Chemical gradients in the Milky Way from unsupervised chemical abundances measurements of the RAVE spectroscopic data set

Corrado Boeche

Potsdam 2010

Astrophysikalisches Institut Potsdam

Chemical gradients in the Milky Way from unsupervised chemical abundances measurements of the RAVE spectroscopic data set

Dissertation
zur Erlangung des akademischen Grades
"doctor rerum naturalium"
(Dr. rer. nat.)
in der Wissenschaftsdisziplin "Astrophysik"
der Universität Potsdam

eingereicht an der
Mathematisch–Naturwissenschaftlichen Fakultät
der Universität Potsdam

von
Corrado Boeche

Potsdam, October 2010

Contents

Allgemeinverständliche Zusammenfassung

Die vorliegende Doktorarbeit wurde im Rahmen des RAdial Velocity Experiment (RAVE) angefertigt. Ihr Ziel ist es, chemische Elementhäufigkeiten an RAVE-Spektren zu messen und zur Untersuchung chemischer Gradienten in der Milchstrassenebene zu benutzen, um verschieden Szenarien der Galaxienentstehung einzugrenzen.

RAVE ist eine grosse spektrokopische Durchmusterung, deren Ziel es ist, bis zum Ende des Jahres 2012 insgesamt 10^6 Sterne zu spektroskopieren, um deren Radialgeschwindigkeiten, sternatmosphärische Parameter und chemische Häufigkeiten zu messen. Das Projekt benutzt das UK Schmidt Teleskop am Australian Astronomical Observatory (AAO) in Siding Spring, Australien, welches mit dem Multiobjekt-Spektrographen 6dF bestückt ist. Bis heute hat RAVE die Spektren von mehr als 450,000 Sternen gesammelt und untersucht. Die Genauigkeit, mit der die Elementhäufigkeiten abgeschätzt werden können, hängt von der Zuverlässigkeit der verwendeten Parameter, (insbesondere der Oszillatorstärken der Absorptionslinien sowie von der effektiven Temperatur, Schwerebeschleunigung und der Metallizität des gemessenen Sterns) ab. Daher identifizierten wir zunächst 604 Absorptionslinien im Wellenlängenbereich von RAVE und verbesserten deren Oszillatorstärken durch eine inverse Spektralanalyse. Dann wurden die stellaren Parameter von RAVE verbessert, indem die RAVE Pipeline und die stellaren Parameter, auf denen sie beruht, modifiziert wurden. Die Änderungen eliminierten einen Teil der systematischen Fehler von stellaren Parametern, die im Laufe dieser Arbeit gefunden wurden. Um Elementhäufigkeiten zu bestimmen, haben wir zwei verschiedene Prozessierungs-Pipelines entwickelt. Beide berechnen die Elementhäufigkeiten unter der Annahme von Sternatmosphären im lokalen thermischen Gleichgewicht (*local thermal equilibrium*, LTE). Die erste Pipeline berechnete Elemenhäufigkeiten anhand der Äquivalentbreiten von Absorptionslinien. Da diese Methode eine geringe Empfindlichkeit für die Elementhäufigeiten relativ zu Eisen erreichte, wurde sie ersetzt. Die neue Pipeline benutzt χ^2-Fits von Modellspektren an die beobachteten Spektren. Dank Ihrer Präzision wurde diese für die Erstellung des RAVE-Katalogs von Elementhäufigkeiten verwendet. Diese Pipeline liefert Elementhäufigkeiten mit einer Genaugkeit von $\sim$0.2dex, während für Spektren mit 20$\leq$S/N$\leq$40 immerhin noch $\sim$0.3dex Genauigkeit erreicht werden. Für die vorliegende Arbeit wurden für 217.358 Sterne die Häufigkeiten von sieben chemischen Elementen bestimmt.

Mit diesen Daten wurde der radiale chemische Gradient unserer Milchstrasse untersucht. Wir finden, dass Sterne mit kleinen vertikalen Geschwindigkeiten |W|, die also nahe der galaktischen Ebene bleiben, einen radialen Gradienten der Eisenhäufigkeit zeigen, der mit früheren Studien übereinstimmt (~ -0.07 dex Kpc^{-1}), während Sterne mit grossen |W|, also solche, die grössere galaktische Höhen erreichen, einen progressiv flachere Gradienten zeigen. Die Gradienten der anderen Element folgen dem gleichen Trend. Das lässt darauf schliessen, dass entweder die Durchmischung der galaktischen dicken Scheibe effizient arbeitet oder aber dass die dicke Scheibe aus interstellarer Materie gebildet wurde, die chemisch recht homogen war. Speziell fanden wir hunderte von Sternen, die zwar kinematisch als zur dicken Scheibe zugehörig klassifiziert werden können, die aber die typische chemische Zusammensetzung der dünnen Scheibe aufweisen. Einige wenige dieser Sterne wurden bereits von anderen Autoren entdeckt, aber ihre Herkunft bleibt immer noch unklar. Eine Möglichkeit ist, dass die Sterne der dünnen Scheibe kinematische geheizt werden, sodass sie effizienter radial gemischt werden, was die chemischen Gradienten verwischt und auch flacher macht. Alternativ dazu könnten diese Sterne einer "Übergangspopulation" angehören, welche hinsichtlich der Scheibenevolution die Verbindung zwischen der dünnen und der dicken Scheibe darstellt. Unsere Untersuchung zeigt, dass sich diese beiden Erklärungen gegenseitig nicht ausschliessen. Künftigte Nachspektroskopierung mit hoher Auflösung wird die Rolle dieser Sterne in der Entwicklungsgeschichte der galaktischen Scheibe aufklären.

Abstract

The present thesis was born and evolved within the RAdial Velocity Experiment (RAVE) with the goal of measuring chemical abundances from the RAVE spectra and exploit them to investigate the chemical gradients along the plane of the Galaxy to provide constraints on possible Galactic formation scenarios.

RAVE is a large spectroscopic survey which aims to observe spectroscopically $\sim 10^6$ stars by the end of 2012 and measures their radial velocities, atmospheric parameters and chemical abundances. The project makes use of the UK Schmidt telescope at Australian Astronomical Observatory (AAO) in Siding Spring, Australia, equipped with the multiobject spectrograph 6dF. To date, RAVE collected and measured more than 450,000 spectra. The precision of the chemical abundance estimations depends on the reliability of the atomic and atmosphere parameters adopted (in particular the oscillator strengths of the absorption lines and the effective temperature, gravity, and metallicity of the stars measured). Therefore we first identified 604 absorption lines in the RAVE wavelength range and refined their oscillator strengths with an inverse spectral analysis. Then, we improved the RAVE stellar parameters by modifying the RAVE pipeline and the spectral library the pipeline rely on. The modifications removed some systematic errors in stellar parameters discovered during this work. To obtain chemical abundances, we developed two different processing pipelines. Both of them perform chemical abundances measurements by assuming stellar atmospheres in Local Thermodynamic Equilibrium (LTE). The first one determines elements abundances from equivalent widths of absorption lines. Since this pipeline showed poor sensibility on abundances relative to iron, it has been superseded. The second one exploits the χ^2 minimization technique between observed and model spectra. Thanks to its precision, it has been adopted for the creation of the RAVE chemical catalogue. This pipeline provides abundances with uncertains of about $\sim$0.2dex for spectra with signal-to-noise ratio S/N$\geq$40 and $\sim$0.3dex for spectra with 20$\leq$S/N$\leq$40. For this work, the pipeline measured chemical abundances up to 7 elements for 217,358 RAVE stars.

With these data we investigated the chemical gradients along the Galactic radius of the Milky Way. We found that stars with low vertical velocities |W| (which stay close to the Galactic plane) show an iron abundance gradient in agreement with previous works

(~ -0.07 dex kpc^{-1}) whereas stars with larger |W|, which are able to reach larger heights above the Galactic plane, show progressively flatter gradients. The gradients of the other elements follow the same trend. This suggests that an efficient radial mixing acts in the Galaxy or that the thick disk formed from homogeneous interstellar matter. In particular, we found hundreds of stars which can be kinetically classified as thick disk stars exhibiting a chemical composition typical of the thin disk. A few stars of this kind have already been detected by other authors, and their origin is still not clear. One possibility is that they are thin disk stars kinematically heated, and then underwent an efficient radial mixing process which blurred (and so flattened) the gradient. Alternatively they may be a "transition population" which represents an evolutionary bridge between thin and thick disk. Our analysis shows that the two explanations are not mutually exclusive. Future follow-up high resolution spectroscopic observations will clarify their role in the Galactic disk evolution.

Introduction

The Milky Way is an ideal place to investigate the mechanisms of galaxy formation and evolution. Due to its proximity we can study in detail the physical properties of its stars (temperature, mass, age), how they move around the centre of the Galaxy (kinematics) and what are their chemical composition (metallicity, chemical abundances). To obtain such data we need to observe stars bright enough to secure high resolution, high signal-to-noise spectra. This cannot be done for stars dimly shining in distant galaxies but only for stars of our Galaxy and few nearby galaxies. In particular, to study the Milky Way we need all sky surveys to observe its stars which lie all around us, since we sit inside it. In the last decades the use of dedicated telescopes in combination with the growth of the computational power, enabled us to collect huge amounts of data and opened the way to all sky surveys targeting the Milky Way. Ground-based projects have measured millions of stars by scanning the entire sky in different photometric bands, such as 2MASS (Cutri et al., 2003), DENIS[1], SDSS (Abazajian et al., 2003), Skymapper[2]. Space-based projects collected photometric and astrometric data for millions of stars (Hipparcos mission, Perryman et al. 1997, Tycho2, Høg et al. 2000), a number that will raise to a billion with the up-coming mission Gaia (Perryman et al., 2001) which will also provide spectra for tens of millions of stars. The recent Multi Object Spectroscopy (MOS) technology, which makes use of optical fibers combined with wide field ground-based telescopes, enables to collect spectra for hundreds or thousands stars per exposure, resulting in large spectroscopic surveys like SDSS-SEGUE (Yanny et al., 2009), LAMOST[3] and RAVE (Steinmetz et al., 2006). Such a surveys provide a vast amount of homogeneous data (because observed and reduced with the same procedure), free from pre-selection biases and therefore suitable of statistical investigations. For a successful exploitation of these surveys the data has to be reduced and processed in a reasonable amount of time by an unsupervised computer code, which must be able to measure and evaluate the quality of the data thousands times faster than a human operator. This is one of the today' s astronomy challenge.

The present thesis was born and evolved in the RAVE project with the goal of developing new methods to measure chemical abundances from the RAVE spectra and exploit

---

[1]http://cdsweb.u-strasbg.fr/denis.html
[2]http://www.mso.anu.edu.au/skymapper/index.php
[3]http://www.lamost.org/website/en

them to increase our comprehension of the Milky Way's evolution. RAVE (RAdial Velocity Experiment) is a large spectroscopic survey which aims to observe spectroscopically $\sim 10^6$ stars by the end of 2012. The project makes use of the UK Schmidt telescope at Australian Astronomical Observatory (AAO) in Siding Spring, Australia, equipped with the multiobject spectrograph 6dF (Watson et al., 2000) which allows up to 150 scientific spectra to be collected per exposure. To date RAVE has gathered radial velocities and stellar parameters for ∼450,000 stars in our Galaxy and, thanks to the results of this thesis, ∼200,000 of those now have chemical abundances estimations. The RAVE stars have astrometric measurements from several sources (Tycho2, PPMX, UCAC2) which, together with the RAVE radial velocities and distances (inferred from the RAVE stellar parameters, Breddels et al. 2010, Zwitter et al. 2010) locate the RAVE stars in 6-dimensional phase space. RAVE has today the largest existing database for the Milky Way's stars, a rich source of data with which to study our host galaxy.

Motivations

Since the seminal work by Eggen, Lynden-Bell and Sandage (1962, hereafter ELS) the Galaxy has been recognized to be composed of parts which kinematics and chemical abundances correlate tightly: population I stars (rich in heavy elements) belongs to the disk, have low velocities and nearly circular orbits whereas population II stars (poor in heavy elements) belongs to the halo, have high velocities and eccentric orbits. The ELS's original picture of a galaxy composed of only two populations became more complex after the discovery that the disk actually has two components (Gilmore & Reid 1983). In their work, Gilmore & Reid recognized a variation in stellar density moving far from the Galactic plane which can be well described as a sum of two exponential laws with scale heights of ∼300 pc and ∼1450 pc (more recent estimates suggest ∼1000 pc for the thick disk, for instance Veltz et al., 2008). These two disk components are called the *thin* and *thick* disks, respectively. In later works more differences were identified. The thin disk is composed of stars having low vertical velocity and nearly circular orbits. The thick disk is composed of stars having high vertical velocity and eccentric orbits. The thin disk stars have an age distribution spanning from old (~ 10 Gyrs) to young stars (less than 1 Gyrs) whereas the thick disk appear to be composed of only old stars (~ 10-12 Gyrs, Freeman & Bland-Hawthorn 2002). Detailed chemical analysis shows that these two populations of stars differ in chemical composition as well. The thick disk stars are poor in heavy elements (metal poor) but are "α enhanced" which means that their content of α-elements (O, Mg, Si, Ca, Ti) appear high with respect to iron. On the other hand thin disk stars are more metal rich and not α enhanced. The reason for these differences lies in the origin and formation of the two disks. The thin disk is believed to form from the collapse of a protogalactic cloud within hierarchically assembled dark matter haloes. The haloes gain angular momentum from the inhomogeneous distribution of the surrounding matter and from merging events. The baryonic part of the cloud cool down and shrinks under gravitational force. This gen-

erates a centrifugally supported, thin rotating disk (Fall & Efstathiou 1980, Mo, Mao & White 1998) where star formation can take place. While for the thin disk there is a general formation process which is commonly accepted (despite far from being fully understood), for the thick disk there is no such consensus. Among several possible thick disk formation mechanisms proposed in the last decades, we cite here four of them (Sales et al. 2009) :

- *dynamical heating*: the thick disk is a primordial thin disk puffed up by an accreted satellite (Villalobos & Helmi 2008)

- *accretion*: the thick disk is a relic of disrupted satellites accreted by the Galaxy (Abaldi et al. 2003)

- *radial migration*: the thick disk is composed of thin disk stars which have migrated radially from the inner region of the Galaxy due to the gravitational action of the bar and spiral arms (Schönrich & Binney, 2009)

- *gas rich merger*: the thick disk is composed of stars formed in situ from gas accreted by the Galaxy (Brook et al. 2004, 2005)

All these mechanisms are able to create a thick disk, but whether a combination of several or only one is responsible for the existence of the thick disk we observe in the Galaxy is still a debated question.

To explain the differences in chemical composition between the thin and thick disks, stellar evolution mechanisms becomes precious tools of investigation. We know that at very early times the universe was composed of hydrogen, helium and small amount of lithium. The heavier elements were synthesized in the cores of the first stars and eventually released via supernova events. Every generation of stars therefore enriched with heavy elements the interstellar matter from which a new generation of stars will form. This process repeats for each generation. Thus, every generation of stars holds the chemical imprint of the star formation history experienced in that region up to the time of its birth and, by knowing the quantities of elements released by the supernovae (yields), it is in principle possible to trace back the star formation history of a galaxy.

In order to understand how the interstellar matter is enriched by supernova events we need to know the yields of the supernovae and the relevant timescales. The progenitors of supernovae type II (SNII) are massive stars ($\gtrsim 6M_\odot$). They explode after a few Myrs releasing light and heavy elements and they are known to be the major producer of α elements. The progenitors of supernovae type Ia (SNIa) are white dwarfs which are bound in binary systems. At the end of their main sequence life the remaining He white dwarfs accrete matter from the companion star until their mass reaches the Chandrasekar limit ($\sim 1.4~M_\odot$). When this limit is overcome (after ~ 1 Gyr) the stars explode as SNIa and release mostly Fe.

The different timescales of these supernovae and their different contributions in terms of chemical enrichment allow us to infer the star formation history when the quantities of chemical elements in a large sample of stars is known. Astronomers used to refer to this quantity as the *chemical abundance* of a star. It is written as [X/H] for an X element and can be mathematically expressed as

$$[X/H] = \log \frac{N_X^{star}}{N_H} - \log \frac{N_X^{Sun}}{N_H}$$

where N_X is the number of atoms of the element X, N_H is the number of atoms of hydrogen H I and $\log(N_H)$=12.0. In other words, the abundance gives (in logarithmic form) the number of atoms of the element X for every 10^{12} atoms of H I for a star with respect to the Sun, so that [X/H]=0.0 dex means that the quantity of the X element is equal to that of the Sun, [X/H]=–1.0 dex means that the quantity is one tenth that of the Sun, [X/H]=–2.0 dex means one hundredth, and so on.

An important indicator is the ratio between the elements X and Fe abundances [X/Fe], in particular if this element is an α element. Often, we refer to [α/Fe] as the average of the abundances of all α elements with respect to Fe. A high value of [α/Fe] indicates that the star formed from interstellar matter enriched by more SNII than SNIa. Because of the shorter timescale for SNII, it follows that the star with high [α/Fe] has formed during or just after the time of intense star formation. If [α/Fe] of a stars is low, it means that it formed during a long star formation period, long enough ($\geq$1 Gyrs) to allow SNIa to explode, enrich the interstellar matter of Fe and thus lower the ratio [α/Fe]. As a first approximation, we can think of [α/Fe] as an indicator of the ratio of supernovae events SNII/SNIa experienced by the interstellar matter from which the star formed.

If we could measure the chemical abundances of the stars populating our Galaxy with very high precision and plot them in chemical space (n-dimensional space where the n dimensions represent n chemical abundances) we would see stars tracing paths of evolution, because the abundances of every generation of stars depends on the previous generation and lie a step after it. In other words, there must exist a sort of "genealogy tree" of stars written into the chemical abundances which, in principle, can be traced back in time (Freeman & Bland-Hawthorn 2002).

The unravelling of the Galaxy's "genealogy tree" would greatly improve our understanding of how the Galaxy formed and evolved. Unfortunately, the precision in chemical abundances required for a sample of stars large enough ($> 10^6$) to trace part of the tree is not yet possible. In fact, the number of stars having high resolution (R$\geq$20000) spectroscopic observations today is of the order of thousands, most of them pre-selected for specific studies and therefore not suitable for statistical studies. But the path is open. The up-coming survey like HERMES will obtain high resolution spectra (R$\sim$30000) of a million of objects. The GREAT project is planning high resolution surveys as Gaia's follow up. Future projects like LAMOST and the main Gaia mission promise to provide tens of millions of star spectra at medium resolution (R$\sim$10000). Although at this resolution the "genealogy

tree" will have fuzzy borders and blended branches, the shear number of objects will permit unprecedented statistical investigations.

Today the RAVE project is pioneering in the field by providing hundreds of thousands of stellar spectra. Despite the medium resolution (R∼7500) of its spectra, they represent a huge contribution to the data, gathered up to now by the astronomical community. In fact, at the time of the first RAVE data release (2006), the number of radial velocities available amounted to ∼70000. With the RAVE first data release (Steinmetz et al. 2006) the radial velocities available were raised up of ∼30% and today the RAVE internal data release is 6 times times larger, counting ∼450,000 radial velocities. The status of the chemical abundances available in literature today is not better. For instance, one of the larger homogeneous sample available is given by Valenti&Fisher (2010) who measured the abundances of 5 elements in 1040 nearby F, G and K stars. Many other works present abundances with several (more than 10) element abundances but for sample which are often composed by few hundreds stars or less (Edvardsson et al.,1993, Fuhrmann 1998, 2008, Luck & Heiter 2006, 2007, Reddy et al. 2006, among others). Larger star samples of chemical abundances can be build up by collecting different literature sources like done by Soubiran & Girard (2005) and Venn et al. (2004) who collected 743 and 821 stars, respectively. Unlike RAVE, these samples are very often selected to study particular populations, for instance thin or thick disk, and are therefore kinematically biased.
In this thesis we tackled and successfully accomplished the challenge to create a vast and homogeneous catalogue by measuring the chemical abundances of the RAVE stars with an automatized unsupervised procedure capable of measuring 10^4 spectra per day on a modern single core processor computer. To date we measured chemical abundances up to the 7 elements (Mg, Al, Si, Ca, Ti, Fe and Ni) of ∼ 217,000 RAVE stars, which increases of tens of times the actual numbers of abundances available in literature.

A new code to measure the chemical abundances of stars

The classical method to measure chemical abundances of stars is to measure the equivalent width (EW) of an absorption line of an element X and infer the abundance [X/H] through its curve of growth (COG), a function which bounds the two quantities. The EW measurement is usually done directly on the spectrum by hand and the abundance [X/H] obtained with software for spectral line analysis. Because this method requires handling the spectra one by one, it is not suitable for the measurement of a large number of spectra. Moreover, EW measurement can be done only for isolated (not blended) lines to be sure there is no pollution from other lines. Isolated lines are found more easily in high resolution spectra, where the lines have low chance of being instrumentally blended. At low resolution the number of lines which are not blended diminishes dramatically. At RAVE resolution the

number of isolated lines are few (no more than 20 in favorable cases) making this method useless for most of the spectra. This problem can be solved when the abundances are measured by using a χ^2 minimization technique to match a synthetic spectrum with the observed one. This method requires spectra with different abundances to be synthesized until a good match between the synthetic and observed spectra is found. This way suffers from the disadvantage of requiring more computational power and time to analyze a large number of spectra.

For the RAVE spectra we developed a hybrid technique. We construct a template spectrum in which absorption lines change EWs by following their COGs (which are functions of the abundances [X/H]) and employs the χ^2 technique to match it with the observed spectrum. The construction is not a real synthesis but the resulting template is close to a synthetic spectrum and the time required is very short. The method is fully explained in Chapter 4.

Scientific potential of the RAVE chemical catalogue

The uncertainties of the stellar parameters ($T_{\mathrm{eff}} \sim$400K, $\log g \sim$0.5dex, [m/H] $\sim$0.2dex) and abundances uncertainties ($\sim$0.2dex) of the RAVE data are due to the medium resolution, the small wavelength coverage ($\sim$390 Å), the low/medium signal-to-noise ratios and the limited number of absorption lines (most of them blended) visible in that small wavelength window. Obviously, they cannot compete with the precision reached by the high resolution spectroscopy (typical errors: $T_{\mathrm{eff}} \sim$70K, $\log g \sim$0.1dex, [m/H] and abundances $\leq$0.1dex) which can counts on wide wavelength ranges (usually they can cover 3000-5000Å), high signal-to-noise ratios and a list of well tested absorption lines chosen among thousands visible in the spectrum. Still, the RAVE chemical pipeline provides stars' abundances correctly sorted in a common linear scale. This means that the measured difference in abundance between different stars is reliable, and that their differences with respect to the absolute zero point (the Sun's abundance) is merely a constant.

The strength of the RAVE project is the sheer number of observed objects, unbiased by pre-selection and homogeneously measured, which makes the RAVE catalogue optimal for systematic and statistical investigations of the Milky Way. A number of the possible research directions are listed here:

- identify groups of stars sharing similar chemical characteristics in order to discover disrupted open clusters (star formation history, galactic dynamic)

- identify stars with peculiar abundances belonging to rare class of stars (stellar evolution)

- identify of moving groups created by gravitational actions of Galactic spiral arms and/or bar (galactic dynamics)

- disentangle thin from thick disk stars, study the properties of such populations and search for new characteristics (Galactic formation and evolution)

- verify theoretical predictions of galaxy chemical models such as the abundances at different Galactic radii, and vertical and radial gradients for different elements, in order to give new constraints to the models (Galactic formation and evolution)

- statistical studies on the abundance distributions compared with the models to give constraints to the theoretical SNe yield predictions (stellar evolution)

- verify the relations between chemical abundances and Galactic orbital parameters (such as mean radius, eccentricity, angular momentum) to infer past dynamical events like mergers and accretion of disrupted satellites (interaction of the Galaxy with the cosmic environment)

- evaluate the impact of the radial migration on the observed abundance distributions and test the current dynamical models of the Galaxy (galactic dynamics)

All these studies can also lead to follow-up observations at higher resolution to increase the precision of the stellar parameters, chemical abundances (when needed) and obtain precise age determination. The RAVE database can be, therefore, a self-standing source of science research as well as a useful tool to select sub-samples of stars of particular interest.

We begun this thesis work by checking and refining (when necessary) the atomic oscillator strengths with an inverse spectral analysis of the Sun and Arcturus spectra (Chapter 2), in order to have trustworthy chemical abundances. The resulted oscillator strengths have been adopted in all the following analysis. We revised part of the RAVE pipeline (Chapter 3) to remove some systematic errors came to light during the work and improve the RAVE stellar parameters which are of fundamental importance for reliable abundances estimations. We wrote a first processing pipeline (Chapter 4) based on the "classical" method to estimate the abundances: measure the Equivalent Widths (EWs) of the absorption lines and compute the abundances through a Local Thermodynamic Equilibrium (LTE) analysis (Chapter 4). This method turned out to be not fully satisfactory because most of the absorption lines are instrumentally blended and the measurement of the individual EWs resulted of insufficient quality. Therefore we developed an alternative processing pipeline based on the minimization of the χ^2 between the observed and model spectra (Chapter 5). The model is reconstructed by using a library of EWs from which the curve of growth of the lines are inferred (assuming LTE) and the lines reconstructed. By varying the EWs of the lines (hence the abundances) the pipeline reconstructs several spectra until the best match with the observed spectrum is found. We tested the reliability and accuracy of this pipeline by testing it on samples of synthetic and real spectra (Chapter 6). After the pipeline passed the quality checks we run it on the RAVE spectra and create the present RAVE chemical catalogue which holds chemical abundances up to 7 elements for 217,358

RAVE stars (Chapter 7). We used this catalogue to investigate the chemical gradients along the Galactic radius of the Milky Way (Chapter 8) with new interesting results involving thick disk stars.

Chapter 1

Constraining the atomic oscillator strengths in the RAVE wavelength region

The measurement of the chemical abundances is an indirect measurement which requires the knowledge of physical processes occurring in a stellar atmosphere. The abundance of an element in a stellar atmosphere is inferred from the intensity of one or more absorption lines formed under the physical conditions of the atmosphere itself. Once these conditions (temperature, pressure, opacity) are determined, the intensity of the lines are function of atomic parameters such as excitation potential and oscillator strength (see Eq. 1.6). The oscillator strength (gf, often expressed as a logarithm $\log gf$) is the most problematic and uncertain. An incorrect $\log gf$ can affect significantly the inferred abundance.

In this chapter we compile a list of absorption lines identified in the RAVE wavelength range, check and (if needed) constrain their $\log gf$-values. The resulting line list will be used for all the following operation of chemical abundance measurements.

1.1 Introduction

Despite the largest spectroscopic surveys planned (Gaia) and ongoing (RAVE) cover the near infrared region 8400-8800$\mathring{A}$, this is still a poorly known region: according with the Vienna Atomic Line Database VALD (1999) only 11 Fe lines have accurate gf-value laboratory measurements (Blackwell et al., 1986, Bard et al., 1991, Bard & Kock, 1994), while other hundreds rely to theoretical and semi-empirical calculations (Kurucz, 1995). Some of these values has been proved to be imprecise with errors as large as 1 dex for important lines belonging to Fe and Si (Bigot & Thévenin, 2006). The Van der Waals interaction

(responsable of the line broadening effect) lack of accurate values as well: empirical correction by applying enhancement factors to the Unsöld's approximation[1] (1955) are commonly used. Despite some effort have been done recently to fill this lack with theoretical computation (Barklem et al., 2000), they appear to be insufficient, with computed values of only 50 lines (7 Fe lines) over hundreds present in the 8400-8800$\mathring{A}$ region. A more detailed and systematic investigation of the atomic parameters in this region would be necessary in order to fulfil the requirements of precision of large spectroscopic survey like RAVE and Gaia.

The usual way to check and correct the gf-values is the inverse spectral analysis: a line in the solar spectrum is synthesize (assuming the solar abundances be known), and its $\log gf$ value is changed until the best match with the observed spectrum is reached. The match can be reached by the equivalent width of the lines (among others 1993, Edvardsson et al.), their intensities (Gurtovenko & Kostik 1981, for instance), or by the line profile (among others 1989, Thévenin, Thévenin 1990). Past works focused mainly in the optical region or in the near infrared 0.98-1.57 μm (Borrero et al., 2003), leaving the RAVE/Gaia region unchecked. Recently, a study of this region has been performed by Bigot & Thévenin (2006, hereafter BT06) who have taken into account 30 lines among the strong and isolated Fe I and Si I lines (synthesized with a 3D Radiative Hydrodynamics (RHD) model). They found differences up to 1 dex with the literature values, whereas the lines which have laboratory gf-values were found to be more trustworthy (better than 0.1dex).

In this chapter we describe our work aimed at improving the oscillator strength values for 604 absorption lines observed in the spectra of the Sun and Arcturus over the 8400-8800$\mathring{A}$ wavelength range. In Section 1.2 the preparation of the linelist is described, while Section 1.3 outlines the basic idea that solves the $\log gf$ corrections for isolated and blended lines and the correction routine. Section 1.4 explains the procedure to correct the $\log gf$ values and simultaneously obtain the Arcturus' chemical abundances. Finally Section 1.5 presents the tests performed to estimate our accuracy and Section 1.6 summarizes our results.

1.2   Linelist

In the present work we use MOOG for the spectral synthesis, a standard LTE line analysis and synthesis code (Sneden, 1973). This code, as well as other similar codes, relies on a linelist containing the spectral lines that have to be synthesized. To match the observed spectra, our linelist is composed of lines we identified in the Solar and Arcturus spectra. The resulting linelist contains 604 lines belonging to the following species: N, O, Mg, Al, Si, S, Ca, Ti, Cr, Fe, Co, Ni, Zr and the CN molecule. The existence of these lines has been confirmed by looking at the high S/N Sun and Arcturus spectra (Hinkle et al., 2000,

---

[1]The Unsöld approximation is a formula which simplify the computation of the Van der Waals interactions.

see also Sec. 1.3.4), and identified by cross checking the line identification by Moore et al. (1966) and Hinkle et al. (2000). Some previously unidentified lines have been added in order to match features that are clearly visible on the Sun and Arcturus spectra. The lines have been selected from Kurucz data (Kurucz, 1995) and identified manually. To do so, we took into account the excitation potential (χ) of the transition for the atomic or molecular species already identified in these stars. Nevertheless, many features remain unidentified. The atomic and molecular data are taken from Kurucz 1995: the excitation potential χ are considered of sufficient accuracy for this work, as the wavelengths are known to have precisions of the order of few mÅ (Borrero et al., 2003, BT06). It is commonly accepted that the line broadening due to Van der Waals interaction obtained from Unsöld's approximation is too weak and needs to be corrected by applying an enhancement factor E_γ to the damping constant γ; we adopted the enhancement factor E_γ following Chen et al. (2000).

1.3 Method

1.3.1 Isolated and blended lines

Previously astrophysical gf-value corrections have been performed on identified isolated lines and their equivalen widths (EWs) measured in the Solar spectrum. Isolation is necessary in order to avoid uncontroled contamination by other lines. The line is then synthesized using a Solar model atmosphere where all the atmospheric parameters and abundances are fixed. The $\log gf$ value is then changed until a good match between the observed and synthesized line is reached (see Fig. 1.1). The match can be performed on the line profile or the EW (which are equivalent if the spectra are continuum corrected and normalized) and a reliable $\log gf$ value is given by the curve of growth equation (see Gray, 2005, formula 16.4)

$$\log \frac{EW}{\lambda} = \log C + \log A + \log(gf \cdot \lambda) - \theta\chi - \log \kappa_\nu, \tag{1.1}$$

where C and θ are functions[2] of the temperature T, and κ_ν is the opacity at the frequency ν. This formula relates EW and gf-value in the weak-line approximation. In the case of a weak line affected by Poisson noise or continuum correction, the $\log gf$ value correction

---

[2] C and θ are expressed as follow

$C = constant \cdot \frac{\pi e^2}{mc^2} \frac{N_j/N_E}{u(T)} N_H$
$\theta = 5040/T$

where T is the temperature, N_j/N_E is the fraction of atoms of the j-th stage of ionization with respect to the number of atoms of the element considered, N_H is the number of hydrogen atoms per unit volume and $u(T)$ is the partition function.
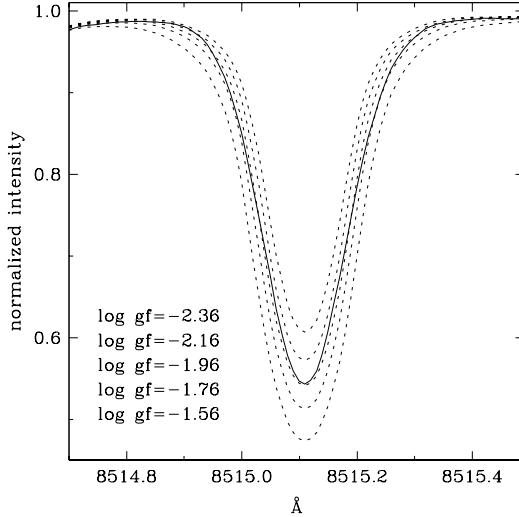
Figure 1.1: The solid line represents the observed line Fe I at λ=8515.110$\mathring{A}$ for the Sun. The dotted lines are synthetic lines with different $\log gf$ values. The best matching line is $\log gf$=–2.02.

would also be affected. This situation can be improved by using a second known spectrum with different atmospheric parameters, where the lines have a larger EW. In this case, the $\log gf$ value is modified until the synthetic line matches both spectra. In order to correct the more problematic blended lines, we exploit the different behaviour the lines have in different physical conditions. For instance, a line will be stronger in an atmosphere with a higher abundance of that species. For lines of high excitation potential (χ), a change in $\log gf$ will lead to a larger change in EW in an atmosphere with higher temperature.

An example is given in Fig. 1.2 where the solid lines represent the spectral feature composed of the two blended lines S I (χ=7.8eV) and Fe I (χ=2.4eV), both located at 8678.927$\mathring{A}$, on the Sun (top) and Arcturus (bottom). The dotted lines are the synthetic spectra with the corrected $\log gf$ values. A change of +0.5 in $\log gf$ value of the SI line (dashed lines) lead to a large change in EW for the Sun whereas the change in Arcturus is negligible. This means that S I is less sensitive to $\log gf$ value changes in Arcturus than in the Sun, i.e. most of the line's EW in Arcturus is due to the Fe I. This difference allows us to find two unique $\log gf$ values which are able to match both spectra simultaneously and resolve the blends.

This idea has mathematical support. When all the atmosphere model parameters are fixed (including abundances), the curve of growth equation (1.1) can be written as a function, f, of $\log gf$ only:

$$EW = f(\log gf) \tag{1.2}$$

The equivalent width of the feature composed of two blended lines $EW_{blend} = EW_{SI} + EW_{FeI}$ is known, whereas the two individual equivalent widths EW_{SI} and EW_{FeI} are in
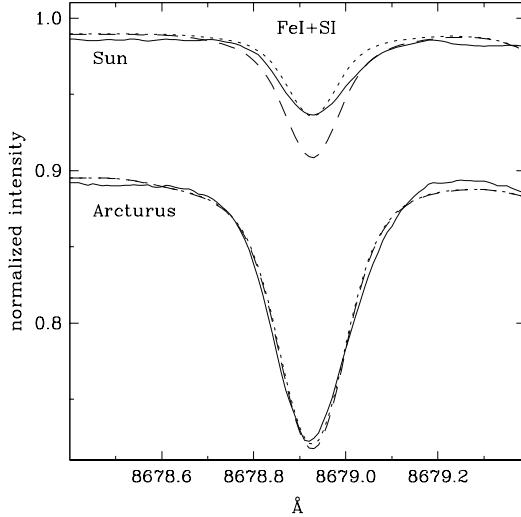
Figure 1.2: Solid lines: observed spectra of the Sun and Arcturus around the feature composed of a S I and Fe I lines at λ=8678.927$\mathring{A}$; Arcturus has been shifted down by 0.1 for clarity. Dotted lines: synthetic spectra with corrected $\log gf$ values. Dashed lines: synthetic spectra with +0.5 added to the S I $\log gf$ value. In Arcturus the change is negligible, i.e. the EW is mainly driven by the Fe I line.

general unknown. Therefore, we can write

$$
\begin{aligned}
EW_{blend}^{Sun} &= f^{Sun}(\log gf_{SI}) + f^{Sun}(\log gf_{FeI}) \\
EW_{blend}^{Arc} &= f^{Arc}(\log gf_{SI}) + f^{Arc}(\log gf_{FeI})
\end{aligned}
\tag{1.3}
$$

where the two equations and their parameters refer to the Sun and Arcturus respectively. The values EW_{blend}^{Sun} and EW_{blend}^{Arc} are known, and are in general different. The two equations represent a solvable system with the unknown $\log gf_{SI}$ and $\log gf_{FeI}$. The existence and uniqueness of the solution is assured even for lines where the weak-line approximation does not hold, thanks to the first derivative of the curve of growth which is always positive. In the case of three blended lines, three different stars must be used to close the system.

1.3.2 The correction routine

The large number of lines in the RAVE/Gaia regime does not allow us to perform the correction by hand. Therefore, we had to resort to an automatic procedure. Here we give a short explanation of the algorithm followed by the correction routine:

1. Synthesize the Solar and Arcturus spectra from our linelist

2. Begin with the first line from the linelist

3. Integrate the absorbed flux of the observed and synthesized spectra over an interval spanning $0.4\mathring{A}$ around the centre of the line (roughly 1 FWHM, which takes into account most of the absorbed flux) and calculate the normalized EW residual (NEWR) defined as follows

$$NEWR = \frac{(flux_{synt} - flux_{obs})}{flux_{obs}} \tag{1.4}$$

for the Sun ($NEWR_{Sun}$) and Arcturus ($NEWR_{Arc}$)

4. If $NEWR_{Sun} < -0.05$ then add $+0.01$ to the $\log gf$, else if $NEWR_{Sun} > 0.05$ then add -0.01 to the $\log gf$ (i.e. match the Sun first). Go to step 7.

5. If $|NEWR_{Sun}| < 0.05$ and $NEWR_{Arc} < -0.05$ then add -0.01 to the $\log gf$, else if $|NEWR_{Sun}| < 0.05$ and $NEWR_{Arc} > 0.05$ then add $+0.01$ to the $\log gf$ (eg. if the Sun matches then try to match Arcturus). Go to step 7.

6. If $|NEWR_{Sun}|$ and $|NEWR_{Arc}|$ are both < 0.05 then a valid $\log gf$ has been reached.

7. Select the next line from the linelist and return to step 3.

The routine is run several times until no more improvement is seen. It tries to match the Sun and Arcturus simultaneously, looking for a compromise between the two spectra to within 5% of the real flux with respect to the Sun. The Sun is our reference point and has a higher weight during the matching process.

For lines with intensities smaller than 0.1 in the Sun we chose a tolerance of 0.3 in NEWR integrating only the three central pixels (i.e. match the intensities), in order to avoid absurdly large results when these lines are blended with unidentified lines or have a bad continuum correction. For some ten lines, a good match could not be reached for several reasons: bad continuum correction, lines located over the Ca II lines where the fit is not good, misidentified line or unrecognized blends. For these lines, we did a visual check and manually corrected the $\log gf$ values. The routine proved able to disentangle moderate blends (i.e. larger than $0.1\mathring{A}$). For closer blends where a good match was not reached, we performed the correction by hand (of course a minimization routine could also be used): the uniqueness of the solution is assured by equation 1.3. The feature centered at $8446.37\mathring{A}$ deserves a special mention, where lines of three different species (O I, Fe I and CN) are blended together. A particular procedure needed to be followed and is explained in Sec. 1.3.5. Zirconium (Zr) also needed a specific treatment, which we explain in Sec. 1.3.6.

The correction is good enough to derive gf-values that yield abundance errors per measured line smaller than 0.2 dex in average for both the Sun and Arcturus. Nevertheless, keep in mind that if a line is in the lower part of its curve of growth on both the Sun and Arcturus (the EW has a small gradient with respect to the $\log gf$ value), its $\log gf$ correction could be poor. In principle, this can be detected in another star where the EWs have a larger gradient.

1.3.3 Multiplets treatment

In the RAVE/Gaia wavelength range there are features that belong to atomic multiplets. An interesting characteristic of these multiplets is that they share the same excitation potential χ and have small differences in wavelength. In this case, obtaining a proper $\log gf$ for the individual lines is problematic. We work around the problem in the following way: for instance, consider the two Al I lines at $\lambda_1 = 8773.896\mathring{A}$ and $\lambda_2 = 8773.897\mathring{A}$ that share the same excitation potential $\chi = 4.021947eV$. The two lines have unknown equivalent widths EW_1 and EW_2 but the sum

$$EW_{Tot} = EW_1 + EW_2, \tag{1.5}$$

is measured and well known. We can write the curve of growth as

$$\frac{EW}{\lambda} = \frac{C \cdot A_{Al} \cdot (gf) \cdot \lambda}{10^{\theta\chi} \cdot \kappa_\nu}, \tag{1.6}$$

where A_{Al} is the aluminium abundance and the other quantities are as in Eq. 1.1. Substituting Eq. 1.6 in Eq. 1.5 and adopting the approximation $\lambda_1 \simeq \lambda_2 = \lambda$ we obtain

$$\frac{EW_{Tot}}{\lambda} = \frac{EW_1 + EW_2}{\lambda} = \frac{C \cdot A_{Al} \cdot (gf_1 + gf_2) \cdot \lambda}{10^{\theta\chi} \cdot \kappa_\nu}. \tag{1.7}$$

Using the logarithmic form, this equation then becomes

$$\log \frac{EW_{Tot}}{\lambda} = \log C + \log A + \log((gf_1 + gf_2) \cdot \lambda) - \theta\chi - \log \kappa_\nu, \tag{1.8}$$

which suggests that an appropriate functional form for a "dummy" variable $\log gf_d$ would be $\log(gf_1 + gf_2)$. Even if this does not allow us to disentangle the two $\log gf$s, we can consider the multiplet as if it were one single line and correct the $\log gf_{fic}$; it will be a "dummy" value but it results in giving the correct abundance of Al. This sort of correction has been applied also to three Mg I multiplets, centred at wavelengths 8717.814, 8736.020 and $8773.896\mathring{A}$.

