

# Gaia FGK benchmark stars: Spectral library and abundances of $\alpha$ and Fe-peak elements of the third version

L. Casamiquela<sup>1,\*</sup>, C. Soubiran<sup>2</sup>, P. Jofré<sup>3</sup>, S. Vitali<sup>3</sup>, S. Blanco-Cuaresma<sup>4,5</sup>, N. Lagarde<sup>2</sup>,  
D. Slumstrup<sup>6,7</sup>, U. Heiter<sup>8</sup>, J. T. Palmerio<sup>9</sup>, N. Brouillet<sup>2</sup>, S. Elgueta<sup>3</sup>, A. Rojas-Arriagada<sup>10,11,12</sup>,  
C. Aguilera-Gómez<sup>13</sup>, I. Hernández-Araya<sup>13</sup>, O. L. Creevey<sup>14</sup>, L. Balaguer-Núñez<sup>15,16,17</sup>, and R. Carrera<sup>18</sup>

- <sup>1</sup> LIRA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CY Cergy Paris Université, CNRS, 92190 Meudon, France  
<sup>2</sup> Laboratoire d'Astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France  
<sup>3</sup> Instituto de Estudios Astrofísicos, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile  
<sup>4</sup> Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA  
<sup>5</sup> Faculty of Psychology, UniDistance Suisse, Brig, Switzerland  
<sup>6</sup> GRANTECAN, Cuesta de San José s/n, 38712 Breña Baja, La Palma, Spain  
<sup>7</sup> Instituto de Astrofísica de Canarias, 38205 La Laguna, Tenerife, Spain  
<sup>8</sup> Observational Astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden  
<sup>9</sup> Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191 Gif-sur-Yvette, France  
<sup>10</sup> Departamento de Física, Universidad de Santiago de Chile, Av. Víctor Jara 3659, Santiago, Chile  
<sup>11</sup> Millennium Institute of Astrophysics, Avenida Vicuña Mackenna 4860, Macul, Santiago 82-0436, Chile  
<sup>12</sup> Center for Interdisciplinary Research in Astrophysics and Space Exploration (CIRAS), Universidad de Santiago de Chile, Santiago, Chile  
<sup>13</sup> Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile  
<sup>14</sup> Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France  
<sup>15</sup> Institut d'Estudis Espacials de Catalunya (IEEC), Esteve Terradas, 1, Edifici RDIT, Campus PMT-UPC, 08860 Castelldefels (Barcelona), Spain  
<sup>16</sup> Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB), Martí i Franquès 1, 08028 Barcelona, Spain  
<sup>17</sup> Departament de Física Quàntica i Astrofísica (FQA), Universitat de Barcelona (UB), Martí i Franquès 1, 08028 Barcelona, Spain  
<sup>18</sup> INAF-Osservatorio di Astrofisica e Scienza dello Spazio, via P. Gobetti 93/3, 40129, Bologna, Italy

Received 18 April 2025 / Accepted 6 November 2025

## ABSTRACT

**Context.** Accurate determination of the chemical abundances in stars plays a pivotal role in understanding stellar structure and evolution, nucleosynthesis, and the chemical enrichment history of the Milky Way. Benchmark stars with precise and accurate atmospheric parameters and abundances are indispensable for calibrating spectroscopic surveys and testing stellar atmosphere models.

**Aims.** This study focuses on the compilation of high-quality spectra and the determination of chemical abundances of iron-peak and  $\alpha$  elements for the third version of the Gaia FGK benchmark stars (GBSv3).

**Methods.** We compiled spectra of the GBSv3 from public archives and complemented these with our observations. We used fundamental atmospheric parameters to perform a spectroscopic analysis using the public code iSpec and derived the chemical abundances.

**Results.** We compiled a homogeneous spectral library of high-resolution (42 000) and high signal-to-noise ( $>100$ ) normalised spectra for 202 stars, including the 192 GBSv3, nine stars with indirect measurement of the angular diameter from previous Gaia FGK benchmark versions, and the Sun. Using four radiative transfer codes, we derived chemical abundances of 13 chemical species (Fe I, Fe II, Mg I, Si I, Ca I, Ti I, Ti II, Sc II, V I, Cr I, Mn I, Co I, Ni I). We performed an in-depth study of several sources of error.

**Conclusions.** The GBSv3 contributes to the legacy samples of spectroscopic reference stars through improved statistics and homogeneity. This work offers the community a homogeneous spectral library and robust reference abundances for iron-peak and  $\alpha$  elements, supported by an extensive analysis of the associated uncertainties.

**Key words.** standards – stars: abundances – stars: atmospheres

## 1. Introduction

Over the past decade, significant observational efforts have been dedicated to uncovering the history of our Galaxy through the analysis of large samples of stars with abundances obtained from

spectroscopic surveys. A lesson learned from the first generation of high-resolution spectroscopic surveys of the Milky Way is that the abundance scales can differ significantly depending on the analysis methods (Jofré et al. 2017; Jönsson et al. 2018; Worley et al. 2020; Hegedűs et al. 2023), implying that different scientific conclusions may be reached depending on which spectroscopic dataset is used. It is then of paramount

\* Corresponding author: [laia.casamiquela@obspm.fr](mailto:laia.casamiquela@obspm.fr)

importance to have comprehensive validation samples that can put the different survey data onto the same scale, thus ensuring the data from all surveys are scientifically consistent (Jofré et al. 2019). An effort has already been made to stitch together surveys via data-driven approaches (Ho et al. 2017; Wheeler et al. 2020; Nandakumar et al. 2022; Thomas et al. 2024), such as The Cannon (Ness et al. 2015), among others. But large benchmark samples are still needed to ensure a robust homogenisation using such tools. At the same time, these calibration samples are useful for a variety of purposes, including as science verification of new instruments (Adibekyan et al. 2020; Strassmeier et al. 2018), for the testing of systematic uncertainties in analysis techniques (Blanco-Cuaresma 2019), for calibration of line lists (Heiter et al. 2021), and as a means of pushing the limits of our understanding of stellar atmospheres, structure, and evolution (Amarsi et al. 2022; Creevey et al. 2024).

The Gaia FGK benchmark stars (GBS) were assembled as a key reference set for this purpose, and they provide robust, homogeneous determinations of stellar parameters (Heiter et al. 2015); chemical compositions (Jofré et al. 2014, 2015; Hawkins et al. 2016); and high-quality spectral libraries (Blanco-Cuaresma et al. 2014). The GBS were selected to span F, G, and K (and M) spectral types across various luminosities and metallicities with sufficient observational data (particularly interferometry) to allow for determination of their effective temperature and surface gravity independently of spectroscopy, achieving a precision of 1–2%. The fact of having fundamental atmospheric parameters for these stars imposes strict constraints on spectroscopic analyses, enabling the determination of accurate reference values for chemical abundances.

The first GBS sample consisted of 30 stars (Heiter et al. 2015; Jofré et al. 2014), and the second version (Jofré et al. 2018) revised these stars and included metal-poor candidates analysed by Hawkins et al. (2016) to construct a second set of 35 stars. Recently, a third version of the GBS sample (GBSv3) has been presented by Soubiran et al. (2024), providing fundamental atmospheric parameters for a set of 192 stars. Compared to the previous versions, aside from increasing the number of stars in the sample, the homogeneity and accuracy of the fundamental  $T_{\text{eff}}$  and  $\log g$  are significantly improved thanks to the high quality of interferometric data and of the Gaia photometric and astrometric data. This paper follows the spirit of the previous series of papers and presents the chemical analysis of iron-peak and  $\alpha$ -capture elements: Fe I, Fe II, Sc II, V I, Cr I, Mn I, Co I, Ni I, Mg I, Si I, Ca I, Ti I, and Ti II. To do so, we performed systematic queries in different available public archives of high-resolution ( $R > 42\,000$ ) optical spectrographs complemented by our own observations. We compiled a total of 2539 individual spectra, which resulted in 522 combined spectra from the different instruments of the full sample of 192 GBS v3, including the Sun, and for ten additional stars that were part of the previous GBS versions but are not considered GBS because their angular diameters have an indirect measurement. We analysed these spectra using different methods to provide the community with a complete spectral library and detailed abundances for these reference stars.

This paper is organised as follows. We describe the source of the observational data and the pre-processing of the data to obtain a homogeneous spectral library in Sect. 2 and the spectroscopic analysis in Sect. 3. In Sect. 4, we provide details on the line selection used to derive the mean elemental abundances and the results obtained for the solar spectra. In Sect. 5, we describe the resulting Fe abundances of the full sample of stars, and Sect. 6 presents the results of the individual element

abundances. In Sect. 7, we analyse the different sources of uncertainty in our abundances. Finally, Sect. 9 contains the details of the accessibility of the data, and conclusions are described in Sect. 9.

## 2. Observational data

The GBS are nearby and bright stars. Thus, in general, they have been observed extensively with different high-resolution spectrographs. That is why the main source of our observational material are spectra retrieved from public archives. We have queried the archives of the following high-resolution optical spectrographs ( $R > 42\,000$ ): HARPS, UVES, and FEROS using the ESO phase3 spectroscopic specialised interface<sup>1</sup>; NARVAL and ESPaDOnS using the PolarBase database<sup>2</sup> (Petit et al. 2014); and the ELODIE archive via its public webpage<sup>3</sup> (Moultaka et al. 2004). From these public archives, we obtained 2335 spectra corresponding to 182 stars.

We also performed our own observational runs to complement this sample, with the UVES, FIES, HERMES and CAFE spectrographs. We particularly observed the remaining 20 stars, but we also re-observed several stars for comparison purposes. In total, we have obtained 204 spectra corresponding to 51 different stars. These observations allowed us to complete the observations of the full GBSv3 sample of 201 stars presented in Soubiran et al. (2024). We add to our sample of spectra the high-S/N solar spectra used in the first version of the GBS (Paper I): four from HARPS, one from NARVAL and one from UVES.

### 2.1. Instruments and data reduction

HARPS (Mayor et al. 2003) is an ESO facility mounted at the 3.6 m telescope in La Silla, and it is mainly dedicated to the measurement of radial velocities for exoplanet search. The spectral range covered is 378–691 nm with a resolving power of 115 000.

UVES (Dekker et al. 2000) is the high-resolution optical spectrograph of the Very Large Telescope (VLT), located at its second unit. We have used the 580 setup which gives as a result the spectrum divided in two parts covering from 476 to 580 nm (lower part) and from 582 to 683 nm (upper part). The resolving power of the instrument depends on the slit used, going from  $\sim 40\,000$  to  $\sim 100\,000$ .

FEROS (Kaufer et al. 1999) is a spectrograph installed at the MPG/ESO 2.2-metre telescope located at ESO's La Silla Observatory. It covers the full optical spectrum from the near-ultraviolet to the near-infrared  $\sim 350$ – $920$  nm with a resolving power of around 48 000.

Reduced, calibrated, extracted and merged 1D (unnormalised) spectra from HARPS, UVES and FEROS are provided in the ESO spectroscopic archive, obtained with the ESO science grade pipelines. We additionally obtained a spectrum with a dedicated DDT program (ID 112.26WR) for HIP 98624, in this case we used the public ESO UVES v6.4.1 pipeline to obtain the 1D spectrum.

NARVAL (Aurière 2003) and ESPaDOnS (Donati et al. 2006) are two twin spectro-polarimeters mounted at the 2.03 m Telescope Bernard Lyot (TBL) at the Pic du Midi Observatory, and at the 3.6 m Canada France Hawaii Telescope

<sup>1</sup> [https://archive.eso.org/wdb/wdb/adp/phase3\\_spectral/form](https://archive.eso.org/wdb/wdb/adp/phase3_spectral/form)

<sup>2</sup> <http://polarbase.irap.omp.eu/>

<sup>3</sup> <http://atlas.obs-hp.fr/elodie/>

(CFHT) located atop the summit of Mauna Kea. They both are capable of obtaining a full optical spectrum from 370 to 1050 nm with a resolving power between 68 000 and 80 000, depending on the observation mode. Spectra from both instruments are automatically reduced, calibrated and extracted with dedicated data reduction pipelines based on the Libre-ESPRIT software (Donati et al. 1997). We have used our own pipeline, based on iSpec, to merge the spectra orders by performing a weighted mean in the overlapping regions.

ELODIE (Baranne et al. 1996) was an echelle spectrograph with resolution of  $\sim 42\,000$  used at the 1.93 m telescope of Observatoire de Haute Provence (OHP) between late 1993 and mid 2006. The 1D spectra provided in the ELODIE dedicated archive (Moultaka et al. 2004) are resampled in wavelength with a constant step of 0.005 nm covering the range 400–680 nm and given in “instrumental” flux, homogeneously treated by a procedure developed by Prugniel & Soubiran (2001).

FIES (Telting et al. 2014) is a high-resolution echelle spectrograph mounted at the 2.5 Nordic Optical Telescope (NOT) at the Roque de los Muchachos Observatory (ORM), with a maximum spectral resolution of  $\sim 67\,000$  covering the entire spectral range 370–830 nm. The reduction, extraction calibration and order merging of the obtained spectra was done automatically in the telescope with the FIESTool software.

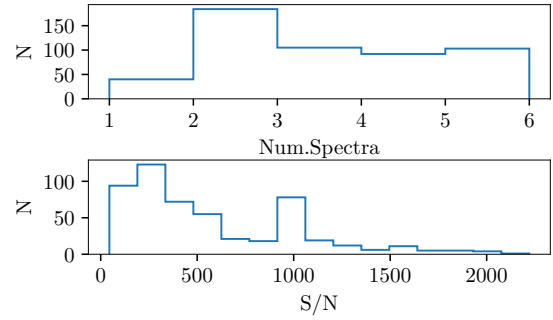
HERMES (Raskin et al. 2011) is a highly efficient echelle spectrograph, which became the workhorse instrument of the 1.2 m Mercator telescope on the ORM, covering the entire wavelength range from 380 to 900 nm with a spectral resolution of 85 000. The obtained spectra were automatically reduced, extracted, calibrated and order-merged with the HERMESDRS pipeline.

CAFE (Aceituno et al. 2013) is a high-resolution (60 000) spectrograph covering the visible range (400–920 nm), mounted at the 2.2 m telescope of the Centro Astronómico Hispano en Andalucía (CAHA) or Calar Alto Observatory. The spectra were reduced with the CAFE data reduction pipeline (Lillo-Box et al. 2020), which reduces, extracts, calibrates and merges the orders obtained with the spectrograph.

## 2.2. Combined spectra

We queried the aforementioned archives, selecting individual spectra with a minimum signal-to-noise ratio (S/N; coming from the data reduction pipelines) of 20, and we co-add all the unnormalised spectra from the same instrument (see Sect. 3.1) to reach a minimum S/N  $\sim 100$ . We obtained a total of 522 combined spectra for the 202 stars (including the Sun). Figure 1 shows the distribution of the number of spectra from the same instrument that were used for each combined spectrum and the S/N distribution of all the combined spectra. Most of the stars are observed with more than one instrument (not shown in the figure), and only 30 stars have a single observation. In particular, three stars (HIP 12114, HIP 57757 and HIP 79672) have been observed with six instruments.

Most of the combined spectra have a S/N well above 100, with the mean S/N being  $\sim 600$  and the maximum being  $\sim 2200$  for HIP 37826 (Pollux). By construction of the query of spectra in the public archives, there is a peak of the S/N distribution at 1000. This is because there are several stars for which we found a particularly large number of spectra of the same instrument available, generally coming from exoplanet search projects. To avoid downloading and combining too many spectra, we only selected the best S/N ones up to when we could reach a combined S/N of 1000.



**Fig. 1.** Distribution of the number of combined spectra per instrument and star (top), and S/N distribution of these combined spectra (bottom).

## 3. Data handling and analysis

We analysed the full sample of spectra using an automated pipeline based on iSpec (version 2020.10.01; Blanco-Cuaresma et al. 2014; Blanco-Cuaresma 2019). iSpec is a Python software designed to perform operations on stellar spectra and to compute radial velocities, atmospheric parameters, and chemical abundances using different atmospheric models and radiative transfer codes. The used pipeline consists of a first pre-processing of the data to ensure a homogeneous treatment of the spectra from different instruments and a subsequent common strategy for the spectroscopic analysis.

