

# Exploring the M-dwarf Luminosity - Temperature - Radius Relationships using Gaia DR2

Sam Morrell<sup>★</sup> and Tim Naylor

*School of Physics, University of Exeter, Exeter, EX4 4QL, UK*

Accepted 2019 August 03. Received 2019 August 03; in original form 2019 May 10

## ABSTRACT

There is growing evidence that M-dwarf stars suffer radius inflation when compared to theoretical models, suggesting that models are missing some key physics required to completely describe stars at effective temperatures ( $T_{\text{SED}}$ ) less than about 4000K. The advent of Gaia DR2 distances finally makes available large datasets to determine the nature and extent of this effect. We employ an all-sky sample, comprising of >15 000 stars, to determine empirical relationships between luminosity, temperature and radius. This is accomplished using only geometric distances and multiwave-band photometry, by utilising a modified spectral energy distribution fitting method. The radii we measure show an inflation of 3 – 7% compared to models, but no more than a 1 – 2% intrinsic spread in the inflated sequence. We show that we are currently able to determine M-dwarf radii to an accuracy of 2.4% using our method. However, we determine that this is limited by the precision of metallicity measurements, which contribute 1.7% to the measured radius scatter. We also present evidence that stellar magnetism is currently unable to explain radius inflation in M-dwarfs.

**Key words:** techniques: photometric – stars: low-mass – stars: late-type – stars: fundamental parameters

## 1 INTRODUCTION

It has been clear for a long time that something may be seriously wrong with our models of main-sequence M-dwarfs. Historically it has been colour-magnitude diagrams (CMDs) of open clusters (e.g. Stauffer et al. 2007; Bell et al. 2012; Thompson et al. 2014) which have demonstrated these problems, showing that the observed colours for stars below  $0.7M_{\odot}$  can be up to half a magnitude redder than the model colours. Problematically, CMDs give little insight as to where in the models the problems lie, because from a physical point of view colours and magnitudes are far removed from the fundamental parameters predicted by models. The primary outputs from these models are the radius ( $R$ ), effective temperature ( $T_{\text{eff}}$ ) and luminosity ( $L$ ) for a star of a given mass. Hence measuring these quantities more directly is a promising way of testing the models, and in recent years such techniques have become sufficiently precise to be useful in this context.

### 1.1 Measuring the radii of M-dwarfs

Currently there are three commonly used methods to measure the radii of M-dwarfs, to which this paper seeks to add a fourth. The double-lined eclipsing binaries (DEBs) provide a direct, fundamental test of the models because they yield radii at a given mass. Again it has long been known that the eclipsing binaries show that M-star

radii appear to be inflated for their mass (see for example Figure 2 of Chen et al. 2014; Chaturvedi et al. 2018; Parsons et al. 2018; Mann et al. 2019), though Torres (2013) highlighted that the effect is subtle. Furthermore there can be significant discrepancies between parameter determinations for the same binary, see for example the discussion of T-Cyg1-12664 in Han et al. (2017), or the difference in measured radius for the secondary in PTFEB132.707+19.810/AD 3814 (Kraus et al. 2017; Gillen et al. 2017). Nevertheless, at face value the failure of the models to match the eclipsing binaries in the mass-radius plane implies that the stellar structure models are incorrect. However, there is a concern that binarism may affect stellar structure; specifically that if stars are tidally locked their high rotation rate will inflate them, and indeed some authors find evidence for this (e.g. the careful work of Kraus et al. 2011). Hence the inflated radii measured for eclipsing binaries may not indicate a problem in single star models.

An alternative test for radius inflation is to obtain radii for the bright M-stars which are accessible to interferometers. Given astrometric distances, interferometry provides a direct measure of the radius, but no direct measure of the mass. Instead we must use one of the remaining parameters ( $T_{\text{eff}}$  or  $L$ ) as a surrogate for mass, and it is probably  $L$  which most closely follows mass and has little dependence on radius. The luminosity is dictated by the thermodynamic properties in the core, and whilst this is dependent to a small degree on the radius, practically the central concentration of the star ensures that the dependence is weak. Conversely the temperature is decided by the outer layers of the star and so has a strong depen-

<sup>★</sup> E-mail: smorrell@astro.ex.ac.uk

dence on radius. Hence [Boyajian et al. \(2012\)](#) measure luminosities (from broad-band photometry) and interferometric radii for a sample of 21 M-stars, and conclude that in the radius-luminosity plane the stars are inflated by about 5 percent, though again numbers are small.

The work of [Mann et al. \(2013\)](#) and [Mann et al. \(2015\)](#) overcomes the small-number-statistics problem by taking a sample of 183 M-stars with known distances and measuring their luminosities (again from broad-band photometry). They use optical spectra to measure temperatures, which can then be converted into radii using the definition of effective temperature. Their temperatures are derived by matching the lines and bands in their observed spectra to those of model atmospheres, hence we refer to this as a spectroscopic temperature,  $T_{\text{sp}}$ . This relies on the detailed opacities of individual lines being correct, as well as the structure of the model atmosphere. Such temperature measurements can depend on the wavelength range and chemical species from which they are determined, which is especially problematic for M-dwarfs. For example determinations of temperature from optical ([Mann et al. 2015](#); [Cortés-Contreras et al. 2017](#)) and near infrared [Newton et al. \(2014\)](#) spectra can differ by up to 1.5 spectral sub-types; around 5 – 10% in  $T_{\text{sp}}$ . At the root this technique relies on  $T_{\text{eff}}$  being equal to  $T_{\text{sp}}$ . That said, when plotted in the luminosity-temperature plane there is evidence either that the binaries are inflated with respect to the [Mann et al. \(2015\)](#) sample (which is itself inflated with respect to the models) or there is a mismatch in the temperature scale.

In summary there are three methods to relatively directly measure M-dwarf radii, each suffering to greater or lesser degree from problems of small-number statistics, rapid rotation or questions of how closely  $T_{\text{sp}}$  matches  $T_{\text{eff}}$ . Hence, in this paper we introduce a fourth method of measuring the radius, using the overall shape of the spectral energy distribution (SED) to measure what we will call  $T_{\text{SED}}$ . By fitting the broad-band photometry of a star with synthetic photometry of a model atmosphere we can derive both  $T_{\text{SED}}$  and, given the distances from Gaia DR2, the luminosity. This allows us to create a sample of 15 274 single stars which gives us the statistics to show that single stars and stars in binaries show a similar degree of inflation with respect to the models. However, whether  $T_{\text{SED}}$  and  $T_{\text{sp}}$  are in agreement remains for debate.

In addition this large sample allows us to address three further problems.

## 1.2 Is there a correlation between [M/H] and radius?

Metallicity is well known to have an effect on the structure of stars, and as we will discuss in Section 4.2.5 theoretical models predict that there will be a 16 percent change in M-dwarf radius per dex change in [M/H] at a given luminosity. The observational evidence for this dependency is weak, though it does seem to preclude a dependency much stronger than predicted, with the data in Figure 9 of [Parsons et al. \(2018\)](#) ruling out changes of more than about 10% dex<sup>-1</sup> as a function of mass (which should correspond roughly to luminosity). [Mann et al. \(2015\)](#) find a correlation of only about 4% dex<sup>-1</sup> if the change in radius is measured as a function of absolute magnitude (which again should correspond roughly to luminosity). The situation is very different if the radius change is measured as a function of effective temperature, with the same [Mann et al. \(2015\)](#) data yielding 40% dex<sup>-1</sup>, which is similar to the result of [Rabus et al. \(2019\)](#), with an even larger dependence being found for the lowest metallicity M-dwarfs by [Kesseli et al. \(2019\)](#). In fact these results simply show that from a physical point of view measuring the radius change as a function of temperature confuses the issue,

since changing the radius at a fixed luminosity will of itself change the temperature. In this work when addressing metallicity we will, therefore, measure the radius change as a function of luminosity, and using this will show in Section 4.2.5, that our technique shows no relationship between radius and metallicity, but our upper limit is consistent with the theoretical predictions.

## 1.3 What causes the radius inflation?

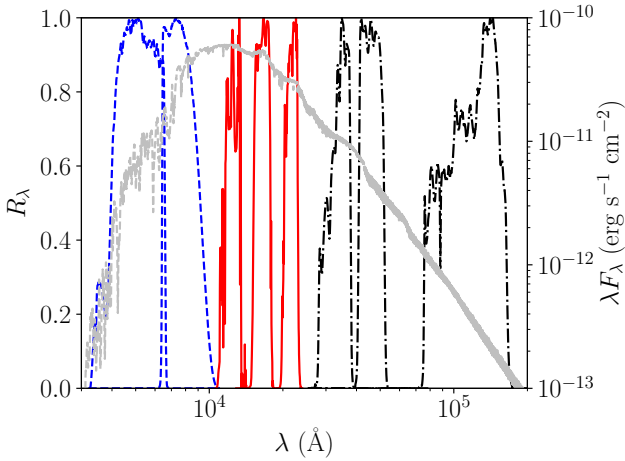
Given that convective flows are the dominant means of energy transport out of the star, many have made the reasonable assertion that inhibiting convection would cause an increase in the stellar radius, and hence explain the inflation of M-dwarfs. This inhibition could be accomplished by the strong magnetic fields ([Mullan & MacDonald 2001](#); [MacDonald & Mullan 2017](#)) that are thought to be the fundamental drivers of both magneto-convection and the delay of convective onset (e.g. [MacDonald & Mullan 2014](#)). Hence there has been much theoretical work over the last decade to develop consistent stellar evolution models that correctly account for dynamo effects and stellar magnetism, including the resulting starspots ([Mullan & MacDonald 2001](#); [Feiden & Chaboyer 2013](#); [Somers & Pinsonneault 2016](#); [MacDonald & Mullan 2017](#)). Recently these effects have been modelled in 1D stellar structure models by [Ireland & Browning \(2018\)](#), who adopt a depth dependent mixing length theory parameter  $\alpha_{\text{MLT}}$ ; emulating the effect of convective inhibition. In their work they observe radii inflated by 10 - 15% when compared to models that do not treat convection in this way, though they caution that such treatments of magnetic inhibition are highly uncertain and may be difficult to calibrate.

Despite this success, the hypothesis that magnetic fields cause inflation has been brought into question by [Kesseli et al. \(2018\)](#), who used a  $v\sin(i)$  technique to show that the radii of samples of rapidly-rotating stars are consistent (to within 5%) with those of slowly-rotating stars, and the inflation in eclipsing binaries consistent with single stars. In addition [Kochukhov & Shulyak \(2019\)](#) conclude that magnetic inflation models do not support their observations of the DEB YY Gem.

Theories which invoke magnetism to explain the radius inflation of M-dwarfs make two predictions we will test in this paper. The first is that unless M-dwarfs all have very similar fields, at a given mass (or luminosity) they should have radii which range between the predictions of the non-magnetic models and some maximum inflation. Our sample is sufficiently large to show that any spread in M-dwarf radii is much less than 1 – 2% (Section 4.2), apparently ruling out the magnetic inflation models. The second prediction is that the degree of radius inflation should correlate with magnetic activity indicators ([López-Morales 2007](#)). Again we find no such relationships (Section 4.2.4), but for indicators such as X-ray activity this could be explained if the surface field is unrelated to the magnetic field in the bulk of the star (see [Brun & Browning 2017](#), for a discussion of how under-developed models of M-stars are). However this criticism cannot be levelled at the absence of a correlation between rotation period (or Rossby number) and inflation we find, which again suggests magnetic fields may not be responsible for radius inflation.

## 1.4 Radii of planets

Whilst understanding the structure of M-stars is a significant problem in its own right, it is currently being given added impetus by the need for precise predictions of M-dwarf radii to understand the



**Figure 1.** The system responses used to generate the synthetic photometry. The photometry is comprised of magnitudes from Gaia (blue dashed), the Two Micron All-Sky Survey (red solid) and WISE (black dot-dashed). For reference, the model spectrum for an M-dwarf star with an effective temperature  $T_{\text{eff}} = 3300\text{K}$  is included (grey dashed).

structure of rocky exoplanets. Their small size makes M-stars the current target-of-choice for rocky exoplanet hunting, since it maximises both the radial velocity and transit signal from a planet of a given mass. However, to use the transit depth to obtain planetary radii and hence densities requires precise M-star masses. The technique developed in this paper gives just such precise radii which can be obtained from archive data, though it can be improved by the addition of a spectroscopic metallicity.

## 2 METHOD

Given a precise parallax one can integrate the area beneath the SED to find the luminosity of the star, while the shape of the SED is a function of temperature. With these, the radius of the star can be calculated. Importantly, both are a function of only the stellar photosphere; allowing the fitting process to be uncoupled from the model interiors. Hence using just synthetic photometry from model atmospheres we have developed a method that uses broadband photometry, readily available from public surveys, to sample the SED and infer these properties. The bands that we used, which comprised broadband photometry from optical, near-infrared and mid-infrared surveys, are detailed in Table 1. The system responses used to generate our synthetic photometry are plotted in Figure 1, along with a model spectrum of an M-dwarf star. This helps justify our motivation for the choice of data and photometric systems. First, all photometric systems correspond to all sky surveys whose system responses are well understood. This means that not only can we draw a sample of stars from the entire sky, but the characterisation of the system responses means that our folded photometry will closely replicate the original photometry. Using these surveys, we also have excellent coverage of the stellar SED, with the optical Gaia prism photometry (Gaia Collaboration et al. 2016, 2018) sampling blue-ward of the blackbody peak, 2MASS NIR photometry (Skrutskie et al. 2006) sampling the peak itself and AllWISE photometry (Wright et al. 2010) constraining the Rayleigh-Jeans tail down to fluxes 3 orders of magnitude lower than the peak. This combination means that we have the ability to accurately constrain the colour of

the objects, and thus the  $T_{\text{SED}}$ , and make a robust estimate of the luminosity of a star from its entire SED.

### 2.1 Target Selection

Each catalogue used is shown in Table 1 along with the relevant publications. We began by selecting all stars from Gaia DR2 with  $r_{\text{est}} < 100\text{pc}$  from Bailer-Jones et al. (2018). We chose this distance cut because the systematic uncertainties in the Gaia DR2 parallaxes generally remain below 0.1mas (Lindegren et al. 2018). This constraint guarantees that our distances are not affected by Gaia’s astrometric systematics while remaining good to 1% uncertainty in luminosity. This preliminary sample comprised 138 279 sources. Each photometric catalogue was then cross matched with this preliminary sample and the quality cuts shown in Table 1 applied. We omitted the  $W_4$  band from the fitting process due to its poor signal-to-noise, and because  $W_3$  does an adequate job of characterising the Rayleigh-Jeans tail of even the coolest stars in the sample. The final input catalogue is then constructed by combining the photometry for only the stars common to all source catalogues. The number of sources remaining in the final input catalogue are 15 765.

### 2.2 Model Grid

The synthetic photometry was produced using the BT-Settl CIFIST stellar atmosphere grid<sup>1</sup> (Allard et al. 2012). These atmospheres are provided in units of mean disc intensity at the stellar surface  $I_\lambda$ , which includes the effect of limb darkening.  $I_\lambda$  is related to the flux observed at the surface of the Earth  $F_\lambda$  by

$$F_\lambda = I_\lambda \frac{R^2}{d^2}, \quad (1)$$

where  $R$  is the stellar radius, which we fitted, and  $d$  is its distance from the Earth (e. g. Girardi et al. 2002). Strictly, this equation should also have a factor that would account for extinction  $A_\lambda$ . However, as our entire sample is within 100pc, the effect of extinction on our photometry is negligible. To reduce computational time we downsampled the wavelength scale of the input stellar atmospheres to  $\approx 1\text{\AA}$  bins using a flux-conserving algorithm. We then performed a bilinear interpolation, first in  $\log(g)$  then in  $T_{\text{eff}}$ , on the 4 nearest downsampled models to produce the final  $I_\lambda$ . We folded  $I_\lambda$  through the system responses to yield the mean apparent synthetic magnitude of a unit surface area of the model atmosphere at its surface within the  $i^{\text{th}}$  band  $Z_i$  with

$$\begin{aligned} m_{i,\text{syn}} &= -2.5 \log_{10} \left[ \frac{\int_\lambda F_\lambda S_{\lambda,i} d\lambda}{\int_\lambda f_{\lambda,i}^\circ S_{\lambda,i} d\lambda} \right] + m_i^\circ \\ &= -2.5 \log_{10} \left[ \frac{\int_\lambda I_\lambda S_{\lambda,i} d\lambda}{\int_\lambda f_{\lambda,i}^\circ S_{\lambda,i} d\lambda} \right] - 5 \log_{10} \left[ \frac{R}{d} \right] + m_i^\circ \\ &= Z_i - 5 \log_{10} \left[ \frac{R}{d} \right], \end{aligned} \quad (2)$$

where  $S_{\lambda,i}$ ,  $f_{\lambda,i}^\circ$  and  $m_i^\circ$  are the system response, zero point flux and zero point of the  $i^{\text{th}}$  band.

