EFFECT OF WATER IN COOL ATMOSPHERES.

MASTER THESIS
Written by *David John Willumsen*

Supervised by
Uffe Gråe Jørgensen

UNIVERSITY OF COPENHAGEN

UNIVERSITY OF
COPENHAGEN

Name of Institute:	University of Copenhagen
Name of Department:	Niels Bohr Institute
Author(s):	David John Willumse (jbq725)
Email:	jbq725@alumni.ku.dk
Title and subtitle:	Effect of Water in Cool Atmospheres.
Supervisor(s):	Uffe Gråe Jørgensen
Handed in:	February 15, 2022
Defended:	March 7, 2022

Name: David John Willumsen

Signature ______________________________

Date: 2.15.2022

Abstract

Water is one of the most abundant species that have been identified in observed exoplanets. Yet, there are several existing theoretical line lists that, when applied to model spectra, give quite different interpretations of the observed spectra. The project concerns comparing results from the different line lists and estimating which are in best agreement with observations. The compared line lists, SCAN, BT2, POKAZATEL, and HITEMP are implemented in cool atmospheric models using MARCS to compute synthetic spectra. The spectra are compared to medium-resolution spectra of M-dwarfs from the SpeX Prism Library and an SWS-ISO spectrum of M-giant SV Peg. All line lists show good fits with the giant spectrum, while the dwarf synthetic spectra from BT2, POKAZATEL, and HITEMP models show long-wavelength and short-wavelength regions, that can not fit the observed spectra simultaneously. I discuss the possible explanations for their inability to reproduce these spectra, and what this means for the use of line lists in stellar modeling.

Contents

1 Introduction

The answer to the question of whether there can be life on planets outside our Solar System is no longer far out of reach. Although, for alien life to sustain itself for long enough that we can hope to observe it, a great number of circumstances need to be in place. Our capacity for discovering terrestrial exoplanets is increasing, and it is becoming more apparent that rocky planets (of radius between $0.5R_{\oplus} - 1.5R_{\oplus}$), within the habitable zone, are abundant throughout the galaxy. Various definitions are used to describe such a habitable zone (HZ), but in general, it is defined as a region around a star, where a rocky planet with an atmosphere of Earth-like (H_2O, CO_2, N_2) composition, can have liquid surface water. When adding the many effects that can influence the exact parameters of the HZ, and taking into account the time that a planet has spent, or will spend, inside the HZ during its star's lifetime, it becomes tricky to generalize the HZ. Kupparapu et al. [2021] define a conservative habitable zone with inner and outer limits determined by greenhouse effects at different solar fluxes. Beyond the inner limit, the oceans would vaporize from the high temperature. Beyond the outer limit, maximum heating from the greenhouse effect would not be sufficient in keeping liquid water on the surface. Bryson et al. [2021], with this definition, estimate a frequency of habitable zone rocky exoplanets around a star of temperature 4800 K - 6300 K, $\eta_{\oplus} = 0.37^{+0.48}_{0.21}$. This would place the nearest one only ≈6 pc away from us. One crucial missing part is our characterization of exoplanets' atmospheres. With the advent of telescopes PLATO, E-ELT, and the recently launched JWST, it will be possible to obtain spectra that can be used to characterize atmospheres. With a limited resolution and a broad span of possible compositions with each their contribution to the shape of the spectrum and the structure of the atmosphere, we need something to compare that data to. This warrants the modification, and use of, comprehensive atmospheric modeling programs in the low temperature, high opacity regime of cool substellar objects and, eventually, exoplanets. It is tempting to base such a model on the Earth. However, this would be a mistake, as we would base the model on several fixed and observed parameters, not allowing for a self-consistent physical solution. The currently most common method for chemical analysis of exoplanets is transmission spectroscopy, where planetary transits, observations of a planet passing into the line-of-sight to the star in their orbit, are recorded in different wavelengths and compared. This is useful since the transit depths will differ in each wavelength, as the opacity is wavelength-dependent. The MARCS program is a self-consistent stellar atmosphere modeling program written in Fortran. Its creation, and a detailed description, are given in Gustafsson et al. [2008]. MARCS allows for self-consistent atmosphere models, and calculating spectra to compare with observed spectra. Over the years, features and updates have been added, along with the possibility to drastically lower the model temperatures and the accompanying changes to the chemistry and physics.

When generating synthetic spectra for low-mass stars and exoplanets, it is important to consider the behavior of water vapor. This includes determining the line absorption of water at a wide range of temperatures, and its impact on the atmospheric structure. This project compares lists of line absorption parameters of water, both theoretically and empirically based, in the context of fitting cool dwarfs. This comparison is done to highlight the complexities of modeling the behavior of water, and potentially, determining the best lists to use for accurate modeling in different contexts. Specifically, I use modern available line lists of the main isotopologue, $^{1}H_2{}^{16}O$, in cool MARCS atmospheric models to generate spectra and compare these to three late-type M-dwarfs, and one M-giant. A similar published comparison has been carried out by Jones et al. [2003] for older popular line lists. Most other comparisons are concerned with fitting to observations of laboratory experiments, such as Rothman et al. [2010], Tashkun [2006]. Modeling becomes important where phenomena such as element depletion and photochemistry complicate spectroscopic observations [Juncher et al., 2017]. Water is an asymmetric triatomar molecule with a complicated vibrational-rotational spectrum compared to other triatomar molecules [Barber et al., 2006]. This causes the list of frequencies at which water can absorb light to be vast, while its abundance causes a veil of those lines to dominate the large wavelength regions of the spectra of cool bodies such as late-type M-dwarfs. This makes correct modeling of water a very important sub-goal in accurately generating synthetic spectra.