### 3.1. Pre-processing

The spectra were obtained with nine different instruments, each one with different technical specifications and different conventions (e.g. the presence of an explicit error spectra, and specific wavelength handling). This is why, to construct a spectral library and to perform a homogeneous spectroscopic analysis, we used a common pipeline developed in Blanco-Cuaresma & Fraix-Burnet (2018) and later adapted in Casamiquela et al. (2020), which performs a pre-processing of the spectra described as follows:

- Each spectrum is cut to a restricted common wavelength range (480 to 680 nm) and downgraded to a common resolution (42 000). A homogeneous resampling of 0.001 nm steps is also performed.
- Negative fluxes are set to zero, and errors are estimated in a common manner using the S/N with the dedicated native iSpec function (which essentially divides the flux by the S/N).
- Telluric regions are identified by a cross-correlating with a telluric line list and masked, and the fluxes are set to zero.
- For the particular case of UVES spectra, the red- and blue-arm spectra for the same star are sometimes provided in separate files; in these cases, the spectra are merged into one, inserting a gap (of zero flux) between them.
- The different spectra from the same star and instrument are co-added to increment the final S/N. To do so, one spectrum is selected as reference, then we perform a cross-correlation of all other spectra of the same star against the reference. Finally, we sum up the absolute fluxes of all aligned spectra.
- Radial velocity determination is performed with the co-added spectra from cross-correlation with a high S/N solar spectrum from NARVAL, as described in Blanco-Cuaresma et al. (2014). The resulting radial velocities are used to correct the co-added spectra to laboratory wavelengths, ready for the spectroscopic analysis. Heliocentric radial velocities



are derived from this procedure after applying the barycentric correction. However, for spectra obtained with HARPS and ELODIE, we adopt the mean of the individual radial velocities available from public archives, as these values are determined by dedicated velocimetric pipelines expected to provide the highest precision.

- The continuum is determined homogeneously, computing a synthetic spectrum using the fundamental  $T_{\text{eff}}$  and  $\log g$  derived by Soubiran et al. (2024). The observed and synthetic spectra are divided, and the result is smoothed with a median filter to determine the continuum placement. Finally, the spectra are then normalised with the determined continuum using the native iSpec function.

Our pipeline is not optimised to analyse spectroscopic binaries, but our algorithm to derive the radial velocity (which uses the iSpec native function *cross\_correlate\_with\_template*) is able to automatically detect double-line spectroscopic binaries. Once the velocity profile was constructed from the cross-correlation process, local minima were searched in the smoothed profile (of twice the resolution). The mean velocity was calculated by fitting a Gaussian profile near the main peak. If other peaks are found with a probability above 50%, other Gaussian profiles are fitted. We have found a star in the GBS v3 sample which shows a secondary peak in the cross-correlation function (CCF): HIP 109176. It is an F dwarf with roughly solar metallicity for which we have a NARVAL spectrum (with very high S/N: 1683) and 38 HARPS spectra of various S/N ranging from 60 to 230. The CCF of the NARVAL spectrum and of most of the individual HARPS spectra show a faint secondary peak, indicating that two sets of lines at different radial velocities are visible. When we visually inspected the two co-added spectra, we saw that they differ significantly, even though we could identify common lines from the main component. This is possibly due to the different phases in which the double-line spectroscopic binary was observed and to the fact that the HARPS spectrum results from the co-addition of 38 spectra taken in different phases. We flag the abundance results of this star in our final tables.

There are another 16 stars in the GBSv3 sample identified as spectroscopic binaries in the literature: HIP 171, HIP 27913, HIP 37279, HIP 3850, HIP 4151, HIP 45343, HIP 46853, HIP 5336, HIP 5445, HIP 61317, HIP 67927, HIP 71683, HIP 72567, HIP 81693, HIP 86614, HIP 92512. We did not see any relevant feature in the CCFs nor the spectra of these stars, indicating that the contribution of the companion is possibly too faint to be detected.

### 3.2. Spectroscopic analysis

We employed the spectral synthesis method to compute the atmospheric parameters and individual chemical abundances using four different available radiative transfer codes: SPECTRUM (version 2.77; Gray & Corbally 1994), Spectroscopy Made Easy (SME; version 574; Valenti & Piskunov 1996; Piskunov & Valenti 2017), MOOG (version November 2019; Sneden et al. 2012), and TURBOSPECTRUM (version 19.1; Alvarez & Plez 1998; Plez 2012; see also Gerber et al. 2023). All of the codes assume local thermodynamic equilibrium (LTE) and solve the equation of radiative transfer within a one-dimensional hydrostatic stellar atmosphere. Atmospheric models with plane-parallel geometry are used for stars with a surface gravity of  $\log g=3.0$  or higher, while spherical models are used at lower  $\log g$  values. In the latter case, the radiative transfer calculations are done in spherically symmetric geometry for two of the codes

(SME and TURBOSPECTRUM), and in plane-parallel geometry for the others (SPECTRUM and MOOG). Further differences between the codes are in part of technical<sup>4</sup> and numerical nature. For example, the algorithms used for solving the equation of radiative transfer range from an attenuation operator method with a Bézier spline approximation for the source function in the case of SME to a Feautrier scheme including scattering in the case of TURBOSPECTRUM. Further aspects of line formation which may be treated differently in each code concern the calculation of the continuous opacity and the computation of molecular equilibrium. A subtle difference between MOOG and the others, already mentioned and discussed in Jofré et al. (2017) and Blanco-Cuaresma (2019), is that it does not recompute electron densities, but uses the values provided by the input model atmosphere. Concerning the line opacity, differences may arise from the “computation radius” used for each line<sup>5</sup>, and the calculation of the line profile including treatment of radiative and collisional broadening. Finally, the calculation of the surface flux from specific intensities may be implemented in different ways, e.g. by integration over the stellar disk or applying a limb darkening correction. For details the reader is directed to the publications referred to for each code and Blanco-Cuaresma (2019), as an in-depth description would go beyond the scope of this article.

For all codes, we use the MARCS<sup>6</sup> atmospheric models (Gustafsson et al. 2008), the solar abundances by Grevesse et al. (2007), and the sixth version of the line list from the Gaia-ESO survey (Heiter et al. 2021). In the case of TURBOSPECTRUM, in addition to the atomic lines, we use the Gaia-ESO molecular line lists as compiled by Thomas Masseron (private communication), which are also publicly accessible<sup>7</sup> (Gerber et al. 2023). This molecular information represents more than 3 million transitions, in contrast to the ~140 000 atomic transitions in the Gaia-ESO atomic line list. TURBOSPECTRUM is specifically optimized to handle such a vast number of transitions efficiently. The other radiative transfer codes, however, are not designed for molecular line lists of this magnitude, and their inclusion would lead to computationally prohibitive processing times. Therefore, the use of these molecular data was restricted to the TURBOSPECTRUM analysis.

For all four codes, spectral fitting was performed by comparing observed fluxes (weighted by their uncertainties) to synthetic spectra of the star’s  $T_{\text{eff}}$  and  $\log g$ , computed at the observed resolution (42 000) for pre-selected line regions. Atmospheric parameters and chemical abundances were varied in two separate steps until convergence was reached using a least-squares algorithm. The fitting is performed only in a set of pre-selected line regions provided by the iSpec default installation, selected to give accurate chemical abundances for the Sun. Two line selections are provided for each code: one optimized for determining atmospheric parameters and a second to be used for a line-by-line determination of individual chemical abundances (see Blanco-Cuaresma 2019, for further details). For each line region, a cross-correlation with the synthetic template is performed to allow small line shifts, and afterwards, the line mask is automatically adjusted to avoid nearby overlapping lines. A basic quality control of the linemasks is done, for instance, rejecting lines

<sup>4</sup> SPECTRUM and SME are written in the C programming language, while TURBOSPECTRUM and MOOG are written in Fortran.

<sup>5</sup> This implies a decision on which transitions to include at each wavelength point, see e.g. <https://www.appstate.edu/~grayro/spectrum/spectrum276/node23.html>.

<sup>6</sup> <http://marcs.astro.uu.se/>

<sup>7</sup> <https://keeper.mpg.de/d/6aeecbf95b88448f98a4/>

detected at too high line shifts (there is a risk of line misidentification), lines for which the flux at the detected peak is null, linemasks which overlap with the peak of a telluric line, lines with too poor fits (high root mean square) or for which the peak of the flux is too different with respect to the peak in the synthetic spectrum.

As for the atmospheric parameters,  $T_{\text{eff}}$  and  $\log g$  were fixed to the fundamental values derived by Soubiran et al. (2024), and each code infers independently  $[M/H]$ ,  $[\alpha/M]$ , as well as the microturbulence parameter  $v_{\text{mic}}$ . The overall metallicity fitted in this step is inferred from the atmosphere models: a zeropoint to the solar mixture, which is obtained by fitting at the same time all linemasks. Since we do not only use Fe lines in this procedure, the  $[M/H]$  is not necessarily equal to the  $[\text{Fe}/H]$ , even though they are similar. The  $[\alpha/M]$  abundance obtained in this step depends on the previous overall metallicities, following the predefined  $[\alpha/M]$  enhancement at low metallicities from the MARCS models. In this work, we do not consider these  $[M/H]$  and  $[\alpha/M]$  values, since we obtain the elemental abundances in a second step instead. Concerning the broadening effects, the projected equatorial rotational velocity  $v \sin i$ , the macroturbulence parameter, and the spectral resolution are degenerated and difficult to disentangle. We applied a similar strategy as in Blanco-Cuaresma (2019): we used a fixed value for  $v \sin i$  of  $1.6 \text{ km s}^{-1}$  and the macroturbulence (fixed to  $0 \text{ km s}^{-1}$ ), and only the spectral resolution at which the synthetic spectra are generated was left free, accounting for all broadening effects. For the Sun, we used the fixed atmospheric parameters  $T_{\text{eff}} = 5771 \text{ K}$  and  $\log g = 4.44 \text{ dex}$ , based on the GBSv1 recommended values of Heiter et al. (2015).

The absolute chemical abundances of individual lines are measured in a second step using the atmospheric parameters fixed to the values resulting from the previous step. To ensure reliable measurements, we only consider lines whose reduced equivalent widths (REWs<sup>8</sup>) fall within the range  $-6.7 < \text{REW} < -4.5$ . This criterion helps exclude lines that are either too weak, which can be affected by noise, or too strong, which can suffer from saturation effects. The choice of these limits is not straightforward in a sample with such a variety of stellar parameters and metallicities, but it is a trade-off between the number of lines detected and the accuracy of the resulting abundances. In particular, the lower cut in REW is motivated by the fact that we are able to measure very weak lines for high S/N spectra, which is particularly crucial for metal-poor stars. However, it tends to select very weak lines in moderately noisy spectra, which do not give precise abundances (particularly for solar and metal-rich stars). Thus, to not compromise accuracy when the S/N is not that high, we have added an additional cut in the depth of the line with respect to the continuum, compared with the overall S/N:  $\text{depth} > 3\sigma$ , where  $(\sigma = \frac{1}{\sqrt{N}})$ .

After the previous filtering, for each spectrum, we use the line selection detailed in Sect. 4.1, and we apply a  $2.5\sigma$  clipping to discard lines which may be affected by cosmics or other issues in particular spectra. This is only done for spectra and elements which have more than 5 lines measured. Finally, we compute the average absolute abundance per spectrum and element, and we use the solar abundances for each code derived in Sect. 4 to compute solar-scaled  $[X/H]$  values. We also compute abundances per star (for each code) obtained by averaging the values obtained by the different spectra and the scatter among the spectra.

## 4. Solar abundances and line selection

In this section, we detail the analysis of the line-by-line chemical abundances obtained for the six solar spectra from UVES, HARPS, and NARVAL using the different radiative transfer codes. We explain the criteria we used for the line selection, and we later compare the abundance results with literature.

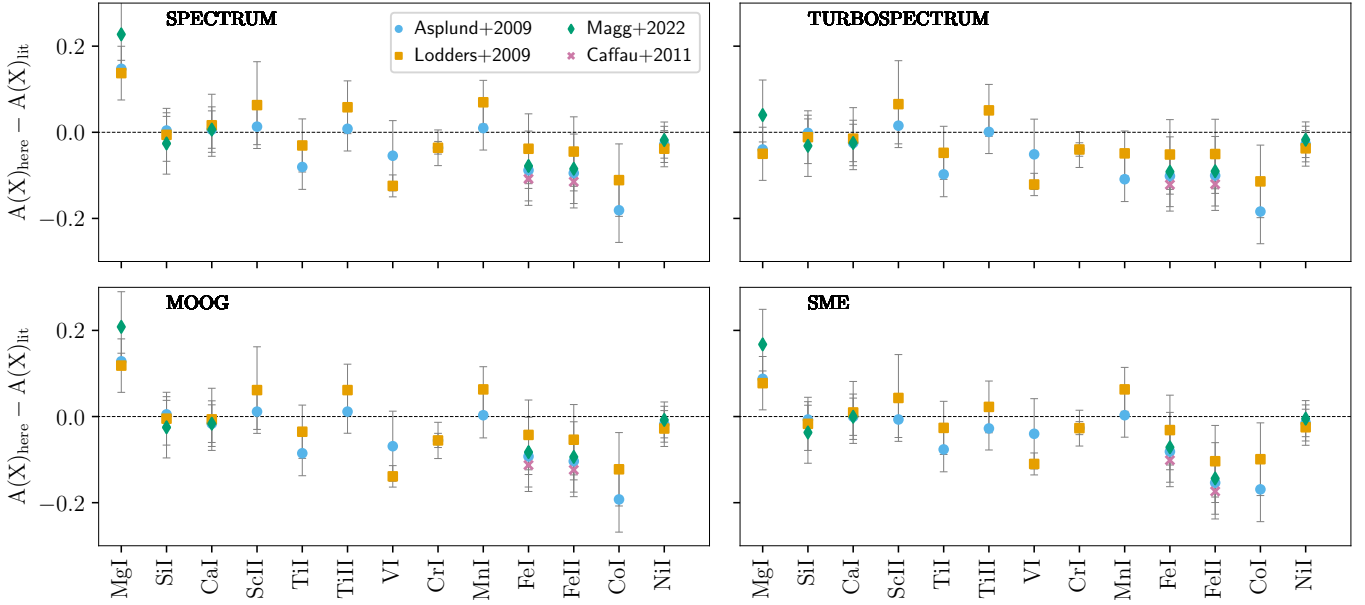
### 4.1. Line selection

We used the solar spectra to perform a selection of lines per element and code based on the line coherence given by the six solar spectra. We based our selection on the flags provided by the Gaia-ESO survey line list study (Heiter et al. 2021). For each studied line, a *loggflag* and a *synflag* were provided, with possible values: ‘yes’ (Y), ‘undetermined’ (U), and ‘no’ (N). These flags indicate the level of accuracy of the atomic  $\log gf$  value and the level of blending from neighbouring lines, respectively. We excluded from our selection all lines with *loggflag*=N for having an unreliable determination of their  $\log gf$ .

In this work, we focus on the analysis of iron-peak and  $\alpha$  elements, which have a minimum of one line flagged as *loggflag*=Y and *synflag*=Y (YY). We based the selection of lines using as reference the YY lines. We then proceeded by keeping additional lines that give abundances similar to the reference lines (up to  $\pm 0.05 \text{ dex}$ ) and that give similar dispersions as the reference lines among the six spectra (up to  $2\sigma$ ). We highlight the following particular cases:

- The two YY lines for Si I (569.0425 and 570.1104 nm) are very often affected by telluric regions in our full sample of spectra. Thus, for homogeneity reasons, we discard those lines, and instead, we keep the seven lines which are compatible with absolute abundances from the literature for the Sun (see next subsection) and compatible among them. They all have *loggflag*=Y.
- The two YY lines for Sc II (565.7896, 566.7149 nm) give consistent abundances among them but 0.15 dex larger than literature values. We were unable to determine the cause of this discrepancy; consequently, we decided to use the YU/YN lines 523.9813, 552.679, 566.9042, and 568.4202 nm, which give very coherent abundances and are consistent with literature values.
- For Mg I the two YY lines correspond to the Mg triplet, which are too strong and out of the REW range considered in this work. There are five other Mg lines measured: on one side, the 550.9597 and 578.5313 do not have any defined flag and are both in a crowded region; and on the other side, we find the 571.1088 (UY), 631.8717 (UU), 631.9237 (UY) with defined flags. 631.9237 is slightly blended, and both 631.9237 and 631.8717 provide very high absolute abundances with respect to the literature for all four radiative transfer codes, around 7.75 compared to usual literature values of 7.45–7.5 (e.g. Magg et al. 2022). This issue seems to come from a small offset between the continuum of the synthesised spectra and the observed spectra in the small region [631.6, 632.2] nm. Since the continuum of the synthetic spectra is slightly higher than the observed one, a larger abundance is needed to provide a good fit at the centre of the line. This effect is particularly relevant for dwarfs because these two lines are much weaker than for the giants. We have checked that performing the fit with a higher resolution (HARPS spectra at 115 000) does not solve the problem and that this difference in the continuum is not instrument-dependent. Additionally, this region is in

<sup>8</sup>  $\text{REW} = \log_{10}(\text{EW}/\lambda)$  where EW is the equivalent width and  $\lambda$  is the wavelength.



