A&A 564, A133 (2014)
DOI: 10.1051/0004-6361/201322440

Astronomy & Astrophysics

Gaia FGK benchmark stars: Metallicity⋆,⋆⋆

P. Jofré[1,2], U. Heiter[3], C. Soubiran[2], S. Blanco-Cuaresma[2], C. C. Worley[1,4], E. Pancino[5,6], T. Cantat-Gaudin[7,8], L. Magrini[9], M. Bergemann[1,10], J. I. González Hernández[11], V. Hill[4], C. Lardo[5], P. de Laverny[4], K. Lind[1], T. Masseron[1,12], D. Montes[13], A. Mucciarelli[14], T. Nordlander[3], A. Recio Blanco[4], J. Sobeck[15], R. Sordo[7], S. G. Sousa[16], H. Tabernero[13], A. Vallenari[7], and S. Van Eck[12]

[1] Institute of Astronomy, University of Cambridge, Madingley Rd, Cambridge CB3 0HA, UK
e-mail: pjofre@ast.cam.ac.uk
[2] LAB UMR 5804, Univ. Bordeaux – CNRS, 33270 Floirac, France
[3] Department of Physics and Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden
e-mail: ulrike.heiter@physics.uu.se
[4] Laboratoire Lagrange (UMR7293), Univ. Nice Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, 06304 Nice, France
[5] INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy
[6] ASI Science Data Center, via del Politecnico s/n, 00133 Roma, Italy
[7] INAF, Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, Padova, 35122 Italy
[8] Dipartimento di Fisica e Astronomia, Università di Padova, vicolo Osservatorio 3, 35122 Padova, Italy
[9] INAF/Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy
[10] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
[11] Instituto de Astrofísica de Canarias, 38200 La Laguna, Tenerife, Spain
[12] Institut d'Astronomie et d'Astrophysique, Univ. Libre de Bruxelles, CP 226, Bd du Triomphe, 1050 Bruxelles, Belgium
[13] Dpto. Astrofísica, Facultad de CC. Físicas, Universidad Complutense de Madrid, 28040 Madrid, Spain
[14] Dipartimento di Fisica & Astronomia, Universitá degli Studi di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[15] Department of Astronomy & Astrophysics, University of Chicago, Chicago IL 60637, USA
[16] Centro de Astrofísica, Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal

Received 2 August 2013 / Accepted 24 January 2014

ABSTRACT

Context. To calibrate automatic pipelines that determine atmospheric parameters of stars, one needs a sample of stars, or "benchmark stars", with well-defined parameters to be used as a reference.
Aims. We provide detailed documentation of the iron abundance determination of the 34 FGK-type benchmark stars that are selected to be the pillars for calibration of the one billion *Gaia* stars. They cover a wide range of temperatures, surface gravities, and metallicities.
Methods. Up to seven different methods were used to analyze an observed spectral library of high resolutions and high signal-to-noise ratios. The metallicity was determined by assuming a value of effective temperature and surface gravity obtained from fundamental relations; that is, these parameters were known a priori and independently from the spectra.
Results. We present a set of metallicity values obtained in a homogeneous way for our sample of benchmark stars. In addition to this value, we provide detailed documentation of the associated uncertainties. Finally, we report a value of the metallicity of the cool giant ψ Phe for the first time.

Key words. standards – techniques: spectroscopic – surveys – stars: fundamental parameters

1. Introduction

Unlike in the field of photometry or radial velocities, stellar spectral analyses have lacked a clearly defined set of standard stars that span a wide range of atmospheric parameters up until now. The Sun has always been the single common reference point for spectroscopic studies of FGK-type stars. The estimate of stellar parameters and abundances by spectroscopy is affected by inaccuracies in the input data, the assumptions made in the model atmospheres, and the analysis method itself. This lack of reference stars, other than the Sun, makes it very difficult to validate and homogenize a given method over a larger parameter space (e.g. Lee et al. 2008a,b; Allende Prieto et al. 2008b; Jofré et al. 2010; Zwitter et al. 2008; Siebert et al. 2011).

This is particularly important for the many Galactic surveys of stellar spectra under development (RAVE, Steinmetz et al. 2006); (LAMOST, Zhao et al. 2006); (APOGEE, Allende Prieto et al. 2008a); (HERMES, Freeman 2010); (*Gaia*, Perryman et al. 2001); (*Gaia*-ESO, Gilmore et al. 2012; Randich et al. 2013). Each of these surveys has developed its own processing pipeline for the determination of atmospheric parameters and abundances, but the different methodologies may lead to a nonuniformity of the parameter scales. This is particularly problematic for the metallicities and chemical abundances, which are important for Galactic studies that performed via star counts. It is thus necessary to define a common and homogeneous scale

⋆ Based on NARVAL and HARPS data obtained within the *Gaia* DPAC (Data Processing and Analysis Consortium) and coordinated by the GBOG (Ground-Based Observations for *Gaia*) working group and on data retrieved from the ESO-ADP database.
⋆⋆ Tables 6–76 are only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/564/A133

to link different spectroscopic surveys probing every part of the Galaxy.

Kinematical and chemical analyses have been used to study the Milky Way for over a century (e.g. Kapteyn & van Rhijn 1920; Gilmore et al. 1989; Ivezić et al. 2012), which has provided, for example, the evidence for the existence of the Galactic thick disk (Gilmore & Reid 1983). This population contains stars, which have different spatial velocities (e.g., Soubiran 1993; Soubiran et al. 2003), different chemical abundance patterns (Bensby et al. 2004; Ramírez et al. 2007), and ages (for example, the works of Fuhrmann 1998; Allende Prieto et al. 2006) than the thin disk stars. Similarly, much of our knowledge about the Milky Way halo comes from these kind of studies (see review of Helmi 2008). A halo dichotomy similar to that of the disk has been the subject of discussion (Carollo et al. 2007; Schönrich et al. 2011; Beers et al. 2012), where the outer halo has a net retrograde rotation and is metal-poor, which is contrary to the inner halo and is slightly more metal-rich. Moreover, the inner halo is composed mainly of old stars (e.g., Jofré & Weiss 2011), although a number of young stars can be observed. The latter may be the remnants of later accretion of external galaxies. Evidence for these remnants have been found in stellar surveys like those by Belokurov et al. (2006). Schuster et al. (2012) found two chemical patterns in nearby halo stars and claim that they have an age difference, which supports the halo dichotomy scenario.

The analyses of stellar survey data are thus a crucial contribution to the understanding of our Galaxy. The problem arises when one wants to quantify the differences, for example, in chemical evolution and time of formation of all Galactic components, which are needed to understand the Milky Way as a unique body. A major obstacle in solving this problem is that each study, like those mentioned above, chooses their own data sets and methods. Homogeneous stellar parameters are therefore a fundamental cornerstone with which to put the different Galactic structures in context. The iron abundance ([Fe/H]) is of particular importance because it is a key ingredient for the study of the chemical evolution of stellar systems. Relations between the elemental abundance ratios [X/Fe] versus [Fe/H], where X is the abundance of the element X, are generally used as tracers for the chemical evolution of galaxies (e.g., Chiappini et al. 1997; Pagel & Tautvaisiene 1998; Reddy et al. 2003; Tolstoy et al. 2009; Adibekyan et al. 2012, 2013, to name a few). Thus, a good determination of the iron abundance is of fundamental importance.

A major contribution in the study of the Milky Way is expected from the *Gaia* mission (Perryman et al. 2001). In particular, the *Gaia* astrophysical parameter inference system (Apsis, Bailer-Jones et al. 2013) will estimate atmospheric parameters of one billion stars. The calibration of Apsis relies on several levels of reference stars with the first one being defined by benchmark stars. Some of these stars were chosen to cover the different spectral classifications and to have physical properties known independently of spectroscopy. This has motivated us to search for stars of different FGK types, which we call *Gaia* FGK Benchmark Stars (GBS). Knowing their radius, bolometric flux, and distance allows us to measure their effective temperature directly from the Stefan-Boltzmann relation and their surface gravity from Newton's law of gravity. Our sample of GBS consists of 34 stars covering different regions of the Hertzsprung-Russell diagram, which thereby represents the different stellar populations of our Galaxy. It is important to make the comment that our set of FGK GBS includes some M giant stars. We have decided to include them in the complete analysis described in this paper because we have been successful in analyzing them with our methods in a consistent way with respect to rest of the FGK stars of our benchmark sample. However, they should be treated with caution as benchmarks for FGK population studies.

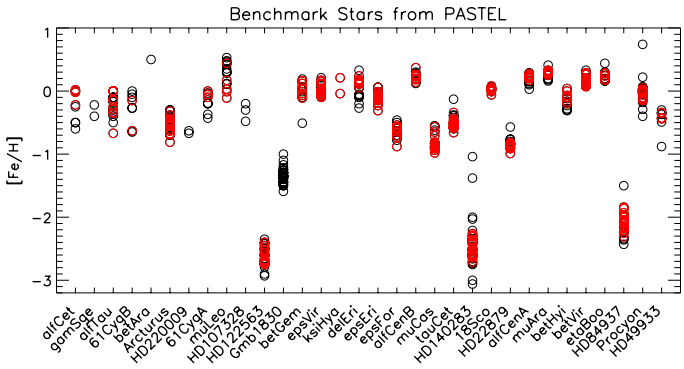

Fig. 1. Spectroscopic metallicities reported for the FGK GBS in the literature between 1948 to 2012, as retrieved from the PASTEL database (Soubiran et al. 2010). Black circles: all measurements. Red circles: only those measurements where $T_{\rm eff}$ and $\log g$ reported by those works agreed within 100 K and 0.5 dex with the fundamental values considered by us (see Table 1).

In Heiter et al. (in prep., hereafter Paper I), we describe our selection criteria and the determination of the "direct" effective temperature and surface gravity. In Blanco-Cuaresma et al. (2014, hereafter Paper II), we present our spectral data of these GBS and how we treat the spectra to build spectral libraries. This article describes the determination of the metallicity using a library of GBS that are compatible with the pipelines developed for the parameter estimation of the UVES targets from the *Gaia*-ESO public spectroscopic survey. For this purpose, up to seven different methods were employed to perform this spectral analysis, which span from methods using equivalent widths (EWs) to synthetic spectra. Since the aim of this work is to provide a metallicity scale based on the fundamental $T_{\rm eff}$ and $\log g$, we homogenized our methods by using common observations, atmospheric models, and atomic data.

Although the direct application of the reference metallicity is for the homogenization and evaluation of the different parameter-determination pipelines from the *Gaia*-ESO Survey and the calibration of Apsis, the final set of GBS parameters and their spectral libraries provides the possibility to calibrate spectroscopic astrophysical parameters for large and diverse samples of stars, such as those collected by HERMES, SDSS, LAMOST, and RAVE.

The structure of the paper is as follows: in Sect. 2, we review the metallicity values available in the literature for the GBS. In Sect. 3, we describe the properties of the spectra, while the methods and analysis structure are explained in Sect. 4. Our results are presented in Sect. 5 with an extensive discussion on the metallicity determination in Sect. 6. The paper concludes in Sect. 7.

2. The metallicity of GBS: reviewing the literature

The criteria to select the 34 GBS discussed in this paper can be found in Paper I. Due to their brightness and proximity, almost every star previously has been studied spectroscopically and has accurate HIPPARCOS parallax. Based on the recently updated PASTEL catalogue (Soubiran et al. 2010), metallicity values have been reported in 259 different works until 2012, which vary from 57 [Fe/H] measurements in the case of HD 140283 to

Table 1. Initial parameters and data information for the GBS.

Star ID	[Fe/H]LIT	σ[Fe/H]	N	$T_{\rm eff}$	$\sigma T_{\rm eff}$	$\log g$	$\sigma \log g$	$v \sin i$	Ref $v \sin i$	Source	R (k)	S/N	Extra spectra
18 Sco	0.03	0.03	15	5747	39	4.43	0.01	2.2	Saar	N	80	380	H
61 Cyg A	–0.20	0.11	5	4339	27	4.43	0.16	0.0	Benz	N	80	360	–
61 Cyg B	–0.27	0.00	2	4045	25	4.53	0.04	1.7	Benz	N	80	450	–
α Cen A	0.20	0.07	9	5840	69	4.31	0.02	1.9	Br10	H	115	430	U, H*
α Cen B	0.24	0.04	7	5260	64	4.54	0.02	1.0	Br10	H	115	460	–
α Cet	–0.26	0.23	8	3796	65	0.91	0.08	3.0	Zama	N	80	300	H, U
α Tau	–0.23	0.3	15	3927	40	1.22	0.10	5.0	Hekk	N	80	320	H
Arcturus	–0.54	0.04	11	4247	37	1.59	0.04	3.8	Hekk	N	80	380	H, U, U.P
β Ara	0.5	0.00	1	4073	64	1.01	0.13	5.4	Me02	H	115	240	–
β Gem	0.12	0.06	5	4858	60	2.88	0.05	2.0	Hekk	H	115	350	–
β Hyi	–0.11	0.08	6	5873	45	3.98	0.02	3.3	Re03	U.P	80	650	N, H, U
β Vir	0.13	0.05	11	6083	41	4.08	0.01	2.0	Br10	N	80	410	H
δ Eri	0.13	0.08	13	5045	65	3.77	0.02	0.7	Br10	N	80	350	H, U, U.P
ϵ Eri	–0.07	0.05	17	5050	42	4.60	0.03	2.4	VF05	U.P	80	1560	H, U
ϵ For	–0.62	0.12	9	5069	78	3.45	0.05	4.2	Schr	H	115	310	–
ϵ Vir	0.12	0.03	3	4983	61	2.77	0.01	2.0	Hekk	N	80	380	H
η Boo	0.25	0.04	9	6105	28	3.80	0.02	12.7	Br10	N	80	430	H
γ Sge	–0.31	0.09	2	3807	49	1.05	0.10	6.0	Hekk	N	80	460	–
Gmb 1830	-1.34	0.08	17	4827	55	4.60	0.03	0.5	VF05	N	80	410	–
HD 107328	–0.30	0.00	1	4590	59	2.20	0.07	1.9	Mass	N	80	380	H
HD 122563	–2.59	0.14	7	4608	60	1.61	0.07	5.0	Me06	N	80	300	H, U, U.P
HD 140283	–2.41	0.10	10	5720	120	3.67	0.04	5.0	Me06	N	80	320	H, U, U.P
HD 220009	–0.67	0.00	1	4266	54	1.43	0.10	1.0	Me99	N	80	380	–
HD 22879	–0.85	0.04	16	5786	89	4.23	0.03	4.4	Schr	N	80	300	–
HD 49933	–0.39	0.07	5	6635	91	4.21	0.03	10.0	Br09	H	115	310	–
HD 84937	–2.08	0.09	13	6275	97	4.11	0.06	5.2	Me06	H	115	480	N, U, U.P
ξ Hya	0.21	0.00	1	5044	38	2.87	0.01	2.4	Br10	H	115	370	–
μ Ara	0.29	0.04	12	5845	66	4.27	0.02	2.3	Br10	U	105	420	
μ Cas A	–0.89	0.04	14	5308	29	4.41	0.02	0.0	Luck	N	80	280	U
μ Leo	0.39	0.10	4	4433	60	2.50	0.07	5.1	Hekk	N	80	400	–
Procyon	–0.02	0.04	18	6545	84	3.99	0.02	2.8	Br10	U.P	80	760	N, H, U
ψ Phe	–	–	0	3472	92	0.62	0.11	3.0	Zama	U	70	220	–
Sun	0.00	0.00	0	5777	1	4.43	2E-4	1.6	VF05	H	115	350	H, N, U**
τ Cet	–0.53	0.05	17	5331	43	4.44	0.02	1.1	Saar	N	80	360	H

Notes. Column description: [Fe/H]LIT corresponds to the mean value of the metallicity obtained by works between 2000 and 2012 as retrieved from PASTEL (Soubiran et al. 2010), where σ[Fe/H] is the standard deviation of the mean and N represents the number of works considered for the mean calculation (see Sect. 2). Effective temperature, surface gravity and their respective uncertainties are determined from fundamental relations as in Paper I, and the rotational velocity $v \sin i$ is taken from literature with Ref representing the source of this value. The column Source indicates the instrument used to observe the spectrum in the 70 k library (see Sect. 4.2), where N, H, U, and U.P denote NARVAL, HARPS, UVES and UVES-POP spectra, respectively. The column R and S/N represent the resolving power and averaged signal-to-noise ratio of the spectra of the original library (see Sect. 4.2), respectively. For stars repeated in the complete 70 k library (see Sect. 4.2), the extra source are indicated in the column labeled as "extra spectra". (*) Two spectra in HARPS are available for this star with different wavelength calibrations. (**) There are many spectra of the Sun taken from different asteroids for HARPS and NARVAL (see Paper II for details of the library).

References. (Saar) Saar & Osten (1997); (Benz) Benz & Mayor (1984); (Br10) Bruntt et al. (2010); (Zama) Zamanov et al. (2008); (Hekk) Hekker & Meléndez (2007); (Me02) De Medeiros et al. (2002); (Re03) Reiners & Schmitt (2003); (VF05) Valenti & Fischer (2005); (Schr) Schröder et al. (2009); (Mass) Massarotti et al. (2008); (Me06) De Medeiros et al. (2006); (Me99) De Medeiros & Mayor (1999); (Br09) Bruntt (2009).

only one measurement for β Ara (Luck 1979), and no measurement at all for ψ Phe. Figure 1 shows those metallicity values taken from PASTEL for each GBS, where we show all metallicities in black and only those where the $T_{\rm eff}$ and $\log g$ values agree within 100 K and 0.5 dex, respectively, with the respective values adopted in Paper I in red. Note that the Sun and ψ Phe are not included in Fig. 1 because they are not in PASTEL.

Recent studies that have analyzed at least ten GBS are Allende Prieto et al. (2004), Valenti & Fischer (2005, hereafter VF05), Luck & Heiter (2006), Ramírez et al. (2007, hereafter R07), Bruntt et al. (2010), and Worley et al. (2012, hereafter W12), but none of them have analyzed the complete sample. The literature value for [Fe/H] that we adopt is the average of the most recent determinations after 2000 as listed in PASTEL. Table 1 gives the mean [Fe/H] with a standard deviation and the number of values considered after 3σ clipping of all references found in PASTEL after 2000. For β Ara the reported value is the only one available, by Luck (1979).

Figure 1 shows how metallicity varies from reference to reference. It is common to have differences of up to 0.5 dex for one star. Although the scatter significantly decreases when one considers those works with temperatures and surface gravities that agree with our values, there are still some stars, which present $\approx$0.5 dex difference in [Fe/H], such as Arcturus and the metal-poor stars HD 140283, HD 122563, and HD 22879. Note that Gmb 1830, γ Sge, and HD 107328 do not have $T_{\rm eff}$ and $\log g$ that agree with those of Paper I.

The stars are plotted in order of increasing temperature with α Cet being the coldest star and HD 49933 as the hottest one of our sample. Note that ψ Phe is colder than α Cet but is not plotted in the figure for the reasons explained above. Cold stars have more scattered metallicity literature values than hot stars.

This could be caused by the fewer works reporting metallicity for cold stars than for hot stars in PASTEL.

There are many sources of uncertainties that can slightly affect the results and ultimately produce these different [Fe/H] values in the literature. The methods of determining [Fe/H] in the literature are *highly inhomogeneous*, as they have been carried out by many groups using different assumptions, methodologies, and sources of data; some of them are briefly explained below. An extensive discussion of how these different aspects affect the determined parameters of giant stars can be found in Lebzelter et al. (2012) and for solar-type stars in Torres et al. (2012). The primary aspects are:

- *Methods*: the analysis of the observed spectra can be based on EWs (e.g. Luck & Heiter 2006; Sousa et al. 2008; Tabernero et al. 2012, R07) or fits to synthetic spectra (e.g. from VF05, Bruntt et al. 2010). Other methods that are different from EWs or fits can be used for deriving [Fe/H], such as the parametrisation methods based on projections (Jofré et al. 2010; Worley et al. 2012). Moreover, each method uses a different approach to find the continuum of the spectra.
- *Atomic data*: for each method, the line list can be built using atomic data from different sources; that is, Bruntt et al. (2010) and VF05 used the VALD database (Kupka et al. 1999), whereas R07 adopted the values given in the NIST[1] database (Wiese et al. 1996). There are also methods where the atomic data is adjusted to fit a reference star, which is typically the Sun (e.g., Santos et al. 2004; Sousa et al. 2008).
- *Observations*: for the same star, different observations are taken and analyzed. For example, Allende Prieto et al. (2004) and R07 studied spectra from the two-coudé instruments (Tull et al. 1995) at the McDonald Observatory and from the FEROS instrument (Kaufer et al. 2000) in La Silla. The VF05 work used spectra from the spectrometer HIRES (Vogt et al. 1994) at Keck Observatory, UCLES (Diego et al. 1990) at the Siding Spring Observatory and the Hamilton spectrograph (Vogt 1987) at Lick Observatory. Worley et al. (2012) used FEROS spectra. These spectra differ in wavelength coverage, resolution, flux calibrations, and signal-to-noise ratios.
- *Atmospheric models*: MARCS (Gustafsson et al. 2008, and references therein) and Kurucz atmospheric models are both used throughout the literature and can produce abundance differences of up to 0.1 dex for identical input parameters (Allende Prieto et al. 2004; Pancino et al. 2011). In addition, some groups have started to use three-dimensional (3D) hydrodynamical atmospheric models, which can lead to different stellar parameters as compared to when using one-dimensional (1D) hydrostatic models(e.g, Collet et al. 2007).
- *Solar abundances*: over the past years, the abundances of the Sun have been updated and, therefore, metallicities are provided using different solar abundances. Edvardsson et al. (1993), for example, considered the solar chemical abundances of Anders & Grevesse (1989) while Meléndez et al. (2008) refered to the solar abundances of Asplund et al. (2005). A change in solar composition affects the atmospheric models and, therefore, the abundances.
- *Nonlocal thermodynamical equilibrium*: the NLTE effects can have a severe impact on the abundance determinations, especially for the neutral lines of predominantly singly-ionized elements, like Fe I (Thévenin & Idiart 1999; Asplund 2005; Asplund et al. 2009). The effect is typically larger for metal-poor and giant stars (Thévenin & Idiart 1999; Bergemann et al. 2012; Lind et al. 2012). Only a few methods make corrections to the abundances due to these effects (e.g. Thévenin & Idiart 1999; Mishenina & Kovtyukh 2001).

