

Annual Review of Astronomy and Astrophysics
**Accuracy and Precision of
Industrial Stellar Abundances**

Paula Jofré,¹ Ulrike Heiter,² and Caroline Soubiran³

¹Núcleo de Astronomía, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Santiago, Chile; email: paula.jofre@mail.udp.cl

²Observational Astrophysics, Department of Physics and Astronomy, Uppsala University, 75120 Uppsala, Sweden

³Laboratoire d'Astrophysique de Bordeaux, CNRS B18N, Univ. Bordeaux, allée Geoffroy Saint-Hilaire, F-33615, Pessac, France

Annu. Rev. Astron. Astrophys. 2019. 57:571–616

First published as a Review in Advance on
July 5, 2019

The *Annual Review of Astronomy and Astrophysics* is
online at astro.annualreviews.org

<https://doi.org/10.1146/annurev-astro-091918-104509>

Copyright © 2019 by Annual Reviews.
All rights reserved

Keywords

stellar spectroscopy, Milky Way, catalogs

Abstract

There has been an incredibly large investment in obtaining high-resolution stellar spectra for determining chemical abundances of stars. This information is crucial to answer fundamental questions in astronomy by constraining the formation and evolution scenarios of the Milky Way as well as the stars and planets residing in it.

We have just entered a new era, in which chemical abundances of FGK-type stars are being produced at industrial scales, and in which the observations, reduction, and analysis of the data are automatically performed by machines. Here, we review the latest human efforts to assess the accuracy and precision of such industrial abundances by providing insights into the steps and uncertainties associated with the process of determining stellar abundances.

We also provide a description of current and forthcoming spectroscopic surveys, focusing on their reported abundances and uncertainties. This allows us to identify which elements and spectral lines are best and why. Finally, we make a brief selection of main scientific questions the community is aiming to answer with abundances.

- Uncertainties in abundances need to be disentangled into random and systematic components.
- Precision can be increased by applying differential or data-driven methods based on accurate data.

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

- High-resolution and signal-to-noise spectra provide fundamental data that can be used to calibrate lower-resolution and signal-to-noise spectra of millions of stars.
- Different survey calibration strategies must agree on a common set of reference stars to create data products that are consistent.
- Data products provided by individual groups must be published using standard formats to ensure straightforward applicability.

Contents

1. INTRODUCTION	572
2. FROM SPECTRA TO ABUNDANCES: STEPS AND ISSUES	573
2.1. Science: Selection of Stellar Sample and Chemical Elements.....	573
2.2. Observed Spectra	574
2.3. Physical Description	577
2.4. Analysis Methods	578
2.5. Line List.....	580
2.6. Stellar Parameters.....	584
3. ASSESSING THE ABUNDANCE ERROR BUDGET	586
3.1. Random Uncertainties.....	586
3.2. Systematic Uncertainties	588
3.3. Reference Objects and Biases	590
3.4. Improving Precision.....	592
3.5. Combination of Uncertainties	593
4. THE PERIODIC TABLE AS SEEN FROM SPECTRAL ANALYSES	593
4.1. Catalogs of Stellar Abundances from High-Resolution Studies.....	594
4.2. Chemical Abundances of Spectroscopic Surveys.....	596
4.3. Discussion on Individual Elements.....	605
5. RECONSTRUCTING THE HISTORY OF THE GALAXY WITH ABUNDANCES	607
5.1. Nucleosynthesis Channels and Chemical Dimensions.....	607
5.2. From Chemical Tagging to Galactic Phylogenetics.....	608
5.3. New Challenges for Chemical Evolution with High-Precision Abundances...	608
5.4. Masses and Ages from Stellar Abundances	609
6. DISCUSSION AND CONCLUDING REMARKS.....	610

1. INTRODUCTION

The elemental abundances of FGK-type stars provide key pieces of information for characterizing the stellar populations of our Galaxy. Different stellar populations have different chemical patterns, and the foundation for explaining these differences is well established: Chemical elements are created in a variety of nucleosynthesis channels inside stars and are distributed into the Galaxy through either supernovae or stellar winds. New stars are born from this enriched material, creating new elements that are then sent back to the interstellar medium (ISM). This cycle has been repeating from the formation of the first stars until today.

The outcome of Galactic chemical evolution is more complex than what is implied by the simple description above, considering the variety of stellar masses and therefore lifetimes, and the

diversity of physical processes taking place inside stars. Therefore, accurate and precise abundances of large samples of stars are required to constrain chemical evolution models. The productions of elements (yields) are different for stars at different masses and metallicities; the amount of the enriched material recycled into new stars depends on the total mass of the Galaxy because it must be able to keep the gas bound to form new stars. Because the masses and the sizes of galaxies change with time, so does the star-formation rate and the subsequent chemical enrichment. Finally, we know that galaxies experience inflow and outflow of material due to, e.g., accretion of other galaxies, which have different chemical enrichment histories and stellar populations with other chemical patterns [see, e.g., Kobayashi et al. (2006) for a description of the ingredients in chemical evolution models]. FGK-type low-mass stars live long enough and have shallow convective zones such that the information on the chemical make-up of the gas from which they formed is retained in their spectra. Hence, their abundances are the best fossil records we can use to constrain the cosmic matter cycle. However, these fossils move about in the Galaxy. With the help of a dynamical model and the ages of the stars, it might be possible to find their original site of formation (Freeman & Bland-Hawthorn 2002). Then the fossils of a stellar population might be found and the ingredients of its chemical evolution constrained (Feltzing & Chiba 2013, and references therein).

First works putting these pieces together were limited by the lack of good measurements of distances, which did not allow probing the distribution of chemical elements in the Galaxy. Even if stellar abundances were believed to be of reasonable accuracy, it was not possible to constrain a chemodynamical model with the scarcity of data on distances, kinematics, and ages. However, these very struggling scientists provided the motivation for the projects that are responsible for the wealth of stellar data we have today, starting from the revolutionary *Gaia* mission (Gaia Collab. et al. 2018b) and followed by the large spectroscopic and asteroseismic surveys. This industrial revolution in Galactic astronomy is only beginning, as more data releases from *Gaia* are approaching and more spectroscopic and seismic surveys are planned.

Newer generations of scientists have the opportunity to work with these ready-to-use data products. Now that the major challenge of good measurements of distances is largely solved thanks to *Gaia*, do we believe that stellar abundances are of sufficient accuracy? High-resolution multi-object spectrographs are restricted to point toward the sky each from a different spot on the surface of the Earth, with different instruments. It is natural that the data products from different surveys will differ, but how is that limiting our capacity to unravel the structure and formation of our Galaxy? This review answers some of these questions, starting with an overview of the major steps involved in the derivation of stellar abundances in Section 2, followed by Section 3, in which we suggest standard ways to quantify the uncertainties in the results. In Section 4, we summarize the large data sets with abundances available today, discussing why we know more about some elements than others. We continue by reviewing the progress the field has experienced thanks to stellar abundances in Section 5, and finish with a discussion answering these questions and some thoughts on the future in Section 6.

2. FROM SPECTRA TO ABUNDANCES: STEPS AND ISSUES

We complement the brief but comprehensive review by Allende Prieto (2016b) by illuminating the main steps in the process of deriving stellar abundances that are illustrated in **Figure 1**. Publicly available tools and material are listed in **Table 1**.

2.1. Science: Selection of Stellar Sample and Chemical Elements

The scientific question determines the type of stars to study and dictates the properties of the spectra, in particular their wavelength coverage. In some cases, a carefully selected sample of few

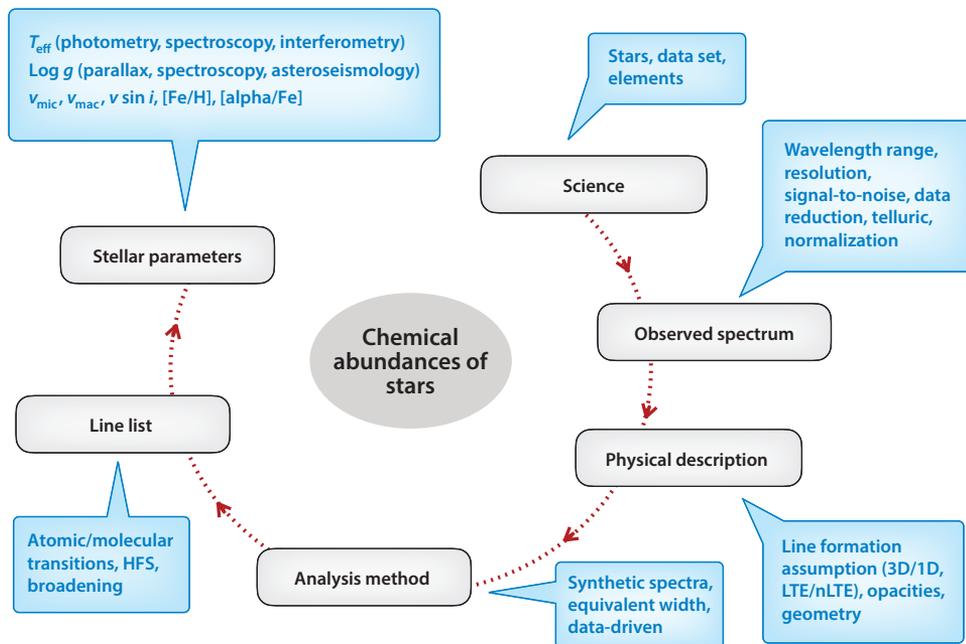


Figure 1

Illustration of the steps needed for obtaining abundances of chemical elements in stars. Each of these steps implies uncertainties in the derived abundances that might affect different science cases to different degrees. Abbreviations: HFS, hyperfine structure; LTE, local thermodynamic equilibrium; nLTE, non-local thermodynamic equilibrium.

stars is sufficient to produce revolutionary scientific results (e.g., Fuhrmann 1998, Meléndez et al. 2009, Nissen & Schuster 2010); the increase in computing power and efficiency of data storage has been driving the field to evolve toward a more industrial scale. This is especially the case for studies of the Milky Way structure and evolution, which is the objective of several ongoing large-scale spectroscopic surveys. Ruchti et al. (2016) present an interesting discussion of how to define a spectral data set that will meet these science goals.

2.2. Observed Spectra

The properties of the spectra, as well as the manipulations on the spectra that are required to analyze them, affect the final results. Here, we explain a selection of examples, which is not exhaustive.

2.2.1. Resolution, signal-to-noise ratios, and time dependencies. Spectral lines will be resolved if the instrumental broadening is less than the broadening mechanisms in the stellar atmosphere caused primarily by Doppler broadening due to temperature, turbulence, and rotation (see Nissen & Gustafsson 2018, their section 2). If many stars need to be analyzed, the instrumental resolution does not need to be much higher than that corresponding to the intrinsic stellar one for the purpose of determining abundances as this saves significant observing time. However, a higher resolution allows one to investigate effects on line profiles such as star spots, asymmetries due to convection, variations due to nonradial pulsations, or blends. **Figure 2** compares spectra with different resolutions for ϵ Eri. The resolving power, signal-to-noise

Table 1 Material and tools for spectral analyses that are publicly available^a

Material	Reference	Comment
SVO Montes	Spectral libraries http://svo2.cab.inta-csic.es/theory/libtest/index.php https://webs.ucm.es/info/Astrof/invest/actividad/spectra.html	Public libraries Compilation
MARCS ATLAS9 STAGGER CO5BOLD	Model atmospheres Gustafsson et al. (2008) Castelli & Kurucz (2003) Magic et al. (2013) Freytag et al. (2012)	1D spherical geometry 1D plane-parallel geometry 3D 3D
Turbospectrum MOOG SYNTHE SPECTRUM DETAIL/SIU	Radiative transfer codes Plez et al. (1992) Sneden (1973) Kurucz (1993) Gray & Corbally (1994) e.g., Bergemann et al. (2012) via http://nlte.mpia.de/	LTE LTE LTE LTE Non-LTE
VALD NIST ASD Sneden et al. Linemake ExoMol BRASS Barklem Kurucz VAMDC	Line lists Ryabchikova et al. (2015) https://www.nist.gov/pml/atomic-spectra-database https://www.as.utexas.edu/~chris/lab.html https://github.com/vmplacco/linemake Tennyson et al. (2016) Laverick et al. (2018) Barklem et al. (2015) Kurucz (2011) http://www.vamdc.eu	Literature compilation Literature compilation Bibliography and molecular line lists Synthetic spectrum lists in MOOG style Very cool objects Centralization of sources Broadening cross-sections Atomic data Electronic infrastructure
AMBRE STAGGER 3D-non-LTE Balmer APOGEE POLLUX	Grids of synthetic spectra de Laverny et al. (2012) Chiavassa et al. (2018) Amarsi et al. (2018b) Mészáros et al. (2012) Palacios et al. (2010)	Optical high resolution CaII triplet centered Balmer lines centered IR Database
SME iSpec FERRE GALA DOOp ARES The Cannon	Automatic codes for the determination of abundances Piskunov & Valenti (2017) Blanco-Cuaresma et al. (2014a) García Pérez et al. (2016) Mucciarelli et al. (2013) Cantat-Gaudin et al. (2014) Sousa et al. (2015) Ness et al. (2015)	With non-LTE on the fly Python wrapper for various tools Match models to data EW code Wrapper for EWs Automatic EWs Label transfer from a training set
INSPECT MPIA	Non-LTE abundance corrections http://inspect-stars.com/ http://nlte.mpia.de/	Line-by-line corrections Line-by-line corrections

^aThis compilation is restricted to tools that are regularly updated and available on the web. Other codes not listed here might be equally suitable and available upon request to their authors. Abbreviations: AMBRE, Archéologie avec Matisse Basée sur les Archives de l'ESO; APOGEE, Apache Point Observatory Galaxy Evolution Experiment; ASD, Atomic Spectra Database; EW, equivalent width; LTE, local thermodynamic equilibrium; NIST, National Institute of Standards and Technology; SVO, Spanish Virtual Observatory; VALD, Vienna Atomic Line Database; VAMDC, Virtual Atomic and Molecular Data Centre.

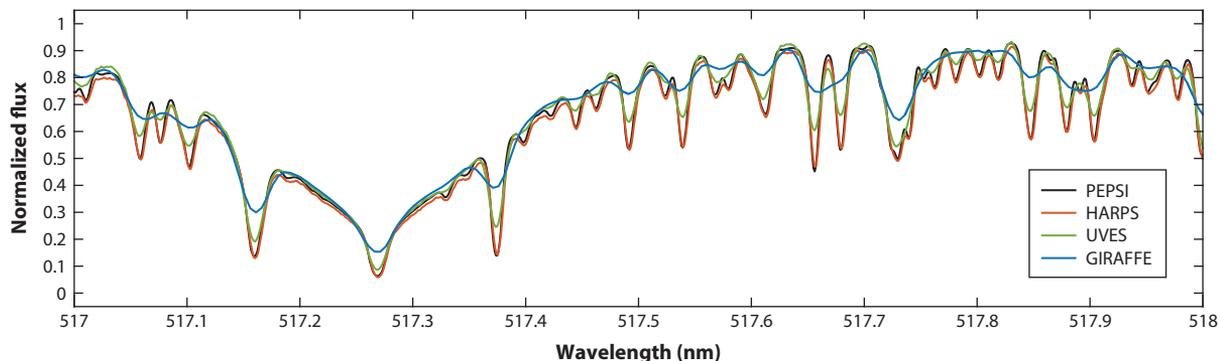


Figure 2

Spectra for the *Gaia* benchmark star ϵ Eri near one of the MgI b lines, obtained with the PEPSI (Potsdam Echelle Polarimetric and Spectroscopic Instrument), HARPS (High Accuracy Radial velocity Planet Searcher), UVES (Ultraviolet and Visual Echelle Spectrograph), and GIRAFFE spectrographs.

(S/N), and reference for each spectrum are, respectively, $\sim 25,900/330/Gaia$ -ESO DR3¹ for GIRAFFE, $\sim 47,000/173/Gaia$ -ESO DR3 for UVES (Ultraviolet and Visual Echelle Spectrograph), $\sim 115,000/474/Blanco$ -Cuaresma et al. (2014b) for HARPS (High Accuracy Radial Velocity Planet Searcher), and $\sim 200,000/1350/Strassmeier$ et al. (2018) for PEPSI (Potsdam Echelle Polarimetric and Spectroscopic Instrument).

Figure 2 shows how the number of spectral features grows with increasing spectral resolution. For example, at around 517.9 nm a broad feature is visible in the GIRAFFE spectrum, which is resolved into two components in the UVES spectrum, whereas the HARPS and PEPSI spectra show a blend of at least four lines. These are identified as being due to TiI, FeI, VI, NiI, and several MgH lines by comparison with a synthetic spectrum.

A signal-to-noise ratio of $S/N > 200$ might be desirable to determine abundances with high confidence, but in general a S/N of 60–100 is good enough for pipelines to derive accurate abundances of most elements for FGK-type Population I and II stars. Below $S/N \sim 40$, abundances become more uncertain, and below 20 they are usually considered to be unreliable (see, e.g., Heiter et al. 2014, Smiljanic et al. 2014). The latter may be the case for a considerable fraction of current spectroscopic surveys (see Section 4.2). Spectra of very low resolution or S/N are rarely able to provide information for a large variety of abundances, although new techniques implementing machine-learning approaches are promising (Leung & Bovy 2018, Ness 2018). They still rely on a training set for which abundances are known from more classical methods calibrated with high-resolution and high S/N spectra. It is important to keep in mind that S/N depends on wavelength, where the blue part of the spectrum often has much lower S/N than the red part.

2.2.2. Data reduction issues. Modern high-resolution spectra are commonly taken with cross-dispersed echelle spectrographs, which provide efficient access to an extended wavelength coverage. However, reducing and extracting the science spectrum from such images is challenging. A seemingly obvious requirement for science-ready spectra is a proper wavelength calibration, but recurring cases of failure or inaccuracies in wavelength calibration have been reported, which may affect abundance analyses (e.g., Hinkel et al. 2016, their section 5.2). The central wavelengths of spectral features must be accurate enough to achieve a match with the line list. Automatic

¹Data obtained from *Gaia*-ESO data release (DR) 3, available at http://archive.eso.org/wdb/wdb/adp/phase3_main/form?collection_name=GAIAESO&release_name=DR3.

procedures must be able to deal with an imperfect wavelength solution by applying a wavelength-dependent radial-velocity correction or a sufficiently wide search window for identifying lines.

Merging the orders is another major challenge in extracting echelle spectra. Piskunov & Valenti (2002) provide a detailed description of this process. Orders are curved, thus deviating from the straight lines defined by the pixel rows or columns of a CCD detector. It is not possible to merge them into a 1D spectrum by interpolating among adjacent orders, as this adds correlations that affect the extracted spectrum. The response of each order to the blaze function varies, and the consequent variation of count levels along and across orders can be prohibitively large. Also, if the resolution is high, some orders are very narrow, making it difficult to deblaze the spectra owing to the presence of strong lines, such as the Balmer lines. Unfortunately, there is no standard or perfect way to merge orders. Some of these challenges are examined extensively by Prugniel & Soubiran (2001).

2.2.3. Removal of telluric features. Part of the stellar light passing through the Earth's atmosphere is absorbed, causing the so-called telluric lines in the spectrum. They can be very strong and have fixed positions in wavelength, but not all features are identified. Atlases of telluric standards or models exist (see discussion by, e.g., Bertaux et al. 2014) and are cross-correlated with the science spectrum to identify and, if possible, remove these lines (Sameshima et al. 2018). There are regions in the spectrum that are more affected by telluric absorption, notably toward the red part of the spectrum (Kos et al. 2017, their figure 16).

2.2.4. Normalization. Because flux calibration is very challenging for high-resolution spectra, it has become customary to determine stellar abundances from spectra normalized to the continuum flux, effectively using the relative strengths of the absorption lines. There is no established standard way to normalize a spectrum to the continuum, although procedures such as the `continuum` task in the IRAF software system are very popular. In general, the pseudocontinuum is determined after fitting a spline or a polynomial to a set of regions that are believed to be free of absorption lines (see, e.g., Prugniel & Soubiran 2001). We stress the term *believed*, because it is not certain that areas of free absorption exist at all across a given spectrum. Examples in which continuum normalization is especially complicated are very cool stars, which have spectra crowded with molecular features, spectra with too low S/N, or spectra whose orders are not properly merged. Other regions that are particularly challenging for FGK-type stars are around the Balmer lines, especially for high-resolution echelle spectra. The definition of the continuum may in fact be responsible for the largest fraction of the uncertainty in abundances (e.g., Jofré et al. 2017c).