1.3.4 Observed and synthesized spectra

In order to perform a good correction of the gf-values we need high resolution and high signal-to-noise ratio spectra as references. We chose the Sun and Arcturus spectra by Hinkle et al. (2000) in the region 8400-8800$\mathring{A}$ (instrumental resolution R$\simeq$86000 at 8600$\mathring{A}$, dispersion=0.01 $\mathring{A}$/pix, S/N$\simeq$1000). The spectra are continuum corrected, but after inspection we found that the Sun's continuum was slightly higher (some thousandths of the normalized continuum) than the synthesized one over the wavelength ranges 8400-8489$\mathring{A}$ and 8603-8647$\mathring{A}$. We corrected it by lowering the continuum in these regions by subtracting a linear

spline. We use the atmosphere model by Castelli (http://wwwuser.oat.ts.astro.it/castelli/) for the Sun, which assumes an effective temperature T_{eff}=5777K, gravity $\log g$=4.44, microturbulence V_t=1.0 km/s and metallicity [M/H]=0.00. Arcturus has been synthesized with an atmosphere model obtained by interpolation of the grid ATLAS9 model atmospheres by Castelli & Kurucz (2003). The atmospheric parameters are taken from Earle Luck & Heiter (2005) (T_{eff}=4340K, $\log g$=1.93, V_t=1.87 km/s [M/H]=–0.55). All the syntheses have been performed assuming solar chemical abundances by Grevesse & Sauval (1998) (hereafter GS98) and in the following all the abundances reported will have this reference unless indicated otherwise.

During the work we observed that MOOG could not correctly synthesize the profile of the H I lines and Ca II lines, and could not accurately reproduce the wings of some strong Fe I, Si I and Mg I lines. It is known that such undesirable features come from the unrealistic LTE assumption and 1-D model atmosphere, which require other parameters (like microturbulence) in order to take into account the 3D velocity fields present in stellar atmospheres. Therefore, we dealt with problematic lines as follows:

- *H I lines*: the lines of the Paschen series are weak and wide on the Sun. The $\log gf$ value correction of these lines is not the subject of our work, but we need to correct the blended continuum around the lines because of its effect on the neighbouring lines. Therefore, we modify the $\log gf$ values and the damping constants until a good match with the observed spectrum is reached. In the case of the H I line at 8750.478$\mathring{A}$ we include it twice (with different $\log gf$ values and damping constant) in order to reach a good fit. Of course the values found are "dummy" and cannot be used for any scientific purpose.

- *Ca I lines*: we modify the $\log gf$ values of the three Ca II lines until a good match of the wings is obtained with the Solar spectrum. The cores are neglected. Like the H I lines, those $\log gf$ values are used only to avoid effecting neighbouring lines.

- *Mg I, Si I, Fe I lines*: the strongest lines belonging to these elements show large wings (though not as large as for H I and CaII) which MOOG also fits poorly. For these lines the method described in this chapter is good enough to compute the proper correction. In fact, the match is obtained with the equality of the EWs integrated over an interval of 0.4$\mathring{A}$ centred on the line, whereas the wings are inferred by the adopted line profile. The differences in the wings represents a fraction of the total EW small enough to allow us to reach our goal: an error in abundance smaller than 0.2dex in abundance per measured line.

1.3.5   Oxygen correction

The Oxygen $\log gf$ value correction requires special treatment. We have three O I lines close to each other at wavelengths 8446.247, 8446.359 and 8446.758$\mathring{A}$. The first two lines,

together with a CN and a Fe I line, form one blended feature. To solve this problem we must include one more template spectrum. We chose the Procyon spectrum from Allende Prieto et al. (2004): because of this star's high effective temperature, only the O I lines are visible and their $\log gf$ values can be corrected without interference from the other lines. For these lines, we followed the procedure:

1. correct the $\log gf$ value of the O I line λ=8446.758$\mathring{A}$ until the synthetic line matches the observed one in the Sun (as this line is relatively isolated, its EW is not strongly affected by neighbouring lines)

2. change Arcturus's O I abundance until its synthetic line at λ=8446.758$\mathring{A}$ matches the observed spectrum

3. change Procyon's O I abundance until its synthetic line at λ=8446.758$\mathring{A}$ matches the observed spectrum

4. correct the $\log gf$ values of the O I lines at 8446.359 and 8446.758$\mathring{A}$ until the synthetic one matches Procyon

5. correct the $\log gf$ value of the Fe I and CN lines until the synthetic ones match the Sun and Arcturus (they have no effect on Procyon)

The result is the $\log gf$ value correction and the O I estimation for Procyon and Arcturus, which turns out to be [O/H]=+0.12 for Procyon and [O/H]=+0.17 for Arcturus.

1.3.6   Zirconium correction

Only one Zirconium line belongs to our linelist. It is identified in the Arcturus spectrum, whereas it is nearly invisible in the Solar spectrum: this makes our correction unsuitable for this element. Hence, we adopted the "classical" method: we took 5 isolated Zirconium lines in the optical region with their EWs as reported by Moore (1966) and by using the Sun atmosphere model we corrected the $\log gf$-values until the abundance provided by MOOG matched the solar one. With EWs of these lines hand-measured in Arcturus and the new $\log gf$-values, we obtained Arcturus' Zr abundance, [Zr/H]=-1.0 ± 0.2. This abundance remained fixed during the entire correction procedure, therefore the final Zr $\log gf$-values are dependent on the abundance adopted.

1.4   Procedure

During the $\log gf$ correction, all the atmospheric parameters are fixed for the reasons explained in Sec. 1.3.1. The Sun's parameters are well known and they are considered

correct in this work. The Arcturus parameters are not complete: Earle Luck & Heiter (2005) did not measure the abundances of the elements O, S and Zr which belong to our linelist. Moreover, in the literature there are few elements which have a measured abundance in Arcturus, and they show a significant scatter between different authors (see Tab. 1.1). We resolve this problem by assuming an initial abundance for all the elements and correct the abundances during the procedure. In fact, if we had correct abundances for Arcturus, and the adopted $\log gf$ values match reasonably well with the Sun, the NEWRs expected in Arcturus must have an average close to zero and the scatter would be due to the $\log gf$ values uncertainties. Therefore, a general offset of the NEWRs is an indication of incorrect abundances. The first synthesis of Arcturus is done assuming an initial abundance of –0.55 for all the elements. The algorithm we use during the procedure is the following:

1. synthesize the Sun and Arcturus spectra

2. correct the $\log gf$ values by running the correction routine until no more improvements are possible

3. check the distribution of the NEWRs for each element separately: if a positive (negative) offset is present, decrease (increase) its abundance

4. synthesize the Sun and Arcturus spectra with the new $\log gf$s and abundances

5. return to step 2.

This algorithm is repeated until there are no more improvements. Fig. 1.3 shows the distribution of the NEWRs for 3 elements and the CN molecule at different stages of the procedure. The first column of Fig. 1.3 shows the NEWRs after the first synthesis. The first $\log gf$ correction is applied and the NEWRs distribution at this stage is represented in the second column. The offset of Arcturus's NEWRs indicate that the initial abundances for the three elements [Si/H]=[Ti/H]=[Fe/H]=–0.55 are too low. After some abundances and $\log gf$ corrections, the final NEWRs distribution is reached (third column). It must be noted that in a stellar atmosphere, CN follows a molecular equilibrium that depends on the C and N abundances individually. Because there are no atomic lines of such elements in our linelist to allow us to get a reliable estimation of them separately, we assumed the abundances [C/H]=[N/H]. The final abundances are reported in Tab. 1.1, the final corrected $\log gf$ values are outlined in Tab. 1.2 and the comparison between Kurucz's and corrected $\log gf$ values are plotted in Fig. 1.4.

1.5   Discussion

In order to test the accuracy of our work, we integrated the synthetic and observed flux of the Sun over an interval of $0.4\mathring{A}$ centered on the lines and computed the residuals. We
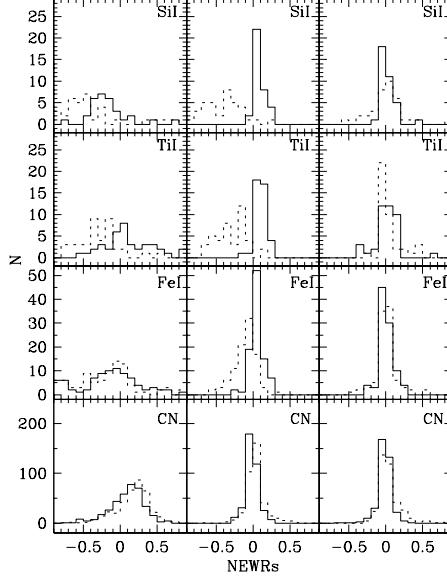
Figure 1.3: Normalized EWs residuals (NEWRs) for the elements Si I, Ti I, Fe I and the molecule CN for the Sun (solid line) and Arcturus (dashed line) during different stages of the procedure. 1st column: the NEWRs distribution using the original Kurucz $\log gf$ values (initial conditions); 2nd column: after the first $\log gf$ correction; 3rd column: after all the abundances and $\log gf$ values correction (final conditions). The bins have a size of 0.1 in NEWR.

used both the original Kurucz $\log gf$ values and the corrected ones. The distributions of the residuals are shown in Fig. 1.5. The small standard deviation of the corrected $\log gf$ values (σ=1.3m$\overset{\circ}{A}$) indicates that the synthetic absorbed flux differs by a few m$\overset{\circ}{A}$ from the observed flux for most of the lines, a clear improvement from those based on the Kurucz $\log gf$ values (σ=7.3m$\overset{\circ}{A}$). To test our internal consistency, we measured by hand the EWs of 100 atomic lines. Fig. 1.6 presents the difference in abundances obtained by using the Kurucz gf-values (open points) and the corrected ones (solid points). Our correction reduced the standard deviations of the abundance residuals from σ=0.63 dex to σ=0.15 dex and their averages from +0.16 to -0.05.

In Fig.1.7 we compare the $\log gf$ values by BT06 for Fe I and Si I lines (this work assume Asplund et al. 2005 solar abundances, which are 0.05 dex and 0.04 dex lower with respect to GS89 for Fe and Si respectively) to our corrected values, the corrected $\log gf$ values from Kurucz (1995), and the laboratory $\log gf$ values from Bard et al. (1991) and Bard & Kock (1994). There is good general agreement with the laboratory measurements and the 3D RHD simulations by BT06 with a small offset of about +0.1dex with respect to the latter. This result indicates that our correction is a significant improvement on the Kurucz values.

Nevertheless, the corrected $\log gf$ values are the result of a procedure based on several assumptions and approximations: the EWs of the lines are measured using only their
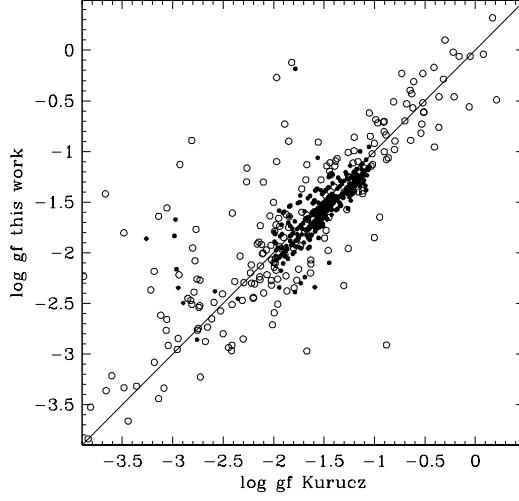
Figure 1.4: Corrected $\log gf$ values by this work compared with the original Kurucz values. Open points: atomic lines. Solid points: molecular lines.

abundance	this work	ELH	T98	F07
[O/H]	+0.17		−0.30	−0.10
[Mg/H]	−0.03	+0.02	−0.45	−0.10
[Al/H]	−0.27	−0.28	−0.40	−0.17
[Si/H]	−0.15	−0.14	−0.30	−0.15
[S/H]	−0.13			
[Ca/H]	−0.26	−0.56	−0.20	−0.28
[Ti/H]	−0.22	−0.39	−0.20	−0.20
[Cr/H]	−0.37	−0.55	−0.20	
[Fe/H]	−0.52	−0.55	−0.40	−0.54
[Co/H]	−0.17	−0.36	−0.20	
[Ni/H]	−0.45	−0.48	−0.35	
[Zr/H]	−1.00			

Table 1.1: Arcturus' abundances compared with the results by Earle Luck & Heiter (2005), Thévenin (1998) (T98), Fulbright et al. (2007) (F07)

central region ($\simeq$2 FWHM) and inferring the wings by using a line profile. The poorer the line profile fit, the poorer the inferred EW. Also, we used a LTE code and a 1D atmosphere model, an approximation that we know to be unrealistic and therefore likely to increase the intrinsic errors. Literature values for the Sun and Arcturus atmosphere parameters have known uncertainties; some lines or blends might be misidentified or not recognized. Therefore the "corrected" $\log gf$ values are dependent on all these uncertainties as well as on numerical (software-dependent) errors we used to obtain them. The same considerations have to be applied to the $\log gf$ by Kirby et al. (2008) who corrected the

wavelength $\mathring{A}$	atomic specie	χ (eV)	$\log gf$ Kurucz	$\log gf$ this work
...	...	...	...	...
8632.412	26.0	4.103650	−1.95	−2.55
8632.448	607.0	0.933757	−1.69	−1.49
8633.124	16.0	8.409819	−0.06	−0.56
8633.933	20.0	4.450947	−0.81	−0.51
8634.128	607.0	1.178822	−1.47	−1.24
...	...	...	...	...

Table 1.2: Linelist with our final corrected $\log gf$ values. By column: wavelength, atomic species, excitation energy (χ), $\log gf$ values by Kurucz (1995), $\log gf$ value corrected by this work. Only 5 lines are showed, and the full table is available in electronic form.
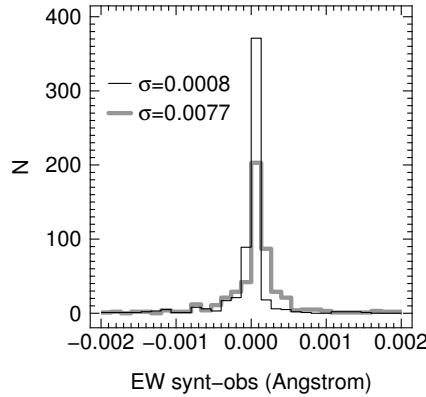
Figure 1.5: Distribution of the residuals between synthetic and observed EWs measured for the Sun, over an interval of $0.002\mathring{A}$ centred over all the lines. The gray line represents the distribution of the residuals between observed and synthetic Solar spectra by using the Kurucz $\log gf$ values; the black line is the same distribution with the corrected $\log gf$ values.

gf-values by applying a similar method, i.e., by synthesizing the Sun and Arcturus and comparing the synthetic spectra with the observed ones. A comparision between our and their results is showed in Fig. 1.8 for the 89 corrected lines in common (belonging to all the elements but O): in average our $\log gf$ are lower of −0.05dex, very likely due to different solar abundances adopted: Kirby adopt the Anders & Grevesse (1989) abundances which are in average higher than Grevesse & Sauval (1998) adopted by us. However, from the consistency tests we performed, we are confident our computed $\log gf$ values are accurate and will lead to robust and reliable estimated abundances.
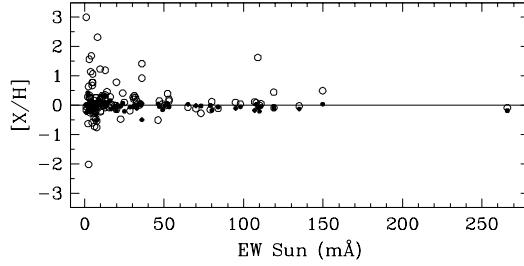
Figure 1.6: Solar abundances obtained from 100 atomic lines for which EWs (x-axis) have been measured by hand from the Solar spectrum. The open points represent the abundances obtained by using the Kurucz gf-values, whereas the corrected gf-values are represented by the solid points. The respective dispersions are σ=0.63dex (open points) and σ=0.15dex (solid points).

1.6   Conclusions

We performed an astrophysical correction of the oscillator strengths of 604 atomic and molecular lines in the 8400-8800$\mathring{A}$ wavelength region by using a LTE synthesis code and 1D atmosphere models. These corrections are made based on simultaneously matching the synthesized spectra to more than one star. We use the different physical conditions present in the different stellar atmospheres to perform a better correction on Solar weak lines and to disentangle blended lines, relying upon their different curves of growth. The $\log gf$ value correction has been performed using the Grevesse & Sauval (1998). The abundance analysis performed with the corrected $\log gf$-values lead to uncertainties not larger than 0.2dex.
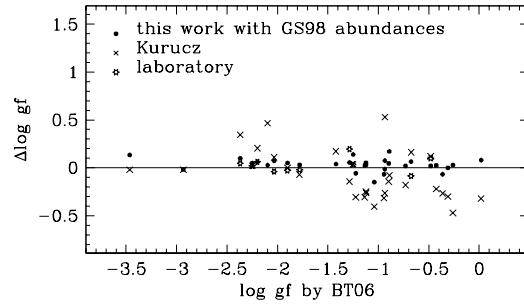
Figure 1.7: Differences between $\log gf$ values by BT06 and the Kurucz's ones, the ones obtained by this work, and the laboratory values available in the literature for 22 Fe I and 10 Si I lines.
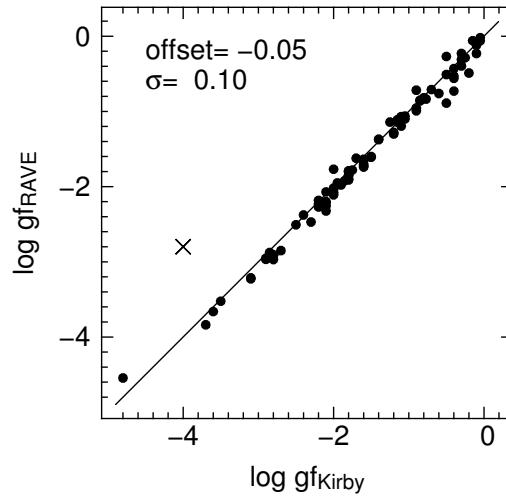

Figure 1.8: Comparision between corrected $\log gf$s by Kirby et al. 2008 (x-axis) and RAVE (y-axis). The statistics on the panel refers to points after the crossed outlier has been removed.

Chapter 2

RAVE processing pipeline: improving the stellar parameters estimation

The RAVE pipeline is the main code used to analyze the RAVE spectra. It is responsible of the automatic spectra extraction, radial velocity and stellar parameters (T_{eff}, $\log g$, [m/H]) measurements (Zwitter et al., 2008). During the development of the chemical pipeline and the analysis of its results, some systematic errors on stellar parameters came to light. In the following we outline the systematic errors found and analyse their origin. Some of them have been successfully removed by modifying part of the RAVE pipeline. Others cannot be removed but are known and under control.

2.1 Short introduction to the RAVE pipeline

We give here a short explanation on how the pipeline works, focusing on the issues that caused the systematic errors. Full details of the pipeline are outlined in the RAVE second data release paper (DR2, Zwitter et al. 2008).
After the extraction and wavelength calibration of the spectra, the RAVE pipeline performs a continuum normalization on them by using the IRAF's task *continuum* with an asymmetric rejection of pixels (low rejection=1.5σ and high rejection=3.0σ) so as to consider the absorption lines present under the continuum level. From this stage onward the pipeline uses the wavelength interval $8499< \lambda(\text{Å})<8746$ of the spectrum (788 pixels). The radial velocity (RV) is measured by using synthetic spectra of the Munari's synthetic spectra library (Munari et al., 2005) as templates at RV=0. The spectra are radial velocity corrected to RV=0 and passed to the subroutine which performs the stellar parameters estimation. As before, the stellar parameters estimation relies on the Munari's library: this is a vast grid of synthetic spectra with 6 different variables: effective temperature T_{eff}, gravity $\log g$, metallicity [m/H], abundance in α-elements [α/Fe], rotational velocity V_{rot} and microturbulence ξ. The pipeline estimates these parameters by using a penalized χ^2
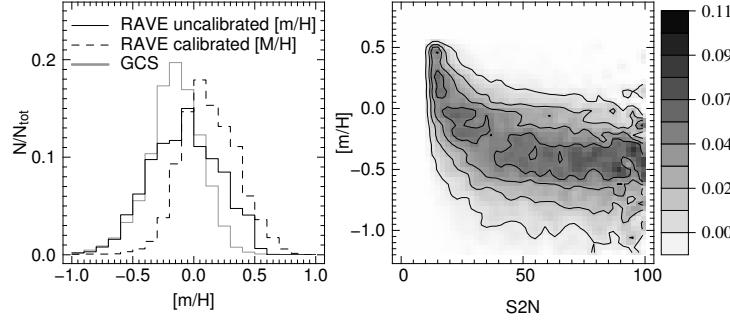
Figure 2.1: Left: distributions of the uncalibrated and calibrated metallicity for 9848 RAVE dwarf stars of the DR2 data release having signal-to-noise S2N>20 (black line) and 15077 dwarf stars of the Geneva-Copenhagen survey (gray line). The dashed line represent the metallicity distribution obtained when the calibrated metallicity [M/H] is used. Right: distribution of 22375 RAVE stars of the DR2 data release on the plane (S2N,[m/H]). The gray levels represent the density of the stars normalized for the number of stars at the S2N bin they belong to.

technique to construct a synthetic spectrum matching the observed one. The synthetic spectrum is constructed as a weighted sum of a sample of template spectra with known parameters and it is assumed that the stellar parameters follow the same weight relation. The sample of template spectra consists on 300 synthetic spectra of the Munari's library having the lowest χ^2. We call χ_{300} the highest χ^2 value of the sample (it corresponds to χ_{lim} in Zwitter et al.); in the next sections we will see that this limit plays a role during the stellar parameters estimation.

The performance of the pipeline was checked on a sample of standard stars having high precision measurements from high resolution and high S/N spectroscopic observations (Figure 2.11). Part of the sample (black points) comprised of 164 stars observed by J.Fulbright at Mc Donald observatory. The remaining 104 stars are from the Soubiran&Girard (2005) compilation of abundance studies. This test highlights systematic errors in [m/H] (too low) and T_{eff} (too low for dwarf stars with $T_{\mathrm{eff}} > 6000$K).

The systematic in [m/H] was already discussed in the DR2 paper and a calibrated metallicity ([M/H]) was proposed in order to correct it (Formula 20 in the DR2 paper). This formula applied to the data generates an odd metallicity distribution skewed to high [m/H] (the dashed line in Figure 2.1, left panel). This is due to a systematic error that was not known at that time: the correlation between [m/H] and S/N (see Figure 2.1 right panel), where the lower the S/N the higher the [m/H]. Therefore, the Formula 20 in the DR2 paper is biased as it was computed using stars with high S/N spectra. For this reason in the following we will not consider the calibrated metallicity [M/H] but we will refer only to the uncalibrated metallicity [m/H] as given by the pipeline.

Before we move on, we wish to place a small note about the variables S/N, SNR, S2N and STN which all indicate signal-to-noise ratios. RAVE has had more than one way to estimate the signal-to-noise ratio: the first SNR was introduced in the first data release
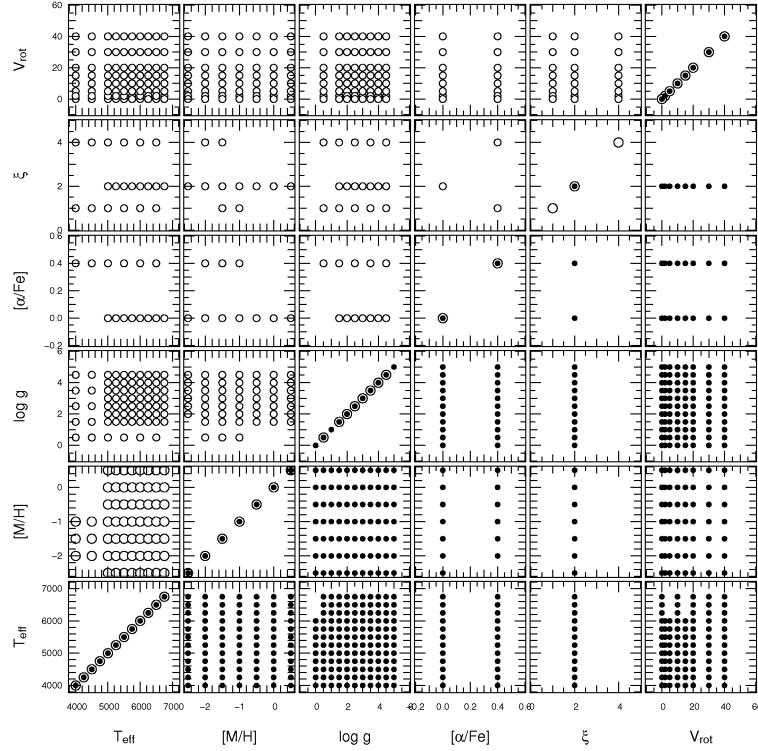
Figure 2.2: Stellar parameters covered by the grid of synthetic spectra adopted by RAVE (Munari et al., 2005). Solid points denote the synthetic spectra computed with the new ODF opacities. Open points denote the old ODF opacities

(DR1, Steinmetz et al., 2006) and replaced by the more precise S2N in the DR2 (Zwitter et al., 2008). In Sec. 2.6.1 we introduce the new STN which is consistent with the S2N. Every time we refer to no specific signal-to-noise ratio but only at its meaning we will call it S/N.

2.2 On the completeness of the synthetic library

The RAVE pipeline adopts the library of synthetic spectra by Munari et al. (2005). It has been synthesized by using atmosphere models based on the new opacity distribution function (ODF) by Castelli & Kurucz (2003) and the old ODF models (1997). They cover different range of parameters and partially overlap (Figure 2.2). RAVE adopted the new ODF models and where they do not cover the desired stellar parameters the old ODF model are used. The resulting grid is not uniform: when the pipeline considers microturbulence ξ=2 km/sec it uses the new ODF models with T_{eff} range 3500<T_{eff} (K)<10000 which has fairly uniform grid in the other parameters. When it considers ξ=1 and 4 km/sec it uses

the old ODF models in the range 3500<T_{eff} (K)<50000 but having only [α/Fe]=0.4dex
and [m/H] <−1.0dex (see Figure 2.2). We have found that these irregularities generate
systematic errors because the χ^2 hypersurface contour (at χ_{300}) is biased by the irregular
coverage of the grid. For this reason we re-defined the grid by dropping the old ODF
models and adopting only the new ODF models for 3500<T_{eff} (K)<10000, which have a
more regular grid coverage and reduce the systematics for most of the spectra. In fact,
98.5% of the RAVE stars have temperature in this T_{eff} range. For the remaining 1.5% of
the stars having T_{eff} >10000K we adopt the old ODF models which cover the temperature
range 10000<T_{eff} (K)<50000.

2.3   Systematic errors due to continuum normalization.

The main systematic error affecting the RAVE data is visible in the shape of the calibrated
metallicity [M/H] distribution of dwarf stars. The left panel of Figure 2.1 shows that the
[M/H] distribution (dashed line) is in average too metal rich with respect to the dwarf
stars' distribution of the Geneva-Copenhagen survey (2004) and its shape is asymmetric
with a long tail on the high [M/H] side while there are few or no stars with [m/H] <-0.6dex.
As previously stated, this was caused mainly to the calibration formula given in the paper
DR2. If we instead look at the distribution of the uncalibrated metallicity [m/H] we see
the expected long tail toward the low [m/H] but still too many metal rich stars. This
excess number of stars is due to the (unwanted and unexpected) correlation between S2N
and [m/H] shown in the right panel of Figure 2.1: the lower the S2N the higher the average
[m/H]. In the next sections we identify the continuum normalization as the main cause of
this correlation and propose solutions.

2.3.1   Correlation between [m/H] and S/N

For S/N<10 noise dominates the spectra and the majority of the metallic lines are no
longer visible. In this case the [m/H] measurement diverges to the highest allowed [m/H]
(+0.5dex) as this gives the lowest χ^2. Since results for spectra with such a low S/N are
not released (as they are not reliable) this case will not be further considered.
For spectra with S/N>10 we see in Figure 2.1 a steep and continuous descent of the average
metallicity which flattens at high S/N. A detailed analysis of the data revealed that this
trend is due to the continuum normalization of the spectra which is correlated with the
S/N. By plotting the averages of the residuals between observed and template spectra as
function of S/N (Figure 2.3, top left panel) one can highlight the correlation between them,
i.e., the continuum is estimated as too high for low S/N but is well estimated for S/N>80.
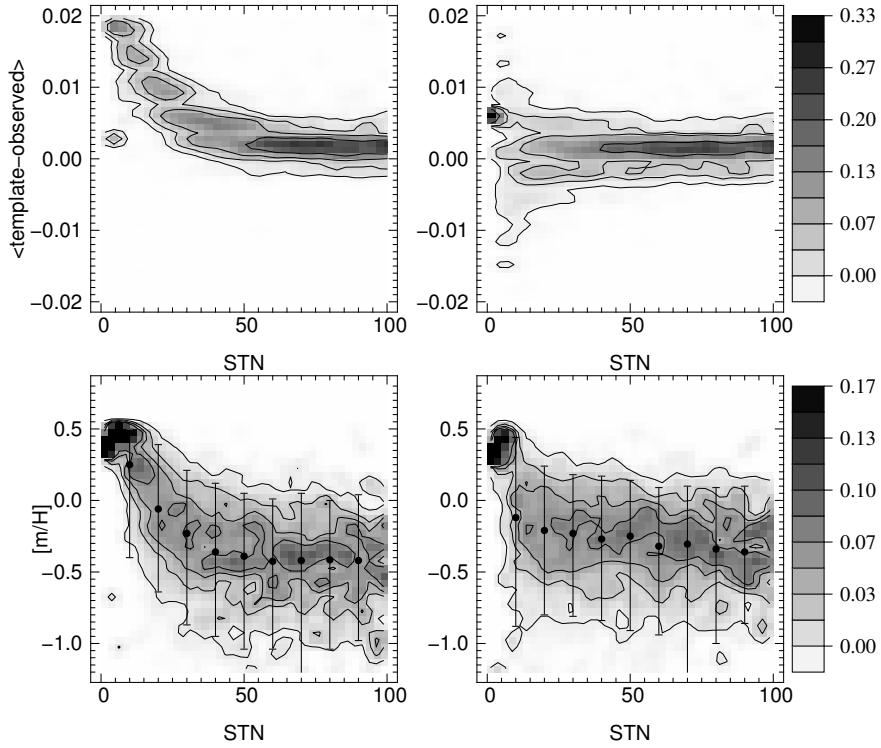
Figure 2.3: Top: distributions of the average residuals (template-minus-observed) for 4684 RAVE stars as function of STN before (left) and after (right) the application of equation 2.1. Bottom: correlation between [m/H] and S/N of the same sample before (left) and after (right) the application of equation 2.1. Black point represent the median and the bars the 90th percentile.

The cause of this correlation is the *continuum* IRAF command used by the pipeline to normalize the spectrum, which uses an asymmetric rejection of the pixels (low rejection=1.5σ and high rejection=3.0σ) under the assumption that the spikes under the continuum are mostly absorption lines and the ones above are noise. This is true at high S/N whereas at low S/N the spikes are mostly noise. Therefore, at low S/N this IRAF command generates a too high estimation of the continuum and causes a too high [m/H] estimation by the RAVE pipeline.

We worked around the problem by using a variable low rejection value having an exponential form

$$low_{rej} = 1.5 + 0.2 \exp(-\frac{STN^2}{2 \cdot 16^2}) \tag{2.1}$$

where STN is the signal-to-noise ratio as computed by the code explained in section 2.6.1 and the constants have been determined through a try-and-error process until the residuals between the observed and template spectra showed no longer correlation (Figure 2.3, top right panel). With this correction the continuum is properly estimated at low S/N (Figure 2.3, top right panel) and the correlation between S/N and [m/H] is strongly reduced (Figure 2.3 bottom).

2.3.2   Continuum errors due to weight normalization

To estimate the stellar parameters, the pipeline searches for the minimum χ^2 (the best match) between the observed spectrum and a synthetic spectrum (model) with known stellar parameters. The model is constructed as weighted sum of 300 templates with weights ω_i following the constrain $\sum \omega_i = 1$. We found that a significant fraction (25%) of the estimations had $\sum \omega_i < 0.997$, leading to wrong parameters: [m/H], for instance, can be underestimated by up to $\simeq 0.3$dex. To correct for this systematic we put more stringent constraints to the minimization process by changing the Lagrange multiplicators involved in the minimum search. With this change we have $0.9999 < \sum \omega_i < 1.0001$ for 97% of the spectra.

2.4   Systematic errors due to degeneracy

To estimate the stellar parameters of an observed spectrum the pipeline computes the χ^2 between it and all the templates of the synthetic library (which can be seen as a χ^2 hypersurface in the parameters space). It is worth noting that the results depend on the shape of this hypersurface and that the shape depends on the quantity of information held by the observed and synthetic spectra. If the observed spectrum is very noisy, has a very small window in wavelength or if in this window there are no usable lines then the χ^2 surface will have a large and flat minimum as in this case all the templates look very similar to each other. With this degeneracy the determination of the stellar parameters can become very uncertain. This means that the quality of the results depends on the quantity of information held in the chosen wavelength range as well as the S/N of the observed spectrum. The RAVE wavelength range generates a valley in the χ^2 surface which is wider at low metallicity or low S/N. The pipeline constructs the best matching template by averaging a sample of 300 templates having the lower χ^2. Most (or all) of them lie in the valley (degeneracy area) where we also expect to find the true parameters we are looking for. If the valley has a symmetric shape then the estimated and the true parameters will be (in average) close to each other and to the center of the valley where the minimum is. But when the shape is asymmetric then the estimated parameter will lie close to the centre of the valley, but the true parameters together with the minimum will be to the side. In principle this would not happen if the weighted average was done by using all the templates in the parameter space, but in order to avoid a too long computing time we must select a limited number (300) of templates and these give the systematic offset. In Figure 2.4 we analyse 4 slices of the χ^2 surface computed for a spectrum with S/N=100 and parameters T_{eff} =6250K, $\log g$ =4.5, [m/H] =–1.0dex. The circle represent the templates with $\chi^2 < \chi^2_{300}$ selected by the pipeline. The plus symbol represents the
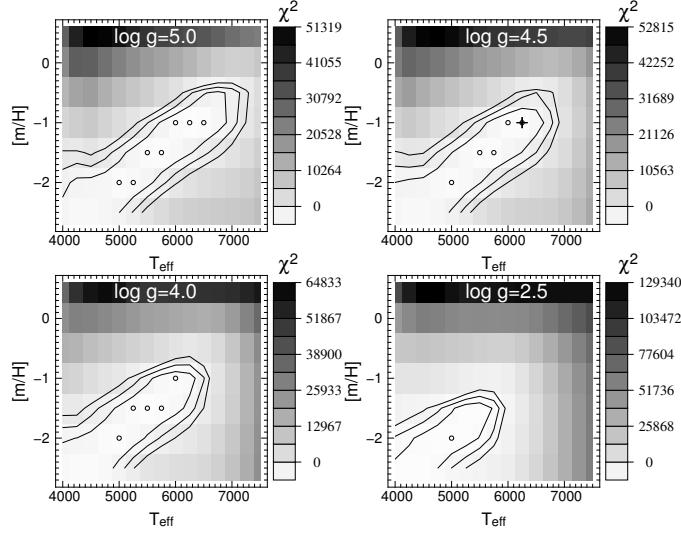
Figure 2.4: Four slices of the χ^2 hypersurface at $\log g$ =5.0, 4.5, 4.0, 2.5 for a spectrum with $T_{\rm eff}$ =6250K, $\log g$ =4.5, [m/H] =-1.0dex with wavelength range 8499–8746Å. The gray level represent the χ^2 value. The black isocountour have $\chi^2 = 2\chi^2_{300}, 3\chi^2_{300}, 4\chi^2_{300}$; the circles represent the parameters of the templates with $\chi^2 \leq \chi^2_{300}$ used to estimate the stellar parameters (many of them overlap with each other since they differ in other parameters like V_{rot} and $[\alpha/Fe]$). The plus symbol represent the position of the true parameters.
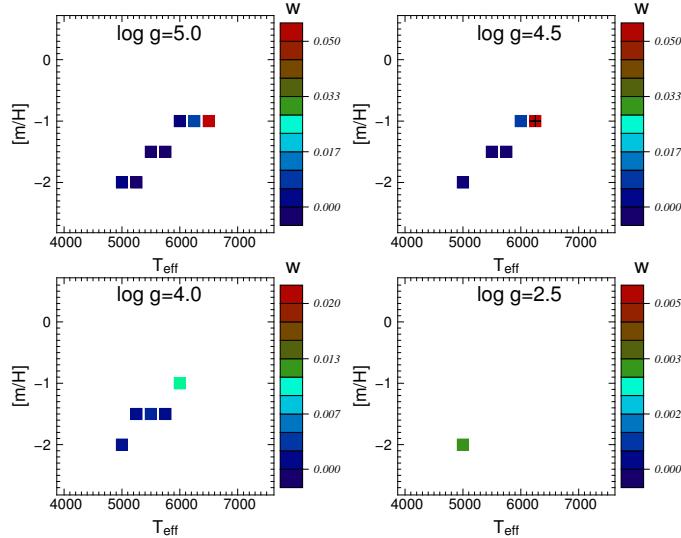
Figure 2.5: As in Figure 2.4, the squares represent the parameters of the templates with $\chi^2 \leq \chi^2_{300}$ having $[\alpha/H]$=0.0 dex and V_{rot}=30 Km/s. The colours codes their weights as showed by the coloured bars. The RAVE pipeline gives the highest weight to the true parameters template, marked with a plus symbol in the top-right panel.
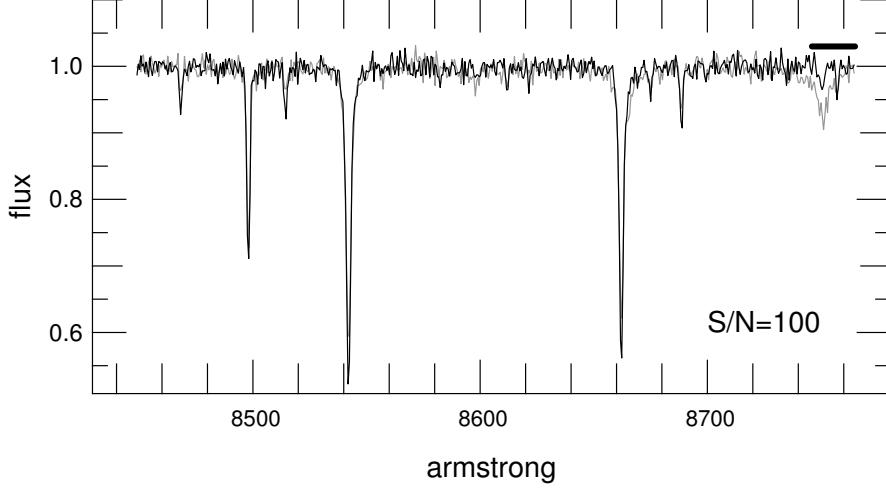
Figure 2.6: At S/N=100 the template with T_{eff} =6250K, $\log g$ =4.5, [m/H]=-1.0dex (gray line) cannot be easily distinguished from the template with T_{eff} =4750K, $\log g$ =2.0, [m/H]=-2.0dex (black line) in the RAVE wavelength range. By extending the wavelength range to include the Paschen line at 8750Å(overlined by the thick black line) the two spectra become different enough to be distinguished by a χ^2 analysis.

position of the true value. Note that it lies on the minimum value of the χ^2 and it has also the higher weight (Figure 2.5). Nonetheless, because the asymmetric shape of the χ^2 surface, the true value lie on the border of the selected area and the estimated parameters are shifted to the centre (T_{eff} =5937K $\log g$ =4.2 [m/H] =–1.40). For such cases all the parameters are affected by systematics.

2.4.1 T_{eff} underestimation for T_{eff} >6000K

This problem affects mainly dwarf stars and it is highlighted in Figure 2.7 where we compare standard dwarf stars of several sources with their expected T_{eff}: for T_{eff} >6000K the RAVE temperature appears underestimated with respect to the reference. In the following we give a possible explanation of this error.

