

Recalibrating Fundamental Parameters of Low-Mass Stars for Current and Future Planet Hunting Surveys

Christopher A. Theissen (PI), Adam J. Burgasser, Christian Aganze

Summary

We request a total of **8,314,035 SUs** on the PSC machine BRIDGES-2 Regular Memory nodes to re-characterize all low-mass stars that are being targeted with current and future planet hunting surveys. We will use well-tested ensemble samplers built on the EMCEE code to fit stellar photosphere models to photometry from the all-sky surveys. To store the chains that will result from the posterior distributions of these fits, we request **7,454 GB** of storage on the BRIDGES-2 Ocean system.

1 Introduction and Scientific Background

The field of exoplanetary science has grown rapidly since the first Nobel-winning discovery of 51 Peg b over two decades ago (Mayor & Queloz, 1995). With the advent of space-based planet hunting satellites (e.g., *Kepler*, Borucki et al. 2010; *Transiting Exoplanet Survey Satellite* or *TESS*, Ricker et al. 2015), and ground-based transit surveys (e.g., the Kilodegree Extremely Little Telescope or KELT, (Pepper et al., 2007); MEarth, (Nutzman & Charbonneau, 2008); the TRAnsiting Planets and PlanetesImals Small Telescope or TRAPPIST, Jehin et al. 2011), the demographics of exoplanetary systems has expanded dramatically. Low-mass stars ($M \lesssim 0.8 M_{\odot}$, $T_{\text{eff}} \lesssim 5200 \text{ K}$, spectral types K0 and later) comprise $\gtrsim 80\%$ of all stars in the Galaxy (Bochanski et al., 2010; Henry et al., 2018) and its members have extremely long and stable main sequence lifetimes ($\gtrsim 10^{13}$ yrs; Laughlin et al. 1997). In addition to their abundance and long lifetimes, low-mass stars are important targets for exoplanet searches, as their habitable zones (HZs) are sufficiently close in ($\sim 0.05\text{--}0.8 \text{ au}$) to facilitate efficient detection with transit and radial velocity (RV) techniques (Triaud et al., 2013).

Low-mass stars are also abundant terrestrial exoplanet hosts. Planets are increasingly common among lower mass stars ($M \lesssim 0.6 M_{\odot}$), with the terrestrial planet occurrence rate reaching as high as 3–8 planets/star for stars with $M \approx 0.5 M_{\odot}$ (Hardegree-Ullman et al., 2019; Hsu et al., 2020). Nearly all terrestrial planets orbiting within the HZ of a host star have been identified around low-mass stars¹ (Dressing & Charbonneau, 2015), as is shown in Figure 1.

¹<http://phl.upr.edu/projects/habitable-exoplanets-catalog>

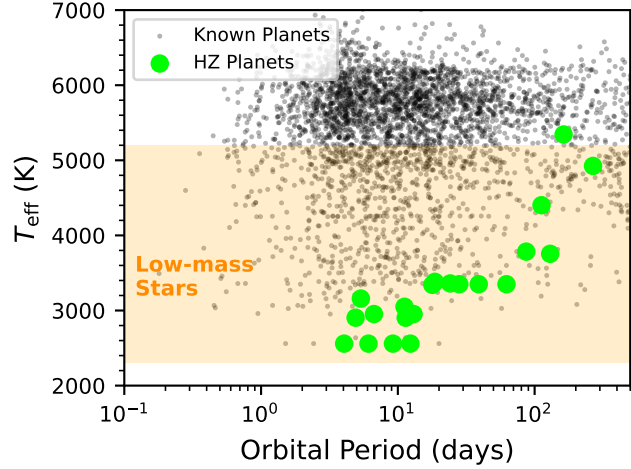


Figure 1 – Orbital period vs. host star effective temperature for known exoplanets (black points; data from <http://exoplanet.eu>). HZ planets are shown as large colored circles. The temperature region encompassing low-mass stars (orange) is indicated. Essentially all the known HZ planets orbit low-mass stars.

A core principle in exoplanetary science is that in order to know a planet, one must know its host star. One essential aspect of stellar characterization for transiting exoplanets is stellar size, as the planet radius is directly inferred from the stellar radius. The most common approach to this is through the Stefan-Boltzmann Law: $L = 4\pi R_*^2 \sigma_{\text{SB}} T_{\text{eff}}^4$. Luminosity can be inferred from broad-band spectral energy distribution measurements and an accurate distance. Temperature can be inferred from empirical calibrations or model fits to measured colors or spectra. The *Kepler* Input Catalog (KIC; Brown et al. 2011), K2 Ecliptic Plane Input Catalog (EPIC; Huber et al. 2016), and *TESS* Input Catalog (TIC; Stassun et al. 2018, 2019)² rely on this procedure to determine luminosities, temperatures and radii for hundreds of thousands of stars. However, these procedures can fail for peculiar stars or for those with extreme temperatures or colors, including low-mass stars. Biases in inferred temperature can be particularly acute when incomplete information is available (e.g., no measured spectrum or parallax). Since the inferred stellar radius—and by extension, exoplanet radii—scales linearly with distance and quadratically with temperature, accurate determination of both of these parameters is crucial for accurate exoplanet characterization.