**Fig. 2.** Literature comparison (with different colours and symbols) of the mean solar abundances obtained here for the four radiative transfer codes (in different panels). Each point represents the mean abundances obtained for all solar spectra, averaging all the measured lines.

the middle of strong telluric bands. Thus, we speculate that this continuum difference may be due to telluric contamination or a lack of information on the underlying line list. Since iSpec’s best results are provided for 571.1088, we decided to keep only this line. Still, we warn the reader that this line gives compatible solar abundances for the literature only for TURBOSPECTRUM (see Fig. 2) and an overestimation of the Mg abundance for the other three codes. We highlight that the abundances obtained from the Mg triplet fit are very close to the literature values, but, as mentioned before, these two lines do not satisfy the established upper limit in REW and thus are discarded from our final results.

#### 4.2. Comparison with literature

We present in Figs. A.1 and A.2 a line-by-line comparison of the absolute<sup>9</sup> abundances obtained in this work for the previously selected lines, with several solar values from literature (Asplund et al. 2009; Caffau et al. 2011; Lodders et al. 2009; Magg et al. 2022). Fig. 2 presents the overall literature comparison of the mean solar abundance values obtained for each radiative transfer code. Each point represents the average among the six solar spectra of the mean line abundances. The displayed errors in the  $y$ -axis are computed as  $\sqrt{\delta A_{\text{here}}^2 + \delta A_{\text{lit}}^2}$ , with  $\delta A_{\text{here}}$  being the error computed from the product of posterior distribution functions as explained in Sect. 7.4, and  $A_{\text{lit}}$  the errors quoted in the literature.

As seen in Fig. 2, we see an overall good agreement with the literature, with a slight tendency of our solar abundances to be underestimated for all radiative transfer codes (overall average of  $-0.04$  dex). In particular, for Fe I, we obtain a mean difference of  $-0.08$  with all literature values and all codes (very similar value for Fe II, except for SME which produces a larger offset), though it is compatible at  $1\sigma$  of the uncertainties. We highlight the case

of Mg for which the selected line is in agreement with the literature only for TURBOSPECTRUM. SPECTRUM, MOOG and SME obtain an overestimation of the absolute abundances of around  $0.10$ – $0.15$  dex for this line. The other Mg lines cited in Sect. 4.1 produce even larger abundance offsets compared to the literature.

We detail the mean solar abundances obtained with the four codes in Table A.1. Our uncertainties tend to be small, particularly those computed from the dispersion among the different spectra, which are a reflection of the high S/N of all solar spectra. The uncertainties coming from the line dispersions are more realistic but are not available for Mg, which only has one single line analysed. Finally, the mean of the quoted errors is always larger than the dispersion among spectral lines (see discussion in Sect. 7.1 for a thorough explanation).

#### 5. Fe abundances

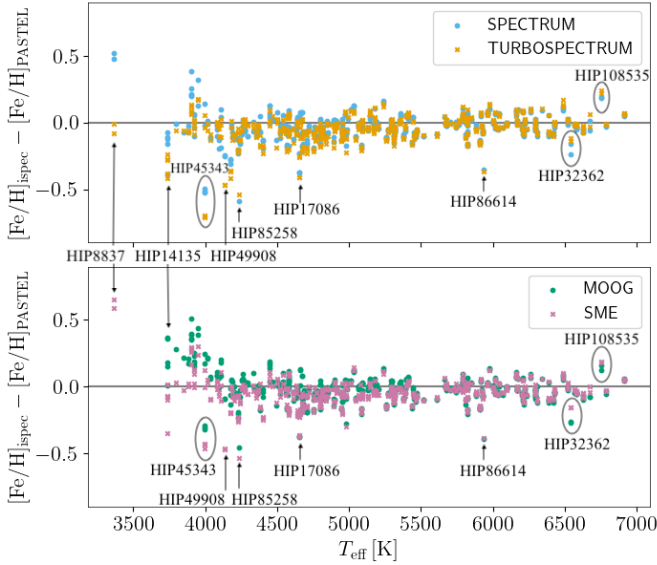
In this section, we study the Fe abundance results per spectrum for all the stars using the mean of the Fe I abundances and subtracting the solar Fe abundance obtained for each radiative transfer code. We compare our results with those adopted by Soubiran et al. (2024), which are a compilation from the literature mainly based on the PASTEL catalogue (Soubiran et al. 2016).

We plot the comparison of the Fe abundance as a function of effective temperature in Fig. 3, for each radiative transfer code. We observe a general good consistency between the literature and the values that we have obtained in this work with the four radiative transfer codes. The mean offsets with respect to the literature are slightly negative, of around  $-0.04$  dex.

A few interesting trends can be seen in Fig. 3, particularly for the coolest stars. We obtained an increasingly large dispersion in the comparison for stars cooler than  $T_{\text{eff}} \sim 5000$  K. In the case of SPECTRUM, MOOG, and SME, a clear trend towards higher  $[\text{Fe}/\text{H}]$  compared to PASTEL is visible at  $T_{\text{eff}} < 4500$  K. This is possibly due to the increasing importance of undetected or poorly modelled blends in cool stars and, in particular, the

<sup>9</sup> Absolute abundances expressed in the form of  $\log(N_X/N_H) + 12$ , with  $N_X$  and  $N_H$  the number of absorbers of the element X, and of hydrogen, respectively.





**Fig. 3.** Abundance of Fe per spectrum obtained in this study compared with the one from the literature adopted for the determination of fundamental  $T_{\text{eff}}$  and  $\log g$  in Soubiran et al. (2024) as a function of  $T_{\text{eff}}$  (fundamental value). The two panels and the different colours show the values from the different radiative transfer codes. Some outliers discussed in the text are indicated (the MOOG value of HIP 8837 is outside of the  $y$ -axis limits at  $\sim 1.3$ ).

presence of molecules. Consequently, the behaviour of TURBOSPECTRUM, which we use for synthesising molecular lines, is significantly different to the other three codes and shows no overall trend. However, we obtain a negative zero point of  $-0.1$  dex for stars at  $T_{\text{eff}} < 4500$  K for TURBOSPECTRUM, which for SPECTRUM and SME is  $-0.05$  and  $-0.03$  respectively, and for MOOG is positive (0.08 dex). A similar zero point for TURBOSPECTRUM was already found in previous GBS versions (Jofré et al. 2014), where it was run by the ULB method, which used BACCHUS (Masseron et al. 2016).

There are a few remarkable cases in the comparison to literature, which we highlight in the plot and comment on below (from left to right):

- HIP 8837 (psiPhe) is the coolest star in our sample, a metal-poor M giant of 3362 K, for which we have a HARPS and a UVES spectrum (the same as GBSv1). It is a particularly difficult case, and the Fe abundance determinations of SPECTRUM, SME and MOOG<sup>10</sup> seem highly overestimated. Additionally, the differences obtained between the two instruments are also notable, even though they are compatible within the quoted uncertainties, which are always  $>0.5$  dex. Interestingly, the results from TURBOSPECTRUM seem more consistent, and they differ by only 0.1 dex with the only determination of  $[\text{Fe}/\text{H}]$  reported for that star in PASTEL, which is the GBSv1 value by Jofré et al. (2014).
- HIP 14135 (alfCet) is the second coolest giant in the sample (3738 K) for which five spectra are available. The absolute discrepancies with the PASTEL values are overall smaller than in the previous case, though the scatter among instruments is large for SPECTRUM, MOOG and SME (0.12–0.15 dex). Again, TURBOSPECTRUM seems to give a more consistent Fe abundance for this star with 0.07 dex scatter among the five spectra. We note that the nine literature

values reported in PASTEL also show a large dispersion of  $T_{\text{eff}}$  and Fe abundance, with  $[\text{Fe}/\text{H}]$  ranging from  $-0.60$  to  $+0.02$  dex, which reflects the difficulty in obtaining robust abundances for cool stars.

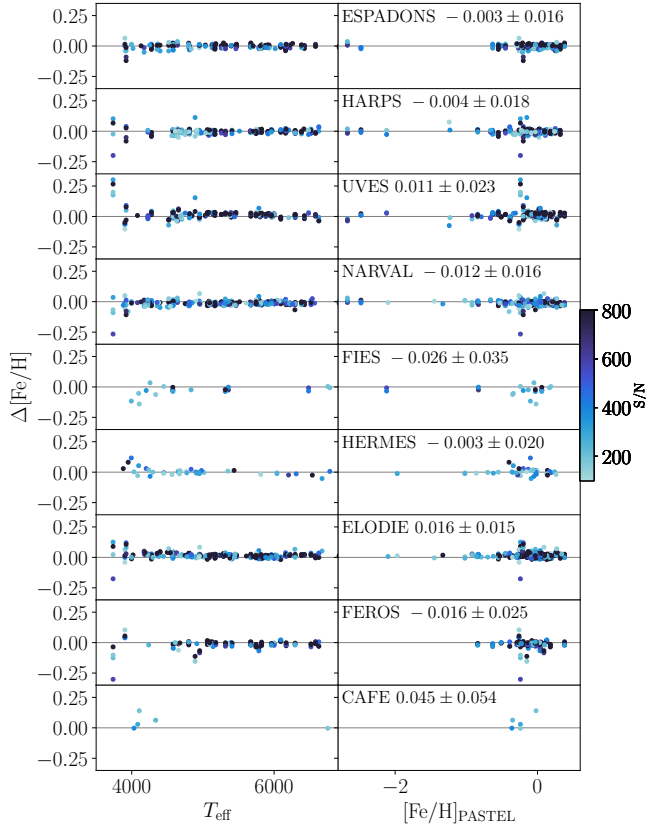
- HIP 45343 is the coolest dwarf in the sample (3997 K), for which we have three spectra that give overall consistent abundances. Depending on the code, we obtain Fe abundance between  $[\text{Fe}/\text{H}] \sim -0.15$  and  $-0.4$ , with TURBOSPECTRUM giving the largest discrepancy with the two literature values from PASTEL, which are  $[\text{Fe}/\text{H}] \sim 0.17$  and  $[\text{Fe}/\text{H}] \sim 0.34$ .
- HIP 49908 is the second coolest dwarf in the sample (4132 K), for which PASTEL gives a mean Fe abundance of  $+0.21 \pm 0.06$  dex. We have spectra from NARVAL and ESPADONS, for which SPECTRUM and MOOG obtain  $[\text{Fe}/\text{H}]$  at around 0, and TURBOSPECTRUM and SME around  $-0.2$ , in all cases with uncertainties of 0.25 dex.
- HIP 85258 (betAra) is a red supergiant at 4232 K for which we have a HARPS spectrum, the same as GBSv1. For all codes, we obtain similar  $[\text{Fe}/\text{H}]$  of around  $\sim 0$  with relatively large uncertainties  $\sim 0.15$ , which is consistent with the value obtained in GBSv1 and contrasts with the only other literature value in PASTEL of  $+0.5 \pm 0.1$  dex (with  $T_{\text{eff}} = 4582$  K, 300 K higher than our value).
- HIP 17086 is a red giant at 4652 K for which we have HARPS and HERMES spectra, and for which all codes give a consistent difference of  $-0.35$  dex with the single literature value listed in PASTEL (with  $T_{\text{eff}} = 4905$  K, 250 K higher than our value).
- HIP 86614 is part of a binary star with the two components being resolved by high-resolution spectroscopy. We have observed the A component, a subgiant of 5936 K for which we obtain  $[\text{Fe}/\text{H}] = -0.4$  (from a single NARVAL spectrum). The five literature values in PASTEL are in good agreement, with a mean of  $[\text{Fe}/\text{H}] = +0.01$ , but with corresponding  $T_{\text{eff}}$  about 400 K larger than our fundamental value.
- We have two spectra (HERMES and NARVAL) from HIP 32362, an F dwarf of 6537 K for which we obtain slightly metal-poor values while the literature quotes a single value of  $0.14 \pm 0.1$  (Yee et al. 2017). This star has very broad spectral lines because it is a high rotator (66  $\text{km s}^{-1}$ , Schröder et al. 2009), and thus our  $[\text{Fe}/\text{H}]$  uncertainties are also of the order of 0.1, larger than other stars at the same temperature.
- HIP 108535 is also a moderately high rotator at 6754 K (26  $\text{km s}^{-1}$ , Schröder et al. 2009). In this case, the abundances that we obtain with the two spectra (FIES and CAFE) give consistent values of  $[\text{Fe}/\text{H}] \sim -0.05$ , which are 0.2 dex more metal-rich than the single literature value from PASTEL with a  $T_{\text{eff}}$  about 600 K lower than our fundamental value.

### Dependence on the instrument

We used the sample of stars that were observed with more than one instrument to check the internal consistency of the Fe abundance and any possible instrumental bias. For each instrument, we compute the difference between the  $[\text{Fe}/\text{H}]$  obtained for the spectrum of that instrument and the value for the same star observed with other instruments. In the plot, we only show the TURBOSPECTRUM values; see below.

Figure 4 shows the results obtained for the Fe abundance for different instruments. We see a very high consistency of the Fe abundance in general, better than 0.03 dex. The plot against

<sup>10</sup> The point of the MOOG comparison for this star is out of the plot limits in Fig. 3, the difference is around  $+1.2$  dex.



**Fig. 4.** Difference in Fe abundance per instrument (computed as  $[\text{Fe}/\text{H}]_{\text{mean}} - [\text{Fe}/\text{H}]_{\text{instr}}$ ) as a function of effective temperature (left) and PASTEL  $[\text{Fe}/\text{H}]$  (right). The colour depicts the S/N of the spectra. The average and spreads of the differences for stars above 4000 K are written in each panel.

effective temperature clearly shows the difficulty in obtaining reliable abundances for  $T_{\text{eff}} \lesssim 4000$  K, without dependence on the instrument or on the S/N. We include in Figure 4 the average and spread of these Fe abundance differences for stars with  $T_{\text{eff}}$  above 4000 K. The mean values are always well below the spreads and, in most cases, of the order of 0.01 dex. The highest differences are found for FIES and CAFE ( $\sim 0.04$  dex), probably due to the low number of stars, but, in any case, below the obtained spread. The right-hand plots in Figure 4 show the differences against external Fe abundance (from PASTEL), which do not show any particular bias with  $[\text{Fe}/\text{H}]$ . These results show that there are no significant differences in the Fe abundance obtained with the different instruments, highlighting the quality of the homogenisation process.

We found that the mean differences among instruments shown in Fig. 4 are very similar when we use the Fe abundances obtained with the other radiative transfer codes. However, for stars with  $T_{\text{eff}}$  below 4000 K, there is a significant improvement in the consistency between instruments when we use TURBOSPECTRUM (see Figure A.4).

## 6. Element abundances

As for iron, we computed the mean abundances for each radiative transfer code for all stars, using the respective solar abundances. Given the fact that TURBOSPECTRUM solar abundances seem to be more compatible with literature (Sect. 4), and the Fe abundance of cool stars are also more consistent with

literature (Sect. 5) as well as more consistent among the different spectra of the same star (Figure A.4), we limit the analysis and plots of this section to the abundances produced for TURBOSPECTRUM. The results for all codes are available at the CDS.

Figure 5 shows the comparison between the abundances derived here, per spectrum, with respect to the values provided in the GBSv1 (Jofré et al. 2015). We see a general consistency in the comparison, with the largest offsets ( $>0.05$  dex) being for Ti, V and Mn. These are mainly due to large differences obtained for the three coolest stars (HIP 21421, HIP 98337 and HIP 14135), which also tend to be outliers in other elements (see Mg, Si, Fe II and Co). This is probably partly because in GBSv1, Jofré et al. (2015) performed a differential analysis and considered abundances from different radiative transfer codes. In their Fig. 7, Jofré et al. (2015) shows that there are few offsets between absolute and differential. Large differences with Ti I compared to Ti II for the three cool stars can also be due to the differential approach that was used in v1 compared to here, or to a different selection of lines which may affect more stars of a certain  $T_{\text{eff}}$ .

