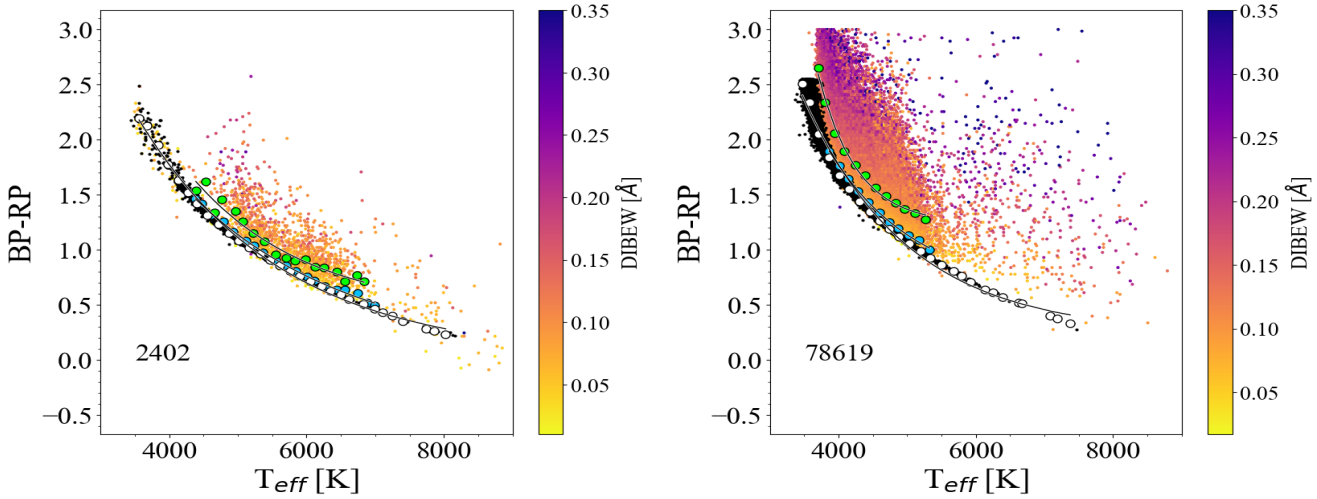


Fig. 19. Trend of (BP-RP) colour with GSP-Spec effective temperature produced by the MatisseGauguin workflow for dwarfs (*left panel*) and giants (*right panel*). The colour code indicates the estimated DIB EW, which increases with interstellar absorption (the DIB flag has been imposed to be equal to zero). Blue circles show the median values of the distribution for the stars with a DIB EW lower than 0.05 Å. Green circles are the median values for stars whose DIB EW is equal to the median value of the distribution (0.07 Å for dwarf stars on the *left panel*, and 0.12 Å for giants on the right panel), plus a dispersion of ±0.01 Å. Black dots (and white circles) are the values (and their median) predicted by the Casagrande et al. (2021) relation, assuming no extinction.

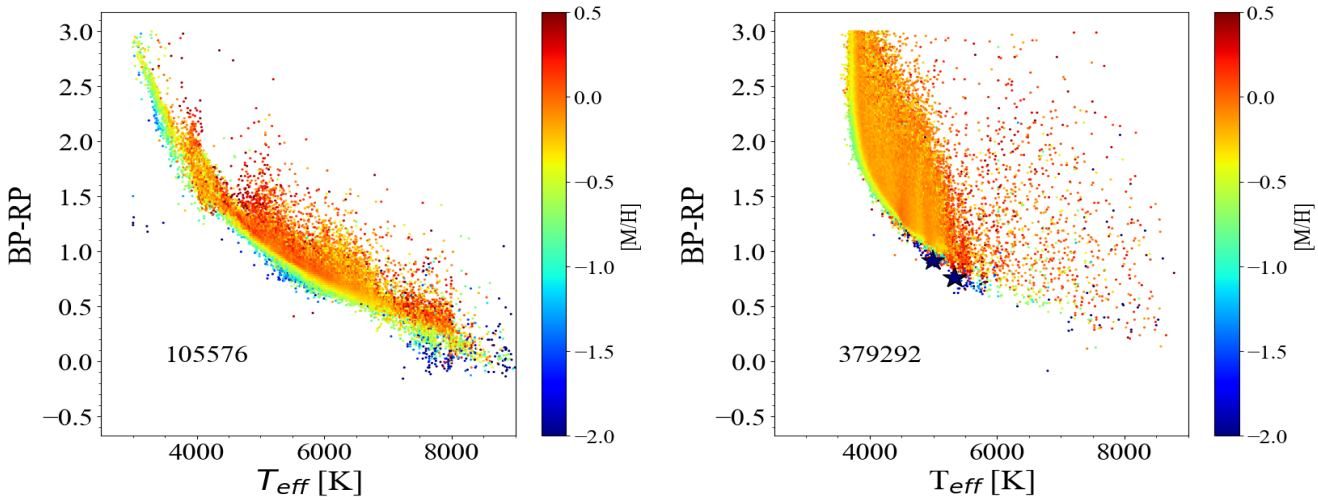


Fig. 20. Same as Fig. 19 but using the estimated stellar metallicity [M/H] as colour code. The selected stars have the first 13 quality flags in the gspspec flagging chain equal to zero. Two extremely metal-poor stars, discussed in Sect. 10.5, are indicated by star symbols. The number of stars is indicated in each panel.

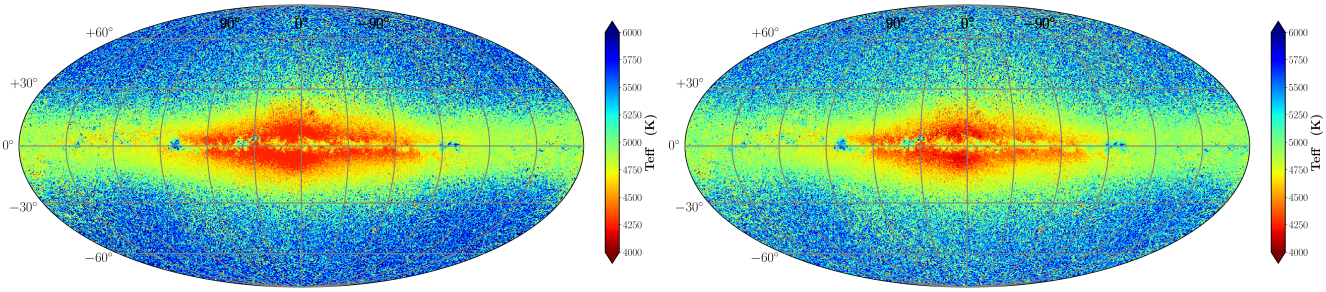


Fig. 21. Milky Way as revealed by the GSP-Spec effective temperature estimated by MatisseGauguin (*left*) and ANN (*right*). These HEALPix maps in Galactic coordinates have a spatial resolution of 0.46°. The colour code corresponds to the median of $T_{\rm eff}$ in each pixel.

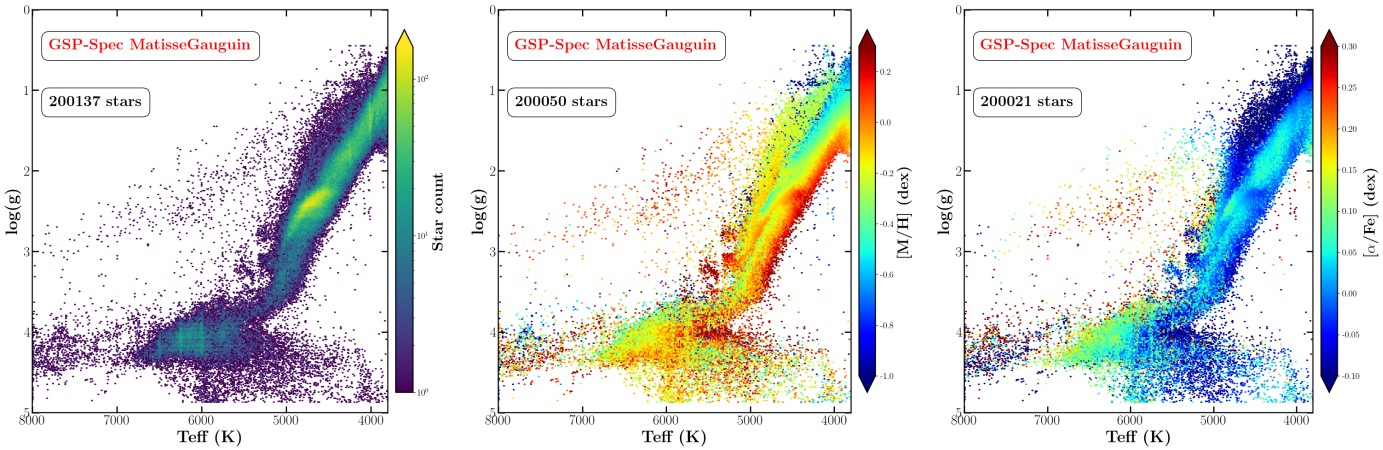


Fig. 22. Kiel diagrams for the MatisseGauguin output parameters (stored in the main DR3 astrophysical parameters table) for high-quality spectra ($S/N > 150$) and excluding high-rotating stars ($vbroadT = vbroadG = vbroadM = 0$) and possibly misclassified very cool giants ($KMtypestars = 0$). The colour codes of the different panels show the stellar density (*left panel*) and the median of [M/H] and [α/Fe] per point (*central and right panels*, respectively). The proposed log(g) and [α/Fe] calibrations are applied.

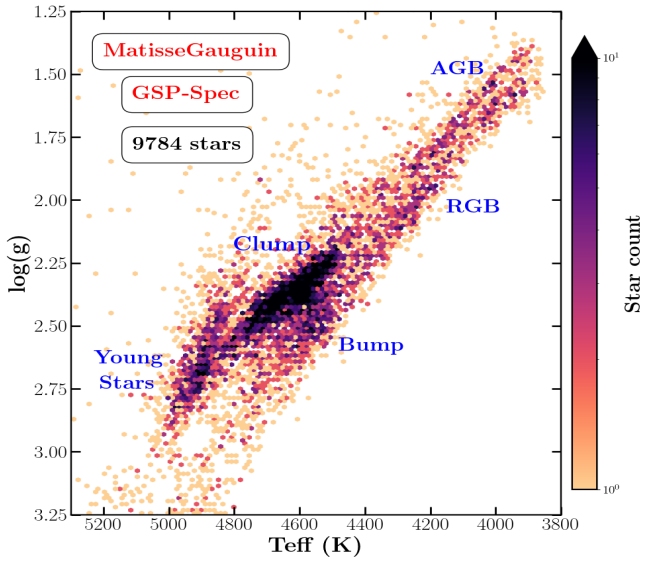


Fig. 23. Zoom onto the MatisseGauguin Kiel diagram for stars in a very restricted metallicity domain, $-0.05 < [M/H] < 0.00$ dex, and with high-quality spectra. The RGB and AGB sequences appear as two resolved parallel tracks. The very close-by RGB bump and HB clump are also isolated. A sequence of young stars (with ages of less than ~1 Gyr) can be identified in the hotter side, with an overdensity at around $T_{\rm eff} \sim 5000$ K. These are second red clump objects, i.e. massive stars burning He in their core.

(5 576 282 stars, left panel) and their ANN effective temperature (5 524 387 stars, right panel). Both figures show the giant star population dominating the Galactic disc and bulge regions. The in-plane interstellar extinction pattern can also be noticed by its effect on the underlying parameterised populations: in higher extinction regions, cool giant stars observable at large distances become too faint in the RVS wavelength domain, and the median of the temperature distribution becomes hotter. Finally, nearby fainter dwarf stars in the foreground dominate the regions above and below the Galactic plane, increasing the median $T_{\rm eff}$ values. It can be observed that ANN provides lower temperatures for these stars. For more details on the GSP-Spec selection function, we refer to Gaia Collaboration (2023a) (see their Sect. 3).

10.4. Atmospheric parameters of high-S/N spectra

To illustrate the GSP-Spec atmospheric parameter estimates in the high-S/N regime, we selected all the stars with $S/N > 150$, excluding high-rotating stars and potentially misclassified cool giants (imposing $vbroadT$=$vbroadG$=$vbroadM = 0$ and $KMtypestars = 0$, respectively). This selects a sample of nearly 202 000 stars.

Figure 22 presents the MatisseGauguin parametrisation of the selected objects in different Kiel diagrams colour coded according to stellar density (left panel), [M/H] (middle panel), and [α/Fe] (right panel). We applied the log(g) calibration proposed in Sect. 9 and the [α/Fe] calibration reported in Table 4 (fourth-order polynomial, without applying the suggested cuts in log(g) in order to show a complete Kiel diagram). The parameters precision can be assessed from the well-defined evolutionary sequences. For instance, the clearly distinguishable red clump presents a metallicity dependence that is independent from that of the red giant branch, as expected. Additionally, younger, more massive stars populate the hotter metal-rich sequence with logg $\lesssim 3$. These stars are located in the Milky Way spiral arms (cf. Gaia Collaboration 2023a). It is worth noting that, in the high-S/N regime, the algorithm shows overfitting patterns (overdensity features at the reference grid points). This can be observed in the left panel of Fig. 22 for the $T_{\rm eff}$. The log(g) values are not affected in this figure because they have been calibrated.

The precision of the Matisse-Gauguin atmospheric parameters (without any use of astrometric inputs) can also be appreciated from Fig. 23, which shows a zoom into the Kiel diagram of the stars in a very restricted metallicity domain, $-0.05 < [M/H] < 0.00$ dex (defined using the upper and lower confidence values in the form `mh_gspspec_upper`< 0.00 dex and `mh_gspspec_lower`> -0.05 dex). In addition, only stars with $T_{\rm eff}$>3750 K, $KMgiantPar = 0$ and `logchisq_gspspec`<–3.75 were selected so as to avoid classification problems at the very cool end of the giant branch. It can be appreciated that the RGB bump[18] is resolved as an overdensity feature at $T_{\rm eff} \sim 4600$ K and log(g) ~ 2.5. The very high

[18] The RGB bump corresponds to the arrival of the narrow burning H-shell to the sharp chemical discontinuity in the H-distribution profile caused by the penetration of the convective envelope.

