

SPECTRO-THERMOMETRY OF M DWARFS AND THEIR CANDIDATE PLANETS:  
TOO HOT, TOO COOL, OR JUST RIGHT?ANDREW W. MANN<sup>1,2</sup>, ERIC GAIDOS<sup>3</sup>, MEGAN ANSDELL<sup>1</sup>*ApJ in review, revised February 27, 2014*

## ABSTRACT

We use moderate-resolution spectra of nearby late K and M dwarf stars with parallaxes and interferometrically determined radii to refine their effective temperatures, luminosities, and metallicities. We use these revised values to calibrate spectroscopic techniques to infer the fundamental parameters of more distant late-type dwarf stars. We demonstrate that, after masking out poorly modeled regions, the newest version of the PHOENIX atmosphere models accurately reproduce temperatures derived bolometrically. We apply methods to late-type hosts of transiting planet candidates in the *Kepler* field, and calculate effective temperature, radius, mass, and luminosity with typical errors of 57 K, 7%, 11%, and 13%, respectively. We find systematic offsets between our values and those from previous analyses of the same stars, which we attribute to differences in atmospheric models utilized for each study. We investigate which of the planets in this sample are likely to orbit in the circumstellar habitable zone. We determine that four candidate planets (KOI 854.01, 1298.02, 1686.01, and 2992.01) are inside of or within  $1\sigma$  of a conservative definition of the habitable zone, but that several planets identified by previous analyses are not (e.g. KOI 1422.02 and KOI 2626.01). Only one of the four habitable-zone planets is Earth sized, suggesting a downward revision in the occurrence of such planets around M dwarfs. These findings highlight the importance of measuring accurate stellar parameters when deriving parameters of their orbiting planets.

*Subject headings:* stars: fundamental parameters — stars: late-type — planetary systems — astrobiology

## 1. INTRODUCTION

M dwarf stars have become prime targets in the search for potentially habitable planets in large part because they are much cooler and smaller than Sun-like stars, allowing smaller planets in the habitable zone (HZ) to be detected. Observations have shown that the occurrence of small planets – including Earth- and super Earth-sized planets – increases with decreasing stellar mass (Howard et al. 2012) (although also see Fressin et al. 2013), and that small planets are nearly ubiquitously around late-type stars (Swift et al. 2013). Since over 70% of stars in the solar neighborhood are M dwarfs (Henry et al. 1994; Chabrier 2003; Reid et al. 2004) it is likely that the vast majority of potentially habitable planets near our Sun orbit around late K and M dwarfs.

For a fixed planet size and radius, the Doppler signal is inversely related to stellar mass,  $M_*$ , and the planet's orbital period,  $P$  ( $\propto M_*^{-2/3} P^{-1/3}$ ), while the transit depth is inversely related to the stellar radius,  $R_*$  ( $\propto R_*^{-2}$ ). The transit probability scales inversely to the planet's orbital period ( $\propto R_* M_*^{-1/3} P^{-2/3}$ ). An early M dwarf has about half the radius and mass of a Sun-like star. A planet in the HZ (the range of stellar irradiances bracketed by the runaway greenhouse on one end and the first CO<sub>2</sub> condensation at the other) of an early M dwarf has  $\sim 1/6$ th the period of a HZ planet around a Sun-like

star (Selsis et al. 2007; Kopparapu et al. 2013b). Thus a planet in the HZ of an early M dwarf will have  $\sim 3$  times the Doppler signal,  $\sim 4$  times the transit depth, and  $\sim 2$  times the probability of transiting compared to a HZ planet around a solar-type star.

The *Kepler* mission (Borucki et al. 2010) has found more than 2000 planet candidates (Borucki et al. 2011; Batalha et al. 2013), enabling more robust studies of exoplanet statistics through large samples. Although *Kepler* target stars are mostly F, G, and early K dwarfs (Batalha et al. 2010), there are also  $\sim 4000$  *Kepler* stars cooler than 4000 K (Dressing & Charbonneau 2013), which collectively harbor more than 100 detected candidate exoplanets. This is roughly an order of magnitude more planets than have been found around M dwarfs using ground-based Doppler (Bonfils et al. 2013) or transit (Berta et al. 2013) surveys. As a result, *Kepler* has been critical in constraining the occurrence of planets around M dwarfs (Mann et al. 2012; Dressing & Charbonneau 2013; Swift et al. 2013), the habitability of planets orbiting M dwarfs (Dressing & Charbonneau 2013; Kopparapu 2013; Gaidos 2013), the metallicity distribution of M dwarfs with or without detected planets (Muirhead et al. 2012a; Mann et al. 2013b), and the mass and radius distributions of M dwarf planets (Gaidos et al. 2012; Morton & Swift 2013).

As is the case for most *Kepler* discoveries, accurate physical parameters for planets are largely limited by uncertainties its host star parameters. For example, the transit depth gives only the ratio of the planet radius to that of the star; one must know the size of the star in order to accurately determine the size of the planet. Further, whether a planet resides in the HZ depends on

<sup>1</sup> Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822

<sup>2</sup> Harlan J. Smith Fellow, Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA

<sup>3</sup> Department of Geology & Geophysics, University of Hawai'i, 1680 East-West Road, Honolulu, HI 96822

the irradiance ( $S$ ) that the planet receives from its host star, which in turn depends sensitively on the host star’s  $T_{\text{eff}}$ .

Total stellar irradiation is not the only measure of planetary habitability (Gaidos et al. 2005) and planets in the HZ of M dwarfs will, at the same total stellar irradiation, experience a different environment than their counterparts orbiting solar-type stars (Tarter et al. 2007). The spectral energy distribution of M dwarfs is distinct from that of G stars. Because the reflectivity of water (ice) surfaces is wavelength-dependent, this alters the efficacy of the ice albedo feedback that can destabilize the climates of Earth-like planets (Joshi & Haberle 2012; Shields et al. 2013). Most oxygenic photosynthesis on Earth requires light bluer than (more energetic than)  $\sim 700$  nm, and thus equivalent life around M dwarfs would have comparatively less light to harvest.

Planets around late M dwarfs with persistent magnetic and chromospheric activity will experience elevated ultraviolet (UV) emission and stellar wind particle fluxes. These could heat and erode atmospheres (Lammer et al. 2007) as well as drive atmospheric chemistry (Segura et al. 2010). Although tidally induced synchronization of the rotation of planets in the HZ of M dwarfs may not lead to atmospheric collapse on the cold night side (Joshi et al. 1997), it could lead to sequestration of less volatile substances such as water ice (Menou 2013). Finally, impacts by planetesimals will be more energetic in the HZs of M dwarfs because orbital velocities are higher; this may remove volatiles such as water from the planet altogether (Lissauer 2007). Although these concerns suggest caution in designating any planets in the HZ as actually “habitable,” they have not dampened enthusiasm for identifying such planets and estimating their occurrence (Dressing & Charbonneau 2013; Kopparapu 2013; Gaidos 2013).

Constraining the physical parameters of M dwarfs is significantly more difficult than for their warmer FGK counterparts. Their cooler temperatures enable the appearance of molecular bands (e.g., CaH, TiO) that dominate their spectra, create line confusion, and obscure the continuum level (Mould 1976; Kirkpatrick et al. 1991). The poorly constrained optical parameters and reaction constants makes these bands difficult to model, rendering spectral synthesis unreliable for M dwarfs, although improvements are ongoing (e.g., Bean et al. 2006; Önehag et al. 2012).

Nonetheless, the radii of bright, nearby M dwarfs can be measured using parallaxes and long-baseline interferometry (e.g., Demory et al. 2009; von Braun et al. 2011, 2012). Boyajian et al. (2012, henceforth B12) use the CHARA interferometer to obtain angular diameters for a set of K and M dwarfs. They utilize it Hipparcos parallaxes (van Leeuwen & Fantino 2005; van Leeuwen 2007) to determine  $R_*$  and combine photometry with template spectra to determine bolometric fluxes and infer  $T_{\text{eff}}$  with high precision (1-2% error in  $R_*$ ,  $< 1\%$  error in  $T_{\text{eff}}$ ). They derive  $L_*$  by combining the bolometric flux with the distance (parallax), and determine  $M_*$  from empirical relations between absolute  $K$ -band magnitude ( $M_K$ ) and  $M_*$  (Henry & McCarthy 1993; Delfosse et al. 2000). As a result, the B12 sample makes an ideal set of “calibration” stars to develop and calibrate techniques for

measuring the physical characteristics of late K and M dwarfs. These empirical techniques can then be applied to more distant stars beyond the reach of interferometry or precision astrometry (parallaxes).

Here we present observations and analysis of the B12 late-type stars to derive empirical relations between fundamental physical parameters and moderate-resolution spectra. We apply these relations to *Kepler* M dwarf planet hosts, (re)estimate planet radii and stellar irradiances, and determine which of their planets reside in the HZ. First we describe our samples of B12 calibration stars and *Kepler* targets in Section 2, then we summarize our moderate-resolution visible and near infrared (NIR) spectroscopy of both samples in Section 3. In Section 4 we use our spectra to re-derive bolometric fluxes, and hence physical parameters, of the B12 sample, and then derive new empirical relations between  $T_{\text{eff}}$  and several other stellar parameters. In Section 5 we calibrate the  $T_{\text{eff}}$  values based on comparing models to observed spectra. We apply our techniques to late-type *Kepler* planet candidate hosts in Section 6. We use these newly derived stellar parameters in Section 7 to calculate the irradiances and radii of the candidate planets, which we utilize to investigate which *Kepler* exoplanets orbit in the HZ. We conclude in Section 8 by briefly commenting on the complications and ramifications of our work.

In this paper we will refer to *Kepler* exoplanet candidates as planets, both for simplicity and to avoid confusion with candidate HZ planets (planets which are potentially in the HZ). For specific systems we will refer to them using the *Kepler* object of interest (KOI) number, with whole numbers referring to the host star (e.g., KOI 854), and decimals to denote the planets (e.g., KOI 854.01).

## 2. SAMPLE

For our calibration sample, we selected B12 stars with  $T_{\text{eff}} < 4800$  K and  $-45^\circ < \delta < 70^\circ$  (observable with the telescopes and instruments specified in Section 3). This cut contains 23 stars spanning  $3100 \text{ K} < T_{\text{eff}} < 4800 \text{ K}$ , which covers the expected spectral types and  $T_{\text{eff}}$  of our *Kepler* planet host sample (defined below).

Following the criterion of Mann et al. (2013b), we selected *Kepler* planet hosts with  $K_P - J > 1.85$ , excluding three suspected false positives (see below). This color cut includes all dwarfs with  $T_{\text{eff}} < 4100$  K, with some as warm as 4500 K (Mann et al. 2012), and at least one as cool as 3200 K (Muirhead et al. 2012b). We included all planet candidates listed in the NASA exoplanet archive as of 2013 February (Batalha et al. 2013).

In our *Kepler* sample, we excluded three planets that are likely to be false positives: KOI 977 is a giant star (Muirhead et al. 2012a; Mann et al. 2012), KOI 1902 is an eclipsing binary (Mann et al. 2013b), and KOI 256 is a white dwarf-M dwarf binary (Muirhead et al. 2013). The final sample includes 123 late-type dwarfs harboring 188 detected planets.

## 3. OBSERVATIONS AND REDUCTION

Optical spectra of both the calibration sample and the *Kepler* planet host sample were obtained with the SuperNova Integral Field Spectrograph (SNIFS, Aldering et al. 2002; Lantz et al. 2004) on the University of Hawaii 2.2m telescope atop Mauna Kea. SNIFS

uses a dichroic mirror to separate the incoming light onto blue (3200-5200 Å) and red (5100-9700 Å) spectrograph channels. The spectral resolution,  $R$ , is  $\simeq 800$  in the blue channel and  $\simeq 1000$  in the red. Integration times for stars in the calibration sample varied from 5 to 40 s, which was sufficient to achieve high signal-to-noise ( $S/N > 200$  per resolving element) in the red channel while avoiding the non linear regime of the detector. For *Kepler* stars, which are fainter, total integration times were between 5 minutes and 2 hr. For targets with integration times longer than 30 m, integrations were divided into three or more exposures  $\leq 30$  minutes long, and then stacked (using the weighted mean) to mitigate contamination from cosmic rays and other artifacts. The resulting  $S/N$  for *Kepler* stars is typically  $> 60$  (per resolving element) in the red channel. For three of the *Kepler* targets (KOI 2306, 1879, and 2418) we were unable to extract a reasonable spectrum from the blue channel. However, the blue channel data is not used for our analysis of the KOIs.

The SNIFS pipeline (Bacon et al. 2001; Aldering et al. 2006) performed dark, bias, and flat-field corrections and cleaned the data of bad pixels and cosmic rays. The pipeline wavelength calibrated the data based on arc lamp exposures taken at the same telescope pointing and time as the science data. Over the course of each night, we obtained spectra of the EG131, Feige 66, Feige 110, BD+284211, or BD+174708 spectrophotometric standards (Bessell 1999; Bohlin et al. 2001; Hamuy et al. 1992; Oke 1990). We combined the standard star observations with a model of the atmosphere above Mauna Kea from Buton et al. (2013) to correct each spectrum for instrument response and atmospheric extinction and to remove telluric lines. We shifted each spectrum in wavelength to the rest frames of their source stars by cross-correlating each spectrum to a spectral template of similar spectral type from the Sloan Digital Sky Survey (Stoughton et al. 2002; Bochanski et al. 2007) using the IDL routine *xcorl* (Mohanty & Basri 2003; West & Basri 2009). More details on our data reduction can be found in Mann et al. (2012) and Lépine et al. (2013).

Multiple observations of the same spectrophotometric standards demonstrate that SNIFS spectra have an error term of  $\simeq 1\%$  in addition to the expected (mostly Poisson) measurement noise (Mann et al. 2011; Buton et al. 2013). For our calibrator stars this source of error is larger than that from the Poisson noise, and is therefore included in our analysis.

