A&A 512, A54 (2010)
DOI: 10.1051/0004-6361/200913204

Astronomy & Astrophysics

An absolutely calibrated $T_{\rm eff}$ scale from the infrared flux method

Dwarfs and subgiants⋆

L. Casagrande[1], I. Ramírez[1], J. Meléndez[2], M. Bessell[3], and M. Asplund[1]

[1] Max Planck Institute for Astrophysics, Postfach 1317, 85741 Garching, Germany
e-mail: luca@mpa-garching.mpg.de
[2] Centro de Astrofísica da Universidade do Porto, Rua das Estrelas 4150-762 Porto, Portugal
[3] Research School of Astronomy and Astrophysics, Mount Stromlo Observatory, Cotter Rd, ACT 2611, Australia

Received 28 August 2009 / Accepted 6 January 2010

ABSTRACT

Various effective temperature scales have been proposed over the years. Despite much work and the high internal precision usually achieved, systematic differences of order 100 K (or more) among various scales are still present. We present an investigation based on the infrared flux method aimed at assessing the source of such discrepancies and pin down their origin. We break the impasse among different scales by using a large set of solar twins, stars which are spectroscopically and photometrically identical to the Sun, to set the absolute zero point of the effective temperature scale to within few degrees. Our newly calibrated, accurate and precise temperature scale applies to dwarfs and subgiants, from super-solar metallicities to the most metal-poor stars currently known. At solar metallicities our results validate spectroscopic effective temperature scales, whereas for [Fe/H] $\lesssim -2.5$ our temperatures are roughly 100 K hotter than those determined from model fits to the Balmer lines and 200 K hotter than those obtained from the excitation equilibrium of Fe lines. Empirical bolometric corrections and useful relations linking photometric indices to effective temperatures and angular diameters have been derived. Our results take full advantage of the high accuracy reached in absolute calibration in recent years and are further validated by interferometric angular diameters and space based spectrophotometry over a wide range of effective temperatures and metallicities.

Key words. stars: fundamental parameters – stars: abundances – stars: atmospheres – infrared: stars – techniques: photometric

1. Introduction

The determination of effective temperatures ($T_{\rm eff}$) in F, G and K type stars has a long and notable history. Because of their long lifetimes these stars retain in their atmospheres a fossil record of the chemical elements in the interstellar medium at the time of their formation. The stellar effective temperature is of paramount importance for reliable abundance analyses and thus for improving our understanding of Galactic chemical evolution.

Stellar abundances are now routinely derived from high resolution spectra, model atmospheres, and spectrum synthesis. While each of these ingredients have their own issues regarding systematic uncertainties, the dominant source of error is in many cases the adopted $T_{\rm eff}$ of the star. Several indirect methods of $T_{\rm eff}$ determination have been devised to avoid the complications introduced by the measurement of stellar angular diameters, which are necessary to derive $T_{\rm eff}$ from basic principles (e.g. Hanbury Brown et al. 1974; van Belle & von Braun 2009). Thus, most published values of $T_{\rm eff}$ are model-dependent or based on empirical calibrations that are not free from systematics themselves.

It is therefore not surprising to find discrepancies among published $T_{\rm eff}$ values. The ionization and excitation balance of iron lines in a 1D LTE analysis is routinely used to derive effective temperatures as well as $\log g$ and [Fe/H]. While for a sample of stars with similar properties this method can yield highly precise relative physical parameters (Meléndez et al. 2009a; Ramírez et al. 2009, see Sect. 3 for its use on solar twins), non-LTE effects and departures from homogeneity can seriously undermine effective temperature determinations, especially in metal-poor stars (e.g. Asplund 2005). Similarly, the line-depth ratio technique has high internal precision, claiming to resolve temperature differences of order 10 K (e.g. Gray & Johanson 1991; Gray 1994; Kovtyukh et al. 2003) but it is not entirely model independent (e.g. Caccin et al. 2002; Biazzo et al. 2007) and the uncertainty on its zero point can be considerably large. Another popular method for deriving $T_{\rm eff}$ in late-type stars is provided by the study of the hydrogen Balmer lines, in particular Hα and Hβ (e.g. Nissen et al. 2007; Fuhrmann 2008). For H lines uncertainties related to observations and line broadening (Barklem et al. 2002), non-LTE (Barklem 2007) and granulation effects (Asplund 2005; Ludwig et al. 2009; Pereira et al. 2010) all influence the estimation of effective temperatures.

In such a scenario, an almost model independent and elegant technique for determining effective temperatures was introduced in the late 70's by D. E. Blackwell and collaborators (Blackwell & Shallis 1977; Blackwell et al. 1979, 1980) under the name of InfraRed Flux Method (hereafter IRFM). Since then, a number of authors have applied the IRFM to determine effective temperatures in stars with different spectral types and metallicities (e.g. Bell & Gustafsson 1989; Alonso et al. 1996a; Ramírez & Meléndez 2005a; Casagrande et al. 2006; González Hernández & Bonifacio 2009). The main ingredient of the IRFM is infrared photometry, with the homogeneous and all-sky coverage provided by 2MASS being the *de facto* choice nowadays.

⋆ Table 8 is only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/512/A54

As such, the IRFM can now be readily applied to many stars, making it ideal to determine colour-temperature-metallicity relations spanning a wide range of parameters. The effective temperatures determined via IRFM are often regarded as a standard benchmark for other techniques. Whilst they have high internal accuracy and are essentially free from non-LTE and granulation effects (Asplund & García Pérez 2001; Casagrande 2009; Ramirez et al. in prep.), the reddening and absolute flux calibration adopted in such a technique can easily introduce a systematic error as large as 100 K (Casagrande et al. 2006).

The effective temperatures of dwarfs and subgiants are still heavily debated with various $T_{\rm eff}$ scales behaving very differently depending on colours and metallicities. One of the most critical discrepancies occur at the metal-poor end, for [Fe/H] $\lesssim -2.5$. In their work on the determination of effective temperatures via IRFM, Ramírez & Meléndez (2005a) found temperatures significantly hotter than those previously published, in particular those determined using the excitation equilibrium method. Differences up to 500 K for the hottest ($T_{\rm eff} \simeq 6500$ K) most metal-poor ([Fe/H] $\lesssim -3.0$) stars were reported (e.g., Meléndez & Ramírez 2004; Meléndez et al. 2006b). In this regime, the recent IRFM investigation by González Hernández & Bonifacio (2009) still supports a temperature scale significantly hotter than excitation equilibrium and Balmer lines, but ∼90 K cooler than Ramírez & Meléndez (2005a).

The abundance pattern measured in metal-poor stars is important for our quest to understand Galactic chemical evolution and Big Bang nucleosynthesis: two notable examples are the oxygen abundance and the lithium trend with metallicity, both of which crucially depend on the adopted $T_{\rm eff}$ scale. For example, a change of +100 K in $T_{\rm eff}$ would decrease the [O/Fe] ratio in turn-off metal-poor stars by ∼0.08 dex when using the OI triplet and FeII lines (Meléndez et al. 2006a), while the same change in $T_{\rm eff}$ would increase the Li abundance by ∼0.07 dex (e.g. Meléndez & Ramírez 2004; Meléndez et al. 2009b, 2010).

At higher metallicities, which encompass most of the stars in the solar neighbourhood, the situation is also uncertain, with spectroscopic effective temperatures in rough agreement with the IRFM scale of Casagrande et al. (2006). The latter is then about 100 K hotter than the IRFM temperatures of Ramírez & Meléndez (2005b) whilst the recent implementation of González Hernández & Bonifacio (2009) falls in between these two extremes. These differences are somewhat puzzling considering that all recent works on the IRFM have used 2MASS photometry. Effective temperature calibrations are also crucial in the context of deriving reliable colours for theoretical stellar models, which apart from few notable exceptions (e.g. VandenBerg & Clem 2003) have to resort entirely to theoretical flux libraries.

The aim of this work is to uncover the reason(s) behind such a confusing scenario and provide a solution to different IRFM effective temperature scales currently available in literature. As we discuss throughout the paper, this ambitious task is accomplished by using solar twins which allow us to set the absolute zero point of the $T_{\rm eff}$ scale. This result is further validated using interferometric angular diameters and space-based spectrophotometry.

The paper is organized as follows. In Sect. 2 we compare the results obtained from different authors, focusing in particular on two independent implementations of the IRFM (Ramírez & Meléndez 2005a; Casagrande et al. 2006) when the same input data are used. This approach allows us to precisely identify where different $T_{\rm eff}$ scales originate from. A cure to such an impasse is then provided in Sect. 3. The validation of our results, together with the new both precise and accurate effective temperature scale are presented in Sects. 4 to 6. We finally conclude in Sect. 7.

2. Comparing different versions

In this paper we use an updated version the IRFM implementation described in Casagrande et al. (2006) to nail down the reasons behind different $T_{\rm eff}$ scales. Our implementation works in the 2MASS system and fully exploits its high internal consistency thus making it well suited to the purpose of the present investigation. The core of the present study is to carry out a detailed comparison with the Ramírez & Meléndez (2005a) implementation when the same input data are used. For the sake of precision, notice that hereafter, when we refer to a $T_{\rm eff}$ determined by Ramírez & Meléndez (2005a) we are referring to the effective temperatures determined using that implementation and not the original values given in that paper. This is because of the updated (and more consistent) input data used here and also because some of the stars presented in this work do not have IRFM $T_{\rm eff}$ values published yet. In fact, in order to reveal trends with metallicity and/or effective temperature, our sample is specifically built to cover as wide a range as possible in those parameters (Fig. 1).

2.1. Input sample

The main ingredient of the IRFM is optical and infrared photometry. The technique depends very mildly on other stellar parameters, such as metallicity and surface gravity, which are needed to interpolate on a grid of model atmospheres (see Sect. 2.2). Below we present the papers from which we gathered [Fe/H] and $\log g$ for all our stars and we also give references to the photometric sources.

The metal-rich dwarfs come from Casagrande et al. (2006) who also provide homogeneous and accurate $BV(RI)_{\rm C}$ photometry while additional metal-rich dwarfs and subgiants are from Ramírez & Meléndez (2005a). We complement the sample with a number of moderately metal-poor stars from the study of Fabbian et al. (2009) and metal-poor turn off stars from Hosford et al. (2009). To investigate the metal-poor end of the $T_{\rm eff}$ scale in more detail, stars with reliable input data from Ramírez & Meléndez (2005a), Bonifacio et al. (2007) and Aoki et al. (2009) were added. Finally, to explore for the first time the hyper-metal-poor regime via IRFM the subgiants HE0233-0343 ([Fe/H] $\lesssim -4$ García Pérez et al. 2008) and HE1327-2326 ([Fe/H] ≤ -5 Frebel et al. 2005; Aoki et al. 2006; Frebel et al. 2008; Korn et al. 2009) were included.

New $UBV(RI)_{\rm C}$ photometric observations for some of the metal-poor stars in the aforementioned papers were conducted by Shobbrook & Bessell (1999; private communication) and are given in Table 1. For the remaining stars, optical Johnson-Cousins photometry was taken either from Beers et al. (2007) or the General Catalogue of Photometric Data (Mermilliod et al. 1997).

Infrared $JHK_{\rm S}$ photometry for the entire sample is available from the 2MASS catalogue (Skrutskie et al. 2006) which also includes the uncertainty for each observed magnitude ("j_", "h_" and "k_msigcom"). The infrared median total photometric error of our sample is 0.07 mag (i.e. "j_"+"h_"+"k_msigcom"= 0.07) and never exceeds 0.14 mag. Such an accuracy in the infrared photometry implies a mean (maximum) internal error in $T_{\rm eff}$ of 25 K (50 K). Notice that the effective internal accuracy

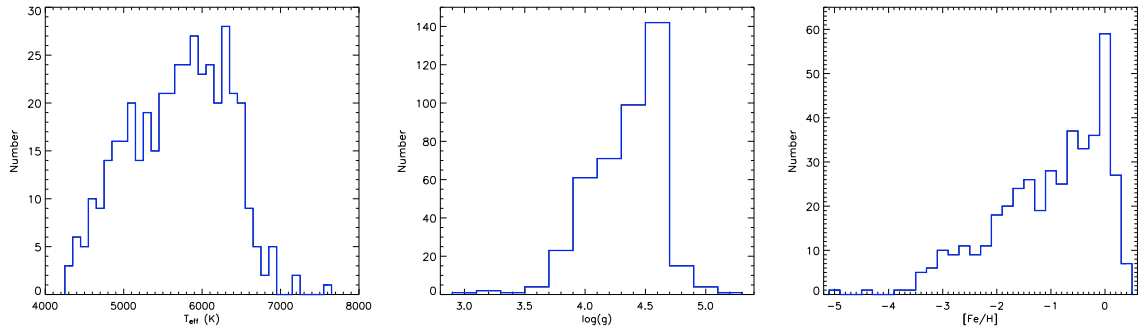

Fig. 1. Distribution of $T_{\rm eff}$, $\log g$ and [Fe/H] for the 423 stars in our sample.

Table 1. New Johnson-Cousins photoelectric photometry obtained for some of the metal-poor stars in the sample.

Name	U	B	V	$R_{\rm C}$	$I_{\rm C}$
HD 3567	9.556	9.695	9.240	8.941	8.631
HD 16031	10.004	10.197	9.770	9.484	9.184
HD 19445	8.207	8.503	8.026	7.737	7.394
HD 34328	9.683	9.903	9.416	9.106	8.773
HD 45282	8.659	8.672	8.010	7.610	7.196
HD 59392	10.048	10.217	9.761	9.457	9.142
HD 64090	8.762	8.951	8.295	7.935	7.536
HD 64606	8.277	8.140	7.412	6.994	6.561
HD 74000	9.880	10.071	9.656	9.381	9.080
HD 84937	8.485	8.702	8.306	8.047	7.759
HD 94028	8.421	8.640	8.202	7.917	7.585
HD 102200	9.009	9.189	8.739	8.449	8.141
HD 106038	10.431	10.627	10.153	9.857	9.529
HD 108177	9.874	10.082	9.647	9.362	9.052
HD 110621	10.230	10.385	9.932	9.628	9.313
HD 114762	7.738	7.833	7.283	6.967	6.629
HD 116064	9.099	9.282	8.833	8.520	8.189
HD 122196	9.055	9.212	8.753	8.444	8.112
HD 132475	8.983	9.100	8.563	8.216	7.855
HD 134169	8.115	8.193	7.663	7.342	7.011
HD 134439	10.033	9.881	9.118	8.661	8.220
HD 140283	7.502	7.692	7.205	6.876	6.522
HD 160617	9.014	9.188	8.740	8.431	8.108
HD 163810	10.185	10.272	9.660	9.280	8.897
HD 179626	9.601	9.710	9.188	8.849	8.502
HD 181743	9.911	10.140	9.683	9.375	9.062
HD 188510	9.303	9.452	8.851	8.486	8.100
HD 189558	8.214	8.299	7.740	7.392	7.034
HD 193901	9.049	9.183	8.644	8.307	7.964
HD 194598	8.666	8.844	8.356	8.055	7.739
HD 199289	8.660	8.803	8.287	7.972	7.643
HD 201891	7.740	7.908	7.390	7.081	6.737
HD 213657	9.869	10.063	9.646	9.368	9.068
HD 215801	10.272	10.471	10.038	9.732	9.418
HD 219617	8.425	8.621	8.153	7.845	7.525
HD 284248	9.407	9.650	9.208	8.927	8.608
HD 298986	10.316	10.506	10.062	9.774	9.470
BD+17 4708	9.718	9.922	9.476	9.183	8.854
BD+02 3375	10.174	10.414	9.944	9.635	9.297
BD-04 3208	10.203	10.375	9.977	9.709	9.417
BD-13 3442	10.529	10.655	10.266	9.994	9.704
CD-30 18140	10.155	10.365	9.946	9.663	9.353
CD-33 3337	9.436	9.581	9.109	8.814	8.490

Notes. Each measurement comprises an average of 4 observations per star. The rms of individual observations are 0.02 for the V magnitude, 0.015 for the $U-B$ colour and 0.008 mag for $B-V$, $V-R$, $R-I$, $V-I$ colours.

is slightly worse because of additional uncertainties stemming from the optical photometry, [Fe/H] and $\log g$. Altogether our final sample consists of 423 stars: all have $BVJHK_{\rm S}$ photometry while more than half have also $(RI)_{\rm C}$ magnitudes available[1].

Proper reddening corrections are crucial to determine $T_{\rm eff}$ via IRFM. We have tested that 0.01 mag in $E(B-V)$ translates into an IRFM effective temperature roughly 50 K hotter. Reddening is usually zero for stars lying within the local bubble $\lesssim$70 pc from the Sun (e.g. Leroy 1993; Lallement et al. 2003) and so we have adopted $E(B-V)=0$ for all stars having *Hipparcos* parallaxes (van Leeuwen 2007) and satisfying this requirement on the distance. For the remaining stars we updated the reddening corrections in Ramírez & Meléndez (2005a) based on various extinction maps and, in particular for metal-poor stars when archive high resolution spectra were available, using interstellar NaD absorption lines (Meléndez et al. 2010). In broad-band photometry the definition of the effective wavelength of a filter ($\lambda_{\rm eff}$) shifts with the colour of the star (e.g. Bessell et al. 1998; Casagrande et al. 2006). Therefore a given $E(B-V)$ colour excess must be scaled according to the intrinsic colour of the source under investigation. From the reddening $E(B-V)$, we computed the extinction in each band adopting the reddening law of O'Donnell (1994) for the optical and Cardelli et al. (1989) for the infrared, using the improved estimation of the stellar intrinsic flux obtained at each iteration to bootstrap the computation of the correct $\lambda_{\rm eff}$ in our IRFM code.

2.2. The IRFM: pros and cons

The basic idea of the IRFM is to compare the ratio between the bolometric flux $\mathcal{F}_{\rm Bol}$(Earth) and the infrared monochromatic flux $\mathcal{F}_{\lambda_{\rm IR}}$(Earth), both measured at the top of Earth's atmosphere (the so-called observational $R_{\rm obs}$ factor) to the ratio between the surface bolometric flux ($\sigma T_{\rm eff}^4$) and the surface infrared monochromatic flux $\mathcal{F}_{\lambda_{\rm IR}}(T_{\rm eff}, {\rm [Fe/H]}, \log g)$ determined theoretically for any given set of stellar parameters. The latter is called the theoretical $R_{\rm theo}$ factor. For stars hotter than about 4200 K, infrared photometry longward of $\sim$1.2 μm ensures we are working in the Rayleigh-Jeans part of a stellar spectral energy distribution, a region largely dominated by the continuum which linearly depends on $T_{\rm eff}$ and thus only mildly on model atmospheres (Fig. 2). An extension of the technique to cooler effective temperatures using near-infrared photometry is possible, as shown by Casagrande et al. (2008), but this is outside the purpose of the present paper.

[1] Other than being available only for a limited number of stars, we did not use U magnitudes because of the little flux emitted in this region and the high uncertainties related to the absolute calibration and standardization of this passband in both observed and synthetic photometry (e.g. Bessell 2005, and references therein).

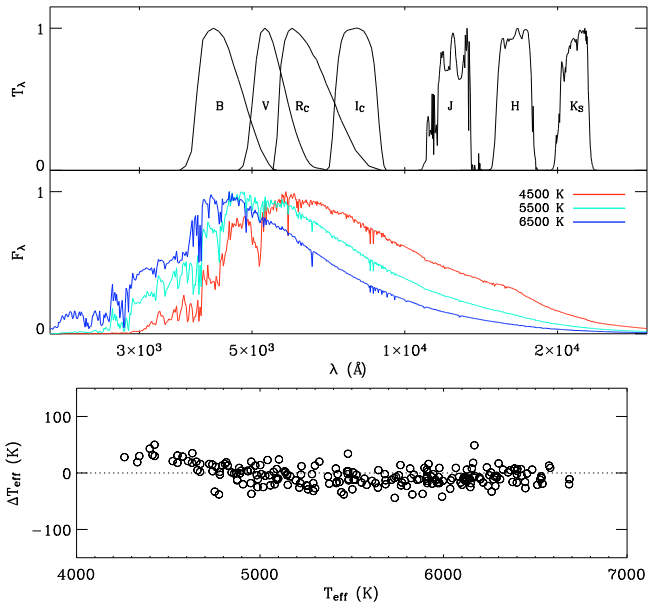

Fig. 2. *Top panel*: Johnson-Cousins-2MASS filter sets used in this work. *Middle panel*: synthetic solar metallicity spectra at different $T_{\rm eff}$. For the sake of comparison all curves have been normalized to unit. *Bottom panel*: difference in effective temperatures with – without using $(RI)_{\rm C}$ magnitudes to recover the bolometric flux.

