Astronomy & Astrophysics manuscript no. paper˙feh˙testv16˙final˙referee
October 10, 2011

Metallicity of M dwarfs

II. A comparative study of photometric metallicity scales*

V. Neves[1,2,3], X. Bonfils[2], N. C. Santos[1,3], X. Delfosse[2], T. Forveille[2], F. Allard[4], C. Natário[5,6], C. S. Fernandes[5], and S. Udry[7]

[1] Centro de Astrofísica, Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
email: `vasco.neves@astro.ua.pt`

[2] UJF-Grenoble 1 / CNRS-INSU, Institut de Planétologie et d'Astrophysique de Grenoble (IPAG) UMR 5274, Grenoble, F-38041, France.

[3] Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal

[4] Centre de Recherche Astrophysique de Lyon, UMR 5574: CNRS, Université de Lyon, École Normale Supérieure de Lyon, 46 Allée d'Italie, F-69364 Lyon Cedex 07, France

[5] Centro de Astronomia e Astrofísica da Universidade de Lisboa, Campo Grande, Ed. C8 1749-016 Lisboa, Portugal

[6] Leiden Observatory, Leiden University, The Netherlands

[7] Observatoire de Genève, Université de Genève, 51 Chemin des Maillettes, 1290 Sauverny, Switzerland

Received XXX; accepted XXX

DELETE
PARAGRAPH
CHANGE TO LOWER CASE / & lc
CHANGE TO UPPER CASE ≡ & cap
CHANGE WORD ORDER
ADD BETWEEN
ADD TO THE END OF A WORD
ADD TO THE BEGINNING
CLOSE UP SPACING
A SPACE #

ABSTRACT

Context. Stellar parameters are not easily derived from M dwarf spectra, which are dominated by complex bands of diatomic and triatomic molecules and not well described at the line by line level by atmospheric models. M dwarf metallicities are therefore most commonly derived through less direct techniques. Several recent publications propose calibrations that provide the metallicity of an M dwarf from its V band absolute magnitude and its V−K color, but disagree at the ±0.1 dex level. We compare these calibrations on a sample of 23 M dwarfs, which we select as wide (> 5 arcsec) companions of F-, G-, or K- dwarfs with metallicites measured on a homogeneous scale, and which we require to have V band photometry measured to better than ∼0.03 magnitude. We find that the Schlaufman & Laughlin (2010) calibration has lowest offsets and residuals against our sample, and use our improved statistics to marginally refine that calibration. With more strictly selected photometry than in previous studies, the dispersion around the calibration is well in excess of the [Fe/H] and photometric uncertainties. This suggests that the origin of the remaining dispersion is astrophysical rather than observational.

Aims. .

Methods. .

Results. .

Conclusions. .

Key words. stars: fundamental parameters – stars: binaries - general – stars: late type – stars: atmospheres – stars: planetary systems

1. Introduction

M dwarfs are the smallest and coldest stars of the main sequence. Long lived and ubiquitous, M dwarfs are of interest in many astrophysical contexts, from stellar evolution to the structure of our Galaxy. Most recently, interest in M dwarfs has been further increased by planet search programs. Planets induce ~~larger~~ higher reflex velocities and deeper transits when they orbit and transit M dwarfs rather than larger FGK stars, and the habitable zone of the less luminous M dwarfs are closer-in. Lower mass, smaller, and possibly habitable planets are therefore easier to find around M dwarfs, and are indeed detected at an increasing pace (e.g. Udry et al. 2007; Mayor et al. 2009).

Interesting statistical correlations between the characteristics of exoplanets and the properties of their host stars have emerged from the growing sample of exoplanetary systems (e.g. Endl et al. 2006; Johnson et al. 2007; Udry & Santos 2007, Bonfils et al. 2011 in prep.). Of those, the planet-metallicity correlation was first identified and remains the best established: a higher metal content increases, on average, the probability that a star hosts Jovian planets (Gonzalez 1997; Santos et al. 2001, 2004; Fischer & Valenti 2005). Within the core-accretion paradigm for planetary formation, that correlation reflects the ~~larger~~ higher mass of solid material available to form protoplanetary cores in the protoplanetary disks of higher-metallicity stars. The correlation is then expected to extend to, and perhaps be reinforced in, the cooler M dwarfs: cap to counterbalance the lower overall mass of their protoplanetary disks, those disks need a higher fraction of refractory material to form similar populations of the protoplanetary core. Wh~~ereas~~ then the planet-metallicity correlation that seems to vanish for Neptunes and lower-mass planets around FGK stars (Sousa et al. 2008; Bouchy et al. 2009), persists for Neptune-mass planets around M dwarfs is still an open question. ok?

Our derivation of the first photometric metallicity calibration for M dwarfs (Bonfils et al. 2005) was largely motivated by probing their planet-metallicity correlation, though only two M-dwarf planetary systems were known at the time. A few planet detections later, a Kolmogorov-Smirnov test of the metallicity distributions of M dwarfs *with* and

⋆ Based on observations collected with the FEROS spectrograph at la Silla observatory under ESO programs 073.D-0802(A), 074.D-0670(A), 078.D-0760(A), and with the ELODIE and SOPHIE spectrographs at the cap observatoire de Haute Provence.

without known planets indicated that they only had a $\sim$ 11% probability ~~to~~ of be ing drawn from a single parent distribution (Bonfils et al. 2007). With an improved metallicity calibration and a larger sample of M dwarf planets, Schlaufman & Laughlin (2010) lower~~ed~~ s the probability that M-dwarf planetary hosts have the same metallicity distribution as the general M dwarf population to $\sim$ 6%. This result is in line with expectations for the core accretion paradigm, but is only significant at the $\sim$2 σ level. lc Finding planets around additional M dwarfs and measuring metallicity more precisely will key both help characterize this correlation~~, or~~ and the possibl~~y~~ a lack thereof. Here we explore the second avenue.

Measuring accurate stellar parameters from the optical spectra of M dwarfs unfortunately is not easy. As the abundances of diatomic and triatomic molecules (e.g. TiO, VO, H_2O, CO) in the photospheric layers increases with spectral subtype, their forest of weak lines eventually erases the spectral continuum, and makes a line-by-line spectroscopic analysis difficult for all but the earlier M subtypes. Woolf & Wallerstein (2005, 2006) measured atomic abundances from the high-resolution spectra of 67 K and M dwarfs through a classical line-by-line analysis, but had to restrict their work to the earliest M subtypes (T_{eff} > 3500 K) and to mostly metal-poor stars (median [Fe/H]= -0.89 dex). They find that metallicity correlates with CaH and TiO band strengths, but do not offer a quantitative calibration.

Although the recent revision of the solar oxygen abundance (Asplund et al. 2009; Caffau et al. 2011) has greatly improved the agreement between model atmosphere prediction and spectra of M dwarfs observed at low-to-medium-s resolution (Allard et al. 2010), many visual-to-red spectral features still correspond to molecular bands ~~which~~ that are missing or incompletely described in the opacity databases that underly the atmospheric models. At high spectral resolution, many individual molecular lines in synthetic spectra are additionally displaced from their actual position. Spectral synthesis, as well, has therefore had limited success in analy~~s~~ z ing M dwarf spectra (e.g. Valenti et al. 1998; Bean et al. 2006). In this context, less direct techniques have been developed to evaluate the metal content of M dwarfs. Of those, the most successful leverage~~s~~ s the photometric effects of the very molecular bands ~~which~~ that complicate spectroscopic analyses. Increased TiO and VO abundances in metal-rich M dwarfs shift radiative flux from the visible range, where these species dominate the opacities, to the near infrared. For a fixed mass, an increased metallicity also reduces the bolometric luminosity. Those two effects of metallicity work together in the visible, but, in the [Fe/H] and T_{eff} range of interest here, they largely cancel out in the near-infrared. As a result, the absolute V magnitude on an M dwarf is very sensitive to its metallicity, while its near infrared magnitudes are not (Chabrier & Baraffe 2000; Delfosse et al. 2000). Position in a lc Color/lc Absolute lc Magnitude diagram that combines visible and near-infrared bands is therefore a sensitive metallicity probe, but one ~~which~~ that needs external calibration.

