Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **471,** 770–791 (2017)
Advance Access publication 2017 June 15

doi:10.1093/mnras/stx1433

Fundamental parameters and infrared excesses of Tycho–*Gaia* stars

I. McDonald,[1]★ A. A. Zijlstra[1,2] and R. A. Watson[1]

[1]*Jodrell Bank Centre for Astrophysics, Alan Turing Building, Manchester M13 9PL, UK*
[2]*Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong*

Accepted 2017 June 7. Received 2017 June 6; in original form 2017 February 14

ABSTRACT

Effective temperatures and luminosities are calculated for 1475 921 Tycho-2 and 107 145 *Hipparcos* stars, based on distances from *Gaia* Data Release 1. Parameters are derived by comparing multi-wavelength archival photometry to BT-SETTL model atmospheres. The 1σ uncertainties for the Tycho-2 and *Hipparcos* stars are $\pm$137 and $\pm$125 K in temperature and $\pm$35 and $\pm$19 per cent in luminosity. The luminosity uncertainty is dominated by that of the *Gaia* parallax. Evidence for infrared excess between 4.6 and 25 μm is found for 4256 stars, of which 1883 are strong candidates. These include asymptotic giant branch (AGB) stars, Cepheids, Herbig Ae/Be stars, young stellar objects and other sources. We briefly demonstrate the capabilities of this data set by exploring local interstellar extinction, the onset of dust production in AGB stars, the age and metallicity gradients of the solar neighbourhood and structure within the Gould Belt. We close by discussing the potential impact of future *Gaia* data releases.

Key words: circumstellar matter – stars: fundamental parameters – Hertzsprung–Russell and colour–magnitude diagrams – stars: mass-loss – solar neighbourhood – infrared: stars.

1 INTRODUCTION

Modern precision astrometry has recovered distances to large samples of nearby stars, the pinnacles of which are the catalogues returned by the *Hipparcos* (Perryman 1989) and *Gaia* satellites (Perryman et al. 2001; Gaia Collaboration et al. 2016a,b). These catalogues provide the basic measurements of colour, brightness and parallactic distance. They do not contain fundamental parameters, such as temperature or luminosity. Hence, 'value added' catalogues are often computed (e.g. Anderson & Francis 2012; McDonald, Zijlstra & Boyer 2012a) for the *Hipparcos* data set. The latter of these papers provides a catalogue of stellar fundamental parameters, which is replicated here using the *Gaia* satellite's Data Release 1 (DR1).[1]

Gaia DR1 is based on the first six months of *Gaia* operations. It lists parallaxes for 2057 050 stars contained in the *Hipparcos* Tycho-2 catalogue (Høg et al. 2000; Michalik, Lindegren & Hobbs 2015). We use spectral energy distribution (SED) fitting of pre-existing photometry to place those stars on the true Hertzsprung–Russell (H–R) diagram. We also identify the stars among them with infrared excess: i.e. excess flux in the mid-infrared (mid-IR, ∼3–30 μm) when compared to the SED from a stellar atmosphere model.

SED fitting to determine stellar parameters has its advantages and limitations. Compared to simple, single-colour bolometric corrections, it can be more robust against bad photometric data. It can also be more accurate, due to the larger number of data points included, and it can be effective over a wider range of stellar effective temperatures. Secondary effects, such as binary companions or reprocessing of stellar light, can sometimes be identified where simple bolometric corrections would not be able to do so. Both bolometric corrections and SED fitting are equally limited by prior assumptions of stellar metallicity, surface gravity and interstellar extinction, which determine the properties of the stellar atmosphere models that the stars are compared against. Stellar temperatures and luminosities from SED fitting are most accurate if both the short- (Wien) and long-wavelength (Rayleigh–Jeans) tails of the SED are covered with good-quality photometry.

Spectroscopic temperature determinations generally have greater accuracy than those obtained through SED fitting. They can also measure metallicity and surface gravity, and are not affected by extinction. However, SED fitting is observationally and computationally much cheaper, allowing it both to be used on fainter stars and to more effectively survey a larger number of stars. SED fitting provides a more accurate luminosity than can be derived via spectroscopic measurements. This allows SED fits to be used to select targets for more expensive follow-up campaigns.

Both photometric colours and spectroscopy often fail to identify infrared excess. Infrared excess is typically caused by warm dust in the circumstellar environment. It is therefore a good tracer of objects at both ends of stellar evolution: young and pre-main-sequence stars that have yet to clear their circumstellar environments of their protoplanetary discs, and evolved stars that are undergoing the terminal process of stellar mass-loss (e.g. Cotten & Song 2016). Other mass-losing or mass-gaining stars can also be identified, such as

★E-mail: iain.mcdonald-2@manchester.ac.uk
[1] http://gea.esac.esa.int/archive/

interacting binary stars containing an accretion disc, Wolf–Rayet stars and Herbig A[e]/B[e] stars. Unlike simple photometric colours, infrared excess can also trace unresolved, non-interacting binary companions and physically separate line-of-sight binary stars, if the contrast ratio is sufficiently close to unity and the colours are sufficiently different.

In this paper, we cross-reference catalogues of multi-wavelength literature photometry to construct SEDs for stars in the Tycho-2 and *Hipparcos* catalogues (Høg et al. 2000; van Leeuwen 2007), supplemented by the Tycho–*Gaia* astrometric solution from *Gaia* DR1. These are compared against stellar atmosphere models to derive effective temperatures for each star. When combined with the parallax information from *Gaia* DR1, this allows us to derive the luminosity of each star (Section 2) and to place it on the H–R diagram. The H–R diagram is presented, and the uncertainties in individual measurements discussed (Section 3). A catalogue of stars that likely exhibit excess infrared flux is presented, and their categorization and location in the H–R diagram is discussed (Section 4). Here, we also explore dust production by evolved stars. Further applications and details are presented in the online appendices that accompany this paper.

2 THE SED-FITTING PROCESS

2.1 Methodology

2.1.1 Cross-referencing photometric source data

This section describes the methodology used to create the SEDs and fit the data. The practical application is detailed in Section 2.2.

A cross-reference catalogue was intended to form part of the *Gaia* DR1 but was not provided with the data release itself. For this paper, photometric data were collected using the CDS 'X-Match' cross-matching service,[2] which provides fast, effective cross-matching across a variety of photometric catalogues.

While fast and efficient, the VizieR cross-matching service contains some limitations. For example, in the following analysis, SDSS Data Release 7 was used in preference to Data Release 9: although DR9 is more complete, the VizieR implementation also matches child objects[3] instead of their parents, resulting in improper photometric matches. Flagging data from DR7 was not passed to the cross-matching service,[4] meaning (e.g.) saturated stars cannot automatically be removed.

A further limitation is that source proper motion is not accounted for during the cross-matching process. Already, many nearby stars are not in the *Gaia* DR1 sample due to their proper motion cut-off of 750 mas yr^{-1}. Unfortunately, this lack of accounting for proper motion appears to remove considerably more. The effect depends both on the 3σ tolerance and the temporal spacing between catalogues. For recent ($\sim$2012) catalogues like AllWISE, comparison to the $\sim$1991 Tycho photometry with a limit of 1.2 arcsec risks removing any object with proper motion greater than 57 mas yr^{-1}, or 5 per cent of the combined Tycho–*Hipparcos* sample.

From this compiled list, we removed stars where the photometric parallax is too uncertain to obtain a meaningful luminosity. We dictated this to be when the uncertainty in the parallax ($\delta\varpi$)[5] led to a factor of 2 uncertainty in the stellar luminosity, i.e. when $\delta\varpi/\varpi > 0.414$. This reduced the number of Tycho–*Gaia* sources from 2057 050 to 1535 006. We explicitly note that the parallax cut-off we have made means that this is not a volume-selected or volume-limited sample. It should not be considered complete for any given set of stars, and retains the biases and limitations present in the *Gaia* and *Tycho* catalogues, and the other photometric catalogues used later.

The bespoke, iterative methods by which we removed bad data from the compiled SEDs are detailed later, in Section 2.3.2 and the online appendix.

We stress that this sample of stars is subject to the Lutz–Kelker bias (Lutz & Kelker 1973). The fractional parallax uncertainty we have used is still relatively lax, and we encourage users to adopt stricter criteria for volume-limited samples. The minimum suggested criterion we can recommend is the $\delta\varpi/\varpi < 0.2$ limit we use in parts of our analysis below (cf. Bailer-Jones 2015). Further discussion on Lutz–Kelker-related effects can be found in Section 3.2.2.

2.1.2 SED-fitting methodology

Once the source data are collated to provide an SED for each star, the fitting procedure can determine the best-fitting spectral model and derive the stellar temperature and luminosity. The GETSED SED-fitting pipeline used here was first described in McDonald et al. (2009) and updated in McDonald et al. (2012a). The pipeline has been altered slightly for this paper to improve efficiency and reduce artefacts in the final H–R diagrams caused by discrepant data. The following provides an account of the fitting procedure, including these alterations.

The pipeline begins with an SED from observed photometry in the form of λ, F_ν. Required metadata are the (*Gaia*) distance, the interstellar extinction to the star and the stellar metallicity. Unless stated otherwise, in the following discussion, we use an assumption of solar metallicity and zero extinction.

Step 1. The best-fitting blackbody is calculated to provide a first estimate of stellar parameters. Each filter is reduced to a single, representative wavelength. The flux of a blackbody at these wavelengths is calculated for a grid of temperatures with 400 K spacing over the range 2600–7400 K. The blackbody is normalized to the wavelength-integrated (bolometric) flux of the observed SED, and a χ^2 minimum is computed. This and later χ^2 minima are determined in magnitudes, rather than fluxes, to avoid giving undue weight to points around the SED peak. If the best-fitting temperature is 7400 K, the temperature range is extended up to 20 000 K, then 60 000 K. A sub-grid is defined at $\pm$200 K from the best-fitting temperature, and a χ^2 minimum computed, then iterated down to 100 and 50 K, thus fitting a blackbody temperature between 2250 and 60 350 K with 50 K resolution.

The apparent bolometric flux of the blackbody fit is used in combination with the input distance to determine the luminosity of the fitted blackbody. This identifies whether the star is a main-sequence star or a giant. A mass is estimated using the procedure

[2] http://cdsxmatch.u-strasbg.fr/xmatch

[3] Sources that SDSS notes as resolved or overlapping are assigned a parent object, then deblended and decomposed into child objects. This process can also occur with saturated stars and artefacts associated with them. Further details are given on the SDSS webpages: http://www.sdss.org/dr12/algorithms/deblend/

[4] We thank the staff at Centre de Données astronomiques de Strasbourg for later including these on our suggestion.

[5] In the remainder of this work, we use δ to denote the uncertainty on an individual object, and σ to denote the standard deviation, uncertainty or any other noted derivative of variance in a statistical ensemble.

described in McDonald et al. (2012a), and this mass is used to obtain a surface gravity, $\log(g)$. The temperature change caused by an imperfect mass and $\log(g)$ estimate is small compared to the total error budget (Section 3.2), provided the mass is within a factor of ∼10 of the true value. For main-sequence and red giant branch (RGB) stars, we expect our masses to be correct to well within a factor of 2, and for asymptotic giant branch (AGB) stars within a factor of 4–10 (depending on their luminosity).

Step 2. Unlike previous implementations, we now repeat this process with a grid of model atmospheres. For this paper, we use the BT-SETTL models of Allard et al. (2003). We use these in preference to the more widely used MARCS models (Decin et al. 2004; Gustafsson et al. 2008) because of their greater completeness. While there are substantial and astrophysically important differences between these models, tests performed in McDonald et al. (2012a) showed that the choice of model atmosphere has negligible impact on the final temperature derived for a variety of types of star.

Each model in the grid is reddened, using the procedure described in McDonald et al. (2009, see also Section 3.2.3), and convolved with a list of filter transmission functions. The flux that would be observed in each filter, and the relative reddening in that filter (A_λ/A_V), is tabulated.

Models are selected from the grid, bracketing the star's assumed metallicity and $\log(g)$. This creates a selection of four models at each temperature point. A two-dimensional, linear interpolation is made to obtain a single photometric flux for each band at each gridded temperature point. The luminosity of each model is then normalized to the luminosity of the SED, and a χ^2 minimum performed to determine the best-fitting temperature. A new value for $\log(g)$ is determined.

Step 3. We interpolate within the now-one-dimensional temperature model grid, modify $\log(g)$ and iterate to a solution. This last two-stage interpolation is the most computationally expensive part of the analysis: unlike before, this interpolation is performed for each point on each filter transmission function, therefore better accounting for wavelength-dependent effects such as molecular band strength changes and interstellar reddening. The two stages of this interpolation are as follows.

(a) We begin our initial temperature interpolation by computing two models, above and below the best-fitting temperature. The deviation above and below is taken as the largest power of two that is numerically less than the temperature grid spacing of the stellar atmosphere models: e.g. if the grid spacing is 100 K, the models are computed at the gridded best-fitting temperature ± 64 K; if the grid spacing is 250 K, a deviation of ± 128 K is applied. If one of these interpolated models is a better χ^2 fit than the original, its temperature is adopted as the new best fit, otherwise the old best-fitting temperature remains. Models are computed at the new best-fitting temperature ± half the previous value, and the process iterated. In our example, that is namely ± 32 K, then ± 16, ± 8, ± 4, ± 2 and ± 1 K, allowing the new best-fitting temperature to deviate from the original by up to 127 K.

(b) A new $\log(g)$ is now determined, and the temperature iteration begun again. To optimize the system, the process begins at the smallest power of two above the deviation from the original value. For example, a star may be initially fitted at 5800 K, and interpolated to 5776 K, the difference being 24 K. The interpolation would then start by interpolating new models at 5776 ± 32 K, rather than ± 64 K as previously.

These two steps (a and b) are iterated until a solution is found. In a small fraction of cases, the solution can oscillate between two solutions, or run towards zero or infinity. To prevent this, the starting deviation of each interpolation is tapered. It is allowed to run at the initial value for three times, then is limited by half at each step. In our example, this limits the interpolation to a maximum deviation to ± 64, 64, 64, 32, 16, 8, 4, 2 and 1 K on subsequent iterations. This allows our example model to deviate by no more than 255 K from its initial best-fitting value (for a grid spacing of 100 K). Investigation showed that this was sufficient to account for any difference in temperature caused by a revised $\log(g)$.

Step 4. Once a best-fitting temperature, luminosity and $\log(g)$ have been determined, the final interpolated model atmosphere is integrated in frequency and a final luminosity produced. The normalized χ^2 minimum is calculated. For each of the n observed filters, the ratio of the observed to modelled flux ($R_n = F_o/F_m$) is computed. A goodness-of-fit metric (Q) is calculated, based on the number of points (n):

$$Q = \sum_n \frac{\left(R_n^* - 1\right)}{n}, \tag{1}$$

where $R_n^* = R_n$ if $R_n > 1$ or R_n^{-1} otherwise. This metric gives $Q = 0$ for a perfectly fitted data set and (e.g.) reaches $Q = 1$ for a data set where the average deviation from the model fit is a factor of 2.

2.2 Data analysis

The data were divided into two subsets, the first corresponding to stars in the original Tycho-2 astrometric and proper motion catalogue, the second to stars in the mission's primary *Hipparcos* catalogue, which also includes parallax data of its own. This separation was motivated by the comparative optical brightness of the *Hipparcos* stars and the greater accuracy in their *Gaia* DR1 parallax.

2.2.1 The Tycho-2 data set

We used the original Tycho-2 catalogue as the astrometric reference, as it is temporally closer to the epoch of the surveys we cross-reference against. A number of catalogues were cross-correlated against the Tycho-2 catalogue, allowing matches within an initial tolerance of 5 arcsec.

For certain catalogues, a 5 arcsec tolerance allows one or more spurious sources to be wrongly matched to the Tycho-2 source. To circumvent this, each matched catalogue was sorted by the distance of the match from the Tycho-2 position, and the 1σ deviation in distance was determined, corresponding to the matching radius at which 68.3 per cent of the sources cross-matched at 5 arcsec tolerance were included. For each catalogue, cross-matches were retained if they fell within 3σ of the Tycho-2 source. The cross-matched source catalogues and their adopted 3σ tolerances (in parentheses[6]) are given below.

(i) The American Association of Variable Star Observers (AAVSO) Photometric All-Sky Survey Data Release 9 (1.65 arcsec; released as VizieR catalogue II/336/apass9: Henden et al., in preparation).[7]

(ii) The Sloan Digital Sky Survey (SDSS) Data Release 7 (1.94 arcsec; Abazajian et al. 2009).

[6] Tolerances for IPHAS and *IRAS* are set manually, rather than using the 3σ cut-off.

[7] http://www.aavso.org/apass

(iii) The Isaac Newton Telescope (INT) Photometric Hα Survey of the Northern Galactic Plane (IPHAS) Data Release 2 (0.70 arcsec; Barentsen et al. 2014).

