Calibration of the General Stellar Parametrizer algorithms using ground-based observations

Definition of calibration data needs and candidate reference stars

prepared by: U. Heiter, C. Soubiran, A. Korn
approved by: C.A.L. Bailer-Jones

reference: GAIA-C8-TN-UAO-UH-001-1
issue: 1
revision: 1
date: 2008-11-12
status: Issued

Abstract

This document serves as a reference for 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 covers some aspects of the first three tasks of that work package. These are the first steps towards the two-fold aim of the calibrations work package: (1) establish an optimal list of *reference stars* that will be observed by Gaia to calibrate the GSP algorithms, (2) provide libraries of empirical spectra and related APs to correct synthetic spectra (*training spectra*). Some thoughts on the calibration procedures for the algorithms as well as the expected final accuracy of the AP determinations are included as well. A major part of this document is devoted to possible source catalogues for selection of AP reference stars. We focus on recent high-spectral-resolution surveys of bright field stars as well as faint open cluster stars. Globular clusters and large catalogues of calibrated spectrophotometric studies are briefly discussed as well.

Document History

Issue	Revision	Date	Author	Comment
1	1	2008-11-11	UH	Final revisions according to Caroline; SDSS-DR6
D	12	2008-10-17	UH	Minor revisions according to Andreas
D	11	2008-10-16	UH	Added source catalogue info from Caroline
D	10	2008-10-15	UH	Source catalogue tables; histogram figure; conclusions
D	9	2008-10-07	UH	Restructuring; major revision of OC section
D	8	2008-10-02	UH	Minor revision of abstract; removed GBOG section
D	7	2008-09-07	UH	Revised open cluster tables
D	6	2008-05-20	UH	Changed title, added abstract, revised introduction
D	5	2008-02-08	UH	Modified primary grid tables
D	4	2008-02-01	UH	Revision according to Andreas
D	3	2008-01-10	UH	Added text; revisions according to Caroline
D	2	2007-12-11	UH	Added text; revisions according to Andreas
D	1	2007-11-23	UH	Creation

Acronyms

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

Acronym	**Description**
AGB	Asymptotic Giant Branch (star)
AO	Announcement of Opportunity
AP	Astrophysical Parameter
DPAC	Data Processing and Analysis Consortium
DSC	Discrete Source Classification (Classifier)
ESO	European Southern Observatory
ESP	Extended Stellar Parametriser
GBOG	Ground-Based Observations for Gaia (DPAC)
GC	Globular Cluster
GOG	Gaia Object Generator
GSP	Generalised Star Parametrisation (Parametriser)
GSPphot	Generalised Stellar Parametriser PHOTometry
GSPspec	Generalised Stellar Parametriser SPECtroscopy
NOT	Nordic Optical Telescope
OC	Open Cluster
RAVE	RAdial Velocity Experiment
RVS	Radial Velocity Spectrometer
SDSS	Sloan Digital Sky Survey
SNR	Signal-to-Noise Ratio (also denoted SN and S/N)

Contents

1 Introduction

One of the goals of CU8 is to provide the astrophysical parameters (APs), $T_{\rm eff}$ (effective temperature), log g (surface gravity) and [M/H] (metallicity), for those sources which the Discrete Source Classifier (DSC) identifies as single normal stars. The related parametrisation algorithms (General Stellar Parametrisers GSP-phot and GSP-spec) rely on templates which are built from grids of synthetic spectra. The calibration of these algorithms requires that Gaia obtains good data ("good" meaning high S/N, no overlap between spectra, etc.) for a set of stars representative of the whole parameter space, for which the APs are well known in advance. We refer to these stars hereafter as *AP reference stars*. Moreover, as the parametrisation is based on stellar atmosphere models which are not perfect, empirical spectra of calibration stars (hereafter referred to as *training spectra*) will be used to test and improve synthetic spectra.

The objective of work package GWP-S-811-20000 – "Provide calibrations of training data" – is the selection of the best stars that can be used as AP reference stars. This activity also deals with the compilation of data – spectra and atmospheric parameters – to be used for homogeneous and detailed characterization of the AP reference stars, and as training spectra. For the selection of AP reference stars and the search for useful data, we have to consider the resolution, spectral interval and magnitude ranges of the Gaia spectroscopy with the BP/RP and RVS instruments. We also have to take into account the final AP accuracy expected to be obtained by the algorithms.

In this document, we present the basic idea for a possible calibration procedure, which is to build a *primary* grid of AP reference stars, calibrated on *benchmark* stars, as well as a *secondary* grid, calibrated on the *primary* grid stars. The analysis will be based on existing data as a starting point, but new ground-based observations are mandatory.

2 Selection criteria for AP reference stars

2.1 Parameter space

2.1.1 Parameter ranges

For now, this document deals with cool stars – F,G,K stars (it might be extended to M stars in the future). Stars of these types will be the most numerous among the objects observed by Gaia, and they span a large range in stellar parameters such as metallicity and age. Therefore, their AP calibration will be crucial for achieving the mission goal of studying the structure and evolution of the Galaxy. We assume here that such stars are covered by the ranges in the parameters $T_{\rm eff}$, $\log g$ and [M/H] given in Table 1. The lower limit for $\log g$ corresponds to a rough estimate of the Eddington limit and depends slightly on $T_{\rm eff}$. The metallicity limits are preliminary and correspond to the metallicity range covered by available stellar catalogues. Stars with very low metallicity might require follow-up ground-based observations, or special treatment by the parametrization algorithms.