This work attempts to reduce the inhomogeneities found in the parameters of our sample of stars. This is done by re-estimating the metallicity using the same technique for all stars.

3. Observational data

The spectra used in this work have a very high signal-to-noise (S/N) and high resolution. Since the GBS cover the northern and southern hemisphere, it is not possible to obtain the spectra of the whole sample with one single spectrograph. For that reason, we have compiled a spectral library collecting spectra from three different instruments: HARPS, NARVAL, and UVES.

The HARPS spectrograph is mounted on the ESO 3.6 m telescope (Mayor et al. 2003), and the spectra were reduced by the HARPS data reduction software (version 3.1). The NARVAL spectrograph is located at the 2 m Telescope *Bernard Lyot* (Pic du Midi, Aurière 2003). The data from NARVAL were reduced with the Libre-ESpRIT pipeline (Donati et al. 1997). The UVES spectrograph is hosted by unit telescope 2 of ESO's VLT (Dekker et al. 2000). Two sources for UVES spectra are considered, the Advanced Data Products collection of the ESO Science Archive Facility[2] (reduced by the standard UVES pipeline version 3.2, Ballester et al. 2000), and the UVES Paranal Observatory Project UVES-POP library (Bagnulo et al. 2003, processed with data reduction tools specifically developed for that library). More details of the observations and properties of the original spectra can be found in Paper II.

To have an homogeneous set of data for the metallicity determination, we have built a spectral library as described in Paper II. The spectra have been corrected to laboratory air wavelengths. The wavelength range has been reduced to the UVES 580 setup, which is from 476 to 684 nm, with a gap from 577 to 584 nm between the red and the blue CCD. We have chosen this range because it coincides with the standard UVES setup employed by the *Gaia*-ESO Survey, and our methods are developed to work in that range. Two libraries of spectra are considered: The first one with $R = 70\,000$, which is the highest common resolution available in our data, and the second one that retains the original resolution ($R > 70\,000$), which is different for each spectrum and is indicated in Table 1. Finally, each method used the best way to identify the continuum.

4. Method

For consistency, we have used common material and assumptions as much as possible, which are explained below. In this section, we also give a brief description of each metallicity determination method considered for this work.

4.1. Common material and assumptions

The analysis is based on the principle that the effective temperature and the surface gravity of each star is known. These values (indicated in Table 1) are obtained independently from the spectra using fundamental methods by taking the angular diameter and bolometric flux to determine the effective temperature, the

[1] http://physics.nist.gov/PhysRefData/ASD/lines_form.html

[2] http://archive.eso.org/eso/eso_archive_adp.html

distance, angular diameter and mass to determine surface gravity. In our analysis, we fix $T_{\rm eff}$, $\log g$ values, and rotational velocity (values also indicated in Table 1). The latter were taken from the literature, for which the source is also indicated in Table 1. For those methods where a starting value for the metallicity is needed, we set [Fe/H] = 0.

We used the line list that has been prepared for the analysis of the stellar spectra for the *Gaia*-ESO survey (Heiter et al., in prep., version 3, hereafter GES-v3). The line list includes simple quality flags like "yes" (Y), "no" (N), and "undetermined" (U). These were assigned from an inspection of the line profiles, and the accuracy of the $\log gf$ value for each line is based on comparisons of synthetic spectra with a spectrum of the Sun and of Arcturus. If the profile of a given line is well reproduced and its gf value is well determined, then the line has "Y/Y". On the contrary, if the line is not well reproduced (also due to blends) and the gf value is very uncertain, the line is marked with the flag "N/N". We considered all lines, except those assigned with the flag "N" for the atomic data or the line profile. Finally, all methods used the 1D hydrostatic atmosphere models of MARCS (Gustafsson et al. 2008), which consider local thermodynamical equilibrium (LTE) and plane-parallel or spherically symmetric geometry for dwarfs and giants, respectively. These atmospheric models were chosen to be consistent with the spectral analysis of the UVES targets from the *Gaia*-ESO Survey.

4.2. Runs

Three main analyses were made, as explained below. These runs allow us to study the behavior of our results under different methods, resolutions and instruments.

1. *Run-nodes*: one spectrum per star at $R = 70\,000$, where the "best" spectrum was selected by visual inspection for stars with more than one spectrum available in our library. The evaluation was mainly based on the behavior of the continuum but also considered the S/N, the amount of cosmic ray features and telluric absorption lines. The source of the spectra used for this test is indicated in Table 1. Hereafter, we call this set of data the "70 k library". The purpose of this run was to have a complete analysis and overview of the performance of different methods for a well-defined set of spectra.
2. *Run-resolutions*: the same selection of spectra as in *Run-nodes* but using the original resolution version of the library. This value is indicated in Table 1. This run allowed us to make a comparative study of the impact of resolution on the accuracy of the final metallicity. This set of spectra is hereafter called the "Original library".
3. *Run-instruments*: all available spectra obtained with several instruments, which are convolved to $R = 70\,000$, i.e. several results for each star. The source of the available spectra for each star (when applicable) is indicated in the last column of Table 1. Hereafter we call this data set the "complete 70 k library". This run gave us a way to study instrumental effects and to assess the internal consistency of the metallicity values with regard to the spectra being employed.

4.3. Nodes method description

In this section, we explain the methods considered for this analysis. They vary from fitting synthetic spectra to observed spectra to classical EW methods. Since this analysis was based on 1D hydrostatic atmospheric models, the microturbulence parameter also needed to be taken into account. We considered the value of $v_{\rm mic}$ obtained from the relations of M. Bergemann and V. Hill that derived for the analysis of the targets from the *Gaia*-ESO Survey (hereafter GES relation). Some of the methods determine this parameter simultaneously with [Fe/H] using the GES relation as an initial guess, while others kept $v_{\rm mic}$ fixed to the value obtained from the relation. In the following, we briefly explain each method individually.

4.3.1. LUMBA

Code description: the LUMBA-node (Lund, Uppsala, MPA, Bordeaux, ANU[3]) uses the SME (Spectroscopy Made Easy, Valenti & Piskunov 1996; Valenti & Fischer 2005) code (version 298) to analyze the spectra. This tool performs an automatic parameter optimization using a chi-square minimization algorithm. Synthetic spectra are computed by a built-in spectrum synthesis code for a set of global model parameters and spectral line data. A subset of the global parameters is varied to find the parameter set, which gives the best agreement between observations and calculations. In addition to the atmospheric models and line list as input, the SME method requires masks containing information on the spectral segments that are analyzed, the absorption lines that are fitted, and the continuum regions that are used for continuum normalization. The masks have to be chosen so that it is possible to homogeneously analyse the same spectral regions for all stars. To create the masks, we plotted the normalized fluxes of all GBS and looked for those lines and continuum points that are present in all stars. The analysis of the LUMBA node was mainly carried out by P. Jofré, U. Heiter, C. Soubiran, S. Blanco-Cuaresma, M. Bergemann, and T. Nordlander.

Iron abundance determination: we made three iterations with SME: (i) determine only metallicity starting from [Fe/H] = 0 and fixing $v_{\rm mic}$ and macroturbulence velocity ($v_{\rm mac}$) to the values obtained from the GES relations. (ii) Determine $v_{\rm mic}$ and $v_{\rm mac}$ by fixing the [Fe/H] value obtained in the previous iteration (see below). (iii) Determine [Fe/H], including a final correction of radial velocity for each line, which accounts for residuals in the wavelength calibration or line shifts due to thermal motions (Molaro & Monai 2012), by using those values obtained in the previous iterations as starting points. To validate the ionization balance in our method, we built two sets of masks for Fe I and Fe II, separately.

Broadening parameters: we estimated the microturbulence and macroturbulence parameters in an additional run with SME. For that, we created a mask including all strong neutral lines with $-2.5 > \log gf > -4.0$ in the spectral range of our data. This value was chosen because lines in this $\log gf$ regime are sensitive to $v_{\rm mic}$ with SME (Valenti & Piskunov 1996). To determine the broadening parameters, we considered the initial values obtained from the GES relation and fixed with SME $T_{\rm eff}$ $\log g$ and [Fe/H].

Discussion: special treatment was necessary for the metal-poor stars with [Fe/H] ≤ -0.6 and for the cold stars with $T_{\rm eff} \leq 4100$ K. In the case of the metal-poor stars, a significant number of lines from the line masks were not properly detected, which resulted in the spectra being incorrectly shifted in radial velocity. Since the library is in the laboratory rest frame, we decided not to make a re-adjustment of the radial velocity for these

[3] Lund: Lund Observatory, Sweden; Uppsala: Uppsala University, Sweden; MPA: Max-Planck-Institut für Astrophysik, Germany; Bordeaux: Laboratoire d'Astrophysique de Bordeaux, France; ANU: Australian National University, Australia.

stars. Cold stars needed a special line mask. In many segments, molecular blends were very strong, making it impossible to obtain a good continuum placement and also a good fit between the observed and the synthetic spectra. Moreover, determining iron abundances of blended lines with molecules that are not included in our line list results in an incorrect estimation of the true iron content in the atmosphere. We looked at each spectrum individually and selected the unblended iron lines.

4.3.2. Nice

Code description: the pipeline is built around the stellar parameterization algorithm MATISSE (MATrix Inversion for Spectrum SynthEsis), which has been developed at the Observatoire de la Côte d'Azur primarily for use in *Gaia* RVS[4] stellar parameterization pipeline (Recio-Blanco et al. 2006) but also for large scale projects such as AMBRE (Worley et al. 2012; De Laverny et al. 2012) and the *Gaia*-ESO Survey. The algorithm MATISSE simultaneously determines the stellar parameters (θ: $T_{\rm eff}$, $\log g$, [M/H], and [α/Fe][5] of an observed spectrum $O(\lambda)$ by the projection of that spectrum onto a vector function $B_\theta(\lambda)$. The $B_\theta(\lambda)$ functions are optimal linear combinations of synthetic spectra $S(\lambda)$ within the synthetic spectra grid. For this work, we adopted the synthetic spectra grid built for the *Gaia*-ESO survey by using the same line list and atmosphere models as the other nodes and the GES relation for the microturbulence. A full documentation on how this grid is computed is found in De Laverny et al. (2012). The analysis done by the Nice group was mainly carried out by C. C. Worley, P. de Laverny, A. Recio-Blanco and V. Hill.

Iron abundance determination: the wavelength regions selected for this analysis were based on the Fe line mask used by LUMBA. Continuum regions of minimum 8 Å were set about each accepted Fe line or group of lines.

Broadening parameters: since this method is restricted to fit synthetic spectra from a pre-computed grid, $v_{\rm mic}$ was determined from the best fit of spectra computed using the GES relation.

Discussion: holding $T_{\rm eff}$ and $\log g$ constant and allowing metallicity to vary, is not fundamentally possible for MATISSE in the current configuration as MATISSE converges on all the parameters simultaneously. The algorithm MATISSE does accept a first estimate of the parameters, which were set in this case to the fundamental $T_{\rm eff}$, $\log g$, solar [M/H], and [α/Fe]. However, MATISSE then iterates freely through the solution space to converge on the best fit stellar parameters for each star based on the synthetic spectra grid.

Additionally, a direct comparison of the normalized observed spectrum to the synthetic spectra by χ^2-test was carried out. The synthetic spectra were restricted to the appropriate constant $T_{\rm eff}$ and $\log g$ with varying [M/H] and [α/Fe]. This test did not require the MATISSE algorithm and only provided grid point stellar parameters. However, it was useful as a confirmation of the MATISSE analysis and also a true test for which $T_{\rm eff}$ and $\log g$ could be held constant by allowing metallicity to vary. In addition, this is a useful analysis as a validation of the grid of synthetic spectra available for the *Gaia*-ESO Survey.

This configuration of considering only regions around Fe lines, which performed well for metal-rich dwarfs but was more problematic for low-gravity and metal-poor stars. Three potential reasons are a) the poor representation of the ionization balance due to the small number of Fe II lines; b) the strong lines from the regions where the wings are typically good gravity indicators; and c) the normalization issues for these small spectral regions around Fe lines.

Even for the problematic stars, where the $\log g$ $B_\theta(\lambda)$ functions did show a lack of strong sensitivity due to a lack of strong features and the regions of reasonable $\log g$ sensitivity (~5000 Å to 5200 Å) were difficult to normalize accurately, MATISSE found the solution for each star that best fits this configuration of the synthetic grid. This was confirmed in most cases by the χ^2-test. Note that the final provided solutions here do not represent those favoured by a full-MATISSE analysis because of the a-priori fixed $T_{\rm eff}$, $\log g$, and the selection of only iron lines in the spectral windows. Some consequences of this fixed analysis for MATISSE are discussed below.

4.3.3. ULB (Université Libre de Bruxelles)

Code description: the ULB node uses the code BACCHUS (Brussels Automatic Code for Characterising High accUracy Spectra), which consists of three different modules designed to derive abundances, EWs, and stellar parameters. The current version relies on an interpolation of the grid of atmosphere models using a thermodynamical structure, as explained in Masseron (2006). Synthetic spectra are computed using the radiative transfer code TURBOSPECTRUM (Alvarez & Plez 1998; Plez 2012). This analysis was carried out mainly by T. Masseron and S. Van Eck.

Iron abundance determination: the iron abundance determination module includes local continuum placement (adopted from spectrum synthesis using the full set of lines), cosmic and telluric rejection algorithms, local S/N estimation, and selection of observed flux points contributing to the line absorption. Abundances are derived by comparison of the observation with a set of convolved synthetic spectra with different abundances using four different comparison methods: χ^2 fitting, core line intensity, synthetic fit, and EWs. A decision tree is constructed from those methods to select the best matching abundances.

Broadening parameters: microturbulence velocity was determined in an iterative way with the iron abundances. For that, a new model atmosphere was taken into account for the possible change in metallicity by adjusting the microturbulence velocity. Additionally, a new convolution parameter for the spectral synthesis encompassing macroturbulence velocity, instrument resolution of 70 000, and stellar rotation was determined and adopted if necessary.

4.3.4. Bologna

Code description: the analysis is based on the measurement of EW. This was done using DAOSPEC (Stetson & Pancino 2008), which is run through DOOp (Cantat-Gaudin et al. 2014), a program that automatically configures some of the DAOSPEC parameters and makes DAOSPEC run multiple times until the input and output FWHM[6] of the absorption lines agree within 3%. The analysis of the Bologna method was mainly carried out by E. Pancino, A. Mucciarelli, and C. Lardo.

[4] Radial Velocity Spectrometer.

[5] The metallicity [M/H] is derived using spectral features of elements that are heavier than helium, while the [α/Fe] determination uses spectral features of α-elements.

[6] The DAOSPEC method uses the same FWHM (scaled with wavelength) for all lines; thus, an input FWHM is required from the user to be able to more easily separate real lines from noise (which generally have FWHM of 1–2 pixels). Later, the code refines the FWHM and determines the best value from the data, which produces an output FWHM.

Iron abundance determination: the abundance analysis was carried out with GALA (Mucciarelli et al. 2013), an automatic program for atmospheric parameter and chemical abundance determination from atomic lines, which is based on the Kurucz suite of programs (Sbordone et al. 2004; Kurucz 2005). Discrepant lines with respect to the fits of the slopes of Fe abundance versus EW, excitation potential, and wavelength were rejected with a 2.5σ cut, as lines with too small or to large EW depending on the star were as well.

Broadening parameters: we looked for the best $v_{\rm mic}$ whenever possible by looking for the solution, which minimized the slope of the [Fe/H] vs. EW relation. If it was not possible to converge to a meaningful value of $v_{\rm mic}$ (mostly because not enough lines in the saturation regime were measurable with a sufficiently accurate Gaussian fit) for some stars, we used the GES relations, which provided a flat [Fe/H] vs. EW relation.

Discussion: some of the stars, which have deep molecular bands or heavy line crowding, had to be remeasured with an exceptionally high order in the polynomial fit of the continuum (larger than 30). The stars, which needed a fixed input $v_{\rm mic}$, were 61 Cyg A and B, β Ara, ϵ Eri, and Gmb 1830.

4.3.5. EPINARBO

Code description: the EPINARBO-node (ESO-Padova-Indiana-Arcetri-Bologna[7]) adopts a code, FAMA (Magrini et al. 2013), based on an automatization of MOOG (Sneden 1973, version released on 2010), which is based on EWs that are determined in the same way as in the Bologna method (see Sect. 4.3.4)[8]. The analysis of this node was mainly carried out by T. Cantat-Gaudin, L. Magrini, A. Vallenari, and R. Sordo.

Iron abundance determination: for the purpose of determination of metallicity only, we fixed the effective temperature and surface gravity and computed $v_{\rm mic}$ with the adopted formulas of the GES relation. By keeping these three atmospheric parameters fixed, we obtained the average of both neutral and ionized iron abundances, discarding those abundances which are discrepant with one-σ clipping.

Broadening parameters: with the value of metallicity obtained as described above, we recomputed $v_{\rm mic}$, which is set to minimize the slope of the relationship between the Fe I abundance and the observed EWs. Iteratively, we repeated the analysis with the new set of atmospheric parameters and, with one σ clipping, we obtained the final values of Fe I and Fe II abundances.

4.3.6. Porto

Code description: this method is based on EWs, which are measured automatically using ARES[9] (Sousa et al. 2007). These are then used to compute individual line abundances with MOOG (Sneden 1973). The analysis of the Porto node was carried out by S.G. Sousa.

Iron abundance determination: for this exercise, we assumed that the excitation and ionization balance is present. In every iteration, we rejected outliers above 2σ. We find the final value of [Fe/H] when the input [Fe/H] of the models is equal to the average of the computed line abundances.

Broadening parameters: for giants, we computed the microturbulence because it depends on [Fe/H], which is a parameter that we initially set to [Fe/H] = 0 for all stars. This was done by determining [Fe/H] and $v_{\rm mic}$ simultaneously requiring excitation balance. For dwarfs, we utilized the value obtained from the GES relation, since it is independent of the [Fe/H] of the star.

4.3.7. UCM (Universidad Computense de Madrid)

Code description: the UCM node relies on EWs. An automatic code based on some subroutines of StePar (Tabernero et al. 2012) was used to determine the metallicity. Metallicities are computed using the 2002 version of the MOOG code (Sneden 1973). We modified the interpolation code provided with the MARCS grid to produce an output model readable by MOOG. We also wrote a wrapper program to the MARCS interpolation code to interpolate any required model on the fly.

Iron abundance determination: the metallicity is inferred from any previously selected line list. We iterate until the metallicity from the Fe lines and metallicity of the model are the same. The EW determination of the Fe lines was carried out with the ARES code (Sousa et al. 2007). In addition, we performed a 3σ rejection of the Fe I and Fe II lines after a first determination of the metallicity. We then reran our program again without the rejected lines. This analysis was carried out by J. I. González-Hernández, D. Montes, and H. Tabernero.

Broadening parameters: for the van der Waals damping prescription, we use the Unsöld approximation. As in the Porto method, we determined $v_{\rm mic}$ only for giants, while we fixed $v_{\rm mic}$ by the values obtained from the GES relation for dwarfs.

5. Results

In this section, we discuss the metallicity obtained from the three runs described in Sect. 4.2. This allows us to have a global view of how the different methods compare to each other. We further discuss the impact that our stellar parameters have on the ionization balance, and finally, we present the NLTE corrections.

5.1. Comparison of different methods

Table 2 lists the results obtained from *run-nodes*, where every node has determined the metallicity of one spectrum per GBS. The value indicates the result obtained from the analysis of Fe I lines under LTE. The table also lists the mean $v_{\rm mic}$ value obtained by the different nodes, with $\sigma v_{\rm mic}$ representing the standard deviation of this mean. In Fig. 2, we show the difference between the result of each node and the mean literature value as a function of GBS in increasing order of temperature. The name of the star is indicated at the bottom of the figure with its corresponding fundamental temperature at the top of it.

For warm stars (i.e. $T_{\rm eff}$ > 5000 K), the values of metallicity obtained by the different methods have a standard deviation of 0.07 dex. Moreover, these values agree well with the literature with a mean offset of +0.04 dex. The standard deviation increases notably for cooler stars, which are typically on the order of 0.1 dex with a maximum of 0.45 for β Ara. Note that this star

[7] European Southern Observatory; Osservatorio Astronomico di Padova, Italy; Indiana University, USA; Osservatorio Astrofisico di Arcetri, Italy; Istituto Nazionale di Astrofisica, Italy.

[8] These measurements were carried out independently from the Bologna ones with slight differences in the configuration parameters (continuum polynomial fit order, input FWHM, starting radial velocity, and so on), leading to mean differences that are generally on the order or ±1%, except for a few stars which could have a mean difference up to ≃3%.

[9] The ARES code can be downloaded at http://www.astro.up.pt/

Table 2. Metallicity of GBS obtained individually by each method by analyzing neutral iron abundances and assuming LTE.