2.3. Physical Description

Very nice summaries of various aspects of the physical description of line formation theory can be found in reviews by Asplund (2005), Allende Prieto (2016b), and Nissen & Gustafsson (2018). In this section, we disentangle these aspects and discuss their importance in abundance determination.

2.3.1. Line formation. A spectral line is usually studied with the help of a radiative transfer calculation and a model for the atmosphere. It is clear that 3D-non-local thermodynamic equilibrium models are the way forward for obtaining accurate absolute abundances. However, one should be aware that these models focus on improving certain aspects of atmospheric physics (geometry and statistical equilibrium), whereas other aspects, such as the treatment of opacities by sampling and binning, can still be quite uncertain (see below). Asplund (2005) makes the interesting point that

just because local thermodynamic equilibrium (LTE) is the standard method of analysis does not mean that departures from LTE only occur occasionally. Yet, the simplistic 1D-LTE models are still the most-used ones when abundances are derived industrially.

The reason might be that we now have a reasonable understanding of the applicability and failure of 1D-LTE. This is mostly thanks to the progress in 3D-non-LTE modeling, which enables quantification of the differences with respect to 1D-LTE models. There has been impressive progress in our abilities to determine abundances of stars considering 3D and non-LTE. Examples of such works are the series of papers on modeling Fe in full 3D-non-LTE in late-type stars (Lind et al. 2017, and references therein), the studies focused on characterizing ultra-metal-poor stars (e.g., Ezzeddine et al. 2017), or the series of works characterizing homogeneously the Milky Way halo with non-LTE (Mashonkina et al. 2017). Non-LTE corrections for abundances of hundreds of lines of several elements, as well as grids of 3D atmospheric models and synthetic spectra, are publicly available (see **Table 1**). In 1D-LTE, the parameters micro- and macroturbulence account for the turbulent motions of particles and need to be specified for modeling the lines. They are not needed when a full hydrodynamical simulation of the atmosphere is performed. With 3D models it is possible to find empirical relations for these parameters as a function of stellar parameters (Steffen et al. 2013). Such resources are important, as they allow identification of the conditions (and lines) for which the differences between 1D-LTE and 3D-non-LTE are minimal. Thus, spectral analyses can be calibrated to yield accurate results even under the assumption of 1D-LTE. The great difficulty is that, for many elements, especially those that produce few lines, no “3D-non-LTE free lines” are available. To achieve high precision and accuracy by taking advantage of all available lines in the spectrum, it is thus of paramount importance to work toward providing the prerequisites for modeling lines of more species in 3D and non-LTE.

2.3.2. Geometry and opacities. In 1D models (and 3D stellar surface simulations of the “box in a star” type), the geometry can be plane-parallel or spherical for each layer of the atmosphere. Both models are of comparable accuracy for dwarfs, but for giants and supergiants, which have extended atmospheres, curvature needs to be taken into account. Abundances have been compared for 1D models with different geometry for several elements and lines by Heiter & Eriksson (2006), who found that strong lines of high excitation potential are most affected.

Another central issue in modeling atmospheres is accounting for all possible opacity sources. They can be divided in continuous (produced by bound-free and free-free transitions) and line (produced by bound-bound transitions) opacities. The contributors are hydrogen atoms, metal atoms, and molecules. In cool stars, molecules are especially problematic, because they are poorly known [see Plez et al. (1992) and Masseron et al. (2014) for details]. Gustafsson et al. (2008) present further discussion on this subject; they test the structural effect on MARCS atmosphere models for different temperatures and optical depths when including and excluding opacities due to H, metals, and molecules. The effect in stellar abundances using solar or nonsolar opacities has been studied by Jofré et al. (2017c) and Saffé et al. (2018). Although the effects are very small, they might be important in high-precision abundance studies.

2.4. Analysis Methods

The classical and most common methods to determine abundances are based on the measurement of equivalent widths (EWs) or the computation of synthetic spectra of absorption lines of the chemical element in question. Recently, machine-learning approaches for measuring abundances have been introduced and applied to stellar surveys; they are further discussed in Section 4.2. As of today, EWs and syntheses are still the dominant methods to determine abundances, especially

ANALYSIS METHODS: EQUIVALENT WIDTH OR SYNTHESIS?

Both types of methods are good competitors, and it is not clear which of them performs best. Reports on comparisons of these methods can be found within the *Gaia*-ESO framework (Smiljanic et al. 2014), as well as in the series of works on the *Gaia* benchmark stars (e.g., Jofré et al. 2015) and work by Hinkel et al. (2016), or by Casamiquela et al. (2017). Using EWs or synthesis with high-resolution, high S/N spectra of solar-like stars often seems to be a decision of personal preference. Syntheses might have more applicability in crowded spectral regions, or in stars with broad lines. Today, computers can quickly synthesize spectra, so the computing time is not the limiting factor as it used to be a decade ago. It is possible that, in the next decades, syntheses will become the preferred way to measure abundances, but EWs should not be set aside completely, as they are the simplest tool to measure the strength, and hence understand the nature, of the lines under analysis.

because machine-learning methods continue to rely on training sets of stars with well-known abundances, which most likely are measured or calibrated from EWs or synthesis methods. (See the sidebar titled Analysis Methods: Equivalent Width or Synthesis?)

2.4.1. Equivalent width. EWs are obtained from either fitting a Gaussian profile for weak lines and Voigt profiles for stronger lines, or just by integrating over the line profile. The latter becomes more accurate when lines have a boxy-shaped profile due to, e.g., hyperfine structure (HFS) components (see Section 2.5). The EW is thus a measure of the strength of the line, which can be directly related to the chemical abundance of the element in a star given the stellar parameters based on the so-called curve of growth (CoG): For weak lines there is a linear increase of abundance with EWs (in a logarithmic sense). Stronger lines lie on the flat part of the CoG: They are saturated, and thus there is no direct relationship of the abundance with the EW. Note that the definition of weak and strong lines might vary from star to star, and therefore one has to select the lines that lie in the linear part of the CoG in each case individually.

The dominant source of uncertainty in EW methods is the placement of the continuum. Today, typical automatic codes still cannot identify the continuum as precisely as can be done by hand using, e.g., the `sp1ot` task of IRAF, especially when the spectra are crowded with stellar features or artifacts due to data reduction. Experienced spectroscopists may be able to identify the continuum for such challenging lines “by eye,” making this process rather more an art than an objective task. Measurements by hand are usually limited to high-precision abundances of small samples of stars (Nissen et al. 2017, Bedell et al. 2018). In this era of industrial stellar abundances, EWs are best measured with automatic pipelines. The uncertainties are probably larger than for the manual measurements, but they can be reproduced and quantified, and so can their effect on derived abundances. A serious limitation of determining abundances from EWs is that, if lines are blended, the abundance will be overestimated. Thus, EWs work best for very high-resolution and high S/N spectra. Likewise, intrinsic broadening of lines contributes to blending, and so EW methods work best for relatively warm stars with slow rotation. Most spectroscopic surveys are designed to obtain spectra of stars in which these conditions are met, and in these cases it is safe to use the EW method.

2.4.2. Synthesis. The abundance of the element is varied until the best fit of a synthetic line profile with respect to the observation is found. Syntheses can be computed on-the-fly for each line and star until the best fit is obtained. It is also possible to use precomputed grids of synthetic spectra with varying abundances for different sets of stellar parameters (García Pérez et al. 2016).

Syntheses on-the-fly have the advantage that they can be easily adapted to different spectra and lines. This freedom allows one also to identify stars with unusual chemical abundances. The disadvantage is that, when large samples of stars need to be analyzed, the analysis can be very time-consuming. This might be especially inefficient when the stars are very similar to each other, like those targeted by spectroscopic surveys. Ting et al. (2018) present a solution to overcome this problem by interpolating between models.

Syntheses are the preferred method when spectra are crowded with absorption features, which is the case for cool stars. In addition, they are the only way to measure abundances from molecules (Roederer et al. 2014) or from very blended lines, for example. This is because the wavelength region to be fitted can be set to intervals of arbitrary size; therefore, abundances are not restricted to being measured from individual lines that have a well-behaved shape. The disadvantage with respect to EWs is that they depend on the instrumental profile (e.g., the spectral resolution needs to be known), and every pixel is fitted, which means that the results are sensitive to an imperfect wavelength calibration (Hinkel et al. 2016, Jofré et al. 2017c), for example.

2.5. Line List

When deriving abundances from absorption lines, it is assumed that the line strength is directly related to the abundance of the element whose transition produces the measured line. The wavelengths and transition probabilities, as well as the properties of the atomic states responsible for these transitions, are stored in a line list. The accuracy of the atomic parameters has become one of the major sources of uncertainty in abundance determination. Significant efforts are being dedicated by laboratory spectroscopists and theorists to provide the needed data for transitions of many elements and species. This is tedious and challenging work, exemplified by the fact that only about half of the lines in the optical wavelength range (480 to 680 nm) that are often used for abundance analysis of solar-type stars have good laboratory transition probabilities; that is, they have typical uncertainties of 10% or better. Furthermore, current lists of lines with good wavelengths contain only half of the lines observed in good quality solar spectra (Kurucz 2014). The situation becomes especially problematic at cool temperatures, where molecular lines dominate over atomic lines in the spectra. The line data are less complete for wavelength ranges outside the optical, such as the UV and the IR. Here, we discuss a selection of issues related to the line list that are important when deriving abundances.

2.5.1. Transition data. The most fundamental information in the line list is the transition probability, often presented in the form of gf -values (product of statistical weight and oscillator strength). When these values are not known accurately, it is common to perform an astrophysical calibration: deriving the oscillator strength for a line by setting the abundance of an element to a reference value and fitting a synthetic to an observed spectrum by varying the gf -value. Usually this is done for the solar spectrum, for which the chemical composition is known with the highest accuracy. Boeche & Grebel (2016) present a detailed discussion on calibrating gf -values based on several *Gaia* benchmark stars. From a comparison with accurate laboratory measurements, they conclude that the final calibrated values may be subject to systematic uncertainties caused by normalization, line fitting procedures, 3D-non-LTE effects, errors in the stellar parameters, and the solar abundances adopted. Although using astrophysically calibrated atomic data has been shown to improve the precision of stellar abundance results on several occasions, it is not obvious that these results are accurate. Calibrating atomic data in this way offers a temporary solution until direct and accurate measurements in the laboratory become available for all lines in stellar spectra.

Experimental and theoretical data for atomic and molecular transitions are made available through online collections and databases, such as those by R.L. Kurucz, at the NIST

(National Institute of Standards and Technology), or the VALD (Vienna Atomic Line Database) (cf. **Table 1**). A major step toward standardized access and distribution of atomic data is done by the VAMDC (Virtual Atomic and Molecular Data Centre) Consortium, which maintains an electronic infrastructure providing access to about 30 databases simultaneously, together with tools and policies that aim to enhance the citation rate of individual data producers.

These databases contain further data that are needed to calculate synthetic spectra, in particular parameters that describe line broadening [see Barklem (2016) for a recent review]. Apart from the natural broadening due to the finite lifetimes of atomic states, the most important broadening process is collisions with neutral hydrogen, which can be described with different recipes. This includes the approximate formulation based on the van der Waals potential from the 1940s and 1950s (Unsöld recipe), and the more detailed theory by Anstee, Barklem & O'Mara from the 1990s [ABO theory, Barklem et al. (2015); further discussion can be found in Barklem (2016)]. An example of the effect on abundances of using the Unsöld recipe versus the ABO theory is given by Sobek et al. (2007). For 58 Cr I lines with a mean EW of 40 mÅ, the change in the mean solar Cr I abundance was 0.02 dex.

An additional complication arises from the presence of HFS components in individual atomic lines for species with odd baryon numbers (nonzero nuclear spin; for Solar System isotopic abundances these correspond mostly to elements with odd atomic numbers). The HFS parameters from which the exact positions of the components in wavelength can be calculated represent another type of atomic input data, whereas the relative intensities of the components are directly computed from quantum numbers. When unresolved, HFS can be regarded as an additional broadening mechanism, changing both the shape of the line profile and the total line intensity. The effect is larger for strong lines, because they may be desaturated. There is extensive literature studying the effects of HFS on abundances (see Battistini & Bensby 2015 and Jofré et al. 2017c for some examples).

Similarly, for atoms with several stable isotopes, the different atomic masses split the energy levels and, thus, a given transition into several components with a different wavelength for each isotope. In this case, the relative intensities of the components only depend on the isotopic composition under consideration. At Solar System composition, there is typically one dominating isotope for each element, thus the effect is mostly negligible, with the notable exception of Cu (with about two-thirds of ^{63}Cu and one-third of ^{65}Cu).

Finally, line lists need to include transition data for molecules as well as atoms. For molecules, we rely to a greater extent on theoretical calculations than for atoms, with correspondingly larger uncertainties in data quality. In G- and K-type stars, the transitions of diatomic molecules play an important role. For example, for the *Gaia*-ESO survey (GES; see Section 4.2.2), data for twelve different molecules of this kind are provided (27 isotopologs, mainly hydrides and carbon-bearing species). The main purpose of including molecular lines in the abundance analysis is to identify and account for blends affecting atomic lines. However, for some elements, in particular C, N, and O, molecular features are also used for abundance determination and to determine isotopic ratios. Masseron et al. (2014) illustrate the effect of including transitions of CH in calculated spectra at wavelengths bluer than ~ 450 nm, for the Sun and four metal-poor stars, showing a significant improvement when comparing to observed spectra.

2.5.2. Line selection. Ideally, one should select lines that have a wide range in strength and are spread out over the spectrum, i.e., at different wavelengths and excitation potentials. This helps to avoid systematic effects of any variations in spectral response and to probe different parts of the atmosphere. Furthermore, one should select lines at different ionization stages, as these show different sensitivity to changes in atmospheric pressure. If the analysis is accurate, the abundances

derived from every line should be consistent, allowing one to provide an average of the results obtained for each line as the final abundance. In reality, in many cases few lines are available and an average might not be accurate.

For example, Hawkins et al. (2016b) provide a comparison of titanium abundances from different lines in spectra from the APOGEE (Apache Point Observatory Galaxy Evolution Experiment) survey (Section 4.2.3). The main result is illustrated in **Figure 3**, where the $[\text{Ti}/\text{Fe}]$ abundances as a function of $[\text{Fe}/\text{H}]$ are shown for four different lines. Note that one of them was not detected

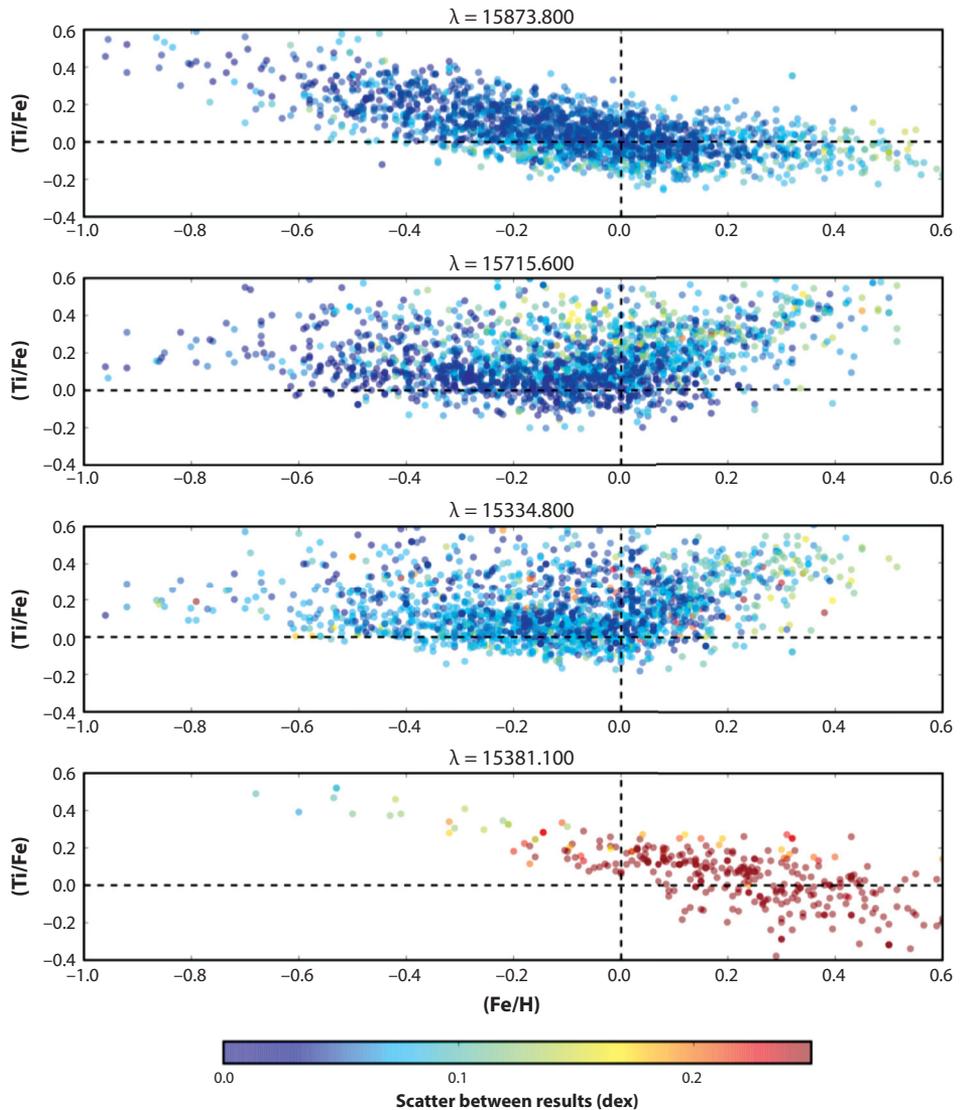


Figure 3

$[\text{Ti}/\text{Fe}]$ ratios as a function of metallicity determined from different Ti lines selected from APOGEE (Apache Point Observatory Galaxy Evolution Experiment) spectra. Colors indicate the scatter of four different methods used to derive abundances. Each panel shows a different trend, showing that line selection plays a crucial role in the final abundances. Adapted from Hawkins et al. (2016b) with permission © ESO.

for a large portion of the stars. The colors in **Figure 3** show the scatter among the different methods that were employed to derive the abundances. Among the three lines that were detected in the bulk of the stars, only one shows the expected trend with $[\text{Fe}/\text{H}]$, similar to that of other α elements, whereas the trends of the other two lines are very different. As a possible explanation, the authors mention non-LTE or saturation effects, as both lines are very strong. The titanium abundances published in the APOGEE data releases are based on a different line list, with astrophysically calibrated atomic data, and a different analysis method (see, e.g., Shetrone et al. 2015, Holtzman et al. 2018, and Section 5.12 in Jönsson et al. 2018). Therefore, the findings by Hawkins et al. (2016b) cannot be directly applied to assess the abundance data from the APOGEE survey.

Chromium abundances in metal-poor stars are also quite sensitive to line selection. Lawler et al. (2017) used ~ 40 CrI lines and 75 CrII lines to derive the abundance of the metal-poor star HD 84937. The mean abundance of CrI was almost 0.1 dex lower and the dispersion about twice as large compared to CrII. The discrepancy decreased to less than 0.05 dex when the 6 CrI resonance lines were removed, with a corresponding improvement in dispersion (becoming similar to that of CrII). The authors note that half of the CrI resonance lines (the triplet at $\lambda \sim 4275$ Å) have often been employed in abundance studies of metal-poor stars. The remaining discrepancy between CrI and CrII line abundances, mainly seen at wavelengths >4000 Å, can be ascribed to non-LTE effects in CrI, as studied by Bergemann & Cescutti (2010), who did not include the resonance lines.

In addition to issues in either atomic data or physical assumptions for the line formation, misidentification of the continuum or unidentified blends may also affect the selected lines. For a given observing time, it is more difficult to obtain good S/N in the blue parts of the spectrum than in the redder parts. Furthermore, the blue region contains more absorption lines (hence blending is more severe for more metal-rich stars in that region). The wavelength coverage of selected lines thus might have a strong dependency on stellar type and metallicity.

Even though the line selection may follow the reasoning discussed above, different criteria for “problematic lines” may be defined by different methodologies. An example for the variation of line selection is given in **Figure 4**, which visualizes the line selection within the GES (Section 4.2.2) for CaI. All groups performing the analysis were provided with the same set of observed spectra and the same line list containing 31 CaI lines. Nevertheless, the number of stars for which each group determined an abundance for each line varied significantly. There are a few lines that were consistently used (e.g., 5513, 6166 Å) or discarded (e.g., 6463, 6799 Å) by all groups, whereas others were employed by only a subset of the groups (e.g., 5260, 5582 Å).