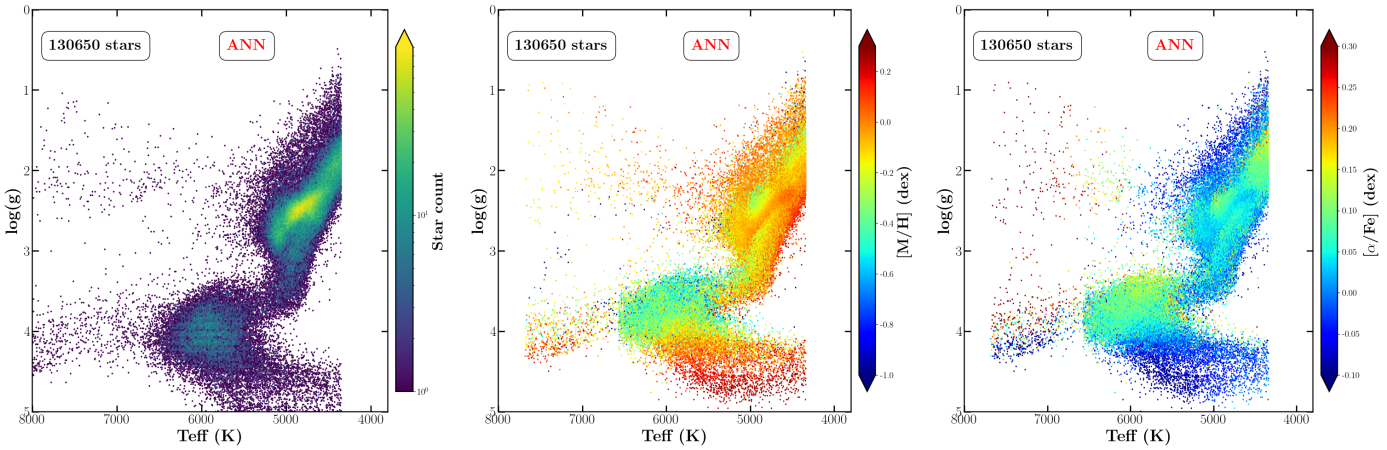


Fig. 24. Same as Fig. 22 but for the ANN output parameters (stored in the supplementary DR3 astrophysical parameters table). In this case, we imposed that the first eight quality flags in Sect. 8 be equal to zero. The calibrations proposed in Sect. 9.2 were applied. Although a larger dispersion is observed with respect to MatisseGauguin, a general agreement exists, supporting the coherence of the two methodologically independent analysis.

parameter precision allows us to separate this RGB feature from the nearby horizontal branch clump visible as a narrow elongated feature between $4500 < T_{\rm eff} < 4800$ K and $2.20 < \log(g) < 2.50$. Moreover, Fig. 23 shows the capability of these very high-quality GSP-Spec parameters to disentangle the extremely close-by red giant branch and asymptotic giant branch sequences, which appear as two parallel tracks for $\log(g) < 2.25$. Finally, the overdensity located around $T_{\rm eff} \sim 5000$ K and $2.50 < \log(g) < 2.80$ corresponds, as mentioned above, to the evolutionary sequence of young stars of about less than 1 Gyr (cf. Gaia Collaboration 2023a for a more detailed analysis of these stars tracing the disc spiral arms). This will put important constraints on stellar evolution models, and specifically on the mass and metallicity dependencies of the red clump, the RGB bump, and the RGB and AGB behaviours.

Figure 24 shows the ANN results for the same stars, after imposing that the first eight quality flags in Sect. 8 be equal to zero and the four parameter calibrations proposed in Sect. 9.2. On one hand, a general agreement is observed with respect to MatisseGauguin. In particular, a well-defined Red Clump and a comparable metallicity trend for giant stars are observed, although with a higher dispersion and an underabundance of metal-poor stars in ANN results ([M/H] ≤ -1.0 dex, partly explainable by the temperature cut in the cool regime due to calibration boundaries). On the other hand, the metallicity and [α/Fe] distributions differ from the MatisseGauguin one for dwarf stars, presenting an unexpected trend with gravity. Despite the higher dispersion of ANN parameters, its overall agreement with MatisseGauguin brings support to the coherence of the two methodologically independent analyses.

Additionally, Fig. 25 shows the [α/Fe] versus [M/H] distribution from the MatisseGauguin analysis for the selected high-S/N spectra, applying the suggested cuts in $\log(g)$ for the [α/Fe] calibration and imposing $T_{\rm eff} \leq 6000$ K and $vbroad \leq 10$ km s^{-1}. These last two filters help to control the quality of the [α/Fe] by reducing second-order temperature trends and refining the filtering performed by the $vbroadT$, $vbroadG$, and $vbroadM$ flags. The halo and disc sequences can be observed, with the thick disc sequence joining the thin disc one at a metallicity of around -0.4 dex. It is also worth noting that, as expected from chemical evolution models, the thin disc sequence continues to decrease at supersolar metallicities. As shown in

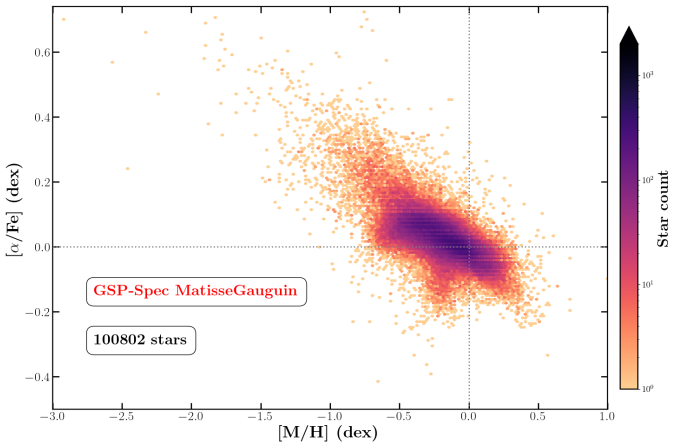


Fig. 25. [α/Fe] versus [M/H] for the same MatisseGauguin stars as in Fig. 22 but applying the recommended gravity interval for the calibration (Table 4) and imposing $T_{\rm eff} \leq 6000$ K and $vbroad \leq 10$ km s^{-1}.

Gaia Collaboration (2023a), the [α/Fe] clearly correlates with the kinematical properties of stellar populations. Moreover, the *Gaia*-Enceladus sequence of accreted stars (Helmi et al. 2018) is also distinguishable (lower [α/Fe] values than those for typical thick discs and halos in the metal-poor regime). Finally, a group of low-[α/Fe] stars at a metallicity of about [M/H] ~ -0.4 dex is also visible. This corresponds to young massive stars in the spiral arms (for a discussion about the chemical properties of these stars see, Gaia Collaboration 2023a). It is worth noting that, as mentioned in Sect. 9, GSP-Spec [α/Fe] estimates are dominated by the [Ca/Fe] abundance. We refer to Gaia Collaboration (2023a) for a detailed illustration of individual α-element abundances, including Mg, Ca, Si, S, and Ti, as well as other chemical species including N, iron-peak elements, and heavy elements.

10.5. Parametrisation of extremely metal-poor stars

Metal-poor stars are relics of the most ancient formation epochs of the Milky Way, and in particular, they are crucial for disentangling the sequence of satellite mergers contributing to the Galaxy build-up (e.g. Helmi 2020). For this reason, they are

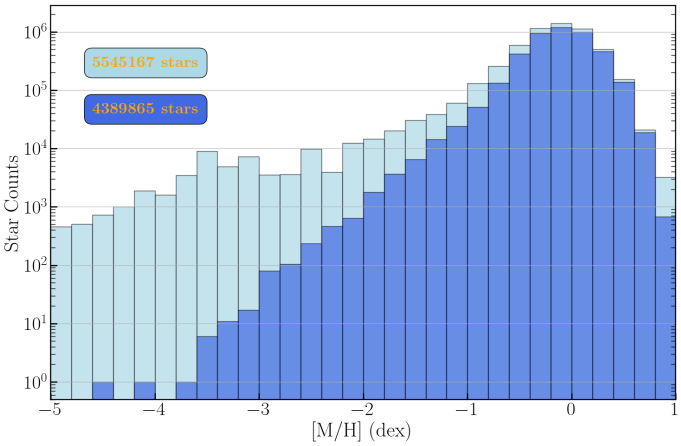


Fig. 26. Metallicity distributions for the MatisseGauguin parametrised stars. The light-blue histogram refers to the whole sample without any filtering. The medium-blue histogram presents a very strict filtering selecting stars with the best derived metallicities (see associated text for more details).

the priviledged targets of several spectroscopic surveys from the ground like Pristine (Starkenburg et al. 2017). However, the lack of spectral signatures in metal-poor spectra reduces the information on the stellar parameters, increasing the uncertainties and making their parametrisation challenging. In the following, we illustrate the GSP-Spec capabilities to parametrise not only metal-poor, but also ultra-metal-poor ([M/H] < −3.0 dex) stars, providing suggestions on the necessary filters to apply.

Figure 26 shows the GSP-Spec MatisseGauguin metallicity distribution in a logarithmic scale. The light-blue histogram refers to the complete sample without any quality filtering. In this distribution, which is useful for a rough stellar selection, there are about 66 000 stars with [M/H] ≤ −2.0 dex. It is nevertheless observed that the profile of the histogram is unexpectedly flat in the ultra-metal-poor regime. This is due to the $T_{\rm eff}$ limitations of the GSP-Spec reference spectra grid ($T_{\rm eff} \leq 8000$ K) inducing a $T_{\rm eff}$–[M/H] degeneracy. This problem can be satisfactorily resolved, as shown by the medium-blue histogram, by (i) disregarding the [M/H] values for stars with the first six characters of the GSP-Spec flagging chain ≥2, which limits the parameterisation biases, (ii) filtering out the metallicities of stars with *extrapol* flag (eighth character of the flagging chain) ≥3, which limits the extrapolation issues, (iii) eliminating the ultra-metal-poor stars hotter than 6000 K, which conservatively filters out metallicities with unreliable uncertainties due to border effects (in $T_{\rm eff}$ and [M/H]), and (iv) filtering out possible remaining GSP-Spec misclassifications of very hot stars with stellar types O and B as estimated by the Extended Stellar Parameteriser of Hot Stars (ESP-HS) and reported in the *AstrophysicalParameters* table as *spectraltype_esphs*. To complete the previous filters, which are optimised for the very metal-poor regime, the medium-blue histogram of Fig. 26 filters out the metallicities of stars with $T_{\rm eff} < 3500$ or $\log(g) > 4.9$ or *KMgiantPar* > 0. The filtering implemeted by the *KMgiantPar* flag, which controls the quality of the parameterisation of very cool giants, can be slightly extended, as reported in Gaia Collaboration (2023a), disregarding the metallicity of stars with $T_{\rm eff} < 4150$ K and $2.4 < \log(g) < 3.8$. Thanks to these different quality filters, the medium-blue histogram presented in Fig. 26 recovers the expected decrease in the number of stars in the very metal-poor regime, reporting only very reliable results within the corresponding uncertainties. Among these, there are about 300 with [M/H] < –2.5 dex and about 40 stars with [M/H] < –3.0 dex.

To confirm that GSP-Spec is indeed able to correctly estimate the parameters of very metal-poor and ultra-metal-poor stars, we show the RVS spectrum of two of them, which were randomly chosen among the highest S/N spectra (Figs. 27 and 28). The very few lines present in these spectra are extremely weak (except those of Ca II) and, as a consequence, no individual abundances were derived for both stars. [M/H] is therefore estimated only from very few available weak calcium or iron lines. The careful visual inspection of the synthetic spectra fit of the corresponding RVS spectra corroborates the very metal-poor nature of both stars.

We first validated the atmospheric parameters of *Gaia* DR3 6268770373590148224 (Fig. 27, [M/H] = –3.19 dex) by computing several synthetic spectra with parameters found within the upper and lower MatisseGauguin uncertainties, i.e. between 5097 and 5456 dex for $T_{\rm eff}$ and [2.38, 2.76] for $\log(g)$. The parameter values are confirmed within the uncertainties. The global fit is excellent and [M/H] is indeed found between the published range [−3.6, −2.7]. We note again that this rather large range of [M/H] is caused by the quasi absence of lines in its RVS spectrum. We also checked the literature for this star and found it to be known as the peculiar star HD 140283, already studied in several articles, the first one going back about 70 years (Chamberlain & Aller 1951). Its most recent published parameters seem to converge towards a slightly hotter star with the majority of published [M/H] being found around −2.5 dex, in agreement with our estimates taking into account the literature error bars. Its [α/Fe] value of 0.56 dex and its kinematical parameters (a Galactic azimuthal velocity of ∼29 km s^{-1}, as taken from Gaia Collaboration 2023a) make this star a typical halo representative.

The second example, *Gaia* DR3 6477295296414847232, is shown in Fig. 28. This star has [M/H] = −3.52 dex (and confidence values between −3.94 dex and −3.17 dex), a $T_{\rm eff}$ of 4994 K (with confidence values between 4781 K and 5071 K) and a $\log(g)$ of 2.13 (with confidence values between 1.5 and 2.26). Again, this parameterisation is confirmed by visual inspection of the spectrum fit in Fig. 28. Additionally, to exclude [M/H] values higher than the reported upper confidence level ([M/H] =–3.07 dex), Fig. 28 includes a synthetic spectrum (in orange) corresponding to [M/H] = –2.5 dex, which clearly overestimates the CaII and Fe lines depth and width and confirms that a lower metallicity is required, as in the one estimated by GSP-Spec. Finally, the literature inspection of this object, identified as the peculiar star HD 2000654, leads to 23 different metallicity estimates with a median value of −2.9 dex ± 0.25 dex, in agreement with our results. It is interesting to note that Roederer et al. (2014) and Hansen et al. (2018) classify this star has an *r*-process enhanced object of *r–I* type, with a metallicity of −3.13 dex and −2.91 dex, respectively. In addition, based on high-resolution spectra, Roederer et al. (2014) reports a [Ca/Fe] = 0.47 dex ± 0.17 in agreement with our calibrated [α/Fe] = 0.62 dex, a carbon abundance of [C/Fe] = 0.31 dex, and a nitrogen abundance of [N/Fe] = −0.4 dex. More recently, Placco et al. (2018), using medium-resolution spectra, classify this star as a CEMP-II object, with a carbon enhancement of [C/Fe] = 0.71 dex. It is important to remark that no sign of carbon enhancement, perturbing the metallicity estimate, is present in the RVS spectra due to the absence of CH molecular lines. In summary, the literature results again validate the GSP-Spec parameterisation of this ultra-metal-poor star, including its [α/Fe] estimates.

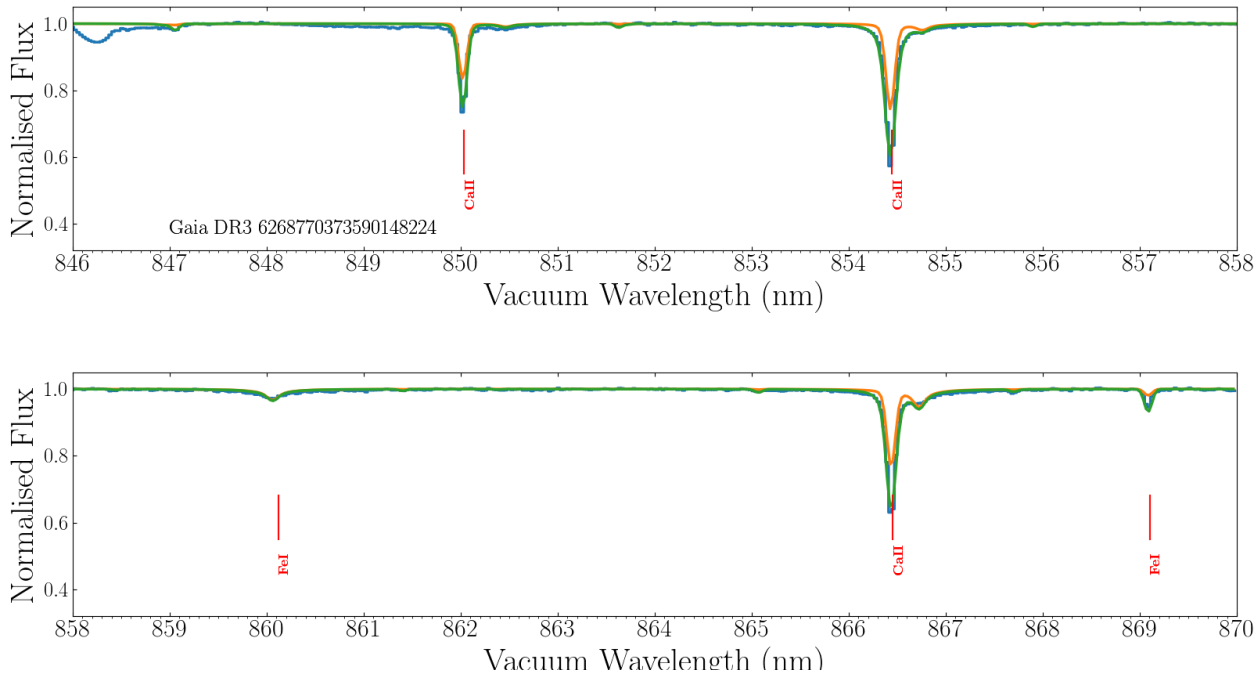


Fig. 27. RVS spectrum (blue histogram) of the very metal-poor star *Gaia* DR3 6268770373590148224 whose MatisseGauguin atmospheric parameters are $T_{\rm eff} = 5331$ K, $\log(g) = 2.54$, [M/H] = −3.19 dex, and [α/Fe] = 0.56 dex (S/N = 419). The model spectra correspond to the lower and upper [M/H] values (−3.60 and −2.71 dex in orange and green, respectively). No rotational profile was applied (suspected low-rotating star). See text for more details.

