Primary AP reference stars for the calibration of the General Stellar Parametrizer algorithms

Selection, available spectra

prepared by: C. Soubiran, U. Heiter

reference: GAIA-C8-TN-LAB-CS-008-1
issue: D
revision: 0
date: 2010-08-12
status: Draft

Abstract

This document describes one of the tasks of the work package "Provide calibrations of training data" (GWP-S-811-20000) within the CU8 work package "Training data" (GWP-S-811). The purpose of that work package is to provide a basis for the calibration of the General Stellar Parametrizer algorithms (GSP-phot and GSP-spec). This document recalls the principles of the calibration procedure proposed in Bailer-Jones (2009) and of the three level grid for reference stars as described in Heiter et al. (2009). We then focus on the selection of primary AP reference stars and the archives where relevant high quality spectra can be found for them. Some thoughts on how we will analyse this large amount of observational material are also given.

Document History

Issue	Revision	Date	Author	Comment
D	5	2010-08-12	CS	comments by C. Bailer-Jones and A. Vallenari taken into account
D	5	2010-06-07	UH+CS	final version
D	4	2010-04-02	CS	section on PASTEL and available spectra changed
D	3	2009-11-14	UH	changed the three spectrum figures
D	2	2009-11-12	UH	more revisions of text, added three figures and one table
D	1	2009-11-10	UH	Minor revisions of text
D	0	2009-07-28	CS	Creation

Acronyms

The following table has been generated from the on-line Gaia acronym list:

Acronym	**Description**
AP	Astrophysical Parameter
CCD	Charge-Coupled Device
CDS	Centre de Donnes astronomiques de Strasbourg
CNO	Carbon-Nitrogen-Oxygen cycle (star)
DPAC	Data Processing and Analysis Consortium
ESO	European Southern Observatory
ESP	Extended Stellar Parametriser
FEROS	Fiber-fed Extended Range Optical Spectrograph
FITS	Flexible Image Transport System
FLAMES	Fiber Large Array Multi-Element Spectrograph (VLT)
GC	Globular Cluster
GSP	Generalised Star Parametrisation (Parametriser)
GSPphot	Generalised Stellar Parametriser PHOTometry
GSPspec	Generalised Stellar Parametriser SPECtroscopy
HARPS	High Accuracy Radial velocity Planet Searcher
HIDES	HIgh Dispersion Echelle Spectrograph
LMC	Large Magellanice Cloud (special, high-density area on the sky)
OC	Open Cluster
OHP	Observatoire de Haute Provence (France)
RAVE	RAdial Velocity Experiment
RV	Radial Velocity
RVS	Radial Velocity Spectrometer

SDSS	Sloan Digital Sky Survey
SEP	South Ecliptic Pole
SME	Spectroscopy Made Easy
SNR	Signal-to-Noise Ratio (also denoted SN and S/N)
SPSS	Spectro-Photometric Standard Star
TBL	Telescope Bernard Lyot
TNG	Telescopio Nazionale Galileo
UVES	UV-Visual Echelle Spectrograph (VLT)
WP	Work Package

Contents

1 Introduction

The goal of the General Stellar Parametrisers, GSP-phot and GSP-spec, is the estimation of Astrophysical Parameters (APs) for those objects identified as stars. The relevant intinsic AP are effective temperature, $T_{\rm eff}$, surface gravity, log g, metallicity, [Fe/H], and to some extent the abundance of α elements, [α/Fe]. The line-of-sight interstellar extinction, Av, which is not an intrinsic property, will be estimated on star by star basis by GSP-phot. Other APs of interest which could be determined for bright stars by GSP-spec include CNO abundance anomalies, activity and rotation.

The GSP algorithms are trained on synthetic spectra which are very convenient because grids of any resolution, spectral range and combination of APs can be built. However synthetic spectra have a major drawback : they do not reproduce correctly real spectra, causing systematic errors in the results of the AP estimation.

The aim of the external calibration of GSP algorithms is to correct as far as possible the synthetic spectra mismatch problem and provide realistic error bars across the AP space.

There are several ways to reduce the synthetic spectra mismatch, using real stars. One can improve the physics used to provide the synthetic spectra, model atmospheres and line data, as constrained by benchmark stars with very well-known properties. Directly comparing high resolution spectra or photometric fluxes of such benchmark stars to computed ones helps understanding where the models are deficient, as demonstrated for instance by Edvardsson (2008). Another way is to directly correct synthetic grids to make them fit artificially BP/RP-like and RVS-like spectra of benchmark stars. These two methods to make synthetic spectra more realistic are under study in the work package “Provide calibrations of training data“(GWP-S-811-20000).

Another way is to use Gaia BP/RP and RVS observations of AP reference stars to correct for the synthetic spectra mismatch. Such correction is performed through the GSP algorithms as an external calibration. Although this process can only start after launch, we are preparing the list of the most suitable stars to be used. This document gives an overview of the related tasks.

2 External calibration of GSP algorithms

2.1 Calibration procedure

A procedure to calibrate algorithms for estimating parameters from spectra, applicable to BP/RP data, has been proposed by Bailer-Jones (2009). It is based on the concept of forward modelling, described in details by Bailer-Jones (2010). The basic idea is to use Gaia observations of AP reference stars to modify the fluxes of the synthetic spectra, resulting in an hybrid synthetic-real grid to be used for the training. There are four fundamental assumptions behind this procedure, the fourth one being directly related to the choice of AP stars. It implies that we can obtain accurate APs of a set of stars which Gaia will observe, covering a broad range of the APs of interest with a sampling sufficient to map the large scale variation of flux with APs.

The accurate APs used for this process must be obtained in advance from conventional methods using ground-based high quality spectra, line data and model atmospheres. One has thus to keep in mind that whatever method is used to derive the APs depends ultimately on a set of stellar models.

