Uniform Forward-Modeling of Benchmark L-Type Dwarfs

SPENCER A. HURT,1 ZHOUJIAN ZHANG,2 AND MICHAEL C. LIU2

1Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA
2Institute for Astronomy, University of Hawai`i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

ABSTRACT

1. INTRODUCTION

Non-irradiated, self-luminous brown dwarfs (\( \sim 13 - 70 \) \( M_J \); Spiegel et al. 2011; Dupuy & Liu 2017) show many properties that overlap with those of directly imaged massive planets, such as temperature or surface gravity. However, atmospheric characterization through emission (e.g. Line et al. 2016; Samland et al. 2017), transmission (e.g. Deming et al. 2013; Kreidberg et al. 2014), and reflection spectroscopy (e.g. Parmentier et al. 2016; MacDonald et al. 2018) is often challenging for exoplanets given their low star-planet contrasts as well as data hindered by inadequate wavelength coverage. In contrast, brown dwarfs are readily observed through high-quality emission spectroscopy and can help improve our understanding of various physical and chemical processes similar to those taking place in exoplanet atmospheres. Ultimately, studying brown dwarf atmospheres may help reveal how planetary-mass objects form and evolve (Burrows et al. 2001; Fortney et al. 2008; Marley & Robinson 2015).

Atmospheric studies of brown dwarfs have traditionally relied on forward-modeling (e.g. Saumon et al. 2000; Geballe et al. 2001; Burgasser et al. 2006; Cushing et al. 2008; Stephens et al. 2009; Cushing et al. 2011; Liu et al. 2011; Manjavacas et al. 2016; Bonnefoy et al. 2014, 2018), where the observed spectrum is compared to a grid of precomputed theoretical model spectra based on several free parameters (typically effective temperature, surface gravity, and sometimes metallicity) and self-consistent assumptions (often including radiative-convective equilibrium, equilibrium and disequilibrium chemistry, and cloud properties). Forward-modeling analysis conventionally compares an observed spectrum with the synthetic spectrum at each model grid point and derives the best-fit parameters with least-squares minimization. These analyses tend to assume flux measurements follow Gaussian statistics and that residuals between different wavelengths are uncorrelated, resulting in a covariance matrix only with inverse weights (typically flux uncertainties squared) along the diagonal. Additionally, because the model grids are so coarse, this method requires interpolation between models to estimate parameters between the grid points. Typically, this is accomplished with linear interpolation.

This conventional method of forward-modeling fails to account for several large sources of uncertainty. First, linear interpolation between model grid points can introduce large parameter uncertainties yet only returns a single model spectrum with no uncertainties. This often leads to underestimated error bars for the inferred physical parameters and may bias the posteriors towards model grid points (e.g. Cottaar et al. 2014; Czekala et al. 2015; Zhang et al. 2020b). Second, because spectrographs over-sample the instrument’s resolution, residuals between the data and model at different wavelengths are correlated. This is further exacerbated by the fact that the number of free parameters in a grid model is too small to fully describe the data. Instead, grid models are simplified through standard assumptions about processes such as cloud sedimentation or vertical mixing. While excluding these uncertainties leads to artificially narrow posteriors for the derived physical parameters, many analyses attempt to account for this by adopting fractions of the model grid spacing as the final parameter uncertainties (e.g. Leggett et al. 2007; Cushing et al. 2008; Stephens et al. 2009; Zhang et al. 2020b).

Czekala et al. (2015) developed a Bayesian framework for forward-modeling that accounts for these often neglected uncertainties, implemented through the Python package Starfish. Rather than linearly interpolating between model grid points, Starfish uses a spectral emulator that generates a probability distribution of model spectra for a given set of parameters, allowing interpolation uncertainties to be propagated into the inferred parameters. Additionally, Starfish includes a covariance matrix with off-diagonal components that account for correlated residuals due to instrumental properties and model imperfections. Zhang et al. (2020a) applied this Bayesian framework to a sample of benchmark ultra-cool T dwarfs, testing cloudless Sonora-Bobcat atmospheric models against evolutionary models. In a companion paper, Zhang et al. (2021) extend this analysis to a sample of 55 ultra-cool T dwarfs, providing a systematic examination of ultra-cool model atmospheres. Both studies found that compared to traditional forward-modeling analy-
ses, the Starfish methodology returns more realistic posteriors.