$R_{\rm obs}$ and $R_{\rm theo}$ can be immediately rearranged to determine $T_{\rm eff}$, effectively reducing the entire problem to properly recover $\mathcal{F}_{\rm Bol}$(Earth) and $\mathcal{F}_{\lambda_{\rm IR}}$(Earth). Both quantities are determined from photometric observations, but an iterative procedure is adopted to cope with the mildly model dependent nature of the bolometric correction. In our case we use the fluxes predicted by the Castelli & Kurucz (2004) grid of model atmospheres starting with an initial estimate of the effective temperature and interpolating at the appropriate [Fe/H] and $\log g$ until convergence in $T_{\rm eff}$ is reached within 1 K. By doing so, we also obtain a synthetic spectrum tailored to the effective temperature empirically determined via IRFM.

Though we interpolate at the proper [Fe/H] and $\log g$ of each star, the dependence of the technique on such parameters is minor (e.g. Ramírez & Meléndez 2005a; Casagrande et al. 2006). This feature makes the IRFM superior to any spectroscopic methods to determine $T_{\rm eff}$ – provided the reddening is accurately known – since in the latter the effects of $T_{\rm eff}$, $\log g$ and [Fe/H] are usually strongly coupled and the model dependence is much more important.

The errors are estimated using realistic observational uncertainties in a Monte Carlo simulation plus the systematics arising from the adopted absolute calibration, as described in Casagrande et al. (2006). With the improved absolute calibration used in this paper, systematics amount to 15 K in $T_{\rm eff}$ and 0.3% in bolometric flux (Sect. 3.2). For stars approximately cooler than 5000 K, $(RI)_{\rm C}$ photometry is crucial to properly compute the bolometric flux. This can be appreciated in the lower panel of Fig. 2: below this temperature a trend appears using $BVJHK_{\rm S}$ magnitudes only. Missing the peak of the energy distribution clearly leads one to underestimate the bolometric flux thus returning cooler effective temperatures. We have linearly fitted the trend below 5000 K to remove such differences in both $T_{\rm eff}$ and $\mathcal{F}_{\rm Bol}$ when $(RI)_{\rm C}$ photometry was not available. For $T_{\rm eff} > 5000$ K no obvious trend appears: constant offsets of merely 7 K in $T_{\rm eff}$ and 0.15% in bolometric flux have been found, consistent with the effect that the absolute calibration in $(RI)_{\rm C}$ can introduce. For the sake of homogenizing the stellar parameters derived in this work, also these small offsets have been corrected for stars with no $(RI)_{\rm C}$ photometry.

The effective temperature can be determined from any infrared photometric band, in our case $JHK_{\rm S}$ from 2MASS. Ideally all bands should return the same $T_{\rm eff}$, but photometric errors and zero point uncertainties in the absolute calibration of each band introduce random plus systematic differences. In the case of 2MASS, those amount to few tens of K as we show later.

The magnitude in a given band ζ is converted into a physical flux (i.e. erg cm^{-2} s^{-1} Å^{-1}) via

$$\mathcal{F}_{\zeta}(\text{Earth}) = \mathcal{F}_{\zeta}^{\rm std}(\text{Earth})10^{-0.4(m_{\zeta}-m_{\zeta}^{\rm std})} \quad (1)$$

which depends on the zero point ($m_{\zeta}^{\rm std}$) and the absolute flux calibration ($\mathcal{F}_{\zeta}^{\rm std}$) of the standard star defining the photometric system under use[2].

Most of the photometric systems, including Johnson-Cousins and 2MASS, use Vega as the zero point standard. Vega's flux and magnitudes in different bands have been notoriously difficult to measure with sufficient accuracy (e.g. Gray 2007, and references therein). The problem is only apparently resolved when resorting to $R_{\rm obs}$: in the ideal case of a unique template spectrum for Vega the choice of its absolute calibration would cancel out in the ratio. In practice, the situation is far from this since the pole-on and rapidly rotating nature of this star imposes the use of a composite absolute calibrated spectrum for different wavelength regions (e.g. Casagrande et al. 2006, and references therein). Such complication does not disqualify Vega as a spectrophotometric standard, but it makes its use more problematic. From Eq. (1) it can be immediately noticed that a change of 0.01 mag corresponds to a change of about 1% in flux. Since it is possible to interchangeably operate on both zero points and fluxes, for the sake of our discussion it is their composite effect that must be considered, though in the following we shall usually refer to fluxes.

Recently, HST spectrophotometry for Vega has provided a unique calibrated spectrum extending from 3200 to 10 000 Å with 1–2% accuracy (Bohlin 2007). In the infrared, once the zero points newly determined from Maíz-Apellániz (2007) are used, this result is also in broad agreement with the 2MASS absolute calibration provided by Cohen et al. (2003). Rieke et al. (2008) have also recently reviewed the absolute physical calibration in the infrared, substantially validating the accuracy of 2MASS: their recommended 2% increase of flux in $K_{\rm S}$ band is in fact compensated by their newly determined zero point for Vega, thus implying an effective change in the overall $K_{\rm S}$ calibration of only 0.2%. We have tested all these different possibilities; with respect to the HST and 2MASS calibration adopted in Casagrande et al. (2006) the derived $T_{\rm eff}$ are affected at most by 20 K. Such difference is thus within the aforementioned global 2% uncertainty which allows for systematics in $T_{\rm eff}$ of order 40 K. Our zero points and absolute fluxes are essentially identical to those adopted in Casagrande et al. (2006) except for a small fine-tuning which will be further discussed in Sect. 3.

[2] We point out that Eq. (1) holds exactly for a heterochromatic measurement, while for computing a monochromatic flux from the observed photometry, an additional correction (the so called q-factor) must be introduced to account for the fact that the zero point of the photometric system is defined by a standard star, which usually has a different spectral energy distribution across the filter window with respect to the problem star (e.g. Alonso et al. 1996a; Casagrande et al. 2006).

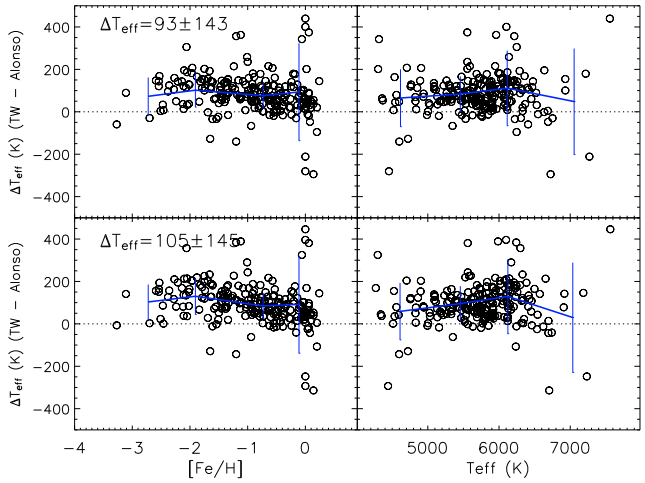

Fig. 3. Difference between the effective temperatures obtained in this work (TW) and those reported in Alonso et al. (1996a) for 220 stars in common. In case of reddening, only stars with values of $E(B-V)$ equal to within 0.02 mag have been plotted. Thick continuous lines connect the means computed in equally spaced bins of [Fe/H] and $T_{\rm eff}$. Error bars are the standard deviation in each bin. Top panels: when Kurucz (1993) models are used in our version of the IRFM. Bottom panels: when the new Castelli & Kurucz (2004) models are used instead. Below [Fe/H] = −1.5 the new models support $T_{\rm eff}$ hotter by 20 to 40 K.

Despite the recent increasing concordance in establishing absolute fluxes, the uncertainties which have historically plagued Vega are crucial in the context of understanding the effective temperatures determined via IRFM by various authors. We have tested that uncorrelated changes of a few percent in the absolute calibration of optical bands (needed to recover the bolometric flux) can introduce spurious trends with $T_{\rm eff}$ and [Fe/H] up to few tens of K. Similar changes in the absolute calibration of infrared bands have only minor impact on the bolometric flux, but as already mentioned, $T_{\rm eff}$ is very sensitive to them since they enter explicitly in the definition of $R_{\rm obs}$: increasing all of them by 2% translates into a decrease of approximately 40 K in $T_{\rm eff}$. Considering that differences of few percent in the adopted zero points and fluxes are commonly present among various IRFM implementations, it can be immediately realized that they are responsible for systematic differences among various authors.

2.3. Alonso et al. (1996) scale

One of the most extensive applications of the IRFM to Pop I and II dwarfs is that of Alonso et al. (1996a), which was based on the infrared photometry collected at the TCS (Telescopio Carlos Sanchez, Alonso et al. 1994b) and absolutely calibrated using a semi-empirical approach relying on (mostly) giant stars with measured angular diameters to determine the reference absolute fluxes (Alonso et al. 1994a). The comparison between our $T_{\rm eff}$ and those by Alonso et al. (1996a) is shown in Fig. 3. Despite the scatter arising from the different input data we used, there is a clear offset with our scale being systematically hotter. No obvious trends in $T_{\rm eff}$ and [Fe/H] appear. This offset is easily explained in terms of the absolute calibration underlying the two different photometric systems adopted. This involves the transformation from TCS to 2MASS system (see also the discussion in Casagrande et al. 2006), which could in principle introduce additional noise (see Sect. 2.4). A more detailed description of the absolute calibration (and angular diameters) employed by Alonso and a comparison with our own is presented in Appendix A.

An area of particular interest is the determination of effective temperatures in very metal-poor, turn-off stars. We have tested the effect of using the new Castelli & Kurucz (2004) model atmospheres in the IRFM instead of the Kurucz (1993) adopted by Alonso et al. (1996a). The IRFM is known to be little model dependent (e.g. Asplund & García Pérez 2001; Casagrande 2009) and in fact there are no big differences except at the lowest metallicities, where Castelli & Kurucz (2004) support effective temperatures hotter by ∼40 K. The reason for such a discrepancy stems from the new models returning higher flux below ∼4000 Å, a region where the most metal-poor, turn-off stars commence emitting non negligible amounts of energy. Since we do not have UV photometry (and its standardization would be uncertain), we must rely on model atmospheres to determine the flux over this region (Fig. 4). The latest model atmosphere calculations show excellent agreement as we checked that nearly identical $T_{\rm eff}$ are obtained when the new MARCS models (Gustafsson et al. 2008) are used instead of those by Castelli & Kurucz (2004) (also Sect. 5.3.1), but see Edvardsson (2008) for a discussion of the performance of model atmospheres in the blue and ultraviolet.

2.4. Ramírez & Meléndez (2005) scale

A revision of the Alonso et al. (1996a) implementation of the IRFM was carried out by Ramírez & Meléndez (2005a) based on the TCS (for the computation of $R_{\rm theo}$) and Johnson's (for the computation of the bolometric fluxes) *JHK* photometric systems (Alonso et al. 1994b; Bessell & Brett 1988). Here we replicate the $T_{\rm eff}$ determination by Ramírez & Meléndez (2005a) for comparison purposes. When running their implementation, we transformed the 2MASS photometry into TCS using their equations. However, when comparing the transformed and original *JHK* values for these stars we found zero point differences at the level of 0.01 mag: these offsets are within the photometric uncertainties and smaller than the scatter in the fits leading to the transformation equations, but they introduce changes in the derived $T_{\rm eff}$ values up to few tens of K (see Sect. 2.2). Therefore we took those into account to precisely transform 2MASS data into the TCS system.

The Ramírez & Meléndez (2005a) bolometric fluxes were determined using the K-band bolometric correction calibration by Alonso et al. (1995), which depends only on the Johnson $(V-K)$ colour index and the stellar metallicity[3]. This calibration is internally accurate within its ranges of applicability and one would expect that extrapolations slightly outside these ranges would still provide reliable results at low metallicities. This approach was followed by Ramírez & Meléndez (2005a). With regards to the absolute flux calibration in the infrared, Ramírez & Meléndez (2005a) adopted that of Alonso et al. (1994a), which is valid for TCS *JHK* photometry while we use an update of Cohen et al. (2003) for the $JHK_{\rm S}$ 2MASS system (see also Sect. 3).

The difference between our results and Ramírez & Meléndez (2005a) when the same input data and reddening values are adopted is illustrated in the top panels of Fig. 5. Some of the scatter arise from transforming 2MASS magnitudes into TCS, but

[3] We have also tested that in the context of computing bolometric fluxes for this work, the updated J. Carpenter transformations from 2MASS to Johnson available online at: `http://www.astro.caltech.edu/~jmc/2mass/v3/transformations` are instead accurate enough and insensitive to small zero point changes.

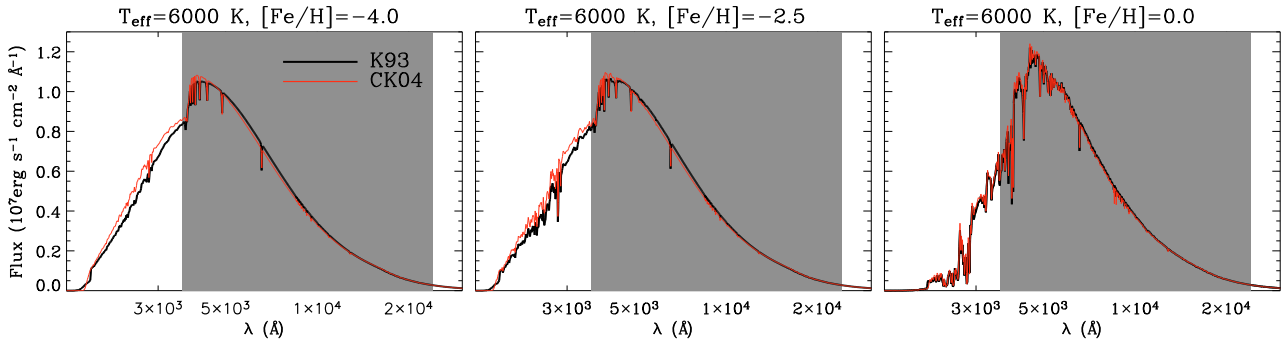

Fig. 4. Comparison between Kurucz (1993) (thick line) and Castelli & Kurucz (2004) (thin line) synthetic spectra at different metallicities for an assumed $\log g = 4.0$. Shaded area is the wavelength region covered by our multiband photometry. The difference in the UV flux gets more prominent when going to more metal-poor stars, but for the sake of the IRFM is entirely negligible at solar metallicity.

clear trends with both with $T_{\rm eff}$ and [Fe/H] are present. For the bulk of the stars with [Fe/H] > -2.0 and $4800 < T_{\rm eff} < 6200$ K a roughly constant offset of about 100 K is observed, our stars being hotter. In the metal-rich regime such an offset is present also for hotter stars ($T_{\rm eff} > 6200$ K), but reduces somewhat for the coolest metal-rich dwarfs, reaching a minimum of about 50 K at $T_{\rm eff} \simeq 4500$ K. A steep trend is seen for moderately metal-poor dwarfs ($-2.0 <$ [Fe/H] < -1.0) below 4800 K, a region with few or no calibrating stars in Alonso et al. (1995). For the warmer, most metal-poor stars in the sample, the differences decrease sharply with increasing $T_{\rm eff}$ and decreasing [Fe/H], quickly becoming negative i.e., Ramírez & Meléndez (2005a) temperatures become warmer, reaching a maximum value of about -100 K at $T_{\rm eff} \simeq 6500$ K and [Fe/H] $\simeq -3.5$.

To investigate the source of these differences, we recalculated the IRFM temperatures of Ramírez & Meléndez (2005a) using our bolometric fluxes instead of the calibration formulae adopted by Ramírez & Meléndez (2005a). This choice is perfectly legitimate, since what is crucial in the IRFM are the infrared fluxes which appear explicitly in the definition of $R_{\rm obs}$, while $T_{\rm eff}$ depends only mildly on the bolometric flux (Sect. 2.2). Therefore, adopting our bolometric fluxes is substantially independent of the underlying temperature scale, i.e. the Ramírez & Meléndez (2005a) scale is still recovered despite now using the new bolometric fluxes determined in the present work. The result of this exercise is shown in the bottom panels of Fig. 5. The major trends caused from extrapolating the Alonso et al. (1995) bolometric formulae now disappear with a constant offset $\Delta T_{\rm eff} = 85 \pm 13$ K above 5000 K. The small trend that remains below this temperature corresponds to the threshold where Ramírez & Meléndez (2005a) stop using the J band to determine $T_{\rm eff}$, which in the TCS system usually returns slightly cooler $T_{\rm eff}$ than H and K bands.

From this comparison it is clear that Ramírez & Meléndez (2005a) temperatures for the metal-poor turn-off stars are warmer due to the use of a photometric calibration to derive the bolometric fluxes. In fact, we realize that the Alonso et al. (1995) formula is robust down to [Fe/H] $\simeq -2.5$ and up to $T_{\rm eff} \simeq 6500$ K but only a few calibrating stars more metal-poor or warmer exist in their sample. Ramírez & Meléndez (2005a) use of this formula in regions where the calibration is uncertain (and in some cases outside of the ranges of applicability) has resulted in the very high temperatures of the more metal-poor turn-off stars. The extrapolation is, of course, not a valid procedure, even though one might expect the [Fe/H] dependence of the calibration not to be so important at low metallicity. However, as can be seen from Fig. 4 in Alonso et al. (1995), at these relatively high

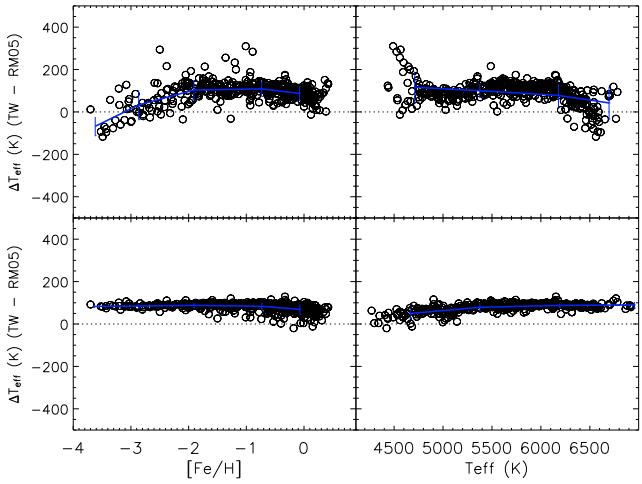

Fig. 5. *Top panels*: difference between the effective temperatures of this work (TW) and those obtained when the same input data are used in the Ramírez & Meléndez (2005a) implementation (RM05). *Bottom panels*: as in the *top panels* but for the Ramírez & Meléndez (2005a) temperatures re-determined using the bolometric fluxes obtained in this work.

temperatures, the effect of [Fe/H] is very important and such extrapolations should not be performed.

The difference that remains after adopting consistent bolometric fluxes between this work and Ramírez & Meléndez (2005a) (lower panels of Fig. 5) is mostly due to the use of different infrared absolute flux calibrations. In fact, by lowering the absolute fluxes adopted by Ramírez & Meléndez (2005a) by about 4%, the mean difference reduces to almost zero. We thus conclude that our and Ramírez & Meléndez (2005a) IRFM implementations can be made perfectly compatible if the same input parameters and flux calibration are used.