At high temperature the lines of many elements (Fe in particular) become weak, extending the degeneracy area in the parameters space. For example, in Figure 2.6 we compare the synthetic spectra (degraded to S/N=100) of a hot metal poor dwarf star and a cool metal poor giant star: in the RAVE wavelength range the two spectra look very similar even if they are located far from each other in the parameter space. This generates a long degeneracy area where all the templates lying in between have a low χ^2 and cause a too low [m/H] and T_{eff} estimation. Figure 2.6 suggest a possible solution: by extending the wavelength range up to 8765Å to include Paschen line H I at 8750Å we can distinguish the two cases. This reduces the degeneracy area and [m/H] and T_{eff} underestimations are
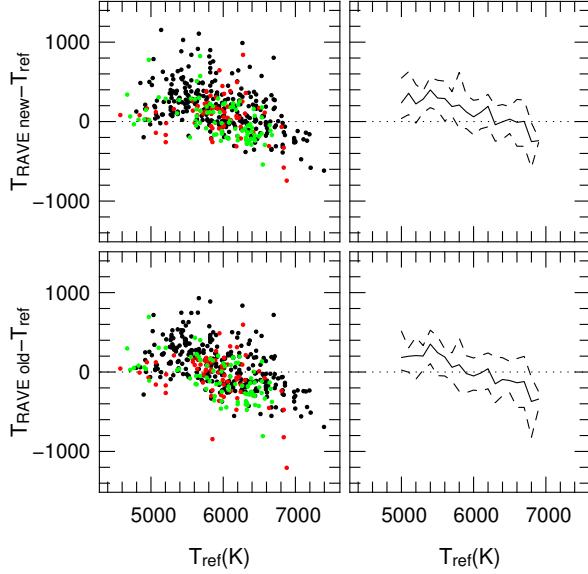
Figure 2.7: **Right panels**: residuals between T_{eff} obtained with the old (bottom) and new (top) RAVE pipeline and reference T_{eff} for 280 GCS stars (black points), 70 stars observed by J. Fulbright (red points) and 86 SG05 stars (green points). **Left panels**: average (solid line) and standard deviation (dashed lines) of the distributions in the right panels.

reduced as well. Unfortunately with the pipeline architecture this correction is applied to all the stars of any T_{eff} , also to giant stars whose T_{eff} estimation was good previously. Consequent tests showed that with the extended wavelength range the pipeline introduces offsets in T_{eff} (+150K) and $\log g$ (+0.2dex) and an unnatural overdensity of stars for stellar parameters around $\log g \simeq 4.9$ and $T_{\text{eff}} \simeq 6500$K. Between the two systematics the RAVE collaboration decided that was better to have the one having the smallest impact in the general statistics and kept the original wavelength range. The systematic error that the extended wavelength range would introduce is discussed in Sec. 2.8.

2.4.2 [m/H] underestimation

The circles in Figure 2.4 show that for a test spectrum with T_{eff}=6250K, $\log g$=4.5 and [m/H]=−1.0dex, none of the templates with [m/H]>−1.0dex belong to the 300 templates used for the weighted average because the grid points at [m/H]≥−0.5dex have $\chi^2 > \chi^2_{300}$. Figure 2.8 shows that χ^2 increases faster[1] at higher [m/H] than lower, therefore the high [m/H] templates are less likely to belong to the sample of 300 templates used for the

---

[1]Since χ^2 measures the difference of absorbed flux between observed and templates spectra it increases roughly exponentially because the equivalent width of the lines (absorbed flux) follow an exponential law (in the first part of their curve-of-growth) as function of the metallicity.
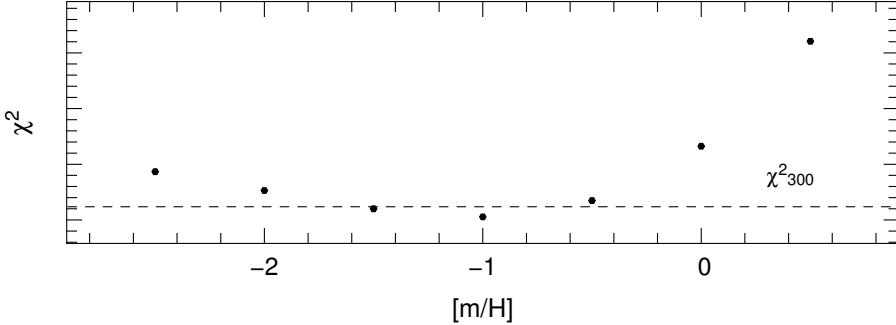
Figure 2.8: Values of the χ^2 of the grid points for T_{eff} =6250K and $\log g$ =4.5. The dashed line represents the χ^2_{300} value.

weighted average. This favours lower [m/H] templates because the sample of templates chosen follow the asymmetric shape of the degeneracy area. To reduce this asymmetry we would need more grid points on the high [m/H] side. For instance, if the synthetic library had templates at [m/H]=–0.9 dex they would have $\chi^2 < \chi^2_{300}$ (see Figure 2.8) and the selection of the 300 templates would be less biased. Here the problem is caused by the step of the grid which is linear in [m/H] whereas χ^2 is roughly exponential as function of [m/H]. Another effect caused by the adopted grid is the overdensity of stars at [m/H]=– 0.05 dex and $\log g$ =4.2 (see Figure 2.13) where stars with metallicity slightly higher than [m/H]=0.0 dex lack templates with [m/H]=0.5 dex among the ones having $\chi^2 < \chi^2_{300}$; these stars are therefore stuck at [m/H]<0.0 dex.

The metallicity steps should gives a (roughly) linear χ^2 through the grid, i.e. the grid steps should be finer at higher metallicity. Therefore, we added new grid points at higher metallicity, effectively raising the probability that higher metallicities belong to the templates sample. We produced new grid points at [m/H]=–0.8, –0.6, –0.4, –0.2, +0.2, +0.4 dex through polynomial interpolation of the existing grid and added them to the library used for the stellar parameters estimations.

2.5   Rotational velocity V_{rot}

Due to the RAVE instrumental resolution, a sharp absorption line in a RAVE spectrum has a FWHM∼1.4Å, which correspond to ∼30 km/sec of rotational velocity V_{rot} in the Munari's spectral library. This means that stars with rotation between zero and 30 km/sec will result to have V_{rot}=30 km/sec. Thus, templates with such rotational velocities are redundant in the grid used by the pipeline. For this resons we decided to drop the rotational velocities V_{rot}=0, 2, 5, 15, 20, 40 km/sec and consider only the templates with V_{rot}=10, 30, 50, 75, 100, 150, 200, 250, 300, 400, 500 km/sec for the new grid.

2.6 New codes implemented in the pipeline

In order to perform the new continuum correction (as described in Sec. 2.3.1) we wrote a new piece of code which gives a new S/N estimation before the continuum correction takes place. This S/N estimation is called STN. To give information on the quality of the analyzed spectrum we also implemented a routine which gives the fraction of the spectrum affected by cosmic rays, fringing and other continuum defects. We explain both codes below.

2.6.1 STN estimation

Initially, the pipeline made use of the signal-to-noise ratio SNR described in Steinmetz et al. (2006), which underestimates the real S/N. Therefore a new signal-to-noise ratio S2N was introduced (see DR2 paper, Zwitter et al. 2008). S2N is obtained after the stellar parameters estimation, by using the residuals between observed and template spectra. However, this means that it cannot be used by the new continuuum correction code (outlined in section 2.3.1) which needs the S/N *before* the parameters estimation. Thus, we developed an algorithm to measure the signal-to noise S/N of a spectrum in which no flux information is used. This is necessary for RAVE spectra as the counts are not normalized through spectrophotometry. The algorithm is as follows:

1. Create a fitting line $f(i)$ of the observed spectrum $s(i)$ with N number of pixels by using a smoothing box car 3 pixels wide (the fluxes of the pixels in the box are averaged)

2. Compute the residuals $R_N(i) = f(i) - s(i)$. σ_N is then the standard deviation of $R_N(i)$

3. Construct the local standard deviation vector $\sigma_l(i) = \sigma_r(i)$ where $\sigma_r(i)$ is the standard deviation of the pixels $r(i-1)$, $r(i)$, $r(i+1)$ of the R_N sample

4. Exclude the pixels having residual larger than $2 \cdot \sigma_N$

5. Repeat the last step. The sample of pixels left is called R_n and counts n pixels ($n < N$). It has standard deviation σ_n and represents what we call *fine noise*

6. The signal-to-noise is then $STN = median(s(i)/\sigma_l(i))/1.62$

This procedure removes those features of the spectrum that are due to absorption lines or consecutive pixels having positive or negative residuals because of noise fluctuations. Tests with generated synthetic noise showed that the standard deviation of the fine noise σ_n is proportional to the standard deviation of the noise σ_N with a ratio $\frac{\sigma_N}{\sigma_n} = 1.62$. Expressing the signal-to-noise ratios as S/N=$\frac{1}{\sigma_N}$ and STN=$\frac{1}{1.62 \cdot \sigma_n}$ they follow a 1:1 relation (Figure 2.9
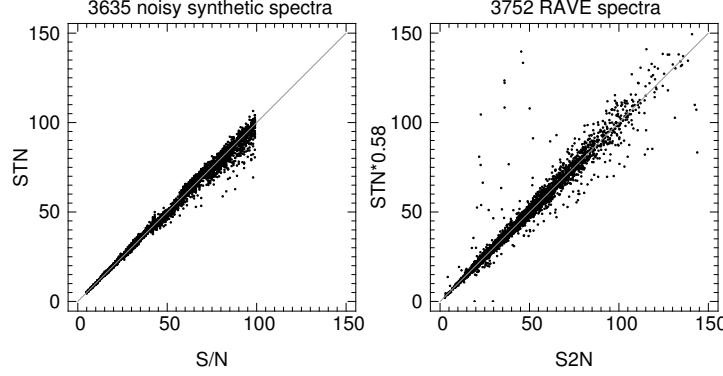
Figure 2.9: Left panel: comparison between the estimated STN and the input synthetic S/N for 3635 noisy synthetic spectra. Right panel: comparison between the estimated STN (multiplied by 0.58) and S2N.

left panel). For real noise in the RAVE spectra, STN is proportional to the signal-to-noise estimation S2N (as defined in Zwitter et al., 2008) with the following ratio $\frac{S2N}{STN} = 0.58$ (Figure 2.9, right panel). The difference is due to the different sources of errors affecting the real spectra with respect to the synthetic ones like flat fielding errors, light ghosts and read-out-noise of the CCD.

The method proved to be robust and gives a signal-to-noise estimation with an error of $\sim 10\%$.

2.6.2   Spectrum quality determination: the code MASK

Roughtly $\simeq 20\%$ of the RAVE spectra have part of the spectrum (usually smaller than 50Å) affected by defects like fringing, cosmics rays etc. which cannot be fully removed by continuum correction. In order to determine the location and extension of the defect we wrote an algorithm which provides this information. Given the flux residual between the observed and template spectrum for the i-th pixel

$$r(i) = flux_{template}(i) - flux_{obs}(i) \tag{2.2}$$

the following steps outline the algorithm:

1. consider the interval $I_j = [j - 10, j + 10]$ centered on the pixel j. The following

quantities are computed

$$\tilde{\chi}^2(j) = \frac{1}{\max(I_j) - \min(I_j)} \cdot \sum_{i=\min(I_j)}^{\max(I_j)} (\frac{r(i)}{\sigma})^2$$

$$\psi(j) = \frac{1}{\max(I_j) - \min(I_j)} \sum_{i=\min(I_j)}^{\max(I_j)} r(i)$$

where $\tilde{\chi}^2(j)$ is the reduced χ^2 and $\psi(j)$ is the estimation of the area between the observed and reconstructed spectra. σ is computed as the inverse of the signal-to-noise ratio.

2. if $\tilde{\chi}^2(j) > 2$ and $\psi(j) > 2 \cdot \sigma$ then the pixels in the interval I_J are labelled as 0, otherwise 1

The process is repeated for all the pixels of the spectrum, obtaining an array of 0 and 1 (the "mask"), each of them referring to the wavelength intervals represented by the pixels. Figure 2.10 displays the result where the pixels labelled with 0 are represented as gray points. The code generates the mask for all spectra and the fraction of each spectrum affected by continuum distortion. We suggest excluding those results for spectra having more than 30% of the pixels labelled with "1".

2.7 The new modified pipeline

With the changes outlined above the new modified pipeline has new characteristics. The grid of synthetic spectra now has

- templates with new ODF models for $3500 < T_{\text{eff}}$ (K) < 10000 and old ODF models for $10000 < T_{\text{eff}}$ (K) < 50000

- only templates with rotational velocities $V_{\text{rot}} = 10, 30, 50, 75, 100, 150, 200, 250, 300, 400, 500$ km/sec

- new points in [m/H] with [m/H]= –2.5, –2.0, –1.5, –1.0, –0.8, –0.6, –0.4, –0.2, 0.0, 0.2, 0.4, 0.5 dex

All the other parameters steps are the same as the original grid. The new pipeline has the following new features:

- new signal-to-noise (STN) estimation
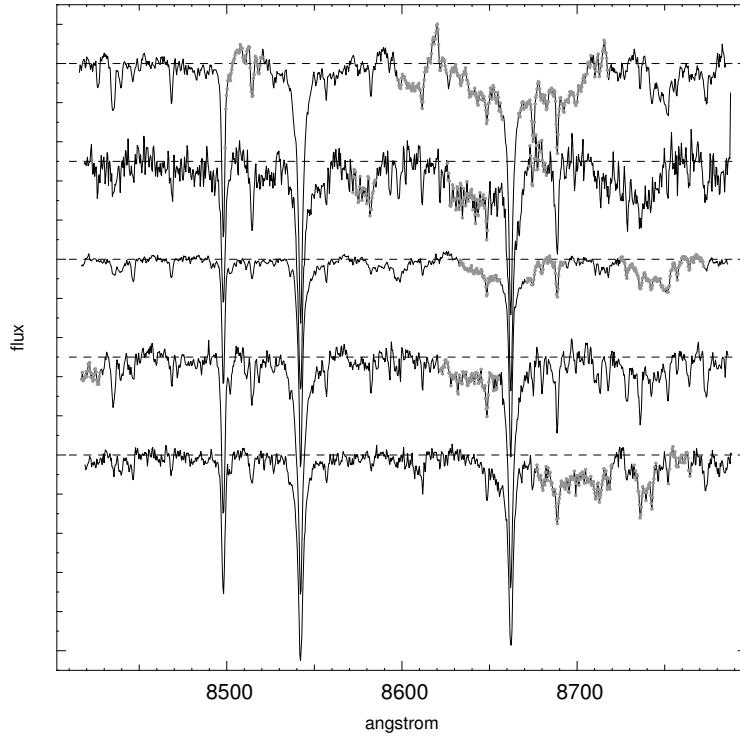
Figure 2.10: 5 RAVE normalized spectra. The grey points are the pixels with bad continuum normalization and recognized by the MASK code.

- new continuum normalization

- new Laplace multiplier

- number of templates used for the averaging process changed from 300 to 150

- every pixels having normalized flux>1.5 has weight=10^{-6} in order to remove pixels with cosmic rays from the χ^2 evaluation

- STN is used instead of SNR during the χ^2 evaluation

- detection of bad matching parts of the spectrum due to ghosts/fringing is performed by the code MASK

2.7.1   Tests on real spectra

To test the above modifications we ran the pipeline on standard stars observed by J. Fulbright and from the work by Soubiran & Girard 2005 (reported in section 2.1). The residuals of the RAVE measured parameters from the reference parameters is plotted as a function of the reference parameters in Figure 2.12. For comparison the results from the

old pipeline are shown in Figure 2.11. By comparing the two figures, we see that the modified pipeline reduces the underestimation of [m/H]. On the other hand, there is a small general overestimation of $T_{\rm eff}$ ($\simeq$+49K) and $\log g$ ($\simeq$+0.08) for the new pipeline. The RMS of the residuals are similar for the two pipelines. Overall, the distribution of RAVE stars in parameter space is improved with the overdensity at [m/H]$\sim$0.0 dex removed and with the [m/H] distribution (Figure 2.14, right-bottom panel) matching the Geneva-Copenhagen Survey (GCS) data (Nordtröm et al. 2004) better that the old pipeline (Figure 2.13, right-bottom panel). The underestimation at $T_{\rm eff}$ >6000K for dwarf stars is nearly unchanged, as shown in Figure 2.7.

2.8 Discussion and conclusions

The modifications of the RAVE pipeline have led in many improvements in the stellar parameters results. Foremost of these is the improvement of the [m/H] distribution, which is mainly due to the continuum normalization (asymmetric rejection of the points as function of S/N and more stringent Lagrange multipliers). Also the unnatural overdensity of stars at [m/H] =–0.05dex and $\log g$ =4.2 in Figure 2.13 was removed (see Figure 2.14) as a result of the new points in [m/H] added to the grid. The systematic underestimation of the uncalibrated metallicity [m/H] is reduced from $\simeq$–0.3 to $\simeq$–0.15. However, the correction of the systematic errors in [m/H] generates new systematics in $T_{\rm eff}$ and $\log g$, with both parameters marginally overestimated by the new pipeline. We found that this overestimations became stronger for the extended wavelength range. This is contrary to the expectation as the extended wavelength range gives more information and therefore we would expect better parameter estimation. Such a systematic may occur if the the Paschen lines H I were too strong in the model spectra. To verify this hypothesis, we compare some real and synthetic spectra. In Figure 2.15 we plot the RAVE solar spectrum and the nearest match synthetic spectrum synthesized with the MARCS model[2] atmospheres (Sordo et al., 2009) and Castelli & Kurucz model atmospheres (Munari et al., 2005). The H I line at 8665Å is too strong in the Castelli& Kurucz model (from RAVE's library), while the MARCS model matches it better. We also looked at a hotter temperature ($T_{\rm eff}$ =6250, $\log g$ =4.5 and [m/H] =–1.0dex) where the two H I lines at 8665 and 8750Å of the Castelli& Kurucz model look again too strong with respect to the MARCS model (Figure 2.16). This would lead to an underestimation of $\log g$ and $T_{\rm eff}$ for any spectra having significative intensities of the Paschen H I lines.

Another characteristic of the RAVE spectrum shown in Figure 2.15 (red line) is the too strong Fe lines compared to the Sun (black line) and to the MARCS model (blue line). This is partially due to difference in microturbulence ξ (which is ξ=1 km/sec for the Sun and the MARCS model and ξ=2 km/sec for the RAVE spectrum) as a higher ξ causes the

---

[2]MARCS model spectra have been kindly provided by B. Edvardsson, private communication

desaturation of the lines and make the EWs larger. This difference cannot be removed because for $\xi=1$ km/sec the grid is not complete, leading to an underestimation of the [m/H].

We conclude that the new pipeline can measure stellar parameters with uncertains of $\sim$300K in T_{eff}, $\sim$0.5 dex in $\log g$ and $\sim$0.2 dex in [m/H] with few known systematic errors: a shift of -0.15 dex in [m/H] and an underestimation in T_{eff} for dwarf stars with T_{eff}>6000K.
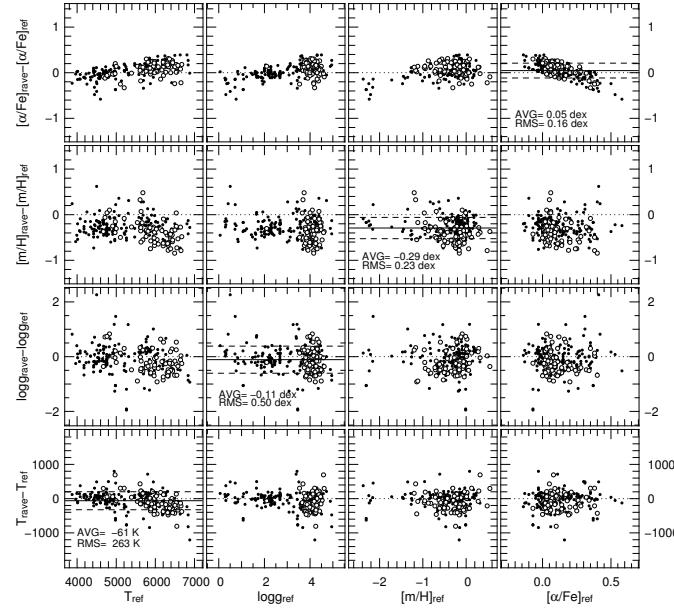
Figure 2.11: Residual of measured-minum-expected versus expected parameters for 162 RAVE spectra of stars observed at high resolution by J. Fulbright (solid points) and 70 RAVE spectra of stars from the Soubiran & Girard catalogue (open points) obtained with the RAVE pipeline. Average (AVG) and standard deviation (RMS) are reported.
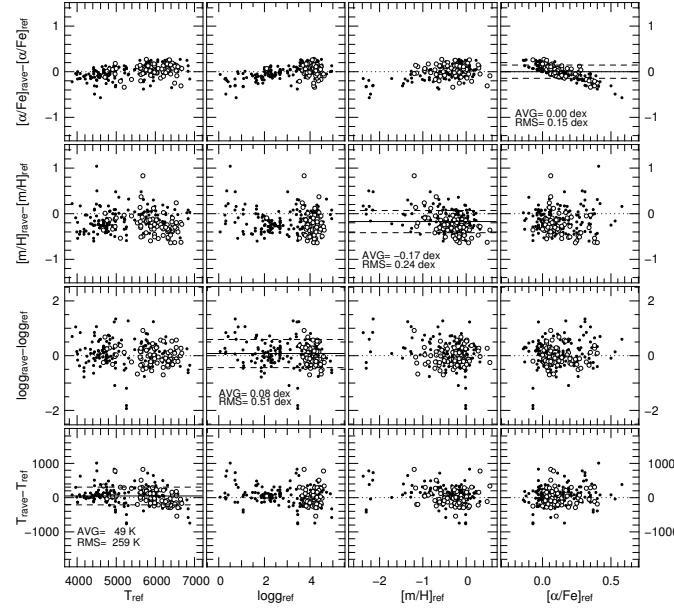

Figure 2.12: As Figure 2.11 but for the new modified pipeline.
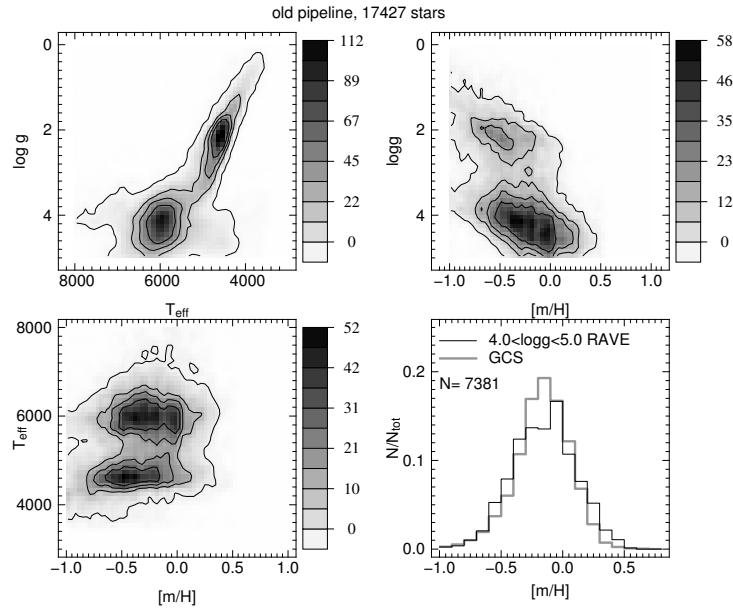
Figure 2.13: Stellar parameters distributions of a sample of 17427 RAVE stars from the second data release (original RAVE pipeline). Bottom right panel: comparison between RAVE (black line) and GCS (gray line) metallicity distributions for dwarf stars.
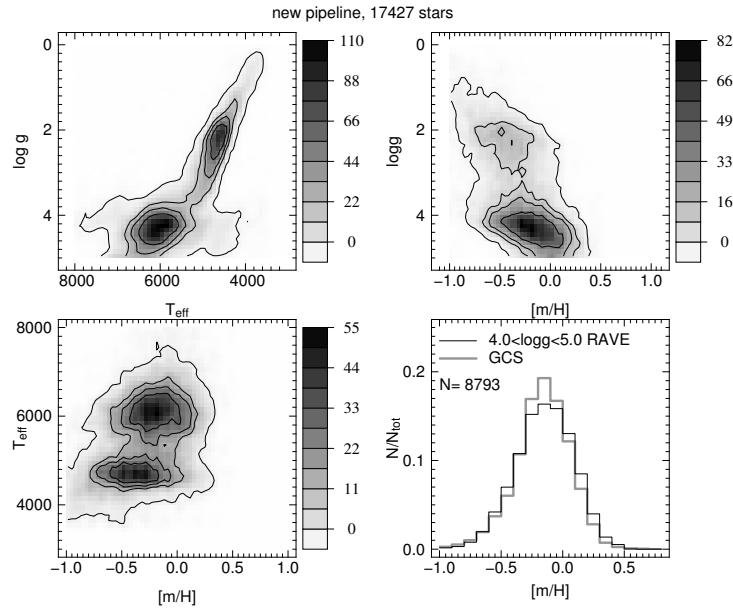

Figure 2.14: As in Figure 2.13 but for the stellar parameters estimated with the new modified pipeline.
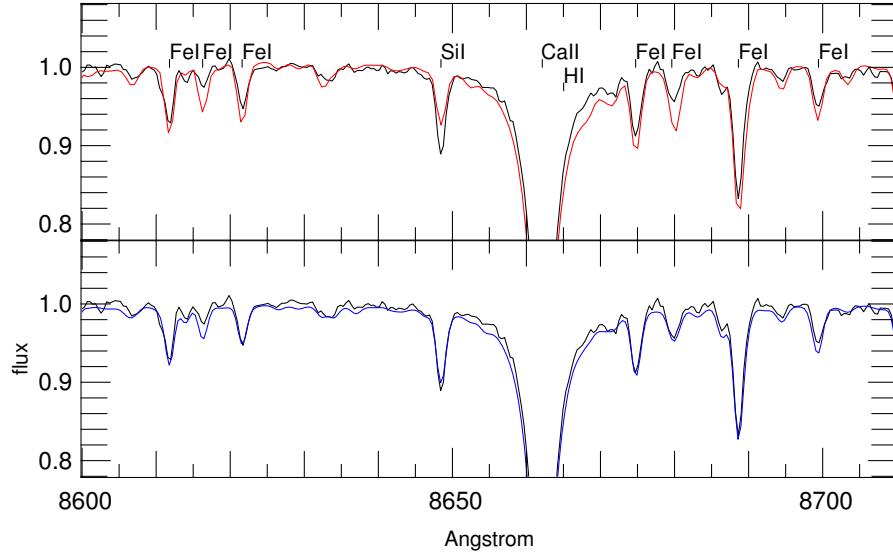
Figure 2.15: Comparison between the RAVE spectrum of the Sun (black line) and the closest synthetic spectra of the RAVE synthetic library (red line, top panel, T_{eff} =5750K, $\log g$ =4.5, [m/H] =0.0dex, [α/Fe]=0.0, ξ=2.0 km/sec) and the closest synthetic spectra by Sordo et al. (blue line, bottom panel, T_{eff} =5750K, $\log g$ =4.5, [m/H] =0.0dex, [α/Fe]=0.0, ξ=1.0 km/sec).
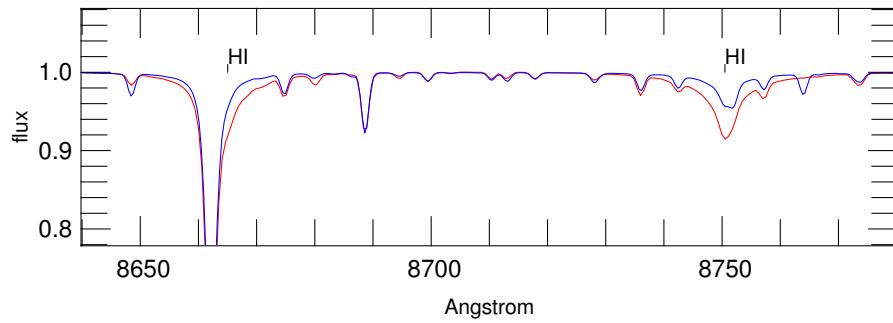

Figure 2.16: Comparison between synthetic spectra with parameters T_{eff} =6250K, $\log g$ =4.5, [m/H] =–1.0dex, [α/Fe]=0.0, ξ=2.0 km/sec of the RAVE library (red line) and the synthetic spectrum by Sordo et al. (blue line, courtesy of B. Edvardsson).

Chapter 3

Determining elements abundances from equivalent widths of absorption lines

In this chapter we describe how we wrote the first processing pipeline in order to measure thousands of RAVE spectra. The pipeline is written in perl programming language and makes use of some auxiliary codes written *ad hoc* for this task. However, this first version has been superseded by a new one (see next chapter) more efficient and sensible to the α elements. In the next sections we describe this first version of the pipeline, show its capabilities and explain why a new pipeline was needed.

3.1 The processing pipeline

The first version of the processing pipeline automatically does what one would do to measure chemical abundances of a star: take a spectrum, measure the equivalent widths of the absorption lines and pass them to a line analysis code (in this case MOOG) which returns the abundances. On this very basic procedure we built the pipeline in perl programming language. It combines several subroutines and Fortran77 codes which do the job and take care of more specific tasks depending on the star's spectral features. The MOOG code plays a key role as it computes the final chemical abundances, therefore we must provide all the information it needs. These are i) an atmosphere model from which MOOG takes information about the physical condition of the star's atmosphere such as temperature, gas pressure, electron density etc. ii) a list of absorption lines we measure (the line list is described in Chapter1), some physical quantities such as wavelength of the lines, excitation potential, oscillator strength iii) equivalent width of the lines in the observed spectrum. These basic requirements drive the main steps of the pipeline:

- create a proper atmosphere model from the stellar parameters of the star (Sec. 3.1.1)

- select an appropriate line list of the absorption lines to be measured (Sec. 3.1.3)

- measure the equivalent width of the lines using the code EWFIND (Sec. 3.1.5)

- pass all these data to MOOG which returns the chemical abundances.

This straightforward procedure is made more sophisticated by other subroutines to improve the performance in accuracy and speed. One subroutine creates the line list by using only lines (of the lines archive) which are strong enough to be visible in the spectrum under analysis (they are functions of the stellar parameters) making the analysis shorter (see Sec.3.1.3); for instance the molecular lines are left out from the line list for hot stars where they are not present. One subroutine estimates the individual equivalent widths of lines blended together (see Sec. 3.1.6) which are numerous in medium resolution spectra; this increases the number of lines used for the abundances measurement. One subroutine takes into account the large number of molecular lines which affect the spectra of late type stars and estimates first the molecular abundance to allow better accuracy of the atomic abundances computed afterwards (see Sec. 3.1.7). Other subroutines are concerned with the reconstruction of the spectrum (which reproduce the absorption features in agreement with their estimated equivalent width), signal-to-noise ratio estimation, detection and rejection of bad data (see Sec. 3.1.8, and 4.2.4). The writing of the results concludes the pipeline procedure for the considered spectrum.

3.1.1   Start and creation of the atmosphere model

The processing pipeline starts by considering one star's spectrum and its stellar parameters such as effective temperature, gravity, metallicity (hereafter T_{eff}, $\log g$ and [m/H], respectively) and signal-to-noise ratio (S/N) taken from the RAVE data archive. In order to perform the next steps (i.e. select a suitable line list and obtain chemical abundances) we need an atmosphere model as close as possible to the atmosphere of the star. We do it by using the atmospheres grid ATLAS9 by Castelli & Kurucz (2003). The Castelli & Kurucz's atmospheres are computed by steps of 250K in temperature, 0.5dex in gravity and metallicity and are (linearly) interpolated in order to build an atmophere model matching the stellar parameters stored in the RAVE archive. The microturbulence ξ is computed using the formula by Allende Prieto et al. (2004)

$$\begin{aligned}
\xi &= 1.645 + 3.854 \times 10^{-4}(T_{eff} - 6387) - 0.6400(\log g - 4.373) \\
&\quad -3.427 \times 10^{-4} \times (T_{eff} - 6387)(\log g - 4.373) \text{ km s}^{-1},
\end{aligned} \tag{3.1}$$

which is a function of the stellar parameter temperature T_{eff} and gravity $\log g$. Following this formula, stars in the temperature range $4000 < T_{\mathrm{eff}}$ (K) < 7000 have microturbulences which vary between 0.0 and 3-4 km s^{-1}. It must be said that tests performed on real

spectra (see Sec. 5.2.3) shown no significant differences in resulting abundances if ξ is kept constant at $\xi = 2.0$ km s^{-1} or if formula (3.1) is employed[1].

3.1.2 LTE analysis with the MOOG code

MOOG is a standard Local Thermodynamic Equilibrium (LTE) line analysis and synthesis code (Sneden, 1973). This code, as well as other similar codes, relies on an atmosphere model and a line list of absorption lines. Among several tasks (like spectra synthesis, curve of growth creation and others) it can calculate the EWs expected by a given atmosphere model or compute the chemical abundances from the EWs. These two features are the ones we use, as they allow us to obtain the chemical abundances and build the list of the visible lines we measure in the spectra (see Sec. 3.1.3). Since this software has been conceived for interactive use, some changes have been necessary in order to adapt it for batch mode use. The changes we made concerned the input data prompts and choices about stellar parameters, which in this case are fixed because taken from the RAVE data archive; during the chemical abundances estimation MOOG prompts for an extra computation for the CN molecular equilibrium. We removed this option since the molecular equilibrium computation needs to have individual abundances of C and N which we cannot measure: there are no C lines in our line archive and the 3 N lines present (see also Sec. 3.1.7) are not always visible and reliable.

3.1.3 The line archive and the line list

After a proper atmosphere model has been created, a list of absorption lines to be measured must be built. The RAVE lines archive holds all the lines collected and corrected as explained in Chapter 1. Early tests had shown that for low S/N the processing pipeline overestimates the abundances when we measure all the lines of the archive. This happened because noise spikes were mistaken for lines which were actually too weak to be visible. Therefore, we decided to measure only the lines strong enough to be visible through the noise. The line list (which is different for different spectra because the difference in stellar parameters) is built from the RAVE lines archive in the following way: the atmosphere model is passed to MOOG with the whole lines archive and it computes the theorically expected EWs; the lines whose intensities are enough to be visible through the noise will constitute the line list for the spectrum under consideration. A line is considered "visible" through the noise if its EWs satisfies the following constraint:

$$EW(m\mathring{A}) > 2 \cdot (\frac{1}{S/N}) \cdot d \cdot 1000, \tag{3.2}$$

---

[1]This is not true for the second chemical pipeline where different ξ give significant differences in abundances.

where d is the spectral dispersion (for RAVE $d = 0.4\mathring{A}$/pixel) and $\frac{1}{S/N}$ represents the fluctuation in normalized flux we expect due to the noise in one pixel.

We apply a specific treatment for the 8 H I and 3 Ca II lines present in the line archive. These lines are considered "strong", because they have very large EWs and wide wings. They are not used for abundances computation, but they are fitted and subtracted from the spectrum in order to remove their contribution which affects the EW measurement of the neighboring lines; this task will be performed by the code described in the next section.

3.1.4   Strong lines and continuum re-normalization

Before the EW measurements, the spectrum is strong lines subtracted and renormalized. The process is explained in the following and illustrated in Fig.3.1

1. the metallic lines are reconstructed with Gaussian profiles having EWs predicted by MOOG for the RAVE metallicity (Fig. 3.1 panel a, red line) and subtracted from the spectrum

2. the strong Ca II and H I lines are fitted respectively with Lorentzian and Gaussian profiles (blue line in panel b) on the result of the previous step (black line in panel b) through a χ^2 minimization routine and subtracted from the spectrum

3. the residuals left (black line in panel c) are due to the continuum shape and noise. They are fitted with a line which represents the continuum level where the EWs are later estimated (violet line in panel c)

The violet line is fitted using the following procedure

1. consider an interval of 31 pixels centered on the i-th pixel of the residuals

2. fit a parabola with a weighted χ^2 minimization process

3. assigns to the i-th pixel of the continuum fit the value of the parabola at i-th pixel and repeat it for all the pixels

As we have already seen in section 2.6.2 even an already normalized spectrum can still have uncorrected continuum features. Because the precision of the EW measurement depends on the continuum position, it is very important to have a reliable estimate for it. This part of the code fits the shape of the continuum from the residuals left after the subtraction of the strong lines and metallic lines. The EWs of the lines are measured over the line composed by the continuum level plus the strong lines (violet line and blue lines respectively) which is represented by the green line in panel d of Fig. 3.1.
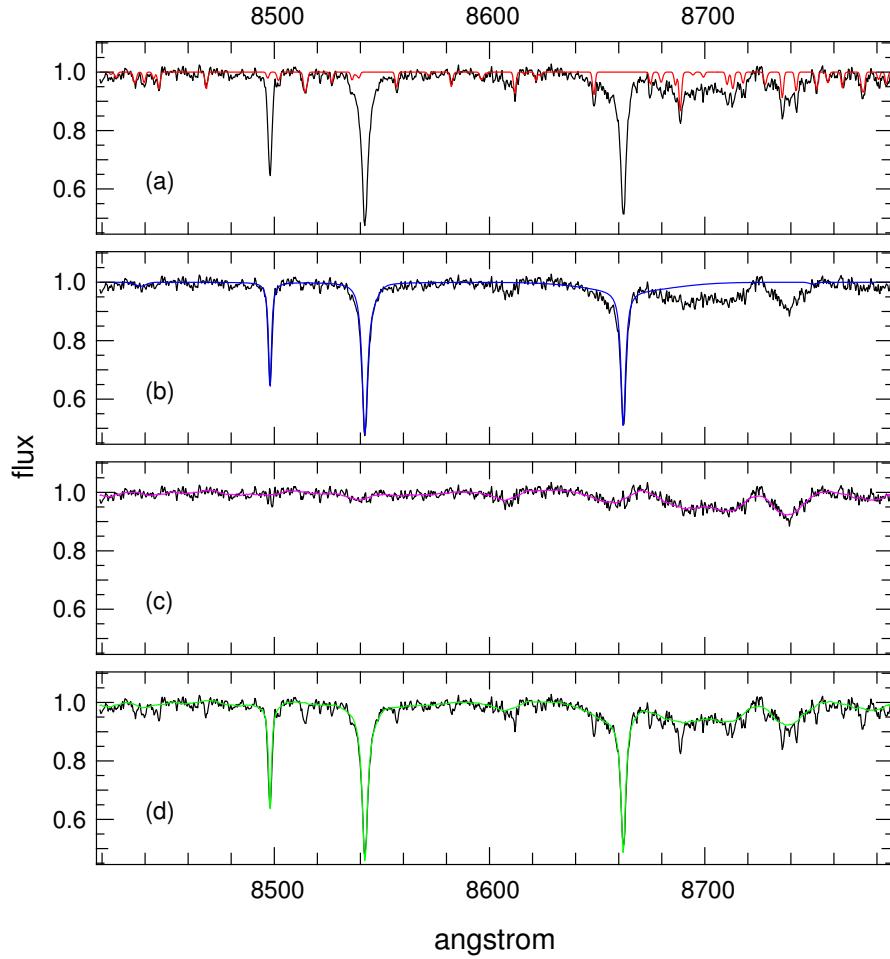
Figure 3.1: Black line: RAVE spectrum of the star HD149996. Red line: metallic lines fitting. Blue line: strong line fitting. Violet line: continuum fitting. Green line: blue and violet line summed together. On this last fitting line the chemical abundances will be estimated.

3.1.5 The EWs measuring code: EWFIND

EWFIND is a dedicated Fortran77 code we wrote to measure the EWs of the absorption lines. Unlike other codes which fit the absorption lines with a given profile, EWFIND integrates the observed absorbed flux of the lines. The code disentangles the lines by solving a linear system; what follows explains the main idea.

Consider the absorbed flux f_i integrated over the interval I_i centered on the i-th line which has a central wavelength λ_i. Assuming a Gaussian profile for the line, the line's fractional area A_i under the interval I_i is

$$A_i = \int_{I_i} g(\lambda_i, \sigma) d\lambda, \tag{3.3}$$

where $g(\lambda_i, \sigma)$ is a Gaussian function with centre λ_i and standard deviation σ. Its EW_i can be expressed as

$$EW_i = \frac{f_i}{A_i}.$$

(3.4)

This idea can be applied to all the lines (isolated and blended) by solving the linear system

$$f_i = \sum_j EW_j \cdot A_{ij},$$

(3.5)

where A_{ij} is the fractional area of the Gaussian function $g(\lambda_j, \sigma)$ computed under the i-th interval:

$$A_{ij} = \int_{I_i} g(\lambda_j, \sigma) d\lambda.$$

(3.6)

Eq. 3.5 states that the measured flux f_i within the interval I_i is the contribution of all the lines by an amount that depends on their EW and on their distance from the i-th line. It is a linear system which can be analytically solved in the ideal case. In case of real spectra, the noise makes the system analytically unsolvable and a minimization routine must be used to find a solution.

For its use in the RAVE pipeline, the EWFIND code relies on data like

- the line list and the number of lines hold in it

- the number of strong lines

- the number of pixels which compose the spectrum

- the S/N ratio of the spectrum

- the spectral resolution expressed as FWHM

During the processing the FWHM has been considered constant; this is usually true, although not for all the spectra, because the spectrometer focus is not always optimized. Our attempt to estimate the FWHM spectrum-by-spectrum by writing a dedicated routine was abandoned because it was not always possible to find one (or more) isolated line strong enough to permit a reliable FWHM estimation. Therefore, after many FWHM estimations (with a Gaussian fitting using the IRAF task *splot*) of several spectra, we decided to assume a FWHM=1.36Å for all the spectra.

3.1.6 The deblending routine

When two or more lines are closer than 1 FWHM, the EWFIND code cannot separate them and it outputs the EW of the blended feature. This happens often in the RAVE

spectra. Because of the medium resolution most of the lines are instrumentally blended and few (or no) isolated lines are left. Therefore we cannot follow the procedure usually adopted with high resolution spectra where only isolated lines are used and blends are avoided because we would be left with few or no lines to measure. The pipeline exploits the information held by the blended feature: the EW of the blend is the upper limit of the EWs of the lines which compose it (therefore an upper limit to the abundances). For instance, be EW_{tot} the equivalent width of a blend composed by the two lines l_1 and l_2 having unknown EW_1 and EW_2. Since we know the stellar parameters of the spectrum under analysis the only parameters which drive the EWs of the lines are the abundances. Thus, the value EW_{tot} is the upper limit of EW_1 and EW_2. If we have a rough estimate of the individual abundances, we can compute the expected EW_1 and EW_2 which, compared with the observed EW_{tot}, help us to get an abundance estimation better than a mere upper and lower limit. We developed a deblend procedure which assumes, as a first guess, that the metallicity [m/H] is equal to the individual abundances [X/H]. Then the individual abundances are inferred from the ratio expected between EW_1 and EW_2 at [X/H]=[m/H] and the constraint $EW_1 + EW_2 = EW_{tot}$. We describe here the procedure we developed to deblend blended lines in the case of two lines l_1 and l_2 belonging to different species and having different abundances. The algorithm procedes as follows:

1. Consider the blend feature with equivalent width EW_{tot} composed of two lines having unknown EW_1 and EW_2 so that $EW_{tot} = EW_1 + EW_2$.