We continue this work by applying the Starfish methodology to a large sample of benchmark ultra-cool L dwarfs using the BT-Settl grid models (Allard et al. 2012) at solar metallicity. Because many L dwarfs share a parameter space with young, dusty planets, these objects offer the opportunity to test cloudy models such as BT-Settl for planetary-mass objects (e.g. Marois et al. 2008; Bowler et al. 2017; Chauvin et al. 2017).

In Section 2, we summarize the targets in our sample and our observations. Next, we review the BT-Settl grid model in Section 3. We then construct and validate our forward-modeling framework for the BT-Settl atmospheric models, apply this framework to infer the properties of our sample, and investigate the systematics in these model atmospheres in Section 4. In Section 5, we additionally derive the physical parameters for each benchmark using evolutionary models and then compare these properties to the results from our spectral fitting in Section 6. Finally, we summarize our results and future work in Section 7.

2. BENCHMARKS AND OBSERVATIONS

2.1. PSO J318.5338-22.8603

PSO J318.5338-22.8603 (PSO J318.5-22) is an extremely red late-L dwarf (L7 ± 1) first discovered by Liu et al. (2013). The free-floating object is located at a distance of 22.2 ± 0.8 pc. Many parallels have been drawn between PSO J318.5-22 and young dusty planets such as those imaged around HR 8799 (Marois et al. 2008) with overlapping colors, magnitudes, spectra, luminosities, and masses.

As a likely member of the β Pictoris moving group (BPMG; Liu et al. 2013; Allers et al. 2016), PSO J318.5-22’s age is comparatively well-constrained. While literature spanning over the past two decades gives a wide range of BPMG ages (∼10 – 40 Myr; Mamajek & Bell 2014), more recent estimates have consistently given values near ∼20 Myr. Combining lithium depletion boundary (LDB) ages from Binks & Jeffries (2014) and Malo et al. (2014) with isochrone analyses of pre-main sequence F- and G-type stars, Mamajek & Bell (2014) find a median age of 23 ± 3 Myr. Also using LDB, Messina et al. (2016) obtain an age of 25 ± 3 Myr.

And using Gaia Data Release 2 astrometry, Miret-Roig et al. (2020) find a dynamical age of 18.5 ± 2.0 Myr. Given that all of these studies find somewhat consistent ages, we adopt the age estimate of 23/24 Myr from Mamajek & Bell (2014). For simplicity, the L dwarf’s membership in the BPMG also implies a nearly-solar metallicity.

2.2. 2MASS J02495639-0557352 c

2MASS J02495639-0557352 c (2MASS J0249-0557 c) is an L2 ± 1 dwarf orbiting a tight ultra-cool binary at a distance of 1950 ± 200 AU (40") discovered by Dupuy et al. (2018). Parallax measurements place the bound triple system at 48.9 ± 4.4 pc away. 2MASS J0249-0557 c has very similar properties to the planet β Pictoris b (Lagrange et al. 2009), with curiously similar masses, magnitudes, and spectral types. Interestingly, both L dwarfs are likely members of the BPMG. Consequently, we assume the same age distribution (23 ± 3 Myr) and solar metallicity for 2MASS J0249-0557 c as we did for PSO J318.5-22 in Section 2.1.

2.3. Others (Not Yet Determined)

2.4. Observations

The near-infrared spectra for each benchmark were measured using the SpeX Spectrograph (Rayner et al. 2003), mounted on the 3 m NASA Infrared Telescope Facility (IRTF). Both spectra were obtained through the SpeX Prism Library (Burgasser 2014), with PSO J318.5-22 being observed by Liu et al. (2013) and 2MASS J0249-0557 c by Dupuy et al. (2018). These observations were made either with the SpeX 0.8” × 15” slit or the 0.5” × 15” slit and have spectral resolutions of R ≈ 50 – 160 and R ≈ 80 – 250, respectively. Before analysis, spectra are flux-calibrated using the object’s $F_{\text{MKS}}$ magnitudes, the WFCAM H-band filter response (Hewett et al. 2006), and the zero-point flux (Lawrence et al. 2007). The photometric and astrometric properties of each target are given in Table 1.