The amount of calibration data required for this procedure depends on how much synthetic data deviate from real data, on how smooth is the variation of the flux with APs and on the spectral resolution. It has been estimated that for the BP/RP resolution, at least some 550 stars would be necessary. [α/Fe] adds another dimension, implying an extension of the reference set, but it is only relevant for cool stars.

The Gaia data of the AP reference stars must be of a sufficient SNR to be used for the training. It implies a G magnitude interval of $\sim$ 6-10 for RVS spectra , and $\sim$10-13 for BP/RP. It is not yet clear if the activation of gates at G$<$10 for BP/RP will have an influence on GSP-phot.

2.2 Three levels for reference stars

The basic idea developed in Heiter et al. (2009) is to have 3 levels of reference stars. At the first level, the benchmark stars are very bright and well-known stars which have been extensively studied before. Very-high-resolution spectroscopy and spectrophotometry is available for them or will be obtained. Ideally their distance and radius is accurately known also. Such stars are used to constrain model atmospheres and synthetic spectra. The list of benchmark stars is available in a dedicated GaiaWiki page[1]. Most of the benchmark stars are too bright to be observed by Gaia.

On the second level, we define a “primary grid” of reference stars, which will be studied in de-

[1] http://www.rssd.esa.int/SA-general/Projects/GAIA/wiki/index.php?title=CU8:_Benchmark_Stars

tail, based on high-resolution spectroscopy. The *primary grid* stars will be calibrated using the results obtained for the *benchmark stars*. These are the stars to be used in the calibration procedure described above. We will have to determine their APs as accurately and homogeneously as possible.

At a third level we will use larger sets of secondary reference stars, more for validation or tests than for calibration. Such stars have their APs determined at a level of precision similar to the one expected from Gaia (ex SDSS, RAVE,...). We will have to recalibrate their APs to the same scale defined by bootstrap from primary reference stars and benchmark stars.

Our current activity concerning the primary grid of reference stars is to select stars from bibliographical compilations and to search for high quality spectra in existing databases. If necessary we will obtain new observations for AP combinations poorly covered. The final step will be to re-determine the AP of the primary reference stars homogeneously from the gathered material.

3 Selection of primary AP reference stars

3.1 Criteria

As explained above the primary reference stars must be in the appropriate magnitude range $6<G<13$. This bright magnitude range may not include a sufficient number of intrinsically faint stars like M dwarfs, or rare stars in the solar neighbourhood like metal-poor stars. In order to improve the sampling of the primary grid in these peculiar areas of the AP space, the faint magnitude limit is increased to $G\simeq15$. As far as possible we will avoid spectroscopic binaries and fast rotating stars.

The primary reference stars must cover the whole (Teff, logg, [Fe/H]) space. It is worth noting that the Extended Stellar Parametrizer (ESP) is in charge of peculiar stars which include in particular hot OB stars, chemically peculiar A stars and very cool dwarfs. The ESP WP is thus looking for appropriate reference stars in these AP ranges. That is why we mainly focus on late-type stars (F,G,K,M).

Since the calibration process will directly use Gaia observations of the primary reference stars, we expect the RVS or BP/RP spectra of the selected stars to be of good quality. Together with the magnitude range which results in the SNR, the number of transits and the existence of other objects in the close neighbourhood producing overlapping spectra has to be considered. The expected number of transits over 5 years for BP/RP and/or RVS can be estimated for any stars with GaiaTools. It will then be possible to focus on stars which present the highest number of transits. The question of overlapping spectra has been already studied in the case of RVS for the choice of standard stars to set the zero point of radial velocities as described in Jasniewicz et al. (2009). A RVS spectrum being about 75 arcsec long, it was concluded that the standards should have no neighbour with a magnitude difference smaller than 5 within a distance of 80 arcsec, to allow for all the possible scanning directions. The search of disturbing neighbours is performed automatically within the USNO-B1 catalogue, as available online at the CDS. In the case of BP/RP, this question has been considered for the selection of the standard stars for the absolute flux calibration of BP/RP spectra (SpectroPhotometric Standard Stars or SPSS). Stars with close companions are only accepted if the companions are at least 5 mag fainter (E. Pancino, private communication). We apply the same criterion for AP reference stars. Since the BP/RP windows are much smaller than the RVS one (the size of the longer side, along scan, is 60 pixels of 59 mas), the minimum distance for bright neighbours can be less restrictive. We adopt a minimum separation of 4 arcsec. The pertubation produced by stars belonging to the conjugate Gaia field is difficult to assess. There is currently no tool available for an accurate simulation of this effect. A simple way of mitigating this effect is to ignore the transits for which the conjugate Gaia field falls on the bulge or the disk of the Galaxy when calculating the total number of transits.

There is apparently no requirement to have the reference stars well distributed on the sky.

Finally, in general the primary AP reference stars must have their AP already known which means that their (Teff, logg, [Fe/H]) must have been determined and published in some catalogue or paper. The accuracy of these literature APs is not essential at this stage. We just want to know where approximately stars lie in the AP space in order to perform a good sampling. For all selected star, APs will be redetermined homogeneously afterwards with high quality spectra. We have two main sources to select primary reference stars with known APs : the PASTEL catalogue and compilations of star clusters (see below).

In addition, another set of reference stars will be observed at the Ecliptic Poles as part of the programme for creating astrometric, photometric and spectroscopic reference fields for Gaia. The Ecliptic Pole catalogues Altmann & Bastian (2009) are required for the initial testing of Gaia and its performance to be conducted immediately after launch. During this initial phase, Gaia will adopt a polar scanning law differing from its normal one - thus covering the two ecliptic poles on each orbit. Although the GSP algorithms will not be tested at this early stage of the mission, the high number of transits in these two fields make them ideal to choose reference stars. Magnitudes and positions are currently acquired with several instruments, while a spectroscopic survey at the Southern Ecliptic Pole has started on ESO/FLAMES with the brightest stars being observed with UVES at high resolution.

