SF2A 2013
L. Cambrésy, F. Martins, E. Nuss and A. Palacios (eds)

STELLAR PARAMETERS AND STELLAR PHYSICS FROM GAIA AND RELATED SPECTROSCOPIC SURVEYS

Alejandra Recio-Blanco[1]

Abstract. Automated stellar parameters are the bedrock of Galactic spectroscopic surveys science. They allow a rapid and homogeneous processing of extensive data sets, necessary for an efficient scientific return. Present Galactic Surveys, including the Gaia mission, have developped a wealth of mathematical approaches that are currently used today. These fundamental stellar parameters will provide important constraints for many stellar physics research fields.

Keywords: methods: data analysis, surveys, stars: abundances, stars: fundamental parameters, Galaxy: abundances

1 Introduction

A suite of ground-based vast stellar surveys mapping the Milky Way and culminating in the ESA Gaia mission, is revolutionizing the empirical information about Galactic stellar populations. In particular, in the recent years, the number of stars analysed with high enough spectroscopic resolution to provide detailed chemical diagnostics has increased from a few hundreds to several tens of thousands. Until the end of 2003, most of the information about the Milky Way was confined to small local samples, for which high-resolution spectroscopic data was obtained. In 2004, the Geneva Copenhagen Survey (Nordström et al. 2004) collected the first large spectro-photometric sample of around 16 000 stars, as part of a Hipparcos follow-up campaign (hence, also confined to 100 pc from the Sun). More recently, optical spectroscopic low-resolution surveys, such as SEGUE (Yanny et al. 2009) and RAVE (Steinmetz et al. 2006), have extended the studied volume to distances of a few kpc from the Sun (mainly in the range 0.5-3 kpc), and increased the numbers of stars with chemo-kinematical information by more than an order of magnitude ($>$ 200 000 spectra for SEGUE and $>$ 500 000 spectra for RAVE).

This effort is now complemented by new vast high-resolution spectroscopic surveys: the Gaia-ESO Survey (GES, 300 nights with the ESO/VLT), the Gaia/Radial Velocity Spectrograph (RVS) survey (part of the Gaia cornerstone mission), the Australian HERMES/GALAH survey, the LAMOST/LEGUE survey and APOGEE (part of the Sloan Digital Sky Survey III and After-SDSSIII).

All the above mentioned Galactic stellar population projects rely on the success of automated techniques of spectral analysis and parameterisation, capable to perform a rapid and homogeneous processing of the data and to allow an efficient scientific return. Figure 1 shows the increase of stars with available spectra in the next five years, as expected from the announced data releases of the different Milky Way surveys. One of the main challenges of Galactic Archaeology will be the correct and rigorous treatment of all those data sets, including the stars' fundamental parameters and the chemical characterization.

Therefore, those projects, including Gaia, the billion stars surveyor, are not only crucial for Galactic physics, but also for stellar physics. They will revolutionize many stellar physics research topics because they will provide, among other things: i) absolute magnitudes with unprecedented precision, crucial for stellar ages estimation, ii) atmospheric parameters and individual element abundances for an extraordinary high number of stars, including rare objects, and iii) precise proper motions and radial velocities.

---

[1] Laboratoire Lagrange (UMR7293), Université de Nice Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, BP 4229,F-06304 Nice cedex 4, France