TABLE 1: Ranges in the parameters $T_{\rm eff}$ (effective temperature), $\log g$ (surface gravity) and [M/H] (metallicity) assumed to cover the regime of FGK stars. Ranges are given in the form [lower limit:upper limit]. The ranges in $(B-V)_0$ are for guidance only (and used for reference star selection below), and the lower and upper limits are set very generously.

$T_{\rm eff}$	log$T_{\rm eff}$	$(B-V)_0$	$\log g$	[M/H]
[4000:5000]	[3.60:3.70]	[1.6:0.9]	[−0.5:5.0]	[−3.0:+0.5]
[5000:6000]	[3.70:3.78]	[0.9:0.6]	[+0.0:5.0]	[−3.0:+0.5]
[6000:7000]	[3.78:3.85]	[0.6:0.2]	[+0.5:5.0]	[−3.0:+0.5]

2.1.2 Step sizes in parameter grid

Contrary to grids of synthetic spectra, grids of real stars cannot be built with exact regular steps in the three parameters. However, we can define a parameter grid and try to find stars which fall near the grid nodes. For setting the step sizes in this parameter grid we have to take into account several kinds of uncertainties:

1. The uncertainty inherent to the reference APs. We can distinguish stars whose APs rely on high resolution, high S/N spectroscopy, and low resolution spectroscopy or photometry.
2. The typical uncertainty of the APs determined in the Gaia data processing. They will mainly depend on the V (G) magnitude and instrumental effects (i.e. S/N) and details of the algorithms.

3. The required accuracy of the APs to achieve the Gaia science goals.

As the data processing algorithms are currently under development, definitive values for the uncertainties of the second kind are not yet available. Also, there is no official requirement for the AP accuracies (“as good as possible”, item 3). Thus, for the purpose of this document, we try to consolidate different values which have been mentioned in various documents. These values are listed in Table 2, together with the values adopted here. In the “Proposal for the Gaia Data Processing” (April 2007, DPAC; aka “AO response document”) , one can find two references to the expected accuracies of the astrophysical parameters, based on preliminary evaluations of parametrisation performance. In the Gaia Concept and Technology Study Report (July 2000), the section on “photometric requirements” states: “To be able to reconstruct Galactic formation history the distribution function of stellar abundances must be determined to $\sim$ 0.2 dex, while effective temperatures must be determined to $\sim$ 200 K”. No requirements for the precision were found in the Mission Requirements Document, ESA Gaia Project (2006). Thus, we adopt the following goals for the precision of the parameter determination in this document: 2% for $T_{\rm eff}$, 0.3 dex for $\log g$, and 0.2 dex for the metallicity of most stars. Note that to achieve the stated accuracy for $\log g$ at all magnitudes will presumably require the use of astrometric data.

In particular for studies of the Galactic halo, it will be important to derive the abundance of α elements, [α/Fe], from Gaia data. The target precision for this parameter is around 0.2 dex. However, we leave the details concerning the calibration of this parameter to a forthcoming document. For the interstellar extinction, the “AO response document” gives values for the estimated precision of between 0.05 and 0.3 mag. The current document will not discuss the calibration of this parameter either.

TABLE 2: Values of required/expected precision for parameter determination found in various documents (AO response . . . “AO response document” (April 2007); Concept Study . . . Gaia Concept and Technology Study Report, July 2000). Values for $\log g$ and [M/H] are in dex.

Document	Section	pages	$T_{\rm eff}$	$\log g$	[M/H]
AO response	2.3.1.2[a]	41	1.5%	0.25	0.1
AO response	5.5.5[b]	163 (Fig. 40), 165 (Tab. 7)	4.3%	0.65	0.42
Concept Study	2.3.1	137	200 K	–	0.2
Adopted	–	–	2%	0.3	0.2

[a]Figure of merit evaluation of parametrisation performance, predicted errors for an F-type sub-giant star, G=17, solar metallicity

[b]Estimates of precision using a flexible nonlinear, multidimensional regression algorithm, for cool stars, G=15

R=11500		R=5500	
V	SNR	V	SNR
8	420	12	100
9	260	13	55
10	160	14	25
11	95	15	12

TABLE 3: End-of-mission signal-to-noise ratios (SNR) per sample for RVS spectra, estimated for a K1V star for various V magnitudes (David Katz, 2007-11-14, priv. comm.).

2.2 Visual magnitudes of reference stars - bright vs. faint stars

The Gaia instruments will obtain observations of stars in different magnitude ranges. The Gaia photometers will saturate at around 10th magnitude, while the radial-velocity spectrograph will provide data useful for AP determination only for stars brighter than about 11th to 13th magnitude (depending on resolution, see Table 3). Stars brighter than 6th magnitude will not be observed by Gaia. This results in two different sets of stars for AP calibration purposes, *bright stars* with $6 < V < 13$ and *faint stars* with $10 < V < 18$. Stars within both magnitude ranges will be considered within this document. Selected calibration stars may fall outside these magnitude limits, notably very bright stars. These will provide an indirect calibration of parametrization methods for Gaia via improvement in the theory of stellar atmospheres (see Sect. 3 below). We refer to such stars as *benchmark stars* and the corresponding ground-based data can contribute to the set of training spectra. Stars which fall within the magnitude limits (the reference stars) can be used for direct calibration of the parametrization methods after the launch of Gaia, by comparing APs determined by the Gaia algorithms and independently, and possibly by replacing input spectra used for training the algorithms with actual observations.

