Gaia spectroscopy overview and comparative spectrum modeling for cool giants

U Heiter\(^1\) and T Lebzelter\(^2\)

\(^1\)Department of Physics and Astronomy, Uppsala University, Box 516, SE-75120 Uppsala, Sweden
\(^2\)Department of Astronomy, University of Vienna, Türkenschanzstrasse 17, A-1180 Vienna, Austria

E-mail: ulrike.heiter@physics.uu.se

Abstract. The Gaia mission will provide spectroscopic and spectrophotometric observations for a vast number of stars, allowing a characterization in terms of stellar parameters along with the astrometric measurements. The pipeline for astrophysical parameter determination relies on libraries of model spectra for mapping stellar parameters to flux distributions. To maximize the reliability of the parametrization, adequate efforts are required to assess the realism of the model spectra and to calibrate the analysis methods. We describe a spectrum modelling experiment, in which high-resolution optical and infrared spectra of four cool giant stars were analysed by 14 different groups. The resulting variance in derived stellar parameters illustrates the need to be cautious when comparing or combining stellar parameters from different model atmospheres and analysis strategies. The main causes for the differences in parameters derived by different groups seem to lie in the physical input data and in the details of the analysis method.

1. Introduction
The ESA cornerstone mission Gaia\(^1\) is due for launch in 2012. This ambitious astrometric mission will revolutionize our understanding of the formation and evolution of the Milky Way Galaxy. Gaia is a scanning satellite that will repeatedly survey the whole sky to obtain positions, parallaxes and proper motions to \(\mu\)as precision for all of the \(10^9\) objects brighter than 20 mag. Compared to its predecessor Hipparcos, Gaia will achieve a substantial improvement in terms of astrometric accuracy and completeness limit: parallax and proper-motion accuracy will be 100 times better and the number of stars is increased by a factor 10000.

The added value of Gaia follows from the non-astrometric instruments on-board. These are the dispersed photometric instruments – the blue photometer at 330–680 nm and the red photometer at 640–1050 nm (RP and BP) with a spectral resolution of about 4–30 nm per pixel, and the medium resolution spectrograph – the RVS or Radial Velocity Spectrometer at 847–874 nm. The RVS wavelength range includes the prominent Ca II IR triplet lines for cooler stars, and the H-Paschen lines for hotter stars, as well as a number of lines of Fe-peak and \(\alpha\) elements. The nominal spectral resolution of the RVS is \(R = \lambda/\Delta\lambda = 11500\), which will be reduced to about 5500 for magnitudes fainter than a certain limit around \(V = 12\). These instruments will enable the accurate simultaneous measurement of radial velocities as well as the determination of astrophysical parameters (APs) – \(T_\text{eff}\), log \(g\), metallicity [Fe/H], and interstellar extinction.

\(^1\) http://sci.esa.int/gaia/
2. Gaia photometric and spectroscopic instruments and data analysis

The building and assembly of the Gaia spacecraft and instruments is in an advanced stage. All instrument parts will be mounted on a torus with about 3 m diameter, made of SiC. Illustrations and images of the Gaia instruments can be found on the image gallery at the Gaia web page maintained by ESA. Useful examples are a schematic layout of the RVS instrument\(^2\), an image of the complete Gaia torus with the ‘Folding-Optics Structure’ ready to be mounted, which will support among others the RVS\(^3\), and the RVS module with its grating plate and several prismatic lenses\(^4\).

The responsibility for processing and analyzing Gaia data lies with the Gaia Data Processing and Analysis Consortium (DPAC). This consortium of over 300 European scientists and software engineers is developing the software system which will convert the raw data recorded by the Gaia satellite into scientifically meaningful data, for immediate use by the astronomical community. The preparations for data analysis rely on simulations of the data expected from the Gaia instruments. [1] show in their Fig. 3 a number of examples for simulated RP/BP data, for solar metallicity and various combinations of \(T_{\text{eff}}\) and \(\log g\). A large variation with \(T_{\text{eff}}\) is clearly visible in both bands. Differences in \(\log g\) will be more difficult to discern, except at the lowest \(T_{\text{eff}}\) values around 3000 K. Synthetic spectra calculated for the RVS wavelength range and parameters corresponding to late-type giants are shown in Fig. 1. The calculations were done with the spectrum synthesis code \texttt{sme.synth}, which is part of the spectrum analysis package SME \([2; 3]\), using MARCS model atmospheres \([4]\) and line data from VALD \([5; 6]\). The sensitivity of this spectral range to stellar metallicity is apparent.