This project aims to improve the characterization of the physical properties of cool dwarf stars and brown dwarfs ($T_{\text{eff}} \lesssim 5200$ K). By integrating updated empirical calibrations and *Gaia* astrometry into the input catalog, this project would result in the first *uniform* catalog of robust temperatures, luminosities and radii for low-mass dwarfs from different planet-hunting surveys.

2 Research Objectives - Improved Physical Characterization of the Cool Dwarfs

We have one main overarching research objective for our proposal, the complete re-characterization (temperature, surface-gravity, metallicity, luminosity, and radius) of all

²<https://tess.mit.edu/science/tess-input-catalogue/>

the cool dwarfs within a well-defined sample. Biases in the physical parameters of the lowest-mass stars have already been demonstrated in the TIC through the construction of the *TESS* Cool Dwarf Catalog (CDC; Muirhead et al. 2018) for stars with $T_{\text{eff}} \lesssim 4000$ K. The CDC makes use of color- T_{eff} relations (Mann et al., 2015, 2016), which are specifically calibrated for dwarf stars with $2700 \leq T_{\text{eff}} \leq 4100$ K (spectral types K7–M7). This distinct calibration results in significant shifts in T_{eff} for individual sources, in some cases by up to 1000 K; and a systemic trend toward lower T_{eff} s in the CDC compared to the TIC ($\Delta T_{\text{eff}} \approx -83$ K; Figure 2). However, the CDC is itself incomplete, as it was constructed prior to the availability of precise astrometry from *Gaia*. Figure 2 shows that the addition of *Gaia* eDR3 parallax measurements reveal clear contaminants in the CDC (e.g., white dwarfs), which undoubtedly result in incorrect determinations of luminosity, temperature, and radius for those sources.

The *K2* EPIC displays similar temperature biases for its coolest stars. Figure 3 compares temperature estimates as a function of $G - G_{RP}$ color for sources within 100 pc for EPIC and *TESS* CDC, both cross-matched with *Gaia* eDR3. The majority of EPIC sources are characterized by stellar color relations, and these fail to follow the linear decline of temperature with increasing color evident in the CDC data. This temperature bias suggests that many low-mass stars in *K2* may have been overlooked as a result. Improve-

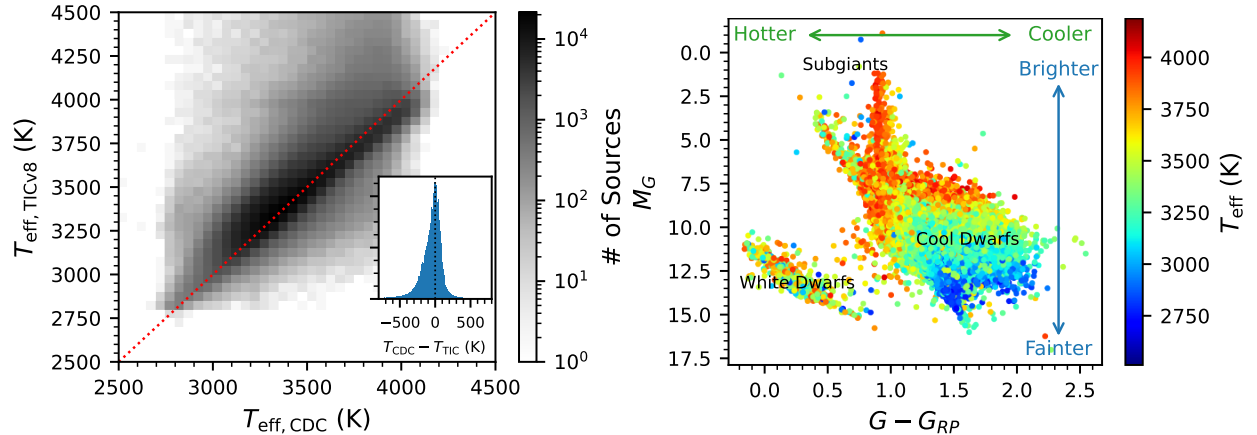


Figure 2 – *Left*: Comparison of T_{eff} values between all sources in the *TESS* CDC and the most recent version of the TIC (v8; Stassun et al. 2019). The former makes use of empirical relations calibrated specifically for the lowest-mass and lowest-temperature stars. The red dotted line denotes equal values. The inset plot shows the distribution of the temperature differences. There are significant outliers, in some cases differing by over 1000 K; and a systemic offset of -83 K between TIC and CDC *Right*: *Gaia* absolute magnitude-color ($M_G/G - G_{RP}$) diagram for the CDC, color coded by T_{eff} . With the inclusion of *Gaia* data, it is now obvious that there are numerous contaminants in the CDC, which requires revision for accurate characterization of low-mass stars and their planetary companions.