3.2 The PASTEL catalogue

The PASTEL catalogue by Soubiran et al. (2010) is an update of the [Fe/H] catalogue by Cayrel de Strobel et al. (1997) and Cayrel de Strobel et al. (2001). It is a bibliographical compilation of stellar atmospheric parameters providing (Teff,logg,[Fe/H]) determinations obtained from the analysis of high resolution, high signal to noise spectra. PASTEL also provides determinations of the one parameter Teff based on various methods. Its current version (March 2010) includes 30151 AP determinations for more than 16649 different stars, corresponding to 865 bibliographical references. Nearly 6000 stars have a determination of the three parameters (Teff,logg,[Fe/H]) with a high quality spectroscopic metallicity. PASTEL can be queried through a dedicated Web interface : `http://pastel.obs.u-bordeaux1.fr`. It is also available in electronic form at the Centre de Données Stellaires in Strasbourg through VizieR (`http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=B/pastel`).

For the selection of the primary AP reference stars, we now focus on the 3691 stars in the F0 to M5 temperature range (7500 - 3000K), fainter than G=6, with known logg and [Fe/H]. We split this sample into the bright one (6≤G<10, 2744 stars) and faint one (G>10, 947 stars). We use the polynomial relationship G=f(V,B-V) from Jordi (2009). Figs. 1 and 1 show their respective distribution in the plane logg vs Teff with color codes corresponding to [Fe/H] intervals (multiple APs available for some stars have been averaged).

It is clear that the catalogue mainly suffers from a lack of cool dwarfs with Teff<4300K, albeit the most numerous in the Galaxy, because such stars are intrinsically faint, thus difficult to observe at high spectral resolution. Also their analysis in terms of atmospheric parameters is still a challenge. It is also clear that the bright sample is dominated by stars of nearly solar metallicity while the faint one mainly includes metal-poor stars.

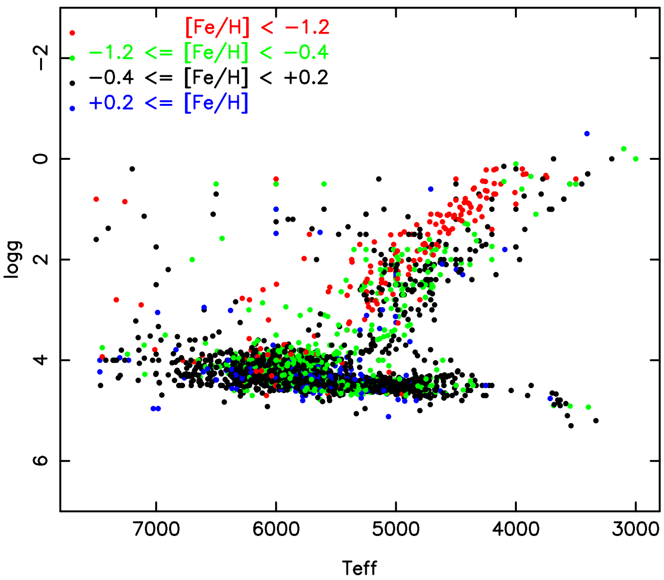

FIGURE 1: Teff vs. logg in several regimes of metallicity for the bright candidates AP reference stars extracted from PASTEL (6≤G<10).

It is recalled that we average the APs compiled from the literature only to select the stars. Such

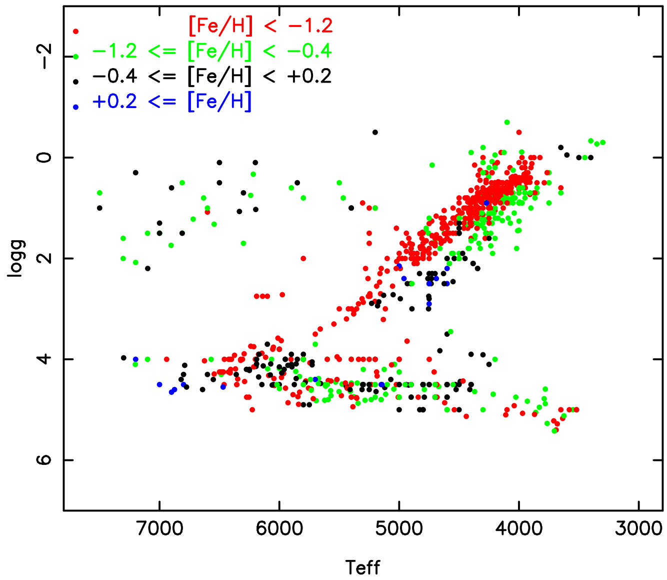

FIGURE 2: Teff vs. logg in several regimes of metallicity for the faint candidates AP reference stars extracted from PASTEL (G≥10).

parameters are inhomogeneous. An homogeneous re-determination of the APs of the most suitable stars will be organized from high-quality spectra, either available in existing databases or obtained from new observations. It is the subject of Sect. 4.

3.3 Primary AP reference stars in clusters

Stars in open and globular clusters are excellent candidates to make primary reference stars. The chemical composition is supposed to be identical for all members (exept a few GC with composite populations). OC and GC span the full range of metallicity. [Fe/H] and Av are available from the literature for a good fraction of them (see Magrini et al. (2009) for open clusters and the Harris catalogue at `http://physwww.mcmaster.ca/~harris/Databases.html`, Harris (1996), for globular clusters). Moreover many FLAMES and UVES spectra are available. The biggest challenge for using individual cluster stars as reference stars is to establish their cluster membership with high confidence. Our current activities are to compile, assess and homogenize metallicities of open clusters, in order to make a selection of AP reference stars in open clusters based on solid grounds. We have started a collaboration with E. Paunzen (Vienna) in order to make the best use of the WEBDA database on open clusters `http://www.univie.ac.at/webda/`. The selection of clusters from which to draw reference stars has to aim at a balance between old/young ages and far/close distances, in order to cover the required ranges in stellar parameters as well as magnitudes. Concerning parameters and spectra of open cluster stars, we also collaborate with DPAC members active in the field of open cluster spectroscopy (A. Bragaglia, S. Randich).