2.5. González Hernández & Bonifacio (2009) scale

The most recent work on the IRFM is that by González Hernández & Bonifacio (2009), which is also based on 2MASS photometry. The main difference between theirs and our implementation is the different absolute calibration and zero points adopted for Vega. They based their work on the Castelli & Kurucz (1994) model and McCall (2004) magnitudes instead of the HST (Bohlin & Gilliland 2004; Bohlin 2007) and 2MASS (Cohen et al. 2003) values that we use. Although such differences are within the current observational errors, in the infrared the combined effect of their fluxes and

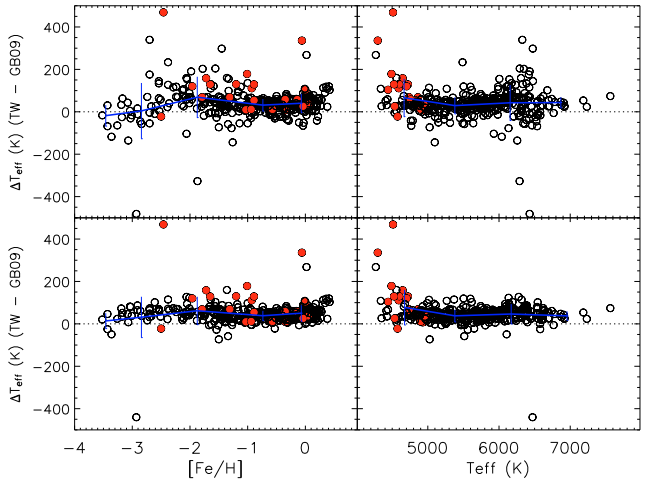

Fig. 6. *Top panels*: difference between the effective temperatures of this work (TW) and those in González Hernández & Bonifacio (2009) (GB09) for 380 stars in common. Filled circles are stars with $T_{\rm eff}$ < 5000 K without $(RI)_{\rm C}$ photometry in GB09. *Bottom panels*: as in the *top panels*, but when the same reddening corrections are used.

zero points is on average 1.5−2.0% higher than ours, implying effective temperatures cooler by 30−40 K (see Appendix A). This can be immediately appreciated in Fig. 6, which indeed shows a constant offset of this magnitude for stars in common, thus confirming the offset noticed by González Hernández & Bonifacio (2009) for stars in common with Casagrande et al. (2006).

The very steep trend at the lowest metallicities is due to the different reddening corrections we adopt with respect to theirs. When the same $E(B-V)$ values are adopted (bottom panels in Fig. 6), the offset remains constant throughout the entire [Fe/H] and $T_{\rm eff}$ range, except for few outliers due to the different input data (mostly optical photometry) adopted. This clearly stresses the importance of proper reddening correction for determining effective temperatures via IRFM in stars outside of the local bubble. For the most metal-poor stars in the sample, we use interstellar NaD lines to achieve higher precision (Sect. 2.1) while González Hernández & Bonifacio (2009) resorted to reddening maps scaled by the distance and the galactic latitude of the star and scale height of the dust layer. The trend towards cooler effective temperatures that we obtain in this regime thus stem entirely from better reddening corrections. Finally, we suspect that the trend for $T_{\rm eff}$ < 5000 K is due to the absence of $(RI)_{\rm C}$ colours in González Hernández & Bonifacio (2009) (Sect. 2.2, bottom panel of Fig. 2).

3. Resolving different versions

It is clear from the discussion above that we now understand where different $T_{\rm eff}$ scales originate from and the crucial role played by the absolute calibration. Our approach has been to adopt the latest calibration available for each photometric system: currently those are accurate at the 2% level, implying possible systematic uncertainties of order 40 K. Here we want to improve upon this uncertainty using an independent verification of the absolute calibration adopted.

3.1. Solar twins

The use of solar-type stars to calibrate photometric systems has a long and noble history, which relies on taking absolutely calibrated measurements of the Sun and computing synthetic colours to compare with other solar-type stars (e.g. Johnson 1965; Campins et al. 1985; Rieke et al. 2008). This rationale can be extended to other physical properties, namely using the solar effective temperature $T_{\rm eff,\odot}$ = 5777 K as the average value for solar-type stars (e.g. Masana et al. 2006). This technique is well established and goes under the name of solar analogs method, but there is some sort of *petitio principii* in the underlying $T_{\rm eff}$ scale adopted and/or the solar colours assumed to select solar analogs in first instance.

A way to break such a degeneracy is provided by solar twins, i.e. stars with spectra indistinguishable from the Sun (Cayrel de Strobel & Bentolila 1989; Porto de Mello & da Silva 1997). Our twins were drawn from an initial sample of about 100 stars broadly selected to be solar like: the identification of the best ones was based on a strictly differential analysis of high-resolution ($R \sim 60\,000$) and high signal-to-noise ($S/N \gtrsim 150$) spectra with respect to the solar one reflected from an asteroid and observed with the same instrument. Within this initial sample, the selection criterion adopted to identify the best twins did not assume any a priori effective temperature or colour, but was based on the measured relative difference in equivalent widths and equivalent widths vs. excitation potential relations with respect to the observed solar reference spectrum and thus entirely model independent (Meléndez et al. 2006a; Meléndez & Ramírez 2007). Since the spectra of the solar twins match so closely the solar one, exceedingly accurate differential spectroscopic analysis with respect to $T_{\rm eff,\odot}$, $[{\rm Fe/H}]_\odot$ and $\log g_\odot$ is possible (Meléndez et al. 2009a; Ramírez et al. 2009).

Ten stars were identified as most closely resembling the Sun and are given in Table 2, including HIP56948, the best solar twin currently known (Meléndez & Ramírez 2007; Takeda & Tajitsu 2009). A crucial requirement for these stars is to have accurate and homogeneous photometry in order to derive reliable $T_{\rm eff}$ via IRFM. While this is possible in the infrared because of 2MASS[4], optical photometry is also important to properly recover the bolometric flux where these stars emit most of their energy. Johnson-Cousins photometry would be the ideal choice, but unfortunately is not available for all these targets. To overcome this limitation, in the optical we used the Tycho2 $B_{\rm T}V_{\rm T}$ system which uniformly and precisely covers the entire sky in the magnitude range of our interest (Høg et al. 2000). Notice that we did not transform $B_{\rm T}V_{\rm T}$ into BV but instead implemented our IRFM code to work directly on the Tycho2 system. Also, as discussed in Sect. 2.2 the absence of $(RI)_{\rm C}$ photometry is not relevant for stars hotter than 5000 K. All twins are closer than 72 pc, where reddening is expected to be zero or negligible: nearly all of them have Strömgren photometry (Meléndez et al. in prep.) and the Schuster & Nissen (1989) reddening calibration confirms indeed such a conclusion.

3.2. A finely tuned absolute calibration

As for the Johnson-Cousins system, we based the absolute calibration of the Tycho2 system on Vega (Bohlin & Gilliland 2004; Bohlin 2007), adopting the $B_{\rm T}V_{\rm T}$ zero points of

[4] In fact, the other well known solar twin 18 Sco (Porto de Mello & da Silva 1997) has saturated 2MASS colours.

Table 2. Tycho2 and 2MASS photometry for our solar twins sample.

HIP	B_T	σ_B	V_T	σ_V	J	σ_J	H	σ_H	K_S	σ_K	$T_{\rm eff}^{\rm spec}$ ±20 K	log g ±0.04 dex	[Fe/H] ±0.022 dex	$T_{\rm eff}^{\rm IRFM}$ (K)
30502	9.483	0.019	8.706	0.013	7.474	0.029	7.139	0.029	7.069	0.024	5745	4.47	−0.01	5760 ± 28
36512	8.498	0.015	7.786	0.011	6.517	0.020	6.213	0.027	6.154	0.024	5755	4.53	−0.08	5763 ± 26
41317	8.613	0.015	7.868	0.010	6.610	0.023	6.289	0.038	6.206	0.024	5740	4.49	−0.02	5739 ± 27
44935	9.522	0.021	8.783	0.015	7.548	0.019	7.260	0.034	7.171	0.024	5800	4.41	0.07	5803 ± 30
44997	9.122	0.017	8.378	0.012	7.107	0.021	6.888	0.051	6.764	0.026	5790	4.52	0.03	5791 ± 30
55409	8.793	0.017	8.066	0.011	6.811	0.019	6.493	0.042	6.419	0.021	5760	4.52	−0.01	5758 ± 26
56948	9.462	0.017	8.748	0.012	7.477	0.019	7.202	0.026	7.158	0.018	5782	4.38	0.01	5801 ± 25
64713	10.048	0.029	9.280	0.021	8.086	0.018	7.771	0.026	7.707	0.034	5815	4.52	−0.01	5853 ± 36
77883	9.532	0.023	8.820	0.018	7.476	0.021	7.176	0.038	7.125	0.034	5695	4.39	0.04	5660 ± 35
89650	9.708	0.023	8.996	0.017	7.781	0.029	7.506	0.034	7.431	0.033	5855	4.48	0.02	5864 ± 35

Notes. Spectroscopic parameters are from Meléndez et al. (2009a) and the effective temperatures are determined via IRFM. For the latter, the errors are those arising from the photometry alone, not including the 15 K uncertainty in the zero point of our temperature scale. All twins have "A" quality flag and Read 1 mode in all 2MASS bands.

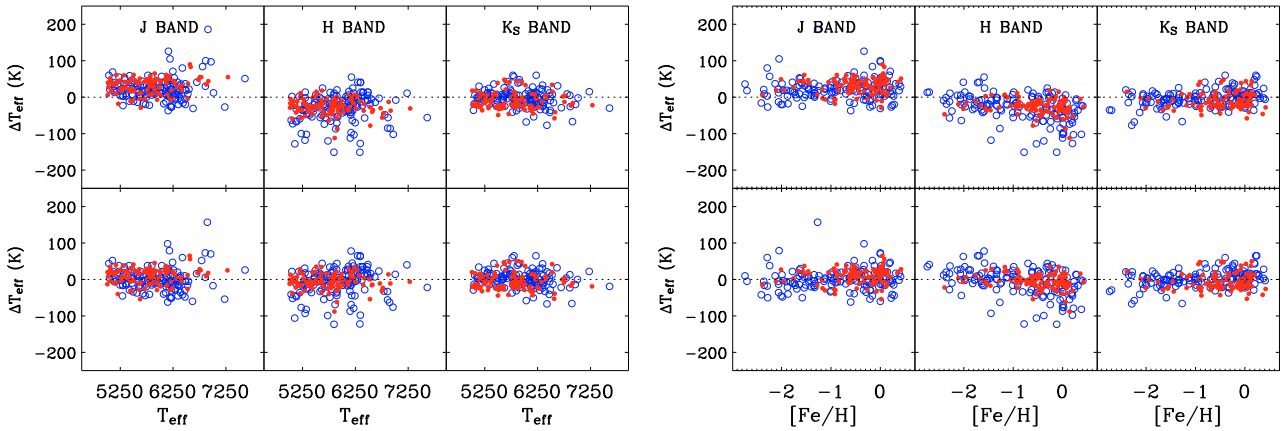

Fig. 7. *Top panels*: difference between $T_{\rm eff}$ and the effective temperature determined in each infrared band before tuning the absolute calibration. Full circles are stars with quality flag "A", Read 1 mode and total 2MASS photometric errors <0.07 mag while open circles are for all other stars. *Bottom panels*: as in the *top panels*, but with the adjusted absolute calibration. H band photometry has usually slightly higher error than J and K_S and the final temperature is the weighted average of that obtained in each band.

Maíz-Apellániz (2007) and the corresponding filter transmission curves of Bessell (2000).

In the first instance, we determined $T_{\rm eff}$ via IRFM for each of the twins in Table 2: their average effective temperature turned out to be 5782 K, remarkably close to $T_{\rm eff,\odot}$, thus confirming the high accuracy achieved using the HST and 2MASS absolute calibration. Based on Monte Carlo simulations with the photometric errors in Table 2, the uncertainty in $T_{\rm eff}$ determined via IRFM is of order 30 K for single stars. Imposing the mean effective temperature of all solar twins to equal $T_{\rm eff,\odot}$ we estimate the uncertainty on the zero point of our temperature scale to be 15 K based on a bootstrap procedure with one million re-samples. At the same time, for HIP56948 we also recover $T_{\rm eff,\odot}$ within 1σ.

Though the solar twins test confirms the global reliability of the adopted absolute calibration, for all stars in Sect. 2.1 having Tycho2 photometry and $T_{\rm eff} > 5000$ K we further required each infrared band to return on average the same $T_{\rm eff}$ as the others (Fig. 7). By imposing such a consistency we improve upon small systematic trends which could arise when determining effective temperatures in stars with $T_{\rm eff}$ and [Fe/H] very different from our solar twins. This led to a decrease of the absolute calibration by 1.6% in the J band and an increase by 1.5 and 0.3% in the H and K_S bands, respectively (see also Appendix A). In terms of synthetic magnitudes these differences make H and K_S redder by 0.016 and 0.003 and J bluer by 0.017, thus removing almost entirely the infrared colour offsets found by Casagrande et al. (2006) when comparing observed and synthetic photometry. We cannot entirely rule out whether these systematic differences arise from the adopted synthetic library or the absolute calibration, but since the IRFM depends only marginally on model atmospheres and the infrared spectral region is relatively easy to model, we are strongly in favour of the second possibility. From a pragmatic point of view, this improves the consistency in determining $T_{\rm eff}$. Also, such changes are within the 2MASS quoted errors and for the K_S band we remark the agreement with the 0.2% increase found by Rieke et al. (2008) and discussed in Sect. 2.2. As expected, stars with the best 2MASS pedigree also return better agreement in all bands (full circles in Fig. 7). We have also checked that the increasing scatter in Fig. 7 is primarily due to photometric errors. We recall that Rieke et al. (2008) found a 2% offset between Read 1 and Read 2 mode in 2MASS[5], though they were not able to derive a universal correction for this effect. All our solar twins have Read 1 mode and the absence of a universal correction suggests that while Read mode 2 can decrease the precision of $T_{\rm eff}$ the overall accuracy of our calibration remains valid.

[5] This mode indicates which readout is used to derive photometry http://www.ipac.caltech.edu/2mass/releases/allsky/doc/sec3_1b.html.

With the fine-tuning discussed above, the median (mean) effective temperature of our solar twins is 5777 (5779) K. Restricting only to the twins having $T_{\rm eff,\odot}$ within the observational errors, still confirm such conclusion. As a further independent test, we applied our IRFM to the list of solar analogs used by Rieke et al. (2008) and determined their median (mean) $T_{\rm eff}$ to be 5791 (5786) K, thus confirming the reliability of the zero point of our temperature scale, which has an uncertainty of 15 K. Such a value implies possible systematics in the absolute calibration at the 1% level. The systematic error in recovering the bolometric luminosity is however smaller since infrared fluxes enter twice in $R_{\rm obs}$, thus partly compensating their uncertainty.

The corrections in the infrared absolute calibration discussed here have been used also in determining $T_{\rm eff}$ for stars in Sect. 2.1. Since for those stars we are using Johnson-Cousins photometry, there could still be small differences arising from the absolute calibration in the optical: for stars in common a mean systematic of 8 K in $T_{\rm eff}$ and 0.15% in bolometric flux was found and corrected.

4. Validating the proposed temperature scale

The IRFM determines $T_{\rm eff}$ in an almost model independent way, primarily recovering the bolometric flux $\mathcal{F}_{\rm Bol}$(Earth) of the star under investigation. From the basic definition linking those two quantities the stellar angular diameter $\theta_{\rm IRFM}$ can be obtained self-consistently and this was actually one of the driving reasons for developing the technique (Blackwell & Shallis 1977). In what follows, we use this information to further validate our results.

4.1. Interferometric angular diameters

An independent test of accuracy for the zero point of our effective temperature scale involves the comparison with the angular diameters measured using interferometric techniques (corrected for limb-darkening, hereafter denoted by $\theta_{\rm LD}$). In our case, angular diameters are a natural consequence of the $T_{\rm eff}$ determination procedure and for each star the $T_{\rm eff}, \mathcal{F}_{\rm Bol}, \theta_{\rm IRFM}$ values are self-consistent, i.e., they represent a unique solution for a given set of input data. We also prefer to compare angular diameters directly (i.e. $\theta_{\rm IRFM}$ vs. $\theta_{\rm LD}$) since the effective temperatures reported in various interferometric works would be more heterogeneous because of the adopted bolometric corrections.

Given the difficulties involved in the measurement of the small angular diameters of dwarfs and subgiants (even the nearest ones have angular diameters below 10 milli-arcsec), only a relatively small group of such stars has been observed to date for that purpose (see also Appendix A for a discussion of the angular diameters used by Alonso et al. 1994a). We performed a literature search for interferometrically determined angular diameters with precision better than 5% (which corresponds to an accuracy of 2.5% in effective temperatures, roughly 150 K at solar temperature, assuming no error in the bolometric flux) and found data for 28 stars, 16 of which have $\theta_{\rm LD}$ measured to better than 2% (Table 3). The efforts made by the interferometry community in the last few years are commendable given that the number of stars with reliable $\theta_{\rm LD}$ has nearly doubled since 2005 (cf. Ramírez & Meléndez 2005a).

Unfortunately, all dwarfs and subgiants with reliable $\theta_{\rm LD}$ are brighter than $V \simeq 6$, implying infrared magnitudes $\lesssim$5 where 2MASS photometry has large observational errors and starts to saturate[6]. Therefore we cannot apply our IRFM directly on them to get $\theta_{\rm IRFM}$. Instead, we adopt an indirect approach using the photometric $T_{\rm eff}$:colour and $\mathcal{F}_{\rm Bol}$:colour relations presented in Sect. 6. Using the photometry of our sample stars (i.e. those used in the construction of the calibrations and therefore with $T_{\rm eff}$ directly determined via IRFM), we checked that the zero point of our $T_{\rm eff}$ and $\mathcal{F}_{\rm Bol}$ scales is correctly reproduced by the calibration formulae presented in Sect. 6, independently of the apparent magnitudes of the stars. Also, for the two stars having HST spectrophotometry (next section) we checked that our calibration formulae reproduce nearly the same results as directly applying the IRFM. We were careful about propagating all possible sources of random error such as uncertainties in the input photometry, metallicity, and the reliability of the colour calibrations, as quantified by the standard deviation of each polynomial fit (Tables 4 and 5). For most of the stars with reliable $\theta_{\rm LD}$ (i.e. better than 2%), only BV photometry was available, while for the remaining $BV(RI)_{\rm C}$ was used. Metallicities were adopted from the updated version of the Cayrel de Strobel et al. (2001) [Fe/H] catalog by Meléndez (in prep.), which nearly triples the number of entries in the original catalog.

The comparison of the angular diameters measured interferometrically with those derived using our IRFM colour calibrations is shown in Fig. 8 (see also Table 3). Stars that have $\theta_{\rm LD}$ determined with accuracy better than 2% are shown with full symbols. Using only the latter, the average difference in

Table 3. Stars with measured interferometric angular diameters.