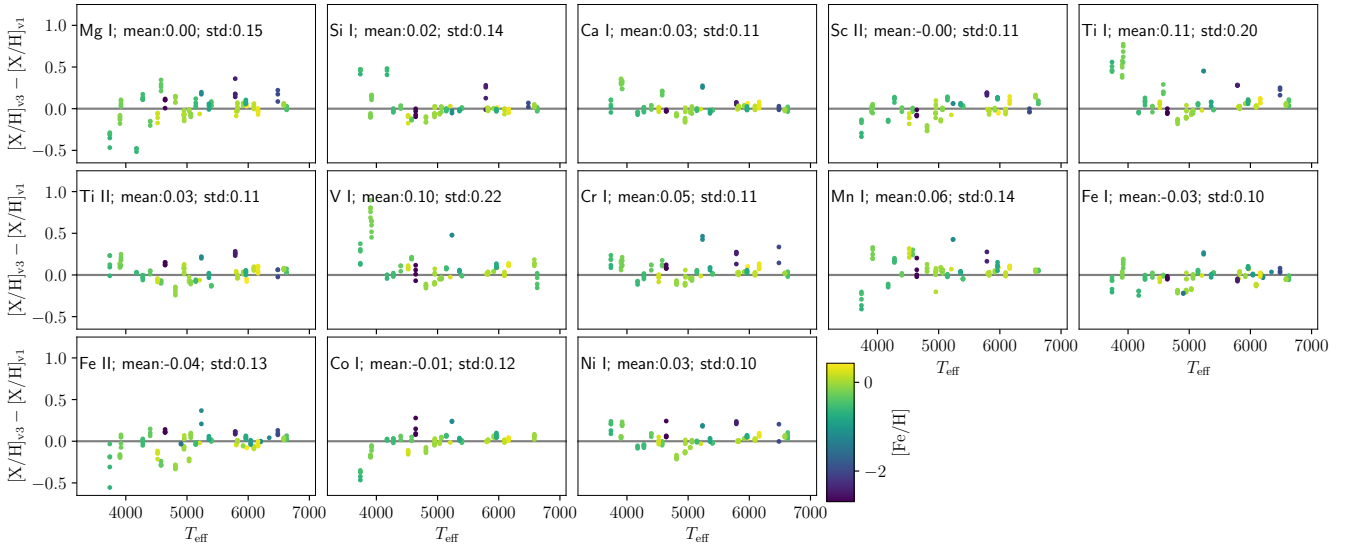
We have several stars in common with APOGEE DR17 and GALAH DR4, which are mainly stars that were included in GBSv1. We obtained an overall agreement between the elemental abundances of these surveys and our values (Figure A.3), particularly with GALAH, where all median offsets are smaller than 0.05 dex except for V and Co (both around 0.14 dex with larger scatter than the other elements). The persistent outlier in the GALAH comparison is psiPhe, for which GALAH values are much more metal-rich ( $-0.05$  dex) than our determination ( $-1.3$  dex). Interestingly, this star was found to be more metal-poor in the previous GALAH release ( $-1.2$  dex). We have more common stars with APOGEE, and the comparison shows larger offsets with respect to GALAH's one, sometimes above 0.1 dex, and a marked trend for Mg and Si towards the metal-poor stars. This trend is a known issue (see e.g. Soubiran et al. 2022) and seems to be also present in other large surveys, highlighting the difficulty in automated analysis and the importance of having well-characterised metal-poor benchmarks to understand the systematics of the surveys.

We also plot in Figure 6 the scatter in abundance obtained among the different spectra of the same star as a function of  $T_{\text{eff}}$ . We generally obtain very low dispersions with mean values around 0.02 dex, particularly in the region between 5000 and 6000 K. Mn, Ti II and Mg are the elements with the largest scatters, which is a direct consequence of the lower number of lines detected for them (1 for Mg, and between 3 and 4 for Mn and Ti II, see Figure A.5), and which makes the determinations of the abundances intrinsically more uncertain. The difference in the number of detected lines is also probably the cause of the changes we see as a function of  $T_{\text{eff}}$ , for instance the increase of the scatter for hot stars  $T_{\text{eff}} > 6000$  K (particularly in Ti, V, Cr, Co) and for cool stars  $T_{\text{eff}} \lesssim 5000$ –4500 K. Very metal-poor stars also tend to give larger scatters for the same reason, and they particularly stand out in the plot for Cr and Ni (see also the three most metal-poor stars in Fig. 5). The coolest star in the sample HIP 8837 (psiPhe) often appears as an extreme outlier in the plots, which is not surprising given its low effective temperature, as we already mentioned in the discussion of Sect. 5.

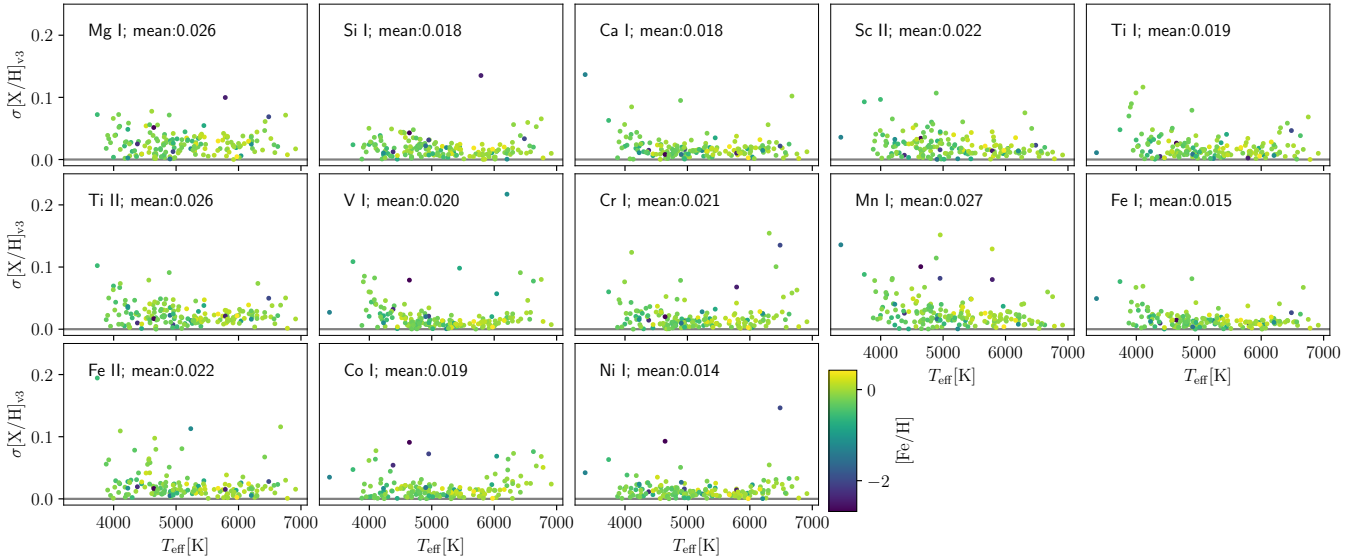
## 7. The error budget

In this section, we aim to analyse the different sources of uncertainty that can affect the abundance values, following the





**Fig. 5.** Comparison of the abundances in this study and the values from GBSv1. The mean difference and standard deviation are indicated.



**Fig. 6.** Scatter in abundance obtained among the different spectra of the same star as a function of  $T_{\text{eff}}$  coloured by the mean  $[\text{Fe}/\text{H}]$  abundance.

suggestion of Jofré et al. (2019). Primarily, each fit to a line has a quoted abundance computed by iSpec, which depicts the standard error of the line fit (Sect. 7.1). When quantifying the uncertainty associated with the abundance of a given element and star, one can also take into account the abundance dispersion given by the different lines (Sect. 7.2). This gives a measure of the differences that may arise from the characterisation of each line, which can be due to inaccuracies from the line list parameters (systematic uncertainties) or to the differences in the goodness of fits (random uncertainties). Finally, the abundances depend to a large extent on the values of the atmospheric parameters adopted. In this study, we have not determined  $T_{\text{eff}}$  and  $\log g$ , but we have assumed their values from the fundamental parameters determined independently from spectroscopy. These values have associated uncertainties, which we propagate to assess their contribution to the global uncertainty budget in Sect. 7.3. In the following subsections, we describe these three sources of uncertainty, and we describe the computation of final error values that we quote in Sect. 7.4.

### 7.1. Errors coming from the line fits

The uncertainties associated with individual spectral line fits in iSpec originate from the least-squares minimisation process employed to derive the best-fitting abundance. The individual line abundance error stems from the covariance matrix estimated during the fitting procedure using the Levenberg–Marquardt (LM) algorithm, implemented in MPFIT<sup>11</sup>.

For each spectral line, the code models the observed flux by fitting a synthetic spectrum generated using a given set of stellar parameters. The fitting process minimizes the residuals between the observed and synthetic fluxes by adjusting the abundance of the chemical element under consideration. The uncertainty of the line abundance is derived from the standard error of the fit,

<sup>11</sup> Originally, the routine (named LMFIT) was part of the MINPACK-1 FORTRAN package. It was converted to IDL by Craig B. Markwardt, translated to Python by Mark Rivers, and later optimized by Sergey Koposov (see <https://github.com/segasai/astrolibpy/tree/master/mpfit>).

which is computed as follows:

$$\delta_j = \sqrt{C_{jj}}, \quad (1)$$

where  $C_{jj}$  is the diagonal element of the covariance matrix corresponding to the abundance parameter of the given spectral line. This covariance matrix was approximated as the inverse of the Hessian matrix,  $\mathbf{H}$ , which in turn is estimated from the Jacobian matrix,  $\mathbf{J}$ :

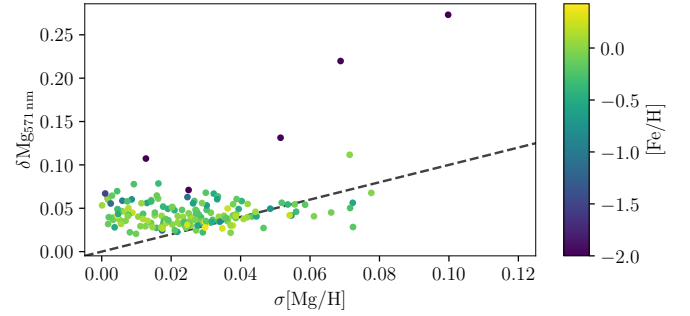
$$\mathbf{H} \approx \mathbf{J}^T \mathbf{J}. \quad (2)$$

However, these error estimates often appear overestimated when compared to the standard deviation of abundances derived from multiple lines of the same element. This discrepancy arises due to several contributing factors:

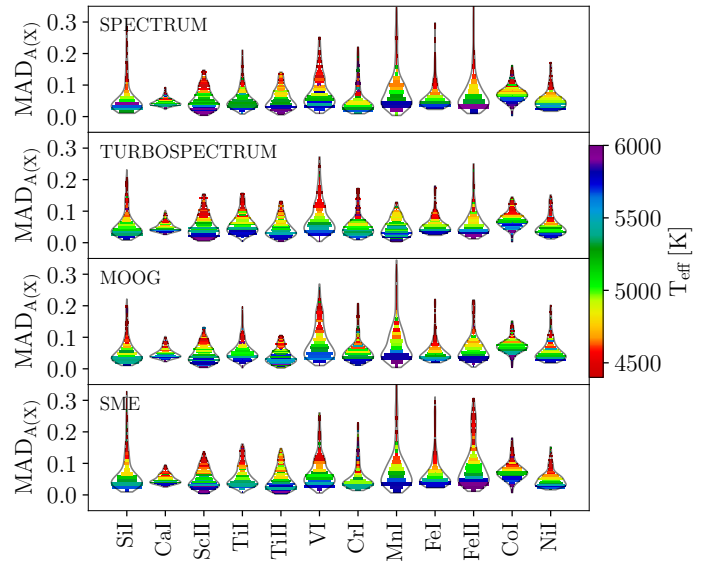
- Correlated noise: the LM algorithm assumes that the residuals are independent and normally distributed, but in practice, spectral noise can be correlated (e.g. after resampling, resolution degradation via convolution, continuum normalisation), leading to an overestimation of uncertainties.
- Numerical approximation in the covariance matrix: the covariance matrix is estimated numerically from finite differences, which can introduce small systematic deviations from the true error structure.
- Non-linearity effects: the covariance matrix is evaluated at a single point (the best-fit parameters), whereas the true function may exhibit significant curvature beyond this local region, leading to nonlinearity effects.

Consequently, while the LM algorithm provides a statistically sound method for estimating parameter uncertainties via the covariance matrix derived from the Jacobian, simplifying assumptions and systematic errors in the data may lead to overestimated uncertainties for an individual line. A more accurate estimation of uncertainties would require a full Bayesian treatment or Monte Carlo simulations, both of which are more computationally expensive. Alternatively, when possible, a more realistic assessment of abundance uncertainties for the mean abundances of an element is often obtained by computing the standard deviation of abundances derived from multiple lines of the same element, which inherently includes both random and systematic effects (see next subsection). Understanding these discrepancies is crucial when comparing the uncertainty of a single line fit to the overall abundance precision obtained from multiple spectral lines.

In the case of only being able to measure one single line of a given element (like Mg in the case of this study, where we only use the 571 nm line), the dispersion of the abundances derived from different spectra provides a measurement of at least the random uncertainties affecting the line fit. We plot in Figure 7 the distribution of the quoted error to the Mg line fit ( $\delta\text{Mg}_{571\text{ nm}}$ ), compared to the scatter in the Mg abundance obtained for stars with more than one spectrum ( $\sigma[\text{Mg}/\text{H}]$ ). A general overestimation is seen with respect to the obtained scatters, particularly when scatters are very small (the errors seem to saturate at 0.05 dex) or for metal-poor stars (probably because the lines are particularly weak compared to the quoted flux noise). However, for a given spectrum and element, the errors coming from the line fits are comparable and correlate very well with the depth of the lines. Thus, we consider that it is safe to use them as an internal relative scale of the quality of the line fit for a given spectrum (see Sect. 7.4).



**Fig. 7.** Mean of the quoted error of the 571 nm Mg line abundance (values from TURBOSPECTRUM) on the  $y$ -axis for stars with more than one spectrum, compared to the scatter in the derived Mg abundances among the different spectra on the  $x$ -axis, colour-coded by the mean Fe abundance derived in this study. The dashed line depicts the 1:1 relation.



**Fig. 8.** Violin plot representing the kernel density estimation of the distribution of the line scatters (computed as the MAD) per element and radiative transfer code. To visualise the dependency of temperature, we overplot corresponding histograms, coloured-coded by the mean  $T_{\text{eff}}$  in each bin.

## 7.2. Errors coming from the line-to-line dispersion

We plot in Figure 8 the distribution of the mean absolute deviation (MAD) per spectrum and element in the form of violin plots. This is a direct measurement of the coherence of the different selected lines for a given element. A vertical histogram is overplotted, showing the dependence in  $T_{\text{eff}}$  of the obtained MAD. Mg is not plotted because only one line is available.

The obtained distributions tend to peak at around 0.05 dex and then show more or less prominent tails towards larger dispersions. V, Mn, Co and Fe II are the elements for which the median of the distributions are larger (around 0.06 dex) for all four codes.