We pioneered that approach in Bonfils et al. (2005), where we anchor~~ing~~ed the relation on a combination spectroscopic metallicities of early-M dwarfs from Woolf & Wallerstein (2005) and metallicities, which we measured for the FGK primaries of binary systems containing a widely separated M dwarf component. That calibration, in terms of the V-band absolute magnitude and the $V-K$ color, results in a modestly significant disagreement between the mean metallicity of solar-neighborhood early/mid-M dwarfs and FGK dwarfs. Johnson & Apps (2009) correctly point~~ed~~s out that M and (at least) K dwarfs have the same age distribution, since both live longer than the age of the universe, and that they are therefore expected to have identical metallicity distributions. They derived an alternative calibration, anchored ~~on~~ in FGK+M binaries that partly overlap the Bonfils et al. (2005) sample, which forces the agreement of the mean metallicities of local samples of M and FGK dwarfs. Most recently, Schlaufman & Laughlin (2010) have pointed out the importance of kinematically matching the M and GK samples before comparing their metallicity distributions, and used stellar structure models of M dwarfs to guide their choice of a more effective parametrization of position in the M_V vs $V-K$ diagram. The difference between the ~~3~~ three calibrations varies slightly across the Herzprung-Russell diagram but, on average, the Johnson & Apps (2009) calibration is 0.2 dex more metal-rich than Bonfils et al. (2005), and Schlaufman & Laughlin (2010) is half-way between those two extremes. Those discrepancies are largely irrelevant when comparing M dwarfs with metallicities consistently measured on any of these ~~3~~ three scales, but they are uncomfortably large in any comparison with external information.

We set out here to test those ~~3~~ three calibrations. For that purpose, we have assembled a sample of 23 M dwarfs with accurate photometry, parallaxes, and metallicity measured from a hotter companion (Sect. 2). We then perform statistical tests of the three calibrations in Sect. 3, and in Sect. 4 we discuss those results and slightly refine the Schlaufman & Laughlin (2010) calibration, which we find works best. Sect~~.~~ion 5 presents our conclusions, and an appendix compares our preferred calibration against metallicities obtained with independent techniques.

2. Sample and Observations

lc

We adopt the now well-established route of measuring the metal content of the primaries of FGK+M binaries through classical spectroscopic methods, ~~and~~ by assuming that it applies to the M secondaries. We searched for such binaries in the ~~3^{rd}~~ third edition of the catalog of nearby stars (Gliese & Jahreiß 1991), the catalog of nearby wide binary and multiple systems (Poveda et al. 1994), the catalog of common proper-motion companions to *Hipparcos* stars (Gould & Chanamé 2004), and the catalog of disk and halo binaries from the revised Luyten catalog (Chanamé & Gould 2004). To ensure uncontaminated measurements of the fainter M secondaries, we required separations of at least 5 arcsec.

That initial selection identified almost 300 binaries. We eliminated known fast rotators, spectroscopic binaries, pairs without a demonstrated common proper motion, as well as systems ~~which~~ that do not figure in the revised *Hipparcos* catalog (van Leeuwen 2007) from which we obtained the parallaxes of the primaries. With very few exceptions, the secondaries have good JHK$_S$ photometry in the 2MASS catalog (Skrutskie et al. 2006), which we therefore adopt as our source of near-infrared photometry. The only exception is Gl 551 (Proxima Centauri), which has saturated K$_S$ 2MASS measurements and for which we use the Bessell (1991) measurements **that we transform into K$_S$ photometry using the equations of Carpenter (2001)**. Precise optical photometry of the secondaries, to our initial surprise, has been less forthcoming, and we suspected noise in their V-band photometry to contribute much of the dispersion seen in previous photometric metallicity calibrations. We therefore applied a strict threshold in our literature search and only retained pairs in which the V band magnitude of the secondary is measured to better than 0.03 magnitude. This criterion turned out to ~~most~~ severely restrict our sample, and we plan to obtain V band photometry for the many systems ~~which~~ that meet all our other requirements, including the availability of a good high-resolution spectrum of the primary. Mermilliod et al. (1997) has been our main source of Johnson-Cousins VRI photometry. For ~~10~~ ten sources, RI photometry was in Weistrop and Kron systems instead of Johnson-Cousins. We therefore applied transformations following Weistrop (1975) and Leggett (1992), respectively. The RIJH photometry was used to calculate metallicity from the Casagrande et al. (2008) calibration, as discussed in the Appendix. Our final sample contains 23 systems, of which 19 have M-dwarf secondaries and ~~4~~ four have K7/K8 secondaries.

We either measured the metallicity of the primaries from high-resolution spectra or adopted measurements from the literature, which are on the same metallicity scale. We obtained spectra for ~~9~~ nine stars with the FEROS spectrograph (Kaufer & Pasquini 1998) on the 2.2m ESO/MPI telescope at La Silla. We used the ARES program (Sousa et al. 2007) to automatically measure the equivalent widths **of the Fe 1 and Fe 2 weak lines (< 200 mÅ) in the Fe line list of Sousa et al. (2008). This list is comprised of 263 Fe 1 and 36 Fe 2 stable lines, ranging, in wavelength, from 4500 to 6890 Å.** Then, we followed the procedure described in Santos et al. (2004): [Fe/H] and the stellar parameters are determined by imposing excitation and ionization equilibrium, using the 2002 version of the MOOG (Sneden 1973) spectral synthesis program with a grid of ATLAS9 plane-parallel model atmospheres (Kurucz 1993).

For ~~3~~ three stars, we used spectra gathered with the CORALIE (Queloz et al. 2000) and SOPHIE (Bouchy & The Sophie Team 2006) spectrographs on the Swiss Euler 1.2 m telescope at la Silla ~~and~~ on the Haute Provence 1.93 m telescope, respectively. For those ~~3~~ three stars, we use metallicities derived from a calibration of the equivalent width of the cross correlation function (CCF) of their spectra with numerical templates (Santos et al.

with

Table 1. Stellar parameters measured from the primaries. The [Fe/H] of the M dwarf secondary is inferred from the primary.

Primary	Secondary	T_{eff} [K]	log g [cm s^{-2}]	ξ_t [km s^{-1}]	[Fe/H]	[Fe/H] source	T_{eff} source
Gl53.1A	Gl53.1B	4705 ± 131	4.33 ± 0.26	0.76 ± 0.25	0.07 ± 0.12	B05	
Gl56.3A	Gl56.3B	5394 ± 47	-	-	0.00 ± 0.10	COR	S08CAL
Gl81.1A	Gl81.1B	5332 ± 22	3.90 ± 0.03	0.99 ± 0.02	0.08 ± 0.02	S08	
Gl100A	Gl100C	4804 ± 81	4.82 ± 0.24	1.25 ± 0.24	-0.28 ± 0.03	New	
Gl105A	Gl105B	4910 ± 65	4.55 ± 0.14	0.77 ± 0.18	-0.19 ± 0.04	New	
Gl140.1A	Gl140.1B	4671 ± 65	4.31 ± 0.15	0.54 ± 0.31	-0.41 ± 0.04	S08	
Gl157A	Gl157B	4854 ± 71	4.75 ± 0.19	1.31 ± 0.20	-0.16 ± 0.03	New	
Gl173.1A	Gl173.1B	4888 ± 72	4.72 ± 0.16	0.97 ± 0.21	-0.34 ± 0.03	New	
Gl211	Gl212	5293 ± 109	4.50 ± 0.21	0.79 ± 0.17	0.04 ± 0.11	B05	
Gl231.1A	Gl231.1B	5951 ± 14	4.40 ± 0.03	1.19 ± 0.01	-0.01 ± 0.01	New	
Gl250A	Gl250B	4670 ± 80	4.41 ± 0.16	0.70 ± 0.19	-0.15 ± 0.09	B05	
Gl297.2A	Gl297.2B	6461 ± 14	4.65 ± 0.02	1.74 ± 0.01	0.03 ± 0.05	New	
Gl324A	Gl324B	5283 ± 59	4.36 ± 0.11	0.87 ± 0.08	0.32 ± 0.07	B05	
Gl559A	Gl551	5857 ± 24	4.38 ± 0.04	1.19 ± 0.03	0.23 ± 0.02	New	
Gl611A	Gl611B	5214 ± 44	4.71 ± 0.06	-	-0.69 ± 0.03	SPO	
Gl653	Gl654	4723 ± 89	4.41 ± 0.24	0.52 ± 0.31	-0.62 ± 0.04	S08	
Gl666A	Gl666B	5274 ± 26	4.47 ± 0.04	0.74 ± 0.05	-0.34 ± 0.02	New	
Gl783.2A	Gl783.2B	5094 ± 66	4.31 ± 0.13	0.30 ± 0.19	-0.16 ± 0.08	B05	
Gl797A	Gl797B	5889 ± 32	4.59 ± 0.06	1.01 ± 0.06	-0.07 ± 0.04	B05	
GJ3091A	GJ3092B	4971 ± 79	4.48 ± 0.15	0.81 ± 0.22	0.02 ± 0.04	S08	
GJ3194A	GJ3195B	5860 ± 47	-	-	0.00 ± 0.10	SOP	S08CAL
GJ3627A	GJ3628B	5013 ± 47	-	-	-0.04 ± 0.10	SOP	S08CAL
NLTT34353	NLTT34357	5489 ± 19	4.46 ± 0.03	0.91 ± 0.03	-0.18 ± 0.01	New	