HD	$\theta_{\rm LD}$ mas	Ref.[a]	$T_{\rm eff}^{\rm IRFM}$ K	[Fe/H] dex	$\theta_{\rm IRFM}$ mas
3651	0.790 ± 0.027	1	5234	0.15	0.756 ± 0.022
6582	0.973 ± 0.009	2	5403	−0.84	0.954 ± 0.021
9826	1.114 ± 0.009	1	6151	0.10	1.121 ± 0.023
10700	2.078 ± 0.031	3	5364	−0.53	2.089 ± 0.026
10780	0.763 ± 0.021	2	5317	0.01	0.806 ± 0.022
19994	0.788 ± 0.026	1	6020	0.18	0.746 ± 0.009
22049	2.148 ± 0.029	3	5056	−0.09	2.200 ± 0.032
23249	2.394 ± 0.029	3	5060	0.08	2.399 ± 0.059
26965	1.650 ± 0.060	3	5188	−0.27	1.482 ± 0.018
61421	5.443 ± 0.030	3	6626	0.00	5.326 ± 0.068
75732	0.854 ± 0.024	1	5282	0.38	0.718 ± 0.025
102870	1.450 ± 0.018	4	6100	0.13	1.426 ± 0.014
117176	1.009 ± 0.024	1	5540	−0.06	0.969 ± 0.021
120136	0.786 ± 0.016	1	6407	0.28	0.840 ± 0.019
121370	2.244 ± 0.019	3	6052	0.26	2.214 ± 0.043
128620	8.511 ± 0.020	3	5772	0.20	8.511 ± 0.079
128621	6.000 ± 0.021	5	5217	0.23	6.151 ± 0.234
131977	1.230 ± 0.030	3	4633	0.04	1.162 ± 0.054
150680	2.397 ± 0.044	3	5780	0.03	2.352 ± 0.055
161797	1.953 ± 0.039	3	5520	0.22	2.004 ± 0.050
185144	1.254 ± 0.012	2	5293	−0.21	1.261 ± 0.029
188512	2.180 ± 0.090	6	5164	−0.18	2.070 ± 0.049
190360	0.698 ± 0.019	1	5564	0.21	0.673 ± 0.017
198149	2.650 ± 0.040	3	4980	−0.16	2.720 ± 0.090
201091	1.775 ± 0.013	3	4429	−0.24	1.706 ± 0.070
209100	1.890 ± 0.020	3	4665	−0.06	1.825 ± 0.021
217014	0.748 ± 0.027	1	5754	0.17	0.698 ± 0.019

Notes. [(a)] 1. Baines et al. (2008); 2. Boyajian et al. (2008); 3. Kervella & Fouqué (2008) (weighted average if more than one measurement was available); 4. North et al. (2009); 5. Bigot et al. (2006); 6. Nordgren et al. (1999).

[6] www.ipac.caltech.edu/2mass/releases/allsky/doc/sec2_2.html#pscphotprop

Table 4. Coefficients and range of applicability of the colour-temperature-metallicity relations.

Colour	[Fe/H] range	Colour range	a_0	a_1	a_2	a_3	a_4	a_5	N	$\sigma(T_{\rm eff})$
$B-V$	$[-5.0, 0.4]$	$[0.18, 1.29]$	0.5665	0.4809	−0.0060	−0.0613	−0.0042	−0.0055	400	73
$V-R_{\rm C}$	$[-5.0, 0.3]$	$[0.24, 0.80]$	0.4386	1.4614	−0.7014	−0.0807	0.0142	−0.0015	201	62
$(R-I)_{\rm C}$	$[-5.0, 0.3]$	$[0.23, 0.68]$	0.3296	1.9716	−1.0225	−0.0298	0.0329	0.0035	211	82
$V-I_{\rm C}$	$[-5.0, 0.3]$	$[0.46, 1.47]$	0.4033	0.8171	−0.1987	−0.0409	0.0319	0.0012	208	59
$V-J$	$[-5.0, 0.4]$	$[0.61, 2.44]$	0.4669	0.3849	−0.0350	−0.0140	0.0225	0.0011	401	42
$V-H$	$[-5.0, 0.4]$	$[0.67, 3.01]$	0.5251	0.2553	−0.0119	−0.0187	0.0410	0.0025	401	33
$V-K_{\rm S}$	$[-5.0, 0.4]$	$[0.78, 3.15]$	0.5057	0.2600	−0.0146	−0.0131	0.0288	0.0016	401	25
$J-K_{\rm S}$	$[-5.0, 0.4]$	$[0.07, 0.80]$	0.6393	0.6104	0.0920	−0.0330	0.0291	0.0020	412	132
$(B-V)_{\rm T}$	$[-2.7, 0.4]$	$[0.19, 1.49]$	0.5839	0.4000	−0.0067	−0.0282	−0.0346	−0.0087	251	79
$V_{\rm T}-J$	$[-2.7, 0.4]$	$[0.77, 2.56]$	0.4525	0.3797	−0.0357	−0.0082	0.0123	−0.0009	272	43
$V_{\rm T}-H$	$[-2.7, 0.4]$	$[0.77, 3.16]$	0.5286	0.2354	−0.0073	−0.0182	0.0401	0.0021	263	26
$V_{\rm T}-K_{\rm S}$	$[-2.4, 0.4]$	$[0.99, 3.29]$	0.4892	0.2634	−0.0165	−0.0121	0.0249	−0.0001	258	18
$b-y$	$[-3.7, 0.5]$	$[0.18, 0.72]$	0.5796	0.4812	0.5747	−0.0633	0.0042	−0.0055	1120	62

Notes. The photometric systems are Johnson-Cousins $BV(RI)_{\rm C}$, 2MASS $JHK_{\rm S}$, Tycho2 $(BV)_{\rm T}$ and Strömgren by. For the latter, additional corrections as function of [Fe/H] and $(b-y)$ apply, as discussed in Sect. 6.1.1. For some indices the calibrations are given down to [Fe/H] = −5.0, meaning that the effective temperatures of such a metal-poor star can be recovered using [Fe/H] = −3.5 in Eq. (3). Notice that only two hyper metal-poor stars are currently known and caution should be used, as discussed in the text. Especially for metal-poor stars, please refer to Fig. 13 to check that the calibration is not extrapolated outside its [Fe/H] range. N is the number of stars employed for the fit after the 3σ clipping and $\sigma(T_{\rm eff})$ is the standard deviation (in Kelvin) of the proposed calibrations. Notice that the standard deviation does not account for the uncertainty in the zero point of the temperature scale, which is of order 15−20 K (Sects. 3.2 and 4.1).

angular diameter (IRFM-LD) is −0.62 ± 1.70% which corresponds to a zero point difference in the effective temperature scale of only +18 ± 50 K at solar temperature. This is also in agreement with the uncertainty on the zero point of our temperature scale discussed in Sect. 3.2. No obvious trends are seen with [Fe/H] (from about −0.8 to +0.3) or $T_{\rm eff}$ (from 4400 to 6600 K). Note, however, that if we exclude the two coolest stars (from the group of those having errors smaller than 2%), a small trend is seen with $T_{\rm eff}$. The trend – if real – appears more clearly for early type stars, with $\theta_{\rm IRFM}$ being underestimated (and therefore the IRFM effective temperatures overestimated) with respect to the interferometric measurements. Interferometry resorts on 1D model atmospheres to correct from the measured uniform-disk angular diameter to the physical limb-darkened disk to which we compare with. Interestingly, 3D models predict less center-to-limb variation than 1D models as moving from K to F type stars (Allende Prieto et al. 2002; Bigot et al. 2006). Reduced limb-darkening corrections imply smaller $\theta_{\rm LD}$: the trend discussed above qualitatively fit into this picture. How well our result agrees quantitatively with this picture we leave to future studies.

Interestingly, Ramírez & Meléndez (2005a) made a similar comparison of angular diameters and also found good agreement with their IRFM $T_{\rm eff}$ scale, which is, however, systematically cooler (by ≃100 K) than the present one for [Fe/H] ≳ −2 (see also Casagrande 2008). We compared the stars with angular diameters in common between Table 4 of Ramírez & Meléndez (2005a, RM05) and the present study (C09, Table 3) and found an average difference (C09-RM05) of 0.1 ± 2.2% in angular diameters, 3.0 ± 3.0% in bolometric fluxes and 40 ± 37 K in $T_{\rm eff}$. Given the large scatter, these numbers are still consistent with the mean differences in $T_{\rm eff}$ and $\mathcal{F}_{\rm Bol}$ from these two studies (Sect. 2.4), however, we would expect our diameters to be roughly smaller by 3%, our fluxes brighter by 1% and our $T_{\rm eff}$ hotter by 100 K (see also Casagrande et al. 2006). While $\mathcal{F}_{\rm Bol}$ and $T_{\rm eff}$ compensate to give almost exactly the same angular diameters, the 40 K offset might be more representative of the difference with the TCS magnitudes used in Ramírez & Meléndez (2005a) (see the discussion on the small zero point differences to convert 2MASS into TCS presented in Sect. 2.4). To gauge further insights, we redetermined the temperatures used by Ramírez & Meléndez (2005a) using their colour calibrations for the same $BV(RI)_{\rm C}$ input data we adopted in this section and found $\Delta T_{\rm eff}$ = 72 ± 52 K. In addition, we adopted our bolometric fluxes lowered by 1%, which corresponds to the average difference we find for our complete sample. In this case the difference in angular diameters sets to −2.4 ± 2.1%, much closer to the expected −3%, offsetting the Ramírez & Meléndez (2005a) scale with respect to interferometric measurements. Since the present work represents an improvement over Ramírez & Meléndez (2005a), in particular the fact that the $T_{\rm eff}, \mathcal{F}_{\rm Bol}, \theta_{\rm IRFM}$ values are a self-consistent and unique solution to each problem star, and given that the number of comparison stars has doubled since 2005 (note also that the $\theta_{\rm LD}$ values of some stars have been re-determined), it is likely that the good agreement found by Ramírez & Meléndez (2005a) was due to a conspiracy of photometric errors which propagated to both $T_{\rm eff}$ and $\mathcal{F}_{\rm Bol}$ determinations and low number statistics. More measurements of stellar angular diameters via interferometry are clearly necessary, and therefore highly encouraged, to better constrain indirectly determined effective temperature scales. However, as this exercise has shown, many critical ingredients enter in the comparison with angular diameters. In particular bolometric corrections and effective temperatures should be determined as self-consistently as possible, also avoiding transformation between photometric systems. It gives us confidence that the zero point uncertainty from solar twins, angular diameters and HST spectrophotometry (next section) returns in all cases independent and very consistent results.

While the angular diameter comparison does not extend below [Fe/H] ≃ −1.0, leaving our results for halo stars "un-tested" in this context, in the next section we use HST spectrophotometry to gauge further insight on the topic.

4.2. HST spectrophotometry

For each star, we obtain a synthetic spectrum tailored at the effective temperature determined via IRFM (Sect. 2.2). Since the

Table 5. Coefficients and range of applicability of the flux calibrations for various $\phi_\zeta = \mathcal{F}_{\rm Bol}({\rm Earth})\, 10^{0.4\, m_\zeta}$.

ϕ_ζ	Colour	[Fe/H] range	Colour range	b_0	b_1	b_2	b_3	b_4	b_5	b_6	N	σ(%)
$B_{\rm T}$	$(B-V)_{\rm T}$	[−2.7, 0.4]	[0.19, 1.43]	2.1904	5.7106	−6.7110	7.4160	−0.6704	−0.1501	−0.0720	260	3.1
$B_{\rm T}$	$V_{\rm T}-J$	[−2.7, 0.4]	[0.51, 2.56]	1.8160	3.2833	−2.3210	1.7358	1.2140	−1.0830	0.0343	261	4.2
$B_{\rm T}$	$V_{\rm T}-H$	[−2.7, 0.4]	[0.53, 3.16]	1.7597	3.1896	−1.8419	0.9465	0.9826	−0.9055	0.0809	255	4.0
$B_{\rm T}$	$V_{\rm T}-K_{\rm S}$	[−2.7, 0.4]	[0.59, 3.29]	1.7202	3.0146	−1.6377	0.8033	0.8591	−0.8644	0.0669	262	3.9
$V_{\rm T}$	$(B-V)_{\rm T}$	[−2.7, 0.4]	[0.19, 1.43]	2.7098	−0.2765	0.1523	0.8122	−0.2261	−0.1789	−0.0413	253	2.7
$V_{\rm T}$	$V_{\rm T}-J$	[−2.7, 0.4]	[0.62, 2.53]	2.1815	0.9268	−0.7701	0.4029	0.1047	−0.2609	−0.0048	249	0.7
$V_{\rm T}$	$V_{\rm T}-H$	[−2.7, 0.4]	[0.68, 3.16]	2.1800	0.8514	−0.5793	0.2235	0.0936	−0.2458	0.0019	261	0.9
$V_{\rm T}$	$V_{\rm T}-K_{\rm S}$	[−2.7, 0.4]	[0.59, 3.29]	2.2565	0.6787	−0.4536	0.1800	0.0785	−0.2407	−0.0011	256	0.9
B	$B-V$	[−5.0, 0.4]	[0.18, 1.22]	1.9571	6.9680	−11.0277	11.4450	−0.4975	−0.1276	−0.0432	331	2.3
B	$V-R_{\rm C}$	[−5.0, 0.3]	[0.24, 0.79]	2.0002	6.6483	−4.6407	25.3881	0.9547	−0.3756	−0.0067	186	1.9
B	$(R-I)_{\rm C}$	[−5.0, 0.3]	[0.23, 0.68]	9.8257	−57.0297	152.2749	−77.6378	4.3253	−1.1377	0.0411	202	3.2
B	$V-I_{\rm C}$	[−5.0, 0.3]	[0.46, 1.47]	4.3948	−6.0713	9.6862	0.2327	0.9298	−0.5392	0.0089	196	1.6
B	$V-J$	[−5.0, 0.4]	[0.50, 2.44]	1.6664	3.5465	−2.5257	1.5310	0.4259	−0.4354	0.0047	332	2.8
B	$V-H$	[−5.0, 0.4]	[0.52, 2.84]	1.6852	3.2925	−1.9206	0.8026	0.2172	−0.1301	0.0346	328	2.9
B	$V-K_{\rm S}$	[−5.0, 0.4]	[0.57, 3.03]	1.5185	3.3566	−1.8830	0.7301	0.2887	−0.2929	0.0240	363	2.8
V	$B-V$	[−5.0, 0.4]	[0.30, 1.03]	1.2581	5.8828	−9.9287	6.8432	0.2290	−0.3935	−0.0420	241	1.9
V	$V-R_{\rm C}$	[−5.0, 0.3]	[0.24, 0.79]	2.6659	−1.6396	3.9243	2.9911	0.0978	−0.2339	−0.0252	177	0.7
V	$(R-I)_{\rm C}$	[−5.0, 0.3]	[0.25, 0.68]	4.9994	−20.1727	49.0418	−27.5918	0.9465	−0.4491	−0.0166	197	1.5
V	$V-I_{\rm C}$	[−5.0, 0.3]	[0.48, 1.47]	3.4468	−3.8760	4.5692	−0.7285	0.1832	−0.2991	−0.0231	184	0.8
V	$V-J$	[−5.0, 0.4]	[0.73, 2.21]	1.8195	1.5562	−1.3322	0.5627	0.1249	−0.3112	−0.0213	314	0.8
V	$V-H$	[−5.0, 0.4]	[0.67, 3.01]	2.0139	1.0845	−0.8071	0.2761	0.0567	−0.2147	−0.0124	369	0.9
V	$V-K_{\rm S}$	[−5.0, 0.4]	[0.93, 3.15]	1.7662	1.4154	−0.9302	0.2726	0.0692	−0.2506	−0.0160	316	0.9
$R_{\rm C}$	$B-V$	[−5.0, 0.3]	[0.35, 1.29]	2.5759	−1.8536	1.3042	0.1015	−0.0130	−0.1229	−0.0142	179	0.8
$R_{\rm C}$	$V-R_{\rm C}$	[−5.0, 0.3]	[0.24, 0.79]	2.7031	−4.2859	6.9274	−2.1959	0.1482	−0.1968	−0.0186	180	0.7
$R_{\rm C}$	$(R-I)_{\rm C}$	[−5.0, 0.3]	[0.23, 0.68]	3.2131	−8.5410	17.3691	−9.1350	0.4602	−0.3054	−0.0171	203	1.0
$R_{\rm C}$	$V-I_{\rm C}$	[−5.0, 0.3]	[0.48, 1.47]	2.9759	−3.2013	2.9454	−0.6516	0.1331	−0.2408	−0.0187	185	0.7
$R_{\rm C}$	$V-J$	[−5.0, 0.3]	[0.86, 2.36]	2.5806	−1.0234	0.4055	0.0107	0.0874	−0.2508	−0.0181	184	0.7
$R_{\rm C}$	$V-H$	[−5.0, 0.3]	[0.93, 2.99]	2.5007	−0.6801	0.1842	0.0176	0.0746	−0.2604	−0.0185	196	0.7
$R_{\rm C}$	$V-K_{\rm S}$	[−5.0, 0.3]	[1.00, 3.13]	2.5606	−0.7448	0.2212	0.0049	0.0665	−0.2509	−0.0184	193	0.7
$I_{\rm C}$	$B-V$	[−5.0, 0.3]	[0.35, 1.29]	2.6765	−3.8643	3.7834	−1.2273	0.0145	−0.0358	−0.0015	200	1.0
$I_{\rm C}$	$V-R_{\rm C}$	[−5.0, 0.3]	[0.24, 0.79]	2.6963	−6.8081	11.3579	−6.1859	0.1798	−0.1418	−0.0112	191	1.1
$I_{\rm C}$	$(R-I)_{\rm C}$	[−5.0, 0.3]	[0.27, 0.68]	3.1500	−10.0132	18.2682	−10.8668	0.3653	−0.2329	−0.0123	177	0.9
$I_{\rm C}$	$V-I_{\rm C}$	[−5.0, 0.3]	[0.48, 1.47]	3.0203	−4.5225	4.0375	−1.1781	0.1371	−0.1866	−0.0122	197	0.8
$I_{\rm C}$	$V-J$	[−5.0, 0.3]	[0.86, 2.36]	2.7912	−2.2548	1.1687	−0.1986	0.0817	−0.1908	−0.0118	186	1.1
$I_{\rm C}$	$V-H$	[−5.0, 0.3]	[0.93, 2.99]	2.7888	−1.8271	0.7688	−0.1061	0.0734	−0.2132	−0.0138	187	1.0
$I_{\rm C}$	$V-K_{\rm S}$	[−5.0, 0.3]	[1.00, 3.13]	2.7797	−1.7014	0.6710	−0.0868	0.0603	−0.1891	−0.0123	193	0.9
J	$B-V$	[−5.0, 0.4]	[0.30, 1.29]	2.2253	−3.5932	2.9303	−0.8741	0.0199	0.0132	0.0057	346	3.4
J	$V-R_{\rm C}$	[−5.0, 0.3]	[0.24, 0.79]	2.5765	−8.1969	12.1713	−6.3037	0.1393	−0.0769	−0.0048	186	2.9
J	$(R-I)_{\rm C}$	[−5.0, 0.3]	[0.27, 0.68]	2.9723	−10.6481	16.3430	−8.5334	0.2971	−0.1854	−0.0095	195	3.0
J	$V-I_{\rm C}$	[−5.0, 0.3]	[0.52, 1.47]	2.8966	−5.1154	4.0119	−1.0879	0.1059	−0.1232	−0.0066	192	2.6
J	$V-J$	[−5.0, 0.4]	[0.82, 2.44]	2.7915	−2.8096	1.2799	−0.2049	0.0479	−0.1059	−0.0054	308	0.8
J	$V-H$	[−5.0, 0.4]	[0.88, 2.99]	2.5885	−2.0262	0.7430	−0.0963	0.0587	−0.1577	−0.0092	303	1.9
J	$V-K_{\rm S}$	[−5.0, 0.4]	[0.93, 3.13]	2.5578	−1.8710	0.6433	−0.0785	0.0457	−0.1326	−0.0078	314	1.7
H	$B-V$	[−5.0, 0.4]	[0.18, 1.29]	2.1337	−3.6473	2.6261	−0.6782	−0.0780	0.1274	0.0179	331	4.6
H	$V-R_{\rm C}$	[−5.0, 0.3]	[0.24, 0.79]	2.5341	−8.8850	12.8801	−6.5281	0.0339	−0.0012	0.0022	184	3.6
H	$(R-I)_{\rm C}$	[−5.0, 0.3]	[0.23, 0.68]	2.9097	−11.1909	16.5901	−8.3344	0.1844	−0.1169	−0.0047	192	3.6
H	$V-I_{\rm C}$	[−5.0, 0.3]	[0.48, 1.47]	2.8833	−5.5447	4.2495	−1.1267	0.0504	−0.0512	−0.0009	195	3.1
H	$V-J$	[−5.0, 0.4]	[0.50, 2.44]	2.5764	−2.6119	1.0580	−0.1490	0.0033	−0.0098	0.0031	353	2.8
H	$V-H$	[−5.0, 0.4]	[0.88, 3.01]	2.4574	−2.0093	0.6768	−0.0808	0.0246	−0.0665	−0.0016	344	1.0
H	$V-K_{\rm S}$	[−5.0, 0.4]	[0.57, 3.15]	2.3732	−1.7778	0.5485	−0.0599	0.0140	−0.0407	0.0005	363	2.2
$K_{\rm S}$	$B-V$	[−5.0, 0.4]	[0.30, 1.29]	2.1537	−3.9640	3.0680	−0.8653	−0.0586	0.1098	0.0163	353	4.9
$K_{\rm S}$	$V-R_{\rm C}$	[−5.0, 0.3]	[0.24, 0.76]	2.5709	−9.5441	14.4103	−7.6430	0.0585	−0.0186	0.0006	190	3.8
$K_{\rm S}$	$(R-I)_{\rm C}$	[−5.0, 0.3]	[0.25, 0.68]	2.8803	−11.4591	17.4060	−9.0419	0.2418	−0.1469	−0.0066	199	3.6
$K_{\rm S}$	$V-I_{\rm C}$	[−5.0, 0.3]	[0.48, 1.47]	2.7928	−5.4377	4.1682	−1.1105	0.0661	−0.0690	−0.0019	201	3.2
$K_{\rm S}$	$V-J$	[−5.0, 0.4]	[0.50, 2.36]	2.6548	−2.8832	1.2411	−0.1878	0.0146	−0.0315	0.0013	328	2.9
$K_{\rm S}$	$V-H$	[−5.0, 0.4]	[0.88, 2.99]	2.4939	−2.1600	0.7593	−0.0946	0.0342	−0.0879	−0.0029	317	2.5
$K_{\rm S}$	$V-K_{\rm S}$	[−5.0, 0.4]	[0.93, 3.03]	2.5097	−2.0732	0.6972	−0.0836	0.0229	−0.0641	−0.0017	328	1.1

Notes. N is the number of stars employed for the fit after the 3 sigma clipping and σ(%) is the standard deviation of the final calibrations in percent. The coefficients of the calibrations b_i are given in units of 10^{-5} erg cm^{-2} s^{-1}.

angular diameter is determined, each synthetic spectrum is absolutely calibrated (i.e. in units of erg cm^{-2} s^{-1} Å^{-1}), and can be used to further test our results. In fact, from F- to early K-type stars, all continuum characteristics approximately longward of the Paschen discontinuity depend almost exclusively on the effective temperature, relatively unaffected by spectral lines and NLTE effects as well as from the treatment of convection.