3. THE BT SETTL MODELS

3.1. Model Description

Throughout this work, we model our spectra with a synthetic grid produced by BT-Settl (Allard et al. 2012), which is distinguished by its treatment of clouds and dust in the atmospheres of cool objects. Using the timescales of condensation, coalescence, mixing, and gravitational settling for 55 grain types, BT-Settl calculates the abundance and size distribution for solids throughout the atmosphere (details of the solids and elements are given in Rajpurohit et al. 2018). The opacities in the spectra are then calculated line-by-line and the overall radiative transfer is handled using the PHOENIX code (Allard et al. 2001), accounting for convection using mixing-length theory and working at hydrostatic and chemical equilibrium. BT-Settl additionally accounts for non-equilibrium chemistry between molecules including CO, CH$_4$, CO$_2$, N$_2$, and NH$_3$. The latest BT-Settl grid (CIFIST2011/2015), which we focus on in this paper, follows modeling techniques introduced by Baraffe et al. (2015), including updated molecular line lists and convective mixing lengths calibrated using multi-dimensional radiative hydrodynamics. Additionally, each grid point is calculated at solar metallicity (Z = 0 dex), as defined by Caffau et al. (2011). Our final grid spans temperatures ($T_{\text{eff}}$) of 1200 – 3000 K and surface gravities (log g) of 3.5 – 5.5 dex.
Figure 1. Comparisons of the cloudy BT-Settl (CIFIST2011/2015) model spectra with varying effective temperature ($T_{\text{eff}}$) and surface gravities ($\log g$). The upper panel compares spectra with effective temperatures of 1200 K (blue), 1900 K (green), and 2600 K (red), all of which have surface gravities of 4.5 dex. The lower panel compares spectra with surface gravities of 3.5 dex (blue), 4.5 dex (green), and 5.5 dex (red), each with a fixed temperature of 1900 K. All model spectra have been degraded to the wavelength-dependent resolution of the 0.5" slit of SpeX prism ($R \approx 80 – 250$) and normalized by their peak K-band fluxes.

3.2. Synthetic Spectra

We investigate the effects that physical parameters have on the resulting BT-Settl spectra by comparing two sets of synthetic spectra with effective temperatures and surface gravities spanning our grid (see Figure 1).

4. ATMOSPHERIC MODEL ANALYSIS

4.1. Traditional Forward-Modeling Analysis

We begin our traditional forward-modeling analysis by downgrading the resolution of each synthetic spectrum in the BT-Settl grid using a Gaussian kernel corresponding to either the 0.5" or 0.8" SpeX slit sizes. This convolution process accounts for the wavelength-dependent resolution of the SpeX prism observing mode (Rayner et al. 2003), returning two grids of theoretical spectra that match the different resolutions of our data. We also trim the model spectra to wavelengths of 0.8 – 2.4 $\mu$m, matching the region of interest in our data.

To explore the full parameter space, we create models for both SpeX slit sizes that take $\log T_{\text{eff}}$ and $\log g$ as inputs, and, by linearly interpolating between the degraded spectra at nearby grid points, returns a predicted flux at each wave-
Figure 2. BT-Settl (CIFIST2011/2015) grid models corresponding to gravities of 2.5 (top left), 3.5 (top right), 4.5 (bottom left), and 5.5 (bottom right) dex. The color of each spectrum corresponds to its effective temperature, which falls between 1200 K (the lower boundary of the grid) and 3000 K. All spectra have been degraded to the wavelength-dependent resolution of the 0.5" slit of SpeX prism ($R \approx 80 – 250$).

We include all 446 BT-Settl grid models in our interpolators, spanning $T_{\text{eff}}$ of [1200, 7000] K with a spacing of 50 K or 100 K and $\log g$ of [2.5, 5.5] dex with a spacing of 0.5 dex. Additionally, our model rescales interpolated spectra using the logarithmic solid angle $\log \Omega = \log \left(\frac{R}{d} \right)^2$, where $R$ is the object's radius and $d$ is its distance. Given the resolution of our data, we find that rotational broadening and Doppler shifting can be left out of our models without biasing our derived atmospheric properties.

We run Markov chain Monte Carlo (MCMC) processes to fit our models and derive the atmospheric properties for each object in our sample. Uninformative priors are placed on $\log T_{\text{eff}}$ and $\log g$ such that they are constrained to the values spanned by the BT-Settl grid. Our final log-likelihood function is

$$L = -\frac{1}{2} \sum_{\lambda} \frac{(f^l(\lambda, \Theta) - f^o(\lambda))^2}{e(\lambda)^2 + \sigma^2} - \ln \frac{1}{2\pi} \left( e(\lambda)^2 + \sigma^2 \right),$$

where $f^l(\lambda, \Theta)$, $f^o(\lambda)$, and $e(\lambda)$ correspond to the model interpolated flux, observed flux, and observed flux uncertainty at wavelength $\lambda$, respectively. $\Theta = \{\log T_{\text{eff}}, \log g, \log \Omega\}$ denotes the parameter set while $\sigma$ is a fitted parameter meant to account for unknown sources of noise or significant residuals between the data and the best-fit models, commonly referred to as jitter. Note that the second term in our log-likelihood function acts as a penalty for large $\sigma$, ensuring that the jitter does not become too large and account for all variability in our data.