References. [B05] Bonfils et al. (2005); [COR] CCF [Fe/H] derived from spectra of the CORALIE Spectrograph; [S08CAL] T_{eff} calibration from Sousa et al. (2008); [S08] Sousa et al. (2008); [New] This paper; [SPO] Valenti & Fischer (2005); [SOP] CCF [Fe/H] taken from spectra of the SOPHIE Spectrograph (Bouchy & The Sophie Team 2006).

2002). We adopted that approach, rather than a standard spectroscopic analysis, because those observations were obtained with a ThAr lamp illuminating the second fiber of the spectrographs for highest radial velocity precision. The contamination of the stellar spectra by scattered ThAr light would affect stellar parameters measured through a classical spectral analysis, but is absorbed (to first order) into the calibration of the CCF equivalent width to a metallicity. That calibration is anchored onto abundances derived with the Santos et al. (2004) procedures, and has been verified to be on the same scale to within 0.01 dex (Sousa et al. 2011).

We adopt 10 [Fe/H] determinations from previous publications of our group (Bonfils et al. 2005; Sousa et al. 2008), which also used the Santos et al. (2004) methods. Finally, we take one metallicity value from Valenti & Fischer (2005). That reference derived its metallicities through full spectral synthesis, and Sousa et al. (2008) find that they are on the same scale as Santos et al. (2004).

Table 1 lists the **adopted** stellar parameters from high - resolution spectra of the primaries. Table 2 lists parallaxes and photometry for the full sample, along with their respective references.

3. Evaluating the photometric metallicity calibrations

To assess the three alternative photometric calibrations, we evaluate d the mean and the dispersion of the difference between the spectroscopic metallicities of the primaries and the metallicities ~~which~~ that each calibration predicts for the M dwarf components. ~~Like~~ As in previous works (Schlaufman & Laughlin 2010; Rojas-Ayala et al. 2010), we also compute d the residual mean square RMS_p and the squared multiple correlation coefficient R^2_{ap} (Hocking 1976).

The f RMS_p is defined as

$$RMS_p = \frac{SSE_p}{n-p}, \qquad SSE_p = \sum (y_{i,model} - y_i)^2, \tag{1}$$

where SSE_p is the sum of squared residuals for a p-term model, n is the number of data points, and p is the number of free parameters of the model. R^2_{ap} is defined as

$$R^2_{ap} = 1 - (n-1)\frac{RMS_p}{SST}, \qquad SST = \sum (y_i - \bar{y})^2. \tag{2}$$

No indent

A low RMS$_p$ means that the model describes the data well, while R^2_{ap} close to 1 signifies that the tested model explains most of the variance of the data. R^2_{ap} can take negative values, when the model under test increases the variance over a null model.

We recall that p should be set to the number of adjusted parameters when a model is adjusted, but instead is zero when a preexisting model is evaluated against independent data. We are, somewhat uncomfortably, in an intermediate situation, with 11, 2, and 12 binary systems in common with the samples ~~which~~ that define the calibrations of Bonfils et al. (2005), Johnson & Apps (2009), and Schlaufman & Laughlin (2010), and some measurements for those systems in common. Our sample therefore is not fully independent, and in full rigor p should take some effective value between zero and the number of parameters in the model. Fortunately, that number, 2 for all three calibrations, is a small fraction of the sample size, 23. The choice of any effective p between 0 and 2 therefore has little impact on the outcome. We **present results** for $p = 0$, except when adjusting an update of the Schlaufman & Laughlin (2010) calibration to the full sample, where we use $p = 2$ as we should.

Table 2. Sample of wide binaries with a FGK primary and a M dwarf secondary. Columns 1 and 3 display the names of the primary and secondary stars, while columns 2 and 4 display their respective spectral types. Column 5 lists the HIPPARCOS parallaxes of the primaries with their associated standard errors. Columns 6 to 11 contain the V(RI)$_C$JHK$_S$ photometry and associated errors. Column 12 contains the bibliographic references for the photometry.

Primary	Sp. Type. primary	Secondary	Sp. Type secondary	π [mas]	V [mag]	R [mag]	I [mag]	J [mag]	H [mag]	K$_s$ [mag]	V/RI/JHK source
Gl53.1A	K4	Gl53.1B	M3	48.20 ± 1.06	13.60 ± 0.02	12.48 ± 0.05	11.01 ± 0.05	9.533 ± 0.039	8.927 ± 0.023	8.673 ± 0.024	W93 / W93 / 2MASS
Gl56.3A	K1	Gl56.3B	K7	37.75 ± 0.95	10.70 ± 0.02	09.84 ± 0.03	9.01 ± 0.03	8.012 ± 0.021	7.369 ± 0.029	7.190 ± 0.020	B90 / B90 / 2MASS
Gl81.1A	G9	Gl81.1B	K7	30.44 ± 0.60	11.20 ± 0.01	10.30 ± 0.01	9.41 ± 0.01	8.413 ± 0.023	7.763 ± 0.021	7.597 ± 0.027	C84 / C84 / 2MASS
Gl100A	K4.5	Gl100C	M2.5	51.16 ± 1.33	12.85 ± 0.01	11.79 ± 0.01	10.43 ± 0.01	9.181 ± 0.027	8.571 ± 0.029	8.347 ± 0.021	C84 / C84 / 2MASS
Gl105A	K3	Gl105B	M4	139.27 ± 0.45	11.66 ± 0.02	10.45 ± 0.05	8.87 ± 0.05	7.333 ± 0.018	6.793 ± 0.038	6.574 ± 0.020	W93 / W93 / 2MASS
Gl140.1A	K3.5	Gl140.1B	K8	51.95 ± 1.16	10.17 ± 0.01	- ± -	- ± -	7.436 ± 0.023	6.828 ± 0.023	6.620 ± 0.040	S96 / - / 2MASS
Gl157A	K4	Gl157B	M2	64.40 ± 1.06	11.61 ± 0.03	- ± -	- ± -	7.773 ± 0.024	7.162 ± 0.033	6.927 ± 0.031	U74 / - / 2MASS
Gl173.1A	K3	Gl173.1B	M3	32.69 ± 1.51	14.19 ± 0.02	13.05 ± 0.05	11.65 ± 0.05	10.263 ± 0.022	9.715 ± 0.028	9.421 ± 0.024	W93 / W93 / 2MASS
Gl211	K1	Gl212	M0	81.44 ± 0.54	09.76 ± 0.01	8.81 ± 0.05	7.76 ± 0.05	6.586 ± 0.021	5.963 ± 0.016	5.759 ± 0.016	HIP / W93 / 2MASS
Gl231.1A	G0	Gl231.1B	M3.5	51.95 ± 0.40	13.27 ± 0.02	12.15 ± 0.05	10.62 ± 0.05	9.088 ± 0.023	8.559 ± 0.042	8.267 ± 0.018	WT81 / WT81 / 2MASS
Gl250A	K3	Gl250B	M2	114.94 ± 0.86	10.08 ± 0.01	09.04 ± 0.01	7.80 ± 0.01	6.579 ± 0.034	5.976 ± 0.055	5.723 ± 0.036	L89 / L89 / 2MASS
Gl297.2A	F6.5	Gl297.2B	M2	44.68 ± 0.30	11.80 ± 0.02	- ± -	- ± -	8.276 ± 0.019	7.672 ± 0.027	7.418 ± 0.016	R04 / - / 2MASS
Gl324A	G8	Gl324B	M4	81.03 ± 0.75	13.16 ± 0.01	11.94 ± 0.05	10.27 ± 0.05	8.560 ± 0.027	7.933 ± 0.040	7.666 ± 0.023	D88 / WT77 / 2MASS
Gl559A	G2	Gl551	M6	772.33 ± 2.42	11.05 ± 0.02	9.43 ± 0.03	7.43 ± 0.03	5.357 ± 0.023	4.835 ± 0.057	4.31 ± 0.03	B90 / B90 / 2MASS+B91
Gl611A	G8	Gl611B	M4	68.87 ± 0.33	14.23 ± 0.02	13.00 ± 0.05	11.38 ± 0.05	9.903 ± 0.021	9.453 ± 0.021	9.159 ± 0.017	W96 / W96 / 2MASS
Gl653	K5	Gl654	M2	93.40 ± 0.94	10.07 ± 0.01	9.10 ± 0.01	7.95 ± 0.01	6.780 ± 0.029	6.193 ± 0.021	5.975 ± 0.026	K02 / K02 / 2MASS
Gl666A	G8	Gl666B	M0	113.61 ± 0.69	08.70 ± 0.01	- ± -	- ± -	7.237 ± 9.999	5.112 ± 0.023	4.856 ± 0.020	E79 / - / 2MASS
Gl783.2A	K1	Gl783.2B	M4	49.04 ± 0.65	14.06 ± 0.02	12.81 ± 0.03	11.20 ± 0.03	9.627 ± 0.018	9.108 ± 0.015	8.883 ± 0.018	DS92 / DS92 / 2MASS
Gl797A	G5	Gl797B	M2.5	47.65 ± 0.76	11.87 ± 0.01	- ± -	- ± -	8.160 ± 0.020	7.645 ± 0.023	7.416 ± 0.016	D82 / - / 2MASS
GJ3091A	K2	GJ3092B	M	33.83 ± 1.00	15.64 ± 0.03	13.81 ± 0.05	11.97 ± 0.05	11.092 ± 0.023	10.540 ± 0.026	10.266 ± 0.021	P82 / E76 / 2MASS
GJ3194A	G4	GJ3195B	M3	41.27 ± 0.58	12.55 ± 0.02	11.49 ± 0.05	10.15 ± 0.05	8.877 ± 0.021	8.328 ± 0.023	8.103 ± 0.029	W96 / W96 / 2MASS
GJ3627A	G5	GJ3628B	M3.5	38.58 ± 2.17	14.10 ± 0.03	12.88 ± 0.05	11.31 ± 0.05	9.828 ± 0.022	9.247 ± 0.021	9.015 ± 0.018	W88 / W88 / 2MASS
NLTT34353	G5	NLTT34357	K7	20.73 ± 1.05	12.41 ± 0.02	11.51 ± 0.03	10.59 ± 0.03	9.595 ± 0.026	8.910 ± 0.026	8.734 ± 0.019	R89 / R89 / 2MASS