<sup>1</sup> [https://phoenix.ens-lyon.fr/Grids/BTSettl/CIFIST2011\\_2015/](https://phoenix.ens-lyon.fr/Grids/BTSettl/CIFIST2011_2015/)

Band	$\lambda_{\text{iso}}$ (Å)	System	Selection Criteria	No. Sources	Coverage	Source
–	–	–	$r_{\text{est}} \leq 100pc$ $(r_{\text{hi}} - r_{\text{lo}})/r_{\text{est}} < 0.01$	138 279	100.0%	Bailer-Jones et al. (2018)
$G_{\text{BP}}$ $G_{\text{RP}}$	5320 7970	Gaia DR2	$G_{\text{BP}} - G_{\text{RP}} > 1.5$ $\pi > 9.0$ $\delta\pi/\pi < 0.02$	99 288	71.8%	Gaia Collaboration et al. (2016) Gaia Collaboration et al. (2018) Passbands: Evans et al. (2018) <sup><math>\alpha, \beta</math></sup>
$J_{2\text{MASS}}$ $H_{2\text{MASS}}$ $K_{s,2\text{MASS}}$	12410 16513 21656	2MASS	ph_qual = "AAA"	55 379	40.0%	Skrutskie et al. (2006)
$W_1$ $W_2$ $W_3$	33792 46293 123338	WISE	ext_flg = 0 cc_flags = "000" ph_qual = A or B $\delta W_3/W_3 < 0.05$	26 182	18.9%	Wright et al. (2010)
All				15 765	11.4%	

**Table 1.** The sources of the photometric data used for the SED fitting described in Section 2.3. To aid in comparison to plots, each filter is listed with its isophotal effective wavelength  $\lambda_{\text{iso}}$ , criteria used to select the photometry, and original source. The bands are grouped into photometric systems and in order of ascending  $\lambda_{\text{iso}}$ . The  $\lambda_{\text{iso}}$  is determined by integrating  $f_{\lambda} = 1$  across the filter, unless directly quoted in the source paper. Due to the poor quality of the photometry, the W4 band has been omitted from the fitting.

<sup>$\alpha$</sup>  Passbands available from [https://www.cosmos.esa.int/web/gaia/iow\\_20180316](https://www.cosmos.esa.int/web/gaia/iow_20180316).

<sup>$\beta$</sup>  As this paper was nearing completion Maíz Apellániz & Weiler (2018) was published which presented further revisions to the Gaia system responses. After careful analysis we found that the differences between the fits resulting from grids consisting of both sets of response curves was contained within the majority of formal uncertainties of our resulting stellar properties.

### 2.3 Free Temperature-Radius Fitting

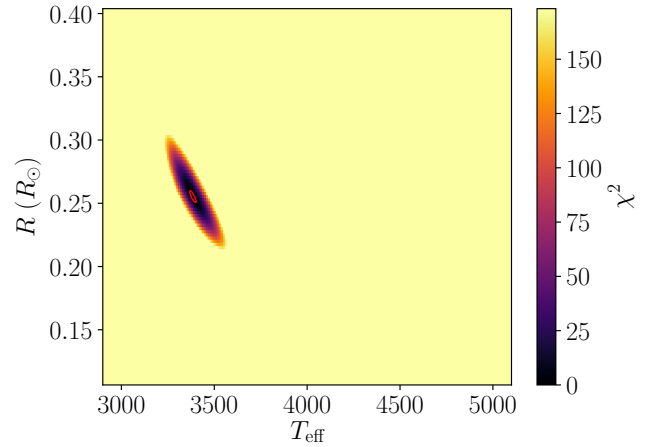
For the fitting, we adapted the spectral energy distribution fit (SEDF) method (e.g. Pecaut & Mamajek 2013; Masana et al. 2006) with an addition to remove the dependence on *a priori* knowledge of the angular radius of the star  $\theta$  (see Section 2.4). This method is similar in principle to the infrared flux method of (Blackwell & Shallis 1977; Blackwell et al. 1979), reviewed in Casagrande (2008). However it is applied across the entire SED, thus generalises well to a wide range of spectral types; including cool stars like M-dwarfs. Essentially we used the shape of the SED to measure the temperature.

We compared our synthetic photometry  $m_{i,\text{syn}}$  to the photometric data  $m_i$  using

$$\chi^2 = \sum_i \left( \frac{m_i - m_{i,\text{syn}}}{\sigma_i} \right)^2 = \sum_i \left( \frac{m_i - Z_i + 5 \log_{10}(R/d)}{\sigma_i} \right)^2, \quad (3)$$

where  $\sigma_i$  is the uncertainty in the  $i$ th photometric band and  $i = G_{\text{BP}}, G_{\text{RP}}, J, H, K_s, W_1, W_2$  and  $W_3$ . For  $\sigma_i$  we adopted a floor value of 0.01 mag, which corresponds to roughly 1%, for all photometric uncertainties in the entire sample. The free parameters for the fit are  $R/d$ ,  $T_{\text{eff}}$  and  $\log(g)$ , which we explored using a simple grid search to generate a 3D cube of  $\chi^2$ . The minimum value of  $\chi^2$  within the grid determines the best fitting solution. We constrained the  $\log(g)$  axis by applying a tophat prior of  $\pm 0.5$  dex around the prescribed model  $\log(g)$ , which was found by matching  $M_G$  of each star with the Baraffe et al. (2015) 4Gyr isochrone.

The grids were first transformed into probability space, using  $P(T_{\text{SED}}, \log(g), R/d) = \exp(-\chi^2/2)$ , and normalised. The 2D  $\chi^2$  space necessary for producing confidence contours was then calculated by marginalising the 3D cube over  $\log(g)$ . We next converted  $R/d$  to  $R$  using the geometric distances of Bailer-Jones et al. (2018), and allowed for uncertainty in distance by convolving each row of constant  $T_{\text{eff}}$  in the 2D  $\chi^2$  space with a Gaussian whose standard deviation was the mean uncertainty from the lower and upper distance



**Figure 2.** The search space for one of our targets whose  $\chi^2$  lies at the median value of the randomly selected uncertainty sample. The red ellipsoid indicates the 68% confidence contour resulting from the process.

bounds; as these are nearly symmetrical. The confidence contours were determined from the resulting 2D PDF by identifying the set of highest probability pixels whose sum was 0.68 and drawing a contour around them. We obtained the uncertainties in  $T_{\text{SED}}$  and  $R$  from a 1D distribution of probability by marginalising along the remaining axis. To illustrate the correlation between  $R$  and  $T_{\text{eff}}$ , we include an example plot of the  $\chi^2$  space resulting from a fit to one of our targets in Figure 2. We randomly picked 158 stars from our input catalogue, 1% of our sample, upon which we performed this full grid search. We used this sample to determine uncertainties characteristic of the full sample. The 68% confidence contours from which those uncertainties are derived are shown in Figure 6.

## 2.4 Analytical Determination of $R$

Performing a full 3D grid search on our entire input catalogue is intractable due to computational time constraints. So for a particular  $T_{\text{eff}}$  and  $\log(g)$  we determined best fitting  $R$  by analytically minimising  $\chi^2$ , effectively making our search space 2D. We accomplished this by differentiating Equation 3 with respect to the  $5 \log_{10}(R/d)$  term and finding the stationary point of the derivative to analytically minimise  $\chi^2$ , yielding

$$\log_{10} \left( \frac{R^2}{d^2} \right) = \log_{10} (\theta^2) = -0.4 \left( \sum_i^N \frac{Z_i - m_i}{\sigma_i^2} \right) / \left( \sum_i^N \frac{1}{\sigma_i^2} \right). \quad (4)$$

This provides the dilution factor  $R^2/d^2$  and hence, by applying the distances from Bailer-Jones et al. (2018), the stellar radius  $R$ . For each  $\log(g)$  and  $T_{\text{eff}}$  we analytically minimised  $\chi^2$  to find the radius in this way. This results in a 2D  $\log(g) - T_{\text{eff}}$  space. An example of the resulting  $\chi^2$  space is shown in Figure 2, and example fits shown in Figure 3.

## 2.5 Flagging

In addition to the initial cuts we performed on the source catalogues, which already produced a stringently constrained sample, we also performed additional post processing to produce flags to be included with the fitted parameters. As the  $G_{\text{BP}}$  and  $G_{\text{RP}}$  fluxes are integrated over a  $3.5 \times 2.5$  arcsec<sup>2</sup> field they are susceptible to contamination from both bright, nearby sources and sky background. Following the method of Evans et al. (2018), we applied the **bad\_phot** photometric contamination flag by  $\sigma$ -clipping sources in the flux excess ratio vs. colour plane. See Figure 4 for the resulting classification after 8 iterations.

The Gaia astrometric data can similarly suffer contamination from crowding and binaries, causing both exaggerated and errant parallaxes (Lindegren et al. 2018). For a similar sample in Lindegren et al. (2018), the criterion

$$u < 1.2 \times \max(1, \exp(-0.2(G_G - 19.5))) \quad (5)$$

was used to clean the sample of poor astrometry. Thus, any source that does not satisfy this expression is flagged with the **bad\_astr** flag in our final catalogue.

There are also those sources for which the fitting could not converge on a reasonable solution, probably because their true  $T_{\text{SED}}$  lies outside the bounds of our sample space. Thus sources lying on either  $T_{\text{eff}}$  bound are flagged as **bad\_teff** in the final catalogue. Finally, those sources which remained unflagged were assigned the **good** flag, meaning their input data should be free of both photometric and astrometric contamination, and they should have a well constrained  $T_{\text{SED}}$ .

The full table, including the input data and resulting parameters, are available at CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?VI/156>, the University of Exeter ORE repository (<https://doi.org/10.24378/exe.1683>) and the GitHub repository of supplementary material for this publication<sup>2</sup>. The table columns are described in Table 2.

<sup>2</sup> <https://github.com/sammorrell/mn19-supplementary-material>

## 3 RESULTS

Of the 15 765 targets in the input catalogue, 15 274 of them are flagged as **good**. We have plotted our sample in  $R - T_{\text{SED}}$  space in Figure 5, along with a selection of solar metallicity isochrones at 1 Gyr and 4 Gyr. The 68% confidence contours in the  $R - T_{\text{SED}}$  space from the subset described in Section 2.3 are shown in Figure 6. These show that we are able to determine the radius to a median statistical uncertainty of  $0.009 R_{\odot}$  (1.6%), ranging to a maximum uncertainty of  $0.025 R_{\odot}$  (2.6%). We found the mean uncertainty in temperature  $T_{\text{SED}}$  was 35K (1.0%), ranging to a maximum of 100K (2.7%).

The gap in the stellar sequence at 4000K in Figure 5 is caused by a discontinuity in the CIFIST BT-Settl model grid where the monotonic relationship between bolometric correction, defined as  $M_{\text{bol}} - M_i$  where  $M_i$  is the absolute magnitude of the  $i^{\text{th}}$  photometric band, and  $T_{\text{eff}}$  breaks down (Figure 7). We found that comparisons between observations and the SED resulting from the atmosphere at 4000K produce a higher  $\chi^2$  than those from the neighbouring atmospheres, causing our fitting to favour the SEDs produced from the atmospheres adjacent to that at 4000K. That this is a property of the model, as opposed to the fitting process, is supported by the fact that when we perform fits with grids derived from different atmospheres, as with the sub and super-solar metallicity grids in Section 4.2.5, the discontinuity disappears.

### 3.1 Radius Inflation

In Figure 5 the purely theoretical models undershoot the median of the radius distribution for all M-dwarf stars within our sample. However the PARSEC 1.2S model, which is calibrated by adopting an empirical  $T - \tau$  relation derived from DEBs as the boundary condition for the stellar interiors, traces the median radius well for the whole sample. When inferring the radius inflation of a sample of stars, the choice of parameters is of vital importance. As Figure 6 shows, the uncertainties in  $T_{\text{SED}}$  and  $R$  are strongly correlated; one cannot simply trace upwards from the theoretical sequence to infer the inflation. As mass is most closely related to luminosity, the radius inflation should in fact be measured in the more fundamental  $L_{\text{SED}} - R$  plane, shown in Figure 8. We measured the radius inflation for each of the models used in this work by picking a point of identical luminosity from the models and finding the difference between the prescribed model radius and our inferred radius.

### 3.2 Temperature - Radius Relation

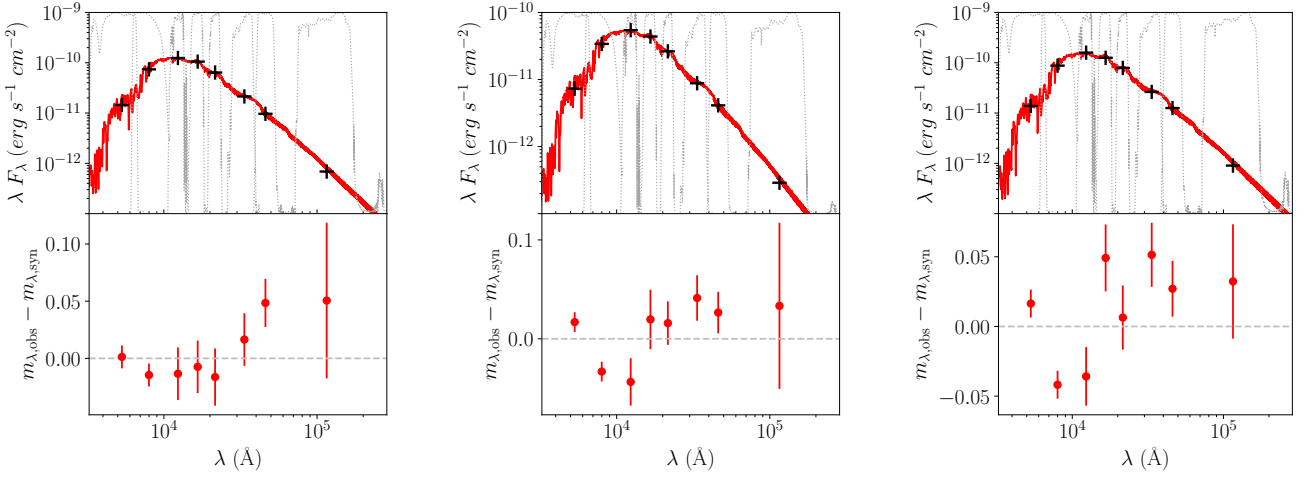
We can use our fits to derive a relation between temperature  $T_{\text{SED}}$  and radius  $R$ . We split the sample into 10K bins and took the median value from each bin. We required a minimum of 11 sources per bin, otherwise the entire bin was ignored. We fitted the medians with a 2<sup>nd</sup>-order polynomial, which is shown in Figure 9 as the blue line, with the black points being the median values. The error bars are the standard deviation of the radius distribution within each bin. The relation that we fit from this sample is

$$R_{\text{fit}}(T_{\text{SED}}) = -3.861 + 2.054 \times 10^{-3} T_{\text{SED}} - 2.335 \times 10^{-7} T_{\text{SED}}^2, \quad 3000\text{K} \leq T_{\text{SED}} \leq 4400\text{K}. \quad (6)$$