The development of the data processing software for the determination of astrophysical parameters has progressed quite far. The software package called ‘Apsis’ (Astrophysical parameters inference system) has two main tasks: to classify all sources, i.e. to assign probabilities for being a star, galaxy, quasar, etc., and to determine astrophysical parameters for stars: \(T_{\text{eff}}\), \(\log g\), metallicity, extinction, and possibly \(\alpha\)-element abundances. The second task will be carried out by two different ‘General Stellar Parametrizer’ modules – GSP-spec and GSP-phot, specializing on RP/BP and RVS data, respectively. Both of them are based on large grids of model stellar spectra to map stellar parameters to flux distributions.

The GSP-phot module employs several numerical algorithms to determine the set of parameters which best represents the RP/BP flux distribution of an unknown stellar object. One of these (ILIUM) is described in detail in [7]. The performance of the algorithm is assessed by generating a test sample of synthetic flux distributions with random stellar parameters, which are analyzed by GSP-phot. The resulting APs are then compared to the known input APs. The performance depends on the stellar magnitude (i.e. the noise level) and the amount of extinction added to the simulated fluxes. In the best case, for 137 simulated stars at \(G=15\) without extinction, the \(T_{\text{eff}}\) residuals are within ±0.2\%, and the \(\log g\) residuals are within ±0.07. However, in general, when introducing variable extinction and including stars with \(G=15\) as well as \(G=18.5\), the \(T_{\text{eff}}\) residuals range from ±3 to 13\%, the \(\log g\) residuals from ±0.3 to 1.1, and the [Fe/H] residuals from ±0.5 to 1.3 dex.

The GSP-spec performance for analysis of RVS data depends strongly on the signal-to-noise ratio (SNR) of the spectra. For example, for FGK stars at a SNR of 50, the error in \(T_{\text{eff}}\) is found to be 100 K on average, i.e. between 2 and 3\%, \(\log g\) can be determined to 0.2, and [Fe/H] to 0.1 dex. \(\alpha\)-element abundance determination might be possible with errors between 0.05 and 0.3 dex for the full range of SNR. More details on the GSP-spec algorithms and their performance can be found in the contribution by A. Recio-Blanco.

The performance tests of Apsis described above are only assessing internal errors stemming

---

\(^2\) http://www.rssd.esa.int/index.php\?project=GAIA\&page=IG_20090923
\(^3\) http://www.rssd.esa.int/index.php\?project=GAIA\&page=IG_20101101b
\(^4\) http://www.rssd.esa.int/index.php\?project=GAIA\&page=IG_20110810a
Besides the prominent Ca II IR triplet lines, a number of weak Fe I, Ti I, and Cr I lines are seen. In the spectrum with higher $T_{\text{eff}}$, there are also some Mg I and Si I lines.

A systematic comparison of synthetic and observed stellar spectra is more difficult, because we do not know the ‘true’ astrophysical parameters of any observed star (except perhaps for the Sun), and because of the difficulties involved in calibrating observed fluxes. A few attempts can be found in the literature, e.g. [9], who compared MARCS model fluxes with three different observed stellar spectral libraries of low to medium resolution. The comparisons for individual stars and models were not conclusive because of the aforementioned problems, but an average over the differences between models and observations for all stars within each library proved to be useful for understanding the performance of the algorithm.
Figure 2. MARCS model atmosphere flux variation with $T_{\text{eff}}$. Upper panel: optical wavelength range, lower panel: RVS range.

be useful. It revealed a few distinctive features in the blue and ultraviolet, pointing to missing or erroneous opacities in the models.

For the external calibration of the Gaia AP determination, we are aiming to assess the realism of the synthetic libraries by comparing model spectra and fluxes with those of a dedicated sample of benchmark stars. The APs of these stars should not be based on atmospheric models as far as possible. We are in the process of establishing a sample of benchmark stars with model-independent AP determinations: $T_{\text{eff}}$ from an angular diameter measurement and the bolometric flux, and $\log g$ from angular diameter, parallax, and mass. The third atmospheric parameter,
Table 1. Characteristics of the benchmark stars and test spectra. $T_{\text{eff}}$ and $\log g$ for $\alpha$ Tau and $\alpha$ Cet are direct determinations from angular diameter, bolometric flux and parallax measurements, and mass determinations, reference [Fe/H] values are taken from the literature (details will be given in Lebzelter et al., in preparation). $R$ is the resolving power, $\Delta \lambda$ the wavelength range, and SNR the signal-to-noise ratio of the spectra.