3.4 The Ecliptic Poles Catalogues

An observing programme is currently scheduled on FLAMES at ESO to obtain high resolution spectral data required for the testing of the relevant instruments on board Gaia (RVS) by a complete knowledge of the point sources in a Gaia sky test field, around the SEP. The approximate limiting magnitude for the high resolution spectroscopy will be around V =16 mag. Three FLAMES fields of 25' diameter will be observed, with long and a short exposure, and two different grisms. For each pointing, the UVES fibres will be used to obtain 7 high resolution, high S/N spectra of some of the brightest targets. It gives the opportunity to compile an additional set of 84 AP reference stars at the SEP. These stars will be mainly Milky Way clump giants, LMC red giants and lower main sequence stars. None of these candidates has been studied in spectroscopy previously.

3.5 Other relevant AP lists

As mentioned previously, the PASTEL catalogue suffers from a lack of stars difficult to observe at high spectral resolution with a high SNR, namely metal-poor stars and cool dwarfs. Candidates belonging to these categories can be found in other catalogues, built for instance from low resolution spectroscopy or photometry.

Traditionally, the HK survey by Beers et al. (1992) and Hamburg-ESO survey (HES) by Christlieb et al. (2008) are the main sources of candidates in searches for metal-poor stars. Both surveys are fairly deep (HK B $<$ 15.5, HES B $<$ 17.5) objective prism, low resolution spectroscopic surveys which provided several thousands of candidates. Follow-up medium resolution (R $\sim$ 2000) spectroscopy was used for confirming metal-poor abundance.

Bonfils et al. (2005) have performed the photometric calibration of M dwarf metallicities and provides estimates for 67 stars. Casagrande et al. (2008) have also extensively studied M dwarfs and determined effective temperatures, radii and metallicities for 343 stars. It is worth noting that no reliable metallicity determination is available for dwarfs cooler than 3300K, roughly corresponding to M3. We then consider that colder stars are in the domain of ESP.

3.6 About interstellar extinction and binarity

The interstellar extinction, Av, is one of the parameters to be determined by GSP-Phot. There is thus a need for calibrating it. The field stars which are currently forseen as primary reference stars are on average close enough from the Sun that their interstellar absorption is negligeable. This is however something that has to be investigated in more details. One method to determine colour excesses is by calibration of intrinsic and observed colour indices versus the spectroscopic effective temperature (see (Kovtyukh et al. 2008) for giants and supergiants). For cluster stars, the reddening is generally better known if the absorption is uniform all over the cluster (some clusters in obscured regions of the Galaxy present variations of reddening across the cluster field, e.g. NGC 2129 Carraro et al. (2006) or NGC 3960 Bonato & Bica (2006)). The reddening values of clusters are available for most of the OC where we have selected reference stars. We will also investigate the reddening of these stars with the method of Kovtyukh et al. We are aware that the list of primary reference stars must be delivered with individual estimations of Av. However, this list will not span the full range of possible exctinctions since highly absorbed stars are generally too faint to be studied at high resolution. Other methods have to be considered to train and calibrate GSP-Phot at high absorption.

A good fraction of stars in the Galaxy are binaries or multiple systems. In spectrocopic binaries, the contribution of the secondary in the spectrum of the primary is most of the time unknown, except in the particular case of double lined binaries. The GSP algorithms will have to deal with this problem. We will take care to eliminate as far as possible known spectroscopic binaries from our sample of reference stars since their atmospheric parameters might be uncertain due the the unknown contamination of the spectrum by the secondary. Note that a good fraction of binaries will be flagged thanks to the Gaia astrometry and treated by CU4.

4 Seeking high resolution high SNR spectra

We now focus on the subset of 3691 candidates primary AP reference stars selected in the PASTEL catalogue. We want to retrieve the best spectroscopic material from existing databases in order to re-determine their APs in the most homogeneous and accurate way. The accurate determination of APs implies in principle the use of very high resolution and high SNR spectra ($R > 80000$, $SNR > 200$), covering a wide spectral range in order to efficiently sample the various spectral diagnostics. In practice, we are more conservative on the resolution and SNR in order to gather a sufficiently large number of spectra, allowing $R \simeq 40000$ or $SNR \simeq 80$.

4.1 The ESO archive (HARPS, UVES, GIRAFFE)

The ESO Science Archive Facility (`http://archive.eso.org/eso/eso_archive_main.html`) delivers all the observations made with the ESO instruments after the one year proprietary period. Images and spectra are stored in the Archive in an unreduced (raw) state together with the calibration and auxiliary data. For a fraction of the archive, reduced data can be queried through the Advanced Data Products Query Form (`http://archive.eso.org/eso/eso_archive_adp.html`. We focus on the two échelle spectrographs HARPS and UVES, together with the multi-object spectrograph GIRAFFE which has observed a quantity of star clusters.

HARPS is an échelle spectrograph fed by a pair of fibres and optimised for mechanical stability and built to obtain very high long term radial velocity accuracy (on the order of 1 m/s). The spectral range covered is 378nm-691nm with a resolving power of 115,000. HARPS is equipped with its own pipeline, providing extracted and wavelength calibrated spectra. The reduced spectra provided by the archive are 1D (orders reconneted) resampled at a constant step of 0.01 Å. HARPS is mounted on the 3.6m telescope on La Silla. Since the end of May 2009, HARPS is equipped with a polarimeter option.