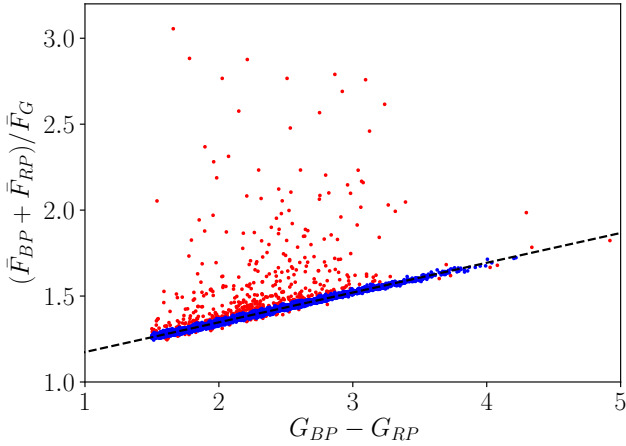
Using the same bins as before, we also drew the points that lie at the 16<sup>th</sup> and 84<sup>th</sup> percentiles to find the value of  $1\sigma$  for each bin. We used these points to fit further 2<sup>nd</sup>-order polynomials to yield upper  $R_{\text{high}}$  and lower  $R_{\text{low}}$  confidence radii for each bin. The upper

Column Name	FITS Name	Column Description
Target ID	id	Internal identifier for the object within our catalogue.
$\alpha$	ra	Right ascension of the source as given in the Gaia DR2 catalogue.
$\delta$	dec	Declination of the source as given in the Gaia DR2 catalogue.
Gaia Source ID	source_id	Unique source identifier of the star within Gaia DR2.
$G_G$	phot_g_mean_mag	G-band mean magnitude of the star in the Gaia DR2 catalogue.
$G_{BP}$	phot_bp_mean_mag	Integrated BP mean magnitude of the star in the Gaia DR2 catalogue.
$G_{RP}$	phot_rp_mean_mag	Integrated RP mean magnitude of the star in the Gaia DR2 catalogue.
$\delta G_{BP}$	phot_bp_mean_mag_err	Uncertainty of the BP mean magnitude of the star in the Gaia DR2 catalogue.
$\delta G_{RP}$	phot_rp_mean_mag_err	Uncertainty of the RP mean magnitude of the star in the Gaia DR2 catalogue.
$F_{BP}$	phot_bp_mean_flux	Integrated BP mean flux of the star in the Gaia DR2 catalogue.
$F_{RP}$	phot_rp_mean_flux	Integrated RP mean flux of the star in the Gaia DR2 catalogue.
$F_G$	phot_g_mean_flux	G-band mean flux of the star in the Gaia DR2 catalogue.
$\chi^2_{al}$	astrometric_chi2_al	Astrometric goodness of fit in the Gaia DR2 catalogue.
$N_{al,good}$	astrometric_n_good_obs_al	No. CCD transits not downweighted in the astrometric solution in the Gaia DR2 Catalogue.
$J$	Jmag_cds	$J$ band magnitude of the star in 2MASS.
$\delta J$	e_Jmag_cds	Uncertainty in $J$ band magnitude of the star in 2MASS.
$H$	Hmag_cds	$H$ band magnitude of the star in 2MASS.
$\delta H$	e_Hmag_cds	Uncertainty in $H$ band magnitude of the star in 2MASS.
$K_S$	Kmag_cds	$K_S$ band magnitude of the star in 2MASS.
$\delta K_S$	e_Kmag_cds	Uncertainty in $K_S$ band magnitude of the star in 2MASS.
$W_1$	W1mag_cds	$W_1$ band magnitude of the star in AllWISE.
$\delta W_1$	e_W1mag_cds	Uncertainty in $W_1$ band magnitude of the star in AllWISE.
$W_2$	W2mag_cds	$W_2$ band magnitude of the star in AllWISE.
$\delta W_2$	e_W2mag_cds	Uncertainty in $W_2$ band magnitude of the star in AllWISE.
$W_3$	W3mag_cds	$W_3$ band magnitude of the star in AllWISE.
$\delta W_3$	e_W3mag_cds	Uncertainty in $W_3$ band magnitude of the star in AllWISE.
Model No.	model_no	Number of the best fitting model in our grid.
$5 \log(R/d)$	norm	Magnitude normalisation calculated by the code to scale $Z_i$ to $m_{i,syn}$ .
$\chi^2$	chisq	Goodness of fit for the best fitting model.
$\chi^2_{\nu}$	reduced_chisq	Goodness of fit per degree of freedom for the best fitting model.
$P$	prob	The corresponding probability of being the best fitting model.
$R^2/d^2$	rsqr_on_dsqr	Measured radius squared on distance squared of the target.
$R$	radius	Our measured radius of the target (in units of $R_{\odot}$ ).
$d$	dist	Estimated distance (in pc) to the target.
$d_-$	dist_lbound	Lower bound on the confidence interval of the estimated distance (in pc).
$d_+$	dist_ubound	Upper bound on the confidence interval of the estimated distance (in pc).
$N$	nphot	The number of photometric data points used for the SED fitting.
$T_{SED}$	Teff	Our measured SED temperature of the target.
$\log(g)$	Log(g)	Our measured surface gravity of the target.
Photometric Contamination Flag	bad_phot	Flag showing whether source is consistent with contaminated photometry in Gaia DR2.
Astrometric Contamination Flag	bad_astr	Flag indicating whether source is consistent with contaminated astrometry in Gaia DR2.
$T_{SED}$ Edge Flag	bad_teff	Flag indicating that the measured $T_{SED}$ is at a grid boundary.
Good Flag	good	True if none of <b>bad_phot</b> , <b>bad_astr</b> or <b>bad_teff</b> are true for the source.
$T_{SED-}$	teff_lbound	Lower bound on the confidence interval of our measured $T_{SED}$ .
$T_{SED+}$	teff_ubound	Upper bound on the confidence interval of our measured $T_{SED}$ .
$\log(g)_-$	logg_lbound	Lower bound on the confidence interval of our measured $\log(g)$ .
$\log(g)_+$	logg_ubound	Upper bound on the confidence interval of our measured $\log(g)$ .
$R_-$	radius_lbound	Lower bound on the confidence interval of our measured radius.
$R_+$	radius_ubound	Upper bound on the confidence interval of our measured radius.
$\chi^2$ 68% Confidence Contour	tau_sq_68_contour	Value of $\chi^2$ within which 68% of the confidence is bounded within the cube.
Uncertainty Sample	uncer_sample	Sample for which the uncertainties correspond. For this publication, it will be 'random'.
$[R - R_{D08}(L)]/R$	d08_inflation	Radius inflation (in %) from the <a href="#">Dotter et al. (2008)</a> 4Gyr solar metallicity isochrone.
$R - R_{fit}(L_{SED})$	d08_lum_corr_radius	Radius residual when corrected for our luminosity - radius relation.
$L_{SED}$	lum	Luminosity measured for the source when correcting the radius for metallicity.
Metallicity Corrected $R$	corrected_radius	Radius of the star when corrected for the effects of metallicity.
Metallicity Corrected $R_-$	corrected_radius_lbound	Lower bound on the confidence interval of our metallicity corrected radius.
Metallicity Corrected $R_+$	corrected_radius_ubound	Upper bound on the confidence interval of our metallicity corrected radius.
$[\text{Fe}/\text{H}]$	fe_h	Measured metallicity of the target.
$[\text{Fe}/\text{H}]$ Source	fe_h_bibcode	The bibcode corresponding to the source of the metallicity measurement.
$F(L_{SED})[\text{Fe}/\text{H}]$	rad_corr	The correction applied to our measured radius to correct for metallicity, measured in $R_{\odot}$ .

**Table 2.** The column descriptions for our output catalogue.



**Figure 3.** Fits resulting from the use of the method presented in Section 2.3 and Section 2.4. The best fitting model spectrum for each target is shown in the top panel (red) with the observed photometry from which it was derived overlaid (black). The appropriate bandpasses are plotted in light grey for reference. In the bottom panel are the residuals and uncertainties in magnitudes for each photometric band. The middle panel corresponds to the median  $\chi^2$  of our randomly selected uncertainty sample, while the left and right panels show the stars at the lower and upper 68% density bounds respectively.



**Figure 4.** This figure illustrates the sigma clipping performed on our final catalogue. The black dashed line shows the final linear fit to the clipped sample after 8 iterations. The red points in this plot are those lying more than  $5\sigma$  away from this line, and are thus flagged as **bad\_phot**. Those points remaining in blue lie within  $5\sigma$  of the fit, and are considered to have uncontaminated Gaia photometry.

and lower bound radii are given by

$$R_{\text{high}}(T_{\text{SED}}) = -3.285 + 1.801 \times 10^{-3} T_{\text{SED}} - 2.040 \times 10^{-7} T_{\text{SED}}^2, \quad (7)$$

$$3000\text{K} \leq T_{\text{SED}} \leq 4400\text{K},$$

$$R_{\text{low}}(T_{\text{SED}}) = -3.288 + 1.691 \times 10^{-3} T_{\text{SED}} - 1.813 \times 10^{-7} T_{\text{SED}}^2, \quad (8)$$

$$3000\text{K} \leq T_{\text{SED}} \leq 4400\text{K}.$$

These functions are also shown in Figure 9. These bounds are separated by 4% at 4400K, increasing to 12% at 3500K and reaching a maximum separation of 30% at the lower temperature limit of 3000K. This scatter is further discussed in Section 4. The  $T_{\text{SED}} - R$  relation is tabulated in Table 3.

$T_{\text{SED}}$ (K)	$R_{\text{fit}}(T_{\text{SED}})$ ( $R_{\odot}$ )	$R_{\text{low}}(T_{\text{SED}})$	$R_{\text{high}}(T_{\text{SED}})$
3000	0.198	0.152	0.283
3001	0.199	0.153	0.284
3002	0.199	0.154	0.285
3003	0.200	0.154	0.285
3004	0.201	0.155	0.286

**Table 3.** A tabulated form of the  $T_{\text{SED}} - R$  relation presented in Equation 6 with bounds from Equation 7 and Equation 8. The full table is available in the online electronic supplementary material.

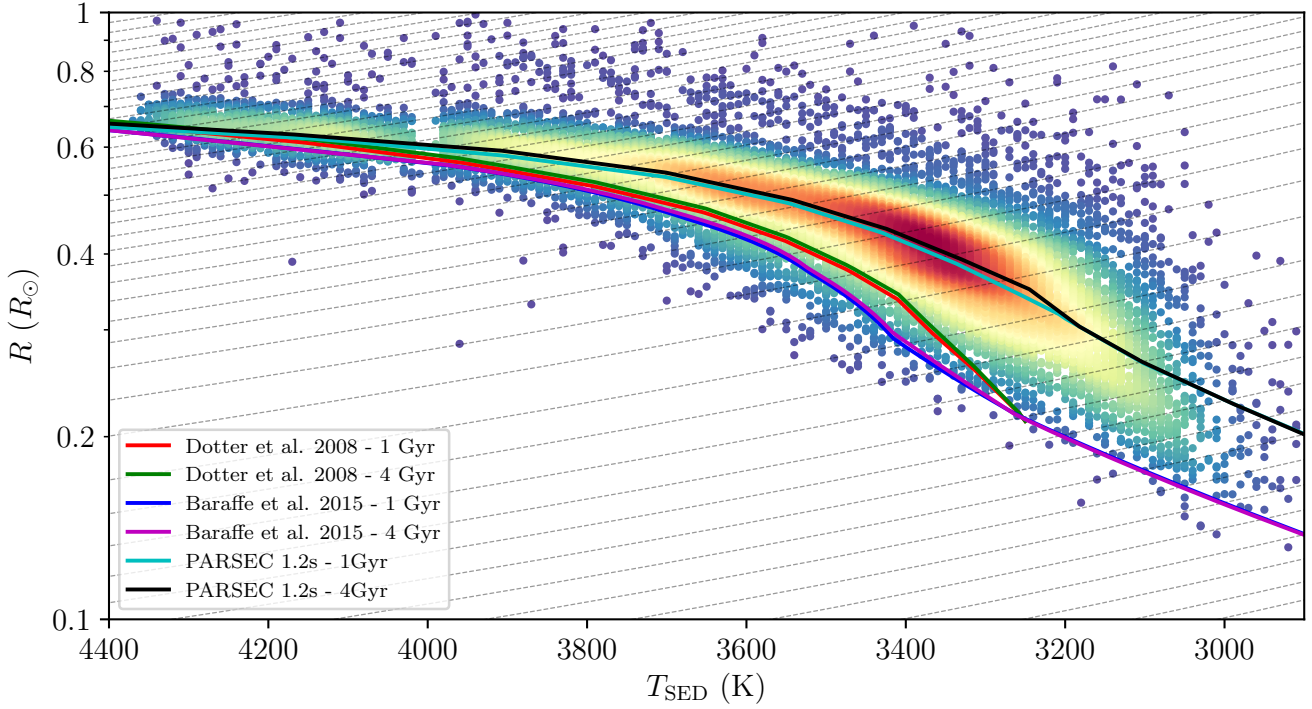
### 3.3 Luminosity - Radius Relation

Temperature - radius relations are useful for the purposes of exoplanet host characterisation, however the stellar modelling community relies on more fundamental parameters when testing models. We have therefore transformed our temperature - radius data into the arguably more fundamental luminosity - radius plane, and used it to fit a relationship using a similar methodology. However, deriving a relation between luminosity  $L$  and radius  $R$  is more problematic, as high order polynomials are required to capture the detail in the relation. Despite falling below the majority of the radius distribution, the isochrones do a good job of predicting the shape of this dataset; suggesting the models capture the physical changes involved. Hence we created our  $L_{\text{SED}} - R$  relationship as corrections to the Dotter et al. (2008) 4Gyr solar metallicity isochrone. We subtracted the radius given by the isochrone from the median, upper and lower bound radius in each bin, leaving the difference between theoretical and observed radii. Then, to get the relation, we simply added our correction to the radius prescribed by the isochrone. This relation holds for values between  $L_{\text{SED}} = 0.003L_{\odot}$  and  $L_{\text{SED}} = 0.1L_{\odot}$ . The correction to the Dotter et al. (2008) isochrone is given by

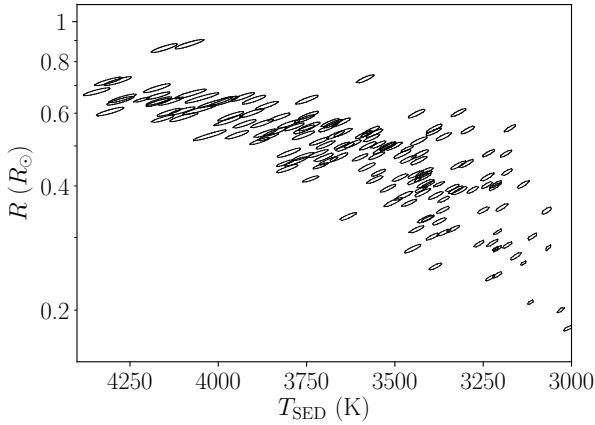
$$R_{\text{fit}}(L_{\text{SED}}) = R_{\text{D08}}(L_{\text{SED}}) + 0.0157 + 0.5670L_{\text{SED}} - 6.0633L_{\text{SED}}^2, \quad (9)$$

$$0.003L_{\odot} \leq L_{\text{SED}} \leq 0.1L_{\odot}$$

where  $L_{\text{SED}}$  is the luminosity derived from our SED fitting and  $R_{\text{D08}}(L_{\text{SED}})$  is the theoretical radius of the star predicted by the



**Figure 5.** The distribution of points from our full sample of 15 274 sources in  $T_{\text{SED}} - R$  space. The colour map results from a kernel density estimation, which is intended to indicate the density of points within the plot. For comparison, we include the [Dotter et al. \(2008\)](#) isochrones for  $z=0.018$  at 1 and 4 Gyr, [Baraffe et al. \(2015\)](#) isochrones for 1 and 4 Gyr and PARSEC 1.2S ([Chen et al. 2014](#); [Marigo et al. 2017](#)) isochrones for 1 and 4 Gyr in red, green, blue, purple, cyan and black lines respectively. Several points are clipped off the sides of the figure. The dashed lines trace constant luminosity, drawn every 100K along the [Baraffe et al. \(2015\)](#) isochrone.