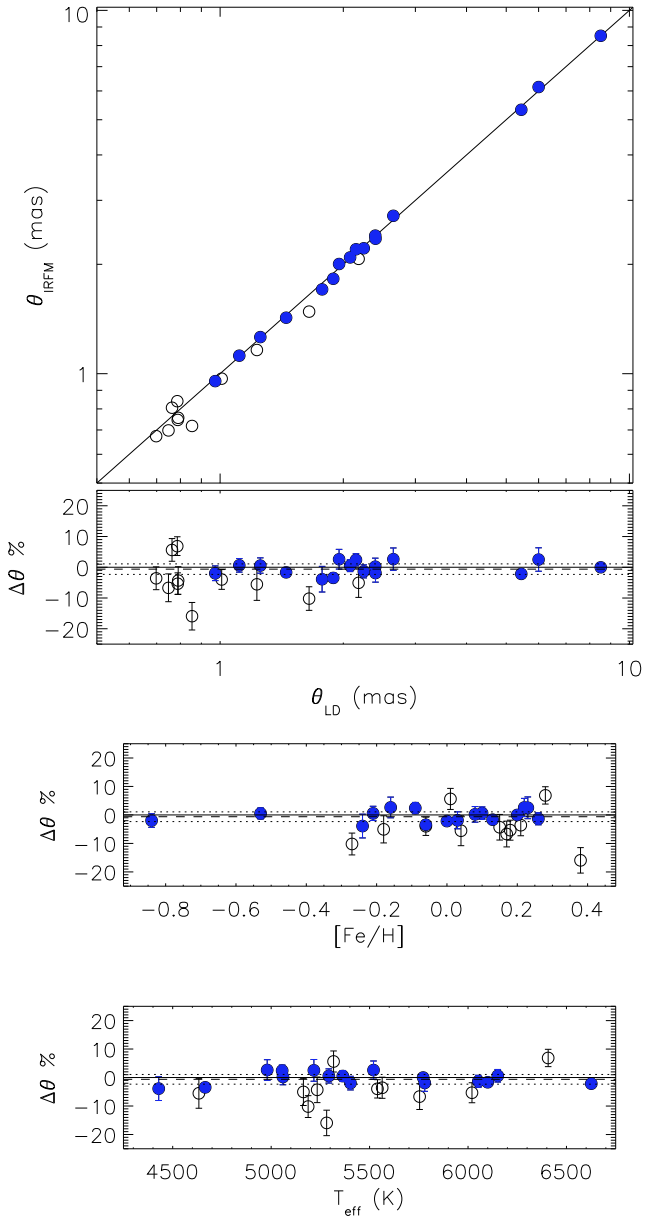

Fig. 8. *Top two panels*: comparison of angular diameters measured interferometrically ($\theta_{\rm LD}$) and via our IRFM photometric calibrations ($\theta_{\rm IRFM}$). Full symbols represent stars that have $\theta_{\rm LD}$ measured with accuracy better than 2%. *Bottom two panels*: difference (in %) between $\theta_{\rm LD}$ and $\theta_{\rm IRFM}$ as a function of stellar parameters. Solid lines represent 1-to-1 correspondence, dashed and dotted lines are the average difference and 1-σ error for the full data point, respectively.

The CALSPEC[7] library contains composite stellar spectra measured by the STIS (0.3−1.0 μm) and NICMOS (1.0−2.5 μm) instruments on board of the HST and used as fundamental flux standard. Free of any atmospheric contamination the HST thus provides the best possible spectrophotometry to date, with 1−2% accuracy, extending from the far-UV to the near infrared. The absolute flux calibration is tied to the three hot, pure hydrogen white dwarfs, which constitute the HST primary calibrators, normalized to the absolute flux of Vega at 5556 Å (Bohlin 2007). Thus, except for the normalization at 5556 Å the absolute fluxes measured by STIS and NICMOS are entirely independent on possible issues regarding Vega's absolute calibration in the infrared and offer an alternative approach to the 2MASS calibration underlying our temperature scale.

Two of the CALSPEC targets are late-type main-sequence dwarfs for which accurate photometry, $\log g$ and [Fe/H] are available: the exoplanet host star HD 209458 (e.g. Charbonneau et al. 2000) and the fundamental SDSS standard BD +17 4708 (e.g. Fukugita et al. 1996; Smith et al. 2002). For each of these targets we computed $T_{\rm eff}$ and derived the corresponding physical flux using the absolute calibration presented in Sect. 3.2. For comparison, we also determined the effective temperatures and the corresponding fluxes when changing our adopted infrared absolute calibration by different amounts up to ±5%, which roughly correspond to ∓100 K in $T_{\rm eff}$. The agreement was quantified using χ^2 statistics between the observed ($\mathcal{F}$) and synthetic ($\tilde{\mathcal{F}}$) spectra at various $T_{\rm eff}$

$$\chi^2 = \sum_{\lambda} \frac{\left(\mathcal{F}_\lambda - \tilde{\mathcal{F}}_\lambda\right)^2}{\sigma_\lambda^2} \qquad (2)$$

where σ_λ^2 is the squared sum of the CALSPEC and our random errors, arising primarily from the photometry and to minor extent [Fe/H] and $\log g$. Angular diameters are needed to scale synthetic spectra into physical units: typical 1% internal accuracy in $\theta_{\rm IRFM}$ implies 2% errors in the derived flux. We decided to use random errors only because the purpose of the test is exactly to verify the range of values allowed once the zero point of the temperature scale is assumed.

Also, the tuning of the absolute calibration in the infrared affects the final $T_{\rm eff}$ but it does not modify in any manner the shape of the synthetic spectrum, which entirely depends on the Castelli & Kurucz (2004) grid interpolated at the proper $T_{\rm eff}$, $\log g$ and [Fe/H]. Notice that we are not searching for the synthetic spectrum which best matches the observation, rather we want to test the effective temperature we derive: while adjustments to [Fe/H] and $\log g$ could improve the agreement in the blue and visible part, the continuum characteristics are more sensitive to $T_{\rm eff}$.

4.2.1. HD 209458

For this target we adopted the spectroscopic [Fe/H] = 0.03±0.02 and $\log g$ = 4.50 ± 0.04 measured from the high precision HARPS GTO sample (Sousa et al. 2008) and used Tycho2 and 2MASS photometry. We obtain $T_{\rm eff}$ = 6113 ± 49 K, $\mathcal{F}_{\rm Bol}$ = $(2.335 \pm 0.025) \times 10^{-8}$ erg cm^{-2} s^{-1} and θ = 0.224 ± 0.004 mas including both random and systematic errors. The latter result is in good agreement with the angular diameters 0.215 ± 0.009 mas obtained using the new *Hipparcos* parallaxes (van Leeuwen 2007) to convert the linear radius measured from exoplanet transit photometry with HST (Brown et al. 2001). Notice that ∼100 K cooler effective temperatures would imply values of θ larger by ∼3.5% in the IRFM.

The comparison between the observed and synthetic spectra at two different $T_{\rm eff}$ is shown in Fig. 9: while they both succeed to capture the main observed features, the continuum of the cooler model is clearly off from the observation. We quantify the agreement between the HST spectrophotometry and the models at various $T_{\rm eff}$ applying χ^2 statistics longward of the Hα line (0.66 μm), the Paschen (0.82 μm) and the Brackett (1.46 μm) discontinuity. These cuts define the beginning of the continuum in a somewhat arbitrary manner, but they all return consistent results thus ensuring that our conclusion is not affected by their choice.

[7] http://www.stsci.edu/hst/observatory/cdbs/calspec.html as of January 2009.

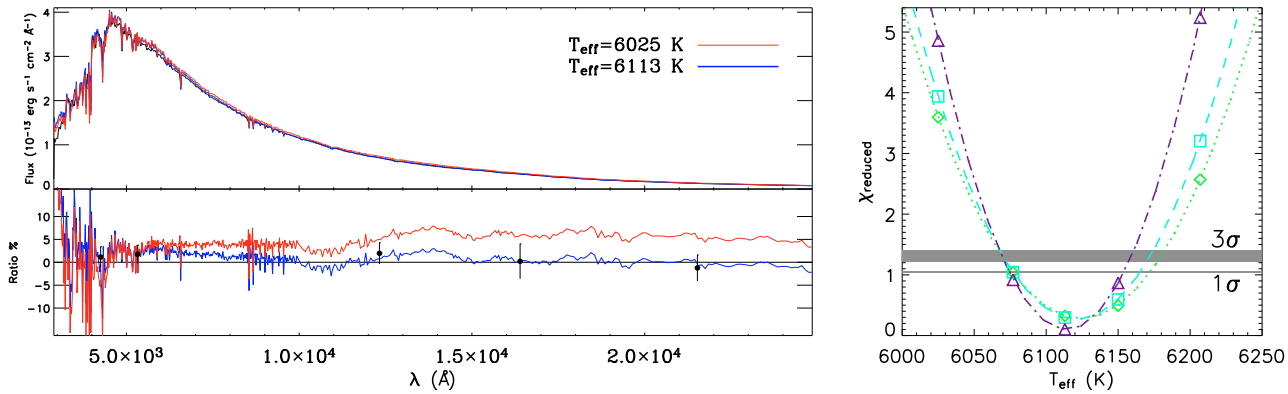

Fig. 9. *Left upper panel*: comparison between the observed HD 209458 CALSPEC spectrum (black line) and the synthetic spectra derived for two different $T_{\rm eff}$, using our preferred absolute calibration (blue line) and increasing the infrared absolute calibration by 5% (red line). *Left lower panel*: ratio of synthetic to observed spectra. Full circles are the ratio between the fluxes obtained once the Vega calibration is used with the observed magnitudes and the fluxes obtained directly from the convolution of the CALSPEC spectrum with the appropriate filter transmission curve. Error bars take into account uncertainty in the Vega calibration and zero points, as well as in the observed magnitudes. *Right panel*: reduced χ^2 for various $T_{\rm eff}$ solutions corresponding to different adopted absolute calibrations. Our choice (Sect. 3.2) always lies very close to the minima obtained fitting a parabola to the data (lines of different style). Different symbols correspond to cut longward of 0.66 μm (diamonds), 0.82 μm (squares) and 1.46 μm (triangles) as explained in the text. The sigma levels have been computed using the incomplete gamma function for the number of degrees of freedom longward of our cuts.

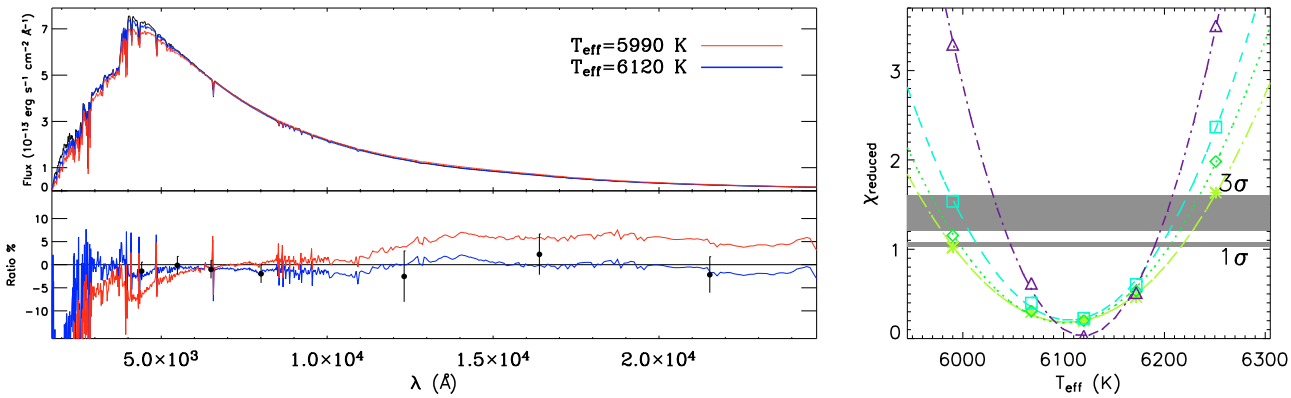

Fig. 10. Same as in Fig. 9 for BD +17 4708. The synthetic spectra have been reddened by $E(B-V) = 0.01$. Different symbols in the right panel correspond to cut longward of 0.50 μm (asterisks) 0.66 μm (diamonds), 0.82 μm (squares) and 1.46 μm (triangles). The maximum wavelength used for computing the reduced χ^2 has been 2 μm to avoid possible contribution from the cool companion.

The reduced χ^2 is lower than 1 in a roughly ±40 K interval effectively centered on our preferred solution. While reduced $\chi^2 < 1$ tells that the size of the errors is still too large to clearly favour a solution within that range, the large number of points used in the test sets low 1σ and 3σ levels, clearly ruling out solutions different by ±100 K.

4.2.2. BD +17 4708

This star is the only subdwarf with well measured absolute flux, thus making it an important benchmark for testing the temperature scale in the metal-poor regime. We adopt the spectroscopic parameters [Fe/H] = −1.74 ± 0.09, [α/Fe] = 0.4 and $\log g$ = 3.87 ± 0.08 from Ramírez et al. (2006) who also derived $T_{\rm eff}$ = 6141 ± 50 K, $\mathcal{F}_{\rm Bol} = (4.89 \pm 0.10) \times 10^{-9}$ erg cm^{-2} s^{-1} and θ = 0.1016±0.0023 mas. We corrected for reddening $E(B-V) = 0.01$ the optical (Table 1) and infrared (2MASS) magnitudes, obtaining $T_{\rm eff} = 6120 \pm 112$ K, $\mathcal{F}_{\rm Bol} = (4.80 \pm 0.04) \times 10^{-9}$ erg cm^{-2} s^{-1} and $\theta = 0.101 \pm 0.003$ all in excellent agreement with the aforementioned analysis. Radial velocities show modulation consistent with the presence of a low mass companion which could influence infrared photometry (Latham et al. 1988). The flags associated with 2MASS indicate excellent quality and no artifact nor contamination in any band, pointing toward a negligible effect, if any. Nonetheless, since the percent contribution of a cool companion increases with increasing wavelength, as safety rule we decided not to use $K_{\rm S}$ in the IRFM though it would change the resulting $T_{\rm eff}$ by only 12 K. For our preferred $T_{\rm eff}$ = 6120 K, shortward of 2 μm there is an outstanding agreement with the CALSPEC observed spectrum, meaning that the solution found represents well the observation at all wavelengths. A moderate increase in the observed with respect to the synthetic flux seems to appear longward of 2 μm, which could be the signature of the cooler companion. On the contrary, cooler solutions overestimate the flux throughout the entire continuum.

Because of the metal-poor nature of this star, the continuum shows up already at bluer wavelengths. We compute the reduced χ^2 in different intervals, starting longward of 0.50 μm: as for the previous star, our solution substantially correspond to the minima of all parabolae, independently of the cut adopted. The random errors associated with this star are larger than in the case of HD 209458, giving shallower minima and thus making it more difficult to discriminate between different solutions. However, differences up to ±100 K are clearly disfavoured (Fig. 10).

Summarizing, CALSPEC data support our temperature scale which provide the best match to the observed spectrophotometry, in both metal-rich and -poor regimes. While differences larger than ±40 K are ruled out for HD 209458, the observational errors for the metal-poor star allow bigger uncertainties. Nonetheless, we have determined the fundamental parameters of both stars with the same procedure and in both cases our solutions are located at the minimum χ^2: we regard such a result as a further indication that our $T_{\rm eff}$ scale is well calibrated over a wide metallicity range.

5. The new effective temperature scale

Our results should be compared with effective temperatures determined employing different methods. First, we focus on large studies which have targeted solar neighbourhood stars, where the vast number of objects imposes the use of fast and efficient techniques, relying on fitting the observed photometry or spectra to their synthetic counterpart. An extensive comparison between the effective temperatures determined from high resolution spectroscopy of solar neighbourhood stars and a version of the IRFM similar to that adopted here has been already carried out in Sousa et al. (2008). For metal-poor stars we restrict the comparison to purely spectroscopic effective temperatures; their validation will be crucial for ongoing and future studies of halo stars which are strongly affected by reddening and often lacking photometry.

5.1. Solar neighbourhood stars

5.1.1. Valenti & Fischer sample

Valenti & Fischer (2005) have presented a uniform catalogue of stellar properties for 1040 nearby F, G and K stars which have been observed by the Keck, Lick and AAT planet search programs. Fitting the observed spectra with synthetic ones, they have obtained effective temperatures, surface gravities and abundances for every star. For 84 objects in common, there is no obvious dependence as a function of $T_{\rm eff}$, except for a drift appearing below 5000 K. However, when $\Delta T_{\rm eff}$ is plotted as function of metallicity the trend becomes clear, with very significant discrepancies at the lowest metallicities (Fig. 11).

5.1.2. Masana et al. sample

Masana et al. (2006) have derived stellar effective temperatures and bolometric corrections by fitting V and 2MASS IR photometry. They calibrate their scale by requiring a set of 50 solar analogs drawn from Cayrel de Strobel (1996) to have on average the same temperature as the Sun.

We have 176 stars in common: there is no obvious trend with effective temperatures, and for metallicities around solar there is an overall good agreement. This is not entirely unexpected considering that both studies have been calibrated to the Sun (though with different approaches): considering [Fe/H] > −1 the mean difference (IRFM − Masana) is $\Delta T_{\rm eff} = -21 \pm 6$ K ($\sigma = 71$ K). However, when focusing on metal-poor stars [Fe/H] < −1 there is a significantly increasing scatter and a trend resulting in our $T_{\rm eff}$ being cooler up to ∼200 K at the lowest metallicities and with a mean difference of -95 ± 22 K ($\sigma = 157$ K).