References. [2MASS] Skrutskie et al. (2006); [B90] Bessell (1990); [B91] Bessell (1991); [C84] Caldwell et al. (1984); [D82] Dahn et al. (1982); [D88] Dahn et al. (1988); [DS92] - Dawson & Forbes (1992); [E76] Eggen (1976); [E79] - Eggen (1979); [HIP] van Leeuwen (2007); [K02] Koen et al. (2002); [L89] Laing (1989); [P82] Pesch (1982); [R04] Reid et al. (2004); [R89] Ryan (1989); [S96] Sinachopoulos & van Dessel (1996); [U74] Upgren (1974); [W88] Weis (1988); [W93] Weis (1993); [W96] Weis (1996); [WT77] Weistrop (1977); [WT81] Weistrop (1981).

We evaluate the uncertainties on the offset, dispersion, RMS_p, and R^2_{ap} through bootstrap resampling. We generated 100,000 virtual samples with the size of our observed sample by random drawing elements of our sample, with repetition. We computed the described parameters for each virtual sample, and used their standard deviation ~~as~~ to estimate ~~of~~ s the uncertainties. ok?

Table 3 displays the defining equations of the various calibrations, their mean offset for our sample, the dispersion around the mean value (rms), the residual mean square (RMS_p), the square of the multiple correlation coefficient (R^2_{ap}), as well as their uncertainties. The M_K from the B05 calibration is the absolute magnitude calculated with the K_S photometric magnitudes and the HIPPARCOS parallaxes. The ΔM from the B05(2) calibration is the difference between the V- and the K-band mass-luminosity relations of Delfosse et al. (2000). In the JA09 calibration, the ΔM_K is the difference between the mean value of [Fe/H] of the main sequence FGK stars from the Valenti & Fischer (2005) catalogue (as defined by a fifth-order polynomial $MS = \sum a_i(V-K_S)^i$, where a = {−9.58933, 17.3952, −8.88365, 2.22598, −0.258854, 0.0113399}), and the absolute magnitude in the K_S band. The $\Delta(V-K)$ in the SL10 and 'This paper' calibrations is the difference between the observed $V - K_S$ color and the ~~5th~~ fifth- order polynomial function of M_{K_S} adapted from the previously mentioned formula from Johnson & Apps (2009). In this case, the coefficients of the polynomial are, in increasing order: (51.1413, −39.3756, 12.2862, −1.83916, 0.134266, −0.00382023).

The upper panel of Fig. 1 depicts the different [Fe/H] calibrations from Bonfils et al. (2005) (a and b), Johnson & Apps (2009) (c), Schlaufman & Laughlin (2010) (d), and the calibration determined in this paper (e). In each panel, ~~the blue/black dots represent~~ the data points. ~~The black line depicts~~ a fit to the data except in panel (a), where the calibrated [Fe/H] is shown as isometallicity contours, and the size of the black dots corresponds to a spectroscopic [Fe/H] value. The lower subpanels show the difference between the calibrated and the spectroscopic metallicity. ~~The black dashed lines represent~~ a null difference, and th~~e red dotted line represents t~~he mean difference for that calibration. Table 4 displays the metallicity values from spectroscopy and the different calibrations, where the individual values for each star can be directly compared.

The bootstrap uncertainties of the parameters (Table 3) s show that the rms values are most robust. The R^2_{ap} parameter, ~~by~~ in contrast, has large uncertainties. With our small sample size, it therefore does not provide an effective diagnostic of the alternative models.