Star	LUMBA	Bologna	EPINARBO	Nice	UCM	ULB	Porto	$v_{\rm mic}$ (Km s^{-1})	$\sigma v_{\rm mic}$
18 Sco	+0.01	+0.03	–0.10	+0.00	–0.02	–0.01	–0.02	1.2	0.2
61 Cyg A	–0.42	–0.35	–0.33	–0.25	–0.40	–0.45	–0.39	1.1	0.04
61 Cyg B	–0.47	–0.35	–0.48	–0.50	–0.34	–0.74	–0.32	1.1	0.36
α Cen A	+0.29	+0.25	+0.14	+0.25	+0.22	+0.14	0.23	1.2	0.07
α Cen B	+0.23	+0.27	+0.06	+0.25	+0.17	+0.21	+0.13	1.1	0.31
α Cet	–0.13	–0.33	–0.39	+0.00	–0.38	–0.64	–	1.4	0.4
α Tau	–0.12	–0.23	–0.31	–0.25	–0.34	–0.43	–	1.4	0.4
Arcturus	–0.52	–0.56	–0.54	–0.50	–0.50	–0.65	–0.46	1.3	0.12
β Ara	+0.35	+0.11	–0.08	+0.00	+0.07	–0.16	–	1.5	0.46
β Gem	+0.05	+0.07	+0.03	+0.00	+0.16	–0.01	0.24	1.1	0.21
β Hyi	–0.04	–0.06	–0.09	–0.25	–0.11	–0.06	–0.09	1.3	0.04
β Vir	+0.17	0.15	+0.10	+0.00	+0.11	+0.11	+0.11	1.4	0.09
δ Eri	+0.06	+0.14	–0.06	+0.00	+0.04	+0.00	+0.00	1.2	0.22
ϵ Eri	–0.10	–0.11	–0.09	–0.25	–0.15	–0.12	–0.19	1.1	0.05
ϵ For	–0.58	–0.59	–0.62	–0.75	–0.68	–0.61	–0.67	1.2	0.13
ϵ Vir	+0.09	+0.09	+0.02	+0.00	+0.24	+0.04	+0.08	1.1	0.25
η Boo	+0.34	+0.30	+0.33	+0.00	+0.08	–0.28	+0.27	1.4	0.19
γ Sge	–0.01	–0.01	–0.09	–0.25	–0.05	–0.39	–	1.4	0.34
Gmb 1830	–1.48	–1.47	–1.62	–1.50	–1.48	–1.80	–1.46	1.1	0.57
HD 107328	–0.20	–0.35	–0.26	–0.25	–0.22	–0.47	–0.10	1.2	0.26
HD 122563	–2.67	–2.76	–2.76	–3.00	–2.75	–2.84	–2.76	1.3	0.11
HD 140283	–2.51	–2.53	–2.44	–2.50	–2.55	–2.54	–2.57	1.3	0.20
HD 220009	–0.82	–0.77	–0.70	–0.75	–0.79	–0.83	–0.79	1.3	0.14
HD 22879	–0.88	–0.87	–0.91	–1.00	–0.95	–0.83	–0.89	1.2	0.19
HD 49933	–0.43	–0.42	–0.43	–0.50	–0.62	–0.39	–0.49	1.9	0.35
HD 84937	–2.22	–2.15	–2.15	–2.00	–2.23	–2.21	–2.21	1.5	0.24
ξ Hya	-0.01	+0.08	+0.10	+0.00	+0.19	+0.06	+0.30	1.1	0.32
μ Ara	+0.36	+0.34	+0.31	+0.25	+0.26	+0.28	+0.32	1.2	0.13
μ CasA	–0.86	–0.82	–0.82	–1.00	–0.89	–0.78	–0.88	1.1	0.29
μ Leo	+0.37	+0.39	+0.31	+0.25	+0.50	+0.23	+0.34	1.1	0.26
Procyon	+0.03	–0.03	–0.08	+0.00	–0.06	–0.01	–0.06	1.8	0.11
ψ Phe	–0.65	–0.57	–0.42	+0.00	–0.40	–0.47	–	1.5	0.33
Sun	+0.03	+0.04	–0.06	+0.00	–0.02	–0.01	–0.03	1.2	0.18
τ Cet	–0.51	–0.49	–0.49	–0.75	–0.56	–0.49	–0.56	1.1	0.28

Notes. The last two columns indicate the mean value for the microturbulence parameter obtained by each method and the standard deviation of this mean.

has a literature value that was determined from photographic plates (Luck 1979) and is thus uncertain. A similar behavior can be seen in Fig. 1 with the values reported in the literature, where [Fe/H] of cold stars present more scatter than hot stars. Obtaining a good agreement in [Fe/H] for cool stars is more difficult than for warm stars, which is mainly due to line crowding and the presence of molecules in the spectra of very cool stars. This means that the iron lines in most of the cases are not well recognized nor well modeled. Moreover, absorption lines in cold stars can be very strong, making the continuum normalization procedure extremely challenging. Also, 3D effects can become important in giants (e.g. Collet et al. 2007; Chiavassa et al. 2010), and our models consider only 1D.

For some stars, like β Ara, 61 Cyg A and B, Gmb 1830, and HD 122563, we obtain a fair agreement in metallicity. The mean value, however, differs significantly from the mean literature value. In Sect.2, we discussed how the [Fe/H] from the different works can differ significantly due to inhomogeneities between the different works. A more detailed discussion of each star, especially those with significant discrepancies compared to the mean literature value, can be found in Sect. 6.2.

When using 1D static models to determine parameters, we need to employ additional broadening parameters (micro- and macroturbulence velocity), which represent the nonthermal motions in the photosphere. Since these motions are not described in 1D static atmosphere models, broadening parameters become important to compensate for the effects of these motions. Figure 3 shows the correlation between [Fe/H] and $v_{\rm mic}$ for the Bologna, LUMBA, ULB, and Porto methods. Nissen (1981) made an analysis of $v_{\rm mic}$ as a function of [Fe/H], $T_{\rm eff}$, and $\log g$ for solar-type dwarfs by obtaining a relation where $v_{\rm mic}$ increases as a function of $T_{\rm eff}$, which agrees with our results of $v_{\rm mic}$ that is shown in Fig. 3 for warm stars ($T_{\rm eff} \geq 5000$ K). This effect has also been noticed in Luck & Heiter (2005) and Bruntt et al. (2012). Metal-poor stars are outliers of the smooth relation, with HD 140283 being the most evident one. These metal-poor stars were not included in the samples of Nissen (1981) and Bruntt et al. (2012). The microturbulence velocity decreases as function of $T_{\rm eff}$ for stars cooler than $T_{\rm eff} \sim 5000$ K, although with a larger scatter than for warm stars. This general behavior agrees with the GES relation (see Sect. 4.3), which is plotted with black dots in Fig. 3.

Although each method shows the same behavior of $v_{\rm mic}$ as a function of temperature, the absolute value of $v_{\rm mic}$ differs. The differences found between methods in $v_{\rm mic}$ help to achieve a better general agreement of [Fe/H].

5.2. Comparison of different resolutions

In Fig. 4, we plot the comparison of the results from LUMBA and UCM obtained for [Fe/H] when considering the 70 k and original library.

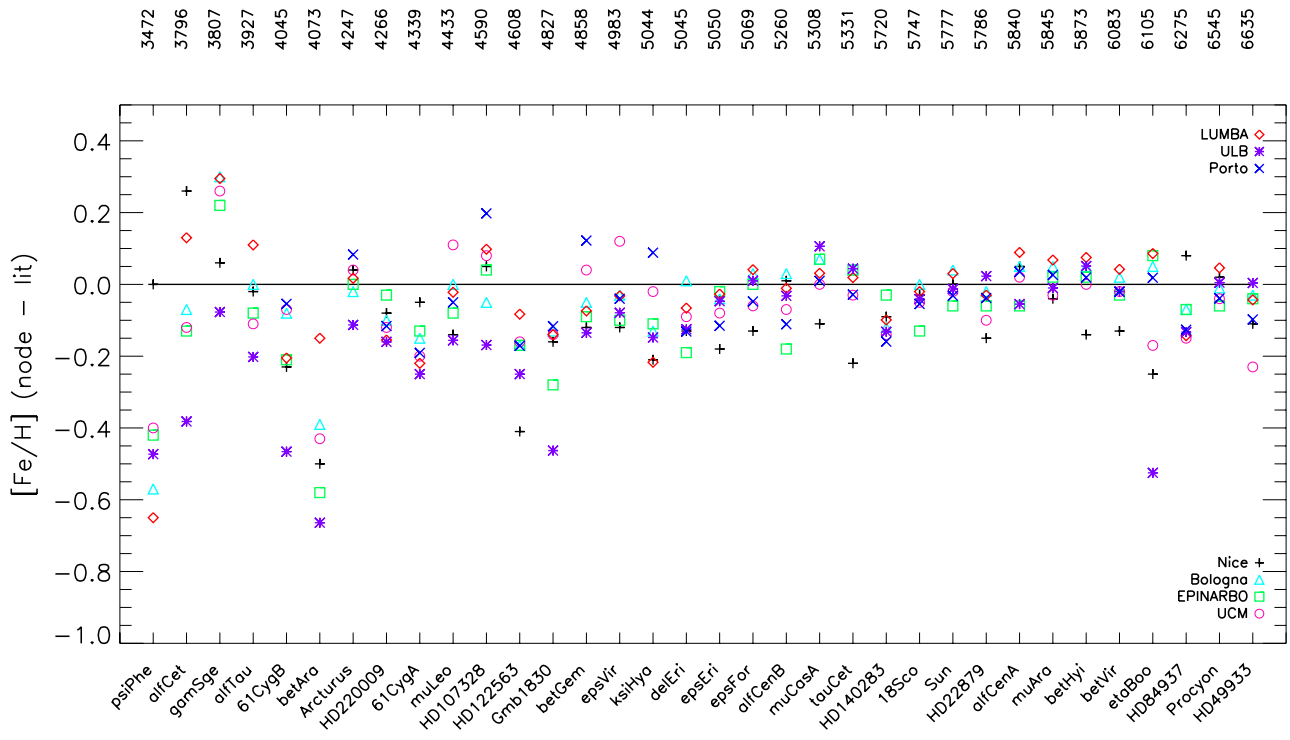

Fig. 2. Difference between the metallicity obtained by each node and the mean literature value (see Sect. 2). Stars are ordered by effective temperature. Different symbols correspond to the different methods, which are indicated in the legend.

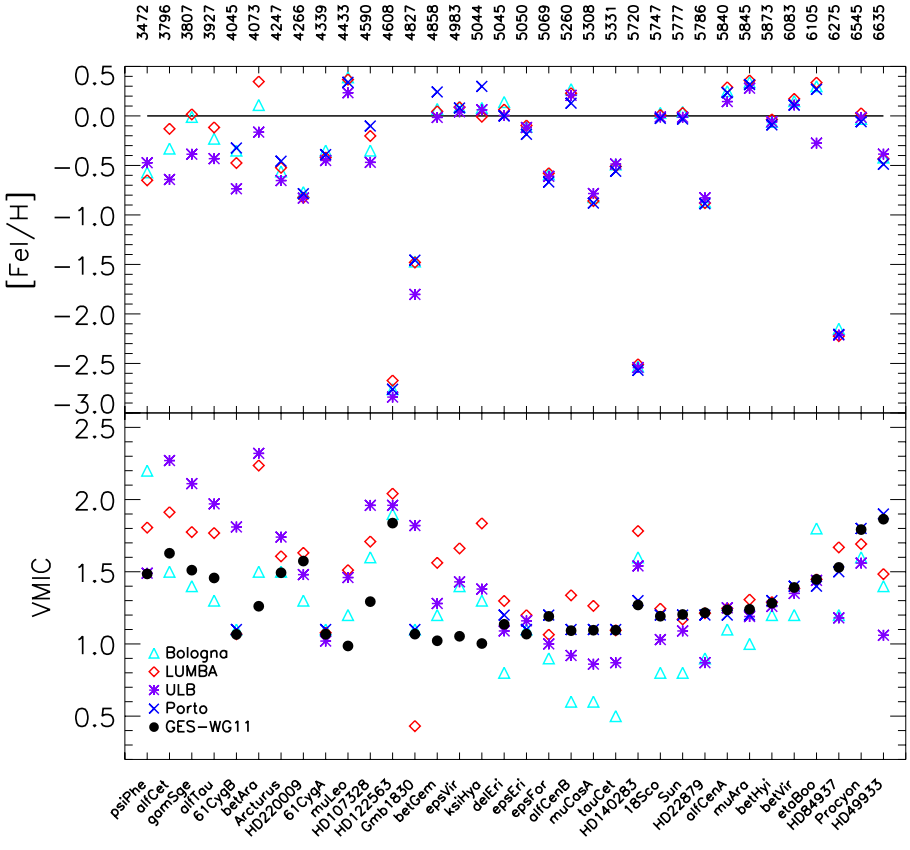

Fig. 3. Metallicity (*upper panel*) and microturbulence velocity (*lower panel*) obtained by different methods for each GBS as a function of temperature. Black dots correspond to the values of $v_{\rm mic}$, as obtained from the GES relation of Bergemann and Hill.

As in previous figures, we illustrate the difference in metallicity as a function of GBS in order of increasing temperature in the upper panel. In the lower panel of Fig. 4, we plotted together the stars observed with the same instrument. Different instruments are separated by the dashed line. The value of the spectral resolution before convolution is indicated at the top of the figure.

It is interesting to comment on the result of ψ Phe, which has the lowest original resolution and is the coldest star, because it shows the greatest difference. In the case of the LUMBA method, the synthetic spectra produced by SME need to have a given resolving power, which is set to be constant along the entire spectral range. In the original spectra, this is not completely

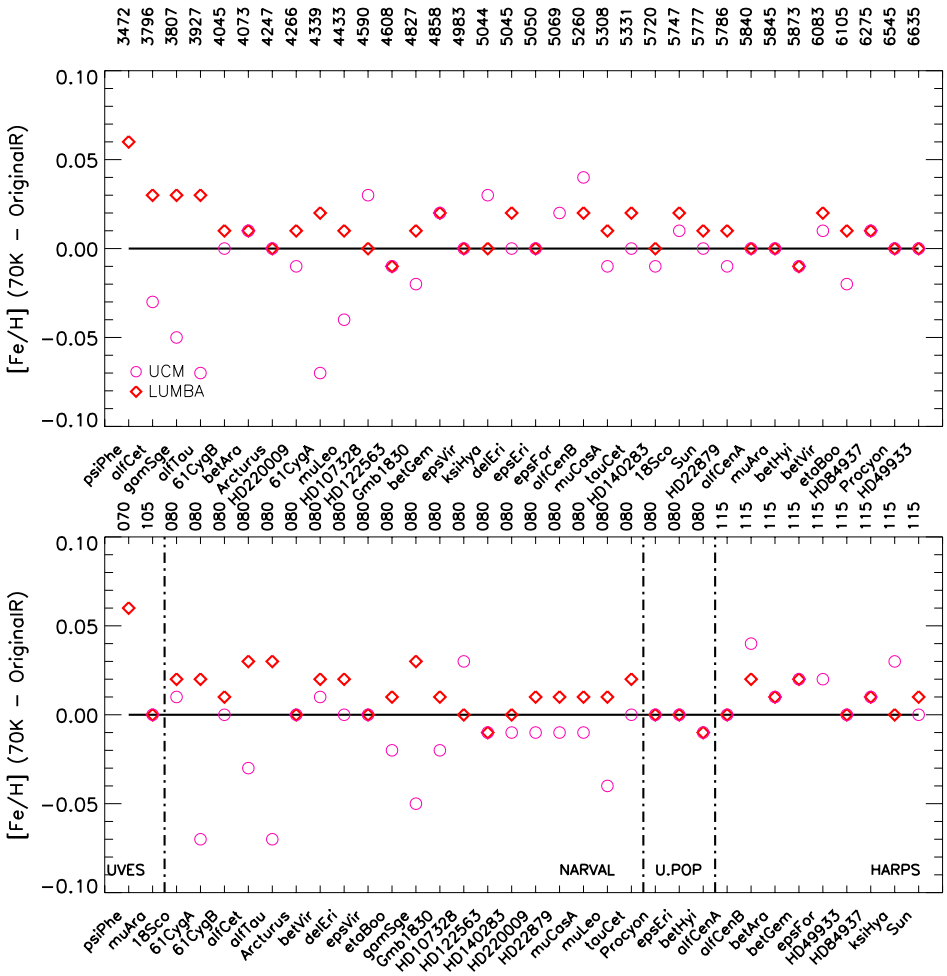

Fig. 4. Difference of metallicity obtained from 70 k and original library for UCM and LUMBA methods. *Upper panel*: difference as a function of GBS temperature. *Lower panel*: difference for stars of same instrument.

true. In this particular case, the upper part of the CCD of the UVES spectrum has a resolution that is lower than 70 000 (see Paper II). In any case, the difference is of about 0.06 dex, which is negligible compared to the uncertainty obtained for this star of about 0.5 dex (see Table 3 and Sect. 6).

The same can happen for the results from the original NARVAL spectra, which we assume to be $R = 80\,000$. As discussed in Paper II, the resolving power of NARVAL might not be exactly 80 000, but it is acceptable to initially assume a constant resolving power of $R = 80\,000$ for all the original spectra for creating the 70 k library. However, when directly analyzing the original spectra with SME, wavelength-dependent deviations from the constant input resolution might affect the results, explaining the scatter around the zero line observed in Fig. 5 for NARVAL spectra. A discussion of the impact of parameters when the exact resolution of spectra is not given can also be found in Wu et al. (2011). The UVES-POP spectra, on the other hand, are well defined in resolving power, and our results agree very well. Finally, the HARPS spectra also have a quite well-established original resolution. It is also the highest resolution of our sample.

It is worth it to comment on the results obtained by UCM for cool stars, where the difference between the original and convolved spectra are larger than for warm stars. This effect can be attributed to the contribution of lines other than Fe that can be better resolved at higher resolution, producing a slightly different measurement of the EW. In general, the differences of less than 0.03 dex are present for both methods when using different resolutions (and S/N), which is within the errors obtained in the abundances (see Sect. 6).

5.3. Comparison of different instruments

For many of the GBS, we have more than one observation. We expect our results to be consistent under different instruments. For that reason, we determined [Fe/H] for each spectrum in the complete 70 k library separately and compared them. The results obtained for the methods of Nice, Bologna, EPINARBO, UCM, and LUMBA are displayed in Fig. 5. The figures present the value of the metallicity as a function of GBS with increasing temperature.

There is a general good agreement when different spectra are analyzed for the same star. Procyon, which has observations in every instrument from our library, agrees well for each method considered here. On general, our results and data are consistent because we do not find a signature of one particular instrument giving systematic differences. In the same way, we do not find the result of one particular star being biased towards one observation. This comparison also shows that the data reduction software of the spectrographs perform correctly.

5.4. Self consistency and ionization balance

Usually, when determining parameters, $T_{\rm eff}$, $\log g$, $v_{\rm mic}$ (and $v_{\rm mac}$ in case of synthetic spectra) and [Fe/H] must be chosen such that the iron abundance obtained from neutral lines agrees with that obtained from ionized lines, which is the so-called ionization balance. Corresponding constraints are used to find the best $T_{\rm eff}$ (a flat trend of Fe I with excitation potential) and $v_{\rm mic}$ (a flat trend of Fe I with EW).

Table 3. Final metallicity of GBSs obtained via combination of individual line abundances of neutral lines corrected by NLTE effects.

Star	[Fe/H]	σ Fe I	Δ ($T_{\rm eff}$)	Δ (log g)	Δ ($v_{\rm mic}$)	Δ (LTE)	Δ (ion)	σ Fe II	N Fe I	N Fe II
Metal-Poor										
HD 122563	–2.64	0.01	0.02	0.00	0.01	+0.10	–0.19	0.03	60	4
HD 140283	–2.36	0.02	0.04	0.02	0.00	+0.07	+0.04	0.04	23	2
HD 84937	–2.03	0.02	0.04	0.02	0.01	+0.06	–0.01	–	20	1
FG dwarfs										
δ Eri	+0.06	0.01	0.00	0.00	0.01	+0.00	+0.04	0.02	156	11
ϵ For	–0.60	0.01	0.01	0.00	0.00	+0.02	+0.09	0.02	148	8
α Cen B	+0.22	0.01	0.01	0.00	0.02	+0.00	+0.09	0.02	147	9
μ Cas	–0.81	0.01	0.01	0.01	0.01	+0.01	+0.01	0.02	145	7
τ Cet	–0.49	0.01	0.00	0.00	0.00	+0.01	+0.01	0.02	148	10
18 Sco	+0.03	0.01	0.01	0.00	0.01	+0.02	+0.00	0.02	158	10
Sun	+0.03	0.01	0.00	0.00	0.00	+0.01	+0.04	0.02	150	9
HD 22879	–0.86	0.01	0.03	0.01	0.01	+0.02	–0.02	0.02	117	10
α Cen A	+0.26	0.01	0.01	0.00	0.00	+0.02	+0.07	0.02	150	12
μ Ara	+0.35	0.01	0.00	0.00	0.00	+0.02	+0.13	0.02	143	13
β Hyi	–0.04	0.01	0.01	0.00	0.00	+0.03	+0.05	0.01	143	12
β Vir	+0.24	0.01	0.01	0.00	0.01	+0.03	+0.06	0.02	148	10
η Boo	+0.32	0.01	0.00	0.00	0.01	+0.02	+0.07	0.03	127	10
Procyon	+0.01	0.01	0.01	0.00	0.00	+0.05	–0.06	0.02	135	12
HD 49933	–0.41	0.01	0.04	0.02	0.02	+0.05	–0.03	0.02	93	6
FGK giants										
Arcturus	–0.52	0.01	0.00	0.00	0.06	+0.01	+0.02	0.04	151	10
HD 220009	–0.74	0.01	0.01	0.00	0.07	+0.01	+0.10	0.03	148	11
μ Leo	+0.25	0.02	0.00	0.00	0.13	–0.01	+0.01	0.08	139	11
HD 107328	–0.33	0.01	0.01	0.00	0.16	+0.01	+0.02	0.03	137	11
β Gem	+0.13	0.01	0.01	0.00	0.13	+0.01	+0.09	0.03	146	13
ϵ Vir	+0.15	0.01	0.02	0.00	0.15	+0.02	–0.03	0.03	139	12
ξ Hya	+0.16	0.01	0.01	0.00	0.17	+0.02	+0.10	0.03	151	11
M giants										
ψ Phe	–1.24	0.24	0.05	0.03	0.30	–0.01	–	–	23	0
α Cet	–0.45	0.05	0.17	0.08	0.34	+0.00	–0.20	0.17	35	3
γ Sge	–0.17	0.04	0.13	0.09	0.22	–0.01	–0.25	0.12	29	4
α Tau	–0.37	0.02	0.02	0.02	0.12	+0.00	+0.06	0.10	76	9
β Ara	–0.05	0.04	0.01	0.04	0.16	+0.00	–0.34	0.08	62	8
K dwarfs										
61 Cyg B	–0.38	0.03	0.01	0.01	0.01	+0.00	–	–	119	2
61 Cyg A	–0.33	0.02	0.00	0.01	0.00	+0.00	–0.29	0.25	138	3
Gmb 1830	–1.46	0.01	0.05	0.03	0.30	+0.00	–0.22	0.10	116	4
ϵ Eri	–0.09	0.01	0.00	0.00	0.00	+0.01	–0.05	0.02	153	11

Notes. The metallicity is associated with different sources or errors: standard deviation of the line-by-line abundance of the selected Fe I lines (σ Fe I); errors due to the uncertainty in $T_{\rm eff}$, log g and $v_{\rm mic}$, (Δ ($T_{\rm eff}$), Δ (log g), Δ ($v_{\rm mic}$), respectively); error due to difference between NLTE and LTE Fe I abundance (Δ (LTE)); error due to difference between Fe I and Fe II abundance Δ (ion); and standard deviation of the line-by-line mean of Fe II abundance (σ Fe II). The last two columns indicate the number of selected lines used for the determination of Fe I and Fe II abundances, respectively.