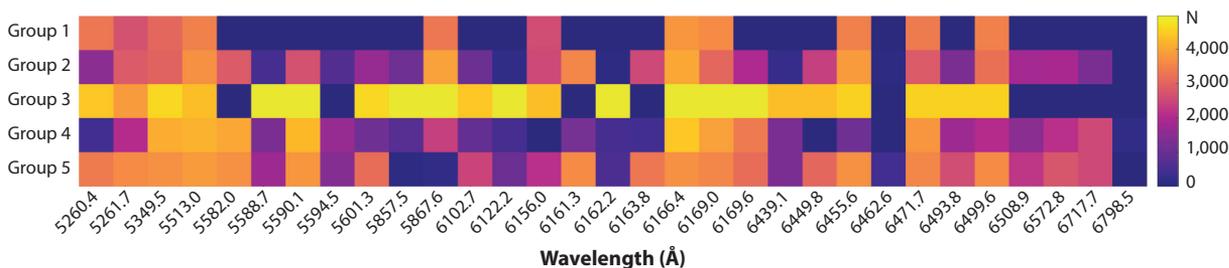


Figure 4

CaI-line selection in the *Gaia*-ESO survey. Color coding represents the number of stars for which an abundance was determined for each line by different analysis groups participating in the internal data release 5. Based on data provided by R. Smiljanic (private communication) reproduced with permission of *Gaia*-ESO.

2.6. Stellar Parameters

The relationships of line strength and abundance depend on the stellar parameters. A natural approach is to determine the parameters “consistently” with the abundances, e.g., from the same spectra, line lists, prescription, method, etc. However, this is not necessary as in some cases accurate parameters can be determined with methods that are independent from spectroscopy. We disentangle the parameters into atmospheric (effective temperature, surface gravity, and metallicity) and other parameters, which account for line broadening and nonsolar abundance scales.

2.6.1. Atmospheric parameters. The PASTEL catalog (Soubiran et al. 2016) is a valuable resource for learning about the complexity of stellar parameter determination. The catalog contains a collection of more than 1,000 bibliographic resources of reported stellar parameters of more than 30,000 stars, determined from any of the methods discussed in **Supplemental Text 1**, and shows the inhomogeneity of stellar parameters resulting from different studies.

An investigation of the values in the PASTEL catalog shows that differences of 200–300 K in effective temperature (T_{eff}) are usual for FGK-type stars analyzed by different methods. For the stars in PASTEL with more than 25 T_{eff} determinations, a typical difference of 50 K is obtained. This suggests that it is today not possible to know the temperature of a star better than this accuracy. Many efforts are invested in obtaining more accurate temperatures of stars, because of the variety of astrophysical applications that depend on a temperature scale; however, no conclusion has yet been reached as to which method should be employed. Methods that are often used are the IR flux method, excitation balance, fitting of Balmer lines, and interferometry, which are explained in detail in **Supplemental Text 1**.

The effect of surface gravity (hereafter $\log g$) on the spectra is weaker compared to T_{eff} , which poses a challenge to constraining this parameter spectroscopically. It is difficult to determine $\log g$ to better than 0.1 dex in FGK-type stars. In the PASTEL catalog, the typical reported errors in the literature are of that size, which agrees with the median difference in $\log g$ obtained from independent works on the same stars. The comparison between APOGEE and LAMOST (Large Sky Area Multi-Object Fiber Spectroscopic Telescope) by Anguiano et al. (2018) shows that $\log g$ has a scatter of 0.25 dex among these surveys. Common methods to derive $\log g$ are the parallax method, ionization balance, fits of strong lines, and asteroseismology. These methods are also explained in **Supplemental Text 1**.

Metals influence both the strength of spectral lines and the continuous opacities in cool stars mainly through the abundance of H^- , which depends on the presence of metallic electron donors. A change in metallicity changes the overall atmospheric structure, which is why metallicity is one of the main stellar atmospheric parameters. Unlike T_{eff} and $\log g$, metallicity can only be measured directly from the analysis of a spectrum. Indirect determinations based on theoretically or empirically calibrated photometry have also been widely used when no spectrum is available. We note that they are affected by the same issues as photometric temperatures (see **Supplemental Text 1**). Metallicity is commonly referred to as $[\text{Fe}/\text{H}]$, because one of the main techniques to estimate this parameter is to determine iron abundances. However, in general, the abundances of other elements may not scale with Fe, which makes the designation of metallicity by $[\text{Fe}/\text{H}]$ imprecise. The stellar metallicity can also be expressed as $[\text{M}/\text{H}]$, usually representing a combination of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$. Whether metallicity refers to $[\text{Fe}/\text{H}]$ or $[\text{M}/\text{H}]$ depends mostly on the method employed to determine this parameter and what assumption is used for the enhancement of α elements of a given star. Like stellar abundances in general, to determine metallicities one must take care of all the steps and issues discussed in this section. There are two main ways to derive metallicities, either by measuring iron abundances from iron lines or by performing a global fitting to the spectra. These methods are discussed with more detail in **Supplemental Text 1**.

2.6.2. Other parameters. To relate $[\text{Fe}/\text{H}]$ and $[\text{M}/\text{H}]$ it is assumed that, at solar metallicities, α elements are solar-scaled and the α -element abundance linearly increases toward lower metallicities, reaching a plateau of $[\alpha/\text{Fe}] = +0.4$ at $[\text{Fe}/\text{H}] = -1$. However, at lower metallicities, variations in C and N might further affect the opacities, and a proper atmosphere model should be adopted to avoid additional uncertainties in abundances (Ezzeddine et al. 2017).

There are line broadening parameters that affect the overall structure of the atmospheres. In 1D modeling, the most notable one is the microturbulence (v_{mic}). It accounts for the small-scale turbulent motions of the particles that lead to excess line broadening. The stronger the line, the larger the effect due to v_{mic} (see **Figure 5** and discussion in Section 3.2.2). In 1D spectral synthesis calculations v_{mic} does not have a physical meaning, but is an ad hoc parameter needed to improve the line shape. Hence, the value of v_{mic} can be slightly different for different methods even when T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ agree. Microturbulence is normally derived by requiring that iron abundances remain the same regardless of the strength of the line. When not enough lines are available, it is possible to use empirical relations that depend on the other stellar parameters. In fact, this is done for most of the surveys (see Section 4.2). In a 1D analysis, v_{mic} counts as a fourth stellar parameter. The value adopted for v_{mic} influences the abundances, so it is important to state which value was considered when abundances are reported.

Further broadening parameters that need to be specified when synthesizing spectral lines are the projected rotational velocity ($v \sin i$) and the macroturbulence (v_{mac}). Similarly to v_{mic} , v_{mac} tries to account for large-scale turbulent motions in the atmospheres, which in 3D modeling are fully incorporated. Carney et al. (2008) present a study of these effects in metal-poor giants. Because v_{mac} and $v \sin i$ have a very similar broadening effect, it is difficult to disentangle both effects directly from the spectra. Carney et al. (2008) performed a Fourier transformation on

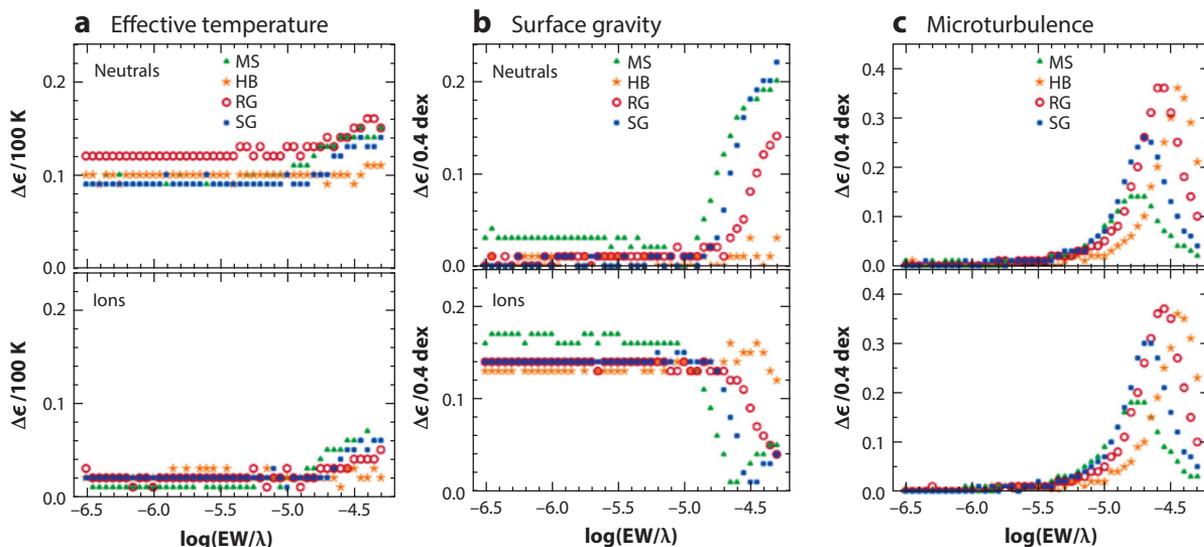


Figure 5

Sensitivity of abundances obtained for neutral and ionized lines on stellar parameters, for metal-poor stars, as a function of reduced equivalent width. Different symbols represent different spectral types. The figure shows (a) the abundance difference when the model temperature is changed by 100 K, (b) the effect when changing $\log g$ by 0.4 dex, and (c) the effect when changing v_{mic} by 0.4 km s^{-1} . Adapted from Roederer et al. (2014) with permission of first author and © AAS. Abbreviations: HB, horizontal branch; MS, main sequence; RG, red giant; SG, subgiant.

high-resolution and high S/N spectra to determine both parameters. However, such analyses are rarely done; rather, it is common to set either v_{mac} or $v \sin i$ to zero and determine a global broadening parameter or to use a value of v_{mac} based on empirical relations like those for v_{mic} . This is especially the case when spectral resolution or S/N are not sufficient to disentangle the effects from the two broadening mechanisms.

3. ASSESSING THE ABUNDANCE ERROR BUDGET

In recent years, large data sets of seemingly homogeneous stellar abundances have appeared on the scene, notably from spectroscopic surveys, moving the production of abundances toward industrial scales. For each data set, the combined effects of the steps discussed in Section 2 on the measurements of abundances of a given star could be interpreted as the ultimate uncertainty. Extensive discussions of such uncertainties can be found in the literature compilations of Suda et al. (2008) and Hinkel et al. (2014). Thus, it becomes increasingly challenging to obtain homogeneous abundances that can be used for a large variety of science cases. Furthermore, combining the abundances from different surveys is nontrivial, often due to correlated uncertainties arising from each step of the abundance analysis procedure. Some of the uncertainties may be amplified when combining results from different groups that employ different data and methods (see, e.g., Smiljanic et al. 2014). As an additional complication, uncertainties are assessed in different ways by different works. It is thus desirable that different catalogs perform similar tests to assess uncertainties, enabling better comparison and combinations. Roederer et al. (2014) provides inspiring work, discussing several sources of uncertainties extensively. The series of works on the *Gaia* benchmark stars (e.g., Jofré et al. 2015) also provide detailed discussions of the matter. Here, we disentangle and briefly discuss different parts of the abundance error budget, dividing the uncertainties into three main categories: random uncertainties, systematic uncertainties, and biases.

3.1. Random Uncertainties

Here, we refer to random uncertainties as uncertainties related to the input material (characteristics of input spectra, uncertainties in laboratory data, data reduction issues, and so on). In order to quantify these, there are some tests that can be performed.

3.1.1. Instrumental error. Using different spectra for the same stars allows one to quantify uncertainties due to the characteristics of the input spectra (S/N, resolution, normalization, instrumental responses in general). The abundance analysis method may be tested using a set of reference stars for which spectra exist in several archives. For example, Roederer et al. (2014) compared EWs from different instruments, and they found that the largest deviations arose for strong lines and low S/N, for which blends could not be identified. However, they demonstrated that, for the typical S/N of their sample, weak lines gave consistent results for different instruments. Another possibility is to use repeated observations of the same star at different S/N. Adibekyan et al. (2016) discuss how abundances are affected when spectra from the same instrument but of different S/N are used. They found an increased significance of abundance trends of $[X/\text{Fe}]$ versus condensation temperature for higher S/N spectra. This implies that, before interpreting such slopes astrophysically (e.g., presence of debris disks or planets), one must carefully assess the instrumental dependencies of the abundances obtained. Such statistical uncertainties are particularly important for high-precision studies. For planning spectroscopic surveys, a key issue is to find the threshold in S/N required for achieving the desired abundance precision for a given set of spectra and the methodology to be used. This uncertainty dictates the size of the data set and the Galactic region sampled.

Table 2 Uncertainty tests performed by different catalogs and surveys^a

	Instrument	Lines	Theory	Params	Methods	Trends	External	Precision
Catalogs								
GBS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Luck	Yes	Yes	Yes	Yes	No	Yes	Yes	No
Bensby	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
AMBRE	Yes	Yes	No	Yes	No	Yes	Yes	No
APOKASC	Yes	Yes	No	Yes	Yes	Yes	No	Yes
HARPS GTO	No	Yes	Yes	Yes	No	Yes	Yes	No
SPOCS	Yes	No	No	Yes	No	Yes	Yes	No
Surveys								
RAVE	Yes	No	No	Yes	No	Yes	Yes	Yes
GES	Yes	Yes	No	Yes	Yes	Yes	Yes	No
APOGEE	Yes	No	No	Yes	No	Yes	Yes	Yes
GALAH	Yes	No	Yes	No	No	Yes	Yes	Yes

^aCatalogs and surveys are sorted as they appear in Sections 4.1 and 4.2, except for GBS, which is discussed in Section 3.3.2. Columns indicate the uncertainty tests described in Section 3: Instrument, uncertainty due to different instrumental responses evaluated; Lines, line-by-line abundance dispersion discussed; Theory, 1D-LTE effects assessed; Params, propagation of stellar parameter uncertainties in final abundances; Methods, different methodologies compared; Trends, consistency of abundances as a function of stellar parameters assessed; External, comparison of results with external sources; Precision, improvement of precision with differential or data-driven methods. Information about uncertainties of abundances for each catalog can be found in the following references: GBS (*Gaia* benchmark stars), Jofré et al. (2015); Luck, Luck (2018); Bensby, Bensby et al. (2014); AMBRE (Archéologie avec Matisse Basée sur les Archives de l'ESO), Mikolaitis et al. (2017); APOKASC (APOGEE–Kepler Asteroseismology Collaboration), Hawkins et al. (2016b); HARPS GTO (High Accuracy Radial Velocity Planet Searcher–Guaranteed Time Observations), Adibekyan et al. (2012); SPOCS (Spectroscopic Properties of Cool Stars), Valenti & Fischer (2005) and Brewer & Fischer (2018); RAVE (Radial Velocity Experiment), Boeche et al. (2011) and Casey et al. (2017); GES (*Gaia*-ESO Survey), Smiljanic et al. (2014); APOGEE (Apache Point Observatory Galaxy Evolution Experiment), Holtzman et al. (2018) and Jönsson et al. (2018); GALAH (Galactic Archaeology with HERMES), Buder et al. (2018a).

An alternative way to quantify uncertainties due to input spectra is to look at the differences obtained in abundances for cluster members, which are expected to have the same abundance pattern (although see, e.g., Liu et al. 2016b). Different stars of the same spectral class essentially should yield the same abundances. Thus, the variation can be attributed to statistical uncertainties. These tests are performed by some surveys (see **Table 2** and Section 4.2). Errors are commonly given as the standard deviation about the mean of the abundances obtained from all measurements.

3.1.2. Uncertainties due to line selection. In general, one can assume that the results will be more accurate the more lines are used for a given element. However, including too many lines might have negative consequences on the results if a considerable number of lines are saturated, too weak, or blended; have poor atomic data, poor HFS treatment, or poor spectra; or are contaminated, etc. A classical way to quantify this uncertainty is providing a line-to-line dispersion (LLD). Uncertainties derived from neutral lines are often observed to be smaller than those from ionized lines, but this may be due to the fact that, in FGK-type stars, more neutral than ionized lines are available for estimating this dispersion. Although this uncertainty is commonly reported, the definition for LLD differs from work to work. It is common, for example, to decrease the uncertainty by adopting a σ -clipping procedure, that is, removing outlier lines whose abundances differ by more than a given value from the mean abundance obtained from all lines. That value can be a factor of σ , with σ representing the standard deviation of the Gaussian distribution of all abundances. The factor varies in the literature; for example, Luck (2018) performs a cut at 2.5σ , whereas Pancino et al. (2011) use 3σ . In some cases (Adibekyan et al. 2012, Mucciarelli et al. 2013, Mikolaitis et al. 2017), the random uncertainty is reported to be the standard error of the mean

(σ/\sqrt{N} , i.e., dividing the LLD by the number of lines employed). This definition of error is obviously much smaller than the LLD, making the two uncertainty estimators incomparable. In many cases, few lines are available per element, and then σ is greatly affected by a single outlier. The median is a more robust estimator of the final abundance, with the interquartile ($q_{75}-q_{25}$) range as its uncertainty (for a good discussion see chapter 3 of Ivezić et al. 2014). Beers et al. (1990) also provide a number of robust and resistant estimators of location and scale that should prove useful.

3.2. Systematic Uncertainties

We refer to systematic uncertainties as those uncertainties that arise from the approach employed to determine abundances, namely the method and line prescription assumed, which might induce different uncertainties in different parts of the parameter space. In this section, we disentangle main sources of such uncertainties.

3.2.1. Theory: 1D-local thermodynamic equilibrium effects. A certain level of uncertainty in the final abundances is caused by approximations in the line-profile prescription. Transitions are affected to various degrees by the assumption of 1D-LTE, which can be quantified and even corrected. The magnitude of these corrections varies across stellar-parameter space. For many elements, model atoms required for non-LTE calculations are available (see **Table 1**). With the corrections at hand, the difference in the final abundance when using LTE and non-LTE results can be evaluated. Quantifying 3D effects is still difficult, because large grids with corrections for lines and elements are not available. However, the following diagnostics can be performed to assess the level of accuracy of the employed line-modeling prescription.

If abundances can be derived from both neutral and ionized lines for the same element, the difference in these results may be attributed to uncertainties in the line-formation calculations. However, this method is not applicable if the stellar parameters have been determined by forcing ionization and excitation balance, as this causes an artificial agreement between abundances derived from neutral and ionized lines. To quantify this uncertainty, it would be ideal to determine T_{eff} and $\log g$ from methods that are less sensitive to 1D-LTE prescriptions (see examples in Section 2.6). Some examples of detailed investigations of this kind for elemental abundances have been published by Sneden et al. (2016) for iron-peak elements, Bergemann et al. (2017) for magnesium, Amarsi et al. (2018a, and references therein) for oxygen, and Amarsi & Asplund (2017) for silicon. These works show that although 1D-LTE modeling can be very uncertain, leading to incorrect measurements of abundances, with a careful selection of lines it is possible to derive accurate abundances. Careful selection of lines would favor ionized lines, for which LTE holds better, and high excitation potential lines, for which 1D modeling is more accurate. This of course depends on the metallicity and overall atmosphere structure. In general, metal-poor stars are most affected (e.g., Ezzeddine et al. 2017, Nordlander et al. 2017).

A selection of accurate 1D-LTE lines might require removing a large variety, if not all, lines in optical spectra [for the case of metal-poor dwarfs, see Sneden et al. (2016) and Roederer et al. (2018)]. It is thus crucial to have very extended wavelength coverage, including the IR to the UV regions, and high resolution, in order to include as many clean lines as possible. Unfortunately, outside the optical window, 3D-non-LTE effects have been investigated for very few elements and lines. One example is Bergemann et al. (2017, and references therein), who investigated the effect of 3D-non-LTE line formation of optical and IR Mg, Si, and Ti lines. Other examples include Zhang et al. (2017), who looked at Mg lines in the H band to quantify this uncertainty in APOGEE stars, and Nordlander & Lind (2017), who quantified the uncertainties due to 3D-non-LTE of Al for a variety of stars and lines in the optical and IR. Regarding UV spectra, we must rely

on observations obtained with the *Hubble Space Telescope* that are competitive, and thus limited data are available. In any case, most stars targeted by surveys have high metallicities and are rather cool. Thus, their UV spectra are so crowded with absorption features that almost no unblended lines can be used (Snedden et al. 2016). To decrease ionization-imbalance uncertainties due to poor modeling, we recommend following the advice of Roederer et al. (2014): Use the same ionization stages for abundance ratios. If Fe I results are to be considered for $[X/\text{Fe}]$, then using the results for other elements from neutral lines will yield more accurate abundance ratios. The same is true for ions.