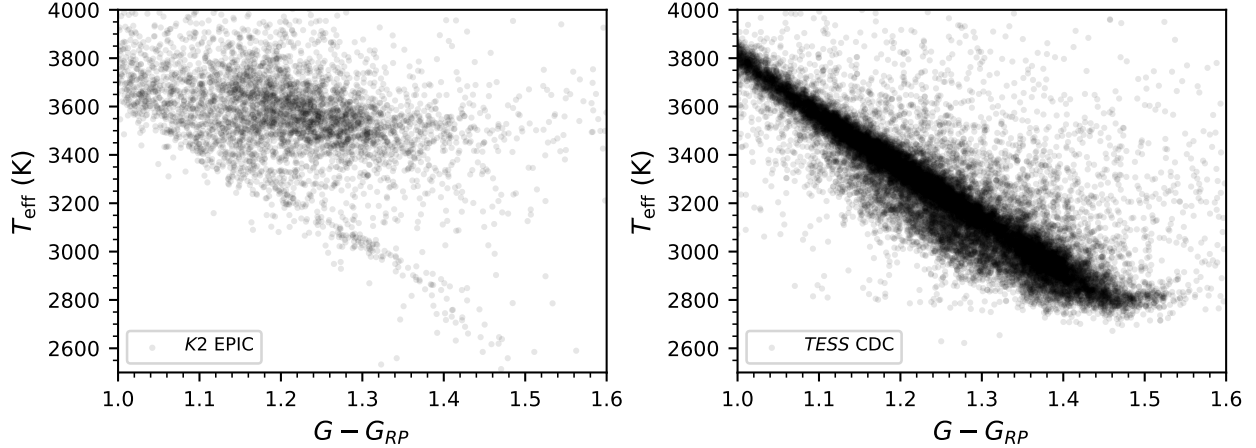


Figure 3 – *Left:* T_{eff} versus *Gaia* $G - G_{RP}$ color for K2 sources within 100 pc. The divergence in inferred T_{eff} s for $G - G_{RP} \geq 1.2$ likely reflects a bias in the color- T_{eff} relation at the lowest temperatures. *Right:* T_{eff} versus $G - G_{RP}$ for TESS CDC sources within 100 pc. Here, T_{eff} is a relatively smooth function of $G - G_{RP}$.

ments to the EPIC have been made with the use of optical spectroscopy and parallax measurements (Hardegree-Ullman et al., 2020); however, these re-calibrations only extend to stars with $T_{\text{eff}} > 3100$ K, thereby excluding the lowest-mass stars.

The objective of our project is to re-evaluate the physical parameters for all cool dwarfs ($T_{\text{eff}} \lesssim 5200$ K) in the *Kepler*, *K2*, and *TESS* input catalogs, as well as targeted sources not contained in the original input catalogs. **We will produce the first uniformly calibrated catalog of cool dwarfs across the *Kepler*, *K2*, and *TESS* missions.** We will specifically identify cool dwarfs missed or mis-identified in these catalogs by performing a cross-match of all sources to *Gaia* eDR3 (the full Data Release 3 is set to be released in June 2022) to robustly identify cool dwarf stars. Additionally, there is considerable evidence that additional members of the lowest-mass stars, the so-called “ultracool dwarfs” (UCDs; $T_{\text{eff}} \leq 2800$ K), remain to be recovered from existing mission data. The KIC contains no stars cooler than ~ 3000 K, while the *K2* EPIC contains 805 dwarfs with $T_{\text{eff}} \leq 2800$ K (spectral types $\geq M6$). The TIC contains 458 dwarfs with $T_{\text{eff}} \leq 2800$ K; however, the CDC contains nearly three times as many sources below this temperature threshold. In total, we estimate that over 2500 UCDs which have not yet been properly characterized have space-based photometric lightcurves.

Gaia provides the mechanism to extract this sample. We will characterize all the cool dwarfs (inclusive of UCDs) up to $T_{\text{eff}} = 5200$ K by combining *Gaia* astrometry and broad-band photometry with the most up-to-date empirical calibrations Mann et al. (2015, 2016, 2019), yielding accurate determinations of luminosity, temperature, radius, and mass for each source. We will compare these physical parameters to those in the existing input

catalogs (including the CDC) and spectroscopic follow-up samples (e.g., Terrien et al., 2014; Newton et al., 2014; Dressing & Charbonneau, 2015; Dressing et al., 2017, 2019) to identify contaminants and assess systematic biases. The final sample will be published and made publicly available (see Data Management Plan).

Our sample (Theissen et al. in prep) is sourced from *Gaia* early Data Release 3 (eDR3; Gaia Collaboration et al. 2021), and cross-matched to the 2-Micron All-Sky Survey (2MASS; Skrutskie et al. 2006), and the *Wide-field Infrared Survey Explorer* (WISE; Wright et al. 2010) CatWISE catalog (Eisenhardt et al., 2020; Marocco et al., 2021). A large number of these stars have been or are being observed by *Kepler/K2*, *TESS*, and other ground-based planet hunting surveys. **In total, our sample contains 14,252,631 sources to be re-characterized.**

3 Resource Usage Plan

Stellar parameters can be determined through broad-band photometry that spans a range of wavelengths which covers both sides of the peak emission within the stellar spectral energy distribution (SED), through comparisons to stellar photosphere models. Although most sensitive to T_{eff} , this method is also able to provide some constraints on surface-gravity and bulk metallicity ($[M/H]$), which are also sensitive to certain photometric bands (e.g., Filippazzo et al. 2015; Schmidt et al. 2016). Additionally, fitting a stellar photosphere model to absolute photometry provides an estimate for the scale factor, or dilution factor, which takes the form of R_*^2/d^2 . Coupled with precise distances from the *Gaia* mission, we are able to estimate the size of the star (R_*), and the best-fit model SED gives us a measure of the star’s bolometric luminosity (L_{bol}).