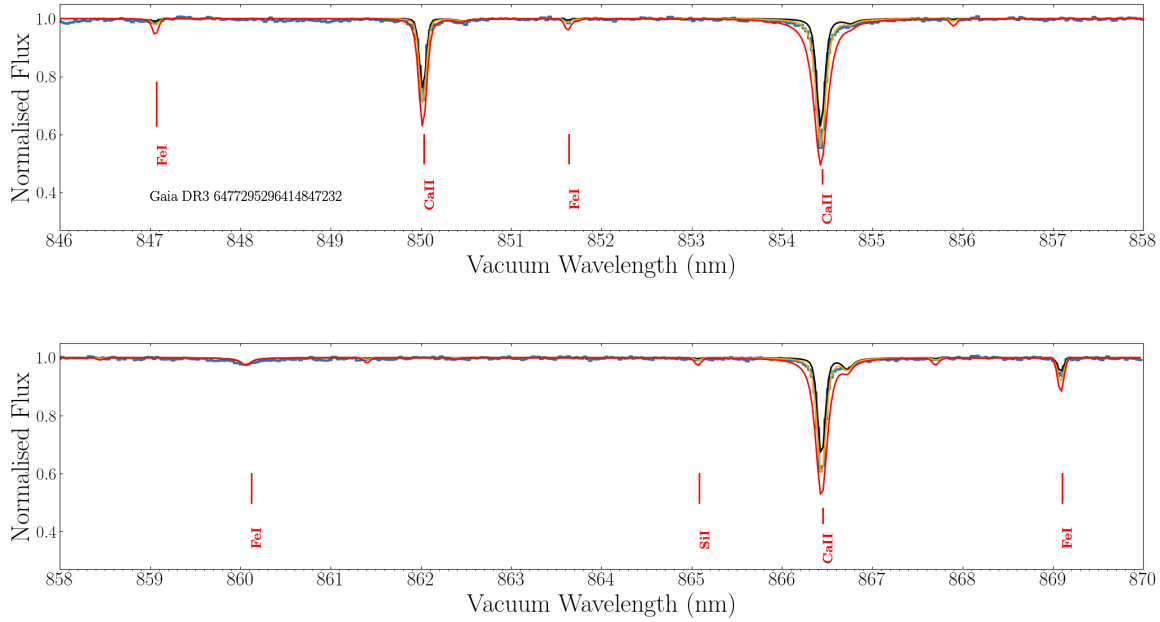


Fig. 28. RVS spectrum (blue histogram) of the ultra-metal-poor star *Gaia* DR3 6477295296414847232 whose MatisseGauguin atmospheric parameters are $T_{\rm eff} = 4994$ K $\log(g) = 2.13$, [M/H] = −3.52 dex, and [α/Fe] = 0.68 dex (S/N = 236). The model spectra correspond to the lower and upper [M/H] values (−3.52 and −3.07 dex in black and orange, respectively). A spectrum with [M/H] = −2.5 dex is also shown (red line) to definitively exclude such higher metallicities. No rotational profiles were considered.

11. Summary and conclusions

Here, we summarise the stellar parametrisation of *Gaia* RVS spectra performed by the GSP-Spec module and published as part of *Gaia* DR3. The goals, the input data, the used methodologies, and the validation are presented in detail. The resulting catalogues are published in the *AstrophysicalParameters* table (for the GSP-Spec MatisseGauguin workflow, including stellar atmospheric parameters, individual chemical abundances, a cyanogen differential EW, and DIB feature parameters), and in the *AstrophysicalParametersSupp* (for the ANN workflow providing atmospheric parameters). The GSP-Spec catalogue flags are also carefully defined and guidance for their use is illustrated with examples. We highly recommend future users of the GSP-Spec parameters to adopt these flags for their specific science cases.

With about 5.6 million stars, the *Gaia* DR3 GSP-Spec all-sky catalogue is the largest compilation of stellar chemo-physical parameters ever published and the first of its kind based on data acquired in space. The extreme homogeneity of the analysis combined with continuous data collection for almost three years enable a careful spectroscopic data reduction, a detailed modelling of systematic errors, and consequently, higher number statistics and a parametrisation quality that is comparable to that of ground-based surveys of higher spectral resolution and wavelength coverage.

GSP-Spec parameters open new horizons in stellar, Galactic, and insterstellar medium studies. In addition to the scientific performance analysis of GSP-Spec data published in Gaia Collaboration (2023a,b), we illustrate the precision of the parameters here with (i) the colour–temperature relation, (ii) the Kiel diagrams and the [α/Fe] vs. [M/H] distribution in the high-S/N regime ($S/N > 150$, more than 2 million stars), (iii) our ability to disentangle different evolutionary stages of giant stars that are extremely close-by in the parameter space (RGB/AGB, bump/clump), and finally, (iv) a demonstration of the capability of GSP-Spec in the challenging parametrisation of metal-poor and extremely metal-poor stars.

Finally, it is worth noting that, as GSP-Spec is one of the parametrisation modules activated at the end of the DPAC analysis chain, this *Gaia* third data release is actually the first GSP-Spec data release. The acquired experience will benefit future releases, for which the number of parametrised stars will be a factor of ten larger (~50 million stars) as a result of the spectra S/N increase with observing time. It is important to note that the present data set is already at least a factor 8 larger than previous individual ground-based catalogues and a factor 3 larger than their very heterogeneous joint compilation. GSP-Spec is therefore exploring Galactic regions that we had previously only hypothesised from models (based on low number statistics). Thanks to the *Gaia* RVS GSP-Spec chemo-physical parametrisation, we now have a privileged view of the sky from beyond Earth.

Acknowledgements. This work presents results from the European Space Agency (ESA) space mission *Gaia* (https://www.cosmos.esa.int/gaia). *Gaia* data are processed by the *Gaia* Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the *Gaia* MultiLateral Agreement (MLA). The *Gaia* archive website is https://archives.esac.esa.int/gaia. Acknowledgments from the financial institutions are given in Appendix H. We sincerely thank the stellar atmosphere group in Uppsala for providing the MARCS model atmospheres, B. Plez for having developped and maintaining the TURBOSPECTRUM package and, M. Bergemann for providing before publication the adopted relation between microturbulent velocity and atmospheric parameters. We also thank A. Bragaglia for her comments on the manuscript and the anonymous referee for very useful comments and suggestions. Finally, part of the calculations have been performed with the high-performance computing facility SIGAMM, hosted by the Observatoire de la Côte d'Azur. We acknowledge financial supports from the french space agency (CNES), Agence National de la Recherche (ANR 14-CE33-014-01) and Programmes Nationaux de Physique Stellaire & Cosmologie et Galaxies (PNPS & PNCG) of CNRS/INSU. ES, ARB, PdL, GK and MS acknowledge funding from the European Union's Horizon 2020 research and innovation program under SPACE-H2020 grant agreement number 101004214 (EXPLORE project).

References

Abdurro'uf, A. K., Aerts, C., Silva Aguirre, V., et al. 2022, ApJS, 259, 35
Allende Prieto, C. 2016, Astron. Nachr., 337, 837
Allende Prieto, C., Beers, T. C., Wilhelm, R., et al. 2006, ApJ, 636, 804
Andrae, R., Fouesneau, A., Sordo, R., et al. 2023, A&A, 674, A27 (*Gaia* DR3 SI)
Bailer-Jones, C. A. L., Andrae, R., Arcay, B., et al. 2013, A&A, 559, A74
Bijaoui, A., Recio-Blanco, A., de Laverny, P., & Ordenovic, C. 2010, in ADA 6- Sixth Conference on Astronomical Data Analysis, J. L. Starck, M. Saber Naceur, & R. Murtagh, 9
Bijaoui, A., Recio-Blanco, A., de Laverny, P., & Ordenovic, C. 2012, Stat. Method. - Elsevier, 9, 55
Birch, K. P., & Downs, M. J. 1994, Metrologia, 31, 315
Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150
Cannon, A. J., & Pickering, E. C. 1918, Ann. Harvard College Obs., 91, 1
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1
Casagrande, L., Lin, J., Rains, A. D., et al. 2021, MNRAS, 507, 2684
Castro-Ginard, A., Jordi, C., Luri, X., et al. 2022, A&A, 661, A118
Catmull, E., & Rom, R. 1974, in Computer Aided Geometric Design, R. E. Barnhill, & R. F. Riesenfeld (Academic Press), 317
Chamberlain, J. W., & Aller, L. H. 1951, ApJ, 114, 52
Contursi, G., de Laverny, P., Recio-Blanco, A., & Palicio, P. A. 2021, A&A, 654, A130
Creevey, O., Sordo, R., Pailler, F., et al. 2023, A&A, 674, A26 (*Gaia* DR3 SI)
Cropper, M., Katz, D., Sartoretti, P., et al. 2018, A&A, 616, A5
Dafonte, C., Fustes, D., Manteiga, M., et al. 2016, A&A, 594, A68
de Laverny, P., Recio-Blanco, A., Worley, C. C., & Plez, B. 2012, A&A, 544, A126
de Laverny, P., Recio-Blanco, A., Worley, C. C., et al. 2013, The Messenger, 153, 18
Forsberg, R., Jönsson, H., Ryde, N., & Matteucci, F. 2019, A&A, 631, A113
Gaia Collaboration (Recio-Blanco, A., et al.) 2023a, A&A, 674, A38 (*Gaia* DR3 SI)
Gaia Collaboration (Schultheis, M., et al.) 2023b, A&A, 674, A40 (*Gaia* DR3 SI)
Gaia Collaboration (Vallenari, A., et al.) 2023c, A&A, 674, A1 (*Gaia* DR3 SI)
Gershman, S. J., & Blei, D. M. 2012, J. Math. Psychol., 56, 1
Gilmore, G., Randich, S., Asplund, M., et al. 2012, The Messenger, 147, 25
Gilmore, G., Randich, S., Worley, C. C.et al. 2022, A&A, 666, A120
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Grevesse, N., Asplund, M., & Sauval, A. J. 2007, Space Sci. Rev., 130, 105
Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, 951
Hansen, T. T., Holmbeck, E. M., Beers, T. C., et al. 2018, ApJ, 858, 92
Helmi, A. 2020, ARA&A, 58, 205
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85
Hinkel, N. R., Timmes, F. X., Young, P. A., Pagano, M. D., & Turnbull, M. C. 2014, AJ, 148, 54
Jofré, P., Heiter, U., & Soubiran, C. 2019, ARA&A, 57, 571
Katz, D., Munari, U., Cropper, M., et al. 2004, MNRAS, 354, 1223
Katz, D., Sartoretti, P., Guerrier, A., et al. 2023, A&A, 674, A5 (*Gaia* DR3 SI)
Kordopatis, G., Recio-Blanco, A., de Laverny, P., et al. 2011, A&A, 535, A106
Kordopatis, G., Hill, V., Irwin, M., et al. 2013, A&A, 555, A12
Kos, J. 2017, MNRAS, 468, 4255
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Manteiga, M., Ordóñez, D., Dafonte, C., & Arcay, B. 2010, PASP, 122, 608
Martell, S. L., Sharma, S., Buder, S., et al. 2017, MNRAS, 465, 3203
Morgan, W. W., Keenan, P. C., & Kellman, E. 1943, An Atlas of Stellar Spectra, With an Outline of Spectral Classification (Chicago: The University of Chicago press)
Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989
Perdigon, J., de Laverny, P., Recio-Blanco, A., et al. 2021, A&A, 647, A162
Placco, V. M., Beers, T. C., Santucci, R. M., et al. 2018, AJ, 155, 256
Plez, B. 2012, Astrophysics Source Code Library [record ascl:1205.004]
Randich, S., Gilmore, G., Magrini, L., et al. 2022, A&A, 666, A121
Recio-Blanco, A. 2014, in Setting the scene for Gaia and LAMOST, eds. S. Feltzing, G. Zhao, N. A. Walton, & P. Whitelock, 298, 366
Recio-Blanco, A., Bijaoui, A., & de Laverny, P. 2006, MNRAS, 370, 141
Recio-Blanco, A., de Laverny, P., Kordopatis, G., et al. 2014, A&A, 567, A5
Recio-Blanco, A., de Laverny, P., Allende Prieto, C., et al. 2016, A&A, 585, A93
Roederer, I. U., Preston, G. W., Thompson, I. B., et al. 2014, AJ, 147, 136
Santos-Peral, P., Recio-Blanco, A., de Laverny, P., Fernández-Alvar, E., & Ordenovic, C. 2020, A&A, 639, A140
Sartoretti, P., Katz, D., Cropper, M., et al. 2018, A&A, 616, A6
Soubiran, C., Brouillet, N., & Casamiquela, L. 2022, A&A, 663A4, 16
Starkenburg, E., Martin, N., Youakim, K., et al. 2017, MNRAS, 471, 2587
Steinmetz, M., Guiglion, G., McMillan, P. J., et al. 2020, AJ, 160, 83
Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645
Tarricq, Y., Soubiran, C., Casamiquela, L., et al. 2021, A&A, 647, A19
Wilkinson, M. I., Vallenari, A., Turon, C., et al. 2005, MNRAS, 359, 1306
Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377
Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, Res. Astron. Astrophys., 12, 723
Zhao, H., Schultheis, M., Recio-Blanco, A., et al. 2021, A&A, 645, A14