3.2.2. Uncertainties due to stellar parameters. The final abundances depend to a large degree on the scale used for the stellar parameters. In **Figure 5**, we show the effect on abundances when varying stellar parameters for neutral and ionized lines of several elements. This study was conducted by Roederer et al. (2014) on metal-poor stars of representative spectral types, namely main sequence, horizontal branch, red giant, and subgiant stars. The figure compares the variation of abundances when changing the atmospheric parameters as a function of line strength. It nicely illustrates that abundances obtained from strong lines are more affected by uncertainties in stellar parameters than those from weak lines. It is also seen that, by changing T_{eff} by 100 K (**Figure 5a**), abundances obtained from neutral lines are affected by 0.1 dex or more, whereas ionized lines change very little except for the strongest lines. The opposite is seen for $\log g$ (**Figure 5b**). When changing the surface gravity in the model, weak ionized lines are more affected than neutral lines, whereas the situation is reversed at the strong-line end. This opposite behavior forms the basis of determining stellar parameters from the combination of ionization and excitation balance. Finally, **Figure 5c** shows that the abundances of strong lines are strongly affected by the adopted value of v_{mic} .

In automatic analyses, it is relatively straightforward to compute abundances using different stellar parameters as input. The error due to stellar parameters can thus be estimated by comparing the difference in abundances obtained when the input parameters are varied according to their uncertainties. In this case, independent errors can be estimated for each parameter, which can be combined as explained in Section 3.5. Alternatively, star cluster members with stellar parameter differences of the order of the errors can be used to estimate the differences in abundances due to uncertainties in stellar parameters, because the abundances should be the same for all cluster stars (although see, e.g., Liu et al. 2016b). In this case, one obtains a single uncertainty accounting for all parameters together.

3.2.3. Using different methods. A comparison of the results obtained from different methods allows one to study the dependency of abundances on the code employed. For example, Casamiquela et al. (2017) perform a systematic study comparing stellar parameters using the EW wrapper GALA with the synthesis wrapper iSpec (see **Table 1**). They use this procedure to show that their conclusions are not affected by the methodology employed in the analysis. At a more industrial scale, most of the error budget in the GES is assessed from the method-to-method dispersion (MMD; Smiljanic et al. 2014), and the same holds for the abundance analysis of the *Gaia* benchmark stars (Jofré et al. 2015). In fact, the strikingly large MMD seen in the GES has motivated the next generation of spectroscopic surveys to rely on one pipeline only (Allende Prieto 2016a). By excluding this type of uncertainty from the total error budget, the results will become more precise, but it will not be possible to investigate the dependency on the methodology employed and to truly assess the accuracy of the results. If many methods are used, a dispersion of the results can be calculated. Like the discussion in Section 3.1.2, we advocate employment of the median and the interquartile range (or the estimates described by Beers et al. 1990) to quantify the dispersion rather than the mean and standard deviation. If only two methods are used, the error can simply be the difference between the two results.

3.3. Reference Objects and Biases

It is important to investigate any overall biases in the results, as this is key to combining different data sets. For this purpose, the results in different parts of the parameter space are compared with external sources.

3.3.1. Trends of abundances with stellar parameters. Star clusters are good laboratories for assessing whether there are systematic uncertainties of the method for stars with different stellar parameters (e.g., dwarfs versus giants). Trends in abundances found as a function of T_{eff} or $\log g$ can be attributed to a systematic uncertainty of the method. In fact, for any stellar sample the behavior of the abundances as a function of stellar parameters should be investigated. If the method is robust for a large range of spectral types and if no effects of stellar evolution are to be expected, then no correlations should be found. If a clear correlation is found, it can be interpreted as a systematic uncertainty. Several works have chosen to apply a correction for such systematics, at least for T_{eff} , by finding an empirical relation that is then used to scale the abundances according to their T_{eff} (Valenti & Fischer 2005, their section 6.4 and figure 10; Adibekyan et al. 2012, their section 3.2 and figure 4).

It is not simple to explain or correct the trends, as they can be caused by a variety of reasons, as discussed in detail by Roederer et al. (2014). In short, if the temperature decreases or the metallicity increases, lines become more affected by blends, which often are not identified. Lines also become stronger and start to saturate, which means that the selection of lines may vary across the parameter space. Thus, systematic differences may simply be the result of a line selection effect instead of being due to variations in stellar parameters. As strongly recommended by Roederer et al. (2014) and clearly demonstrated by, e.g., Nissen (2015), if spurious $[X/\text{Fe}]$ trends exist as a function of stellar parameters, selecting stars from within a small region in parameter space for chemical evolution studies is the most secure way to proceed.

3.3.2. Toward an absolute scale for abundances using reference stars. A comparison of results with external sources helps quantify the overall error budget and understand for what kind of stars the method is most accurate. The catalogs presented in Section 4.1 are widely used for comparison as they are large, enhancing the chance to have a sufficient overlap between data sets and to study differences in a statistically significant way. A standard for reference objects provides a more straightforward link between catalogs. Such reference objects can be either individual stars with well-defined properties or fields of stars with high-quality data available for a large number of stars, such as clusters or asteroseismic fields.

3.3.2.1. The Sun. In terms of stars with well-defined properties the Sun is undoubtedly *the* reference star, considered as the standard reference for cosmic abundances. However, the determination of solar abundances is problematic, because different methods lead to different results. A review on the chemical composition of the Sun is given by Asplund et al. (2009), who also provide recommended solar abundances determined with 3D hydrodynamical atmospheres. These revised abundances are, for some elements—in particular, light elements—significantly lower than those obtained with conventional methods [e.g., the widely used scale of Grevesse & Sauval (1998)]. Although the Sun is not observable in the same way as other stars, it is the most used star for differential studies, often using the reflection of sunlight from a Solar System body. Abundances tell us whether a given star with solar atmospheric parameters has exactly the same chemical composition as the Sun. The literature is very rich in studies looking for the closest solar twin. Several dozens of stars were claimed to be solar twins based on their atmospheric parameters, but when

their detailed chemical compositions are considered, the similarity is less obvious. For instance, Yana Galarza et al. (2016) performed a high-precision analysis that confirmed HIP 100963 to be a good solar twin but with abundances of the s - and r -process elements, as well as Li, slightly enhanced relative to the Sun. Other solar twins studied at high precision, and with a chemical pattern very similar to that of the Sun, include Kepler-11 (Bedell et al. 2017), HIP 76114 (Mahdi et al. 2016), M67-1194 (Liu et al. 2016a), and HIP 114328 (Meléndez et al. 2014), which are all good options to use as a reference star instead of the Sun. Abundances of solar-like stars relative to the Sun can be different due to several factors, such as Galactic chemical evolution, age, or the relative effects of non-LTE on stars with similar, but not exactly the same, stellar parameters.

3.3.2.2. *Gaia* benchmark stars: beyond the Sun. For stars that differ significantly from the Sun, it is not possible to measure abundances differentially to it. For that reason, the sample of *Gaia* benchmark stars was built in order to establish a system of reference stars covering a larger range of atmospheric parameters (Heiter et al. 2015). The sample was designed to provide an anchor to the *Gaia* astrophysical parameter inference system that will estimate atmospheric parameters of one billion stars (Apsis; Bailer-Jones et al. 2013). These stars are fundamental calibrators because their effective temperature and surface gravity can be deduced directly from the accurate knowledge of their radius and flux distribution (see Section 2.6). Determination of their metallicity and abundances is described by Jofré et al. (2014) using a library of high-quality spectra (Blanco-Cuaresma et al. 2014b). Some surveys already use the *Gaia* benchmark stars for their calibration, but the sample is still too small (around 30 stars), the stars are too bright, and they suffer from a deficiency of metal-poor stars. Substantial efforts have been dedicated to extending the sample toward fainter and more metal-poor stars (Hawkins et al. 2016a). Updated information and stellar parameters are provided via the CDS (Centre de Données astronomiques de Strasbourg; Jofré et al. 2018).

3.3.2.3. *Literature sources.* Several surveys use other stars with well-defined properties that can be found in large catalogs such as PASTEL (Soubiran et al. 2016) or Hypatia (Hinkel et al. 2014). Both are bibliographical catalogs, making it possible to find well-studied stars that have been analyzed independently by different groups who found consistent results. This approach may be a way forward toward establishing a common set of reference stars. However, it is important to agree on a common set of procedures and criteria when selecting stars from such catalogs.

3.3.2.4. *Clusters.* Open and globular clusters are convenient reference objects owing to their large number of members sharing in principle the same age and chemical composition. Some clusters, such as M67, have been extensively studied with high-resolution spectroscopy that is available in public archives. Measuring the dispersion of abundances of cluster members obtained by an automatic pipeline is a good way to evaluate the internal precision over a range of stellar parameters. However, it is worth noting that the Hyades, another famous reference cluster, was found to be inhomogeneous in chemical composition at the 0.02-dex level (Liu et al. 2016b). Membership determinations in open and globular clusters have dramatically improved with *Gaia* DR2 (Cantat-Gaudin et al. 2018, Gaia Collab. et al. 2018a), making them promising validation targets in the future.

3.3.2.5. *Asteroseismic fields.* Asteroseismic fields observed by the space missions CoRoT (*Convection, Rotation & Planetary Transits*), *Kepler*, and *Kepler2* (K2) are of great interest because stellar surface gravities and ages can be determined with very high precision from seismic data (Chaplin & Miglio 2013, Stello et al. 2017; see also Section 2.6). This valuable information has led several

surveys to observe these fields that offer a very good opportunity for calibration. Examples are the APOKASC (APOGEE–*Kepler* Asteroseismology Collaboration) sample (Pinsonneault et al. 2014) observing *Kepler* targets with APOGEE, the K2 stars in RAVE (Radial Velocity Experiment; Valentini et al. 2017), and CoRoT targets in GES (Pancino et al. 2017). Spectroscopic data sets also have the potential for intercomparisons (Jofré et al. 2017b), provided that surveys agree on stars to observe in common. Some asteroseismic fields also include a few open clusters, which make them even more interesting for reference purposes (Stello et al. 2016, the case of M67). The use of asteroseismic fields for calibration, training, or validation of automatic pipelines requires the stellar abundances to be determined in those fields with a high level of accuracy and precision. This effort has already started (see, e.g., Hawkins et al. 2016b, Nissen et al. 2017).

3.4. Improving Precision

A homogeneous analysis with significantly reduced uncertainties might be achieved using a single pipeline, either by performing a differential analysis or by applying a data-driven approach. In this section, we discuss these procedures.

3.4.1. Differential analyses. Differential analyses consist of determining abundances in the same fashion for a given star and a reference star. The highest possible precision is achieved if the reference star is similar to the star of interest, because the overlap of suitable lines will be maximized. This reduces the LLD significantly, because uncertainties due to blends, poor atomic data, non-LTE effects, etc., are cancelled out to a certain degree. Furthermore, the continuum-normalized spectra are expected to be similar for similar kinds of stars, thereby reducing systematic uncertainties due to the methodology or due to stellar parameters. Nissen & Gustafsson (2018) provide a complementary review focused on high-precision spectroscopic studies based on the differential technique.

The accuracy of differential abundances fully relies on the abundance accuracy of the reference star. Differential analyses are thus very popular for solar twins (e.g., Tucci Maia et al. 2014, Nissen 2015, Bedell et al. 2018), because the Sun is our most accurate reference star (see Section 3.3.2). Precisions achieved are so high (better than 0.01 dex) that only with such an approach it is possible to study, e.g., the effect of planet formation [see Section 5 and Nissen & Gustafsson (2018) for science applications]. However, differential analyses of stars too different from the Sun require another reference star, because the more different the stars are, the less lines in common are available. Extensive discussions on this matter are provided by Jofré et al. (2015). In giants, Hawkins et al. (2016b) improved the precision of the abundances by performing a differential analysis with respect to Arcturus. In metal-poor stars, Reggiani et al. (2016) performed a high-precision abundance study using as a reference G64-12. In clusters, precision can be improved by using one cluster member as a reference and deriving abundances for the other stars at the same location in the color-magnitude diagram differentially [e.g., Liu et al. (2016b) for the Hyades cluster].

3.4.2. Data-driven approaches. Recently, new revolutionary ways to derive abundances with machine-learning tools have become very popular for the analysis of large data sets of spectra (Leung & Bovy 2018, Ness 2018, Ting et al. 2018). Empirical models or neural networks are built, and a relation between the spectrum and certain labels (abundances) is trained on a previously analyzed subset of spectra. These relations are then applied to large samples of stars, resulting in impressively precise abundances even from data of seemingly rather low quality. Machine-learning methods have been very efficient in transferring the known information from the so-called training sets to entire data sets. However, it is not fully explored to what extent such methods can identify

outliers. As in the case of differential studies, the accuracy of the labels obtained with data-driven methods fully relies on the training (reference) sample.

3.5. Combination of Uncertainties

Table 2 lists the different surveys and catalogs described in Section 4, where we summarize which of the uncertainty assessments discussed here are performed. The abundance tests are separated according to assessing random and systematic uncertainties and biases. We can see that all catalogs carry out at least one test in each of the categories, although they are not always the same. In the listed works, the tests performed might not necessarily be included in the final error budget. This makes the comparison between catalogs, including uncertainties, difficult. Here, we provide guidance as to how the different uncertainties can be combined and standardized to provide for a more straightforward comparison in future catalogs.

Although an assessment of accuracy is provided by the external uncertainty (e.g., overall agreement with reference stars), a conservative measurement of precision for abundance determinations should account for both the random and systematic uncertainties. According to our list, this means combining five different sources of uncertainties. Following standard formulas for error propagation, the total error budget can be obtained considering the variances and covariances of the uncertainties. Let σ_I , σ_L , σ_T , σ_P , and σ_M be the uncertainties of instrument, lines, theory, parameters, and methods, respectively (see **Table 2**). These may be independent from each other, which means that the total error budget can be obtained from adding their variances (σ_i^2), where i represents each of the five above sources.

Determining σ_p^2 can be more complicated, because it might originate from the analysis of the response of abundances to changes of the different stellar parameters separately (see, e.g., Jofré et al. 2015). The appendix of McWilliam et al. (1995) provides a well-structured presentation of a procedure based on a standard formalism for propagation of errors, showing how to obtain final uncertainties based on line-by-line abundance measurements, and how uncertainties in stellar parameters (which are not independent from each other) affect the final results. We discuss a few important conclusions from that work. First, because the uncertainties are correlated, the covariances between uncertainties can be calculated for a few representative stars in the sample and applied to the entire data set. McWilliam et al. (1995) provide the atmospheric parameter variances and covariances for a metal-poor star based on an analysis of optical lines. It would be useful to have such covariances for other types of stars to aid in the homogeneous presentation of abundances and their uncertainties by catalogs. Second, it is shown that increasing the number of lines might reduce the random component of the uncertainty, whereas the systematic component remains constant. This implies that one should not consider σ/\sqrt{N} , where N is the number of lines used, as an estimate of the total uncertainty of an average abundance. Third, for estimating abundance-ratio uncertainties one needs to keep in mind that the adopted atmospheric parameters (and associated uncertainties) are the same for both elements involved in the ratio and that for some element lines the response of the final abundance to the stellar parameter uncertainty will also be very similar, leading to a partial cancellation of the systematic uncertainty. It is thus necessary to compute the covariances between the element abundances if the abundance ratio uncertainties are to be estimated realistically. Barklem et al. (2005, their appendix B) describe a modified version of the formalism of McWilliam et al. (1995) applicable to methods performing a global spectrum fit rather than determining line-by-line abundances.

4. THE PERIODIC TABLE AS SEEN FROM SPECTRAL ANALYSES

We start with an overview of relatively large ($\sim 1,000$ stars) catalogs of stellar abundances, followed by how they have served to build the industrial products from spectroscopic surveys available today

and in the future. That information will help us understand why certain elements are more popular than others, as well as discuss how different surveys compare for common stars.

4.1. Catalogs of Stellar Abundances from High-Resolution Studies

The catalogs of stellar abundances published by individual groups can be divided into two main categories. The first one is compilations of literature sources, and the second one is spectral analyses done independently. Here, we revise these two categories.

4.1.1. Bibliographic compilations. Soubiran & Girard (2005) made an early attempt to combine abundances from different studies in order to build a large catalog for the investigation of abundances and kinematic trends in the Galactic disk. This work resulted in 743 stars with abundances of Fe, O, Mg, Ca, Ti, Si, Na, Ni, and Al in the metallicity range $-1.3 < [\text{Fe}/\text{H}] < +0.5$, with a typical precision of 0.6 dex. SAGA (Stellar Abundances for Galactic Archaeology database; Suda et al. 2008) is another compilation of stellar parameters and abundances for ~ 30 elements from the literature, with the initial motivation being to characterize extremely metal-poor stars, in order to constrain the nature of the first stars. The catalog is now being extended to a greater range of metallicities. It includes more than 1,000 stars of the Milky Way and other nearby galaxies. Efforts are being made to follow-up on LAMOST targets in order to homogenize and complete the SAGA database.

The Hypatia catalog (Hinkel et al. 2014, 2017) is a recent compilation that at the time of writing has collected 278,968 abundance measurements in 171 catalogs for 6,156 FGK stars within 150 pc from the Sun. The main purpose is to evaluate the spread in chemical abundances for nearby stars analyzed by different groups, which allows one to estimate uncertainties when studying the chemical composition of exoplanet hosts, the connection between thick- and thin-disk stars, or stars with different kinematic properties. The content in terms of stellar abundances is shown in the histogram of **Figure 6**. Light and α elements (C, O, Na, Mg, Al, Si, Ca, and Ti), as well as iron-peak elements (Sc, V, Cr, Mn, Fe, Co, and Ni), are very common in the literature. Neutron-capture (Sr, Y, Zr, Ba, La, and Eu) elements are less common, but still quite popular, as they help to answer important scientific questions regarding stellar and chemical evolution. Other elements have very few abundance measurements in FGK-type stars, for reasons that are discussed in **Supplemental Text 2**.

Supplemental Material >

4.1.2. Independent catalogs. Luck (2018, and references therein) has undertaken a large high-resolution spectroscopic abundance study. His data set includes abundances of $\sim 3,000$ dwarfs, subgiants, and giants within ~ 100 pc from the Sun using good quality spectra selected in public archives of echelle spectrographs. Abundances of C, N, O, Li, Na, Mg, Al, Si, S, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Sr, Y, Zr, Ba, La, Ce, Nd, Sm, and Eu were determined with a high level of precision. A smaller (~ 700 stars), yet very widely used, catalog was published by Bensby et al. (2014). They performed the largest ever by-hand EW analysis to provide abundances of O, Na, Mg, Al, Si, Ca, Ti, Cr, Fe, Ni, Zn, Y, and Ba for nearby dwarf stars. Battistini & Bensby (2015, 2016) complemented the catalog with Sc, V, Mn, and Co as well as s - and r -process abundances for a subset of the sample. The study has become a reference for how the trends of $[\text{X}/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ are expected to look for thin- and thick-disk stars in the Solar Neighborhood.