This work will not go into details of quantum chemistry that went into line list construction (outlined in Jørgensen et al. [2001]), but aims to compare the performance of line lists based on their method of overcoming issues related to line list accuracy and completeness. The line lists in question are widely used and have not previously been directly compared in their ability to reproduce stellar spectra in available publications, while they have been compared in their ability to reproduce laboratory results. Laboratory tests are not sufficient for our purposes, since care has to be taken to the very weak lines that govern the very energetic water transitions, as those are very important for correct thermodynamics in upper atmospheres. Line lists can differ in this regard since the approach in the construction of line lists varies. Whereas one of our lists is based on using as accurate calculation as possible, others adjust line positions or band intensities to laboratory data. Some are limited by the computational power or data available at the time of construction. Different limits are set on the maximum energies included, and how to assign these highly energetic transitions, when the quantum states involved do not adhere to the theory used to categorize them. The best approach for line list construction is hard to determine, as there are quite large limitations on which data can be obtained in the laboratory, while calculations based on semi-classical quantum number (QN) assignment breaks down in the high energies. This means that any line list is bound to be flawed. Section 2.7 gives a quick overview of the problems with QN-assignment. The line lists compared in this work are SCAN [Jørgensen, 1995], HITEMP2010 [Rothman et al., 2010], BT2 [Barber et al., 2006] and POKAZATEL [Polyansky et al., 2018], all described in more detail in section 2.

Cloud formation appears in atmospheres at temperatures $T_{eff} < 2700$ K and becomes increasingly significant at lower temperatures. Clouds change the structure and spectrum of the atmosphere and are a crucial part of adapting MARCS to even cooler stars and exoplanets. Juncher et al. [2017] have taken the first steps towards adapting the MARCS code to exoplanets, by implementing a method of merging MARCS with a cloud formation program.

I present a comparison of BT2, HITEMP2010, SCAN, and POKAZATEL in their ability to reproduce the spectra of three cool M-dwarfs observed by 2MASS Skrutskie et al. [2006 February] and the cool giant, SV Peg [Aringer et al., 2002]. Primarily M-dwarfs were chosen around the temperature limit of cloud formation to showcase the application of MARCS in cooler atmospheres, while these spectra are dominated by water in the near-infrared. I omit cloud formation, as it is not necessary for this context, nor strongly present until quite lower temperatures. This is not an oversight in the analysis, as clouds only become very significant at lower temperatures than examined here, while water clouds only appear at very low temperatures. I briefly discuss the calculation of the partition function for water and discuss its importance in the context of line list construction. I also show the immediate effects of temperature and surface gravity on the synthetic spectra, to argue for how these parameters are chosen in fitting. This work is, in part, a continuation of an unfinished article by Diana Juncher and U.G. Jørgensen from 2016. It was carried out under the supervision of Uffe Gråe Jørgensen, creator of the SCAN list and co-author of several contained articles about SCAN and MARCS. My version of MARCS, accompanying code and documentation, and the SCAN line list was provided by Jørgensen.

2 Theory

2.1 Atmosphere Structure

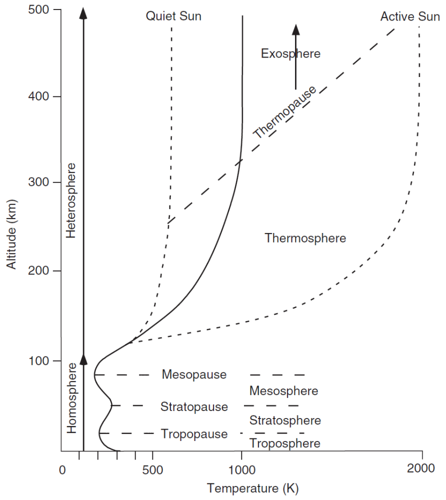

Figure 1: Vertical regions of the Earth's atmosphere. Source: Catling [2017]