There is a clear dependence of the obtained MAD on the  $T_{\text{eff}}$  of the star, with cool stars  $T_{\text{eff}} < 5000$  K always giving the largest dispersions. The coolest stars ( $T_{\text{eff}} < 4500$  K) shape long tails in the distribution of certain elements (e.g. Fe II, Mn, V, Si), which can reach values of more than 0.1 dex. As discussed in previous sections, this is somewhat expected due to the intrinsically larger

**Table 1.** Sample of stars and spectra used to test the effects of the uncertainties on the atmospheric parameters.

| Type          | Star       | $T_{\text{eff}}$ | $\log g$        | $[\text{Fe}/\text{H}]_{\text{PASTEL}}$ | $[\text{Fe}/\text{H}]_{\text{here}}$ | $\sigma_{\text{MC}}[\text{Fe}/\text{H}]$ | Instr. | S/N  |
|---------------|------------|------------------|-----------------|----------------------------------------|--------------------------------------|------------------------------------------|--------|------|
| F dwarfs      | HIP 67927  | $6161 \pm 18$    | $3.82 \pm 0.05$ | $0.25 \pm 0.01$                        | $0.32 \pm 0.04$                      | 0.015                                    | NARVAL | 1073 |
|               | HIP 32851  | $6628 \pm 89$    | $4.19 \pm 0.04$ | $-0.39 \pm 0.03$                       | $-0.45 \pm 0.03$                     | 0.048                                    | HARPS  | 1020 |
| G dwarfs      | HIP 79672  | $5824 \pm 30$    | $4.42 \pm 0.01$ | $0.03 \pm 0.01$                        | $0.06 \pm 0.01$                      | 0.020                                    | ELODIE | 236  |
|               | HIP 76976  | $5788 \pm 45$    | $3.75 \pm 0.09$ | $-2.48 \pm 0.03$                       | $-2.42 \pm 0.04$                     | 0.042                                    | HARPS  | 419  |
| K dwarf       | HIP 104214 | $4398 \pm 34$    | $4.63 \pm 0.01$ | $-0.13 \pm 0.03$                       | $-0.33 \pm 0.14$                     | 0.017                                    | ELODIE | 330  |
| Subgiant      | HIP 95362  | $5268 \pm 18$    | $3.77 \pm 0.07$ | $-0.11 \pm 0.02$                       | $-0.17 \pm 0.02$                     | 0.011                                    | ELODIE | 226  |
| Clump giants  | HIP 37826  | $4810 \pm 14$    | $2.55 \pm 0.03$ | $0.02 \pm 0.03$                        | $-0.07 \pm 0.03$                     | 0.010                                    | ELODIE | 373  |
|               | HIP 92167  | $4902 \pm 30$    | $2.53 \pm 0.02$ | $-1.45 \pm 0.03$                       | $-1.51 \pm 0.04$                     | 0.028                                    | NARVAL | 181  |
| Bright giants | HIP 69673  | $4277 \pm 23$    | $1.58 \pm 0.07$ | $-0.55 \pm 0.01$                       | $-0.55 \pm 0.04$                     | 0.006                                    | HARPS  | 382  |
|               | HIP 68594  | $4642 \pm 35$    | $1.32 \pm 0.03$ | $-2.67 \pm 0.02$                       | $-2.68 \pm 0.05$                     | 0.043                                    | HARPS  | 318  |

**Notes.**  $T_{\text{eff}}$  and  $\log g$  come from fundamental values (Soubiran et al. 2024), and we also list the  $[\text{Fe}/\text{H}]$  quoted in PASTEL and the one obtained here for that star from TURBOSPECTRUM. We indicate the dispersion in Fe abundance obtained from the 30 random realisations ( $\sigma_{\text{MC}}[\text{Fe}/\text{H}]$ ).

random uncertainties implied in the abundance measurement of cool stars. On the contrary, solar-type and hotter stars seem to provide the most consistent abundances among the different lines.

### 7.3. Errors coming from the uncertainties in the atmospheric parameters

It is well known that the choice of the atmospheric parameters has a large impact on the abundance determination. The present analysis assumes that each star's effective temperature and surface gravity are known. These values were obtained independently from the spectra using fundamental methods, i.e. from the limb darkened angular diameter ( $\theta_{\text{LD}}$ ) and bolometric flux ( $F_{\text{bol}}$ ) to determine the effective temperature, and with the parallax ( $\varpi$ ), angular diameter and mass ( $M$ ) to determine the surface gravity. This section aims to assess the impact of fixing the  $T_{\text{eff}}$  and  $\log g$  to their fundamental values in the obtained chemical abundances.

Due to the complexity of the interrelations between each quantity and the different fitting procedures used to determine them (see a schematic view in Fig. A.6), the impact of the uncertainties on the chemical abundance values becomes intractable analytically. However, it can be estimated via Monte Carlo propagation, which does not require knowledge of the correlation matrix. We thus generated random realisations of  $T_{\text{eff}}$  and  $\log g$  while taking into account the form of the original equations (see sections 3 and 4 from Soubiran et al. 2024), which make the two atmospheric parameters a function of the following variables:

$$T_{\text{eff}}(\theta_{\text{LD}}, F_{\text{bol}}); \log g(\theta_{\text{LD}}, \varpi, M) \quad (3)$$

At the same time, the mass is not an independent variable because it has been obtained with the SPInS code (Lebreton & Reese 2020) in the  $T_{\text{eff}}$ -luminosity plane, assuming the  $[\text{Fe}/\text{H}]$  from the PASTEL catalogue. Two resulting masses were obtained using two grids of stellar evolution models, and the mean was taken to compute the fundamental  $\log g$ . So, in the end, mass is a function of  $M(\theta_{\text{LD}}, F_{\text{bol}}, \varpi, [\text{Fe}/\text{H}]_{\text{PASTEL}})$ . Additionally, we point out that  $F_{\text{bol}}$  also depends on the parallax (through the extinction) and on the  $[\text{Fe}/\text{H}]_{\text{PASTEL}}$ . However, its dependence on the photometry and the used models is much more important for these stars, particularly because their parallax uncertainties are very small, and because the input  $[\text{Fe}/\text{H}]$  used has a small impact in the SED fitting (see discussion in sect 3.4

of Soubiran et al. 2024). So, for this study, we consider  $F_{\text{bol}}$  an independent variable.

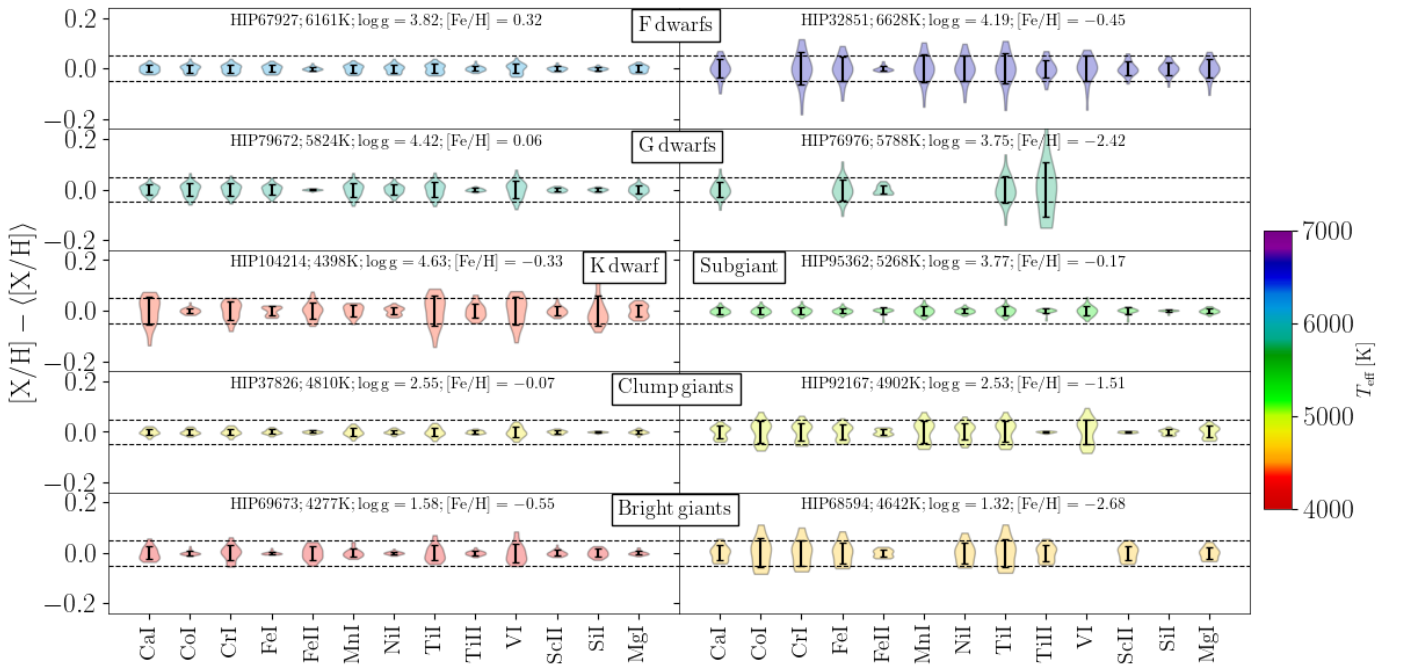
We thus generated a set of random realisations in the independent parameters ( $\theta_{\text{LD}}, F_{\text{bol}}, \varpi, [\text{Fe}/\text{H}]_{\text{PASTEL}}$ ) assuming gaussian uncertainties, and we recomputed  $T_{\text{eff}}$  and  $\log g$  with the same procedure as done in Soubiran et al. (2024). Using this sampling on the atmospheric parameters, we re-did the chemical abundance computation with iSpec.

Due to the high computational cost of running the full pipeline described in Sect. 3, we have limited the random sampling to 30 pairs of  $T_{\text{eff}}$  and  $\log g$ , generated for ten representative stars of the sample. These stars are listed in Table 1, and have been chosen to provide good coverage of the  $T_{\text{eff}}$ ,  $\log g$  and  $[\text{Fe}/\text{H}]$  range of the full sample. Also for computational reasons, we have limited this experiment to one radiative transfer code, TURBOSPECTRUM.

The resulting abundance distributions for each star and element (scaled to the mean element abundance for visualisation purposes) are shown as violin plots in Fig. 9. The obtained scatter for each star and element is indicated with an error bar, and the scatter value for  $[\text{Fe}/\text{H}]$  ( $\sigma[\text{Fe}/\text{H}]_{\text{MC}}$ ) is listed in Table 1. In general, the obtained distributions show small scatters of the order of 0.01–0.03 dex, particularly for cases with solar  $[\text{Fe}/\text{H}]$ . The metal-poor stars HIP 76976, HIP 92167, and HIP 68594 show the largest dispersions, highlighting the difficulty of determining their abundance due to the smaller number of available lines. The largest dispersions are found for HIP 32851, a slightly metal-poor F dwarf, but which has the largest uncertainty in  $T_{\text{eff}}$ , resulting in a large range in the input  $T_{\text{eff}}$  values (6344–6785 K), thus impacting the abundance range. Remarkably for this star (and also for the other F and G dwarfs), the dispersion obtained of Fe II is much smaller than for the rest of the elements. This is because of the high sensitivity of the Fe I lines to  $T_{\text{eff}}$  (see discussion of sec 5.3 in Amarsi et al. 2022).

Overall, we consider that for the GBS sample, the uncertainties coming from the atmospheric parameters are smaller than the dispersions found among the different lines (see the uncertainty in the  $[\text{Fe}/\text{H}]$  value of the spectrum listed in the table, and discussion in Sect. 7.2). This is a consequence of the high precision in the  $T_{\text{eff}}$  and  $\log g$  determinations. Exceptions to this may be the cases highlighted above, particularly when the quoted uncertainties in effective temperature are large ( $\sim 80$  K). For the GBSv3 sample, the  $T_{\text{eff}}$  uncertainties range from 5 to 183 K with a median value of 44 K and 15% of the stars having an error larger than 80 K.





**Fig. 9.** Distributions of the elemental abundances obtained from perturbing the atmospheric parameters within their uncertainties (see Table 1), coloured by the  $T_{\text{eff}}$  of the star. The standard deviation for each violin plot is shown as an error bar centred in zero (the mean abundance), and the dashed lines depict the  $\pm 0.05$  dex level.

#### 7.4. Adopted abundances and uncertainties per spectrum and star

We compute, for each spectrum and chemical species, the weighted mean of the absolute abundances for all selected lines, using the inverse of the errors coming from the line fits as weights. We give as an error of this mean abundance the unbiased weighted standard deviation, computed as

$$\sigma^2 = \frac{\sum w_i (x_i - \mu)^2}{\sum w_i - \frac{\sum w_i^2}{\sum w_i}}, \quad (4)$$

where  $w_i$  are the weights (inverse of the error of each line fit  $1/\delta_i$ ),  $x_i$  are the line abundances, and  $\mu$  is the weighted mean. This strategy allows for taking into account the goodness of the line fits to compute the mean abundance without overestimating the error budget (see discussion in Sect. 7.1). For the spectra where only one line of the element is measured, the error coming from the line fit is the only uncertainty estimation that we have, even though we warn the reader that it might be difficult to interpret and compare among the different stars. This is the case of Mg, but particularly for metal-poor stars it can also be the case for other elements for which we only managed to measure one line. For each spectrum and element, we indicate the number of lines measured in the final table. Finally, the  $[X/H]$  values are computed using the solar abundances obtained respectively for each code, and the uncertainties in  $[X/H]$  are computed as the quadratic sum of the previously mentioned error and the solar abundance error.

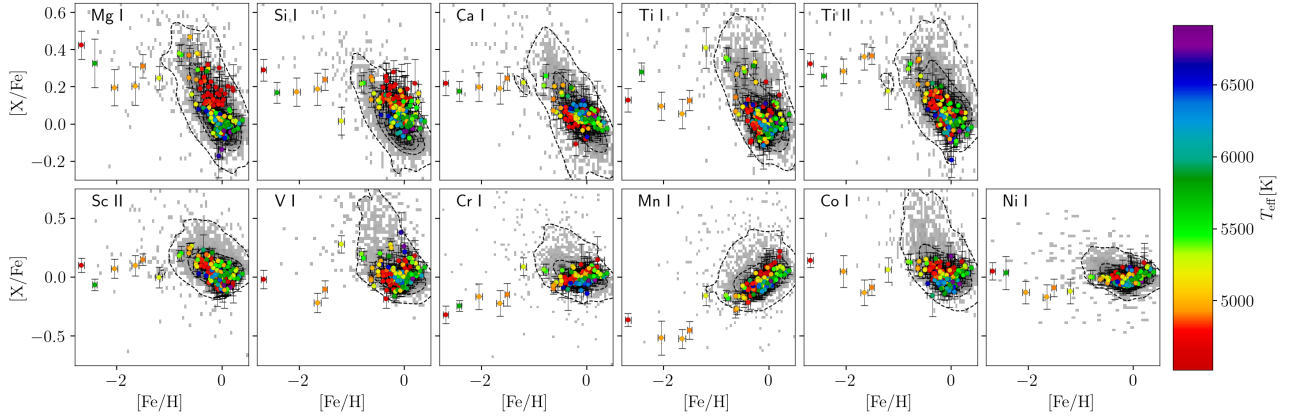
For stars with more than one observed spectrum, we combine the solar-scaled abundances of the different spectra using the product of experts method (Hinton 1999). In other words, we assume that the different measurements of  $[X/H]$  follow a Gaussian distribution whose sigma is given by the uncertainty. Then, we perform the product of the different distributions

and we take the mean of the resulting distribution as the final abundance per star, and the standard deviation of the distribution as the uncertainty of the abundance.

## 8. Final abundances

We explore in Figure 10 the  $[X/Fe]$  ratios vs  $[Fe/H]$  for all the elements we have analysed in this paper. We only plot here the 192 GBS with direct measurement of the angular diameter (Soubiran et al. 2024). To make the plot clearer, we exclude the stars with  $T_{\text{eff}} \lesssim 4500$  K, for which the abundances tend to be less reliable, and the few stars which have very large uncertainties in  $[X/Fe]$  ( $> 0.2$ ). We observe similar general behaviours of all the elements when using the other three radiative transfer codes. We overplot the Gaia-ESO survey stars (Hourihane et al. 2023) (type: “GE\_MW”) observed with the UVES setup U580 using a 2D histogram and contour lines. We have transformed their absolute abundances to solar-scaled using the solar abundances from Grevesse et al. (2007).