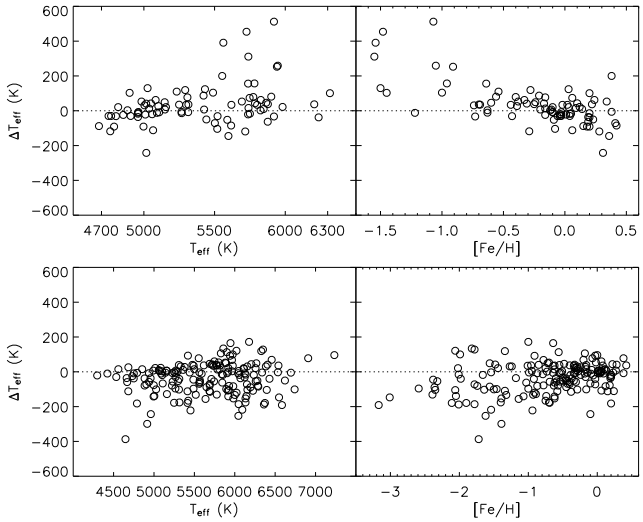

Fig. 11. *Upper (lower) panels*: comparison between the effective temperatures determined in this work and those obtained by Valenti & Fischer (2005) (Masana et al. 2006). $\Delta T_{\rm eff}$ are this − other works in all panels.

5.2. Metal-poor, halo stars

5.2.1. Temperatures from fits to hydrogen line profiles

The wings of hydrogen lines are strongly sensitive to the effective temperature of the star and only mildly dependent on the other stellar parameters, other than being unaffected by reddening. Such approach is particularly effective with metal-poor stars, given the lack of severe line blending affecting the hydrogen lines. Thus, provided a proper continuum normalization is applied, which can be non-trivial in some cases (e.g. Barklem et al. 2002), these lines can be used to determine $T_{\rm eff}$. Although significant progress has been made in the last few years, the modeling of hydrogen lines (e.g., the Balmer line profiles) is still quite uncertain (Barklem et al. 2000; Barklem 2007). Nonetheless, the relative $T_{\rm eff}$ values derived in this manner can be very precise (e.g. Nissen et al. 2007).

We remark that there is no such thing as one Balmer line $T_{\rm eff}$ scale, but instead each study depends upon the adopted prescriptions: LTE vs. NLTE, broadening recipes, mixing-length parameter and even the details on how lines are fitted. Also, the thermal structure of the model atmosphere is crucial for the Balmer temperatures: as concerns 1D models, studies relying on OS- instead of ODF-model atmosphere determine hotter $T_{\rm eff}$ (Grupp 2004).

Aware of the complexity of the picture, in the upper panels of Fig. 12 our IRFM effective temperatures are compared with those derived from fits to the Balmer lines in two different studies, which we regard as representative of the LTE and NLTE approach, respectively. Circles refer to the comparison with Fabbian et al. (2009) who used the Hβ lines. There is an obvious offset, the IRFM returning $T_{\rm eff}$ hotter by 84 ± 13 K ($\sigma = 66$ K), but the small scatter between these two sets further strengthen the conclusion that both techniques have high internal precision. A similar conclusion holds also from the comparison with the effective temperatures reported in Bergemann (2008, and references therein) who used both Hα and Hβ line profiles. In this case the difference (IRFM − H lines) is 21 ± 23 K ($\sigma = 72$ K) with a possible trend suggesting excellent agreement roughly below 6000 K (one star, HD 25329 with $T_{\rm eff} = 4785$ K and $\Delta T_{\rm eff} = -15$ K is not shown in the upper left panel of Fig. 12).

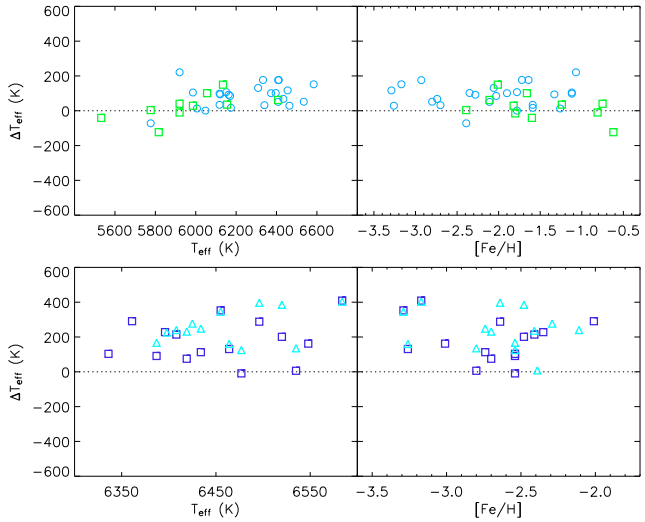

Fig. 12. *Upper panels*: comparison between the effective temperatures determined in this work and those obtained from the Hβ (Fabbian et al. 2009, circles) and Hα plus Hβ (Bergemann 2008, squares) line profiles. *Lower panel*: comparison with respect to the excitation equilibrium temperatures determined by Hosford et al. (2009). Two sets of data points are shown because Hosford et al. (2009) temperatures are sensitive to the uncertain $\log g$ values of metal-poor stars; squares (triangles) represents $T_{\rm eff}$ derived assuming the star to be on the main-sequence (sub-giant branch). $\Delta T_{\rm eff}$ are this − other works in all panels.

5.2.2. Excitation equilibrium temperatures

An important number of iron lines are present in the spectra of cool dwarfs, even the metal-poor ones. In an ideal case, the iron abundances determined from each of those lines should be consistent with each other. In practice, however, given an initial set of stellar parameters, the line-by-line abundances show trends with excitation potential (EP) and/or reduced equivalent width. By tuning the stellar parameters, these trends can be eliminated. The EP trend is particularly sensitive to $T_{\rm eff}$, given the strong dependence of the atomic level populations on temperature, and therefore $T_{\rm eff}$ determined by removing the abundance vs. EP trend are often referred to as "excitation equilibrium" temperatures. Because of its nature, this method of $T_{\rm eff}$ determination is highly model-dependent. Not only it does require realistic model atmospheres and spectrum synthesis, but also accurate atomic data and, ideally, a non-LTE treatment of the line formation. The advantage of such method is that it is independent of interstellar reddening and can be applied to stars with uncertain or unavailable photometry.

Recently, Hosford et al. (2009) have determined LTE excitation equilibrium temperatures for a sample of metal-poor stars. The difference found between their temperatures and ours is illustrated in Fig. 12 (HD 140283 with $E(B-V) = 0.000$, $T_{\rm eff} = 5777$ K and $\Delta T_{\rm eff} = 8$ K is not shown in the lower left panel). Because the excitation temperatures are somewhat sensitive to $\log g$ and surface gravities of metal-poor stars are difficult to determine due to uncertain/unavailable parallaxes, they provide two sets of $T_{\rm eff}$ values, one assuming the star to be on the main-sequence (MS) and another one assuming the star to be on the subgiant branch (SGB). We remark that for HD 140283 parallax and Balmer jump rule out the main-sequence stage; our fit (Mike Bessell) of the MILES fluxes using Munari et al. (2005) spectral library provide $T_{\rm eff} = 5812/5875$ K and $\log g = 3.75$ for $E(B-V) = 0.000/0.017$, respectively.

The IRFM temperatures are significantly hotter than the excitation temperatures by 177 ± 33 K ($\sigma = 122$ K) (for their MS temperatures) and 240 ± 32 K ($\sigma = 116$ K) (SGB). In particular, the large scatter suggests a decreased relative precision when applying excitation equilibrium to very metal-poor stars, so that the further investigation of non-LTE effects will be highly desirable (Hosford et al. in prep.).

5.3. The most metal-poor stars in the Galaxy

Despite theoretical uncertainties on the exact mass range under which the first stars formed, it is likely that the most metal-poor objects currently observed in the Milky Way halo are second generation stars. In case of dwarfs/subgiants, their abundance patterns carry direct information on the first stars ever formed in the Galaxy (e.g. Frebel et al. 2005) and/or on still poorly known long time-scale processes which might take place below the surface or deep into stellar interior (e.g. Venn & Lambert 2008; Korn et al. 2009).

Determining their effective temperature and evolutionary status (i.e. $\log g$) is crucial to derive reliable abundances and constrain different scenarios. At the same time, such a quest is in stark contrast with the many practical limitations associated with hyper-metal-poor stars: parallaxes are not available to help constrain their surface gravities and even when spectra with sufficient resolution and S/N are obtained, the model atmospheres used for the analysis are not yet fully tested at such low metallicities. Rigorous analyses should also take into account 3D (Frebel et al. 2008) and NLTE (Aoki et al. 2006) effects, which are expected to be considerable in this regime. Determining $T_{\rm eff}$ in a way mostly unaffected by the above limitations is not only desirable, but also necessary to put spectroscopic analyses on firmer grounds.

5.3.1. HE1327-2326

For this star the IRFM returns $T_{\rm eff} = 6250 \pm 60$ K in agreement within the errors with the spectroscopic value of 6120 ± 150 K obtained from the NLTE analysis of the Balmer lines (Korn et al. 2009), roughly with an offset of the same order of that discussed in Sect. 5.2.1. As we already pointed out, the IRFM depends only weakly on the adopted surface gravity: changing it by ±0.5 dex affects $T_{\rm eff}$ by approximately ±25 K. In our case, we used $\log g = 3.7$ as recently determined by Korn et al. (2009). The exact metallicity of HE1327-2326 is also uncertain: although it is well established that its [Fe/H] < -5.0, estimates range from −5.9 to −5.4 depending on the adopted stellar parameters and 1D/3D LTE/NLTE analysis performed (Aoki et al. 2006; Frebel et al. 2008). The IRFM is known to depend very little on the metallicity and we verified this being particularly true (at least in this $T_{\rm eff}$ regime) for the featureless spectra of this hyper-metal-poor star: increasing [Fe/H] by 1 dex in the IRFM affects the derived $T_{\rm eff}$ by less then 10 K. This conclusion supports the suggestion that for hyper-metal-poor stars colour−temperature calibration of normal very-metal-poor stars can be used instead (see discussion in Sect. 6).

When running the IRFM for this star we used the new grid of MARCS model atmosphere (Gustafsson et al. 2008) which extend down to [Fe/H] = −5.0 and this value was used in our implementation. Because of the weak metallicity dependence discussed above, very similar results are obtained if the Castelli & Kurucz (2004) grid (which stops to [M/H] = −4.0) is used instead. For the sake of ensuring our results do not depend too

much on the adopted spectra library, we also checked that for stars with higher metallicities MARCS or ATLAS9 models return very similar results, with differences usually well within 10 K and at most of order 20 K (see also Casagrande et al. 2006).

We feel the major source of possible systematic error stems from reddening, which is very high for this star. We used $E(B-V) = 0.076$ based on both extinction maps and interstellar absorption lines (Aoki et al. 2006; Beers et al. 2007) but it should be kept in mind that a change of ±0.01 mag in $E(B-V)$ affects $T_{\rm eff}$ by ±50 K.

5.3.2. HE0233-0343

Though the exact metallicity of this star is still uncertain, it seems well secured as having [Fe/H] $\lesssim -4.0$ (García Pérez et al. 2008, García Pérez private communication). Its evolutionary status is also ambiguous, with spectroscopic estimates of $\log g$ varying from 3.5 to 4.5. Also in this case, the exact values of $\log g$ and [Fe/H] are not crucial for the IRFM and we checked that changing them even considerably affects $T_{\rm eff}$ by an amount similar to that discussed for HE1327-2326. We adopt [Fe/H] $= -4.0$ and $\log g = 4.0$ from which we derive $T_{\rm eff} = 6270 \pm 80$ K, without accounting for possible systematics arising from $E(B-V) = 0.025$ (Beers et al. 2007). As we point out in Sect. 6 there might be some issue with the $R_{\rm C}$ photometry for this star. Were we to exclude this band when running the IRFM, $T_{\rm eff}$ would increase by 25−35 K depending on the surface gravity assumed. Spectroscopic $T_{\rm eff}$ estimates for this star are still uncertain, primarily because of its uncertain $\log g$. Were its subgiant status to be confirmed, our effective temperature would be in good agreement with the spectroscopic one (García Pérez et al. 2008).

6. Empirical calibrations

The effective temperatures and the bolometric luminosities derived via IRFM for our sample allow us to build calibrations relating those quantities to the measured colours and metallicities. As discussed in Sect. 2, to correctly account for reddening is crucial though fortunately, for the sake of deriving colour relations, reddening affects both the observed photometry and the derived fundamental stellar parameters, thus making such relations – built using dereddened colours – independent on the adopted $E(B-V)$ in first approximation.

In the following we give the functional form of these calibrations, together with the number of stars used, the standard deviation obtained in the fitting process and the range of applicability. The results presented here usually match Casagrande et al. (2006) within the limits of those calibrations, but extend over a wider range now and thus supersede the previous work. Though our sample has been assembled explicitly to cover a parameter space as large as possible in effective temperature and metallicity, the detection and observation of stars with [Fe/H] $\lesssim -2.5$ is still strongly biased around $T_{\rm eff} \sim 6500$ K. Even if the formal range of applicability of the calibrations extend well below [Fe/H] < -3, the number of known metal-poor stars considerably decreases as one moves away from the aforementioned $T_{\rm eff}$ (see Fig. 13). In particular, for metallicities below −4, only two stars are currently known, a number clearly inadequate to give fits. Fortunately, at these temperatures, calibrations at about −3.5 seem adequate for even more metal-poor stars, as we discuss further in Sects. 6.1 and 6.2. Nonetheless, we advocate particular caution when using these calibrations in poorly sampled regions of Fig. 13. On the contrary for [Fe/H] $\gtrsim -2$, typical for most of the stellar population observed in the solar neighbourhood and Galactic star clusters, our calibrations are robust and can be readily used for a number of purposes.

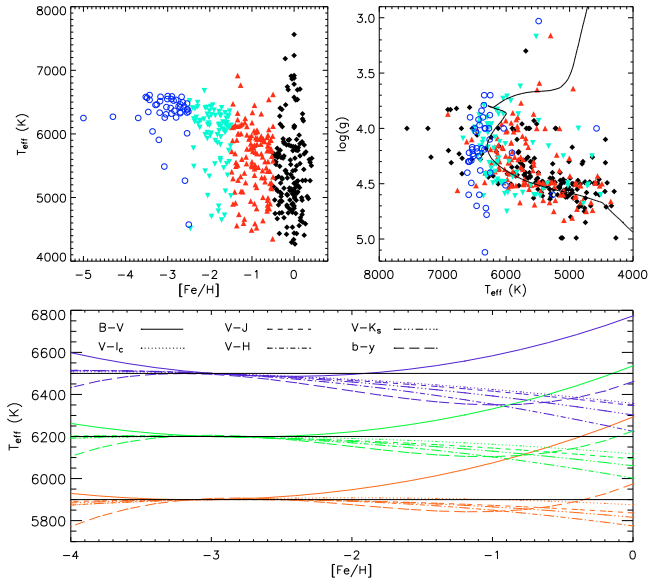

Fig. 13. *Upper left panel*: metallicities and effective temperatures of our sample. All stars have 2MASS and Johnson-Cousins photometry. *Upper right panel*: effective temperatures and gravities of our sample. Symbols for different metallicity bins are the same as in the *left panel*. Overplotted for reference is a 3 Gyr solar isochrone from Bertelli et al. (2008). *Lower panel*: metallicity sensitivity of our colour−temperature calibration in different bands for stars having $T_{\rm eff} = 6500$ K (*top*), 6200 K (*middle*) and 5900 K (*lower*) at [Fe/H] = −3.0.

The core of the present work is to accurately define the zero point of the temperature scale in many standard photometric systems; we caution however that in some cases real systems might not exactly reproduce standard systems, especially in the case of the faintest sources (Bessell 2005). Users of our calibrations should always keep this in mind: although the zero point of the $T_{\rm eff}$ scale is now well defined, in gathering photometry from heterogeneous sources there might be small zero point issues between different authors, and this observational uncertainty – if present – will introduce small systematic errors to our accurate empirical calibrations.

6.1. Colour-temperature-metallicity

To reproduce the observed $T_{\rm eff}$ versus colour relation and take into account the effects of metallicity, the usual fitting formula has been adopted (e.g. Alonso et al. 1996b; Ramírez & Meléndez 2005b; Casagrande et al. 2006; González Hernández & Bonifacio 2009)

$$\theta_{\rm eff} = a_0 + a_1 X + a_2 X^2 + a_3 X[{\rm Fe/H}] + a_4[{\rm Fe/H}] + a_5[{\rm Fe/H}]^2 \quad (3)$$

where $\theta_{\rm eff} = 5040/T_{\rm eff}$, X represents the colour and $a_i\,(i = 0,\ldots,5)$ are the coefficients of the fit obtained iteratively, discarding points departing more than 3σ.

The IRFM depends only very mildly on the adopted $\log g$ (Sect. 2.2) but certain colours could be more affected: for all indices we have checked the residual of our calibration and did not find any obvious trend with $\log g$. Nevertheless, a dependence on the gravity could be built into the calibrations, since $\log g$ decreases as one moves from cool dwarfs to hotter turn-off stars (Fig. 13).

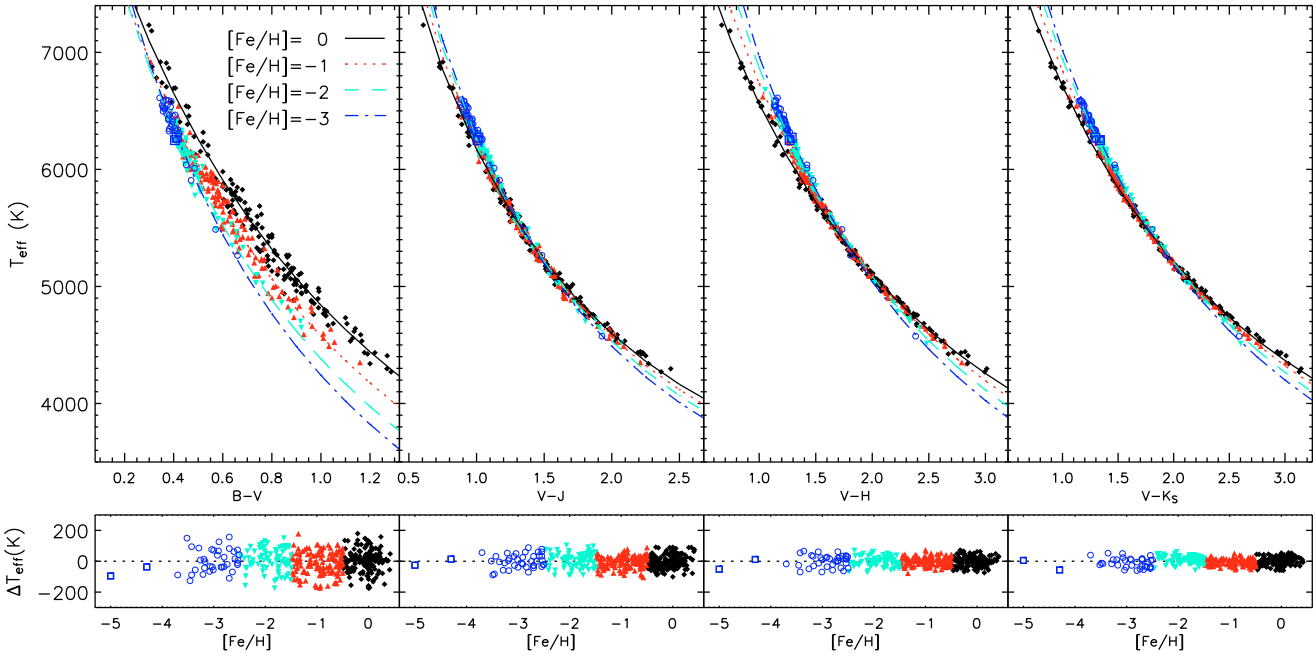

Fig. 14. *Upper panels*: empirical colour-temperature-metallicity calibrations in the metallicity bins $-0.5 < [Fe/H] \leq 0.5$ (filled diamonds), $-1.5 < [Fe/H] \leq -0.5$ (upward triangles), $-2.5 < [Fe/H] \leq -1.5$ (downward triangles) and $[Fe/H] \leq -2.5$ (open circles). Open squares are for the hyper metal-poor stars HE0233-0343 and HE1327-2326. *Lower panels*: residual of the fit as function of metallicity. For the two hyper-metal-poor stars, the residual is with respect to the fit at $[Fe/H] = -3.5$.