A planet's atmosphere controls its climate through the greenhouse effect and its chemical composition. The temperature and pressure vary with altitude, and many physical processes take place at different layers in the atmosphere. Whereas Earth is unique among observed planets in having a stratosphere, caused by a *temperature inversion* due to the ozone layer, meaning that temperatures rise with altitude, the remaining four layers have analogs in most observed planetary atmosphere. In order of altitude, Earth's five layers are the *troposphere, stratosphere, mesosphere, thermosphere* and *exosphere* (see Figure 1). A troposphere is mainly characterized by convection, caused by radiation heating the surface, heating the air near it, which is lifted upwards, while air parcels move to lower pressures, where they expand and cool until they sink to the surface again. Where the convection stops, a *radiative-convective* boundary is found, whereafter radiation becomes the main source of upwards heat transport. This is because atmospheres become less opaque to IR-radiation somewhat below the *tropopause*, "losing" heat from the surface. Convection ceases after the rate of temperature decrease with altitude exceeds a critical value. The tropopause is generally found in planetary atmospheres around 0.1 bar regardless of composition [Catling, 2017]. a stratosphere, characterized by temperature inversion may exist above it if there is a strong presence of UV-absorbing molecules. Above the ozone layer on Earth, cooling resumes due to efficient radiation by CO_2 in the mesosphere until a new global minimum at the *mesopause*. At this minimum, pressure is so low that outwards radiation dominates the energy transport. This is generally found around 0.1 Pa in planetary atmospheres in the Solar System. Above it, in every planetary atmosphere, is an extremely thin thermosphere, fluctuating in temperature in response to extreme ultraviolet radiation. Photolysis and the very large mean free path, making collisions very rare, mean that an *ionosphere* of free ions forms. The ions and free atoms absorb highly energetic photons and convert them to kinetic energy, so the only heat transfer downwards happens when atoms and ions collide. For planets containing more CO_2 than Earth, radiative cooling in the upper atmosphere cools the thermosphere, and the upper atmosphere will generally be cool. The region outside the thermosphere is so thin that light atoms can escape. Since the thinning is gradual, there is no clear boundary to space. This general region is called the exosphere. There is a further distinction between regions where diffusive separation according to mass dominates atomic composition rather than turbulent mixing. On Earth this boundary is around ≈ 100 km, dividing the atmosphere into the *homosphere* below, and the *heterosphere,* above. The general structure of a stellar atmosphere is quite simple in comparison. It is determined by temperature, composition, and surface gravity. It is divided into the photosphere, the visible part, from where almost all our observations come from, a hotter chromosphere, dominated by emission lines, a somewhat elusive transition region, and finally, a very hot, thin region, called the corona. While the photosphere temperature decreases with altitude, the temperature starts to increase again, until it reaches the extremely hot corona. The reason for the temperature rise after the photosphere is not fully determined, but it is essentially caused by magnetic processes or small bursts of ultra-hot plasma happening constantly, and heating the corona [NASA, 2014][NASA, 2015]. Atmosphere modeling and much of our knowledge of stellar atmospheres are, however, limited to the photosphere [Gray, 2005].

Atmosphere structure is a term used for temperature change with altitude. We can examine this for example by studying the temperature as a function of pressure since pressure generally decreases with altitude due to hydrostatics. The dominant processes responsible for heat transport are radiation and convection [Catling, 2017].

2.2 Basics of Atmosphere Modeling

2.2.1 Assumptions

The atmosphere is assumes to be divided into homogeneous plane-parallel layers, which is an assumption based on the radial extension of the atmosphere being very small compared to the radius of the star. This means we do not obtain any insight into the actual homogeneity of the atmosphere. Figure 2 shows the structure of such a plane-parallel model. T and P are the temperature and pressure of each layer, θ is the angle of an incident light ray and the plane normal, and τ is the optical depth, equal to the number of mean free paths taken to travel through a layer. We also assume that the atmosphere is in a steady-state, avoiding all time-dependent phenomena. We assume hy-

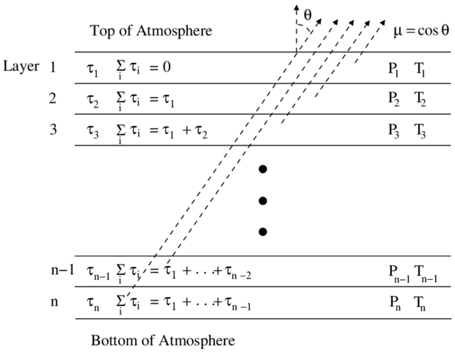

Figure 2: Structure of a plane-parallel atmosphere model. P and T are the temperature and pressure of each layer, and τ is the *optical depth*, a dimensionless measure of the opacity. Source: Hesman et al. [2005]

drostatic equilibrium overall. A possibly large effect ignored from this is magnetic forces on large scales. We assume that the atmosphere is flux-constant, meaning that all heat produced by the stellar body must leave the atmosphere in order to obey energy conservation [Mihalas, 1978].

2.2.2 Chemical Equilibrium

The assumption of local thermal equilibrium allows for calculating gas composition from minimization of the Gibbs free energy $G = U + PV - TS$. In order to calculate the abundances of molecules, the model program must calculate multiple chemical equilibria at once. First, we need an expression for the change in the Gibbs free energy, since energy is minimized when molecules reach a chemical equilibrium [Woitke et al., 2018]. By looking at the first law of thermodynamics in a closed system with possible change of composition and the definition of G:

$$dU = TdS - PdV + \sum_i \mu_i N_i, \qquad G = U + PV - TS. \tag{1}$$

Comparing the two, we can define dG as [Liliana Mammino, 2015]:

$$dG = VdP - SdT + \sum_i \mu_i dN_i, \quad T = T_0,\ P = P_0 \implies dG = \sum_i \mu_i dN_i. \tag{2}$$

$$\text{After integration:} \quad G = \sum_i \mu_i N_i. \tag{3}$$

In the above, G is the Gibbs free energy, V volume, S entropy, P pressure, μ_i the chemical potential of molecule N_i and T temperature. $T = T_0, P = P_0$ defines the isobaric, isothermal case. It follows that in the isothermal, isobaric case, the chemical potential can be interpreted as the Gibbs free energy per molecule. The number of molecules is not generally constant in a system undergoing constant chemical reactions. Using the matrix γ_{ij}, we can express the number of atoms N_j residing in molecules N_i [Woitke et al., 2018]:

$$N_j = \sum_i \gamma_{ij} N_i. \tag{4}$$

Equations (3) and (4) provide means for examining basic chemical reactions on the form:

$$c_1 X_1 + c_2 X_2 \rightleftharpoons k_1 Y_1 + k_2 Y_2. \tag{5}$$

To solve a system equilibrium, meaning obtaining a constant abundance of each molecule N_i, each equation of the form (4) and (5) must be solved while varying each molecules number N_i, to minimize the Gibbs free energy. The difference in G can be calculated from an equilibrium constant K_{eq} [Woitke et al., 2018], which depends on partial pressures of products and reactants weighed by their stoichiometric coefficients $c_{1,2}, k_{1,2}$. ΔG will then be calculated, favoring $\Delta G < 0$, until equilibrium is reached at $\Delta G = 0$. It should be noted that a real atmosphere deviates from this picture due to element depletion from cloud formation and photolysis in the upper atmosphere.

2.2.3 Net Energy Transport

Generally in physics, we assume conservation of energy. In the case of atmospheres, this is expressed through the total flux leaving and entering the atmosphere. This can happen through convection or radiation. An effective temperature T_{eff} is defined for the star, according to the Stefan-Boltzmann law [Gray, 2005]:

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(F_{rad}+F_{conv}\right)=0, \qquad F_{total}=\sigma T_{eff}^4, \tag{6}$$

where σ is the Stefan-Boltzmann constant. The picture is more complicated for a planet, which emits IR radiation, while it is heated from above by its parent star. This requires careful treatment in future atmospheric modeling, but the assumption remains, that an equivalent total flux F_{tot} must leave the atmosphere.

2.2.4 Hydrostatic Equilibrium

The balance between outward total pressure and the inward gravity of a body is described by hydrostatic equilibrium. In terms of the total pressure P, gravitational acceleration, g, and gas density ρ:

$$\frac{\mathrm{d}P}{\mathrm{d}r} - \rho g. \tag{7}$$

The total pressure, in theory, has contributions from radiation, convection, magnetic pressure, and gas pressure. The magnetic and turbulent contributions, however, are negligible in MARCS:

$$P_{tot}=P_{rad}+P_{gas}, \qquad P_{rad}=\frac{4}{3}\frac{\sigma}{c}T^4, \qquad P_{gas}\approx\frac{k_B T}{\mu m_u}\rho. \tag{8}$$

The gas pressure is the sum of partial pressures of every molecule species, atom, and ion. In equation (7), we express it in terms of the ideal gas law, which is valid for the low density of stellar atmospheres [Mihalas, 1978]. μ is the average particle mass, m_u is the atomic mass constant and c, σ, c are natural constants.

2.2.5 Radiative Transfer

Radiative transfer describes how emission, absorption, and scattering are treated for a light beam traveling through a medium. The time-independent equation for radiative transfer of a light beam traveling along a direction $\hat{s}$ is expressed in terms of wavelength-dependent radiation intensity I_λ, wavelength-dependent absorption coefficient κ_λ, density of the medium ρ, and a wavelength-dependent source function $S_\lambda = \frac{\text{emission}}{\text{absoprtion}}$:

$$\frac{\partial I_\lambda}{\partial s} = -\kappa_\lambda \rho \left(I_\lambda - S_\lambda\right). \tag{9}$$

Assuming local thermal equilibrium (LTE), which is common in atmospheric modeling, and negligible scattering, the source function can be expressed as a Planck function [Gray, 2005]:

$$S_\lambda \approx B_\lambda(T) = \frac{2hc^2}{\lambda^5}\frac{1}{\exp\left(hc/\lambda k_B T\right)-1}, \tag{10}$$

where λ is wavelength, T is temperature, h is Planck's constant, c is the speed of light in vacuum and k_B is Boltzmann's constant. For a plane-parallel model atmosphere with one-dimensional layers, the radiative transfer equation becomes [Mihalas, 1978]:

$$\mu \frac{\mathrm{d}I_\lambda}{\mathrm{d}\tau_\lambda} = I_\lambda - B_\lambda(T), \tag{11}$$

where $\mu = \cos\theta$, θ being the angle between an incoming light beam and the plane normal, and τ_λ is the monochromatic optical depth: $d\tau_\lambda = \kappa_\lambda \rho \cos\theta \, ds$.

2.2.6 Convection

The rate at which the temperature decreases with altitude is called the *lapse rate*, $\Gamma = -\frac{\mathrm{d}T}{\mathrm{d}z}$. If a parcel of air ascends or descends, such that it warms or cools without energy exchange with its surroundings through radiation or conduction, its temperature change is *adiabatic*. Convection occurs when the lapse rate of a convection cell (parcel of air that is hotter or cooler than its surroundings) exceeds that of the adiabatic lapse rate [Catling, 2017]:

$$\text{Convection:} \quad \Gamma = \left(\frac{\mathrm{d}T}{\mathrm{d}z}\right) > \Gamma_a = \left(\frac{\mathrm{d}T}{\mathrm{d}z}\right)_a . \tag{12}$$

In atmosphere modeling, convection is handled by defining a *mixing length*, $l = \alpha H$, where H is the atmosphere scale height, and α is a chosen mixing length parameter. The mixing length defines the average length of travel for a convection cell before it absorbs or releases energy to its surroundings. This energy flux depends on gravitational acceleration g, medium density, specific heat capacity, C_p temperature, and the temperature difference between the cell and the medium [Gray, 2005][Gustafsson et al., 2008]. Convection in atmosphere models causes the deeper layers of the star to lose heat more effectively in the convective region.