(iv) The United Kingdom InfraRed Telescope (UKIRT) Infrared Deep Sky Survey Large Area Survey Data Release 9 (4.62 arcsec).

(v) The Deep Near Infrared Survey of the Southern Sky (DENIS) Third Data Release (1.15 arcsec; released as VizieR catalogue B/denis);

(vi) The Two Micron All-Sky Survey (2MASS) all-sky catalogue (0.71 arcsec; Cutri et al. 2003).

(vii) The *AKARI*/Infrared Camera all-sky survey (2.34 arcsec; Ishihara et al. 2010).

(viii) The *Wide-Field Infrared Survey Explorer AllWISE* all-sky catalogue (abbreviated WISE; 1.20 arcsec; Skrutskie et al. 2006).

(ix) The *InfraRed Astronomical Satellite* (*IRAS*) all-sky survey (5 arcsec; Neugebauer et al. 1984).

2.3 The *Hipparcos* data set

This procedure was broadly repeated for the *Hipparcos* data. Here, parallaxes were taken from the Tycho–*Gaia* DR1 catalogue if they had been updated, or the 'new' *Hipparcos* reduction of van Leeuwen (2007) if they had not. In the combined catalogue, 88 417 objects had revised parallaxes, while 18 915 parallaxes come from the original data set. This includes objects with high proper motions and very red colours, which are known to be missing from the *Gaia* data set (Section 2.1.1). Objects were removed if they had negative parallaxes, or if they had parallax uncertainties greater than $\delta\varpi/\varpi > 0.414$, totalling 6399 objects.

The *Hipparcos* stars are typically much brighter than the Tycho-2 stars, resulting in severe saturation problems that rendered several catalogues unusable. A significant number of brighter stars have insufficient photometry to make a good fit: often only Tycho B_T and V_T, and the *Hipparcos* H_p data, which together do not cover a sufficiently large range of wavelengths to constrain the SED. For this reason, we have incorporated a number of additional optical and infrared catalogues of bright stars. This increased data set makes us more robust against bad data (as it is easier to flag), at the expense of maintaining a homogeneous catalogue between the *Hipparcos* and Tycho-2 stars. The extra catalogues are as follows.

(i) Mermilliod's 'Photoelectric Photometric Catalogue of Homogeneous Means in the UBV System' (see Warren 1991).

(ii) Morel & Magnenat (1978), containing *UBVRIJHKLMN*-band photometry.

(iii) The *Cosmic Background Explorer* (*COBE*) Diffuse Infrared Background Experiment (DIRBE) Point Source Catalogue (Smith, Price & Baker 2004).

(iv) The *Midcourse Space Experiment* Astrometric Catalogue (Egan & Price 1996).

Astrometric matching tolerances for the four catalogues were set respectively to 0.7, 0.47, 0.66 and 5 arcsec. Data were fitted with the SED fitter as above. A detailed discussion of the methods used to remove bad data is listed in the online appendix. We stress again that proper motions have not been taken into account in our simple matching exercise: the limited astrometric matching radius means that photometric data will not always be matched for stars with proper motions that are significant on the ∼15 yr time-scales between the *Hipparcos* observations and the relevant catalogue observations. In many cases, a faint, unrelated source may be matched instead. Care has been taken to remove these from the catalogue where they stand out.

2.3.1 *Interstellar extinction*

The line-of-sight interstellar extinction was estimated using maps from the *Planck* Legacy Archive. *Planck* provides visible extinction maps based on the Draine & Li (2007) dust model in HEALPIX format in Galactic coordinates. To facilitate cross-referencing, the Galactic longitude and latitude for each star in the *Hipparcos* and Tycho-2 catalogues were derived via the VizieR portal, and the PYTHON HEALPY ANG2PIX routine was used to locate HEALPIX pixels corresponding to catalogue positions, providing the extinction for each object.

Without assuming a prior model for Galactic extinction, there is no ready means to tell whether the extinction lies behind or in front of the object of interest. We must therefore compute two estimates, one with zero and one with full line-of-sight extinction, to bracket the possible range of model fits. Further information on the use of these interstellar extinction data is given in Section 3.2.3.

2.3.2 *Removing bad data*

The data quality of the fitted photometry can be tested using both the goodness of fit of individual data points and the overall goodness of fit of a star's SED. These can be used as a basis for removing bad data from the sample. Due to the extensive nature of these tests, and the complex way in which bad data are deleted from the data set, we have moved the detailed discussion of this topic to the online appendices. Sources with three or more remaining photometric points were retained for the catalogue: this reduced the number of fitted stars to 1475 921.

3 THE FINAL CATALOGUE AND H–R DIAGRAM

3.1 The catalogue and diagram

Fig. 1 shows the main H–R diagram of the combined Tycho–*Gaia* and *Hipparcos*–*Gaia* data sets, under the assumption of zero interstellar extinction. The top panel contains the entire data set, while the bottom panel shows a restricted subset of well-fitted objects. These data are tabulated in Tables 1 and 2 , for the Tycho-2 and *Hipparcos* stars, respectively.

The upper panel of Fig. 1 shows several artefacts. The main sequence is broad, reflecting the higher extinction and greater parallax uncertainties in some of the data. Vertical bands of red symbols (poorly fitted stars) in the most luminous regions of the diagram come mainly from *Hipparcos* stars that are not well modelled by a single stellar atmosphere model. The vertical stripe between 3400 and 3500 K on the upper giant branch seems largely occupied by stars that have a combination of high reddening and uncertain distances: these are mostly normal giant branch stars that have been pushed on to this artificial sequence by interstellar reddening.

The giant branch also has a significant overdensity about halfway along its length: this is a real feature, representing the merged features of the RGB bump and red clump.[8]

The lower panel of Fig. 1 shows a subset of same data, but with poor-quality data removed (objects on highly extincted lines

[8] The RGB bump is a concentration of stars on the hydrogen-burning RGB, caused by the transition of the hydrogen-burning shell into material that has previously been convectively mixed. The red clump is the high-mass equivalent of the horizontal branch, and represents the core-helium-burning phase of giant branch evolution (e.g. Karakas & Lattanzio 2014).

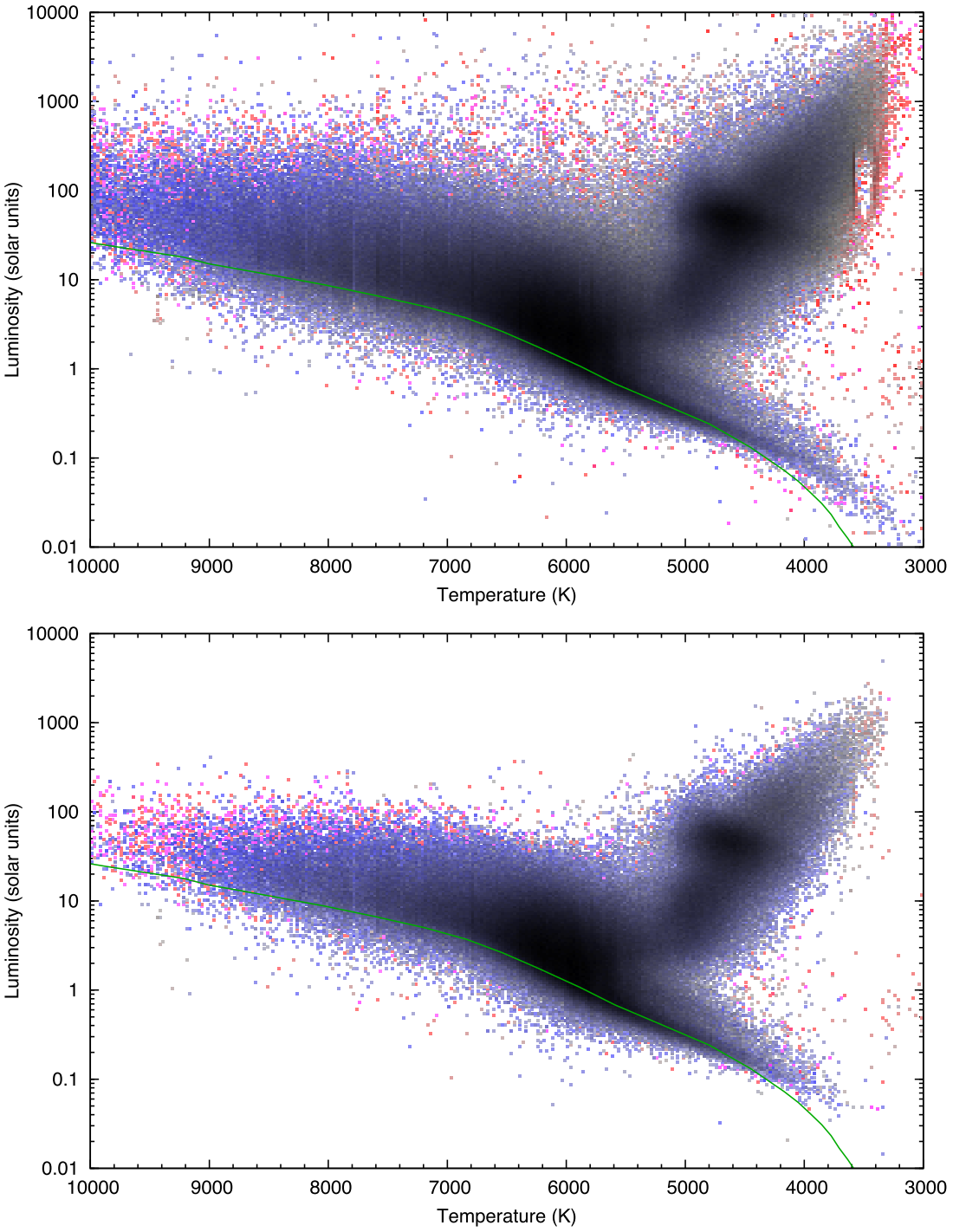

Figure 1. The H–R diagram of nearby stars. Darker points represent a greater density of stars. The average value of $\log(Q)$ for each bin is indicated by colour: blue colours denote the best fits, grey colours denote intermediate fits and red colours denote the worst fits. Systematic deviations from unity can be caused by poor-quality input photometry, or poor fitting by the model atmospheres. The zero-age main sequence is shown in green (Marigo et al. 2008). The bottom panel shows a restricted set (40 per cent) of objects, with <25 per cent parallax uncertainty, line-of-sight $A_V < 3$ mag and goodness-of-fit $Q < 0.5$.

of sight, with large parallax uncertainties, or where the SEDs are not well fitted by a single stellar model). In this lower panel, the main sequence stands out clearly, being best populated for solar-like stars, but with distributions tailing off towards very hot temperatures (rare stars that cannot be well modelled without good UV data and extinction corrections) and towards very low temperatures (faint stars missing due to photometric incompleteness).

Both panels include a zero-age main-sequence (ZAMS) model, derived from the Padova stellar evolution models of Marigo et al. (2008). The lower main sequence, between ∼4600 and ∼5400 K, fits the ZAMS model very well. At temperatures >5400 K, scatter above the ZAMS line indicates the presence of more evolved main-sequence stars, which are approaching the main-sequence turn-off. This can be used to extract age information about the solar

Table 1. Fundamental parameters and infrared excess for Tycho-2 stars. A portion of the online table is shown here, where table columns are numbered for clarity. The columns are described in full in the text, but can briefly be described as (1) Tycho-2 reference number; (2) Tycho-2 right ascension; (3) Tycho-2 declination; (4) and (5) Tycho-2 Galactic latitude and longitude; (6) and (7) distance and fractional uncertainty; (8) and (9) extinction and absolute uncertainty; (10) and (11) effective temperature and absolute uncertainty; (12) and (13) luminosity and fractional uncertainty; (14) implied stellar radius; (15) assumed surface gravity; (16) and (17) fitted temperature and luminosity when full line-of-sight reddening is applied; (18) and (19) fitted temperature and luminosity under the Lutz–Kelker correction of Astraatmadja & Bailer-Jones (2016); (20) fit quality; (21)–(24) number of data points in (respectively) the full SED, and optical, near-IR and mid-IR regions; (25)–(28) average fit deviation in the total SED, and optical/near-IR/mid-IR regions, respectively; (29) mid-IR excess; (30) mid-IR excess with most excessive data point removed; (31) (uncalibrated) significance of the excess; (32) deviation of most excessive data point; (33) luminosity of the infrared excess; (34) fraction of reprocessed infrared light; (35) peak wavelength of infrared excess; (36)–(55) deviation of individual data points; (56)–(75) fluxes of data points used in final fit. Complete versions are to be made available through the Centre de Données astronomiques de Strasbourg (CDS).

(1) TYC	(2) RA (J2000)	(3) Dec. (J2000)	(4) G. Lat. (deg)	(5) G. Long. (deg)	(6) d (pc)	(7) $\delta\varpi/\varpi$	(8) A_V (mag)	(9) δA_V (mag)
1000-1016-1	264.019 440	11.275 677	34.759 265	21.778 061	575.585	0.137	0.897	0.037
1000-1018-1	262.982 107	11.568 592	34.585 083	22.823 855	347.823	0.094	0.816	0.016
1000-1043-1	264.093 473	12.636 898	36.126 451	22.280 018	465.817	0.120	1.365	0.066
...	...	...	...	...	...	...	...	...

(10) $T_{\rm eff}$ (K)	(11) $\delta T_{\rm eff}$ (K)	(12) L ($L_\odot$)	(13) $\delta L/L$	(14) r ($R_\odot$)	(15) log(g) (dex)	(16) $T_{\rm Av}$ (K)	(17) $L_{\rm Av}$ ($L_\odot$)	(18) $T_{\rm ABJ}$ (K)	(19) $L_{\rm ABJ}$ ($L_\odot$)	(20) Q	(21 ··· 24) N				(25 ··· 28) $\Re$			
7182	212	12.524	0.147	2.289	3.881	7818	16.360	0	0.000	0.039	13	6	5	2	1.009	1.014	0.989	1.040
6020	138	1.945	0.104	1.284	4.273	6399	2.384	6020	1.945	0.070	13	6	5	2	1.024	1.012	1.016	1.083
4769	125	57.760	0.131	11.148	2.341	5138	73.964	4769	57.760	0.052	14	5	5	4	1.019	1.002	1.011	1.050
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...		

(29) $X_{\rm MIR}$	(30) $X'_{\rm MIR}$	(31) $S_{\rm MIR}$	(32) $R_{\rm max}$	(33) $L_{\rm XS}$ ($L_\odot$)	(34) $f_{\rm XS}$	(35) $\lambda_{\rm XS,\,peak}$ (μm)	(36) (F_o/F_m) (Jy)	··· {BT ... [25]}	(55) F_o, (Jy)	(56) {BT ... [25]} (Jy)	···	(75) (Jy)
1.037	1.052	1.471	1.090	0.0002	0.000 015	3.4	1.010	···	0.000	139.132	···	0.000
1.068	1.066	1.669	1.269	0.0001	0.000 077	2.3	1.269	···	0.000	50.842	···	0.000
1.043	1.039	1.916	1.135	0.0064	0.000 111	2.2	1.043	···	0.000	306.356	···	0.000
...	...	...	...	...	...	...	...	...	...	...	...	...

neighbourhood. The bottom end of the main sequence is not well fitted by a ZAMS model, but this deviation is substantially reduced in the lower panel. This suggests that it results from a combination of photometric inaccuracy or incompleteness near the sensitivity limit of photometric data bases (including Tycho-2 itself), biased scatter upwards in the diagram due to uncertain parallaxes (possibly a manifestation of the Lutz–Kelker bias; Lutz & Kelker 1973), and (in a limited number of cases) heavy reddening of lower main-sequence stars.

Many cool stars on the upper giant branch are not included in the lower panel of Fig. 1. Several factors contribute to this. (1) Despite their luminosity, these are often red, optically faint stars, which consequently have significant uncertainties in their Tycho-2 positions, hence also in their *Gaia* parallaxes. (2) Being luminous stars, these stars are visible at large distances from the Earth, and congregate in the Galactic plane, so are more often subject to strong interstellar extinction than nearby stars. (3) Variability of stars in this part of the H–R diagram leads to variability-induced motion (see van Leeuwen 2007), which increases the uncertainty in their parallax. Variability also worsens (increases) the SED quality estimator, Q. (4) A substantial fraction of these stars have circumstellar dust, which reprocesses their light from the optical into the infrared, resulting them in being poorly fitted by a simple stellar SED.

3.2 Limitations and uncertainties

For well-fitted stars, the three primary sources of uncertainty in this analysis are (1) random and systematic uncertainties in the source data; (2) Lutz–Kelker effects when converting parallax to distance; (3) systematic 'cooling' of the SEDs caused by interstellar reddening and (4) the effect on the stellar temperature of the unknown metallicity of each star.