We obtained NIR spectra of our B12 calibration sample and selected *Kepler* stars with candidate HZ planets (see Section 7) using the SpeX spectrograph (Rayner et al. 2003) attached to the NASA Infrared Telescope Facility (IRTF) on Mauna Kea. SpeX observations were taken in the short cross-dispersed (SXD) mode using the  $0.3 \times 15''$  slit, yielding simultaneous coverage from 0.8 to  $2.4 \mu\text{m}$  at a resolution of  $R \simeq 2000$ . A target was placed at two positions along the slit (A and B) and observed in an ABBA pattern to accurately subtract the sky background by differencing. For each B12 star we took 6 exposures in this pattern, which gave a  $S/N > 150$  in the  $H$ - and  $K$ -bands for all targets (and typically  $> 200$ ). For the *Kepler* HZ candidate hosts a total of 10-30 exposures were taken to achieve a peak  $S/N \gtrsim 60$  in the  $H$ - and  $K$ -bands. To remove effects from large telescope

slews, we obtained flat-field and argon lamp calibration sequences after each target.

Spectra were extracted using the SpeXTool package (Cushing et al. 2004), which performed flat-field correction, wavelength calibration, sky subtraction, and extraction of the one-dimensional spectrum. Multiple exposures were combined using the IDL routine *xcombpec*. To correct for telluric lines, we observed an A0V-type star within 1 hr and 0.1 airmass of the target observation (usually much closer in time and airmass). In the cases where a target's airmass changed by  $\geq 0.2$  or a single star observation lasted  $> 1$  hr two A0V stars were obtained. A telluric correction spectrum was constructed from each A0V star and applied to the relevant spectrum using the *xtellcor* package (Vacca et al. 2003). We placed each spectrum in its star's rest frame by cross-correlating it with a spectrum of the template M1.5 dwarf HD 36395 from the IRTF spectral library (Cushing et al. 2005; Rayner et al. 2009).

Rayner et al. (2009) find that, when using the  $0.3''$  slit in SXD mode, time-dependent changes in seeing, guiding, and differential atmospheric refraction can cause changes in the slope of SpeX data by as much as  $\simeq 2\%$ . We mitigated this effect by observing at the parallactic angle and minimizing the time between target and standard star observations. We used the change in slope between unstacked observations of the same star as a measure of the size of this error term for each target, which for our stars is typically  $< 1\%$ . This change in slope is significant for observations of our calibration sample, and thus is included in our error analysis.

Wavelengths in this paper are all reported in vacuum.

#### 4. MODIFICATIONS TO THE STELLAR PARAMETERS OF B12

##### 4.1. Bolometric Flux, Effective Temperature, and Luminosity

B12 measure the  $T_{\text{eff}}$  of their sample following the method outlined in van Belle et al. (2008). They use *BVR<sub>I</sub>JHK* photometry to find the best-fit template spectra (or combination of template spectra) from the Pickles (1998) library as well as the normalization constants for the spectra. Spectra for most of the late K and M dwarfs from Pickles (1998) extend to  $\simeq 1.1 \mu\text{m}$ , beyond which van Belle et al. (2008) and B12 extrapolate using NIR (usually *JHK*) photometry. They calculate  $F_{\text{BOL}}$  values from the normalized spectrum using the formula:

$$F_{\text{BOL}} = \int_0^\infty F_\lambda d\lambda, \quad (1)$$

where  $F_\lambda$  is the flux density at wavelength  $\lambda$ . From there B12 calculate  $T_{\text{eff}}$  of a given star using the Stefan-Boltzmann law:

$$T_{\text{eff}} = 2341 \left( \frac{F_{\text{BOL}}}{\theta_{\text{LD}}^2} \right)^{\frac{1}{4}}, \quad (2)$$

where  $T_{\text{eff}}$  is given in Kelvin,  $F_{\text{BOL}}$  is in units of  $10^{-8} \text{ erg cm}^{-2} \text{ s}^{-1}$ , and  $\theta_{\text{LD}}$  is the measured angular diameter of the star corrected for limb darkening and in units of milli-arcseconds.

It is straightforward to calculate the stellar luminosity ( $L_*$ ) for stars with a known distance (parallax) and  $F_{\text{BOL}}$

according to the formula:

$$L = 4\pi d^2 \times F_{\text{BOL}}, \quad (3)$$

where  $d$  is the distance, which B12 draw from *Hipparcos* parallaxes (van Leeuwen & Fantino 2005; van Leeuwen 2007). For all above calculations B12 assume no extinction along the line of site, which is reasonable given that these stars are nearby. Thus we make the same assumption.

Rather than employ template spectra, we calculate  $F_{\text{BOL}}$  from the actual spectra we obtained of each star. Although our combined SNIFS and SpeX spectra cover the wavelengths where K and M dwarfs emit the majority of their flux, SpeX spectra have a small gap at  $\simeq 1.8\mu\text{m}$  and several of our stars have comparatively low S/N ( $< 30$ ) in regions of telluric contamination. We replace these regions using the best-fit PHOENIX BT-SETTL model atmosphere (Allard et al. 2011) following the procedure described in Mann et al. (2013b) and Lépine et al. (2013), and summarized in Section 5.1. The resulting spectra cover 0.2–3.0  $\mu\text{m}$ . Blueward of this region we assume the flux follows Wein’s approximation, and redward we assume it follows the Rayleigh–Jeans law, and which we fit for using the 0.2–0.4 $\mu\text{m}$  and 2.0–3.0 $\mu\text{m}$  regions, respectively. Note that our typical stars have very little flux outside of 0.3–2.4 $\mu\text{m}$ , thus our approximation of the spectrum in this region has a negligible effect on our results.

Variations in atmospheric transparency and observing conditions between target and standard star observations as well as our choice of spectrophotometric standards (e.g., the EG131 standard is slightly variable, Wickramasinghe et al. 1978) can change the overall flux level (the absolute flux calibration) of each spectrum even while preserving the relative flux calibration. An erroneous overall flux level will lead directly to an erroneous  $F_{\text{BOL}}$  value, so we correct for this using visible and infrared photometry from the literature (e.g., Mermilliod et al. 1997). Sources of photometry from each star are listed in Table 1. Following the technique of Rayner et al. (2009), for each photometric point we calculate the ratio of the photometric flux to the equivalent flux synthesized from the spectra,  $C_i$ ;

$$C_i = \frac{f_{\text{zp},i} \times 10^{-0.4m_i}}{(\int_{\lambda} F_{\lambda} \times S_{\lambda,i} d\lambda)}, \quad (4)$$

where  $f_{\text{zp}}$  is the band zero point,  $m_i$  is the apparent magnitude, and  $S_{\lambda,i}$  is the filter transmission for a given photometric band,  $i$ . Errors in  $C_i$  account for (uncorrelated) errors in  $m_i$  and  $F_{\lambda}$ , but  $f_{\text{zp},i}$  and  $S_{\lambda,i}$  are assumed to have negligible uncertainty. The final correction factor ( $C_*$ ) applied to each star’s spectrum is the error-weighted mean of the  $C_i$  values for a given star, which has error:

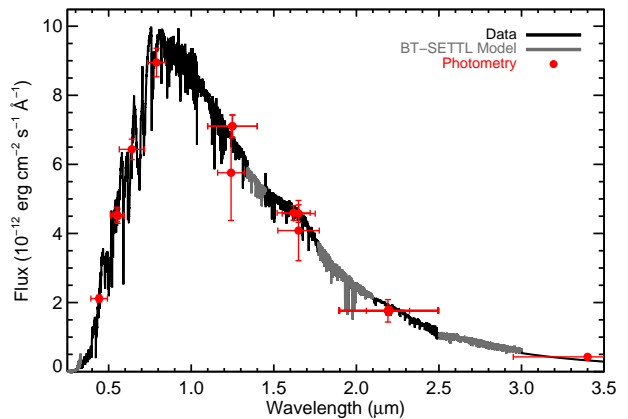
$$\sigma_{C_*}^2 = \frac{1}{\sum_i \sigma_{C_i}^2}. \quad (5)$$

If errors in the photometry or spectroscopy are underestimated, or if different photometric sources are systematically offset from each other, Equation 5 will be an underestimate of the total error in  $C_*$ . For this reason we also calculate the reduced  $\chi^2$  ( $\chi_{\text{red}}^2$ ) values for our correction factors (listed in Table 1). For all but one star

(GJ 725B) we get  $\chi_{\text{red}}^2 < 3$  (and most are  $< 2$ ). This indicates that our derived errors properly account for the scatter in  $C_i$  values for all targets except GJ 725B, which has a  $\chi_{\text{red}}^2$  of 7.4.

We see no obvious cause for the disagreement between the photometric and spectroscopic data for GJ 725B. It is not active (no detectible H- $\alpha$  emission), and the fit for GJ 725A is significantly better ( $\chi_{\text{red}}^2=2.9$ ). Because our choice of which photometry to use for this star significantly affects our derived stellar parameters, we conservatively elect to remove this star from our calibration sample.

We show the observed spectrum and literature photometry of GJ 887 ( $\chi_{\text{red}}^2 = 0.9$ ) in Figure 1 as a demonstration of our procedure. We report revised  $F_{\text{BOL}}$ ,  $T_{\text{eff}}$ , and  $L_*$  for the B12 sample in Table 1. We include the derived stellar parameters for GJ 725B in Table 1, although we caution readers that the errors in this star’s parameters are probably underestimated.



**Figure 1.** Spectrum of GJ 887 from SNIFS, SpeX, and PHOENIX models with photometry shown in red. Errors in the photometry are shown along the Y-axis, errors on the X-axis denote the approximate spectral region covered by the given filter. Regions of significant telluric contamination or gaps in the data are filled in using PHOENIX models, which are shown in gray.

Our  $F_{\text{BOL}}$  values are, on average,  $\simeq 4\%$  higher than those reported by B12 (see Figure 2), resulting in  $\simeq 1\%$  higher  $T_{\text{eff}}$ , and 4% higher  $\simeq L_*$ . These differences are small, but typical errors in  $F_{\text{BOL}}$  values are only 1 – 2% so the offset is significant for many stars, and highly significant when considering the whole sample (see Figure 2).

This discrepancy is in part due to slightly different choices for input photometry between our study and that of B12. If we redo our analysis, but restrict ourselves to photometry reported in B12, the disagreement shrinks to  $\simeq 3\%$ , mostly by moving in some of the largest outlier stars.

We find that majority of the discrepancy can be explained by systematic issues with the Pickles (1998) templates and extrapolation of these templates into the NIR. Of the 20 Pickles (1998) templates K4 and later, only three of them have data past 1.1 $\mu\text{m}$ . These three the spectra are built by combining data from multiple sources, but still have gaps in each spectrum where the data do not overlap. We find that the combined tem-

**Table 1**  
Calibration Sample

Name	$F_{\text{BOL}} (10^{-8})$ erg cm $^{-2}$ s $^{-1}$	$\sigma_{F_{\text{BOL}}}$	$T_{\text{eff}}$ K	$\sigma_{T_{\text{eff}}}$	$R_*$ $R_{\odot}$	$\sigma_{R_*}$	$M_*$ $M_{\odot}$	$\sigma_{M_*}$	$L_*$ $L_{\odot}$	$\sigma_{L_*}$	[Fe/H] <sup>a</sup>	$\chi^2_{\text{red}}$ <sup>b</sup>	Phot Ref <sup>c</sup>
GJ 15A	5.664	0.067	3602	13	0.3863	0.0021	0.405	0.041	0.02256	0.00027	-0.30	2.2	1,2,4,5
GJ 205	6.510	0.133	3850	22	0.5735	0.0044	0.637	0.064	0.06449	0.00139	+0.49	2.8	1,2,6,3,4,7
GJ 380	15.200	0.224	4176	19	0.6398	0.0046	0.711	0.071	0.11174	0.00167	+0.24	0.9	1,2,4,5
GJ 526	4.105	0.052	3646	34	0.4840	0.0084	0.490	0.049	0.03694	0.00051	-0.31	1.6	1,2,6,3,4,5
GJ 687	3.511	0.067	3457	35	0.4183	0.0070	0.403	0.040	0.02228	0.00044	-0.05	1.8	2,8,9
GJ 880	3.572	0.032	3731	16	0.5477	0.0048	0.572	0.057	0.05181	0.00058	+0.21	2.3	1,2,6,3,10,11
GJ 887	11.118	0.225	3695	35	0.4712	0.0086	0.494	0.049	0.03694	0.00075	-0.06	0.9	1,2,3,12,7
GJ 699	3.319	0.032	3238	11	0.1869	0.0012	0.159	0.016	0.00342	0.00003	-0.40	1.8	1,2,6,4,7
GJ 411	10.627	0.146	3532	17	0.3924	0.0033	0.392	0.039	0.02134	0.00030	-0.38	2.0	1,2,3,13,5
GJ 105A	17.302	0.216	4704	21	0.7949	0.0062	0.767	0.124	0.27615	0.00356	-0.28	1.2	1,2,3,14,5
GJ 338A	6.171	0.119	3953	41	0.5773	0.0131	0.630	0.063	0.07316	0.00276	-0.01	2.7	2,15,16
GJ 338B	5.799	0.044	3926	37	0.5673	0.0137	0.617	0.062	0.06875	0.00229	-0.04	2.2	1,15,15,16
GJ 412A	3.044	0.029	3537	41	0.3982	0.0091	0.390	0.039	0.02214	0.00024	-0.37	2.9	1,2,4,5
GJ 436	0.874	0.009	3520	66	0.4546	0.0182	0.447	0.045	0.02834	0.00072	+0.01	1.5	1,17,18
GJ 570A	20.463	0.278	4588	58	0.7390	0.0190	0.740	0.119	0.21609	0.00317	-0.06	2.2	1,2,4,7
GJ 581	0.971	0.011	3487	62	0.2990	0.0100	0.308	0.031	0.01181	0.00021	-0.15	1.1	1,2,6,4,7
GJ 702B	19.922	0.474	4475	33	0.6697	0.0089	0.749	0.075	0.15953	0.00396	+0.01	2.5	2,14,19
GJ 725A	3.989	0.038	3417	17	0.3561	0.0039	0.330	0.033	0.01573	0.00019	-0.23	2.9	1,2,20,8,16
GJ 725B <sup>d</sup>	2.351	0.023	3142	29	0.3232	0.0061	0.257	0.026	0.00927	0.00011	-0.30	7.4	1,2,3,8,16
GJ 809	3.413	0.042	3744	27	0.5472	0.0067	0.573	0.057	0.05250	0.00069	-0.06	2.3	1,2,10,9
GJ 820A	38.436	0.506	4399	16	0.6611	0.0048	0.727	0.073	0.14606	0.00196	-0.27	1.3	1,2,14,11
GJ 820B	22.284	0.323	4025	24	0.6010	0.0072	0.656	0.066	0.08468	0.00125	-0.22	0.9	1,2,14,11
GJ 892	21.141	0.290	4773	20	0.7784	0.0053	0.771	0.124	0.28046	0.00389	-0.23	0.6	1,2,14,11

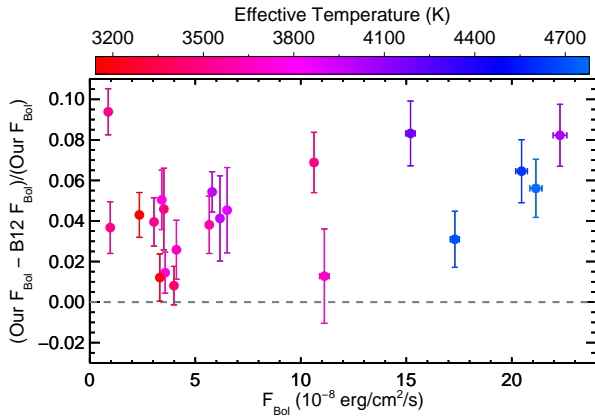
**Note.** — Photometry references: (1) Skrutskie et al. (2006); (2) Morel & Magnenat (1978); (3) Cousins (1980); (4) Johnson (1965); (5) Glass (1975); (6) Koen et al. (2010); (7) Mould & Hyland (1976); (8) Johnson & Morgan (1953); (9) Persson et al. (1977); (10) Erro (1971); (11) Johnson et al. (1968); (12) Bessel (1990); (13) Johnson (1964); (14) Johnson et al. (1966); (15) Cowley et al. (1967); (16) Veeder (1974); (17) Weis (1993); (18) Cutri et al. (2003); (19) Christou & Drummond (2006); (20) Eggen (1979).