2.3 Cool Dwarf Spectra and the Importance of Water

Atmospheric modeling has reached a point where models constructed for stellar atmospheres can be adapted to substellar objects. This is also true for MARCS with the inclusion of a routine called GGchem [Woitke et al., 2018], calculating thermo-chemical equilibrium at each model layer, for each iteration, returning new equilibrium abundances, available to subsequent subroutines of MARCS relying on the molecular composition. Decreasing temperature causes more atoms to combine into molecules. Molecules have many transitions other than electronic, due to vibration and rotation. This means that cooler atmospheres will have an increased opacity per gram of material. When the opacity increases, the radiation pressure will cause the atmosphere to expand, lowering the gas pressure [Jørgensen and Johnson, 1992]. Because ultra-cool brown dwarfs have similar atmospheric physics to exoplanets [Cushing et al., 2011], modeling cool dwarfs atmospheres is a step towards modeling atmospheres for exoplanets, which will prove an invaluable tool when categorizing planets through future spectroscopic data, and will help in the search for habitable planets. While the strongest water absorption bands have the most obvious effect on the spectra, the weaker lines play a very important role, as these transitions have a large effect on the temperature-pressure profile.

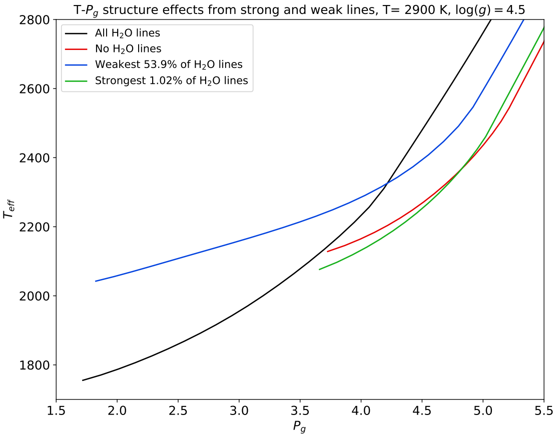

Figure 3: T-P structure of the upper atmosphere of MARCS computed atmospheres with the SCAN line lists, computed with different fractions of included H_2O lines.

The stronger bands have a higher transition probability, and cause higher opacity, meaning bigger "dips" in the spectrum. The weak lines represent high-energy transitions, occurring at high temperatures. They are needed to realistically model an atmosphere's temperature-pressure relationship, altering the structure of the atmosphere, by drastically lowering the gas pressure in the upper layers. This expansion is caused by the strong interaction of the gas with the energy being transported through the gas. Properly estimating and representing the weak lines is difficult. A laboratory experiment consisting of heating gas in a chamber, and measuring its absorption, will sample the sum of the strong lines' absorption. This means that experiments can certainly be used to test the agreement between line lists and empirical data to check the strong lines, however, almost no information about the weak lines will be present. This is because the weak lines constitute a very small fraction of the total integrated absorption coefficient. Stellar spectra are sensitive to the veil of weak lines and how well the computations hold up across broad spectral ranges. Thus, strong and weak lines must both be represented proportionally to their transitions' effect on real spectra. This is a challenge, and different solutions and opinions exist about how to best solve that problem. This is also what makes comparison with observed spectra an excellent test for line lists. The plot in Figure 3 shows the direct effect of the strong and weak lines on the $T - P_g$ structure. Four model atmospheres are computed, one without any water lines (red), one with the total SCAN line list (black), one with only the 1.02% strongest water lines (blue), and one with only the weakest 53.9% of water lines. The plot shows only the upper atmosphere, as the structures' shapes in the deep atmosphere follow the top right of the plot, and can thus be inferred. Not shown on the plot is that the blue and red curves extend slightly "deeper" into the high temperatures. The plot shows that the weakest 53.9% of water lines, corresponding to 0.073% of the total integrated absorption coefficient, cause a huge expansion of the atmosphere, lowering the gas pressure by 2 orders of magnitude, while causing heating in the upper atmosphere, and cooling below, while the 1.02% strongest, corresponding to 88.8% of the total integrated absorption coefficient, cause only a very moderate cooling compared to the "no water"-structure in the upper atmosphere, and smaller heating below. This means that an error of 400% in intensities of this portion of weak lines, would only cause a 0.29% deviation between laboratory and computed integrated absorption coefficients.