UVES is a cross-dispersed échelle spectrograph mounted on the Nasmyth B focus of VLT-UT2. It has several spectroscopic modes which give a resolution up to 110000, over a 400 nm range within 420 nm to 1100 nm. Note that this includes the RVS wavelength range, but until March 2005, the corresponding standard setting (reference wavelength 860 nm) had a gap from 854 to 866 nm. A new setting with reference wavelength 760 nm and continuous wavelength coverage was introduced in April 2005.

GIRAFFE is a medium-high (R=5600-46000) resolution spectrograph for the entire visible range, 370-950 nm. It is one of the components of the FLAMES instrument mounted at the Nasmyth A platform of VLT-UT2, which can access targets over a large corrected field of view (25 arcmin diameter). GIRAFFE is equipped with two gratings and several filters are available to select the required spectral range, giving 39 fixed setups (31 high resolution + 8 low resolution modes). The MEDUSA fibers allow up to 132 separate objects (including sky fibres) to be observed in one go. An archive of reduced GIRAFFE spectra is available at `http://giraffe-archive.obspm.fr/`.

4.2 The SOPHIE archive

SOPHIE is a cross-dispersed échelle spectrograph attached to the 1.93-m telescope at OHP since September 2006. The spectra cover the wavelength range 3872-6943 Å. The instrument is entirely computer-controlled and a standard data reduction pipeline automatically processes the data upon CCD readout. For late-type stars (F,G,K,M) this includes radial velocities by numerical cross-correlation techniques which can yield very accurate velocities (down to 2-3 m/s using simultaneous Th calibration), depending on the signal-to-noise ratio. There are two spectral resolutions : a high-resolution at R=75000 and a high efficiency mode at R=40000 (average resolving powers near 5500 Å).

The SOPHIE Archive (`http://atlas.obs-hp.fr/sophie/`) gives access to all science observations obtained with the spectrograph since it started operations. Data enter the database one day after they are obtained. Normal protection is for one year. Spectra with an extended 5-year protection are still released after one year but with the time information masked. In its present version the database will produce, upon interrogation, a list of all existing observations for a given object or list of objects together with pointers to the extracted (e2ds) and reconnected (s1d) spectra, if public.

Instrument	Resolution	Date	S/N
SOPHIE	75000	2007-01-06	800
ELODIE	42000	1997-01-24	500
NARVAL	65000	2007-03-14	1300
HARPS	115000	2007-11-06	400

TABLE 1: Characteristics of Procyon spectra shown in Figs. 3 and 4.

Status of the Archive in May 2010 : The number of indexed spectra in the database is 38036. There are 3314 distinct object designations, 32649 spectra are publicly available and 5387 are protected.

4.3 The ELODIE archive

ELODIE is a cross-dispersed échelle spectrograph which has been in operation on the 1.93-m telescope at OHP from June 1993 to August 2006. The spectra cover a 3000 Å wavelength range (3850-6800 Å) with a spectral resolution of 42000. The ELODIE archive (`http://atlas.obs-hp.fr/elodie/`) contains nearly 35000 spectra with S/N $>$ 5 in V (5500 Å). Close to 90% of the spectra in the archive are now public. There are more than 7700 distinct object names, but the actual number of distinct objects is probably closer to 3000, due to the use of different aliases for the same star that are not yet resolved. A cross-identification process is underway to remove these aliases.

4.4 The NARVAL archive

Narval is a spectro-polarimetre, mounted on the 2.0m TBL on Pic du Midi, allowing the users to observe simultaneously the spectrum of a point object in 2 polarization states over a spectral range of 370 to 1000 nm, with a resolution of 65000 in polarimetric mode and 65000 or 80000 in spectroscopic mode. NARVAL is thus particularly interesting because its spectral range covers that of RVS. Narval is operational since the end of october 2006. The NARVAL archive can be queried at `http://tblegacy.bagn.obs-mip.fr/narval.html`. In March 2010, it includes 16352 spectra of 449 different stars.

4.5 Query results

To give an idea of the quality of spectra available in the different archives, we show the spectrum of Procyon which is available in some of them over the full spectral range available (Fig. 3), as well in a narrow region around the MgI triplet 516-519nm (Fig. 4). The spectrum characteristics are listed in Table 1.

Table 2 gives the number of stars in the PASTEL selection which have at least one spectrum