3.2.1 Random versus systematic uncertainties

Formal uncertainties for SED fitting of this nature are difficult to determine. The published photometric uncertainties for many of the public surveys can grossly underestimate the true uncertainties involved, within individual catalogues, across catalogues and across different epochs. For example, the 2MASS photometric uncertainties can be as low as a few millimagnitudes, and represent the internal error in the catalogue, yet the photometric zero-points are uncertain by ∼2 per cent.[9] Different surveys take these uncertainties into account in different ways, and to different degrees. Across

[9] http://www.ipac.caltech.edu/2mass/releases/allsky/faq.html

Table 2. Fundamental parameters and infrared excess for *Hipparcos* stars. A portion of the online table is shown here, where table columns are numbered for clarity. The columns are described in full in the text, but can briefly be described as (1) *Hipparcos* reference number; (2) *Hipparcos* right ascension; (3) *Hipparcos* declination; (4) and (5) *Hipparcos* Galactic latitude and longitude; (6)–(35) as Table 1; (36) source of parallax (*Hipparcos*/*Gaia*); (37)–(62) deviation of individual data points; (63)–(90) fluxes of data points used in final fit. Complete tables are to be found at CDS.

(1) HIP	(2) RA (J2000)	(3) Dec. (J2000)	(4) G. Lat. (deg)	(5) G. Long. (deg)	(6) d (pc)	(7) $\delta\varpi/\varpi$	(8) A_V (mag)	(9) δA_V (mag)
3	0.005 024	38.859 279	112.090 026	−22.927 558	350.804	0.344	0.929	0.083
4	0.008 629	−51.893 546	320.793 090	−63.415 309	135.654	0.039	0.124	0.039
5	0.009 973	−40.591 202	337.897 763	−72.861 671	381.080	0.092	0.057	0.019
...	...	...	...	...	...	...	...	...

(10) $T_{\rm eff}$ (K)	(11) $\delta T_{\rm eff}$ (K)	(12) L ($L_\odot$)	(13) $\delta L/L$	(14) r ($R_\odot$)	(15) log(g) (dex)	(16) $T_{\rm Av}$ (K)	(17) $L_{\rm Av}$ ($L_\odot$)	(18) $T_{\rm ABJ}$ (K)	(19) $L_{\rm ABJ}$ ($L_\odot$)	(20) Q	(21 ··· 24) N				(25 ··· 28) $\Re$			
7096	2561	194.076	0.732	9.230	2.642	7261	210.793	7093	210.805	0.618	10	3	4	3	1.281	1.229	0.670	2.147
6777	168	8.373	0.059	2.102	3.930	6834	8.523	6777	8.425	0.058	14	6	5	3	1.015	1.021	0.992	1.042
4885	125	56.536	0.106	10.512	2.364	4897	56.882	4885	55.987	0.039	13	5	4	4	1.021	1.015	0.999	1.050
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...		

(29) $X_{\rm MIR}$	(30) $X'_{\rm MIR}$	(31) $S_{\rm MIR}$	(32) $R_{\rm max}$	(33) $L_{\rm XS}$ ($L_\odot$)	(34) $f_{\rm XS}$	(35) $\lambda_{\rm XS,\,peak}$ (μm)	(36) G/H	(37) ··· (62) $(F_o/F_m)_{\{...\}}$ (Jy)			(63) ··· $F_{o,\,\{...\}}$ (Jy)		(90) (Jy)
2.361	3.206	3.217	3.523	0.0171	0.000 088	17.1	G	0.000	...	0.000	0.000	...	0.000
1.034	1.050	1.584	1.128	0.0001	0.000 013	8.2	G	0.000	...	0.000	0.000	...	0.000
1.042	1.051	2.667	1.103	0.0027	0.000 048	8.6	G	0.000	...	0.000	0.000	...	0.000
...	...	...	...	...	...	...	...	...	...	...	...	...	

catalogues, source blending and astrophysical sky background can become important, particularly in crowded regions and in the infrared. Across different epochs, stellar variability or proper motion can become significant.

This means that quantifying uncertainties on photometry and assigning appropriate weights is non-trivial. For this reason, no weighting was applied to the photometry during the fitting process. This can cause problems, particularly when observations are near the limit of photometric completeness. However, in such cases, fits can generally be improved simply by removing these photometric data points from the catalogue, as described in the online appendices.

For the luminosity measurement, in the vast majority of cases, the largest uncertainty is from the photometric parallax of the star (Fig. 2).

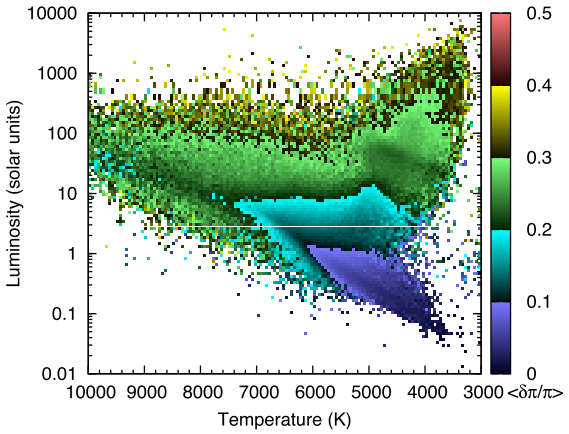

Figure 2. The fractional parallax uncertainty of stars in different regions of the H–R diagram. A binned average is displayed for each pixel.

3.2.2 *Lutz–Kelker effects*

The derived luminosity of a star is subject to the uncertainty in its distance and hence its parallax as $L \propto d^2 \propto \varpi^{-2}$. The probability distribution function (PDF) in parallax is normally expected to be Gaussian (e.g. Lutz & Kelker 1973; Bailer-Jones 2015). However, when inverting parallax to distance, the PDF becomes non-Gaussian and asymmetric. For stars with small fractional uncertainties, this is a relatively minor effect, but at large uncertainties it manifests itself in a variety of phenomena that can be broadly termed Lutz–Kelker effects, after Lutz & Kelker (1973).

The full range of Lutz–Kelker effects are complex, and there is no definitively appropriate way to correct for them. The magnitude by which Lutz–Kelker effects affect quantities derived from this data set varies according to the sub-sample chosen, particularly with respect to any limiting fractional parallax uncertainty ($\delta\varpi/\varpi$).

To account for the Lutz–Kelker effect, we present two sets of temperatures and luminosities. In the first, we present temperatures and luminosities derived from a simple inversion of parallax to obtain distance ($T_{\rm naive}$, $L_{\rm naive}$). For comparison, we also present temperatures and luminosities derived from distances quoted by Astraatmadja & Bailer-Jones (2016), who model the Lutz–Kelker effects on the *Gaia* DR1 sample using a population model of the Milky Way ($T_{\rm ABJ}$, $L_{\rm ABJ}$). We strongly advise the reader to explore which of these is most appropriate for their individual application, and to use the difference between the 'naïve' and 'ABJ' parameters

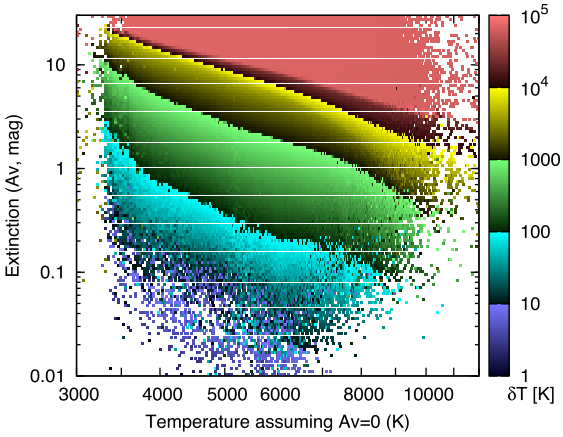

Figure 3. The correction (δT) applied to a star of given photometrically derived temperature behind a given column density of extinction. Stars with $\delta T \gtrsim T_{\rm eff}$ cannot reliably be fitted, even if the extinction is known.

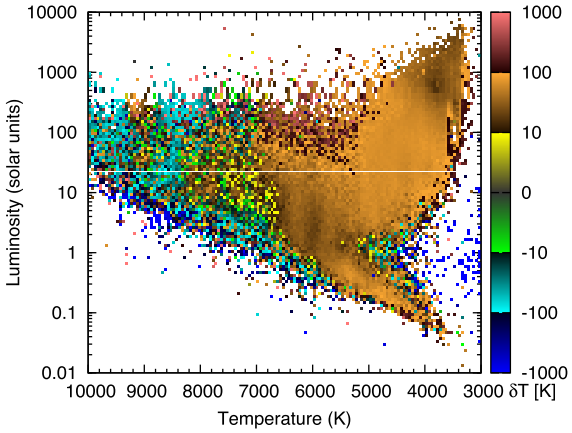

Figure 4. Difference in temperature between models assuming [Fe/H] = 0 and −0.5 dex, coded in the sense that positive numbers (red colours) denote cooler stellar temperatures in the metal-poor models, while negative numbers (blue colours) denote warmer stellar temperatures in the metal-poor models.

as a qualitative estimate of how much the Lutz–Kelker bias could affect their data.

A detailed comparison of these two sets of data is presented in online Appendix C. In summary, roughly 35 per cent of our stars are estimated to suffer some level of Lutz–Kelker bias in their naïve distances. The corrected luminosities for the remainder are almost all only modestly (a few per cent) different from the naïve assumptions. Barring a small number of stars, the corrections are all negligible compared to the luminosity uncertainties applied from other sources. The resulting distance changes also affect the assumed stellar gravity and (in many cases) stellar mass, resulting in a marginally different temperature distribution that is generally within the temperature uncertainties of the source in question and, for the vast majority of stars, within 200 K of the naïve estimate. While a detailed comparison of the two approaches is beyond the scope of this work, the corrected distances from Astraatmadja & Bailer-Jones (2016) result in either no clear improvement or a slightly *worse* fit to specific features on the H–R diagram; therefore, we retain the naïve estimates for use in the remainder of this paper.

3.2.3 *Interstellar reddening*

The interstellar reddening towards each star is unknown. The *Planck* data we use provide the line-of-sight reddening, which will be partly in front of and partly behind the star. To estimate the uncertainty this creates, we have dereddened the input photometry, assuming that the full *Planck* line-of-sight reddening is in front of the star, and rerun the SED-fitting code. For stars with large reddening, we also compute fits for $A_{\rm V} = 1, 2$ and 3 mag. The photometry is dereddened using the Milky Way $R_{\rm V} = 3.1$ extinction curve of Draine (2003). Dereddening is performed for each point in the model SED, before it is convolved with the filter transmission functions, ensuring accurate dereddening for sources with high extinction.

Fig. 3 shows the increase in temperature that must be applied to a star that is subject to a given amount of interstellar reddening. Taking the whole data set, the average star is 6000 K and lies in a line of sight with an extinction $A_{\rm V} = 1.0$ mag. If we assume half of this extinction to lie between us and the star, the *average* underestimation of the temperature for these stars is ∼240 K.

For most stars, this value should be conservatively large. At higher extinctions, there is a progressively greater chance that the star will be made too faint to be found in the Tycho-2 catalogue. The significant majority of stars in the Tycho-2 catalogue are below the completeness limit.[10] Due to the steep increase in number of stars per magnitude ($N{\rm d}V_{\rm T} \propto V_{\rm T}^9$), the vast majority of stars suffering significant extinction ($A_{\rm V} \gtrsim 1$ mag) will be reddened out of the Tycho-2 catalogue. This corollary should hold strongest for stars that are optically faint, hence stars of later spectral types (which need less correction) and more distant stars (which are likely to suffer from more reddening anyway). Therefore, the average star in our final catalogue should have a reddening correction that is >240 K. However, care should be taken for luminous stars and hot stars, where larger corrections could be required.

Further discussion on interstellar extinction and its spatial correlation can be found in Appendix D (online version only).

3.2.4 *Metallicity*

Fig. 4 shows the correction to our fitted stellar temperatures that must be applied to stars of [Fe/H] = –0.5 dex. Note that the BT-SETTL elemental abundance ratios also change during this step, from [α/Fe] = 0 to +0.2 dex. The majority of stars below ∼6500 K require a temperature adjustment of between −10 and −100 K if the metallicity is decreased to [Fe/H] = −0.5 dex. The majority of stars warmer than ∼6500 K require a temperature change of +10 to +100 K. Stars lying outside the main regions of the H–R diagram tend to be stars that are poorly fitted. Here, temperature changes of 1000 K are not uncommon, as a better fit can often result from relatively minor changes to the poorly constrained SED.

Different studies using differing methods yield different metallicity distributions for stars in the local neighbourhood (e.g. Taylor & Croxall 2005; Reid et al. 2007; Bensby, Feltzing & Oey 2014; Hinkel et al. 2014). The large majority of stars fall in the range $-0.3 \lesssim$ [Fe/H] $\lesssim +0.2$ dex, although significant tails make substantial contributions to $-0.9 \lesssim$ [Fe/H] $\lesssim +0.6$ dex. While age plays a factor in this spread, it is also location dependent, with metal-poor stars being further from the Galactic plane. It is expected that the

[10] The 90 per cent completeness limit in $V_{\rm T}$ is ∼11.5 mag, and 86 per cent of stars are fainter than this. The 10 per cent completeness limit is reached about a magnitude below this, and few stars are found at $V_{\rm T} > 12.5$ mag.

typical star in this sample requires a metallicity correction to its temperature of <100 K, and much less than this in most cases.

3.2.5 *Comparison to literature data*

In order to better estimate the combined uncertainties inherent in our temperatures, we compare to published literature measurements. One of the most accurate sets of stellar temperatures comes from the exoplanet community: radial velocity confirmations of exoplanets require high signal-to-noise spectra, and measurements of exoplanet properties require accurate stellar classification. To construct a sample of exoplanet host parameters, we used the Exoplanet Orbit Database (EOD; Wright et al. 2011),[11] which was used in Chandler, McDonald & Kane (2016) to validate temperatures derived from the *Hipparcos* sample of stars. From a selection of 5454 catalogued exoplanets, coordinates and $T_{\rm eff}$ were returned for 2616 unique hosts. Of these, 591 could be matched with stars in the Tycho–*Gaia* catalogue. Of those, 150 have measurable parallaxes and are present in our final catalogue.

Among the 150 measured stars, the EOD quotes a literature stellar mass of 1.06 ± 0.43 $M_\odot$ (st. dev.) and a metallicity of [Fe/H] = 0.05 ± 0.24 dex (st. dev.). The average spectroscopic temperature was quoted as 5960 K. These parameters provide a good match to typical stars in our sample.

A comparison of the photometric and spectroscopic temperatures of these 150 stars is shown in Fig. 5 . The average photometric temperature is 73 ± 200 K (1.2 ± 3.4 per cent) lower than the spectroscopic temperature. For comparison, the median difference is slightly less, 52 K lower, and the 68th centile interval is −245 to 61 K, showing that the uncertainties are inflated by a number of poorly fitted outliers.

Warmer stars have their temperature underpredicted more frequently, and the scatter is greater towards underpredicted temperatures (1σ = 193 K) than overpredicted temperatures (1σ = 113 K). Scatter on the underpredicted side of the median will still be affected by interstellar reddening. However, the scatter on the overpredicted side of the median (113 K) should approximate the 1σ uncertainty in the results.

The same comparison was performed against the *Hipparcos* data set, where 359 stars could be matched against stars present in our final catalogue. Among those stars, the average stellar mass (with standard deviation) is 1.19 ± 0.37 $M_\odot$, the average metallicity is [Fe/H] = 0.09 ± 0.28 dex and the average spectroscopic temperature is 5396 ± 658 K. The *Hipparcos* exoplanet hosts are typically cooler, yet very slightly more massive, due to the larger fraction of evolved stars. They lie at a much closer average distance ($\langle d \rangle$ = 66 pc, cf. $\langle d \rangle$ = 270 pc for the Tycho-2 hosts). The average photometric temperature is 64 ± 163 K (1.2 ± 3.1 per cent) lower than the spectroscopic temperature. The median difference is marginally greater, at 69 K lower; however, the 68th centile interval is considerably smaller, at −153 to 37 K, providing a scatter of $^{+106}_{-84}$ K.

The magnitude of the systematic offsets and scatter for both data sets are typical: other studies have made previous comparisons of these methods on small fields, over which interstellar reddening is both known and constant (McDonald, Johnson & Zijlstra 2011b; Johnson et al. 2015; Chandler, McDonald & Kane 2016). Based on these studies, the global systematic offset of ∼50–70 K probably represents an artificial difference in modelling approach, either in the fine detail of the model atmospheres used, few-per-cent differences in the zero-points and colour terms in the underlying photometric catalogues, or the effects of atmospheres that are out of local thermodynamic equilibrium (see, e.g., discussions in Lapenna et al. 2014; Johnson et al. 2015). Meanwhile, the scatter of ∼100 K likely contains contributions from the uncertainty in the spectroscopic temperature (∼50 K), errors from the assumed stellar metallicity (∼30 K; Fig. 4), remaining scatter from the interstellar reddening (∼10 K, based on the difference between the median Tycho-2 and *Hipparcos* temperature offsets) and errors from the assumed stellar gravity (∼50 K). The remainder (∼60 K for the *Hipparcos* stars and ∼80 for the Tycho-2 stars, if added in quadrature) probably comes from random uncertainties in the input photometry. We stress, however, that these estimated uncertainties are meant for indicative purposes only. They are not derived from an unbiased, random sample of the data, and should not be applied directly to any single star without great care. Our final adopted uncertainties (Section 3.2.6, below) are slightly inflated from these values to be conservative, regarding these values as a lower limit.