[1] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Nice, France
[2] CIGUS CITIC – Department of Computer Science and Information Technologies, University of A Coruña, Campus de Elviña s/n, A Coruña 15071, Spain
[3] CIGUS CITIC, Department of Nautical Sciences and Marine Engineering, University of A Coruña, Paseo de Ronda 51, 15071 A Coruña, Spain
[4] Royal Observatory of Belgium, 3 Avenue circulaire, 1180 Brussels, Belgium
[5] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
[6] CNES Centre Spatial de Toulouse, 18 Avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[7] ATG Europe for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[8] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
[9] National Observatory of Athens, I. Metaxa and Vas. Pavlou, Palaia Penteli 15236, Athens, Greece
[10] Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[11] INAF – Osservatorio Astrofisico di Catania, Via S. Sofia 78, 95123 Catania, Italy
[12] Telespazio for CNES Centre Spatial de Toulouse, 18 Avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[13] Dpto. de Matemática Aplicada y Ciencias de la Computación, Univ. de Cantabria, ETS Ingenieros de Caminos, Canales y Puertos, Avda. de los Castros s/n, 39005 Santander, Spain
[14] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 place Jules Janssen, 92190 Meudon, France
[15] Centre for Astrophysics Research, University of Hertfordshire, College Lane, AL10 9AB Hatfield, UK
[16] INAF – Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
[17] Institut d'Astrophysique et de Géophysique, Université de Liège, 19c, Allée du 6 Août, 4000 Liège, Belgium
[18] Theoretical Astrophysics, Division of Astronomy and Space Physics, Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden
[19] European Space Agency (ESA), European Space Astronomy Centre (ESAC), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[20] Data Science and Big Data Lab, Pablo de Olavide University, 41013 Seville, Spain
[21] Department of Astrophysics, Astronomy and Mechanics, National and Kapodistrian University of Athens, Panepistimiopolis, Zografos, 15783 Athens, Greece
[22] Dipartimento di Fisica e Astronomia "Ettore Majorana", Università di Catania, Via S. Sofia 64, 95123 Catania, Italy
[23] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92190 Meudon, France
[24] Université Rennes, CNRS, IPR (Institut de Physique de Rennes) – UMR 6251, 35000 Rennes, France
[25] Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark
[26] DXC Technology, Retortvej 8, 2500 Valby, Denmark
[27] Aurora Technology for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[28] IPAC, Mail Code 100-22, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[29] IRAP, Université de Toulouse, CNRS, UPS, CNES, 9 Av. colonel Roche, BP 44346, 31028 Toulouse Cedex 4, France
[30] Thales Services for CNES Centre Spatial de Toulouse, 18 Avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[31] Dpto. de Inteligencia Artificial, UNED, c/ Juan del Rosal 16, 28040 Madrid, Spain
[32] INAF – Osservatorio astronomico di Padova, Vicolo Osservatorio 5, 35122 Padova, Italy
[33] Institute of Global Health, University of Geneva, Geneva, Switzerland
[34] Applied Physics Department, Universidade de Vigo, 36310 Vigo, Spain
[35] Sorbonne Université, CNRS, UMR7095, Institut d'Astrophysique de Paris, 98bis bd. Arago, 75014 Paris, France

Appendix A: *GSP-Spec* and radial velocities

The number of stars missing radial velocities ($V_{\rm Rad}$) for different *GSP-Spec* parameters are provided in Table A.1. Parameters not listed are not missing any $V_{\rm Rad}$.

Table A.1. *GSP-Spec* and radial velocity statistics.

GSP-Spec parameter	# stars missing $V_{\rm Rad}$
teff_gspspec	95
logg_gspspec	95
mh_gspspec	84
alphafe_gspspec	83
fem_gspspec	9
sife_gspspec	6
cafe_gspspec	6
mgfe_gspspec	2
feiim_gspspec	2
sfe_gspspec	6

Appendix B: Atomic lines selected for the chemical analysis

This Appendix introduces the list of selected lines used in the determination of the chemical abundances. In Table B.1, we summarise the reference wavelength value of each atomic line, as well as the wavelength ranges considered for its abundance determination and second normalisation windows. We note that the reference wavelength (col. 2) can differ from the vacuum wavelength of the analysed atomic line in case of multiplets or broad lines. For instance, for the Ca II IR triplet transitions at 850.036, 854.444, and 866.452 nm, two Ca abundances have been derived from the wings of each line to avoid the line core that could not be well modelled. In those cases, col. 2 refers to one of the Ca II wings.

The following lines (denoted by an asterisk in Table B.1) have multiple lines within the same abundance determination window: (855.913, 855.916) for Si I; (867.082, 867.258, 867.297, 867.366) and (869.632, 869.701) for S I; (852.037, 852.069) for Ti I; (848.283, 848.296, 848.431), (851.641, 851.745, 851.751), (852.738, 852.901, 853.020) and (868.916, 869.101) for Fe I[19]; (851.368, 851.381, 851.375) for Ce II. For the Fe II line measured in hot star spectra (see Sect. 8.7.2), some blends of weak Fe I transitions may be present in cooler star spectra.

Table B.1. List of the atomic lines adopted for the determination of individual chemical abundances by *GSP-Spec*. Col. 2 refers to the reference wavelength of the analysed lines (see text for details). The abundance determination window corresponds to the interval $[\lambda^-_{ab}, \lambda^+_{ab}]$ (third and fourth column, respectively) while the refined normalisation window includes the wavelength range $[\lambda^-_{norm}, \lambda^+_{norm}]$ (fifth and sixth column, respectively). All the wavelengths are in nanometres and in the vacuum.

Elt	λ	λ^-_{ab}	λ^+_{ab}	λ^-_{norm}	λ^+_{norm}
N I	863.161	863.071	863.281	862.891	863.371
N I	868.579	868.489	868.699	868.309	868.939
Mg I	847.602	847.512	847.692	847.212	847.812
Si I	853.851	853.731	853.941	853.371	854.961
*Si I	855.916	855.856	856.036	855.376	856.156
Si I	868.872	868.782	868.992	868.602	869.232
*S I	867.258	866.988	867.378	866.898	867.998
*S I	869.701	869.551	869.821	869.281	869.971
Ca I	863.631	863.511	863.691	863.361	863.931
Ca II	849.856	849.706	849.976	849.586	850.276
Ca II	850.216	850.156	850.276	849.886	850.306
Ca II	854.264	854.114	854.384	853.544	854.864
Ca II	854.624	854.564	854.744	854.294	854.804
Ca II	866.272	866.152	866.332	866.002	866.572
Ca II	866.632	866.512	866.692	866.302	866.782
*Ti I	852.069	851.979	852.129	851.799	852.249
Ti I	857.209	857.119	857.269	856.999	857.359
Ti I	869.472	869.382	869.562	869.292	869.832
Cr I	855.118	855.058	855.208	854.878	855.478
Cr I	864.567	864.447	864.627	864.207	864.867
*Fe I	848.296	848.206	848.446	847.666	848.896
*Fe I	851.641	851.551	851.851	851.281	852.001
*Fe I	852.901	852.691	853.081	852.481	853.321
Fe I	857.416	857.296	857.506	856.876	858.166
Fe I	858.462	858.312	858.612	858.132	858.762
Fe I	862.397	862.277	862.517	862.127	862.697
Fe I	867.713	867.593	867.863	867.443	868.013
*Fe I	869.101	868.891	869.191	868.441	869.821
Fe II	858.794	858.764	858.824	858.254	859.274
Ni I	863.937	863.847	864.027	863.697	864.147
Zr II	852.748	852.658	852.838	852.388	853.018
*Ce II	851.375	851.285	851.465	851.015	851.555
Nd II	859.389	859.299	859.479	859.209	859.689

[19] Only the strongest transitions are provided for iron.

Appendix C: Definition of the *GSP-Spec* flags

The following tables include the detailed definition of the individual characters in the *GSP-Spec* quality flag chain presented in Table 2. In addition, Fig. C.1 illustrates the implemented modelling of parameter biases induced by rotational broadening, leading to the definition of *vbroad*T, *vbroad*G and *vbroad*M quality flags (cf. Sec. 8.1 and Table C.1). The particular case of effective temperature biases is illustrated. Finally, Fig. C.2 presents the validation flow chart associated with the definition of quality flags for the DIB parametrisation (cf. Sect. 8.9 and Table C.13).

Table C.1. Definition of the parameter flags considering potential biases due to rotational velocity and/or macroturbulence. These flags are part of the *flags_gspspec* string chain defined in Table 2.

Flag name	Condition	Flag value
vbroadT	ΔTeff>2000 K	Filtered all
	500< ΔTeff≤2000 K	Flag 2
	250< ΔTeff≤500 K	Flag 1
	ΔTeff≤250 K	Flag 0
vbroadG	Δlogg>2 dex	Filter all *except Teff & DIB if Teff>7000 K* Flag 9
	1< Δlogg≤2 dex	Flag 2
	0.5< Δlogg≤1 dex	Flag 1
	Δlogg≤0.5 dex	Flag 0
vbroadM	Δ[M/H]>2 dex	Filter [M/H] & [X/Fe] Flag 9
	0.5< Δ[M/H]≤2 dex	Flag 2
	0.25< Δ[M/H]≤0.5 dex	Flag 1
	Δ[M/H]≤0.25 dex	Flag 0

Table C.2. Same as Table C.1 but for the potential biases due to uncertainties in the radial velocity shift correction (see also Table 2).

Flag name	Condition	Flag value
vradT	ΔTeff>2000 K	Filter all
	500< ΔTeff≤2000 K	Flag 2
	250< ΔTeff≤500 K	Flag 1
	ΔTeff≤250 K	Flag 0
vradG	Δlogg>2 dex	Filter all *except Teff & DIB if Teff>7000 K* Flag 9
	1< Δlogg≤2 dex	Flag 2
	0.5< Δlogg≤1 dex	Flag 1
	Δlogg≤0.5 dex	Flag 0
vradM	Δ[M/H]>2 dex	Filter [M/H] & [X/Fe] Flag 9
	0.5< Δ[M/H]≤2 dex	Flag 2
	0.25< Δ[M/H]≤0.5 dex	Flag 1
	Δ[M/H]≤0.25 dex	Flag 0

Table C.3. Definition of the parameter flags considering potential biases due to uncertainties in the RVS flux (MatisseGauguin parametrisation; see also Table 2).

Flag name	Condition	Flag value
fluxNoise	σTeff>2000 K or σlogg>2 dex	Filter all Flag 9
	σTeff≤2000 K and σlogg≤2 dex and σ[M/H]>2 dex	Filter [M/H] & [X/Fe] Flag 5
	σTeff≤2000 K and σlogg≤2 dex and σ[M/H]≤2 dex and σ[α/Fe]>0.8 dex	Filter [α/Fe] & [X/Fe] Flag 4
	500< σTeff≤2000 K and 1< σlogg≤2 dex and 0.5< σ[M/H]≤2 dex and 0.2< σ[α/Fe]≤0.8 dex	Flag 3
	250< σTeff≤500 K and 0.5< σlogg≤1 dex and 0.25< σ[M/H]≤0.5 dex and 0.1< σ[α/Fe]≤0.2 dex	Flag 2
	100< σTeff≤250 K and 0.2< σlogg≤0.5 dex and 0.1< σ[M/H]≤0.25 dex and 0.05< σ[α/Fe]≤0.1 dex	Flag 1
	σTeff≤100 K and σlogg≤0.2 dex and σ[M/H]≤0.1 dex and σ[α/Fe]≤0.05 dex	Flag 0

Table C.4. Same as Table C.3 but for the ANN parametrisation.

ANN fluxNoise	
Condition	**Value**
$\sigma T_{\rm eff}$> 525 K or σlog(g)>0.9 or σ[M/H]>0.5 or σ[α/Fe]< 0.16 dex	Flag 9
225< $\sigma T_{\rm eff}$≤525 K and 0.4< σlog(g)≤0.9 or 0.19< σ[M/H]≤0.5 dex or 0.09< σ[α/Fe]≤0.16 dex	Flag 3
140< $\sigma T_{\rm eff}$≤225 K and 0.3< σlog(g)≤0.4 or 0.12< σ[M/H]≤0.19 dex or 0.06< σ[α/Fe]≤0.09 dex	Flag 2
100< $\sigma T_{\rm eff}$≤140 K and 0.24< σlog(g)≤0.3 or 0.09< σ[M/H]≤0.12 dex or 0.05< σ[α/Fe]≤0.06 dex	Flag 1
$\sigma T_{\rm eff}$≤100 K and σlog(g)≤0.24 and σ[M/H]≤0.09 dex and σ[α/Fe]≤0.05 dex	Flag 0

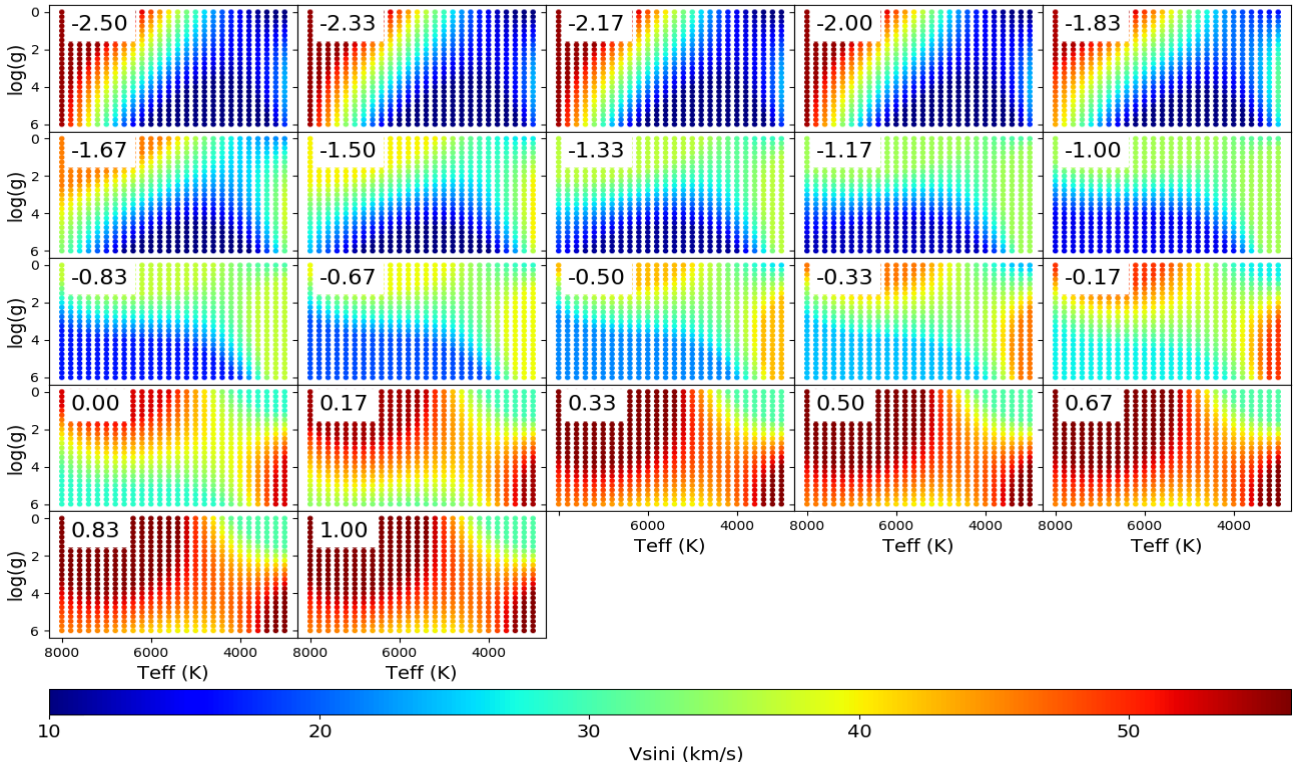


Fig. C.1. Limiting V sin i values (colour code) leading to a bias of 250< $\Delta T_{\rm eff}$≤500 K in the *GSP-Spec* parametrisation. This has been used to estimate the third-order polynomial with $T_{\rm eff}$, log(g), and [M/H] as variables used to define the *vbroad*T flag (equal to 1 in this example). The [M/H] values for each panel are indicated in their upper right corner.