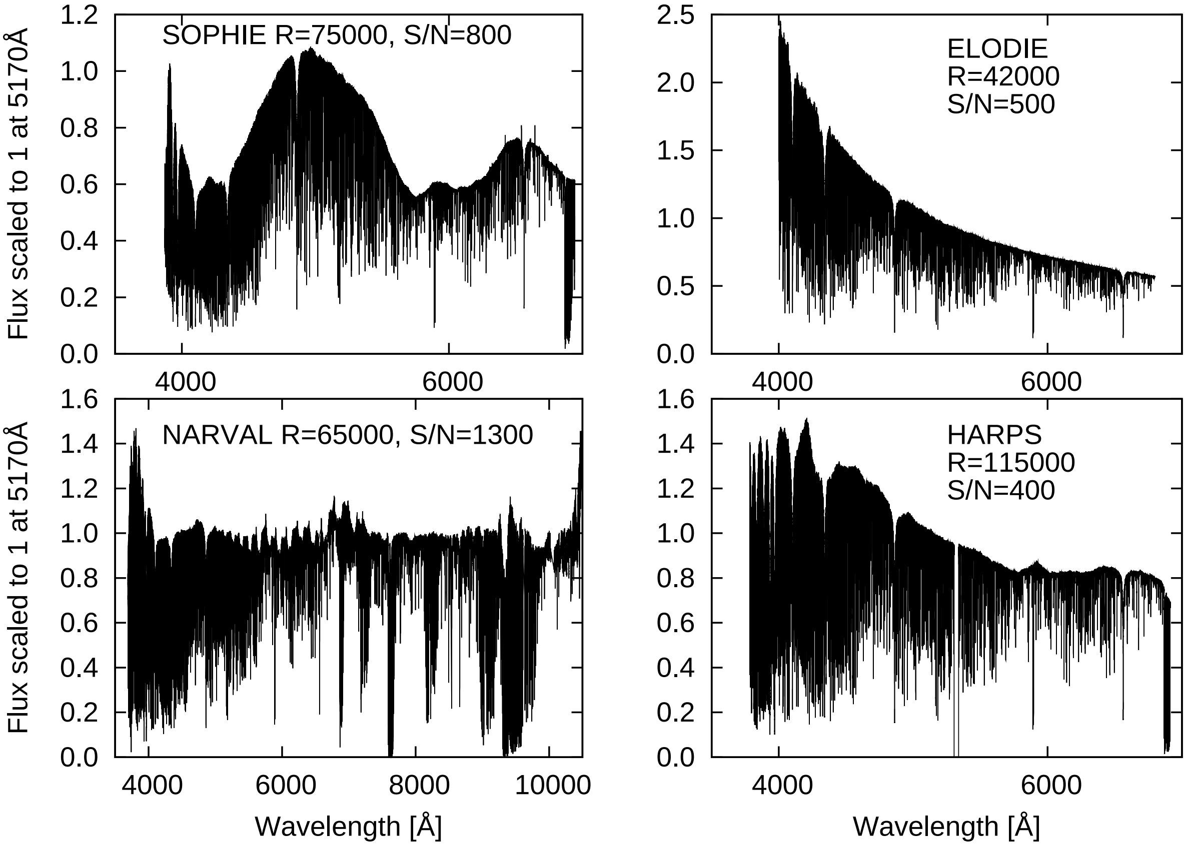

FIGURE 3: Reduced high-resolution spectra of Procyon extracted from four on-line archives.

available in each archive. These numbers are first estimates. Because PASTEL and the spectrograph archives are updated several times per year, these numbers are expected to grow.

4.6 Other relevant datasets

A number of on-line spectroscopic databases have been produced by groups around the world studying stars in the solar neighborhood. The website of the *Nearby Stars Project* (`http://bifrost.cwru.edu/NStars/`) gives access to reduced high-resolution, high signal-to-noise spectra and astrophysical parameters for 253 dwarf stars published in Luck & Heiter (2006). The spectra are taken with the Sandiford Echelle Spectrograph attached to the 2.1m Otto Struve Telescope at McDonald Observatory. They cover a wavelength range of 4800–6900 Å at a resolution of 60000.

The on-line database of the *Spectroscopic Survey of Stars in the Solar Neighborhood* (S4N, `http://hebe.as.utexas.edu/s4n/`) provides reduced high-resolution spectra for 118

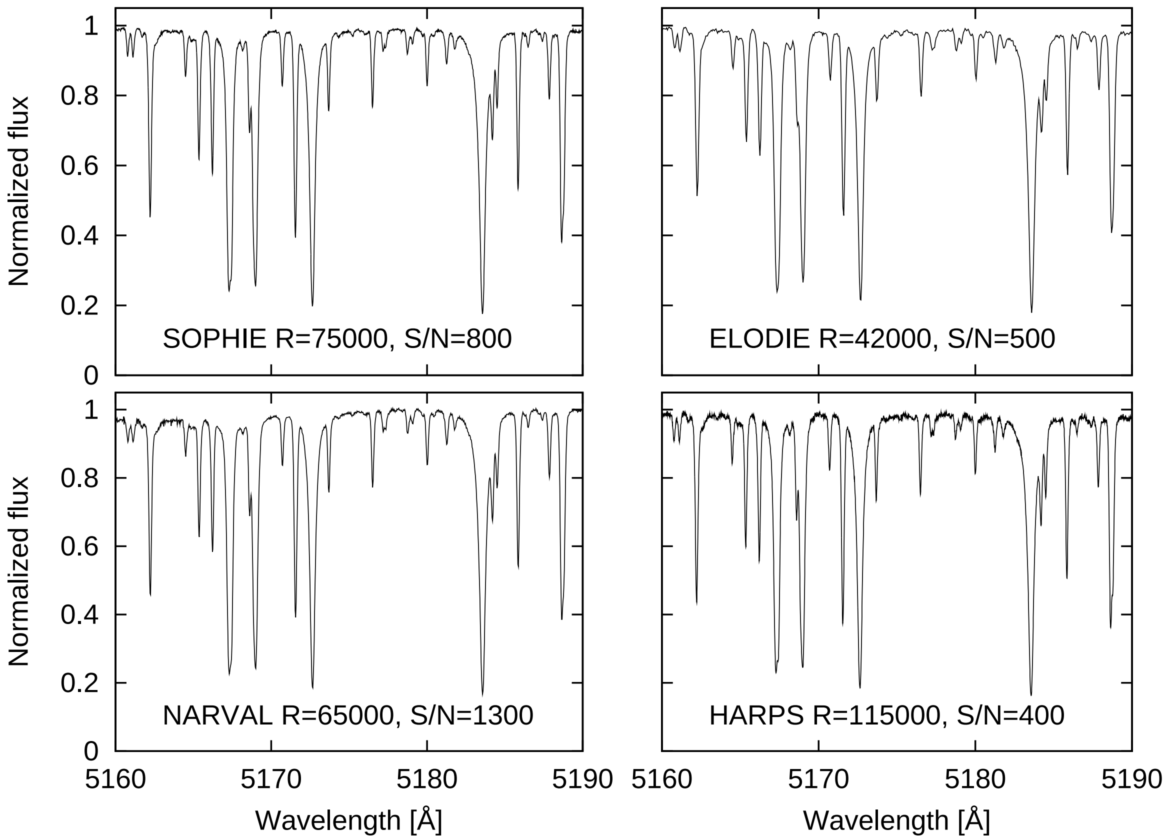

FIGURE 4: Same as Fig. 3, but for a narrow region around the MgIb triplet (516-519 nm).

nearby stars. These spectra, taken with the 2dcoudé spectrograph at the 2.7m Harlan J. Smith Telescope at McDonald Observatory and with the FEROS spectrograph at the ESO 1.52m telescope, cover a wavelength range of 3600–9200 Å, thus including the RVS range, at a resolution of 50000. Details of the data acquisition and reduction, as well as an analysis of the kinematics and chemical abundances are published in Allende Prieto et al. (2004).