The most robust method to determine the aforementioned parameters is from model-fitting to precise photometry ($\sim 1\%$ uncertainty). Although many methods exist to fit photometry to stellar models (e.g., amoeba method), we have chosen to use a Markov chain Monte Carlo method for its ability to fully explore the parameter space while returning informed uncertainties in the form of a posterior distribution which encompasses all measurement errors and potential systematic errors in the models. Here we describe our MCMC methodology for fitting stellar models (Theissen & West, 2017).

We use the EMCEE package (Foreman-Mackey et al., 2013), a Python implementation of the Goodman & Weare (2010) affine invariant sampler, for our MCMC sampler. For stellar models, we have chosen the PHOENIX-ACES-AGSS-COND-2011 stellar photosphere models (Husser et al., 2013). As stellar models are not continuous across the parameter space (i.e., T_{eff} , $\log g$, $[M/H]$), we interpolated between grid points using the Spectral Modeling Analysis and RV Tool (SMART) package (Hsu et al., 2021a,b). For each step in

the MCMC, the log-likelihood is given as,

$$\ln \mathcal{L}(\Theta | \mathbf{X}, \sigma) = -\frac{1}{2} \sum_{n=1}^N \left[\left(\frac{(\Theta_n \times 10^S) - X_n}{\sigma_n \times C_\sigma} \right)^2 + \ln(2\pi(\sigma_n \times C_\sigma)^2) \right], \quad (1)$$

where Θ is a vector of length N containing the model predicted fluxes for a given set of stellar parameters (T_{eff} , $\log g$, and $[M/H]$), \mathbf{X} is a vector containing the observed fluxes, σ is a vector containing the measurement errors for the observed fluxes, S is an exponential scaling factor where $10^S = (R_*/d)^2$, C_σ is a scaling factor on the uncertainties, and the length N pertains to the number of bands in which data were available. Uniform priors were chosen across the parameter space, and assumed all the parameters were normally distributed within the prior ranges listed in Table 1.

The need for a scaling factor on the uncertainties (C_σ) is primarily driven by unknown uncertainties in the models themselves (e.g., incomplete line lists, unknown molecular opacities) which can make synthetic photometry derived from the models deviate from measured photometry. Additionally, the extremely high precision of some photometric data (*Gaia* eDR3 claims precision as high as 0.00067%) makes it suspect, especially considering many low-mass stars have photometric variability on the order of 1% in the optical (Winters et al., 2017). Therefore, we allow each photometric system uncertainty to be scaled by differing factors dependent on their relative precisions (see Table 1).

Description	Symbol	Bounds
Stellar Effective Temp.	T_{eff}	(2300, 5400) K
Stellar Surface Gravity	$\log g$	(2.2, 5.5) dex
Stellar Bulk Metallicity	$[M/H]$	(-4, 1) dex
log(Scale Factor)	$\log(10^S)$	(-30, -10) dex
Noise Scale Factor (<i>Gaia</i>)	$C_{\sigma_{\text{Gaia}}}$	(1, 10^4)
Noise Scale Factor (2MASS)	$C_{\sigma_{\text{2MASS}}}$	(1, 2)
Noise Scale Factor (<i>WISE</i>)	$C_{\sigma_{\text{WISE}}}$	(1, 100)

Table 1 – MCMC Prior Ranges

Each MCMC sampler is initialized with 100 walkers, and utilizing the kernel-density estimator (KDE) from Farr et al. (2014), which has been shown to have quick convergence for similar type problems (e.g., Theissen et al. 2022). Initial walker positions are chosen based on relationships that map T_{eff} to photometric colors (e.g., Pecaut & Mamajek 2013). While these relationships tend to have high scatter they offer an informed starting point for the walkers. Additionally, the precise trigonometric parallaxes (distances) from *Gaia* eDR3, combined with relationships that map T_{eff} to R_* (e.g., Theissen & West

2017) provide informed initial positions for the scale factor ($10^S = R_*^2/d^2$). We show an example walker plot for Teegarden’s Star (Teegarden et al., 2003), a known low-mass exoplanet host (Zechmeister et al., 2019), in Figure 4. Although this walker plot shows longer chains (1000 steps) to our proposed chains (600 steps, justified below), it illustrates the strong convergence of the chains in a relatively short number of steps. The SED for Teegarden’s Star is shown in Figure 5, where we have pulled 100 models from the posterior distribution to highlight the range of models which fit the observational data. Based on our posterior model parameters and the precise *Gaia* eDR3 distance of 3.832 ± 0.093 pc, we estimate a stellar radius of $R_* = 0.12 \pm 0.01 R_\odot$.