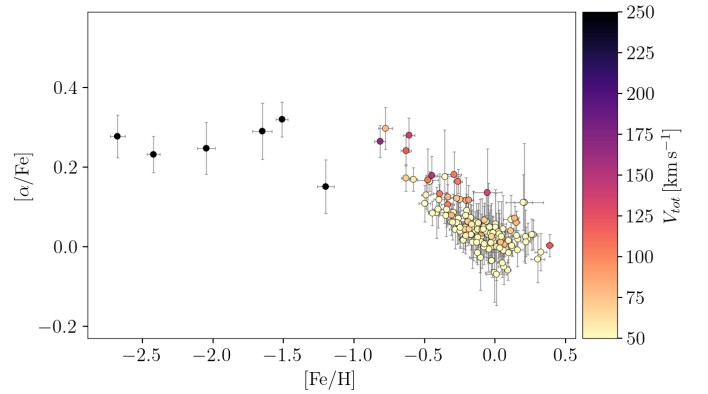
For all the  $\alpha$  elements (top row of Figure 10), we observe a decreasing trend with Fe abundance for the range  $[Fe/H] > -0.1$  and a plateau for metal-poor stars. This trend is consistent with what has been historically seen with solar neighbourhood samples in the literature (Fuhrmann 1998), and it is also seen for the Gaia-ESO stars. It is produced by the progressive enrichment in iron of the interstellar medium due to thermonuclear supernovae and is well reproduced by Galactic chemical evolution models. The expected  $[\alpha/Fe]$  vs  $[Fe/H]$  trend appears clearer for the GBS when we use Ca or Ti II. This is probably because these two elements are those for which the line-to-line scatters are the smallest and have less dependence on effective temperature (see Figure 8). For Si and Ti I, the tails of the line-to-line scatter distributions of the GBS spread towards larger values (particularly for Si) and are populated by moderately cool stars. If we do a restrictive cut in effective temperature ( $> 5000$  K), the trends in



**Fig. 10.**  $[X/Fe]$  vs  $[Fe/H]$  for all the elements analysed in this paper, coloured by effective temperature. The UVES data from the Gaia-ESO survey are overplotted with a 2D histogram and contour lines. Notice the change of the Y scale between the top and bottom panels.

these two elements are more clear, but we lose a large fraction of the sample. In some cases, we see a mild dependence on  $T_{\text{eff}}$ , particularly in Mg and Si ( $>0.1$  dex amplitude), which is probably due to the line strength dependence on the atmospheric parameters. This is why, for Galactic archaeology studies, it is important to take into account these types of effects, or use stars of the same spectral type. Alternatively, this could be mitigated with a differential analysis. For the Gaia-ESO stars, the increasing  $\alpha$  trend seems to extend to larger  $\alpha$  values than for the GBS, particularly for Ca, Si and Ti. This could be due to the particular selection of stars.

Concerning the Fe-peak elements' dependence with  $[Fe/H]$  (bottom row of Figure 10), we highlight several behaviours. Sc shows a decreasing tendency starting at  $-0.5$  dex towards solar values, reproduced by the Gaia-ESO stars, similar to what is seen for  $\alpha$  elements, and has been discussed in the literature (e.g. Battistini & Bensby 2015). We observe increasing trends of  $[Mn/Fe]$  (which is also visible for the Gaia-ESO stars) and  $[Cr/Fe]$ , which can partly be due to non-LTE (NLTE) effects. Indeed, the NLTE corrections for Mn can go up to 0.5 dex at  $[Fe/H] \sim -2$  (Bergemann & Gehren 2008), and around 0.3 dex at  $[Fe/H] \sim -2$  for Cr (Bergemann & Cescutti 2010). It remains as future work to perform a full 3D NLTE analysis on the GBSv3 spectra to quantify these trends, which still should be expected to exist due to the differences of supernova yields (Andrews et al. 2017; Kobayashi et al. 2020). A tail of high  $[V/Fe]$ , up to more than 0.7 dex (with large scatter) towards moderately subsolar  $[Fe/H]$  is seen in the Gaia-ESO stars, which we do not reproduce with the GBS. Our V values have a large scatter at around solar  $[Fe/H]$ , which can blur the trend, if it is there. At metallicities smaller than  $-1$ , other studies find  $[V/Fe]$  close to Solar: e.g. (Battistini & Bensby 2015) their fig. 6, and see also Ishigaki et al. (2013), which would be in agreement with our three metal-poor stars. For  $[Co/Fe]$ , Gaia-ESO stars also show a large decreasing trend towards moderately subsolar  $[Fe/H]$ , which we also see with the GBS but at smaller amplitude. This smaller amplitude seems to be in agreement with what other works find (Battistini & Bensby 2015,  $\sim 0.2$  dex at  $[Fe/H] = -1$  dex). They also find a plateau of  $[Co/Fe] \sim 0.2$  at lower metallicities in their fig. 6. Because of the scatter and the few GBS stars at  $[Fe/H] < -1$ , it is difficult to say if our values are compatible with such a plateau. Finally, a flat trend followed by a slightly increasing trend at super-solar metallicity is seen in  $[Ni/Fe]$ , not evident for the Gaia-ESO stars, but that has been reported in the literature (e.g. Mishenina et al. 2006).



**Fig. 11.**  $[\alpha/Fe]$  vs  $[Fe/H]$  coloured by total velocity and with similar restrictions as in Fig. 10, i.e. not-flagged stars with  $T_{\text{eff}} > 4500$  K and uncertainty on  $[\alpha/Fe]$  lower than 0.2 dex.

In order to illustrate how the GBS sample is representative of the local stellar populations, we show in Fig. 11 the distribution of  $[\alpha/Fe]$  vs  $[Fe/H]$  coloured by the total galactic velocity ( $V_{\text{tot}} = \sqrt{U^2 + V^2 + W^2}$ ). The global  $\alpha$  abundance has been estimated by computing a weighted mean of Si, Ca and Ti (Ti being the weighted mean of Ti I and Ti II). We have not considered Mg because it is not available for all stars, and it has larger uncertainties. Galactic velocities (U, V, W) have been computed by combining positions, parallaxes and proper motions from Gaia DR3 (Gaia Collaboration 2023b) (or Hipparcos for the brightest stars) with radial velocities from our spectra. Figure 11 clearly shows that the 6 stars with  $[Fe/H] < -1.0$  have kinematics typical of the halo (total velocity  $>300 \text{ km s}^{-1}$ ). The  $[\alpha/Fe]$  distribution of these six stars is rather flat with a small dispersion around a mean value of  $+0.25$  dex, consistent with the well-known plateau observed in many spectroscopic investigations of the halo (e.g. Fuhrmann (1998)). This demonstrates the reliability of our abundances. The rest of the sample extends from about  $-0.80$  to  $+0.40$  in  $[Fe/H]$  with a decreasing trend of  $[\alpha/Fe]$  in two parallel sequences. The upper sequence, with higher  $\alpha$  abundances, shows, in general also hotter kinematics characteristic of the thick disc, while the bulk of the sample in the lower alpha sequence shows smaller velocities typical of the thin disc. There are a few exceptions to the general behaviour of the three populations. For instance, the most metal-rich star (HIP 95447=31 Aql,  $[Fe/H]=+0.39$ ,  $[\alpha/Fe]=+0.01$ ) has a total

velocity of  $121 \text{ km s}^{-1}$  which would make it a good thick disc candidate. However, the large velocity is only due to its radial component  $U$ , which, combined with the super-metallicity with no  $\alpha$ -enhancement, would suggest that the star was formed in the inner disc and migrated outward in the solar neighbourhood afterwards. Another notable case is HIP 25993, which has thin disc kinematics together with a significant  $\alpha$  enhancement  $[\alpha/\text{Fe}] = +0.15$ , but also large abundance uncertainties. The high  $\alpha$  and low  $\alpha$  subpopulations in the disc and their correlation with kinematics have been commonly reported in the past from various surveys and high-resolution samples (see, for instance Soubiran & Girard (2005); Bensby et al. (2014); Gaia Collaboration (2023a); Imig et al. (2023)). Known chemodynamical features are well reproduced by our study and therefore demonstrate that the GBSv3 sample is representative of the stellar populations of the solar neighbourhood.

## 9. Conclusions

The third version of the GBS recently released by Soubiran et al. (2024) provides a step forward in the characterisation of reference stars. With respect to the previous versions, the GBSv3 sample has improved statistics and better homogeneity and accuracy in the atmospheric parameters  $T_{\text{eff}}$  and  $\log g$ . The present study, the eighth paper of a series, provides a complete spectral library containing 522 normalised spectra with excellent S/N covering the wavelength range of 480–680 nm, for the sample of 192 GBSv3 and for ten additional stars, including the Sun, that were part of the previous GBS versions. Using the fundamental atmospheric parameters, we performed a chemical analysis of 13 Fe-peak and  $\alpha$ -capture elements, with the goal of providing to the community reference values for the chemical abundances of these stars.

Our analysis was done with the MARCS model atmospheres (Gustafsson et al. 2008) and the sixth version of the Gaia-ESO survey line list (Heiter et al. 2021). We obtained 1D LTE line-by-line abundances using spectral synthesis fitting with four radiative transfer codes widely used in the literature. We performed a line selection based on the flags provided with the line list and compared the line abundances obtained from the Sun with literature values.

We discussed how our results for the abundances compare with the literature (Fig. 3, 5) and showed that, in general, we have small offsets ( $<0.05$  dex on average). We did find a larger scatter and a different behaviour for cool stars ( $<4500$  K) with the four different codes. While SPECTRUM, MOOG, and SME tend to overestimate the Fe abundance for these stars with a marked trend with  $T_{\text{eff}}$ , TURBOSPECTRUM seems to give abundances more consistent with the literature and no trends, possibly due to the inclusion of a molecular linelist. This is why in the final table we provide the TURBOSPECTRUM values as recommended. The elements that show the largest offsets with the literature are Ti, V, and Mn, and this is in general caused by a large difference obtained for the coolest stars in our sample ( $<4000$  K).

We analysed the different sources of error that affect our abundances. We find that uncertainties due to the errors in atmospheric parameters are, in general, smaller (of the order of 0.01–0.03 dex) than the line-to-line dispersions, which is a consequence of the high precision in the fundamental  $T_{\text{eff}}$  and  $\log g$ . Exceptions to this might be stars with the largest uncertainties in  $T_{\text{eff}}$ . We observed that the errors quoted for each line fit generally seem overestimated and difficult to compare among the different stars. However, since they correlate well with the line depth, we

used them as weights to compute the mean standard deviation among lines in order to quote an error per element and spectrum. For the analysed elements, we obtained typical uncertainties of 0.05 dex, and there is a clear dependence of the uncertainty on  $T_{\text{eff}}$ , with cool stars always giving larger values. Some of these uncertainties can reach values of more than 0.1 dex in certain elements (particularly Fe II, Mn, V, and Si).

We explored the  $[\text{X}/\text{Fe}]$  trends of the GBSv3 in comparison with the Milky Way stars from the Gaia-ESO survey. In general, we observed a good overlap in the distributions, with some differences that can be explained by the small selection of stars in the GBS compared to the field stars. Notably, we reproduced the expected  $[\alpha/\text{Fe}]$  versus  $[\text{Fe}/\text{H}]$  trends in the GBS sample for all  $\alpha$  elements, with Ca and Ti II being the elements with the clearest trends. Sc showed a similar trend to the  $\alpha$  elements: a decreasing tendency starting at -0.5 dex towards solar values, reproduced by the Gaia-ESO stars and other literature works. We noted steep trends in  $[\text{Mn}/\text{Fe}]$  (similar to the Gaia-ESO stars) and  $[\text{Cr}/\text{Fe}]$  for metal-poor stars, and we speculate that NLTE effects can be the cause. V, Co, and Ni showed increasing trends for super-solar metallicity stars, but of much smaller amplitudes.

We also explored the  $[\alpha/\text{Fe}]$  versus  $[\text{Fe}/\text{H}]$  distribution coupled with the total galactic velocity. We reproduced the expected behaviour of a local stellar population, with metal-poor  $\alpha$ -enhanced stars having kinematics typical of the halo and a decreasing trend of  $[\alpha/\text{Fe}]$  at  $[\text{Fe}/\text{H}]$  between -0.8 and +0.4, in two sequences with a correlation with kinematics that are typical of the thick and thin disks.

The resulting abundances per star, spectrum, and line are available at the CDS. As future steps, we are going to provide abundances of light and neutron-capture elements, and we plan to do an NLTE analysis, which will complement the set of LTE abundances provided in this work.

## Data availability

The data obtained in this paper are available at the CDS via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/705/A167>. We publish four tables resulting from our abundance analysis:

- The first table contains the final abundances per star for each of the four radiative transfer codes, computed as explained in Sect. 7.4. We provide a column with a flag indicating if the star has an indirect measurement of the angular diameter (not recommended by Soubiran et al. 2024), or if it is a double-line spectroscopic binary (see discussion in Sect. 3.1) and thus also not recommended as a spectroscopic reference.
- A second table details the mean abundances per spectrum for each radiative transfer code.
- We also make available a table with the absolute abundances and equivalent widths measured for all lines selected in the different spectra and radiative transfer codes, with the references for the gf-values.
- Finally, we provide a table with the recommended abundances for each star, the mean  $\alpha$  abundances and radial velocities measured from our collection of spectra.

The first table (per star) and the normalised spectral library at the original resolution of each instrument, and degraded to the common resolution (42 000), are also available on the website <https://www.blancocuaresma.com/s/benchmarkstars>.

**Acknowledgements.** L.C. and C.S. acknowledge financial support from Observatoire Aquitain des Sciences de l'Univers (OASU). P.J., C.A.G., I.H., S.E.



and S.V. acknowledge support from FONDECYT Regular 1231057. N.L. acknowledges financial support from “Programme National de Physique Stellaire” (PNPS) and from the “Programme National Cosmology et Galaxies (PNCG)” of CNRS/INSU, France. A.R.A. acknowledges support from DICYT through grant 062319RA and from ANID through FONDECYT Regular grant No. 1230731. C.A.G. acknowledges support from FONDECYT Iniciación 11230741. U.H. acknowledges support from the Swedish National Space Agency (SNSA/Rymdstyrelsen). This work was (partially) supported by the Spanish MICIN/AEI/10.13039/501100011033 and by “ERDF A way of making Europe” by the “European Union” through grant PID2021-122842OB-C21; the Institute of Cosmos Sciences University of Barcelona (ICCUB, Unidad de Excelencia ‘María de Maeztu’) through grant CEX2019-000918-M; the Horizon Europe HORIZON-CL4-2023-SPACE-01-71 SPACIOUS project funded under Grant Agreement no. 101135205; and the Spanish MICIN/AEI/10.13039/501100011033 through grant RED2022-134612-T. The preparation of this work has made extensive use of Topcat (Taylor 2005), of the Simbad and VizieR databases at CDS, Strasbourg, France, and of NASA’s Astrophysics Data System Bibliographic Services. We warmly thank Carme Jordi for her contribution to the observations of the GBS as well as Thomas Nordlander and Clare Worley for the fruitful discussions.