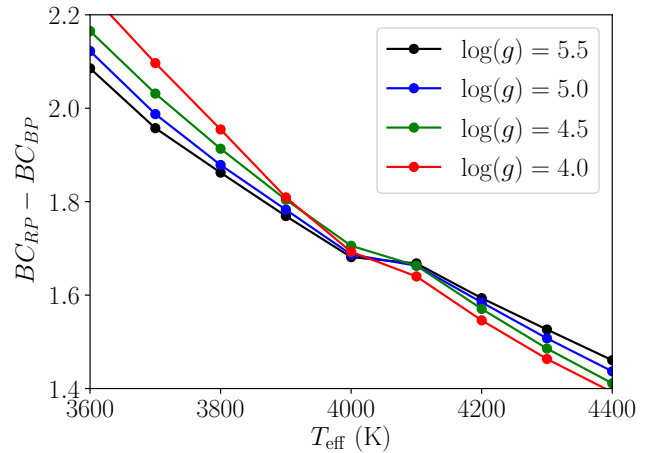


**Figure 6.** The 68% confidence contours resulting from the full 3D grid search we performed on 1% of our sample (158 stars) in [Section 2.3](#).

[Dotter et al. \(2008\)](#) isochrone at the given luminosity. The upper and lower bounds for the relation are given by

$$R_{\text{high}}(L_{\text{SED}}) = R_{\text{D08}}(L_{\text{SED}}) + 0.0294 + 0.6833L_{\text{SED}} - 4.5443L_{\text{SED}}^2, \quad 0.003L_{\odot} \leq L_{\text{SED}} \leq 0.1L_{\odot} \quad (10)$$

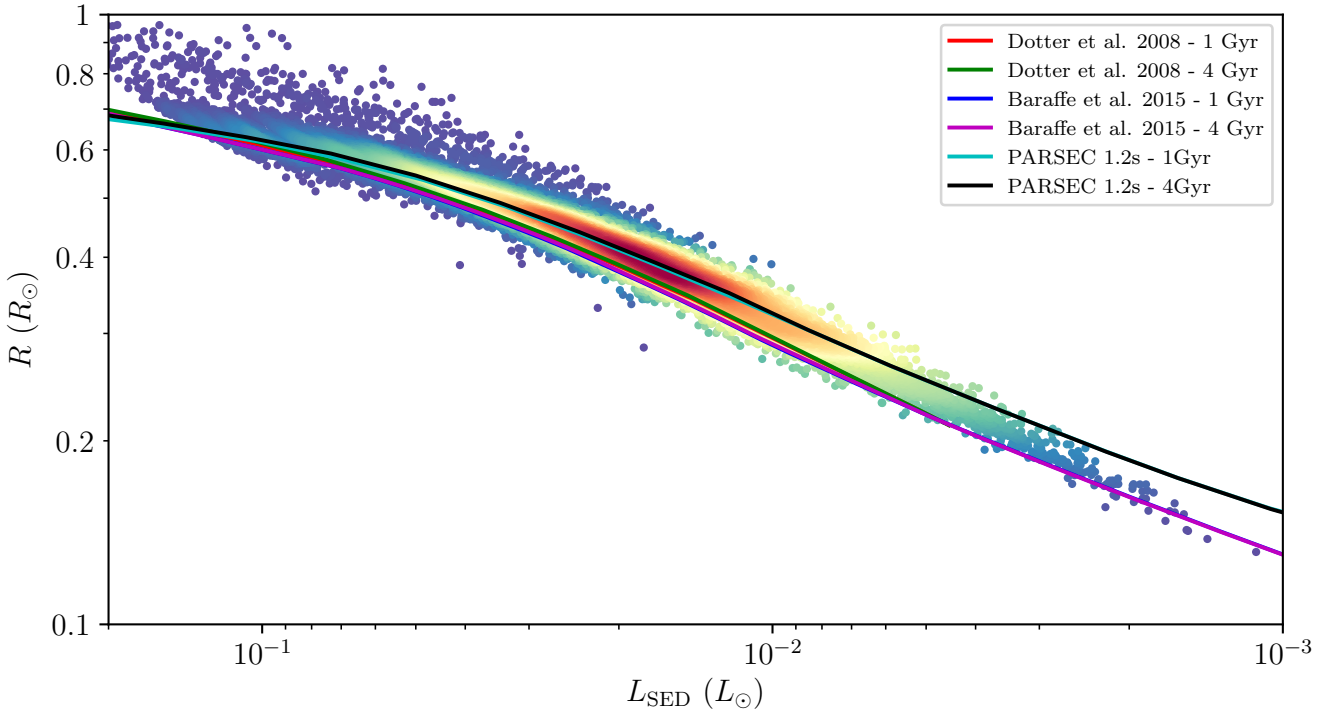
$$R_{\text{low}}(L_{\text{SED}}) = R_{\text{D08}}(L_{\text{SED}}) + 0.0033 + 0.2634L_{\text{SED}} - 4.0079L_{\text{SED}}^2, \quad 0.003L_{\odot} \leq L_{\text{SED}} \leq 0.1L_{\odot}. \quad (11)$$



**Figure 7.** The gap in the stellar sequence at 4000K evident in [Figure 5](#) is due to a discontinuity in the CIFIST BT-Settl model grid. By plotting the bolometric corrections from the Gaia DR2 bands this effect is clearly seen. The plot shows  $\log(g) = 4.0, 4.5, 5.0$  and  $5.5$  (black, blue, green and red respectively) to demonstrate that this discontinuity affects the entire span of  $\log(g)$  sampled by our grid.

The correction and final relation are shown atop the data in [Figure 11](#) and [Figure 10](#) respectively. The  $L_{\text{SED}} - R$  correction and the sum of the [Dotter et al. \(2008\)](#) isochronal radius with the correction are tabulated in [Table 4](#).





**Figure 8.** The **good** sample of our dataset in the  $L_{\text{SED}} - R$  plane. Due to the strong correlation between the axes in the  $T_{\text{SED}} - R$  plane, this plane is preferable for accurately measuring radius inflation in our sample. Accompanying the data are the same isochrones as in Figure 5. The Dotter et al. (2008) and Baraffe et al. (2015) isochrones trace a similar sequence in this space, resulting in the Dotter et al. (2008) isochrones being hidden behind the Baraffe et al. (2015) isochrones.

$L_{\text{SED}} (L_{\odot})$	$\Delta R$	$\Delta R_{\text{low}}$	$\Delta R_{\text{high}}$	$R_{\text{fit}} (R_{\odot})$	$R_{\text{low}}$	$R_{\text{high}}$
0.003	0.017	0.004	0.031	0.195	0.182	0.209
0.004	0.018	0.004	0.032	0.219	0.206	0.233
0.005	0.018	0.004	0.033	0.240	0.226	0.254
0.006	0.019	0.005	0.033	0.258	0.244	0.273
0.007	0.019	0.005	0.034	0.275	0.260	0.289

**Table 4.** A tabulated form of the  $L_{\text{SED}} - R$  relationship presented in Equation 9 as the sum of the Dotter et al. (2008) isochronal radius with a correction  $\Delta R$ . We also include the bounds from Equation 10 and Equation 11 and the applied correction. The full table is available in the online electronic supplementary material.

## 4 DISCUSSION

### 4.1 Comparison with Literature Radii

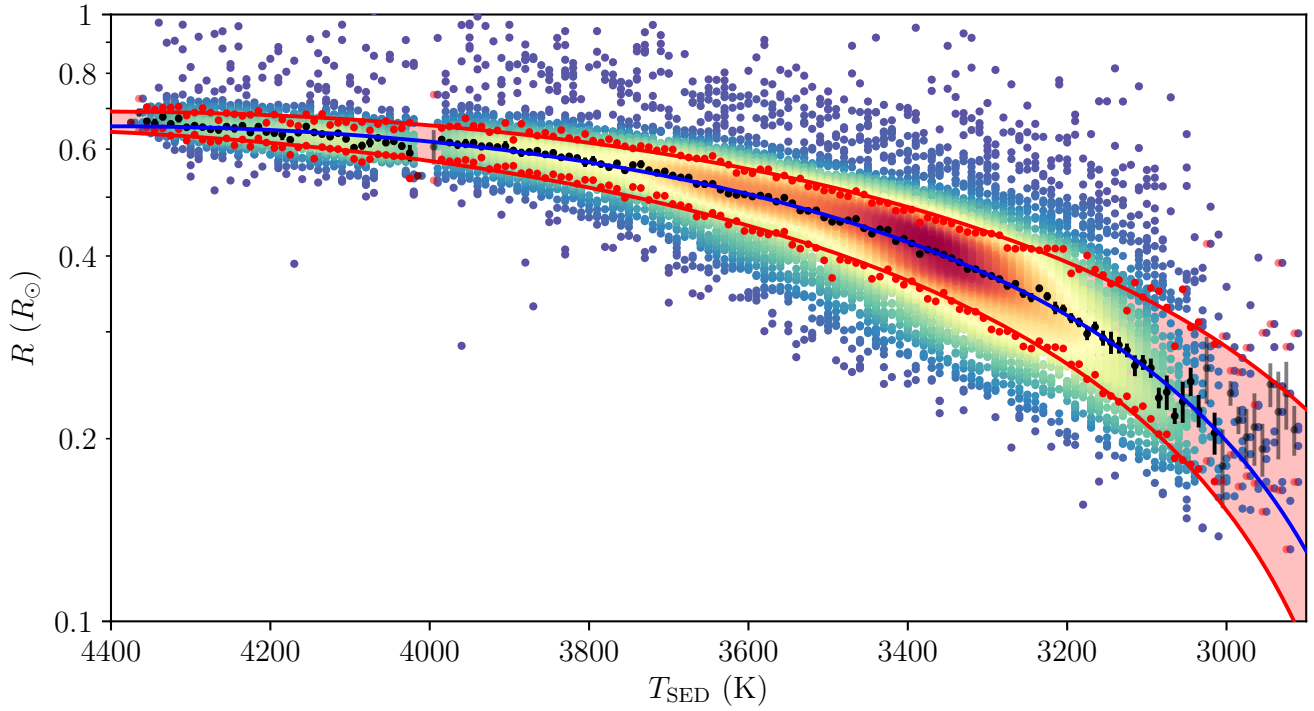
To compare our measure of radius inflation with that from DEBs, interferometry and  $L_{\text{SED}} + T_{\text{sp}}$  (Mann et al. 2015) we limited ourselves to the range 3400K - 4400K, where all methods are well sampled. The mode of all methods coincides at between 3 – 7% inflated compared to the models. In Figure 12 we plotted the distribution of relative radius residual from Equation 9 (see Section 3.3). The DEB and interferometric samples show medians at around 3% lower than ours. The  $L_{\text{SED}} + T_{\text{sp}}$  median corresponds well with our own, however it does exhibit a second peak at 7% under inflation, though it is difficult to be sure whether this is a genuine feature of the population. There is also a long tail of outliers on the high

inflation wing of our distribution; we suggest that these are a small number of binaries that have leaked into our sample.

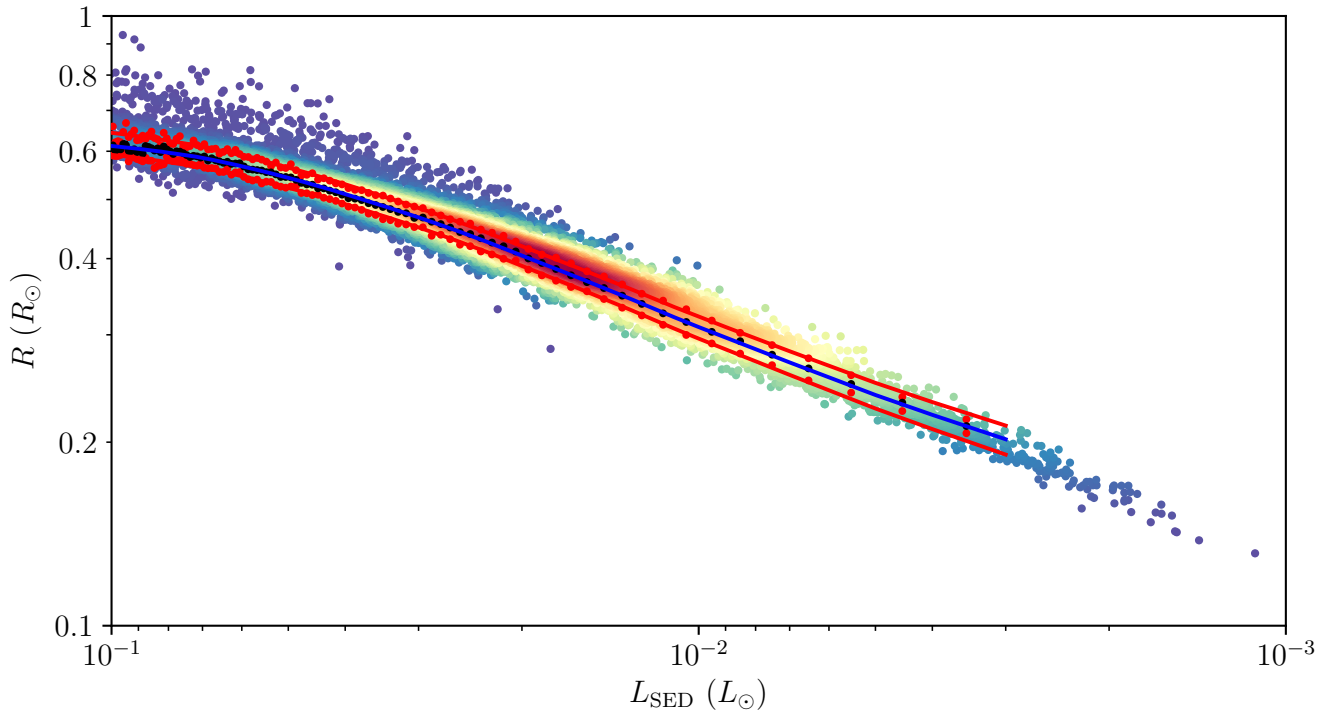
In Figure 13 we compare the interferometric, eclipsing binary and  $L_{\text{SED}} - T_{\text{sp}}$  datasets with our own as a function of luminosity, as we require that the datasets have another physical quantity in common in addition to the radius. For the reasons outlined in Section 4.2 luminosity is the best abscissa to use. Although the eclipsing binaries are normally viewed as a mass-radius dataset, the eclipsing binaries to which we are comparing also have temperatures derived from the spectra or photometric surface brightness, which in combination with the radius allows us to calculate a luminosity. The natural plane for the interferometric data is the luminosity-radius plane. The  $L_{\text{SED}} + T_{\text{sp}}$  dataset can again be converted into  $L_{\text{SED}}$ -radius. The data presented in this paper are derived in the  $T_{\text{SED}}$ -radius plane, and so can be converted into the  $L_{\text{SED}} - R$  plane. We emphasize that all of these comparisons can only be made assuming  $T_{\text{SED}} = T_{\text{sp}} = T_{\text{eff}} = T_{\text{br}}$ , where  $T_{\text{br}}$  is the brightness temperature measured for some eclipsing binaries. As Figure 13 shows, all datasets show a radius inflation with respect to the models.

For the luminosity range for which our  $L_{\text{SED}} - R$  relation is valid, there are a total of 29 DEBs, 16 interferometric stars and 154  $L_{\text{SED}} + T_{\text{sp}}$  stars. We determined which side of our median radius these samples fall. We found the 66% of the DEBs, 88% of the interferometric stars and 80% of the  $L_{\text{SED}} + T_{\text{sp}}$  sample lie below our median; meaning that all 3 other methods yield a lower median radius than ours. This skew to lower radii across this range supports the distributions shown in Figure 12.