When using an MCMC framework, it is important to obtain convergence of all parameters, and to do it within a computationally tractable number of steps. A commonly used test of convergence is the Gelman–Rubin convergence diagnostic (Gelman & Rubin,

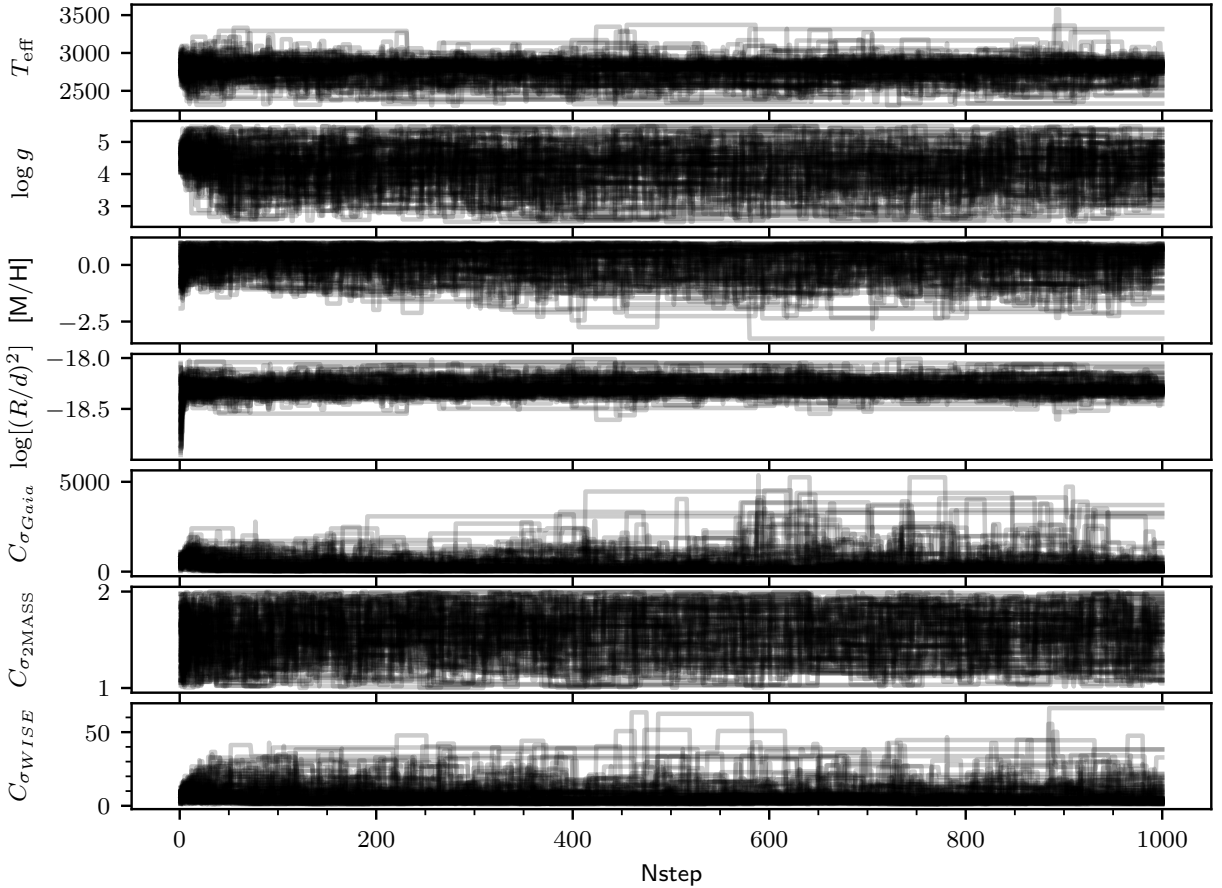


Figure 4 – Walker plot for Teegarden’s Star, a known exoplanet host (Zechmeister et al., 2019). Initial positions are chosen based on a relationship between photometric color and T_{eff} . Shown are 100 walkers over 1000 steps.

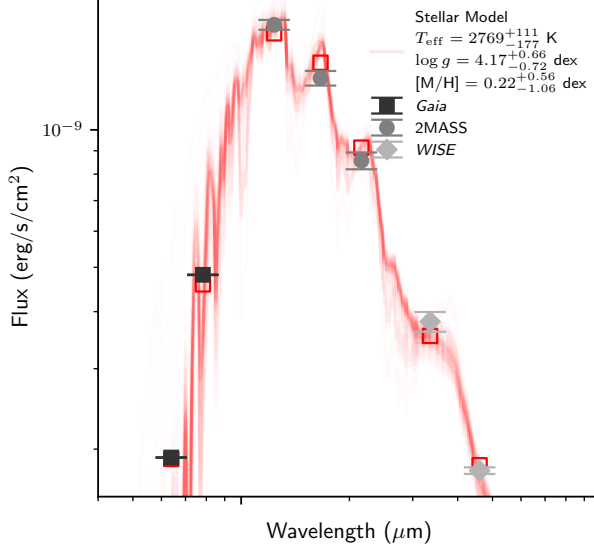


Figure 5 – Example fit using our MCMC method. Best-fit stellar parameters are shown in the top right corner. Measured photometry is illustrated with different markers and errorbars. We show 100 models pulled from the posterior distribution (red transparent lines). Also shown are synthetic photometry from the median model for each measured photometric band (red squares).

1992; Brooks & Gelman, 1998), which is a measure of intra-chain variance vs. inter-chain variance. The relevant statistic is the potential scale reduction factor (PSFR), given as

$$\hat{R} = \sqrt{\frac{\text{Var}(\hat{\theta})}{W}}, \quad (2)$$

where $\text{Var}(\hat{\theta})$ is the posterior variance (inter-chain), and W is the within-sequence variance (intra-chain). This value should approach 1, indicating that the chains have asymptotically reached a stable value. A typical estimate of convergence is when $\hat{R} \lesssim 1.1$. In Figure 6 we show the PSFR for the chains from our fit for Teegarden’s Star. The PSFR falls below convergence criteria within ~ 400 steps, which supports our choice of 600 steps to ensure adequate burn-in to obtain high-fidelity posterior distributions.