Brown dwarfs' spectra are dominated by broad and overlapping features from molecules and condensate, whose strengths depend on gas pressure, abundance, temperature, and available dust, which are related to surface gravity, metallicity, and T_{eff} [Manjavacas et al., 2014]. Due to differences between solar-type stars and very cool- to substellar objects, new spectral features have to be modeled in the spectra. Brown dwarfs have (progressively late) spectral classes M, L, T, and Y, the latter three being exclusive to substellar objects, while the M class can contain brown dwarfs, as long as their estimated mass is below the stellar limit of 65 M_{JUP}. The L class is characterized by a strong lithium line, which should otherwise not be present in a star older than 100 Myr. Their spectra are dominated by H_2O, FeH, and CO absorption in the NIR [Manjavacas et al., 2014]. The T class exhibits methane absorption at 2 μm, not observed in any main sequence stars, collision-induced absorption (CIA) due to H_2 and typically deep H_2O absorption bands around 1.3-1.5 μm [Manjavacas et al., 2014]. The Y class is defined by Cushing et al. [2011],Kirkpatrick et al. [2012], as a combination of absorption effects and very low temperature. They mention NH_3-bands as an indicator for this new class since it surpasses N_2 in abundance when $T \leqslant 700$ K, although, there is an overlap with strong H_2O bands, clashing in particular with water ice absorption centered at ≈ 1.5 μm. In cool enough brown dwarfs ($T \leqslant 350$ K), water clouds are expected to form, causing water absorption to dominate parts of the spectrum [Skemer et al., 2016]. The age of substellar objects can be inferred from the fact that they contract and cool with time. The formation and sedimentation of the iron and silicate grain dust clouds Below $T_{eff} \approx 2600K$ is also influenced by surface gravity. The formation will be more efficient in an extended cool gas, and convection and mixing should be more efficient [Manjavacas et al., 2014]. The SpeX late M-dwarf spectra that are used in this project are not expected to be substellar, so their spectra are expected to be dominated by absorption from TiO in 0.7-1.0 μm, and then by water, almost undisturbed, longward of 1.3 μm. Details about SpeX is found in section 3.5. In cool dwarfs, as with exoplanets, the total pressure of the hydrostatic equilibrium equation (eq. 7) is dominated by the gas pressure ($P_{tot} \approx P_{gas}$).

2.4 Line Lists

For our purposes, a *line* can be defined as a transition between two values of J (total angular momentum) of an atom or molecule. These J-values have corresponding degeneracy $g_n = 2J_n + 1$, and the transition has an intensity, or "strength", S, equivalent to the transition probability. The *Einstein coefficient* A_{ji} is a good, unambiguous measure of line strength. It represents the rate of decay in a two-level system:

$$A_{ji} = \frac{16\pi^3 \nu_3 S}{3\varepsilon_0 h c^3 g_i}, \qquad [A_{ji}] = \mathrm{s}^{-1}. \tag{13}$$

where ν is frequency, and ε_0 (vacuum permittivity), h and c are natural constants. Astronomers typically use a related quantity, called gf-values for line lists [Bernath, 2019]. Line lists come in a variety of formats, units, and sizes. In general, they must purvey information about the states and transitions, most importantly, line positions, line intensities, a set of quantum numbers, and the energy of the lower state. A handy way to obtain a measure for the strength of a transition in terms of gf-values, using the Einstein coefficient and the states' quantum numbers is through the *absorption oscillator strength*, f_{ij} (here, j is the higher energetic state). f_{ij} is simply put an absorption rate divided by the absorption rate of a classical, single-electron oscillator :

$$f_{ij} = -\frac{g_j}{g_i} f_{ji} = \frac{g_j}{3g_i} \frac{A_{ji}}{\gamma_{cl}}, \tag{14}$$

where γ_{cl} is the emission rate of a classical single-electron oscillator at frequency ω_{ji}, and A_{ji} is the emission rate in question (the Einstein coefficient for emission). The factor 1/3 ensures that if $J_j = 1$, $J_i = 0$, only absorption can occur:

$$g_j = 2J_j + 1 = 3, \qquad g_i = 2J_i + 1 = 0 \implies f_{ij} = 1, \qquad f_{ji} = -\frac{1}{3}. \tag{15}$$

In order to calculate f_{ij}, we need the definition of γ_{cl}:

$$\gamma_{cl} = \frac{e^2\omega_{ji}^2}{6\pi\varepsilon_0 m_e c^3}, \qquad \text{and conversion of frequency:} \qquad (\nu_{ji}^2 = 4\pi^2\omega_{ji}^2), \tag{16}$$

where e and m_e are electron charge and mass, respectively. We now obtain an expression we can use directly with quantities obtained from our line lists:

$$f_{ij} = \frac{g_j}{g_i}\frac{\varepsilon_0 m_e c^3}{2\pi e^2 \nu_{ji}^2} A_{ji}. \tag{17}$$

To get our desired gf-value, we simply multiply by g_i: $gf = g_i f_{ij} = -g_j f_{ji}$. If needed, we also have the handy conversion to the integrated line strength, s_0, [Hilborn, 1983]:

$$s_0 \text{ [cm/molecule]} = 5.33129 \times 10^{11} gf. \tag{18}$$

For reproducible results, line lists should also include a partition function for the listed molecule, and lastly, a choice has to be made on a line shape function, before one can compute a synthetic spectrum. Sometimes, information about the line shape is included with the line list. The standard choice of line shape is a combination of a Gaussian function and a *Voigt* function to represent Doppler- and pressure broadening, respectively. The Gaussian broadening is quite simple, whereas the Voigt function, which depends on composition and temperature, needs estimates for values that are not usually known from experiment [Bernath, 2019]. The need for a partition function is clarified in section 2.7.