The AMBRE (Archéologie avec Matisse Basée sur les aRchives de l'ESO) project consists of the automatic parameterization of large sets of ESO high-resolution archived spectra from FEROS (Fiber-fed Extended Range Optical Spectrograph; Worley et al. 2012), HARPS (De Pascale et al. 2014), and UVES (Worley et al. 2016). Guiglion et al. (2016) determined abundances of Li for for

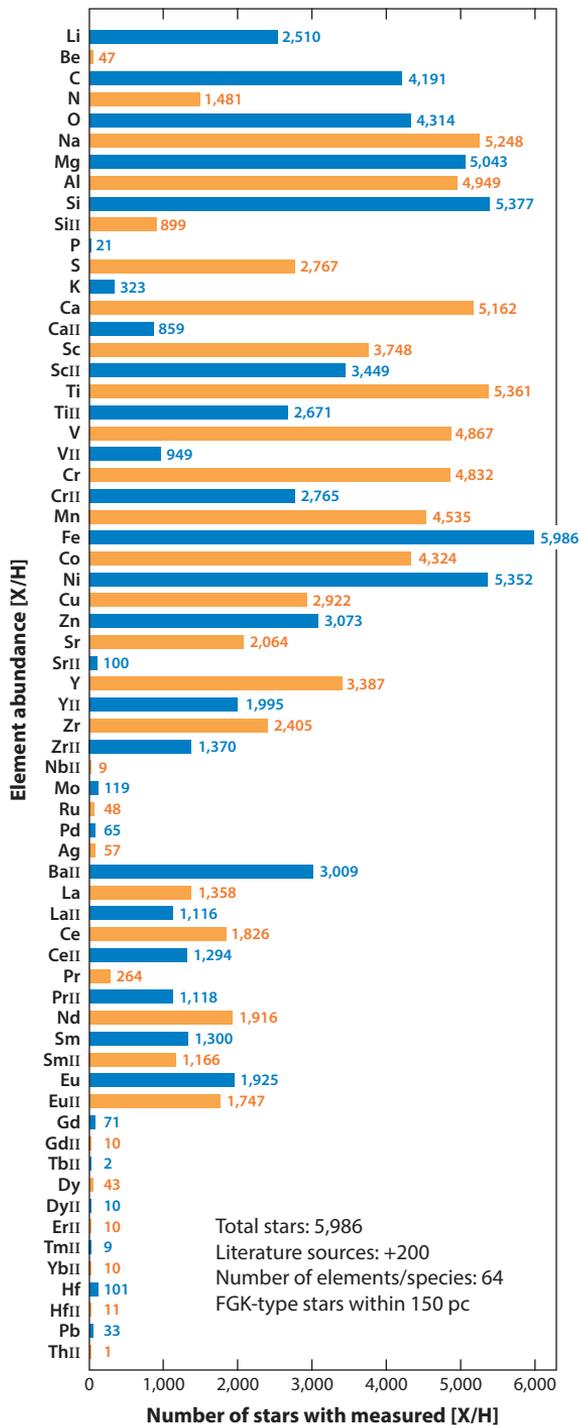


Figure 6

Overview of abundances of FGK-type stars in the solar vicinity as included in the Hypatia catalog. Adapted from Hinkel et al. (2017) with permission of first author and © AAS.

7,300 AMBRE stars, and Mikolaitis et al. (2017) derived Mn, Fe, Ni, Cu, Zn, and Mg abundances for 4,666 stars.

Hawkins et al. (2016b) published abundances of C, N, O, Mg, Ca, Si, Ti, S, Al, Na, Ni, Mn, Fe, K, V, P, Cu, Rb, Yb, Co, and Cr for a sample of $\sim 2,000$ *Kepler* giant stars that have IR spectra from APOKASC (Pinsonneault et al. 2014; see also Section 4.2.3). The stars, as targeted by *Kepler*, benefit from asteroseismic data that allow one to better constrain the surface gravity. These data are used to provide a catalog that is self-consistent, precise, and accurate.

Planet search programs with radial velocity follow-up are actively generating large spectroscopic catalogs with more than a thousand stars, with high-quality spectroscopy and homogeneous analyses. The characterization of exoplanets requires the properties of the host star to be well-known, and this is why several studies have provided the stellar parameters and abundances of the targets of these observing programs. It is worth noting that these programs are dominated by dwarfs. Adibekyan et al. (2012) provided chemical abundances of Na, Mg, Al, Si, Ca, Ti, Cr, Ni, Co, Sc, Mn, and V for 1,111 FGK-type stars of the HARPS GTO (Guaranteed Time Observations) planet search program (see also Sousa et al. 2011). Although its aim was to characterize planet host stars, this sample has provided insights into stellar populations, making it an additional reference for $[X/Fe]$ trends as a function of $[Fe/H]$ in Galactic studies. This work has been further completed with the determination of more chemical elements; in particular neutron-capture elements have been determined by Delgado Mena et al. (2017) for the entire sample. A subset of stars also have measurements of light element abundances, namely Li, Be, C, N, O (see discussion in, e.g., Delgado Mena et al. 2017). The SPOCS (Spectroscopic Properties of Cool Stars) catalog (Valenti & Fischer 2005) contains abundances of Na, Si, Ti, Fe, and Ni for 1,040 nearby F, G, and K stars that have been observed by the Keck, Lick, and AAT planet search programs. The California-Kepler Survey (CKS; Petigura et al. 2017) is a follow-up program developed to characterize stars with transiting planets detected by *Kepler*. The catalog provides HIRES (High Resolution Echelle Spectrometer) spectra that were analyzed by Brewer et al. (2016) to generate a catalog of C, N, O, Na, Mg, Al, Si, Ca, Ti, V, Cr, Mn, Fe, Ni, and Y abundances. The catalog of Spectroscopic Parameters and atmospheric ChemIstriEs of Stars (SPECIES; Soto & Jenkins 2018) is built from public spectra and includes Na, Mg, Al, Si, Ca, Ti, Cr, Mn, Ni, Cu, and Zn abundances for about 1,000 planet host stars. Note that the catalogs above have a significant number of stars in common, allowing for comparisons.

4.2. Chemical Abundances of Spectroscopic Surveys

The catalogs presented above have shown that chemical abundances of FGK-type stars provide key information on their formation processes and sites. The next step is to construct large chemical maps of the Galaxy to constrain models of its formation and evolution. In addition, stars need to have a well-defined selection function in order to probe Galaxy models properly. To that aim, stars are surveyed with multiobject spectrographs over several years, and automatic pipelines to measure abundances are designed.

The first efforts in the field of massive spectroscopy were dedicated to the search of metal-poor stars. This started with large objective-prism surveys that produced hundreds of candidates followed-up at medium or high resolution. This pioneering work, essential to the later expansion into the era of industrial abundances, is reviewed by Beers et al. (1985). The first very ambitious project aiming at determining spectroscopic abundances at industrial scales is the Sloan Digital Sky Survey (SDSS) with its Sloan Extensions for Galactic Understanding and Exploration (SEGUE; Yanny et al. 2009), which provided moderate-resolution ($\sim 1,800$) spectra for well over 500,000 unique stars. The SEGUE Stellar Parameter Pipeline (SSPP; Lee et al. 2008a)

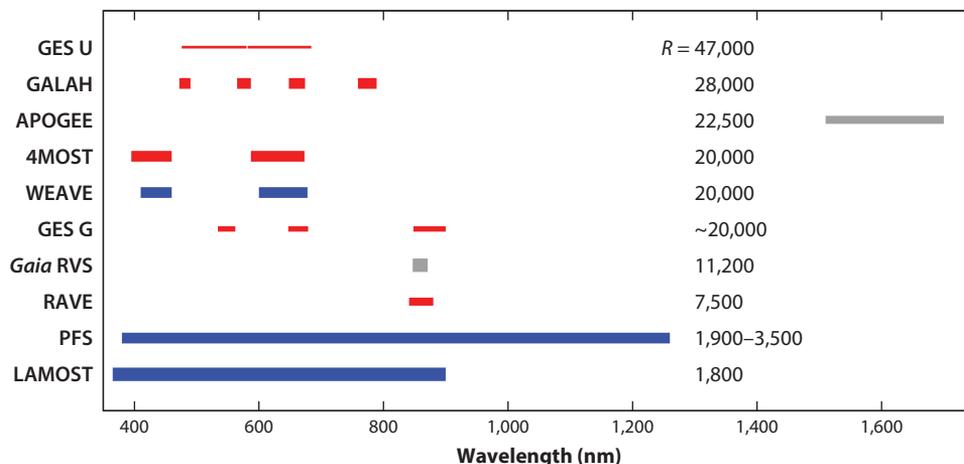


Figure 7

Overview of ongoing and future spectroscopic surveys (see Section 4.2), sorted by spectral resolution (see *labels to the right*). Horizontal lines show the covered wavelength intervals. Northern, southern, and all-sky surveys are represented by blue, red, and gray colors, respectively. Broader lines indicate larger numbers of target stars. Abbreviations: 4MOST, 4-metre Multi-Object Spectroscopic Telescope; APOGEE, Apache Point Observatory Galaxy Evolution Experiment; GALAH, Galactic Archaeology with HERMES; GES G, *Gaia*-ESO GIRAFFE; GES U, *Gaia*-ESO UVES; LAMOST, Large Sky Area Multi-Object Fiber Spectroscopic Telescope; PFS, Prime Focus Spectrograph; RAVE, Radial Velocity Experiment; RVS, Radial Velocity Spectrometer; UVES, Ultraviolet and Visual Echelle Spectrograph; WEAVE, the next generation wide-field spectroscopy facility for the William Herschel Telescope.

was developed specifically to obtain large-scale estimates of T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ and was later extended to $[\text{C}/\text{Fe}]$ and $[\alpha/\text{Fe}]$ determinations. The SSPP pioneered the use of multiple techniques to determine stellar parameters, as well as validation with open and globular clusters (Lee et al. 2008b) and estimation of errors by comparison to parameters from high-resolution studies (Allende Prieto et al. 2008). The impact of SDSS on the understanding of the Milky Way stellar populations is reviewed by Ivezić et al. (2012).

In this section, we describe the main ongoing and future spectroscopic surveys that deliver abundances of at least five individual elements. We discuss the targeted accuracy, methods, performances, and calibration strategies that they implement in their latest data releases following our listing in **Table 2** for uncertainties of the abundances. Their main characteristics are shown in **Figure 7**. The abundances determined by each of the ongoing surveys are marked for each survey with a different color in the periodic table of **Figure 8**. (See the sidebar titled Intercomparisons of Surveys.)

4.2.1. RAVE. The Radial Velocity Experiment (RAVE; Steinmetz et al. 2006) was the first large survey that provided abundances of several individual elements. The observations were performed with the 6-degree field multi-object spectrograph on the 1.2-m UK Schmidt Telescope of the Anglo-Australian Observatory. DR5 (Kunder et al. 2017) contains 457,589 stars in the magnitude range of $9 < I < 12$, observed between 2003 and 2013. The spectra will be made public with the last data release (DR6).

The chemical pipeline is described by Boeche et al. (2011) and obtains abundances of Mg, Al, Si, Ca, Ti, Fe, and Ni for about 300,000 stars based on EWs (cf. Section 2.4) from 604 absorption lines identified in spectra of the Sun and Arcturus. Stellar parameters are derived from

Elements potentially detected in spectroscopic surveys of the Milky Way

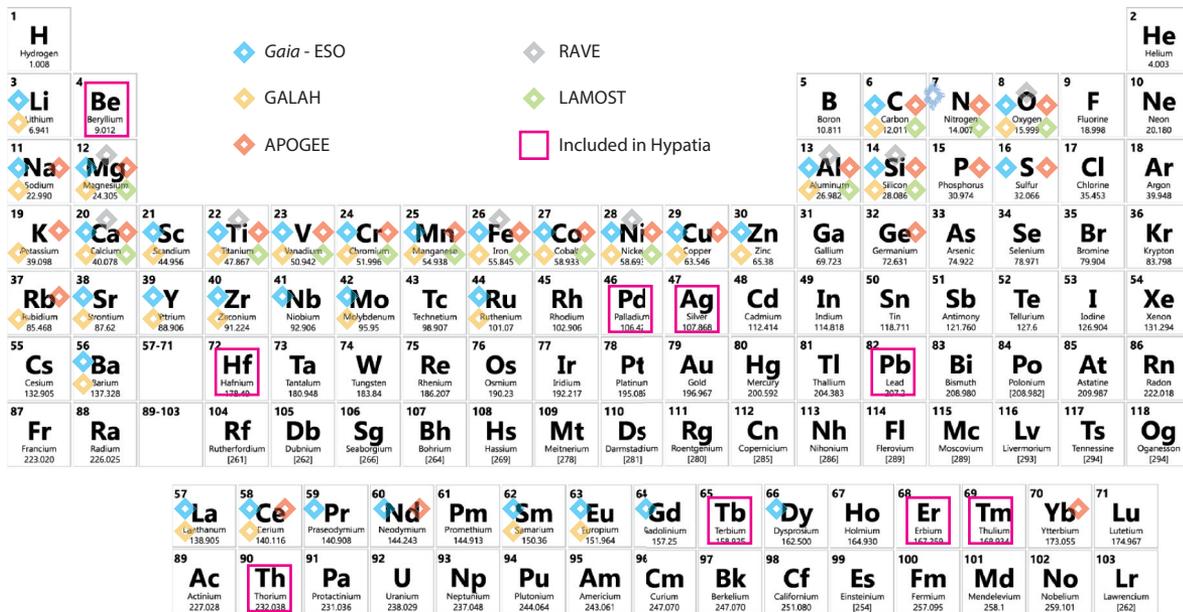


Figure 8

Periodic table annotated with the surveys for which lines of a given element can be detected. Elements in magenta squares are included in Hypatia, but there are no plans for studying them in spectroscopic surveys.

the spectra by fitting to a grid of synthetic spectra, which was built using MARCS models, the Turbospectrum radiative transfer code, and astrophysical gf -values. The best model is found from a combination of a decision-tree algorithm (DEGAS; Bijaoui et al. 2010) and a projection method of models to data (MATISSE; Recio-Blanco et al. 2006). The final parameters are calibrated by a modification of the parameters with second-order polynomials obtained from the comparisons against samples of reference stars. Metallicities are calibrated considering sets of reference stars that are a combination of literature sources of results obtained with high-resolution optical spectra

INTERCOMPARISONS OF SURVEYS

Due to their different characteristics, the spectroscopic surveys and catalogs have only several hundreds of stars in common, or a few thousand for the largest ones. Their intercomparison is mandatory to track systematic differences, understand their origin, and put all survey products onto the same scale.

One primary ambition in the field is to have stellar properties and abundances on the same scale, as this allows the community to straightforwardly merge the data sets from different surveys for their specific science case. This is particularly crucial when the samples are chosen from the *Gaia* database, which covers the entire sky. The efforts made by the surveys to calibrate their stellar parameters using benchmark objects are devoted to that goal. However, the situation is not yet satisfactory, because systematic differences become apparent when comparing the different surveys and catalogs. For the time being, this is preventing the community from making optimal use of the huge chemical information that is available. A few systematic studies comparing surveys have been published recently (e.g., Jofré et al. 2017b, Anguiano et al. 2018, Jönsson et al. 2018), and we expect that more such studies will become available in the coming years.

(PASTEL, *Gaia* benchmark stars; Soubiran & Girard 2005, among others). Surface gravities are further calibrated using 72 giants with seismic information from K2 (Valentini et al. 2017), and T_{eff} is calibrated with photometric relations. To assess the errors in stellar parameters, extensive comparisons with external data sets are performed (see below).

4.2.1.1. *Abundance uncertainties.*

- **Random/Instrument:** Synthetic spectra with different levels of added noise and repeated observations.
- **Random/Lines:** Not reported.
- **Systematic/Theory:** Not reported.
- **Systematic/Parameters:** Effect on abundances from uncertainties in parameters.
- **Systematic/Methods:** Not reported.
- **Bias/Trends:** Investigated as a function of temperature.
- **Bias/External:** Comparison with Soubiran & Girard (2005) and Ruchti et al. (2011).
- **Final error budget:** The combination of these different tests showed the typical accuracy to be 0.2 dex depending on S/N and atmospheric parameters.

Together with the standard DR5 pipeline, RAVE spectra have been reanalyzed by Casey et al. (2017) with the data-driven method The Cannon (Ness et al. 2015), providing abundances of O, Mg, Al, Si, Ca, Fe, and Ni for red giant stars in a complementary catalog called RAVE-on. The training set was built with RAVE stars that are also in APOGEE (see Section 4.2.3). The typical precision, estimated from repeated observations, is much better (0.07 dex) than it is for the standard pipeline, but the sample is smaller owing to the challenge of finding suitable training sets for the entire parameter space covered by RAVE.

Kunder et al. (2017) compare stellar parameters of stars in common between RAVE and other surveys. Their table 5 is a nice summary of the situation, showing, in general, agreements for $S/N > 50$ spectra in T_{eff} on the order of 100 K, and in $\log g$ and $[\text{Fe}/\text{H}]$ of about 0.1 to 0.2 dex. Recent comparisons of the different sets of RAVE parameters (DR4, DR5, RAVE-on) to independent determinations by Placco et al. (2018) in the low-metallicity regime show significant discrepancies well above the errors mentioned earlier. Casey et al. (2017) compare the RAVE-on abundances with GES (see Section 4.2.2) for 30 stars in common. The total differences vary from 0.06 dex for $[\text{Al}/\text{H}]$ and $[\text{Mg}/\text{H}]$ to 0.26 dex for $[\text{Si}/\text{H}]$.

4.2.2. *Gaia-ESO Survey.* The GES (Gilmore et al. 2012, Randich et al. 2013) targets 10^5 stars in different populations of the Milky Way, as well as in a large sample of open clusters of different characteristics with FLAMES (Fibre Large Array Multi-Element Spectrograph) on the 8-m VLT. One special feature of this survey is that it targets stars of wider spectral ranges than FGK-type owing to different science cases for which groups within the consortium have developed their own analysis methods. The GIRAFFE spectrograph was used with several wavelength ranges, depending on the stellar type, providing spectra for stars down to $V = 19$. In parallel, UVES spectra were obtained in each field for brighter stars. Observations took place between 2011 and 2018. The latest data release (DR3) includes observations and data processing between 2011 and 2014. The information about the release can be found in an ESO document (ESO Public Surv. Program 2017). In total, 12 FLAMES configurations were used, the wavelength intervals and resolutions of which are given by Pancino et al. (2017), and **Figure 7** shows those used for FGK-type stars. Different working groups share the spectral analysis task depending on the spectral type. Within each of them several teams participate in the analyses. For FGK-type stars, they span the entire range of methods described in Section 2.4, namely from fitting synthetic spectra

to determination of EWs. In the latest data release, abundances of Li, C, N, O, Na, Mg, Al, Si, Ca, Sc, Ti, V, Cr, Fe, Co, Ni, Zn, Y, Zr, Ba, La, Ce, Nd, and Eu are included.

Common inputs were adopted for the analyses, including a set of MARCS model atmospheres and a grid of synthetic spectra computed with Turbospectrum following de Laverny et al. (2012). Several teams did not employ the grid of model spectra and used other radiative transfer codes such as SME or MOOG. For the line list, the best atomic data available in the literature were collected, excluding astrophysically calibrated *gf*-values. In addition, lines used to determine abundances were flagged according to their laboratory data quality, as well as according to the amount of blending estimated from a comparison of observed and synthetic line profiles for the Sun and Arcturus. Teams were encouraged to use this flagging system to choose the best lines for their analysis. However, this did not impede a different line selection among them (see **Figure 4** in Section 2.5.2).

4.2.2.1. Abundance uncertainties.

- **Random/Instrument:** Repeated observations of the *Gaia* benchmark stars, other observations across S/N, and cluster stars.
- **Random/Lines:** For the UVES analysis, the line-by-line scatter is utilized to perform line selection for final results. Furthermore, if a method does not treat HFS properly for a line that requires it, their result for that line is rejected.
- **Systematic/Theory:** The SME method included non-LTE calculations for some elements for UVES spectra. They are, however, weighted in the same way as other methods.
- **Systematic/Parameters:** Cluster stars to assess scatter. Final results on abundances are weighted according to their errors in stellar parameters.
- **Systematic/Methods:** Most of the decisions that determine the final abundances are based on agreement between methods.
- **Bias/Trends:** Trends of the abundances derived from each line by all methods are checked as a function of T_{eff} and $\log g$. Each line is excluded from or included in the final result depending on the trend and scatter. Final abundances are checked with cluster stars.
- **Bias/External:** The abundances of the benchmark stars and of the Sun (Grevesse et al. 2007) are used as priors to help establish the scales. No other catalog is used to compare abundances.
- **Final error budget:** Estimated from Bayesian modeling to infer the typical errors of the parameters and abundances from the distributions of values provided by the teams, and is typically between 0.1 and 0.2 dex, depending on S/N and spectral type. We note that published abundances are provided for stars selected to have reliable parameters only, namely with uncertainties of less than 5% in T_{eff} , 0.4 dex in $\log g$, and 0.2 dex in [Fe/H]. These criteria resulted in 2,000 to 9,000 stars with abundances in DR3, depending on the element.

Due to the many different methodologies, spectra, and stellar types targeted by the survey, GES dedicates substantial effort to understanding any systematic differences. A complex communication strategy has been put in place to provide feedback between the data reduction and the analysis teams, as well as the homogenization group producing the final abundances. The size of the data set increases with each data release, and the abundances improve thanks to the various intercomparisons of results between teams and spectral setups the survey considers.

The final data release will include spectra of all observations, the stellar parameters will use the *Gaia*-DR2 astrometric information as priors, and the calibration strategy will include the *Gaia* benchmark stars as well as a sample of seismic targets from CoRoT and K2. Everything (individual results from each methodology and spectra) will be made publicly available.