2.3 Position of reference stars – ecliptic-poles fields

Two fields of one square degree around the north and south ecliptic poles will be observed extensively during the commissioning and testing phase of the mission. As many as possible of the AP reference stars should be located within these fields. Thus, when selecting stars from existing catalogues or when planning new ground-based observations, the positions of the stars should be taken into account in addition to magnitudes and physical parameters. Also, a cross-correlation with the Ecliptic-poles star catalogue, which is being built up by M. Altmann (CU3) has to be carried out [1].

[1] http://www.rssd.esa.int/SA-general/Projects/GAIA/wiki/index.php?title=CU3:_Auxiliary_Data:_Ecliptic-poles_star_catalogue

3 Calibration procedure - a three-level approach

A simple combination of the values in Sect. 2.1 results in almost 10000 stars. Homogeneous observations at high spectral resolution and high signal-to-noise ratio do not exist for such a large number of stars, and cannot be obtained within the program of ground-based observations for Gaia. Even if this were possible, detailed spectrum analyses are neither available nor feasible before the launch of Gaia. Therefore we propose a *three-level approach* for calibration.

On the first level, we define a set of "benchmark stars". This consists of a small number (order of 10) of carefully selected, well-known bright stars. For the *benchmark stars*, a set of homogeneous data comprising very-high-resolution spectroscopy, spectrophotometry[2] and photometry will be obtained. These data will be analysed using the best available atmospheric models and input data. Standard analyses similar to those used for the *primary grid* (see next paragraph) will also be done in order to determine the necessary corrections for the results of such analyses. This work is done by the Gaia-SAM group and is related to Gaia work package GWP-S-811-10020 ("Expert Panel on Stellar Physics"). For more details we refer to the Gaia-SAM webpage[3] and the GaiaWiki page on Benchmark stars[4].

On the second level, we define a "primary grid" of reference stars, which will be studied in detail, based on high-resolution spectroscopy. The *primary grid* stars will be calibrated using the results obtained for the *benchmark stars*. Published spectroscopic data and standard analyses found in the literature will be used as far as available and will be complemented by new ground-based observations and new analyses. To link the *primary grid* to the *secondary grid* (see next paragraph), we also need spectrophotometric data with a similar resolution and photometric data in the same systems for all *primary grid* stars. We define the following steps in parameters for the *primary grid*: 7% for $T_{\rm eff}$, 1.0 dex for $\log g$, and 0.5 dex for the metallicity. This results in a managable number of 560 stars. The sources considered for selection of the *primary grid* stars are discussed in Sect. 4.1.

On the third level, we have the large grid of stars corresponding to our adopted values of parameter ranges and steps, which we call the "secondary grid". This grid will be used to calibrate the synthetic spectra used for algorithm training. The *secondary grid* itself will be calibrated using the results obtained for the *primary grid* stars. Stars will be selected from published analyses and calibrations of large surveys (10 000s of stars), like that of Allende Prieto et al. (2006). Ideally, the *primary grid* would be a subset of the *secondary grid*, but this may be difficult to realize (see currently considered sources for star selection in Sect. 4.2). The photometric or spectrophotometric data already available for these stars will be used to link the physical parameters to the synthetic spectra used for algorithm training. This requires to compute synthetic observables from the corresponding models.

[2] the term "spectrophotometry" is used as a synonym for low-resolution spectroscopy ($R \approx 2000 \ldots 5000$)

[3] `http://www.astro.uu.se/~ulrike/GaiaSAM/index.html`

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

4 Sources for AP reference stars

4.1 Primary grid

4.1.1 Field stars

TABLE 4: Source catalogues for selection of *primary grid* stars: type of data and references.

Catalogue name	Catalogue description	d_{max} [pc]	R	S/N	$[\lambda_{\mathrm{min}}:\lambda_{\mathrm{max}}]$ [nm]	References
PASTEL	Compilation, 450 references, 1990 and later					Soubiran et al. (in preparation), based on Cayrel de Strobel et al. (2001)
Taylor	Compilation, 340 references, 1960 and later					Taylor (1999), Taylor (2005)
Twarog spec	Compilation, see text					Twarog et al. (2007)
Gaia ESP cool	Compilation, see text					A. Lanzafame (priv. comm.)
Fuhrmann	Volume-limited	25	60 000	200	[420:750]	Fuhrmann (1998, 2000[a], 2004, 2008)
Luck & Heiter	Volume-limited dwarf[b], giant[c] and planetary host stars	[b]15, [c]100	60 000	>150	[475:685]	Luck & Heiter (2006, 2007)
Mishenina	Clump giants	361	42 000	130–230	[440:680]	Mishenina et al. (2006)
Ramírez	Nearby F–K stars	150	45 000–120 000	100–600	[450:780]	Ramírez et al. (2007)
S4N	Volume-limited	14.5	50 000	150–600	[362:921]	Allende Prieto et al. (2004)
SPOCS	Nearby F–K stars from planet search programs	202	70 000	300	[483:618]	Valenti & Fischer (2005); Takeda et al. (2007)
Takeda	Nearby F–K dwarfs and subgiants	53	70 000	100–500	[500:700]	Takeda et al. (2005)

[a] http://www.ing.iac.es/~klaus/

Bright stars within the *primary grid* will be selected from various catalogues of analyses of field stars. A list of source catalogues under consideration is given in Tables 4 and 5. Abstracts

TABLE 5: Source catalogues for selection of *primary grid* stars: number of stars and parameter ranges. In addition to the data for the full catalogue, the number of stars and corresponding parameter ranges are listed for stars fainter than V=6 (RVS).