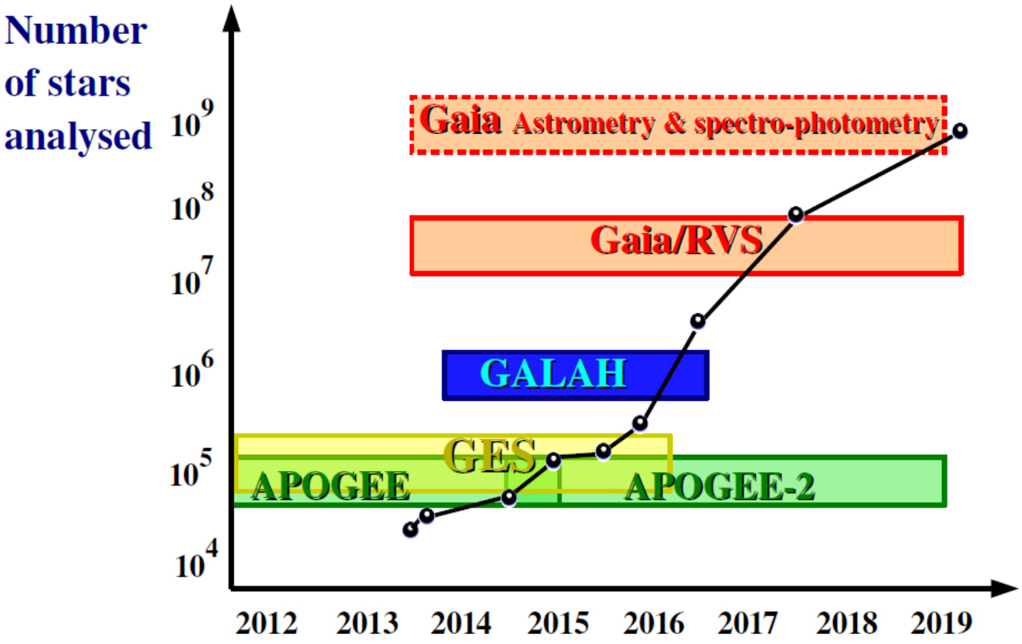


Fig. 1. Number of stars with available spectra in the next years, as expected from the announced data releases of the different Milky Way surveys.

2 Automated parameterization methods

The physical parameterization can be applied when the physics of the studied objects is enough well known, and modeled through continuous variables. For instance, the stellar effective temperature, the surface gravity, the global metallicity and the individual element abundances are more appropriate to describe a stellar spectra than spectral types and luminosity classes. parameterization algorithms use reference data to define the mapping between the observed targets and the models. Those models, usually synthetic spectra, constitute a N-dimensional grid, where N is the number of parameters to determine. There are three main mathematical parameterization approaches: optimization methods, projection methods and classification. All of them try to find the absolute minimum of the distance function, with different techniques (Recio-Blanco 2014).

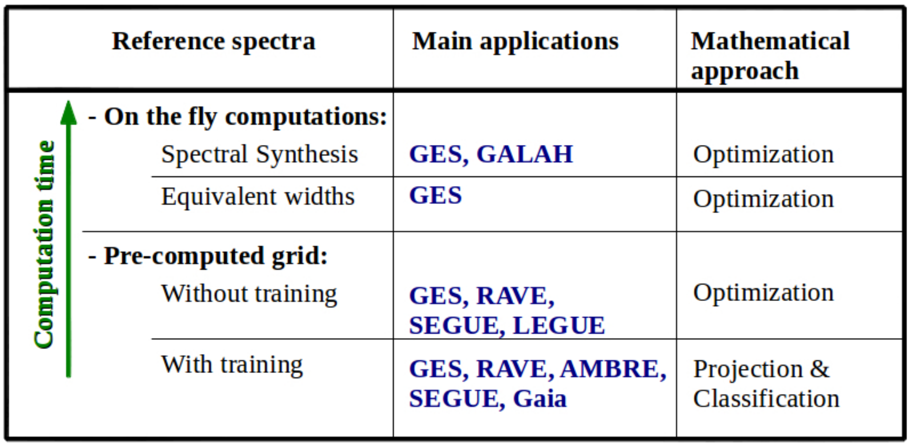
Figure 2 shows the different types of automated parameterisation methods found in the literature, currently used by Milky Way spectroscopic surveys and Galactic archaeology projects. The algorithms can be divided depending on the way in which the reference models are computed and used: on the fly computations, pre-computed grid of reference models used without and with training. This has an important influence on the computation time and, ultimately, it depends on the implemented mathematical approach. In addition, Figure 2 shows what approaches are used by the main spectroscopic surveys and projects, including the European Space Agency Gaia mission.

2.1 Methods using a pre-computed grid of reference synthetic spectra

The use of a pre-computed grid of synthetic spectra reduces the computing time of the algorithms application. The methods using this kind of approach are divided into those without and with a phase of algorithm training.