<sup>a</sup> Errors on [Fe/H] are  $\simeq 0.07$  dex for all stars (since measurement errors are negligible compared to calibration errors).

<sup>b</sup>  $\chi^2_{\text{red}}$  from our fit of the spectra to published photometry.

<sup>c</sup> Some of the photometry for these stars comes from the General Catalogue of Photometric Data (Mermilliod et al. 1997).

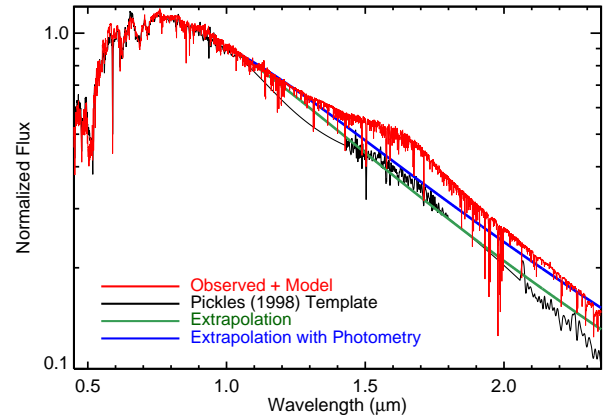
<sup>d</sup> GJ 725B has a  $\chi^2_{\text{red}}$  of 7.4. The next highest  $\chi^2_{\text{red}}$  is 2.9 (GJ 725A), suggesting that either the photometry or spectroscopy for GJ 725B is inaccurate, or that the star is variable. For this reason we removed GJ 725B from our analysis.



**Figure 2.** Fractional difference in our estimates of  $F_{\text{BOL}}$  compared to those from B12. The dashed line indicates no difference. GJ 702B is not included because blended photometry prevents B12 from calculating an accurate  $F_{\text{BOL}}$  for this star.

plates do not properly conserve the relative flux calibration. An example can be seen in Figure 3, where we show a Pickles (1998) template and our spectrum of GJ 380. Although the spectra are well matched in the optical, the NIR data are systematically offset. The result is differences in  $F_{\text{BOL}}$  of 4%–12% between templates and our combined spectra. The situation improves if we ‘fix’ the template by applying a constant offset to the template NIR data to better match our own spectrum, but offsets of 2%–3% in the derived  $F_{\text{BOL}}$  values persist, similar to

the difference between our own  $F_{\text{BOL}}$  values and those from B12.



**Figure 3.** Observed spectrum (red) with the best-match Pickles (1998) template star (black). We also show the extrapolation of the optical data into the NIR by fitting the 0.9–1.1  $\mu\text{m}$  region with a Planck function (green) and the fit if we include the NIR photometry (blue).

The remaining 17 templates lack NIR data, and instead B12 must extrapolate beyond 1.1  $\mu\text{m}$ . We investigate how this affects the result by fitting our spectra in the 0.9–1.1  $\mu\text{m}$  region using a Planck function. The fit for GJ 380 is shown in Figure 3. Generally this method significantly

underestimates the overall  $F_{\text{BOL}}$  level (by 5%–10%). The discrepancy decreases if we include available NIR ( $JHK$ ) photometry in the fit (resulting fit shown in Figure 3). However, even in this case the fit often fails to match the  $H$ -band “bump”, resulting in discrepancies of 3%–5% in  $F_{\text{BOL}}$ , which is similar to the offset between our results and that of B12. Late K and M dwarfs have spectra which are rich and irregular, and therefore cannot be fit with a simple function.

#### 4.2. Masses

B12 derive  $M_*$  using the relation between  $M_K$  and  $M_*$  from Henry & McCarthy (1993). They choose not to use the slightly more precise relation from Delfosse et al. (2000) because the Delfosse et al. (2000) sample includes only late K and M dwarfs, while B12 include many early K dwarfs. Because of our  $T_{\text{eff}}$  cut, our calibration sample has only three stars that are outside the limits of the Delfosse et al. (2000) relation. Therefore we derive  $M_*$  using the relation from Henry & McCarthy (1993) for those three stars, and the Delfosse et al. (2000) relation for the rest of the sample. Considering the conservative 10% error assumed by Delfosse et al. (2000) all  $M_*$  values reported by B12 are within  $1\sigma$  of our own. Resulting values of  $M_*$  and errors are reported in Table 1.

#### 4.3. Relations between $T_{\text{eff}}$ and fundamental parameters

Our revision of  $F_{\text{BOL}}$ ,  $T_{\text{eff}}$ , and  $L_*$  values for the B12 stars changes the empirical relations B12 derive between  $T_{\text{eff}}$ ,  $M_*$ ,  $R_*$ , and  $L_*$ . We use the Levenberg-Marquardt least-squares minimizer MPFIT (Markwardt 2009) to derive the following formulae:

$$R_* = -16.883175 + 1.1835396 \times 10^{-2} T_{\text{eff}} - 2.70872 \times 10^{-6} T_{\text{eff}}^2 + 2.105028 \times 10^{-10} T_{\text{eff}}^3, \quad (6)$$

$$L_* = -0.78107639 + 7.3973011 \times 10^{-4} T_{\text{eff}} - 2.4911 \times 10^{-7} T_{\text{eff}}^2 + 2.947987 \times 10^{-11} T_{\text{eff}}^3, \quad (7)$$

$$M_* = -22.296508 + 1.5446387 \times 10^{-2} T_{\text{eff}} - 3.488452 \times 10^{-6} T_{\text{eff}}^2 + 2.64961 \times 10^{-10} T_{\text{eff}}^3, \quad (8)$$

where  $T_{\text{eff}}$  is given in Kelvin and  $M_*$ ,  $L_*$ , and  $R_*$  are given in solar units. Data and fits for these equations are shown in Figure 4 along with the corresponding fits from B12 (a  $T_{\text{eff}}$ - $M_*$  relation is not shown because B12 do not derive a corresponding fit for comparison). The  $\chi_{\text{red}}^2$  for all three fits are  $< 1$ , suggesting that some of the errors may be slightly overestimated.

Errors for each fit are estimated using a Monte Carlo (MC) approach. Because of correlations between values (e.g., a change in  $F_{\text{BOL}}$  alters  $T_{\text{eff}}$  and  $L_*$  in a coherent way) we perform our MC error estimate over all steps in our analysis. Specifically, we perturb the spectra, literature photometry, parallaxes, and angular diameters randomly according to their errors. We assume all sources of error are Gaussian, except for the additional error term in the SpeX data (see Section 3), which affects the overall slope of the spectrum. Values of  $M_*$  are assumed to have a correlated (but still Gaussian) error term, i.e., that derived  $M_*$  for all stars may increase or decrease together for each MC run (due to possible systematics). This is treated separately from the non-correlated errors in the measurement of  $M_K$  arising from errors in the photometry and parallax.

We perform the full analysis laid out in Section 4 using the perturbed values. In addition to perturbing the values we also randomly remove one calibrator star from the analysis (jackknifing), to account for sensitivity to our exact set of calibration stars. We repeat this process (perturbing values and randomly removing one star) 1000 times, each time we record the resulting polynomial fits between  $T_{\text{eff}}$ ,  $R_*$ ,  $M_*$ , and  $L_*$  and  $L_* M_*^{-2/3}$ . MC error estimation is also done for a relation between  $T_{\text{eff}}$  and  $L_* M_*^{-2/3}$  because this quantity is used in our HZ calculations (see Section 7), and errors in  $L_*$  and  $M_*$  are correlated. The adopted errors for each relation are assumed to be the standard deviation of the fitted  $R_*$ ,  $M_*$ , and  $L_*$  values at a given  $T_{\text{eff}}$ . Errors as a function of  $T_{\text{eff}}$  for each relation are shown in Figure 5. As expected, the errors grow significantly at  $T_{\text{eff}} < 3300$  K where our calibration sample contains only a single star.

In Figure 6 we show the difference between our fits of  $R_*$  and  $L_*$  and those from B12 relative to our estimated errors. Note that for most temperatures, the differences are not significant. However, the offset is systematic and a function of temperature, and thus could be important. The higher discrepancy at cool temperatures ( $T_{\text{eff}} < 3400$ ) is not due to our updated parameters, but instead due to the exclusion of GJ 725B from our data set. GJ 725B has an unusually large radius and luminosity for its temperature (assuming the assigned temperature is correct), which drives the B12 fits to larger radii and higher luminosities at low temperatures.

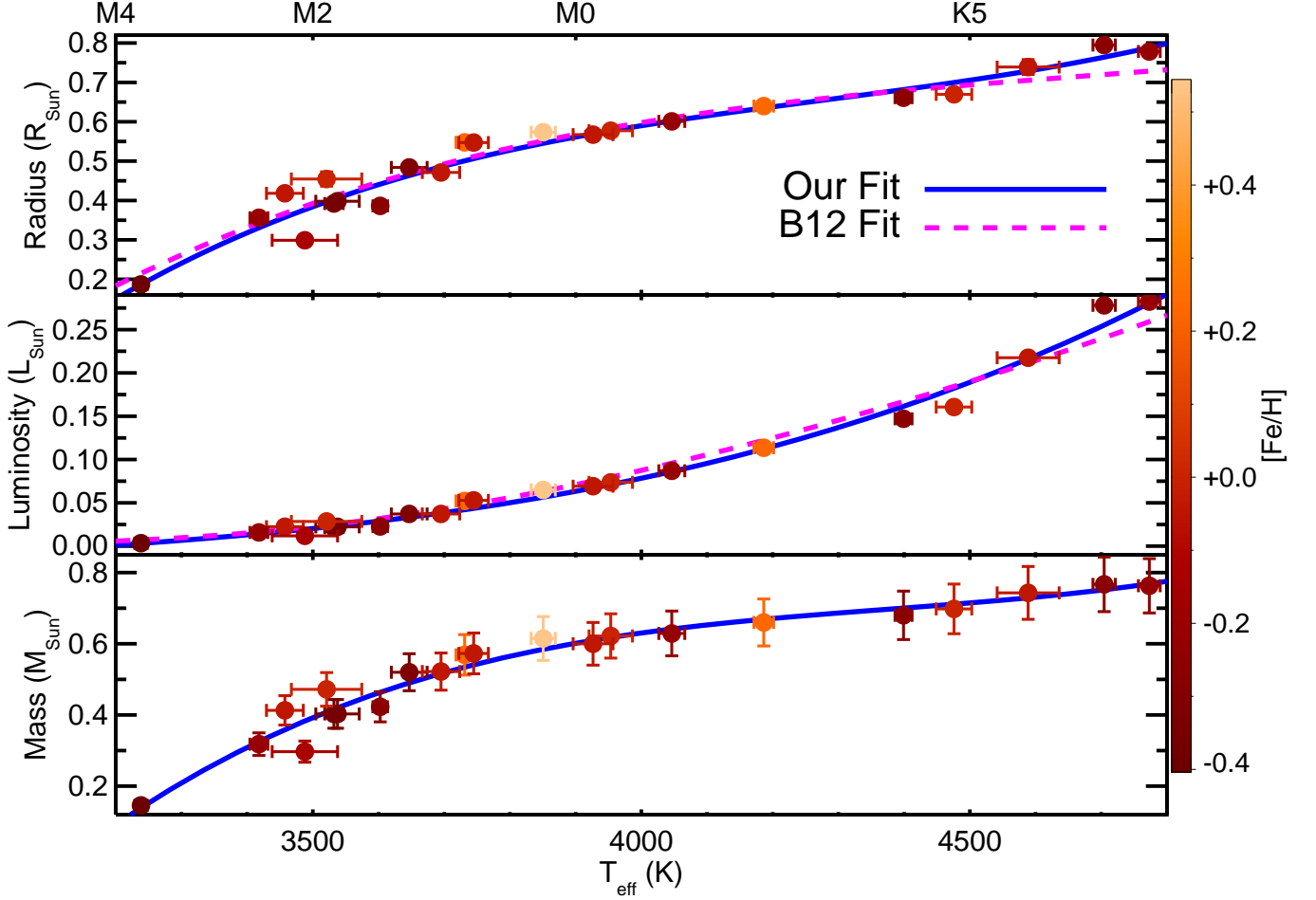
#### 4.4. Metallicity

B12 finds that radius is insensitive to metallicity, in contrast to model predictions (Baraffe et al. 1998; Dotter et al. 2008). However, such a conclusion requires accurate metallicities for late-type stars, which are notoriously difficult to determine. B12 draw metallicities for their K and M dwarfs from Edvardsson et al. (1993), Anderson & Francis (2011), Rojas-Ayala et al. (2012), and Neves et al. (2012). These sources use different techniques to determine metallicity, and likely suffer from systematic offsets from each other like those noted in Mann et al. (2013a). Further, abundance analysis techniques from Edvardsson et al. (1993) and Anderson & Francis (2011) are poorly calibrated for cool stars.