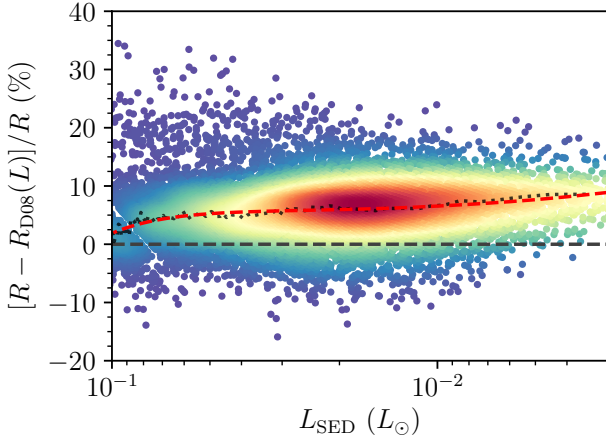
This difference between the methods might be due to starspots. First, any method which measures a spectroscopic temperature, es-



**Figure 9.** The  $T_{\text{SED}} - R$  relationship derived for our sample. The final relationship, given by Equation 6, is the solid blue line along with its 68% confidence intervals shown in red. The black dots show the stars used to perform the fit of the relationship. The red points bordering the upper and lower bound lines also show the stars used to fit them. The points in bins flagged as not good, and thus not used during fitting, are shown with semi-transparent markers.



**Figure 10.** The luminosity - radius relationship plotted atop the stars from our sample. As with Figure 9, we show the relation (blue), upper and lower bound lines (red) and the points from which each of the lines were fit in the corresponding colour.



**Figure 11.** The radius inflation of our data from the 4 Gyr [Dotter et al. \(2008\)](#) solar metallicity isochrone  $[R - R_{D08}(L)]/R$ . The median radius inflation within each luminosity bin is shown as a black point. The luminosity - radius correction to this same isochrone is overlaid as a dashed red line.

pecially in the optical, will not be sensitive to spots, because the immaculate photosphere spectrum will dominate the spot spectrum, leading to an over-prediction of the  $T_{\text{eff}}$  and under-prediction of radius. This affects both the eclipsing binary and  $L_{\text{SED}} + T_{\text{sp}}$  sample and would result in radii smaller than those measured by our method. Although the  $L_{\text{SED}} + T_{\text{sp}}$  and interferometric methods integrate under the SED of the star to find the luminosity, as we do, many of their objects saturate in WISE and so lack mid-IR coverage. Not sampling the region of the SED where the spotted photosphere has its strongest contribution relative to the immaculate photosphere, and may make a non-negligible contribution to the overall stellar flux, may cause the measured luminosity to be too low, again resulting in radii that are too small. In summary, we found the radii measured by all methods to be inflated above the theoretical sequence, albeit we measure radii that are larger than the other methods.

## 4.2 Contributions to the Radius Scatter

[Figure 13](#) demonstrates that for a given luminosity there is a 3 – 7% spread in measured radius. It is important to determine whether this radius spread is real, and if so what effects contribute to it.

### 4.2.1 Could observational uncertainties contribute to the scatter?

We first wished to establish whether the uncertainties in our radius determination could explain the spread before searching for a physical origin. [Figure 6](#) shows 68% confidence contours for the uncertainty in  $R$  and  $T_{\text{SED}}$  for an unbiased sub-sample of our catalogue, which we discussed in [Section 2.3](#). The mean uncertainty in this selection of stars is (1.6%); much less than the 3 – 7% spread that is observed. Furthermore we find a similar (but slightly larger) spread in the literature radii, with 62% of the DEBs, 50% of interferometric stars and 53% of the  $L_{\text{SED}} + T_{\text{sp}}$  sample lying within our relation’s 68% confidence bounds. Thus we conclude that the spread is not the result of observational uncertainties.

### 4.2.2 Does flux contamination from faint counterparts contribute to the scatter?

[Wilson & Naylor \(2017\)](#) show that the AllWISE bands can suffer contamination due to faint, hidden sources falling within the large PSFs of brighter stars. As well as making accurate catalogue matching problematic in crowded fields, this can also cause contamination to AllWISE photometry. Even a modest flux contamination from a stray faint source within the WISE PSF has the potential to cause a large discrepancy in both the retrieved  $T_{\text{SED}}$  and  $R$ . Fortunately, [Wilson & Naylor \(2018\)](#) provide a catalogue of Gaia DR2 - WISE matches in the galactic plane, which allows us to assess the effect of contamination on our sample. According to their work, of the 2334 sources that match between the catalogues, fewer than 4% were likely to be contaminated by 10% or more. The rest of the sky is less prone to crowding, so we suggest that 4% of sources suffering contamination is the upper limit for our entire sample. Were our AllWISE photometry affected by contamination due to the presence of an unseen counterpart, we would expect the contaminated sources to exhibit more inflated radii than clean sources. However, we found no correlation between radius inflation and predicted flux contamination or probability; indicating that contamination in the WISE bands is unlikely to contribute towards the spread.

### 4.2.3 Do starspots contribute towards the radius spread?

Starspots introduce a second, cooler component to the SED; effectively diverting some of the luminosity of the star away from the immaculate photosphere. [Jackson & Jeffries \(2014\)](#) used a polytropic model including starspots to reproduce the radii of pre-main sequence (PMS) stars in the Pleiades and NGC 2516, which required spot coverages of 35 - 51%. Using these models, they were able to find an 8% inflation in the stars when compared to the [Baraffe et al. \(2015\)](#) stellar models. [Higl & Weiss \(2017\)](#) were also able to explain observed radii with starspots by covering large percentages (up to 44%) of the surface of their models with spots.

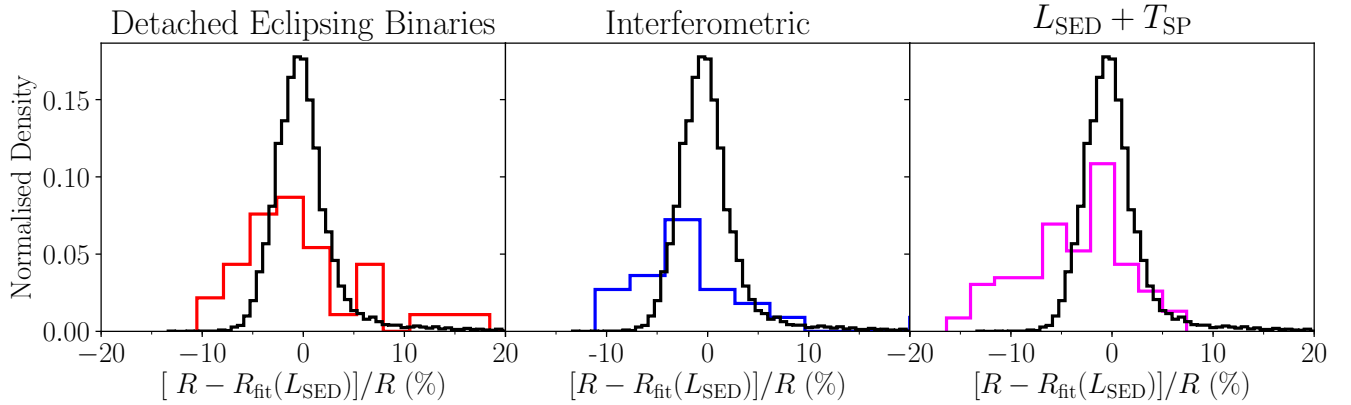
Our fitting assumed that the entire surface of the star is a single temperature. To examine the effect of starspots on our measured radii we produced a catalogue of simulated magnitudes of stars with spots and ran this catalogue through the same fitting process used for our observed sample. We simulated the input catalogue by sampling from a grid with  $\log(g) = 5.0$  and varying spot filling factor  $\gamma$  between  $\gamma = 0.0 - 1.0$ ; whereas in all grids used for fitting data we have assumed  $\gamma = 0.0$  (no spots). The composite photosphere consists of an immaculate and a spotted photosphere with temperatures  $T_{\text{imac}}$  and  $T_{\text{spot}}$  respectively. In determining the temperature of the spotted photosphere, we make the reasonable assumption (see [Berdyugina 2005](#)) that  $T_{\text{spot}} = 0.8 T_{\text{imac}}$  for  $5000\text{K} < T_{\text{imac}} < 3000\text{K}$ . The synthetic magnitudes  $Z_i$  in this grid are given by

$$Z_{\lambda, \text{syn}} = -2.5 \log_{10} \left[ \frac{\int_{\lambda} ((1 - \gamma) I_{\lambda, \text{imac}} + \gamma I_{\lambda, \text{spot}}) S_{\lambda, i} d\lambda}{\int_{\lambda} f_{\lambda}^{\circ} S_{\lambda, i} d\lambda} \right] + m_{\lambda}^{\circ}, \quad (12)$$

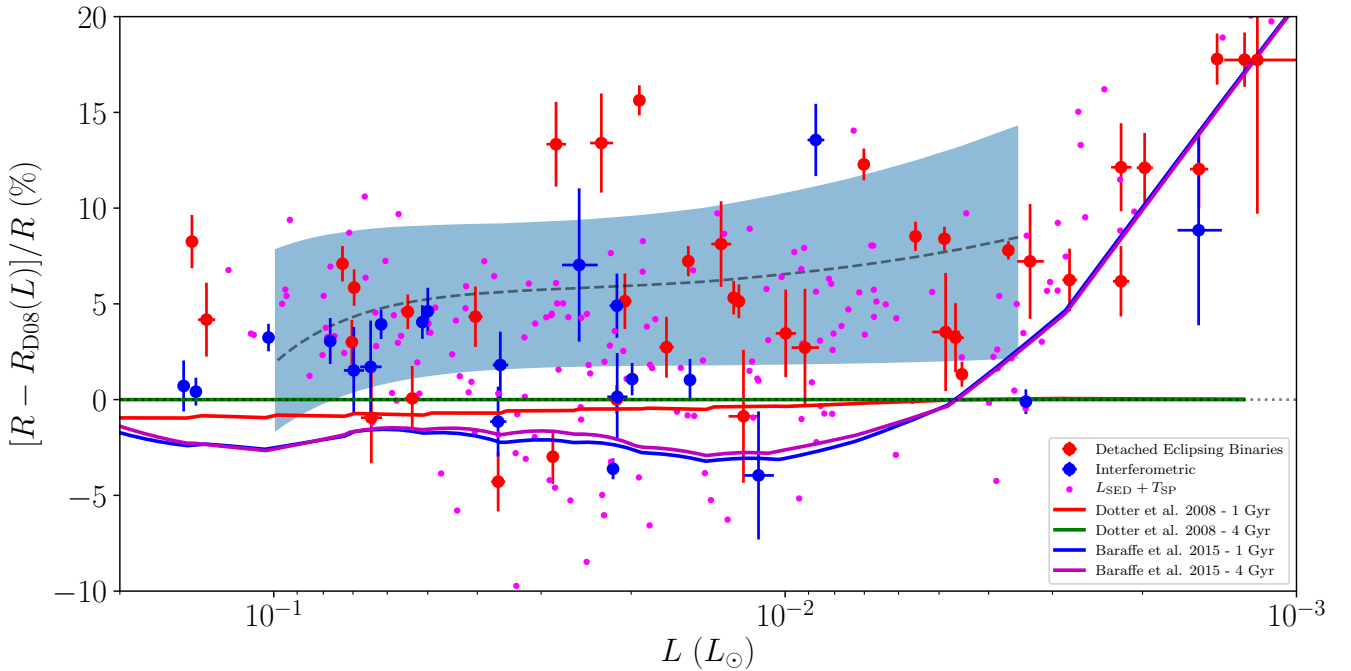
where  $I_{\lambda, \text{imac}}$  and  $I_{\lambda, \text{spot}}$  are the intensity of the immaculate and spotted photosphere respectively. The effective temperature  $T_{\text{eff}}$  of these spotted models becomes

$$T_{\text{eff, spotted}} = \left( (1 - \gamma) T_{\text{imac}}^4 + \gamma T_{\text{spot}}^4 \right)^{\frac{1}{4}}. \quad (13)$$

To produce the simulated input catalogue we iterated through immaculate photosphere temperature, adopting the effective temperature of the combined photosphere as given in [Equation 13](#).



**Figure 12.** A comparison between the distributions of the relative residual of our measured radius  $R$  with respect to our  $L_{\text{SED}} - R$  relation  $R_{\text{fit}}(L_{\text{SED}})$ ; defined in Equation 9. Our sample is shown in black, with each of the others overplotted. Both the interferometric and DEB sample have median values about 3% below our own. The Mann et al. (2015) sample exhibits a bimodal structure, with the more pronounced peak occurring close to our median.

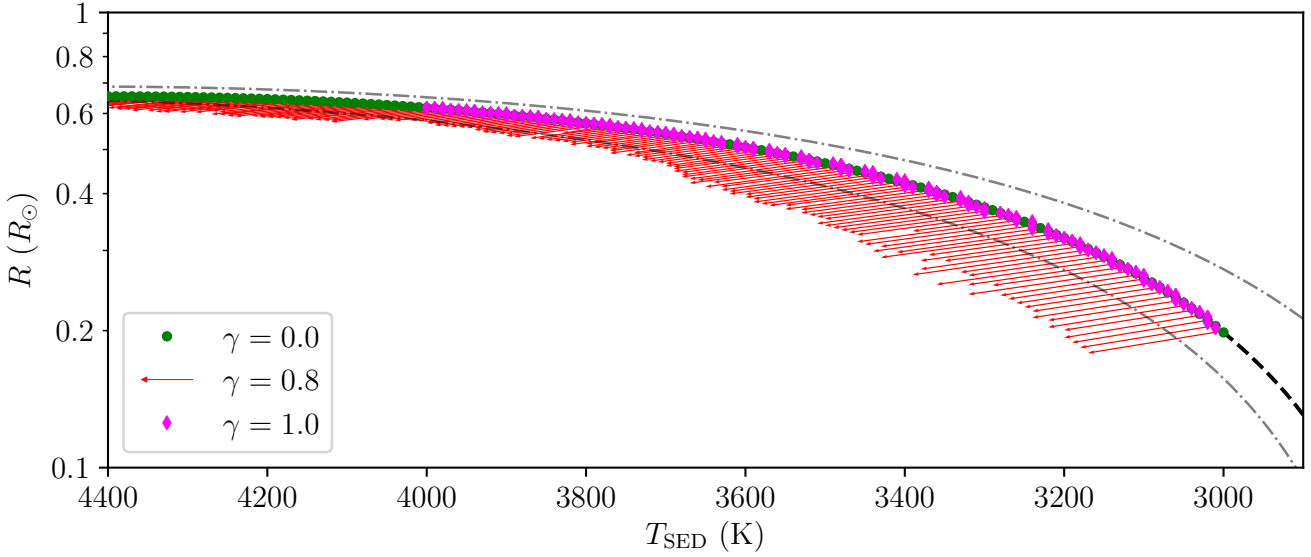


**Figure 13.** The radius inflation with respect to the Dotter et al. (2008) 4 Gyr isochrone ( $[R - R_{\text{D08}}(L)]/R$ ) obtained by the four different methods. Inflation is plotted as a function of luminosity. The red points show detached eclipsing binaries (Southworth 2015; Parsons et al. 2018). The blue points show interferometric stars (Boyajian et al. 2012). The pink points show stars from Mann et al. (2015). The error bars for this sample have been omitted for reasons of clarity. The luminosity - radius relation derived in this paper (grey dashed line) with its associated 68% density bounds is also shown. For comparison, model isochrones for 1Gyr and 4Gyr ages are shown for Dotter et al. (2008) at  $z = 0.018$  and Baraffe et al. (2015). The original sources for (Southworth 2015) are Welsh et al. (2015); Torres et al. (2014); Torres & Ribas (2002); Terrien et al. (2012); Stassun et al. (2004); Rozyczka et al. (2009, 2007); Pietrzyński et al. (2013); Morales et al. (2009b,a); Kraus et al. (2017, 2011); Hełminiak et al. (2015); Hartman et al. (2018). The original sources for Parsons et al. (2018) are Pyrzas et al. (2012); Parsons et al. (2018, 2016, 2012a,c,b, 2010).