- *Without training*:

  This category of methods is based on optimization approaches. It includes the Nelder-Mead method (Nelder & Mead 1965) implemented by Allende Prieto et al. (2006). This non-linear downhill simplex method was already used for the SDSS-SEGUE SSPP pipeline for the derivation of both the atmospheric



Reference spectra	Main applications	Mathematical approach
- On the fly computations:		
Spectral Synthesis	**GES, GALAH**	Optimization
Equivalent widths	**GES**	Optimization
- Pre-computed grid:		
Without training	**GES, RAVE, SEGUE, LEGUE**	Optimization
With training	**GES, RAVE, AMBRE, SEGUE, Gaia**	Projection & Classification

Fig. 2. Different types of automated abundance analysis methods currently used by Milky Way spectroscopic surveys and galactic archaeology projects.

parameters and the [α/Fe] (Lee et al. (2011)) from low resolution stellar spectra. It is also part of the methods used by the Gaia-ESO Survey and it is the core of the APOGEE ASPCAP pipeline.

The penalized χ^2 method of Zwitter et al. (2008) is also in this category of algorithms. It has been used for the RAVE survey first and second data releases for the derivation of the iron abundance. The RAVE survey also developed the Boeche et al. (2011) method for the individual abundance analysis of the high signal-to-noise data (third data release). It is a minimum of distances method using of a grid of pre-computed equivalent widths.

The UlySS method of Koleva et al. (2009) implements a full spectrum fitting and a parametric minimization using χ^2 maps. It was part of the SEGUE SSPP pipeline and it is actually integrated in the LASP LAMOS pipeline for the analysis of the LEGUE survey data.

The GAUGUIN method (Bijaoui et al. 2012) uses the Gauss-Newton algorithm for the determination of the global metallicity simultaneously with stellar atmospheric parameters. It can also be used for the derivation of individual abundances in a second step of the spectrum analysis. The GAUGUIN algorithm is already applied to GES data and it is been prepared for its integration in the Apsis pipeline for the individual abundance analysis of the Radial Velocity Spectrograph (RVS) data, collected by the Gaia mission of the European Space Agency.

- *With training*:

The methods with a faster application are those relying on a training phase. They are based on projection and classification approaches. The neural network algorithms of Re Fiorentin et al. (2007) is an example of this kind of methods. It is part of the SEGUE SSPP pipeline for the derivation of the iron abundance, simultaneously with the effective temperature and the surface gravity. It implements a global and non-linear regression mapping.

The MATISSE and DEGAS methods are part of the algorithms developped by the Nice group. The MATISSE algorithm is a local multi-linear regression method (see Recio-Blanco et al. 2006). The stellar parameters are determined through the projection of the input spectra on a set of vectors, calculated during a training phase. The DEGAS method is based on an oblique k-d decision tree and uses the pattern recognition approach for stellar parameterization. The MATISSE and DEGAS methods have been used in Kordopatis et al. (2011) for a study of the Thick Disc outside the solar neighbourhood (700 stars analysed) and for the last data release (DR4) of the RAVE Galactic Survey (Kordopatis et al. 2013, submitted, 228 060 spectra). These two applications share the same wavelength domain and resolution of the RVS one. In addition, MATISSE is the core method of the AMBRE project. AMBRE (de Laverny et al. 2012, see), under agreement between the European Southern Observatory (ESO) and the Observatoire de la Côte d'Azur, aims at determining the parameters ($T_{\rm eff}$, $\log g$, [M/H] and [α/Fe]) of the high resolution

Acronym	name
DSC	Discrete Source Classifier
ESP	Extended Stellar Parametrizer:
-CS	ESP Cool Stars
-ELS	ESP Emission Line Stars
-HS	ESP Hot Stars
-UCD	ESP Ultra Cool Dwarfs
FLAME	Final Luminosity Age and Mass Estimator
GSP-Phot	Generalized Stellar Parametrizer Photometry
GSP-Spec	Generalized Stellar Parametrizer Spectroscopy
MSC	Multiple Star Classifier
OA	Outlier Analysis
OCA	Object Clustering Algorithm
QSOC	Quasar Classifier
TGE	Total Galactic Extinction
UGC	Unresolved Galaxy Classifier

Table 1. Modules of the Gaia Astrophysical Parameters Inference System (Apsis Bailer-Jones et al. 2013).

stellar spectra contained in the ESO archive. This concerns the FEROS, HARPS, UVES and FLAMES spectrographs. The results of the AMBRE project are presented in Worley et al. (2012) (FEROS data analysis), De Pascale et al. (2013, in preparation; HARPS data analysis) and Worley et al. (2013, in preparation, UVES data analysis). MATISSE has also been used for the characterization of several disc fields observed by the CoRoT mission (Gazzano et al. 2010, 2013). In addition, the MATISSE algorithm is part of the methods used for the stellar parameterization of FGK type targets of the Gaia-ESO Large Public Survey. In particular, the first data release of GES parameters for the FGK-type stars observed with the GIRAFFE spectrograph includes the MATISSE results for those data.