Reduced spectra for 160 F-K type stars, which have been collected with the HIDES spectrograph attached to the 188cm reflector at Okayama Astrophysical Observatory, are available for download at `ftp://dbc.nao.ac.jp/DBC/ADACnew/J/other/PASJ/57.13/`. These spectra have a resolution of 70000 and cover three wavelength ranges (5000–6200Å, 5800–7000Å, 7600–8800Å). See Takeda et al. (2005) for more details.

In Fig. 5), we compare the spectrum for Procyon from the S4N and Takeda databases in the RVS spectral range with a spectrum from the NARVAL archive (the same as in Fig. 3).

Further spectral data are in principle available from the archives of several observatories, usu-

Archive	bright	faint
SOPHIE	301 (1875)	0
ELODIE	538 (3110)	0
NARVAL	50 (985)	0
HARPS	249 (2804)	36 (1309)
UVES	518 (5362)	163 (2809)
GIRAFFE	0	33 (111)

TABLE 2: Number of bright and faint candidates having a spectrum available in each of the considered archives. In brackets, the total number of spectra available for them. For HARPS the S/N of the spectra is not given in the query results so that we don't know whether all the available spectra are of good quality. For GIRAFFE, we have considered the configurations leading to a resolution higher than 16000.

ally in raw format. Here, we mention the archive for data taken with the 3.58m TNG on La Palma, which can be queried through a web interface at `http://wwwas.oats.inaf.it/IA2/index.php?option=com_wrapper&Itemid=27`. In particular, choosing the instrument option "SRG", one can retrieve spectra taken with the SARG spectrograph, which is designed for the spectral range from 370 up to 900 nm, and for resolutions from R=29000 up to R=164000. To date (2009-11-12), 17985 SARG object spectra are stored in the archive. Unfortunately, the web interface doesn't allow to query for lists of stars and it doesn't include the information on the resolution. Therefore it is of limited use in its current form.

5 Storage and future spectral analyses

5.1 The CU6/CU8 database

The spectra gathered in the different archives will be stored in the dedicated CU6/CU8 database (restricted access) : `http://gaia.aip.de/cu6a8/`. Required for both CU6 and CU8, the database is aimed at storing the astrophysical parameters (for CU8), the radial velocities (for CU6) and the raw and reduced data (both CU6 and CU8) from ground-based observations for a large number of reference stars that will be observed during the mission. The database consists of a set of four tables which can be accessed and updated via a web interface, plus a local disk storage. The Basic Data Table stores external information like identifiers, astrometry, photometry, spectral classification, variability information. The RV measurements Table stores radial velocities and associated quantities/description to assess the radial velocity stability for both stars and asteroids. The Stellar parameters measurement Table stores measured stellar parameters for a given set observation/method, either from the literature (PASTEL) or from our own determinations. The Observation Table contains a basic description of the observation and associated spectra. The full description of the observation can be accessed via the web interface from the FITS headers of each file. There is also an Asteroid ephemeris table. Calibrated and