Each  $T_{\text{eff,spotted}}$  was mapped onto the corresponding stellar radius given by Equation 6. We utilised a Monte Carlo method to account for the uncertainties, which could potentially add to the spread. For each band we generated a CDF for the uncertainties in the observed catalogue and Monte Carlo sampled it for each simulated star. With this catalogue we then performed a fitting using the unspotted grid.

The fits resulting from this process is shown in Figure 14. What

this makes clear is that spot coverage can contribute towards the perceived spread in radius. At  $\gamma = 0.0$  and  $\gamma = 1.0$ , we recover the original  $T_{\text{SED}} - R$  relation as would be expected. For  $0.0 < \gamma \leq 0.8$  the stars scatter to lower radii at increased  $T_{\text{SED}}$ , conserving the overall luminosity. However, for  $0.8 < \gamma < 1.0$  we see the effects of the spotted photosphere in the SED and the fits begin retrieving temperatures closer to those of the spots and the measured radii become



**Figure 14.** The points show the radius and temperature retrieved by our fitting a catalogue of simulated synthetic photometry with varying spot filling factor  $\gamma$ . We show 0% ( $\gamma = 0.0$ ) and 100% ( $\gamma = 1.0$ ) spot coverage, which lie along the relation, along with  $\gamma = 0.8$  which lies at the extremity of the spread. We show the 68% density bounds of our  $T_{\text{SED}} - R$  relation as grey dotted lines.

closer to the relation. This has the effect of producing a scatter in our relation which closely corresponds to the 68% confidence lines resulting from our  $T_{\text{SED}} - R$  relation presented in Section 3.2.

#### 4.2.4 Correlations with activity

Although we have shown that the effect of starspots on our measurement technique is able to explain our scatter, this hypothesis would necessitate a correlation between measured radius and magnetic activity. To probe magnetic field strength we checked for correlations between radius inflation and markers of magnetic activity, the most reliable of which is rotation period  $P_{\text{rot}}$ . We investigated a correlation with rotation period by assembling a sample of periods from McQuillan et al. (2013) and McQuillan et al. (2014) observed using Kepler. Unfortunately, there are only 21 targets in common with our sample, so we chose to supplement these catalogues with rotation periods determined from Gaia DR2 lightcurves in Lanzafame et al. (2018) which are based on sparser lightcurves, but add an appreciable number of stars; resulting in a final sample of 189 stars for which we have rotation periods. We used a theoretical expression for convective turnover time  $\tau_c$ , provided by Cranmer & Saar (2011), to determine the equivalent Rossby number  $\text{Ro} = P_{\text{rot}}/\tau_c$  for each star. To avoid spurious correlations with spectral type, we first performed a linear fit on both  $P_{\text{rot}}$  and  $\text{Ro}$  vs  $T_{\text{SED}}$  and corrected for it in our final correlations, which are shown alongside the fits in Figure 15. Rotation alone is adequate to show that our sample lacks appreciable correlations between magnetism and radius inflation, and is the most fundamental because more rapid rotation rates presumably promote a stronger dynamo action within the stellar interior. However, Wright et al. (2011) and Newton et al. (2017) show that X-ray luminosity and  $H_\alpha$  excess correlate well with rotation period in their unsaturated regimes. When unsaturated both can be used as additional markers of magnetic fields emergent at the stellar surface. It has been shown both observationally (Wright et al. 2011; Newton et al. 2017) and theoretically (Reiners et al. 2009)

that activity saturates at around  $\text{Ro} \approx 0.1$ , meaning that 31% of our rotation sample would be in the unsaturated regime. Although rotation itself does not saturate, we have split our rotation periods into saturated and unsaturated sub-samples around this threshold to aid comparison with the following samples, which do.

We investigated the correlation with X-ray luminosity by cross-matching with DR6 of the XMM-Newton Serendipitous Source Catalog (3XMM DR6; Rosen et al. 2016), yielding 95 stars. This sample was divided into saturated and unsaturated sub-samples using the threshold  $L_X/L_{\text{bol}} \approx 3 \times 10^{-4}$ , defined from the lower limit of the spread around the saturated sample of Wright et al. (2018); meaning 53% of our stars are unsaturated. To avoid a spurious correlation with spectral type, and ensure that we only probe excess X-ray emission due to activity, as before we performed a linear fit to the data in the  $T_{\text{SED}} - L_X/L_{\text{bol}}$  plane and used it to correct the values of  $L_X/L_{\text{bol}}$ . Both the fit and final correlation with relative radius residual are shown in Figure 16.

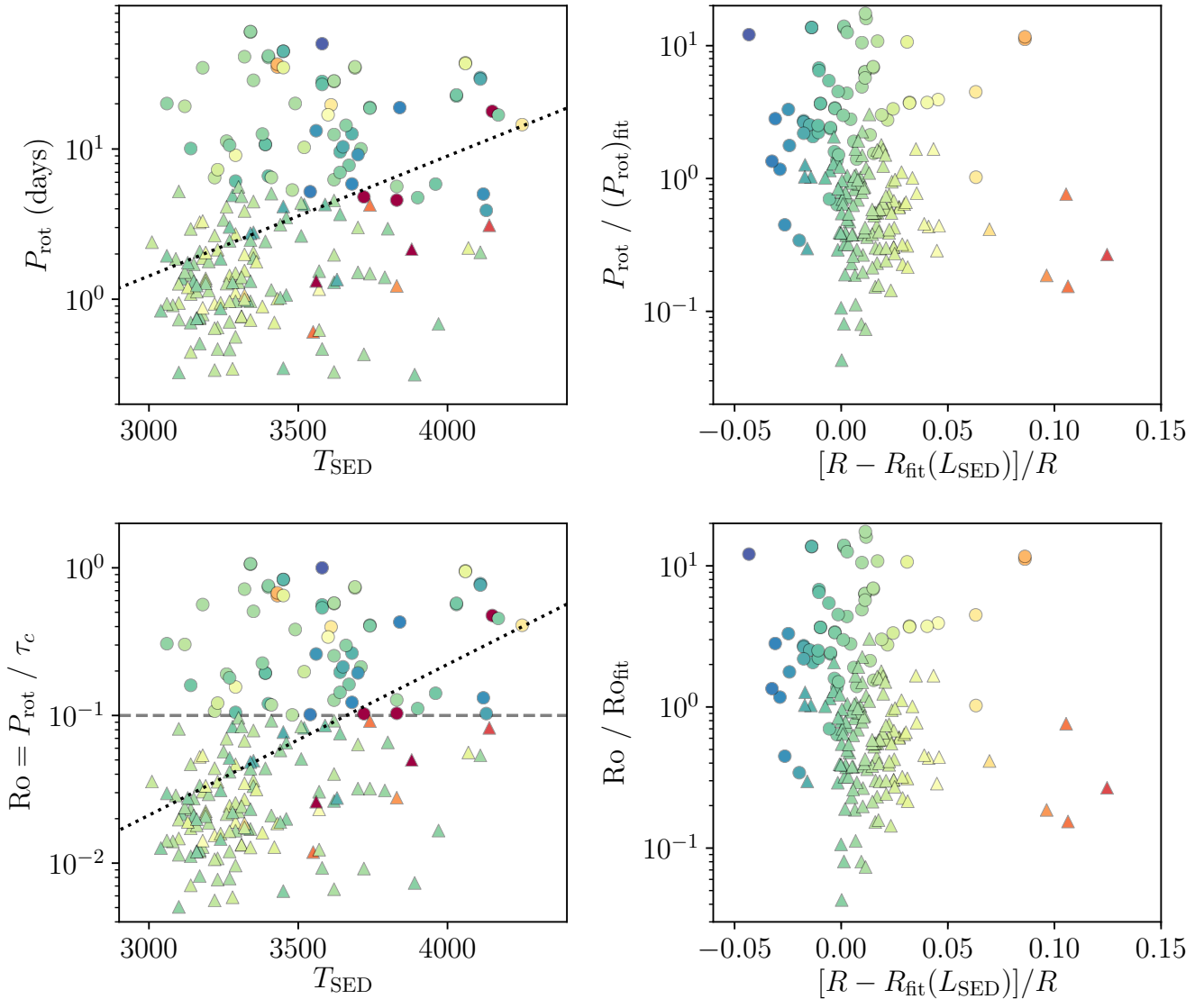
Finally, we investigated the correlation with  $H_\alpha$  by matching our catalogue with both DR2 of the INT Photometric  $H_\alpha$  Survey of the Northern Galactic Plane (IPHAS2; Drew et al. 2005; Barentsen et al. 2014) and the VST Photometric  $H_\alpha$  Survey of the Southern Galactic Plane and Bulge (VPHAS+; Drew et al. 2014). This results in a total of 573 stars for which we have  $H_\alpha$  magnitudes. We have presented this sample in terms of  $L_{H_\alpha}/L_{\text{bol}}$  using

$$\frac{L_{H_\alpha}}{L_{\text{bol}}} = W_{H_\alpha} \chi, \quad (14)$$

where  $\chi$ , introduced in Walkowicz et al. (2004), was interpolated from the tabulated values from Table 8 of Douglas et al. (2014). The equivalent width of the  $H_\alpha$  line  $W_{H_\alpha}$  due to activity was determined by measuring the excess flux across the  $H_\alpha$  band, of width  $\Delta\lambda_{H_\alpha}$ , using

$$W_{H_\alpha} = \Delta\lambda_{H_\alpha} \left[ 10^{0.4(m_r - m_{H_\alpha}) - (BC_{H_\alpha} - BC_r)} - 1 \right], \quad (15)$$

where  $m_r$  and  $m_{H_\alpha}$  are the observed magnitudes, and  $BC_r$  and



**Figure 15.** The correlation between our measured relative radius residual  $[R - R(L)]/R$  and rotation period  $P_{\text{rot}}$  (left) and Rossby number  $\text{Ro}$  (right). The sample is divided between saturated (triangles) and unsaturated (circles) activity samples. The transition is marked by a grey dashed line and determined to be  $\text{Ro} = 0.1$ , both observationally (Newton et al. 2017) and theoretically (Reiners et al. 2009). We removed spurious correlations with spectral type from the right pane by performing a linear fit, shown as a black dotted line, in the accompanying left pane. Several stars at extreme inflations are not visible in the plot. The colour map denotes  $[R - R_{\text{fit}}(L_{\text{SED}})]/R$ .

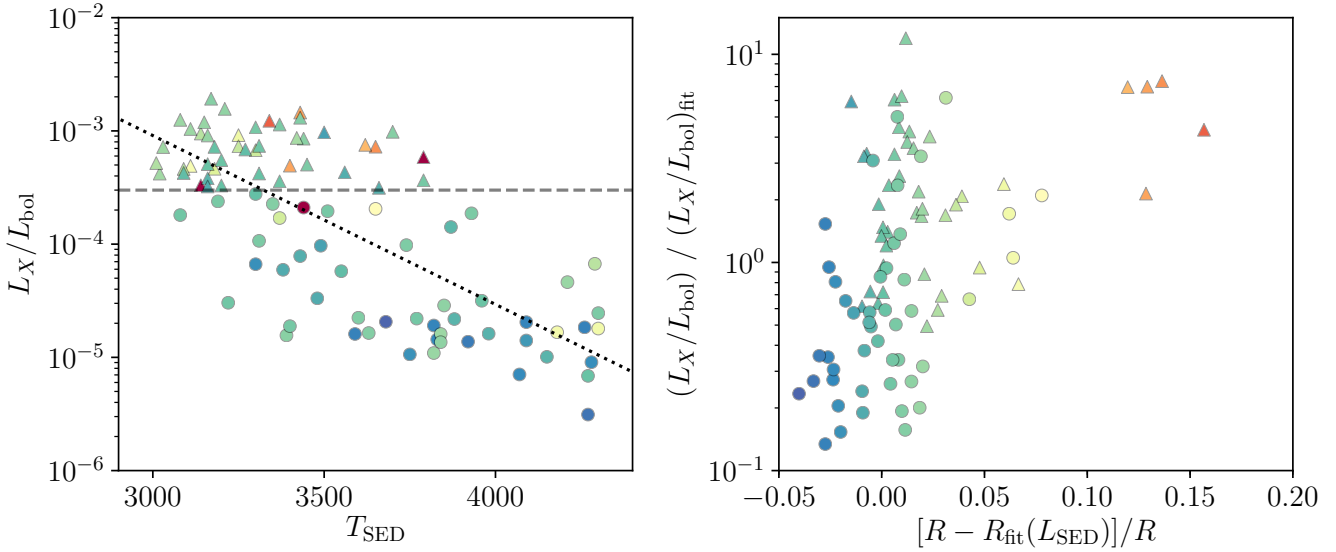
$BC_{H_\alpha}$ , the inactive model bolometric corrections. This sample is shown in Figure 17, along with the saturation threshold of  $L_{H_\alpha}/L_{\text{bol}} = 10^{-4}$  (Newton et al. 2017; Douglas et al. 2014). We note that some 53 of our  $W_{H_\alpha}$  measurements are mildly negative, which indicates quiescence, thus we count these among our unsaturated sample. However due to resulting in negative  $L_{H_\alpha}/L_{\text{bol}}$  they do not appear on Figure 17. This demonstrates that our  $H_\alpha$  sample spans both the saturated and unsaturated regimes, with around 25% of our sample being unsaturated.

To summarise, we have found that markers of both interior field strength ( $P_{\text{rot}}$ ) and surface field strength ( $L_X/L_{\text{bol}}$  and  $L_{H_\alpha}/L_{\text{bol}}$ ) show no appreciable correlation with radius inflation for M-dwarfs. All three markers are sampled in both the saturated and unsaturated

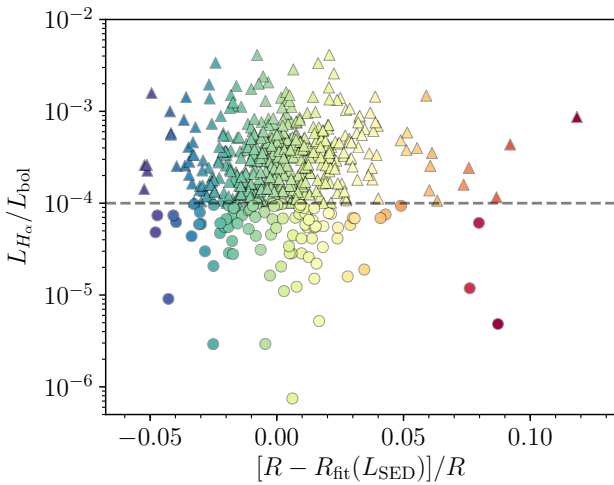
activity regime, with between a quarter and a half of each being unsaturated.

#### 4.2.5 How does metallicity affect the radius spread?

Differences in stellar metallicity could also cause a scatter in radii (Berger et al. 2006). A reduction in metallicity, and thus opacity, would allow the star to more efficiently radiatively dissipate internal energy, resulting in a smaller radius at the same luminosity (Berger et al. 2006). Stars within the solar neighbourhood show a metallicity spread with  $\sigma = 0.2$  dex (Boone et al. 2006) with the lower extremity at  $[M/H] \approx -0.6$  (Neves et al. 2013). Comparing the Dotter et al. (2008) solar metallicity isochrones with those of  $[M/H] = \pm 0.25$  in luminosity - radius space, gives a difference in radius of



**Figure 16.** Our sample of stars with X-ray luminosity  $L_X/L_{\text{bol}}$  against relative radius residual  $[R - R_{\text{fit}}(L_{\text{SED}})]/R$ . The sample is divided into stars that lie in the saturated (triangles) and unsaturated (circles) regimes. This transition is marked by a grey dashed line drawn at  $L_X/L_{\text{bol}} \approx 3 \times 10^{-4}$ , which corresponds to the lower limit from Wright et al. (2018). The left pane shows a spurious correlation in our sample between  $L_X/L_{\text{bol}}$  and  $T_{\text{SED}}$ , which is corrected for in the right plot using a linear fit (black dotted line). Several stars are omitted from the high inflation wing of this plot.