Finally, the spectrum analysis of the Gaia mission is based on this type of algorithms with a training phase. The computational time is a crucial constraint for the Gaia pipeline, as several tenths of millions of RVS spectra will be analysed in cycles of 6 months.

3 The Gaia astrophysical parameters inference system

In addition to its astrometry, Gaia will obtain optical BP/RP low-resolution spectro-photometry for all one billion of its target sources, as well as higher resolution RVS spectra (for the targets brighter than G$\sim$17). These spectra are used to calculate a chromatic calibration of the astrometry, and to estimate the stellar radial velocities, respectively. But they also provide valuable information on the physical properties of the sources.

Figure 3 outlines the architecture of the data processing system developed by DPAC/CU8 for this purpose. Known as Apsis, the Astrophysical Parameters Inference System, it comprises multiple software modules (boxes in the diagram), each charged with a specific task (see Table 1). GSP-Phot, for example, estimates the astrophysical parameters effective temperature, line-of-sight interstellar extinction, metallicity, and surface gravity, for all stars. Other modules examine other types of objects, possibly using the outputs from previous modules (e.g. the GSP-Phot outputs are used by GSP-SPEC, FLAME, ESP, and TGE). The coloured bars show which Gaia data are used for the various modules.

Multiple methods are used for many types of stars (Bailer-Jones et al. 2013), producing multiple results for the end user according to different models and assumptions. Prior to its application to real Gaia data the accuracy of these methods cannot be assessed definitively. But as an example of the current performance, GSP-phot can attain internal accuracies (RMS residuals) on F,G,K,M dwarfs and giants at G = 15 (V = 1517) for a wide range of metallicites and interstellar extinctions of around 100 K in effective temperature (Teff), 0.1 mag in extinction (A0), 0.2 dex in metallicity ([Fe/H]), and 0.25 dex in surface gravity (log g). GSP-spec estimates Teff , log g, global metallicity [M/H], al- pha element abundance [α/Fe], and some individual chemical abundances for single stars using continuum-normalized RVS spectra. GSPspec internal accuracies can attain, at G = 13, 70 K for effective temperature, 0.12 dex for log g and 0.09 dex for metallicity. The individual abundances of several elements (Fe,Ca, Ti, Si) will be measured for brighter stars with an expected internal precision of 0.1

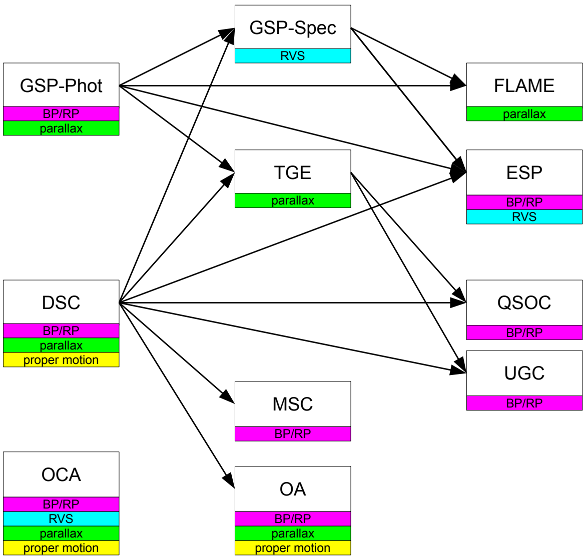
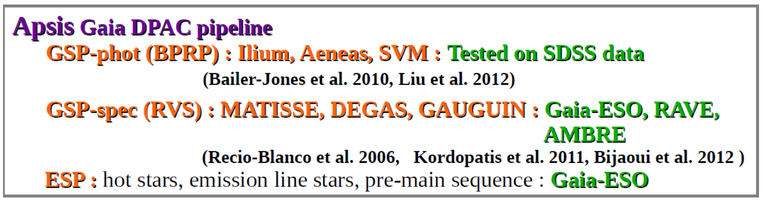


Fig. 3. Architecture of the Gaia Apsis data processing. The coloured bars show which Gaia data are used for the various modules.