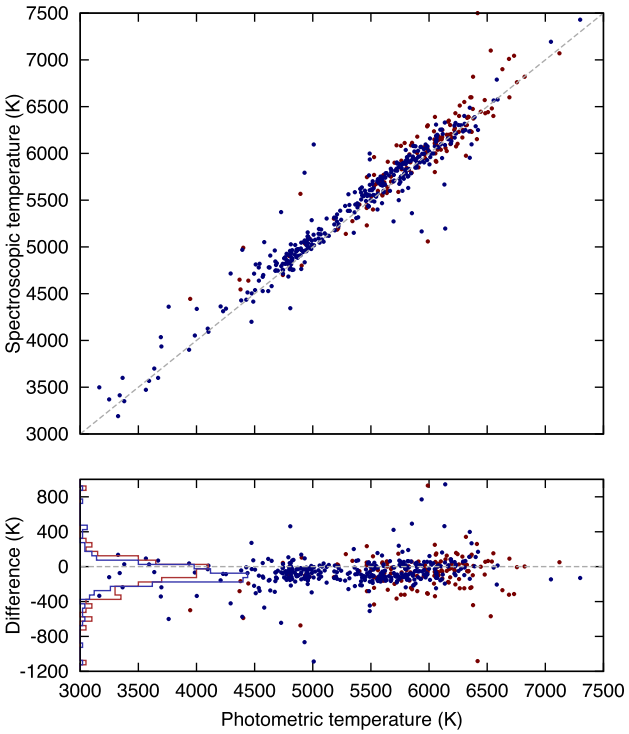

Figure 5. Comparison between literature temperatures of exoplanet hosts (mostly derived from spectroscopy) compared with temperatures derived from photometry in this work. The top panel plots both temperatures, while the bottom panel shows the difference between them, with a histogram shown on the left. Blue points represent *Hipparcos* stars, while red points represent Tycho-2 stars.

3.2.6 *Adopted uncertainty on the derived temperature*

To construct an error estimate that takes into account both the systematic offset and random scatter in Fig. 5, we adopt the 68th centile of the distribution of *absolute* deviations, as a measure that best reflects the uncertainty assigned to a typical star. For the Tycho-2 stars, this is $\sigma_{\rm T}$ = 137 K. For the *Hipparcos* stars, $\sigma_{\rm T}$ = 125 K. These uncertainties should be appropriate for a star with typical fit

[11] http://exoplanets.org

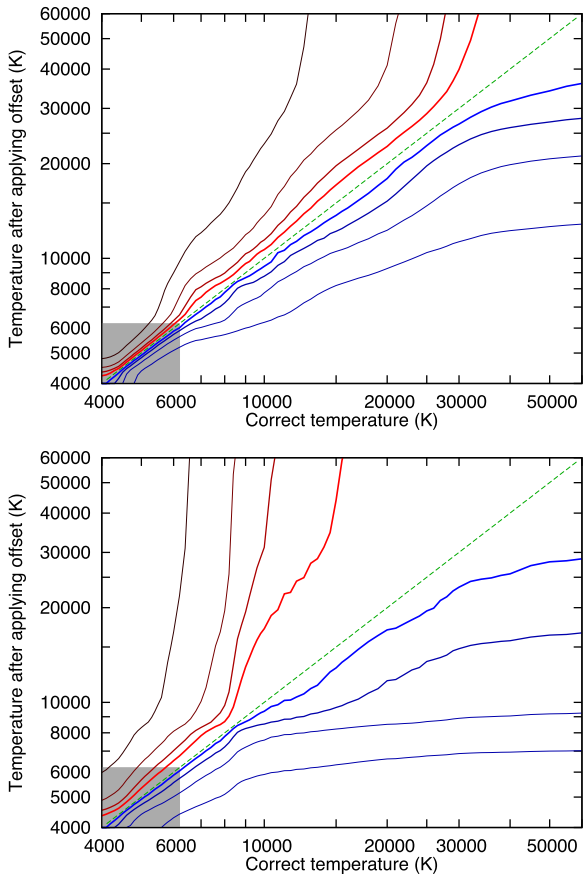

Figure 6. Temperature error arising from photometric inaccuracies in short-wavelength bands. Top panel: for data with U-band photometry. Red lines (above the dashed green line) show the effect of underestimating the $U-B$, $U-B_{\rm T}$, $B-V$ or $B_{\rm T}-V_{\rm T}$ colour (whichever is the most constraining) by 0.05, 0.1, 0.2 and 0.4 mag. Blue lines (below the dashed green line) show the effects of overestimating the colour by the same amount. Bottom panel: for data without U-band photometry, showing the effects of only the $B-V$ and $B_{\rm T}-V_{\rm T}$ colours. The grey box denotes temperatures below 6200 K, where redder colours provide better constraint on the temperature than those included here.

uncertainties ($Q=0.052$ and 0.042, respectively), provided the temperature is below ∼6200 K. The uncertainty should scale roughly with Q.

In hotter stars, there are no points sufficiently far down the Wien tail of the SED to accurately confine the stellar temperature. This limit is reached at ∼6200 K for photometry limited by the Johnson B, Tycho $B_{\rm T}$, or especially Sloan g' filters (depending on the stellar gravity and metallicity). However, for some *Hipparcos* stars, photometry extends to the Sloan u' or Johnson U filters. The magnitude of the Balmer jump, covered by these filters, can provide accurate temperatures up to a little over 10 000 K.

Absolute flux calibration of the shortest wavelength bands is particularly important here. Fig. 6 shows how the derived temperature departs from the mean for hot stars with and without U-band photometry, for a range of different photometric errors. For example, a 0.1 mag uncertainty in the $u'-B_{\rm T}$ colour of a 10 000 K star can result in a temperature uncertainty of the order of ±600 K, as will a 0.05 mag uncertainty in the $U-B$ colour. Equivalent uncertainties on a 12 000 K star result in a range in temperatures from 11 000 to 19 000 K, meaning stars with temperatures above 10 000 K cannot be accurately placed on the H–R diagram via the SED method without UV photometry. In such cases, correctly accounting for interstellar extinction becomes extremely important (see Fig. 3).

We assign an uncertainty on the derived temperature for *Hipparcos* with U-band or u'-band photometry, given by the largest out of the following options:

(i) $\delta T = 125$ K;
(ii) $\delta T = 125(Q/0.051)$ K;
(iii) $\delta T = \Delta_Q$ K, as described below, if $T > 6250$ K (see note below);
(iv) $\delta T = \Delta_R$ K, as described below, if $T > 6250$ K (see note below).

The first option denotes a minimum standard error. The second option accounts for badly fitted stars: roughly 68 per cent of stars have $Q < 0.051$; thus, we can expect this to be the approximate threshold above which stars exceed the typical 125 K error calculated in the previous section.[12]

The third option accounts for hot stars. Here, Δ_Q is the difference between the 'correct' and 'offset' temperatures in the top panel of Fig. 6 for an offset of $\sqrt{2}Q$. For stars with $6250 < T < 10\,500$ K, this effect is brought in gradually, such that

$$\delta T = \Delta_Q \frac{T-6250}{10\,500-6250}\,\mathrm{K}. \quad (2)$$

This accounts for the fact that some constraint is still applied by the longer wavelength filters below 10 500 K.

The fourth option accounts for hot stars that are otherwise well fitted, but where the short-wavelength photometry is poorly fitted. It substitutes the offset of $\sqrt{2}Q$ for an offset of $R_{\rm U}$ or $R_{\rm u'}$ as appropriate. These options also account (to first order) for temperature uncertainties caused by circumstellar or interstellar reddening for both hot and cool stars. For *Hipparcos* stars without u'-band or U-band photometry, we use the lower panel of Fig. 6 for the third option, and $R_{\rm B}$ or $R_{\rm BT}$ for the fourth option. As with Δ_Q, Δ_R is brought in gradually between 6250 and 10 500 K for stars without u'-band or U-band photometry, and 'instantaneously' at 10 500 K for those with either of these bands observed.

Similarly, we assign an uncertainty for Tycho-2 stars as the largest out of the following options:

(i) $\delta T = 137$ K;
(ii) $\delta T = 137(Q/0.060)$ K;
(iii) $\delta T = \Delta_Q$ K, as described below, if $T > 6250$ K;
(iv) $\delta T = \Delta_R$ K, as described below, if $T > 6250$ K.

Since the Tycho-2 sample lacks reliably matched U-band or u'-band photometry, the lower panel of Fig. 6 is always used for the third option, and $R_{\rm B}$ or $R_{\rm BT}$ is always used for the fourth option.

For both Δ_Q and Δ_R, we round up to the nearest 0.01 mag in Q and R, and round up to the temperature grid point above the derived temperature (this is almost universally more uncertain than the grid point below). This provides a fairly conservative estimate of the random uncertainty applied by both the photometry and fitting procedure to the temperature assigned to the star. It does not fully include uncertainties due to interstellar or circumstellar reddening, which are detailed in Section 3.2.3. We stress that none of these uncertainties is a formal uncertainty measure, but instead simply an estimate of the 1σ uncertainty that can be assigned to the stellar

[12] For comparison, the 68th centile for the planet hosts is comparable, at $Q = 0.053$.

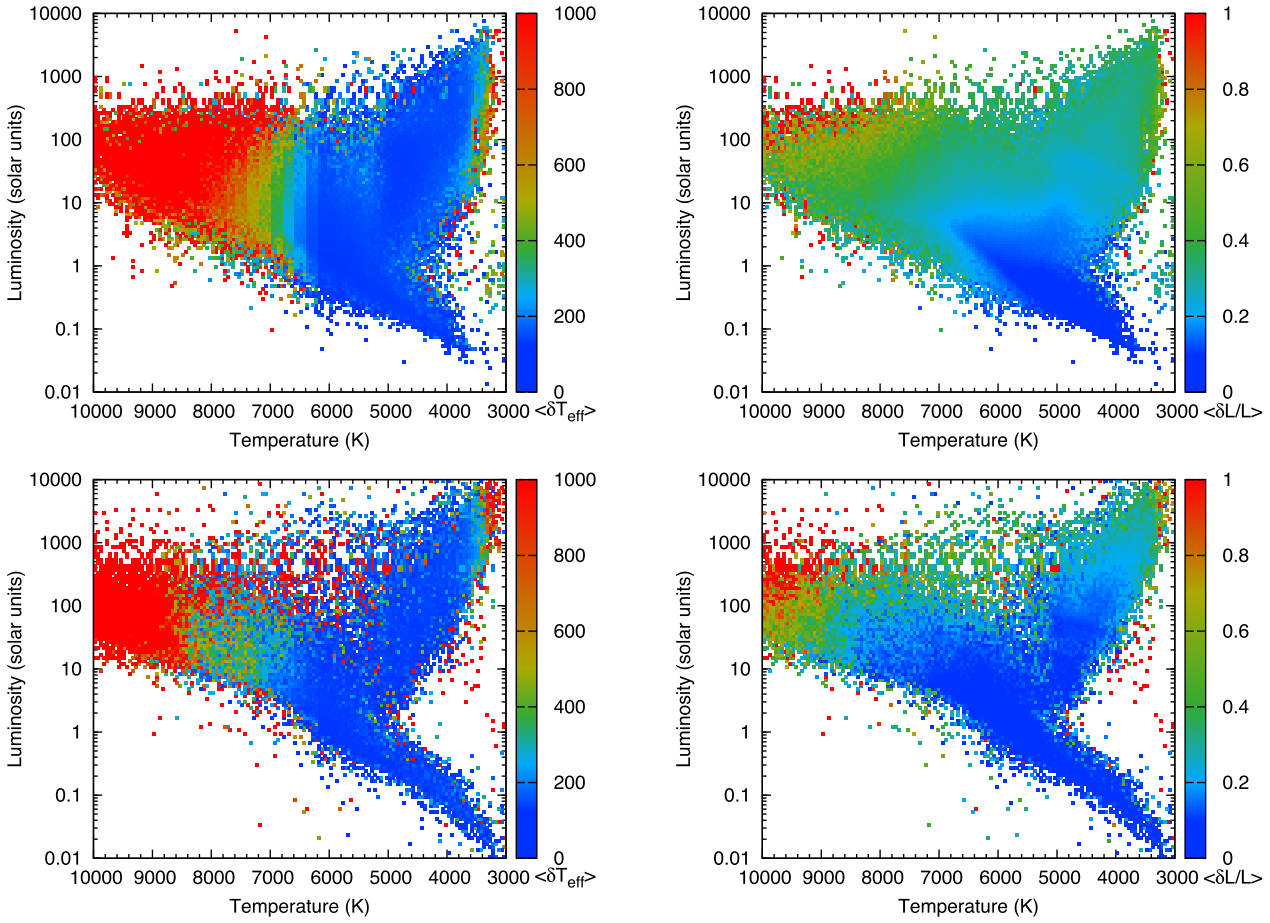

Figure 7. Uncertainties in temperature (left, absolute error) and luminosity (right, fractional error) averaged across the binned H–R diagram. The Tycho-2 sample is shown in the top panels, and the *Hipparcos* sample in the bottom panels.

temperature. These uncertainties are listed in Tables 1 and 2 and mapped on to the H–R diagram in Fig. 7 .

3.2.7 *Adopted uncertainty on the derived luminosity*

The contribution of photometric uncertainty to the uncertainty in derived luminosity is discussed with case studies in McDonald et al. (2011a). Photometric uncertainty affects temperature and luminosity in different ways, depending on the wavelength in question. Overprediction of flux at wavelengths bluer than the SED peak leads to overprediction in effective temperature and overprediction in luminosity, while overprediction of flux at redder wavelengths leads to *under*prediction of the effective temperature and *under*prediction of the luminosity. The greatest luminosity change that can normally be effected is $\delta L/L = 4\delta T/T$, since (for a blackbody) $L \propto T^4$. The combination of the above effects means that the power law is shallower than this, but not normally by much. Therefore, $\delta L/L = 4\delta T/T$ represents a fairly good estimate, yet also a conservative one. For example, an underprediction of temperature of 137 K on a 4500 K star leads to an overestimation of its luminosity by $\delta L/L =$ 12.1 per cent.

The uncertainty in luminosity has a reasonably strong correlation with the uncertainty in temperature, but that correlation and its direction depend on the photometric data causing the uncertainty. Optical data that are overly bright will lead to overestimated temperature and luminosity; overestimated infrared data will lead to *under*estimated temperature but still overestimated luminosity. Photometric uncertainties are usually fractionally larger at longer wavelength (due to the thermal or astrophysical background, or sensitivity issues). Hence, there is more usually an anti-correlation between the photometric and luminosity uncertainties.

For hot stars, uncertainties in luminosity correlate with uncertainties in temperature, scaling as[13] $\delta L/L = 3\delta T/T$. The aforementioned ∼600 K uncertainty in the temperature of a 10 000 K star results in a 24 per cent uncertainty in luminosity.

In most cases, the photometric contribution to the luminosity uncertainty is exceeded by the distance uncertainty to the star. The average parallax uncertainty on our Tycho–*Gaia* sample is σ_ϖ/ϖ = 16.4 per cent, leading to an uncertainty in luminosity of $\sigma_L/L =$ 32.8 per cent. For the *Hipparcos*/*Hipparcos–Gaia* sample, they are $\sigma_\varpi/\varpi = 7.6$ and $\sigma_L/L = 15.1$ per cent, respectively.

Our final luminosity uncertainty (see also Fig. 7) is given as

$$\delta L/L = \sqrt{(n\delta T/T)^2 + (\delta\varpi/\varpi)^2}, \quad (3)$$

[13] In hot stars, the uncertainty is driven by the short-wavelength filters: the flux of the Rayleigh–Jeans tail is observationally well constrained. However, the flux at a wavelength on a blackbody's Rayleigh–Jeans tail varies linearly with temperature. If poor-quality optical photometry leads to an overestimation in optical flux, the derived temperature increases. Accordingly, the derived surface area then decreases as $R \propto T^{-2}$. Thus, by $L \propto R^2T^4$, the luminosity relation is to the third power, rather than the fourth.

where $n = 4$ if $T < 6200$ K, $n = 3$ if $T > 10\,500$ K and $n = 4 - (10\,500 - T)/(10\,500 - 6200)$ in between. These uncertainties are listed in Tables 1 and 2. We again stress that these are not formal uncertainties.