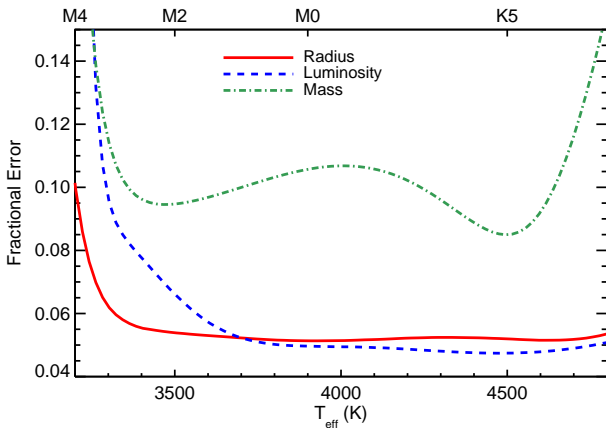
We estimate  $[\text{Fe}/\text{H}]$  for the 23 stars in our sample following the methods of Mann et al. (2013a), which are tailored to data from SNIFS and SpeX. Mann et al. (2013a) provide empirical metallicity calibrations based on visible,  $J$ -,  $H$ -, and  $K$ -band spectra. We calculate the metallicity of each star using the weighted mean from each of the four calibrations, accounting for both measurement errors and systematic errors in the calibration as reported by Mann et al. (2013a).

We report derived metallicities in Table 1. In Figure 7 we show a comparison of our derived metallicities with those reported in B12. On average, our  $[\text{Fe}/\text{H}]$  values are 0.05 dex lower than those reported in B12. The difference in metallicity for 16 of the 23 stars (69%) is  $< 1\sigma$ . Given the difficulty in measuring M dwarf metallicities, we consider the differences to be minor.

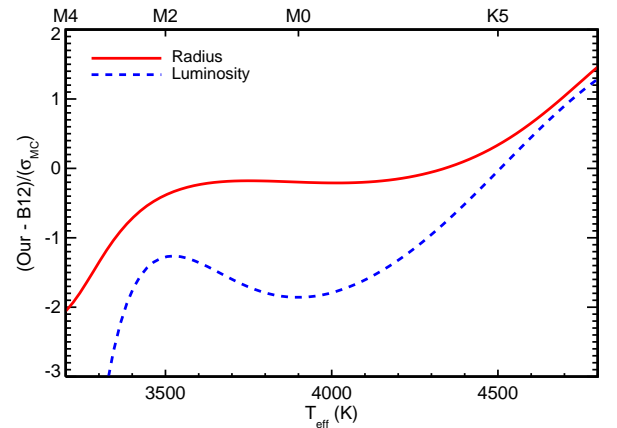
Adding metallicity as a parameter in Equations 6–8 does not improve the  $\chi_{\text{red}}^2$  in any significant way (for an



**Figure 4.**  $R_*$ ,  $L_*$ , and  $M_*$  of stars in our calibration sample as a function of their  $T_{\text{eff}}$ . Stars are shaded (colored) based on their metallicities. Approximate spectral types are given for reference. Error bars are shown, although errors on  $R_*$  and  $L_*$  are sometimes smaller than the plotting points. Blue lines indicate our best-fit relations, described by Equations 6–8. The magenta dashed line shows the fits derived in B12 (B12 do not provide a  $T_{\text{eff}}-M_*$  relation).



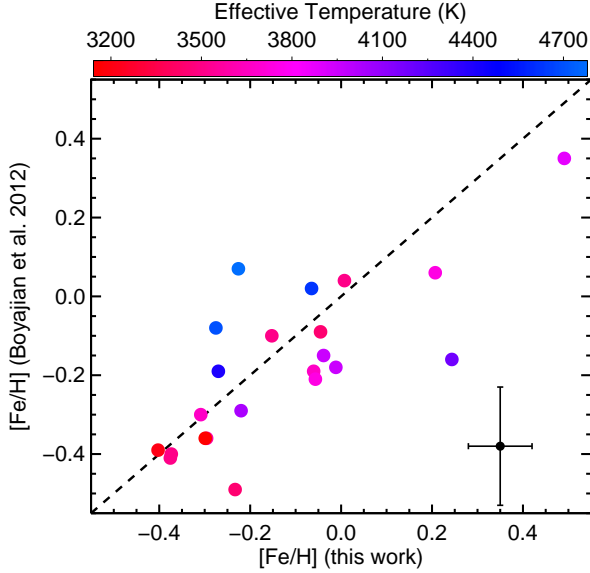
**Figure 5.** Fractional error as a function of  $T_{\text{eff}}$  for Equations 6–8:  $\text{refeqn:RM}$ . Approximate spectral types are shown for reference. Errors are calculated by a Monte Carlo process explained in Section 4.3 accounting for both the scatter in the data points around the fit and the uncertainty in the fit itself. Note that this error estimate does not include additional errors from measuring  $T_{\text{eff}}$ . Rather, this represents the floor of what errors are possible using this technique.



**Figure 6.** Difference between  $T_{\text{eff}}-R_*$  and  $T_{\text{eff}}-L_*$  relations from Equations 6–7 and those from B12 divided by the errors calculated from our MC/jackknife estimate (see Figure 5). Difference considers errors from our analysis only (not those reported by B12).

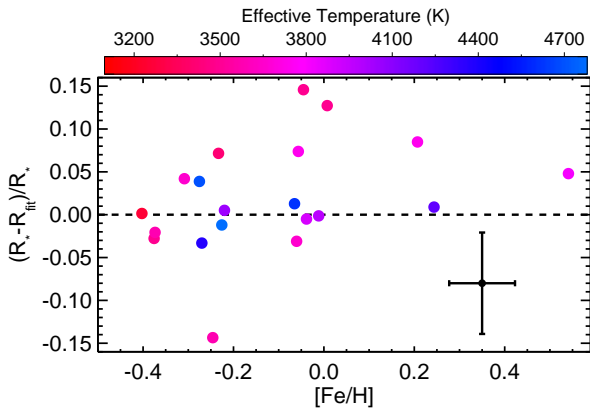
example see Figure 8). The difference between a measured parameter (e.g.,  $R_*$ ) and the fitted value of that parameter shows no significant correlation with  $[\text{Fe}/\text{H}]$





**Figure 7.** Comparison of  $[\text{Fe}/\text{H}]$  values derived from our analysis with those reported in B12 (see references within for original sources). A typical error bar is shown in the bottom right. Points are color-coded by their  $T_{\text{eff}}$ . Note that errors in B12 metallicities vary based on the original literature source.

for  $M_*$ ,  $R_*$ , or  $L_*$  based on the Spearman rank test (probability of a correlation  $< 70\%$  in all cases). M dwarf models predict a difference in the  $T_{\text{eff}}-R_*$  relation that should be detectable given the precision of our stellar parameters (see B12 for a further discussion of this discrepancy).



**Figure 8.** Fractional deviation of  $R_*$  from the relation given by Equation 6 as a function of  $[\text{Fe}/\text{H}]$ . Points are color-coded by their  $T_{\text{eff}}$ . A typical error bar is shown in the bottom right. The error estimate includes errors arising from extrapolating  $T_{\text{eff}}$  to  $R_*$  (i.e., errors in  $T_{\text{eff}}$  and errors in Equation 6). Although a weak correlation may be present, it is not statistically significant.

## 5. TECHNIQUES TO DETERMINE $T_{\text{eff}}$

Equations 6–8 provide a means to estimate other stellar properties from  $T_{\text{eff}}$ . For nearby stars with parallaxes, we can determine  $F_{\text{BOL}}$  from Equation 1 and  $L_*$  from Equation 3 and then constrain  $T_{\text{eff}}$  by inverting Equation 7. However, we need a reliable means to measure

$T_{\text{eff}}$  for stars without parallaxes (e.g., most *Kepler* stars). For this reason we use the calibration sample to develop and test techniques to measure  $T_{\text{eff}}$  using primarily visible (but also NIR) spectra that can then be applied to more distant stars in a homogenous manner.

### 5.1. Model Spectrum-fitting

We compare the SNIFS spectra of our calibration stars to a grid of models of K and M dwarf spectra from the BT-SETTL version of the PHOENIX code (Allard et al. 2011, 2012). We test the set of models that use the Asplund et al. (2009) solar abundances (AGSS), and those using the updated Caffau et al. (2011) revised solar abundances (CIFIST, Allard et al. 2013; Rajpurohit et al. 2013). The AGSS models have been used extensively in previous studies of M dwarfs (e.g., Mann et al. 2012; Muirhead et al. 2012a; Lépine et al. 2013), but the CIFIST models use a revised TiO line list (Plez 1998). We download synthetic spectra for both model grids spanning  $T_{\text{eff}}$  of 2700–5200 K in 100 K increments,  $\log g$  values of 4.0, 4.5, and 5.0, and  $[\text{M}/\text{H}]$  values of  $-1.0$ ,  $-0.5$ ,  $0.0$ ,  $0.3$ , and  $0.5$ .  $[\alpha/\text{Fe}]$  is taken to be solar for  $[\text{M}/\text{H}] \geq -0.5$  and  $+0.2$  for  $[\text{M}/\text{H}] \leq -1.0$ .

For a given stellar spectrum, we restrict our comparison to the subset of models with the two metallicity values immediately above and below the metallicity calculated for the star (see Section 4.4). For example, for a star with  $[\text{Fe}/\text{H}] = 0.1$  we restrict our comparison to models with  $[\text{M}/\text{H}] = 0.0$  or  $[\text{M}/\text{H}] = +0.3$ . We assume the difference between  $[\text{Fe}/\text{H}]$  and  $[\text{M}/\text{H}]$  is negligible, which is reasonable for stars near solar metallicity. We place no restriction on models based on assumed  $\log g$  or  $T_{\text{eff}}$  other than those imposed by the range of grid points available.

Following Lépine et al. (2013) we convolve the models with a Gaussian with full width at half maximum of  $7\text{\AA}$  ( $R \sim 1000$ ) using the IDL code *gaussfold*. Although SNIFS spectra are already shifted to vacuum and rest frames, we correct for any additional wavelength differences by cross-correlating the observed and model spectra using the IDL routine *xcorl*. The size of this correction is typically less than one resolution element.

M dwarf atmospheric models contain unknown systematic errors, in large part due to poorly modeled molecular absorption (e.g., TiO bands, Reylé et al. 2011). As a result, the distribution of (data-model)/error values will not be normally distributed and  $\chi^2$  does not apply. Instead, we use the technique from Cushing et al. (2008) and find the best-fit model by minimizing a goodness-of-fit statistic  $G_K$ , which is analogous to  $\chi^2$ .  $G_K$  is defined as:

$$G_K = \sum_{i=1}^n \left( \frac{w_i (F_i - C_K F_{K,i})}{\sigma_i} \right)^2, \quad (9)$$

where  $w_i$  is the weight of the  $i$ th wavelength bin,  $F_i$  and  $F_{K,i}$  are the data and model flux densities, respectively, and  $\sigma_i$  is the error on  $F_i$ .  $C_K$  is a normalization constant that, for absolute flux calibrated spectra, is equal to  $R_*^2/d^2$ . Although we know  $R_*^2/d^2$  for our calibration stars, neither the radius nor the distance are known generally. Instead we set  $C_K$  such that the mean of  $F$  and  $F_K$  are the same.

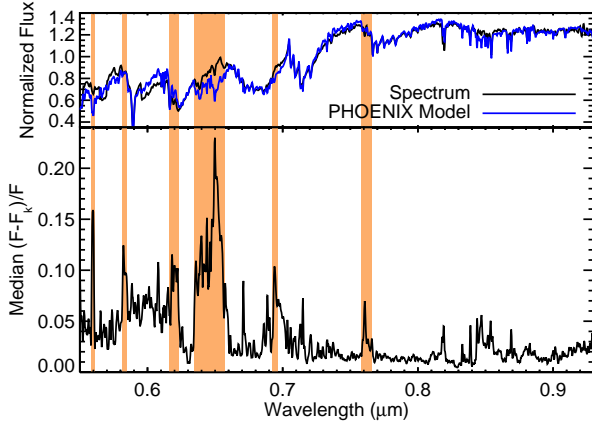
Weights blueward of  $5500\text{\AA}$  and redward of  $9300\text{\AA}$  are



set to 0 because of low S/N for fainter M dwarfs in the blue region and due to contamination from variable  $\text{H}_2\text{O}$  lines on the red end.  $\text{H}_2\text{O}$  variability is not a problem for the calibration stars because the exposure times are short (a few seconds), but this can be a major source of error for faint stars where the exposure times are often  $\geq 30$  minutes.

The rest of the weights are assigned either 0 or 1 based on how accurately models reproduce the real spectra. We first find the best-fit model for each of the calibration stars (described below) assuming  $w_i = 1$  for all  $i$ . We compute the residuals from the best-fit model for each star, and then compute the median fractional deviation between the data and model for all stars at a given wavelength. We show the final median spectral residuals in Figure 9. From this stacked residual we identify regions with a median deviation of  $> 10\%$  ( $> 1$  order of magnitude greater than our measurement errors) that are larger than  $10 \text{ \AA}$  wide. For these regions we set  $w_i = 0$ . All other weights are set to  $w_i = 1$ .

We then repeat the process of fitting the model spectra to the data, computing the median residuals and adjusting the weights. After two iterations the weights do not change. We adopt this final list of weights for our model-fitting procedure. We also mask out a small region around  $0.76 \mu\text{m}$  due to concerns about accurate removal of the telluric  $\text{O}_2$  line. We show the regions with  $w_i = 0$  (shaded) in Figure 9. One part of the spectrum that stands out is the  $0.64\text{--}0.66 \mu\text{m}$  section, which contains the poorly modeled TiO absorption band (Reyl   et al. 2011).