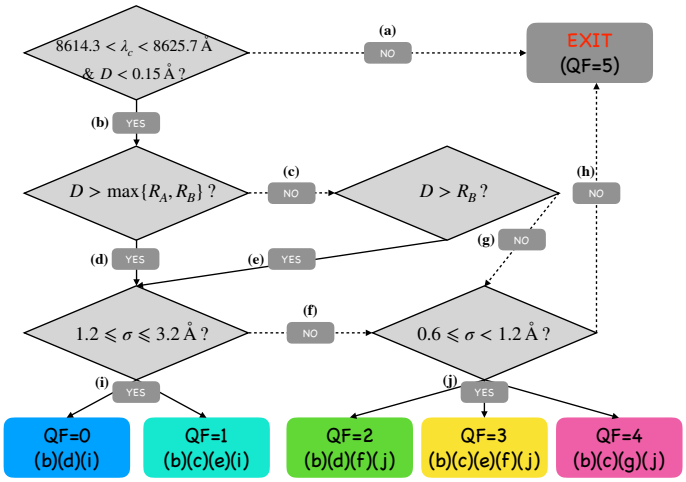


Fig. C.2. Flow chart of the different values for the DIB quality flag. See associated text in Sect. 8.9.

Table C.5. Definition of parameter flags considering potential biases due to extrapolated parameters (MatisseGauguin parametrisation, see also Table 2).

Flag name	Condition	Flag value
extrapol	gof=NaN and (Teff>9000 K or Teff<2500 K or logg>6 or logg<-1)	Filter all *except DIB if Teff>7000 K* Flag 9
	gof=NaN and 2500≤Teff≤9000 K and -1≤logg≤6 and ([M/H]<-6 dex or [M/H]>1.5 dex)	Filter [M/H],[X/Fe] Flag 4
	gof=NaN and 2500≤Teff<9000 K and -1≤logg≤6 and -6≤[M/H]≤1.5 dex and [α/Fe] out from standard by ± 0.8	Filter [X/Fe] Flag 3
	gof=NaN and 2500≤Teff≤9000 K and -1≤logg≤6 dex and -6≤[M/H]≤1.5 and [α/Fe] within ± 0.8 from standard	Flag 2
	gof ≠NaN and (Teff≥7625 or Teff≤3500 K or logg≥4.75 or logg≤0.25 or [M/H]≤-3 or [M/H]≥0.75 dex or [α/Fe] out from standard by ± 0.35)	Flag 1
	gof ≠NaN and 3500<Teff<7625 K and 0.25<logg<4.75 and -3<[M/H]<0.75 dex and [α/Fe] within ± 0.35 from standard	Flag 0

Table C.6. Same as Table C.5 but for the ANN parametrisation.

ANN extrapol	
Condition	**Value**
$T_{\rm eff}$>8320 K or $T_{\rm eff}$<3680 K or log(g)>6.0 or log(g)<-1.0 or [M/H]>1.3 dex or [M/H]<-5.3 dex or [α/Fe]>1.0 dex or [α/Fe]<-0.6 dex	Flag 9
7680<$T_{\rm eff}$≤8320 K or 3680≤$T_{\rm eff}$<4320 K or 4.9<log(g)≤6.0 or -1.0≤log(g)<0.06 or 0.6<[M/H]≤1.4 dex or -5.3≤[M/H]<-4.6 dex or 0.6<[α/Fe]≤1.0 dex or -0.5≤[α/Fe]<0.2 dex	Flag 1
4320≤$T_{\rm eff}$≤7680 K and 0.06≤log(g)≤4.9 and -4.6≤[M/H]≤0.6 dex and -0.2≤[α/Fe]≤0.6 dex	Flag 0

Table C.7. Definition of parameter flags considering RVS flux issues or emission line probability (see also Table 2).

Flag name	Condition	Flag value
nanFlux	Flux=NaN	Filter all *except DIB if Teff>7000 K* Flag 9
emission	*CU6_is_emission*	Filter all *except DIB if Teff>7000 K* Flag 9
negFlux	> 2 *wlp* with flux<0	Filter all *except DIB if Teff>7000 K* Flag 9
	1 or 2 *wlp* with flux<0	Flag 1
	flux>0	Flag 0
nullFluxErr	σTeff=0 K or σlogg=0 dex or σ[M/H]=0 dex or σ[α/Fe]=0 dex	Filter all Flag 9

Table C.8. Definition of the parameter flag considering problems in the paramerisation of KM-type giants. $F_{\rm min}$ is the minimum flux value in the corresponding RVS spectrum (see also Table 2).

Flag name	Condition	Flag value
KM-type stars	$T_{\rm eff}$<4000 K and log(g)<3.5 and (gof >-3.0 and $F_{\rm min}$ >0.22)	Flag 2 Filter [α/Fe]
	$T_{\rm eff}$<4000 K and log(g)<3.5 and [(-3.4< gof <-3.0) or (gof >-3.0 and $F_{\rm min}$ <0.22) or (gof <-3.4 and $F_{\rm min}$ >0.22)]	Flag 1 Filter [α/Fe]
	[$T_{\rm eff}$<4000 K and log(g)<3.5 and (gof <-3.4 or $F_{\rm min}$ <0.22)] or $T_{\rm eff}$>4000 or log(g)>3.5	Flag 0

Table C.9. Definition of individual abundance upper limit flags (see Table 2). Xfe_gspspec_upper is the upper confidence value of the abundance (corresponding to the 84th quantile of the Monte-Carlo distribution). σ[X/Fe] is the 84th-16th interquantile abundance uncertainty. XfeUpperLimit is the mean value of the abundance upper limit for the considered lines of the X-element in the spectrum (depending on the mean S/N in the line *wlp* and the stellar parameters). X_MAD_UpperLimit is the median absolute deviation of upper limit in the line *wlp*. Finally, the c-coefficients are reported in Table C.11.

Flag name	Condition	Flag value
XUpLim	vbroadT≥2 or vbroadG≥2 or vbroadM≥2 or σ[X/Fe] = 0 or (Xfe_gspspec_upper - Xfe_gspspec) = 0 or Teff≤c1 or Teff≥c2 or logg≤c3 or logg≥c4 or ((2 - XfeUpperLimit) / σ[X/Fe])≤c5 or (S/N≤c6 and gof≥c7) or (Xfe_gspspec + [M/H])≤c8	Flag 9 Filter [X/Fe]
	vbroadT<2 and vbroadG<2 and vbroadM<2 and σ[X/Fe] != 0 and (Xfe_gspspec_upper - Xfe_gspspec) != 0 and c1<Teff<c2 and c3<logg<c4 and ((2 - XfeUpperLimit) / σ[X/Fe])>c5 and (S/N>c6 or gof<c7) and (Xfe_gspspec + [M/H])<c8 and ((Xfe_gspspec - XfeUpperLimit) / (1.48·X_MAD_UpperLimit))<1.5	Flag 2
	vbroadT<2 and vbroadG<2 and vbroadM<2 and σ[X/Fe] != 0 and (Xfe_gspspec_upper- Xfe_gspspec) != 0 and c1<Teff<c2 and c3<logg<c4 and ((2 - XfeUpperLimit) / σ[X/Fe])>c5 and (S/N>c6 or gof<c7) and (Xfe_gspspec+[M/H])<c8 and 1.5≤((Xfe_gspspec - XfeUpperLimit) / (1.48·X_MAD_UpperLimit))<2.5	Flag 1
	vbroadT<2 and vbroadG<2 and vbroadM<2 and σ[X/Fe] != 0 and (Xfe_gspspec_upper - Xfe_gspspec) != 0 and c1<Teff<c2 and c3<logg<c4 and ((2 - XfeUpperLimit) / σ[X/Fe])>c5 and (S/N>c6 or gof<c7) and (Xfe_gspspec + [M/H])<c8 and ((Xfe_gspspec - XfeUpperLimit) / (1.48·X_MAD_UpperLimit))≥2.5	Flag 0

Table C.10. Definition of individual abundance uncertainty flags (see Table 2). Xfe_gspspec_upper is the upper confidence value of the abundance (corresponding to the 84th quantile of the Monte-Carlo distribution). σ[X/Fe] is the 84th-16th interquantile abundance uncertainty. XfeUpperLimit is the mean value of the abundance upper limit for the considered lines of the X-element in the spectrum (depending on the mean S/N in the line *wlp* and the stellar parameters). Finally, the c-coefficients are reported in Table C.11.

Flag name	Condition	Flag value
XUncer	vbroadT≥2 or vbroadG≥2 or vbroadM≥2 or σ[X/Fe]=0 or (Xfe_gspspec_upper - Xfe_gspspec) = 0 or Teff≤c1 or Teff≥c2 or logg≤c3 or logg≥c4 or ((2-XfeUpperLimit) / σ[X/Fe])≤c5 or (S/N≤c6 and gof≥c7) or (Xfe_gspspec + [M/H])≤c8	Flag 9 Filter [X/Fe]
	vbroadT<2 and vbroadG<2 and vbroadM<2 and σ[X/Fe] != 0 and (Xfe_gspspec_upper - Xfe_gspspec) != 0 and c1<Teff<c2 and c3<logg<c4 and c5<((2 - XfeUpperLimit) / σ[X/Fe])<7 and (S/N>c6 or gof<c7) and (Xfe_gspspec + [M/H])<c8	Flag 2
	vbroadT<2 and vbroadG<2 and vbroadM<2 and σ[X/Fe] != 0 and (Xfe_gspspec_upper - Xfe_gspspec) != 0 and c1<Teff<c2 and c3<logg<c4 and 7 ≤ ((2 - XfeUpperLimit) / σ[X/Fe]) < 10 and (S/N>c6 or gof<c7) and (Xfe_gspspec + [M/H])<c8	Flag 1
	vbroadT<2 and vbroadG<2 and vbroadM<2 and σ[X/Fe] != 0 and (Xfe_gspspec_upper - Xfe_gspspec) != 0 and c1<Teff<c2 and c3<logg<c4 and ((2 - XfeUpperLimit) / σ[X/Fe]) ≥ 10 and (S/N>c6 or gof<c7) and (Xfe_gspspec + [M/H])<c8	Flag 0

Table C.11. Coefficients for individual chemical abundance filtering (see [X/Fe] upperLimit flag and σ[X/Fe] quality flag in Table C.9 and C.10, respectively).

Chemical abundance	c1	c2	c3	c4	c5	c6	c7	c8
[N/Fe]	4200	8000	0.0	5.5	4.5	100	-3.6	99
[Mg/Fe]	3500	8000	-1.0	5.5	5.5	80	-3.5	99
[Si/Fe]	4000	8000	-1.0	5.5	6.0	110	-3.8	99
[S/Fe]	5500	8000	3.0	5.5	5.0	120	-3.7	99
[Ca/Fe]	3500	8000	-1.0	5.5	10.0	60	-3.2	99
[Ti/Fe]	4000	6500	-1.0	5.5	6.0	110	-3.65	99
[Cr/Fe]	3500	6000	-1.0	5.5	6.0	1000	-3.65	1.5
[Fe I/M]	3500	8000	-1.0	5.5	5.0	1000	-3.4	1.5
[Fe II/M]	5700	8000	3.5	5.5	5.0	70	-3.5	1.5
[Ni/Fe]	4000	6500	-1.0	5.5	6.0	100	-3.6	1.5
[Zr/Fe]	3500	8000	-1.0	5.5	1.0	100	-3.4	99
[Ce/Fe]	3500	8000	-1.0	5.5	5.0	100	-3.5	99
[Nd/Fe]	3500	5500	-1.0	5.5	2.0	100	-3.5	99

Table C.12. Definition of the quality flag of the CN differential EW with respect to the solar C and N abundances.

Flag name	Condition	Flag value
DeltaCNq	vbroadT≥1 or vbroadG≥1 or vbroadM≥1 or CN_EW_err = 0 or S/N ≤ 80 or gof≥-3.5 or $T_{\rm eff}$≥4800 K or log(g)≥3.8 or |CN_p1 - 849.037|≥0.05 or CN_p2≥0.25	Flag 9 Filter CN
	vbroadT<1 and vbroadG<1 and vbroadM<1 and CN_EW_err != 0 and S/N > 80 and gof <-3.5 and $T_{\rm eff}$<4800 K and log(g)<3.8 and |CN_p1 - 849.037|<0.05 and CN_p2<0.25	Flag 0

Table C.13. Definition of the quality flag for the DIB parameterisation. See Sect. 8.9 for the definition of *Ra* and *Rb*.

Flag name	Condition	Flag value
DIBq	S/N≤70 or Vrad_err > 5 km.s^{-1} or $T_{\rm eff}$<3 500 K or $T_{\rm eff}$>10^5 K or Σ_i RVS_Flux$_i$ < 0 or p$_0$ < 3/(S/N)	Flag 9 Filter DIB
	p_1 < 861.66 nm or p_1 > 862.81 nm or p_0 > 0.015 nm	Flag 5
	861.66 nm < p_1 < 862.81 nm and p_0 < R_b and 0.6 < p_2 < 1.2	Flag 4
	861.66 nm < p_1 < 862.81 nm and p_0>R_b and 0.6<p_2<1.2	Flag 3
	861.66 nm < p_1 < 862.81 nm and p_0 > $max(R_a, R_b)$ and 0.6<p_2<1.2	Flag 2
	861.66 nm < p_1 < 862.81 nm and p_0 > R_b and 1.2<p_2<3.2	Flag 1
	861.66 nm < p_1 < 862.81 nm and p_0 > $max(R_a, R_b)$ and 1.2<p_2<3.2	Flag 0

Appendix D: Bias comparisons per survey for MatisseGauguin parameters

Here, we perform a similar analysis to that shown in Sect. 9.1.1, but we investigate how the individual surveys compare with *GSP-Spec*. First of all, Figure D.1 presents a cumulative histogram of the RVS spectra S/N for the selected comparison samples between *GSP-Spec* MatisseGauguin and RAVE-DR6, GALAH-DR3, and APOGEE-DR17. As expected from the selection functions of the different ground-based surveys, RAVE-DR6 targets have RVS spectra with higher S/N values than GALAH-DR3 or APOGEE-DR17.

Figure D.2 is the equivalent to Fig. 11, showing only the 99 and 66 percent contour lines for RAVE-DR6 (in black), GALAH-DR3 (in red), and APOGEE-DR17 in blue. Table D.1

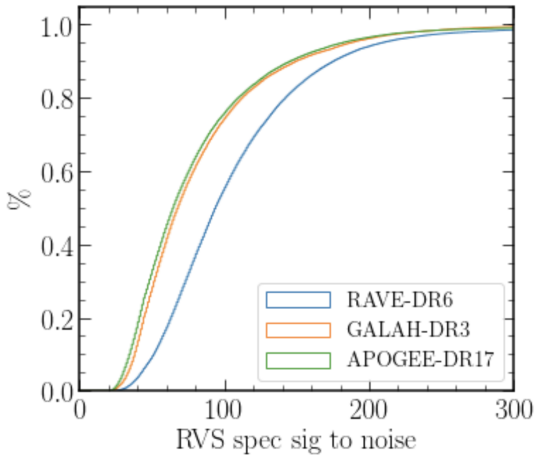


Fig. D.1. Cumulative histogram of RVS S/N for the selected comparison sample between *GSP-Spec* MatisseGauguin and three ground-based surveys: RAVE-DR6, GALAH-DR3, APOGEE-DR17. Table D.1 provides the median values and standard deviation per survey.

quantifies the comparisons, by showing the median offset (*GSP-Spec* – reference) as well as the robust sigma before and after the calibration for each survey. One can see that trends are similar no matter the reference catalogue, and that the biases are significantly decreased when using the calibrated values. It is important to note here that RVS-RAVE targets benefit from a higher S/N with respect to those of RVS-GALAH and RVS-APOGEE.