We also show the case where our initial T_{eff} and R/d estimates are significantly off from optimal values in Figure 7. There is still strong convergence within ~ 600 steps, again supporting our choice of minimal chain length to assure proper convergence.

4 Proposed Computations and Justification of Requested Resources

Our full sample of stars includes 14,252,631 sources. Based on our previous usage of BRIDGES-2 Regular Memory (see PERFORMANCE AND SCALING), the time per source on a single core is 2,100 seconds. We can compute the total Run Time and number of SUs (core-hours) for completing our sample as follows:

$$\text{Run Time (hours)} = \frac{1}{3600} \left(\frac{N_{\text{sources}} \times t_{\text{source}}}{N_{\text{cores}}} \right) \quad (3)$$

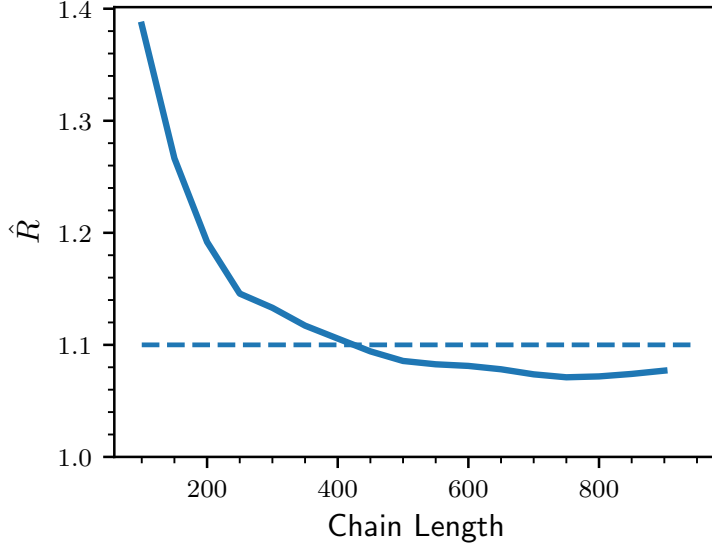


Figure 6 – Gelman–Rubin diagnostic (\hat{R} ; Gelman & Rubin 1992) as a function of chain length. Our limit for convergence, $\hat{R} = 1.1$, is indicated with the dashed line. Convergence typically occurs within 400-500 steps.

$$\text{SUs} = \text{Run Time (hours)} \times N_{\text{cores}} \quad (4)$$

where N_{sources} is the total number of sources, t_{source} is the time to finish per source (2,100 sec), and N_{cores} is the number of cores allocated. We use $N_{\text{cores}} = 6,400$, which is the maximum allowable cores on BRIDGES-2 Regular Memory nodes. From these equations, we can calculate **the total run time to be 1300 hours, or approximately 54 days, and 8,314,035 core-hours (SUs).**

We will store the compressed chains for each source on BRIDGES-2 Ocean storage system. For each source, **the compressed chains take approximately 523 KB of storage, which equates to a total of 7.454 TB.** Once our runs have completed, we will move all data over to our group’s network attached storage for archival purposes.

5 Resource Appropriateness

Our proposed project capitalizes on high scalability. BRIDGES-2 offers the highest amount of scalability at the lowest cost for our project. We have done testing using startup allocations with BRIDGES-2 Regular Memory nodes (see PERFORMANCE AND SCALING), and have determined it is optimal for our proposed estimation of stellar parameters for our entire sample.

6 Other Computational Resources

Our group has access to two servers running AMD Ryzen Threadripper 3990X processors (64 cores each). While these machines are adequate for testing small batches of our full sample, the wall-time required to run our full sample is longer than 3 years, which makes these inadequate resources to perform our proposed computations. Additionally, our group has a network attached storage with 27 TB of redundant storage capacity. This is