| Object   | $T_{\text{eff}}$ [K] | $\log(g)$ [cm/s$^2$] | [Fe/H] | $R$ | $\Delta \lambda$ [nm] | SNR |
|----------|----------------------|-----------------------|--------|-----|------------------------|-----|
| $\alpha$ Tau | 3930±40              | 1.2±0.1               | −0.22±0.11 | 80000 | 490–975 | > 200 |
| $\alpha$ Cet | 3800±60              | 0.9±0.1               | +0.02±0.03 | 80000 | 490–975 | > 200 |
| ‘Star 3’ | 4257                 | 1.47                  | −0.4    | 50000 | 1546–1567 | 125 |
| ‘Star 4’ | 3280                 | 0.06                  | +0.1    | 50000 | 1546–1567 | 125 |

Metallicity or chemical composition, must be determined from spectroscopy, and thus will always be model-dependent. However, through carefully designed tests we can assess the influence of various modeling assumptions and approximations, as well as the details of the analysis method on the metallicity determination. In the second part of this contribution, we report on a test campaign for two specific benchmark stars.

3. Comparative spectrum modeling for cool giants
The experiment was conducted around a workshop supported by GREAT-ESF in Aug 2010. Observed optical spectra of two benchmark stars and two simulated H-band spectra were analyzed before the workshop by 14 groups using different models and analysis approaches. The analysis details and results were presented and discussed during the workshop. Discussions and follow-up tests continued after the workshop. The participating groups were provided with stellar spectra with $R=80\,000$ and SNR > 200 from 490 to 975 nm for two stars of undisclosed identities. The stars were in fact two well-known giant stars with $T_{\text{eff}} \approx 4000$ K ($\alpha$ Tau and $\alpha$ Cet). Spectra for ‘Star 3’ and ‘Star 4’ from 1546 to 1567 nm were provided as well. These had been calculated with $R=50\,000$ and artificial noise was added to simulate a SNR of 125. In addition, approximate broad-band colors were given for all four objects. The characteristics of the test stars and spectra are given in Table 1.

The participating groups were asked to calculate model spectra for these objects, matching the observations, and to submit the corresponding best-fit atmospheric parameters to the experiment organizers for comparison. Some information on the groups and their methods is summarized in Table 2. The groups were from institutes all over Europe, as well as from the US, Canada, India, Australia, and Japan. About half of the groups based their analysis on MARCS atmospheric models and related spectrum synthesis codes, a considerable number used the ATLAS suite of codes [10; 11], and the three software packages PHOENIX [12], CODEX [13] and Tsuji08 [14] were employed by one group each. Most of the groups carried out spectrum synthesis at the resolution of the observed spectra or for degraded spectra, while three groups measured equivalent widths from the provided spectra which they compared to calculated equivalent widths.

As an example for the results, we show in Fig. 3 the best-fit spectrum for $\alpha$ Tau in the RVS region from one of the groups using MARCS models. The parameters of the model are $(T_{\text{eff}}, \log g, [\text{Fe/H}])=(4050\,\text{K}, 1.5, 0.0)$, for spherical geometry with one solar mass and a microturbulence of 2 km s$^{-1}$. These were determined in this case by $\chi^2$-minimization for the RVS region, varying only $T_{\text{eff}}$ and $\log g$. Changing the metallicity to $-0.25$ dex, an equally good fit is obtained for the same $T_{\text{eff}}$, but $\log g=1.0$. An inspection of the $\chi^2$-surface shows that it does not vary significantly with $T_{\text{eff}}$, and values between 3800 and 4300 K fit the spectrum almost as well. This wavelength region seems to be most sensitive to $\log g$ variations, and somewhat to variations in [Fe/H]. The
Table 2. Groups participating in the spectrum modeling experiment. Model suites employed by the groups are the following: M – MARCS, A – ATLAS, P – PHOENIX, C – CODEX, T – Tsuji08. EW in the Model column indicates that equivalent widths rather than synthetic spectra were used. The numbers in the column headed ‘Stars’ indicate which of the four test stars were analyzed. † . . . symbols used in Figs. 4 and 6.

| Contact  | Institute      | Model | Stars | Contact  | Institute      | Model | Stars |
|----------|----------------|-------|-------|----------|----------------|-------|-------|
| Nowotny  | Vienna         | M1    | 12    | Maldonado| Madrid         | A2,EW| 12    |
| Plez     | Montpellier    | M2    | 12    | Neilsen  | Bonn           | A3   | 34    |
| Worley   | Nice           | M3    | 12    | Peterson | UCO/Lick       | A4   | 1     |
| Eriksson | Uppsala        | M4    | 1     | Goswami  | India          | A5,EW| 1     |
| Abia     | Granada        | M5    | 34    | Short    | Halifax        | P    | 12    |
| Merle    | Nice           | M6    | 1     | Ireland  | Sydney         | C    | 2 4   |
| Wahlgren | GSFC           | A1    | 1 3   | Tsuji    | Tokyo          | T,EW | 4     |

same group analysed two further wavelength regions, 515–520 nm and 610–640 nm, and found the latter to be most sensitive to $T_{\text{eff}}$ variations, while the former was not very sensitive to either parameters. Note that in Fig. 3, even for the best-fit spectrum, differences between observations and calculations reach up to 10%, which is similar to the differences between different synthetic spectral libraries found by [8]. In the blue-most region investigated, the deviations reach up to 30% for the best-fit model.