3.3 'Sanity checking' of local population and interstellar extinction

3.3.1 Galactic thick- and thin-disc populations

Fig. 8 shows the H–R diagram for stars at a fixed range of distances (300–400 pc) at differing Galactic latitudes.[14] The solar-metallicity thin-disc population dominates at these scaleheights. Stars are recovered down to the main-sequence turn-off in all cases, and extinction does not yet severely affect star counts in the Galactic plane (however, see discussion on the Gould Belt below). Without performing a detailed population model, it is still clear that completeness declines markedly below $\sim$3 $L_\odot$ at all latitudes.

At high latitudes, few stars at ages <3 Gyr are seen. The red clump appears both young and luminous if at solar metallicity.[15] Martig et al. (2016) determined a median age of $\sim$5 Gyr for red clump stars at scaleheights of $\sim$300 pc. Even at high latitudes, we expect approximately solar abundances, as solar metallicity was reached by the time star formation ceased in the Galactic thick disc, $\sim$10 Gyr ago (Bensby, Feltzing & Lundström 2004). A significant component from the thick disc is not expected until scaleheights of >500 pc (e.g. Gilmore & Reid 1983; Kong & Zhu 2008). Along with our completeness limitations, this combination of factors explains the lack of stars lying below the solar-metallicity main sequence. However, the luminosity of the RGB bump is also strongly metallicity dependent (cf. Boyer et al. 2009; McDonald et al. 2011a), so including an old, metal-poor population that reduces the average abundance to slightly sub-solar metallicities ($\sim$0.2 dex) allows the RGB bump to be fitted reasonably well.

Fig. 9 shows the H–R diagram for high-latitude stars between 600 and 800 pc from the Sun (520–800 pc from the plane). Sensitivity declines rapidly below $\sim$6 $L_\odot$, limiting inclusion to main-sequence turn-off stars $\lesssim$5 Gyr in age. Few stars are younger than $\sim$3 Gyr, or hotter than >6500 K. A significant shift in the temperature of the giant branch and red clump indicates stars are metal poor: a crude estimate places them at [Fe/H] ~ -0.5 dex, as expected from chemical studies (e.g. Masseron & Gilmore 2015).

3.3.2 The Galactic plane and Gould Belt

The Gould Belt is an elliptical structure of young stars and star forming regions, with major and minor axes roughly 400 × 300 pc. It is centred approximately on the α Per moving group, but presents on the terrestrial sky with a roughly constant 20° inclination with respect to the Galactic plane. The Sun lies close to its inner edge, as traced by the Scorpius–Centaurus OB association (e.g. Herschel 1847; Olano 1982, 2001; de Zeeuw et al. 1999; Ward-Thompson et al. 2007). *Gaia* DR1 records distances to individual stars with sufficient accuracy that membership of associations can be made within a few hundred pc of the Sun, covering roughly the nearer half of the Gould Belt. This region is presented in Fig. 10 and mapped on to the sky in Fig. 11 . In the further half of the Gould Belt, parallax uncertainties become large and smearing of associations in the radial direction and the associated Lutz–Kelker effects restrict detailed analysis of this region.

The majority of structures in the western part of the Gould Belt ($150° < l < 360°$) are located within 300 pc, and the majority of the structures in the eastern part ($60° < l < 150°$) are between

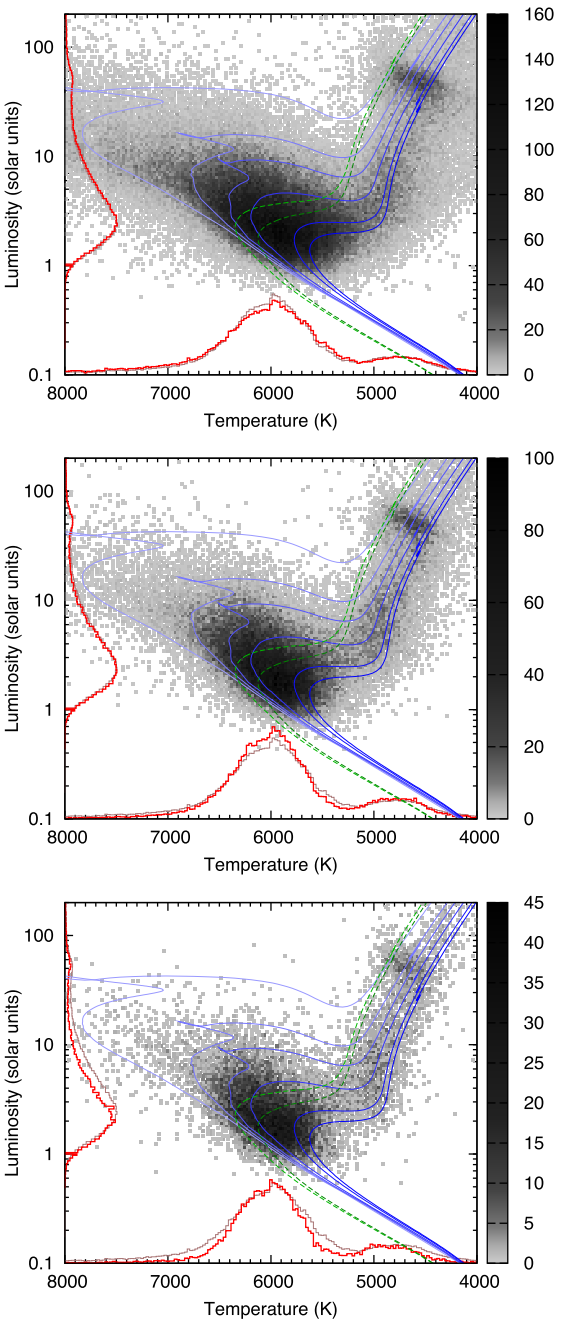

Figure 8. Density-coded (Hess) H–R diagram of stars between 300 and 400 pc from the Sun. The panels show (top to bottom) Galactic latitudes ±0°–30°, 30°–60° and 60°–90°, representing distances 0–200, 150–350 and 260–400 pc from the Galactic plane. Thick red lines show histograms of sources in that plot, compared to the lighter lines of sources at all latitudes. Overlaid on the H–R diagrams are isochrones from Marigo et al. (2008), showing (in blue, top to bottom) isochrones for solar-composition stars at 1, 2, 3, 5, 10 and 13 Gyr. The dashed, green lines show 10 and 13 Gyr isochrones at [Fe/H] = −1 dex and [α/Fe] = +0.2 dex.

[14] A mild Lutz–Kelker bias exists at these distances, which is latitude dependent due to the changing density of objects.

[15] The metallicity correction in this region is typically <100 K dex^{-1} in metallicity (Fig. 4).

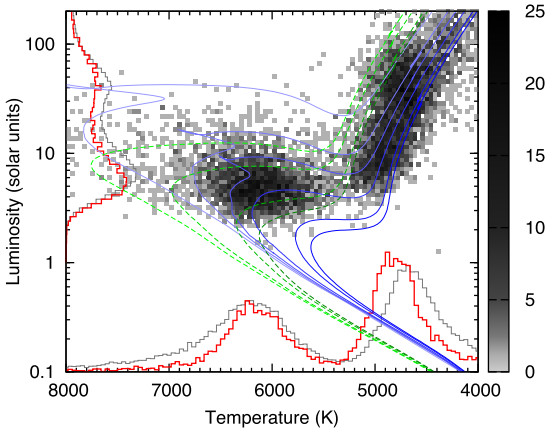

Figure 9. As the bottom panel of Fig. 8, but for stars between 600 and 800 pc from the Sun at Galactic latitudes ±60°–90°. The thinner, grey histogram shows stars in the range ±10°–50°. Metal-poor isochrones are shown (green, dashed lines) for 3, 5, 10 and 13 Gyr, as well as the solar-metallicity isochrones from Fig. 8. Note the warmer giant branch.

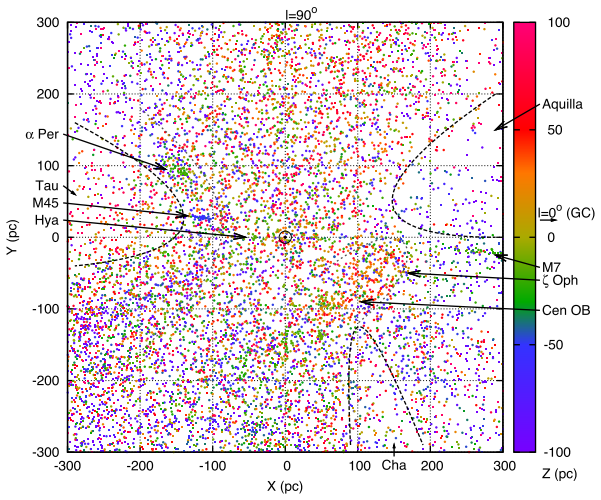

Figure 10. Hot stars (>8000 K) within 100 pc of the Galactic plane, colour-coded by height above/below the plane. Major known features are identified. The Galactic Centre is to the right, and Galactic longitude (l) increases anti-clockwise.

300 and 600 pc, as in the studies cited above. However, at high resolution, the belt breaks up into the more discontinuous features of Fig. 10. Fig. 11 also shows the regions affected by large interstellar dust clouds. The three primary offenders (Aquila, Taurus and Chameleon) are shown in Fig. 10. Stars in these regions suffer several magnitudes of visual extinction, so are either reddened sufficiently that they no longer appear to be above 8000 K (cf. Fig. 3) or were otherwise rendered entirely invisible to the *Hipparcos* and Tycho instruments. The presence of the Gould Belt is also traced by the distribution of stars with infrared excess in Fig. 13, indicating the large number of young stars (pre-main-sequence and Herbig Ae/Be stars) in this region.

4 INFRARED EXCESS

4.1 Criteria for defining infrared excess

A definition of infrared excess must take into account all the above factors. We start with two assumptions.

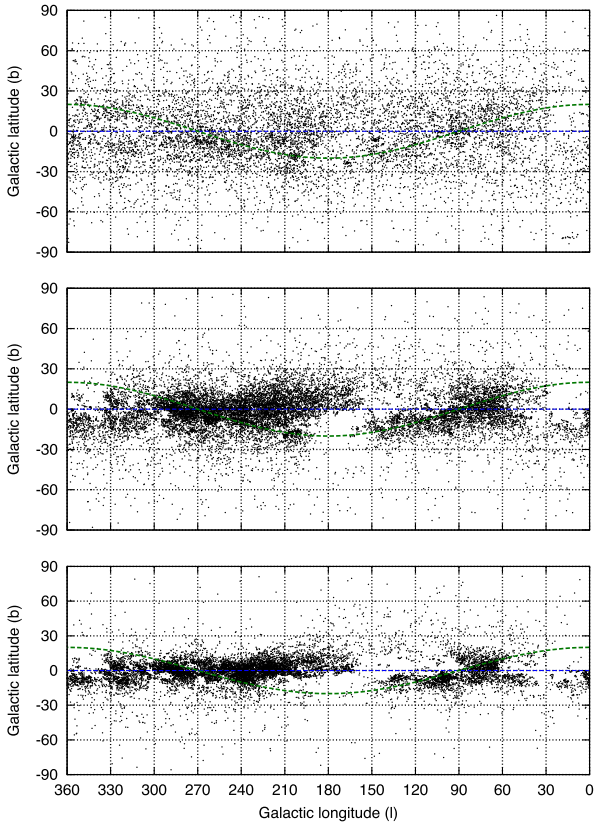

Figure 11. Hot stars (>8000 K) in the solar neighbourhood, showing the Gould Belt (dashed line). From top to bottom, the panels represent stars in the ranges 0–300, 300–600 and 600–900 pc from the Sun. Note that the bottom plot in particular will suffer from a strong Lutz–Kelker bias from stars at greater distances. Note the absence of stars along various Galactic plane sightlines, indicating strong interstellar extinction.

(i) The region <4.3 μm defines the stellar continuum. This region should be relatively free from circumstellar emission.

(ii) The region ≥4.3 μm defines the regime in which infrared excess occurs.

The factors behind these assumptions are detailed in Appendix E (online only).

To help quantify infrared excess, we define the following statistics, using the individual observed/modelled flux ratios (F_o/F_m) and the overall quality of fit (Q) described in Section 2.1.2.

(i) $\Re_{opt}$ defines the average value of F_o/F_m over the optical filters (*UBVR*, *ugr*).

(ii) Similarly, $\Re_{NIR}$ defines the average of F_o/F_m over the near-infrared (near-IR) filters (*IJHK*$_s$*L*, *iz* and *WISE* [3.4]).

(iii) Also, $\Re_{MIR}$ defines the average of F_o/F_m over the mid-IR filters (longwards of *L* and [3.4]).

(iv) N_{opt}, N_{NIR} and N_{MIR} denote the number of optical, near-IR and mid-IR data points, respectively, which contribute to the above.

(v) The combined $\Re_{opt+NIR}$ and $N_{opt+NIR}$ represent the same quantities as $\Re_{opt}$ and N_{opt}, but computed over the full *U* through [3.4] range.

(vi) $\Re'_{MIR}$ provides an alternative version of $\Re_{MIR}$, removing the point with the maximum *R* from the mid-IR data.

(vii) X_{MIR} provides a statistic of overall mid-IR excess, calculated as

$$X_{MIR} = \Re_{MIR}/\Re_{opt+NIR}. \quad (4)$$

This statistic should be most sensitive to faint mid-IR excess if the host star is unreddened. If it is substantially reddened, or contains a single bad mid-IR data point, then

$$X'_{\rm MIR} = \Re'_{\rm MIR}/\Re_{\rm NIR} \quad (5)$$

should provide a more accurate value. Robustness of the detection is therefore increased where both $X_{\rm MIR}$ and $X'_{\rm MIR}$ are significantly above unity.

(viii) $S_{\rm MIR}$ provides a statistic of the significance of mid-IR excess, calculated as

$$S_{\rm MIR} = (\Re_{\rm MIR} - 1)\sqrt{N_{\rm MIR}}/Q. \quad (6)$$

This approximates the signal-to-noise statistic of the infrared excess. Note that this will generally be an overestimate for stars with little excess: scatter due to photometric errors will typically be much greater in the infrared than the optical and near-IR, meaning that the fit quality parameter, Q, will be an underestimate for the 'noise' component in this equation. For stars with significant excess, this will generally be an underestimate, as the infrared excess artificially inflates the Q parameter. We also note that this significance statistics does not exclude objects such as stars heavily reddened by interstellar extinction. This statistic is therefore presented for guidance only and should be used in combination with the others in this section to define whether a source has a significant excess.

(ix) To determine the amount of light emitted in the infrared excess, we construct a trapezoid integral, interpolated in the (log F_ν)–(log λ) plane. This (respectively) provides the total luminosity and fraction of the stellar flux re-emitted into the infrared:

$$L_{\rm XS} = \int_{\nu=0}^{7\times10^{13}\,{\rm Hz}} (F_\nu - F_*)\,{\rm d}\nu \quad (7)$$

and

$$f_{\rm XS} = \frac{\int_{\nu=0}^{7\times10^{13}\,{\rm Hz}} (F_\nu - F_*)\,{\rm d}\nu}{\int_{\nu=0}^{\infty} F_*\,{\rm d}\nu}, \quad (8)$$

where we assume that the infrared excess beyond 1 mm is zero,[16] and that the stellar flux ($F_\star$) is the modelled flux ($F_{\rm m}$) multiplied by $\Re_{\rm NIR}$. The cut-off of 7×10^{13} Hz corresponds to 4.3 μm. This is a lower limit to the fraction of reprocessed light, since the SED fitting partially takes into account the optical absorption and infrared emission from this reprocessing.

(x) Finally, we use these data to extract the wavelength at which the peak flux (F_ν) of the infrared excess occurs, $\lambda_{\rm peak,\,XS}$, which is defined by the point at which ($F_\nu - F_\star$) reaches a maximum.

4.2 An H–R diagram of infrared excess

Fig. 12 shows the H–R diagram of *Hipparcos* and Tycho-2 stars, colour-coded by infrared excess, while Fig. 13 shows the distribution of sources across the sky. Sources are only included in these figures if $N_{\rm opt} + N_{\rm NIR} > 0$ (i.e. they have optical and infrared data), $N_{\rm MIR} > 1$ (i.e. they have more than one mid-IR data point) and if the parallax uncertainty $\delta\varpi/\varpi < 0.2$. Fig. 12 is also limited by $A_{\rm V} < 1.5$ mag.

The majority of these 600 667 stars are well fitted. The standard deviation of $X_{\rm MIR}$ is 0.185; however, this is dominated by a small number of stars with large infrared excesses. If we take the central 68 per cent around the median of Med($X_{\rm MIR}$) = 1.024, the scatter is reduced to $\sigma_X = ^{+0.025}_{-0.027}$.