4.2.3. APOGEE. APOGEE (Majewski et al. 2017), one of the SDSS surveys, was optimized to explore the dust-hidden populations in the Milky Way. Using the 2.5-m Sloan Telescope, APOGEE has been collecting spectra since 2011. Recently, an identical spectrograph has been installed at the 2.5-m du Pont Telescope in Chile with the goal of extending the survey to the Southern Hemisphere. Observations on that telescope started in February 2017. Spectra cover the range from 1.514 to 1.696 μm at $R \simeq 22,500$. The latest SDSS data release (DR14) described by Abolfathi et al. (2018) corresponds to $\sim 260,000$ stars observed until 2016. The release includes abundances of C, N, O, Na, Mg, Al, Si, P, S, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, and Rb. The stars are predominantly red giants and other luminous postmain-sequence stars situated in the obscured parts of the Galactic disk and bulge.

The APOGEE Stellar Parameter and Chemical Abundances Pipeline, ASPCAP, is described by García Pérez et al. (2016). ASPCAP is based on the FERRE code (see **Table 1**), which finds the best fit between observed and synthetic spectra. Holtzman et al. (2018) describe the details of the spectral analyses of the latest data releases of APOGEE (DR13 and D14). The synthetic grid was built with ATLAS9 models (MARCS models are included for the coolest M giants), and Turbospectrum and has many dimensions (T_{eff} , $\log g$, $[M/H]$, $[\alpha/Fe]$, v_{mic} , and v_{mac} ; $[C/H]$ and $[N/H]$ for giants; and rotational velocity for dwarfs). Micro- and macroturbulence are determined from empirical relations that depend on stellar parameters. The final stellar parameters are then empirically calibrated. Similar to RAVE, T_{eff} is calibrated with the help of photometric temperatures and $\log g$ with stars that benefit from seismic observations. The latter uses $\sim 2,000$ stars from the APOKASC catalog (Pinsonneault et al. 2014), which is a joint effort between APOGEE and *Kepler* for the purpose of this calibration.

Elemental abundances are determined by fitting parts of the spectra within spectral windows located around features of each element. These windows are constructed using weights for each spectrum pixel proportional to the change of the flux with the abundance at the corresponding wavelength. The line list employed comes mainly from the National Institute of Standards and Technology (NIST) and is described by Shetrone et al. (2015). For the latest data releases it has been improved by adding HFS for Al and Co, as well as molecular data for H_2O , which is important for the coolest stars. As laboratory data for IR lines are more scarce than in the optical, gf -values and damping constants were astrophysically calibrated (see Section 2.5.1) for $\sim 20,000$ lines by fitting synthetic line profiles to observed ones for the Sun and Arcturus.

4.2.3.1. Abundance uncertainties.

- **Random/Instrument:** Evaluated using stars in 23 globular and open clusters; scatter in abundances is provided as a function of S/N.
- **Random/Lines:** Because abundances are determined from a simultaneous fit of all absorption features, this uncertainty is not given. However, differences compared to optical regions are studied by Jönsson et al. (2018).
- **Systematic/Theory:** Not reported.
- **Systematic/Parameters:** Evaluated using clusters; scatter in abundances is provided as a function of stellar parameters.
- **Systematic/Methods:** Only one method used.
- **Bias/Trends:** Investigated using clusters, finding trends as a function of metallicity and temperature. Abundances are calibrated with polynomials as a function of metallicity and temperature.
- **Bias/External:** $[X/M]$ was calibrated shifting the zero point to force the mean abundance ratios of all stars with $-0.1 < [M/H] < 0.1$, $-5^\circ < b < 5^\circ$, and $70^\circ < l < 110^\circ$ to have solar abundance ratios based on the catalogs of Bensby et al. (2014).

- **Final error budget:** For stars with $T_{\text{eff}} = 4,500$ K, $[M/H] = 0$, and $S/N = 100$ typical uncertainties vary between 0.02 and 0.1 dex. The global uncertainty (the scatter of all cluster stars) is about 0.02 dex larger.

APOGEE is the first ambitious project to collect near-IR spectra at massive scales, opening a new window of spectroscopy and pushing the progress in modeling of spectral lines that for a long time have been essentially unexplored. The spectra have, in general, $S/N > 100$ and are public, which has produced several complementary data sets of parameters and stellar abundances with alternative methods. APOGEE data have become a favorite playground for developing new machine-learning tools to derive abundances at industrial scales. The first application has been The Cannon (Ness et al. 2015). The results on abundances of a modified version (Casey et al. 2016) are part of DR14, which uses a subset of ASPCAP labels to train the model for providing The Cannon labels. Comparisons of the results from The Cannon and ASPCAP are given by Holtzman et al. (2018).

The Payne has been introduced by Ting et al. (2018), deriving parameters and 15 elemental abundances from APOGEE spectra. The method fits the data to models in a special way that allows precise and quick determination of many labels simultaneously. It is based on neural networks with gradient spectra (change in model spectra as each stellar label is varied by a small amount). Its performance in terms of abundance precision is competitive with The Cannon, but offers the possibility of building a parameter-complete training set because it is based on synthetic spectra. Leung & Bovy (2018) showed how abundances of APOGEE spectra can be derived with deep learning using artificial neural networks. The results, trained on a selection of reliable ASPCAP labels, are as precise as those from The Cannon. Stars are also analyzed extremely quickly, showing the potential of this method for future big data sets of stellar abundances.

An early attempt to compare parameters and α -element abundances of ~ 200 stars in common between GES and APOGEE is provided by Jofré et al. (2017b), who discuss the main outputs of a workshop held with key developers of survey chemical pipelines. T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$, and $[\alpha/\text{Fe}]$ agree within 120 K, 0.27, 0.15, and 0.14 dex, respectively. However, they explain that $[\alpha/\text{Fe}]$ should not be directly compared (or merged between surveys!), since the values from GES and APOGEE are based on absorption features that are produced by different α elements.

Jönsson et al. (2018) compared stellar parameters and abundances from APOGEE DR13 and DR14, which were determined with the ASPCAP pipeline and by The Cannon, and included independent analyses focused on the optical. They selected five studies in the literature with high-quality parameters and abundances that had at least 100 stars in common with APOGEE, including GES. For most of the elements (C, Na, Mg, Al, Si, S, Ca, Cr, Mn, and Ni), the DR14 ASPCAP analysis showed systematic differences with the comparison samples of less than 0.05 dex (median) and random differences of less than 0.15 dex (standard deviation). Fe, Mg, and Ni are the elements that show the best agreement with the reference values.

4.2.4. GALAH. The Galactic Archaeology with HERMES (GALAH) survey is a large high-resolution spectroscopic survey using the High-Efficiency and Resolution Multi-Element Spectrograph (HERMES) on the 3.9-m Anglo-Australian Telescope. The HERMES spectrograph provides spectra for ~ 400 stars simultaneously over a 2-degree field of view (De Silva et al. 2015). The goal is to observe up to 10^6 stars and to measure 30 individual chemical element abundances per star from Li to Eu with errors below 0.1 dex.

GALAH DR2 (Buder et al. 2018a) provided abundances of Li, C, O, Mg, Si, Ca, Ti, Na, Al, K, Sc, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Rb, Sr, Y, Zr, Ba, La, Ru, Ce, Nd, and Eu for 342,682 stars observed between January 2014 and 2018. The spectra are not public. The pipeline has two

steps: (a) A training set is defined that is analyzed with SME and (b) The Cannon is applied to the entire data set. The spectral analysis of the training set considers MARCS models and the *Gaia*-ESO line list (see Section 4.2.2), complemented by following the same procedure for the spectral ranges outside GES. An interesting characteristic of this survey is that the training set has been analyzed in non-LTE (using SME), for elements for which this is possible (Li, O, Na, Mg, Al, Si, and Fe). The training set was built to be representative of the parameter space, with stars having relevant information from the literature, such as the *Gaia* benchmark stars, asteroseismic targets, stars with known parallaxes, members of open and globular clusters, and further stars used as references in other projects [TESS (*Transiting Exoplanet Survey Satellite*), K2, APOGEE]. This resulted in a total of 10,605 stars, although not all of them have abundance measurements for all elements. Validation tests include repeated observations, *Gaia* benchmark stars, open and globular clusters, and K2 asteroseismic targets. To assess any biases, GALAH performs leave-out tests, i.e., randomly selecting 80% of the stars from the training set for training, and testing on the other 20%. For the production run, however, the entire training set is used.

4.2.4.1. Abundance uncertainties.

- **Random/Instrument:** Analysis of repeated observations in the field and scatter of abundances found in M67.
- **Random/Lines:** Because SME performs a simultaneous fit of all lines, this is not reported.
- **Systematic/Theory:** The SME method performs non-LTE calculations for some elements. Comparisons are reported by Buder et al. (2018b) and Gao et al. (2018).
- **Systematic/Params:** Not reported.
- **Systematic/Methods:** Not reported.
- **Bias/Trends:** Investigated as a function of $\log g$.
- **Bias/External:** Investigated using leave-out tests.
- **Final error budget:** Estimated with stars belonging to the open cluster M67. Uncertainties were found to range from the highest precisions of 0.04–0.08 dex (Fe, Al, Sc, Ti, V, and Cu), over high precisions of 0.08–0.12 dex (C, Na, Si, Cr, and Mn) and intermediate precisions of 0.12–0.16 dex (O, Mg, K, Ca, Co, Ni, Zn, and Y), to low precisions >0.16 dex (Li, Ba, La, and Eu).

4.2.5. LAMOST. The LAMOST Experiment for Galactic Understanding and Exploration (LEGUE) survey of Milky Way stellar structures is conducted at the 4-m Guo Shoujing Telescope in China (Zhao et al. 2006, Deng et al. 2012). Using a modified Schmidt telescope, LAMOST can observe up to 4,000 objects simultaneously over a 20-square-degree field of view. LAMOST DR3 published spectra for more than 3 million stars. Despite low resolution (see **Figure 7**), several groups have managed to measure elemental abundances from this vast set of spectra. Li et al. (2016) developed a template-matching technique to measure $[\alpha/\text{Fe}]$ ratios with an accuracy better than 0.1 dex for $S/N > 20$. With The Cannon, Ho et al. (2017a) performed a label transfer from APOGEE to LAMOST and measured $[\alpha/M]$ for 454,180 giants, 20% of the LAMOST DR2 and the largest catalog of $[\alpha/M]$ for giant stars to date. Xiang et al. (2017) estimated stellar atmospheric parameters, absolute magnitudes, and elemental abundances ($[M/H]$, $[\alpha/M]$, $[C/H]$, $[N/H]$) from the LAMOST spectra with Kernel-based principal component analysis using an algorithm trained with stars in common with other catalogs (*Hipparcos*, *Kepler*, APOGEE). They obtained a precision of 0.1 dex for $[Fe/H]$, $[C/H]$, and $[N/H]$, and better than 0.05 dex for $[\alpha/M]$. Boeche et al. (2018) obtained $[\alpha/H]$ abundances for 1,097,231 stars. The typical precision is ~ 0.15 dex in $[Fe/H]$ and ~ 0.1 dex in $[\alpha/Fe]$ for spectra with $S/N > 40$, with some differences between dwarf and giant stars. Ting et al. (2017) measured 14 elemental abundances (C, N, O, Mg, Al, Si, Ca, Ti, V, Cr,

Mn, Fe, Co, and Ni) for objects with $S/N > 30$ using The Payne, with a training set made of ~ 500 cross-matched objects between the APOGEE DR13 and LAMOST DR3 catalogs with a LAMOST $S/N > 200$. The typical precision is 0.1 dex.

Lee et al. (2015) applied a special version of the SSPP to LAMOST spectra (LSSPP), thus obtaining $[\alpha/\text{Fe}]$ and $[\text{C}/\text{Fe}]$, and compared their results to the parameters obtained with the regular LAMOST pipeline, and to those from RAVE, APOGEE, and SEGUE. SEGUE and LAMOST are found to be not on the same abundance scale, with an offset of 0.15 in metallicity. APOGEE and SEGUE are in very good agreement for $[\alpha/\text{Fe}]$, whereas LSSPP seems to underestimate it.

The parameters of APOGEE and LAMOST have also been compared by Anguiano et al. (2018). Using $\sim 40,000$ stars in common between APOGEE DR14 and LAMOST DR3, they evaluated mean $[\text{Fe}/\text{H}]$ discrepancies as a function of S/N , T_{eff} , $[\text{Fe}/\text{H}]$, and $\log g$. Even though on average the metallicity offset between the two surveys was found to be low (0.03 or 0.06 dex depending on the LAMOST pipeline used) and the scatter reasonable (0.13 dex), complex dependencies between the parameters were found. They report significant discrepancies of 0.10 to 0.15 dex among metal-poor stars and also show that the differences increase with decreasing T_{eff} .

LAMOST is the perfect data set to test the new generation of data-driven methods, because, on one hand, the spectra are of low resolution and do not allow one to perform the standard methods for abundance determination. On the other hand, the data set is very large, so fast methods are needed.

4.2.6. Forthcoming industrial abundances. The era of large spectroscopic surveys has just begun. Several even bigger projects are planned for the next decade. The next future survey is WEAVE (Dalton et al. 2012), a new multiobject survey spectrograph for the 4.2-m William Herschel Telescope (WHT) at the Observatorio del Roque de los Muchachos on La Palma (Canary Islands). The facility will be capable of obtaining about 1,000 spectra over a two-degree field of view in a single exposure starting in 2019. WEAVE's fibre-fed spectrograph comprises two arms, one optimized for the blue and one for the red, and offers two possible spectroscopic resolutions, 5,000 and 20,000, respectively.

Gaia has a spectrograph on board (RVS or Radial Velocity Spectrometer) covering a wavelength interval around the Ca IR triplet with a resolving power of $R \sim 11,500$, which is similar to the RAVE or the GIRAFFE HR21-setup spectra (Recio-Blanco et al. 2016). We know from RAVE that, from such spectra, it is possible to derive abundances for a limited number of elements (see Section 4.2.1). The third *Gaia* data release, expected in the first half of 2021, will release millions of stellar parameters, abundances, and spectra.²

The 4-m Multi-Object Spectroscopic Telescope project (4MOST) is the next ESO spectroscopic survey facility on the VISTA (Visible and Infrared Survey Telescope for Astronomy) telescope, scheduled to start observations in 2022. With its large field of view it will be able to simultaneously obtain spectra of $\sim 2,400$ objects. Feltzing et al. (2018; see also Bensby et al. 2019, Chiappini et al. 2019, Christlieb et al. 2019, de Jong et al. 2019, Helmi et al. 2019) present an overview of the science goals, spectral properties, and the design of the chemical pipeline. From the high-resolution spectra it will be possible to measure chemical abundance ratios to better than 0.1 dex for Fe, Mg, Si, Ca, Ti, Na, Al, V, Cr, Mn, Co, Ni, Y, Ba, Nd, and Eu, and better than 0.2 dex for Zr, La, and Sr (Caffau et al. 2013). This precision comes from the number of lines in simulated spectra at different S/N and excludes systematic uncertainties related to stellar parameters or atomic data. Caffau et al.'s table 1 shows the number of lines for each element that is expected to be detected with 4MOST. However, Hansen et al. (2015) discuss the possibility

²<https://www.cosmos.esa.int/web/gaia/release>.

that, from the bluest arm of 4MOST, it might be possible to detect new elements for certain stars, these being heavy elements such as Pb, Th, Dy, Ce, and Sm. Similar abundances should be obtained from WEAVE due to their spectral similarities (see **Figure 7**). 4MOST will be succeeded by another ESO facility on the 8-m VLT (Very Large Telescope), the Multi-Object Optical and Near-infrared Spectrograph (MOONS; Cirasuolo et al. 2014). It will combine a wide field of view (~ 500 square arcmin) with a large degree of multiplexity and wavelength coverage (1,000 fibers, optical to near-IR). MOONS has a medium-resolution ($R = 5,000$) and a high-resolution ($R = 20,000$) mode, the latter focused on the J and H bands.

The Prime Focus Spectrograph (PFS; Takada et al. 2014) is the next generation facility instrument on the 8.2-m Subaru Telescope. It is a very wide-field, massively multiplexed optical and near-IR spectrograph that will dedicate a portion of its time to observing 10^6 stars in the Galactic thick disk, halo, and tidal streams for magnitudes down to $V \sim 22$. A medium-resolution mode with $R = 5,000$ to be implemented in the red arm will enable the measurement of α -element abundances. Finally, the Maunakea Spectroscopic Explorer (MSE; McConnachie et al. 2016), a rebirth of the 3.6-m Canada-France-Hawaii Telescope on Maunakea, is a proposed 11.25-m wide-field (1.5 square-degree) telescope, equipped with multi-object spectrographs, that will obtain for each pointing more than 4,000 optical and near-IR spectra of low, intermediate, and high resolution.

4.3. Discussion on Individual Elements

Widely measured elements can be identified in the histogram of **Figure 6** as those that have more than $\sim 4,000$ measurements and are covered by at least all surveys with high-resolution and large wavelength coverage (GALAH, GES, and APOGEE; see **Figure 8**). These are C, O, Na, Mg, Al, Si, Ca, Ti, V, Cr, Mn, Fe, Co, and Ni. Cu and Nd are common elements targeted by surveys but not as common as the others in the Hypatia catalog.

To evaluate the precision of abundances for the common elements (see Section 4.3), let us take as an example the abundances published by GES to illustrate the situation in the optical wavelength region. We choose GES because of its unique strategy of combining several results using different methods based on a common list of atomic data with special effort in flagging “good” and “bad” lines (see Section 4.2.2). This allows one to assess the method-to-method dispersion (MMD) or LLD for different species. The precision of the abundances evaluated using the MMD was found to be highest (< 0.15 to 0.2 dex) for the species NaI, MgI, AlI, SiI, CaI, ScII, TiI, VI, FeI, ZrI, MoI, and BaII (Smiljanic et al. 2014, p. 23). The elements with precise abundances in common among all setups are AlI, SiI, and CaI. In the IR, abundances more precise than 0.05 dex in APOGEE DR14 (Holtzman et al. 2018) are those of C, O, Mg, Al, Si, Ca, Ti, Cr, Mn, Fe, and Ni. In summary, popular elements in spectroscopic surveys in the optical and IR that are also precise are Mg, Al, Si, Ca, Ti, and Fe. These elements all have agreements and biases between optical and IR surveys within the uncertainties (Jönsson et al. 2018). The works of Jofré et al. (2015), Hawkins et al. (2016b), and Jönsson et al. (2018) help us to investigate in depth which are the best lines used for abundance analyses in the IR and optical. We exclude Fe from this discussion as being the element with the largest number of lines available in FGK-star spectra, which forms the basis for the metallicity parameter. (See the sidebar titled Most Precise and Accurate Elements.)

Common elemental abundances that can be precisely measured from IR spectra, but not from optical spectra, are C, O, and Mn. C and O abundances can be derived from molecular features, whereas few clean Mn lines are suitable. In the optical the situation is more complicated, as there is only one C I and one O I line that have good atomic data and are free of blends. The C I line is very weak, and O I lies in telluric regions. The O triplet at 7770 \AA is popular, and is included in GALAH, but it is subject to strong non-LTE effects (Amarsi et al. 2018a, and references therein).