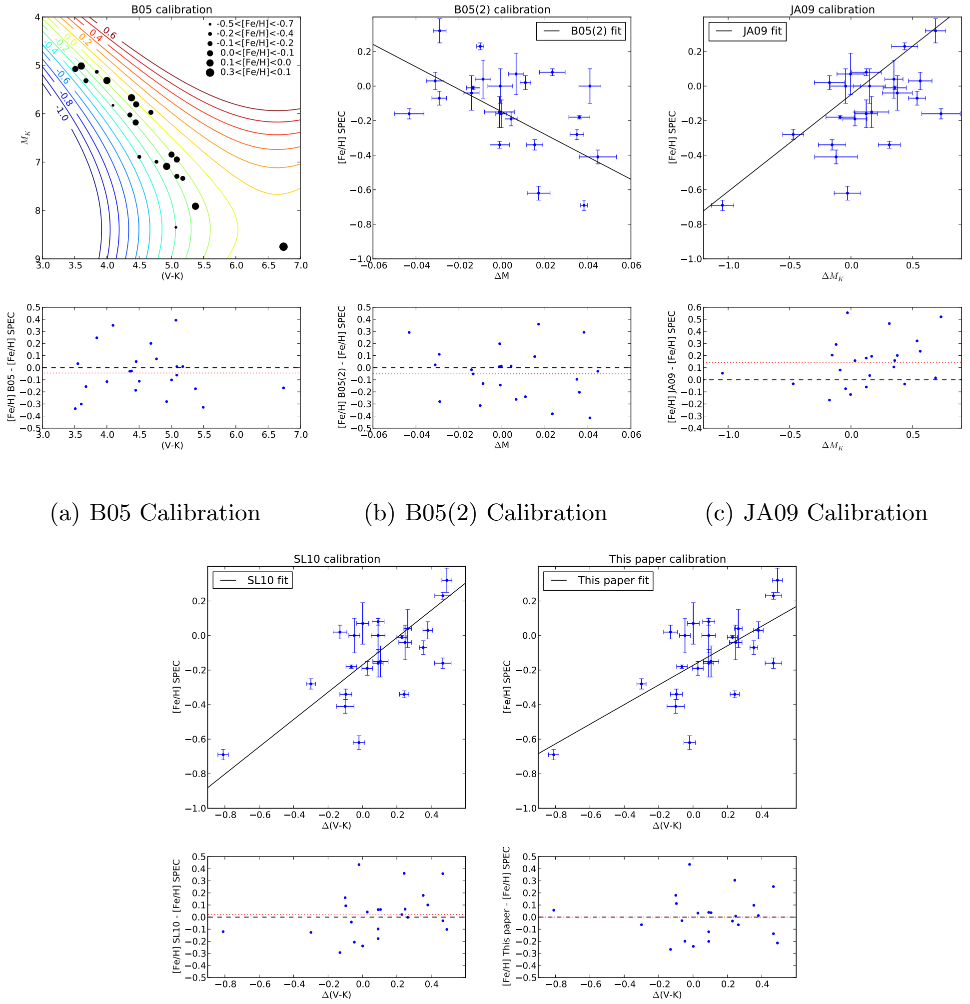
Table 3. Average offset, rms, residual mean square (RMS$_P$), and adjusted square of the multiple correlation coefficient (R$^2_{ap}$) of the calibrations of Bonfils et al. (2005), Casagrande et al. (2008), Johnson & Apps (2009), and Schlaufman & Laughlin (2010), and of this paper applied to our data.

Calibration Source + equation	offset [dex]	rms [dex]	RMS$_P$ [dex]	R$^2_{ap}$
B05 : $[Fe/H] = 0.196 - 1.527M_K + 0.091M_K^2 + 1.886(V-K) - 0.142(V-K)^2$	-0.04 ± 0.04	0.20 ± 0.02	0.04 ± 0.01	0.31 ± 0.22
B05(2) : $[Fe/H] = -0.149 - 6.508\Delta M, \Delta M = Mass_V - Mass_K$	-0.05 ± 0.04	0.22 ± 0.02	0.05 ± 0.01	0.21 ± 0.34
JA09 : $[Fe/H] = 0.56\Delta M_K - 0.05, \Delta M_K = MS - M_K$	0.14 ± 0.04	0.24 ± 0.04	0.06 ± 0.02	0.03 ± 0.51
SL10 : $[Fe/H] = 0.79\Delta(V-K) - 0.17, \Delta(V-K) = (V-K)_{obs} - (V-K)_{iso}$	0.02 ± 0.04	0.19 ± 0.03	0.04 ± 0.01	0.41 ± 0.29
This paper : $[Fe/H] = 0.595\Delta(V-K) - 0.158$	0.00 ± 0.04	0.17 ± 0.03	0.03 ± 0.01	0.43 ± 0.23

Table 4. Spectroscopic metallicity of the primaries and metallicities predicted for the secondary by the calibrations of Bonfils et al. (2005), Johnson & Apps (2009), Schlaufman & Laughlin (2010), and of this paper.

as above

Primary	Secondary	[Fe/H] [dex]					
		Spectroscopic	B05	B05(2)	JA09	SL10	This paper
Gl53.1A	Gl53.1B	0.07	-0.21	-0.19	-0.05	-0.17	-0.17
Gl56.3A	Gl56.3B	0.00	-0.34	-0.42	-0.07	-0.21	-0.20
Gl81.1A	Gl81.1B	0.08	-0.22	-0.30	0.02	-0.10	-0.12
Gl100A	Gl100C	-0.28	-0.39	-0.38	-0.31	-0.41	-0.34
Gl105A	Gl105B	-0.19	-0.18	-0.18	-0.03	-0.15	-0.15
Gl140.1A	Gl140.1B	-0.41	-0.38	-0.44	-0.12	-0.25	-0.23
Gl157A	Gl157B	-0.13	0.04	0.13	0.36	0.20	0.10
Gl173.1A	Gl173.1B	-0.33	-0.27	-0.25	-0.14	-0.25	-0.23
Gl211	Gl212	0.04	-0.08	-0.09	0.15	0.04	-0.02
Gl231.1A	Gl231.1B	0.01	-0.11	-0.06	0.15	0.01	-0.04
Gl250A	Gl250B	-0.15	-0.18	-0.14	0.04	-0.09	-0.11
Gl297.2A	Gl297.2B	0.03	0.00	0.05	0.27	0.13	0.05
Gl324A	Gl324B	0.32	-0.01	0.04	0.34	0.22	0.11
Gl559A	Gl551	0.23	0.06	-0.08	0.19	0.20	0.10
Gl611A	Gl611B	-0.69	-0.30	-0.40	-0.64	-0.81	-0.64
Gl653	Gl654	-0.62	-0.27	-0.26	-0.07	-0.19	-0.18
Gl666A	Gl666B	-0.34	-0.09	-0.14	0.12	0.02	-0.03
Gl783.2A	Gl783.2B	-0.16	-0.15	-0.15	0.02	-0.10	-0.12
Gl797A	Gl797B	-0.07	-0.02	0.04	0.25	0.11	0.03
GJ3091A	GJ3092B	0.02	-0.15	-0.22	-0.15	-0.27	-0.25
GJ3194A	GJ3195B	0.00	-0.19	-0.14	0.04	-0.10	-0.12
GJ3627A	GJ3628B	-0.04	-0.10	-0.06	0.16	0.03	-0.03
NLTT34353	NLTT34357	-0.18	-0.34	-0.38	-0.10	-0.22	-0.21



(a) B05 Calibration (b) B05(2) Calibration (c) JA09 Calibration

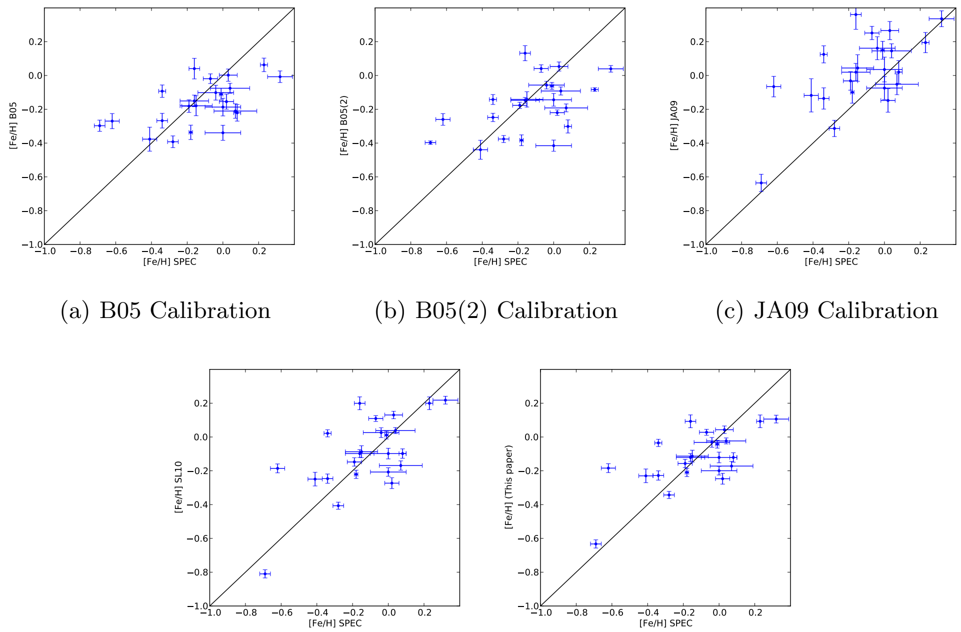
(d) SL10 Calibration (e) This paper

(title)

Fig. 1. The upper panel depicts the different [Fe/H] calibrations from Bonfils et al. (2005) (a and b), Johnson & Apps (2009) (c), Schlaufman & Laughlin (2010) (d), and the calibration determined in this paper (e). In each panel, the blue/black dots represent the data points. The black line depicts a fit to the data except in panel (a), where the calibrated [Fe/H] is shown as isometallicity contours. The lower subpanels show the difference between the calibrated and the spectroscopic metallicity. The black dashed lines represent no ~~a null~~ difference, and the red dotted line represents the mean difference for that calibration.) ok ?