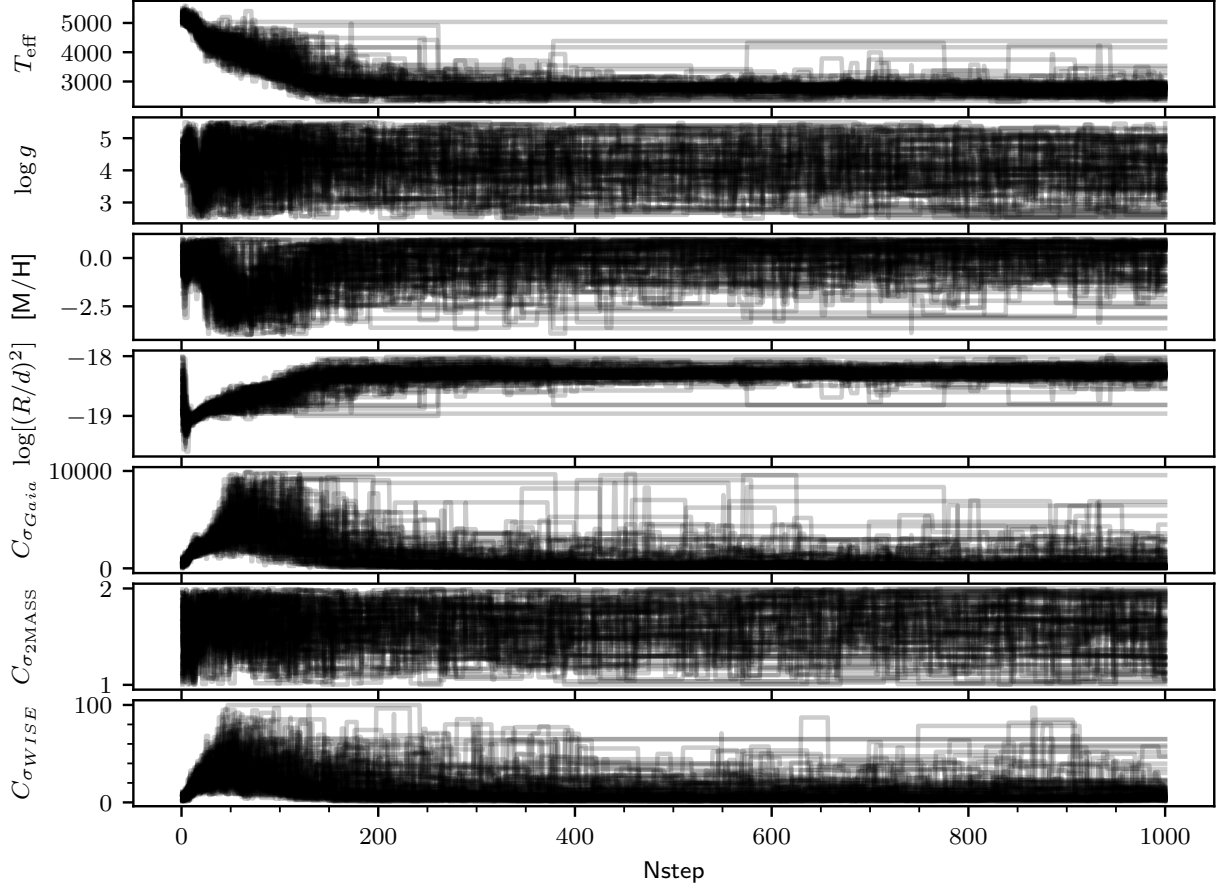


Figure 7 – Same as Figure 4, except starting from a T_{eff} position far from the optimal value. Even in this case, we see that convergence happens within 600 steps (our nominal value for all fits).

adequate to host the chains from our sample, but the high-performance read/write speed of BRIDGES-2 Ocean is superior for hosting the data during our runs. We will move all finished data to our NAS once the runs have completed.

References and Citations

- Bochanski, J. J., Hawley, S. L., Covey, K. R., et al. 2010, *AJ*, 139, 2679,
Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977,
Brooks, S. P., & Gelman, A. 1998, *Journal of Computational and Graphical Statistics*, 7, 434,
Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, *AJ*, 142, 112,
Dressing, C. D., & Charbonneau, D. 2015, *ApJ*, 807, 45,
Dressing, C. D., Newton, E. R., Schlieder, J. E., et al. 2017, *ApJ*, 836, 167,
Dressing, C. D., Hardegree-Ullman, K., Schlieder, J. E., et al. 2019, *AJ*, 158, 87,
Eisenhardt, P. R. M., Marocco, F., Fowler, J. W., et al. 2020, *ApJS*, 247, 69,
Farr, B., Kalogera, V., & Lijten, E. 2014, , 90, 024014,
Filippazzo, J. C., Rice, E. L., Faherty, J., et al. 2015, *ApJ*, 810, 158,
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306,
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, *A&A*, 649, A1,
Gelman, A., & Rubin, D. B. 1992, *Statistical Science*, 7, 457,
Goodman, J., & Weare, J. 2010, *Communications in Applied Mathematics and Computational Science*, 5, 65
Hardegree-Ullman, K. K., Cushing, M. C., Muirhead, P. S., & Christiansen, J. L. 2019, *AJ*, 158, 75,
Hardegree-Ullman, K. K., Zink, J. K., Christiansen, J. L., et al. 2020, *ApJS*, 247, 28,
Henry, T. J., Jao, W.-C., Winters, J. G., et al. 2018, *AJ*, 155, 265,
Hsu, C.-C., Theissen, C., Burgasser, A., & Birky, J. 2021a, SMART: The Spectral Modeling Analysis and RV Tool, v1.0.0, Zenodo,
Hsu, C.-C., Burgasser, A. J., Theissen, C. A., et al. 2021b, *ApJS*, 257, 45,
Hsu, D. C., Ford, E. B., & Terrien, R. 2020, *MNRAS*, 498, 2249,
Huber, D., Bryson, S. T., Haas, M. R., et al. 2016, *ApJS*, 224, 2,
Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, 553, A6,
Jehin, E., Gillon, M., Queloz, D., et al. 2011, *The Messenger*, 145, 2
Laughlin, G., Bodenheimer, P., & Adams, F. C. 1997, *ApJ*, 482, 420,
Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, *ApJ*, 804, 64,
—. 2016, *ApJ*, 819, 87,
Mann, A. W., Dupuy, T., Kraus, A. L., et al. 2019, *ApJ*, 871, 63,
Marocco, F., Eisenhardt, P. R. M., Fowler, J. W., et al. 2021, *ApJS*, 253, 8,
Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355,
Muirhead, P. S., Dressing, C. D., Mann, A. W., et al. 2018, *AJ*, 155, 180,
Newton, E. R., Charbonneau, D., Irwin, J., et al. 2014, *AJ*, 147, 20,
Nutzman, P., & Charbonneau, D. 2008, *PASP*, 120, 317,
Pecaut, M. J., & Mamajek, E. E. 2013, *ApJS*, 208, 9,
Pepper, J., Pogge, R. W., DePoy, D. L., et al. 2007, *PASP*, 119, 923,
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 014003,
Schmidt, S. J., Wagoner, E. L., Johnson, J. A., et al. 2016, *MNRAS*, 460, 2611,
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, 131, 1163,
Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2018, *AJ*, 156, 102,
Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, *AJ*, 158, 138,
Teegarden, B. J., Pravdo, S. H., Hicks, M., et al. 2003, *ApJ*, 589, L51,
Terrien, R. C., Mahadevan, S., Deshpande, R., et al. 2014, *ApJ*, 782, 61,
Theissen, C. A., Konopacky, Q. M., Lu, J. R., et al. 2022, *ApJ*, 926, 141,
Theissen, C. A., & West, A. A. 2017, *AJ*, 153, 165,
Triaud, A. H. M. J., Gillon, M., Selsis, F., et al. 2013, *ArXiv e-prints*.