**Figure 9.** Top: sample spectrum (black) and best-fit PHOENIX (BT-SETTL) model (blue). Bottom: median fractional deviation of the 22 spectra from the best-fit PHOENIX model as a function of wavelength. Regions that are larger than  $10 \text{ \AA}$  with a deviation larger than  $10\%$  are masked out as part of our fitting process (shaded regions). The area around  $0.76 \mu\text{m}$  is excluded because of concerns about our ability to accurately remove the telluric  $\text{O}_2$  line. Non-shaded regions are given full weight ( $w_i = 1$ ) in our model fits.

A more nuanced weighting scheme would be to weight each interval according to how consistent it is with the models, or test different weighting schemes to see which gives the best agreement with the bolometric temperatures. However regions with modest agreement between the real and synthetic spectra may contain more temperature information than regions with slightly better

matches, and we have no a priori information about what spectral regions are the most temperature sensitive. Further, binary (0 or 1) weights are sufficient to get excellent agreement between the bolometric and model temperature scale.

After calculating  $G_K$  for the relevant models, we select the seven models with the lowest  $G_K$  values and construct 10,000 linear combinations from them (L  pine et al. 2013). The  $T_{\text{eff}}$  value assigned to each combination is the equivalent linear combination of the  $T_{\text{eff}}$  values of the components. As before, we calculate  $G_K$  for these 10,000 models. We adopt the temperature of the best-fit model from these linear combinations as the PHOENIX temperature ( $T_{\text{PH}}$ ).

We compare  $T_{\text{PH}}$  from each model grid (AGSS and CIFIST) to the  $T_{\text{eff}}$  derived from bolometric flux and angular diameter ( $T_{\text{BOL}}$ ) in Figure 10. Using either grid  $T_{\text{PH}}$  accurately reproduces  $T_{\text{BOL}}$  for stars warmer than 4000 K. For cooler stars the AGSS models predict significantly cooler temperatures than those derived bolometrically. We find a best-fit linear relation of:

$$T_{\text{eff,BOL}} = 279.(\pm 43) + 0.942(\pm 0.013)T_{\text{eff,AGSS}}. \quad (10)$$

Errors (quoted in parentheses) are determined via MC. Specifically, we randomly perturb  $T_{\text{BOL}}$  values and the visible-wavelength spectra by reported errors. We then re-fit the spectra using by minimizing  $G_K$  as explained above. We repeat this 1000 times, and report the  $1\text{-}\sigma$  errors. The difference between  $T_{\text{PH}}$  and  $T_{\text{BOL}}$  is small for warm stars, but results in a systematic offset of more than 75 K at 3500 K. We find that the slope of  $T_{\text{PH}}$  vs.  $T_{\text{BOL}}$  using the AGSS models is inconsistent with 1 at  $5\sigma$ , and conclude that this slope cannot be due to random noise. If, instead, we adopt the B12 temperatures for our calibration stars (but still use the AGSS models) we derive the relation:

$$T_{\text{eff,B12}} = 366.(\pm 41) + 0.907(\pm 0.011)T_{\text{eff,AGSS}}. \quad (11)$$

Thus the difference between temperatures derived using the AGSS models has a more significant slope ( $8\sigma$ ) when we use the B12 temperatures.

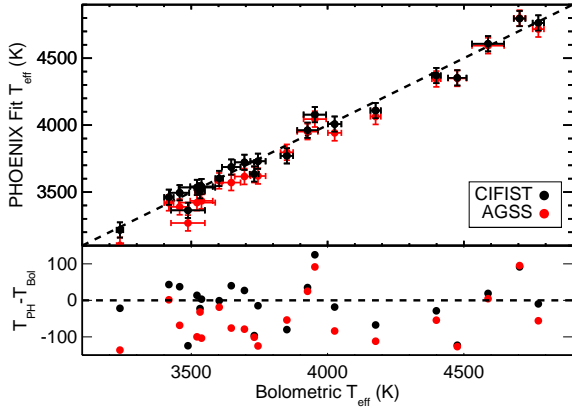
We note a significant slope between CIFIST and B12 temperatures. We derive a best fit relation of:

$$T_{\text{eff,B12}} = 192.(\pm 41) + 0.960(\pm 0.012)T_{\text{eff,CIFIST}}. \quad (12)$$

Although this slope is still  $4\sigma$  inconsistent with 1, CIFIST models are still better at reproducing the B12 temperature scale than AGSS models.

Effective temperatures derived using the CIFIST grid are in excellent agreement with our derived  $T_{\text{BOL}}$  values, with a median difference of  $1 (\pm 13) \text{ K}$ , and no statistically significant slope ( $0.99 \pm 0.01$ ), as can be seen in Figure 10. The rms deviation of  $T_{\text{PH}}$  from  $T_{\text{BOL}}$  when using the CIFIST models is just 57 K (better than when using AGSS even after correcting for the offset), which we adopt as the error from our model fitting procedure. Because of the superior performance of the CIFIST models in reproducing our  $T_{\text{BOL}}$  scale (and even the B12  $T_{\text{BOL}}$  scale) we use only CIFIST models for the rest of the paper.

## 5.2. Temperature Sensitive Indices



**Figure 10.** Comparison of  $T_{\text{eff}}$  determined from fitting PHOENIX models (AGSS and CIFIST) to visible-wavelength spectra to those calculated from bolometric values calculated with Equation 2. Red points correspond to fits from the AGSS models, while black points correspond to fits from the CIFIST models. The bottom plot shows the residuals assuming a 1:1 relation. The dashed line indicates equality (for both plots). The rms deviation of  $T_{\text{PH}}$  from  $T_{\text{BOL}}$  is when using the CIFIST models is 57 K, which we adopt as the error on our model spectrum-fitting procedure.

In addition to our model spectrum-fitting procedure we develop four novel temperature indices, one for each wavelength regime (visible,  $J$ -,  $H$ -, and  $K$ -band). We define an index as:

$$\text{Index} = \frac{\langle \text{Band1} \rangle / \langle \text{Band2} \rangle}{\langle \text{Band2} \rangle / \langle \text{Band3} \rangle}, \quad (13)$$

where  $\langle \rangle$  denotes the median flux level within a given wavelength region. An index is a measure of the curvature of the spectrum in a particular region. This definition is identical to that given in Covey et al. (2010) although they define bands to measure the level of  $\text{H}_2\text{O}$  absorption in NIR spectra of M dwarfs, while we are interested in an empirical method to measure  $T_{\text{eff}}$  with no preference for any particular absorption band.

We search for the best performing (lowest  $\chi^2$ ) band definitions for each wavelength regime: visible (0.3 - 1.0  $\mu\text{m}$ ),  $J$ - (1.0 - 1.45  $\mu\text{m}$ ),  $H$ - (1.45 - 1.85  $\mu\text{m}$ ), and  $K$ -band (1.85 - 2.4  $\mu\text{m}$ ). We vary the central wavelengths of each band ( $\lambda_{c,1}$ ,  $\lambda_{c,2}$ ,  $\lambda_{c,3}$ ) in increments of 5  $\text{\AA}$ , and the width of the bands ( $\delta\lambda$ ) from 25  $\text{\AA}$  to 100  $\text{\AA}$  in increments of 5  $\text{\AA}$ . We exclude bands which would overlap with features that are identified by Mann et al. (2013a) as significantly metal sensitive.

For every iteration we calculate the index (Equation 13) for each calibration star. We fit a second-order polynomial between the index measurement and the  $T_{\text{BOL}}$  values and calculate  $\chi^2$ . The best  $\lambda_i$  and  $\delta\lambda$  values are taken to be the ones that give the smallest  $\chi^2$ . We use the final rms of the fitted temperature vs. the  $T_{\text{BOL}}$  as an estimate of the error for each relation.

We report the  $\lambda_i$  and  $\delta\lambda$  values, the polynomial coefficients for calculating  $T_{\text{eff}}$  from these indices, and the adopted errors from each relation in Table 2. We also show the indices and  $T_{\text{eff}}$  values for each wavelength regime in Figure 11. The relation for visible-wavelength spectra gives the smallest error, although it is still not as precise as the model spectrum-fitting tech-

nique. Most likely the model spectrum-fitting procedure performs better because it is using more spectral information. However, these relations have the advantage of being model independent, and can be calculated easily from low or moderate-resolution spectra.

Screening out metal-sensitive regions from Mann et al. (2013a) reduces the chance of systematics with metallicity in our  $T_{\text{eff}}$ -index relations. However, Mann et al. (2013a) detects metal sensitive atomic and molecular lines, and do not investigate the role of metallicity on continuum shape. To test the effect of metallicity on our indices we perform a Spearman rank test on  $T_{\text{eff}} - T_{\text{eff,fit}}$  vs.  $[\text{Fe}/\text{H}]$  for all relations. We find a correlation probability of 5%, 9%, 20%, and 66% for the visible,  $J$ -,  $H$ -, and  $K$ -band relations, respectively. We conclude that metallicity is a minor effect on these indices for the metallicity regime covered by our sample.

## 6. APPLICATION TO *Kepler* PLANET HOSTS

We apply our methods to calculate the  $T_{\text{eff}}$ ,  $M_*$ ,  $R_*$ , and  $L_*$  of the sample of *Kepler* planet hosts described in Section 2. We use our model spectrum-fitting procedure (using the CIFIST models) to determine  $T_{\text{eff}}$  because this gives the most precise  $T_{\text{eff}}$  values. One immediate concern is that the S/N of the visible spectra for our calibration sample is extremely high ( $\text{S/N} > 200$ ) compared to that of the *Kepler* target sample ( $\text{S/N} \approx 80$ ). We test the effect of lower S/N by adding Gaussian noise to the spectra of our calibration sample. We then repeat our process described in Section 5.1. We find that the rms is essentially unchanged ( $\sigma_{T_{\text{eff}}} = 58$  instead of  $\sigma_{T_{\text{eff}}} = 57$ ) and no significant systematic offset is present at the S/N of the *Kepler* target sample. Only at  $\text{S/N} \leq 50$  does the error in our model spectrum-fitting procedure begin to rise steeply, most likely because above this S/N systematic errors dominate over measurement errors.

We report  $T_{\text{eff}}$ ,  $M_*$ ,  $R_*$ , and  $L_*$  values for each of the late-type *Kepler* planet hosts in Table 3. Quoted errors consider both errors in Equations (5)-(7) (modeled by MC) and errors arising from our measurement of  $T_{\text{eff}}$ .

We derive a  $T_{\text{eff}}$  of  $5071 \pm 57$  K for KOI 940, which is significantly warmer than the next warmest star (KOI 2417,  $T_{\text{eff}} = 4582 \pm 57$ ), and makes it an early K dwarf. This temperature is outside the range spanned by our calibration sample, and thus the derived stellar parameters should be viewed skeptically. KOI 940 makes our  $K_P - J$  color cut (Section 2) most likely due to an erroneous  $K_P$  magnitude, or because of significant reddening. This illustrates a drawback of using  $K_P - J$  to define our sample instead of the full spectral energy distribution.

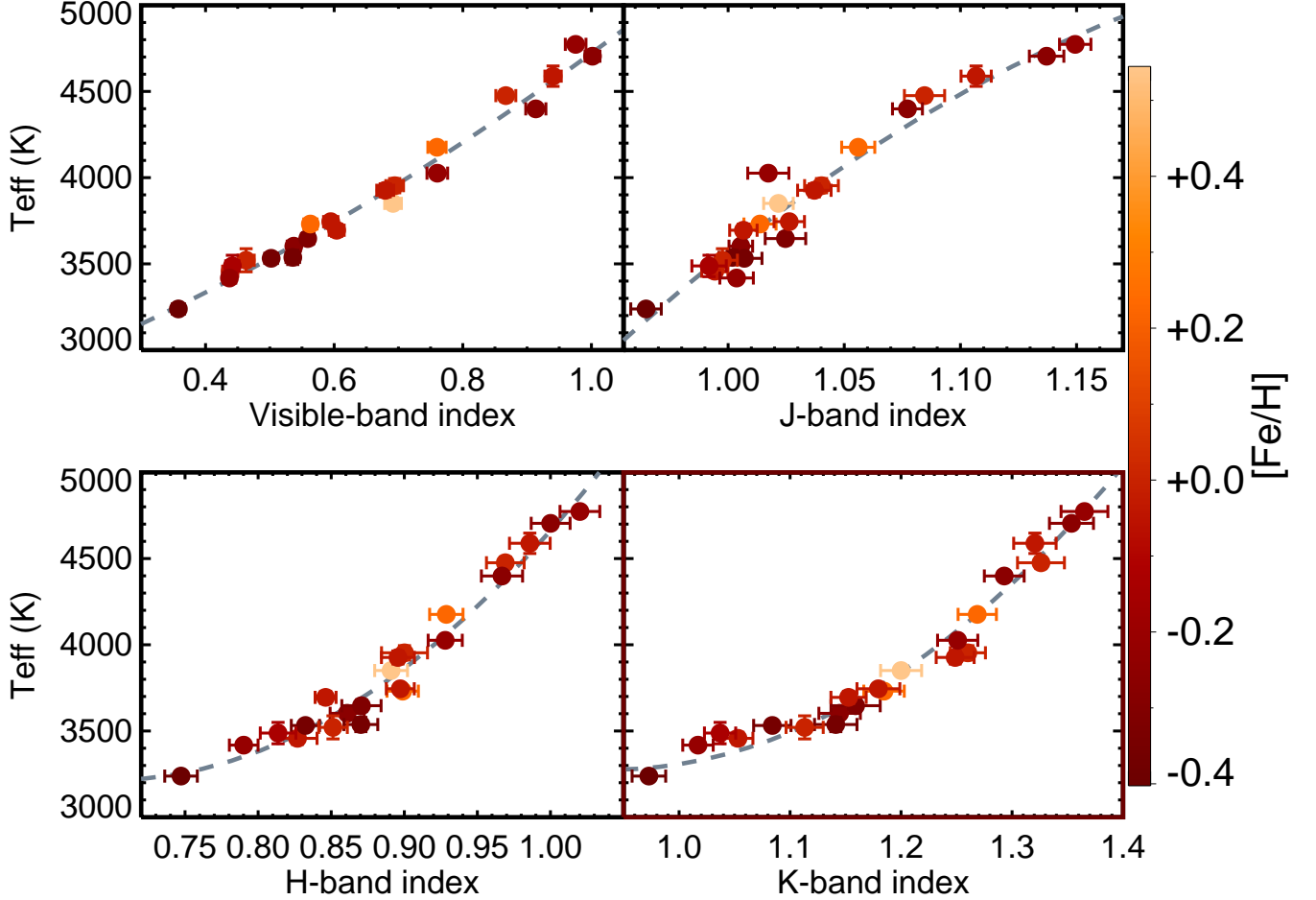
The coolest star in the sample is KOI 2704 with a  $T_{\text{eff}}$  of  $3157 \pm 57$  K. This star is likely even cooler than KOI 961 (Kepler-42,  $T_{\text{eff}} = 3241 \pm 57$ ). Unfortunately, KOI 2704's  $T_{\text{eff}}$  is slightly below the  $T_{\text{eff}}$  range of our calibrators (the coolest is GJ 699 with a  $T_{\text{eff}}$  of  $3238 \pm 11$ ), which makes it difficult to constrain its physical parameters using the methods outline in this paper. For example, our empirical relations yield a  $L_*$  below 0 for this  $T_{\text{eff}}$ , which is obviously unphysical. Instead, we conservatively assign  $1\sigma$  upper limits to the parameters of this star corresponding to the coolest calibrator star in the sample (GJ 699). Additional, cooler calibration stars are needed to better constrain the parameters of this star.