Using emcee (Foreman-Mackey et al. 2013), we fit our object’s 0.8 – 2.4 $\mu$m spectra with 15 walkers, beginning with a 5000 step burn-in period. We continue to run the MCMC process until the number of iterations exceeds 50 times the autocorrelation time for each parameter (typically after $\sim 10^4$ steps), confirming that the chain has converged.

Uncertainties in the photometry used to flux-calibrate our spectra contribute to systematic errors in $\log \Omega$. Following Zhang et al. (2020a) and Zhang et al. (2021), we adopt a systematic uncertainty of $0.4\sigma_{\text{MKO}}$ in the inferred $\log \Omega$. We draw errors for each sample in our chain from a normal distribution centered at zero with a standard deviation equal to...
the adopted systematic uncertainty and add those errors to the corresponding chain values, effectively inflating the uncertainties in log Ω.

From this modified chain, we then derive several new properties. First, we assign a distance drawn from a normal distribution centered on the median value with a standard deviation equal to the uncertainty in the distance to each sample in the chain. Using these distances and the posterior for log Ω, we then derive the radius posterior for each object. From the radii and surface gravities, we then calculate the posterior for each object’s mass. We additionally divide each σ value in the chain by the average flux uncertainty, providing a reference point for how large the jitter is and allowing a quick evaluation of how well the model fits to the data.

Throughout our traditional forward-modeling analysis, we find that the starting positions for our MCMC processes have significant effects on the resulting posteriors, particularly for spectra where the best-fitting models show large residuals. Simple algorithms—such as starting the MCMC at whichever grid point has the lowest χ² statistic—tend to get walkers stuck in small regions with very high surface gravities and lower than expected temperatures, leading to large overestimates in radius and mass. The χ² values for these objects mapped over the grid points do not show a well-constrained region of low χ² values, indicating BT-Settl is unable to adequately fit the data, a behavior consistently observed for earlier type objects. As an example, the χ² map for 2MASS J04433761+0002051 is shown in Figure 3. To account for this in our fitting, we begin our MCMC processes near the weighted average T eff and log g, where the weights correspond to the p-value for each grid point. This allows the walkers to avoid immediately getting stuck in a region of high probability and explore more of the allowed parameter space.

4.2. Starfish Forward-Modeling Analysis

We begin our Starfish analysis by creating the spectral emulators that allow us to construct model spectra at arbitrary effective temperatures and surface gravities, one for each SpeX slit size. This process closely follows those in Zhang et al. (2020a) and Zhang et al. (2021), however there are several key differences. First, we are unable to build a spectral emulator that covers the full range of BT-Settl parameters relevant to ultra-cool dwarfs. As discussed previously, including models with T eff below 1900 K introduces significant systematics into our emulator reconstruction. Additionally, we find that including low-surface gravity models below 3.5 dex causes additional systematics, although this region of the grid is less relevant to most objects in our sample. Consequently, we train our spectral emulators using all 60 grid points spanning T eff of [1900, 3000] K with a spacing of 100 K and log g of [3.5, 5.5] dex with a spacing of 0.5 dex.
Figure 4. The mean spectrum ($\xi_\mu$), standard deviation spectrum ($\xi_\sigma$), and 11 eigenspectra ($\xi_i$) of the 60 BT-Settl models used to train the 0.5" SpeX slit Starfish spectral emulator. All spectra have been scaled using the same factor and are offset from one another by constants.
Figure 5. Each panel contains a BT-Settl model used to train the 0.5” SpeX slit spectral emulator, marked with a black line, along with the distribution of spectra generated by the spectral emulator at the same grid point given in red. The corresponding $T_{\text{eff}}$ and $\log g$ values are given in the upper right corner of the panel. Both models are normalized using the J-band peak of the original BT-Settl model. Residuals between the original and reconstructed models are shown on the same scale at the bottom of each panel. Generally, the spectral emulator is able to reconstruct the original model spectra well. However, residuals are consistently present near 1 µm, occasionally exceeding 1% of the peak J-band flux. While not ideal, we quantify the effects of such systematics later on.
ACKNOWLEDGMENTS