The coefficients for various colour indices are given with their range of applicability in Table 4 and a comparison between the polynomial fits and our sample of stars is shown in Fig. 14. We remark that the functional form of Eq. (3) may return non-physical values when extrapolated to very low metallicities, as extensively discussed by Ryan et al. (1999) for the calibration of Alonso et al. (1996b) below $[Fe/H] \sim -2.5$. We have considerably increased the number of very metal-poor (turnoff) stars and our calibration behaves as one would expect, i.e. it shows a decreasing sensitivity on [Fe/H] when moving from −2 to −3, where the metallicity sensitivity vanishes in all bands (Fig. 13). Moving to $[Fe/H] = -4$ (or lower), the diverging behaviour in Fig. 13 reflects the form of the fitting function and the values of the coefficients rather than the characteristics of metal-poor turnoff stars. In Fig. 14 the two hyper metal-poor stars (represented by open squares) clearly follow the same trend of other iron deficient stars with similar effective temperatures. Using Eq. (3) at a fixed $[Fe/H] = -3.5$ recovers their IRFM $T_{\rm eff}$ within the typical accuracy of the calibration. This is always true for HE1327-2326, and also for HE0233-0343 except when using the $R_{\rm C}$ index, possibly indicating a photometric issue in this band for the latter star. This comparison thus warrants the applicability of our calibrations for hyper-metal-poor stars if $[Fe/H] = -3.5$ is assumed and a typical $T_{\rm eff} \sim 6200$ K is obtained. How well this holds at other effective temperatures is still unknown.

The calibration presented here applies till late K-type dwarfs. Those interested in M dwarfs, can instead refer to Casagrande et al. (2008): though in that work the zero point has not been constrained using solar twins, the absolute calibration adopted was similar to that used here, resulting in effective temperatures approximately on the same scale. Nonetheless, if a link between the two scales is needed, we advise users to a careful case-by-case study, also considering that the calibration for M dwarfs has a different functional form and does not include any metallicity term.

6.1.1. Strömgren calibration

The Strömgren index $b - y$ deserves a separate discussion. It is often used as a $T_{\rm eff}$ indicator, but because of its very nature has a strong sensitivity on the metallicity and a proper functional form is not trivial. Alonso et al. (1996b) excluded the coolest dwarfs, where the dependence of $b - y$ upon $T_{\rm eff}$ possibly flattens out. Yet, for the most metal poor stars that calibration diverges to unphysical values, as discussed in Ryan et al. (1999).

For $b - y$ we have verified that a calibration of the form of Eq. (3) has strong residuals as function of both colour and metallicity and used polynomial fits to correct such trends, i.e. $T_{\rm eff} = 5040/\theta_{\rm eff} + P([Fe/H], b - y)$. To this purpose, we have increased the sample with more than 1000 stars from the GCS catalogue (Nordström et al. 2004) all having Strömgren photometry, spectroscopic metallicities from an updated version of the Cayrel catalogue (Meléndez, in prep.) and for which the IRFM could be applied directly using Tycho2 and 2MASS (Casagrande et al. in prep.).

We checked that a third order polynomial in both colour and metallicity was enough; the calibration before and after adopting such a correction is shown in Fig. 15 and the coefficients, given in the form $P([Fe/H], b - y) = \sum_{i=0}^{3} M_i [Fe/H]^i + \sum_{i=0}^{3} C_i (b - y)^i$ are $M_0 = -1.9$, $M_1 = 130.4$, $M_2 = 125.7$, $M_3 = 27.4$, $C_0 = -1003.7$, $C_1 = 7325.9$, $C_2 = -17207.4$, $C_3 = 12977.7$. Notice that the form of these corrections can lead to unphysical values if extrapolated and should never be applied outside of the colour and [Fe/H] ranges of Fig. 15.

6.2. Colour-flux-metallicity

We adopt the same definition of Casagrande et al. (2006) to define the bolometric correction in a given ζ band, where

$$BC_\zeta = m_{\rm Bol} - m_\zeta \quad (4)$$

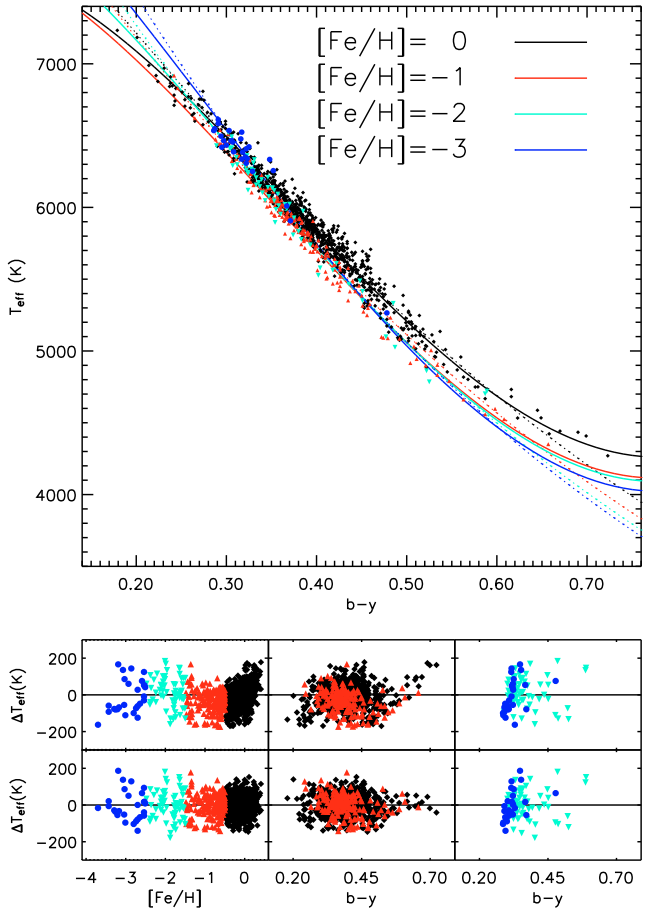

Fig. 15. *Upper panel*: empirical colour-temperature-metallicity calibration in $b-y$ before (dotted) and after (continuous lines) the polynomial correction. *Central and lower panels*: residuals before and after the polynomial corrections.

and the zero point of the $m_{\rm bol}$ scale is fixed by choosing $M_{\rm Bol,\odot} = 4.74$. Empirical bolometric corrections in various bands can thus be readily computed using Eq. (4) and dereddening the observed magnitudes given in Table 8.

A complementary way of deriving stellar integrated flux via photometric indices is given in the form of Casagrande et al. (2006), using the coefficients given in Table 5

$$\mathcal{F}_{\rm Bol}({\rm Earth}) = 10^{-0.4\,m_\zeta}\Big(b_0 + b_1 X + b_2 X^2 + b_3 X^3 + b_4 X[{\rm Fe/H}] + b_5[{\rm Fe/H}] + b_6[{\rm Fe/H}]^2\Big). \quad (5)$$

As for the temperature calibrations, also in this case the fluxes of the two hyper metal-poor stars can be recovered adopting [Fe/H] $= -3.5$ in Eq. (5), though we caution that the license of this approach for considerably bluer or redder indices is still unknown.

6.3. Colour-angular diameters

Limb-darkened angular diameters can be readily derived from the basic definition involving effective temperatures and bolometric fluxes, using the calibrations given in Sects. 6.1 and 6.2. Nonetheless, very tight and simple relations exist in the J band and in Table 6 we give them in the form of Casagrande et al. (2006)

$$\theta = c_0 + c_1\sqrt{\phi(m_J, X)} \quad (6)$$

where

$$\phi(m_J, X) = 10^{-0.4\,m_J} X \quad (7)$$

for a given colour index X. These relations show remarkably small scatter and no metallicity dependence, thus proving ideal to build a network of small calibrators for interferometric measurements, for characterizing extrasolar planet transits or microlensing events.

Table 6. Coefficients and range of applicability of the angular diameter calibrations.

Colour	Colour range	c_0	c_1	N	σ(%)
$V-J$	[0.73, 2.44]	0.00015	4.65293	394	2.2
$V-H$	[0.88, 3.01]	0.00004	4.15613	389	1.8
$V-K_{\rm S}$	[0.93, 3.15]	0.00020	4.05037	394	1.9
$V_{\rm T}-J$	[0.72, 2.56]	0.00218	4.44568	270	2.1
$V_{\rm T}-H$	[0.87, 3.14]	0.00227	3.98945	255	1.5
$V_{\rm T}-K_{\rm S}$	[0.92, 3.27]	0.00286	3.88433	268	1.5

Notes. N is the number of stars employed for the fit after the 3σ clipping and σ(%) is the standard deviation of the calibrations.

6.4. The colours of the Sun

The interpolation of Eq. (3) at $T_{\rm eff} = 5777$ K and [Fe/H] $= 0$ returns the colours of the Sun, which are given in Table 7. For the Tycho2 and 2MASS system, those can be readily compared with the averaged ones from the twins of Sect. 3.1: not unexpectedly there is good agreement, all but one within few millimag, which usually (at maximum) correspond to few (20) K in $T_{\rm eff}$. We have also checked that fitting our twins as function of $T_{\rm eff}$ and [Fe/H] returns colours almost identical to their average, further confirming that our sample of twins is homogeneously distributed in temperature and metallicity around the colours of the Sun inferred from our scale.

In recent years, there has been considerable work in order to determine the colours of the Sun (e.g. Sekiguchi & Fukugita 2000; Ramírez & Meléndez 2005b; Holmberg et al. 2006; Pasquini et al. 2008). One of the most extensive analysis is that of Holmberg et al. (2006): the remarkably good agreement we have in the optical colours can be understood from the dependence of these indices on both $T_{\rm eff}$ and [Fe/H]. The approximately 100 K cooler effective temperature scale adopted by Holmberg et al. (2006) favours bluer colours, which are grossly compensated to the red by the underestimation of ∼0.1 dex in the GCS photometric metallicities with respect to spectroscopic ones selected to be consistent with our temperature scale (Holmberg et al. 2009). Our $B-V = 0.641$ is also in very good agreement with the $B-V = 0.649 \pm 0.016$ found studying solar twins in M 67 (Pasquini et al. 2008). For this cluster, using our colour-temperature relation to compare $V-K_{\rm S}$ photometry with theoretical isochrones shows remarkably good agreement (Vandenberg, private communication).

Infrared indices derived inverting Eq. (3) depend almost exclusively on the adopted $T_{\rm eff}$ scale, which is responsible for our much redder colours than those of Holmberg et al. (2006). Our $V-J$, H and $K_{\rm S}$ are in good agreement with those reported in

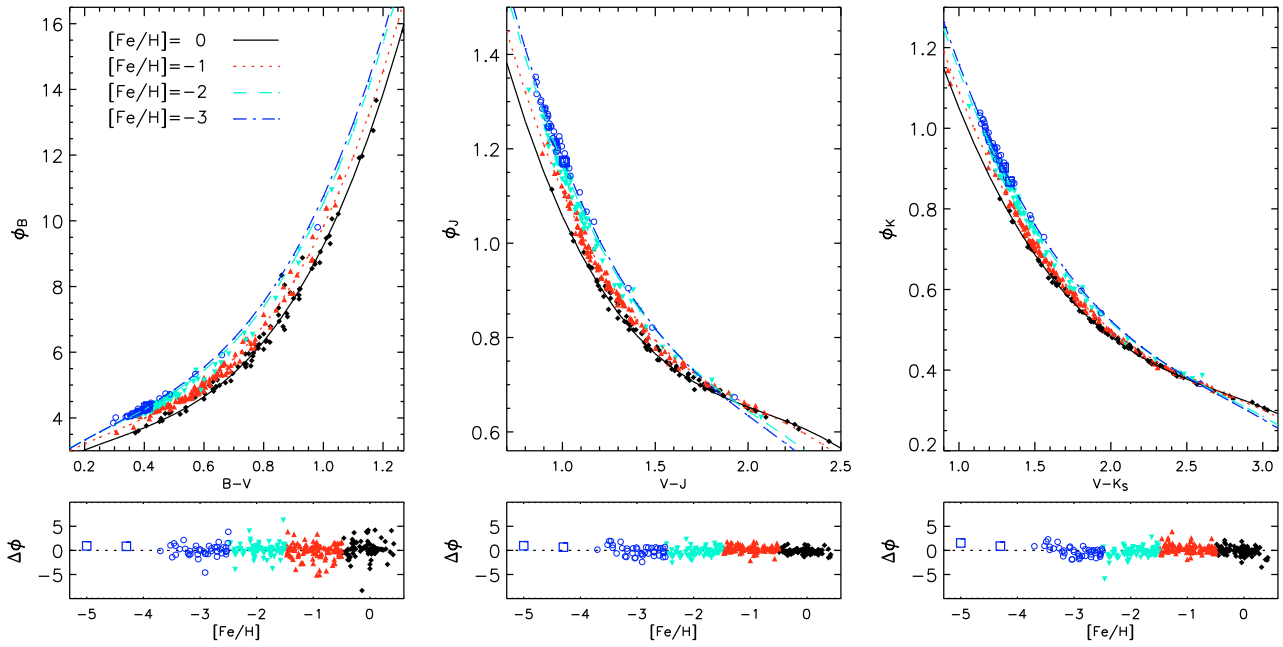

Fig. 16. Same as Fig. 14, but for the colour-flux-metallicity calibrations. The reduced flux in different bands $\phi_\zeta = \mathcal{F}_{\rm Bol}({\rm Earth})\, 10^{0.4\, m_\zeta}$ is plotted as function of different colour indices in units of $10^{-0.5}$ erg cm^{-2} s^{-1}.

Table 7. The colours of the Sun.

	$T_{\rm eff}$ scale (rand. + syst. errors)	MARCS	ATLAS9	Twins
$B-V$	0.641 ± 0.024 ± 0.004	0.622	0.645	
$V-R_{\rm C}$	0.359 ± 0.010 ± 0.003	0.357	0.358	
$(R-I)_{\rm C}$	0.333 ± 0.010 ± 0.002	0.347	0.349	
$V-I_{\rm C}$	0.690 ± 0.016 ± 0.004	0.704	0.707	
$V-J$	1.180 ± 0.021 ± 0.007	1.171	1.180	
$V-H$	1.460 ± 0.023 ± 0.010	1.458	1.479	
$V-K_{\rm S}$	1.544 ± 0.018 ± 0.010	1.543	1.553	
$J-K_{\rm S}$	0.362 ± 0.029 ± 0.003	0.372	0.373	
$(B-V)_{\rm T}$	0.730 ± 0.031 ± 0.006	0.723	0.750	0.735 ± 0.024[a]
$V_{\rm T}-J$	1.254 ± 0.022 ± 0.008	1.240	1.250	1.254 ± 0.041[a]
$V_{\rm T}-H$	1.534 ± 0.019 ± 0.011	1.527	1.550	1.549 ± 0.048[a]
$V_{\rm T}-K_{\rm S}$	1.619 ± 0.013 ± 0.011	1.612	1.623	1.623 ± 0.040[a]
$b-y$	0.409 ± 0.010 ± 0.002			0.409 ± 0.003[b]

Notes. For the indices obtained inverting our $T_{\rm eff}$ scale, random errors are from the dispersion of the fits in Table 4. The uncertainty on the zero point of our $T_{\rm eff}$ scale is of order 15 K (Sect. 3.2), which usually implies systematic errors considerably smaller than the random ones. The only exception is for optical-infrared indices which are very sensitive to $T_{\rm eff}$ and show small intrinsic scatter, of the order of the aforementioned zero point uncertainty. Also shown for comparison are the averaged colours and standard deviation of the solar twins, as well as the synthetic colours computed from the ATLAS9 (Castelli & Kurucz 2004) and MARCS (Gustafsson et al. 2008) models.
[(a)] Average and standard deviation of the colours in Table 2.
[(b)] From Meléndez et al. in prep., fitting solar twin colours as a function of $T_{\rm eff}$, [Fe/H] and log g.

Rieke et al. (2008) and obtained from solar-type stars or computed convolving various solar spectra with the appropriate filter curves and using their revised absolute physical calibration.

The empirical colours in Table 7 are also in agreement with the synthetic ones, computed using the same zero points and absolute calibration for Vega used in the IRFM to derive our $T_{\rm eff}$ scale. Therefore, the uncertainty in the zero points used to generate synthetic colours is at the smallest level possible, yet of the order of 0.01 mag (Sect. 3.2), allowing us to address the reliability of the models at this level of precision. While using a theoretical spectra of Vega may (partly) compensate model inaccuracies in the process of setting the zero points, the approach adopted here allows us to focus on the quality of the solar synthetic spectra. The agreement is remarkable, on the order of 0.01 mag and never exceeding 0.02, which is also of the same size of the difference between those synthetic models.

7. Conclusions

The primary goal of this work has been to provide a new absolute effective temperature scale. An unprecedented accuracy of few tens of Kelvin in the zero point of our scale has been achieved using a sample of solar twins. For these stars the high degree of resemblance to the Sun has been determined entirely model independently, without any prior assumption on their physical parameters, most importantly $T_{\rm eff}$. Notice that by calibrating our results via solar twins we are entirely unaffected from possible issues and uncertainties related to Vega. Nonetheless, we regard as comforting that our findings are in close agreement with the latest absolute fluxes (Cohen et al. 2003; Bohlin 2007; Rieke et al. 2008). We further took advantage of such a promising situation by fine-tuning the adopted fluxes so as to improve the consistency of the effective temperatures determined from each band used in the IRFM. This methodology gives us confidence that the stellar parameters we determined are well calibrated not only around the solar value, but over a wide range in $T_{\rm eff}$ and [Fe/H]. Notice that the IRFM is little model dependent and certainly not at the solar value because of our calibration procedure. Small spurious trends arising from the adopted library at different temperatures and metallicities can not be entirely ruled out, but should be small. Though the zero point of our new $T_{\rm eff}$ scale is entirely set by solar twins, it agrees within few degrees with independent verifications conducted via interferometric angular diameters and HST spectrophotometry in the metal-poor and -rich regimes.

In the process of establishing the zero point of the effective temperature scale via IRFM, we nailed down the differences with respect to other implementations of the same technique. We have used two independent IRFM versions to study the discrepancies among various temperature scales that appeared in literature over the years and proved that the absolute calibration of the photometric systems used was responsible for explaining most of the differences. At solar temperatures and metallicities the long-standing dichotomy between photometric and spectroscopic $T_{\rm eff}$ is easily explained once it is understood that the IRFM can in principle accommodate any temperature scale since its zero point depends on the absolute calibration of the photometry adopted. The main goal of the present paper has been exactly to tackle this issue using the best constraint available to date.

The improved bolometric fluxes determined for metal-poor stars have also been used to put on firmer ground the temperature scale in this rather unexplored regime. For metallicities typical of halo stars our $T_{\rm eff}$ scale is roughly 100 K hotter than those determined from the Balmer lines and 200 K hotter than those obtained from the excitation equilibrium. While spectroscopic effective temperature determinations have considerable model dependence and are degenerate with other stellar parameters (namely $\log g$ and [Fe/H]), the IRFM offers a powerful alternative, free from any of the above limitations. However, relying on the photometry, the IRFM is influenced by reddening, which becomes a considerable source of uncertainty when targeting objects outside of the local bubble. For our sample of metal-poor stars we have been cautious in determining reddening as best we could. Our improved determination of $E(B-V)$ also explain the remaining discrepancies with other $T_{\rm eff}$ scales. We think the effective temperatures determined for our sample of stars will serve to better calibrate spectroscopic $T_{\rm eff}$ determinations. This will be particularly relevant when large spectroscopic surveys targeting different stellar populations in the Galaxy start operating: support from the existing or forthcoming photometric surveys will be possible only if reddening will be determined on a star-by-star basis. We feel this will not be possible in many cases and stellar parameters will have to rely on spectroscopy only.