4. The latest metallicity measurements and calibrations

In this section we 'll ~~will~~ discuss the three photometric metallicity calibrations in turn, and examine their agreement with our spectroscopic sample. Fig. we 2 plots the [Fe/H] obtained from each calibration against the spectroscopic [Fe/H], and it ~~will~~ guides us **through** that discussion.



(a) B05 Calibration (b) B05(2) Calibration (c) JA09 Calibration

(d) SL10 Calibration (e) This paper

Fig. 2. [Fe/H] estimated from the the calibrations versus spectroscopic metallicity.The blue dots with error bars represent the data points. The black line depicts a one-to-one relationship.

4.1. Bonfils et al. (2005) calibration

As recalled in the introduction, B05 first calibrated position in a $\{(V-K)-M_K\}$ color-magnitude diagram into a useful metallicity indicator. That calibration is anchored, on the one hand, on spectroscopic metallicity measurement of early metal-poor M-dwarfs by Woolf & Wallerstein (2005), and on the other, on later and more metal-rich M dwarfs which belong in multiple systems for which B05 measured the metallicity of a hotter component. The B05 calibration has a ∼0.2 dex dispersion. Then, they used the calibration to measure the metallicity distribution of a volume-limited sample of 47 M dwarfs, which they found to be more metal-poor (by 0.07 dex[1]) than ~~compared to~~ 1000 FGK stars, with a modest significance of 2.6 σ. As mentioned above, Bonfils et al. (2007) used that calibration to compare M dwarfs with and without planets, and found that planet hosts are marginally metal-rich.

For our sample, the B05 calibration is offset by -0.04 ± 0.04 dex and has a dispersion of 0.20±0.02 dex. The negative offset is in line with SL10 finding (see Section 4.3) that B05 generally underestimates the true [Fe/H]. Correcting from this -0.04 offset almost eliminates the metallicity difference between local M dwarfs and FGK stars.

SL10 also report that the B05 calibration has a very poor R^2_{ap}, under 0.05, and that their own model explains almost an order of magnitude more of the variance of their

[1] erroneously quoted as a 0.09 dex difference in Johnson & Apps (2009)

calibration sample. In Sect. 3, we noted, however, that R^2_{ap} is a noisy diagnostic for small samples.

In addition to their more commonly used calibration, B05 provide an alternative formulation for [Fe/H]. That second expression, labeled B05(2) in Table 3, works from the difference between the V- and K-band mass-luminosity relations of Delfosse et al. (2000). The two B05 formulations perform essentially equally for our sample, with B05(2) having a marginally higher dispersion. In the remain~~ing~~der of this paper we therefore no longer discuss B05(2). (OR: rest)

4.2. Johnson & Apps (2009) calibration

Johnson & Apps (2009) argue~~d~~s that local M and FGK dwarfs should have the same metallicity distribution, and accordingly chose to fix their mean M dwarf metallicity to the value (−0.05 dex) for a volume-limited sample of FGK dwarfs from the (Valenti & Fischer 2005) sample. They defined a sequence representative of average M dwarfs in the $\{(V-K)-M_K\}$ color-magnitude diagram, and used the distance to that main sequence along M_K as a metallicity diagnostic. They note~~d~~s that the inhomogeneous calibration sample of B05 is a potential source of systematics, and consequently chose to calibrate their scale from the metallicities of just ~~6~~ six metal-rich M dwarfs in multiple systems with FGK primary components.

JA09 present~~ed~~s two observational arguments for fixing the mean M dwarf metallicity. They first measured [Fe/H] for 109 G0-K2 stars (4900<T$_{eff}$<5900 K) and found no significant metallicity gradient over this temperature range, from which they conclude that no diference is to be expected for the cooler M dwarfs. We note, however, that a linear fit to their G0-K2 data set ([Fe/H]$= 9.74 \times 10^{-5}(T_{\rm eff} - 5777) - 0.04$) allows for a wide metallicity range when extrapolated to the cooler M dwarfs (2700<Teff<3750, for M7 to M0 spectral type, with [Fe/H] $= -0.24$ allowed at the 1 σ level for an M0 dwarf and signficantly lower than the [Fe/H] offset in B05. More importantly, they measured a large (0.32 dex) offset between the B05 metallicities of ~~6~~ six metal-rich M dwarfs in multiple systems and the spectroscopic metallicities which they measured for their primaries. This robustly points to a systematic offset in the B05 calibration for metal-rich M dwarfs, but does not directly probe the rest of the (T$_{eff}$, [Fe/H]) space. We do find that the JA09 calibration is a good metallicity predictor for our sample at high metallicities, where its calibrator ~~were~~ as chosen. With decreasing metallicity, that calibration however, increasingly overestimates the metallicity, as previously pointed out by SL10 (see below). Quantitatively, we measure for our sample a $+0.14 \pm 0.04$ dex offset and a dispersion of 0.24 ± 0.04.

4.3. Schlaufman & Laughlin (2010) calibration

Schlaufman & Laughlin (2010) improve upon B05 and JA09 in two ways. They first point out that, for M and FGK dwarfs to share the same mean metallicity, matched kinematics is as important as volume completeness. Since the various kinematic populations of our Galaxy have very different mean metallicities, the mean metallicity of small samples fluctuates very significantly with their small number of stars from the metal-poor populations. To overcome this statistical noise, they draw from the Geneva-Copenhagen Survey volume-limited sample of F and G stars a subsample that kinematically matches the volume limited sample of M dwarfs used by JA09. They find a $\simeq -0.14 \pm 0.06$ dex mean metallicity for that sample, 0.09 dex lower than adopted by JA09. However, they only used that sample to verify ~~the fact~~ that the mean metallicity of M dwarfs in the solar neighborhood is well defined. Finally, the M dwarfs within a sample of binaries with a FGK primary that they used to fix their calibration are not volume-limited or kinematically-matched, but their mean metallicity ([Fe/H] $= -0.17 \pm 0.07$) is statistically indistinguishable from the mean metallicity of the volume-limited and kinematically-matched sample.

?

Secondly, they use stellar evolution models to guide their parametrization of the color-magnitude space. Using [Fe/H] isocontours for the Baraffe et al. (1998) models, they show that in a $\{(\mathrm{V-K})-\mathrm{M}_K\}$ diagram, changing [Fe/H] affects (V−K) at an essentially constant M_K. [Fe/H] is therefore best parametrized by (V−K), and their calibration uses a linear function of the (V−K) distance from a nominal sequence in the $\{(\mathrm{V-K})-\mathrm{M}_K\}$ diagram. They do not force any specific mean metallicity, but verify *a posteriori* that it matches expectations.

We measure a 0.14 ± 0.02 dex dispersion for the SL10 sample against their calibration, but that calibration has a significantly higher dispersion of 0.19 ± 0.03 for our validation sample. That increased dispersion reflects our sample probing a wider metallicity range than SL10, as verified by computing the dispersion of an 18 stars sub-sample that matches the metallicity range of the SL10 sample. That dispersion is 0.15 ± 0.03 dex, and indistinguishable from 0.14 ± 0.02 dex for the SL10 sample. The increased dispersion for a wider metallicity range suggests that a linear function of (V−K) does not fully describe metallicity. We also measure an offset of 0.02 ± 0.04 dex. Offset and rms both improve over either of the B05 and JA10 calibrations.

4.4. Refining the Schlaufman & Laughlin (2010) calibration

We produced updated coefficients for the SL10 prescription, using the RMS_p free parameter $p = 2$ (see Section 3). The expression for the new metallicity calibration is

$$[Fe/H] = 0.57\Delta(V - K_s) - 0.17, \tag{3}$$
$$\Delta(V - K_s) = (V - K_s)_{obs} - (V - K_s)_{iso},$$

where $(V-K_s)_{obs}$ is the observed $V-K_s$ color and $(V-K_s)_{iso}$ is a 5^{th} fifth- order polynomial function of M_{K_s} that describes the mean main sequence of the lc Solar neighborhood from the Valenti & Fischer (2005) catalog. This expression is adopted from Schlaufman & Laughlin (2010), who adapted a n M_{K_s} vs $(V-K_s)$ formula from Johnson & Apps (2009).