Catalogue name	V range	N	$T_{\rm eff}$ range	log g range[a]	[M/H] range
PASTEL database	[0:15]	4245	[2710:32600]	[−1:5.4]	[−4.8:+1.1]
(FGK)	[0:15]	3792	[4000:7000]	[−1:5.1]	[−4.8:+0.8]
(RVS)	[6:15]	2490	[4000:7000]	[−1:5.1]	[−4.8:+0.6]
Taylor – dwarf stars	[0:14]	1331	[4050:6875]	[IV:V]	[−1.8:+0.5]
(RVS)	[6:14]	1004	[4050:6850]	[IV:V]	[−1.8:+0.5]
Twarog spec	[0:11]	1801	[4130:7375]	[3.0:5.1]	[−2.0:+0.6]
(RVS)	[6:11]	1502	[4130:7200]	[3.1:5.1]	[−2.0:+0.6]
Gaia ESP cool stars	[4:16]	150	[3800:6025]	—	—
Fuhrmann	[3:10]	296	[4800:6700]	[3.0:4.7]	[−2.1:+0.5]
(RVS)	[6:10]	183	[4800:6350]	[3.1:4.7]	[−2.1:+0.5]
Luck & Heiter (2006)	[0:10]	217	[4100:7300]	[3.1:5.0]	[−1.3:+0.6]
(RVS)	[6:10]	108	[4100:7100]	[3.7:5.0]	[−1.3:+0.6]
Luck & Heiter (2007)	[0: 8]	298	[4300:8000]	[1.7:4.4]	[−0.6:+0.3]
(RVS)	[6: 8]	86	[4600:8000]	[2.3:4.3]	[−0.6:+0.3]
Mishenina	[3: 9]	177	[4300:5500]	[1.5:3.2]	[−0.7:+0.3]
(RVS)	[6: 9]	48	[4600:5300]	[1.7:2.9]	[−0.6:+0.2]
Ramírez	[0:11]	523	[4300:6600]	[2.0:4.7]	[−1.5:+0.5]
(RVS)	[6:11]	392	[4700:6400]	[3.2:4.7]	[−1.5:+0.3]
S4N	[0: 7]	118	[4160:7650]	[1.9:4.9]	[−0.9:+0.5]
(RVS)	[6: 7]	27	[4420:5350]	[4.5:4.7]	[−0.9:+0.3]
SPOCS	[0:11]	1039	[4700:6600]	[3.1:5.1]	[−1.9:+0.5]
(RVS)	[6:11]	901	[4700:6600]	[3.1:5.1]	[−1.9:+0.5]
Takeda	[3: 7]	160	[5000:7000]	[3.2:4.9]	[−1.3:+0.5]
(RVS)	[6: 7]	54	[5100:6700]	[3.6:4.9]	[−1.3:+0.5]

[a]or luminosity class range

of the references can be accessed from UH's ADS private library "AP reference stars"[5]. The PASTEL database contains all (non-compilation) catalogues listed in Table 4. The Taylor (2005) compilation contains Allende Prieto et al. (2004) and Fuhrmann (2004). The *Twarog spec* compilation is based on the SPOCS catalogue – Valenti & Fischer (2005) – and 25 other surveys (including those listed in Table 4 except for Fuhrmann 2000, 2008, Mishenina et al. 2006 and Ramírez et al. 2007). The parameter determinations for all stars in *Twarog spec* are transformed to the SPOCS $T_{\rm eff}$, log g, [M/H] scales. Quoting Twarog et al. (2007), "the majority of the final abundances should have uncertainties below 0.03 dex, and all should be below 0.06 dex". This compilation is used for the uvbyβ metallicity calibration of the Twarog catalogue referred to in

[5]`http://adsabs.harvard.edu/cgi-bin/nph-abs_connect?library&libname=AP+reference+stars&libid=470e474905`

Sect. 4.2. The *Gaia ESP cool* compilation is a list of reference star candidates for calibration of the Gaia algorithm "Extended Stellar Parametrizer", more specifically the module dealing with cool, chromospherically active stars.

The individual catalogues listed below the compilations were selected according to the following criteria: (1) large number of stars studied, (2) spectroscopic analyses based on high-resolution spectra, (3) analyses of "good quality", i.e. based on well-known theoretical models with a critical review of methods and results included in the publication (4) spectral data available from the authors, for possible homogeneous reanalysis of the final AP reference star sample. The list of catalogues is preliminary and will be extended in the future, mainly towards lower gravities and lower metallicities. Detailed parameter-parameter diagrams and histograms for each catalogue will be presented in a forthcoming document.

4.1.2 Open cluster stars

Faint stars within the *primary grid* will be selected from open clusters (OCs). We start from clusters with known metallicity and select candidate AP reference stars covering the $T_{\rm eff}$ and $\log g$ grid parameters given in Sect. 3 from each cluster.

The clusters are selected following the procedure described below. The selected clusters are listed in Table 6 along with the cluster parameters metallicity, distance and age, as well as observability. They span a range in metallicity from -0.4 to $+0.4$ around solar. The clusters have well-known distances, reddening and ages.

The list of clusters is preliminary and will be extended and revised according to more recent metallicity determinations, e.g. Sestito et al. (2008); Bragaglia et al. (2008); Sestito et al. (2007, 2006); Randich et al. (2006). Additional criteria may also have to be taken into account, e.g. cluster radius (and therefore stellar density on the sky) to ensure optimal observing conditions.