Apsis Gaia DPAC pipeline

GSP-phot (BPRP) : Ilium, Aeneas, SVM : Tested on SDSS data

(Bailer-Jones et al. 2010, Liu et al. 2012)

GSP-spec (RVS) : MATISSE, DEGAS, GAUGUIN : Gaia-ESO, RAVE, AMBRE

(Recio-Blanco et al. 2006, Kordopatis et al. 2011, Bijaoui et al. 2012)

ESP : hot stars, emission line stars, pre-main sequence : **Gaia-ESO**

Fig. 4. Algorithms integrated in the stellar parameterisation modules of the Apsis pipeline. In green, the ground-based Galactic projects that allowed the testing of the methods, or that include the same algorithms in their data analysis pipelines.

dex for G < 13. The accuracy is a strong function of the parameters themselves, varying by a factor of more than two up or down over this parameter range.

Finally, the algorithms integrated in the Apsis pipeline for the stellar parameterisation are already being used for the analysis of ground based Galactic projects as the Sloan Digital Sky Survey, the RAVE survey, the Gaia-ESO Survey and the AMBRE project. This is shown in Fig. 4, together with the names of the integrated methods for each stellar parameterisation module.

4 Conclusions

The study of the stellar fundamental parameters is rapidly evolving, and will definitly be revolutionized by the Gaia data. In particular, unprecedented constraints to stellar structure and evolution models from very precise distances and homogeneous parameters will be possible, including stars in rapid evolutionary phases and rare objects. In addition, many other stellar science cases are at the core of the Gaia project and other Galactic spectroscopic surveys as the Gaia-ESO Survey: stellar activity, the interplay between stellar dynamics and evolution in dense environements as globular clusters, brown dwarfs and white dwarfs evolution, stellar variability across the HR diagramme, binary system, binary systems studies, open clusters evolution,... The Gaia era will reveal the synergy between the stellar populations community and the stellar physics research.

References

Allende Prieto, C., Beers, T. C., Wilhelm, R., et al. 2006, ApJ, 636, 804

Bailer-Jones, C. A. L., Andrae, R., Arcay, B., et al. 2013, ArXiv e-prints

Bijaoui, A., Recio-Blanco, A., de Laverny, P., & Ordenovic, C. 2012, Statistical Methodology

Boeche, C., Siebert, A., Williams, M., et al. 2011, AJ, 142, 193

de Laverny, P., Recio-Blanco, A., Worley, C. C., & Plez, B. 2012, A&A, 544, A126

De Pascale, M., Recio-Blanco, A., de Laverny, P., & Worley, C. 2013, in preparation

Gazzano, J.-C., de Laverny, P., Deleuil, M., et al. 2010, A&A, 523, A91

Gazzano, J.-C., Kordopatis, G., Deleuil, M., et al. 2013, A&A, 550, A125

Koleva, M., Prugniel, P., Bouchard, A., & Wu, Y. 2009, A&A, 501, 1269

Kordopatis, G., Recio-Blanco, A., & de Laverny, P. 2013, submitted

Kordopatis, G., Recio-Blanco, A., de Laverny, P., et al. 2011, A&A, 535, A107

Lee, Y. S., Beers, T. C., An, D., et al. 2011, ApJ, 738, 187

Nelder, J. & Mead, R. 1965, Comp. Journal, 7, 308

Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989

Re Fiorentin, P., Bailer-Jones, C. A. L., Lee, Y. S., et al. 2007, A&A, 467, 1373

Recio-Blanco, A. 2014, in IAU Symposium, Vol. 298, Setting the Scene for Gaia and LAMOST, ed. S. Feltzing, G. Zhao, N. Walton, & P. Whitelock

Recio-Blanco, A., Bijaoui, A., & de Laverny, P. 2006, MNRAS, 370, 141

Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645

Worley, C., Recio-Blanco, A., de Laverny, P., & De Pascale, M. 2013, in preparation

Worley, C. C., de Laverny, P., Recio-Blanco, A., et al. 2012, A&A, 542, A48

Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377

Zwitter, T., Siebert, A., Munari, U., et al. 2008, AJ, 136, 421