2. Compute the expected theoretical EW_1^e, EW_2^e and $EW_{tot}^e = EW_1^e + EW_2^e$ by using MOOG and the proper atmosphere model (this mean that all the abundances are assumed equal to the general metallicity [m/H] of the star).

3. Compute the ratios $R_1^e = EW_1^e/EW_{tot}^e$ and $R_2 = EW_2^e/EW_{tot}^e$.

4. Obtain the "deblended" EWs as $EW_1 = EW_{tot} \cdot R_1^e$ and $EW_2 = EW_{tot} \cdot R_2^e$.

The main drawback of this method is that the resulting abundances are not fully independent from the first guess [X/H]=[m/H] . As a consequence, the difference |[X/H]-[m/H]| can be underestimated.

3.1.7   The CN routine

Most of the cold giant stars have their spectra dominated by CN molecular lines (376 out of 604 lines of the line archive are CN lines) and the medium spectral resolution of RAVE spectra makes them blended with most of the atomic lines. This affects the measurements of the latter. A good estimation of the line intensities improves the work of the deblending routine and gives more accurate atomic abundances. Therefore we wrote the CN routine which searches for proper C and N abundances and helps the deblending routine to give

more reliable results. We apply the CN routine only to the spectra which show more than 50 CN lines. It works as follow:

1. after the deblending routine run, MOOG performs the first abundances estimation

2. we take [C/H] and [N/H] from the first abundances estimation, insert them in the atmosphere model and run again the deblending routine. MOOG performs the second abundaces estimation

3. we take [C/H] and [N/H] from the second abundances estimation, insert them in the atmosphere model and run again the deblending routine. MOOG performs the third abundaces estimation

The CN routine iteratively approximate the intensity of the CN lines. As a result, the measurement of the metallic lines will not be affected by the presence of the CN lines.

3.1.8 Spectrum reconstruction

After all the EWs and the consequent abundances have been obtained, we reconstruct the spectrum from the EWs of the lines we measured, by using Lorentzian line profiles for Ca II and Gaussian profiles for H I and other metallic lines. The reconstructed and observed spectra are used to compute the χ^2 statistic as parameter of goodness of fit.

3.2 Accuracy and reliability

During the development of this work we performed detailed tests to estimate the quality of the results provided by this pipeline. Because it has been superseded by a new pipeline (see Chapter 4), we chose to illustrate here only part of them (the most representative ones) to show the quality of the results expected and to keep the text free of unnecessary complications. We present here one test on synthetic spectra and one test on real spectra.

3.2.1 Test on synthetic spectra

We synthesized a sample of 1353 synthetic spectra and added three intensity levels of artificial noise in order to test the accuracy at S/N=100, 40, 20. To make the sample as realistic as possible the spectra have been synthesized with distributions of T_{eff} and $\log g$ from a mock sample of RAVE observations created by using the Besançon model (M. Williams, private communication). The T_{eff} vs $\log g$ distributions of the sample is shown of Figure 3.2. Each spectrum of the sample has the chemical abundances of one of the
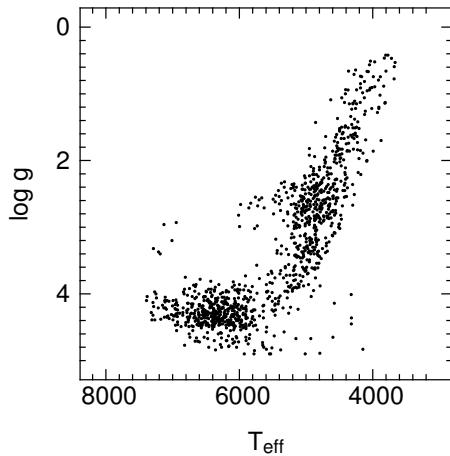
Figure 3.2: Density distribution of the synthetic spectra sample on the stellar parameters plane $T_{\rm eff}$ and $\log g$.

stars fromthe Soubiran&Girard catalogue (2005); this ensures a realistic distribution of the chemical abundances. Since our line list has only 8 elements in common with the Soubiran&Girard catalogue, we assigned the abundance [X/H]=[Fe/H] to the element Cr, Co, Ni, Zr and [X/H]=[α/H] to the C, N, O, Mg, Al, Si, S, Ca, Ti every time the measurement of one of these elements is missing. The spectra have been synthesized using the code MOOG at resolution 0.01Å/pix and degraded to RAVE resolution (0.4Å/pix, 1.4Å FWHM). We run the pipeline on these spectra at S/N=100, 40, 20 to check the performances at different S/N regimes.

Results at S/N=100. Figure 3.3 and Figure 3.4 compare the expected values (x-axis) versus the measured ones (y-axis) for [X/H] and [X/Fe] respectively. The pipeline is capable of measurements with error of $\simeq$0.1dex at S/N=100 with small offsets. The worrying feature is visible in Figure 3.4 where the enhancements of the elements with respect to Fe are systematically underestimated. This is an undesired feature because the enhancement of the α elements is an indicator distinguishing thin from thick disk stars which is important during the data analysis and interpretation.

3.2.2   Test on real spectra

Because RAVE observes stars in the magnitude range 9<I<12, there is no star from that survey having precedent high precision chemical abundances estimation to compare with ours. Therefore, we observed for this purpose 104 stars chosen from the Soubiran& Girard work (2005, hereafter SG05) which collected hundred of high precision abundances measurements from the literature. We used these stars as "standard stars" for our abundances

measurements. Since these stars are particularly bright (6<V<9) most of the spectra have S/N>80, therefore the results represent what can be obtained at high S/N regime. On the other hand, by comparing our results with other's results (even if high precision results) we get errors which are larger than the one given by the pipeline itself, because the SG05 errors sum quadratically with ours. The comparison is limited to only the 8 elements we have in common with SG05: O, Mg, Al, Si, Ca, Ti, Fe, Ni.

The results on real spectra confirm what has been seen before: errors in abundance [X/H] are small ($\simeq$0.1-0.2dex) but the enhancement with respect to Fe are underestimated (Figure 3.5 and Figure 3.6).

3.3   Conclusions

The experience gathered with the first version of the RAVE chemical pipeline highlight the difficulties in measuring chemical abundances on medium resolution spectra. At that resolution, most of the lines are instrumentally blended and the individual EWs can be measured in very few cases, making the usual measuring method unpractical for a survey like RAVE. The deblending technique developed in this work is not fully satisfactory since it causes the underestimation of the elements abundances with respect to iron so that, for instance, stars enhanced of α elements appear as they were not enhanced.

Beside, this technique is strongly dependent to the RAVE [m/H] estimation and its uncertains. We conclude that for resolution R$\simeq$10000 or lower the EWs measurement technique must be abandoned and the abundances estimation must be performed by comparing the spectra with models of known parameters with a χ^2 technique. We applied this solution to the second version of RAVE chemical pipeline described in the next chapter.
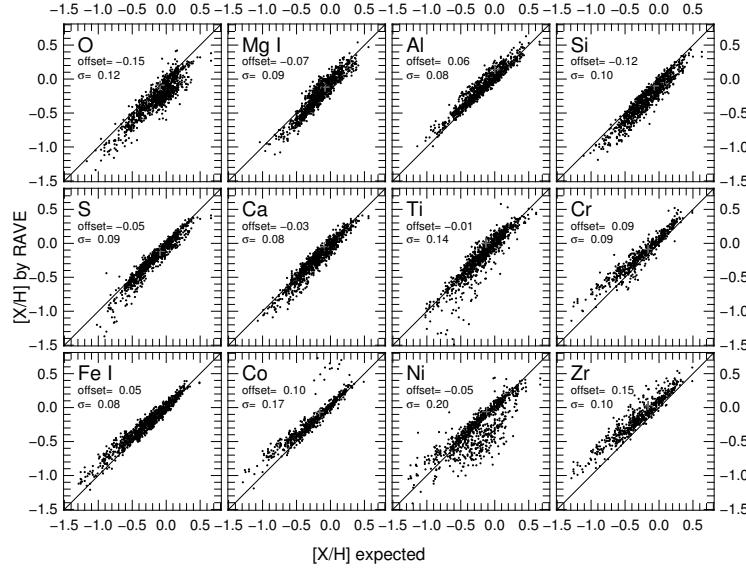
Figure 3.3: Expected abundances [X/H] versus RAVE abundances for the synthetic sample of spectra at S/N=100. The offset and standard deviation σ from the expected values are reported in the panels for each element.
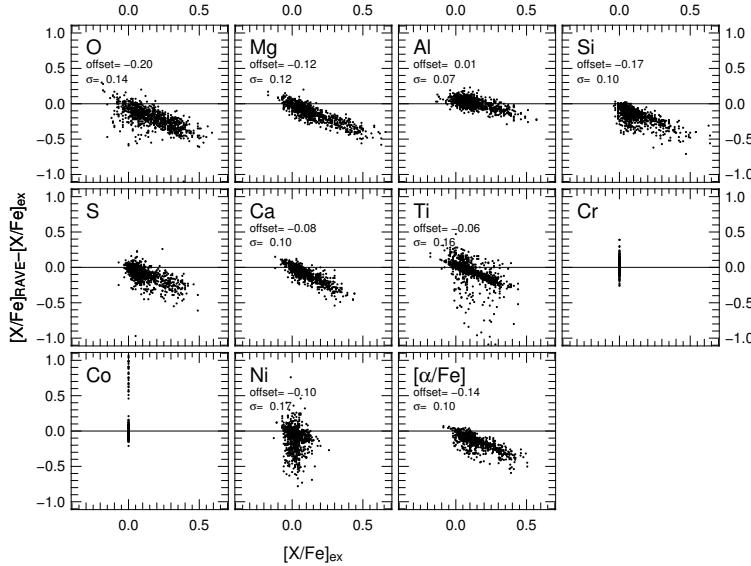

Figure 3.4: Expected relative abundances [X/Fe] versus residuals between RAVE and expected abundances for the synthetic sample of spectra at S/N=100. The offset and standard deviation σ from the expected values are reported in the panels for each element.
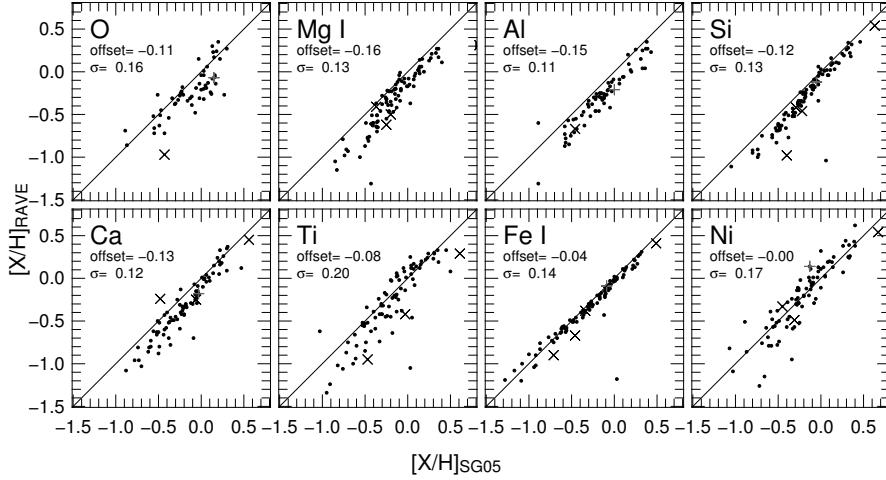
Figure 3.5: Comparison between abundances by Soubiran& Girard (SG05) and RAVE for a sample of 104 stars. We here adopted SG05 stellar parameters. Crosses are stars with S/N<30.
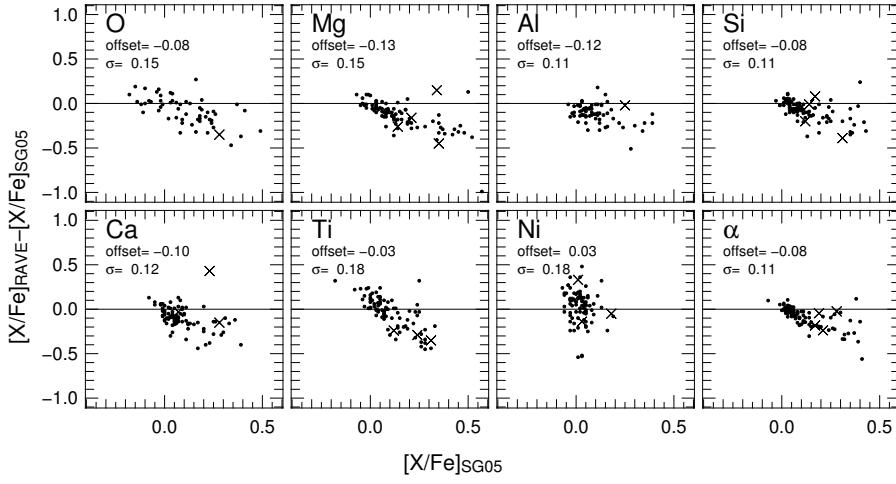
Figure 3.6: Residuals of abundances relative to Fe (y-axis, RAVE–SG05) and expected by SG05 (x-axis) for a sample of 104 stars. We here adopted SG05 stellar parameters. Crosses are stars with S/N<30. α is the average between Mg and Si.

Chapter 4

Determining elements abundances from χ^2 minimization technique

We here present a new chemical pipeline. It was born from the experience and know-how we gathered after the refinement of the RAVE pipeline (see chapter 2) and the previous chemical pipeline. It is based on a totally different concept with respect to the old one, and it provides better results. This is the chemical pipeline used to construct the RAVE chemical catalogue presented in Chapter 6.

4.1 The concept

The best method to estimate the abundances would be by synthesis: synthesize spectra with different chemical abundances until the χ^2 between synthetic and observed spectrum is minimized. The flaw of synthesis is that it is very expensive in terms of computational time for large wavelength ranges (more than few Å). On the other side, it can deal easily with blended lines because the intensity of the lines are inferred from their curve of growth (COG). This is a very desirable feature since at RAVE resolution most of the lines are instrumentally blended and very few lines are isolated and available for a precise EW measurement.

This version of the chemical pipeline makes use of the COGs of the lines. It reconstructs them by assuming Gaussian profiles with FWHM of the instrumental profile and EWs according to their COGs. The spectrum is reconstructed by summing up all the lines. Then, the abundances are varied until the best match with the observed spectrum is reached.

The COGs of the lines are inferred by exploiting one of the MOOG tasks: *ewfind*. Given the stellar parameters, MOOG computes the expected EWs of the spectrum lines at different abundances. With EWs as function of the abundances, we reconstruct the COGs.

This new method is much faster than a synthesis but shares a flaw with the old method: the

opacity of the neighbouring lines is not taken in account. Therefore the EWs of physically blended lines are computed as they were isolated, leading to an underestimation of the resulting abundances. The quality tests have shown that this effect is small for most of the lines (some exceptions will be outlined later) and under control. The main payback of this method is that it fully resolves the instrumentally blended lines since the EWs of the lines of a specific element are all function of the unknown we are looking for: the abundance.

In the following we will outline in details the operations performed by the pipeline.

4.2   The new pipeline

This pipeline relies on an EWs library (holding the EWs of the RAVE lines archive for a grid of stellar atmospheres) and on the stellar parameters (T_{eff} $\log g$ [m/H]) of the RAVE archive. It also makes use of some auxiliary codes to find the STN and normalize the spectrum. The former is the same used for the RAVE pipeline (Sec. 2.6.1), the latter has been rewritten for this new pipeline. The pipeline procedure can be outlined as follows:

1. upload the observed spectrum, the RAVE stellar parameters $T_{\mathrm{eff}}^{\mathrm{RAVE}}$ $\log g^{\mathrm{RAVE}}$ [m/H]$^{\mathrm{RAVE}}$ and the RAVE line archive

2. compute the S/N and consider only the lines whose EW at RAVE stellar parameters are large enough to be visible through the noise (Formula 3.2)

3. fit the strong lines and correct the continuum

4. estimate the chemical abundances and infer the metallicity [m/H]$^{\mathrm{chem}}$ and α element enhancements [α/Fe]

In the following we outline the procedure step-by-step.

4.2.1   EWs library and COGs reconstruction

In order to infer the COGs of the absorption lines, we used the code MOOG to create an EW library of 145080 files holding the EWs of the RAVE line list for 30640 different atmosphere models. The atmosphere models are linear interpolation of the Castelli&Kurucz atmosphere models ATLAS9 (2003) and constitute a grid covering the T_{eff} range 3600-7600K in step of 100K, gravity range 0.0-5.0dex in step of 0.2dex, [m/H] range –2.5-0.5dex in step of 0.1dex. We computed for each atmosphere model the EWs for 5 levels of abundances enhancement [X/Fe]= –0.4,–0.2, 0.0, 0.2, 0.4dex assuming [Fe/H]=[m/H]. Every
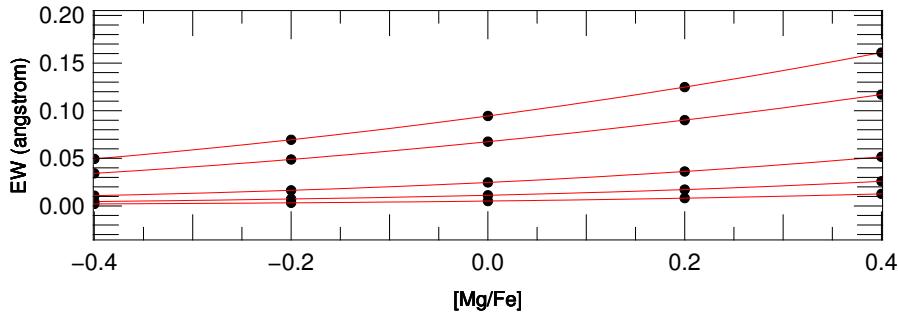
Figure 4.1: Curves of growth of 5 Mg lines reconstructed by fitting the EWs at 5 levels of enhancement (points) with a 3rd degree polynomial function (solid lines).

time we need EWs for stellar parameters which lie between the points of the grid, the closest points are linearly interpolated to obtain the EWs at the wanted parameters. Before the abundance measurement, these 5 EWs points are fitted by a 3rd order polynomial function which will serve as COG of the line in the range $-0.4 \leq [X/Fe] \leq +0.4$ (Figure 4.1).

4.2.2 Microturbulence

For every atmosphere model the microturbulence ξ must be fixed since the EWs depend on it. In high resolution spectroscopy ξ is determined by measuring the EWs of Fe I lines and then adjusting ξ until the abundances inferred from the individual lines are equal. We cannot follow this procedure because we do not measure EWs, therefore we must rely on a formula which gives ξ as function of the stellar parameters. Such formulas have been derived by several authors including Edvardsson et al. (1993), Reddy et al., (2003), Allende Prieto et al., (2004) where ξ is a function of $T_{\rm eff}$ and $\log g$, but unfortunately their formulas cover specific parameter ranges (hot dwarfs or cold giants). Thus, we derived our own formula in order to cover $T_{\rm eff}$ and $\log g$ ranges as wide as possible. Making use of high resolution spectroscopic data from several works (Earle-Luck & Heiter 2006 2007, Bensby et al. 2005, Fuhrmann 1998, Fulbright et al. 2006, Allende Prieto et al. 2004) which report stellar parameters and ξ of their star samples, we collected 712 giant and dwarf stars covering a wide range of $T_{\rm eff}$, $\log g$ and ξ (see Figure 4.2). We fit the data with the following 3rd degree polynomial function

$$\xi_{poly}(\text{km s}^{-1}) = \sum_{\substack{i,j=0,1,2,3}}^{i+j\leq 3} a_{ij}T^i(\log g)^j \qquad (4.1)$$
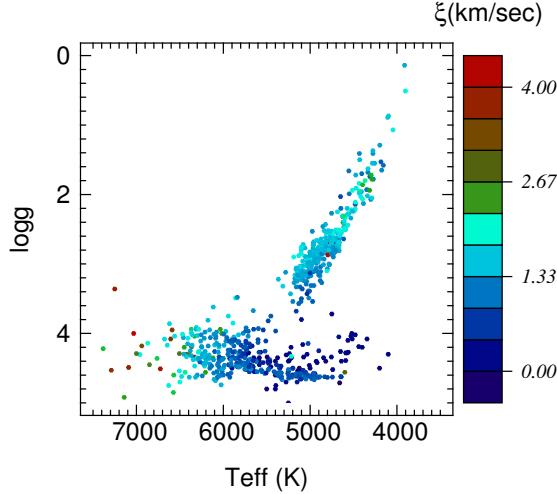
Figure 4.2: Distribution on $(T_{\text{eff}},\log g)$ plane of 712 high resolution spectroscopically observed stars. The colors code the microturbulence ξ.

and coefficients a_{ij}

$$
\begin{aligned}
a_{00} &= -10.3533 \\
a_{01} &= 2.59492 \\
a_{02} &= 0.161863 \\
a_{03} &= 0.176579 \\
a_{10} &= 0.00509193 \\
a_{11} &= -0.00157151 \\
a_{12} &= -3.99782 \cdot 10^{-7} \\
a_{20} &= -0.000361822 \\
a_{21} &= 3.82802 \cdot 10^{-7} \\
a_{30} &= -5.18845 \cdot 10^{-11}
\end{aligned}
\tag{4.2}
$$

A comparison between the law of Allende Prieto (2004) and our polynomial law is showed by the residuals between measured and computed ξ by the Allende Prieto's law (top panel, Figure 4.3) and ours ξ_{poly}(bottom panel). With an error of σ_ξ=0.32 km s^{-1} by our polynomial law we expect a consequent error in abundances smaller than 0.04dex for dwarf stars (Reddy et al. 2003, Mishenina et al. 2003).
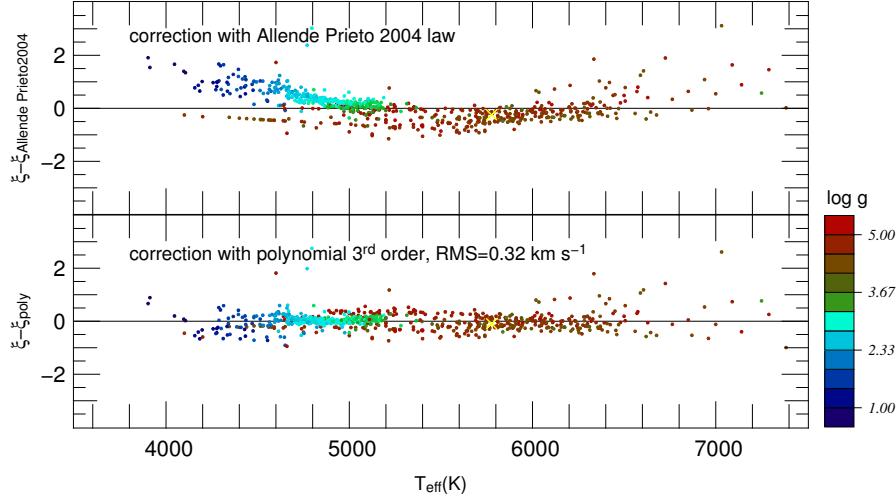
Figure 4.3: Top panel: residuals between measured and computed ξ by the Allende Prieto's formula for 712 stars. Bottom panel: residuals between measured ξ and computed by a 3rd degree polynomial function. The colours code the gravity $\log g$. The yellow cross represents the Sun.

4.2.3   Metallicity estimation

The pipeline performs a first metallicity estimation which serves for the continuum correction explained in the next section. In fact, in order to infer the spectral continuum we need to know the intensity of the absorption lines. Subtracting them from the spectrum, what remains is the continuum. On the other hand, to measure the intensity of a line we must know where the continuum lay, which leads to a vicious circle. Therefore we assume that the normalization performed by the RAVE pipeline is, in average, correct[1] and measure the intensities of the lines as done for the chemical abundances estimation (all the details are explained in Sec. 4.2.5) with the difference that here all the abundances vary together as one variable $[X/H]=[m/H]^{chem}$. Once the best match has been found, the value $[m/H]^{chem}$ will be used to reconstruct the metallic lines (first step described in Sec.4.2.4) and subtract them from the observed spectrum. This avoids pollution due to the metallic lines during the strong line fitting (second step described in Sec.4.2.4). During the $[m/H]^{chem}$ estimation the large wings of the Ca II and H I lines affect the results of the χ^2 analysis. Therefore we perform the analysis by rejecting wavelengths in the interval ~ 20Å centered on such a line.

---

[1]This is a good assumption because experience tells us that the continuum can be only locally wrong, i.e. on intervals smaller than 50Å .

4.2.4   Strong line fitting, continuum correction and statistic alarm

After the computation of needed quantities like STN and ξ the spectra have to be prepared before the chemical abundances measurement operations. The measurements do not make use of strong lines such as Ca II triplet or H I Paschen because they are affected by NLTE effects. Therefore these lines are fitted with a Lorentzian profile and removed from any further measurement. Besides, a new continuum normalization is applied because the previous normalization with a 3rd order spline function proved to be not suitable for abundance measurements. The strong line fitting and the continuum correction are performed in four steps as described by the 4 panels in Figure 4.4 from top to bottom:

1. the subroutine performs first a metallicity estimation. This estimation is done as described in Sec. 4.2.3 but avoiding the strong Ca II and H I lines. The metallic lines (red curve) are subtracted.

2. the strong lines are fitted with a Lorentzian profile for Ca II and Gaussian profile for H I and subtracted (blue curve, second panel from the top in Figure 4.4)

3. the continuum is the smooth curve obtained by a smoothing box car from what is left at this point (purple curve)

4. the strong lines and the continuum are summed to obtain the green curve on which the abundances measurement will be performed

   Around 25% of the RAVE spectra have part of the spectrum affected by ghost or fringing. The previous procedure can partially correct these undesired characteristics but the original intensities of the absorption lines may be affected. These spectra are labelled by the code "MASK" (explained in Section 2.6.2) which computes the fraction of spectrum having a particularly distorted continuum. This is also used as statistical alarm: when more than 30% of the spectrum suffer of these kinds of distortions, it is advisable to reject the spectrum.

4.2.5   Chemical abundances estimation

The pipeline uses the parameters $T_{\mathrm{eff}}^{\mathrm{RAVE}}$, $\log g^{\mathrm{RAVE}}$ and $[\mathrm{m/H}]^{\mathrm{RAVE}}$ for the abundances estimation. The spectrum is reconstructed by adding the absorption lines (Gaussian profile is assumed) to the estimated continuum (which include the strong lines, green curve in Figure 4.4). The pipeline computes the COGs of the lines for the RAVE stellar parameters and reconstructs the spectrum at $[\mathrm{X/H}]=[\mathrm{m/H}]^{\mathrm{RAVE}}$ as first guess. Then it computes the χ^2 beween the observed and reconstructed spectra and, through a minimization process, it changes the abundances $[\mathrm{X/H}]$ until the best match (minimum χ^2) with the observed
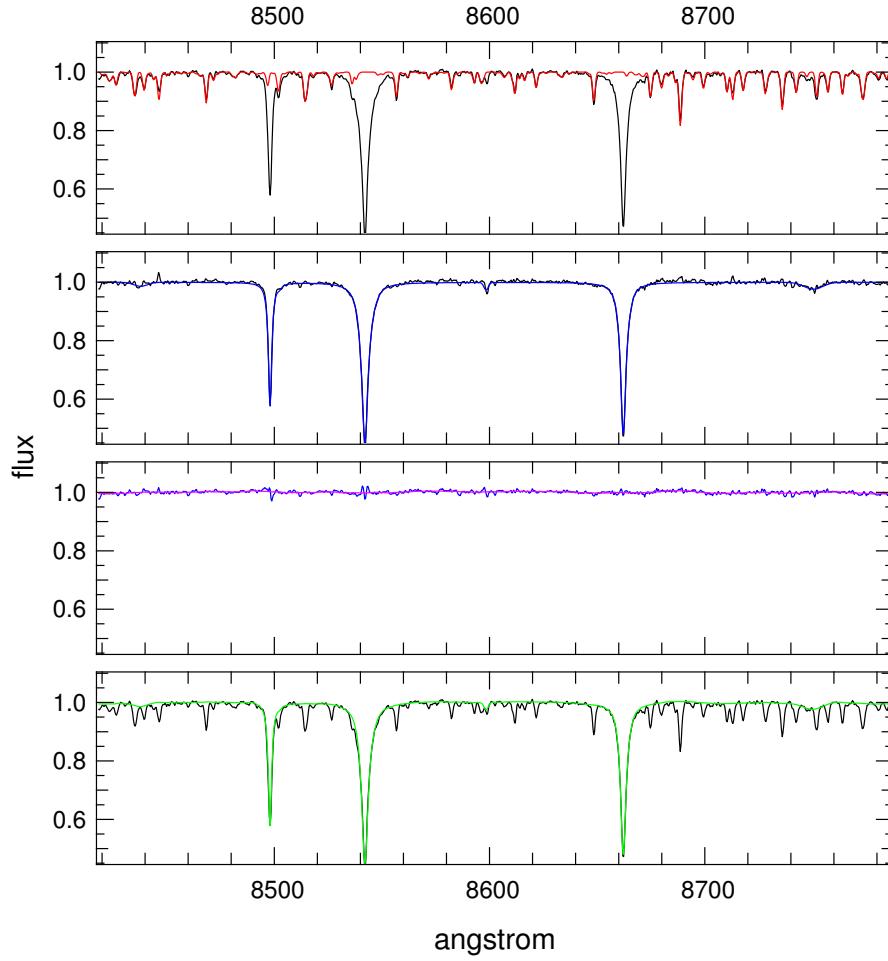
Figure 4.4: Black curve: RAVE spectrum of the Sun. Red curve: metallic lines fitting. Blue curve: strong line fitting. Purple curve: continuum fitting. Green curve: blue and purple curves summed together. On this curve the chemical abundances will be estimated.

spectrum is reached. The minimization process deals with 15 variables: 13 element abundances, one molecule and the instrumental Full Width Half Maximum (FWHM) of the absorption lines. The main minimum can easily be found because the single elements have only one minimum and they are (quasi) independent from one another. Figure 4.5 shows the best matching reconstructed spectrum of the Sun (green curve) versus the observed one (black curve). The three spectra on the top gives a view of how the reconstructed spectrum is built up: the spectra of three elements (Fe, Si, Mg) at parameters $T_{\mathrm{eff}} = 5861\mathrm{K}$ and $\log g = 4.54$ are reconstructed according to the estimated abundances and summed. The result is represented by the green curve.

As said before, this method neglects the opacity of the neighboring lines, leading to an underestimation of the abundances. In order to minimize this systematic error we apply a
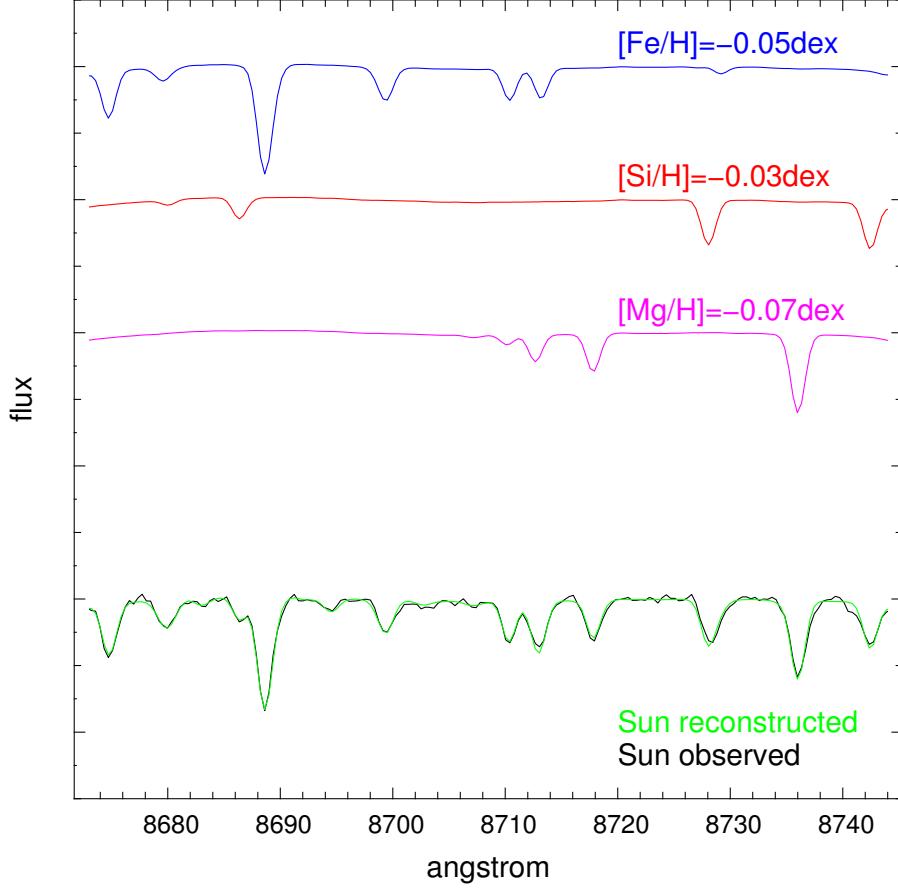
Figure 4.5: Once the abundances of Fe, Si and Mg are fixed their absorption lines can be reconstructed (blue, red and magenta curves respectively). By summing them together we obtain the green curve which is the reconstructed spectrum of the Sun. The black curve is the Sun' spectrum observed by RAVE.

correction coefficient to reduce the EWs of lines which are physically blended with other lines and are therefore affected by their opacity. The corrected EW is

$$EW_{\mathrm{corr}} = EW \cdot coeff \cdot cont \tag{4.3}$$

where the coefficient

$$coeff = 1 - \sum_{i}^{neighbor < 0.2\overset{\circ}{A}} EW_i / 2.50 / dispersion \tag{4.4}$$

is computed by considering all the lines closer than 0.2Å to the line considered (the dispersion is expressed in Å/pix) and *cont* is the value of the continuum at the central λ of the line. The latter takes into account the opacity due to strong lines (like Ca II) for absorption lines close to them. This correction reduces the systematic error which is estimated
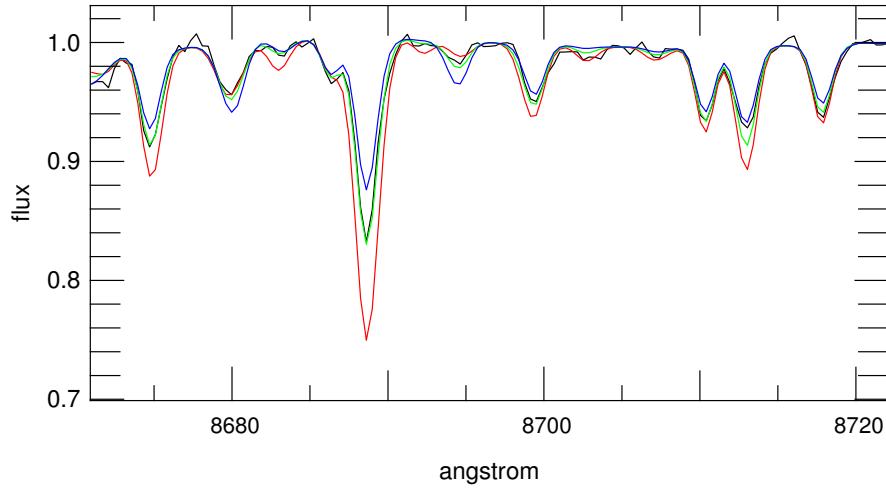
Figure 4.6: Spectrum of the Sun (black line) fitted with reconstructed spectra with [m/H] =–0.05dex, $\log g$ =4.54 and T_{eff} =5871K (green line), T_{eff} =5871+500K (blue line) and T_{eff} =5871-500K (red line).

to ∼–0.1 dex or less for most of the elements (more details are given in the next sections where quality check tests are outlined).

4.2.6 T_{eff} and $\log g$ estimation

The pipeline can estimate the parameters T_{eff} and $\log g$ by using the variation of the EWs of the lines as these parameters change. (In the following we use T_{eff} as an example and the line of reasoning is extended to $\log g$ and the general metallicity [m/H] as well.) This is in principle possible because different metallic lines have different behaviour under different physical conditions: for instance one Fe line with low excitation potential χ is stronger at low T_{eff} than one Fe line which has higher χ, and vice versa for high T_{eff}. This means that we can infer the T_{eff} by comparing their intensities. This varation can be described with a COG-like function inferred from the EW library as we did for the usual COG. In this case all the parameters but the one we want to measure have to be fixed. Let's consider T_{eff} as example: $\log g$ and [m/H] are therefore fixed and $T_{\mathrm{eff}}^{\mathrm{RAVE}}$ is the RAVE temperature estimation which is used as first guess. We compute the EWs of the lines at 5 temperatures $T_{\mathrm{eff}}^{\mathrm{RAVE}}$–500K, $T_{\mathrm{eff}}^{\mathrm{RAVE}}$–250K, $T_{\mathrm{eff}}^{\mathrm{RAVE}}$, $T_{\mathrm{eff}}^{\mathrm{RAVE}}$+250K, $T_{\mathrm{eff}}^{\mathrm{RAVE}}$+500K and infer a COG-like function for them, where now the variable is T_{eff} . The COG-like function is computed by fitting a 3rd order polynomial function to the 5 temperature points. The same procedure is used to estimate the gravity and the general metallicity [m/H] during the continuum correction described in Sec. 4.2.4.

In Figure 4.6 is shown how the intensities of the lines changes as function of T_{eff} when [m/H] and $\log g$ are fixed. The different behaviours of the lines with the variation of T_{eff}

assure the existence of a minimum χ^2 between the reconstructed and observed spectrum, i.e. the existence of a best T_{eff} estimation. At the moment this capability is not used to improve the stellar parameters since it needs more testing and developement. Moreover it works only for high S/N where the weak metallic lines are visible.

Chapter 5

Error estimation, accuracy, and reliability

The measurement of the chemical abundances, like any other physical measurement, is affected by errors. In this chapter we will identify the error sources and quantify the effects in order to estimate the accuracy of our results. After the identification of the internal and external errors sources, we will run the processing pipeline on synthetic and real spectra of well studied stars to better understand and highlight the effect of the errors acting together.

The measurement of a chemical abundance is an indirect measurement. From the spectrum of a star we determine the stellar parameters ($T_{\rm eff}$, $\log g$, [m/H]), which define the physical conditions of the stellar atmosphere, and then the absorbed flux of the lines (expressed as equivalent width, EW). From this information we obtain the chemical abundance through a model, i.e. we need to model the stellar atmosphere, to solve the equations relating all the quantities, and to compute how abundant the elements should be to match the observed spectrum. Therefore, the final errors depend on observational errors, on our incomplete knowledge of atomic process (such as transition probabilities), and on approximations in the atmosphere modelling.

We distinguish the errors as *internal* errors due to the assumptions made by the processing pipeline, and *external* errors due to instrumental effects, imprecision of physical quantities, and data used as input which do not depend on the pipeline's structure.

5.1   Internal errors

We call *internal errors* the uncertaines due to the specific procedure followed by the processing pipeline. They are outlined in the following.

Opacity due to neighboring lines affects the chemical abundance estimation. When one absorption line is close and physically overlap another line, its contribution to the local opacity affects the latter, weakening the intensity. The local opacity is fully considered during a spectrum synthesis but not in the RAVE chemical pipeline where lines are simply summed up. As a consequence, our abundances are slightly underestimated. To reduces the underestimation we applied the Formula 4.3, which efficiency will be quantified together with all the internal errors in the tests on synthetic spectra (see Sec. 5.2.1).

Stellar continuum normalization can be a difficult task in wavelength regions where numerous absorption lines are blended together and modify the continuum. If we do not know a priori the metallicity of the stars to infer the continuum level, the abundances are often underestimated because the continuum appears to be lower than it really is.

Strong lines like the Ca II triplet and the Paschen H I lines are not used for abundances computation but are simply subtracted from the spectra in order to better measure the neigbouring lines affected by their large wings. Since these strong lines show a Voigt profile, a Lorentzian profile (easier to implement than the Voigt function) does not always fit them best: this imprecision is minimized by the continuum renormalization applied to the spectra before the measurement. The effects of the subtraction of the strong lines is not easy to quantify in a real spectrum, because these lines undergo modifications of their profile by several causes (stellar rotation, core emission, asymmetries due to velocity fields in the stellar photosphere etc.). It will be evaluated in the following, when we perform abundance measurements on synthetic and real spectra.

Non-LTE effects affect the abundances measurement. They are not taken in account by this pipeline. Non-LTE effect depends from $T_{\rm eff}$, $\log g$ and [m/H] of the star as well as on the excitation potential of the lines (see Asplund 2005). In some cases they can affect the abundances by as much as ±0.3-0.5 dex. We are aware that our results are affected by these systematic errors (as are most of the literature works which rely on the LTE analysis) and they cannot be corrected until new Non-LTE codes capable to measure efficiently hundreds of lines will be available.