Based on our sample of dwarfs and subgiants, a set of homogeneously calibrated colours versus temperatures, bolometric fluxes and angular diameters have also been determined. A number of problems of interest to stellar and Galactic Chemical evolution depend on the assumption made in these relations and our results will permit those problems to be tackled with greater confidence.

Acknowledgements. We thank Ana García-Pérez for preliminary results on HE0233-0343 and interesting discussion on excitation temperatures in metal-poor stars. We are also indebted to Maria Bergemann and Andreas Korn for many insights on determining effective temperatures from Balmer lines. Martin Cohen is acknowledged for useful correspondence on the absolute calibration and Gerard van Belle for enlightening discussions on interferometry at various times. Don Vandenberg is kindly acknowledged for useful correspondence and for a preliminary comparison of our temperature scale with M 67. We thank an anonymous referee for relevant comments and suggestions that helped to strength the presentation and clarify the results. J.M. is supported by a Ciência 2007 contract (FCT/MCTES/Portugal and POPH/FSE/EC) and acknowledges financial support from PTDC/CTE-AST/65971/2006 (FCT/Portugal). This research has made use of the General Catalogue of Photometric data operated at the University of Lausanne and the SIMBAD database, operated at CDS, Strasbourg, France. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

Table A.1. Characteristic parameters of the 2MASS and TCS photometric systems.

	2MASS	TCS	Δ(%)
$\lambda_{\rm eff}$	12 285	12 790	
m_J	−0.001	−0.013	
$\mathcal{F}_J$	3.079	2.912	
$\mathcal{F}_{\rm eff}$	3.076	3.303	−6.9%
$\lambda_{\rm eff}$	16 385	16 483	
m_H	+0.019	−0.005	
$\mathcal{F}_H$	1.150	1.192	
$\mathcal{F}_{\rm eff}$	1.170	1.211	−3.4%
$\lambda_{\rm eff}$	21 521	21 869	
m_K	−0.017	−0.029	
$\mathcal{F}_K$	0.430	0.426	
$\mathcal{F}_{\rm eff}$	0.423	0.439	−3.6%

Notes. Wavelengths are in Å and the Vega's monochromatic absolute fluxes in 10^{-10} erg cm^{-2} s^{-1} Å^{-1}. 2MASS effective wavelengths ($\lambda_{\rm eff}$), magnitudes and fluxes are from Casagrande et al. (2006), where the latter has been modified by −1.6, +1.5 and +0.3 percent in J, H and $K_{\rm S}$ band respectively as described in Sect. 3.2. The TCS values are from Alonso et al. (1994b). For the TCS system $\mathcal{F}_{\rm eff}$ has been computed by shifting the value at the 2MASS effective wavelength (Fig. A.1). Δ(%) is the percent decrease of the TCS $\mathcal{F}_{\rm eff}$ needed to match the 2MASS values.

Appendix A: Comparing the TCS and 2MASS absolute calibration

The absolute calibration of Alonso et al. (1994a) was obtained applying the IRFM to a sample of stars for which direct measurements of angular diameters were available. Because of the difficulties involved in achieving milli-arcsecond resolution, that sample was almost entirely composed of giants with angular diameters measured via Lunar Occultations and Michelson Interferometry ($T_{\rm eff} < 5000$ K) or Intensity Interferometry ($T_{\rm eff} > 6000$ K). One of the intriguing results of that analysis was the impossibility of setting the same zero point of the absolute calibration using angular diameters measured by Lunar Occultations (White & Feierman 1987; Ridgway et al. 1980) and Michelson Interferometry (Hutter et al. 1989; di Benedetto & Rabbia 1987; Mozurkewich et al. 1991) with those measured by Intensity Interferometry (Hanbury Brown et al. 1974). The absolute calibration (in the Johnson system) proposed by Alonso et al. (1994a) is a weighted average from their table 10 and it is interesting to notice that the one derived from Intensity Interferometry alone is 4.8 (J) 1.3 (H) and 4.0 (K) percent lower than the averaged, proposed one. As we have discuss throughout the paper[8], lower infrared fluxes support higher $T_{\rm eff}$ (in this case, the average difference in Johnson system would be 3.4% supporting $T_{\rm eff}$ approximately hotter by 70 K), so it is not surprising our effective temperature scale provides good agreement with interferometric measurements despite being considerably hotter than most of the previous IRFM analyses.

To gauge a further insight into the problem, here we directly compare the TCS (given in Alonso et al. 1994b) and 2MASS absolute calibration. Such an exercise, however is not straightforward since the absolute calibration in different photometric system is obtained using different filter transmission curves and therefore is associated to different effective wavelengths. In addition, for the sake of the IRFM, in any given band ζ, it is the

[8] We have verified using our IRFM implementation that a 1% increase in infrared fluxes correspond to a decrease of 20 K in $T_{\rm eff}$.

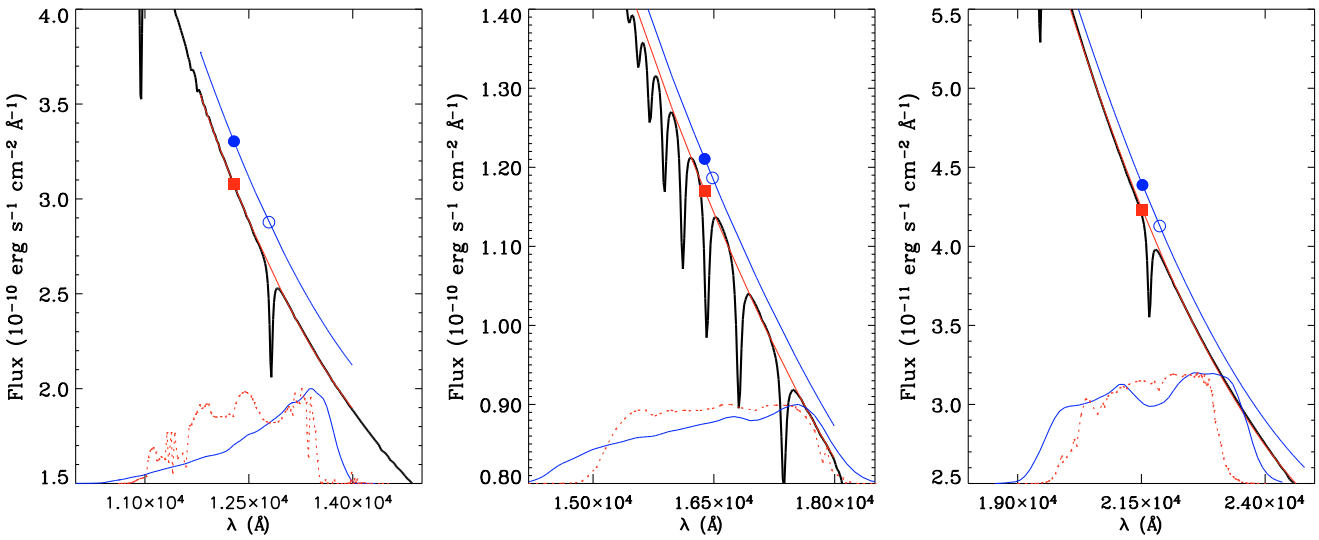

Fig. A.1. Comparison between $\mathcal{F}_{\rm eff}$ in 2MASS (filled squares) and the TCS before (open circles) and after (filled circles) correcting for the same $\lambda_{\rm eff}$. The correction has been done shifting the TCS values along the continuum of Vega, obtained by fitting a second order polynomial to the observed spectral energy distribution (from Bohlin 2007). Overplotted for comparison are the 2MASS (dotted lines) and TCS (continuous lines) filter transmission curves.

composite effect of Vega's magnitudes and fluxes which matters, i.e. $\mathcal{F}_\zeta 10^{0.4\ m_\zeta}$. Therefore, for a meaningful comparison we need to refer everything to a common wavelength, the 2MASS one being the natural choice in this case. This is done in Table A.1 by computing $\mathcal{F}_{\rm eff}$ i.e the composite effect of magnitudes and fluxes shifted to the 2MASS $\lambda_{\rm eff}$ in the case of TCS (Fig. A.1).

The 2MASS absolute calibration is on the average lower than the TCS by 4.6% (a value qualitatively in agreement with the difference in the Johnson system discussed above), thus returning $T_{\rm eff}$ on average hotter by ~90 K, and explaining the bulk of the differences discussed in Sect. 2.3 when comparing the sample stars directly. Similar conclusions can be drawn when comparing with the absolute fluxes and magnitudes of Vega used in Table 1 of González Hernández & Bonifacio (2009). In this case the photometric system is the same (2MASS) and one can directly compare $\mathcal{F}_{\rm eff}$: the difference is −3.3% (J), +1.3% (H) and −2.8% ($K_{\rm S}$) thus giving an average of −1.6% which correspond to ~30 K, again in line with the differences discussed in Sect. 2.5.

References

Allende Prieto, C., Asplund, M., García López, R. J., & Lambert, D. L. 2002, ApJ, 567, 544
Alonso, A., Arribas, S., & Martinez-Roger, C. 1994a, A&A, 282, 684
Alonso, A., Arribas, S., & Martinez-Roger, C. 1994b, A&AS, 107, 365
Alonso, A., Arribas, S., & Martinez-Roger, C. 1995, A&A, 297, 197
Alonso, A., Arribas, S., & Martinez-Roger, C. 1996a, A&AS, 117, 227
Alonso, A., Arribas, S., & Martinez-Roger, C. 1996b, A&A, 313, 873
Aoki, W., Frebel, A., Christlieb, N., et al. 2006, ApJ, 639, 897
Aoki, W., Barklem, P. S., Beers, T. C., et al. 2009, ApJ, 698, 1803
Asplund, M. 2005, ARA&A, 43, 481
Asplund, M., & García Pérez, A. E. 2001, A&A, 372, 601
Baines, E. K., McAlister, H. A., ten Brummelaar, T. A., et al. 2008, ApJ, 680, 728
Barklem, P. S. 2007, A&A, 466, 327
Barklem, P. S., Piskunov, N., & O'Mara, B. J. 2000, A&AS, 142, 467
Barklem, P. S., Stempels, H. C., Allende Prieto, C., et al. 2002, A&A, 385, 951
Beers, T. C., Flynn, C., Rossi, S., et al. 2007, ApJS, 168, 128
Bell, R. A., & Gustafsson, B. 1989, MNRAS, 236, 653
Bergemann, M. 2008, Physica Scripta Volume T, 133, 014013
Bertelli, G., Girardi, L., Marigo, P., & Nasi, E. 2008, A&A, 484, 815
Bessell, M. S. 2000, PASP, 112, 961
Bessell, M. S. 2005, ARA&A, 43, 293
Bessell, M. S., & Brett, J. M. 1988, PASP, 100, 1134
Bessell, M. S., Castelli, F., & Plez, B. 1998, A&A, 333, 231
Biazzo, K., Frasca, A., Catalano, S., & Marilli, E. 2007, Astron. Nachr., 328, 938
Bigot, L., Kervella, P., Thévenin, F., & Ségransan, D. 2006, A&A, 446, 635
Blackwell, D. E., & Shallis, M. J. 1977, MNRAS, 180, 177
Blackwell, D. E., Shallis, M. J., & Selby, M. J. 1979, MNRAS, 188, 847
Blackwell, D. E., Petford, A. D., & Shallis, M. J. 1980, A&A, 82, 249
Bohlin, R. C. 2007, in The Future of Photometric, Spectrophotometric and Polarimetric Standardization, ed. C. Sterken, ASP Conf. Ser., 364, 315
Bohlin, R. C., & Gilliland, R. L. 2004, AJ, 127, 3508
Bonifacio, P., Molaro, P., Sivarani, T., et al. 2007, A&A, 462, 851
Boyajian, T. S., McAlister, H. A., Baines, E. K., et al. 2008, ApJ, 683, 424
Brown, T. M., Charbonneau, D., Gilliland, R. L., Noyes, R. W., & Burrows, A. 2001, ApJ, 552, 699
Caccin, B., Penza, V., & Gomez, M. T. 2002, A&A, 386, 286
Campins, H., Rieke, G. H., & Lebofsky, M. J. 1985, AJ, 90, 896
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Casagrande, L. 2008, Physica Scripta Volume T, 133, 014020
Casagrande, L. 2009, Mem. Soc. Astron. Ital., 80, 727
Casagrande, L., Portinari, L., & Flynn, C. 2006, MNRAS, 373, 13
Casagrande, L., Flynn, C., & Bessell, M. 2008, MNRAS, 389, 585
Castelli, F., & Kurucz, R. L. 1994, A&A, 281, 817
Castelli, F., & Kurucz, R. L. 2004, Modelling of Stellar Atmospheres, ed. N. Piskunov et al., IAU Symp., 210 [arXiv:astro-ph/0405087]
Cayrel de Strobel, G. 1996, A&A Rev., 7, 243
Cayrel de Strobel, G., & Bentolila, C. 1989, A&A, 211, 324
Cayrel de Strobel, G., Soubiran, C., & Ralite, N. 2001, A&A, 373, 159
Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJ, 529, L45
Cohen, M., Wheaton, W. A., & Megeath, S. T. 2003, AJ, 126, 1090
di Benedetto, G. P., & Rabbia, Y. 1987, A&A, 188, 114
Edvardsson, B. 2008, Physica Scripta Volume T, 133, 014011
Fabbian, D., Nissen, P. E., Asplund, M., Pettini, M., & Akerman, C. 2009, A&A, 500, 1143
Frebel, A., Aoki, W., Christlieb, N., et al. 2005, Nature, 434, 871
Frebel, A., Collet, R., Eriksson, K., Christlieb, N., & Aoki, W. 2008, ApJ, 684, 588
Fuhrmann, K. 2008, MNRAS, 384, 173
Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748
García Pérez, A. E., Christlieb, N., Ryan, S. G., et al. 2008, Physica Scripta Volume T, 133, 014036
González Hernández, J. I., & Bonifacio, P. 2009, A&A, 497, 497
Gray, D. F. 1994, PASP, 106, 1248
Gray, D. F., & Johanson, H. L. 1991, PASP, 103, 439
Gray, R. O. 2007, in The Future of Photometric, Spectrophotometric and Polarimetric Standardization, ed. C. Sterken, ASP Conf. Ser., 364, 305
Grupp, F. 2004, A&A, 426, 309
Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, 951
Hanbury Brown, R., Davis, J., & Allen, L. R. 1974, MNRAS, 167, 121
Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27
Holmberg, J., Flynn, C., & Portinari, L. 2006, MNRAS, 367, 449

Holmberg, J., Nordström, B., & Andersen, J. 2009, A&A, 501, 941
Hosford, A., Ryan, S. G., García Pérez, A. E., Norris, J. E., & Olive, K. A. 2009, A&A, 493, 601
Hutter, D. J., Johnston, K. J., Mozurkewich, D., et al. 1989, ApJ, 340, 1103
Johnson, H. L. 1965, Communications of the Lunar and Planetary Laboratory, 3, 73
Kervella, P., & Fouqué, P. 2008, A&A, 491, 855
Korn, A. J., Richard, O., Mashonkina, L., et al. 2009, ApJ, 698, 410
Kovtyukh, V. V., Soubiran, C., Belik, S. I., & Gorlova, N. I. 2003, A&A, 411, 559
Kurucz, R. L. 1993, ATLAS9 Stellar Atmosphere Programs and 2 km s^{-1} grid. Kurucz CD-ROM No. 13 (Cambridge, Mass.: Smithsonian Astrophysical Observatory), 13
Lallement, R., Welsh, B. Y., Vergely, J. L., Crifo, F., & Sfeir, D. 2003, A&A, 411, 447
Latham, D. W., Mazeh, T., Carney, B. W., et al. 1988, AJ, 96, 567
Leroy, J. L. 1993, A&A, 274, 203
Ludwig, H., Behara, N. T., Steffen, M., & Bonifacio, P. 2009, A&A, 502, L1
Maíz-Apellániz, J. 2007, in The Future of Photometric, Spectrophotometric and Polarimetric Standardization, ed. C. Sterken, ASP Conf. Ser., 364, 227
Masana, E., Jordi, C., & Ribas, I. 2006, A&A, 450, 735
McCall, M. L. 2004, AJ, 128, 2144
Meléndez, J., & Ramírez, I. 2004, ApJ, 615, L33
Meléndez, J., & Ramírez, I. 2007, ApJ, 669, L89
Meléndez, J., Dodds-Eden, K., & Robles, J. A. 2006a, ApJ, 641, L133
Meléndez, J., Shchukina, N. G., Vasiljeva, I. E., & Ramírez, I. 2006b, ApJ, 642, 1082
Meléndez, J., Asplund, M., Gustafsson, B., & Yong, D. 2009a, ApJ, 704, L66
Meléndez, J., Ramírez, I., Casagrande, L., et al. 2009b, Ap&SS, in press [arXiv:0910.5845]
Meléndez, J., Casagrande, L., Ramírez, I., Asplund, M., & Schuster, W. I. 2010, A&A, submitted
Mermilliod, J.-C., Mermilliod, M., & Hauck, B. 1997, A&AS, 124, 349
Mozurkewich, D., Johnston, K. J., Simon, R. S., et al. 1991, AJ, 101, 2207
Munari, U., Sordo, R., Castelli, F., & Zwitter, T. 2005, A&A, 442, 1127
Nissen, P. E., Akerman, C., Asplund, M., et al. 2007, A&A, 469, 319
Nordgren, T. E., Germain, M. E., Benson, J. A., et al. 1999, AJ, 118, 3032
Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989
North, J. R., Davis, J., Robertson, J. G., et al. 2009, MNRAS, 393, 245
O'Donnell, J. E. 1994, ApJ, 422, 158
Pasquini, L., Biazzo, K., Bonifacio, P., Randich, S., & Bedin, L. R. 2008, A&A, 489, 677
Pereira, T., Asplund, M., Trampedach, R., & Collet, R. 2010, A&A, submitted
Porto de Mello, G. F., & da Silva, L. 1997, ApJ, 482, L89
Ramírez, I., Allende Prieto, C., Redfield, S., & Lambert, D. L. 2006, A&A, 459, 613
Ramírez, I., & Meléndez, J. 2005a, ApJ, 626, 446
Ramírez, I., & Meléndez, J. 2005b, ApJ, 626, 465
Ramírez, I., Meléndez, J., & Asplund, M. 2009, A&A, 508, L17
Ridgway, S. T., Joyce, R. R., White, N. M., & Wing, R. F. 1980, ApJ, 235, 126
Rieke, G. H., Blaylock, M., Decin, L., et al. 2008, AJ, 135, 2245
Ryan, S. G., Norris, J. E., & Beers, T. C. 1999, ApJ, 523, 654
Schuster, W. J., & Nissen, P. E. 1989, A&A, 221, 65
Sekiguchi, M., & Fukugita, M. 2000, AJ, 120, 1072
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Smith, J. A., Tucker, D. L., Kent, S., et al. 2002, AJ, 123, 2121
Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, A&A, 487, 373
Takeda, Y., & Tajitsu, A. 2009, PASJ, 61, 471
Valenti, J. A., & Fischer, D. A. 2005, ApJS, 159, 141
van Belle, G. T., & von Braun, K. 2009, ApJ, 694, 1085
van Leeuwen, F. 2007, A&A, 474, 653
VandenBerg, D. A., & Clem, J. L. 2003, AJ, 126, 778
Venn, K. A., & Lambert, D. L. 2008, ApJ, 677, 572
White, N. M., & Feierman, B. H. 1987, AJ, 94, 751