Finally, we investigated how *GSP-Spec* MatisseGauguin and the literature uncertainties compare to the observed parameter differences. To this purpose, Figure D.3 shows, in blue, the histograms of the $T_{\rm eff}$ (left column), log(g) (middle column), and [M/H] (right column) differences with respect to RAVE-DR6 (upper row), GALAH-DR3 (middle row), and APOGEE-DR17 (lower row). *GSP-Spec* MatisseGauguin log(g) and [M/H] values are calibrated. These parameter differences are normalised by the total uncertainty (defined as the quadratic sum of the *GSP-Spec* and the survey's uncertainties). The dotted histograms show the same distributions inflating the reported uncertainties by a factor of 4. Additionally, the red curves show a normal distribution of unit dispersion and zero mean. An unbiased parameter estimation with correct uncertainties should follow this distribution. Regarding the effective temperature, the reported uncertainties from both *GSP-Spec* and the literature seem to correspond to the observed differences (with some uncertainty overestimation for the RVS-GALAH sample). Regarding log(g), the situation differs from one survey to another. While the agreement between *GSP-Spec* MatisseGauguin and GALAH is good and the reported uncertainties appear overestimated again, the comparison with RAVE and APOGEE suggests that the reported uncertainties are underestimated by a factor of 2 or 3 (the factor 4 is excluded by the normal distribution). Finally, the right column histograms show that [M/H] uncertainties are coherent with the observed differences between *GSP-Spec* and RAVE. However, [M/H] uncertainties from *GSP-Spec* or the GALAH/APOGEE reference or both seem underestimated by about a factor of 4. While in these examples, we only illustrate the impact of artificially inflating *GSP-Spec* uncertainties (through the dotted histograms), it cannot be excluded that the disagreement with respect to the normal distribution is caused by an underestimation of the uncertainties reported by the literature, as possibly suggested by the variety of situations that exist, for the same atmospheric parameter, when comparing to different surveys.

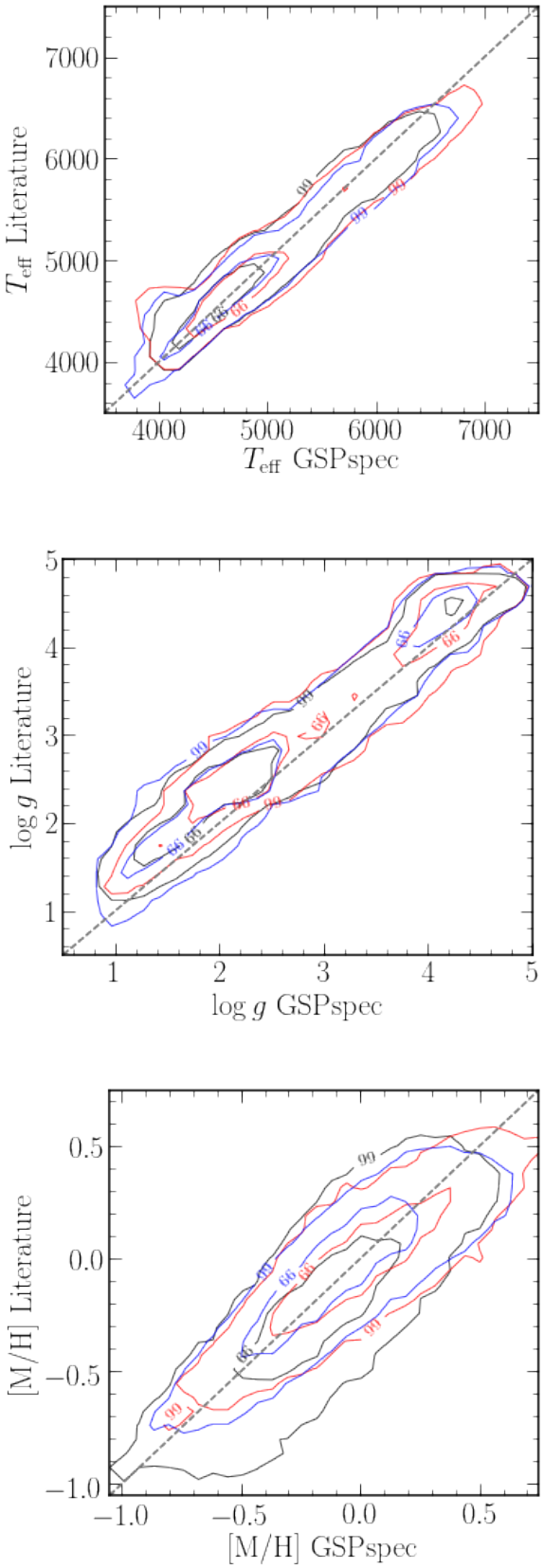


Fig. D.2. Similar to Fig. 11 but showing only the contour lines of the 99th and 66th percentiles for RAVE-DR6 (black), GALAH-DR3 (red), and APOGEE-DR17 (blue).

This analysis illustrates the complexity of comparing parametrisation results from different sources, each of them with its own uncertainty definitions, methodological and theoretical trends, and underlying selection functions. Once again, the importance of using a homogeneous catalogue for scientific purposes rather than a compilation of different sources (even after re-calibrations) is highlighted.

Table D.1. Median offsets and robust sigma between *GSP-Spec* and individual surveys.

	$T_{\rm eff}$	$\log(g)$	[M/H]	$\log(g)_{\rm calibrated}$	[M/H]$_{\rm calibrated}$	RVS S/N
RAVE-DR6	(-12; 93)	(-0.28; 0.19)	(-0.05; 0.11)	(-0.003; 0.18)	(-0.05; 0.09)	(94; 64)
GALAH-DR3	(20;87)	(-0.26; 0.21)	(0.01; 0.10)	(0.003; 0.18)	(-0.001; 0.10)	(68; 53)
APOGEE-DR17	(-32; 86)	(-0.32; 0.17)	(0.04; 0.12)	(-0.005; 0.15)	(0.06; 0.12)	(65; 80)

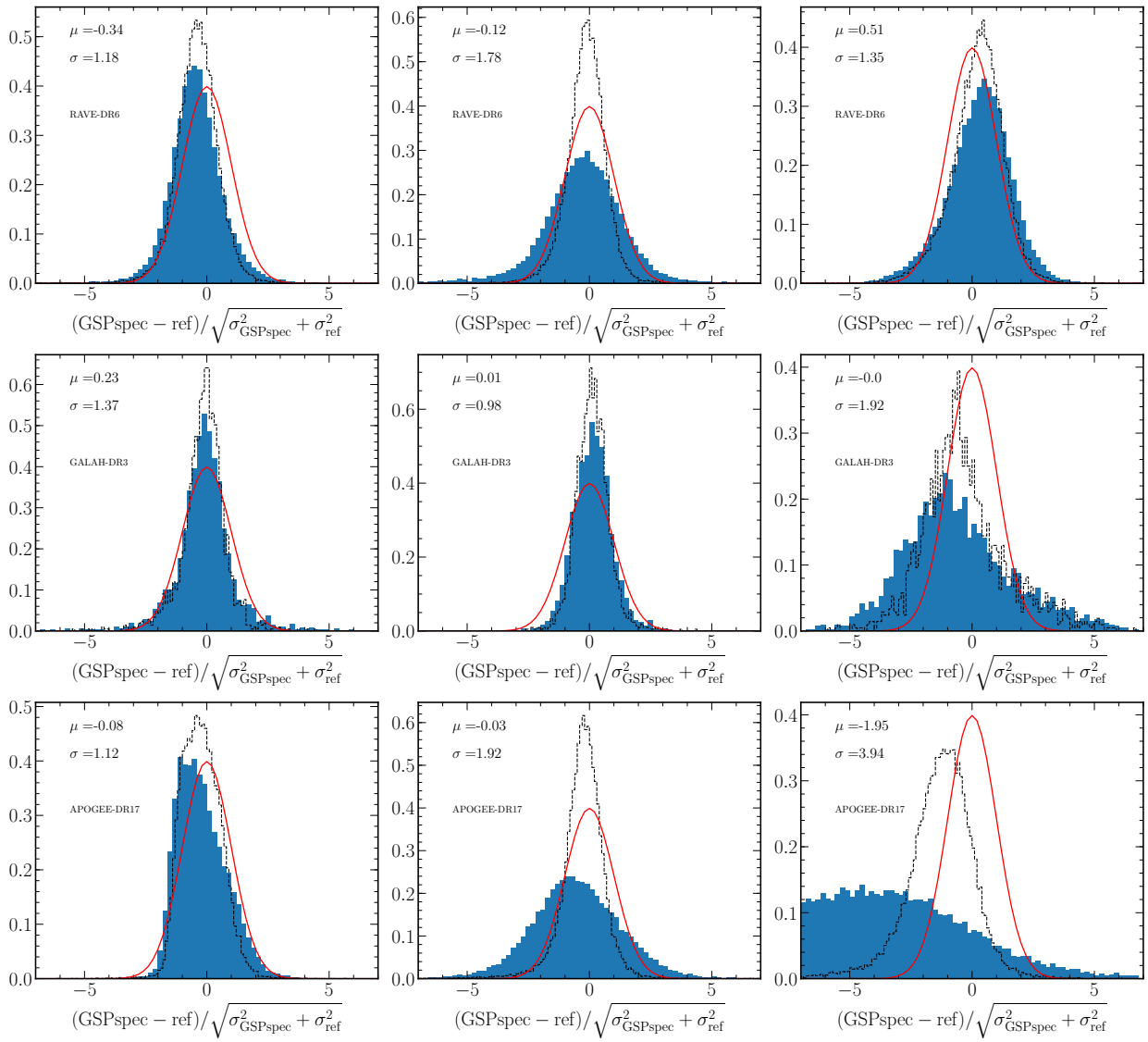


Fig. D.3. Distributions of parameter differences, normalised with respect to the reported *GSP-Spec* and literature uncertainties. From left to right: $T_{\rm eff}$, $\log(g)$, and [M/H] differences. From up to bottom: Differences with respect to RAVE-DR6, GALAH-DR3, and APOGEE-DR17. Dotted histograms correspond to the same distributions inflating the uncertainties (*GSP-Spec* and literature) by a factor of 4. The red curve shows a normal distribution of unit dispersion and zero mean. An unbiased parameter estimation with correct uncertainties should follow this distribution.

Appendix E: Illustration of polynomial corrections for MatisseGauguin chemical abundances and quantification of uncertainties

Figures E.1 and E.2 illustrate the calibrations for individual chemical abundances and the comparison with literature data presented in Sect. 9.1.2. In addition, the uncertainties in the polynomial coefficients p_1, p_2, p_3, and p_4 provided in Table 4 are presented in Table E.1.

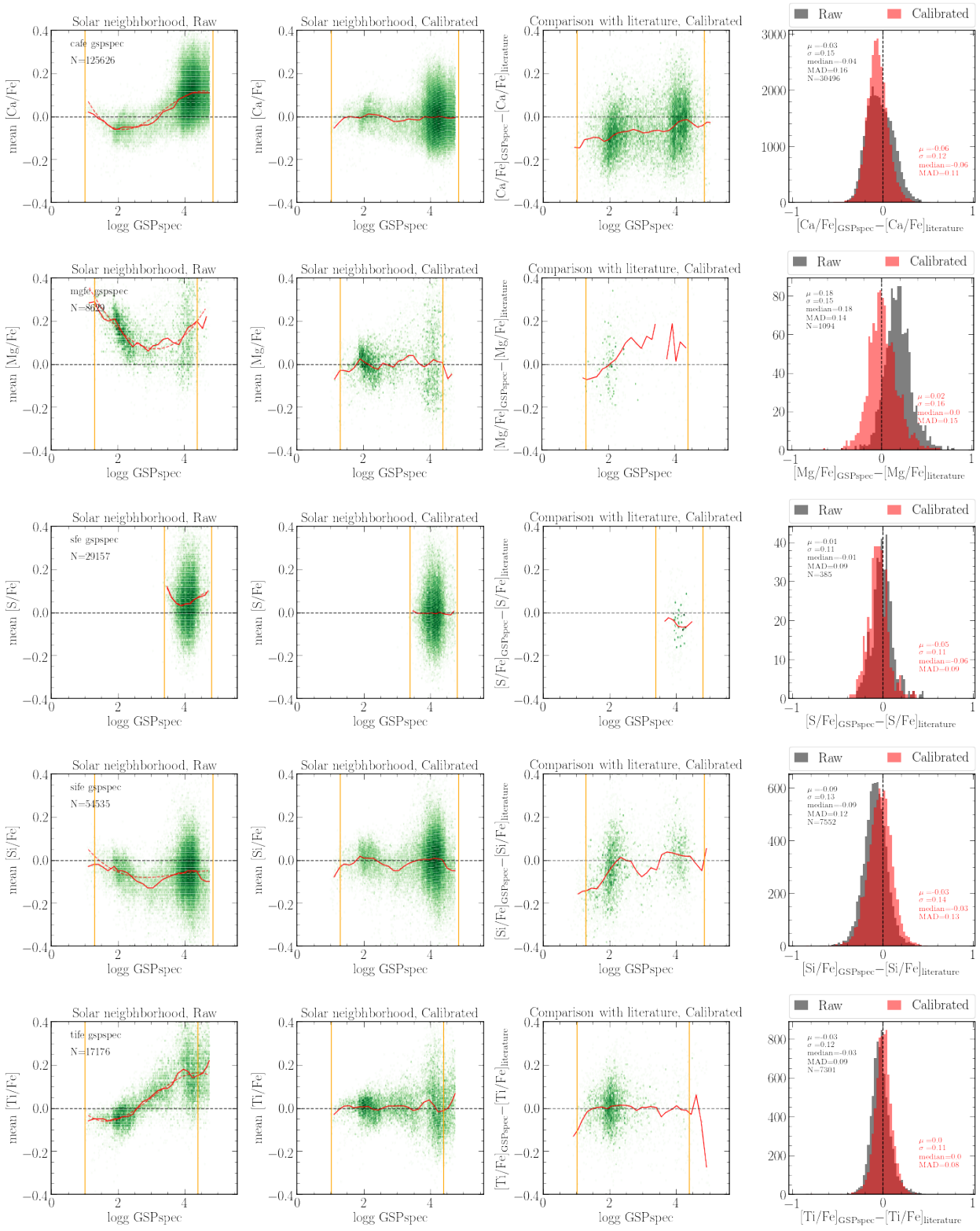


Fig. E.1. Same as Fig. 14, but for individual α-elements.