The authors thank Aaron Do, Michael Bottom, Robert Jedecke, Benjamin Shappee, and Xudong Sun for their helpful discussions and insight. S.A.H. acknowledges support from the Research Experience for Undergraduates program at the University of Hawaii Institute for Astronomy funded through NSF grant 2050710 and thanks University of Hawaii for their hospitality. Finally, the authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

APPENDIX

This appendix details components from the project that normally would not be included in a paper. The Python package used for our forward-modeling analysis, Starfish, was recently upgraded from version 0.2 to version 0.3. This most recent version has not yet been used in a publication and is structured very differently from v0.2, requiring additional development and testing. This included building instrument profiles that account for the wavelength-dependent resolution of SpeX prism data, incorporating an interface for the BT-Settl grid models, and fixing the normalization of emulator models so that the scaling factor between models and data, $\Omega$, can be physically interpreted as

\[ \Omega = \frac{R^2}{d}, \]

where $R$ is the object's radius and $d$ is its distance. There also appeared to be a problem in the code that improperly scales the off-diagonal covariance matrix, causing the inferred posteriors to be flatter than expected. Further investigation revealed that these were likely numerical and computational problems and were resolved by introducing a scaling factor to the data and models when calculating the log-probability.

No additional writing has been added to this report since the one due ~3 weeks ago. As mentioned in several reports from the REU directors, this is a large project that has pushed us for time, and rather than focus on writing, we felt it was better to focus on our analysis and the final presentations. We will, however, submit a final paper to AAS journals and plan to share a draft with the REU directors sometime in the upcoming months. In this appendix, we present most of the figures produced in our current analysis and a brief discussion of our results.

We first compare results from the Starfish and traditional fitting methodologies in Figure 8. The two parameters are generally well-correlated, take several exceptions. Objects with very different Starfish and traditional parameters are mostly those with complex $\chi^2$ surfaces that are poorly captured by conventional MCMC processes. Future work will likely account for these complex surfaces using more sophisticated sampling techniques such as nested sampling.

We recently finished running our forward-modeling and evolutionary model fitting for 60 of the objects in our sample of YMG members and candidates. As can be seen in Figure 7, BT-Settl is able to replicate spectra for later-type, dust-rich and earlier-type, dust-poor objects. This is somewhat surprising, given that dusty objects are expected to be the most difficult to model. Patterns in the residuals emerge for objects located near the M/L transition boundary; specifically, BT-Settl consistently underestimates the flux in the H-band and K-band. We are currently considering what the causes of this discrepancy could be, such as inaccuracies in the line lists used to calculate the BT-Settl spectra or systematics in the modeling techniques used.

In our MCMC fits, we include error inflation terms to account for systematic noise in our modeling, similar to jitter terms used to account for stellar noise when modeling radial velocities. We reparameterize these error inflation terms as percentages of the peak flux in each observed spectrum ($\epsilon_{\text{max}}$). When comparing $\epsilon_{\text{max}}$ to spectral type in Figure 9, we see a peak near the M/L transition boundary, again confirming that the BT-Settl models systematically underperform for objects in this regime.

We also find that BT-Settl struggles to replicate low surface gravity features for many objects. For example, rather than matching the flat K-band shape seen in several of our spectra, BT-Settl returns a triangular shape indicative of high surface
This suggests that BT-Settl consistently overestimates the surface gravity, which is confirmed by a cluster of log g values between 5.0 dex and and 5.5 dex (see Figure 9 and Figure 11).

We calculate the evolutionary model parameters for each object from their ages and bolometric luminosities using the rejection sampling procedure in Dupuy & Liu (2017). Given that evolutionary masses and radii are consistent with dynamical masses and transit radii, these parameters are considered reliable benchmarks against which we can compare the BT-Settl parameters. In Figure 10, we plot the BT-Settl model spectra corresponding to the evolutionary model parameters. Very few of the models match the data well, indicating discrepancies between the BT-Settl and evolutionary parameters. In most cases, the evolutionary model parameters over-estimate the flux in the H-band and K-band, contrary to the parameters derived from BT-Settl. Additionally, the evolutionary model parameters return significant residuals in the J-band. For later-type objects, the models rarely overlap with the data, indicating that BT-Settl and the evolutionary models disagree on radius for these objects.