Microturbulence, ξ, makes the absorption lines broader and stronger. It represents the effect of photospheric velocities due to turbulence present on a stellar atmosphere. Classical LTE-analysis codes can approximate this effect by using the ξ parameter, which spans from 0 to 5 km s^{-1} for stars with $T_{\rm eff}{<}7000$K. We assume microturbulence as function of the stellar parameters as given in Eq. 4.1.

5.2 External errors

We call *external errors* the uncertainties which do not depend on the processing pipeline itself but on errors in the input data. The spectra have been observed with the 6dF instrument (see Steinmetz et al., 2006) and undergo a reduction process during which they are extracted, calibrated and normalized. Afterwards, stellar parameters are measured by the RAVE pipeline as explained in Zwitter et al., (2008). All these procedures introduce uncertainties that we consider "external" to the chemical processing pipeline.

The **focus** of the telescope and the resulting **spectral resolution** affect the measurements of the absorption lines because a bad focus makes the instrumental FWHM larger i.e. it decreases the resolution. It may happen that spectra are only partially out of focus, showing a different resolution at different wavelengths. Because it is not possible to consider different FWHMs in one spectrum, the pipeline automatically estimates one FWHM that best fits all the lines of the spectrum.

The **noise** in spectra is due to photon shot noise (which is Poissonian) and read-out noise (Gaussian) due to the detector devices. The noise is quantified in the parameter S/N as explained in the previous chapter. It affects the lines' EW as well as the determination of stellar parameters. Its impact on the abundance determination is evaluated through tests on synthetic and real spectra and discussed below.

The **stellar continuum** can be affected by optical defects such as fringing or different local response function of the detector at different wavelengths. The RAVE pipeline (explained in Steinmetz et al., 2006) corrects most of these defects but sometimes a satisfactory correction is not possible. The chemical pipeline can partially correct these defects thanks to the continuum renormalization (Sec. 4.2.4) and uses a statistic alarm (the MASK code) to determine where and how much the spectrum is affected.

Stellar parameters such as T_{eff}, $\log g$ and [m/H] are taken from the RAVE database. Wrong stellar parameters lead to wrong abundances. These errors are quantified in tests with synthetic spectra by measuring abundances after assuming stellar parameters with artificial errors.

Atomic parameters like oscillator strengths and damping constants can affect the abundances estimations. As we have seen in Chapter 1 oscillator strengths introduce errors in abundances of the same order as their own uncertains (0.2 dex or less). Damping constants are obtained from Unsöld's approximation multiplied by an enhancement factor E_γ, as suggested by Chen et al. (2000). Uncertainties on damping constants lead to smaller errors: Chen and collaborators showed that a change of 50% of E_γ leads to a change in

abundances by 0.01 dex.

Rotational velocity, V_{rot}, makes the line's FWHM larger and its core shallower even if the EW is conserved. Because of the medium resolving power of RAVE spectra, the V_{rot} obtained by RAVE are not quantitatively trustworthy, but it can be used as indicator of the width of the lines: large lines can be fast rotators or double-lined spectra. We want to avoid them since the pipeline cannot deal with double peaked lines. Spectra classified with $V_{rot} <$50 km/sec have a FWHM which is equal to the instrumental resolution. Therefore by choosing this constraint we avoid errors due to V_{rot}.

5.2.1   Chemical abundances accuracy from synthetic spectra

To see if the pipeline can recover the right abundances, we run the pipeline on a sample of synthetic spectra synthesized with known stellar parameters and chemical abundances. The sample of synthetic spectra is the same used to test the old pipeline (described in Section 3.2.1). For these tests we adopted the stellar parameters of the synthesis.

Results at S/N=100. Figure 5.1 shows that the pipeline can measure abundances with an average error smaller than 0.1dex at S/N=100. As expected, the abundances are slightly underestimated by $\simeq$0.1 dex or less for most of the elements but Ni, N and O which show underestimation of $\sim$0.2 dex or larger. These underestimations are due to the neglected opacity of the neighboring lines and to the continuum normalization which overcorrect wavelength regions crowded with lines. Since most the elements are affected at a similar degree (but the three elements just mentioned), the measurement of the enhancement [X/Fe] is still good and reliable, as showed in Figure 5.2.
Results at S/N=40. At this S/N the pipeline can only estimate the abundances of elements which exibith stronger lines like Fe, Mg, Al, Si (see appendix Figure A.3) and their enhancement with respect to Fe (see appendix Figure A.4). The expected error is $\simeq$ 0.1-0.2 dex. The systematics seen at S/N=100 are still present but less evident because of the larger dispersion due to the noise. A new systematic error appears for Ti, which is underestimated for $T_{eff} >$5000K.
Results at S/N=20. At this S/N the pipeline can still estimate abundances of Fe, Al and Si (Figure A.5 in appendix) and the relative enhancement to iron [X/Fe] (Figure A.6 in appendix), even with an error of $\simeq$0.2-0.3 dex. Other elements like Mg and Ti show systematic errors. An α-elements estimation computed as average of Mg and Si yields reliable results with an error of $\sigma \simeq$ 0.2 dex. Most of the elements cannot be measured because the lines are too weak to be detected in the noise.

In Figure 5.3 we study the **errors in abundances as function of the stellar pa-**

rameters at S/N=100. We found that abundances are underestimated for lower T_{eff} and $\log g$. As said before, this is due to the opacity of the neighboring lines and continuum overcorrected for spectra crowded of lines like for cold giants stars. For S/N<100 the trends are the same with larger scatter (see appendix for plots with S/N=40, 20).

5.2.2 Effect of stellar parameters errors on the accuracy of abundances

In the previous section we used spectra with known stellar parameters in order to evaluate the errors due to the pipeline only. We now want to estimate the errors in abundance when wrong stellar parameters are provided to the pipeline. This is what happens with real spectra because the RAVE stellar parameters are affected by errors. To simulate the RAVE errors we added to the correct parameters a random Gaussian error with standard deviation $\simeq$300K in T_{eff}, 0.5 dex in $\log g$ and 0.3 dex in [m/H].

Results at S/N=100. With errors added to the stellar parameters the abundances [X/H] have errors of $\simeq 0.15$ dex for the most important elements like Mg, Al, Si, Fe, Ni (see Figure 5.4). The elements N, O, S and Ti show systematic errors, being overestimated at high abundance and understimated at low abundance. The enhancement relative to Fe [X/Fe] (see Figure 5.5) has errors lower than 0.2 dex and no systematic errors for most of the elements.
Results at S/N=40. The abundance errors are in average $\simeq$0.15 dex (see Figure A.12 in appendix) and show the same systematics seen for S/N=100. The relative abundances [X/Fe] (see Figure A.13 in appendix) are reliable even with a large error for the most important elements Mg, Al, Si and Fe.
Results at S/N=20. At this S/N the pipeline can still estimate abundances of Fe, Al and Si and the relative enhancement to iron [X/Fe] even with an error of $\simeq$0.2-0.3 dex. Other elements like Mg and Ti show systematic errors. An α-elements estimation computed as average of Mg and Si yields reliable results with an error of $\sigma \simeq 0.2$ dex. Elements such as Ca and Ni cannot be reliably measured because the lines are too weak to be detected through the noise.

Figure 5.6 highlights the correlation between abundances and the errors on stellar parameters, showing that most of the element abundances are correlated with T_{eff}. Weak or no correlations with $\log g$ and [m/H] errors are visible.
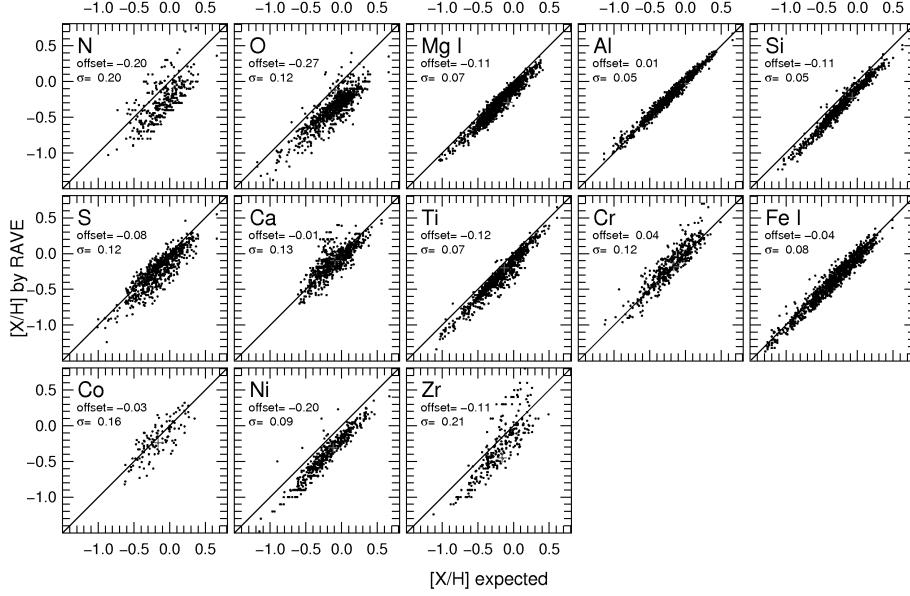
Figure 5.1: Expected abundances [X/H] versus RAVE abundances for the sample of synthetic spectra at S/N=100. The offset and standard deviation σ from the expected values are reported in the panels for each element. (See Appendix for plots with S/N=40, 20.)
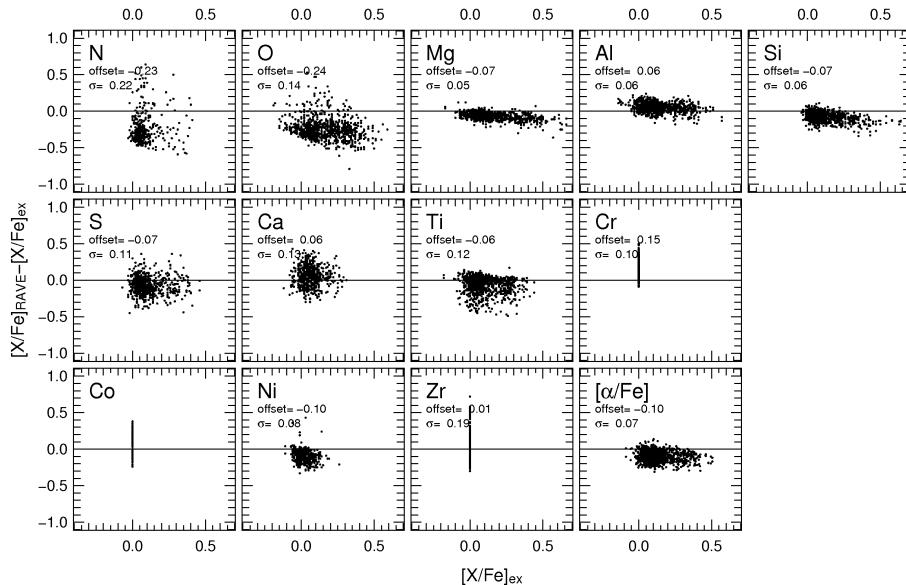

Figure 5.2: Expected relative abundances [X/Fe] versus residuals between RAVE and expected abundances for the sample of synthetic spectra at S/N=100. The offset and standard deviation σ from the expected values are reported in the panels for each element.(See Appendix for plots with S/N=40, 20.)
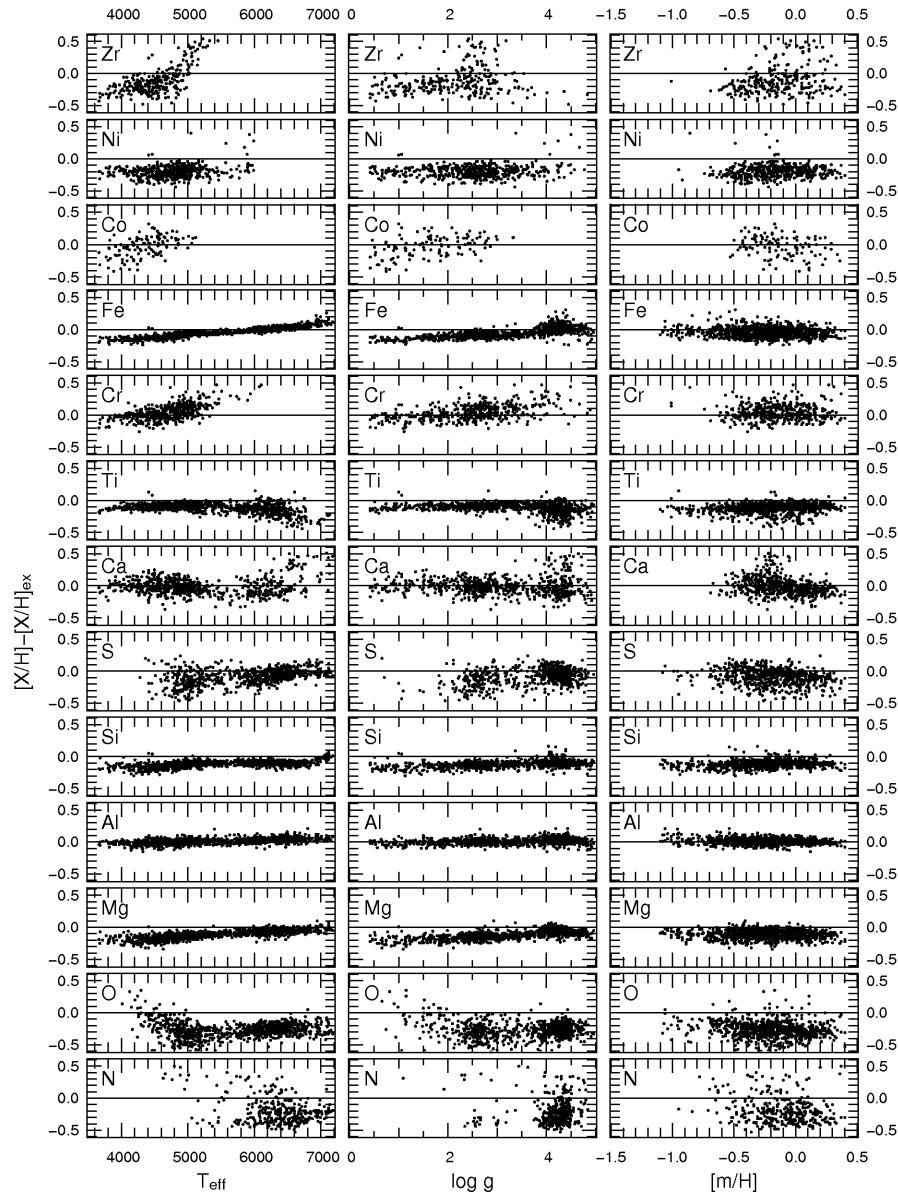
Figure 5.3: Correlation between abundance residuals (measured-minus-expected) and stellar parameters T_{eff} , $\log g$, [m/H] at S/N=100. See appendix for plots at S/N=40, 20.
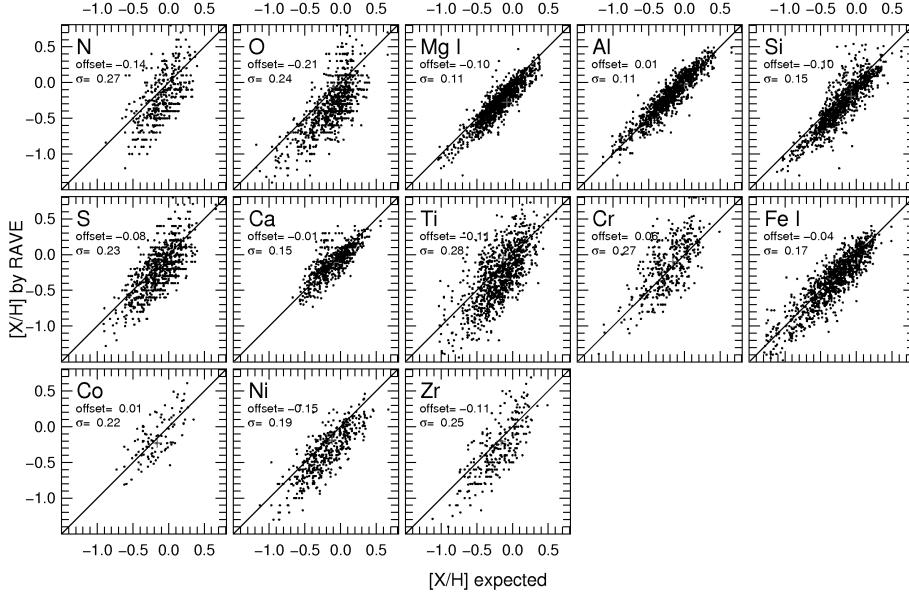
Figure 5.4: Comparison between expected (y-axis) and measured abundances (x-axis) at S/N=100 with "wrong" stellar parameters to simulate the RAVE stellar parameter errors. See appendix for plots at S/N=40, 20.
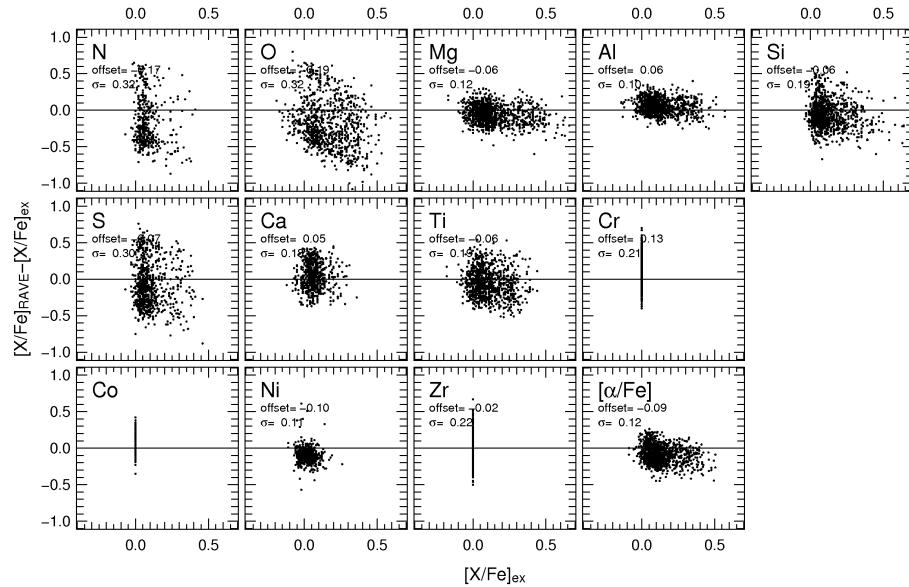
Figure 5.5: Like Figure 5.4 but for relative abundances [X/Fe]. See appendix for plots at S/N=40, 20.
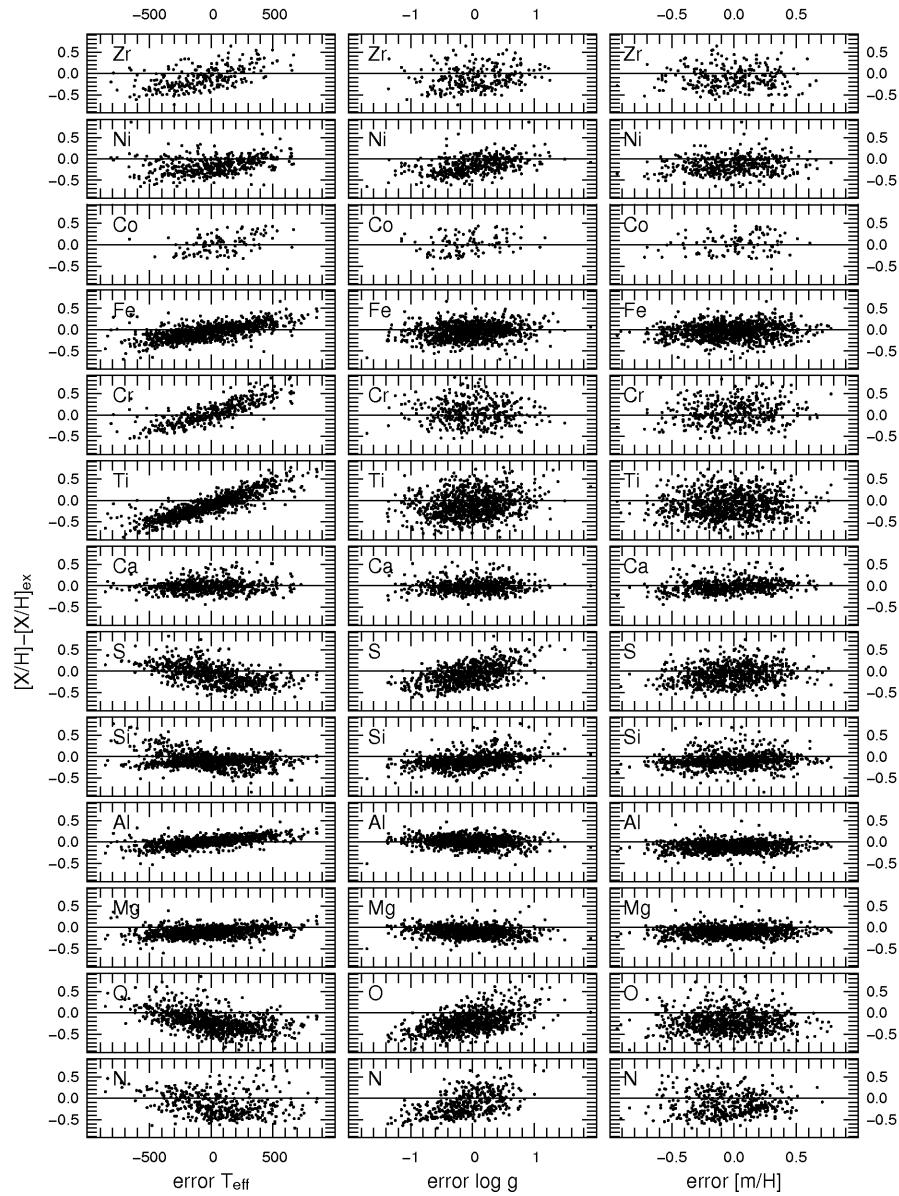
Figure 5.6: Correlation between abundance errors (y-axis) and parameter errors (x-axis) at S/N=100.

5.2.3   Chemical abundances accuracy from real spectra

In the RAVE input catalogue there are no stars whose chemical abundaces are known with high accuracy and that could be used as comparison. Therefore, we observed 104 stars chosen from the Soubiran & Girard work (2005) (hereafter SG05) which is a collection of high precision abundance measurements from the literature. We also used 207 spectra of 167 RAVE stars observed by J. Fulbright (JF) at the Apache Point Observatory for calibration purposes. Since they are particularly bright ($6<V<9$), most of the spectra have S/N>80 and the results represent what we can obtain at high S/N. For the SG05 stars the comparison is limited to only 8 elements in common with RAVE (O, Mg, Al, Si, Ca, Ti, Fe, Ni) whereas for the JF stars we can compare only Fe and α abundances (i.e. the average of the α elements abundances). To evaluate the errors due to propagation of the errors of the stellar parameters, we first analyze the result obtained by using the SG05 stellar parameters (which are considered more precise); then we repeat the analysis with the RAVE stellar parameters.

The results on the SG05 spectra obtained by using the stellar parameters by SG05 (Figure 5.7 and Figure 5.8) confirm the general underestimation of $\sim$0.1 dex found on tests with synthetic spectra (Figure 5.7). The errors are $\sim$0.1-0.2 dex, slightly larger than the ones obtained with synthetic spectra. We believe this is due to the SG05 errors which sum quadratically with the chemical pipeline errors. The pipeline can recover the enhancements [X/Fe] with a slight ($\sim$0.1dex) underestimation at high [X/Fe].
When we use the RAVE stellar parameters, the residuals (see Figure 5.9 and Figure 5.10) become slightly larger; in particular for Ti and Fe they are larger by $\sim$0.1 dex. The relative abundaces [X/Fe] appear to be underestimated compared to when SG05 stellar parameters are used.

The results on JF spectra obtained by using RAVE stellar parameters confirm the precision reached with the SG05 stars ($\sim$0.2 dex) (Figure 5.11). The slight abundance underestimation for giant stars observed in tests with syntetic spectra is here confirmed in real spectra (Figure 5.12) for Fe and α elements. Iron abundance and α elements enhancement is well recovered but for few cold metal poor supergiants stars ($\log g <1$) for which RAVE T_{eff} and $\log g$ appear too high.
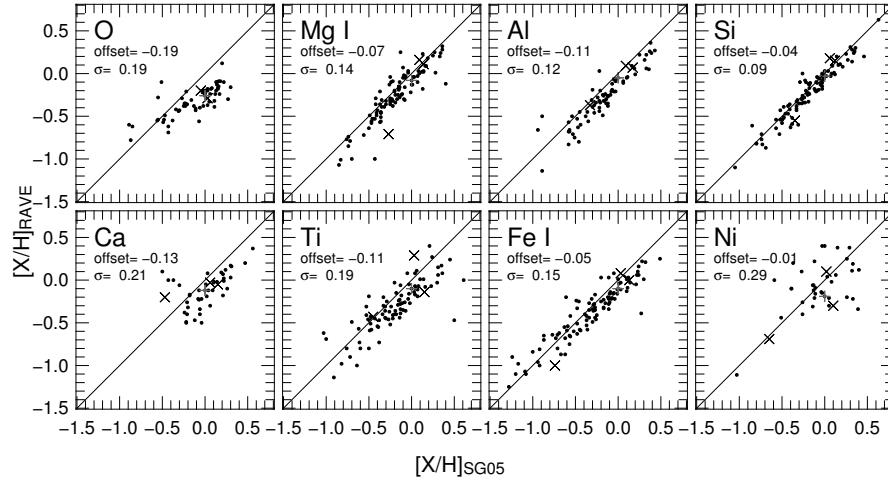
Figure 5.7: Comparison between abundances by Soubiran & Girard (SG05) and RAVE for a sample of 104 stars. We here used SG05 stellar parameters. Crosses are stars rejected because of continuum normalization problems.
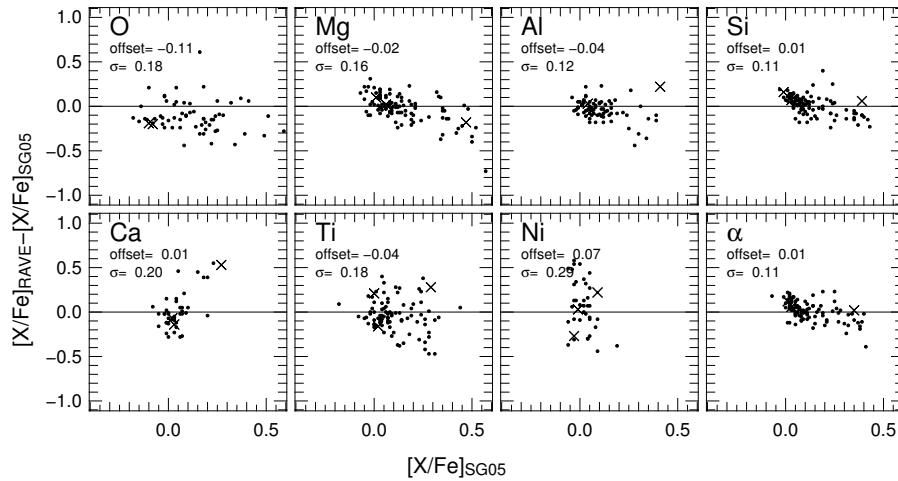

Figure 5.8: Residuals of abundances relative to Fe (y-axis, RAVE-minus-SG05) and expected by SG05 (x-axis) for a sample of 104 stars. We here used SG05 stellar parameters. Crosses are stars rejected because of continuum normalization problems. α is the average of Mg and Si.
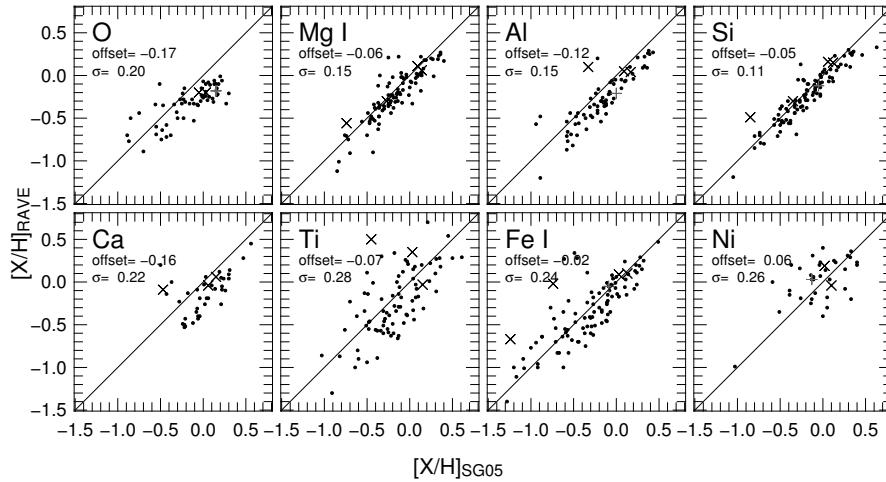
Figure 5.9: Like Figure 5.7 but using RAVE stellar parameters.
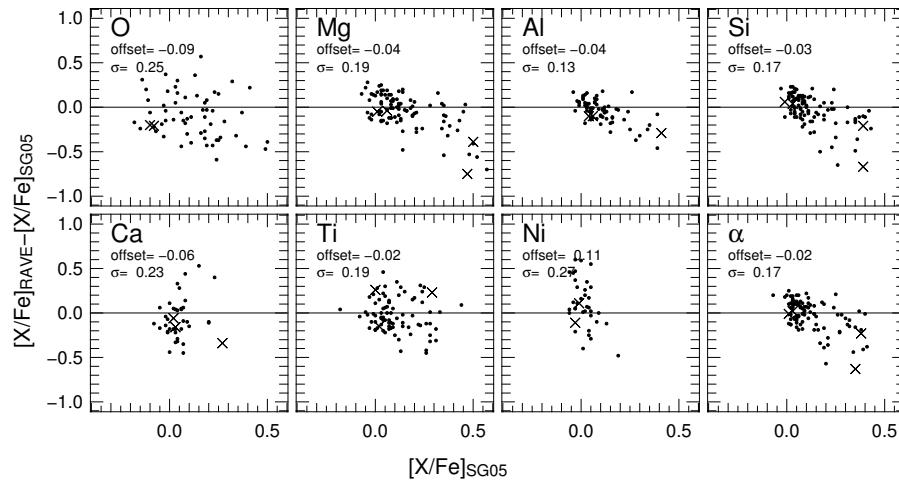
Figure 5.10: Like Figure 5.8 but using RAVE stellar parameters.
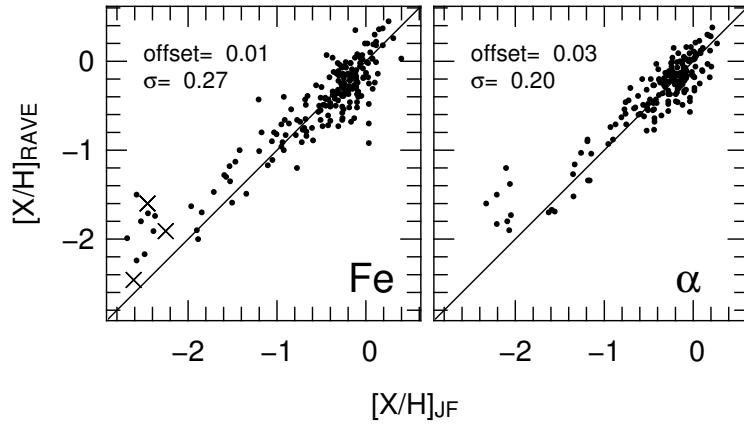
Figure 5.11: Comparison between abundances by J. Fulbright (JF) and RAVE for 207 spectra of 167 RAVE stars. We here used RAVE stellar parameters. Crosses are stars rejected because of continuum normalization problems.
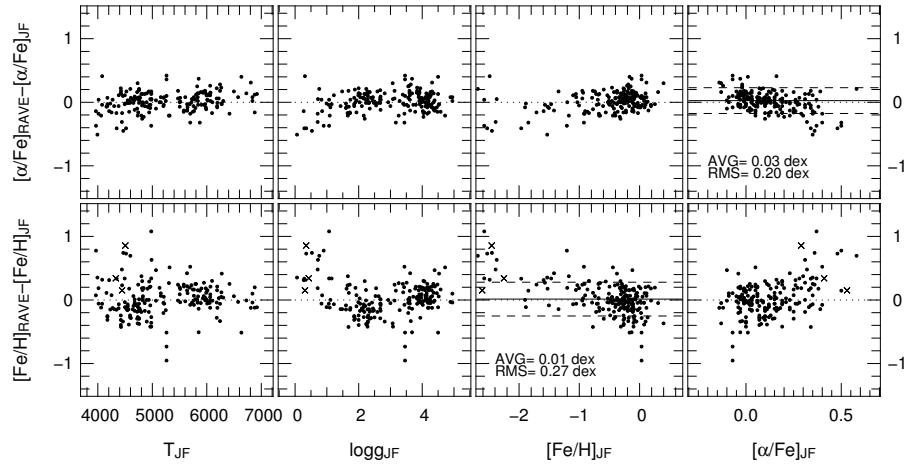

Figure 5.12: Abundance residuals RAVE-JF for Fe and α elements as function of the stellar parameters. Crosses are stars rejected because of continuum normalization problems.
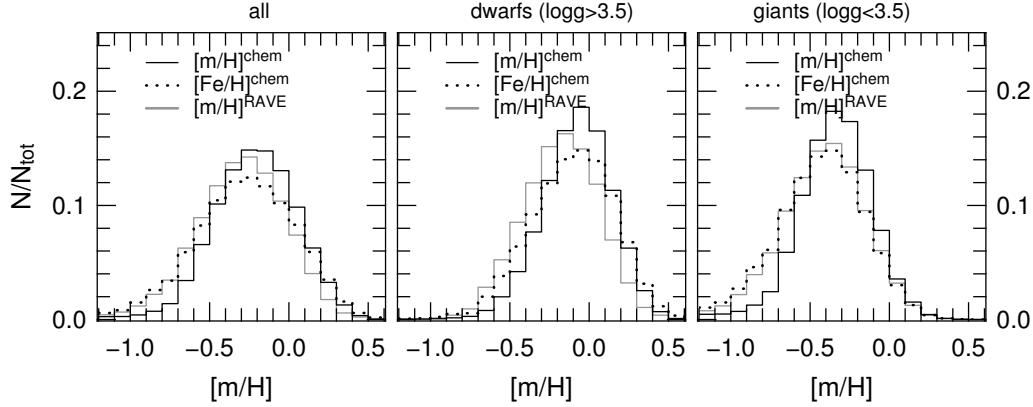
Figure 5.13: Comparison between $[m/H]^{chem}$ (black line) $[Fe/H]^{chem}$ (dotted line) and $[m/H]^{RAVE}$ (gray line) distributions for the whole sample (215422 stars, left panel), dwarf (middle panel) and giant stars (right panel).

5.3 $[m/H]^{RAVE}$ vs. $[m/H]^{chem}$: a comparison

In Figure 5.13 we compare the distributions of $[m/H]^{RAVE}$, $[m/H]^{chem}$ and $[Fe/H]^{chem}$ for 215422 RAVE stars with STN>20. The metallicity $[m/H]^{chem}$ is inferred from the chemical abundances with the formula given by Salaris et al. (1993)

$$[m/H]^{chem} = [Fe/H] + \log(0.638 \cdot 10^{[\alpha/Fe]} + 0.362) \qquad (5.1)$$

The α enhancement is computed as

$$[\alpha/Fe] = \frac{[Mg/H]+[Si/H]}{2} - [Fe/H] \qquad (5.2)$$

where the abundances $[Mg/H]$, $[Si/H]$ and $[Fe/H]$ come from the chemical pipeline.
The $[m/H]^{RAVE}$ is ~ 0.1 dex lower than $[m/H]^{chem}$ (a similar result has been found in Chapter 2 when we compared $[m/H]^{RAVE}$ with standard stars). The $[m/H]^{RAVE}$ distribution's shape is fairly similar to the $[m/H]^{chem}$ distribution for dwarf stars but different for giants. Moreover, $[m/H]^{RAVE}$ seems to match better $[Fe/H]^{chem}$ than $[m/H]^{chem}$, in particular for giants. This can be due to the α enhancement: among giants stars there are more thick disk stars (α enhanced) than among dwarfs stars which belong mostly to the thin disk (not α enhanced). Recalling that the RAVE pipeline is unable to measure the α elements (Zwitter et al., 2008), our conclusion is that $[m/H]^{RAVE}$ actually follow $[Fe/H]$, in particular for giant stars. Apart from the shift of 0.1 dex in metallicity already cited, $[m/H]^{RAVE}$ and $[m/H]^{chem}$ distributions appear in fair agreement.

5.4 Zero point of the RAVE abundance scale

We have seen in the previous chapters that the abundance measurements are indirect measurements, because they are inferred from the comparison between the lines' intensity seen in real spectra and the lines' intensity expected from our stellar atmosphere models. Since the models are different for different stellar parameters, we raise the questions of weather all the models yield abundances which refer to the same zero point (i.e., the origin of the internal abundance scale), and if this zero point refers to the same zero point of the real spectra (i.e., comparison between the internal and external scales). The latter question has a prompt answer: we do not have the external scale, since nobody has ever probed real stellar atmospheres, and all the abundance measurements refer to models. Therefore, we can only check the consistency of the internal scale. This can be done by comparing the measured abundances of a sample of synthetic spectra, as done in Fig. 5.1. In this plot, the points align along a straight line with slope roughly equal to one. This means that the measured differences in abundance between different spectra are the same as the differences expected, i.e. they refer to the same zero point. The linearity of the scale is confirmed by the distribution of the points, which align along a straight line. The offsets express the difference between the zero point of the measured abundances and the expected abundances due to the models. This difference is constant for any stellar parameter and can change for different elements.

We conclude that the abundance scale is consistent, which means that it is linear and has one zero point, which is common for any stellar parameter. Even if we cannot know the difference between the internal and external (real) zero points (which is merely a constant), we can safely state that the RAVE chemical pipeline correctly sorts stars as function of their chemical abundance.

5.5 Discussion and conclusions

In this chapter we studied the error sources that contribute to the chemical abundances uncertainties. The errors act together in non-linear relations, therefore it is not always possible to disentangle their effects. We must then estimate the errors globally.

The reliability of the abundance of each element depends on the number of lines strong enough to be visible in the noise. At the same time the intensity of the lines depend on the stellar parameters and abundances of the star. This means that for a fixed S/N, the precision of the element abundance is a function of the stellar parameters and the abundance itself. The interplay of these factors makes the accuracy estimate difficult.

At high S/N the precision varies from element to element because of the number and the intensity of the lines. For instance, at S/N=100 sulfur lines can be measured only in

stars with T_{eff} larger than $\sim$4800K. Chromium is measurable for T_{eff} lower than $\sim$5200K. Ca I has few weak lines which can be measured only in stars with [m/H] $>$–0.5 dex. And so on. Some elements like N, Co and Zr have very few estimations because of the weakness of their lines.

Everything we have just said applies even more to low S/N, because the pipeline measures even fewer lines, only the ones strong enough to be visible through the noise. As a consequence, the number of measured elements diminishes and uncertains increase. In the last section we tested the pipeline down to S/N=20 to see if at such a low S/N the measurements are still trustworthy. The result seems to suggest we can use such data, but with care. The noise generates selection effects: spectra having low [m/H] or high T_{eff} do not have lines strong enough to overcome the noise and they go through the pipeline unmeasured. The reduced number of visible lines can generate systematic errors because the abundance of an element depends on few measured lines (or only one) which depend on the precision of few the oscillator strengths (or only one). Further systematics can origin from the χ^2 fitting technique. When the noise is strong with respect to the intensity of the lines (i.e., low S/N), the minimization routine can be unable to find a minimum because the χ^2 hyper surface is nearly flat. For such a case the pipeline does not really measure the abundance but it merely render it equal to the first guess given at the input, that means $[\text{X/H}]=[\text{m/H}]^{\text{RAVE}}$.

Nonetheless, the tests performed with synthetic spectra at S/N=20 show an unexpected capability to estimate the abundances of the elements having strong lines like Fe, Mg, Al and Si. Errors are as large as $\sim$0.2-0.3 dex as expected at such a low S/N, but they do not show important systematics and the residuals expected-minus-measured are on average close to zero.

Our conclusion is that we can trust the abundances down to S/N=40 for 7 elements (details will follow in the next chapter) whereas between S/N=20 and 40 we can trust the abundance [Fe/H] and (to stay on the safe side) the abundance [α/H] which is the average of Mg and Si. Other uses of the abundances at low S/N has to be done only for exploration purposes until new comparison stars will be available at such S/N to estabilish the reliability of our results.

We summarize in the following the expected errors on abundances. This has to be considered as a general guideline.

The errors on abundances seem to be 0.1 dex larger in real spectra than in synthetic spectra. For real spectra we used abundances given in literature as reference abundances which are affected by errors as well. Therefore the standard deviation σ of the residuals between ours and the reference abundances is actually the quadratic sum of the errors due to the chemical pipeline and the errors of the reference abundances. In fact, looking at the

standard deviation of the [X/Fe] distributions in Figure 5.14 we see that the errors must be $\sim$0.15 dex for Mg and Si, and $\sim$0.25 dex for the others at S/N>40, very close to the errors estimated with synthetic spectra. This suggest that the errors obtained in tests with synthetic spectra are closer to the real abundances errors.