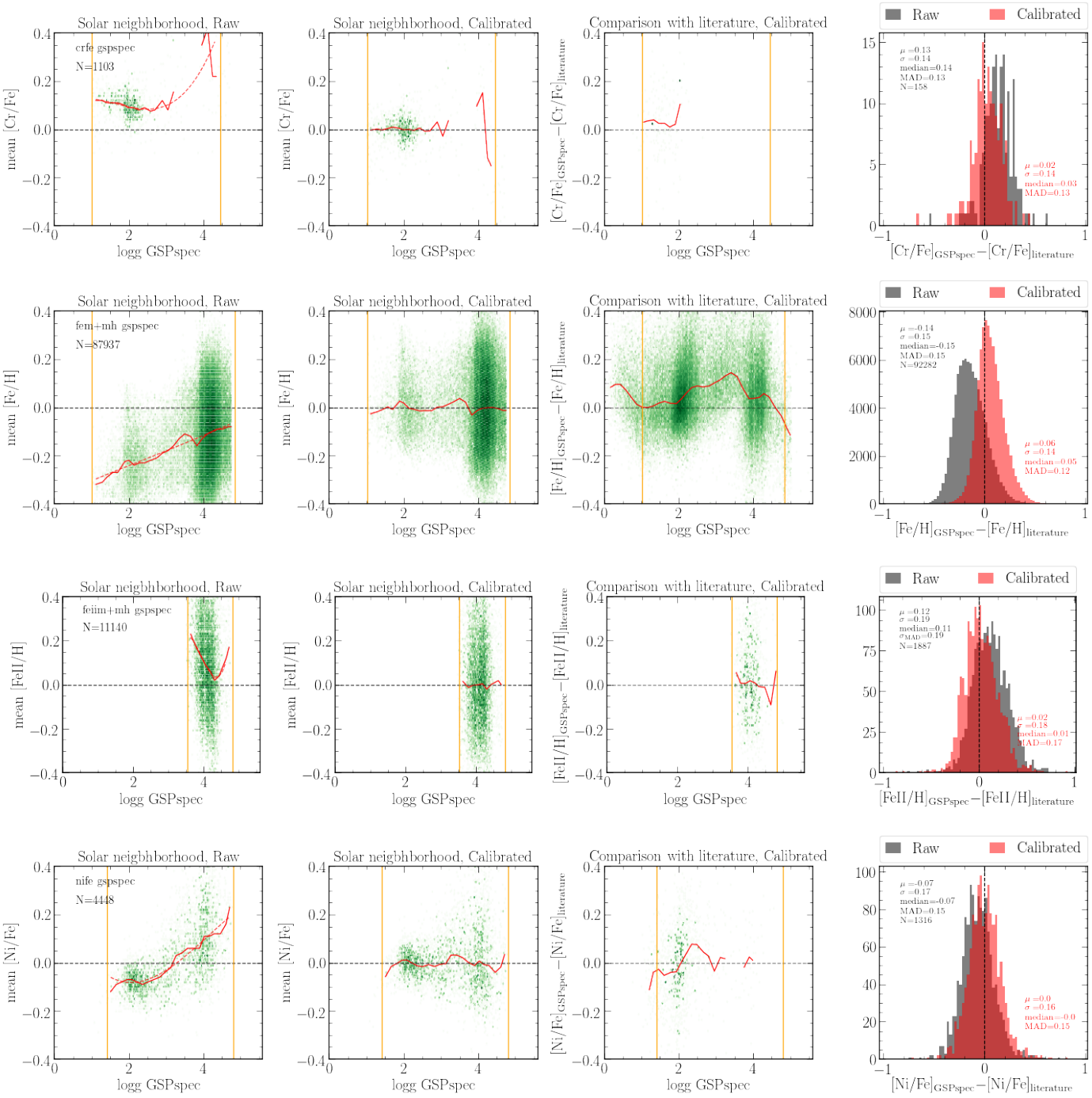


Fig. E.2. Same as Fig. 14, but for individual iron-peak elements.

Table E.1. Uncertainties on the polynomial coefficients of Table 4.

Element	δp_0	δp_1	δp_2	δp_3	δp_4	*extrapol* flag
	As a function of log(g)					
[α/Fe]	0.0965	0.0960	0.0302	0.0031	0.0000	0
[Ca/Fe]	0.0750	0.0747	0.0236	0.0024	0.0000	0
[Mg/Fe]	0.2002	0.2252	0.0819	0.0095	0.0000	0
[S/Fe]	6.5357	4.7804	1.1639	0.0943	0.0000	0
[Si/Fe]	0.1491	0.1562	0.0522	0.0056	0.0000	0
[Ti/Fe]	0.0729	0.0807	0.0286	0.0032	0.0000	0
[Cr/Fe]	0.0951	0.1310	0.0571	0.0077	0.0000	0
[Fe I/H]	0.1254	0.1260	0.0401	0.0041	0.0000	0
[Fe II/H]	13.9985	10.1731	2.4608	0.1981	0.0000	0
[Ni/Fe]	0.1829	0.1979	0.0692	0.0078	0.0000	0
[N/Fe]	0.0580	0.0716	0.0283	0.0035	0.0000	0
[α/Fe]	0.0646	0.0971	0.0524	0.0119	0.0010	$\leq$1
[Ca/Fe]	0.1000	0.1419	0.0726	0.0158	0.0012	$\leq$1
	As a function of t=$T_{\rm eff}$/5750					
[α/Fe]	0.5852	1.8825	2.0028	0.7049	0.0000	$\leq$1
[Ca/Fe]	0.7270	2.3284	2.4674	0.8655	0.0000	$\leq$1
[S/Fe]	0.0323	0.0300	0.0000	0.0000	0.0000	$\leq$1

Appendix F: Validation of ANN biases and uncertainties as a function of S/N

As explained in Sect. 7, the ANN algorithm is trained with noisy spectra to optimise the parametrisation in different S/N regimes. For this reason, it is important to validate the correct behaviour of internal and external errors as a function of S/N.

To study the internal biases and uncertainties, a parametrisation test with a random sample of 10 000 synthetic spectra in the three S/N regimes listed in Table 1 was performed. First of all, we studied the global behaviour of the bias as a function of S/N and a possible dependency of the bias on the parameters themselves by fitting the obtained parameter X^{ANN} as a function of the true parameter X^{Syn} (Fig. F.1). To model this behaviour, we use three different functions: a simple straight line, a parabola, and a piecewise first-order polynomial function (two, three, and five degrees of freedom, respectively) selecting as the best function the one with the smallest Bayesian information criterion (BIC). This process is repeated for each S/N.

In addition, for each S/N, the internal uncertainty (σ_{inter}) on each parameter was estimated from the standard deviation of the distribution $X^{ANN} - X^{Syn}$. The internal uncertainty trends with S/ N are shown in Figure F.2 together with the function that best fits these points. To find this best-fit function, two possible functional relationships were considered: simple parabolic and an inverse square root of the S/N, selecting once again the function with the minimum BIC. It is worth noting that the preferred function is the inverse square root function in all S/N bins, confirming the consistency of the estimations and leading to the following equations:

$$\sigma_{inter_Teff} = -505 + 4763/\sqrt{S/N}, \quad \text{(F.1)}$$

$$\sigma_{inter_logg} = -0.9 + 8/\sqrt{S/N}, \quad \text{(F.2)}$$

$$\sigma_{inter_[M/H]} = -0.6 + 5/\sqrt{S/N}, \quad \text{(F.3)}$$

$$\sigma_{inter_[\alpha/Fe]} = -0.6 + 5/\sqrt{S/N}. \quad \text{(F.4)}$$

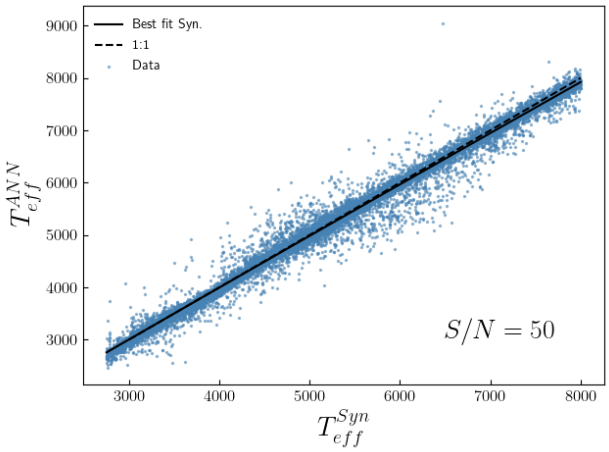


Fig. F.1. Illustration of the ANN tests with synthetic spectra to evaluate internal biases and uncertainties, for S/N$_{ANN}$ =50. The estimated $T_{\rm eff}{}^{ANN}$ as a function of the true parameter $T_{\rm eff}{}^{Syn}$ is shown, including the polynomial fit modelling the observed behaviour. Similar analyses were performed for log(g), [M/H], and [α/Fe].

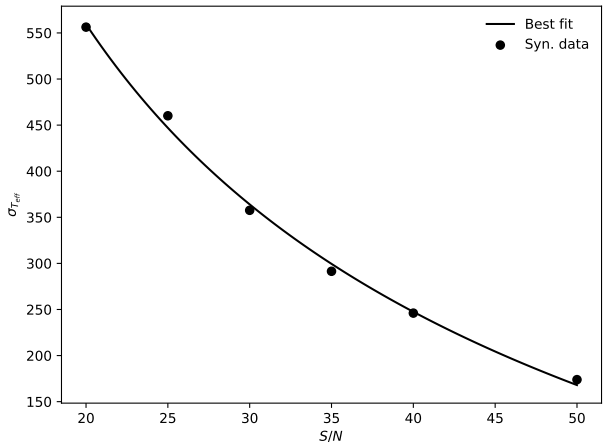


Fig. F.2. Illustration of the estimated trends on the internal ANN T_{eff} uncertainty with S/N. The best fit to this trend is also shown. A similar analysis was performed for log(g), [M/H], and [α/Fe].

Table F.1 summarises the estimated internal biases and uncertainties as a function of S/N. As expected, internal biases are negligible.

Finally, to complete the previous validation of the trend of the ANN estimates with S/N, the differences with respect to the literature (see Sect. 9.2) were examined. As a significant proportion of results from the three reference surveys have an important S/N dependence, with the lower resolution RAVE survey dominating for brighter sources in the high-S/N regime, we decided to validate the uncertainty behaviour with S/N with APOGEE DR16 and GALAH DR3 exclusively[20]. Figure F.3 illustrates the distribution of $T_{\rm eff}$ differences with respect to the literature for the five S/N regimes of the ANN training. Similar analyses were performed for the other three atmospheric parameters and the estimated biases and mean absolute deviations are reported in Table F.2. The expected increase in the spread for lower S/N regimes can be seen, validating the S/N optimisation of the ANN algorithm.

[20] These two surveys, as expected from their higher wavelength coverage and resolution, also show a better agreement with ANN parameters for sources with S/N>50.

Table F.1. ANN internal biases and uncertainties (from the mean absolute deviation) in the different S/N regimes considered for the ANN training.

S/N		$T_{\rm eff}$	log(g)	[M/H]	[α/Fe]
50	Bias	-19 K	-0.02 dex	0.03 dex	-0.01 dex
	MAD	102 K	0.16 dex	0.13 dex	0.06 dex
40	Bias	30 K	-0.07 dex	-0.01 dex	0.02 dex
	MAD	149 K	0.25 dex	0.18 dex	0.09 dex
35	Bias	39 K	-0.01 dex	0.07 dex	0.02 dex
	MAD	179 K	0.31 dex	0.21 dex	0.11 dex
30	Bias	-22 K	-0.01 dex	-0.05 dex	0.01 dex
	MAD	233 K	0.41 dex	0.28 dex	0.13 dex
25	Bias	-2 K	-0.01 dex	0.03 dex	0.01 dex
	MAD	318 K	0.56 dex	0.38 dex	0.16 dex

Table F.2. ANN biases and mean absolute deviations with respect to the literature, in the different S/N regimes considered for the ANN training.

S/N		$T_{\rm eff}$	log(g)	[M/H]	[α/Fe]
50	Bias	-162 K	-0.13 dex	-0.29 dex	0.07 dex
	MAD	155 K	0.21 dex	0.15 dex	0.05 dex
40	Bias	-89 K	-0.10 dex	-0.26 dex	0.10 dex
	MAD	139 K	0.29 dex	0.13 dex	0.07 dex
35	Bias	-38 K	0.005 dex	-0.25 dex	0.12 dex
	MAD	148 K	0.32 dex	0.15 dex	0.08 dex
30	Bias	-6 K	-0.04 dex	-0.32 dex	0.11 dex
	MAD	192 K	0.42 dex	0.19 dex	0.11 dex
25	Bias	-98 K	-0.02 dex	-0.43 dex	0.09 dex
	MAD	253 K	0.47 dex	0.25 dex	0.12 dex

Table F.3. Polynomial coefficients for the $T_{\rm eff}$ calibration at different S/N$_{ANN}$ values.

	S/N_{ANN}	p_0	p_1	p_2	p_3
$T_{\rm eff}$	30	-834	3.13E-1	-2.53E-5	
	35	-1344	4.95E-1	-4.26E-5	
	40	3182	-2.34	5.38E-4	-3.84E-8
	50	12816	-8.1	1.65E-3	-1.07E-7

After the analysis of the uncertainties, we realised that there is a direct relation between $T_{\rm eff}$ and S/N and so we decided to propose different calibrations depending on the S/N ranges defined in Table 1. Furthermore, we observed that the number of stars with $T_{\rm eff}$ > 6000 K in the literature is statistically insignificant, and so the calibration beyond this limit should not be applied. For log(g), [M/H], and [α/Fe], although there is an intrinsic relation with S/N, the global calibration proved to be the best solution. We provide the calibration of $T_{\rm eff}$ for S/N$_{ANN}$ ~50 in Section 9.2 and we give the polynomial coefficients for lower S/N in Table F.3.

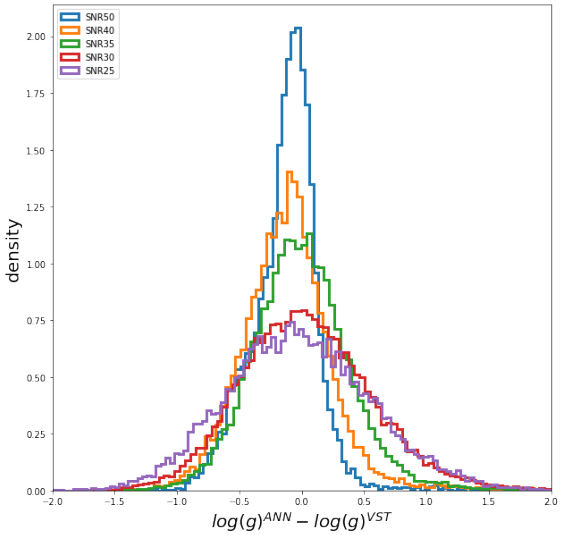


Fig. F.3. Error distributions for ANN estimations with respect to the literature (Validation Source Table - VST).