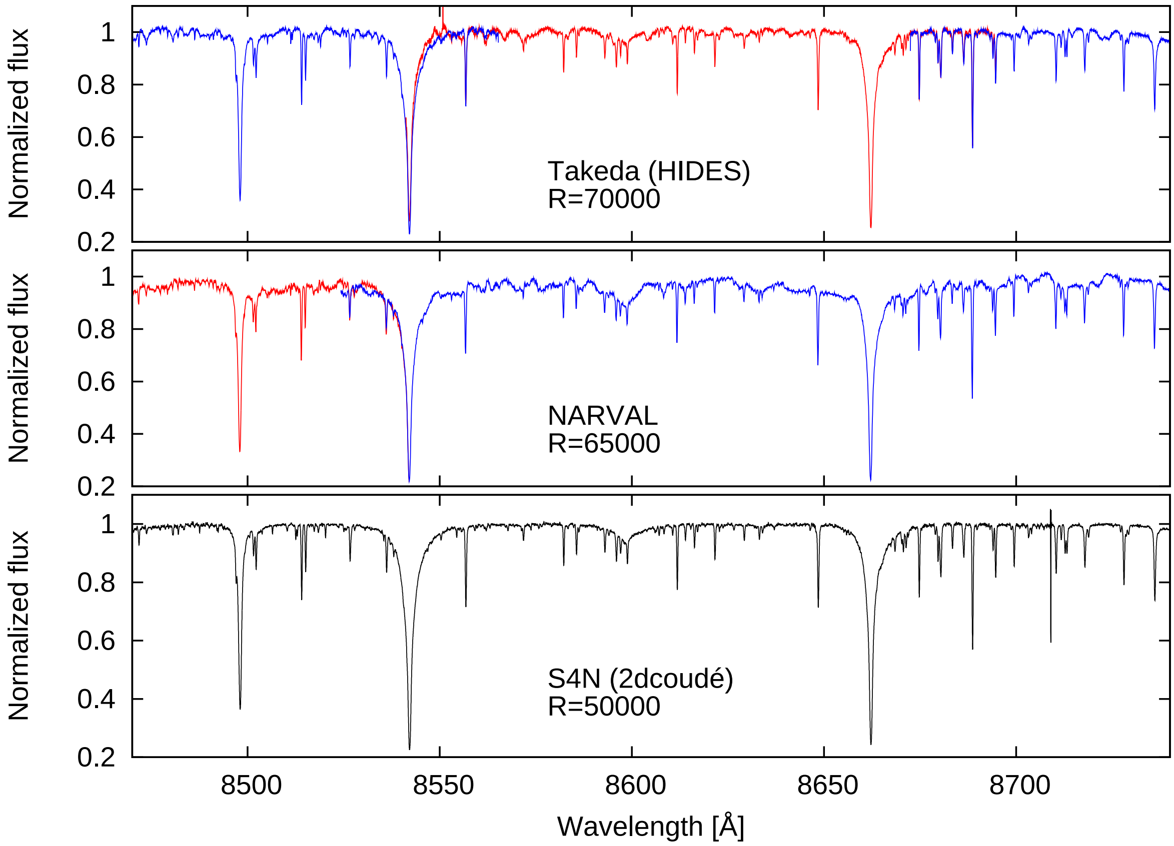

FIGURE 5: Reduced high-resolution spectra of Procyon extracted from the NARVAL archive, from the S4N, Allende Prieto et al. (2004), and from the Takeda et al. (2005) databases, for the RVS spectral range (847–874 nm).

raw ground-based spectra are stored in a local disk storage.

5.2 SME : Spectroscopy Made Easy

Since we will have several hundreds of spectra to analyse, we will have to use a quasi-automated method to determine homogeneously their APs. The basis will be the Spectroscopy Made Easy (SME) code of Valenti & Piskunov (1996) which uses a non-linear least squares parameter estimation algorithm to match observed spectra with synthetic spectra in selected comparison wavelength regions. We will use the best model atmospheres available to compute the synthetic spectra and the most appropriate line lists, following the conclusions of the expert panel on stellar physics within GWP-S-811-10000 ”Provide synthetic stellar spectra”. The SME code takes all relevant input data (spectrum, line list, initial guess stellar parameters and abundances) and provides optimised parameters without any need for user interaction. The main tasks of the interface software are, in a completely automated fashion, to extract relevant spectral regions appropriate for the line list, to identify continuum points and normalise the spectra relative to the continuum, to make small adjustments (within the error of the wavelength calibration) to line central wavelengths such that they match the observed line centres, and to reject lines polluted by artifacts such as cosmic ray hits. The atmospheric parameters and abundances are solved using a list of spectral features, atomic and molecular data and comparison regions assembled for the stars of interest. The results are then extracted, error estimates calculated and valid detections assigned, again by an automated procedure. Since we want to obtain homogeneous APs from different sets of spectra, we will take care to have stars in common with the various instruments, especially benchmark stars. The method will be calibrated with the benchmark stars which will define our parameter scale.

6 Conclusions

We have presented our strategy to provide a sample of primary reference stars for the external calibration of GSP algorithms. It consists in three steps : (i) select bright and faint stars in the PASTEL database and other relevant sources, (ii) gather available high-resolution, high signal-to-noise spectra, (iii) provide homogeneous APs of the selected stars with the available spectra using the SME software. We are now in the phase where we explore the archives of several échelle spectrographs, while an observing programme in on-going on NARVAL. We have identified nearly 4000 stars in PASTEL with known atmospheric parameters.

Acknowledgments

We thank the Bordeaux 1 University for providing U. Heiter's a one-month position in 2009 and 2010. We thank the french Programme National de Physique Stellaire for supporting our observing programme on NARVAL. We thank the observatories which have developed a spectroscopic archive to provide a quantity of high quality science ready stellar spectra.

References

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

Altamnn M., Bastian U., 2009, Ecliptic Poles Catalogue Version 1.1, URL http://www.rssd.esa.int/llink/livelink/open/2885828

Bailer-Jones C., 2009, A procedure to calibrate algorithms for estimating parameters from spectra, URL http://www.rssd.esa.int/llink/livelink/open/2878794

Bailer-Jones C. 2010, MNRAS, 403, 96

Beers, T. C., Preston, G. W., & Shectman, S. A. 1992, AJ, 103, 1987

Bonato C. & Bica E. 2006, A&A, 455, 931

Bonfils, X., Delfosse, X., Udry, S., Santos, N. C., Forveille, T., & Ségransan, D. 2005, A&A, 442, 635

Carraro, G.; Chaboyer, B.; Perencevich, J. 2006, MNRAS, 365, 867

Casagrande, L., Flynn, C., & Bessell, M. 2008, MNRAS, 389, 585

Cayrel de Strobel G., Soubiran C., Ralite N., 2001, A&A, 373, 159

Cayrel de Strobel G., Soubiran C., Friel E.D., Ralite N., Francois P., 1997, A&AS, 124, 299

Christlieb, N., Schörck, T., Frebel, A., Beers, T. C., Wisotzki, L., & Reimers, D. 2008, A&A, 484, 721

Edvardsson, B. 2008, Physica Scripta Volume T, 133, 014011

Harris, W. E. 1996, AJ, 112, 1487

Heiter U., Soubiran C., Korn A., 2009, Calibration of the General stellar parameterizer algorithms using ground-based observations, URL http://www.rssd.esa.int/llink/livelink/open/2864468

Jasniewicz G and contributors, 2009, DU640 Radial Velocity Zero-point Software Design Document, URL http://www.rssd.esa.int/llink/livelink/open/2776876

Jordi C., 2009, Photometric relationships between Gaia photometry and existing photometric systems, URL http://www.rssd.esa.int/llink/livelink/open/2760608

Kovtyukh, V. V.; Soubiran, C.; Luck, R. E. et al., 2008, MNRAS, 389, 1336

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

Magrini, L., Sestito, P., Randich, S., & Galli, D. 2009, A&A, 494, 95

Soubiran C., Le Campion J.-F. , Cayrel de Strobel G. , Caillo A., 2010, A&A, in press.

Takeda, Y., Ohkubo, M., Sato, B., Kambe, E., & Sadakane, K. 2005, PASJ, 57, 27

Valenti, J. A., & Piskunov, N. 1996, A&AS, 118, 595