Table 3 shows limited differences between this new fit and the original SL10 calibration. The dispersion of the new fit is tighter by just 0.02 dex (0.17 ± 0.03 dex instead of 0.19 ± 0.03), and the offset is now 0.00 ± 0.04, as expected. The R^2_{ap} value ~~are~~ is similar (0.43 ± 0.23 vs 0.41 ± 0.29), and uncertain. Readjusting the coefficients therefore produces at best a marginal improvement.

The dispersion, in all panels of Fig. 1, is well above the measurement uncertainties. Those therefore contribute negligibly to the overall dispersion, which must be dominated by other sources.

As can be seen in Fig. 2, B05 or B05(2) tend to underestimate [Fe/H], while the JA09 calibration clearly overestimates [Fe/H] except at the highest metallicities.

5. Summary

We have assembled a sample of M dwarf companions to hotter FGK stars, where the system has an accurate parallax and the M dwarf component has accurate V and K_S-band photometry. Using the metallicities of the primaries, newly measured or retrieved from the literature, and the assumption that the two components have identical initial compositions, we compared the dispersions of the Bonfils et al. (2005), Johnson & Apps (2009), and Schlaufman & Laughlin (2010) photometric metallicity calibrations. We find that the Schlaufman & Laughlin (2010) scale, which is intermediate between Bonfils et al. (2005) and Johnson & Apps (2009), has the lowest dispersion. We slightly refine that relation, by readjusting its coefficients on our sample.

We find that our tight selection of binaries with accurate parallaxes and photometry sample has insignificantly reduced the dispersion of the measurements around the calibration compared to looser criteria. This suggests that the dispersion, ~~and therefore~~ hence the random errors of the calibration, is not defined by measurement uncertainties but ~~rather~~ instead reflects intrisic astrophysical dispersion. Non-linearities in the metallicity dependence of the $V - K_s$ color are likely to contribute, as suggested both by atmospheric models (Allard, private communication) and by the increased dispersion ~~which~~ that we measure over a wider metallicity range. They are however unlikely to be the sole explanation, since we see dispersion even in narrow areas of the color-magnitude diagram. Stellar evolution cannot significantly contribute, since early-M dwarfs evolve rapidly to the main sequence, and they remain there for much longer than a Hubble time, but rotation and magnetic activity could play a role. Unless, or until, we develop a quantitative understanding of this astrophysical dispersion, the photometric calibration approach may therefore have reached an intrinsic limit. Those calibrations also have the very practical inconvenient ce of needing an accurate parallax. This limits their use to the close lc Solar neighborhood, at least until the GAIA catalog becomes available in a decade.

Alternative probes of the metallicities of M dwarfs are therefore obviously desirable. One obvious avenue is to work from higher spectral resolution informations, and to identify spectral elements that are most sensitive to metallicity and others ~~which~~ that are most sensitive to effective temperature. We are pursuing this approach at visible wavelengths (Neves et al. in prep.), as do Rojas-Ayala et al. (2010, see Appendix A.2) in the near infra-red, with encouraging results in both cases.

Acknowledgements. We would like to thank Luca Casagrande for kindly providing ~~us~~ s the metallicities calculated from his calibration. We acknowledge the support by the European Research Council/European Community under the FP7 through Starting Grant agreement number 239953. NCS also acknowledges the support from Fundação para a Ciência e a Tecnologia (FCT) through program Ciência 2007 funded by FCT/MCTES (Portugal) and POPH/FSE (EC), and in the form of grant reference PTDC/CTE-AST/098528/2008. VN would also like to acknowledge the support from the FCT in the form of the fellowship SFRH/BD/60688/2009.

References

Allard, F., Homeier, D., & Freytag, B. 2010, ArXiv e-prints. ID: 1011.5405. To appear in the proceedings of Cool Stars 16

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARAA, 47, 481

Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, A&A, 337, 403

Bean, J. L., Benedict, G. F., & Endl, M. 2006, ApJ, 653, L65

Bessell, M. S. 1990, AAPS, 83, 357

Bessell, M. S. 1991, AJ, 101, 662

Blackwell, D. E. & Shallis, M. J. 1977, MNRAS, 180, 177

Bonfils, X., Delfosse, X., Udry, S., et al. 2005, A&A, 442, 635

Bonfils, X., Mayor, M., Delfosse, X., et al. 2007, A&A, 474, 293

Bouchy, F., Mayor, M., Lovis, C., et al. 2009, A&A, 496, 527

Bouchy, F. & The Sophie Team. 2006, in Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies, ed. L. Arnold, F. Bouchy, & C. Moutou, 319–325

Caffau, E., Ludwig, H.-G., Steffen, M., Freytag, B., & Bonifacio, P. 2011, SOLPHYS, 268, 255

Caldwell, J. A. R., Spencer Jones, J. H., & Menzies, J. W. 1984, MNRAS, 209, 51

Carpenter, J. M. 2001, AJ, 121, 2851

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

Casagrande, L., Portinari, L., & Flynn, C. 2006, MNRAS, 373, 13

Chabrier, G. & Baraffe, I. 2000, ARAA, 38, 337

Chanamé, J. & Gould, A. 2004, ApJ, 601, 289

Dahn, C. C., Harrington, R. S., Kallarakal, V. V., et al. 1988, AJ, 95, 237

Dahn, C. C., Harrington, R. S., Riepe, B. Y., et al. 1982, AJ, 87, 419

Dawson, P. C. & Forbes, D. 1992, AJ, 103, 2063

Delfosse, X., Forveille, T., Ségransan, D., et al. 2000, A&A, 364, 217

Eggen, O. J. 1976, ApJS, 30, 351

Eggen, O. J. 1979, ApJS, 39, 89

Endl, M., Cochran, W. D., Kürster, M., et al. 2006, ApJ, 649, 436

Fischer, D. A. & Valenti, J. 2005, ApJ, 622, 1102

Gliese, W. & Jahreiß, H. 1991, Preliminary Version of the Third Catalogue of Nearby Stars, Tech. rep.

Gonzalez, G. 1997, MNRAS, 285, 403

Gould, A. & Chanamé, J. 2004, ApJS, 150, 455

Hocking, R. R. 1976, Biometrics, 32, 1

Johnson, J. A. & Apps, K. 2009, ApJ, 699, 933

Johnson, J. A., Butler, R. P., Marcy, G. W., et al. 2007, ApJ, 670, 833

Kaufer, A. & Pasquini, L. 1998, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3355, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. S. D'Odorico, 844–854

Koen, C., Kilkenny, D., van Wyk, F., Cooper, D., & Marang, F. 2002, MNRAS, 334, 20

Kurucz, R. 1993, ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993., 13

Laing, J. D. 1989, South African Astronomical Observatory Circular, 13, 29

Leggett, S. K. 1992, ApJS, 82, 351

Mayor, M., Bonfils, X., Forveille, T., et al. 2009, A&A, 507, 487

Mermilliod, J., Mermilliod, M., & Hauck, B. 1997, A&AS, 124, 349

Pesch, P. 1982, PASP, 94, 345

Poveda, A., Herrera, M. A., Allen, C., Cordero, G., & Lavalley, C. 1994, Rev. Mexicana Astron. Astrofis., 28, 43

Queloz, D., Mayor, M., Weber, L., et al. 2000, A&A, 354, 99

Reid, I. N., Cruz, K. L., Allen, P., et al. 2004, AJ, 128, 463

Rojas-Ayala, B., Covey, K. R., Muirhead, P. S., & Lloyd, J. P. 2010, ApJ, 720, L113

Ryan, S. G. 1989, AJ, 98, 1693

Santos, N. C., Israelian, G., & Mayor, M. 2001, A&A, 373, 1019

Santos, N. C., Israelian, G., & Mayor, M. 2004, A&A, 415, 1153

Santos, N. C., Mayor, M., Naef, D., et al. 2002, A&A, 392, 215

Schlaufman, K. C. & Laughlin, G. 2010, A&A, 519, A105+

Sinachopoulos, D. & van Dessel, E. L. 1996, AAPS, 119, 483

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Sneden, C. 1973, Ph.D. Thesis, Univ. of Texas