We show a comparison of our  $T_{\text{eff}}$  values to those from

**Table 2**  
Index Band Definitions:  
 $T_{\text{eff}} = a + b(\text{Index}) + c(\text{Index}^2)$

Name	Coverage $\mu\text{m}$	$\lambda_{c,1}$ $\mu\text{m}$	$\lambda_{c,2}$ $\mu\text{m}$	$\lambda_{c,3}$ $\mu\text{m}$	$\delta\lambda$ $\text{\AA}$	$a$	$b$	$c$	$\sigma_{T_{\text{eff}}}$ K
<i>V</i>	0.52 - 0.95	0.5700	0.6500	0.6850	30	$2.683 \times 10^3$	$1.354 \times 10^3$	$6.850 \times 10^2$	62
<i>J</i>	1.00 - 1.45	1.0130	1.1010	1.1630	85	$-2.277 \times 10^4$	$4.199 \times 10^4$	$-1.565 \times 10^4$	101
<i>H</i>	1.45 - 1.81	1.4640	1.6630	1.7960	30	$1.087 \times 10^4$	$-2.196 \times 10^4$	$1.568 \times 10^4$	80
<i>K</i>	2.00 - 2.35	2.0260	2.2765	2.3385	35	$1.051 \times 10^4$	$-1.543 \times 10^4$	$8.228 \times 10^3$	73

**Note.** —  $\text{Index}_i$  is defined by Equation 13 using the bands defined in this table.  $\lambda_{c,i}$  denotes the central wavelength of a given band in  $\mu\text{m}$ ,  $\delta\lambda$  denotes the width of a given band in  $\text{\AA}$ .



**Figure 11.** Relation between our empirical indices and  $T_{\text{eff}}$  for the B12 sample. See Table 2 for index definitions. Note that many errors are smaller than the data points. Each point is colored according to its metallicity (see legend) to demonstrate that there is no obvious trend with metallicity.

Muirhead et al. (2012a) and Dressing & Charbonneau (2013) in Figure 12. Because the  $\text{H}_2\text{O}$  index is only sensitive to  $T_{\text{eff}}$  below 4000 K (Rojas-Ayala et al. 2012; Dressing & Charbonneau 2013; Mann et al. 2013b), we exclude stars warmer than 4000 K from Muirhead et al. (2012a) in our comparison sample. Our median  $T_{\text{eff}}$  is  $43^{+12}_{-8}$  higher than those from Dressing & Charbonneau (2013) and  $72^{+9}_{-12}$  K higher than those from Muirhead et al. (2012a) (errors based on bootstrap resampling).

Most of this disagreement can be explained by the models utilized by the different groups. Muirhead et al. (2012a)  $T_{\text{eff}}$  values are based on a calibration of the  $\text{H}_2\text{O}$ -K2 index to AGSS models (Rojas-Ayala et al. 2012).  $T_{\text{eff}}$  from Dressing & Charbonneau (2013) are based on matching photometry from the *Kepler* Input Catalog (Brown et al. 2011) and the Two-Micron All Sky Survey (Skrutskie et al. 2006) to Dotter et al. (2008) models, which are based on an earlier version of the PHOENIX code that utilizes the Grevesse & Sauval (1998) solar

**Table 3**  
Late-type *Kepler* planet host parameters

KOI #	KID #	$T_{\text{eff}}^{\text{a}}$ K	$R_*$ $R_{\odot}$	$\sigma_R$ $R_{\odot}$	$M_*$ $M_{\odot}$	$\sigma_M$ $M_{\odot}$	$L_*$ $L_{\odot}$	$\sigma_L$ $L_{\odot}$
227	6185476	4093	0.615	0.035	0.653	0.070	0.095	0.012
247	11852982	3852	0.547	0.034	0.586	0.065	0.057	0.008
248	5364071	3970	0.583	0.034	0.623	0.068	0.074	0.010
249	9390653	3548	0.413	0.039	0.428	0.058	0.024	0.005
250	9757613	4049	0.604	0.034	0.643	0.070	0.087	0.011
251	10489206	3770	0.517	0.035	0.553	0.062	0.047	0.007
254 <sup>b</sup>	5794240	3820	0.536	0.034	0.574	0.064	0.053	0.008
314 <sup>c</sup>	7603200	3871	0.553	0.034	0.593	0.065	0.060	0.008
940 <sup>d</sup>	9479273	5085	0.938	0.076	0.885	0.284	0.416	0.038
952 <sup>e</sup>	9787239	3801	0.529	0.035	0.566	0.063	0.051	0.008
961 <sup>f</sup>	8561063	3241	0.189	0.055	0.143	0.073	0.003	0.003
1361 <sup>g</sup>	6960913	4158	0.630	0.035	0.665	0.070	0.107	0.013
2704 <sup>h</sup>	9730163	3157	<0.187	...	<0.159	...	<0.0034	...

**Note.** — (This table is available in its entirety in a machine-readable form in the online journal.)

<sup>a</sup> Error on  $T_{\text{eff}}$  58 K for all stars (see Section 5.1).

<sup>b</sup> Consistent with  $T_{\text{eff}} = 3820 \pm 90$  K  $R_* = 0.55 \pm 0.11$  and  $M_* = 0.59 \pm 0.06$  from Johnson et al. (2012).

<sup>c</sup> Consistent with  $R_* = 0.54 \pm 0.05$ ,  $M_* = 0.57 \pm 0.05$  from Pineda et al. (2013).

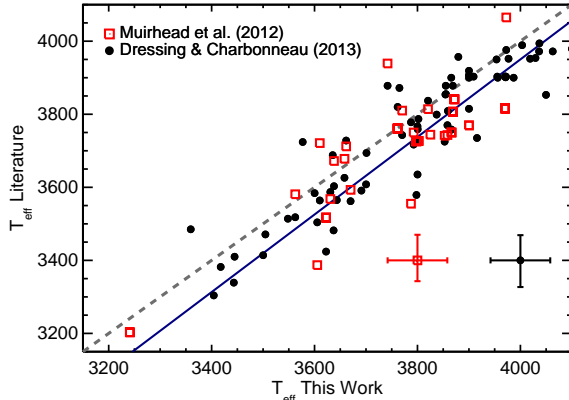
<sup>d</sup> KOI 940 and KOI 2704 have  $T_{\text{eff}}$  values outside our grid of calibration stars.

<sup>e</sup> Kepler-32. Consistent with  $T_{\text{eff}} = 3793^{+80}_{-74}$ ,  $R_* = 0.53 \pm 0.02$ ,  $M_* = 0.54 \pm 0.02$  from Swift et al. (2013).

<sup>f</sup> Kepler-42. Consistent with  $T_{\text{eff}} = 3068 \pm 174$ ,  $R_* = 0.17 \pm 0.04$ ,  $M_* = 0.13 \pm 0.05$  from Muirhead et al. (2012b).

<sup>g</sup> Kepler-61. Consistent with  $T_{\text{eff}} = 4017^{+68}_{-150}$ ,  $R_* = 0.62^{+0.02}_{-0.05}$ , and  $M_* = 0.635 \pm 0.037$  from Ballard et al. (2013).

<sup>h</sup> KOI 2704 has a  $T_{\text{eff}}$  cooler than our grid of calibration stars. We conservatively assign upper limits to this star based on the calibration star with the best matched temperature (GJ 699).



**Figure 12.**  $T_{\text{eff}}$  values from Dressing & Charbonneau (2013) (black circles) and Muirhead et al. (2012a) (red squares) vs. those found in this work. The gray dashed line indicates equality. The solid blue line denotes equality if we used the AGSS models instead of CIFIST (see Section 5.1). Typical error bars are shown in the bottom right color-coded by their corresponding source.

abundances (Hauschildt et al. 1999a,b). Had we used AGSS models in our model-fitting procedure, the median differences would have been  $0^{+7}_{-7}$  K and  $6^{+6}_{-8}$  (both consistent with no difference) for Dressing & Charbonneau (2013) and Muirhead et al. (2012a), respectively.

## 7. PLANETS IN THE HABITABLE ZONES OF *M DWARFS*

Revised estimates of  $T_{\text{eff}}$ ,  $R_*$ , and  $L_*$  allow us to determine the irradiance,  $S$ , and radius,  $R_P$ , of the 188 planets around these stars and determine their position with respect to the HZs of their host stars (Dressing & Charbonneau 2013; Kopparapu 2013; Gaidos 2013). We estimated  $S$  experienced by each of the 191 KOIs around these late-type dwarfs using the  $L_*$

**Table 4**  
*Kepler* planet parameters

KOI #	KID #	$R_P$ $R_{\oplus}$	$\sigma_{R_P}$ $R_{\oplus}$	$S$ $S_{\oplus}$	$\sigma_S$ $S_{\oplus}$
227.01	6185476	2.853	0.114	7.151	0.897
247.01	11852982	1.791	0.127	6.421	0.900
248.01	5364071	2.036	0.083	19.0	2.5
248.02	5364071	2.991	0.072	10.9	1.5
248.03	5364071	2.036	0.073	74.9	10.0
248.04	5364071	2.163	0.083	5.373	0.717
2704.01 <sup>a</sup>	9730163	0.835	0.667	3.7	...
2704.02 <sup>a</sup>	9730163	0.576	0.672	7.0	...

**Note.** — (This table is available in its entirety in a machine-readable form in the online journal.)

<sup>a</sup> KOI 2704 has a  $T_{\text{eff}}$  below our coolest calibration star. Therefore only upper limits are given.

values above, the orbital periods as from *Kepler*, and the  $M_*$  from our effective  $T_{\text{eff}}$ - $M_*$  relation (see Section 4.3). Assuming near-circular orbits, the orbit-averaged  $S$  in terrestrial units is given by:

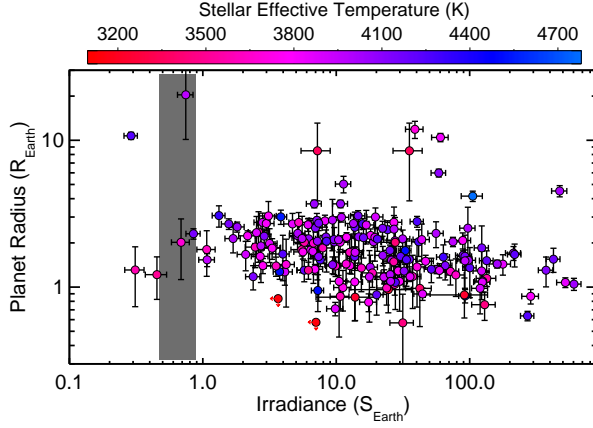
$$\frac{S}{S_{\oplus}} \simeq \frac{L_*}{L_{\odot}} \left( \frac{P}{365 \text{ days}} \right)^{-4/3} \left( \frac{M_*}{M_{\odot}} \right)^{-2/3}. \quad (14)$$

We estimated the planet radius using  $R_P/R_*$  from Dressing & Charbonneau (2013), where available, and from the KOI catalog in the MAST database<sup>4</sup>, otherwise. Figure 13 shows  $R_P$  vs.  $S$  with the points color-coded by the  $T_{\text{eff}}$  of the host star. Errors in radius account for the errors in  $R_*$  and  $R_P/R_*$ ; the latter are taken from MAST as Dressing & Charbonneau (2013) do not report these values. Errors in  $R_P/R_*$  dominate for the smaller planets. Estimated  $R_P$  and  $S$  are given in Table 4.

We adopt a conservative definition of the HZ as the

<sup>4</sup> <http://archive.stsci.edu/kepler/koi/search.php>



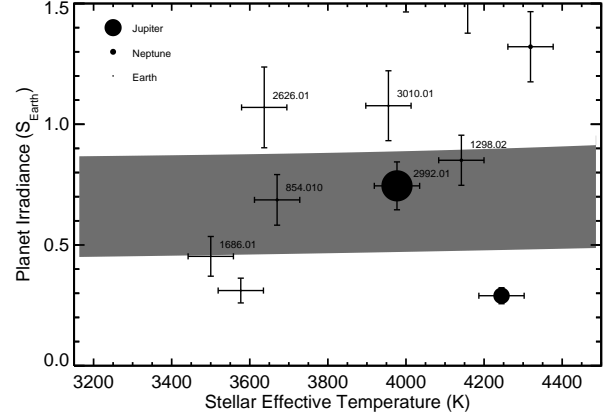


**Figure 13.**  $R_P$  of *Kepler* planets around M dwarf stars vs. estimated  $S$  (both in terrestrial units). Points are color-coded by their stellar  $T_{\text{eff}}$  following the same color scheme as Figures 2, 5, and 6. The  $R_P$  and  $S$  values of KOIs 2704.01 and 2704.02 are assigned upper limits because the late M host stars is outside our range of calibrated stellar parameters. The bounds of the HZ for a  $T_{\text{eff}}=3800$  K star are shown as a shaded region (Kopparapu et al. 2013b).