Since we do not change $T_{\rm eff}$ and log g in this particular work, the simultaneous determination of the other parameters becomes the dominant means for approaching ionization and excitation balance. For methods based on EWs, $v_{\rm mic}$ helps to obtain abundances in a line-to-line approach that does not depend on the reduced EW or wavelength range. For methods based on synthetic spectra, $v_{\rm mic}$ and $v_{\rm mac}$ are treated as broadening parameters that help to improve the fit of the synthesis to observed line profiles.

Since $T_{\rm eff}$ and log g are taken from fundamental relations and are independent of spectral modeling, ionization balance and the mentioned relations tell us how well our models are able to reproduce our observations. Figure 6 displays the iron content obtained from neutral and ionized lines for the GBS using EPINARBO, UCM, Bologna and LUMBA methods. The stars have been plotted with increasing temperature, and each symbol represents one method. Open and filled symbols indicate Fe I and Fe II abundances, respectively.

Generally, all nodes show a significant difference between Fe I and Fe II abundances for HD 122563, Gmb 1830, and μ Ara. For other cases, such as β Gem, only some methods show large differences while others show an agreement. Cool stars like α Tau or α Cet are also problematic because the available Fe II lines are often blended by molecules, and it becomes difficult to model them with our current theoretical input data. It was impossible to create a Fe II line mask for ψ Phe when analyzed with the LUMBA method. The Fe II abundances obtained for the coolest stars by any method can thus be unreliable. To be able to obtain reliable Fe II abundances for such stars, the synthesis methods would need to have a list of molecules capable of reproducing those blends.

Figure 7 shows the trends of the iron abundance as a function of EW and excitation potential for the Sun (a good case) and HD 122563 (an unbalanced case), as obtained by the Bologna node (see also Sect. 4.3.4). Black and red dots correspond to neutral and ionized iron abundances, respectively. The figure shows

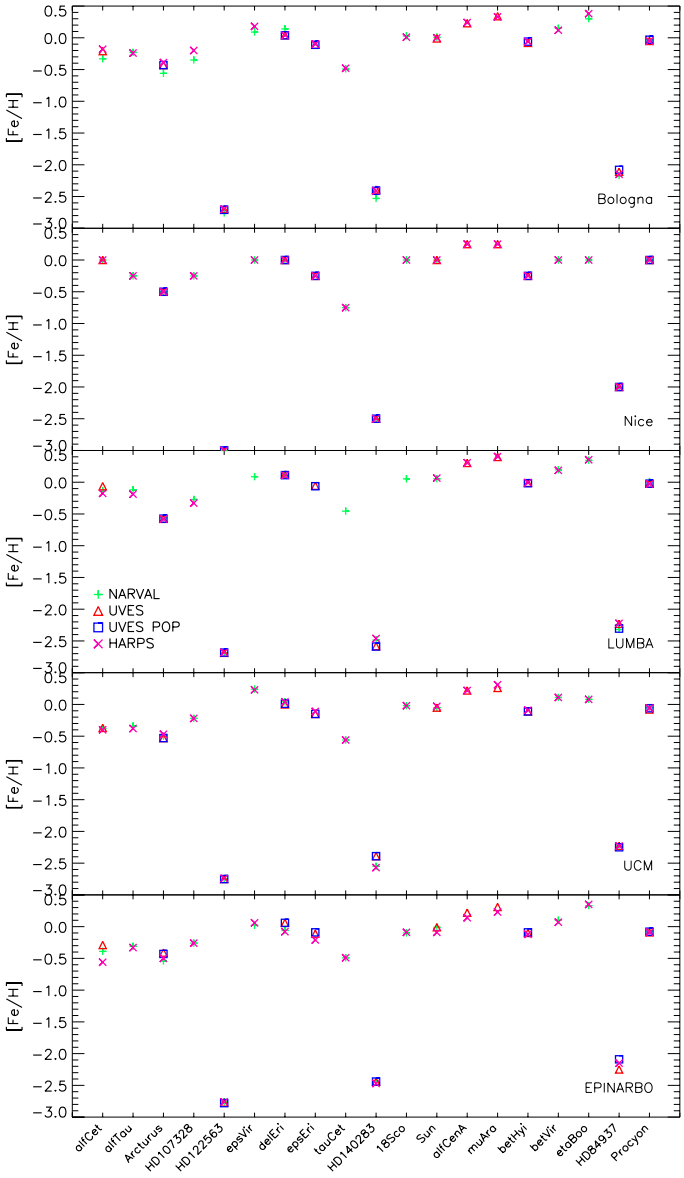

Fig. 5. Metallicity of GBS as a function of effective temperature. Symbols represent different instruments (see legend). Each panel shows the result of one method, as indicated in each panel.

that a perceptible difference between Fe I and Fe II abundances results when using $\log g$ from Table 1 and also a trend of iron abundance with excitation potential appears when using the $T_{\rm eff}$ from the same table. If the parameters were let free, as in the traditional EW-based method, both gravity and temperature would have to be re-adjusted to obtain self-consistent results.

Even in the good cases, where the abundances of neutral and ionized iron are well determined, a small difference between the two can appear and it is often difficult to reconcile Fe I and Fe II abundances. In their attempt to review the fundamental parameters of Arcturus with a method very similar to the one presented in this work, Ramírez & Allende Prieto (2011) obtained a difference of 0.12 dex between Fe I and Fe II abundances. This is explained as a limitation of the 1D-LTE models, which cannot reproduce the data well enough. Similarly, Schuler et al. (2003) reported problems in their analysis of the open cluster M 34, where $T_{\rm eff}$ and $\log g$ were kept fixed to values obtained from the color-magnitude diagram and the final iron abundance from ionized and neutral Fe lines did not fully satisfy ionization balance, especially in the case of the coldest K dwarfs. An extensive discussion on this subject can be found in Allende Prieto et al. (2004), who analyzed field stars in the solar neighborhood. Their Fig. 8 shows the differences obtained from neutral and ionized lines of iron and calcium, where differences can reach 0.5 dex in the most metal-rich cases. They argue that dramatic modifications of the stellar parameters are necessary to satisfy ionization balance, which would be translated to unphysical values. All aforementioned works explain this effect as due to departures from LTE, surface granulations, incomplete opacities, chromospheric and magnetic activity, and so on. For an extensive discussion on this issue for five of our GBS (the Sun, Procyon, HD 122563, HD 140283, HD 84937, and HD 122563), see also Bergemann et al. (2012).

We performed an additional abundance analysis by simultaneously determining $T_{\rm eff}$ and $\log g$, [Fe/H], and $v_{\rm mic}$ on the 70 k library. Our idea was to quantify the amount that $T_{\rm eff}$ and $\log g$ be altered in order to obtain excitation and ionization balance in each method. The results of this "free" analysis are illustrated in Fig. 8, where the difference between the "fixed" (determination of [Fe/H] via fixing $T_{\rm eff}$ and $\log g$) and the "free" analysis are shown for each GBS. Metallicity, temperature, and surface gravity are plotted in the upper, middle, and lower panels of Fig. 8, respectively.

As expected, the metallicity obtained when forcing ionization equilibrium for 1D LTE models is different from that obtained with the fundamental $T_{\rm eff}$ and $\log g$. The median difference in metallicity for solar-type stars is smaller than for the coldest, hottest, and metal-poor stars. The differences obtained are usually related to larger deviations in $T_{\rm eff}$ and $\log g$ from the fundamental value, as seen in Fig. 8 and discussed in Allende Prieto et al. (2004) and Ramírez & Allende Prieto (2011). In Gmb 1830, for example, the results of $T_{\rm eff}$ and $\log g$ from the free spectral analysis agree more with what has been reported in PASTEL (Soubiran et al. 2010), which is more than 250 K above the fundamental value. The object HD 140283 is another case where the free temperature and surface gravity are 200 K and 0.7 dex smaller than the fundamental value, resulting in a [Fe/H] that is ∼0.2 dex more metal-poor than the fixed case. On the other hand, the smallest differences in [Fe/H] are related to small deviations in $T_{\rm eff}$ and $\log g$. Examples of this cases are μ Cas A, α Cen A, α Cen B, and the Sun.

In general, when looking at the results of individual methods, a difference of up to 200 K in $T_{\rm eff}$ and 0.25 dex in $\log g$ would be necessary to restore excitation and ionization balance in the problematic GBS. This would introduce a change of ∼0.1 dex in metallicity as well. It is important to comment that this test is just an illustration of the effects of freeing $T_{\rm eff}$ and $\log g$ to retrieve ionization balance but does not represent the real performance of the different methods when determining three parameters. Here, we are only concentrating in the analysis of iron lines and not the analysis of other important spectral features that can affect the determination of $T_{\rm eff}$ and $\log g$. This can have important consequences for methods based on SME or MATISSE, for example. A full explanation of the performance of the methods in the parametrization of UVES spectra will be found in Smiljanic et al. (in prep.).

5.5. NLTE corrections

Recently, Bergemann et al. (2012) presented a thorough investigation of the Fe I-Fe II ionization balance in five of the GBS included here (Sun, Procyon, HD 122563, HD 84937, and

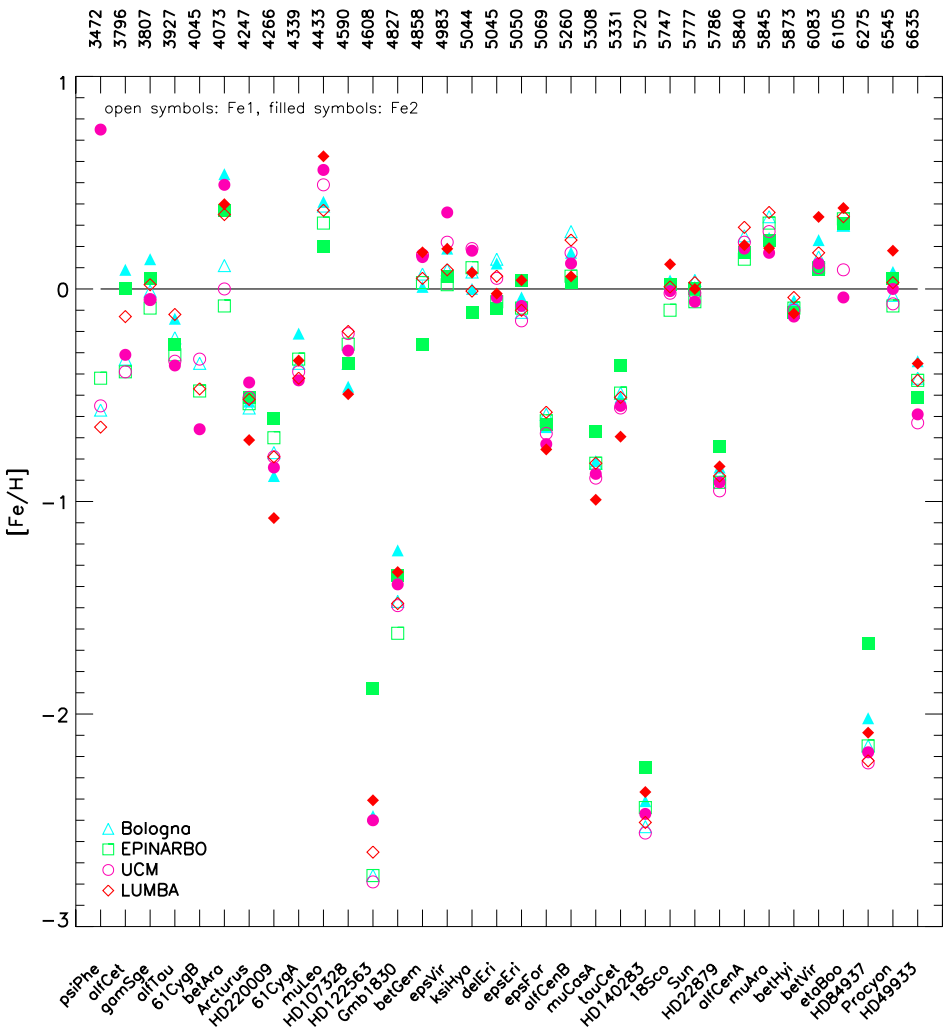

Fig. 6. Neutral and ionized iron abundances obtained for GBS as a function of effective temperature by different methods (see legend). Open symbols represent Fe I abundances, while filled symbols represent Fe II abundances.

HD 140283) and one more extremely metal-poor star (G64-12). In particular, they utilized an extensive Fe model atom and both traditional 1D and spatially and temporally averaged 3D hydrodynamical models to assess the magnitude of NLTE effects on Fe line formation. Bergemann et al. (2012) concluded that only very minor NLTE effects are needed to establish ionization balance at solar metallicities, while very metal-poor stars imply effects on the order of +0.1 dex on Fe I lines. The Fe II lines are well modeled everywhere by the LTE assumption.

The NLTE calculations were extended by Lind et al. (2012) to cover a large cool star parameter space. Here, we interpolated within the grid of NLTE corrections by Lind et al. (2012) the stellar parameters adopted for each GBS as taken from Table 1. Each Fe line used in the final [Fe/H] determination was corrected individually. When an NLTE correction was not available for a specific line, we used the median of the corrections computed for all other lines. This is possible to do as the corrections for all lines of a particular star are very similar, as shown by Bergemann et al. (2012). The difference between the final Fe abundances for single and ionized lines is visualized in Fig. 9 for each star (see Sect. 6 for details of how the final abundances are determined). The stars are plotted in order of increasing effective temperature. Black indicates that the iron abundance is determined from Fe I lines while red indicates that the abundance is determined from Fe II lines. Dots and square symbols indicate the LTE and NLTE abundances, respectively. The error bars are plotted only for the LTE abundances, as they do not change after NLTE corrections. The errors considered in this plot correspond to the sum of the scatter found for the line-by-line abundance determination and the errors obtained by considering the associated uncertainties in the fundamental parameters (see Sect. 6 for details).

In general, NLTE corrections can vary between −0.10 to +0.15 dex for individual lines, but the departures of NLTE affect the metallicity by <0.05 dex for all stars on average. Exceptions are the hottest stars and the most metal-poor ones, which can differ up to 0.1 dex. Since the corrections due to NLTE effects are small, even when looking at the final NLTE abundances in Fig. 9, we still find cases where ionization imbalance is significant, especially for the cold stars. We conclude that neglecting NLTE effects is not a likely explanation for the ionization imbalance.

6. The metallicity determination

Since each method and corresponding criterium gives a final [Fe/H] value, we combine our results by looking at individual abundances in a line-by-line approach. Since the Nice method is based on a global fitting of a whole section of the spectrum, abundances of individual lines for that method are not provided. We note that the setup employed by the LUMBA node for this analysis performed a simultaneous fit of all pixels contained in the specified line mask, and thus it did not provide abundances of individual lines per se. However, LUMBA employed a post-processing code that determined best-fit $\log gf$ values for each line. This is equivalent to determining best-fit abundances. The resulting $\log gf$ deviation from the nominal value is then added

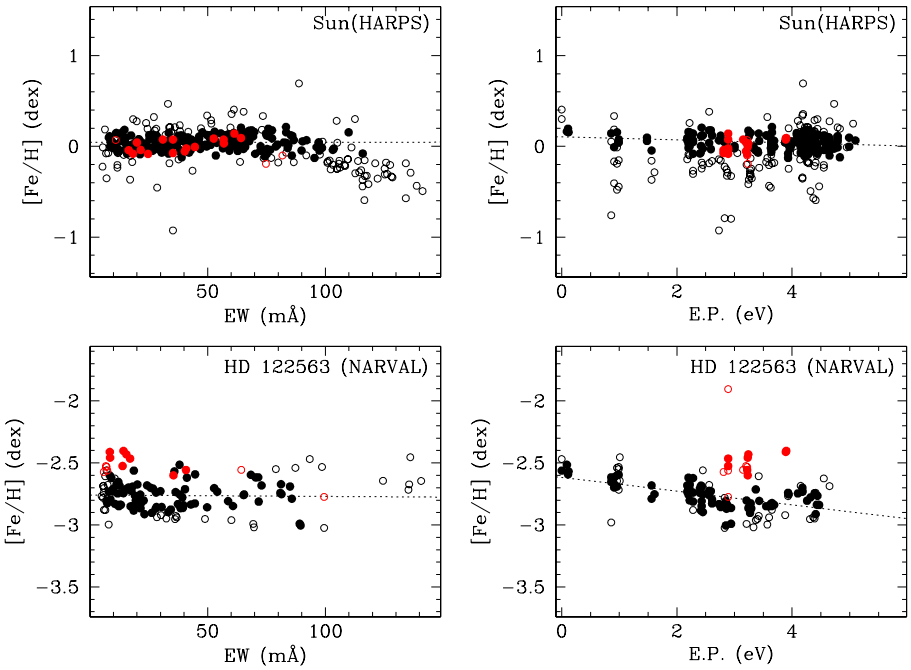

Fig. 7. GALA outputs of the Bologna method for the Sun (HARPS, *upper panels*) and HD 122563 (NARVAL, *lower panels*) for the *run-nodes* test. In all panels, black symbols refer to Fe I and red ones to Fe II, while empty symbols refer to rejected lines (see Sect. 4.3.4) and solid ones to lines effectively used for the analysis. A dotted line shows the result of a linear fit to the used Fe I lines in all panels.

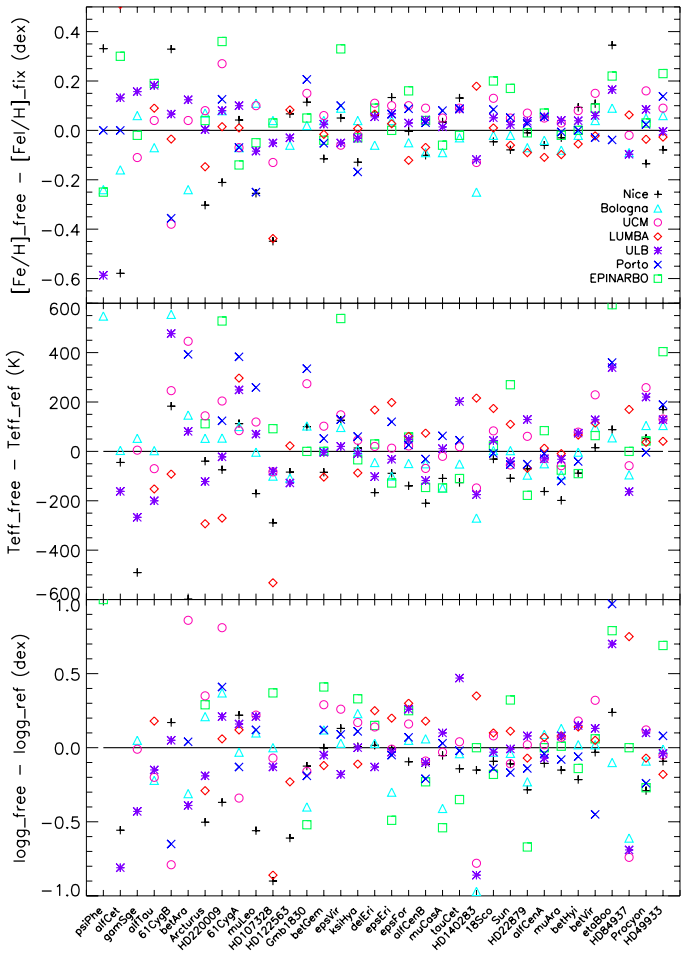

Fig. 8. Difference in metallicity (*upper panel*), effective temperature (*middle panel*) and surface gravity (*lower panel*) of GBS as obtained by different methods between free and fixed analysis (see text).

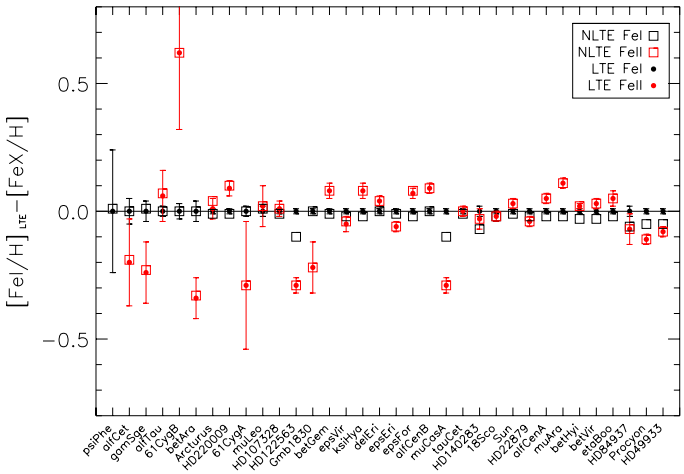

Fig. 9. Difference of final [Fe I/H](black) and [Fe II/H] (red) for each GBS. Squares show the abundances after NLTE corrections. Error bars represent the uncertainties coming from the line-to-line scatter and the uncertainties coming from the associated uncertainties in $T_{\rm eff}$, $\log g$ and $v_{\rm mic}$ (Sect. 6).

to the global metallicity of each star derived by SME in order to reconstruct individual line abundances.