Cluster selection

Starting point was the database for open clusters, WEBDA[6]. The section on cluster metallicities in WEBDA[7] contains lists of metallicity determinations or "FeH catalogues", by Twarog et al. (1997), Gratton (2000), Chen et al. (2003) and a list of "new values of the [Fe/H] iron abundances" with references to many works, including those contained in Friel (1995). These catalogues overlap in part. They were combined to a single catalogue by simply taking the values for clusters in common from the first catalogue in which these clusters appear, in the order given above. This results in a list of 136 clusters.

In addition, the web page of the "Open Clusters and Galactic Structure" project, cf. Dias et al.

[6] http://www.univie.ac.at/webda/
[7] http://www.univie.ac.at/webda/description.html

TABLE 6: Preliminary list of open clusters considered for selection of faint *primary grid* AP reference stars. The cluster names contain hyperlinks to the corresponding page in the on-line WEBDA database. Columns 2–4 give metallicity, distance and age. The "Observability" column gives observatories and periods where/when these clusters can be observed (odd . . . April to September, even . . . October to March).

Name	[M/H]	d [pc]	t [Myr]	Observability	Reference for [M/H]
NGC 2506	−0.4	3315	1648	ESO/NOT even	Twarog et al. (1997)
NGC 2660	−0.2	2859	1351	ESO even	Twarog et al. (1997)
Melotte 111	+0.0	86	522	ESO/NOT odd/even	Gratton (2000)
IC 4756	+0.0	415	674	ESO/NOT odd	Twarog et al. (1997)
IC 2395	+0.0	792	15	ESO odd/even	Clariá et al. (2003)
NGC 2682	+0.0	820	4093	ESO/NOT even	Twarog et al. (1997)
NGC 6819	+0.0	2188	2172	NOT odd	Twarog et al. (1997)
Berkeley 18	+0.0	4772	4271	NOT even	Gratton (2000)
NGC 6791	+0.2	4418	7850	NOT odd	Twarog et al. (1997)
Melotte 25	+0.2	42	708	ESO/NOT even	Twarog et al. (1997)
NGC 6067	+0.2	1676	95	ESO odd	Twarog et al. (1997)
NGC 6253	+0.4	1567	3949	ESO odd	Gratton (2000)

TABLE 7: Same as Table 6, in simple format (e.g. for presentations). d(nearby) $<$ 1 kpc $<$ d(far), t(young) $<$ 1 Gyr $<$ t(old).

Name	[M/H]	Distance	Age	North/South
NGC 2506	−0.4	far	old	N/S
NGC 2660	−0.2	far	old	S
Melotte 111	+0.0	nearby	young	N/S
IC 4756	+0.0	nearby	young	N/S
IC 2395	+0.0	nearby	young	S
NGC 2682	+0.0	nearby	old	N/S
NGC 6819	+0.0	far	old	N
Berkeley 18	+0.0	far	old	N
NGC 6791	+0.2	far	old	N
Melotte 25	+0.2	nearby	young	N/S
NGC 6067	+0.2	far	young	S
NGC 6253	+0.4	far	old	S

(2002), was consulted. Version 2.7 (27 Oct 2006) contains a list of metallicities for 143 clusters[8]. This list contains 19 clusters which are not in the WEBDA metallicity compilation.

[8] http://www.astro.iag.usp.br/~wilton/refsmetallicities.txt

Further, Paunzen & Netopil (2006) recently published a list of 72 open clusters with most accurately known parameters (age, reddening and distance). This list of "standard open clusters" and the list of clusters with metallicities have 40 clusters in common. None of these lie in one of the ecliptic pole fields.

Finally, from these 40 clusters, 12 clusters with roughly equispaced values of metallicity (step size 0.2 dex) and different distances and ages (to cover the desired magnitude range) were selected (nine of these are observable from the south, see Table 7).

We verified that the selected clusters are located in regions of the sky for which Gaia will obtain high-quality data, i.e. with a SNR of 100 or more. This SNR estimation is based on the calculations presented in Carrasco et al. (2006). The estimation will be repeated in the future with specific calculations using the Gaia Object Generator (GOG).

Selection of stars from open clusters

Candidate AP reference stars covering the $T_{\rm eff}$ and $\log g$ grid parameters given in Sect. 3 are selected from each of the selected clusters as follows. We base the selection on the UBV data compiled in WEBDA (because they are the most complete data available). The first step is to select stars according to V and $B-V$ criteria, taking into account reddening: $9 \leq V \leq 18$ and $0.1 \leq (B-V)_0 \leq 1.7$. For each star, $T_{\rm eff}$ and $\log g$ is then estimated from the photometry: $T_{\rm eff}$ from $B-V$ (Castelli & Kurucz models), and $\log g$ from V, cluster distance, the bolometric correction (estimated from Castelli & Kurucz models), and $T_{\rm eff}$, assuming 1 solar mass and taking into account cluster reddening.

A grid in $T_{\rm eff}$ and $\log g$ is defined, with smaller step sizes than adopted for the *primary grid*. The $T_{\rm eff}$ values are 4000,4200,4400,4600,4800,5000,5400,5800,6200,6600,7000 K. The $\log g$ values range from 1.0 to 5.0 with a step of 0.5. Together with the metallicity range defined by the cluster selection, placing one star at each node would result in 495 stars. Next, for each grid node a star with estimated parameters deviating by at most 100 K in $T_{\rm eff}$ and 0.1 in $\log g$ is selected, if possible. For those clusters for which membership probability data is available, only stars with a probability larger than 60% are included (all stars with unknown membership probability are included as well).