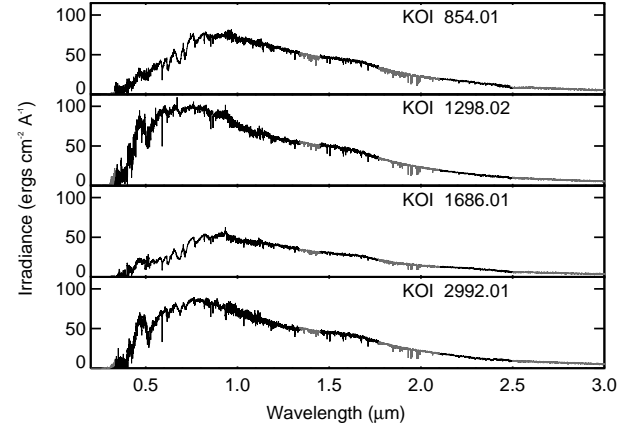
range of  $S$  bracketed by the runaway greenhouse limit and the first  $\text{CO}_2$  condensation limit for an Earth-like atmosphere on a one Earth-mass ( $M_\oplus$ ) planet. We use the empirical relations described in Kopparapu et al. (2013b), corrected in Kopparapu et al. (2013a) and available in the accompanying online calculator. These calculations do not account for water clouds other than a decrease in the effective planetary albedo. Moreover, these criteria will not apply to planets with a very un-Earth-like atmosphere (Pierrehumbert & Gaidos 2011) or much more massive planets where the surface gravity (and hence pressure, for the same optical depth) is much higher.

Figure 14 plots  $S$  (in terrestrial units) vs.  $T_{\text{eff}}$  for planets near or in the HZ. Because estimates of the  $L_*$  and  $M_*$  of stars are covariant, errors in the quantity  $L_* M_*^{-2/3}$  were calculated by MC (Section 4.3). We ignore errors in  $P$ , which are comparatively small. Four candidate planets (KOIs 854.01, 1298.02, 1686.01, and 2992.01) lie within or within one standard deviation of the HZ. Based on the  $R_P/R_*$  from MAST, we estimate the radius of KOI 2992.01 to be  $\simeq 20R_\oplus$  and thus would not be a planet. However, the impact parameter given in MAST is 1.28, suggesting possible problems with the analysis of the *Kepler* light curve. Our upper limit on the  $L_*$  of the late M dwarf system KOI 2704 means that either or both of two of its candidate planets (components 0.01 and 0.02) might be in the HZ, although we are unable to make reliable assignments at this time. Two other candidate planets (KOI 2626.01 and KOI 3010.01) lie marginally ( $1.2$  and  $1.5\sigma$ ) interior to the HZ.

We find that two of the four Earth-size ( $0.5\text{--}1.4R_\oplus$ ) planets identified by Kopparapu (2013) as orbiting in the HZ are actually outside the HZ: KOIs 1422.02, 2418.01, and 2626.01. Although KOI 2626.01 is only  $1.2\sigma$  outside the HZ, KOI 1422.02 is significantly ( $4.1\sigma$ ) interior to the HZ. A third planet, KOI 2418.01, is significantly ( $2.7\sigma$ ) outside the  $\text{CO}_2$  condensation definition of the limit of the HZ, but falls within the ‘maximum  $\text{CO}_2$  greenhouse definition of the outer limit used by Kopparapu (2013).



**Figure 14.** Estimated  $S$  of *Kepler* planets around M dwarf stars (terrestrial units) vs. the  $T_{\text{eff}}$  of the host star. The upper and lower bounds of the HZ (shaded region) as defined by the runaway greenhouse and first  $\text{CO}_2$  condensation limits are calculated based on Kopparapu et al. (2013b). Data points are scaled in size by the estimated planet size. For the smallest planets the plotted points are smaller than the thickness of the error bars, while for the largest planets the error bars are smaller than the point. Objects within  $2\sigma$  of the HZ are labeled. Most planets are off scale above the top of the plot.

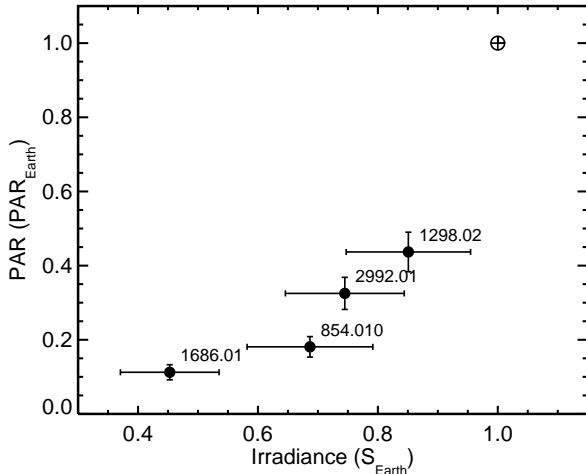


**Figure 15.** Spectra of the stellar irradiance experienced by four candidate planets in (or within  $1\sigma$ ) of the HZ, scaled using the estimated  $L_*$  of each host star. Gray points interpolate across observationally inaccessible wavelength regions using the best-fit PHOENIX model.

These differences are primarily due to our revised estimates of  $T_{\text{eff}}$  and indicate the importance of accurate  $T_{\text{eff}}$  estimates for stars without parallaxes. Removal of these two objects from the tally of Earth-size HZ planets (our additions of 854.01, 1298.02, and 2992.01 have radii of  $\geq 2R_\oplus$ ) reduces the Kopparapu (2013) estimated value of  $\eta_\oplus$ .

Figure 15 plots the stellar  $S$  spectra experienced by the four candidate HZ planets, scaled using the estimated  $L_*$ . We interpolate across gaps in wavelength regions outside the observationally accessible windows using the best-fit PHOENIX models (gray regions).

The spectra of M dwarfs are characteristically less intense at blue wavelengths relative to the Sun. On a planet in the HZ of an M dwarf there would be comparatively less photosynthetic active radiation (PAR, 400–700 nm) available for hypothetical life capable of oxygenic photosynthesis. Wolstencroft & Raven (2002)



**Figure 16.** Photosynthetic active radiation (top of the atmosphere) in terrestrial units vs. total irradiance (top of the atmosphere) in terrestrial units for *Kepler* planet candidates in the HZs of late K and early M dwarfs. Earth is shown for reference.

show that a planet at the inner edge of the HZ of an M0 dwarf would receive about one third of PAR photons compared to an equivalent planet around a G2 dwarf. Using our spectra we calculated the incident normal flux in PAR received at the orbit of each candidate HZ planet. We normalized by the total solar irradiance in PAR ( $530.1 \text{ W m}^{-2}$ ) based on the 2000 ASTM Standard Extraterrestrial Spectrum REference E-490-00<sup>5</sup>. We find that these candidate HZ planets receive anywhere from 10% to 45% of the PAR as the Earth (Figure 16); this varies somewhat from the of Wolstencroft & Raven (2002) in part because these stars cover a range of spectral types. However, lest this dim prospects for plants on M dwarf planets, the reduced PAR is still orders of magnitude higher than the minimum light requirement of terrestrial oxygen-evolving photosynthetic life (Cockell et al. 2009). Indeed, operation of the photosystem II moiety of oxygenic photosynthesis can be inhibited by high light conditions (photoinhibition, Long et al. 1994), a phenomenon which would be relieved in the HZ of an M dwarf. Moreover, two chlorophylls (*d* and *f*) have been recently identified as capable of harvesting light as red as 750 nm for oxygenic photosynthesis (Chen & Blankenship 2011). Such adaptations may be important for life around a red dwarf.

## 8. DISCUSSION

Making use of nearby, well-characterized K and M dwarfs, we develop and calibrate techniques to determine the physical characteristics of late-type dwarfs from moderate-resolution spectra ( $1000 \leq R \leq 2000$ ). Our method is most effective for faint and/or distant stars lacking parallaxes, and is observationally inexpensive compared to techniques that require high-resolution spectra (e.g., Pineda et al. 2013).

As part of our analysis, we take spectra of and re-derive  $T_{\text{eff}}$ ,  $M_*$ , and  $L_*$  for a sample of stars with interferometrically measured radii. Our spectra of these stars are available online. Our derived  $F_{\text{BOL}}$  values are systematically higher than those previously estimated by

B12 (Figure 2). Although these differences are small, they are statistically significant compared to the random errors ( $\lesssim 2\%$  in  $F_{\text{BOL}}$ ) for the calibration stars. The higher  $F_{\text{BOL}}$  values result in 1% higher  $T_{\text{eff}}$  and 4% higher  $L_*$  compared to B12 and corresponding changes in the derived empirical relations between  $T_{\text{eff}}$  and other physical parameters. We show the relations derived by B12 for  $T_{\text{eff}}-R_*$  and  $T_{\text{eff}}-L_*$  next to our own in Figure 4 to illustrate the size of these differences. Because of other sources of error (e.g., sample size), utilizing B12 values for calibration stars would result in only minor changes in our assigned astrophysical parameters for *Kepler* targets. When we re-derive the parameters for *Kepler* planet hosts assuming B12 values for the calibration stars we find that most  $R_*$  values change by  $< 2\sigma$ . Nonetheless the changes are systematic and thus should not be ignored.

We compare PHOENIX model spectra to the visible and NIR spectra of the calibrator stars, and compare the best-fit PHOENIX  $T_{\text{eff}}$  values to the bolometric values. We show that the AGSS version of the PHOENIX models significantly under-predicts the temperatures of M dwarfs, amounting to a difference of 75 K at 3500 K. However, we find that by masking out poorly modeled regions (e.g., at 6500Å), the CIFIST version of PHOENIX accurately reproduces the bolometrically derived temperatures. With this technique we are able to measure  $T_{\text{eff}}$  accurate to 57 K using moderate-resolution optical spectra. The excellent performance of the CIFIST models highlights how much M dwarf atmospheric models have improved in just the last few years.

We estimate the parameters of late-type *Kepler* planet-hosting dwarfs using comparisons of (CIFIST) PHOENIX model spectra to observations, and our calibration between bolometric  $T_{\text{eff}}$  and PHOENIX  $T_{\text{eff}}$ . We utilize our empirical relations to refine the stars'  $T_{\text{eff}}$ ,  $R_*$ ,  $M_*$ , and  $L_*$ . We estimate the radius of each planet using the revised value of  $R_*$ , and the irradiance it experiences using the revised values of  $L_*$  and  $M_*$ . We determine which of these planets are likely to orbit within the circumstellar HZ. Our spectra of these late-type KOIs are available online.

We find that four planets in or within  $1\sigma$  of the HZ: KOI 854.01, 1298.02, 1686.01, and 2992.01. Only one of these (1686.01) is Earth-size, and KOI 2992.01 has a radius of  $\simeq 20R_{\oplus}$  and may not be a planet. This suggests that  $\eta_{\oplus}$  for M dwarfs may need to be revised downward from the Kopparapu (2013) estimate, although the value depends on the *Kepler* detection completeness for small planets, which is still being assessed.

We calculate the irradiation in the 400-700 nm range (useful for oxygenic photosynthesis) incident at the top of the atmosphere for the four HZ planets. Although this is a factor of 3-5 less than that on Earth, it is still adequate to support photosynthetic life. We include spectra of these four stars to aid with modeling of the climate and habitability of these planets.

Our list of HZ stars differs from that of Gaidos (2013), Dressing & Charbonneau (2013) and Kopparapu (2013). For example, Kopparapu (2013) identify KOI 1422.02, 2418.01, and 2626.01 as planets that orbit in their stars' HZ. We find KOI 2626.01 is only  $1.2\sigma$  outside the HZ, and KOI 1422.02 is  $4.1\sigma$  outside the HZ. KOI 2418.01

<sup>5</sup> <http://redc.nrel.gov/solar/spectra/am0/ASTM2000.html>



could be in the HZ depending on the definition used. These discrepancies are largely due to differences in atmospheric models utilized in calculating stellar parameters, and highlight the importance of accurate determinations of  $T_{\text{eff}}$ ,  $L_*$ , and  $M_*$  when assessing the habitability of orbiting planets.

Only one of the candidate HZ planets is Earth-size (KOI 1686.01). However precise radii require not only accurate estimates of  $R_*$ , but also precise values of the transit depth from *Kepler* data (Dressing & Charbonneau 2013) and corrections for limb darkening (Csizmadia et al. 2013). Our  $T_{\text{eff}}$  estimates should allow more robust selections of limb darkening laws for future analyses of the *Kepler* light curves, which in turn can provide better constraints on  $R_P/R_*$ .

We caution that although our quoted errors for individual *Kepler* systems are quite small (roughly 57 K in  $T_{\text{eff}}$ , 7% in  $R_*$ , 11% in  $M_*$ , and 13% in  $L_*$ ), unquantified errors may remain. For example, our sample of calibration stars contains only 22 targets, and they are not evenly distributed over  $T_{\text{eff}}$ . Twenty-two stars is insufficient for a fourth-order (five parameter) fit to the data. Analysis of a much larger sample of stars, such as those with parallaxes and high-resolution spectra (e.g., Pineda et al. 2013), might shed light on systematic errors present in analyses like ours.

Another issue with our analysis of the *Kepler* targets is that our calibration sample only covers  $-0.4 \lesssim [\text{Fe}/\text{H}] \lesssim +0.5$ . Although none of the stars in our *Kepler* sample are more metal rich than this, some are as metal poor as  $[\text{Fe}/\text{H}] \simeq -0.6$  (Mann et al. 2013b; Muirhead et al. 2012b). We see no statistically significant effect of metallicity on the  $T_{\text{eff}}-R_*$  relation, although the existence of subdwarfs, which are smaller than solar-metallicity stars of the same  $T_{\text{eff}}$ , demonstrates that there must be some dependence of radius on the metallicity at sufficiently low  $[\text{Fe}/\text{H}]$ . It is unclear at what point metallicity becomes an important factor, but this is unlikely to significantly alter our derived stellar parameters for *Kepler* targets because none of them are subdwarfs. However, we caution against using calibrations for stars significantly more metal poor than  $[\text{Fe}/\text{H}] \simeq -0.5$  where they are untested.