As a general trend, stars near the main sequence and lower giant branches tend to be well fitted. Deviations become more apparent as one moves off these two sequences. Particularly noticeable are infrared deficits ($X_{\rm MIR} < 1$) among hot ($\gtrsim$ 8000 K), luminous (>30 $L_\odot$) stars and cool (∼3500–4500 K), luminous (∼100–3000 $L_\odot$) stars.

Among hot stars, this deficit may be due to interstellar reddening. The opacity of interstellar dust has a steeper law than a blackbody's Wien tail in the optical, but a shallower law in the infrared. Reddened hot stars are modelled as cooler stars but, because of this opacity law, tend to be underluminous in the optical and mid-IR, and overluminous in the near-IR.

Reddened cool stars exhibit different qualities. Molecular opacity in the cool-star models has a strong temperature dependence. The opacity is mostly caused by TiO, and has a steeper wavelength dependence [$F \underset{\sim}{\propto} \lambda^6$ over $(U-R)$] than interstellar extinction ($F \underset{\sim}{\propto} \lambda^4$). Consequently, stars that are reddened by interstellar extinction and are fitted by cooler stellar models tend to have a less sharp peak to their SEDs compared to stars intrinsically at that temperature; hence, they tend to be overluminous in the optical and mid-IR, and underluminous in the near-IR, when compared to said models. This causes reddened giant branch stars to congregate around 3600–3700 K and exhibit mid-IR excess (cf. the artefact at this temperature identified in Fig. 1).

Instead, the mid-IR deficit among the giant stars seems to result from a combination of difficulties in accurately modelling the TiO absorption bands in the optical in cool stars, as well as underestimation of flux in the H band due to inaccurate modelling of the H^- opacity peak (see Appendix A, Fig. A12, online only).

4.3 Characteristics of infrared excess across the sky

Small-scale variations of $X_{\rm MIR}$ can be seen across the sky (Fig. 13). Generally speaking, the regions of greatest deficit can be seen towards the Galactic bulge and near the North Galactic Pole (NGP). Towards the bulge, crowding means that only optically brighter (typically hotter) stars are present in the *Hipparcos*/Tycho-2 and *Gaia* observations, which are then reddened. Towards the NGP, a large proportion of stars are old, cool stars. The previous section describes why these stars should be apparently underluminous in the infrared.

Regions of moderate extinction, however, generally show a slight excess overall. This is most notable around the Musca interstellar clouds ($\alpha = 180°$, $\delta = -80°$), the ρ Oph star-forming region ($\alpha = 250°$, $\delta = -20°$) and the Orion star-forming region ($\alpha = 90°$, $\delta = 0°$). Since these are regions of diffuse emission in the mid-IR, it is possible that background light affects some of the observations here at the level of a few per cent. This background light may be from dust heated by the star in question (as seen in the Pleiades) or by other sources in the line of sight.

Stars with substantial infrared excess ($X_{\rm MIR} > 1.15$) also tend to occupy these regions, but are also more widely spread along the Galactic plane.

[16] Dust optical depth typically drops at longer wavelengths, as the emissivity of dust typically has a spectral slope steeper than a blackbody's (e.g. Schöier, Lindqvist & Olofsson 2005). For many objects, other emission mechanisms become important in the sub-millimetre and beyond (e.g. Reid & Menten 1997).

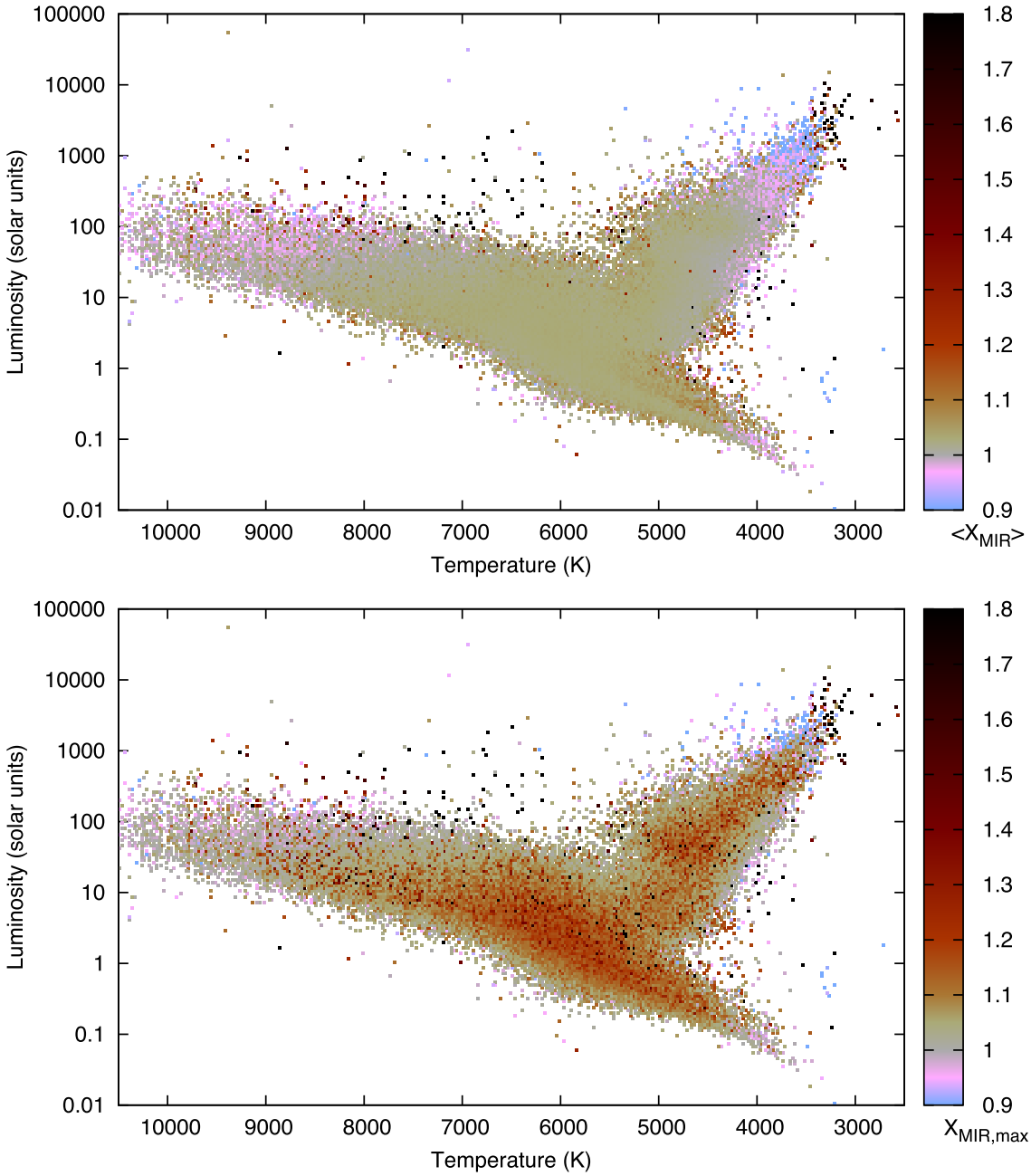

Figure 12. A binned H–R diagram, coloured to show the average mid-infrared excess ($X_{\rm MIR}$) in each bin. Stars are included if $N_{\rm opt} + N_{\rm NIR} > 0$, $N_{\rm MIR} > 1$, $\delta\varpi/\varpi < 0.2$ and $A_{\rm V} < 1.5$ mag. The top panel shows that average excess ($X_{\rm MIR}$) in each bin, with unity being no excess. The bottom panel shows the highest value of $X_{\rm MIR}$ in each bin, to show the most extreme sources.

4.4 Defining criteria to flag infrared excess

We define an infrared excess by two criteria. The first relates to the scatter calculated in Section 4.2. With 600 667 stars, if our distribution of $X_{\rm MIR}$ was Gaussian in nature, we could expect a 5σ threshold to remove random fluctuations in the data; hence, sources with $X_{\rm MIR} > {\rm Med}(X_{\rm MIR}) + 5\sigma_X = 1.15$ should be considered strong candidates for infrared excess. In practice, our distribution has a supra-Gaussian tail of badly fitting points on either side of the distribution; hence, such a cut-off only removes the majority of badly fitting points.

The fraction of stars with $X_{\rm MIR} > 1.15$ is marginally larger towards lines of sight with higher extinction (Fig. 14). Hence, we modify our criterion to remove stars with marginal infrared excess along high-extinction lines of sight. To qualify as a candidate for infrared excess, stars must have $X_{\rm MIR} > 1.15 + A_{\rm V}/100$. This criterion is shown as the dashed line in Fig. 14.

There are 1879 sources from the *Hipparcos* sample that meet these criteria (0.18 per cent), and 2377 sources from the Tycho-2 sample (0.016 per cent). The much lower fraction from the Tycho-2 catalogue is caused primarily by the comparatively poor quality of the infrared photometry available for the Tycho-2 stars, due to their faintness and (in high-extinction lines of sight) the consequent difficulty of extracting them from the diffuse infrared background. Secondary effects include the less certain parallax measurements for the Tycho-2 sample and the propensity for bright (*Hipparcos*)

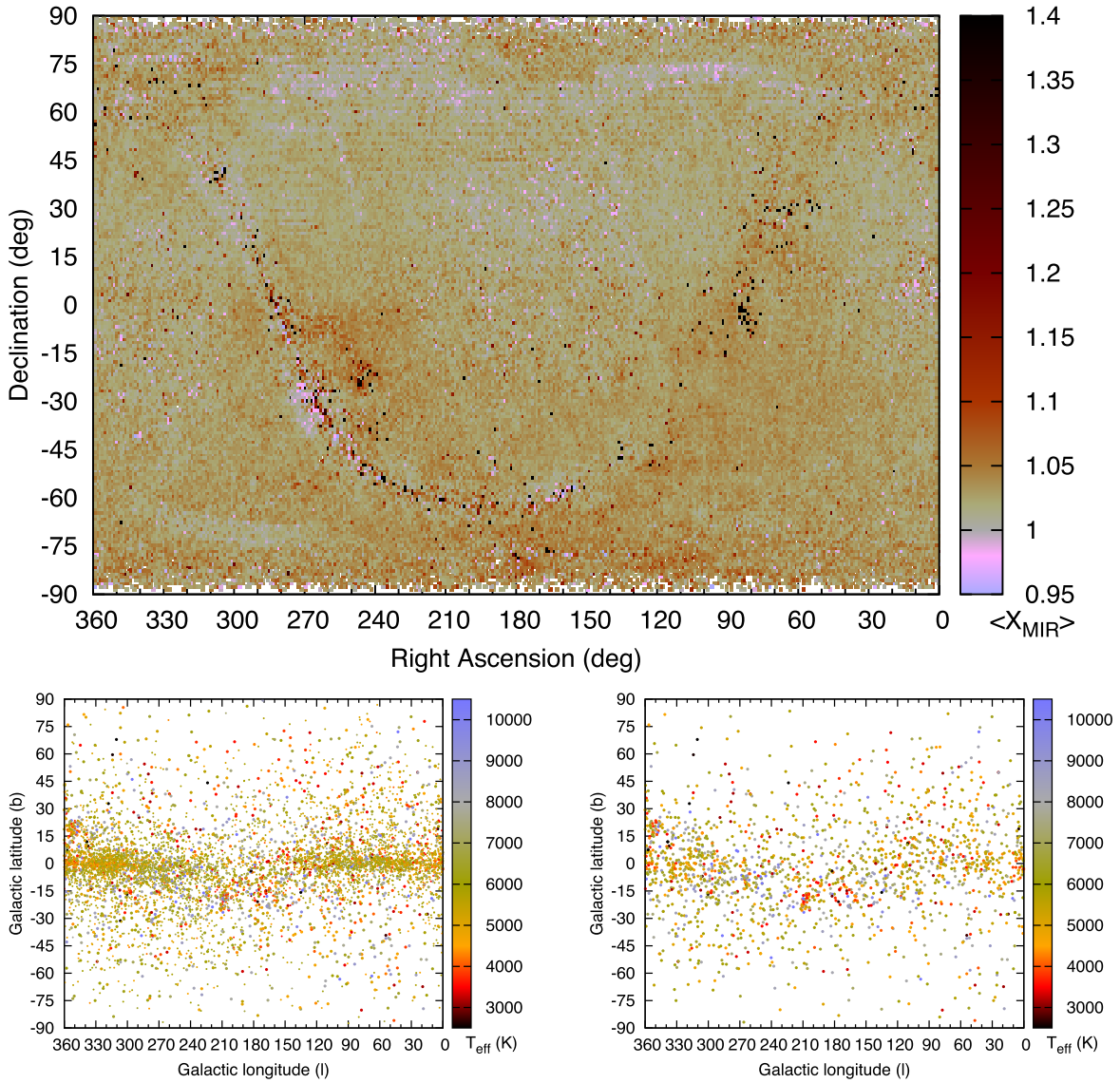

Figure 13. The spatial distribution of infrared excess for Tycho-2 and *Hipparcos* catalogue stars. The bottom-left panel shows individual stars that are candidates for having infrared excess, colour-coded by temperature. The bottom-right panel shows the same plot for stars that are strong candidates (a score of more than three points).

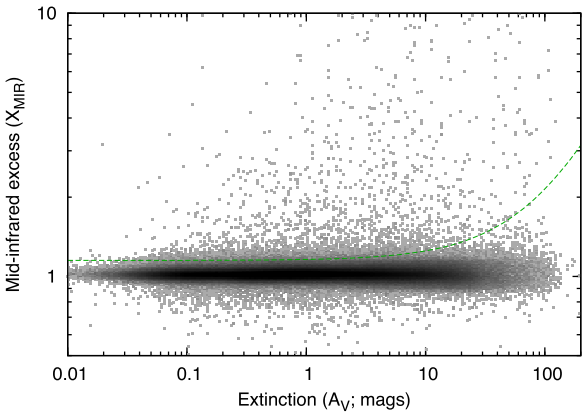

Figure 14. A Hess diagram showing the relationship between infrared excess ($X_{\rm MIR}$) and line-of-sight interstellar extinction ($A_{\rm V}$). The dashed line shows the cut-off used to determine candidacy for infrared excess.

stars to display infrared excess (e.g. Herbig Ae/Be stars, Cepheids, giant branch stars). Improvements in the resolution and depth of the available infrared data bases would substantially improve our ability to extract infrared excess.

We define these 4256 stars as having candidate infrared excess associated with them. We strongly advise users of these data to inspect the associated mid-IR imagery of each object, and cross-check the relevant values of Q, $S_{\rm MIR}$ and $X'_{\rm MIR}$, to help confirm or refute its presence.

4.5 A catalogue of stars with infrared excess

4.5.1 The catalogue and its contents

Table 3 catalogues the objects defined as having infrared excess. The SIMBAD spectral types are listed (Table 4), as well as `otype`

Table 3. Catalogue of stars with candidacy for hosting infrared excess. A portion of the online table is shown here, where table columns are numbered for clarity. The columns are described in full in the text, but can briefly be described as (1) Tycho-2 or *Hipparcos* identifier; (2)–(18) as Table 1; (19) mid-IR excess; (20) mid-IR excess, calculated with the point with the strongest excess removed; (21) uncalibrated significance of the excess; (22) SIMBAD primary name; (23) SIMBAD primary object type; (24) full list of SIMBAD object types; (25) SIMBAD spectral class; (26) point-based quality criterion. Complete tables are to be found at CDS.

(1) Name	··· ··· ···	(18) Q	(19) $X_{\rm MIR}$	(20) $X'_{\rm MIR}$	(21) $S_{\rm MIR}$	(22) SIMBAD Name	(23) SIMBAD `otype`	(24) SIMBAD `otypes`	(25) SIMBAD Class	(26) Quality (points)
HIP 66	···	0.107	1.211	1.278	3.245	HD 224790	*	*,IR	F2V	4
HIP 75	···	0.108	1.153	1.215	3.256	HD 224821	*	*,IR	K4III	3
HIP 122	···	1.324	1.909	2.461	1.257	* tet Oct	*	*,IR	K3III	5
···	···	···	···	···	···	···	···	···	···	
TYC 9529-1698-2	···	0.522	1.940	2.092	2.855	CPD-85 549B	*	**,*,IR	G5	6

Table 4. Summary of spectral types among stars with mid-IR excess. The first count column gives all candidate sources; the second column gives sources with >3 points.

Spectral type	Count		Notes
O	13	11	
B	565	475	Including two DB stars
A	549	355	Including five DA stars
F	382	186	
G	302	132	
K	410	133	
M	124	74	
C	15	11	
S	5	2	

parameters[17] (Table 5), providing a basic categorization of each source. There are 95 entries that did not receive a SIMBAD match. The location of different categories of source on the H–R diagram is shown in Fig. 15 .

The statistics in Tables 4 and 5 are not complete, and each list is not exhaustive. Of the 4161 sources with SIMBAD entries, 3049 have a primary classification of 'star'. Examination of individual records indicates that many of these are known objects of interest (e.g. emission-line stars, late-type giants, etc.) that have not yet been correctly designated as such by SIMBAD. Yet, many of these 3049 objects appear to be new candidates for hosting infrared excess.