MOST PRECISE AND ACCURATE ELEMENTS

- **Magnesium (6318 Å, 6319 Å, 15231 Å, 15693 Å, 15740 Å, 15879 Å, 15954 Å, 16365 Å):** In the GES range, 12 MgI lines are included. Among them, 2 have accurate atomic data and are good lines for a variety of stars and methods. In APOGEE, there are 14 MgI lines, among which 6 are suitable for a large variety of stars. Non-LTE corrections have been determined in the optical by Osorio et al. (2015) and in the IR by Zhang et al. (2017) and Bergemann et al. (2017).
- **Aluminum (5557 Å, 6696 Å, 6699 Å, 8773 Å, 8774 Å, 16763 Å):** No reliable experimental transition probability data exist for the five AlI lines in the GES line list. However, none of them is considered heavily blended, possibly explaining the good precision (MMD) achieved despite the lack of reliable laboratory data. They can be used for a variety of stars, preferentially of solar metallicities. In APOGEE, three AlI lines are present; however, one seems to be more robust for different types of stars. Non-LTE studies of optical and IR lines have been performed by Nordlander & Lind (2017).
- **Silicon (5690 Å, 5701 Å, 5949 Å, 15377 Å, 15888 Å, 16216 Å, 16681 Å, 16828 Å):** Among the 45 SiII lines visible in GES spectra, only 3 have good atomic data and are blend free in dwarfs and giants. The partially blended line at $\lambda 5708$ Å with good atomic data is strong enough to be measured in metal-poor stars. APOGEE has 17 SiII features, among which 5 are good lines to be used for a variety of stars. Improved modeling for SiII lines has been done by Shi et al. (2012) and Amarsi & Asplund (2017) for optical lines and by Zhang et al. (2016) for optical and IR lines.
- **Calcium (5867 Å, 16150 Å, 16157 Å, 16197 Å):** The GES has 31 CaI and 8 CaII lines detectable, among which 12 and 2 have reliable laboratory data, respectively. The lines are highly sensitive to stellar parameters and, therefore, only one can be used to derive abundances for a large variety of spectral types (excluding metal-poor stars). In APOGEE, 4 lines are visible, but only 3 can be used for a wide range in stellar parameters. Non-LTE corrections in the optical are provided for the *Gaia* benchmark stars by Jofré et al. (2015). In the IR, calculations are needed.
- **Titanium (5689 Å, 5702 Å, 6091 Å, 15873 Å):** After FeI, TiI is the element with the most numerous absorption features, including 105 in GES spectra, among which 23 have good atomic data and are largely blend free. Some 3 lines can be used for elemental abundances in a large variety of stars. In APOGEE, 9 Ti lines are detected (but see Section 2.5.2 for a discussion on their applicability). Non-LTE effects have been studied by Bergemann (2011) for optical lines. In the IR, calculations are needed.

For Mn, although many lines are detected in the optical, they are subject to strong HFS, affecting resulting abundances by up to 0.6 dex if HFS is not properly taken into account (Jofré et al. 2017c). Sodium is derived with high precision from optical spectra but has more problems in the IR. There are three NaI lines that are considered largely clean in the optical. Although they are subject to HFS, the effect is small, because the LLD and MMD for the abundances are small (Smiljanic et al. 2014). In the IR, the two available absorption features are in most cases too weak or blended. Non-LTE effects for NaI can be very strong (Lind et al. 2011). Finally, V and Co are elements that are common but more uncertain than the other elements, being subject to HFS. Although V and Co lines are very numerous in the optical, they can be quite weak. They become scarce and too weak in the IR or for metal-poor stars. V and Co abundances are normally derived from blue lines for metal-poor stars (Snedden et al. 2016).

The fact that our periodic table in **Figure 8** has most cells unmarked means that, for most of the elements, it is either very challenging or impossible to detect and model lines for typical spectra of FGK-type stars in order to obtain accurate abundances. From **Figure 8**, the elements that are potentially detected in less than two surveys correspond to Li, P, S, K, Sc, Zn, Ge, Rb, Sr, Y,

PROSPECTS FOR MEASURING ABUNDANCES OF NEW ELEMENTS IN SURVEYS

The most challenging elements are those for which the lines are scarce, too weak, and blended. A number of such challenging lines belong to heavy neutron-capture elements, which are part of families of other nucleosynthesis channels than the ones typically measured from survey spectra. They might provide more dimensions in chemical space, and the information might be available for future machine-learning approaches. We must provide a good training set, including the corresponding abundances, using spectra of high resolution, high S/N, and of very extended wavelength range. This is key for maximizing the number of industrial abundances that will be extracted from future surveys.

Zr, Nb, Mo, Ru, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Dy, and Yb. Most of the abundances provided by surveys for these elements are so uncertain that they are not released publicly. Further discussion on these elements can be found in **Supplemental Text 2**. (See the sidebar titled Prospects for Measuring Abundances of New Elements in Surveys.)

Supplemental Material >

5. RECONSTRUCTING THE HISTORY OF THE GALAXY WITH ABUNDANCES

Freeman & Bland-Hawthorn (2002) discussed the power of using abundances of FGK-type stars in order to find the building blocks of the Galaxy. This is possible if (a) stars are born in groups that share the same chemical composition; (b) the evolution of this chemical composition is the product of a unique combination of star formation and nucleosynthesis, which depends on where and when the stars were born; and (c) FGK-type stars retain the information of their chemical make-up in their atmospheres.

In general, there is evidence that the principles mentioned above are correct. But when we become ambitious and want to recover the history of each star in the Galaxy, the principles discussed above are challenged by secondary effects in the general picture. In this section, we discuss a selection of topics in which active research is ongoing, thanks to the progress in deriving precise abundances for large samples of stars.

5.1. Nucleosynthesis Channels and Chemical Dimensions

Freeman & Bland-Hawthorn (2002) indicate that to reconstruct the expected 10^8 star-formation sites in the disk, a chemical (C) space of at least 10 abundance ratios reflecting different nucleosynthesis channels, at a precision better than 0.05 dex, would be needed. It is good news that many current large catalogs contain abundances of more than 10 elements (see **Figure 8**), but some are correlated owing to their similar production mechanisms (see discussion in, e.g., Liu et al. 2016b). Burbidge et al. (1957) suggested different chemical families according to their nucleosynthesis paths. A detailed description of the nucleosynthesis channels from supernovae is provided by Nomoto et al. (2013). Karakas & Lattanzio (2014) give a similar review regarding nucleosynthesis from AGB stars. Briefly, elements can be divided into five major families: α -capture, iron-peak, odd- Z , light, and neutron-capture. Each of these families contain elements or isotopes that might be produced by different channels (hence environments and timescales!), and so abundance ratios of elements within a family increase the dimensions in C space, serving as diagnostics to study chemical evolution.

Ting et al. (2012) study this dimensionality with a principle component analysis, combining $[X/Fe]$ ratios and using different catalogs from the literature. They found six major components

formed by combinations of elements that are correlated with nucleosynthesis channels. Depending on the catalog (e.g., the amount and precision of abundances) and its overall metallicity distribution, the components are formed by different combinations of abundance ratios. The neutron-capture family was shown to have a large impact on the number of dimensions. Later, Ting et al. (2015) continued the discussion, showing that such C spaces would still allow one to find many prominent groups of stars in the Galaxy (10^3 – 10^4). They noted that the number and the size of the groups (C cells) depends more on the uncertainties of the abundances measured than on the number of elements or number of stars that a survey might have. It remains to be seen if considering abundance ratios between or within families might increase the number of dimensions of the C space for chemical tagging with current spectroscopic surveys and abundance uncertainties. For example, populations have been shown to separate in $[\text{Mn}/\text{Mg}]$ (Hawkins et al. 2015), $[\text{Co}/\text{Cr}]$, and $[\text{Ca}/\text{Mg}]$ (McWilliam et al. 1995), and in $[\text{Ba}/\text{Eu}]$ and $[\text{Ba}/\alpha]$ (Tolstoy et al. 2009).

5.2. From Chemical Tagging to Galactic Phylogenetics

Using the chemical elements to identify the groups of stars that have common origins forms the basis of chemical tagging. As discussed above, this can work given the principle that every group formed at a given place and time in the Universe has a unique chemical pattern. Freeman & Bland-Hawthorn (2002) postulated that this chemical pattern can be attributed to the stellar DNA, such that chemical tagging could allow for temporal sequencing of stars, which is similar to building a family tree through DNA sequencing. This is only possible because the chemical patterns evolve with time and not in a random way. In fact, there is a chemodynamical model based on physical principles describing how stellar generations become more metal rich with time (Kobayashi et al. 2006). The fact that there is a physical process behind the change of the chemical pattern of stars implies that chemical elements carry evolutionary information from one generation of stars to the next. This principle of ancestry forms the basis of phylogenetic studies. In Galactic phylogenetics (Jofré et al. 2017a), the only useful traits are the chemical elements, because no other trait (e.g., kinematics, ages, stellar parameters) carries information that is passed from one generation to the next. Using these traits to construct phylogenetic trees can provide a powerful way to constrain the chemical evolution model underneath, in the same way as many other applications in evolutionary studies.

5.3. New Challenges for Chemical Evolution with High-Precision Abundances

As discussed above, improving precision allows one to detect more C cells. With strictly differential techniques, in which abundances are derived with respect to reference stars of the same spectral type, precisions of 0.01 dex have been achieved (Nissen & Gustafsson 2018). With the help of such precise abundances, clusters might be shown to have chemical inhomogeneities above that level, challenging the first principle mentioned at the start of this section. For example, Liu et al. (2016b) analyzed the Hyades cluster, finding that stars of the cluster can have an abundance dispersion on the order of 0.02 to 0.03 dex.

For typical uncertainties at a more industrial level, cluster stars have a dispersion in abundances that is of the order of the measurement errors (Bovy 2016). It is important to quantify this dispersion, as it is key for prospects of chemical tagging. Hogg et al. (2016) recovered stars from open clusters from precise abundances from APOGEE using k -means clustering. However, Ness et al. (2018) found a small fraction of field stars that have abundances that are indistinguishable from cluster stars within the uncertainties yet have different birth origins. Stars of different origins and the same chemical abundances [called *doppelgängers* by Ness et al. (2018)] should not exist for chemical tagging to work. Accurate ages, kinematics (although see discussions in

Mitschang et al. 2014), and the inclusion of more dimensions in C space with, e.g., neutron-capture elements, are important for understanding the nature of these stars.

Atomic diffusion is another challenge. It is well known that heavy elements sink toward the center of stars owing to gravitational settling. Dotter et al. (2017) studied this effect in the context of chemical tagging, showing that abundances can significantly decrease over the lifetime of a star. For the effects of diffusion to be minimized, stars of the same evolutionary stage, and $[X/Fe]$, instead of $[X/H]$ abundance ratios, should be used. It is, however, still plausible that stars of the same evolutionary phase but different masses (and hence ages) will present small differences in $[X/Fe]$ ratios detectable at the 0.01-dex level of precision, which could explain the problem of the doppelgängers, for example. To truly quantify these differences, better theoretical treatment of atomic diffusion is needed, in particular the combined effect with radiative levitation.

High-precision spectroscopic studies have shown that stellar abundances might encode signatures of planet formation. The encoding typically appears as a trend of $[X/Fe]$ as a function of condensation temperature (T_c). Meléndez et al. (2009) determined precise abundances of 11 solar twins and found that the solar refractory elements were more deficient than the volatile elements when compared to other stars. Recently, Bedell et al. (2018) presented a comprehensive discussion of the chemical homogeneity of Sun-like stars, considering 79 solar twins with 30 measured elements and finding that the Sun has indeed an unusual slope of T_c versus $[X/Fe]$. Whether this implies an unusual formation scenario for the planets of the Solar System is still debated. In any case, it is certainly important to keep in mind that there is a possibility that stars that have formed from the same molecular cloud might present different trends of abundances with T_c . The binary 16 Cyg AB analyzed by Tucci Maia et al. (2014) is an example, although the binary α Cen AB analyzed by Morel (2018) is a contradictory example.

5.4. Masses and Ages from Stellar Abundances

Recent active discussions in the literature show that masses and ages can be determined from stellar spectra. Such discussions were initiated by Masseron & Gilmore (2015), who showed that the C/N of APOGEE DR12 red giants revealed that the thin disk and the thick disk had different formation histories because of their different distributions in stellar masses. They explain that C/N relates to mass because of a very fundamental principle in stellar evolution: As red giants experience the dredge-up, synthesized material from the CNO cycle at their cores is brought outward, which results in an enhancement of the nitrogen surface abundance at the expense of carbon. The amount of mixing depends on the depth of the dredge up, which depends on the mass of the star.

Masseron & Gilmore (2015) warn that translating C/N into mass is complicated by the uncertainties in stellar evolutionary models, especially the initial metallicity and $C+N$ abundances, and the poorly understood effects of mixing-length theory, as well as the role of opacities due to α -element enhancement and extra mixing in evolved giants. Salaris et al. (2015), Masseron et al. (2017), and Lagarde et al. (2017) investigated these complications with dedicated studies of stellar evolution theory and C/N abundances. The complications have, however, not impeded Martig et al. (2016) or Das & Sanders (2019) from building empirical relations for masses, ages, and C/N thanks to the relative large spectroscopic data sets for which ages and masses can be derived from asteroseismology. Such relations have helped to create maps of ages and masses of the Galaxy with data-driven methods (e.g., Ness et al. 2016, Ho et al. 2017b, Das & Sanders 2019) with APOGEE or LAMOST data. This is another example of the importance of having seismic data for Galactic studies.

6. DISCUSSION AND CONCLUDING REMARKS

Astronomy, being one of the oldest sciences of mankind, has been traditionally hampered by the lack of good data. Astronomers have thus developed the habit of blaming the small amount of data available for the unanswered questions. They have been dreaming of having millions of stars with accurate data and that the wealth of data will help us progress in understanding how our own Galaxy is shaped.

This dream is becoming true! Today, Galactic astronomy benefits from exquisite data: Billions of accurate parallaxes from *Gaia* and thousands of high-resolution spectra and asteroseismic data are enabling us to dig into the physics of stars. We are experiencing a unique opportunity to reformulate our understanding about stellar properties because data quantity and quality are not a problem anymore. The variety of data available naturally has led to a variety of analysis methods, which, at this rapidly growing data rate, have the dangerous potential of diverging significantly and so too do their results. Fortunately, we are learning that the best way to maximize the accuracy and precision of our data products is to work together as a community, that surveys need to be compared and complemented. RAVE has provided a crucial playground for learning what we will obtain from the forthcoming millions of similar spectra from *Gaia*-RVS; the GES has been revolutionary in making us aware of the impact of different methodologies on results; APOGEE has shown us the power of moving outside our wavelength range of comfort (from the optical to the IR); GALAH is going one step ahead in pushing for better modeling and propagating the improved parameters with data-driven approaches to an entire data set; and SDSS and LAMOST are demonstrating that we can trace the chemistry of huge volumes of the Galaxy even with low resolution.

We still do not have a best method to determine accurate and precise stellar parameters and abundances for all stars in the Galaxy. But intercomparisons between surveys are starting to become mandatory thanks to new efforts to observe common targets between surveys. We are getting closer to defining standard procedures for comparing and connecting results, being more aware of uncertainties, and having well-defined strategies to improve them. We are also learning to appreciate that we need both high- and low-quality data. High-quality data (high S/N and resolution) are crucial for improving the theory of line formation, as well as to identify and study the chemical signatures that are needed to maximize the size of the chemical-space in the Galaxy. Seemingly low-quality data (low S/N and resolution) still provide the only realistic way to travel across the Galaxy and probe its outskirts. Clearly, the information we can obtain from the latter kind of data fully depends on what we can obtain from the former. Thus, a concerted effort is key for taking full advantage of the treasure that the previous generation of astronomers has given us thanks to their habit of wanting more data.

Working together means also making a serious effort of standardizing our data products, not only in terms of meaningful physical properties (e.g., [Fe/H], [α /Fe], etc.) but also in terms of their format. Catalogs, in particular those created by independent groups intended to be used for reference, must be published in the Virtual Observatory (via CDS). This one extra step in the publishing procedure can lead to productive synergies in the community. Reference stars are the basis for Galactic studies and must be accessible for the entire community if we want pipeline products to truly converge in accuracy. Likewise, we must decide on the most suitable pipeline to yield abundances that are at the same common scale. This is only possible if the data of stars in common are made public, as this will provide researchers with the opportunity to freely test their tools and reproduce results. We should learn from the examples of SDSS, APOGEE, LAMOST, and *Gaia*.

This review attempts to link details of the art of determining abundances of a single line from a single star with the art of propagating this information to millions of stars, as well as how each step and star is crucial for this chemical ladder to work. In Section 2, we described the several

steps involved in abundance determination, and in Section 3, we discussed different tests that help us quantify the uncertainties in abundances. In Section 4, we described how catalogs and surveys apply these steps to provide abundances at industrial scales, discussing which elements are common and which are not. Finally, in Section 5, we discussed a selection of science applications in which these abundances are being used.

The Milky Way, our home Galaxy, harbors stars of a great variety, each of them containing unique information about their present and past environment. The evolution of this environment encodes the laws of physics in an elegant way that we have yet to fully decipher, with the fossil stars there to help us out. Starting from the Sun, we need to find the best way to connect to other reference stars, ensuring that accuracy and precision are maintained. From these reference stars, we can then connect to all other stars of the Milky Way, taking care that we are not overlooking even one single star that might contain the key missing information about the assembly history of our home Galaxy.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We apologize in advance for all the investigations that could not be properly cited here due to space limitations. We thank our collaborators K. Hawkins, P. Das, and M. Tucci Maia for various discussions; S. Buder for clarifying details about the GALAH pipeline; and C. Worley, R. Smiljanic, and G. Gilmore for clarifying details of the GES pipeline. We further thank P.E. Nissen and B. Gustafsson for fruitful discussions and for sharing unpublished material that was needed for writing complementary reviews on stellar abundances as well as T. Beers for his careful revision. P.J. and U.H. thank C. Hidalgo for hosting us at the MIT Media Lab at a critical time of writing for this review. Finally, P.J. warmly thanks T. Madler for unlimited support in all matters.

P.J. acknowledges financial support of FONDECYT Iniciaci3n 11170174. U.H. acknowledges support from the Swedish National Space Agency (SNSA/Rymdstyrelsen). Posted with permission from the *Annual Review of Astronomy and Astrophysics*, Volume 57 © by Annual Reviews, <http://www.annualreviews.org>.

LITERATURE CITED

- Abolfathi B, Aguado DS, Aguilar G, et al. 2018. *Ap. J. Suppl.* 235:42
Adibekyan V, Delgado-Mena E, Figueira P, et al. 2016. *Astron. Astrophys.* 591:A34
Adibekyan VZ, Sousa SG, Santos NC, et al. 2012. *Astron. Astrophys.* 545:A32
Allende Prieto C. 2016a. *Astron. Nachr.* 337:837
Allende Prieto C. 2016b. *Living Rev. Solar Phys.* 13:1
Allende Prieto C, Sivarani T, Beers TC, et al. 2008. *Astron. J.* 136:2070–82
Amarsi AM, Asplund M. 2017. *MNRAS* 464:264–73
Amarsi AM, Barklem PS, Asplund M, et al. 2018a. *Astron. Astrophys.* 616:A89
Amarsi AM, Nordlander T, Barklem PS, et al. 2018b. *Astron. Astrophys.* 615:A139
Anguiano B, Majewski SR, Allende-Prieto C, et al. 2018. *Astron. Astrophys.* 620:A76
Asplund M. 2005. *Annu. Rev. Astron. Astrophys.* 43:481–530
Asplund M, Grevesse N, Sauval AJ, Scott P. 2009. *Annu. Rev. Astron. Astrophys.* 47:481–522
Bailer-Jones CAL, Andrae R, Arcay B, et al. 2013. *Astron. Astrophys.* 559:A74