## References

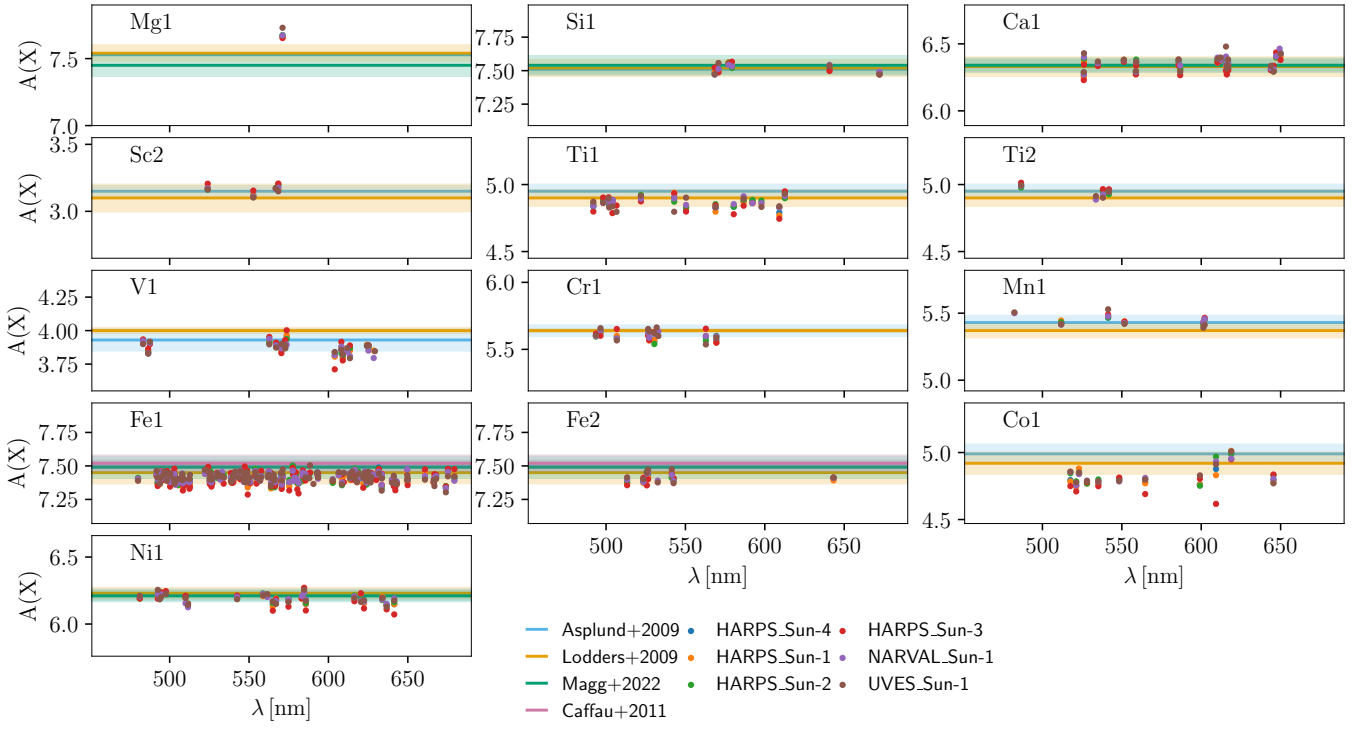
- Aceituno, J., Sánchez, S. F., Grupp, F., et al. 2013, *A&A*, **552**, A31
- Adibekyan, V., Sousa, S. G., Santos, N. C., et al. 2020, *A&A*, **642**, A182
- Alvarez, R., & Plez, B. 1998, *A&A*, **330**, 1109
- Amarsi, A. M., Liljegren, S., & Nissen, P. E. 2022, *A&A*, **668**, A68
- Andrews, B. H., Weinberg, D. H., Schönrich, R., & Johnson, J. A. 2017, *ApJ*, **835**, 224
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Aurière, M. 2003, in *EAS Publications Series*, 9, EAS Publications Series, eds. J. Arnaud, & N. Meunier, 105
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, *A&AS*, **119**, 373
- Battistini, C., & Bensby, T. 2015, *A&A*, **577**, A9
- Bensby, T., Feltzing, S., & Oey, M. S. 2014, *A&A*, **562**, A71
- Bergemann, M., & Gehren, T. 2008, *A&A*, **492**, 823
- Bergemann, M., & Cescutti, G. 2010, *A&A*, **522**, A9
- Blanco-Cuaresma, S. 2019, *MNRAS*, **486**, 2075
- Blanco-Cuaresma, S., & Fraix-Burnet, D. 2018, *A&A*, **618**, A65
- Blanco-Cuaresma, S., Soubiran, C., Heiter, U., & Jofré, P. 2014, *A&A*, **569**, A111
- Caffau, E., Ludwig, H. G., Steffen, M., Freytag, B., & Bonifacio, P. 2011, *Sol. Phys.*, **268**, 255
- Casamiquela, L., Tarricq, Y., Soubiran, C., et al. 2020, *A&A*, **635**, A8
- Creevey, O. L., Cassisi, S., Thévenin, F., Salaris, M., & Pietrinferni, A. 2024, *A&A*, **689**, A243
- Dekker, H., D’Odorico, S., Kaufer, A., Delabre, B., & Kotzlowski, H. 2000, *SPIE Conf. Ser.*, **4008**, 534
- Donati, J. F., Semel, M., Carter, B. D., Rees, D. E., & Collier Cameron, A. 1997, *MNRAS*, **291**, 658
- Donati, J. F., Catala, C., Landstreet, J. D., & Petit, P. 2006, in *Astronomical Society of the Pacific Conference Series*, 358, Solar Polarization 4, eds. R. Casini, & B. W. Lites, 362
- Fuhrmann, K. 1998, *A&A*, **338**, 161
- Gaia Collaboration (Recio-Blanco, A., et al.) 2023a, *A&A*, **674**, A38
- Gaia Collaboration (Vallenari, A., et al.) 2023b, *A&A*, **674**, A1
- Gerber, J. M., Magg, E., Plez, B., et al. 2023, *A&A*, **669**, A43
- Gray, R. O., & Corbally, C. J. 1994, *AJ*, **107**, 742
- 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
- Hawkins, K., Jofré, P., Heiter, U., et al. 2016, *A&A*, **592**, A70
- Hegedűs, V., Mészáros, S., Jofré, P., et al. 2023, *A&A*, **670**, A107
- Heiter, U., Jofré, P., Gustafsson, B., et al. 2015, *A&A*, **582**, A49
- Heiter, U., Lind, K., Bergemann, M., et al. 2021, *A&A*, **645**, A106
- Hinton, G. E. 1999, in *Ninth International Conference on Artificial Neural Networks*, 1999: ICANN ’99 (Conf. Publ. No. 470), 1, IET, 1
- Ho, A. Y. Q., Ness, M. K., Hogg, D. W., et al. 2017, *ApJ*, **836**, 5
- Hourihane, H., François, P., Worley, C. C., et al. 2023, *A&A*, **676**, A129
- Imig, J., Price, C., Holtzman, J. A., et al. 2023, *ApJ*, **954**, 124
- Ishigaki, M. N., Aoki, W., & Chiba, M. 2013, *ApJ*, **771**, 67
- Jofré, P., Heiter, U., Soubiran, C., et al. 2014, *A&A*, **564**, A133
- Jofré, P., Heiter, U., Soubiran, C., et al. 2015, *A&A*, **582**, A81
- Jofré, P., Heiter, U., Worley, C. C., et al. 2017, *A&A*, **601**, A38
- Jofré, P., Heiter, U., Tucci Maia, M., et al. 2018, *RNAAS*, **2**, 152
- Jofré, P., Heiter, U., & Soubiran, C. 2019, *ARA&A*, **57**, 571
- Jönsson, H., Allende Prieto, C., Holtzman, J. A., et al. 2018, *AJ*, **156**, 126
- Kaufer, A., Stahl, O., Tubbesing, S., et al. 1999, *The Messenger*, **95**, 8
- Kobayashi, C., Karakas, A. I., & Lugaro, M. 2020, *ApJ*, **900**, 179
- Lebreton, Y., & Reese, D. R. 2020, *A&A*, **642**, A88
- Lillo-Box, J., Aceituno, J., Pedraz, S., et al. 2020, *MNRAS*, **491**, 4496
- Lodders, K., Palme, H., & Gail, H. P. 2009, *Landolt & Bornstein*, **4B**, 712
- Magg, E., Bergemann, M., Serenelli, A., et al. 2022, *A&A*, **661**, A140
- Masseron, T., Merle, T., & Hawkins, K. 2016, BACCHUS: Brussels Automatic Code for Characterizing High accuracy Spectra, Astrophysics Source Code Library [[record ascl:1605.004](#)]
- Mayor, M., Pepe, F., Queloz, D., et al. 2003, *The Messenger*, **114**, 20
- Mishenina, T. V., Bienaymé, O., Gorbaneva, T. I., et al. 2006, *A&A*, **456**, 1109
- Moultaka, J., Ilovaisky, S. A., Prugniel, P., & Soubiran, C. 2004, *PASP*, **116**, 693
- Nandakumar, G., Hayden, M. R., Sharma, S., et al. 2022, *MNRAS*, **513**, 232
- Ness, M., Hogg, D. W., Rix, H. W., Ho, A. Y. Q., & Zasowski, G. 2015, *ApJ*, **808**, 16
- Petit, P., Louge, T., Théado, S., et al. 2014, *PASP*, **126**, 469
- Piskunov, N., & Valenti, J. A. 2017, *A&A*, **597**, A16
- Plez, B. 2012, Turbospectrum: Code for spectral synthesis, Astrophysics Source Code Library [[record ascl:1205.004](#)]
- Prugniel, P., & Soubiran, C. 2001, *A&A*, **369**, 1048
- Raskin, G., van Winckel, H., Hensberge, H., et al. 2011, *A&A*, **526**, A69
- Schröder, C., Reiners, A., & Schmitt, J. H. M. M. 2009, *A&A*, **493**, 1099
- Sneden, C., Bean, J., Ivans, I., Lucatello, S., & Sobeck, J. 2012, MOOG: LTE line analysis and spectrum synthesis, Astrophysics Source Code Library [[record ascl:1202.009](#)]
- Soubiran, C., & Girard, P. 2005, *A&A*, **438**, 139
- Soubiran, C., Le Campion, J.-F., Brouillet, N., & Chemin, L. 2016, *A&A*, **591**, A118
- Soubiran, C., Brouillet, N., & Casamiquela, L. 2022, *A&A*, **663**, A4
- Soubiran, C., Creevey, O. L., Lagarde, N., et al. 2024, *A&A*, **682**, A145
- Strassmeier, K. G., Ilyin, I., & Weber, M. 2018, *A&A*, **612**, A45
- Taylor, M. B. 2005, in *Astronomical Society of the Pacific Conference Series*, 347, Astronomical Data Analysis Software and Systems XIV, eds. P. Shopbell, M. Britton, & R. Ebert, 29
- Telting, J. H., Avila, G., Buchhave, L., et al. 2014, *Astron. Nachr.*, **335**, 41
- Thomas, G. F., Battaglia, G., Gran, F., et al. 2024, *A&A*, **690**, A54
- Valenti, J. A., & Piskunov, N. 1996, *A&AS*, **118**, 595
- Wheeler, A., Ness, M., Buder, S., et al. 2020, *ApJ*, **898**, 58
- Worley, C. C., Jofré, P., Rendle, B., et al. 2020, *A&A*, **643**, A83
- Yee, S. W., Petigura, E. A., & von Braun, K. 2017, *ApJ*, **836**, 77

## Appendix A: Additional figures and tables

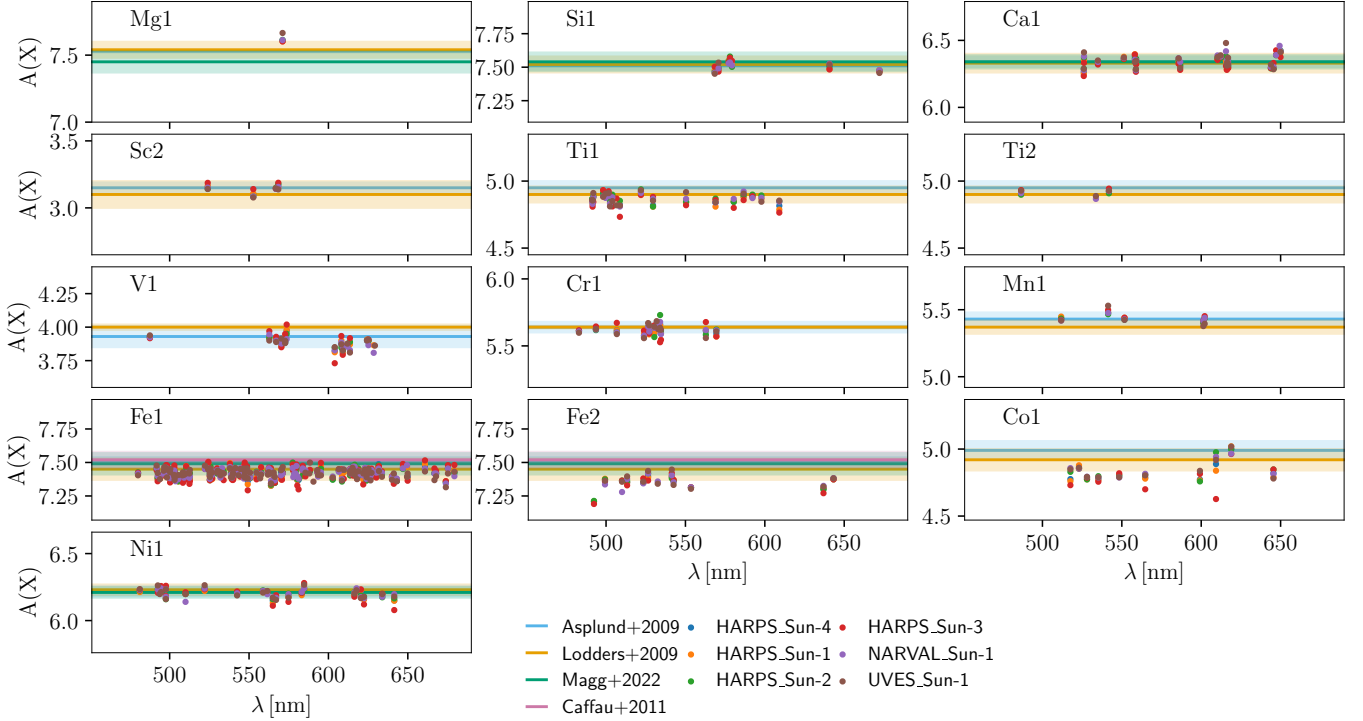
**Table A.1.** Solar abundances obtained for the different radiative transfer codes. We provide the final abundance combining the line-by-line weighted mean values of the six analysed spectra of the Sun, and its uncertainties as computed in Sect. 7.4.

|       | TURBOSPECTRUM |          | SPECTRUM |          | SME      |          | MOOG     |          |
|-------|---------------|----------|----------|----------|----------|----------|----------|----------|
|       | <i>A</i>      | $\delta$ | <i>A</i> | $\delta$ | <i>A</i> | $\delta$ | <i>A</i> | $\delta$ |
| Ca I  | 6.32          | 0.02     | 6.35     | 0.02     | 6.34     | 0.02     | 6.32     | 0.02     |
| Co I  | 4.81          | 0.03     | 4.81     | 0.03     | 4.82     | 0.03     | 4.80     | 0.03     |
| Cr I  | 5.60          | 0.01     | 5.60     | 0.01     | 5.61     | 0.01     | 5.58     | 0.01     |
| Fe I  | 7.40          | 0.01     | 7.41     | 0.01     | 7.42     | 0.01     | 7.41     | 0.01     |
| Fe II | 7.40          | 0.01     | 7.41     | 0.01     | 7.35     | 0.02     | 7.40     | 0.02     |
| Mg I  | 7.49          | 0.01     | 7.68     | 0.02     | 7.62     | 0.02     | 7.66     | 0.02     |
| Mn I  | 5.32          | 0.01     | 5.44     | 0.01     | 5.43     | 0.01     | 5.43     | 0.02     |
| Ni I  | 6.19          | 0.01     | 6.19     | 0.01     | 6.21     | 0.01     | 6.20     | 0.01     |
| Sc II | 3.17          | 0.01     | 3.16     | 0.01     | 3.14     | 0.01     | 3.16     | 0.01     |
| Si I  | 7.51          | 0.01     | 7.51     | 0.01     | 7.50     | 0.01     | 7.52     | 0.01     |
| Ti I  | 4.85          | 0.01     | 4.87     | 0.01     | 4.87     | 0.01     | 4.86     | 0.01     |
| Ti II | 4.95          | 0.01     | 4.96     | 0.01     | 4.92     | 0.01     | 4.96     | 0.01     |
| V I   | 3.88          | 0.02     | 3.88     | 0.02     | 3.89     | 0.02     | 3.86     | 0.01     |