We performed several steps to combine and determine the metallicity of each star. This analysis was mostly carried out by P. Jofré, U. Heiter, J. Sobeck, and K. Lind.

First, we selected the lines with $\log(EW/\lambda) \leq -4.8$. The objective was to use lines, which are on the linear part of the curve of growth, to avoid saturated lines and to mitigate the effect of "wrong microturbulence" and "wrong damping parameters", which affects strong lines. The transition from the linear part to the saturated part of the curve of growth occur at $\log(EW/\lambda) \sim -5.0$, which is more or less independent of stellar parameters (see e.g. Figs. 16.1 to 16.6 of Gray 2005, or Villada & Rossi 1987). The transition point is slightly above for cool models, while slightly below for hot models. In addition, the transition value was checked for each GBS by constructing empirical curves of growth from the output of the Bologna method. For the different kind of stars presented here, the limit

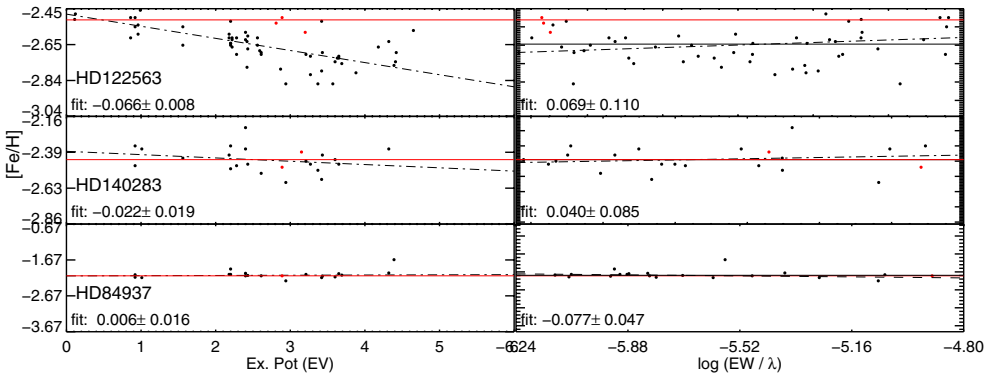

Fig. 10. Trends of abundances as a function of excitation potential (*left panels*) and reduced EW (*right panels*) in the group of metal-poor stars.

of −4.8 seems to be a good compromise between the number of lines and the saturation criterion.

Second, we calculated the mean and standard deviation of all abundances and selected those lines that were analyzed by at least three different groups and for which the values agreed within 2σ with the mean abundance.

Third, we calculated the mean abundance from the different methods for each selected line. For consistency checks on metallicities, each abundance was plotted as a function of wavelength, EW, and excitation potential (E.P.) to account for excitation balance. The relations can be found in Figs. 10–14. Additionally, NLTE corrections were applied individually for each selected line and star (see Sect. 5.5). An extensive discussion is found in Sect. 6.2.

Finally, we computed the final value of Fe I and Fe II abundances from the average of the selected lines. To compute the final metallicity, we considered the value of 7.45 for the absolute solar iron abundance from Grevesse et al. (2007). The final value of [Fe/H] obtained from Fe I lines after corrections by NLTE effects is listed in the second column of Table 3. The third column indicates the standard deviation of the abundances obtained from the selected Fe I lines. The list of lines selected for each star can be found as part of the online material.

6.1. Errors due to uncertainties in $T_{\rm eff}$, log g, and $v_{\rm mic}$

We are basing our analysis on fixed values for $T_{\rm eff}$ and log g, but these values have associated errors that give the metallicity an additional uncertainty. In a similar manner, we want to study the effect on the final metallicity due to the uncertainties in the $v_{\rm mic}$ parameter. To quantify the error of [Fe/H] due to the associated errors in $T_{\rm eff}$, log g, and $v_{\rm mic}$, we performed additional runs determining the iron abundances using the same setup as described for *run-nodes* in Sect. 4 but change the input value of $T_{\rm eff}$, log g, and $v_{\rm mic}$ by considering $T_{\rm eff} \pm \Delta T_{\rm eff}$, log $g \pm \Delta \log g$ and $v_{\rm mic} \pm \Delta v_{\rm mic}$, respectively. The values of $\Delta T_{\rm eff}$ and $\Delta \log g$ can be found in Table 1 and were determined in Paper I, while we considered the scatter found by the different nodes from the standard *run-nodes* for the value of $\Delta v_{\rm mic}$, which can be found in the last column of Table 2.

This analysis gave us six additional runs, which were performed by the methods LUMBA, EPINARBO, Porto, UBL, and UCM. To be consistent with our main results, we determined the iron abundance of only the lines that passed the selection criteria after the main run. The final differences of $([{\rm Fe/H}]_{\Delta^-} - [{\rm Fe/H}]_{\Delta^+})$, where $[{\rm Fe/H}]_{\Delta^\pm}$ correspond to the metallicities obtained, considers the parameters of their errors for $T_{\rm eff}$, log g, and $v_{\rm mic}$ respectively. These values are also listed in Table 3 for each star.

6.2. Discussion

To understand better our results, we divided the stars into five groups: metal-poor stars, FG dwarfs, FGK giants, M giants, and K dwarfs. Each group is discussed separately in the following sections.

6.2.1. Metal-poor stars

This group includes the stars HD 122563, HD 140283, and HD 84937. Our results agree well with an internal scatter in a line-by-line approach of about 0.12 dex before the line selection process described in Sect. 6. A similar differential analysis between the results obtained for atmospheric parameters from EWs and synthetic spectra on high resolution spectra of metal-poor stars was done by Jofré et al. (2010). In that study, 35 turn-off metal-poor stars were analyzed using the same data and line list and different atmosphere models. The general scatter was 0.13 dex in metallicity when log g and $T_{\rm eff}$ were forced to agree by 0.1 dex and 100 K, respectively. Although here we determine only metallicity, it is encouraging to obtain a mean scatter of 0.06 dex when considering the independent results of the seven methods.

The abundances of the selected lines for each metal-poor star as a function of E.P. are shown in the left panels of Fig. 10, while the abundances as a function of reduced EW are shown in the right panels of the figure. Black dots correspond to Fe I abundances, corrected by NLTE effects as described in Sect. 5.5, while the red dots correspond to the Fe II abundances. The solid red and black horizontal lines indicate the averaged Fe II and Fe I abundance, respectively. In addition, we plotted with a dot-dashed line the linear regression fit to the Fe I abundances, where its slope and error are written in the bottom of each panel.

In metal-poor stars the continuum is easy to identify, although other difficulties appear, such as the low number of iron lines detectable in the spectra, especially those of ionized iron. In our case, the common lines that passed the selection criteria explained above can be seen in Fig. 10. The star HD 84937 is the most extreme case, where we have only 1 ionized and 20 neutral iron lines that are used for the final [Fe/H] determination.

The NLTE effects can significantly change the metallicity of metal-poor stars (Thévenin & Idiart 1999; Asplund 2005). After applying NTLE corrections to our selected LTE Fe I abundances, the metallicities increase by up to approximately 0.1 dex, which agree with the investigation of Bergemann et al. (2012) for these three GBS.

The largest difference between Fe I and Fe II abundances is for the metal-poor giant HD 122563. However, one can see a significant slope in the regression fit of −0.066 ± 0.008 in

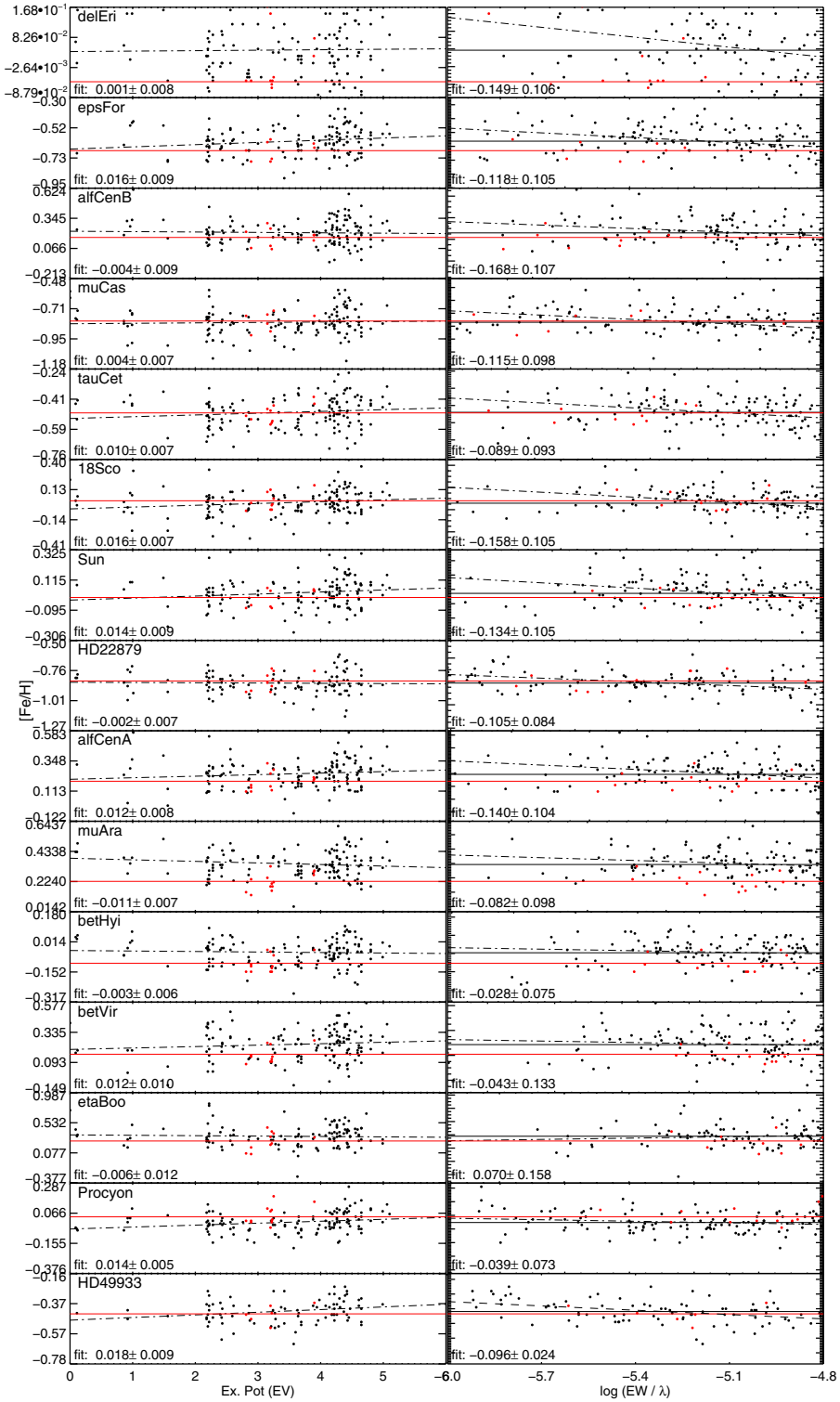

Fig. 11. Trends for group of FG dwarfs.

the distribution of Fe I lines as a function of E.P. The regression fit as a function of EW shows a slope of 0.07 that can be neglected when considering the error of 0.11. Since those fits are obtained after making the NLTE corrections, we attribute this trend to 3D effects, which are most important for cool metal-poor stars (e.g., Asplund et al. 1999; Collet et al. 2007). See also Bergemann et al. (2012) for the study in this regard of HD 122563. The second metal-poor star, HD 140283, also presents a negative slope for Fe I abundances as a function of E.P., although it is less pronounced and its error is larger than the case of HD 122563. It is interesting to see that we obtain a good ionization balance for this metal-poor subgiant. The last metal-poor star of our group, HD 84937, presents quite a flat regression fit when looking at the abundances as a function of E.P or EW and considering the errors. Moreover, Fe I and Fe II abundances agree when the errors due to $T_{\rm eff}$ and $\log g$ are taken into account.

Although one should be aware that there is a large ionization and excitation imbalance for HD 122563, we conclude that we can average the abundances and obtain robust values of

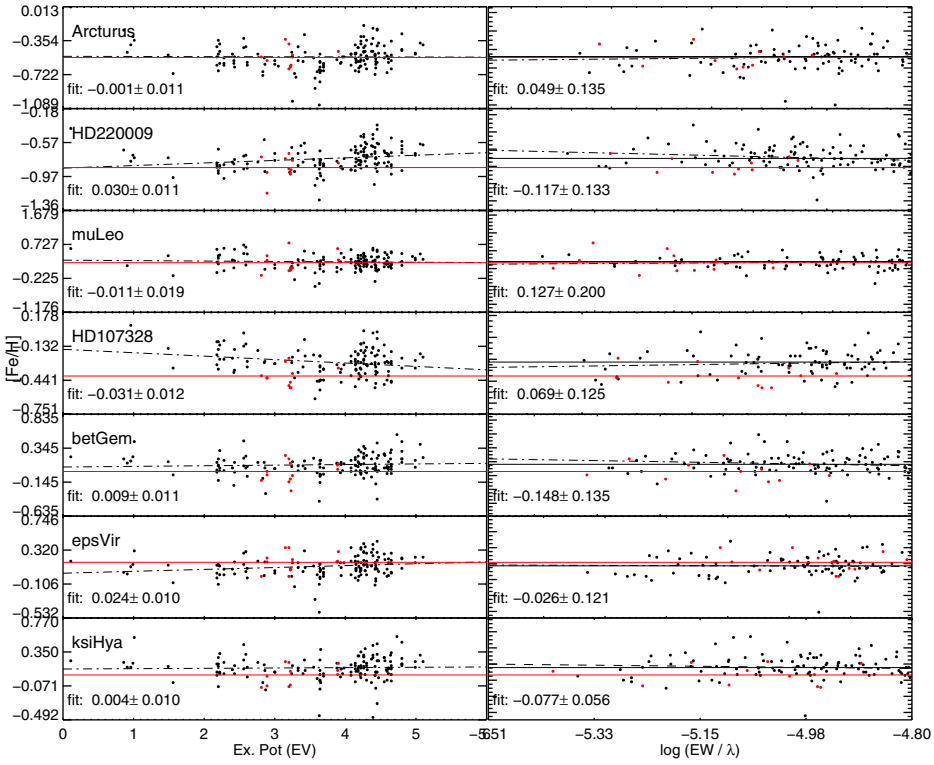

Fig. 12. Trends for group of FGK giants.

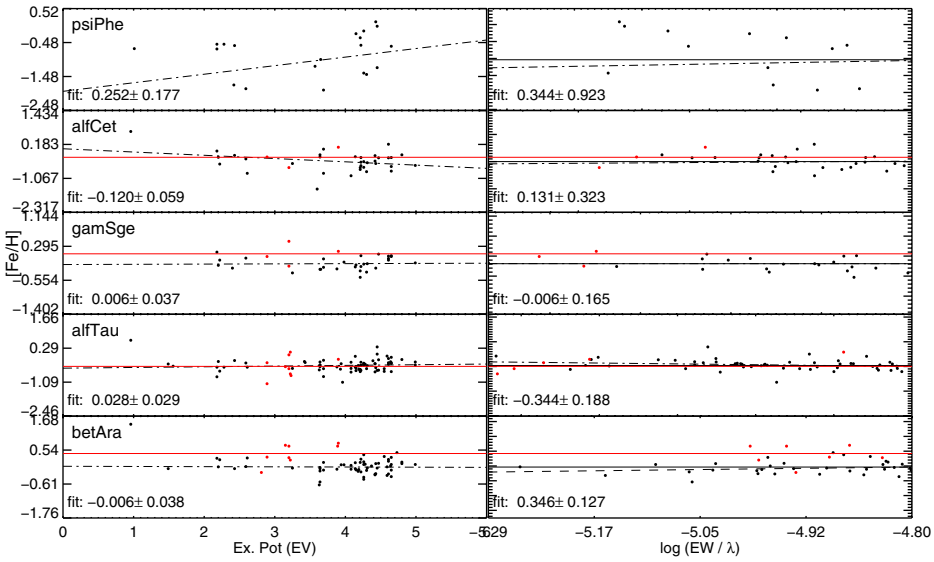

Fig. 13. Trends for group of M giants.

metallicities for metal-poor stars given their fundamental parameters and associated errors.

6.2.2. FG dwarfs

The stars δ Eri, ϵ For, α Cen A & B, μ Cas, τ Cet, 18 Sco, Sun, HD 22879, μ Ara, β Hyi, β Vir, η Boo, Procyon, and HD 49933 belong to this group. The mean internal 1σ scatter of these stars when looking at all abundances of individual lines is of 0.13 dex, while the value when looking at the results of the individual methods is of 0.07 dex. Moreover, our results agree within 0.04 dex with the literature, as seen in Fig. 2. Note that the final line-to-line scatter for these stars is reduced to ∼0.01 from the initial scatter after our selection of lines. The NLTE corrections for these stars are very small, usually less than 0.03 dex, with the exception of Procyon and HD 49933, which are of the order of 0.05 dex (see Fig. 9). These stars have high effective temperatures, which produce greater departures from LTE than cool stars (Bergemann et al. 2012).

As in the case of the metal-poor group, we have plotted the abundances of the selected lines for each star as a function of E.P. and reduced EW in Fig. 11. This group shows that our selected lines are well-behaved, in the sense that excitation and ionization balance are in general satisfied. Usually a difference between ionized and neutral iron abundances is less than 0.1 dex for this group of stars, which can be confirmed with Fig. 9. There are a few exceptions, such as the hot star Procyon, and the solar-type stars, ϵ For, α Cen B, and μ Ara. The latter presents the larger ionization imbalance, which can be explained by the rather large excitation imbalance (with a slope of -0.012 ± 0.008 dex in the regression fit as a function of E.P.). We find no significant trend as a function of $\log(EW/\lambda)$ when considering the errors of the

regression fits. Note that the hot stars Procyon and HD 49933 also present a significant excitation imbalance in the regression fits.

Recently, Torres et al. (2012) made a comparative spectral analysis of FG dwarfs using three different methods to determine parameters. Two of their methods overlap with our own, namely SME (LUMBA) and MOOG (UCM, Porto and EPINARBO). They obtained a systematic difference of 0.068 ± 0.014 dex in metallicity when analyzing 31 stars with these two methods, which is attributed to the different $T_{\rm eff}$ and $\log g$ obtained from the simultaneous analysis, the different way of placing the continuum, and the different lines used by each methods.

We conclude that it is acceptable to average the abundances of our selected lines and that we are able to provide robust results for [Fe/H] for FG dwarfs based on their fundamental temperature and surface gravity.

6.2.3. FGK giants

These are Arcturus, μ Leo, β Gem, ϵ Vir, ξ Hya, HD 220009, and HD 107328. Although the scatter between the nodes is larger than the scatter for dwarfs (see Fig. 2), it is encouraging to obtain an agreement within 0.08 dex for giants considering the different methods. The mean 1σ scatter of all iron abundances for every line is of 0.2 dex, although it is reduced to 0.08 dex when considering only the abundances of the selected lines. The FGK giants are challenging objects to model due to their complex atmospheres and large number of lines, namely, lines that form from molecules. In addition, convection in red giants becomes important, and 1D models can differ from 3D models, which impacts the final abundances, especially for metal-poor stars (Collet et al. 2007). Microturbulence becomes, therefore, a sensitive parameter, which explains the large error in $v_{\rm mic}$ of Table 2.

Typically, the NLTE departures for this group of stars are negligible when compared with the errors obtained for the abundances, which can be seen in Fig. 9. In general, an ionization imbalance of ∼0.1 dex is found for this group of stars, which agrees with the recent conclusion of Ramírez & Allende Prieto (2011). The abundances of the selected neutral and ionized iron lines for each giant are shown in Fig. 12. The dot-dashed lines correspond to the linear regression fits of the Fe I abundances as a function of E.P and $\log(EW/\lambda)$. While no significant trend of abundances as a function of reduced EW is obtained when considering the error of the fit for most of the stars, a significant positive slope in the regression fit as a function of E.P. is found. The change in abundance over the range in E.P. covered by the lines is, however, smaller than the final error. Thus, we are confident that performing a mean on the abundances of our selected lines provides robust results for the [Fe/H] of the benchmark FGK giants.

We obtain typical differences of about ±0.07 dex or less with the literature values, which is within the uncertainties and scatter found by us and by the literature. As exceptional cases, we obtain a slightly lower metallicity of 0.1 dex than the literature value for ξ Hya. The PASTEL catalog has only two works reporting parameters for this star, where McWilliam (1990) obtained [Fe/H] = −0.04 and Bruntt et al. (2010) obtained [Fe/H] = +0.23. In Table 1, we present only the latter due to the restriction on publication year for the extraction from PASTEL (see Sect. 2). Our value of [Fe/H] = 0.12 lies in between those values. For HD 220009, we obtain ∼0.14 dex lower than the literature. The only work in PASTEL after 2000 that reports [Fe/H] = −0.67 is that of Smiljanic et al. (2007). The difference can be explained from the different values for the stellar parameters considered by that work, that is, the effective temperature and surface gravity are 100 K and 0.5 dex, respectively, which is higher than the fundamental values considered by us.

Finally, we comment that we noticed that the effective temperature of HD 107328 had been overestimated by 90 K during the time when this analysis was carried out by our different groups. For that reason, we created a set of line-by-line corrections for HD 107328 to account for the lower temperature. We used the same grid as for the NLTE corrections but used only LTE curves-of-growth. The uncertainties in the metallicity due to associated errors in the other stellar parameters were then determined using the most recent temperature.

6.2.4. M giants

The analysis of this group is the most difficult one, where an averaged line-to-line scatter of 0.5 dex is obtained. It includes the stars ψ Phe, α Cet, β Ara, γ Sge, and α Tau. Note that the spectral class of α Tau is not well established (see Lebzelter et al. 2012, for a discussion), which is in the limit between late K and early M type. Since our results for α Tau are more comparable to those of the M-type than those of FGK group of giants for simplicity, we classify α Tau into the M giant group.