In Figure 11, we directly compare the parameters from BT-Settl to the evolutionary parameters. The effective temperatures show some correlation, however BT-Settl consistently overestimates the $T_{\text{eff}}$ for low-temperature (< 1500 K) objects. For higher, temperature objects, BT-Settl tends to underestimate the $T_{\text{eff}}$. Additionally, the distribution of BT-Settl temperatures shows a large gap centered at 2100 K that is not present in the evolutionary $T_{\text{eff}}$. No strong correlation is visible for the surface gravities, with the majority of BT-Settl log g clustered between 5.0 and 5.5 dex, confirming that BT-Settl poorly fits low surface gravity features.

We also search for relationships between spectral type and temperature. Given that spectral types are determined using spectral features and not physical properties, it is unclear if there is a correlation between the two. In Figure 12, we find that the evolutionary derived temperatures vaguely follow the relationship previously found in Filippazzo et al. (2015). However, the $T_{\text{eff}}$ for L dwarfs derived using BT-Settl do not show much variation. Additionally, the gap in BT-Settl temperatures centered at 2100 K appears near the M/L transition boundary, suggesting BT-Settl’s poor modeling of these objects is related.

We plan to apply this analysis to 30 additional ultracool dwarfs located in the Pleiades, expanding our sample to 90 objects. As mentioned previously, we will also test different sampling techniques to account for poorly-behaved $\chi^2$ surfaces, returning more representative posteriors for each object. And lastly, we will investigate what could cause the discrepancies between BT-Settl, the evolutionary models, and the data and how they might be resolved in future atmospheric models.
Figure 6. Comparison of the parameters obtained using the Starfish and traditional methods of spectral fitting for objects where both fits were possible. Blue error bars correspond to uncertainties in the traditional fitting while red corresponds to the uncertainties in the Starfish fitting.
Figure 7. Spectra plotted alongside 100 BT-Settl models drawn from our MCMC fits for all 60 objects in our sample. Black lines correspond to data, blue lines represent models from the traditional fitting method, and red lines give models fit using Starfish. Dust-rich, late-type objects and dust-poor, early-type objects are best modeled by BT-Settl while the H-band and K-band fluxes of objects near the M/L transition boundary are consistently underestimated.
Figure 8. Residuals between our data and 100 models drawn from our MCMC fits are shown in black for all 60 objects in our sample. The median 1σ, 2σ, and 3σ dispersion drawn from the covariance matrix for each object are shown as blue shadows for traditional fits and as red shadows for Starfish fits. The H-band and K-band are most prominent for objects near the L/M transition boundary.
Figure 9. Normalized error-inflation terms, expressed as $\epsilon_{\text{max}}$, from Starfish (red) and traditional (blue) BT-Settl modeling compared to spectral type (left), $T_{\text{eff}}$ retrieved from spectral fitting (middle), and $\log g$ obtained from spectral fitting (right). In the left-most panel, $\epsilon_{\text{max}}$ peaks near the M/L transition boundary, confirming that BT-Settl systematically produces poor fits for these objects.
Figure 10. Spectra plotted against 100 BT-Settl models calculated using parameters drawn from the evolutionary model posteriors for each object in our sample. Black lines correspond to data, yellow lines represent models calculated using the Bonnefoy et al. (2014) models, and green lines give models calculated using the Saumon & Marley (2008) models. Few models are consistent with the data, suggesting significant discrepancies between the evolutionary and BT-Settl models. For objects near the M/L transition boundary, the evolutionary parameters return model spectra that consistently over-estimate the flux in the H-band and K-band. Later-type L dwarf spectra are consistently on smaller scales than the model spectra, indicating a mismatch between the BT-Settl and evolutionary model parameters.
Figure 11. Comparison of the parameters obtained by fitting BT-Settl models and the evolutionary models. Along the y-axis, blue error bars correspond to uncertainties in the traditional fitting while red corresponds to the uncertainties in the Starfish fitting. Along the x-axis, green error bars represent uncertainties from the Saumon & Marley (2008) models while the yellow error bars give uncertainties from the Baraffe et al. (2015) models.
Figure 12. Effective temperatures derived via BT-Settl (left) and effective temperatures derived via evolutionary models (right) versus spectral type. No apparent trend is visible for the BT-Settl $T_{\text{eff}}$ for L dwarfs, however there is a large gap in the temperatures near the M/L transition boundary, where the BT-Settl models show consistently poor fits. $T_{\text{eff}}$ derived using the evolutionary models vaguely follow the relationship found in Filippazzo et al. (2015).