Thus, for real spectra we estimate errors of $\simeq$0.1-0.2 dex for S/N$\geq$40 and $\simeq$0.3 dex for 20$\leq$S/N<40. There is a general underestimation of the abundances due to the opacity of the neighboring lines and the overcorrection of the continuum for low-T_{eff} stars. Since most of the elements follow this systematic, the abundances relative to iron, [X/Fe], show weak or negligible systematic errors. Errors in stellar parameters generate systematic errors in abundances. The main driver is T_{eff}: elements like Ca, Ti and Fe are overestimated for too high T_{eff} whereas O and S follow the opposite trend. Gravity and metallicity errors do not yield important errors.

In the next chapter we will present the RAVE chemical catalogue and discuss the reliability and confidence intervals of the RAVE abundances element by element, based on the validation tests just discussed.
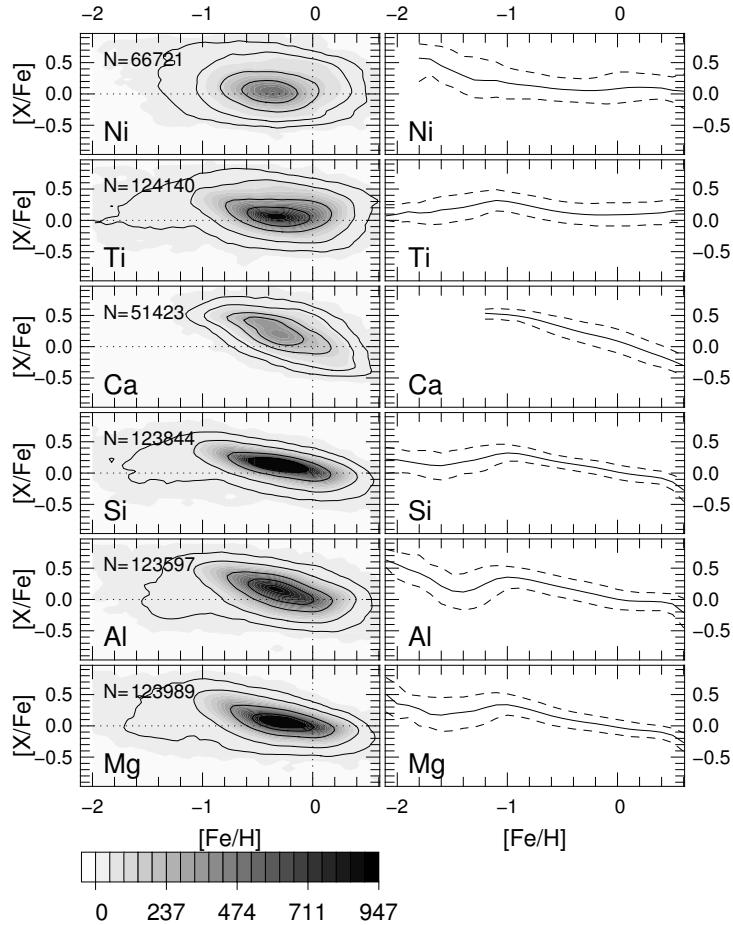
Figure 5.14: Relative abundance distributions of the 6 most important elements of the chemical catalogue with STN>40. Left column: the gray level code the density of the stars. The isocontours hold 34%, 68%, 95% and 99.5% of the stars. Right column: lines represent the mean (solid lines) and the standard deviation (dashed line) of the distributions.

Chapter 6

The RAVE chemical catalogue

We here present the RAVE chemical abundances catalogue, which holds chemical abundances for 217,358 RAVE stars.

6.1 Sample selection

The spectra have been selected from 295,618 RAVE spectra of the internal database by using the following constraints:

- **effective temperature $4000 \leq T_{\text{eff}}(\mathbf{K}) \leq 7000$**: this is the temperature range in which the RAVE line archive has been calibrated. At lower temperature the spectra are characterized by molecular lines other than CN (CH, TiO and other molecule), which are not considered in the line archive. We also avoided higher temperature since the spectra show lines of ionized atoms which are not included in the line archive.

- **signal-to-noise STN>20**: for STN<20 the absorption lines are strongly affected by the noise and stellar parameters and chemical abundances are not reliable.

- **rotational velocity $\mathbf{V}_{rot}$ <50 km s^{-1}**: at higher rotational velocities the lines show a FWHM larger than that due to the spectral resolution (FWHM$\simeq$1.3Å) and they cannot be precisely measured. Moreover, any spectra showing larger lines might be a double lined spectrum, which we want to avoid (we cannot measure abundances of a spectrum which is the result of an overlap of two different star spectra).

Despite this selection, some spectra show emission or double lines or are affected by bad continuum normalization. For such spectra the stellar parameters and chemical abundances are not reliable and they should not be used. Constraints on the parameters χ^2 and *frac*
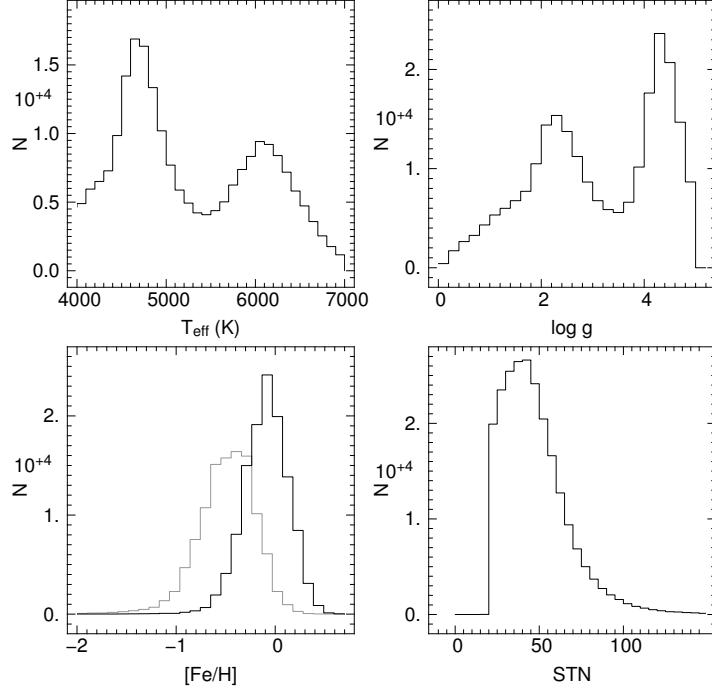
Figure 6.1: Distributions of stellar parameters $T_{\rm eff}$, $\log g$, [Fe/H] and signal-to-noise ratio STN for 215461 stars of the chemical catalogue. In the bottom left panel, the [Fe/H] distribution is given for dwarfs (black line) and giants (gray line), separately.

(= fraction of the spectrum rejected by the MASK code seen in Sec. 4.2.4) help to identify these spectra. For statistic studies we suggest to reject spectra with $\chi^2 > 3000$ and *frac* < 0.7. With these constraints the catalogue contains 215,461 stars. Among them, 11,067 stars had multiple observations. For these stars, the chemical abundances are computed as weighted average of the abundances measured in their spectra.

6.2   Stellar parameters

The distributions of the stellar parameters and STN are presented in Fig. 6.1. The quantities $T_{\rm eff}$ and $\log g$ are computed by the stellar parameters estimation procedure explained in Chapter 2 whereas STN and [Fe/H] have been estimated by the chemical pipeline. The distributions in $T_{\rm eff}$ and $\log g$ show two peaks. These peaks correspond respectively to giants at lower temperature and dwarfs, most of them at higher temperature. The iron abundance distribution peaks at [Fe/H]$\simeq$–0.1 dex for dwarfs and [Fe/H]$\simeq$–0.5 dex for giants.

6.3 Chemical abundances: selection effects due to S/N and [m/H]

The chemical pipeline can potentially measures the 13 elements N, O, Mg, Al, Si, S, Ca, Ti, Cr, Fe, Co, Ni, Zr, and the molecule CN. In practice only some of them are estimated in a spectrum because one or more of these elements have absorption lines too weak to be measured. The absorption lines can be indetectable because of the low [m/H] of the star and/or the low STN of the spectrum. STN and [m/H] together affect the accuracy of the abundances and the total number of abundance estimations.
This is illustrated in Fig 6.2 and Fig 6.3 where the fraction of spectra with [X/H] estimation diminishes with STN and [m/H] . In general, only metal-rich stars have abundances estimation at any STN whereas metal poor stars have estimation only if their spectra have high STN. Similarly, the single elements can or cannot be measurable depending on other variables such as the stellar parameters: titanium, for instance, is better measured in giant stars because the Ti lines in the RAVE wavelenght range have low excitation potential and are thus stronger at lower T_{eff}.

6.4 Accuracy and reliability element by element

In the following we shortly discuss and summarize the reliability of the chemical abundances for each element in the light of the knowledge gathered up to now. There are some general remarks about the measured abundances and their errors:

- the accuracy of the chemical abundances depends on several variables. In particular if T_{eff} is misestimated, the abundances are affected with different degree for different elements. In case of systematics in T_{eff} there will be systematics in [X/H] as well

- there is a general underestimation of [X/H] as function of T_{eff} : stars with $T_{\mathrm{eff}} <5000K$ yield on average underestimated abundances and their abundances are ~ 0.1 dex lower than stars with $>5000K$ whose average abundance is close to zero. This increases the errors when stars with different T_{eff} are analyzed. On the other hand, relative abundances [X/Fe] are nearly unaffected because the trend is similar for all elements

- the error given in the following refer to expected errors for two intervals of STNs at [m/H] ~ 0.0 dex. As already discussed in the previous section, errors can increase for lower STN and lower [m/H]

Nitrogen abundance is estimated from 3 weak lines visible only at $T_{\mathrm{eff}} >6000K$. It is underestimated and has large errors. Its use is not advisable.

Oxygen is a problematic element in our data; its abundance comes from the measurements of a triplet which, at RAVE resolution, is blended together with two Fe lines and two CN lines. This means that the local opacity is affected by the neighbor lines, leading to an underestimation of oxigen abundance. At T_{eff} >5000K it appears to be underestimated by ~0.2 dex and this value reduces to zero at T_{eff} ~4500K. The relative abundance [O/Fe] cannot be properly recovered. Because this strange behaviour its use is not advisable at this stage of the pipeline development[1].

Magnesium yields reliable results on synthetic and real spectra. At any T_{eff} we expect an abundance error σ_{Mg} ≤0.15 dex for STN≥40 and ~0.25 dex for 20≥STN≥40.

Aluminum abundances are obtained from the measurement of only two isolated lines (which are blended together at RAVE resolution). Despite the blending, they are intense and lead to a good precision in tests with synthetic and real spectra. It shows no significant offset. For STN≥40 we expect abundance errors σ_{Al} ≤0.2 dex and ~0.3 dex for 20≥STN≥40.

Silicon is the most reliable element together with Fe. We expect an abundance error σ_{Si} ≤0.15 dex for STN≥40 and ~0.20 dex for 20≥STN≥40.

The sulfur abundance precision has been tested only on synthetic spectra. Its abundance appears reliable only for stars with T_{eff} >5500K. We predict abundance errors of σ_{S} ~0.2 dex for STN≥40 and ~0.4 dex for 20≥STN≥40. Because the lack of verifications with real spectra, its use is not advisable.

The calcium abundance has been obtained from only five weak Ca I lines. It is better measured at higher metallicity and T_{eff} <5000K. Estimated errors are σ_{Ca} ~0.25 dex for STN≥40 and ~0.4 dex for 20≥STN≥40.

Titanium gives reliable abundances at high STN. At STN~40 its abundance appear reliable for T_{eff}<5000K and underestimated for higher temperatures. The correlation with T_{eff} errors is particularly strong (T_{eff} underestimation generates [Ti/H] underestimation and vice versa), leading to larger errors. We expect an abundance error of σ_{Ti} ~0.2 dex at STN≥40 and ~0.3 dex for 20≥STN≥40. We recommend to analyze stars with T_{eff} lower and higher than 5000K separately.

The chromium abundance precision has been tested only on synthetic spectra. It can be measured on spectra with T_{eff} <5000K and abundance high enough to make its lines

---

[1]We believe that it might be possible to properly measure [O/H] by considering the neighbor lines opacity. Investigation in this direction will be done in the near future.

measurable ([Cr/H]$\geq$-0.8). We expect an abundance error of σ_{Cr} $\sim$0.3 dex for STN$\geq$40 and $\sim$0.45 dex for 20$\geq$STN$\geq$40. Because the poor quality of its abundances and the lack of verifications with real spectra, its use is not advisable.

Iron is the most reliable element together with Si. It is measured on spectra with any T_{eff}. We expect an abundance error of σ_{Fe} $\sim$0.1 dex at STN$\geq$40 and $\sim$0.2 dex at 20$\geq$STN$\geq$40.

The cobalt abundance results from three weak lines and has poor precision. Its use is not advisable.

Nickel has six weak lines in the RAVE wavelength visible only at T_{eff} <5000K . In synthetic spectra tests with STN=100 it seems to yield errors comparable to other elements but with an underestimation of $\sim$0.2 dex which become $\sim$0.4 dex at STN=40. It is not measurable for [m/H] <-0.6dex. Tests on real spectra are not meaningful because performed at T_{eff}>5000K. We predict errors of σ_{Ni} $\sim$0.3 dex for STN$\geq$40 and $\sim$0.5 dex for 20$\geq$STN$\geq$40. Because of the poor quality of the Nickel abundances, it is advisable to use them with care.

The zirconium abundance results from one weak line and has a poor precision. Its use is not advisable.

The α enhancement [α/Fe] has been computed as average of [Mg/Fe] and [Si/Fe] (Formula 5.2). It is advised to use it in particular for 20$\geq$STN$\geq$40 on behalf of the single α elements because it is more reliable. We estimate errors of σ_{α} $\sim$0.1 dex for STN$\geq$40 and $\sim$0.2 dex for 20$\geq$STN$\geq$40.

6.5   The data

The RAVE chemical catalogue is provided as ASCII table of 232,134 lines. The version to be published will contain chemical abundances for the elements Mg, Al, Si, Ca, Ti, Fe and Ni, stellar parameters, signal-to-noise STN, object name and other parameters for quality checks as explained in Table 6.1.
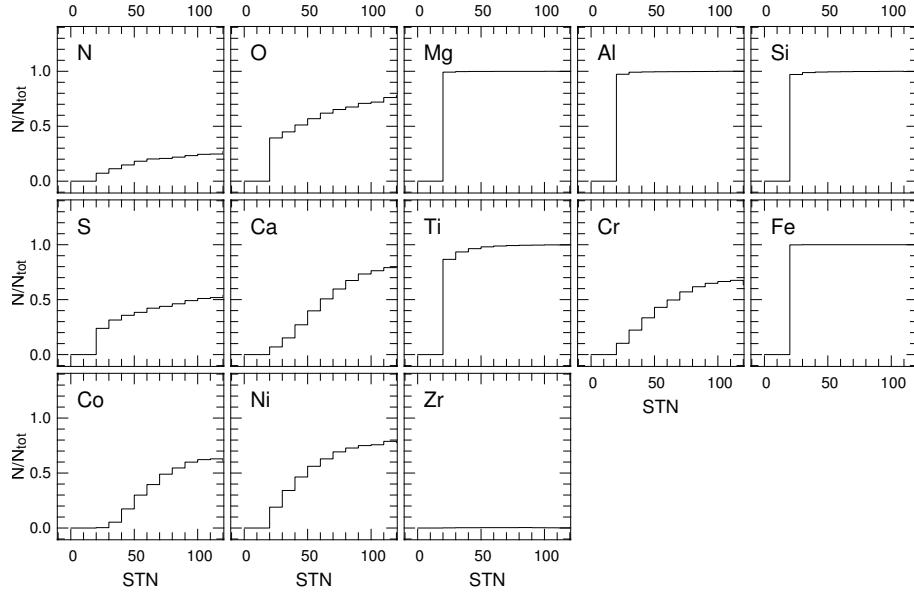
Figure 6.2: Fraction of spectra having abundance estimations (normalized to 1 for each bin) for the 13 elements measured as function of signal-to-noise STN.
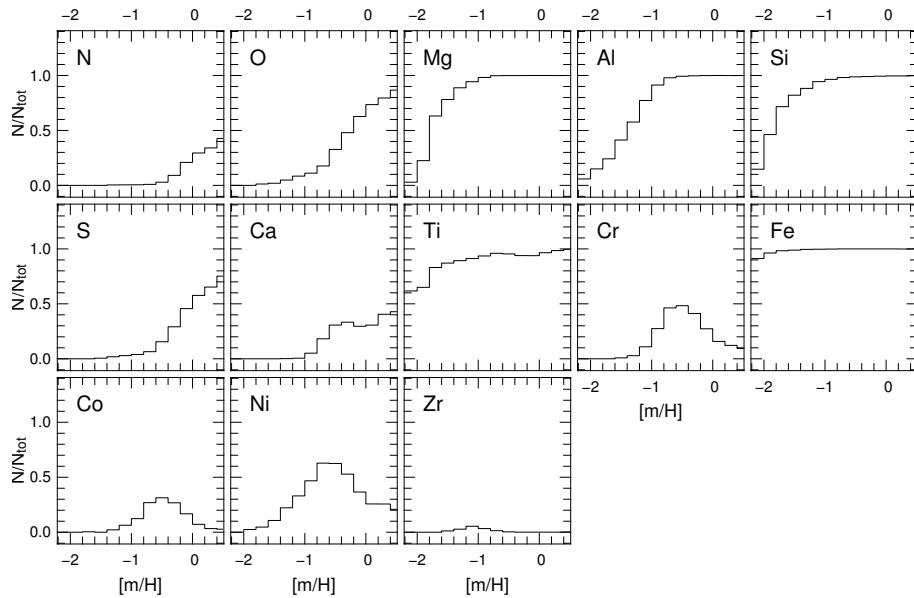

Figure 6.3: Fraction of spectra having abundance estimations (normalized to 1 for each bin) for the 13 elements measured as function of metallicity.

Table 6.1: Catalogue description

Field number	name	null value	description
1	Object ID	...	Identifier of the star
2	[Mg/H]	-9.99	Mg abundance
3	N	...	number of Mg lines measured
4	[Al/H]	-9.99	Al abundance
5	N	...	number of Al lines measured
6	[Si/H]	-9.99	Si abundance
7	N	...	number of Si lines measured
8	[Ca/H]	-9.99	Ca abundance
9	N	...	number of Ca lines measured
10	[Ti/H]	-9.99	Ti abundance
11	N	...	number of Ti lines measured
12	[Fe/H]	-9.99	Fe abundance
13	N	...	number of Fe lines measured
14	[Ni/H]	-9.99	Ni abundance
15	N	...	number of Ni lines measured
16	T_{eff}	...	RAVE effective temperature
17	$\log g$	...	RAVE gravity
18	$[\mathrm{m/H}]^{\mathrm{RAVE}}$	...	RAVE metallicity
19	$[\alpha/\mathrm{Fe}]^{\mathrm{RAVE}}$	...	RAVE $[\alpha/\mathrm{Fe}]^{\mathrm{chem}}$
20	$[\mathrm{m/H}]^{\mathrm{chem}}$	...	metallicity from the chemical pipeline
21	$[\alpha/\mathrm{Fe}]^{\mathrm{chem}}$	...	$[\alpha/\mathrm{Fe}]^{\mathrm{chem}}$ from the chemical pipeline
22	STN	...	signal-to-noise ratio
23	frac	...	fraction of spectrum well matching the template
24	Ntot	...	total number of lines measured
25	χ^2	...	χ^2 between observed and template spectra

Chapter 7

Chemical gradients in the Milky Way

7.1 Introduction

The spatial distribution of the chemical abundances in the Milky Way is an important constraint for our understanding of galaxy formation. The distribution of the chemical species traces the way galaxies formed and evolved, since we expect higher abundances where star formation has been more intense. After a star formation event, the new born stars synthesize heavy elements which are released into the intragalactic space through supernova (SN) events. The more SN events occur, the richer in heavy elements the interstellar matter becomes. Therefore we expect a difference (gradient) of chemical abundances between two locations which experienced different star formation history. Nowadays the existence of a negative chemical gradient along the Galactic radius seems well established, thanks to several dedicated observational studies we describe hereafter.

In order to map the chemical distribution over the Galaxy, one needs to observe objects whose distance can be estimated and which are bright enough to be visible at great distances. Such objects are Cepheids, planetary nebulae, open clusters, blue giants stars and H II regions. Since the number of works dedicated to the gradient measurement is large, we cite a few recent works and invite the reader to look at the references therein. By using Cepheids as tracer, Lemasle et al. (2008) estimated an iron gradient $\frac{d[Fe/H]}{dR_{\mathrm{G}}} = -0.023$ dex kpc^{-1} in the Galactic radius range R_{G}=8–15 kpc; Luck et al. (2006) found –0.068 dex kpc^{-1} in the range R_{G}=4.0–14.6 kpc; Andrievsky et al. (2002) obtained –0.029 dex kpc^{-1} in the range R_{G}=6–12 kpc. Other elements, when measured, follow the negative slope of [Fe/H]. Gradients obtained by using planetary nebulae gave $\frac{d[O/H]}{dR_{\mathrm{G}}} = -0.058$ dex kpc^{-1} (Maciel & Quireza, 1999). Works on open clusters lead to $\frac{d[Fe/H]}{dR_{\mathrm{G}}} = -0.06 \pm 0.01$ dex kpc^{-1} in the range R_{G}=7–16 kpc (Friel et al. 2002); Sestito et al. (2008) found –0.17±0.02 dex kpc^{-1} in the range R_{G}=6–11 kpc and point out that the Friel's sample lead to –0.09±0.02dex kpc^{-1} when only open clusters with R_{G} <11 kpc are used. It has been noted that the
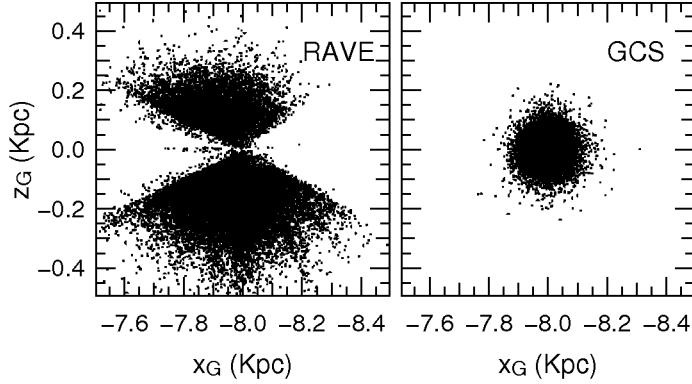
Figure 7.1: Space distribution of the selected 27634 RAVE stars (left) and 12087 stars of the Geneva-Copenhagen Survey (right) on the (x_G, z_G) Galactic plane.

gradients vary with the Galactic radius: Costa et al. (2004) found $\frac{d[O/H]}{dR_G} = -0.05$ dex kpc^{-1} in the range R_G=4–15 kpc which become –0.09dex kpc^{-1} at R_G=4-5 kpc and zero at $R_G >11$ kpc. The same gradient flattening beyond R_G=10–11 kpc has been highlighed by using Cepheids, open clusters and red giant star as well (Yong et al. 2006, Yong et al. 2005, Carney et al. 2005, Andrievsky et al. 2002). Generally speaking, there is a common agreement in finding a steeper gradient at small Galactic radius R_G and flatter at large R_G.
To explain the gradients observed in the Milky Way an inside-out formation of the Milky Way has been suggested (Matteucci & François, 1989) in which the inner parts of the Galaxy experienced a higher star formation rate than the outer parts. The inside-out formation scenario predicts gradients which are generally consistent with the observed ones (Chiappini et al. 1997, Cescutti et al., 2007) but are slightly flatter on average.

In the present work we investigate the chemical gradients of the elements Fe, Al, Mg, Si, Ti over the galactocentric range 5.5–9.5kpc by using the chemical and kinematic data of the RAVE data archive.

7.2   Data

We apply our analysis to two different samples of stars in order to test our investigation method and compare the results obtained from two independent sources. The first sample is composed of RAVE stars and the study of the chemical gradients is extended to 5 different elements. The second sample is selected from the Geneva-Copenhagen Survey (Nördstrom et al., 2004, hereafter GCS) which contains Hipparcos distances, space velocities and Fe abundances.

7.2.1 The RAVE sample

We selected the sample from the RAVE chemical catalogue. Proper motions, radial velocities and stellar atmospheric parameters T_{eff} and $\log g$ are taken from the RAVE data archive whereas the abundances [X/H] for Fe, Al, Mg, Si, Ti come from the RAVE chemical catalogue. Distances derived by Zwitter et al. (2010) enabled galactocentric positions and space velocities determinations. With such data we computed galactic orbits, apocentre (R_a) pericentre (R_p) of the orbits projected on the Galactic plane and maximum distance reached from it along its orbit by the star (Z_{max}). The orbits have been integrated by adopting the potential model n.2 by Dehnen & Binney (1998) which matches best the observed properties of the Galaxy. A trustworthy chemical gradient estimation must be based on reliable chemical abundances and distances. The RAVE chemical catalogue provides abundances with errors $\sigma \sim 0.2$dex for stars with spectra having signal-to noise ratio S/N$\geq$ 40 and $\sigma \sim$0.3dex at 20$\leq$S/N$\leq$40. We limited our investigation to : i) stars having spectra with S/N>40 ii) dwarf stars with gravity $\log g \geq$3.8dex and effective temperatures T_{eff} from 4000 to 7000K (limits given by the RAVE chemical catalogue) iii) stars which have distance uncertains smaller than 30%. This selection left 27634 stars whose spatial distribution is shown in Fig. 7.1. The distance from the Galactic plane $|z|_{\mathrm{G}}$ is smaller than 0.3kpc for most of the sample and it spans the Galactic radius interval R_{G}=7.6–8.3 kpc. Even if this is a small interval we can investigate the local gradient. In addition, we will use two other inferred radii: the mean radius R_{m} computed as $R_{\mathrm{m}} = (R_a + R_p)/2$ and the weighted radius R_{w}, computed as the average of the galactocentric distances $R(t)dt$ as function of time t weighted by the time spent at the radius interval $R(t + dt) - R(t)$

$$R_{\mathrm{w}} = \frac{\oint R dt}{\oint dt}.$$

Since for each star we integrated one complete orbit around the centre of the Galaxy with a constant time step, the radius R_{w} can be computed as follow

$$R_{\mathrm{w}} = \frac{\sum_i R_i \cdot w_i}{\sum_i w_i}$$

where R_i is the galactocentric distance of the star at the time step i and the weight w_i is expressed as

$$w_i = \frac{1}{|\Delta R_i|}$$

where $\Delta R_i = R_{i+1} - R_i$ is the length of the step in R. The radius R_{w} better represents the radius at which the star spends more time (where with high probability the star was born). All the cited radii refer to the galacticentric distance projected on the galactic plane because we want to study the gradients along it. By using R_{m} and R_{w} we can study the gradient at the Galactic radius interval R=5.5-9.5 kpc.
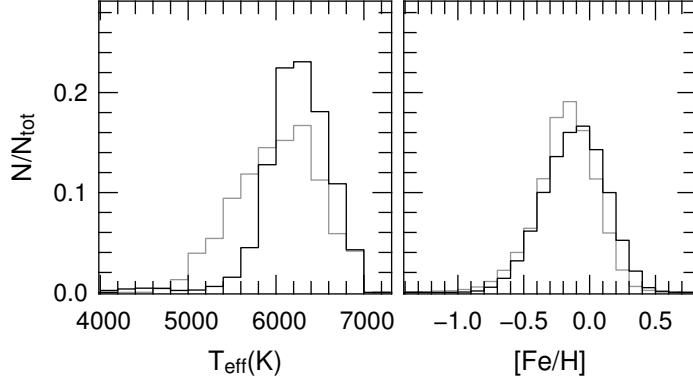
Figure 7.2: Effective temperature (left) and [Fe/H] (right) distributions of the RAVE (black line) and GCS (gray line) samples.

7.2.2   The Geneva-Copenhagen Survey sample

The GCS provides information such as the temperature, metallicity, kinematics, distance and age of 16682 F and G dwarf stars in the solar neighborhood (Nördstrom et al., 2004). The estimated errors are ∼0.07-0.12dex in [Fe/H] and 13% in distance. First, we analysed these data by using samples selected by age intervals following Nördstrom and colleagues, in order to compare their results with the ones obtained by our method. Then, we selected stars by using the same constraints in $T_{\rm eff}$ and galactocentric radii $R_{\rm G}$ used for RAVE stars, in order to keep the characteristics of GCS and RAVE samples as close as possible. With these constraints we are left with 12087 stars. Still, there are some differences. In fact, GCS stars are spherically distributed in the space and most of them do not lie farther than 0.2kpc from the Galactic plane (Fig. 7.1), whereas the RAVE stars have a cone shaped distribution in space and extends up to 0.5kpc from the Galactic plane. Since most of the RAVE sample is inside the scale height of the thin disk (0.3kpc), it is reasonable to assume that the difference in thick/thin disk star ratio between the two samples is negligible. In Fig. 7.2 we compare the $T_{\rm eff}$ and [Fe/H] distributions of the two samples. The GSC sample is richer of stars at $T_{\rm eff}$ ≤5500K with respect to the RAVE sample and its [Fe/H] distribution peak toward lower abundances with respect to the RAVE sample. Despite the highlighted differences, the GCS and RAVE samples show chemical gradients in fair agreement, as we will see later on.

7.3   Method and error estimation

The [Fe/H] distribution has an asymmetric shape (see Fig 7.2, right panel) which can be well fitted by a lognormal function whose mode changes as a function of R. We measure the abundance gradient by fitting the mode as function of R with a straight line. The

fitting procedure employs the maximum likelihood method. In the following we outline the details.

7.3.1   The fitting function

The lognormal probability distribution function is defined for $x > 0$ as

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)\right] \tag{7.1}$$

and $F(x)=0$ otherwise. The parameters σ and μ are the standard deviation and mean of the distribution, respectively. The mode is given by $M = exp(\mu - \sigma^2)$. However, we have to modify the function so that it can match the observed distribution. We invert the x axis as the tail and mode of the observed distribution is for positive x only. We use the change of variable

$$x = [Fe/H]_0 - [Fe/H]$$

where the parameter $[Fe/H]_0$ allows us to restrict x to $\Re^+$.

Because we are looking for a gradient of the mode, we choose to parametrize it by using a linear relation $M = \alpha R + M_0$, where α is the value of the gradient and M_0 the offset of the mode. This induces a relation between μ and σ which is given by

$$\mu = \ln\left(\alpha R + M_0\right) + \sigma^2.$$

Eq. 7.1 then transforms to

$$F(R, [Fe/H]) = \frac{1}{\sqrt{2\pi}\sigma([Fe/H]_0 - [Fe/H])}\cdot$$

$$\cdot \exp\left[-\frac{1}{2}\left(\frac{\ln\left([Fe/H]_0 - [Fe/H]\right) - \ln\left(\alpha R + M_0\right) - \sigma^2}{\sigma}\right)\right],$$

and, for a fixed radius R, the normalization is

$$\int F_R([Fe/H]|[Fe/H]_0, \alpha, \sigma, M_0)d[Fe/H] = 1.$$

The sample being fixed, the complete probability density can be normalised by dividing it by the number of stars so that its integral over the observed space is 1 (i.e., it is a probability), making this relation suitable for a maximum likelihood method.

The mode of the resulting distribution is given by

$$M = [Fe/H]_0 - M_0 - \alpha R.$$

With this relation, the gradient of the mode is $-\alpha$ and the offset is $[Fe/H]_0 - M_0$.

The capabilities of such a function to reproduce an observed distribution are given in the right panels of Fig. 7.3 and Fig. 7.4 where the [Fe/H] distribution of the points is presented together with a match of the modified lognormal distribution of Eq. 7.1. The function can fit properly the distributions of the GCS and RAVE samples.

The model parameters which we are trying to find are $([Fe/H]_0, \alpha, \sigma, M_0)$ given the sample defined by $[Fe/H]_i$ and R_i where i is a running index on the sample. The probability of a star having [Fe/H] and R given the parameters, with $P_j = ([Fe/H]_0, \alpha, \sigma, M_0)$ the parameters for the model j, is then

$$\mathcal{P}_i([Fe/H]_i, R_i | P_j) = \frac{1}{N} \int_\varepsilon F([Fe/H]_i, R_i | P_j) d[Fe/H] dR,$$

where the integral is a convolution of the observational errors and N is a normalisation factor (the integral of F over the volume in [Fe/H] and R considered). Assuming the measurements are perfect (eg. we drop the convolution), the likelihood can be written

$$\mathcal{L} = \Pi_i \mathcal{P}'_{i|j} = \Pi_i \frac{1}{N} F([Fe/H], R | \mathbf{P}_j),$$

where i refers to a given star in the sample and j to the model considered. Maximizing the likelihood is equivalent to minimizing the quantity χ^2 given by

$$\chi^2(i) = -2. \log \mathcal{L} = -2. \sum_i \log \mathcal{P}'_{i|j}.$$

The quantity $\chi^2(i)$ follows a chi square law (hence the notation). Given the sample and the derivatives of the χ^2 function with respect to $\mathcal{P}_j$, the only thing left to do is to employ a minimization routine to find the best parameters in the maximum likelihood sense.

7.3.2   Error estimation

The confidence intervals of the parameters $([Fe/H]_0, \alpha, \sigma, M_0)$ which minimize the χ^2 can be evaluated by using contours of constant $\Delta\chi^2$ as boundaries of the confidence region (as reported in Numerical Recipes, Press et al. 1992). The application of this method gives the internal errors, which are found to be $\simeq 0.0001 \text{dex kpc}^{-1}$ or smaller. Nonetheless, we noted that by removing even only one point of the sample, the maximum likelihood method can give a slope $-\alpha$ which can differ by up to 0.01dex with respect to the full sample. This suggests that external errors are $\simeq 10$ times larger than the internal ones. For this reason, we evaluated the confidence intervals by using the bootstrap resampling method. The upper and lower limits of the 68% confidence interval obtained by the bootstrap method are reported in the following as subscripts and superscripts of the gradients.

7.4   Analysis and results

7.4.1   Gradients from the RAVE sample

We applied the maximum likelihood method just described to the RAVE sample by using their present galactocentric distances R_{G}, R_{m} and R_{w} for the elements Fe, Al, Mg, Si, Ti. We present here only iron, with the other elements discussed in Sec. 7.5.

The resulting gradient for iron at the present position of the stars is $\mathrm{d}[\mathrm{Fe/H}]/\mathrm{d}R_{\mathrm{G}} = -0.057^{+0.013}_{-0.009}$ dex kpc^{-1}. As we said before, R_G may not be representative of the distance at which the stars were born, since their orbits may not be circular. The mean radii R_{m} and R_{w} better represents (probabilistically speaking) the galactocentric distance of the birth place of the star. By using R_{m} and R_{w} as galactocentric distances and the constraints $5.5 < R_m(\mathrm{kpc}) < 9.5$ and $6.0 < R_w(\mathrm{kpc}) < 10.0$ respectively (to avoid the few outlier stars which affect the computation) we obtained the results

$\frac{d[Fe/H]}{dR_m} = -0.070^{+0.004}_{-0.005}$ dex kpc^{-1}

$\frac{d[Fe/H]}{dR_w} = -0.045^{+0.005}_{-0.005}$ dex kpc^{-1}

The difference will be discussed in Sect. 7.5. Interestingly, we obtained different gradients when we selected subsamples as function of Z_{max}. We divided the full sample in three subsamples with different Z_{max} ranges: $0.0 < Z_{\mathrm{max}}$ (kpc)≤ 0.4 (22346 stars, named $Z^{\mathrm{RAVE}}_{0.0}$ sample) $0.4 < Z_{\mathrm{max}}$ (kpc)≤ 0.8 (4142 stars, the $Z^{\mathrm{RAVE}}_{0.4}$ sample) and Z_{max} (kpc)> 0.8 (786 stars, the $Z^{\mathrm{RAVE}}_{0.8}$ sample). For these samples we found

$\frac{d[Fe/H]}{dR_m}(Z^{\mathrm{RAVE}}_{0.0}) = -0.078^{-0.002}_{+0.002}$ dex kpc^{-1}

$\frac{d[Fe/H]}{dR_m}(Z^{\mathrm{RAVE}}_{0.4}) = -0.063^{-0.005}_{+0.005}$ dex kpc^{-1}

$\frac{d[Fe/H]}{dR_m}(Z^{\mathrm{RAVE}}_{0.8}) = 0.000^{-0.009}_{+0.008}$ dex kpc^{-1}

In Fig. 7.3 the gradients are represented by the slope of a straight line (left panels) and the Fe abundance distributions are approximated by the best fitting lognormal function (right panels). We applied the same technique to estimate other element gradients as well as the gradients of the abundances relative to iron by using the radii R_{m} and R_{w}. All the results are outlined in Tab. 7.1 and Tab. 7.2.

7.4.2   Gradients from the Geneva-Copenhagen survey sample

To evaluate the robustness of our results we repeat the same procedure by using the GCS data. First, we verify that our method gives results in agreement with the results derived by Nordström. By selecting stars by ages, the resulting gradients are

$\frac{d[Fe/H]}{dR_m}(\mathrm{age} < 1.5\mathrm{Gyr}) = -0.074^{+0.018}_{-0.005}$ kpc^{-1},
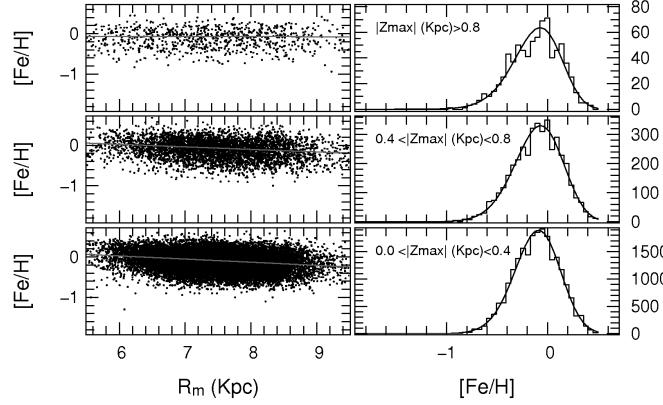
Figure 7.3: **Left**: distribution of 27279 RAVE dwarf stars in the ([Fe/H],R_{m}) plane for three Z_{max} ranges. The gray line describes the mode obtained with the maximum Likelihood method. **Right**: [Fe/H] distributions of the stars (histogram) and their best fitting lognormal functions.

$$\frac{d[Fe/H]}{dR_m}(4<\text{age}<6\text{Gyr}) = -0.084^{+0.012}_{-0.010}\text{dex kpc}^{-1},$$

$$\frac{d[Fe/H]}{dR_m}(\text{age}>10\text{Gyr}) = +0.049^{+0.013}_{-0.011}\text{dex kpc}^{-1},$$

consistent with the results of Nordström.
By using the actual galactocentric radius R_{G} of all the stars with distance estimations and considering only $7.8< R_G <8.2$ kpc we are left with 12641 stars which give a gradient of $d[Fe/H]/dR_{\mathrm{G}} = -0.086^{+0.016}_{-0.003}$ dex kpc^{-1}. By using R_{m} the gradient is $d[Fe/H]/dR_{\mathrm{m}} = -0.023^{+0.004}_{-0.004}$ dex kpc^{-1}, flatter than what we found with the RAVE data.

By following the procedure used for the RAVE sample, we selected subsamples in three different Z_{max} ranges: $0.0 < Z_{\mathrm{max}}$ (kpc)< 0.4 (10884 stars, named the $Z_{0.0}^{\mathrm{GCS}}$ sample) $0.4 < Z_{\mathrm{max}}$ (kpc)< 0.8 (1234 stars, the $Z_{0.4}^{\mathrm{GCS}}$ sample) and Z_{max} (kpc)> 0.8 (395 stars, the $Z_{0.8}^{\mathrm{GCS}}$ sample). These stars satisfy the same constraints in T_{eff} and R_{m} used for the RAVE samples. The resulting gradients are
$\frac{d[Fe/H]}{dR_m}(Z_{0.0}^{\mathrm{GCS}}) = -0.041^{+0.004}_{-0.004}$ dex kpc^{-1},
$\frac{d[Fe/H]}{dR_m}(Z_{0.4}^{\mathrm{GCS}}) = -0.016^{+0.010}_{-0.010}$ dex kpc^{-1},
$\frac{d[Fe/H]}{dR_m}(Z_{0.8}^{\mathrm{GCS}}) = +0.074^{+0.016}_{-0.017}$ dex kpc^{-1}.
The corresponding gradients and [Fe/H] distributions are plotted in Fig. 7.4.

7.4.3   Binary contamination

In using quantities inferred from the orbits such as Z_{max} or R_{m} we must pay attention to the fact that many stars can be undetected binaries. In this case the measured velocity is
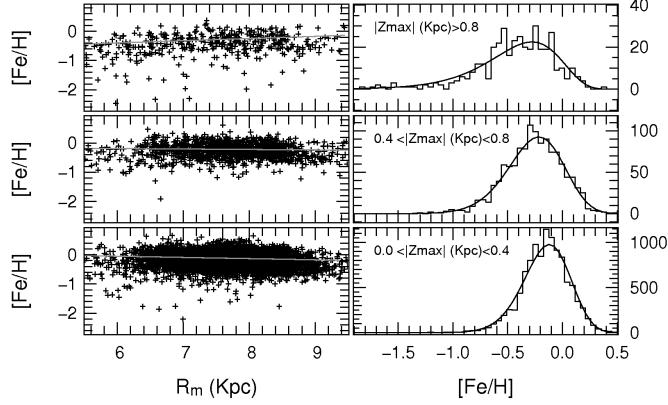
Figure 7.4: **Left**: distribution of 11379 dwarf stars from the Geneva-Copenhagen survey in the ($[M/H]$,R_m) plane for three Z_{max} ranges. The linear regression line (gray line) describes the mode of the distributions. **Right**: [Fe/H] distributions of the stars (histogram) and their best fitting lognormal functions.

affected by the binary orbital velocity and the derived galactic orbit might no longer be reliable. On the other hand, we expect that the binary orbital velocities would change the resulting Z_{max} and R_m randomly and this would add noise to the gradient determination. In the case of the GCS sample, Seabroke et al. (2008) showed that the binary contamination does not significantly affect the inferred velocity distribution of the stars. However the metallicity might be affected because of the overlap of the two spectra. Therefore, in order to test the effect of the binary contamination we performed the gradients measurement by using first the whole sample and then repeated the measurements by removing the flagged binaries. The two resulting gradients changed on the order of 0.001 dex kpc^{-1}, which indicates that a possible pollution in [Fe/H] is negligible.