These cool giants have very challenging spectra, mostly because of the presence of molecules. The strength of TiO and CN absorption bands in the coldest stars is particularly high (Peterson 1976), making it extremely difficult to identify the continuum around most of the iron lines. The blends with molecules can become so dominating that an overestimation of metallicity can be obtained when using a given line which has an unidentified molecular blend (Peterson 1976).

Additionally, the efficiency of convective energy transport and its effect on line-formation reaches its maximum at $T_{\rm eff}$ ∼ 4000 K (Heiter et al. 2002). For that reason, 3D hydrodynamical models are much more suitable for modeling line-formation in such spectra. Such models for stars other than the Sun are not easily available, which is mainly due to the large computing power needed to model them. In particular, red supergiants give rise to large granules that can imprint irregular patterns (Chiavassa et al. 2009, 2010), but the influence of this effect in spectra of such cool stars has not been investigated so far. A detailed discussion on spectral modeling for cold giants can be found in Lebzelter et al. (2012). They determined atmospheric parameters of the GBS α Cet and α Tau using 11 different methods and made a comparative analysis as for this work. In their analysis (employing different line lists and atmosphere models between the methods), the unweighted mean values for metallicity were [Fe/H] = −0.2 ± 0.2 dex for both stars. We obtain a value of −0.45 for α Cet and −0.37 for α Tau, respectively. Although we obtain values that are more metal-poor, they lie within the errors.

The abundances of the selected lines can be visualized in Fig. 13. Because of the reasons explained above, we obtain few unblended and clean lines that pass our selection criteria. In this work, α Tau and α Cet show good ionization and excitation balance, although the scatter of the regression fit and the uncertainties of our results are quite high. The other three stars of this group show, on the other hand, a significant slope of the regression fit as a function of E.P. Note, however, we have no lines at low excitation potentials, making the regression fit not a good representation of the trend. We also obtain significant slopes in the regression fit as a function of reduced EW.

The NLTE effects are very small compared with the uncertainties obtained for the abundances. Ionization balance is, on

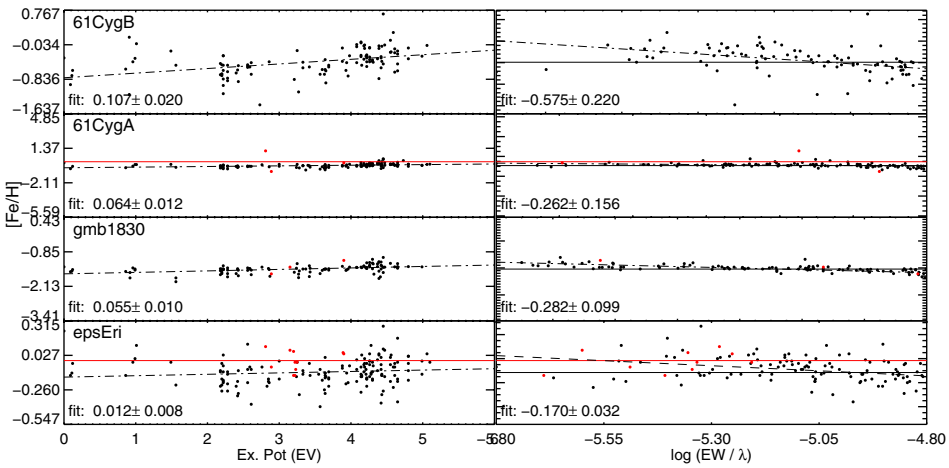

Fig. 14. Trends for group of K dwarfs.

the other hand, unsatisfied for this group, except for α Tau and α Cet, when considering the errors. The most extreme cases are ψ Phe and β Ara. As discussed in Sect. 5.4, it is impossible to find enough clean and unblended Fe II lines in this wavelength domain for such low temperatures, making these Fe II determinations, thus unrealistic or not even possible. In the case of ψ Phe, no line passed our selection criteria.

We recall that we found only one old reference for metallicity in the PASTEL catalog for β Ara (Luck 1979) and no reference for ψ Phe. Being aware of the difficulties in the analysis of these stars, we expect the Fe I abundances obtained by us to be uncertain but finally only one of our methods (Porto) could not provide a final value. Given this, we find it encouraging to obtain errors smaller than 0.3 dex and 0.2 dex for ψ Phe and β Ara, respectively.

6.2.5. K dwarfs

The objects Gmb 1830, 61 Cyg A, 61 Cyg B, and ϵ Eri are the benchmark K dwarfs. As in the previous groups, we plotted the abundances of the selected lines for each star in different panels in Fig. 14. Even after considering the errors, the members of this group do not present a good excitation balance, since significant trends are obtained for the regression fits for both E.P. and EW. The most extreme case is 61 Cyg B, where the Fe I abundances increase as a function of E.P. at a rate of 0.107 ± 0.02 and decrease as a function of reduced EW at a rate of 0.58 ± 0.22. This star is very cold, and therefore, its spectrum is very affected by blends of molecules that are not considered in our line list. A more suitable line list for such cold stars might help in obtaining a better excitation balance.

The objects 61 Cyg A and 61 Cyg B belong to a binary system; therefore, the same metallicity for both stars is expected. We obtained a value of −0.33 dex and −0.38 dex for the A and B components, respectively. The difference of 0.05 dex is within the errors. These values are about 0.15 dex lower than the literature values. We attribute this difference to the different temperature adopted by Luck & Heiter (2005) of 4640 K and 4400 K for the components A and B, respectively. These temperatures are ∼300 K above the values adopted by this work. Note that 61 Cyg B does not present a quantification of the ionization balance. Although we could select two Fe II lines, the mean iron abundance obtained for those lines was of +1.84, which is unphysical. As mentioned above, the reason for such unphysical results comes from the incapacity to detect unblended ionized iron lines for such cool stars. Thus, we do not list a ionization imbalance or line-by-line standard deviation of Fe II lines for 61 Cyg B in Table 3.

During one of the first attempts to determine metallicities for this system, it is worth mentioning that the values of fundamental $\log g$ considered for the analysis were different (4.49 and 4.61 dex) because they were obtained from evolutionary tracks of [Fe/H] = −0.10 and [Fe/H] = −0.30 for the A and B components of 61 Cyg, respectively. At that time, we retrieved a new metallicity of −0.49 and −0.55 dex for 61 Cyg A and B, respectively, which was translated to a difference in $\log g$ of −0.06 and −0.08, respectively. A third iteration on $\log g$ with the newest metallicity and a further iteration on [Fe/H] with the newest surface gravity would be desirable, although we have decided not do to this because of the large errors associated with the mass of this system (see Paper I) and also the errors obtained here for the final [Fe/H].

Note that the metallicity obtained for Gmb 1830 differs from the literature by ∼0.12 dex. The 19 works after 2000 in PASTEL have a mean temperature of 5090 ± 89 K, which is more than 250 K above the fundamental value. Recently, Creevey et al. (2012), who determined the angular diameter used to obtain the temperature in Paper I, obtained a value that is about 200 K less than the classical spectroscopic values. They suggested a revision of the metallicity based on this fundamental value. We have done this here, and we have seen that the consequence is considerable ionization and excitation imbalance for this new temperature. We have also studied the NTLE effects and concluded that they are not significant in this particular star. Moreover, Gmb 1830 is not so cold as to be strongly affected by molecules; however, its rather low metallicity and mass (0.6 $M_{\odot}$) suggests 3D or granulation effects caused by convection. The trends found in Fig. 14 and the ionization imbalance could be partly explained by the use of inaccurate 1D LTE models, but we cannot exclude the possibility that the fundamental temperature might be too low, or, perhaps, there is another effect that has not been investigated so far, such as magnetic fields or other activity process in the atmosphere. We noted that the knee in the curve of growth of this star is located at a slightly lower reduced EW than for other stars. This could lead to an inclusion of lines in the EW method that are too strong, which could bias the $v_{\rm mic}$ measurement and other slopes in this analysis. The difference in iron abundance due to this effect, however, should not be as significant as seen here. If the radius and bolometric flux of this star were accurately determined, then these new stellar parameters would imply that there is a physical process affecting this star in a way that we are not able to quantify. The problem with this hypothesis is that Gmb 1830 is a rather "normal" star, meaning that it has been commonly studied and has "normal" stellar parameters (not too cool, not too low gravity, and not too metal-poor). This makes

us believe that the measurement of its angular diameter might be affected by systematic errors. Until this issue is resolved, we prefer to point out that this object should be treated with caution as GBS.

6.3. Line list: golden lines

In this Section, we give an overview of the Fe I and Fe II line selection and line data, which were used to derive the final metallicity values listed in Table 3. Only the lines, which remained after the selection process described in Sect. 6, were considered. We determined which lines were used in common for each of the groups as defined in Sect. 6.2 and refer to these as the "golden lines". We found that there were significant differences in line selection between individual methods within several star groups, and thus the group definitions were somewhat expanded as explained below. The unique lists of 171 Fe I and 13 Fe II lines occurring in any of the groups can be found in Tables 4 and 5, respectively. The tables give the most relevant atomic data. For the lines identified for each individual group, we give the minimum and maximum standard deviations of the average line abundances and the minimum and maximum number of abundances averaged for each line in the respective column. Additionally, a master table containing a summary of all selected lines for all stars is available in the CDS service in Table 6.

The *metal-poor stars* (Sect. 6.2.1) were divided into dwarfs (HD 84937 and HD 140283) and giants (HD 122563) and designated "MPD" and "MPG", respectively, in Tables 4 and 5. As can be seen in the tables, the number of golden lines is considerably larger for the metal-poor giant (56 Fe I and 4 Fe II lines) than for metal-poor dwarfs (17 Fe I and 1 Fe II lines). Fifteen of the MPD Fe I lines are contained in the MPG list, while the single Fe II line common to the MPDs is different from the MPG Fe II lines. For the two MPDs, the standard deviations of the abundances are rather similar for all lines.

The group of *FG dwarfs* (Sect. 6.2.2) contains four stars for which the Fe I line selection differs from the others. This subgroup is designated "FGDb" in Table 4 and comprises of η Boo, HD 22879, HD 49933, and Procyon. The remaining stars listed in Sect. 6.2.2 are designated "FGDa". In general, the metallicity of the stars in the FGDb group is based on fewer Fe I lines than those in FGDa (see Table 3). However, the number of golden Fe I lines is similar for FGDa and FGDb (79 and 74, respectively) with 51 lines in common between the two subgroups. The four stars in FGDb differ from those in FGDa in various respects, which reduce the number of useful lines; HD 49933 and Procyon have the highest effective temperatures, HD 49933 and η Boo have the largest $v \sin i$; and HD 22879 is a moderately metal-poor star. The Fe II line lists are more homogeneous, resulting in six golden lines for all stars, with two exceptions as noted in Table 5 (column "FGD"). For the FG dwarfs, the abundance dispersions show a large variation from star to star. A more detailed investigation for the FGDa group shows that for Fe I lines the lowest minimum values are mostly due to the star β Hyi and the highest maximum values to β Vir. For most other stars and most Fe I lines, the dispersion is around 0.06 dex. In the FGDb group, Procyon and HD 22879 have the minimum dispersion for half of the Fe I lines each (and HD 49933 for seven lines). The maximum dispersion is mostly due to η Boo (61 lines) and sometimes to HD 49933 or Procyon (12 and 1 lines, respectively). The mode of the dispersion for FGDs is about 0.04 dex for all Fe II lines.

The group of *FGK giants* consists of the stars listed in Sect. 6.2.3. The 101 golden Fe I and 6 Fe II lines identified for this group are marked in column "FGKG" in Tables 4 and 5. For this group, the variation of dispersions is even larger than for FG dwarfs. The maximum dispersion for Fe I lines is mainly seen for μ Leo (82 lines), while the minimum dispersion occurs mainly for HD 220009, HD 107328, and Arcturus (for 54, 16, and 12 lines respectively). For most other stars, the dispersion scatters around 0.08 dex. For the Fe II lines, the largest dispersion is found for μ Leo. The dispersion is in general higher than for FG dwarfs (around 0.12 dex).

The group of *M giants* consists of the stars listed in Sect. 6.2.4 with one exception. The line list for ψ Phe differs significantly from the other stars (23 Fe I lines, of which only six are in common with the others). The 21 golden Fe I and 3 Fe II lines identified for this group are marked in column "MG" in Tables 4 and 5, while ψ Phe is listed in a separate column in Table 4 (no Fe II lines were selected for this star). The minimum abundance dispersion for Fe I lines in M giants is mostly found for γ Sge (13 lines), and the maximum dispersion is equally often in α Cet and β Ara (7 and 8 lines, respectively). The star ψ Phe shows high dispersions in general with the notable exceptions of the Fe I lines at 6219.28 and 6336.82 Å with dispersions of about 0.1 dex.

Finally, the group of *K dwarfs* described in Sect. 6.2.5 was divided into two subgroups with two different lists of golden Fe I lines. These are designated "KDa" (61 Cyg A, ϵ Eri) and "KDb" (61 Cyg B, Gmb 1830) in Table 4 with 127 and 85 Fe I lines, respectively, and 72 lines in common between the two subgroups. The differences in line selection between the two subgroups may be related to the specific parameter combinations ($T_{\rm eff}$, [Fe/H]) of the stars. In the KDa group, the maximum dispersion occurs for 61 Cyg A for 2/3 of the lines. In the KDb group, 61 Cyg B accounts for the maximum dispersion for most of the lines (77). Regarding the Fe II lines, the star ϵ Eri stands out among the group members with the largest number of lines selected (11 compared to 1–4). These are marked in Table 5 in column "KD", which includes a note identifying the lines in common with the other three stars.

6.3.1. Discrepant lines

While selecting the golden lines, it is important to discuss here that we found that the derived abundances by our methods differed significantly in some cases. This can be up to 0.4 dex for FG dwarfs for which we obtain the lowest line-to-line scatter in the final abundance determination (see above). This was surprising, since our golden lines were chosen to be unblended and are located in spectral regions with easy continuum placement. Moreover, our analysis is based on a great effort to have homogeneous atomic data and atmospheric models, making such differences difficult to explain.

Thus, we made a deep investigation of this issue and considered four examples of discrepant lines. This analysis was carried out mainly by M. Bergemann, U. Heiter, P. Jofré, K. Lind, T. Masseron, J. Sobeck, and H. Tabernero. We compared three of the radiative transfer codes (SME, MOOG, Turbospectrum) and found that their profiles were consistent when considering the same stellar parameters (They were set to those of the Sun, Arcturus, HD 84937, and HD 140283.). Naturally, a difference could still be seen due to different prescriptions and treatment of lines and spectrum formation (collisional broadening, radiative broadening, scattering, limb darkening, and spherical geometry, to name a few). However, this did not explain the 0.4 dex of the discrepant line examples.

We concluded that these discrepancies apparently come from a combination of different measured EWs (differing up to 60%), the details of the fitting procedures, the choice of

microturbulence parameter (see Fig. 3 for the different values), and the continuum placement. Understanding the contribution in the final discrepancy of each individual line from each of the aforementioned sources goes beyond the purpose of this paper. Here, we aim to combine abundances of numerous lines and methods homogeneously and provide a reference value for the metallicity of GBS. In general, our results agree very well on a line-by-line basis, and cases such as those discussed here are rare. However, we point out that this problem can arise even after performing analyses focused on homogeneity. Therefore, it is worthwhile to investigate further the sources of these discrepancies.

6.4. The final metallicity and its uncertainties

We have extensively mentioned in this manuscript that our revision of metallicity, using fixed $T_{\rm eff}$ and $\log g$ with values that are independent from spectroscopic analysis, does not necessarily give a [Fe/H] value that satisfies ionization and excitation balance. The differences can be used as a way to quantify the uncertainties in the theoretical assumptions used to compute synthetic line profiles and atmospheric models. This reminds us that the motivation of defining a set of GBS: the importance of having a standard set of stars with stellar parameters that are independent from spectroscopy, as this helps to make improvements to spectroscopic methods and models for stars of different spectral types.

In this work, it is worth discussing that we obtained three different values for metallicity: a value from Fe I lines, another one from Fe II lines, and a final one from Fe I lines after corrections from NLTE effects. Since we aimed to provide a reference metallicity, we chose the metallicity from Fe I lines after NLTE corrections to be our final value. Although we know that neutral iron lines are more sensitive to NLTE effects, we do not obtain a metallicity from ionized iron lines for all stars. Since we aim for homogeneity in this work, we prefer to consider the results from neutral lines for the final value. Finally, we have two main reasons that favor choosing the NLTE values: (i) the ionisation balance is slightly restored after NLTE corrections; and (ii) it is more accurate.

The natural question that arises from our choice is how to use the [Fe/H] of this work when someone has a parametrization method that employes LTE or that obtains Fe II abundance. The answer to this question is that we consider these values as part of the uncertainties of the final value. In other words, we quantify the error associated here with the ionization imbalance and with the NLTE effects by providing the difference between the final value and [Fe II/H] or [Fe I/H] with LTE approximation, respectively. These values are found in Table 3, labeled as Δ(ion) and Δ (LTE). To retrieve the metallicity value that one would obtain using Fe II lines, one needs to determine [Fe/H] − Δ(ion). Similarly, one needs to determine [Fe/H] − Δ(LTE) to retrieve the value obtained under LTE. Moreover, the full information for each selected line and individual method can be retrieved electronically from CDS for iron abundances and EWs in Tables 7−40 and 41−74, respectively.

To finish this section, we summarize that our reference metallicity is the one obtained by averaging the NLTE abundances of the selected Fe I lines. This value is associated with a series of sources of uncertainties, which are (i) the scatter in the line-by-line analysis of the selected lines; (ii) the difference in [Fe/H] when considering the uncertainty in the fundamental $T_{\rm eff}$; (iii) the difference in [Fe/H] when considering the uncertainty in the fundamental $\log g$; (iv) the difference in [Fe/H] when considering the uncertainty in $\upsilon_{\rm mic}$; (v) the difference in [Fe/H] obtained from neutral and ionised iron lines; and (vi) the difference in [Fe/H] obtained from LTE and NLTE analyses. The final value and its six sources of errors are listed in Table 3. In addition, the line-by-line standard deviation from Fe II abundances and the number of lines employed for the determination of metallicity from neutral and ionized iron lines are also indicated in the Table.

7. Summary and conclusions

We have made an extensive study on the determination of metallicity for the sample of 34 FGK GBS that are introduced in Paper I. In this study, we performed a spectral analysis of high S/N and high resolution ($R \geq 70\,000$) spectra taken from the library of GBS, as described in Paper II. Two different libraries were analyzed, one with the spectra at their original resolution and the other one convolved to $R = 70\,000$. In addition, the analysis was done for the same star observed with different instruments.

The analysis consisted of fixing effective temperature and surface gravity to the fundamental values presented in Paper I and determining metallicity and microturbulence velocity simultaneously. Up to seven different methods were used for this analysis; all of them considering the same input material, such as spectra, line list, and atmosphere models.

Three different runs were performed: *run-nodes*, which consists of one spectrum per GBS, and allows a one-to-one comparison between different methods; *run-resolutions*, which consists of the same spectrum of the previous run but uses its version in original resolution. This run allows the study of the impact of the varied resolution. The third run, *run-instruments* consists of the whole library convolved to $R = 70\,000$ and allows us to study instrumental effects. We obtained consistent and robust results, where the final metallicity was not biased either by method, resolution, or instrument.

Since we fixed $T_{\rm eff}$ and $\log g$ by values that are independent of spectroscopy, the metallicity analysis resulted in Fe I and Fe II abundances that did not necessarily agree. The comparison between neutral and ionized iron abundances was discussed with a quantification of how much $T_{\rm eff}$ and $\log g$ would need to deviate from the fundamental value to comply with ionization balance, excitation balance, and line strength balance. This was done by a test of determining $T_{\rm eff}$, $\log g$, and $\upsilon_{\rm mic}$ together with [Fe/H].

To provide a final value of metallicity, we combined our results using a line-by-line approach. Starting from all individual abundances of every method, we only selected those lines, which were analyzed by at least three methods and agreed within 2σ of the average abundance calculated from all lines. The selected lines were then averaged to have only one abundance per line, which was then used to perform NLTE corrections and quality checks, such as ionization and excitation balance. Our final value consists of the iron abundance obtained from Fe I lines after NLTE corrections.

We studied many different sources of errors, which are all reported separately. The first one comes from the consideration of the 1σ scatter of the line-by-line analysis. Then, we determined the uncertainty of the metallicity due to the errors associated with the effective temperature, surface gravity and microturbulent velocity. To do so, iron abundances were calculated by performing six additional runs only on the selected lines; each run fixes $T_{\rm eff}$, $\log g$, and $\upsilon_{\rm mic}$ to the values considering their associated errors. Finally, errors due to ionization imbalance and deviations from NLTE were quantified.

Table 4. List of "golden" Fe I lines for various groups of stars (see text for definition of groups).