This results in 15-40 stars per cluster, in total about 330 stars, which is about two thirds of the number covering the whole parameter space defined above. The coverage of the ($T_{\rm eff}$, $\log g$, [M/H]) space by the selected stars is shown in Fig. 1. One can see that the high-metallicity part of the *primary grid* is covered fairly well by these stars, except for a) stars at the hot edge, b) metal-poor stars at the cool edge and c) low values of gravity (corresponding to bright giants and supergiants, which probably fall outside the magnitude limit for most clusters). It remains to be investigated if these gaps can be filled by scrutinizing the whole list of 40 clusters with metallicity determinations and accurately known parameters. Otherwise, we have to look for reference stars in less well-studied open clusters or in catalogues of field stars (e.g. Sect. 4.2).

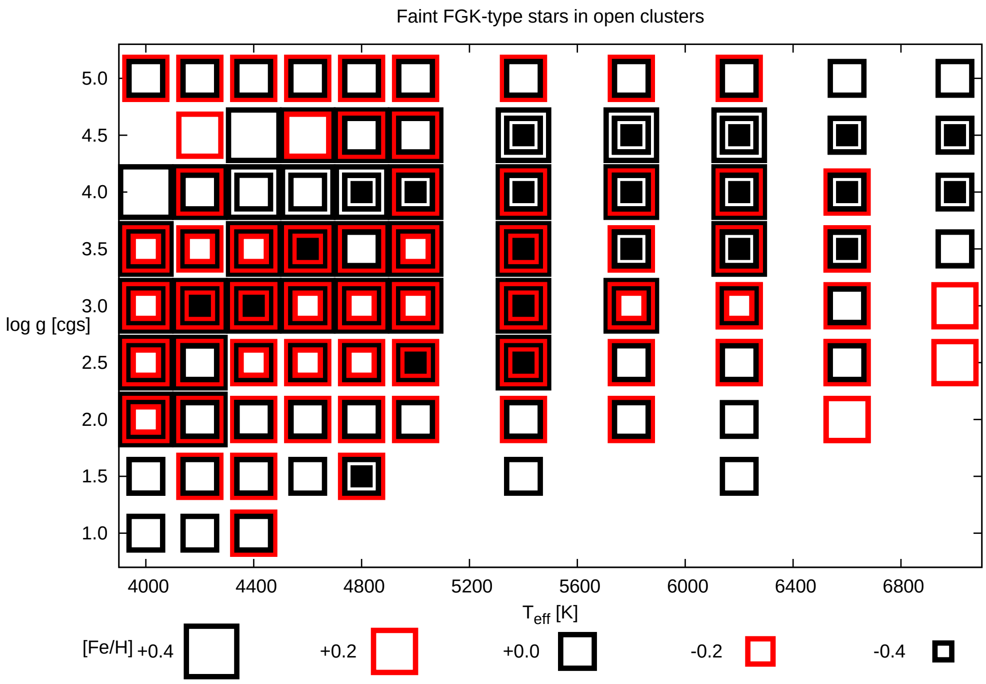

FIGURE 1: Coverage of astrophysical parameters by FGK type stars selected from the open clusters listed in Table 6 as outlined in Sect. 4.1.2.

Note that for each of the selected stars, accurate individual parameters ($T_{\rm eff}$, log g, [M/H]) have to be determined from the literature or from new analyses. A subset of these stars should be observed with high spectral resolution from one observatory (e.g. ESO), after having searched the archives for existing data. New, low-resolution spectra should be obtained with one and the same instrument (e.g. EFOSC2@NTT from ESO) for all stars, because we need a complete, homogeneous set of spectra. The issue of homogeneity is of higher concern for low resolution (flux calibrated) spectra than for high resolution spectra (Rosanna Sordo, priv. comm.).

4.1.3 Globular cluster stars

The selection of stars from globular clusters (GCs) will be done, if possible, in a similar way as for the open clusters. The details of the selection process will be described in a forthcoming version of this document or a separate document.

Globular clusters will be able to supplement the open cluster data as regards parameter-space coverage, in particular in terms of metallicity (extending down to [Fe/H] = -2.4), abundance ratios (α-enhanced populations) and age (halo GCs are systematically older than disk OCs by several Gyr). It is clear that good coverage of the existing parameter space is highly desirable. The price to pay is the apparent faintness of GCs. NGC 6397, one of the GCs with the lowest distance modulus $(m - M)_0 = 12.13 \pm 0.15$ at $E(B - V) = 0.18$, see Reid & Gizis (1998), has turnoff-point stars at apparent magnitudes close to 16.5 in V (M92 is significantly fainter). At the low-metallicity end, additional physical effects could become interesting even at the precision Gaia is expected to achieve: the convective envelopes of cool stars become thinner and less massive which results in apparent abundance anomalies on the 0.2 dex level, see Korn et al. (2007). Targeting unevolved and evolved stars in GCs would allow us to assess the relationship between initial and apparent photospheric abundances for a variety of evolutionary phases (main sequence vs. red giant branch vs. horizontal branch vs. AGB), also in terms of mixing with nuclear burning products dredged up from the stellar core.

4.2 Secondary grid

TABLE 8: Source catalogues for selection of *secondary grid* stars: type of data and references.