Winters, J. G., Sevrinsky, R. A., Jao, W.-C., et al. 2017, *AJ*, 153, 14,
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868,
Zechmeister, M., Dreizler, S., Ribas, I., et al. 2019, *A&A*, 627, A49,

Performance and Scaling

1 Testing with BRIDGES-2 Regular Memory

Our fitting code is built on top of the EMCEE package, which is an affine invariant Markov chain Monte Carlo (MCMC) ensemble sampler. In our RESEARCH PROPOSAL we discuss our choice of fitting procedure, model parameters, and prior limits. Here we highlight scaling tests done on BRIDGES-2 Regular Memory nodes using a startup allocation.

We performed both strong and weak scaling tests for our custom MCMC Python code (see RESEARCH PROPOSAL; 100 walkers; 600 steps) using BRIDGES-2 Regular Memory nodes. Figure 1 shows the results of both of our tests. For our strong scaling, we are very close to ideal even at the maximum allowable cores (6400; 50 nodes). Our weak scaling tests were only done in units of whole nodes (minimum 128 cores). Figure 1 (right) shows that our code is >90% efficient (dotted line) even at 50 nodes (6400 cores). Our weak scaling test results are given in Table 1.

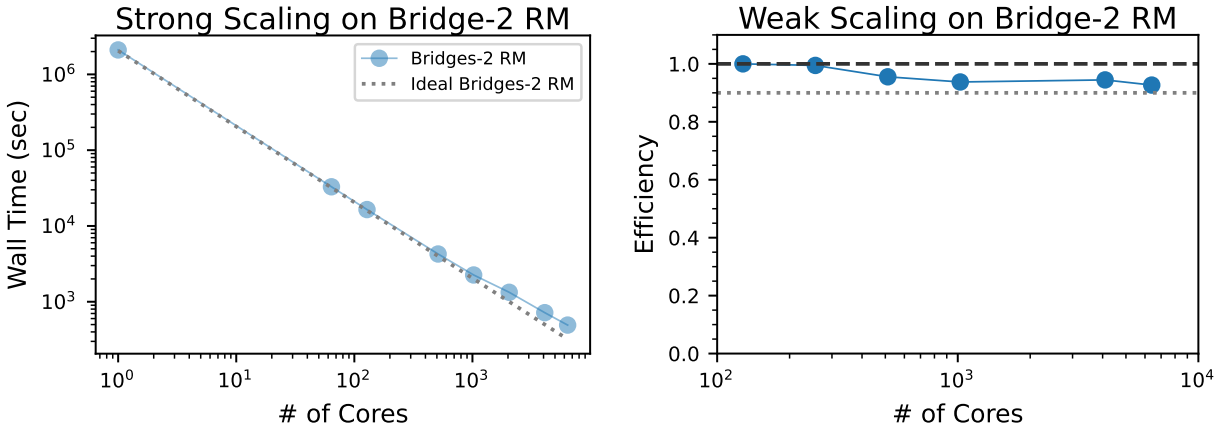


Figure 1 – Strong scaling on BRIDGES-2 Regular Memory nodes. *Left:* Strong scaling on BRIDGES-2 Regular Memory. Our scaling is very close to ideal, even for the maximum allowable cores (6400; 50 nodes). *Right:* Weak scaling on BRIDGES-2 Regular Memory nodes. We only show points in units of whole nodes (factors of 128 nodes). Even for the maximum allowable cores (6400; 50 nodes) our efficiency is higher than 90% (dotted line).

# of Cores	# of Nodes	# of Sources fit	Efficiency
128	1	128	1.0
256	2	256	0.99
512	4	512	0.96
1024	8	1024	0.94
4096	32	4096	0.94
6400	50	6400	0.93

Table 1 – Weak Scaling Test Results