The inhomogeneity of our input data quality means that the confidence on the detection of infrared excess varies. We therefore introduce a point-based quality criterion to judge the likelihood of excess being present. Points are awarded successively if $X_{\rm MIR} > 1.2 + A_{\rm V}/80$, $X_{\rm MIR} > 1.3 + A_{\rm V}/40$ and $X_{\rm MIR} > 1.5 + A_{\rm V}/3.1$; if $X_{\rm MIR} > Q + 1$; if $X'_{\rm MIR} > Q + 1$; or if $S_{\rm MIR} > 1$, giving a maximum possible six points. Examination of individual sources shows that, typically, more than three points are needed to show a high-quality detection of infrared excess: there are 1883 objects with more than three points, 1156 of which have either no SIMBAD classification, or a primary classification of 'star'.

4.5.2 *Types of object with infrared excess*

The statistics in Table 5 show that we detect a variety of stellar types that are expected to host infrared excess. These include Herbig Ae/Be stars, and a variety of young and pre-main-sequence stars, evolved (post-)AGB stars and stars experiencing third dredge-up (S-type stars and carbon stars; see, e.g., Karakas & Lattanzio 2014), and a variety of variable stars that are known to exhibit dust. Also included are a variety of binary stars. Some of these are expected to host circumstellar or circumbinary material, and some are not. In many cases, the infrared excess may simply arise from problems caused by fitting two superimposed stellar SEDs with a single stellar atmosphere model.

There are a variety of other types of object that are not a priori expected to host infrared excess. These are stars in clusters, nebulae and stellar associations. Several of these stars are in regions of known nebulosity, such as the Pleiades and various parts of the Orion star-forming complex. It also includes stars in nearby clusters, but clearly not associated with them, such as HIP 81894. Other causes of infrared excess in such objects may be attributable to stellar blending (e.g. McDonald et al. 2011a).

A number of objects are identified by SIMBAD as extragalactic, but are unlikely to be so. These include TYC 273-677-1 and TYC 705-746-1, where *Gaia* has measured parallaxes of 5.99 ± 0.95 and 2.42 ± 0.31 mas, respectively, and TYC 7415-696-1, which is the T Tauri object Hen 3-1722 (Wray 1966; Stock & Wing 1972; Henize 1976).

4.5.3 *Properties of infrared-excess stars on the H–R diagram*

Fig. 15 places various categories of infrared-excess stars in the H–R diagram. Stars with infrared excess at high confidence are typically found away from the main sequence and giant branch, mostly above the main sequence. Variable stars are found all over the H–R diagram, with no clear sign of the bounds of the instability strip. Likewise, binary stars are found in many locations, although they do not frequent the giant branch due to observational biases against their detection.

Stars associated with clusters or nebulosity scatter above the main sequence, suggesting that source confusion or incorporation of background light into the SED may have occurred. In some cases, these may also be young stars that have yet to descend to the main sequence.

Young stars in the cool end of the H–R diagram tend to lie at varying distances above the main sequence. The majority of the T Tauri stars and Herbig–Haro objects lie in the Hayashi forbidden zone (Hayashi 1961), commensurate with their young age.

By contrast, evolved stars are logically found predominantly near the top of the giant branch. However, a large number of 'evolved' stars are well down the giant branch ($<200\,L_{\odot}$), and there are even some on the main sequence. Such objects include the following.

[17] http://simbad.u-strasbg.fr/simbad/sim-display?data=otypes

Table 5. Summary of common SIMBAD object types among stars with mid-IR excess. Objects may appear more than once in the list. Only those types with ≥3 entries are shown. Purely observational characteristics (e.g. infrared source) are excluded.

Object type	Count	Notes
	Young stellar types and hot stars	
Be*	199	Herbig Be star
Y*O	38	Young stellar object (YSO)
TT*	37	T Tauri star
Ae*	30	Herbig Ae star
Ae?	30	Candidate Ae star
pr*	28	Pre-main-sequence star
Y*?	8	Candidate YSO
HH	5	Herbig–Haro object
bC*	4	β Cepheid variable
	Evolved stellar types	
C*	19	Carbon star
Mi*	8	Mira variable
S*	6	S-type star
AB*	7	AGB star
WD*	5	White dwarf
pA?	7	(Candidate) post-AGB star
	Variable star types	
V*	437	Variable star
LP*	56	Long-period variable (LPV)
Ro*	15	Rotational variable stars
Or*	25	'Orion-type' variable stars
dS*	12	δ Scu star
Pu*	10	Pulsating variable
a2*	9	Rotational (α_2 CVn) variable
LP?	9	Candidate LPV
Ir*	7	Irregular variable
No*	6	Nova
BY*	5	Rotational (BY Dra) variable
V*?	5	Candidate variable
El*	4	Ellipsoidal variable
Ce*	3	Cepheid variable
RI*	3	Rapid, irregular variable
NL*	3	Nova-like star
Fl*	3	Flare star
	Binary star types	
**	425	Binary star
SB*	85	Spectroscopic binary star
i*	20	In multiple star system
Al*	33	Detached (Algol) eclipsing binary
WU*	17	Contact binary (W UMa) stars
bL*	13	Semi-detached (β Lyr) system
RS*	12	RS CVn close binary stars
EB*	11	Eclipsing binary stars
EB?	8	Candidate eclipsing binary
blu	5	Blue straggler
HXB	3	High-mass X-ray binary
	Other types of object	
Em*	290	Emission-line star
iC	72	Star in cluster
iN	44	Star in nebula
EmO	7	Emission object (ISM)
As*	7	Stellar associations
iA	7	Star in association
Pe*	3	Peculiar stars

(i) The carbon star HIP 56551 (HD 100764), which may be an extrinsic carbon star.

(ii) HIP 91260 (CE Lyr), which is a Mira variable, but which suffers from contamination by a nearby star.

(iii) A number of post-AGB objects also fall into this category. They include the post-AGB star HM Aqr, and the candidate post-AGB stars/protoplanetary nebulae TYC 2858-542-1 (IRAS 02529+4350) and TYC 718-517-1 (HD 246299). The remainder appear to either be misclassified Herbig Ae/Be stars or T Tauri stars: HIP 78092 (HD 142527), HIP 78943 (HD 144432), TYC 6679-305-1 (HD 143006) and TYC 6856-876-1 (HD 169142).

Finally, Herbig Ae/Be stars scatter to cooler temperatures than expected for their spectral classifications, as a result of the circumstellar material that surrounds them. Ae stars cluster around 4000 K and 2 $L_\odot$, while Be stars occupy a broader range, between 7000 and 10 000 K, and 100 and 3000 $L_\odot$. Generally speaking, they lie well above the main sequence. Many of the undesignated objects in the same region of the H–R diagram may also be Be stars in their own right.

4.6 Application to mass-losing stars on the giant branch

A useful application of this research is into the minimum luminosity of dusty giant branch stars. This is one of the few places on the H–R diagram where dust production is expected to be confined to a specific region. Fig. 16 shows the upper giant branches of the H–R diagram. Below ∼300 $L_\odot$, source densities are affected by our temperature cut-off at 4400 K. Above ∼300 $L_\odot$, our parallax uncertainty criterion of <20 per cent limits us to nearby sources. This closely matches the bright limit of *Gaia* DR1, so parallaxes of giant stars above 300 $L_\odot$ largely come from the *Hipparcos* mission, and are within ∼1 kpc of Earth. At these distances, all stars will be easily detectable by either *Hipparcos* or *Gaia*, so the source density is not strongly influenced by the easier detectability of luminous stars.

The precise conditions needed to initiate dust production around evolved stars remain unknown. Circumstellar dust around RGB stars is thought to be very rare, though not necessarily impossible (e.g. Groenewegen 2012; McDonald et al. 2012a, 2014; McDonald & Zijlstra 2016). In (metal-poor) globular clusters and the Magellanic Clouds, the onset appears between 700 and 1500 $L_\odot$ (Boyer et al. 2009, 2015; McDonald et al. 2011a,c). While the total mass-loss rate (at least in older stars) does not appear to be strongly linked to metallicity (van Loon, Boyer & McDonald 2008; McDonald & Zijlstra 2015), the onset luminosity is likely to have some metallicity dependence (e.g. McDonald et al. 2010b), as the dust column density should scale approximately with metallicity (van Loon 2006; Groenewegen et al. 2016). However, the onset is hard to trace in solar-metallicity populations due to distance or contamination. Based on the above studies, we can expect the onset of dust production to be traced by a gradual increase in the fraction of stars with infrared excess, starting at some point below the RGB tip.

The RGB tip is present in the upper panel at ∼2000 $L_\odot$. However, it is poorly defined due to a variety of observation and astrophysical factors: primarily the distance uncertainty, which can alter the luminosity by up to ±40 per cent, and the stellar mass and metallicity, which can alter the luminosity by ±20 per cent (e.g. Marigo et al. 2008). For intermediate-age and older populations, the evolutionary speed on the AGB is ∼3–5 times faster than on the RGB; hence, density declines above the RGB tip by a factor of ∼4–6. The inexact position of the RGB tip obfuscates its presence in the source density

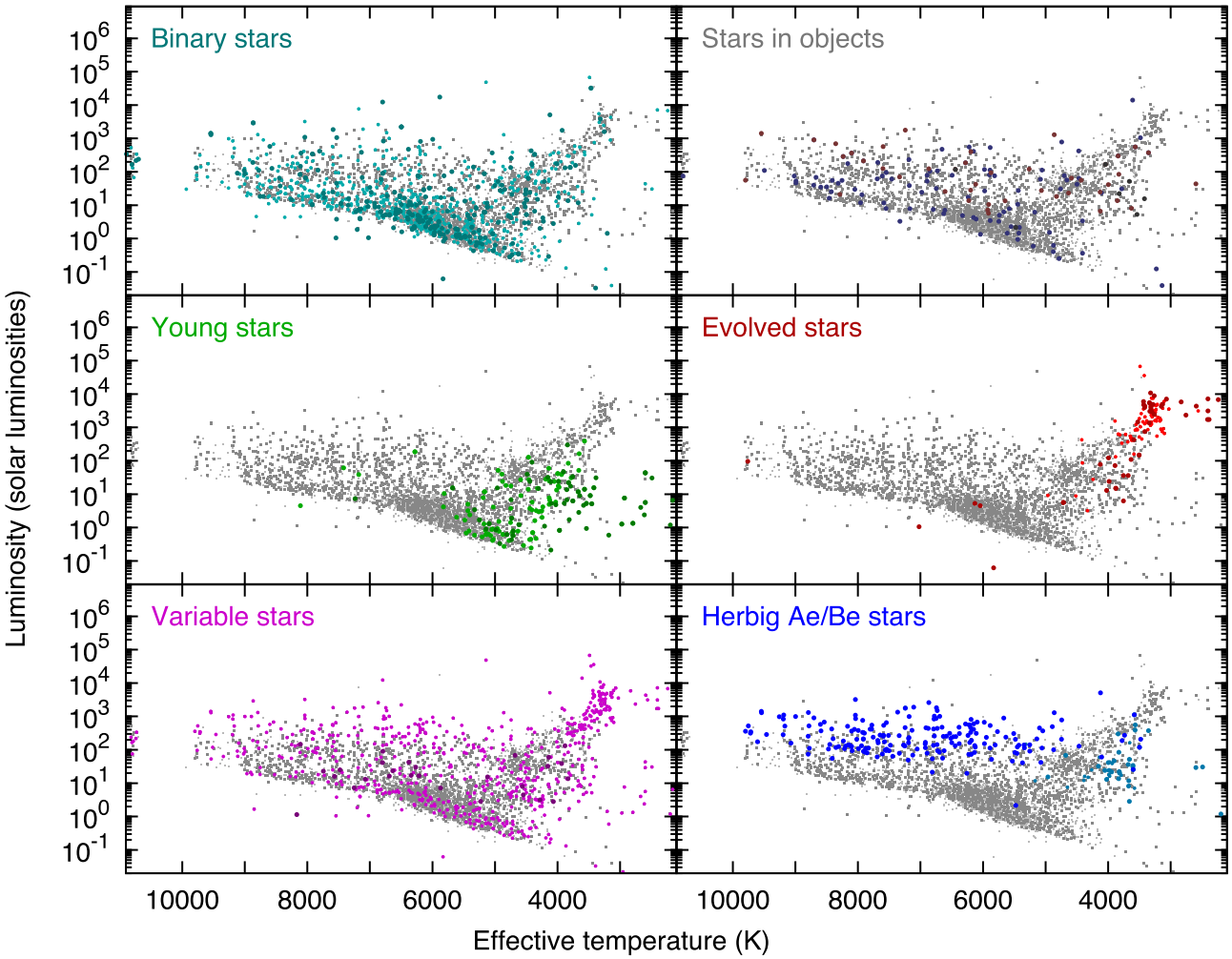

Figure 15. H–R diagrams, showing the locations of different classifications of stars. In each case, the light grey dots show all candidate stars in Table 3, with the slightly darker grey dots showing stars with high confidence (>3 points). Binary stars with no further designator are shown as smaller points with lighter colour. Stars within objects are shaded red to denote nebulae and blue to denote clusters. Young testers are coloured lighter for pre-main-sequence stars and YSOs, and darker for T Tauri stars and Herbig–Haro objects. Evolved stars are coloured light for long-period variables, and dark if their designator provides further information (e.g. Mira variable, carbon star, etc.). Variable stars are shown in larger, darker points if they are known instability strip variables (e.g. Cepheids). Herbig Ae/Be stars are shown in cyan for Ae and blue for Be stars: smaller symbols denote questionable designations (SIMBAD's Ae? and Be?).

plot (the blue line in the bottom panel of Fig. 16), but it can be seen as a small discontinuity between 2000 and 3000 $L_\odot$. Beyond the RGB tip, source density declines sharply as one ascends the upper AGB (the thermally pulsating or TP-AGB).

The limitations in modelling these cool stellar atmospheres become problematic here, however. The median $X_{\rm MIR}$ ratio starts at just above unity near the middle of the giant branch and rises slowly (the offset being largely due to the poor H^- modelling). Beyond the RGB tip, the median $X_{\rm MIR}$ rises more rapidly, until the value becomes stochastic among the most luminous AGB stars.

Simultaneously, the fraction of stars with identified excess rises slowly towards the RGB tip. However, the number of stars with clear-cut excess remains negligible until ∼890 $L_\odot$. Only a handful of giant stars with excess fall below this luminosity: l Vir, Z Peg, FW Vir, HD 68425, SU And (carbon star), RT Boo, AU Peg (W Vir variable), HM Aqr (post-AGB star), HD 100764 (carbon star), DY Boo and RU Crt. With the possible exceptions of l Vir (686 $L_\odot$), RU Crt (664 $L_\odot$) and HD 68425 (483 $L_\odot$), these objects all have very strong infrared excess, are not well modelled by a simple stellar photosphere and do not fall on the giant branch in the H–R diagram. It is likely that the luminosity has been underestimated for these stars. Circumstellar material has been detected from RU Crt (McDonald et al., in preparation), identifying it as the lowest luminosity giant where a dusty outflow has been convincingly detected.

As one progresses above 890 $L_\odot$, there comes a steady list of sources with infrared excess. The fraction of sources is fairly low at first, but increases significantly at the RGB tip (Fig. 16, bottom panel). The luminosity function of sources with strong infrared excess does not change appreciably across the RGB tip, arguing that few (if any) RGB stars exhibit circumstellar dust. All the giant stars that have infrared excess and are near the RGB tip are therefore expected to be AGB stars. The fraction of stars with infrared excess, and the amount of infrared excess they have, both increase with luminosity as stars ascend the AGB.