Catalogue name Method	Data	References
SDSS spectroscopic study	SDSS spectra (381–910 nm, R=2000) and $ugriz$ photometry	Allende Prieto et al. (2006); Allende Prieto (2007)
Stellar parameters derived from fit to spectra and SDSS $g - r$ colors using Kurucz (1993) models and Hubeny & Lanz (2000) SYNSPEC code.		
Twarog catalogue	uvbyβ photometry	Twarog et al. (2007)
uvbyβ metallicity calibration using the *Twarog spec* compilation (see Sect. 4.1).		
RAVE stellar parameters	RAVE spectra (841–880 nm, R=7500)	Zwitter et al. (2008)
Stellar parameters derived from fit to spectra using synthetic spectra from Munari et al. (2005) based on “the latest generation of Kurucz models”.		
Geneva-Copenhagen survey	uvbyβ photometry	Holmberg et al. (2007)
uvbyβ – $T_{\rm eff}$ calibration using $V - K$ colors and Di Benedetto (1998) calibration. uvbyβ – metallicity calibration using five works based on photometric $T_{\rm eff}$s.		
Soubiran clump giants	ELODIE spectra (390–680 nm, R=42 000, S/N≈20)	Soubiran et al. (2008)
Stellar parameters derived from fit to spectra using TGMET code (Katz et al. 1998), based on template spectra observed with ELODIE.		

The *secondary grid* has yet to be defined. A first list of source catalogues for selection of *secondary grid* stars is given in Tables 8 and 9. Further sources are to be considered, e.g. Feltzing et al. (2001); Robinson et al. (2007). Abstracts of the references can be accessed from UH's ADS private library “AP reference stars”[5].

TABLE 9: Source catalogues for selection of *secondary grid* stars: number of stars and parameter ranges.

Catalogue name	V range	N	T_{eff} range	log g range[a]	[M/H] range
SDSS spectroscopic study	[14:20]	94000[b]	[5000:8000]	[0.5:5.0]	[−4.8:+0.7]
Twarog catalogue	[3:14]	35000[c]	[5000:6500]	[V:V]	[−1.0:+0.5]
RAVE stellar parameters	[9:12][d]	21000	[3400:27000]	[0:5]	[−2.0:+1.0]
Geneva-Copenhagen survey	[2:12]	15000[e]	[4500:7000]	[V:V]	[−2.7:+0.8]
Soubiran clump giants	[6:10][f]	523	[4500:5700]	[1.0:4.5]	[−1.5:+0.3]

[a]or luminosity class range
[b]73000 stars with $V \leq 18$
[c]32000 stars with $6 \leq V \leq 13$, 500 stars with $V \geq 10$
[d]I magnitude range
[e]14500 stars with $V \geq 6$, 100 stars with $V \geq 10$
[f]Tycho2 V_{T} magnitude range

The parameter data for the SDSS study have been obtained from Carlos Allende Prieto (2008, priv. comm.). As an example, Fig. 2 shows the distribution of stars over V magnitude, T_{eff}, log g and [M/H] for this data set. The parameter data for the Geneva-Copenhagen survey are available from the CDS database[9]. The parameter data for the Twarog catalogue have been obtained from Bruce Twarog (2008, priv. comm.).

5 Conclusions

The calibration of the GSP-phot and GSP-spec algorithms is a challenging task. In this document, we have started to investigate the possibility of using grids of reference stars to test the algorithms as well as the synthetic spectra which are used as training data. Starting from the parameter regime of FGK stars, we have discussed selection criteria for reference stars, which include considerations of expected accuracy of the algorithms, magnitude ranges, and position on the sky. We have also outlined a possible approach for the calibration procedure. The final decision on the procedure has to be taken after consultation with algorithm providers and providers of synthetic spectra (discussions taking place in late-2008).

A major part of this document is devoted to surveying the possible source catalogues for selection of reference stars. In recent years, a number of studies using high-resolution spectra (resolutions on the order of 60 000) have resulted in precise measurements of atmospheric parameters of numerous nearby, bright stars ($V < 12$, on the order of 1000). The parameters covered are concentrated in the region of F to K dwarfs and the metallicities typically range from −2 dex to +0.5 dex. Fainter reference stars, suitable for calibration of the GSP-phot algo-

[9]`http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/475/519`

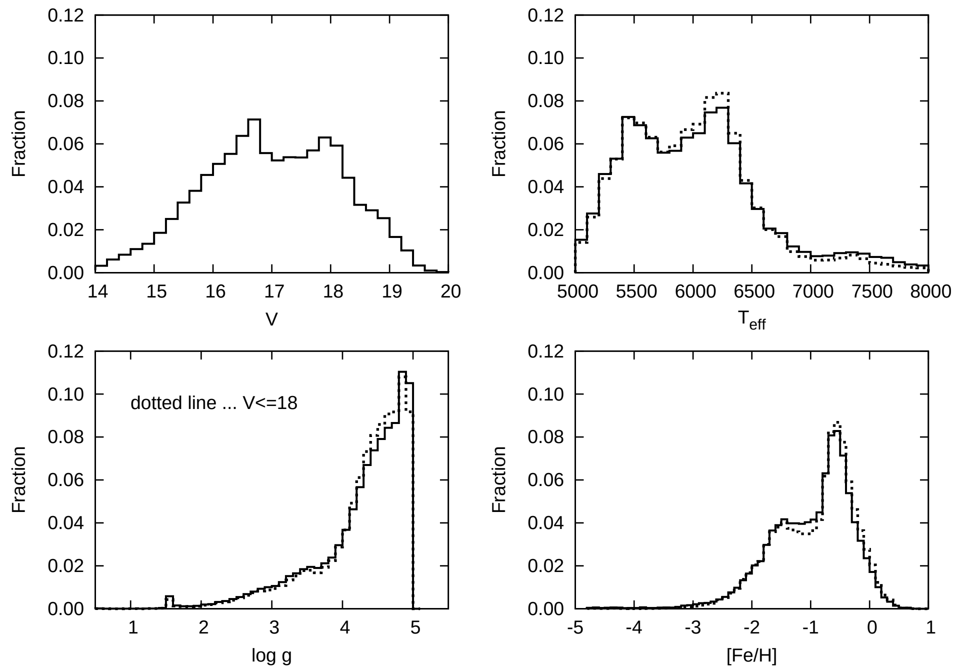

FIGURE 2: Distribution of stars over V magnitude, $T_{\rm eff}$, $\log g$ and [M/H] for SDSS data set (Allende Prieto et al. 2006; Allende Prieto 2007, Allende Prieto 2008, priv. comm.). Note that the distributions include all analysed spectra (about 90 000), and have not been corrected for duplicate observations of objects (a few thousand).