**Figure 17.** Our sample of stars for which we have  $H_\alpha$  luminosity  $L_{H_\alpha}/L_{\text{bol}}$  against relative radius residual  $[R - R_{\text{fit}}(L_{\text{SED}})]/R$ . The sample is divided into stars that lie in the saturated (triangles) and unsaturated (circles) regime, delimited by the dashed line drawn at  $L_{H_\alpha}/L_{\text{bol}} = 10^{-4}$ , corresponding to the value from Douglas et al. (2014) and Newton et al. (2017).

about  $\pm 4\%$ . So, theory suggests the contribution to the spread is minimal. However, Mann et al. (2015) and Rabus et al. (2019) show a correspondence between metallicity and relative residual in radius for a number of M-dwarf stars.

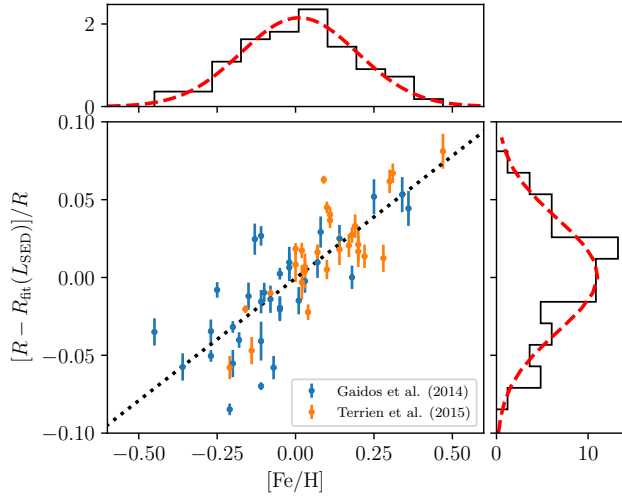
To investigate this in our data, we assembled a sample of stars from our catalogue with measurements of  $[\text{Fe}/\text{H}]$  from Terrien et al. (2015) and Gaidos et al. (2014). The sample with which we matched ranges between  $-0.5 < [\text{Fe}/\text{H}] < 0.5$ , indicating that our catalogue is relatively free of M-subdwarfs ( $[\text{Fe}/\text{H}] < -0.5$ ). We verified this using the Besançon population synthesis model (Robin et al.

2003), which yielded population sizes consistent with our sample and indicated  $\approx 6\%$  of our sample has  $[\text{M}/\text{H}] < -0.5$  and  $\approx 0.5\%$  having  $[\text{M}/\text{H}] \leq -1.0$ , as well as reproducing our distribution of metallicities well for  $[\text{M}/\text{H}] > -0.5$ . We ensured that these targets lie within the valid range of the  $L_{\text{SED}} - R$  relation from Section 3.3 and we avoid the high inflation wing of our sample by choosing inflations  $[R - R_{\text{D08}}(L)]/R < 12\%$ . We established in Section 4.1 that the radii from Mann et al. (2015) are inconsistent with our own, so we chose not to use them. The radius residuals resulting from the luminosity correction are shown as a function of  $[\text{Fe}/\text{H}]$  in Figure 18. The sample spans a residual of  $\pm 6\%$ , which appears to correlate strongly with metallicity, and corresponds well to the spread of metallicity in the solar neighbourhood.

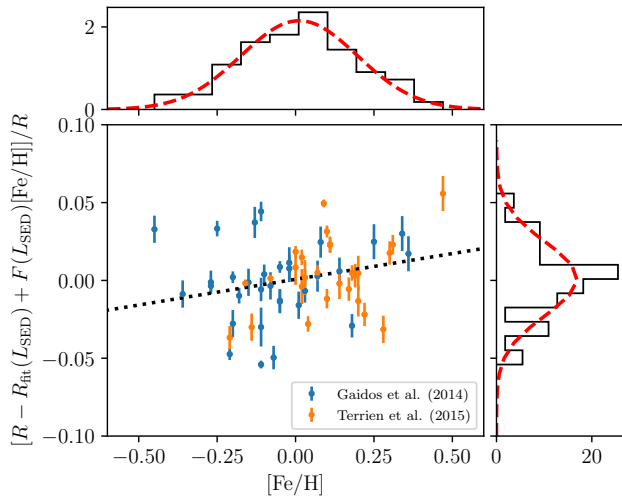
This initially suggested that the spread was physical, and caused by metallicity, however the correlation is steeper than predicted by theory. Furthermore, we had measured all our radii using only solar metallicity atmospheres. To determine the error induced by this we produced grids for  $[\text{M}/\text{H}] = \pm 0.25$  by interpolating synthetic photometry from  $[\text{M}/\text{H}] = \pm 0.5$  and  $[\text{M}/\text{H}] = 0.0$  grids<sup>3, 4</sup>. We fitted each star in our input catalogue with all three metallicities, allowing us to produce a  $L_{\text{SED}} - R$  relation for each metallicity following the procedure from Section 3.3. We found that our technique is able to determine luminosity consistently to within about 1%, regardless of which metallicity atmospheres are used. Thus, using the median radius from each luminosity bin, which we interpolated between bin midpoints, we calculated the difference between the radii measured at each metallicity and the radii measured at solar metallicity as a function of luminosity. Hence, we were able to measure the theoretical luminosity-dependent relationship between

<sup>3</sup> To ensure solar abundance for all metallicities we used the BT Settl AGSS2009 models, which adopts the Asplund et al. (2009) solar abundances, as opposed to CIFIST which uses those of Caffau et al. (2011).

<sup>4</sup> The BT Settl AGSS2009 model atmospheres are available from <https://phoenix.ens-lyon.fr/Grids/BT-Settl/AGSS2009/SPECTRA/>



**Figure 18.** The correlation between radius residual and  $[\text{Fe}/\text{H}]$  for stars within our sample. The metallicities are from Terrien et al. (2015) (orange) and Gaidos et al. (2014) (blue). The radii and uncertainties are from our sample and fit using only solar metallicity atmospheres.



**Figure 19.** As Figure 18 but with the correction from Section 4.2.5 applied. This correction accounts for using only solar metallicity atmospheres to fit a range of metallicities.

measured radius and  $[\text{M}/\text{H}]$  in the form of a linear relationship, with gradient  $F(L_{\text{SED}})$ ; values for which are tabulated in Table 5. This allowed us to determine the correct radius residual for each star, which are shown in Figure 19, using

$$\frac{\delta R}{R} = \frac{1}{R} (R - R_{\text{fit}}(L_{\text{SED}}) + F(L_{\text{SED}}) [\text{Fe}/\text{H}]), \quad (16)$$

where  $R_{\text{fit}}(L_{\text{SED}})$  is the  $L_{\text{SED}} - R$  relation given in Equation 9. The resulting corrected radii were fit with to yield another  $L_{\text{SED}} - R$  relation, which remains consistent to within 1% with the relation presented in Section 3.3 for  $0.015L_{\odot} < L_{\text{SED}} < 0.09L_{\odot}$ . When corrected for metallicity, we found that the correlation between relative radius residual and metallicity was no longer significant (see Figure 19).

Therefore, the correlation of radius with metallicity in Fig-

$L_{\text{SED}} (L_{\odot})$	$F(L_{\text{SED}})$	$L_{\text{SED}}$	$F(L_{\text{SED}})$	$L_{\text{SED}}$	$F(L_{\text{SED}})$
0.0035	-0.0159	0.0365	-0.0744	0.0695	-0.0586
0.0045	-0.0192	0.0375	-0.0836	0.0705	-0.0467
0.0055	-0.0270	0.0385	-0.0799	0.0715	-0.0534
0.0065	-0.0334	0.0395	-0.0645	0.0725	-0.0509
0.0075	-0.0342	0.0405	-0.0819	0.0735	-0.0582
0.0085	-0.0415	0.0415	-0.0717	0.0745	-0.0369
0.0095	-0.0443	0.0425	-0.0784	0.0755	-0.0446
0.0105	-0.0414	0.0435	-0.0792	0.0765	-0.0647
0.0115	-0.0470	0.0445	-0.0636	0.0775	-0.0448
0.0125	-0.0573	0.0455	-0.0820	0.0785	-0.0494
0.0135	-0.0497	0.0465	-0.0671	0.0795	-0.0447
0.0145	-0.0557	0.0475	-0.0887	0.0805	-0.0670
0.0155	-0.0577	0.0485	-0.0674	0.0815	-0.0607
0.0165	-0.0663	0.0495	-0.0743	0.0825	-0.0096
0.0175	-0.0623	0.0505	-0.0751	0.0835	-0.0447
0.0185	-0.0626	0.0515	-0.0644	0.0845	-0.0583
0.0195	-0.0668	0.0525	-0.0639	0.0855	-0.0608
0.0205	-0.0684	0.0535	-0.0706	0.0865	-0.0323
0.0215	-0.0697	0.0545	-0.0784	0.0875	-0.0614
0.0225	-0.0744	0.0555	-0.0466	0.0885	-0.0350
0.0235	-0.0702	0.0565	-0.0672	0.0895	-0.0434
0.0245	-0.0758	0.0575	-0.0552	0.0905	-0.0373
0.0255	-0.0789	0.0585	-0.0724	0.0915	-0.0600
0.0265	-0.0759	0.0595	-0.0540	0.0925	-0.0463
0.0275	-0.0698	0.0605	-0.0846	0.0935	-0.0467
0.0285	-0.0803	0.0615	-0.0533	0.0945	-0.0538
0.0295	-0.0772	0.0625	-0.0613	0.0955	-0.0351
0.0305	-0.0753	0.0635	-0.0601	0.0965	-0.0425
0.0315	-0.0830	0.0645	-0.0611	0.0975	-0.0247
0.0325	-0.0790	0.0655	-0.0384	0.0985	-0.0279
0.0335	-0.0774	0.0665	-0.0544	0.0995	-0.0340
0.0345	-0.0818	0.0675	-0.0766		
0.0355	-0.0696	0.0685	-0.0497		

**Table 5.** The tabulated values for  $F(L_{\text{SED}})$  in Equation 16. In between  $L_{\text{SED}}$  points we linearly interpolate neighbouring values. This table is also available in the online electronic supplementary material.

ure 18 is the result of fitting stars with a spread of metallicities with only solar metallicity atmospheres, which determines  $T_{\text{SED}}$  hence  $R$  incorrectly. Thus the observed correlation between metallicity and radius spread in our sample is not physical and can be corrected for with accurate measurements of the stellar luminosity and metallicity. Mann et al. (2015) also found a correlation between metallicity and radius within their sample, however it only appears as a function of  $T_{\text{sp}}$ , not as a function of  $M_{K_s}$  - a proxy for luminosity. Interestingly, the interferometric and DEB sample show a very similar scatter to our uncorrected radii, however we attribute this to observational uncertainty. The mean uncertainty in metallicity from Gaidos et al. (2014) and Terrien et al. (2012) will result in a 1.7% spread in our measured radius in Figure 19. Given that our median radius uncertainty from the fitting process is 1.6% (see Section 3), this would be consistent with the 2.4% scatter seen in Figure 19. There could also be an intrinsic scatter of the stars in our sample, but to remain consistent with the spread it would have to be at most 1 – 2%. This leads us to conclude that our knowledge of M-dwarf radii is currently limited by the accuracy and precision of the metallicity measurement of the star.



#### 4.2.6 The cause of the scatter - a summary

We have been able to show that faint contaminants in AllWISE photometry were not the cause of our scatter (see Section 4.2.2). We have also established that the uncertainties in our measurements were not able to explain the spread (see Section 4.2.1). We found that extremely high starspot coverage in most of the population could explain the scatter. But since stars with small spot coverages would have lower measured radii, and stars with larger coverages, larger radii, this would imply a measured radius - activity correlation, which we were unable to find (see Section 4.2.3 and Section 4.2.4). However, we found that there is a strong correlation between [Fe/H] and our measured radius. Although this could be misconstrued as a physical spread in radius caused by metallicity, we found that in fact this correlation resulted from using solar metallicity atmospheres to fit the SED of non-solar metallicity stars. When this is corrected for we found that the correlation disappeared, resulting in a scatter of about 2.4% in the subset for which we could find metallicities. This spread is consistent with the 2–3% found by Schweitzer et al. (2019). Uncertainty in [M/H] measurements accounts for about 1.7% of this spread, with our uncertainties in radius explaining the remainder. The accuracy of the measured metallicities also makes determinations of effective temperature problematic, imposing another limit on the accuracy of our radii. Therefore we suggest that the spread of this corrected radius distribution is currently dominated by the accuracy and precision of current metallicity measurements (see Section 4.2.5) and the intrinsic spread in M-dwarf radii is less than 1 – 2%. This well characterised radius residual distribution also places an upper limit on the typical variability of spot filling factors of stars within our sample of < 10%.

### 4.3 Explaining the Radius Inflation

We now address the question as to why the M-dwarf main sequence is inflated from theoretical predictions, and note in passing the problem also appears to apply to the pre-main sequence (Jackson et al. 2018). Section 1.3 summarises our current understanding of inflation - it is clear that the mechanism behind radius inflation in M-dwarfs remains a contentious point. So far the most compelling hypothesis has been dynamo driven magnetic fields, which would inhibit convection and probably modify the specific entropy in the convective region, and hence the internal structure of the star. Stars below a mass of approximately  $0.35M_{\odot}$  are thought to have interiors that are fully convective (Limber 1958), making this a satisfying explanation. If this were to be the case, we might expect to see a spread in radius, from the theoretical non-magnetic models to some level of maximum inflation, provided we sampled a large range of rotation rates. However, we established in Section 4.2.5 that there is a tight (with perhaps a 1 – 2% intrinsic scatter) main sequence for M-dwarf stars, refuting this scenario. One could argue that we have not adequately considered the case that radius inflation has saturated, much like the activity does, for all of the stars in our sample. Figure 15 shows we sample a large range in rotation period, including at least a quarter where activity is unsaturated. For saturation of radius inflation to be consistent with such tight a sequence, it would have to occur at rotation rates much slower than both saturation of photospheric activity indicators (at  $Ro = 0.1$ ) and the slowest rotators in our sample  $Ro \approx 1.0$ .

A further argument against stellar magnetism being responsible for radius inflation comes from the fact that all of the indicators of magnetic activity we studied show no appreciable correlation with measured radius residual (see Section 4.2.4). We

would expect to see correlations between the activity and radius inflation were magnetism responsible. Given these two arguments, we reluctantly conclude that stellar magnetism is currently unable to explain radius inflation in main sequence M-dwarf stars.

#### 4.3.1 Alternate explanations for radius inflation

It is unlikely that magnetic fields are the cause of radius inflation in main sequence M-dwarf stars, so an alternate explanation must be sought. Metallicity is the main driving force behind the measured radius scatter, however in Section 4.2.5 it was shown to be caused by our measurement technique. This renders us unable to explain the inflated radii of main sequence M-dwarfs. However, one avenue of inquiry comes from the PARSEC 1.2S models (Chen et al. 2014; Marigo et al. 2017) which adopts an empirical  $T - \tau$  relation as the boundary condition to their interiors. As Figure 5 and Figure 8 show, this modification does a remarkably good job of characterising the degree of inflation and the intrinsic sequence for early to mid M-dwarf stars.

### 4.4 Determining Accurate M-dwarf Radii

We showed in Section 4.2.5 that metallicity can cause problems when determining M-dwarf radii. Using atmospheres of a different metallicity from the star being fitted causes the retrieved temperature, and thus the radius, to be incorrect. However, we have devised strategies for obtaining accurate M-dwarf radii in the face of these issues.

#### 4.4.1 Without a metallicity measurement

If there is not a measured metallicity for a star, there are two strategies available. Firstly, disregarding the metallicity of the star and fitting indiscriminately with solar metallicity models yields radii accurate to better than 5%; in our case about 3.6%. For statistical samples and less precise applications, this may be adequate. However, our method can measure the luminosity of the star correctly to within a couple percent regardless of metallicity (see Section 4.2.5). Thus, we can use the measured luminosity along with an empirical  $L_{\text{SED}} - R$  relation, such as that in Section 3.3, to obtain correct radii. The scatter about this relation ranges between 3.6% – 4.5%.