- Barklem PS. 2016. *Astron. Astrophys. Rev.* 24:9
- Barklem PS, Anstee SD, O'Mara BJ. 2015. Astrophysics Source Code Library. ascl:1507.007
- Barklem PS, Christlieb N, Beers TC, et al. 2005. *Astron. Astrophys.* 439:129–51
- Battistini C, Bensby T. 2015. *Astron. Astrophys.* 577:A9
- Battistini C, Bensby T. 2016. *Astron. Astrophys.* 586:A49
- Bedell M, Bean JL, Meléndez J, et al. 2017. *Ap. J.* 839:94
- Bedell M, Bean JL, Meléndez J, et al. 2018. *Ap. J.* 865:68
- Beers TC, Flynn K, Gebhardt K. 1990. *Astron. J.* 100:32–46
- Beers TC, Preston GW, Shectman SA. 1985. *Astron. J.* 90:2089–102
- Bensby T, Bergemann M, Rybizki J, et al. 2019. *Messenger* 175:35–38
- Bensby T, Feltzing S, Oey MS. 2014. *Astron. Astrophys.* 562:A71
- Bergemann M. 2011. *MNRAS* 413:2184–98
- Bergemann M, Cescutti G. 2010. *Astron. Astrophys.* 522:A9
- Bergemann M, Collet R, Amarsi AM, et al. 2017. *Ap. J.* 847:15
- Bergemann M, Kudritzki RP, Plez B, et al. 2012. *Ap. J.* 751:156
- Bertaux JL, Lallement R, Ferron S, et al. 2014. *Astron. Astrophys.* 564:A46
- Bijaoui A, Recio-Blanco A, de Laverny P, Ordenovic C. 2010. In *ADA 6 - 6th Conference on Astronomical Data Analysis, in Honor of Albert Bijaoui*, Monastir, Tunisia, May 3–7, ed. JL Starck, M Saber Naceur, F Murtagh, id.9. http://www.aset.org.tn/conf/ADA6/online_presentations.php
- Blackwell DE, Shallis MJ. 1977. *MNRAS* 180:177–91
- Blanco-Cuaresma S, Soubiran C, Heiter U, Jofré P. 2014a. *Astron. Astrophys.* 569:A111
- Blanco-Cuaresma S, Soubiran C, Jofré P, Heiter U. 2014b. *Astron. Astrophys.* 566:A98
- Boeche C, Grebel EK. 2016. *Astron. Astrophys.* 587:A2
- Boeche C, Siebert A, Williams M, et al. 2011. *Astron. J.* 142:193
- Boeche C, Smith MC, Grebel EK, et al. 2018. *Astron. J.* 155:181
- Bovy J. 2016. *Ap. J.* 817:49
- Brewer JM, Fischer DA. 2018. *Ap. J. Suppl.* 237:38
- Brewer JM, Fischer DA, Valenti JA, Piskunov N. 2016. *Ap. J. Suppl.* 225:32
- Buder S, Asplund M, Duong L, et al. 2018a. *MNRAS* 478:4513–52
- Buder S, Lind K, Ness MK, et al. 2018b. *Astron. Astrophys.* 624:A19
- Burbidge EM, Burbidge GR, Fowler WA, Hoyle F. 1957. *Rev. Mod. Phys.* 29:547–650
- Caffau E, Koch A, Sbordone L, et al. 2013. *Astron. Nachr.* 334:197
- Cantat-Gaudin T, Donati P, Pancino E, et al. 2014. *Astron. Astrophys.* 562:A10
- Cantat-Gaudin T, Jordi C, Vallenari A, et al. 2018. *Astron. Astrophys.* 618:A93
- Carney BW, Gray DF, Yong D, et al. 2008. *Astron. J.* 135:892–906
- Casamiquela L, Carrera R, Blanco-Cuaresma S, et al. 2017. *MNRAS* 470:4363–81
- Casey AR, Hawkins K, Hogg DW, et al. 2017. *Ap. J.* 840:59
- Casey AR, Hogg DW, Ness M, et al. 2016. arXiv:1603.03040
- Castelli F, Kurucz RL. 2003. Poster presented at the *Modelling of Stellar Atmospheres, LAU Symposium 210*, Uppsala, Sweden, June 17–21, 2002 (Poster A20)
- Chaplin WJ, Miglio A. 2013. *Annu. Rev. Astron. Astrophys.* 51:353–92
- Chiavassa A, Casagrande L, Collet R, et al. 2018. *Astron. Astrophys.* 611:A11
- Chiappini C, Minchev I, Starkenburg, et al. 2019. *Messenger* 175:30–34
- Christlieb N, Battistini C, Bonifacio P, et al. 2019. *Messenger* 175:26–29
- Cirasuolo M, Afonso J, Carollo M, et al. 2014. In *Ground-based and Airborne Instrumentation for Astronomy V*, ed. SK Ramsay, IS McLean, H Takami, *Proc. SPIE Conf. Ser.* 9147:91470N. Bellingham, WA: SPIE
- Dalton G, Trager SC, Abrams DC, et al. 2012. In *Ground-based and Airborne Instrumentation for Astronomy IV*, ed. IS McLean, SK Ramsay, H Takami, *Proc. SPIE Conf. Ser.* 8446:84460P. Bellingham, WA: SPIE
- Das P, Sanders J. 2019. *MNRAS* 484:294–304
- de Jong RS, Agertz O, Berbel AA, et al. 2019. *Messenger* 175:3–11
- de Laverny P, Recio-Blanco A, Worley CC, Plez B. 2012. *Astron. Astrophys.* 544:A126
- De Pascale M, Worley CC, de Laverny P, et al. 2014. *Astron. Astrophys.* 570:A68

- De Silva GM, Freeman KC, Bland-Hawthorn J, et al. 2015. *MNRAS* 449:2604–17
- Delgado Mena E, Tsantaki M, Adibekyan VZ, et al. 2017. *Astron. Astrophys.* 606:A94
- Deng LC, Newberg HJ, Liu C, et al. 2012. *Res. Astron. Astrophys.* 12:735–54
- Dotter A, Conroy C, Cargile P, Asplund M. 2017. *Ap. J.* 840:99
- ESO Public Surv. Program, Gilmore G. 2017. *Gaia-ESO spectroscopic survey, data release DR3.1*, ESO Science Archive Facility, Munich, Germ. <http://www.eso.org/rm/api/v1/public/releaseDescriptions/92>
- Ezzeddine R, Frebel A, Plez B. 2017. *Ap. J.* 847:142
- Feltzing S, Bensby T, Bergemann M, et al. 2018. *Galactic Archeology with 4MOST*. In *Rediscovering Our Galaxy, IAU Symp. 334*, ed. C Chiappini, I Minchev, E Starkenburg, M Valentini, pp. 225–32. Cambridge, UK: Cambridge Univ. Press
- Feltzing S, Chiba M. 2013. *New Astron. Rev.* 57:80–99
- Freeman K, Bland-Hawthorn J. 2002. *Annu. Rev. Astron. Astrophys.* 40:487–537
- Freytag B, Steffen M, Ludwig HG, et al. 2012. *J. Comput. Phys.* 231:919–59
- Fuhrmann K. 1998. *Astron. Astrophys.* 338:161–83
- Gaia Collab., Babusiaux C, van Leeuwen F, et al. 2018a. *Astron. Astrophys.* 616:A10
- Gaia Collab., Brown AGA, Vallenari A, et al. 2018b. *Astron. Astrophys.* 616:A1
- Gao X, Lind K, Amarsi AM, et al. 2018. *MNRAS* 481:2666–84
- García Pérez AE, Allende Prieto C, Holtzman JA, et al. 2016. *Astron. J.* 151:144
- Gilmore G, Randich S, Asplund M, et al. 2012. *The Messenger* 147:25–31
- Gray RO, Corbally CJ. 1994. *Astron. J.* 107:742–46
- Grevesse N, Asplund M, Sauval AJ. 2007. *Space Sci. Rev.* 130:105–14
- Grevesse N, Sauval AJ. 1998. *Space Sci. Rev.* 85:161–74
- Guiglion G, de Laverny P, Recio-Blanco A, et al. 2016. *Astron. Astrophys.* 595:A18
- Gustafsson B, Edvardsson B, Eriksson K, et al. 2008. *Astron. Astrophys.* 486:951–70
- Hansen CJ, Ludwig HG, Seifert W, et al. 2015. *Astron. Nachr.* 336:665
- Hawkins K, Jofré P, Heiter U, et al. 2016a. *Astron. Astrophys.* 592:A70
- Hawkins K, Jofré P, Masseron T, Gilmore G. 2015. *MNRAS* 453:758–74
- Hawkins K, Masseron T, Jofré P, et al. 2016b. *Astron. Astrophys.* 594:A43
- Heiter U, Eriksson K. 2006. *Astron. Astrophys.* 452:1039–48
- Heiter U, Jofré P, Gustafsson B, et al. 2015. *Astron. Astrophys.* 582:A49
- Heiter U, Soubiran C, Netopil M, Paunzen E. 2014. *Astron. Astrophys.* 561:A93
- Helmi A, Irwin M, Deason A, et al. 2019. *Messenger* 175:23–25
- Hinkel NR, Mamajek EE, Turnbull MC, et al. 2017. *Ap. J.* 848:34
- Hinkel NR, Timmes FX, Young PA, et al. 2014. *Astron. J.* 148:54
- Hinkel NR, Young PA, Pagano MD, et al. 2016. *Ap. J. Suppl.* 226:4
- Ho AYQ, Ness MK, Hogg DW, et al. 2017a. *Ap. J.* 836:5
- Ho AYQ, Rix HW, Ness MK, et al. 2017b. *Ap. J.* 841:40
- Hogg DW, Casey AR, Ness M, et al. 2016. *Ap. J.* 833:262
- Holtzman JA, Hasselquist S, Shetrone M, et al. 2018. *Astron. J.* 156:125
- Ivezić Ž, Beers TC, Jurić M. 2012. *Annu. Rev. Astron. Astrophys.* 50:251–304
- Ivezić Ž, Connelly AJ, VanderPlas JT, Gray A. 2014. *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data*. Princeton, NJ: Princeton Univ. Press
- Jofré P, Das P, Bertranpetit J, Foley R. 2017a. *MNRAS* 467:1140–53
- Jofré P, Heiter U, Buder S. 2017b. In *Proceedings of the International Workshop on Stellar Spectral Libraries 2017*, ed. P Coelho, L Martins, E Griffin, *Astron. Soc. India Conf. Ser.* 14:37–44
- Jofré P, Heiter U, Soubiran C, et al. 2014. *Astron. Astrophys.* 564:A133
- Jofré P, Heiter U, Soubiran C, et al. 2015. *Astron. Astrophys.* 582:A81
- Jofré P, Heiter U, Tucci Maia M, et al. 2018. *Res. Notes Am. Astron. Soc.* 2:152
- Jofré P, Heiter U, Worley CC, et al. 2017c. *Astron. Astrophys.* 601:A38
- Jönsson H, Allende Prieto C, Holtzman JA, et al. 2018. *Astron. J.* 156:126
- Karakas AI, Lattanzio JC. 2014. *Publ. Astron. Soc. Aust.* 31:e030
- Kobayashi C, Umeda H, Nomoto K, et al. 2006. *Ap. J.* 653:1145–71

- Kos J, Lin J, Zwitter T, et al. 2017. *MNRAS* 464:1259–81
- Kunder A, Kordopatis G, Steinmetz M, et al. 2017. *Astron. J.* 153:75
- Kurucz RL. 1993. *SYNTHE Spectrum Synthesis Programs and Line Data*. Cambridge, MA: Smithson. Astrophys. Obs.
- Kurucz RL. 2011. *Can. J. Phys.* 89:417–28
- Kurucz RL. 2014. In *Determination of Atmospheric Parameters of B-, A-, F- and G-Type Stars*, pp. 39–51. Cham: Springer
- Lagarde N, Robin AC, Reylé C, Nasello G. 2017. *Astron. Astrophys.* 601:A27
- Laverick M, Lobel A, Royer P, et al. 2018. *Galaxies* 6:78
- Lawler JE, Sneden C, Nave G, et al. 2017. *Ap. J. Suppl.* 228:10
- Lee YS, Beers TC, Carlin JL, et al. 2015. *Astron. J.* 150:187
- Lee YS, Beers TC, Sivarani T, et al. 2008a. *Astron. J.* 136:2022–49
- Lee YS, Beers TC, Sivarani T, et al. 2008b. *Astron. J.* 136:2050–69
- Leung HW, Bovy J. 2018. *MNRAS* 483:3255–77
- Li J, Han C, Xiang MS, et al. 2016. *Res. Astron. Astrophys.* 16:110
- Lind K, Amarsi AM, Asplund M, et al. 2017. *MNRAS* 468:4311–22
- Lind K, Asplund M, Barklem PS, Belyaev AK. 2011. *Astron. Astrophys.* 528:A103
- Liu F, Asplund M, Yong D, et al. 2016a. *MNRAS* 463:696–704
- Liu F, Yong D, Asplund M, et al. 2016b. *MNRAS* 457:3934–48
- Luck RE. 2018. *Astron. J.* 155:111
- Magic Z, Collet R, Asplund M, et al. 2013. *Astron. Astrophys.* 557:A26
- Mahdi D, Soubiran C, Blanco-Cuaresma S, Chemin L. 2016. *Astron. Astrophys.* 587:A131
- Majewski SR, Schiavon RP, Frinchaboy PM, et al. 2017. *Astron. J.* 154:94
- Martig M, Fouesneau M, Rix HW, et al. 2016. *MNRAS* 456:3655–70
- Mashonkina L, Jablonka P, Sitnova T, et al. 2017. *Astron. Astrophys.* 608:A89
- Masseron T, Gilmore G. 2015. *MNRAS* 453:1855–66
- Masseron T, Lagarde N, Miglio A, et al. 2017. *MNRAS* 464:3021–28
- Masseron T, Plez B, Van Eck S, et al. 2014. *Astron. Astrophys.* 571:A47
- McConnachie A, Babusiaux C, Balogh M, et al. 2016. arXiv:1606.00060
- McWilliam A, Preston GW, Sneden C, Searle L. 1995. *Astron. J.* 109:2757
- Meléndez J, Asplund M, Gustafsson B, Yong D. 2009. *Ap. J. Lett.* 704:L66–70
- Meléndez J, Schirbel L, Monroe TR, et al. 2014. *Astron. Astrophys.* 567:L3
- Mészáros S, Allende Prieto C, Edvardsson B, et al. 2012. *Astron. J.* 144:120
- Mikolaitis Š, de Laverny P, Recio-Blanco A, et al. 2017. *Astron. Astrophys.* 600:A22
- Mitschang AW, De Silva G, Zucker DB, et al. 2014. *MNRAS* 438:2753–64
- Morel T. 2018. *Astron. Astrophys.* 615:A172
- Mucciarelli A, Pancino E, Lovisi L, et al. 2013. *Ap. J.* 766:78
- Ness M. 2018. *Publ. Astron. Soc. Aust.* 35:e003
- Ness M, Hogg DW, Rix HW, et al. 2015. *Ap. J.* 808:16
- Ness M, Hogg DW, Rix HW, et al. 2016. *Ap. J.* 823:114
- Ness M, Rix HW, Hogg DW, et al. 2018. *Ap. J.* 853:198
- Nissen PE. 2015. *Astron. Astrophys.* 579:A52
- Nissen PE, Gustafsson B. 2018. *Astron. Astrophys. Rev.* 26:6
- Nissen PE, Schuster WJ. 2010. *Astron. Astrophys.* 511:L10
- Nissen PE, Silva Aguirre V, Christensen-Dalsgaard J, et al. 2017. *Astron. Astrophys.* 608:A112
- Nomoto K, Kobayashi C, Tominaga N. 2013. *Annu. Rev. Astron. Astrophys.* 51:457–509
- Nordlander T, Amarsi AM, Lind K, et al. 2017. *Astron. Astrophys.* 597:A6
- Nordlander T, Lind K. 2017. *Astron. Astrophys.* 607:A75
- Osorio Y, Barklem PS, Lind K, et al. 2015. *Astron. Astrophys.* 579:A53
- Palacios A, Gebran M, Josselin E, et al. 2010. *Astron. Astrophys.* 516:A13
- Pancino E, Lardo C, Altavilla G, et al. 2017. *Astron. Astrophys.* 598:A5

- Pancino E, Mucciarelli A, Bonifacio P, et al. 2011. *Astron. Astrophys.* 534:A53
- Petigura EA, Howard AW, Marcy GW, et al. 2017. *Astron. J.* 154:107
- Pinsonneault MH, Elsworth Y, Epstein C, et al. 2014. *Ap. J. Suppl.* 215:19
- Piskunov N, Valenti JA. 2017. *Astron. Astrophys.* 597:A16
- Piskunov NE, Valenti JA. 2002. *Astron. Astrophys.* 385:1095–106
- Placco VM, Beers TC, Santucci RM, et al. 2018. *Astron. J.* 155:256
- Plez B, Brett JM, Nordlund A. 1992. *Astron. Astrophys.* 256:551–71
- Prugniel P, Soubiran C. 2001. *Astron. Astrophys.* 369:1048–57
- Randich S, Gilmore G, Gaia-ESO Consort. 2013. *The Messenger* 154:47–49
- Recio-Blanco A, Bijaoui A, de Laverny P. 2006. *MNRAS* 370:141–50
- Recio-Blanco A, de Laverny P, Allende Prieto C, et al. 2016. *Astron. Astrophys.* 585:A93
- Reggiani H, Meléndez J, Yong D, et al. 2016. *Astron. Astrophys.* 586:A67
- Roederer IU, Preston GW, Thompson IB, et al. 2014. *Astron. J.* 147:136
- Roederer IU, Sneden C, Lawler JE, et al. 2018. *Ap. J.* 860:125
- Ruchti GR, Feltzing S, Lind K, et al. 2016. *MNRAS* 461:2174–91
- Ruchti GR, Fulbright JP, Wyse RFG, et al. 2011. *Ap. J.* 737:9
- Ryabchikova T, Piskunov N, Kurucz RL, et al. 2015. *Phys. Scr.* 90:054005
- Saffe C, Flores M, Miquelarena P, et al. 2018. *Astron. Astrophys.* 620:A54
- Salaris M, Pietrinferni A, Piersimoni AM, Cassisi S. 2015. *Astron. Astrophys.* 583:A87
- Sameshima H, Matsunaga N, Kobayashi N, et al. 2018. *Publ. Astron. Soc. Pac.* 130:074502
- Shetrone M, Bizyaev D, Lawler JE, et al. 2015. *Ap. J. Suppl.* 221:24
- Shi JR, Takada-Hidai M, Takeda Y, et al. 2012. *Ap. J.* 755:36
- Smiljanic R, Korn AJ, Bergemann M, et al. 2014. *Astron. Astrophys.* 570:A122
- Sneden C, Cowan JJ, Kobayashi C, et al. 2016. *Ap. J.* 817:53
- Sneden CA. 1973. *Carbon and nitrogen abundances in metal-poor stars*. PhD Thesis, The University of Texas at Austin.
- Sobeck JS, Lawler JE, Sneden C. 2007. *Ap. J.* 667:1267–82
- Soto MG, Jenkins JS. 2018. *Astron. Astrophys.* 615:A76
- Soubiran C, Girard P. 2005. *Astron. Astrophys.* 438:139–51
- Soubiran C, Le Campion JF, Brouillet N, Chemin L. 2016. *Astron. Astrophys.* 591:A118
- Sousa SG, Santos NC, Adibekyan V, et al. 2015. *Astron. Astrophys.* 577:A67
- Sousa SG, Santos NC, Israelian G, et al. 2011. *Astron. Astrophys.* 533:A141
- Steffen M, Caffau E, Ludwig HG. 2013. *Mem. Soc. Astron. Ital. Suppl.* 24:37
- Steinmetz M, Zwitter T, Siebert A, et al. 2006. *Astron. J.* 132:1645–68
- Stello D, Vanderburg A, Casagrande L, et al. 2016. *Ap. J.* 832:133
- Stello D, Zinn J, Elsworth Y, et al. 2017. *Ap. J.* 835:83
- Strassmeier KG, Ilyin I, Weber M. 2018. *Astron. Astrophys.* 612:A45
- Suda T, Katsuta Y, Yamada S, et al. 2008. *Publ. Astron. Soc. Jpn.* 60:1159–71
- Takada M, Ellis RS, Chiba M, et al. 2014. *Publ. Astron. Soc. Jpn.* 66:R1
- Tennyson J, Yurchenko SN, Al-Refai AF, et al. 2016. *J. Mol. Spectrosc.* 327:73–94
- Ting YS, Conroy C, Goodman A. 2015. *Ap. J.* 807:104
- Ting YS, Conroy C, Rix HW, Cargile P. 2018. arXiv:1804.01530
- Ting YS, Freeman KC, Kobayashi C, et al. 2012. *MNRAS* 421:1231–55
- Ting YS, Rix HW, Conroy C, et al. 2017. *Ap. J. Lett.* 849:L9
- Tolstoy E, Hill V, Tosi M. 2009. *Annu. Rev. Astron. Astrophys.* 47:371–425
- Tucci Maia M, Meléndez J, Ramírez I. 2014. *Ap. J. Lett.* 790:L25
- Valenti JA, Fischer DA. 2005. *Ap. J. Suppl.* 159:141–66
- Valentini M, Chiappini C, Davies GR, et al. 2017. *Astron. Astrophys.* 600:A66
- Worley CC, de Laverny P, Recio-Blanco A, et al. 2012. *Astron. Astrophys.* 542:A48
- Worley CC, de Laverny P, Recio-Blanco A, et al. 2016. *Astron. Astrophys.* 591:A81
- Xiang MS, Liu XW, Shi JR, et al. 2017. *MNRAS* 464:3657–78

- Yana Galarza J, Meléndez J, Ramírez I, et al. 2016. *Astron. Astrophys.* 589:A17
Yanny B, Rockosi C, Newberg HJ, et al. 2009. *Astron. J.* 137:4377–99
Zhang J, Shi J, Pan K, et al. 2016. *Ap. J.* 833:137
Zhang J, Shi J, Pan K, et al. 2017. *Ap. J.* 835:90
Zhao G, Chen YQ, Shi JR, et al. 2006. *Chin. J. Astron. Astrophys.* 6:265–80