Improvements could be made in our calibrations with additional interferometric measurements for M dwarfs. In particular, the coolest star in our calibration sample is  $3238 \pm 12$  K, but several *Kepler* planet hosts are near or below that boundary (KOI 2704  $T_{\text{eff}} = 3157 \pm 57$ , KOI 961  $T_{\text{eff}} = 3241 \pm 57$ ). As mentioned above, none of the stars in the calibration sample have metallicities below  $[\text{Fe}/\text{H}] = -0.40$ . However, the number of dwarfs with  $T_{\text{eff}} < 3400$  and/or  $[\text{Fe}/\text{H}] < -0.4$  that are nearby and bright enough for existing facilities like the CHARA array is rather small. Significant improvements in this regime will likely require interferometers with larger apertures or improved instrumentation, e.g. AO for the CHARA array will allow it to reach fainter targets (ten Brummelaar et al. 2012).

We thank the anonymous reviewer for their productive comments on the manuscript. We thank Tabettha Boyajian and Kaspar von Braun for their useful discussions on GJ 725B and other stars with interferometrically determined radii, France Allard and Derek Homeier for their

advice with the PHOENIX models, and Kimberly Aller for her assistance with some of the aesthetic elements of the manuscript.

This work was supported by NSF grant AST-0908419, NASA grants NNX10AI90G and NNX11AC33G (Origins of Solar Systems) to EG. SNIFS on the UH 2.2-m telescope is part of the Nearby Supernova Factory project, a scientific collaboration among the Centre de Recherche Astronomique de Lyon, Institut de Physique Nuclaire de Lyon, Laboratoire de Physique Nuclaire et des Hautes Energies, Lawrence Berkeley National Laboratory, Yale University, University of Bonn, Max Planck Institute for Astrophysics, Tsinghua Center for Astrophysics, and the Centre de Physique des Particules de Marseille. Based on data from the Infrared Telescope Facility, which is operated by the University of Hawaii under Cooperative Agreement no. NNX-08AE38A with the National Aeronautics and Space Administration, Science Mission Directorate, Planetary Astronomy Program. This paper includes data collected by the *Kepler* mission. Funding for the *Kepler* mission is provided by the NASA Science Mission directorate. Some/all of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts.

*Facilities:* IRTF, UH:2.2m, Kepler

## REFERENCES

- Aldering, G., Adam, G., Antilogus, P., et al. 2002, in Proc. SPIE, ed. J. A. Tyson & S. Wolff, Vol. 4836, 61–72
- Aldering, G., Antilogus, P., Bailey, S., et al. 2006, ApJ, 650, 510
- Allard, F., Homeier, D., & Freytag, B. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 448, 16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. C. Johns-Krull, M. K. Browning, & A. A. West, 91
- Allard, F., Homeier, D., Freytag, B., et al. 2013, Memorie della Societa Astronomica Italiana Supplementi, 24, 128
- Allard, F., Homeier, D., Freytag, B., & Sharp, C. M. 2012, in EAS Publications Series, Vol. 57, EAS Publications Series, 3–43
- Anderson, E., & Francis, C. 2011, VizieR Online Data Catalog, 5137, 0
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
- Bacon, R., Copin, Y., Monnet, G., et al. 2001, MNRAS, 326, 23
- Ballard, S., Charbonneau, D., Fressin, F., et al. 2013, ApJ, 773, 98
- Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, A&A, 337, 403
- Batalha, N. M., Borucki, W. J., Koch, D. G., et al. 2010, ApJ, 713, L109
- Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, ApJS, 204, 24
- Bean, J. L., Sneden, C., Hauschildt, P. H., Johns-Krull, C. M., & Benedict, G. F. 2006, ApJ, 652, 1604
- Berta, Z. K., Irwin, J., & Charbonneau, D. 2013, ApJ, 775, 91
- Bessel, M. S. 1990, A&AS, 83, 357
- Bessell, M. S. 1999, PASP, 111, 1426
- Bochanski, J. J., West, A. A., Hawley, S. L., & Covey, K. R. 2007, AJ, 133, 531
- Bohlin, R. C., Dickinson, M. E., & Calzetti, D. 2001, AJ, 122, 2118
- Bonfils, X., Delfosse, X., Udry, S., et al. 2013, A&A, 549, A109
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
- Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, ApJ, 736, 19
- Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012, ApJ, 757, 112

- Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, *AJ*, 142, 112
- Buton, C., Copin, Y., Aldering, G., et al. 2013, *A&A*, 549, A8
- Caffau, E., Ludwig, H.-G., Steffen, M., Freytag, B., & Bonifacio, P. 2011, *Sol. Phys.*, 268, 255
- Chabrier, G. 2003, *PASP*, 115, 763
- Chen, M., & Blankenship, R. E. 2011, *Trends in Plant Sciences*, 16, 427
- Christou, J. C., & Drummond, J. D. 2006, *AJ*, 131, 3100
- Cockell, C. S., Raven, J. A., Kaltenecker, L., & Logan, R. C. 2009, *Plant Ecology & Diversity*, 2, 207
- Cousins, A. W. J. 1980, *South African Astronomical Observatory Circular*, 1, 166
- Covey, K. R., Lada, C. J., Román-Zúñiga, C., et al. 2010, *ApJ*, 722, 971
- Cowley, A. P., Hiltner, W. A., & Witt, A. N. 1967, *AJ*, 72, 1334
- Csizmadia, S., Pasternacki, T., Dreyer, C., et al. 2013, *A&A*, 549, A9
- Cushing, M. C., Rayner, J. T., & Vacca, W. D. 2005, *ApJ*, 623, 1115
- Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, *PASP*, 116, 362
- Cushing, M. C., Marley, M. S., Saumon, D., et al. 2008, *ApJ*, 678, 1372
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *VizieR Online Data Catalog*, 2246, 0
- Delfosse, X., Forveille, T., Ségransan, D., et al. 2000, *A&A*, 364, 217
- Demory, B.-O., Ségransan, D., Forveille, T., et al. 2009, *A&A*, 505, 205
- Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, *ApJS*, 178, 89
- Dressing, C. D., & Charbonneau, D. 2013, *ApJ*, 767, 95
- Edvardsson, B., Andersen, J., Gustafsson, B., et al. 1993, *A&A*, 275, 101
- Eggen, O. J. 1979, *ApJS*, 39, 89
- Erro, B. I. 1971, *Boletín del Instituto de Tonantzintla*, 6, 143
- Fressin, F., Torres, G., Charbonneau, D., et al. 2013, *ApJ*, 766, 81
- Gaidos, E. 2013, *ApJ*, 770, 90
- Gaidos, E., Deschenes, B., Dundon, L., et al. 2005, *Astrobiology*, 5, 100
- Gaidos, E., Fischer, D. A., Mann, A. W., & Lépine, S. 2012, *ApJ*, 746, 36
- Glass, I. S. 1975, *MNRAS*, 171, 19P
- Grevesse, N., & Sauval, A. J. 1998, *Space Sci. Rev.*, 85, 161
- Hamuy, M., Walker, A. R., Suntzeff, N. B., et al. 1992, *PASP*, 104, 533
- Hauschildt, P. H., Allard, F., & Baron, E. 1999a, *ApJ*, 512, 377
- Hauschildt, P. H., Allard, F., Ferguson, J., Baron, E., & Alexander, D. R. 1999b, *ApJ*, 525, 871
- Henry, T. J., Kirkpatrick, J. D., & Simons, D. A. 1994, *AJ*, 108, 1437
- Henry, T. J., & McCarthy, Jr., D. W. 1993, *AJ*, 106, 773
- Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, *ApJS*, 201, 15
- Johnson, H. L. 1964, *Boletín de los Observatorios Tonantzintla y Tacubaya*, 3, 305
- . 1965, *ApJ*, 141, 170
- Johnson, H. L., MacArthur, J. W., & Mitchell, R. I. 1968, *ApJ*, 152, 465
- Johnson, H. L., Mitchell, R. I., Iriarte, B., & Wisniewski, W. Z. 1966, *Communications of the Lunar and Planetary Laboratory*, 4, 99
- Johnson, H. L., & Morgan, W. W. 1953, *ApJ*, 117, 313
- Johnson, J. A., Gazak, J. Z., Apps, K., et al. 2012, *AJ*, 143, 111
- Joshi, M. M., & Haberle, R. M. 2012, *Astrobiology*, 12, 3
- Joshi, M. M., Haberle, R. M., & Reynolds, R. T. 1997, *Icarus*, 129, 450
- Kirkpatrick, J. D., Henry, T. J., & McCarthy, Jr., D. W. 1991, *ApJS*, 77, 417
- Koen, C., Kilenny, D., van Wyk, F., & Marang, F. 2010, *MNRAS*, 403, 1949
- Kopparapu, R. K. 2013, *ApJ*, 767, L8
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013a, *ApJ*, 770, 82
- . 2013b, *ApJ*, 765, 131
- Lammer, H., Lichtenegger, H. I. M., Kulikov, Y. N., et al. 2007, *Astrobiology*, 7, 185
- Lantz, B., Aldering, G., Antilogus, P., et al. 2004, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 5249, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ed. L. Mazuray, P. J. Rogers, & R. Wartmann, 146–155
- Lépine, S., Hilton, E. J., Mann, A. W., et al. 2013, *AJ*, 145, 102
- Lissauer, J. J. 2007, *ApJ*, 660, L149
- Long, S. P., Humphries, S., & Falkowski, P. G. 1994, *Annual Review of Plant Physiology and Plant Molecular Biology*, 45, 633
- Mann, A. W., Brewer, J. M., Gaidos, E., Lépine, S., & Hilton, E. J. 2013a, *AJ*, 145, 52
- Mann, A. W., Gaidos, E., & Aldering, G. 2011, *PASP*, 123, 1273
- Mann, A. W., Gaidos, E., Kraus, A., & Hilton, E. J. 2013b, *ApJ*, 770, 43
- Mann, A. W., Gaidos, E., Lépine, S., & Hilton, E. J. 2012, *ApJ*, 753, 90
- Markwardt, C. B. 2009, in *Astronomical Society of the Pacific Conference Series*, Vol. 411, *Astronomical Data Analysis Software and Systems XVIII*, ed. D. A. Bohlender, D. Durand, & P. Dowler, 251
- Menou, K. 2013, *ApJ*, 774, 51
- Mermilliod, J.-C., Mermilliod, M., & Hauck, B. 1997, *A&AS*, 124, 349
- Mohanty, S., & Basri, G. 2003, *ApJ*, 583, 451
- Morel, M., & Magnenat, P. 1978, *A&AS*, 34, 477
- Morton, T. D., & Swift, J. J. 2013, *ArXiv e-prints*, arXiv:1303.3013
- Mould, J. R. 1976, *A&A*, 48, 443
- Mould, J. R., & Hyland, A. R. 1976, *ApJ*, 208, 399
- Muirhead, P. S., Hamren, K., Schlawin, E., et al. 2012a, *ApJ*, 750, L37
- Muirhead, P. S., Johnson, J. A., Apps, K., et al. 2012b, *ApJ*, 747, 144
- Muirhead, P. S., Vanderburg, A., Shporer, A., et al. 2013, *ApJ*, 767, 111
- Neves, V., Bonfils, X., Santos, N. C., et al. 2012, *A&A*, 538, A25
- Oke, J. B. 1990, *AJ*, 99, 1621
- Önehag, A., Heiter, U., Gustafsson, B., et al. 2012, *A&A*, 542, A33
- Persson, S. E., Aaronson, M., & Frogel, J. A. 1977, *AJ*, 82, 729
- Pickles, A. J. 1998, *PASP*, 110, 863
- Pierrehumbert, R., & Gaidos, E. 2011, *ApJ*, 734, L13
- Pineda, J. S., Bottom, M., & Johnson, J. A. 2013, *ApJ*, 767, 28
- Plez, B. 1998, *A&A*, 337, 495
- Rajpurohit, A. S., Reylé, C., Allard, F., et al. 2013, *A&A*, 556, A15
- Rayner, J. T., Cushing, M. C., & Vacca, W. D. 2009, *ApJS*, 185, 289
- Rayner, J. T., Toomey, D. W., Onaka, P. M., et al. 2003, *PASP*, 115, 362
- Reid, I. N., Cruz, K. L., Allen, P., et al. 2004, *AJ*, 128, 463
- Reylé, C., Rajpurohit, A. S., Schultheis, M., & Allard, F. 2011, in *Astronomical Society of the Pacific Conference Series*, Vol. 448, *16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*, ed. C. Johns-Krull, M. K. Browning, & A. A. West, 929
- Rojas-Ayala, B., Covey, K. R., Muirhead, P. S., & Lloyd, J. P. 2012, *ApJ*, 748, 93
- Segura, A., Walkowicz, L. M., Meadows, V., Kasting, J., & Hawley, S. 2010, *Astrobiology*, 10, 751
- Selsis, F., Chazelas, B., Bordé, P., et al. 2007, *Icarus*, 191, 453
- Shields, A. L., Meadows, V. S., Bitz, C. M., et al. 2013, *Astrobiology*, 13, 715
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, 131, 1163
- Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *AJ*, 123, 485
- Swift, J. J., Johnson, J. A., Morton, T. D., et al. 2013, *ApJ*, 764, 105
- Tarter, J. C., Backus, P. R., Mancinelli, R. L., et al. 2007, *Astrobiology*, 7, 30

- ten Brummelaar, T. A., Sturmann, L., Sturmann, J., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8447, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, *PASP*, 115, 389
- van Belle, G. T., van Belle, G., Creech-Eakman, M. J., et al. 2008, *ApJS*, 176, 276
- van Leeuwen, F. 2007, *A&A*, 474, 653
- van Leeuwen, F., & Fantino, E. 2005, *A&A*, 439, 791
- Veeder, G. J. 1974, *AJ*, 79, 1056
- von Braun, K., Boyajian, T. S., Kane, S. R., et al. 2011, *ApJ*, 729, L26
- . 2012, *ApJ*, 753, 171
- Weis, E. W. 1993, *AJ*, 105, 1962
- West, A. A., & Basri, G. 2009, *ApJ*, 693, 1283
- Wickramasinghe, D. T., Bessell, M. S., & Whelan, J. A. J. 1978, *MNRAS*, 182, 53P
- Wolstencroft, R. D., & Raven, J. A. 2002, *Icarus*, 157, 535