#### 4.4.2 With a metallicity measurement

If the star has an accurate measurement of the metallicity, there are two more avenues open. The simplest option is to perform the fitting with an atmosphere of the appropriate metallicity. However for large samples it can be impractical to generate large grids of synthetic photometry at a number of differing metallicities. So our other suggestion is to follow the method presented in Section 4.2.5. This entails fitting the entire sample with solar metallicity models and correcting for the metallicity. Both strategies are highly dependent upon the accuracy and precision of the metallicity measurement. A metallicity constrained to about 10% induces a scatter of about 1.7% in measured radius. However, for cooler stars, with well constrained photometry and sufficiently accurate metallicities (better than about 3%), radius measurements of better than 1% can be achieved with this method.

## 5 CONCLUSIONS

We have measured the temperature, radius and luminosity of a sample of 15 274 late K and early M-dwarf stars using a modified spectral energy distribution fitting method. This method requires only accurate photometry and precision astrometry and thus adds a fourth method to those used to evaluate the veracity of stellar models. Importantly, this method works natively in the  $T_{\text{eff}} - R$  space which is crucial for the characterisation of exoplanets. We have derived empirical  $T_{\text{SED}} - R$  and  $L_{\text{SED}} - R$  relations, which can be used to characterise exoplanet host stars and validate stellar evolution models (Section 3.2 and Section 3.3). The key conclusions of this work are as follows.

- (i) Currently, none of the purely theoretical stellar models can describe the mean radius inflation of the main sequence at temperatures lower than about 4000K. The measured radii are inflated by 3 – 7% compared to those predicted by models (Section 3.1).
- (ii) We have shown that M-dwarfs lie on a tight sequence (with a scatter smaller than 1 – 2%) (see Section 4.2.5). This is in conflict with magnetic models, which would suggest a spread in radius, from the theoretical sequence for non-magnetic models to a maximum inflation. We have also shown that there is no appreciable correlation between all observational markers of magnetic activity and radius inflation (see Section 4.2.4). This leads us to conclude that stellar magnetism is currently unable to explain radius inflation in main sequence M-dwarf stars (see Section 4.2). Furthermore, this would explain the unexpected result that detached eclipsing binaries are not inflated with respect to their single star counterparts (see Figure 12).
- (iii) We discovered that fitting a distribution of metallicities with only solar metallicity models introduces an apparent correlation between  $[\text{Fe}/\text{H}]$  and  $R$  (Section 4.2.5).
- (iv) However, we found that our technique correctly measures the luminosity regardless of metallicity, meaning that we can correct the measured radii. Without a measured metallicity we achieved a precision of 3.6%. However, this was improved to 2.4% when corrected for metallicity. Given that the uncertainty in  $[\text{M}/\text{H}]$  accounts for 1.7% of this spread, it is clear that the precision of metallicity measurements is currently the limiting factor for this method (Section 4.4).
- (v) In the absence of metallicity measurements we present an empirical  $L_{\text{SED}} - R$  relation which can be used to measure correct radii to a precision of 3.6 – 4.5% (see Section 4.4).

What remains deeply puzzling is how tight the intrinsic sequence appears. Even if the stellar radii were identical at a given luminosity, we show in Section 4.2.3 that small changes in starspot filling factors would give an apparent spread in radius because of our measurement technique. Hence the tight measured sequence implies that the starspot coverage among main sequence M-dwarfs is remarkably consistent, which is physically counter intuitive when considering the diverse levels of activity in even a modest subset of our full sample (see Section 4.2.4).

## ACKNOWLEDGEMENTS

SM is supported by a UK Science and Technology Facilities Council (STFC) studentship. The authors would like to thank Matthew Browning and Lewis Ireland for a critical reading of the manuscript. We would also like to thank Victor See and Sean Matt for useful conversations regarding stellar rotation. We are grateful to the referee for a very useful report, which sharpened our thinking and improved

our presentation, in addition to providing useful background such as the discrepancies in binary parameters referred to in Section 1.1. Finally we would like to thank Stuart Littlefair for prompting us to investigate correlations with activity.

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center / California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

This paper makes use of data obtained as part of the INT Photometric  $H\alpha$  Survey of the Northern Galactic Plane (IPHAS, [www.iphas.org](http://www.iphas.org)) carried out at the Isaac Newton Telescope (INT). The INT is operated on the island of La Palma by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. All IPHAS data are processed by the Cambridge Astronomical Survey Unit, at the Institute of Astronomy in Cambridge. The bandmerged DR2 catalogue was assembled at the Centre for Astrophysics Research, University of Hertfordshire, supported by STFC grant ST/J001333/1.

Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 177.D-3023, as part of the VST Photometric  $H\alpha$  Survey of the Southern Galactic Plane and Bulge (VPHAS+, [www.vphas.eu](http://www.vphas.eu)).

This research has made use of data obtained from the 3XMM XMM-Newton serendipitous source catalogue compiled by the 10 institutes of the XMM-Newton Survey Science Centre selected by ESA.

This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France. We gratefully acknowledge the software that made this work possible, including: STILTS and TOPCAT (Taylor 2006), SCIPY (Jones et al. 2001), NUMPY (Oliphant 2015), MATPLOTLIB (Hunter 2007) and ASTROPY (Astropy Collaboration et al. 2013).

## REFERENCES

- Allard F., Homeier D., Freytag B., Sharp C. M., 2012, in EAS Publications Series, pp 3–43, doi:10.1051/eas/1257001
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Astropy Collaboration et al., 2013, *A&A*, 558, A33
- Bailer-Jones C. A. L., Rybizki J., Foesneau M., Mantelet G., Andrae R., 2018, preprint, (arXiv:1804.10121)
- Baraffe I., Homeier D., Allard F., Chabrier G., 2015, *A&A*, 577, A42
- Barentsen G., et al., 2014, *MNRAS*, 444, 3230
- Bell C. P. M., Naylor T., Mayne N. J., Jeffries R. D., Littlefair S. P., 2012, *MNRAS*, 424, 3178
- Berdyugina S. V., 2005, *Living Reviews in Solar Physics*, 2, 8
- Berger D. H., et al., 2006, *ApJ*, 644, 475
- Blackwell D. E., Shallis M. J., 1977, *MNRAS*, 180, 177
- Blackwell D. E., Shallis M. J., Selby M. J., 1979, *MNRAS*, 188, 847

- Boone R. H., King J. R., Soderblom D. R., 2006, *New Astron. Rev.*, **50**, 526
- Boyajian T. S., et al., 2012, *ApJ*, **757**, 112
- Brun A. S., Browning M. K., 2017, *Living Reviews in Solar Physics*, **14**, 4
- Caffau E., Ludwig H.-G., Steffen M., Freytag B., Bonifacio P., 2011, *Sol. Phys.*, **268**, 255
- Casagrande L., 2008, *Physica Scripta Volume T*, **133**, 014020
- Chaturvedi P., Sharma R., Chakraborty A., Anandarao B. G., Prasad N. J. S. S. V., 2018, *AJ*, **156**, 27
- Chen Y., Girardi L., Bressan A., Marigo P., Barbieri M., Kong X., 2014, *MNRAS*, **444**, 2525
- Cortés-Contreras M., et al., 2017, *A&A*, **597**, A47
- Cranmer S. R., Saar S. H., 2011, *ApJ*, **741**, 54
- Dotter A., Chaboyer B., Jevremović D., Kostov V., Baron E., Ferguson J. W., 2008, *ApJS*, **178**, 89
- Douglas S. T., et al., 2014, *ApJ*, **795**, 161
- Drew J. E., et al., 2005, *MNRAS*, **362**, 753
- Drew J. E., et al., 2014, *MNRAS*, **440**, 2036
- Evans D. W., et al., 2018, *A&A*, **616**, A4
- Feiden G. A., Chaboyer B., 2013, *ApJ*, **779**, 183
- Gaia Collaboration et al., 2016, *A&A*, **595**
- Gaia Collaboration Brown A. G. A., Vallenari A., Prusti T., de Bruijne J. H. J., Babusiaux C., Bailer-Jones C. A. L., 2018, preprint, ([arXiv:1804.09365](https://arxiv.org/abs/1804.09365))
- Gaidos E., et al., 2014, *MNRAS*, **443**, 2561
- Gillen E., Hillenbrand L. A., David T. J., Aigrain S., Rebull L., Stauffer J., Cody A. M., Queloz D., 2017, *ApJ*, **849**, 11
- Girardi L., Bertelli G., Bressan A., Chiosi C., Groenewegen M. A. T., Marigo P., Salasnich B., Weiss A., 2002, *A&A*, **391**, 195
- Han E., et al., 2017, *AJ*, **154**, 100
- Hartman J. D., et al., 2018, *AJ*, **155**, 114
- Helminiak K. G., et al., 2015, *MNRAS*, **448**, 1945
- Higl J., Weiss A., 2017, *A&A*, **608**, A62
- Hunter J. D., 2007, *Computing in Science and Engineering*, **9**, 90
- Ireland L. G., Browning M. K., 2018, *ApJ*, **856**, 132
- Jackson R. J., Jeffries R. D., 2014, *MNRAS*, **441**, 2111
- Jackson R. J., Deliyannis C. P., Jeffries R. D., 2018, *MNRAS*, **476**, 3245
- Jones E., Oliphant T., Peterson P., et al., 2001, SciPy: Open source scientific tools for Python. <http://www.scipy.org/>
- Kesseli A. Y., Muirhead P. S., Mann A. W., Mace G., 2018, *AJ*, **155**, 225
- Kesseli A. Y., et al., 2019, *AJ*, **157**, 63
- Kochukhov O., Shulyak D., 2019, arXiv e-prints,
- Kraus A. L., Tucker R. A., Thompson M. I., Craine E. R., Hillenbrand L. A., 2011, *ApJ*, **728**, 48
- Kraus A. L., et al., 2017, *ApJ*, **845**, 72
- Lanzafame A. C., et al., 2018, *A&A*, **616**, A16
- Limber D. N., 1958, *ApJ*, **127**, 363
- Lindgren L., et al., 2018, *A&A*, **616**, A2
- López-Morales M., 2007, *ApJ*, **660**, 732
- MacDonald J., Mullan D. J., 2014, *ApJ*, **787**, 70
- MacDonald J., Mullan D. J., 2017, *ApJ*, **850**, 58
- Maíz Apellániz J., Weiler M., 2018, *A&A*, **619**, A180
- Mann A. W., Gaidos E., Ansdell M., 2013, *ApJ*, **779**, 188
- Mann A. W., Feiden G. A., Gaidos E., Boyajian T., von Braun K., 2015, *ApJ*, **804**
- Mann A. W., et al., 2019, *ApJ*, **871**, 63
- Marigo P., et al., 2017, *ApJ*, **835**, 77
- Masana E., Jordi C., Ribas I., 2006, *A&A*, **450**, 735
- McQuillan A., Aigrain S., Mazeh T., 2013, *MNRAS*, **432**, 1203
- McQuillan A., Mazeh T., Aigrain S., 2014, *ApJS*, **211**, 24
- Morales J. C., et al., 2009a, *ApJ*, **691**, 1400
- Morales J. C., Torres G., Marschall L. A., Brehm W., 2009b, *ApJ*, **707**, 671
- Mullan D. J., MacDonald J., 2001, *ApJ*, **559**, 353
- Neves V., Bonfils X., Santos N. C., Delfosse X., Forveille T., Allard F., Udry S., 2013, *A&A*, **551**, A36
- Newton E. R., Charbonneau D., Irwin J., Berta-Thompson Z. K., Rojas-Ayala B., Covey K., Lloyd J. P., 2014, *AJ*, **147**, 20
- Newton E. R., Irwin J., Charbonneau D., Berlind P., Calkins M. L., Mink J., 2017, *ApJ*, **834**, 85
- Oliphant T. E., 2015, Guide to NumPy, 2nd edn. CreateSpace Independent Publishing Platform, USA
- Parsons S. G., Marsh T. R., Copperwheat C. M., Dhillon V. S., Littlefair S. P., Gänsicke B. T., Hickman R., 2010, *MNRAS*, **402**, 2591
- Parsons S. G., et al., 2012a, *MNRAS*, **419**, 304
- Parsons S. G., et al., 2012b, *MNRAS*, **420**, 3281
- Parsons S. G., et al., 2012c, *MNRAS*, **426**, 1950
- Parsons S. G., et al., 2016, *MNRAS*, **458**, 2793
- Parsons S. G., et al., 2018, *MNRAS*,
- Pecaut M. J., Mamajek E. E., 2013, *ApJS*, **208**, 9
- Pietrzyński G., et al., 2013, *Nature*, **495**, 76
- Pyrzas S., et al., 2012, *MNRAS*, **419**, 817
- Rabus M., et al., 2019, *MNRAS*, **484**, 2674
- Reiners A., Basri G., Browning M., 2009, *ApJ*, **692**, 538
- Robin A. C., Reylé C., Derrière S., Picaud S., 2003, *A&A*, **409**, 523
- Rosen S. R., et al., 2016, VizieR Online Data Catalog, **9050**
- Rozyczka M., Kaluzny J., Krzeminski W., Mazur B., 2007, *Acta Astron.*, **57**, 323
- Rozyczka M., Kaluzny J., Pietrukowicz P., Pych W., Mazur B., Catelan M., Thompson I. B., 2009, *Acta Astron.*, **59**, 385
- Schweitzer A., et al., 2019, arXiv e-prints,
- Skrutskie M. F., et al., 2006, *AJ*, **131**, 1163
- Somers G., Pinsonneault M. H., 2016, in Kastner J. H., Stelzer B., Metchev S. A., eds, IAU Symposium Vol. 314, Young Stars & Planets Near the Sun. pp 91–94 ([arXiv:1510.01702](https://arxiv.org/abs/1510.01702)), [doi:10.1017/S1743921315006092](https://doi.org/10.1017/S1743921315006092)
- Southworth J., 2015, in Rucinski S. M., Torres G., Zejda M., eds, Astronomical Society of the Pacific Conference Series Vol. 496, Living Together: Planet, Host Stars and Binaries. p. 164 ([arXiv:1411.1219](https://arxiv.org/abs/1411.1219))
- Stassun K. G., Mathieu R. D., Vaz L. P. R., Stroud N., Vrba F. J., 2004, *ApJS*, **151**, 357
- Stauffer J. R., et al., 2007, *ApJS*, **172**, 663
- Taylor M. B., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, Astronomical Society of the Pacific Conference Series Vol. 351, Astronomical Data Analysis Software and Systems XV. p. 666
- Terrien R. C., Fleming S. W., Mahadevan S., Deshpande R., Feiden G. A., Bender C. F., Ramsey L. W., 2012, *ApJ*, **760**, L9
- Terrien R. C., Mahadevan S., Deshpande R., Bender C. F., 2015, *The Astrophysical Journal Supplement Series*, **220**, 16
- Thompson B., Frinchaboy P., Kinemuchi K., Sarajedini A., Cohen R., 2014, *AJ*, **148**, 85
- Torres G., 2013, *Astronomische Nachrichten*, **334**, 4
- Torres G., Ribas I., 2002, *ApJ*, **567**, 1140
- Torres G., Sandberg Lacy C. H., Pavlovski K., Feiden G. A., Sabby J. A., Bruntt H., Viggo Clausen J., 2014, *ApJ*, **797**, 31
- Walkowicz L. M., Hawley S. L., West A. A., 2004, *PASP*, **116**, 1105
- Welsh W. F., et al., 2015, *ApJ*, **809**, 26
- Wilson T. J., Naylor T., 2017, *MNRAS*, **468**, 2517
- Wilson T. J., Naylor T., 2018, *MNRAS*, **481**, 2148
- Wright E. L., et al., 2010, *AJ*, **140**, 1868
- Wright N. J., Drake J. J., Mamajek E. E., Henry G. W., 2011, *ApJ*, **743**, 48
- Wright N. J., Newton E. R., Williams P. K. G., Drake J. J., Yadav R. K., 2018, *MNRAS*, **479**, 2351