rithm, can be selected from open and globular clusters. Based on a survey of the data available in the WEBDA database, we have selected about 330 stars from 12 open clusters, which cover about two thirds of the parameter space of high-metallicity FGK stars. The metallicities of these stars are fairly well known from the cluster metallicities, but need to be confirmed, and further APs ($T_{\rm eff}$, $\log g$) remain to be verified. Globular cluster stars and stars from recently published large catalogues of calibrated spectrophotometric studies provide further promising sources for the selection of reference stars.

6 References

April 2007, DPAC: Proposal for the Gaia Data Processing, URL http://www.rssd.esa.int/llink/livelink/open/2720336

Allende Prieto C., Dec. 2007, In: Bulletin of the American Astronomical Society, vol. 38 of Bulletin of the American Astronomical Society, 1001–+

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

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

Bragaglia A., Sestito P., Villanova S., et al., Mar. 2008, A&A, 480, 79

Carrasco J., Jordi C., Figueras F., et al., September 2006, Toward the selection of standard stars for absolute flux calibration. Signal-to-noise ratios for BP/RP spectra and crowding due to FoV overlapping, URL `http://www.rssd.esa.int/llink/livelink/open/2703304`

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

Chen L., Hou J.L., Wang J.J., Mar. 2003, AJ, 125, 1397

Clariá J.J., Lapasset E., Piatti A.E., Ahumada A.V., Oct. 2003, A&A, 409, 541

di Benedetto G.P., Nov. 1998, A&A, 339, 858

Dias W.S., Alessi B.S., Moitinho A., Lépine J.R.D., Jul. 2002, A&A, 389, 871

ESA, July 2000, GAIA — Composition, Formation and Evolution of the Galaxy, Tech. rep., concept and Technology Study Report, ESA-SCI(2000)4

ESA Gaia Project, November 2006, Mission Requirements Document (MRD), URL `http://www.rssd.esa.int/llink/livelink/open/463164`

Feltzing S., Holmberg J., Hurley J.R., Oct. 2001, A&A, 377, 911

Friel E.D., 1995, ARA&A, 33, 381

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

Fuhrmann K., Jan. 2004, Astronomische Nachrichten, 325, 3

Fuhrmann K., Feb. 2008, MNRAS, 384, 173

Gratton R., 2000, In: Pallavicini R., Micela G., Sciortino S. (eds.) Stellar Clusters and Associations: Convection, Rotation, and Dynamos, vol. 198 of Astronomical Society of the Pacific Conference Series, 225–+

Holmberg J., Nordström B., Andersen J., Nov. 2007, A&A, 475, 519

Hubeny I., Lanz T., 2000, Synspec: A user's guide, Greenbelt:GSFC, URL `http://tlusty.gsfc.nasa.gov`

Katz D., Soubiran C., Cayrel R., Adda M., Cautain R., Oct. 1998, A&A, 338, 151

Korn A.J., Grundahl F., Richard O., et al., Dec. 2007, ApJ, 671, 402

Kurucz R.L., 1993, Kurucz cd-rom 13, atlas9 stellar atmosphere programs and 2 km /s grid, Cambridge: SAO

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

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

Mishenina T.V., Bienaymé O., Gorbaneva T.I., et al., Sep. 2006, A&A, 456, 1109

Munari U., Sordo R., Castelli F., Zwitter T., Nov. 2005, A&A, 442, 1127

Paunzen E., Netopil M., Oct. 2006, MNRAS, 371, 1641

Ramírez I., Allende Prieto C., Lambert D.L., Apr. 2007, A&A, 465, 271

Randich S., Sestito P., Primas F., Pallavicini R., Pasquini L., May 2006, A&A, 450, 557

Reid I.N., Gizis J.E., Dec. 1998, AJ, 116, 2929

Robinson S.E., Ammons S.M., Kretke K.A., et al., Apr. 2007, ApJS, 169, 430

Sestito P., Bragaglia A., Randich S., et al., Oct. 2006, A&A, 458, 121

Sestito P., Randich S., Bragaglia A., Apr. 2007, A&A, 465, 185

Sestito P., Bragaglia A., Randich S., et al., Sep. 2008, A&A, 488, 943

Soubiran C., Bienaymé O., Mishenina T.V., Kovtyukh V.V., Mar. 2008, A&A, 480, 91

Takeda G., Ford E.B., Sills A., et al., Feb. 2007, ApJS, 168, 297

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

Taylor B.J., Feb. 1999, A&AS, 134, 523

Taylor B.J., Dec. 2005, ApJS, 161, 444

Twarog B.A., Ashman K.M., Anthony-Twarog B.J., Dec. 1997, AJ, 114, 2556

Twarog B.A., Vargas L.C., Anthony-Twarog B.J., Nov. 2007, AJ, 134, 1777

Valenti J.A., Fischer D.A., Jul. 2005, ApJS, 159, 141

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