Can we say the same for the RAVE sample? Seabroke et al. suggest that the RAVE sample can be used on an equal footing with the GCS sample. We want to point out that the binaries radial velocities reported by Nordström are the velocities of the centre-of-mass computed by the method of Wilson (1941). This was possible thanks to the multiple observations of the GCS. Thus, the computed heliocentric U, V and W velocities are closer to the centre-of-mass velocities than the ones that can be obtained from a single observation. In the case of RAVE this is not possible because most of the stars are observed just once. Therefore we would expect a velocity distribution affected (widened) by binarity. Since it is impossible to recognize a single-lined spectroscopic binary with only one observation there is no way to remove them from our sample. However, the consistent results obtained by the two samples strongly suggest a negligible effect of the binaries in the gradient determination.

7.5   Discussion

Independent data sets such as RAVE and GCS give similar results. This suggests they are reliable. The weighted radius R_w gives less negative gradients with respect to R_m, because R_w is in general larger that R_m. For circular orbits they are equal but for eccentric orbits the star spend more time close to the apocentre and R_w become larger. In the $(R,[\mathrm{X/H}])$ plane (Fig. 7.3 left panels, for instance) the stars with eccentric orbits move toward higher R at constant $[\mathrm{X/H}]$, making the gradient less negative. Nonetheless, the trend is the same for both radii: the larger Z_max the flatter the gradient.

7.5.1 Z_max <0.4 kpc

In the Galactic radius interval 5.5-9.5 kpc, the stars with Z_max <0.4 kpc exhibit [Fe/H] gradients which are in agreement with the literature works cited in Sec. 7.1. The chemical models predict a range of [Fe/H] gradients, from $\simeq$–0.04dex/kpc (Chiappini et al., 1997, Cescutti et al., 2007), to a steeper $\simeq$–0.09dex/kpc (Schönrich et al. 2009). Our work seems to favour the steeper gradients, with $\mathrm{d[Fe/H]}/\mathrm{d}R_\mathrm{m} = -0.078\mathrm{dex}\ \mathrm{kpc}^{-1}$. The gradients of the other elements do not significantly differ from Fe. Negative gradients suggest a higher star formation rate at smaller galactocentric radii.

7.5.2   0.4< Z_max <0.8 kpc

Gradients and average abundances of this sample lie in between the $Z_{0.0}^\mathrm{RAVE}$ and $Z_{0.8}^\mathrm{RAVE}$ samples, with a progression from negative to more positive gradients and from metal rich to metal poor stars when Z_max increases. This sample appears to be a mixture of the other two Z_max ranges, suggesting a progressive and continuous transition between them.

7.5.3 Z_max >0.8 kpc

The $Z_{0.8}^\mathrm{RAVE}$ sample shows flatter abundance gradients with respect to the $Z_{0.0}^\mathrm{RAVE}$ sample for all the five elements considered. Do these stars belong to different populations? Comparing the two samples in Fig. 7.5 one would give a positive answer: the $Z_{0.0}^\mathrm{RAVE}$ sample has higher V velocities and lower eccentricities as one would expect from stars which belong to the thin disk. On the other hand the $Z_{0.8}^\mathrm{RAVE}$ sample has lower V velocities and higher eccentricities, typical of thick disk stars. Since the samples are composed of stars which lie close to the Galactic plane, constraints in Z_max correspond roughly to constraints in W velocity. In fact, 79% of the $Z_{0.0}^\mathrm{RAVE}$ stars lie at $|\mathrm{W}|$ <15 km s^{-1} , 85% of $Z_{0.4}^\mathrm{RAVE}$ lie at 15< $|W|$ <40 km s^{-1} and 91% of $Z_{0.8}^\mathrm{RAVE}$ stars lie at $|\mathrm{W}|$ >40 km s^{-1}. By using the Schwarzschild ellipsoid velocity distribution we compute the expected fraction of thick disk (TD) stars respect to
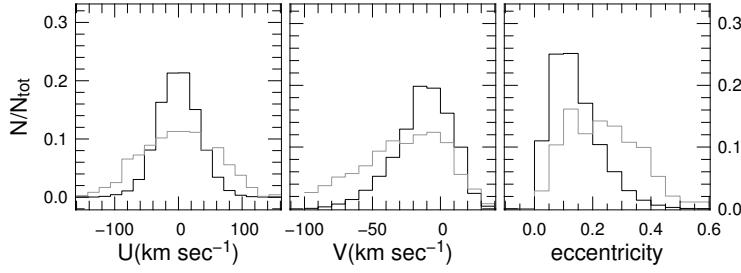
Figure 7.5: Distribution of stars with $Z_{\max}$ <0.4 kpc (black line) and $Z_{\max}$ >0.8 kpc (gray line) in U, V velocities and eccentricity.

the thin disk (D) stars. Since we use only the W velocity we restrict the ellipsoid to one dimension. The number of stars in the velocity interval dW is

$$\rho(W, \sigma_W)dW = \frac{N}{\sqrt{2\pi}\sigma_W} \exp\left(-\frac{W^2}{2\sigma_W^2}\right)dW$$

where ρ is the stars density in the velocity space, N is the total number of stars and σ_W is the velocity dispersion of the population considered with respect to the local standard of rest. Hence, the number of TD stars in the $Z_{0.8}^{\mathrm{RAVE}}$ sample is

$$N_{TD} = f_{TD} \cdot \int_{40}^{\infty} \rho(W, \sigma_W^{TD})dW$$

where f_{TD} is the fraction of TD stars with respect to the whole sample, and the number of D stars is

$$N_D = f_D \cdot \int_{40}^{\infty} \rho(W, \sigma_W^D)dW$$

where f_D is the fraction of D stars with respect to the whole sample. Adopting the velocity dispersions σ_W^D=20 km s^{-1} and σ_W^{TD}=39 km s^{-1} (from Soubiran et al. 2003), the fraction f_{TD}=0.1 (from Jurić et al. 2008) and assuming f_D=1-f_{TD}, the ratio between the number of thin and thick disk stars in the $Z_{0.8}^{\mathrm{RAVE}}$ sample is N_{TD}/N_D=0.74. Here we made the assumption that f_{TD} is constant, as our sample lies close to the Galactic plane. Arguing that the space distribution of the RAVE stars is cone shaped and many stars lie at z_G >0.1kpc (but not beyond 0.5kpc), f_{TD} might be a little higher. By using a more generous f_{TD}=0.15 given by Soubiran et al. (2003) we obtain N_{TD}/N_D=1.18. Kinematically speaking, the $Z_{0.8}^{\mathrm{RAVE}}$ sample is composed by an important fraction of thick disk stars (40-50% of it). One of the discriminating characterictics between the thin and thick disk stars is that the latter are metal poorer and α elements enhanced. Still, by looking at Fig. 7.6 we cannot recognize these features and the $Z_{0.0}^{\mathrm{RAVE}}$ and $Z_{0.8}^{\mathrm{RAVE}}$ samples appear chemically indistinguishable. Nonetheless, these two samples differ in space and chemical distributions since their chemical gradients are not the same.

The information collected up to now does not fit easily in the kinematical and chemical scenario of the Milky Way formation we are used to. We make here two "extreme" hypotheses on the origin of the $Z_{0.8}^{\text{RAVE}}$ sample to try to clarify the picture: 1) this sample is composed only by thick disk stars or 2) this sample is composed only by thin disk stars kinematically heated up. In the first case they are not chemically representative of the thick disk and we would have a large sample of stars having thick disk kinematics and thin disk abundances, a sort of "transition stars" between the thick and thin disk also found by Mishenina et al. (2003), Bensby et al. (2003) and Reddy et al. (2006). The latter found that these stars appear to be confined to a specific part of the velocity space, in particular they seem to prefer positive U velocities: this is not confirmed by our data since the $Z_{0.8}^{\text{RAVE}}$ sample stars cover symmetrically the U velocity space (Fig. 7.5, left panel). High resolution spectroscopic observations of these stars would help to better define their chemical abundances and support or disprove the existence of a "transition population". The flat gradients $d[X/H]/dR_{\text{m}}$ they exibit would suggest that i) the stars formed in a homogeneous medium or ii) the stars experienced an efficent radial mixing (Schönrich & Binney, 2009) and are so well mixed in radius that the gradient has disappeared. In the first case they would have a common origin: if they belong to the thick disk than their age distribution would favour older ages. Since we do not have precise age estimation for RAVE stars, we look at the age distribution of the $Z_{0.8}^{\text{GCS}}$ sample in comparison to the $Z_{0.0}^{\text{GCS}}$ sample: the latter favour stars younger than 4Gyr whereas the former exibit a much flatter distribution in the range 2–15Gyr. This defies the hypothesis they have formed from a homogeneous medium, since it is hard to imagine an interstellar medium where stars form and have the same chemical composition for such a long time. In the second case (efficient radial mixing) the memory of their origin would be lost and they would constitute a random sample. The second hypothesis is that the $Z_{0.8}^{\text{RAVE}}$ sample is composed of thin disk stars kinematically heated up to appear as thick disk stars. The investigation on the GCS by Nordström at al. (2004) and Holmberg et al. (2007) showed that the older the stars the larger the Z_{max} they can reach, due to the kinematic heating of the disk with time. If this is the case, one must explain why these old stars exibit such a high abundances and no α-element enhancement. The two hypotheses made so far may go together: the $Z_{0.8}^{\text{RAVE}}$ sample might be composed by thick disk stars with thin disk abundances (the "transition stars") and by thin disk stars kinematically heated up to be kinematically mistaken by thick disk stars (in this case the radial mixing can play and important role) which contribute to to the high abundances and flatten the gradients.

7.6   Conclusions

We have analysed 27634 stars selected from the RAVE chemical catalogue and 12087 stars selected from the Geneva-Copenhagen survey (GCS) making use of chemical and kinematic information such as chemical abundances, absolute velocities and orbit parameters to measure chemical gradients $d[X/H]/dR_{\text{m}}$ of the elements Fe, Mg, Al, Si, Ti along the
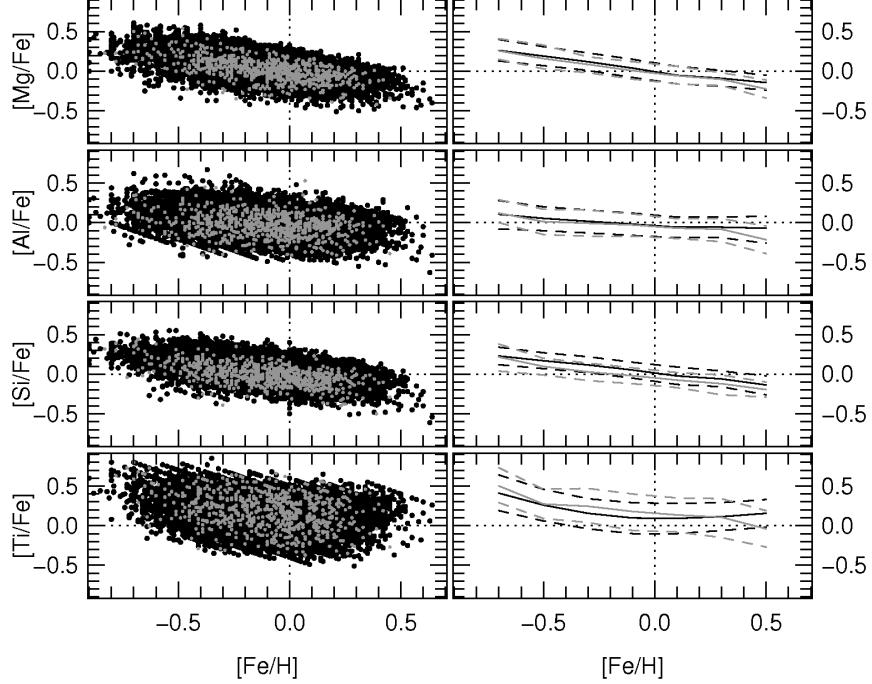
Figure 7.6: **Left**: abundances of the element Mg, Al, Si, Ti relative to iron. Black points represent stars of the $Z_{0.0}$ sample, while gray points represent stars of the $Z_{0.8}$ sample. **Right**: average (solid lines) and 1σ limits of the distribution (dashed lines) for the two samples.

Galactic radius in the galactocentric distance interval 5.5-9.5kpc. The stars of each sample have been divided into three subsamples as a function of the maximum distance from the Galactic plane reached along their orbit Z_{max}: the $Z_{0.0}$ sample is composed of stars with Z_{max} <0.4kpc, the $Z_{0.4}$ sample is composed of stars with 0.4< Z_{max} ≤0.8kpc and the $Z_{0.8}$ sample is composed of stars with Z_{max} ≥0.8kpc. We found that for GCS and RAVE $Z_{0.0}$ samples the gradients are d[Fe/H]/dR_m ∼ −0.04dex kpc^{-1} and d[Fe/H]/dR_m ∼ −0.08dex kpc^{-1}, respectively, in general agreement with chemical models and other observational works in the literature. One of our main findings is that the gradients d[X/H]/dR_m are flatter at higher Z_{max}. This behaviour might be explained by assuming that the $Z_{0.8}^{RAVE}$ sample is composed by a thin disk population kinematically heated up to assume a thick disk kinematics and has undergone an efficient radial mixing process which blurred (and hence flatten) a possible gradient, or by a "transition population" composed of thick disk stars which have thin disk chemical abundances, previously detected by some autors (Mishenina et al. 2003, Bensby et al. 2003 and Reddy et al. 2006). The two hypotheses do not exclude each other. Further high resolution spectroscopic observations would help to constrain the abundances and ages of these stars and clarify the origin of such extreme high velocity thin disk stars and extreme chemically abundant thick disk stars.

	$\frac{d[X/H]}{dR_m}$				
$Z_{\max}$ range (kpc)	Fe	Mg	Al	Si	Ti
$0.0 \leq Z_{\max} < 0.4$	$-0.078^{+0.002}_{-0.002}$	$-0.063^{+0.011}_{-0.010}$	$-0.086^{+0.002}_{-0.002}$	$-0.063^{+0.013}_{-0.001}$	$-0.071^{+0.003}_{-0.003}$
$0.4 \leq Z_{\max} < 0.8$	$-0.063^{+0.005}_{-0.005}$	$-0.052^{+0.004}_{-0.004}$	$-0.067^{+0.005}_{-0.005}$	$-0.049^{+0.004}_{-0.004}$	$-0.054^{+0.005}_{-0.005}$
$Z_{\max} > 0.8$	$+0.000^{+0.008}_{-0.009}$	$-0.013^{+0.007}_{-0.007}$	$-0.011^{+0.010}_{-0.010}$	$-0.002^{+0.007}_{-0.007}$	$-0.027^{+0.011}_{-0.010}$

Table 7.1: Abundance gradients expressed as dex kpc^{-1} for three ranges of $Z_{\max}$. R_{m} is used as galactocentric radius. Intervals of 68% confidence obtained with the bootstrap method are reported as sub and superscrips.

	$\frac{d[X/H]}{dR_w}$				
$Z_{\max}$ range (kpc)	Fe	Mg	Al	Si	Ti
$0.0 \leq Z_{\max} < 0.4$	$-0.051^{+0.006}_{-0.007}$	$-0.069^{+0.009}_{-0.003}$	$-0.061^{+0.008}_{-0.006}$	$-0.060^{+0.005}_{-0.005}$	$-0.049^{+0.004}_{-0.005}$
$0.4 \leq Z_{\max} < 0.8$	$-0.040^{+0.015}_{-0.012}$	$-0.039^{+0.012}_{-0.015}$	$-0.043^{+0.015}_{-0.011}$	$-0.041^{+0.012}_{-0.009}$	$-0.038^{+0.005}_{-0.005}$
$Z_{\max} > 0.8$	$+0.008^{+0.024}_{-0.020}$	$-0.003^{+0.021}_{-0.023}$	$-0.004^{+0.023}_{-0.027}$	$+0.002^{+0.018}_{-0.014}$	$-0.010^{+0.010}_{-0.009}$

Table 7.2: Abundance gradients expressed as dex kpc^{-1} for three ranges of $Z_{\max}$. R_{w} is used as galactocentric radius. Intervals of 68% confidence obtained with the bootstrap method are reported as sub and superscrips.

Summary and outlook

In these work we developed an automatized unsupervised processing pipeline, which efficiently measures chemical abundances from the RAVE spectra, and produced the *largest chemical abundances catalogue existing today*.
In order to obtain robust chemical abundance measurements we checked and refined the oscillator strengths of the absorption lines used for the abundance determinations through an inverse spectral analysis. We also modified the RAVE pipeline to remove some systematic errors in the RAVE stellar parameters (discovered during this work) which could lead to systematic errors in chemical abundances. To measure the chemical abundances, we developed two different processing pipelines, both assuming stellar atmospheres in local thermodynamic equilibrium. The first pipeline measures the equivalent width of the absorption lines and computes the chemical abundances by using a line analysis software. The second pipeline uses the χ^2 minimization technique between observed and template spectra. The second pipeline performs better than the first one in precision and speed, and it has been adopted for the creation of the RAVE chemical catalogue. Despite the pipeline measures the abundances of 13 different atomic species and one molecule, some of them (chromium, cobaltium, zirconium) are not reliable. Some other (nitrogen, oxigen, sulfur and the CN molecule) are promising, but we decided to delay their publication until more tests and comparisons with high precision results are available to prove their reliability. Therefore, the present catalogue contains the chemical abundances up to 7 elements (Mg, Al, Si, Ca, Ti, Fe and Ni) for 217,358 RAVE stars. The expected abundances uncertainties are of about ~ 0.2dex for spectra with signal-to-noise ratio S/N$\geq$40 and ~ 0.3dex for spectra with 20$\leq$S/N$<$40.
We employed the catalogue to measure the chemical gradients along the Galactic radius. We found that stars with low vertical velocities |W| (which stay close to the Galactic plane) show abundance gradients in agreement with previous works (~ -0.07 dex Kpc^{-1}), whereas stars with larger |W|, which are able to reach larger heights above the Galactic plane, show progressively flatter gradients, suggesting that an efficient radial mixing acts in the Galaxy or that the thick disk formed from homogeneous interstellar matter.

Thanks to its measurements homogeneity and unbiased pre-selection, the RAVE chemical catalogue is suitable for statistical investigations. Proper motions, radial velocities and distance estimates locate the stars in the 6 dimensional phase-space, while the chemical

abundances locate them in the chemical space. The catalogue is therefore a rich mine of information to:

- identify groups of stars sharing similar chemical characteristics in order to discover disrupted open clusters (star formation history, galactic dynamic)

- identify stars with peculiar abundances belonging to rare class of stars (stellar evolution)

- identify of moving groups created by gravitational actions of Galactic spiral arms and/or bar (galactic dynamics)

- disentangle thin from thick disk stars, study the properties of such populations and search for new characteristics (Galactic formation and evolution)

- verify theoretical predictions of galaxy chemical models such as the abundances at different Galactic radii, and vertical and radial gradients for different elements, in order to give new constraints to the models (Galactic formation and evolution)

- statistical studies on the abundance distributions compared with the models to give constraints to the theoretical SNe yield predictions (stellar evolution)

- verify the relations between chemical abundances and Galactic orbital parameters (such as mean radius, eccentricity, angular momentum) to infer past dynamical events like mergers and accretion of disrupted satellites (interaction of the Galaxy with the cosmic environment)

- evaluate the impact of the radial migration on the observed abundance distributions and test the current dynamical models of the Galaxy (galactic dynamics)

In the future it will be possible to improve the RAVE stellar parameters and, therefore, the chemical abundances. The limited quantity of information held by a medium resolution, 400Å wide spectrum like the RAVE spectra does not allow the uncertainties to go below $\sim$300K in effective temperature and $\sim$0.5 dex in gravity, unless additional information is provided. Up-coming surveys like Skymapper will provide multi-bands photometry for billions of objects (RAVE stars included), allowing precise effective temperature estimations. These can be used to constrain the gravity of the RAVE stars, and the improved stellar parameters will lead to improved chemical abundances. The RAVE chemical pipeline is also suitable of improvements and developments. In the near future we can modify the EWs library to take in account the opacity of the neighboring lines, making the reconstructed spectrum even closer to a synthesized spectrum. The pipeline has already proved to have stellar parameter estimation capabilities at its present shape. By implementing the Ca II triplet measurements these estimations will become more accurate and reliable, even at low S/N. Last, but not least, the pipeline can easily adjusted to measure high resolution spectra covering any wavelength range.

Bibliography

Abaldi et al., 2003, ApJ 597, 21

Abazajian, K., et al, 2003, AJ 126, 2081

Allende Prieto, C., Barklem, P.S., Lambert, D.L., Cunha, K., 2004, A&A 420, 183

Anders, E., Grevesse, N., 1989, Geochim. Cosmochim. Acta 53, 197

Andrievsky, S.M., et al., 2002, A&A 381, 32

Asplund, M., 2005, ARA&A 43 , 481

Asplund, M., Grevesse, N., Sauval, A. J., 2005, ASPConf.Ser., 336, 25

Bard, A., Kock, A., Kock, M. 1991, A&A, 248, 315

Bard, A., Kock, M. 1994, A&A, 282 ,1014

Barklem, P.S., Piskunov, N., O'Mara, B.J., 2000, A&AS 142, 467

Bensby, T., Feltzing, S., Lundström, I., 2003, A&A 410, 527

Bensby T., Feltzing S., Lundström I., Ilyin I., 2005, A&A 433, 185

Bigot, L., Thévenin, F. 2006, MNRAS 372, 609

Blackwell, D. E., Booth, A. J., Haddock, D. J., Petford, A. D., Leggett, S. K. 1986, MNRAS, 220, 549

Blackwell, D. E., Booth, A. J., Menon, S. L. R., Petford, A. D. 1986, MNRAS, 220, 303

Blackwell, D. E., Booth, A. J., Menon, S. L. R., Petford, A. D. 1986, MNRAS, 220, 289

Boeche, C., and the RAVE collaboration, 2010, in preparation

Borrero, J. M., Bellot Rubio, L. R., Barklem, P. S., del Toro Iniesta, J. C., 2003 A&A 404,749

Breddels, M. A., 2010, A&A 511, 90

Brook et al., 2004, ApJ 612, 894

Brook et al., 2005, ApJ 630,298

Carney, B.W. et al., 2005, AJ 130, 1111

Castelli, F., Kurucz, R. L., 2003, IAUS, 210, P.A20C

Castelli, F., Gratton, R.G. & Kurucz, R. L., 1997, A&A 318, 841

Cescutti, G., et al., 2007, A&A 462, 943

Chen, Y.Q., Nissen, P.E., Zhao, G., Zhang, H.W., Benoni, T., 2000, A&AS 141, 491

Chiappini., C., Matteucci.,F., Gratton, R., 1997, AJ 477, 765

Costa, A.D.D., et al., 2004, A&A 423, 199

Cutri, R. M., et al. 2003, The 2MASS All-Sky Catalog of Point Sources (Pasadena: Caltech)

Dehnen, W., Binney, J., 1998, MNRAS 294, 429

Luck, R.E., Heiter, U. 2005, 2005, AJ 129, 1063

Luck, R.E., Heiter, U. 2006, 2006, AJ 131, 3069

Luck, R.E., Heiter, U. 2007, 2007, AJ 133, 2464

Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D.L., Nissen, P.E., Tomkin, J., 1993, A&A 275, 101

Eggen, O., J., Lynden-Bell, D., Sandage, R., 1962, ApJ 136, 748

Fall, S.M., Efstathiou, G., MNRAS 1980, 193,189

Freeman, K., Bland-Hawthorn, J., 2002, ARA&A 40, 487

Friel, E.D., 2002, AJ 124, 2693

Fuhrmann K., 1998, A&A 338, 161

Fuhrmann, K., 2008, MNRAS 384, 173

Fulbright, J.P., McWilliam, A., Rich, R.M., 2006, AJ 636, 821

Fulbright, J.P., McWilliam, A., Rich, R.M., 2007, AJ 661, 1152

Gilmore, G., Reid, N., 1983, MNRAS 202, 1025

Gray, D. F., 2005, *The Observation and Analysis of Stellar Photospheres*, Third Edition, Cambridge University Press, New York

Grevesse, N., Sauval, A.J. 1998, Space Sci. Rev. 85, 161

Gurtovenko, E.A., Kostik, R. I. 1981, A&AS, 46, 239

Hambly, N. C., et al. 2001b, MNRAS, 326, 1279

Hinkle, K., Wallace, L., Valenti, J.,; Harmer, D., 2000, *Visible and Near Infrared Atlas of the Arcturus Spectrum 3727-9300 Å*, Available at: ftp://ftp.noao.edu/catalogs/arcturusatlas/visual

Høg, E., et al. 2000, A&A, 355, L27

Holmberg, J., et al., 2007, A&A 475, 519

Holweger, H., Bard, A., Kock, A., Kock, M., 1991, A&A 249, 545

Jurić, M. Brooks, A., Lupton, R.H., et al., AJ 673, 874

Kirby, E.N., et al., 2008, AJ 682, 1217

Kurucz, R. L. 1995, ASPConf.Ser., 78, 205

Lemasle, B. et al., 2008, A&A 490, 613

Luck, R.E.,et al., 2006, AJ 132,902

Luck, R.E., Heiter, U., 2006, AJ 131, 3069

Luck, R.E., Heiter, U., 2007, AJ 133, 2464

Maciel, W.J., Quireza C., 1999, A&A 345, 629

Mäckle, R., Holweger, H., Griffin, R., Griffin, R., 1975, A&A 38, 239

Matteucci, F., François. P., 1989, MNRAS 239, 885

Mishenina, T.V., Kovtyukh, V.V., Korotin, S.A., Soubiran, C., 2003, Astron. Zh. 80, 458

Mo, H.J., Mao, S., White S., 1998, MNRAS 295, 319

Moore, C. E., Minnaert, M. G. J., Houtgast, J., 1966, *The solar spectrum 2935 Å to 8770 Å*, National Bureau of Standards Monograph 61

Munari, U., Sordo, R., Castelli, F., & Zwitter, T., 2005, A&A 442, 615

Nordström, B., et al., 2004, A&A 418, 989

Perryman, M.A.C., et al., 1997, *The Hipparcos and Tycho Catalogues* (ESA SP-1200;Noordwijk: ESA)

Perryman, M. A. C., et al., 2001, A&A 369, 339

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992, *Numerical Recipes in FORTRAN*, second edition, Cambridge University Press

Reddy, B.E., Lambert D.L., Allende Prieto C., 2006, MNRAS 367, 1329

Reddy B.,E., Tomkin J., Lambert L., and Allende Prieto C., 2003, MNRAS 340, 304

Rudolph, A.L., Fich, M., Bell, G.R., et al., 2006, AJ 162, 346

Salaris, M., Chieffi, A., Straniero, O., 1993, AJ 414, 580

Sales, L.V., et al., 2009, MNRAS 400, L61

Sauval, A. J., Grevesse, N., 1994, IAUS, 154, 549

Schönrich, R., Binney, J., 2009, MNRAS 399, 1145

Simmons, G. J., Blackwell, D. E., 1982, A&A 112, 209

Sneden, C., 1973, Ph.D. thesis, Univ. Texas ar Austin, software available at: http://verdi.as.utexas.edu/moog.html

Soubiran, C., Girard, P., 2005, A&A 438, 139

Soubiran, C., Bienaymé, O., Siebert, A., 2003, A&A 398, 141

Steinmetz, M., and the RAVE collaboration, 2006, AJ 132, 1645

Seabroke, G.M., et al., 2008, MNRAS 384, 11

Sestito, P., et al., 2008, A&A 488, 943

Sordo, R., et al., 2009, MmSAI 80, 103S

Tamuz, O., Mazeh, T., North, P., 2006, MNARS 367, 1521

Thévenin, F. 1989, A&AS, 77, 137

Thévenin, F. 1990, A&AS, 82, 179

Thévenin, F. 1998, Bull. Inf. CDS 49 (unpublished)

Unsöld, A., 1955, Physik der Sternatmospharen. Springer, Berlin

Kupka, F.,Piskunov, N., Ryabchikova, T.A., Stempels, H.C., Weiss, W.W., 1999, A&AS, 138, 119

Valenti, J.A., Fisher, D.A., AJSS 159,141

Veltz, L., et al., 2008, A&A 480, 753

Venn K.A., Irwin M., Shetrone M.D., Tout C.A., Hill V., Tolstoy E., AJ 128, 1177

Villalobos, A., Helmi, A., 2008, MNRAS 391, 1806

Watson, F.G., et al., 2000, SPIE Conf. Series 4008, 123

Wilson, O.C., 1941, ApJ 93, 29

Yanny, B., et al., 2009, AJ 137, 4377

Yong. D., et al., 2006, AJ 131, 2256

Yong. D., et al., 2005, AJ 130, 597

Zwitter, T., and the RAVE collaboration, 2008, AJ 136, 421

Zwitter, T., et al., 2010, A&A accepted

Appendix A

Tests on synthetic spectra

We report here all the plots about quality tests performed with the chemical pipeline on synthetic spectra.
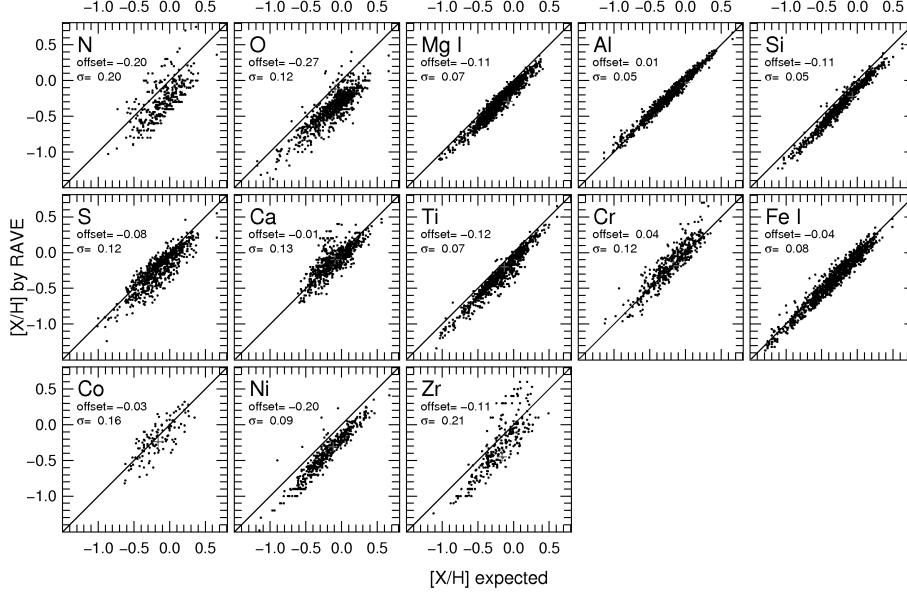
Figure A.1: Expected abundances [X/H] versus RAVE abundances for the sample of synthetic spectra at S/N=100. The offset and standard deviation σ from the expected values are reported in the panels for each element.
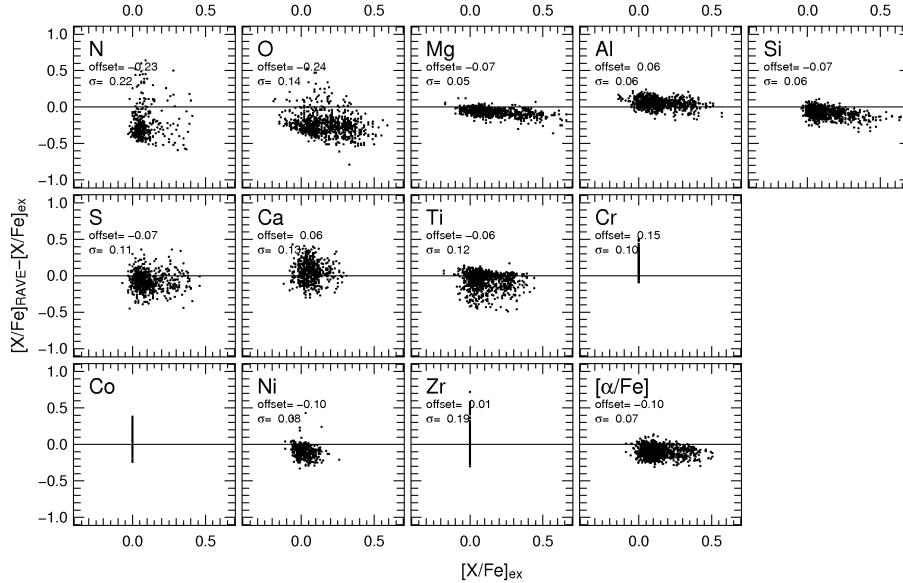
Figure A.2: Expected relative abundances [X/Fe] versus residuals between RAVE and expected abundances for the sample of synthetic spectra at S/N=100. The offset and standard deviation σ from the expected values are reported in the panels for each element.
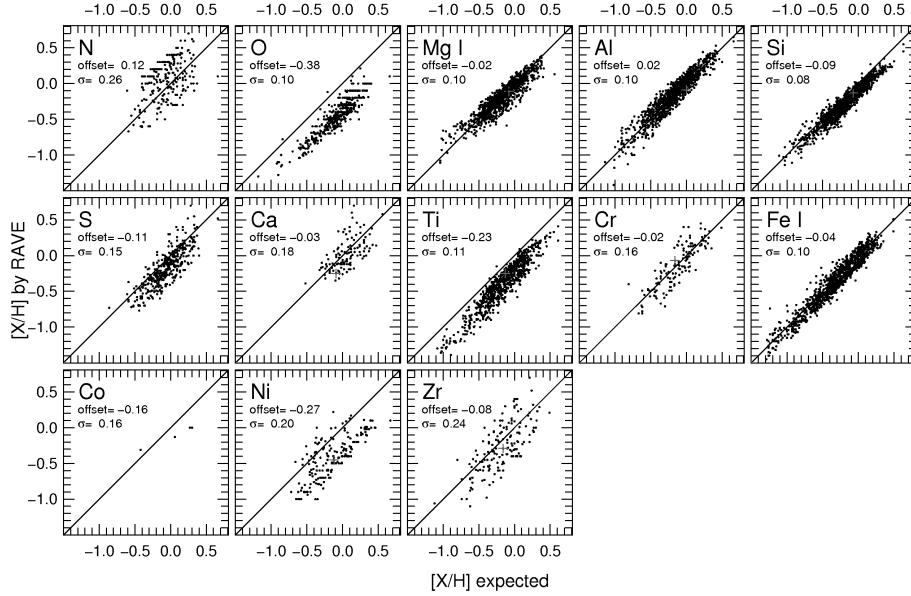
Figure A.3: As in Figure A.1 but for S/N=40.
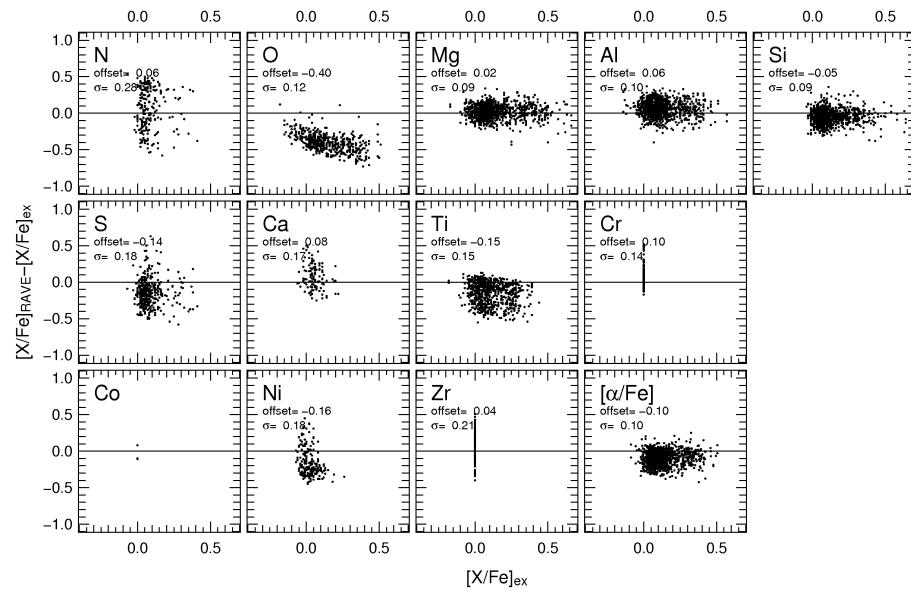
Figure A.4: As in Figure A.2 but for S/N=40.
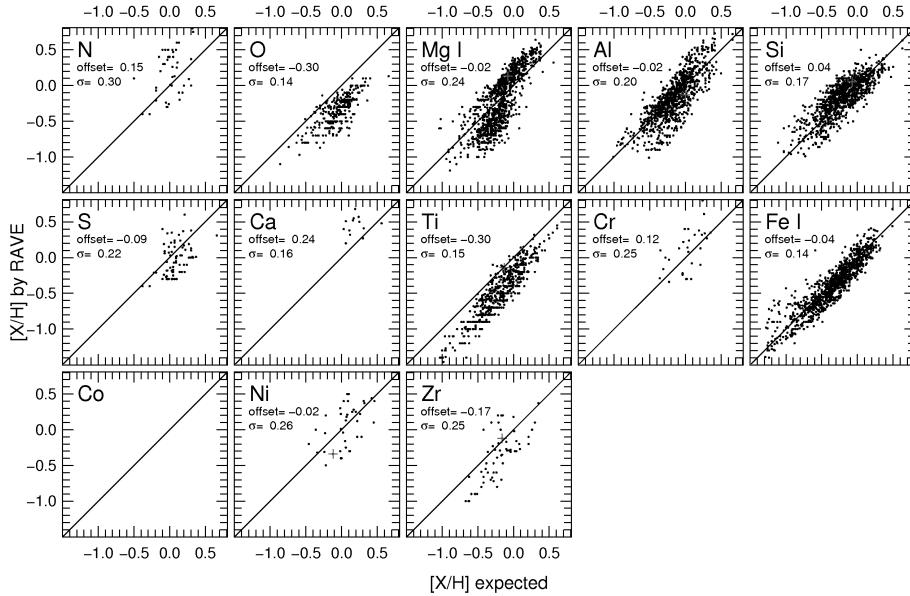
Figure A.5: As in Figure A.1 but for S/N=20.
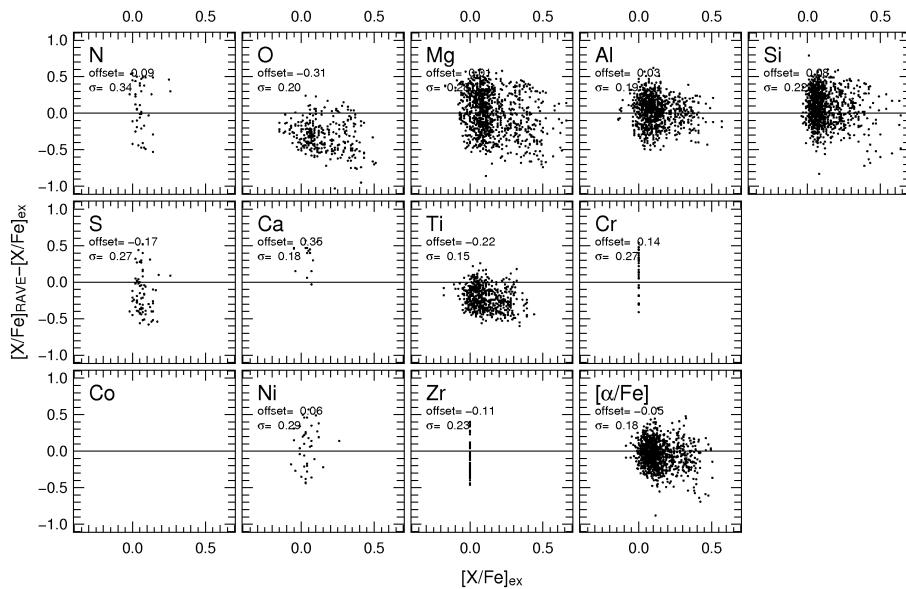