Appendix G: Query examples from the Gaia Archive

G.1. MatisseGauguin parameters from the AstrophysicalParameters table

```
SELECT source_id
FROM user_dr3int6.astrophysical_parameters
WHERE ((teff_gspspec>=3800) OR (logg_gspspec>=3.5))
    AND ((teff_gspspec>=4150) OR (logg_gspspec>=3.6)
    OR (logg_gspspec<=2.4))
```

Listing 1. ADQL query example with simple cuts in the limiting parameters.

```
SELECT source_id
FROM user_dr3int6.astrophysical_parameters
WHERE (teff_gspspec>3500) AND (logg_gspspec>0) AND
    (logg_gspspec<5) AND
    ((teff_gspspec_upper-teff_gspspec_lower)<750)
    AND ((logg_gspspec_upper-logg_gspspec_lower)<1.)
    AND ((mh_gspspec_upper-mh_gspspec_lower)<.5) AND
    (teff_gspspec>=3800 OR logg_gspspec<=3.5) AND
    (teff_gspspec>=4150 OR logg_gspspec<=2.4 OR
    logg_gspspec>=3.6 ) AND ((flags_gspspec LIKE
    "____________0%") OR (flags_gspspec LIKE
    "____________1%")) AND ((flags_gspspec LIKE
    "0%") OR (flags_gspspec LIKE "1%")) AND
    ((flags_gspspec LIKE "_0%") OR (flags_gspspec
    LIKE "_1%")) AND ((flags_gspspec LIKE "__0%") OR
    (flags_gspspec LIKE "__1%")) AND ((flags_gspspec
    LIKE "___0%") OR (flags_gspspec LIKE "___1%"))
    AND ((flags_gspspec LIKE "____0%") OR
    (flags_gspspec LIKE "____1%")) AND
    ((flags_gspspec LIKE "_____0%") OR
    (flags_gspspec LIKE "_____1%")) AND
    ((flags_gspspec LIKE "______0%") OR
    (flags_gspspec LIKE "______1%") OR
    (flags_gspspec LIKE "______2%") OR
    (flags_gspspec LIKE "______3%")) AND
    ((flags_gspspec LIKE "_______0%") OR
    (flags_gspspec LIKE "_______1%") OR
    (flags_gspspec LIKE "_______2%"))
```

Listing 2. ADQL query example including conditions on the parameter flags (c.f. Table 2).

G.2. ANN parameters from the AstrophysicalParametersSupp table

```
SELECT source_id, teff_gspspec_ann, logg_gspspec_ann,
    mh_gspspec_ann, alphafe_gspspec_ann,
    flags_gspspec_ann
FROM user_dr3int6.astrophysical_parameters_supp
WHERE TO_BIGINT(flags_gspspec_ann) < 10000
```

Listing 3. Best quality sources, no S/N dependency (~1.3 M sources).

```
SELECT ann.source_id, teff_gspspec_ann,
    logg_gspspec_ann, mh_gspspec_ann,
    alphafe_gspspec_ann, flags_gspspec_ann,
    rv_expected_sig_to_noise
FROM user_dr3int6.gaia_source as gaia RIGHT JOIN
   (
   SELECT source_id, teff_gspspec_ann,
       logg_gspspec_ann, mh_gspspec_ann,
       alphafe_gspspec_ann, flags_gspspec_ann
   FROM user_dr3int6.astrophysical_parameters_supp
   WHERE TO_BIGINT(flags_gspspec_ann) < 10000
   ) as ann USING(source_id)
WHERE rv_expected_sig_to_noise > 108
```

Listing 4. Best quality sources with S/N > 108 (S/N_{ANN} 50) (~275 k sources).

Appendix H: Acknowledgements

(*Funding*) The *Gaia* mission and data processing have financially been supported by, in alphabetical order by country:
– the Algerian Centre de Recherche en Astronomie, Astrophysique et Géophysique of Bouzareah Observatory;
– the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (FWF) Hertha Firnberg Programme through grants T359, P20046, and P23737;
– the BELgian federal Science Policy Office (BELSPO) through various PROgramme de Développement d'Expériences scientifiques (PRODEX) grants, the Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek) through grant VS.091.16N, the Fonds de la Recherche Scientifique (FNRS), and the Research Council of Katholieke Universiteit (KU) Leuven through grant C16/18/005 (Pushing AsteRoseismology to the next level with TESS, GaiA, and the Sloan DIgital Sky SurvEy – PARADISE);
– the Brazil-France exchange programmes Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Coordenação de Aperfeicoamento de Pessoal de Nível Superior (CAPES) - Comité Français d'Evaluation de la Coopération Universitaire et Scientifique avec le Brésil (COFECUB);
– the Chilean Agencia Nacional de Investigación y Desarrollo (ANID) through Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) Regular Project 1210992 (L. Chemin);
– the National Natural Science Foundation of China (NSFC) through grants 11573054, 11703065, and 12173069, the China Scholarship Council through grant 201806040200, and the Natural Science Foundation of Shanghai through grant 21ZR1474100;
– the Tenure Track Pilot Programme of the Croatian Science Foundation and the École Polytechnique Fédérale de Lausanne and the project TTP-2018-07-1171 'Mining the Variable Sky', with the funds of the Croatian-Swiss Research Programme;
– the Czech-Republic Ministry of Education, Youth, and Sports through grant LG 15010 and INTER-EXCELLENCE grant LTAUSA18093, and the Czech Space Office through ESA PECS contract 98058;
– the Danish Ministry of Science;
– the Estonian Ministry of Education and Research through grant IUT40-1;
– the European Commission's Sixth Framework Programme through the European Leadership in Space Astrometry (ELSA) Marie Curie Research Training Network (MRTN-CT-2006-033481), through Marie Curie project PIOF-GA-2009-255267 (Space AsteroSeismology & RR Lyrae stars, SAS-RRL), and through a Marie Curie Transfer-of-Knowledge (ToK) fellowship (MTKD-CT-2004-014188); the European Commission's Seventh Framework Programme through grant FP7-606740 (FP7-SPACE-2013-1) for the *Gaia* European Network for Improved data User Services (GENIUS) and through grant 264895 for the *Gaia* Research for European Astronomy Training (GREAT-ITN) network;
– the European Cooperation in Science and Technology (COST) through COST Action CA18104 'Revealing the Milky Way with *Gaia* (MW-Gaia)';
– the European Research Council (ERC) through grants 320360, 647208, and 834148 and through the European Union's Horizon 2020 research and innovation and excellent science programmes through Marie Skłodowska-Curie grant 745617 (Our Galaxy at full HD – Gal-HD) and 895174 (The build-up and fate of self-gravitating systems in the Universe) as well as grants 687378 (Small Bodies: Near and Far), 682115 (Using the Magellanic Clouds to Understand the Interaction of Galaxies), 695099 (A sub-percent distance scale from binaries and Cepheids – CepBin), 716155 (Structured ACCREtion Disks – SACCRED), 951549 (Sub-percent calibration of the extragalactic distance scale in the era of big surveys – UniverScale), and 101004214 (Innovative Scientific Data Exploration and Exploitation Applications for Space Sciences – EXPLORE);
– the European Science Foundation (ESF), in the framework of the *Gaia* Research for European Astronomy Training Research Network Programme (GREAT-ESF);
– the European Space Agency (ESA) in the framework of the *Gaia* project, through the Plan for European Cooperating States (PECS) programme through contracts C98090 and 4000106398/12/NL/KML for Hungary, through contract 4000115263/15/NL/IB for Germany, and through PROgramme de Développement d'Expériences scientifiques (PRODEX) grant 4000127986 for Slovenia;
– the Academy of Finland through grants 299543, 307157, 325805, 328654, 336546, and 345115 and the Magnus Ehrnrooth Foundation;
– the French Centre National d'Études Spatiales (CNES), the Agence Nationale de la Recherche (ANR) through grant ANR-10-IDEX-0001-02 for the 'Investissements d'avenir' programme, through grant ANR-15-CE31-0007 for project 'Modelling the Milky Way in the *Gaia* era' (MOD4Gaia), through grant ANR-14-CE33-0014-01 for project 'The Milky Way disc formation in the *Gaia* era' (ARCHEOGAL), through grant ANR-15-CE31-0012-01 for project 'Unlocking the potential of Cepheids as primary distance calibrators' (UnlockCepheids), through grant ANR-19-CE31-0017 for project 'Secular evolution of galaxies' (SEGAL), and through grant ANR-18-CE31-0006 for project 'Galactic Dark Matter' (GaDaMa), the Centre National de la Recherche Scientifique (CNRS) and its SNO *Gaia* of the Institut des Sciences de l'Univers (INSU), its Programmes Nationaux: Cosmologie et Galaxies (PNCG), Gravitation Références Astronomie Métrologie (PNGRAM), Planétologie (PNP), Physique et Chimie du

Milieu Interstellaire (PCMI), and Physique Stellaire (PNPS), the 'Action Fédératrice *Gaia* ' of the Observatoire de Paris, the Région de Franche-Comté, the Institut National Polytechnique (INP) and the Institut National de Physique nucléaire et de Physique des Particules (IN2P3) co-funded by CNES;
– the German Aerospace Agency (Deutsches Zentrum für Luft- und Raumfahrt e.V., DLR) through grants 50QG0501, 50QG0601, 50QG0602, 50QG0701, 50QG0901, 50QG1001, 50QG1101, 50QG1401, 50QG1402, 50QG1403, 50QG1404, 50QG1904, 50QG2101, 50QG2102, and 50QG2202, and the Centre for Information Services and High Performance Computing (ZIH) at the Technische Universität Dresden for generous allocations of computer time;
– the Hungarian Academy of Sciences through the Lendület Programme grants LP2014-17 and LP2018-7 and the Hungarian National Research, Development, and Innovation Office (NKFIH) through grant KKP-137523 ('SeismoLab');
– the Science Foundation Ireland (SFI) through a Royal Society - SFI University Research Fellowship (M. Fraser);
– the Israel Ministry of Science and Technology through grant 3-18143 and the Tel Aviv University Center for Artificial Intelligence and Data Science (TAD) through a grant;
– the Agenzia Spaziale Italiana (ASI) through contracts I/037/08/0, I/058/10/0, 2014-025-R.0, 2014-025-R.1.2015, and 2018-24-HH.0 to the Italian Istituto Nazionale di Astrofisica (INAF), contract 2014-049-R.0/1/2 to INAF for the Space Science Data Centre (SSDC, formerly known as the ASI Science Data Center, ASDC), contracts I/008/10/0, 2013/030/I.0, 2013-030-I.0.1-2015, and 2016-17-I.0 to the Aerospace Logistics Technology Engineering Company (ALTEC S.p.A.), INAF, and the Italian Ministry of Education, University, and Research (Ministero dell'Istruzione, dell'Università e della Ricerca) through the Premiale project 'MIning The Cosmos Big Data and Innovative Italian Technology for Frontier Astrophysics and Cosmology' (MITiC);
– the Netherlands Organisation for Scientific Research (NWO) through grant NWO-M-614.061.414, through a VICI grant (A. Helmi), and through a Spinoza prize (A. Helmi), and the Netherlands Research School for Astronomy (NOVA);
– the Polish National Science Centre through HARMONIA grant 2018/30/M/ST9/00311 and DAINA grant 2017/27/L/ST9/03221 and the Ministry of Science and Higher Education (MNiSW) through grant DIR/WK/2018/12;
– the Portuguese Fundação para a Ciência e a Tecnologia (FCT) through national funds, grants SFRH/BD/128840/2017 and PTDC/FIS-AST/30389/2017, and work contract DL 57/2016/CP1364/CT0006, the Fundo Europeu de Desenvolvimento Regional (FEDER) through grant POCI-01-0145-FEDER-030389 and its Programa Operacional Competitividade e Internacionalização (COMPETE2020) through grants UIDB/04434/2020 and UIDP/04434/2020, and the Strategic Programme UIDB/00099/2020 for the Centro de Astrofísica e Gravitação (CENTRA);
– the Slovenian Research Agency through grant P1-0188;
– the Spanish Ministry of Economy (MINECO/FEDER, UE), the Spanish Ministry of Science and Innovation (MICIN), the Spanish Ministry of Education, Culture, and Sports, and the Spanish Government through grants BES-2016-078499, BES-2017-083126, BES-C-2017-0085, ESP2016-80079-C2-1-R, ESP2016-80079-C2-2-R, FPU16/03827, PDC2021-121059-C22, RTI2018-095076-B-C22, and TIN2015-65316-P ('Computación de Altas Prestaciones VII'), the Juan de la Cierva Incorporación Programme (FJCI-2015-2671 and IJC2019-04862-I for F. Anders), the Severo Ochoa Centre of Excellence Programme (SEV2015-0493), and MICIN/AEI/10.13039/501100011033 (and the European Union through European Regional Development Fund 'A way of making Europe') through grant RTI2018-095076-B-C21, the Institute of Cosmos Sciences University of Barcelona (ICCUB, Unidad de Excelencia 'María de Maeztu') through grant CEX2019-000918-M, the University of Barcelona's official doctoral programme for the development of an R+D+i project through an Ajuts de Personal Investigador en Formació (APIF) grant, the Spanish Virtual Observatory through project AyA2017-84089, the Galician Regional Government, Xunta de Galicia, through grants ED431B-2021/36, ED481A-2019/155, and ED481A-2021/296, the Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC), funded by the Xunta de Galicia and the European Union (European Regional Development Fund – Galicia 2014-2020 Programme), through grant ED431G-2019/01, the Red Española de Supercomputación (RES) computer resources at MareNostrum, the Barcelona Supercomputing Centre - Centro Nacional de Supercomputación (BSC-CNS) through activities AECT-2017-2-0002, AECT-2017-3-0006, AECT-2018-1-0017, AECT-2018-2-0013, AECT-2018-3-0011, AECT-2019-1-0010, AECT-2019-2-0014, AECT-2019-3-0003, AECT-2020-1-0004, and DATA-2020-1-0010, the Departament d'Innovació, Universitats i Empresa de la Generalitat de Catalunya through grant 2014-SGR-1051 for project 'Models de Programació i Entorns d'Execució Paralells' (MPEXPAR), and Ramon y Cajal Fellowship RYC2018-025968-I funded by MICIN/AEI/10.13039/501100011033 and the European Science Foundation ('Investing in your future');
– the Swedish National Space Agency (SNSA/Rymdstyrelsen);
– the Swiss State Secretariat for Education, Research, and Innovation through the Swiss Activités Nationales Complémentaires and the Swiss National Science Foundation through an Eccellenza Professorial Fellowship (award PCEFP2_194638 for R. Anderson);
– the United Kingdom Particle Physics and Astronomy Research Council (PPARC), the United Kingdom Science and Technology Facilities Council (STFC), and the United Kingdom Space Agency (UKSA) through the following grants to the University of Bristol, the University of Cambridge, the University of Edinburgh, the University of Leicester, the Mullard Space Sciences Laboratory of University College London, and the United Kingdom Rutherford Appleton Laboratory (RAL): PP/D006511/1, PP/D006546/1, PP/D006570/1, ST/I000852/1, ST/J005045/1, ST/K00056X/1, ST/K000209/1, ST/K000756/1, ST/L006561/1, ST/N000595/1, ST/N000641/1, ST/N000978/1, ST/N001117/1, ST/S000089/1, ST/S000976/1, ST/S000984/1, ST/S001123/1, ST/S001948/1, ST/S001980/1, ST/S002103/1, ST/V000969/1, ST/W002469/1, ST/W002493/1, ST/W002671/1, ST/W002809/1, and EP/V520342/1.

The GBOT programme uses observations collected at (i) the European Organisation for Astronomical Research in the Southern Hemisphere (ESO) with the VLT Survey Telescope (VST), under ESO programmes 092.B-0165, 093.B-0236, 094.B-0181, 095.B-0046, 096.B-0162, 097.B-0304, 098.B-0030, 099.B-0034, 0100.B-0131, 0101.B-0156, 0102.B-0174, and 0103.B-0165; and (ii) the Liverpool Telescope, which is operated on the island of La Palma by Liverpool John Moores University in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias with financial support from the United Kingdom Science and Technology Facilities Council, and (iii) telescopes of the Las Cumbres Observatory Global Telescope Network.