Sousa, S. G., Santos, N. C., Israelian, G., et al. 2011, A&A, 526, A99+

Sousa, S. G., Santos, N. C., Israelian, G., Mayor, M., & Monteiro, M. J. P. F. G. 2007, A&A, in press

Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, A&A, 487, 373

Udry, S., Bonfils, X., Delfosse, X., et al. 2007, A&A, 469, L43

Udry, S. & Santos, N. 2007, ARAA, 45, 397

Upgren, A. R. 1974, PASP, 86, 294

Valenti, J. A. & Fischer, D. A. 2005, VizieR Online Data Catalog, 215, 90141

Valenti, J. A., Piskunov, N., & Johns-Krull, C. M. 1998, ApJ, 498, 851

van Leeuwen, F. 2007, A&A, 474, 653

Weis, E. W. 1988, AJ, 96, 1710

Weis, E. W. 1993, AJ, 105, 1962

Weis, E. W. 1996, AJ, 112, 2300

Weistrop, D. 1975, PASP, 87, 367

Weistrop, D. 1977, ApJ, 215, 845

Weistrop, D. 1981, AJ, 86, 1220

Woolf, V. M. & Wallerstein, G. 2005, MNRAS, 356, 963

Woolf, V. M. & Wallerstein, G. 2006, PASP, 118, 218

Appendix A: Other methods

A.1. Calibration of Casagrande et al. (2008)

In ~~section~~ Sect. 4 we described in detail the photometric metallicity calibrations. Casagrande et al. (2008) devised a completely different technique, based on **their** previous study of FGK stars using the infrared flux method (Casagrande et al. 2006), to determine effective temperatures and metallicities. The infrared flux method uses multiple photometry bands to derive effective temperatures, bolometric luminosities, and angular diameters. The basic idea of IRFM (Blackwell & Shallis 1977) is to compare the ratio between the bolometric flux and and infrared monochromatic flux, both measured ~~at~~ on Earth, to the ratio between the surface bolometric flux ($\propto \sigma Teff^4$) and the surface infrared monochromatic flux for a model of the star. To adapt this method to M dwarfs, Casagrande et al. (2008) added optical bands, creating the so-called MOITE, Multiple Optical and Infrared TEchnique. This method provides sensitive indicators of both temperature and metallicity. The proposed effective temperature scale extends down to 2100-2200 K, into the L-dwarf limit, and is supported by interferometric angular diameters above $\sim$ 3000K.

Table A.1. Metallicity values from spectroscopy and obtained using the method of Casagrande et al. (2008) (C08 in this Table).

Primary	Secondary	[Fe/H] [dex]	
		Spectroscopic	C08
Gl53.1A	Gl53.1B	0.07	-0.07
Gl56.3A	Gl56.3B	0.00	-0.21
Gl81.1A	Gl81.1B	0.08	-0.08
Gl100A	Gl100C	-0.28	-0.10
Gl105A	Gl105B	-0.19	-0.30
Gl140.1A	Gl140.1B	-0.41	-0.30
Gl157A	Gl157B	-0.13	-0.10
Gl173.1A	Gl173.1B	-0.33	-0.20
Gl211	Gl212	0.04	-0.21
Gl231.1A	Gl231.1B	0.01	-0.28
Gl250A	Gl250B	-0.15	-
Gl297.2A	Gl297.2B	0.03	0.00
Gl324A	Gl324B	0.32	-0.20
Gl559A	Gl551	0.23	-
Gl611A	Gl611B	-0.69	-0.40
Gl653	Gl654	-0.62	-0.30
Gl666A	Gl666B	-0.34	-
Gl783.2A	Gl783.2B	-0.16	-0.30
Gl797A	Gl797B	-0.07	-0.90
GJ3091A	GJ3092B	0.02	-0.30
GJ3194A	GJ3195B	0.00	-0.60
GJ3627A	GJ3628B	-0.04	-0.20
NLTT34353	NLTT34357	-0.18	0.19

Table A.2. Offset, rms, residual mean square (RMS_P), and adjusted square of the multiple correlation coefficient (R^2_{ap}) of the [Fe/H] values obtained using the Casagrande et al. (2008) method.

offset [dex]	rms [dex]	RMS_P [dex]	R^2_{ap}
-0.11 ± 0.07	0.32 ± 0.06	0.10 ± 0.04	-1.09 ± 1.45

Casagrande et al. (2008) obtain metallicities by computing the effective temperature of the star for each colour band ($V(RI)_C JHK_S$) for several trial metallicities, between -2.1 and 0.4 in 0.1 dex steps, and by selecting the metallicity ~~which~~ that minimizes the scatter among the six trial effective temperatures. Casagrande et al. (2008) estimate that their total metallicity uncertainty is 0.2 to 0.3 dex.

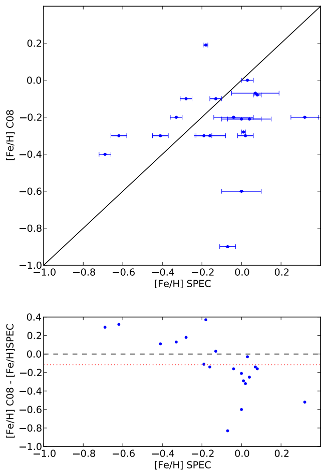


Fig. A.1. [Fe/H] obtained with the Casagrande et al. (2008) method versus the spectroscopic metallicity. The blue dots with error bars represent the data points. The black line depicts a one to one relationship. The metallicity difference between the values of the calibrations and the spectroscopic measurements is shown below each [Fe/H]-[Fe/H] plot. The black dashed line is the zero point of the difference, and the red dotted line represents the average of the metallicity difference.

(reduce ?)

The MOITE method does not summarize into a closed form that can be readily applied by others, but Luca Casagrande kindly computed MOITE [Fe/H] values for our sample (Table A.1). We test this method in Fig. A.1 and Table A.2, which show it has a higher rms and RMS_p and a poorer R^2_{ap} than the three ~~3~~ photometric calibrations, consistently with the high (?) metallicity uncertainty quoted by Casagrande et al. (2008). The negative R^2_{ap} value formally means that this model increases the variance over a constant metallicity model, but as usual R^2_{ap} is a noisy diagnostic.

A.2. Rojas-Ayala et al. (2010) calibration

Rojas-Ayala et al. (2010) have recently published a novel and potentially very precise technique for measuring M dwarf metallicities. Their technique is based on spectral indices measured from moderate-dispersion ($R \sim 2700$) K-band spectra, and it needs neither a V magnitude nor a parallax, allowing measurement of fainter (or/and farther) stars. They analysed 17 M dwarf secondaries with a FGK primary, which ~~that~~ served also as metallicity calibrators, and measured the equivalent widths of the NaI doublet (2.206 and 2.209 μm), and the CaI triplet (2.261, 2.263 and 2.265 μm). With these measurements and a water absorption spectral index sensitive to stellar temperatures, they constructed a metallicity scale with an adjusted multiple correlation coefficient greater than the one of Schlaufman & Laughlin (2010) ($R^2_{ap} = 0.63$), and also with a tighter RMS_p of 0.02 when compared to other studies (0.05, 0.04 and 0.02 for Bonfils et al. (2005), Johnson & Apps

(2009), and Schlaufman & Laughlin (2010) respectively). The metallicity calibration is valid over -0.5 to +0.5 dex, with an estimated uncertainty of ± 0.15 dex.

A test of the Rojas-Ayala et al. (2010) calibration for our full sample would be very interesting, but is not currently possible for lack of near-infrared spectra for most of the stars. Seven of our stars, however, have their metallicities measured in Rojas-Ayala et al. (2010) (Gl 212, Gl 231.1B, Gl 250B, Gl 324B, Gl611B, Gl783.2B, and Gl 797B with predicted [Fe/H] of $0.09, -0.05, -0.04, 0.30, -0.49, -0.19$, and -0.06 dex respectively). We find a dispersion of only 0.08 dex and an offset of 0.04 dex offset between our spectroscopic measurements of the primaries and the Rojas-Ayala et al. (2010) metallicities of the secondaries. These numbers are extremely encouraging, but still have little statistical significance. They will need to be bolstered by testing against a larger sample and over a wider range of both metallicity and effective temperature.