Figure A.6: As in Figure A.2 but for S/N=20.
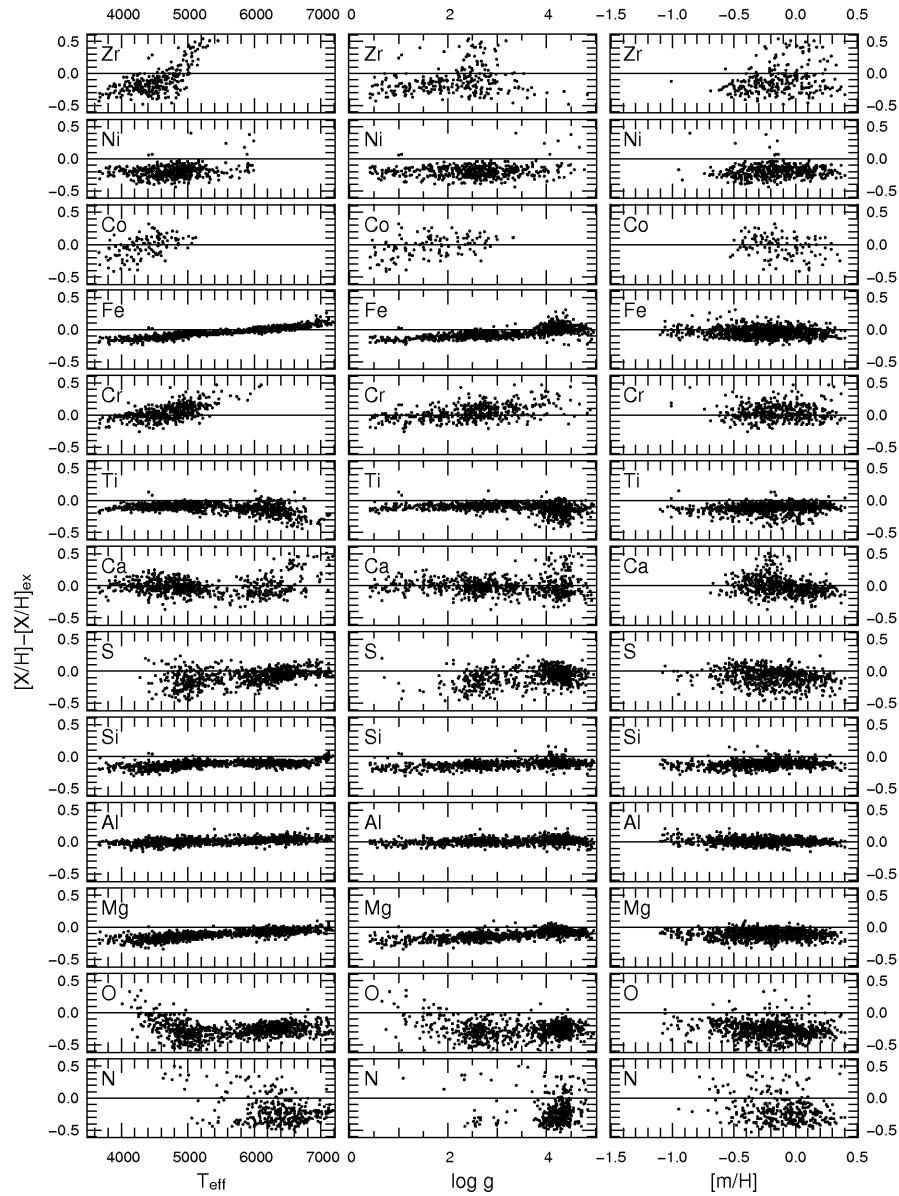
Figure A.7: Correlation between the abundances residuals (measured–expected) and stellar parameters $T_{\rm eff}$, $\log g$, [m/H] at S/N=100.
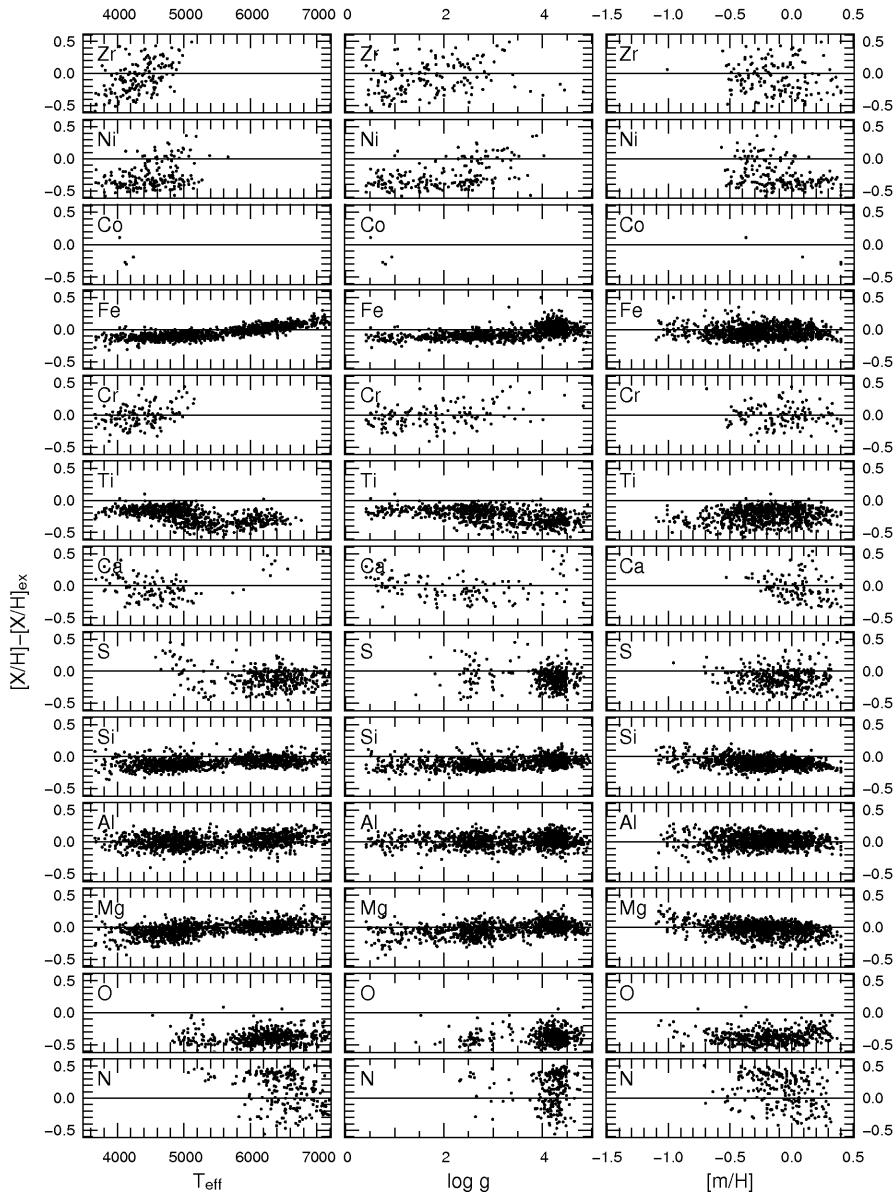
Figure A.8: As in Figure A.7 but for S/N=40.
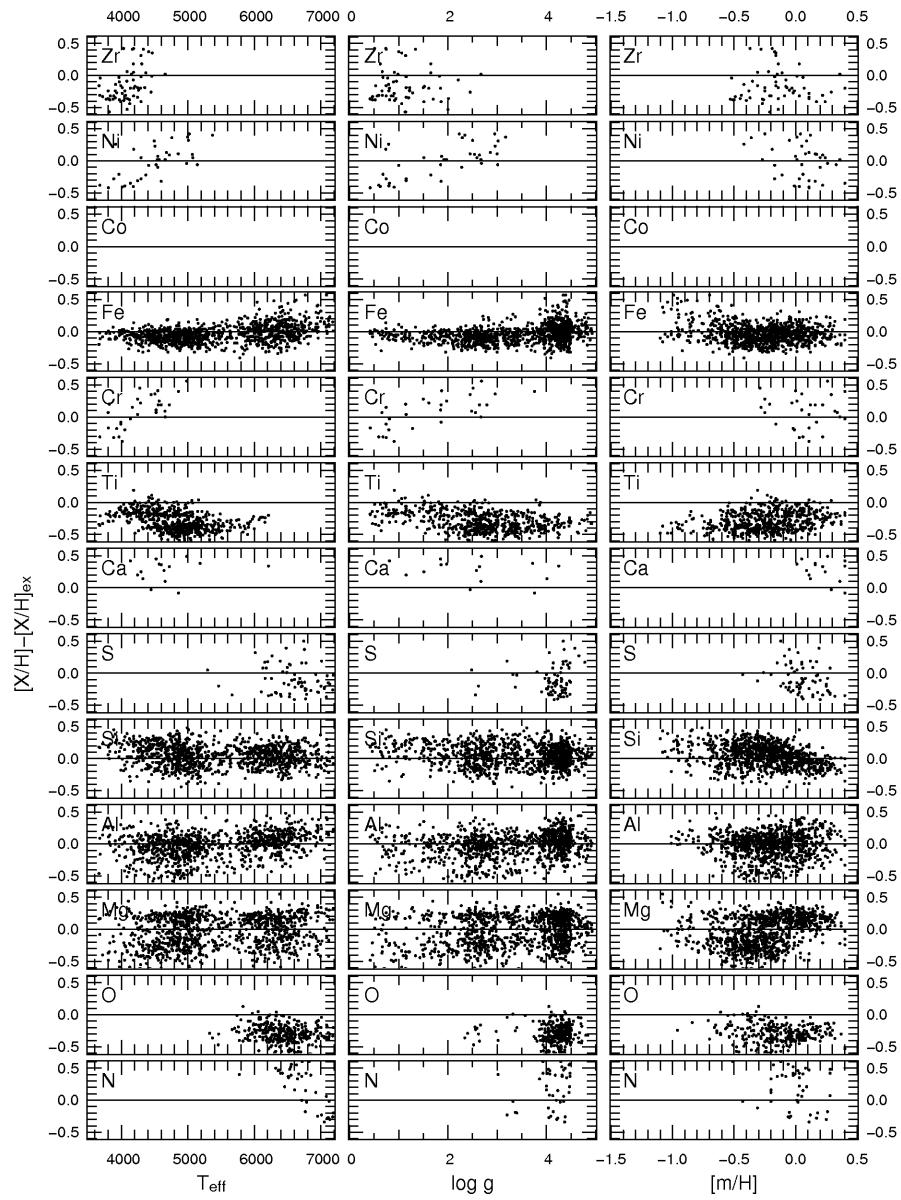
Figure A.9: As in Figure A.7 but for S/N=20.
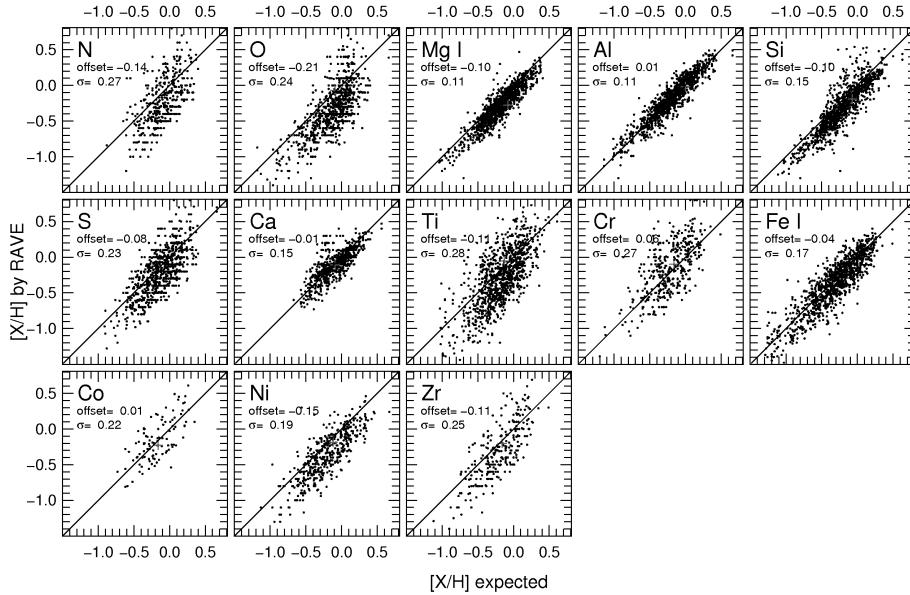
Figure A.10: Comparison between expected (y-axis) and measured abundances (x-axis) at S/N=100 with "wrong" stellar parameters to simulate the RAVE stellar parameters errors.
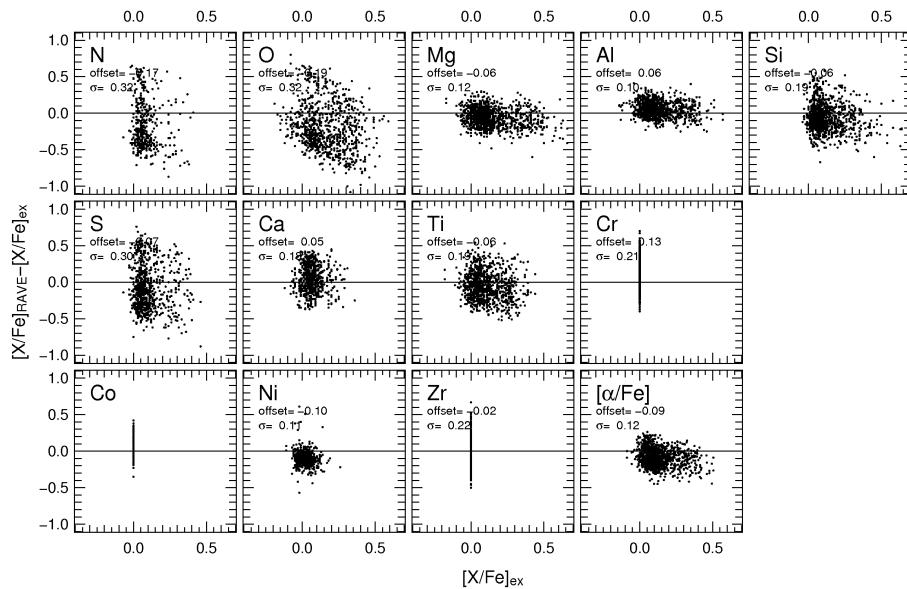
Figure A.11: As in Figure A.10 but for relative abundances [X/Fe].
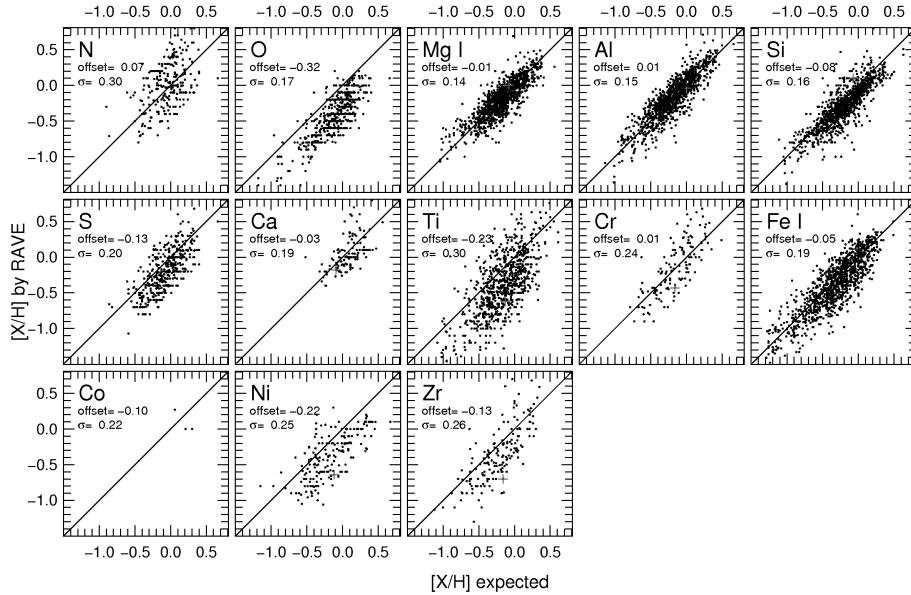
Figure A.12: Comparison between expected (y-axis) and measured abundances (x-axis) at S/N=40 with "wrong" stellar parameters to simulate the RAVE stellar parameters errors.
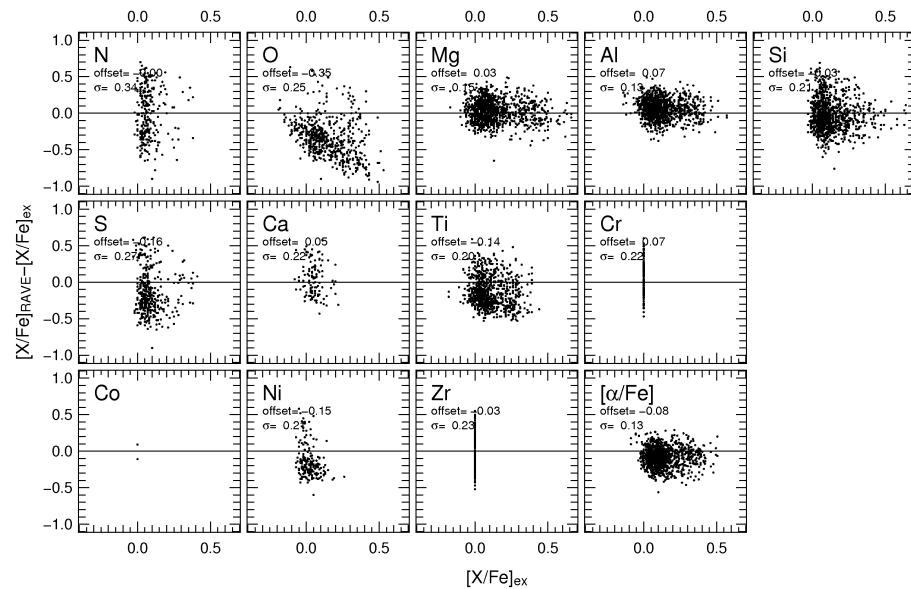
Figure A.13: As in Figure A.10 but for relative abundances [X/Fe].
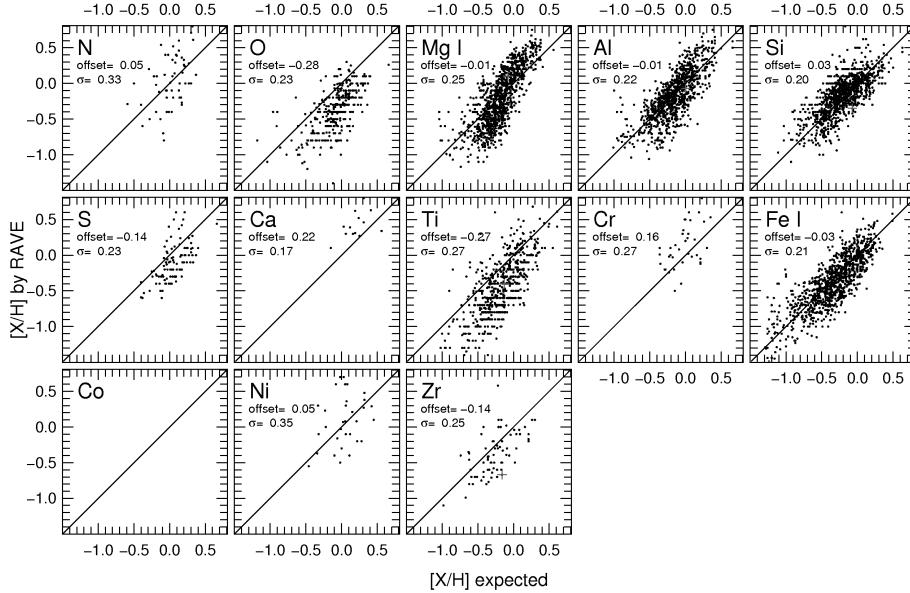
Figure A.14: Comparison between expected (y-axis) and measured abundances (x-axis) at S/N=20 with "wrong" stellar parameters to simulate the RAVE stellar parameters errors.
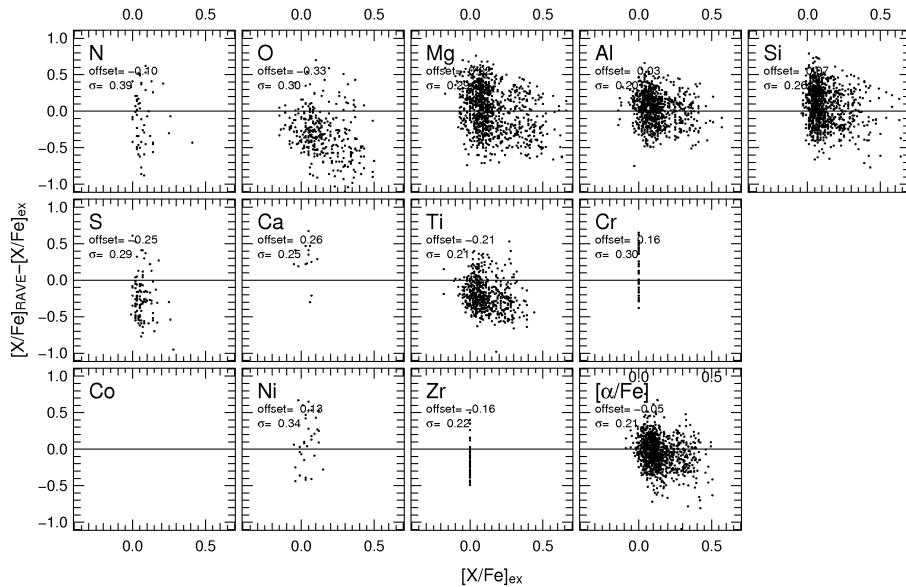
Figure A.15: As in Figure A.10 but for relative abundances [X/Fe].
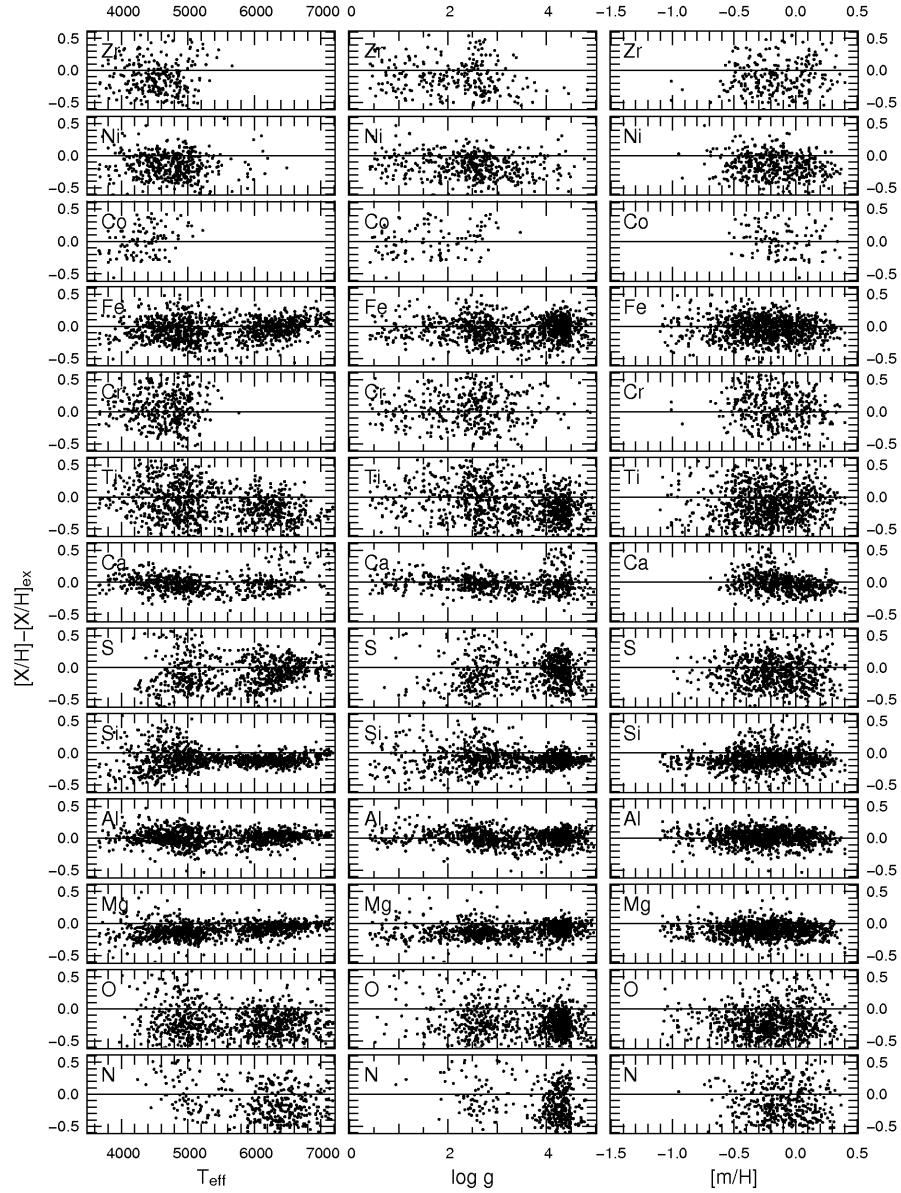
Figure A.16: Correlation between the abundances residuals (measured–expected) and stellar parameters $T_{\rm eff}$, $\log g$, [m/H] at S/N=100 with "wrong" stellar parameters to simulate the RAVE stellar parameters errors.
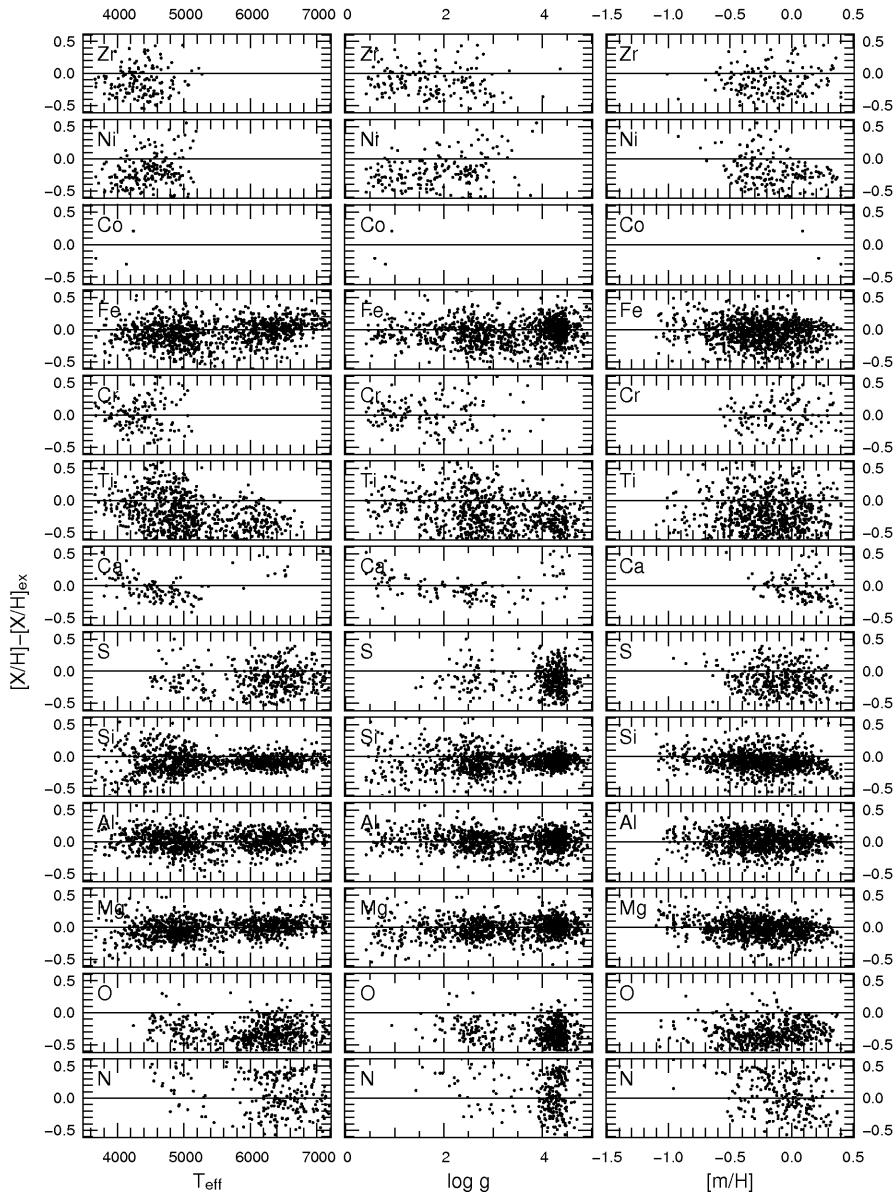
Figure A.17: As in Figure A.16 but for S/N=40.
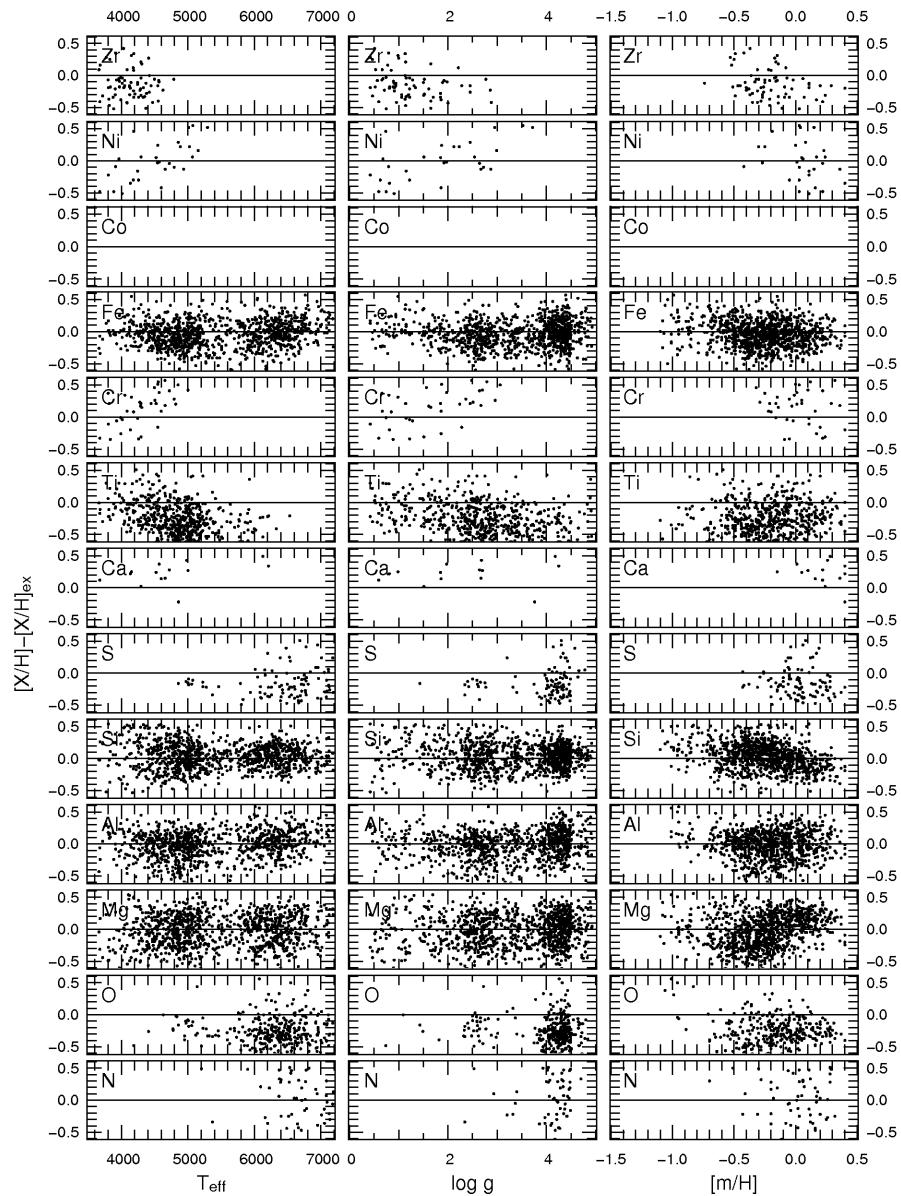
Figure A.18: As in Figure A.16 but for S/N=20.
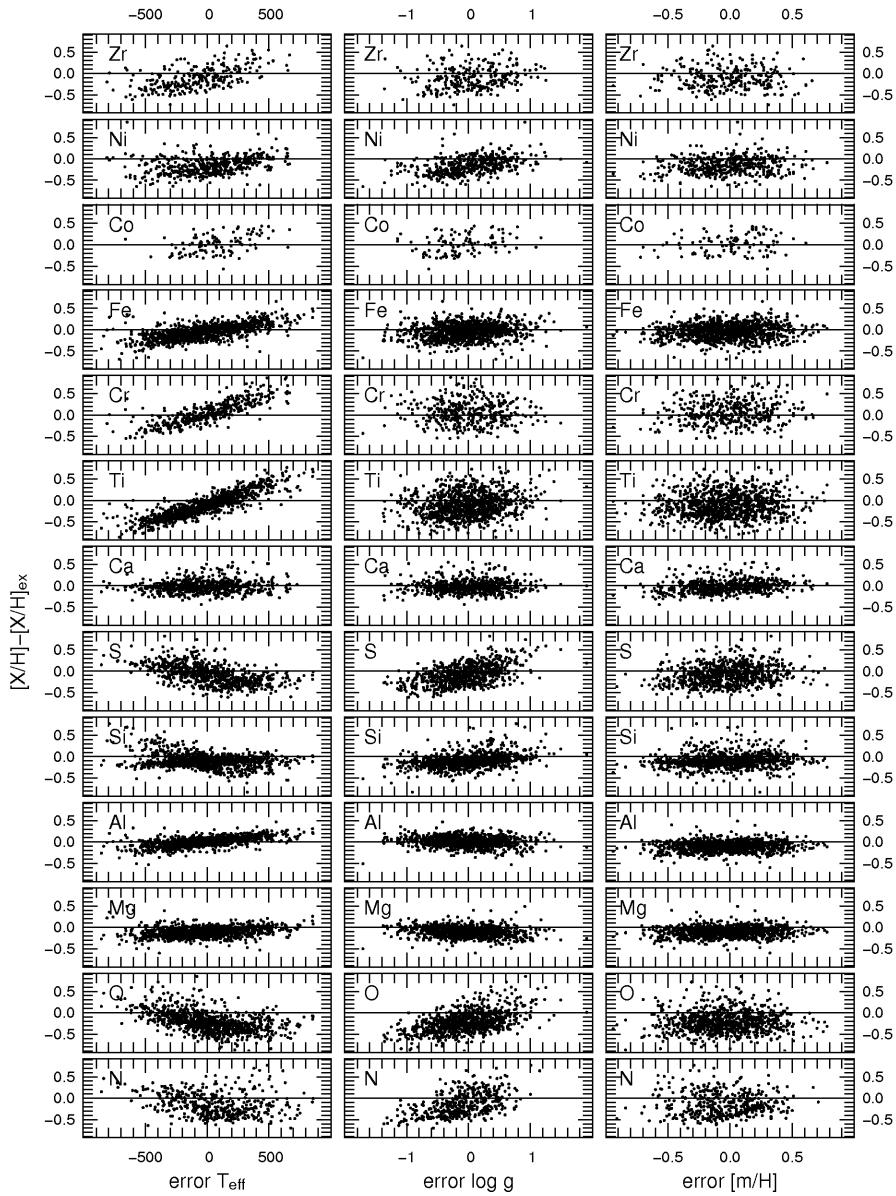
Figure A.19: Correlation between abundances errors (y-axis) and parameters errors (x-axis) at S/N=100.
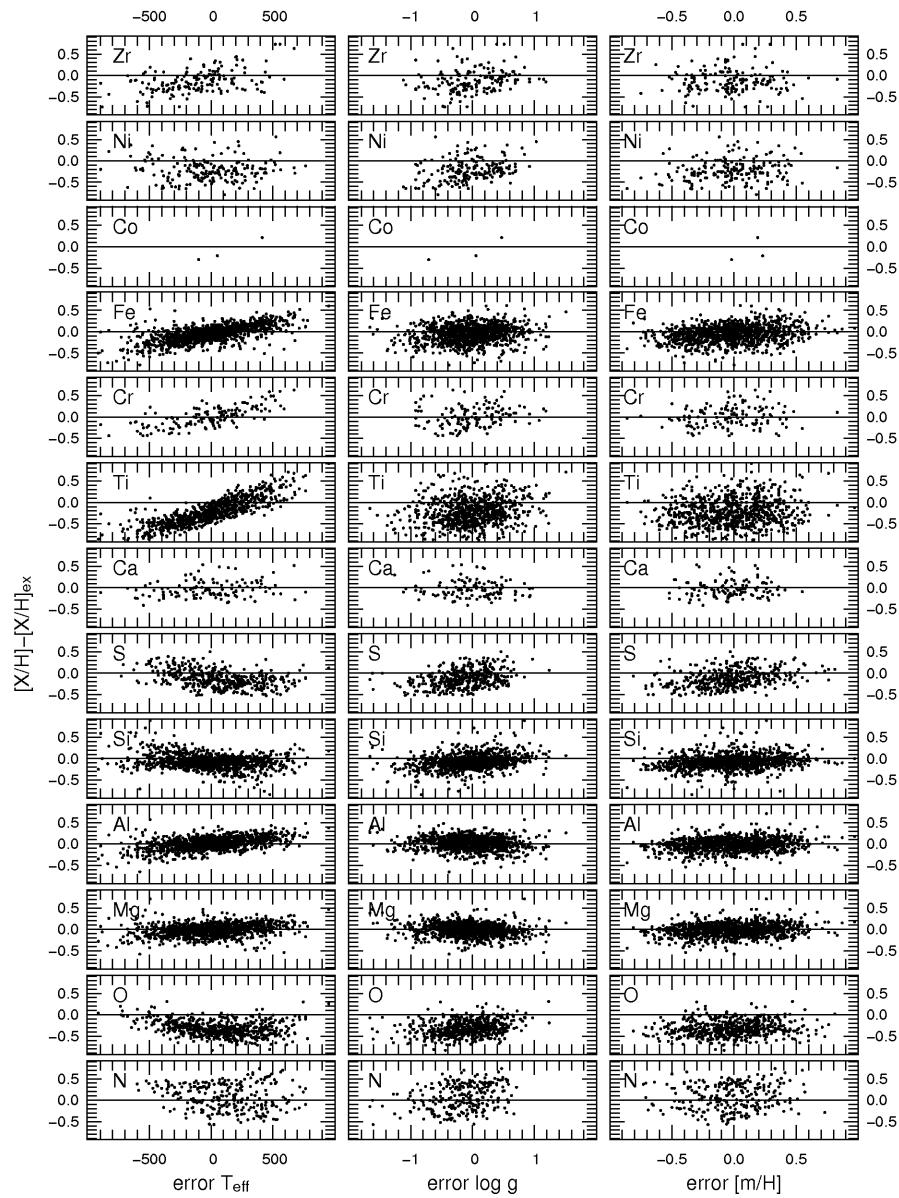
Figure A.20: Correlation between abundances errors (y-axis) and parameters errors (x-axis) at S/N=40.
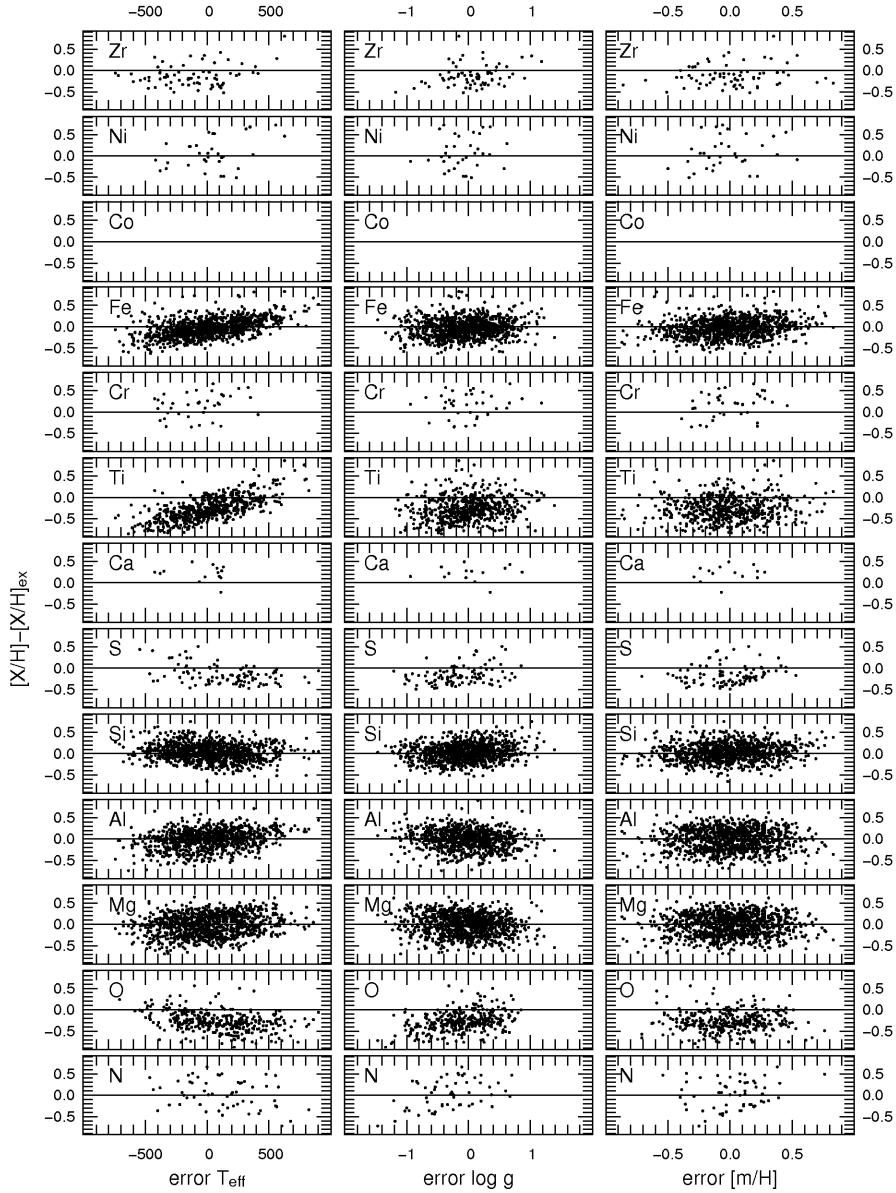
Figure A.21: Correlation between abundances errors (y-axis) and parameters errors (x-axis) at S/N=20.

Acknowledgements