The group using PHOENIX models used a similar $\chi^2$-minimization approach, but for two larger wavelength regions (490 to 680 nm and 800 to 900 nm). In this case, the bluer region proved to be suited for determining a best-fit $T_{\text{eff}}$ value, and to a lesser extent for constraining $\log g$, while the IR region was not sensitive to changes in any parameter. The best-fit model for $\alpha$ Tau had ($T_{\text{eff}}$, $\log g$) = (3910 K, 1.75) for solar metallicity, and the same $\log g$, but $T_{\text{eff}}$ = 3800 K for $[\text{Fe/H}]$ = −0.5. Here, the relative flux differences in the IR between observed and calculated spectra are up to 20% for any parameter combination around the best fit in the blue region, where the deviations reach up to 50% for certain spectral lines.

Details on the analysis of all participating groups as well as an extensive discussion will be presented in a forthcoming publication (Lebzelter et al.). Here, we present a summary of the results in Fig. 4 for the two benchmark stars $\alpha$ Tau and $\alpha$ Cet. We compare the best-fit parameters determined by all groups and the $T_{\text{eff}}$ and $\log g$ values for these stars determined in an independent way (angular diameter, bolometric flux, parallax, mass). The reference $[\text{Fe/H}]$ values are previous metallicity determinations based on high-resolution spectroscopy found in the literature (mean of eight references for $\alpha$ Tau and one reference for $\alpha$ Cet).

The overall conclusion is that the spread in the results is larger than the estimated error bars for each point, but most of them agree within 2$\sigma$. $T_{\text{eff}}$ seems to be better constrained than $\log g$. For $\alpha$ Tau, determinations for all parameters cluster around the reference values, although the $\log g$ results seem to be divided into two groups larger and smaller than the reference value. For $\alpha$ Cet, the optical spectrum fits point towards a lower metallicity than determined before by [15], based on high-resolution spectra in the H-band. We do not see any clear systematic trends with regard to the type of model used, except that groups using ATLAS or PHOENIX models tend to obtain higher $\log g$ values than groups using MARCS models.

In an attempt to pin down the factors determining the differences in the results, we did a direct comparison of synthetic optical spectra calculated for a fixed set of parameters by eight groups, using six different codes or code versions and sets of input line data. Considerable differences in the depths of numerous spectral lines between the different synthetic spectra became apparent over the whole wavelength range. They are largest in the blue region and decrease towards the
Figure 3. One of the best fit spectra for α Tau using MARCS model atmospheres for the Gaia RVS region. For the purpose of presentation, the spectra are sampled at 0.04 nm. Solar abundances, $T_{\text{eff}}=4050$ K, $\log g=1.5$. Bottom: Scaled flux, observed and calculated; top: observed minus calculated flux.

near-IR. For example, in the interval 511 to 512 nm we see typical differences in normalized line depth of 0.4. In the interval 656 to 657 nm differences are mostly below 0.1, with a few outliers of up to 0.4. The main cause for these differences presumably lies in the input line data, both oscillator strengths and broadening parameters, as well as in the methods and data used for calculating the continuum flux.

Besides the model differences, the results of the experiment are clearly affected by different methods and strategies employed for parameter determination. Half of the groups used the provided color information for an initial guess of the parameters or as an additional constraint. Each group selected subsections of the available wavelength range to work with. For the optical spectra, Fig. 5 shows the wavelength ranges used by the 11 groups who analysed these spectra. Two of the groups using the same model and similar wavelength ranges arrived at similar $T_{\text{eff}}$ values, but different $\log g$ and [Fe/H]. The differences may be caused by using wavelength ranges, which are not sensitive to changes in these parameters. One of the groups ran their analysis method twice, for different wavelength regions and degrading the spectra to different resolutions. Using the RVS region at low resolution resulted in higher $T_{\text{eff}}$ values than using the region 490 – 670 nm at higher resolution.