## SPECTRUM



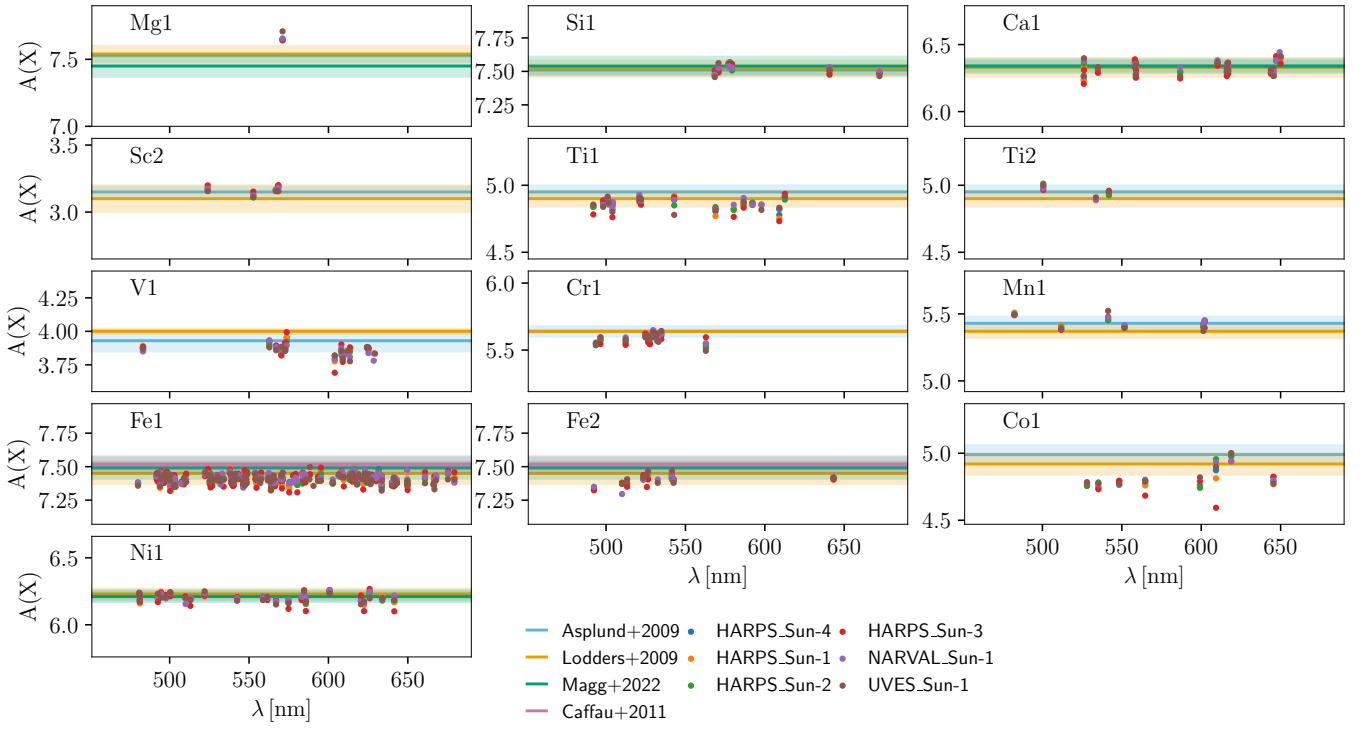
## SME



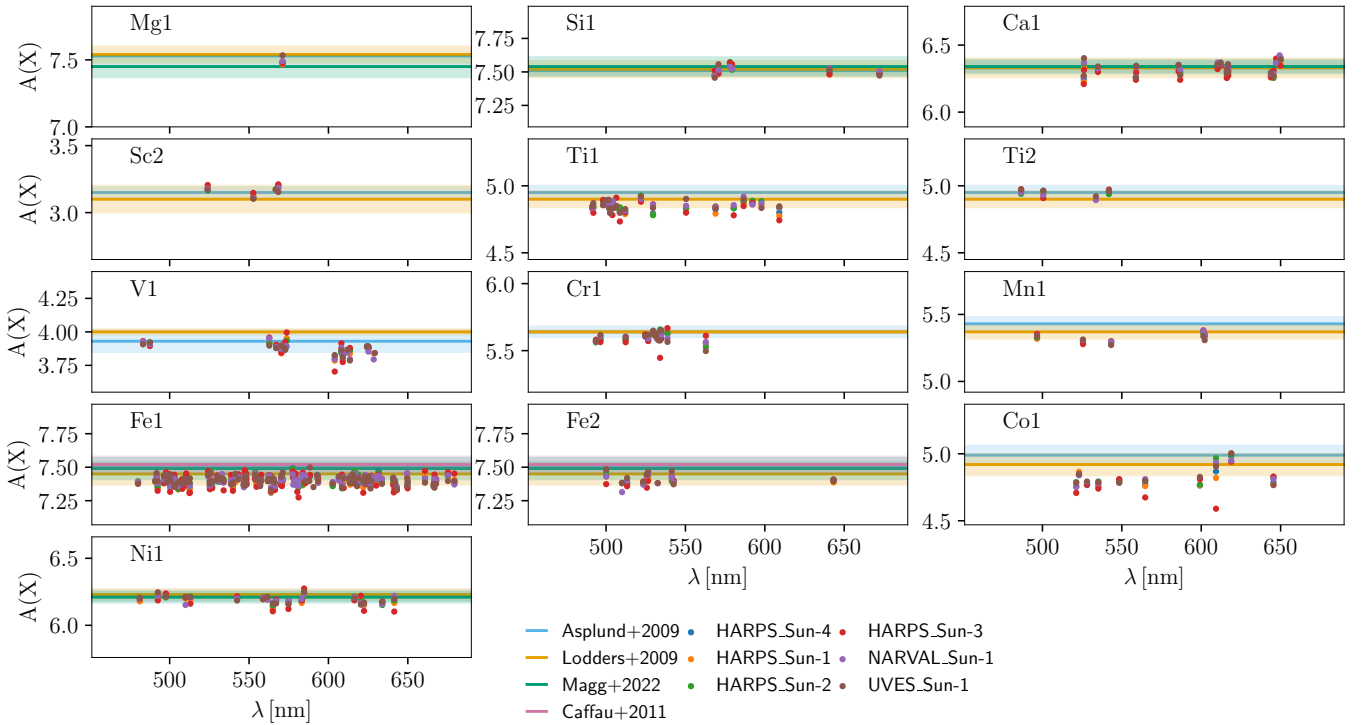
**Fig. A.1.** Line by line absolute abundances of the selection of lines performed in Sect. 4 in the six solar spectra. Horizontal lines represent literature values of the solar abundances, with their uncertainties represented as shadowed regions. Selection done for SPECTRUM (top), and SME (bottom)



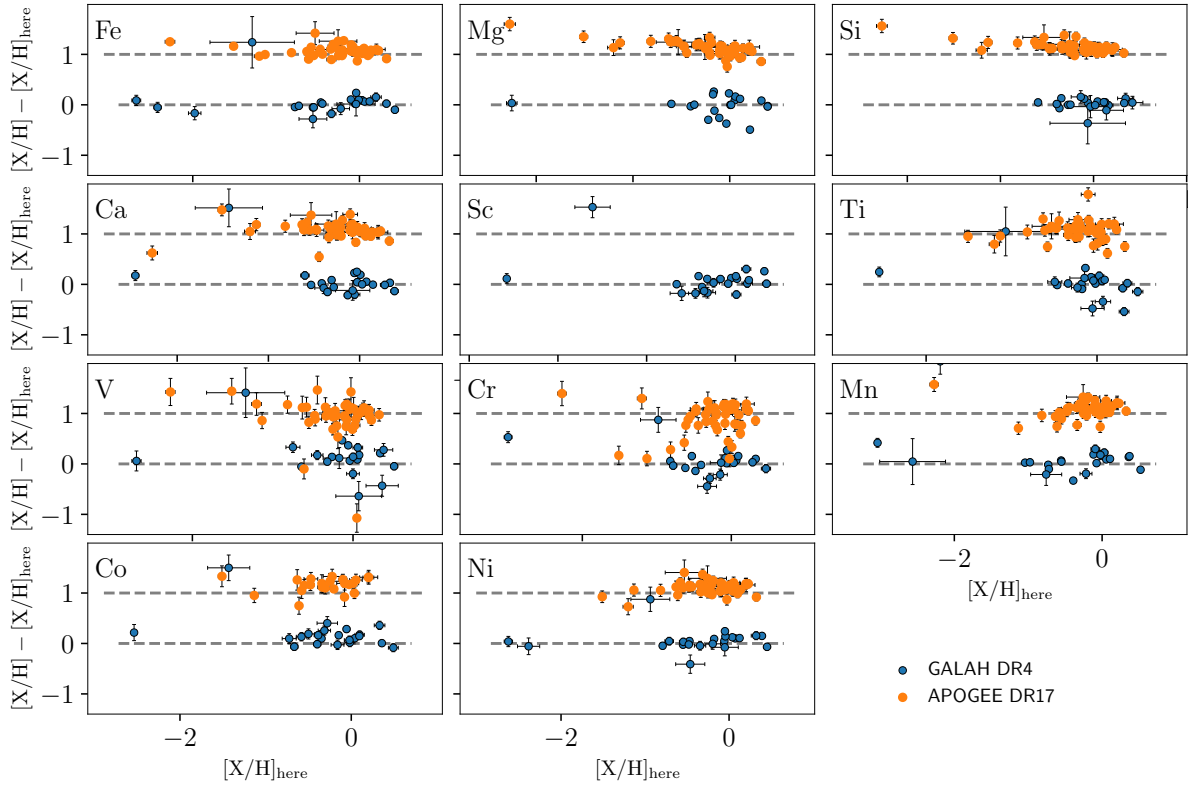
## MOOG



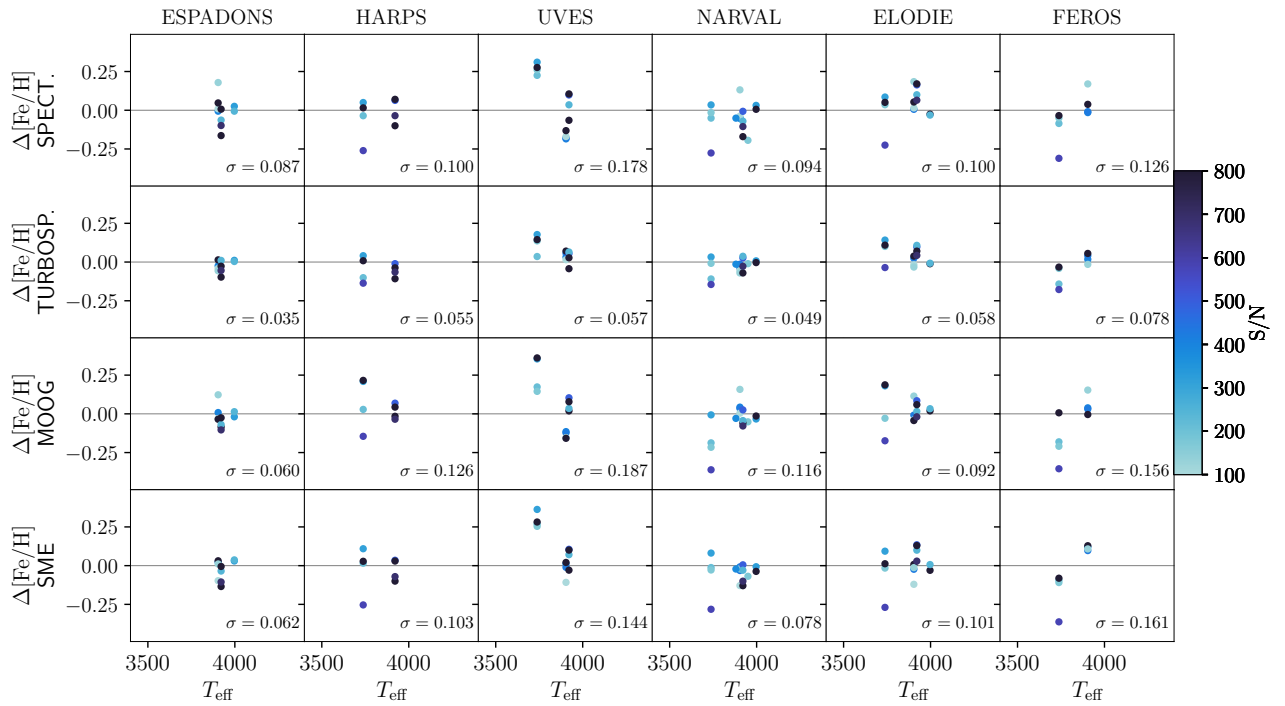
## TURBOSPECTRUM



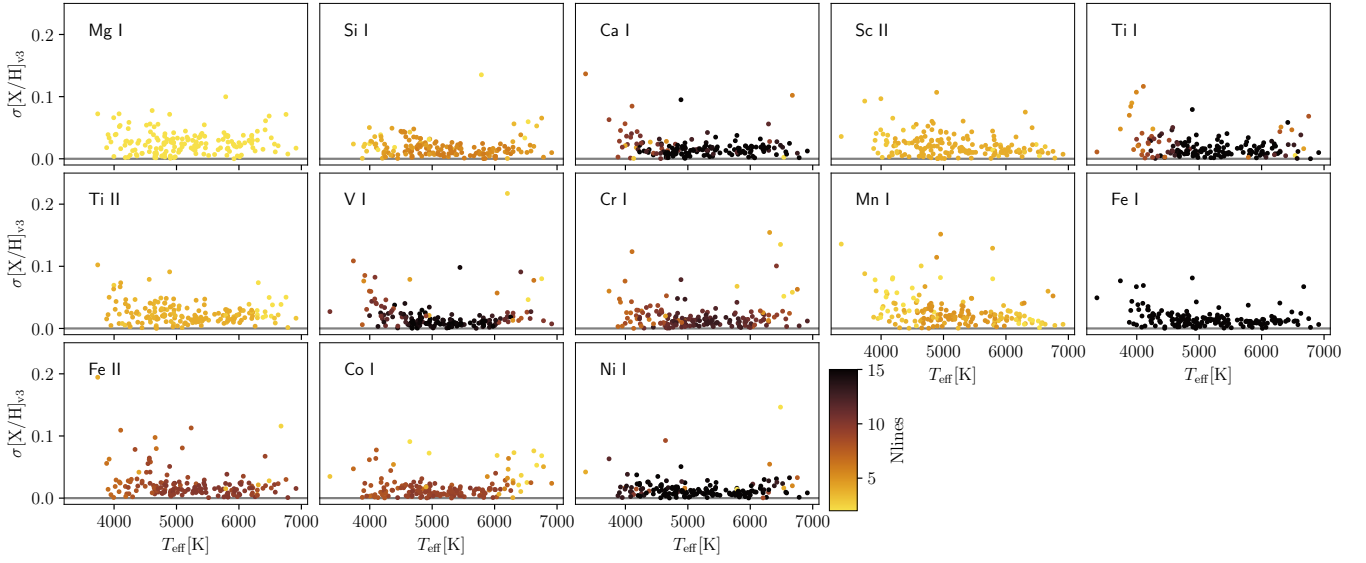
**Fig. A.2.** Same as Fig. A.1 for MOOG (top) and TURBOSPECTRUM (bottom).



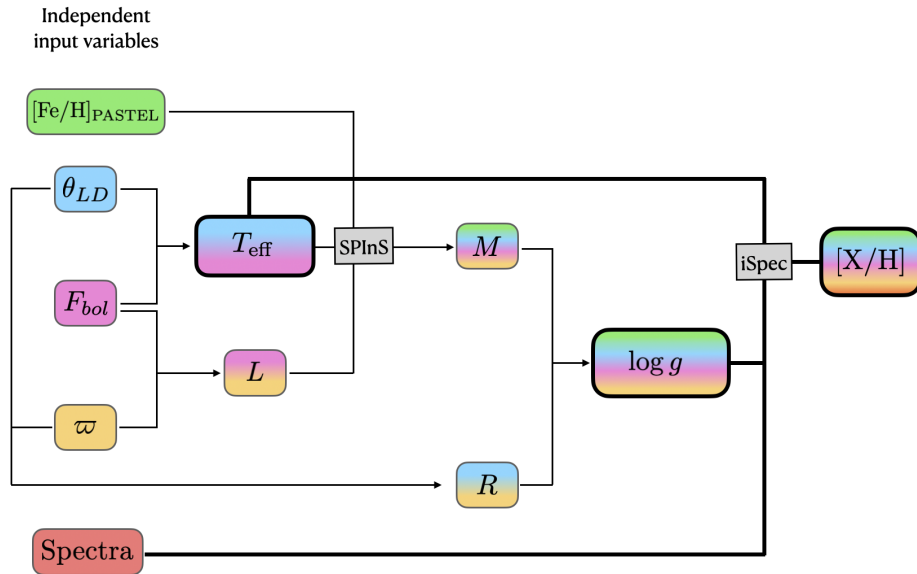
**Fig. A.3.** Comparison of the  $[X/H]$  values for the common stars between GBSv3 and GALAH DR4 (blue) and APOGEE DR17 (orange). Differences in abundances are plotted against the values from this work. APOGEE's comparison is shifted to 1 for clarity.



**Fig. A.4.** Comparison of the Fe abundances obtained with the different instruments (in vertical panels) and with the different codes (horizontal panels) as a function of effective temperature, only for stars below 4000 K. We only show the instruments which have a significant number of repeated cool stars. The colour depicts the S/N of the spectra. The vertical scatter is indicated in each panel.



**Fig. A.5.** Scatter in abundance obtained among the different spectra of the same star as a function of  $T_{\text{eff}}$  coloured by the number of lines detected for each star.



**Fig. A.6.** Schematic diagram representing the dependencies of the  $[X/H]$  on the various input quantities. Each colour corresponds to one independent quantity.