2.4.1 Weak and Strong Lines

There are several important differences between any laboratory setup and a stellar atmosphere. First, is a natural limit on the resolution. If one were to construct a very high-resolution laboratory experiment, on the order of 1000 lines per Å, the individual lines would be impossible to resolve due to crowding and broadening of the lines. A cool atmosphere is a complex structure, and its thermal properties all respond to the absorption and emission in the gas, as opposed to a laboratory setup, where the strongest absorber dominates. Even if laboratory measurements take place under elevated temperatures, the pressure will usually be kept constant, or at least not be able to respond to high energy transitions in a way that mimics an entire atmosphere. As seen in section 2.3, since the weakest lines make up only a small fraction of the integrated absorption coefficient, those that compare their line lists with laboratory measurements should be careful, as a grave error in the weak lines has very little effect on the integrated absorption coefficient. Two line lists for the same species can have almost identical integrated absorption coefficients but different effects on the atmospheric structure. Even if one takes care to not only validate results using the total integrated absorption coefficient, comparing with laboratory data is still not a very good way of evaluating an entire line list, especially if the list has been adjusted with the data that it is being compared to. Laboratory setups will often be limited to a narrow spectral range, while comparison with stellar spectra depends on the full set of computations, and whether the weak lines are handled properly. To quantify how a dense atmosphere becomes opaque at certain frequencies and affects the radiative transfer, an opacity function is used. An opacity function uses information about absorption, such as line strength, elemental abundance, isotopes, and line shapes (possibly dependent on temperature or pressure), and calculates absorption coefficients of all species combined at given temperatures.

The strong lines are close to the ground state and are therefore easier to compute accurately, using the simple harmonic oscillator as eigenfunction. This system breaks down at higher energies, making quantum number assignment useless. As mentioned, line lists are constructed using different approaches. Determining the lists that best reproduce data may shed some light on which assumptions are good when constructing a line list. Three different approaches are prevalent in line list construction: The first is to do the necessary computations as accurately as possible, with no regard to laboratory measurements. This is the philosophy behind the SCAN line list. The second is to adjust those theoretical results, or a part of them, to laboratory results. This is what was used in the Partridge & Schwenke (PS) list [Partridge and Schwenke, 1997], referenced in other papers presenting line lists, but not tested here. The third approach is to use laboratory-measured band intensities, where one believes they are more accurate than the theoretical ones, which is done for the HITEMP2010 list. Correcting the band intensities involves a risk of tampering too much with the weak lines in that spectral range, favoring the accuracy of strong observed lines. One should avoid doing so, or at least be very careful if the list is to be used in atmospheric modeling. Tashkun [2006] compares, in a broad spectral range, BT2, PS, an older version HITEMP, SCAN, and a completely empirically-based list aimed at a more industrial context by the name of TOMSK. Their comparison favors BT2 and shows that SCAN overall performs worst. The comparison is done with a focus on isothermal observations. Since then, HITEMP has been updated, and the ExoMol group has adopted the new line list POKAZATEL.

2.4.2 SCAN

The SCAN line list database by U.G. Jørgensen (Jørgensen and Jensen [1993],Jørgensen et al. [2001]) contains almost 3 billion absorption lines and 98% of energy levels up to the dissociation energy for water. However, for astrophysical calculations, a few tens of millions are sufficient [Jørgensen et al., 2001] The SCAN H_2O line list [Jørgensen et al., 2001] is constructed by attempting to most accurately calculate transitions with no adjustments. These calculations are described in Jørgensen and Jensen [1993]. They compare their version of the SCAN list with integrated absorption coefficients of the original HITRAN database [Rothman et al., 1998], which included only 86 strong lines at low temperature. Jørgensen et al. [2001] showed that a line list consisting of 3 billion lines could be reduced to around 100 million, while still accurately reproducing high-quality spectra in self-consistent models. This reduction was possible when eliminating a number of the ultra-weak lines around the dissociation of water (around 3500 K). Here, they compared their integrated absorption coefficients with that of the HITEMP1995 line list. This 100 m. line version of the SCAN list has a vibrational cutoff at 15000

cm^{-1}, the lowest between the four compared here. This corresponds to 0.66 μm, while our observed spectra have a minimum wavelength of 0.65 μm.

2.4.3 BT2

In the paper presenting BT2, Barber et al. [2006], the authors claim to have addressed problems present in previous line lists, including PS, and SCAN, by having more accurate calculations and a simple increase in computational power allowing for more completeness and accuracy. They also claim that the most accurate previous line list is the PS list. They use an energy cutoff at $J = 50$, $E = 30000$ cm^{-1}, and have 506 m. lines in their list. They note that this cutoff only omits 0.02% of the total partition function. Note, however, that the crucial weak lines that lower the pressure in the upper layers have very little contribution to the total partition function.