λ [Å]	E [eV]	log gf	Waals	Ref	MPD (2)	MPG (1)	FGDa (11)	FGDb (4)	FGKG (7)	MG (4)	ψ Phe	KDa (2)	KDb (2)
4787.83	2.9980	–2.563	818.227	102					0.04/0.41 3/3			0.06/0.07 3/3	
4788.76	3.2370	–1.763	238.249	102			0.02/0.16 3/3	0.03/0.09 3/3	0.06/0.52 3/3			0.05/0.11 3/3	
4802.88	3.6420	–1.514	356.244	102								0.11/0.30 3/4	
4808.15	3.2510	–2.690	297.274	156			0.04/0.23 5/6	0.04/0.16 3/5	0.06/0.31 5/6	0.14/0.25 3/4		0.03/0.04 6/6	0.05/0.06 4/5
4869.46	3.5460	–2.420	246.248	156					0.06/0.65 3/3				
4875.88	3.3320	–1.920	848.231	156			0.07/0.39 4/4	0.08/0.21 3/4	0.04/0.51 4/4			0.13/0.15 3/4	0.08/0.21 3/3
4877.60	2.9980	–3.050	795.230	156					0.04/0.30 3/3			0.05/0.07 3/3	
4907.73	3.4300	–1.840	909.227	129			0.04/0.14 3/3		0.02/0.40 3/3			0.07/0.09 3/3	
4924.77	2.2790	–2.178	360.244	102		0.04 3		0.03/0.10 3/3					0.11/0.11 3/3
4946.39	3.3680	–1.170	848.232	187				0.04/0.13 3/3					
4950.11	3.4170	–1.670	880.228	129		0.02 4	0.07/0.22 3/5	0.03/0.14 4/5	0.06/0.16 3/4			0.14/0.16 4/4	0.05/0.16 3/5
4962.57	4.1780	–1.182	0.000	102			0.02/0.19 3/4	0.01/0.15 3/4				0.01/0.12 4/4	0.01/0.16 3/4
4969.92	4.2170	–0.710	962.279	129				0.01/0.16 3/3					0.03/0.10 3/3
4985.55	2.8650	–1.340	727.238	190		0.03 3							
4994.13	0.9150	–3.002	246.245	102	0.03/0.05 3/3	0.04 3		0.08/0.18 3/4					
5001.86	3.8810	–0.010	725.240	114		0.02 3							
5012.69	4.2830	–1.690	1020.279	156			0.05/0.20 3/4		0.07/0.30 3/4			0.06/0.11 4/4	
5044.21	2.8510	–2.038	713.238	102				0.04/0.11 3/3					
5049.82	2.2790	–1.349	353.239	102		0.01 3							
5058.50	3.6420	–2.830	353.313	167			0.09/0.20 3/5		0.09/0.22 4/5	0.15/0.40 3/4		0.12/0.28 4/4	
5060.08	0.0000	–5.431	0.000	102								0.17/0.29 3/3	0.08/0.08 3/3
5088.15	4.1540	–1.680	810.278	156					0.09/0.52 3/3				
5107.45	0.9900	–3.091	248.245	102		0.04 3							
5107.64	1.5570	–2.358	289.258	102		0.02 3							
5109.65	4.3010	–0.980	980.280	167								0.07/0.15 3/3	0.04/0.40 3/3
5127.36	0.9150	–3.278	243.246	102		0.03 3		0.05/0.16 3/3					
5131.47	2.2230	–2.515	356.274	102				0.03/0.17 3/3				0.11/0.34 3/3	0.09/0.09 3/3
5141.74	2.4240	–2.101	367.251	102		0.09 4	0.03/0.37 3/5	0.04/0.24 4/5				0.08/0.20 3/4	
5194.94	1.5570	–2.021	286.255	102		0.07 3							
5197.94	4.3010	–1.540	925.279	156			0.06/0.20 4/5	0.04/0.14 4/4	0.17/0.57 4/5			0.03/0.19 3/5	
5198.71	2.2230	–2.113	351.271	102		0.03 3		0.03/0.21 3/3					
5215.18	3.2660	–0.871	849.229	102		0.03 3							0.03/0.05 3/3
5217.39	3.2110	–1.116	815.232	102	0.03/0.07 3/3	0.05 4		0.07/0.18 3/4					0.09/0.23 3/4
5223.18	3.6350	–1.783	390.253	102					0.05/0.34 3/3				
5225.53	0.1100	–4.755	207.253	102		0.04 4		0.04/0.27 3/4					0.11/0.21 4/4
5228.38	4.2200	–1.190	809.278	156								0.07/0.22 3/3	
5232.94	2.9400	–0.076	713.238	102	0.06/0.07 5/5	0.08 4							
5242.49	3.6340	–0.967	361.248	102		0.07 3		0.05/0.10 3/4				0.12/0.23 3/4	0.06/0.06 3/4
5243.78	4.2560	–1.050	842.278	156				0.00/0.25 4/4	0.13/0.46 3/4			0.07/0.09 4/4	0.04/0.05 3/4
5247.05	0.0870	–4.975	206.253	102		0.04 3						0.02/0.13 3/3	0.15/0.33 3/3
5250.21	0.1210	–4.918	207.253	102		0.03 3						0.13/0.20 3/3	0.10/0.10 3/3

Notes. Column descriptions: λ: wavelength. $E_{\rm low}$: lower level energy. "Waals": parameters used to calculate line broadening due to collisions with neutral hydrogen. Integer component: broadening cross-section at a velocity of 10^4 m s^{-1} in atomic units. Fractional component: velocity parameter (see Barklem et al. 2000). If zero, the Unsöld approximation was used. "Ref": reference code for the gf-values (see below). The remaining columns are headed by a label for each group defined in the text and the number of stars in parantheses. The columns give the minimum and maximum standard deviations of the average line abundances and the minimum and maximum number of abundances averaged for each line for each group.

References. 102: Bard et al. (1991); Bard & Kock (1994); Blackwell et al. (1979a,b, 1982a,b, 1995); O'Brian et al. (1991). 114: Bridges & Kornblith (1974); Fuhr & Wiese (2006). 129: Garz & Kock (1969); Fuhr et al. (1988). 156: May et al. (1974). 167: Richter & Wulff (1970); Fuhr et al. (1988). 186: Wolnik et al. (1970); Fuhr et al. (1988). 187: Wolnik et al. (1971); Fuhr et al. (1988).

Table 4. continued.

λ [Å]	E [eV]	log gf	Waals	Ref	MPD (2)	MPG (1)	FGDa (11)	FGDb (4)	FGKG (7)	MG (4)	ψ Phe	KDa (2)	KDb (2)
5250.65	2.1980	–2.180	344.268	102	0.06/0.07 3/3	0.01 4		0.04/0.13 3/4	0.09/0.26 3/3			0.24/0.27 3/3	0.11/0.15 3/4
5253.02	2.2790	–3.840	368.253	156			0.03/0.13 4/5		0.06/0.26 4/5		0.51 3	0.06/0.09 5/5	
5253.46	3.2830	–1.573	849.229	102				0.04/0.20 3/3				0.17/0.19 3/3	
5285.13	4.4340	–1.540	1046.282	156					0.05/0.25 3/3			0.02/0.09 3/3	
5288.52	3.6940	–1.508	353.297	102								0.06/0.08 3/3	
5293.96	4.1430	–1.770	0.000	156			0.05/0.15 3/4	0.05/0.12 3/4	0.04/0.23 3/4			0.03/0.04 4/4	
5294.55	3.6400	–2.760	394.237	156			0.04/0.41 4/5		0.05/0.17 5/5			0.04/0.22 3/5	
5295.31	4.4150	–1.590	1014.281	156			0.03/0.19 3/5	0.02/0.11 4/5	0.04/0.21 5/5			0.05/0.09 5/5	0.06/0.18 3/3
5302.30	3.2830	–0.720	835.231	102		0.07 4							
5321.11	4.4340	–1.089	1024.281	102								0.03/0.10 3/3	
5322.04	2.2790	–2.802	341.236	102								0.07/0.10 3/3	0.02/0.04 3/3
5339.93	3.2660	–0.684	815.234	102		0.06 4							
5365.40	3.5730	–1.020	283.261	102		0.15 3		0.05/0.19 3/4			0.39 3	0.08/0.09 3/4	0.06/0.08 3/4
5367.47	4.4150	0.444	972.280	102		0.08 3							
5373.71	4.4730	–0.760	1044.282	156			0.03/0.12 3/4	0.00/0.08 3/4	0.06/0.27 3/4			0.09/0.18 4/4	
5379.57	3.6940	–1.514	363.249	102			0.01/0.13 3/4	0.02/0.25 3/4				0.06/0.08 4/4	0.01/0.11 3/4
5386.33	4.1540	–1.670	930.278	156			0.06/0.21 3/5	0.04/0.16 4/5	0.06/0.30 5/5		0.66 3	0.08/0.13 5/5	0.07/0.22 3/4
5389.48	4.4150	–0.410	959.280	187		0.06 3		0.02/0.07 3/4				0.08/0.09 3/4	0.02/0.04 3/4
5395.22	4.4450	–2.070	995.281	156					0.03/0.19 3/4		0.25 3	0.09/0.09 3/3	
5397.13	0.9150	–1.988	238.249	102	0.10/0.13 3/3								
5398.28	4.4450	–0.630	993.280	156			0.04/0.10 3/5	0.01/0.29 4/5	0.07/0.31 3/5		0.49 3	0.10/0.16 4/4	0.02/0.04 3/5
5412.78	4.4340	–1.716	971.280	102			0.01/0.12 3/4		0.03/0.18 4/4		0.17 3	0.03/0.05 3/4	
5415.20	4.3860	0.643	910.279	102		0.09 3							
5417.03	4.4150	–1.580	944.280	156			0.04/0.30 4/5	0.01/0.11 3/5	0.05/0.30 4/5			0.06/0.14 5/5	0.13/0.59 3/4
5424.07	4.3200	0.520	825.278	186	0.03/0.06 3/3	0.02 3							
5434.52	1.0110	–2.119	243.247	102	0.09/0.22 3/3								
5441.34	4.3120	–1.630	807.278	156			0.02/0.25 4/5		0.06/0.32 4/5	0.17/0.39 3/4		0.05/0.18 5/5	0.10/0.11 3/3
5445.04	4.3860	–0.020	895.279	186									0.03/0.04 3/3
5464.28	4.1430	–1.402	380.250	102					0.06/0.33 3/3			0.07/0.08 3/3	
5466.40	4.3710	–0.630	865.278	187				0.02/0.13 3/4	0.14/0.45 4/4			0.13/0.29 4/4	0.03/0.09 4/4
5470.09	4.4460	–1.710	953.280	156			0.05/0.13 3/4		0.05/0.25 3/4			0.07/0.09 3/4	0.08/0.12 3/3
5473.90	4.1540	–0.790	738.241	114								0.12/0.16 3/3	0.01/0.03 3/3
5483.10	4.1540	–1.406	737.241	102								0.06/0.08 3/3	
5487.15	4.4150	–1.430	908.279	156					0.13/0.37 3/3				
5494.46	4.0760	–1.990	0.000	156			0.01/0.24 3/4		0.08/0.33 4/4				
5522.45	4.2090	–1.450	744.215	156			0.02/0.09 3/4	0.02/0.09 3/4	0.06/0.31 4/4	0.06/0.23 3/4	0.27 3	0.05/0.05 4/4	0.01/0.13 3/4
5539.28	3.6420	–2.560	383.260	156			0.03/0.35 3/4		0.08/0.29 3/4			0.02/0.16 3/4	
5543.94	4.2170	–1.040	742.238	156			0.01/0.10 3/4	0.02/0.14 3/4	0.04/0.33 3/4		0.22 3	0.06/0.06 4/4	0.01/0.21 4/4
5546.51	4.3710	–1.210	825.278	156			0.05/0.10 3/4	0.04/0.17 3/4	0.06/0.33 3/4			0.05/0.06 4/4	
5560.21	4.4340	–1.090	895.278	156			0.03/0.10 3/4	0.03/0.20 3/4	0.06/0.28 3/4	0.02/0.20 3/4		0.05/0.07 4/4	0.01/0.01 3/3
5569.62	3.4170	–0.486	848.233	102	0.03/0.03 3/3	0.08 4							
5576.09	3.4300	–0.900	854.232	156	0.03/0.06 3/3	0.06 4		0.06/0.15 3/4					
5586.76	3.3680	–0.120	817.238	102		0.02 3							
5618.63	4.2090	–1.275	732.214	102			0.02/0.09 5/5	0.02/0.21 4/5	0.04/0.24 4/5			0.05/0.08 5/5	
5619.60	4.3860	–1.600	808.277	156			0.02/0.26 3/4	0.01/0.05 3/3	0.04/0.41 4/4			0.03/0.04 4/4	
5633.95	4.9910	–0.230	635.270	156			0.02/0.18 3/4	0.01/0.18 3/4	0.08/0.42 3/4			0.04/0.07 4/4	0.03/0.13 3/4
5636.70	3.6400	–2.510	368.310	156			0.02/0.33 3/4		0.02/0.27 4/4			0.03/0.04 4/4	
5638.26	4.2200	–0.770	730.235	156			0.03/0.17 4/5	0.03/0.17 4/5	0.09/0.41 4/5			0.05/0.09 4/5	0.02/0.13 5/5

Table 4. continued.

λ [Å]	E [eV]	log gf	Waals	Ref	MPD (2)	MPG (1)	FGDa (11)	FGDb (4)	FGKG (7)	MG (4)	ψ Phe	KDa (2)	KDb (2)
5641.43	4.2560	–1.080	739.234	156			0.01/0.28 3/3					0.03/0.05 3/3	
5649.99	5.0990	–0.820	719.265	156					0.04/0.24 3/3			0.02/0.04 3/3	
5651.47	4.4730	–1.900	898.278	156			0.04/0.17 5/6		0.04/0.18 6/6	0.12/0.20 3/4		0.04/0.05 5/6	
5652.32	4.2600	–1.850	754.210	156			0.05/0.22 4/5		0.02/0.27 4/5			0.04/0.04 5/5	0.10/0.28 3/3
5653.87	4.3860	–1.540	792.277	156			0.01/0.13 3/3		0.01/0.14 3/3			0.02/0.10 3/3	0.04/0.04 3/3
5655.18	5.0640	–0.600	0.000	156								0.08/0.27 3/3	
5661.35	4.2840	–1.756	765.209	102			0.03/0.10 3/4		0.02/0.20 3/4			0.04/0.08 4/4	
5662.52	4.1780	–0.573	724.235	102		0.06 4	0.04/0.22 3/5	0.04/0.24 4/5	0.10/0.25 3/4			0.14/0.28 4/4	0.06/0.14 3/5
5679.02	4.6520	–0.820	1106.291	156			0.04/0.12 5/5	0.03/0.30 4/5	0.08/0.26 5/5			0.05/0.07 5/5	0.04/0.11 4/4
5691.50	4.3010	–1.420	746.231	156								0.02/0.04 3/3	
5696.09	4.5480	–1.720	965.279	102					0.01/0.28 3/4			0.03/0.07 3/3	
5698.02	3.6400	–2.580	385.252	156					0.01/0.22 3/3				
5701.54	2.5590	–2.160	361.237	102		0.06 4	0.05/0.28 4/5		0.10/0.28 3/3			0.09/0.11 4/4	0.06/0.16 4/5
5705.46	4.3010	–1.355	744.231	102			0.04/0.11 4/5		0.03/0.23 5/5	0.12/0.49 3/4	0.48 3	0.04/0.06 5/5	0.03/0.16 4/5
5731.76	4.2560	–1.200	727.232	156			0.02/0.15 5/5	0.02/0.15 4/5	0.09/0.38 4/5		0.58 3	0.03/0.06 5/5	0.03/0.07 4/5
5732.30	4.9910	–1.460	613.275	156					0.02/0.11 4/5	0.11/0.18 3/4		0.04/0.10 4/4	
5741.85	4.2560	–1.672	725.232	102			0.03/0.17 5/6		0.05/0.25 6/6	0.07/0.23 3/4	0.62 3	0.03/0.06 6/6	0.05/0.17 4/5
5760.34	3.6420	–2.390	386.250	156								0.03/0.07 3/3	
5775.08	4.2200	–1.297	720.231	102					0.07/0.34 5/6	0.04/0.39 3/5			0.05/0.07 5/5
5778.45	2.5880	–3.430	361.237	102					0.06/0.27 4/6	0.03/0.44 3/4			0.04/0.10 4/4
5784.66	3.3960	–2.532	796.244	102					0.05/0.42 3/3				
5849.68	3.6940	–2.890	379.305	156					0.02/0.10 3/4	0.09/0.13 3/4		0.02/0.02 3/3	
5853.15	1.4850	–5.180	0.000	156					0.02/0.16 3/4			0.03/0.05 3/3	
5855.08	4.6080	–1.478	962.279	102			0.03/0.17 5/6		0.04/0.33 5/6	0.14/0.49 3/5		0.04/0.11 5/6	
5858.78	4.2200	–2.160	786.278	156					0.01/0.09 3/4	0.02/0.18 3/3		0.02/0.08 3/3	
5883.82	3.9600	–1.260	998.250	156			0.01/0.15 3/4	0.01/0.12 3/3				0.06/0.07 3/4	0.03/0.08 3/3
5902.47	4.5930	–1.710	227.252	156					0.06/0.19 3/4			0.11/0.26 3/3	
5905.67	4.6520	–0.690	994.282	156			0.01/0.16 4/5	0.02/0.15 4/5				0.06/0.21 4/4	0.03/0.16 3/3
5927.79	4.6520	–0.990	984.281	156			0.02/0.09 3/4		0.05/0.16 3/4	0.14/0.46 3/4	0.39 3	0.04/0.05 3/4	0.01/0.08 3/3
5929.68	4.5480	–1.310	864.275	156			0.02/0.14 3/4		0.03/0.76 3/4			0.03/0.05 3/4	0.03/0.04 3/3
5930.18	4.6520	–0.230	983.281	187		0.07 3	0.04/0.19 3/5	0.01/0.13 3/5	0.06/0.17 3/5			0.14/0.19 3/4	0.02/0.05 4/4
5934.65	3.9280	–1.070	959.247	156			0.02/0.18 3/4	0.01/0.20 3/4	0.03/0.30 3/4			0.11/0.15 4/4	0.03/0.07 3/3
5956.69	0.8590	–4.553	227.252	102		0.04 4						0.03/0.10 4/5	0.03/0.16 4/4
6003.01	3.8810	–1.120	898.241	187				0.03/0.20 3/4				0.16/0.31 3/3	0.06/0.10 3/3
6012.21	2.2230	–4.038	309.270	102								0.03/0.05 3/4	0.04/0.12 3/4
6027.05	4.0760	–1.089	380.250	102			0.04/0.10 4/5	0.03/0.11 3/5	0.06/0.14 3/5			0.06/0.27 5/5	0.02/0.16 4/4
6065.48	2.6080	–1.470	354.234	102	0.03/0.05 4/4	0.05 4		0.04/0.29 3/5					0.09/0.34 3/4
6079.01	4.6520	–1.020	920.276	156			0.01/0.18 3/4	0.02/0.19 3/4	0.05/0.11 3/4			0.02/0.03 3/4	0.02/0.08 3/3
6093.64	4.6070	–1.400	866.274	156			0.01/0.06 3/4		0.02/0.07 3/4	0.05/0.20 3/3		0.03/0.04 3/4	
6094.37	4.6520	–1.840	914.276	156					0.03/0.32 3/4			0.03/0.13 3/3	
6096.66	3.9840	–1.830	963.250	156			0.03/0.23 4/5	0.02/0.08 4/4	0.04/0.18 4/5			0.04/0.14 5/5	0.03/0.08 3/4
6127.91	4.1430	–1.399	0.000	102								0.10/0.27 3/3	
6136.99	2.1980	–2.941	280.265	102		0.02 3	0.05/0.37 3/4					0.05/0.12 3/4	0.02/0.07 3/3
6151.62	2.1760	–3.312	277.263	102			0.03/0.10 5/5	0.01/0.15 4/4	0.01/0.83 3/5	0.13/0.78 3/4	0.65 4	0.06/0.07 4/5	0.03/0.14 4/5
6165.36	4.1430	–1.473	380.250	102			0.02/0.18 4/5	0.01/0.07 3/5	0.05/0.64 4/5	0.15/0.32 3/4		0.04/0.06 5/5	0.04/0.16 4/5
6173.33	2.2230	–2.880	281.266	102		0.05 4	0.03/0.12 4/5	0.02/0.13 4/5	0.06/0.21 3/4	0.19/0.78 3/4		0.08/0.20 4/5	0.06/0.12 3/4
6187.99	3.9430	–1.620	903.244	156			0.03/0.22 5/6	0.02/0.20 5/6	0.05/0.65 5/6			0.04/0.21 6/6	0.03/0.08 4/5
6200.31	2.6080	–2.405	350.235	102		0.07 3	0.01/0.13 3/4	0.02/0.09 3/4	0.02/0.20 3/4			0.08/0.10 3/4	0.07/0.29 3/3

Table 4. continued.