5 CONCLUSIONS

In this paper, we have photometrically matched numerous public data bases of stellar photometry against parallactic measurements

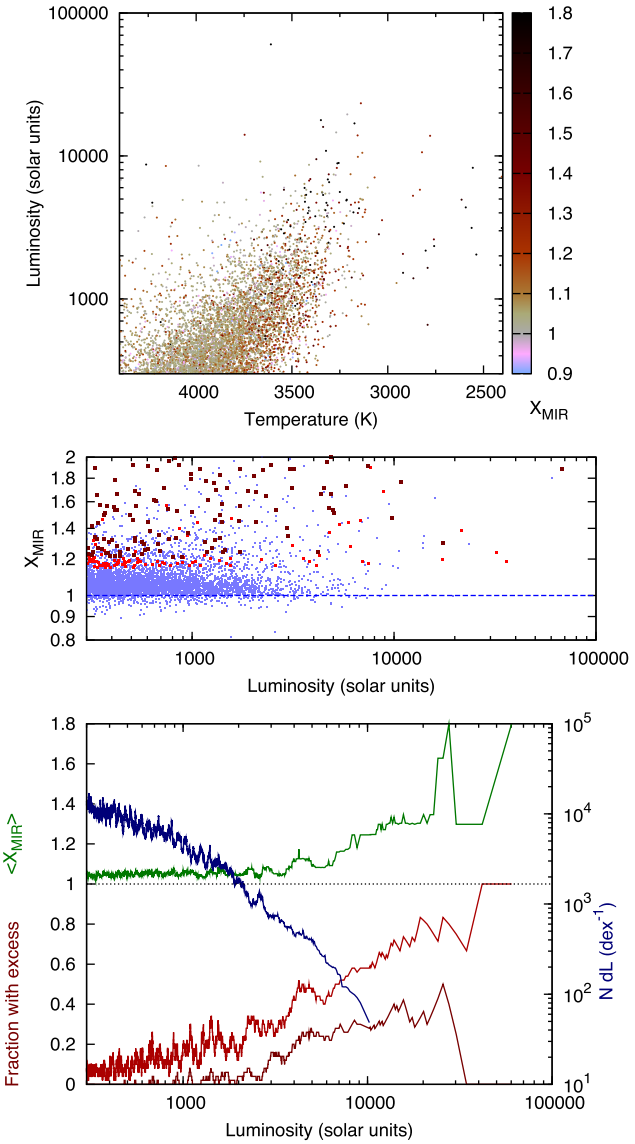

Figure 16. Top panel:the upper giant branches, with infrared excess colour-coded as in the top panel of Fig. 13. Only stars with $A_V < 1.5$ mag and $\delta\varpi/\varpi < 0.2$ are considered. Middle panel: mid-IR excess of individual stars as a function of luminosity (light blue points). Red points show stars with candidate infrared excess; dark red points show those stars with a score of more than three points. Bottom panel: the evolution of stellar properties with luminosity, smoothed by a running mean of 50 stars. From top to bottom (as viewed from the left-hand side of the plot), (1) the blue line shows the number of sources per dex in luminosity. This is shown on the right-hand (logarithmic) scale. (2) The green line shows the median X_{MIR} stars at that luminosity have. (3) The red line shows the fraction of stars meeting our infrared-excess criterion ($X_{MIR} > 1.15 + A_V/100$). (4) The darker red line shows the same plot for $X_{MIR} > 1.5 + A_V/3.1$. The RGB tip lies between around 2400 and 2500 $L_\odot$, for most solar-metallicity stars. A dotted line is placed at unity to guide the eye.

of stellar distances from the *Gaia* satellite's first data release. Modelling of the resulting SEDs has allowed us to derive the temperature and luminosity for 1583 066 unique objects, placing them on the H–R diagram. We report on the goodness of fit of each best-fitting model, and quantify the presence of infrared excess around each star.

We list 4256 stars that are candidates for infrared excess, of which 1883 are qualified as having strong evidence of infrared excess. These objects have been categorized by their literature classifications. A large number of previously identified binary, variable and emission-line stars are recovered, along with a substantial number of potentially new detections.

We briefly explore some of the facets of this data set.

(i) We identify that the vast majority of the *Gaia* DR1 data set exhibits relatively little extinction, although a small but significant number of stars (mainly giant stars) are still considerably affected.

(ii) We explore dust production among nearby giant stars, confirming that little or no dust condensation takes place around RGB stars, but becomes prevalent in AGB stars at an evolution point close to the RGB tip.

(iii) We explore populations at different Galactic scaleheights, identifying that stars with ages <3 Gyr have a strong tendency to be located within ∼200 pc of the Galactic plane, and that the metallicity of nearby stars remains close to the solar value until one exceeds ∼600 pc from the plane.

(iv) We identify hot stars within a few hundred parsecs of the Sun, and use these to map out sites of recent star formation in the solar neighbourhood. Dust clouds and hot stars are presented in three dimensions and basic inferences drawn on their relation to the Gould Belt.

Our closing recommendations for repeating this study on a larger data set, following future *Gaia* data releases, are presented in Appendix F (online only).

ACKNOWLEDGEMENTS

The authors acknowledge support from the UK Science and Technology Facility Council under grant ST/L000768/1. This paper could not have been possible without data from a large of surveys and facilities. Their standard requested acknowledgements are listed below.

(i) This research made use of the cross-match service provided by CDS, Strasbourg.

(ii) This research has made use of the SIMBAD data base, operated at CDS, Strasbourg, France.

(iii) This work has made use of data from the European Space Agency (ESA) mission *Gaia* (http://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, http://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.

(iv) Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the US Department of Energy Office of Science and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, the Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, Lawrence

Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, the Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University and Yale University.

(v) The DENIS project has been partly funded by the SCIENCE and the HCM plans of the European Commission under grants CT920791 and CT940627. It is supported by INSU, MEN and CNRS in France, by the State of Baden-Württemberg in Germany, by DGICYT in Spain, by CNR in Italy, by FFwFBWF in Austria, by FAPESP in Brazil, by OTKA grants F-4239 and F-013990 in Hungary and by the ESO C&EE grant A-04-046.
Jean Claude Renault from IAP was the Project manager. Observations were carried out thanks to the contribution of numerous students and young scientists from all involved institutes, under the supervision of P. Fouqué, survey astronomer resident in Chile.

(vi) Based on data obtained as part of the INT Photometric Hα Survey of the Northern Galactic Plane.

(vii) This work is based in part on data obtained as part of the UKIRT Infrared Deep Sky Survey.

(viii) 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.

(ix) Based on observations with *AKARI*, a JAXA project with the participation of ESA.

(x) 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.

(xi) The *InfraRed Astronomical Satellite* was developed and operated by the Netherlands Agency for Aerospace Programmes (NIVR), the US National Aeronautics and Space Administration (NASA) and the UK Science and Engineering Research Council (SERC).

(xii) This research used the DIRBE Point Source Photometry Research Tool, a service provided by the Legacy Archive for Microwave Background Data at NASA's Goddard Space Flight Center.

(xiii) This research made use of data products from the *Midcourse Space Experiment*. Processing of the data was funded by the Ballistic Missile Defense Organization with additional support from NASA Office of Space Science. This research has also made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

(xiv) Based on observations obtained with *Planck* (http://www.esa.int/Planck), an ESA science mission with instruments and contributions directly funded by ESA Member States, NASA and Canada.

The authors also wish to thank the anonymous referee for their careful scrutiny of the paper and the enlightening comments raised during the ensuing discussion.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Adams J. J. et al., 2013, ApJ, 771, 112
Albrecht A. et al., 2006, preprint (arXiv:astro-ph/0609591)
Allard F., Guillot T., Ludwig H.-G., Hauschildt P. H., Schweitzer A., Alexander D. R., Ferguson J. W., 2003, in Martín E., ed., Proc. IAU Symp. 211, Brown Dwarfs. Astron. Soc. Pac., San Francisco, p. 325
ALMA Partnership et al., 2015, ApJ, 808, L3
Anderson E., Francis C., 2012, Astron. Lett., 38, 331
Andrews S. M. et al., 2016, ApJ, 820, L40
Astraatmadja T. L., Bailer-Jones C. A. L., 2016, ApJ, 833, 119
Bailer-Jones C. A. L., 2015, PASP, 127, 994
Barentsen G. et al., 2014, MNRAS, 444, 3230
Beichmann C. A., Helou G., Walker D. W., 1988, Infrared Astronomical Satellite (IRAS). Catalogs and Atlases. NASA RP (Reference Publication). NASA, Washington
Benitez N. et al., 2014, preprint (arXiv:1403.5237)
Benjamin R. A. et al., 2003, PASP, 115, 953
Bensby T., Feltzing S., Lundström I., 2004, A&A, 421, 969
Bensby T., Feltzing S., Oey M. S., 2014, A&A, 562, A71
Boyer M. L. et al., 2009, ApJ, 705, 746
Boyer M. L. et al., 2011, AJ, 142, 103
Boyer M. L., McDonald I., Srinivasan S., Zijlstra A., van Loon J. T., Olsen K. A. G., Sonneborn G., 2015, ApJ, 810, 116
Broekhoven-Fiene H. et al., 2013, ApJ, 762, 52
Carey S. J. et al., 2009, PASP, 121, 76
Chambers K. C. et al., 2016, preprint (arXiv:1612.05560)
Chandler C. O., McDonald I., Kane S. R., 2016, AJ, 151, 59
Churchwell E. et al., 2009, PASP, 121, 213
Cotten T. H., Song I., 2016, ApJS, 225, 15
Cutri R. M. et al., 2003, The IRSA 2MASS All-Sky Point Source Catalog. NASA/IPAC Infrared Science Archive
de Jong J. T. A. et al., 2013, The Messenger, 154, 44
de Zeeuw P. T., Hoogerwerf R., de Bruijne J. H. J., Brown A. G. A., Blaauw A., 1999, AJ, 117, 354
Decin L., Morris P. W., Appleton P. N., Charmandaris V., Armus L., Houck J. R., 2004, ApJS, 154, 408
Dodson-Robinson S. E., Su K. Y. L., Bryden G., Harvey P. M., Green J. D., 2016, ApJ, 833, 183
Draine B. T., 2003, ARA&A, 41, 241
Draine B. T., Li A., 2007, ApJ, 657, 810
Drew J. E. et al., 2013, The Messenger, 154, 41
Drew J. E. et al., 2014, MNRAS, 440, 2036
Edge A., Sutherland W., Kuijken K., Driver S., McMahon R., Eales S., Emerson J. P., 2013, The Messenger, 154, 32
Egan M. P., Price S. D., 1996, AJ, 112, 2862
Gaia Collaboration et al., 2016a, A&A, 595, A1
Gaia Collaboration et al., 2016b, A&A, 595, A2
Gail H., Sedlmayr E., 1999, A&A, 347, 594
Gilmore G., Reid N., 1983, MNRAS, 202, 1025
Gramajo L. V., Whitney B. A., Gómez M., Robitaille T. P., 2010, AJ, 139, 2504
Groenewegen M. A. T., 2012, A&A, 540, A32
Groenewegen M. A. T. et al., 2016, A&A, 596, A50
Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, A&A, 486, 951
Hayashi C., 1961, PASJ, 13, 450
Henize K. G., 1976, ApJS, 30, 491
Herschel Sir J. F. W., 1847, Results of Astronomical Observations Made during the Years 1834, 5, 6, 7, 8, at the Cape of Good Hope; Being the Completion of a Telescopic Survey of the Whole Surface of the Visible Heavens, Commenced in 1825. Smith, Elder & Co., London

Hinkel N. R., Timmes F. X., Young P. A., Pagano M. D., Turnbull M. C., 2014, AJ, 148, 54
Høg E. et al., 2000, A&A, 355, L27
Ishihara D. et al., 2010, A&A, 514, A1
Ishihara D., Takeuchi N., Kobayashi H., Nagayama T., Kaneda H., Inutsuka S.-i., Fujiwara H., Onaka T., 2017, A&A, 601, A72
Johnson C. I. et al., 2015, AJ, 149, 71
Karakas A. I., Lattanzio J. C., 2014, 31, e030
Kervella P., Homan W., Richards A. M. S., Decin L., McDonald I., Montargès M., Ohnaka K., 2016, A&A, 596, A92
Kong D.-L., Zhu Z., 2008, Chin. Astron. Astrophys., 32, 360
Lapenna E., Mucciarelli A., Lanzoni B., Rosario Ferraro F., Dalessandro E., origlia L., Massari D., 2014, ApJ, 797, 124
Leão I. C., Bright S. N., Chesneau O., 2015, in Lagadec E., Millour E., Lans T., eds, EAS Publ. Ser. Vol. 71, On the Inner Circumstellar Envelopes of R Coronae Borealis Stars. EDP Sciences, Les Ulis, France, p. 289
Lee C.-D., Chen W.-P., Liu S.-Y., 2016, A&A, 592, A130
Lutz T. E., Kelker D. H., 1973, PASP, 85, 573
Lykou F. et al., 2015, A&A, 576, A46
McCabe C., Duchêne G., Pinte C., Stapelfeldt K. R., Ghez A. M., Ménard F., 2011, ApJ, 727, 90
McDonald I., Zijlstra A. A., 2015, MNRAS, 448, 502
McDonald I., Zijlstra A. A., 2016, ApJ, 823, L38
McDonald I., van Loon J. T., Decin L., Boyer M. L., Dupree A. K., Evans A., Gehrz R. D., Woodward C. E., 2009, MNRAS, 394, 831
McDonald I., Sloan G. C., Zijlstra A. A., Matsunaga N., Matsuura M., Kraemer K. E., Bernard-Salas J., Markwick A. J., 2010a, ApJ, 717, L92
McDonald I., van Loon J. T., Dupree A. K., Boyer M. L., 2010b, MNRAS, 405, 1711
McDonald I. et al., 2011a, ApJS, 193, 23
McDonald I., Johnson C. I., Zijlstra A. A., 2011b, MNRAS, 416, L6
McDonald I. et al., 2011c, MNRAS, 417, 20
McDonald I., Zijlstra A. A., Boyer M. L., 2012a, MNRAS, 427, 343
McDonald I., White J. R., Zijlstra A. A., Guzman Ramirez L., Szyszka C., van Loon J. T., Lagadec E., Jones O. C., 2012b, MNRAS, 427, 2647
McDonald I., Zijlstra A. A., Sloan G. C., Kerins E., Lagadec E., Minniti D., 2014, MNRAS, 439, 2618
McDonald I., Zijlstra A. A., Sloan G. C., Lagadec E., Johnson C. I., Uttenthaler S., Jones O. C., Smith C. L., 2016, MNRAS, 456, 4542
VHS CollaborationMcMahon R. G., Banerji M., Gonzalez E., Koposov S. E., Bejar V. J., Lodieu N., Rebolo R., 2013, The Messenger, 154, 35
Malmquist G. K., 1920, Medd. Lunds Astron. Obs. Ser. II, 22, 3
Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883
Martig M., Minchev I., Ness M., Fouesneau M., Rix H.-W., 2016, ApJ, 831, 139
Masseron T., Gilmore G., 2015, MNRAS, 453, 1855
Michalik D., Lindegren L., Hobbs D., 2015, A&A, 574, A115
Minniti D. et al., 2010, New Astron., 15, 433
Miroshnichenko A. S., Bjorkman K. S., Grosso M., Hinkle K., Levato H., Marang F., 2005, A&A, 436, 653
Morel M., Magnenat P., 1978, A&AS, 34, 477
Neugebauer G. et al., 1984, ApJ, 278, L1
Olano C. A., 1982, A&A, 112, 195
Olano C. A., 2001, AJ, 121, 295
Perryman M. A. C., 1989, Nature, 340, 111
Perryman M. A. C. et al., 2001, A&A, 369, 339
Reid M. J., Menten K. M., 1997, ApJ, 476, 327
Reid I. N., Turner E. L., Turnbull M. C., Mountain M., Valenti J. A., 2007, ApJ, 665, 767
Richards A. M. S. et al., 2014, A&A, 572, L9
Ruffle P. M. E. et al., 2015, MNRAS, 451, 3504
Schöier F. L., Lindqvist M., Olofsson H., 2005, A&A, 436, 633
Shanks T. et al., 2013, The Messenger, 154, 38
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Sloan G. C. et al., 2012, ApJ, 752, 140
Sloan G. C. et al., 2016, ApJ, 826, 44
Smith H., 2003, MNRAS, 338, 891
Smith B. J., Price S. D., Baker R. I., 2004, ApJS, 154, 673
Stock J., Wing R. F., 1972, BAAS, 4, 324
Taylor B. J., Croxall K., 2005, MNRAS, 357, 967
van Leeuwen F., 2007, A&A, 474, 653
van Loon J. T., 2006, in Lamers H. J. G. L. M., Langer N., Nugis T., Annuk K., eds, ASP Conf. Ser. Vol. 353, Stellar Evolution at Low Metallicity: Mass Loss, Explosions, Cosmology. Astron. Soc. Pac., San Francisco, p. 211
van Loon J. T., Boyer M. L., McDonald I., 2008, ApJ, 680, L49
Ward-Thompson D. et al., 2007, PASP, 119, 855
Warren W. H. Jr, 1991, Photoelectric Photometric Catalogue of Homogeneous Means in the UBV System (Mermilliod 1991). NASA, Greenbelt, MD
Woods P. M. et al., 2011, MNRAS, 411, 1597
Wray J. D., 1966, PhD thesis, Northwestern University
Wright J. T. et al., 2011, PASP, 123, 412

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table 1. Fundamental parameters and infrared excess for Tycho-2 stars.
Table 2. Fundamental parameters and infrared excess for *Hipparcos* stars.
Table 3. Catalogue of stars with candidacy for hosting infrared excess.
Appendix A. Data Flagging in Tycho-2 Data.
Appendix B. Data Flagging in *Hipparcos* Data.
Appendix C. Exploring the Lutz–Kelker Bias and Related Effects.
Appendix D. Discussion on Interstellar Extinction.
Appendix E. Discussion on Infrared Excess.
Appendix F. Looking Forward to Future *Gaia* Releases.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.