The infrared spectra of ‘Star 3’ and ‘Star 4’ were only analysed by three and four groups,
Figure 4. Best-fit parameters determined by 14 different groups for $\alpha$ Tau (top) and $\alpha$ Cet (bottom). Black symbols represent different groups as defined in Table 2. Not all groups provided error estimates for all parameters. M3 provided two results for different wavelength and resolution set-ups. M4 and P provided two results each for $\alpha$ Tau, for two different adopted metallicity values. The large cross-symbol in red represents atmospheric-model-independent $T_{\text{eff}}$ and $\log g$ determinations and literature [Fe/H] values. respectively. The resulting best-fit values for $T_{\text{eff}}$ and $\log g$ are shown in Fig. 6, compared to the input parameters used to generate the artificial observations. For Star 3, the derived $T_{\text{eff}}$ values are in reasonable agreement with the reference value. The resulting $\log g$ and [Fe/H] values from all three groups are higher than the reference values. It is unclear if this is caused by a systematic difference between the employed models and the model used to generate the test spectra, or if it is due to the analysis methods used. For Star 4, all parameters are close to the input values.

4. Conclusions
The Gaia mission promises great advances for stellar astrophysics, and the hardware and software are nearing completion. To ensure the best possible use of these resources, we have to put adequate efforts into calibrating our data analysis methods.

The spectrum modeling experiment described in Section 3 illustrates the need to be cautious when comparing or combining stellar parameters from different model atmospheres and analysis strategies. As no systematic differences between ‘model families’ (MARCS, ATLAS, etc.) are apparent, we conclude that the main causes for the differences in parameters derived by different
Figure 5. Wavelength ranges used by the different groups analyzing the optical spectra.

Figure 6. Best-fit parameters for ‘Star 3’ and ‘Star 4’. Black symbols represent different groups as defined in Table 2. Not all groups provided error estimates. Asterisk-symbols in red represent the input parameters used to generate the artificial observations.
groups lie in the physical input data (e.g. oscillator strengths) and in the details of the analysis method (e.g. selected wavelength regions). However, due to complexity of problem, we could not derive clear trends for the effect of any given assumption. We note that the experiment represents a typical situation in abundance analysis and provides an overall assessment of the current status of this field, at least for the case of cool giant stars.

Acknowledgments
UH acknowledges support from the Swedish National Space Board, and from the ESF. TL acknowledges support from the Austrian Science Fund (FWF). We thank the groups participating in the spectrum modeling experiment for their work.

References
[1] Jordi C, Gebran M, Carrasco J M, de Bruijne J, Voss H, Fabricius C, Knude J, Vallenari A, Kohley R and Mora A 2010 A&A 523 A48 (Preprint 1008.0815)
[2] Valenti J A and Fischer D A 2005 ApJS 159 141–166
[3] Valenti J A and Piskunov N 1996 A&A 118 595–603
[4] Gustafsson B, Edvardsson B, Eriksson K, Jørgensen U G, Nordlund A and Plez B 2008 A&A 486 951–970 (Preprint 0805.0554)
[5] Heiter U, Barklem P, Fossati L, Kildiyarova R, Kochukhov O, Kupka F, Obbrugger M, Piskunov N, Plez B, Ryabchikova T, Stempels H C, Stütz C and Weiss W W 2008 Journal of Physics Conference Series 130 012011
[6] Kupka F, Piskunov N, Ryabchikova T A, Stempels H C and Weiss W W 1999 A&A 138 119–133
[7] Bailer-Jones C A L 2010 MNRAS 403 96–116 (Preprint 0911.5242)
[8] Vallenari A and Sordo R 2009 Impact on GOG simulations of different input stellar libraries gaia DPAC Technical Note AV-009 URL http://www.rssd.esa.int/l1link/livelink/open/2867451
[9] Edvardsson B 2008 Physica Scripta Volume T 133 014011
[10] Kurucz R 1993 ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993. 13
[11] Kurucz R 1993 SYNTHE Spectrum Synthesis Programs and Line Data. Kurucz CD-ROM No. 18. Cambridge, Mass.: Smithsonian Astrophysical Observatory
[12] Hauschildt P and Baron E 1999 J. Comp. App. Math. 109 41
[13] Ireland M J, Scholz M and Wood P R 2008 MNRAS 391 1994–2002
[14] Tsuji T 2008 A&A 489 1271–1289 (Preprint 0807.4387)
[15] Meléndez J, Asplund M, Alves-Brito A, Cunha K, Barbuy B, Bessell M S, Chiappini C, Freeman K C, Ramírez I, Smith V V and Yong D 2008 A&A 484 L21–L25 (Preprint 0804.4124)