λ [Å]	E [eV]	log gf	Waals	Ref	MPD (2)	MPG (1)	FGDa (11)	FGDb (4)	FGKG (7)	MG (4)	ψ Phe	KDa (2)	KDb (2)
6219.28	2.1980	–2.434	278.264	102		0.02 4	0.02/0.40 3/5	0.04/0.26 3/5	0.02/0.28 3/3		0.10 3	0.09/0.17 3/4	0.05/0.07 4/4
6226.73	3.8830	–2.120	845.244	156			0.01/0.25 4/5		0.03/0.62 4/5			0.03/0.05 3/5	
6240.65	2.2230	–3.203	301.272	102								0.09/0.38 3/3	
6246.32	3.6020	–0.805	820.246	102	0.01/0.05 3/4	0.06 5		0.06/0.18 4/6	0.04/0.21 4/4		0.22 3	0.23/0.34 3/4	0.06/0.31 5/5
6252.56	2.4040	–1.727	326.245	102	0.02/0.05 5/5	0.04 5	0.07/0.26 4/6	0.07/0.14 3/6	0.09/0.26 3/4		0.37 3	0.17/0.19 3/4	0.11/0.35 3/5
6265.13	2.1760	–2.545	274.261	102		0.03 4	0.03/0.34 4/5	0.03/0.29 4/5	0.02/0.25 3/4			0.12/0.17 4/5	0.09/0.65 4/5
6270.22	2.8580	–2.536	350.249	102			0.03/0.29 4/5	0.03/0.17 3/5	0.06/0.76 4/5			0.04/0.05 4/5	0.02/0.27 4/5
6271.28	3.3320	–2.703	720.247	102			0.04/0.28 3/4		0.04/0.20 3/4			0.03/0.07 3/4	
6297.79	2.2230	–2.702	278.264	102		0.04 3		0.04/0.15 3/4				0.15/0.18 4/4	0.08/0.19 3/4
6301.50	3.6540	–0.718	0.000	102		0.09 3							
6315.81	4.0760	–1.610	410.250	156			0.02/0.35 3/4	0.04/0.21 3/3	0.05/0.18 3/4			0.05/0.10 4/4	0.02/0.53 3/4
6322.69	2.5880	–2.448	345.238	102		0.02 4	0.02/0.32 4/5	0.02/0.24 4/5	0.08/0.24 3/4		0.31 3	0.08/0.20 4/4	0.02/0.18 4/5
6335.33	2.1980	–2.177	275.261	102	0.03/0.09 3/3	0.05 4	0.03/0.14 3/5	0.05/0.15 3/5	0.08/0.28 3/3	0.23/0.27 3/3		0.18/0.22 3/3	0.12/0.28 4/4
6336.82	3.6860	–0.856	845.240	102		0.06 5	0.06/0.28 3/6	0.05/0.26 5/6	0.11/0.18 4/5	0.24/0.65 3/4	0.13 3		0.05/0.26 4/5
6393.60	2.4330	–1.504	326.246	102	0.04/0.07 5/5	0.06 5	0.05/0.23 3/5	0.07/0.45 4/5	0.12/0.24 3/4		0.51 3		0.14/0.18 3/4
6411.65	3.6540	–0.656	820.247	102	0.03/0.06 4/5	0.05 5		0.06/0.12 4/6	0.04/0.26 3/4		0.51 3		0.11/0.42 4/5
6430.85	2.1760	–1.976	271.257	102	0.02/0.05 5/5	0.05 5	0.03/0.45 3/6	0.05/0.13 3/6	0.10/0.25 3/4		0.51 3	0.18/0.29 3/4	
6481.87	2.2790	–2.985	308.243	102		0.03 4	0.02/0.42 4/5	0.02/0.41 3/5	0.07/0.28 3/4			0.04/0.10 3/5	0.00/0.07 3/4
6494.98	2.4040	–1.256	321.247	102	0.03/0.08 3/3	0.04 3							
6496.47	4.7950	–0.530	925.279	156					0.07/0.37 3/4			0.03/0.08 3/4	0.01/0.02 3/3
6498.94	0.9580	–4.688	226.253	102		0.06 4						0.11/0.11 4/5	0.04/0.23 3/4
6533.93	4.5580	–1.360	908.277	156					0.01/0.07 3/4			0.02/0.03 3/4	
6574.23	0.9900	–5.013	0.000	102								0.08/0.13 3/4	0.04/0.16 3/4
6593.87	2.4330	–2.394	321.247	102		0.03 3	0.01/0.15 3/4	0.03/0.09 3/4	0.03/0.26 3/3			0.08/0.11 4/4	0.09/0.23 3/3
6597.56	4.7950	–0.970	893.276	156			0.01/0.18 3/4	0.02/0.14 3/4				0.02/0.19 3/4	0.04/0.05 3/3
6609.11	2.5590	–2.676	335.245	102		0.04 3	0.02/0.42 3/4	0.04/0.17 3/4	0.09/0.28 3/3			0.05/0.06 3/3	0.02/0.38 3/4
6627.54	4.5480	–1.580	754.209	156								0.02/0.13 3/3	
6648.08	1.0110	–5.918	229.254	102							0.50 3	0.02/0.04 3/3	
6699.14	4.5930	–2.101	297.273	102					0.05/0.14 4/5			0.04/0.07 3/4	
6703.57	2.7580	–3.060	320.264	156			0.04/0.11 3/4					0.02/0.03 3/3	
6713.74	4.7950	–1.500	857.272	156					0.01/0.44 3/4			0.04/0.05 3/3	
6739.52	1.5570	–4.794	256.244	102					0.01/0.67 3/4			0.01/0.03 3/3	
6750.15	2.4240	–2.604	335.241	102		0.03 3	0.03/0.14 3/4				1.09 3	0.07/0.07 3/4	0.03/0.05 3/3
6810.26	4.6070	–0.986	873.275	102					0.05/0.43 3/4				0.02/0.12 3/3

Table 5. List of "golden" Fe II lines for various groups of stars (see text for definition of groups).

					smallskip				
λ [Å]	E [eV]	log gf	Waals	Ref	MPD (2)	MPG (1)	FGD† (15)	FGKG (7)	MG (4)	KD‡ (1)
4923.93	2.8910	−1.260	175.202	158	0.16/0.25 4/4					
4993.36	2.8070	−3.684	172.220	166		0.04 3	0.00/0.21 3/4	0.13/0.64 3/4		0.29 4
5264.81	3.2300	−3.130	186.300	158						0.09 3
5325.55	3.2210	−3.160	179.252	158						0.08 3
5414.07	3.2210	−3.580	185.303	158			0.01/0.07 3/4	0.07/0.24 3/4		0.08 3
5425.26	3.1990	−3.220	178.255	158		0.14 3	0.01/0.14 4/5	0.09/0.35 4/5	0.17/0.50 3/5	0.11 5
5534.85	3.2450	−2.865	178.239	166						0.06 3
5991.38	3.1530	−3.647	172.221	166			0.01/0.12 3/4			0.04 3
6084.11	3.1990	−3.881	173.223	166				0.08/0.20 3/4		0.03 3
6247.56	3.8920	−2.435	186.272	166						0.04 3
6432.68	2.8910	−3.570	169.204	158		0.07 3	0.01/0.32 3/4	0.02/0.07 3/4	0.10/0.19 3/4	0.05 3
6456.38	3.9030	−2.185	185.276	166		0.14 4	0.03/0.19 4/5	0.04/0.38 4/5	0.14/0.24 4/4	0.03 5

Notes. Column descriptions: λ: wavelength. $E_{\rm low}$: lower level energy. "Waals": parameters used to calculate line broadening due to collisions with neutral hydrogen. Integer component: broadening cross-section at a velocity of 10^4 m s^{-1} in atomic units. Fractional component: velocity parameter (see Barklem et al. 2000). If zero, the Unsöld approximation was used. "Ref": reference code for the gf-values (see below). The remaining columns are headed by a label for each group defined in the text and the number of stars in parantheses. The columns give the minimum and maximum standard deviations of the average line abundances and the minimum and maximum number of abundances averaged for each line for each group. † 5414.07 Å was not used in μ Cas and 5991.38 Å not in HD 49933. ‡ The column marks the lines used for ϵ Eri; of these, only 4993.36 Å and 6456.38 Å were used in 61 Cyg A and Gmb 1830; only 4993.36 Å and 5425.26 Å in 61 Cyg B.

References. 158: Meléndez & Barbuy (2009). 166: Raassen & Uylings (1998).

Generally, we were able to obtain robust values for [Fe/H] for the stars of our sample, which makes this work the first to determine metallicity homogeneously for the complete set of *Gaia* FGK GBS. Our final [Fe/H] values are thus appropriate for use as reference values. When comparing our results with previous studies in the literature, we obtain a good agreement for 28 stars and different values for four stars (HD 220009, 61 Cyg A, 61 Cyg B, and β Ara), which we adopt as a new reference [Fe/H]. In addition, we provide a value for the metallicity of ψ Phe for the first time. Although we obtain very different metallicities for Gmb 1830 compared to the literature, we prefer to caution against defining a new set of reference parameters for this star. We are unable to understand the reason for this discrepancy, and further investigations on its fundamental parameters are needed. The final reference values and their uncertainties are indicated in Table 3. Having well-determined stellar parameters for the GBS will improve the homogeneous analyses of current stellar surveys, which have become a key piece in Galactic studies.

We made a careful study in the selection of candidates to serve as benchmarks for stellar spectra analyses. The accurate distance and angular diameter of these stars provide us with fundamental determinations of effective temperature and surface gravity. Their proximity and brightness provide us with the possibility of having high quality spectra that are suitable for a more precise determination of metallicity. With Paper I and Paper II, this paper on the series of FGK GBS describes and extensively discusses our choice for the reference values of the three main stellar parameters $T_{\rm eff}$, log g and [Fe/H]. We encourage our colleagues to use the spectra of the GBS, and their parameters to evaluate the performance of parametrization methods as a way to relate the data to the *Gaia*-ESO Survey. We can transform our spectra such that they look like the data taken from other spectrographs. Our metallicities can be reproduced as we document each individual value used for its final determination in the online tables. Using this material will allow for the connection of different methods and cross-calibration of surveys, which will lead to a more consistent understanding of the structure and evolution of our Galaxy.

Acknowledgements. We thank all LUMBA members for the rich discussions on the development of SME for automatic analyses of spectra, which were crucial to the development of the setups used for the analysis of this project. P.J. acknowledges the useful comments and proof reading done by T. Mädler. U.H. acknowledges support from the Swedish National Space Board (Rymdstyrelsen). S.G.S. acknowledges the support from the Fundação para a Ciência e Tecnologia (Portugal) in the form of the grants SFRH/BPD/47611/2008. The computations for the AMBRE project have been performed with the high-performance computing facility SIGAMM, hosted by OCA. R.S. acknowledges the support of the ASI (Agenzia Spaziale Italiana) under contracts to INAF I/037/08/0 and I/058/10/0.

References

Adibekyan, V. Z., Sousa, S. G., Santos, N. C., et al. 2012, A&A, 545, A32
Adibekyan, V. Z., Figueira, P., Santos, N. C., et al. 2013, A&A, 554, A44
Allende Prieto, C., Barklem, P. S., Lambert, D. L., & Cunha, K. 2004, A&A, 420, 183
Allende Prieto, C., Beers, T. C., Wilhelm, R., et al. 2006, ApJ, 636, 804
Allende Prieto, C., Majewski, S. R., Schiavon, R., et al. 2008a, Astron. Nachr., 329, 1018
Allende Prieto, C., Sivarani, T., Beers, T. C., et al. 2008b, AJ, 136, 2070
Alvarez, R., & Plez, B. 1998, A&A, 330, 1109
Anders, E., & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197
Asplund, M. 2005, ARA&A, 43, 481
Asplund, M., Nordlund, Å., Trampedach, R., & Stein, R. F. 1999, A&A, 346, L17
Asplund, M., Grevesse, N., & Sauval, A. J. 2005, in Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis, eds. T. G. Barnes, III, & F. N. Bash, ASP Conf. Ser., 336, 25

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Aurière, M. 2003, in EAS Pub. Ser. 9, eds. J. Arnaud, & N. Meunier, 105
Bagnulo, S., Jehin, E., Ledoux, C., et al. 2003, The Messenger, 114, 10
Bailer-Jones, C. A. L., Andrae, R., Arcay, B., et al. 2013, A&A, 559, A74
Ballester, P., Modigliani, A., Boitquin, O., et al. 2000, The Messenger, 101, 31
Bard, A., & Kock, M. 1994, A&A, 282, 1014
Bard, A., Kock, A., & Kock, M. 1991, A&A, 248, 315
Barklem, P. S., Piskunov, N., & O'Mara, B. J. 2000, A&AS, 142, 467
Blanco-Cuaresma, S., Soubiran, C., Jofré, P., & Heiter, U. 2014, A&A, submitted [arXiv:1403.3090]
Beers, T. C., Carollo, D., Ivezić, Ž., et al. 2012, ApJ, 746, 34
Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJ, 642, L137
Bensby, T., Feltzing, S., & Lundström, I. 2004, A&A, 415, 155
Benz, W., & Mayor, M. 1984, A&A, 138, 183
Bergemann, M., Lind, K., Collet, R., Magic, Z., & Asplund, M. 2012, MNRAS, 427, 27
Blackwell, D. E., Ibbetson, P. A., Petford, A. D., & Shallis, M. J. 1979a, MNRAS, 186, 633
Blackwell, D. E., Petford, A. D., & Shallis, M. J. 1979b, MNRAS, 186, 657
Blackwell, D. E., Petford, A. D., Shallis, M. J., & Simmons, G. J. 1982a, MNRAS, 199, 43
Blackwell, D. E., Petford, A. D., & Simmons, G. J. 1982b, MNRAS, 201, 595
Blackwell, D. E., Lynas-Gray, A. E., & Smith, G. 1995, A&A, 296, 217
Bridges, J. M., & Kornblith, R. L. 1974, ApJ, 192, 793
Bruntt, H. 2009, A&A, 506, 235
Bruntt, H., Bedding, T. R., Quirion, P.-O., et al. 2010, MNRAS, 405, 1907
Bruntt, H., Basu, S., Smalley, B., et al. 2012, MNRAS, 423, 122
Cantat-Gaudin, T., Donati, P., Pancino, E., et al. 2014, A&A, 562, A10
Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, Nature, 450, 1020
Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765
Chiavassa, A., Plez, B., Josselin, E., & Freytag, B. 2009, A&A, 506, 1351
Chiavassa, A., Haubois, X., Young, J. S., et al. 2010, A&A, 515, A12
Collet, R., Asplund, M., & Trampedach, R. 2007, A&A, 469, 687
Creevey, O. L., Thévenin, F., Boyajian, T. S., et al. 2012, A&A, 545, A17
De Laverny, P., Recio-Blanco, A., Worley, C. C., & Plez, B. 2012, A&A, 544, A126
De Medeiros, J. R., & Mayor, M. 1999, A&AS, 139, 433
De Medeiros, J. R., Udry, S., Burki, G., & Mayor, M. 2002, A&A, 395, 97
De Medeiros, J. R., Silva, J. R. P., Do Nascimento, Jr., J. D., et al. 2006, A&A, 458, 895
Dekker, H., D'Odorico, S., Kaufer, A., Delabre, B., & Kotzlowski, H. 2000, in SPIE Conf. Ser. 4008, eds. M. Iye, & A. F. Moorwood, 534
Diego, F., Charalambous, A., Fish, A. C., & Walker, D. D. 1990, in SPIE Conf. Ser. 1235, ed. D. L. Crawford, 562
Donati, J., Semel, M., Carter, B. D., Rees, D. E., & Collier Cameron, A. 1997, MNRAS, 291, 658
Edvardsson, B., Andersen, J., Gustafsson, B., et al. 1993, A&A, 275, 101
Freeman, K. C. 2010, in Galaxies and their Masks, eds. D. L. Block, K. C. Freeman, & I. Puerari (Springer Science + Business Media), 319
Fuhr, J. R., & Wiese, W. L. 2006, J. Phys. Chem. Ref. Data, 35, 1669
Fuhr, J. R., Martin, G. A., & Wiese, W. L. 1988, J. Phys. Chem. Ref. Data, 17
Fuhrmann, K. 1998, A&A, 338, 161
Garz, T., & Kock, M. 1969, A&A, 2, 274
Gilmore, G., & Reid, N. 1983, MNRAS, 202, 1025
Gilmore, G., Wyse, R. F. G., & Kuijken, K. 1989, ARA&A, 27, 555
Gilmore, G., Randich, S., Asplund, M., et al. 2012, The Messenger, 147, 25
Gray, D. F. 2005, The Observation and Analysis of Stellar Photospheres (Cambridge University Press)
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
Heiter, U., Kupka, F., van't Veer-Menneret, C., et al. 2002, A&A, 392, 619
Hekker, S., & Meléndez, J. 2007, A&A, 475, 1003
Helmi, A. 2008, A&ARv, 15, 145
Ivezić, Ž., Beers, T. C., & Jurić, M. 2012, ARA&A, 50, 251
Jofré, P., & Weiss, A. 2011, A&A, 533, A59
Jofré, P., Panter, B., Hansen, C. J., & Weiss, A. 2010, A&A, 517, A57
Kapteyn, J. C., & van Rhijn, P. J. 1920, ApJ, 52, 23
Kaufer, A., Stahl, O., Tubbesing, S., et al. 2000, in SPIE Conf. Ser. 4008, eds. M. Iye, & A. F. Moorwood, 459
Kupka, F., Piskunov, N., Ryabchikova, T. A., Stempels, H. C., & Weiss, W. W. 1999, A&AS, 138, 119
Kurucz, R. L. 2005, Mem. Soc. Astron. It. Suppl., 8, 14
Lebzelter, T., Heiter, U., Abia, C., et al. 2012, A&A, 547, A108
Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008a, AJ, 136, 2022
Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008b, AJ, 136, 2050
Lind, K., Bergemann, M., & Asplund, M. 2012, MNRAS, 427, 50
Luck, R. E. 1979, ApJ, 232, 797
Luck, R. E., & Heiter, U. 2005, AJ, 129, 1063
Luck, R. E., & Heiter, U. 2006, AJ, 131, 3069
Magrini, L., Randich, S., Friel, E., et al. 2013, A&A, 558, A38
Massarotti, A., Latham, D. W., Stefanik, R. P., & Fogel, J. 2008, AJ, 135, 209
Masseron, T. 2006, Ph.D. Thesis, Observatoire de Paris, France
May, M., Richter, J., & Wichelmann, J. 1974, A&AS, 18, 405
Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20
McWilliam, A. 1990, ApJS, 74, 1075
Meléndez, J., & Barbuy, B. 2009, A&A, 497, 611
Meléndez, J., Asplund, M., Alves-Brito, A., et al. 2008, A&A, 484, L21
Mishenina, T. V., & Kovtyukh, V. V. 2001, A&A, 370, 951
Molaro, P., & Monai, S. 2012, A&A, 544, A125
Mucciarelli, A., Pancino, E., Lovisi, L., Ferraro, F. R., & Lapenna, E. 2013, ApJ, 766, 78
Nissen, P. E. 1981, A&A, 97, 145
O'Brian, T. R., Wickliffe, M. E., Lawler, J. E., Whaling, W., & Brault, J. W. 1991, J. Opt. Soc. Am. B Opt. Phys., 8, 1185
Pagel, B. E. J., & Tautvaisiene, G. 1998, MNRAS, 299, 535
Pancino, E., Mucciarelli, A., Bonifacio, P., Monaco, L., & Sbordone, L. 2011, A&A, 534, A53
Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, A&A, 369, 339
Peterson, R. 1976, ApJS, 30, 61
Plez, B. 2012, Turbospectrum: Code for spectral synthesis, Astrophysics Source Code Library, ascl: 1205.004
Raassen, A. J. J., & Uylings, P. H. M. 1998, A&A, 340, 300
Ramírez, I., & Allende Prieto, C. 2011, ApJ, 743, 135
Ramírez, I., Allende Prieto, C., & Lambert, D. L. 2007, A&A, 465, 271
Randich, S., Gilmore, G., Gaia-ESO Consortium 2013, The Messenger, 154, 47
Recio-Blanco, A., Bijaoui, A., & de Laverny, P. 2006, MNRAS, 370, 141
Reddy, B. E., Tomkin, J., Lambert, D. L., & Allende Prieto, C. 2003, MNRAS, 340, 304
Reiners, A., & Schmitt, J. H. M. M. 2003, A&A, 412, 813
Richter, J., & Wulff, P. 1970, A&A, 9, 37
Saar, S. H., & Osten, R. A. 1997, MNRAS, 284, 803
Santos, N. C., Israelian, G., & Mayor, M. 2004, A&A, 415, 1153
Sbordone, L., Bonifacio, P., Castelli, F., & Kurucz, R. L. 2004, Mem. Soc. Astron. It. Suppl., 5, 93
Schönrich, R., Asplund, M., & Casagrande, L. 2011, MNRAS, 415, 3807
Schröder, C., Reiners, A., & Schmitt, J. H. M. M. 2009, A&A, 493, 1099
Schuler, S. C., King, J. R., Fischer, D. A., Soderblom, D. R., & Jones, B. F. 2003, AJ, 125, 2085
Schuster, W. J., Moreno, E., Nissen, P. E., & Pichardo, B. 2012, A&A, 538, A21
Siebert, A., Williams, M. E. K., Siviero, A., et al. 2011, AJ, 141, 187
Smiljanic, R., Porto de Mello, G. F., & da Silva, L. 2007, A&A, 468, 679
Sneden, C. A. 1973, Ph.D. Thesis, The University of Texas, Austin
Soubiran, C. 1993, A&A, 274, 181
Soubiran, C., Bienaymé, O., & Siebert, A. 2003, A&A, 398, 141
Soubiran, C., Le Campion, J.-F., Cayrel de Strobel, G., & Caillo, A. 2010, A&A, 515, A111
Sousa, S. G., Santos, N. C., Israelian, G., Mayor, M., & Monteiro, M. J. P. F. G. 2007, A&A, 469, 783
Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, A&A, 487, 373
Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645
Stetson, P. B., & Pancino, E. 2008, PASP, 120, 1332
Tabernero, H. M., Montes, D., & González Hernández, J. I. 2012, A&A, 547, A13
Thévenin, F., & Idiart, T. P. 1999, ApJ, 521, 753
Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, 47, 371
Torres, G., Fischer, D. A., Sozzetti, A., et al. 2012, ApJ, 757, 161
Tull, R. G., MacQueen, P. J., Sneden, C., & Lambert, D. L. 1995, PASP, 107, 251
Valenti, J. A., & Fischer, D. A. 2005, ApJS, 159, 141
Valenti, J. A., & Piskunov, N. 1996, A&AS, 118, 595
Villada, M., & Rossi, L. 1987, Ap&SS, 136, 351
Vogt, S. S. 1987, PASP, 99, 1214
Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, in SPIE Conf. Ser. 2198, eds. D. L. Crawford, & E. R. Craine, 362
Wiese, W. L., Fuhr, J. R., & Deters, T. M. 1996, Atomic transition probabilities of carbon, nitrogen, and oxygen: a critical data compilation (Washington, DC: Am. Chem. Soc.)
Wolnik, S. J., Berthel, R. O., & Wares, G. W. 1970, ApJ, 162, 1037
Wolnik, S. J., Berthel, R. O., & Wares, G. W. 1971, ApJ, 166, L31
Worley, C. C., de Laverny, P., Recio-Blanco, A., et al. 2012, A&A, 542, A48
Wu, Y., Singh, H. P., Prugniel, P., Gupta, R., & Koleva, M. 2011, A&A, 525, A71
Zamanov, R. K., Bode, M. F., Melo, C. H. F., et al. 2008, MNRAS, 390, 377
Zhao, G., Chen, Y.-Q., Shi, J.-R., et al. 2006, Chin. J. Astron. Astrophys., 6, 265
Zwitter, T., Siebert, A., Munari, U., et al. 2008, AJ, 136, 421