2.4.4 HITEMP2010

The former version of HITEMP, tested in Jones et al. [2003], included 1,283,466 lines with an energy cutoff at 15000 cm^{-1}. It is an observationally based list, including high accuracy observation of water lines from laboratory setups (inevitably favoring the strong lines because of limitations in such a setup.) HITEMP has since been updated to HITEMP2010, through a desire to have better agreement with the HITRAN database, more updated calculation, and having a higher cutoff to more accurately represent high-temperature transitions. This culminated in the emergence of a blend of BT2 and HITRAN, with "semi-empirical line positions", called HITEMP2010. The former will be referred to as HITEMP1995. In Rothman et al. [2010], presenting HITEMP2010, they show their integrated absorption to agree better with laboratory data from Coppalle and Vervisch [1986], than the SCAN line list in the range 1.3 μm - 2.3 μm. This is expected, as the HITEMP list comes with values that are adjusted to laboratory data. They were, however, only compared to measurements in that range, and not in high-temperature environments. Like BT2, the vibrational energy cutoff of HITEMP2010 is at 30000 cm^{-1}, although given in different terms at a reference temperature in its paper.

2.4.5 POKAZATEL

POKAZATEL, the line list of Polyansky et al. [2018] has the highest cutoff at 41000 cm^{-1}, $J = 72$ (highest value of J for which all energies lie below the cutoff.) The final list uses empirical energy levels to correct line positions when available. The list has over 5.6 billion transitions and is currently recommended by the ExoMol group. The authors list PS and BT2 as the most widely used line lists and claim PS to be more accurate than BT2 longward 1 μm, and both lists to be unsatisfactory at temperatures > 3000 K. They mention that water has been observed on dwarf stars with temperatures up to 4000 K, as a reason for the need of a water line list that is complete up to these temperatures. Like in Barber et al. [2006], improved computation and improved calculations (fitted to improve line positions) are mentioned as reasons for their improved list, along with nuclear motion calculations at energies up to dissociation. They use observational work by IUPAC [Tennyon et al., March 2013] to obtain very accurate experimental and experimentally derived water energy levels, and claim to have succeeded in their aim to make the first complete rotational-vibrational line list for a polyatomic molecule. Their results are fitted to spectroscopic data from Shirin et al. [2003] up to 25000 cm^{-1}, and data from Matsyutenko et al. [2009] from 27000 cm^{-1} to dissociation, D_0. Bernath [2019] details how their method of calculations retain both the benefits of HITEMP's empirical accuracy and the possibility of easy extrapolation of energy levels.

2.5 Cloud Formation

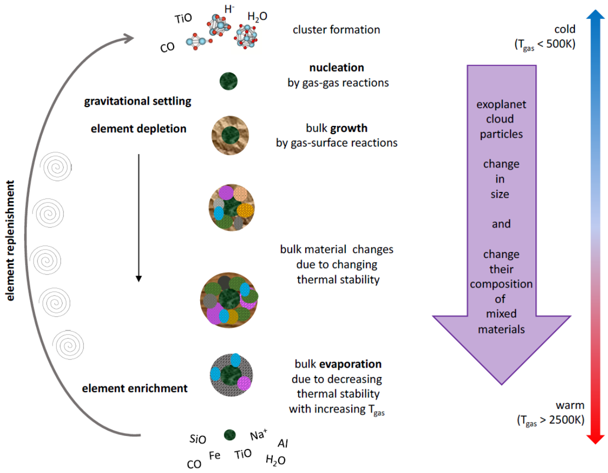

Figure 4: General overview of cloud formation cycle as described in the text. Source: Helling [2019].

Cloud formation is expected in most exoplanets, so it is critical to have good tools for modeling cloud formation in exoplanet atmosphere modeling and understanding its impact in exoplanetary spectroscopy. It is modeled in terms of grain number and size, in the context of radiative transfer and mixing processes. Clouds require *seed particles*, and in the case of gas giants without crust, seed formation is required [Helling et al., 2016]. On these seeds (dust particles), surface reactions that are more efficient than the seed-forming gas chemistry can occur. The cloud structure is determined by nucleation rate J_* [s^{-1}], grain size $\langle a \rangle$ [μm] and number of cloud particles n_d. When cloud particles are formed, they undergo gravitational settling, in which they continue to grow and change composition (as do their surroundings during descent). While growing, the clouds deplete seed-forming elements, skewing the metallicity and C/O ratio in the atmosphere. This depletion is counteracted by convection from lower warm layers. Once replenished, the upper layers can form seeds for new clouds. Friction with surrounding gas during settling determines the coupling between clouds and gas, affecting the descent rate and the specific chemical reactions. If there is a lot of friction, the cloud will be carried along by the wind. If the gas density becomes too low, or the cloud particles too heavy, particles decouple. If this happens on large scale (growth can not keep up with decoupling), rain occurs. The overall likelihood of cloud formation, size, and composition are ultimately determined by T_{gas} and ρ_{gas}, while gas velocity governs vertical mixing [Helling et al., 2016]. The general cloud formation mechanism as described here is shown in Figure 4. The interplay between cloud and atmospheric composition affects both modeling and observation at low temperatures. Clouds have high opacity, causing heating, while they block the line of sight to lower parts of the atmosphere. Their formation and growth also alter chemical abundances, most noticeably for uncommon elements. 3D models reveal zonal patterns similar to that of Jupiter due to the horizontal motion of gas and clouds, and lightning, indicating a global electric circuit (GEC) can link atmospheric processes over long distances. [Helling, 2019].

2.6 The MARCS Code

The MARCS code was constructed in the 1970ies for modeling stellar atmospheres in light of the need for self-consistent models of stellar atmospheres. It has existed in many different versions since then and has been employed in different modeling approaches. A short overview of the MARCS code's history is given in Gustafsson et al. [2008]. In the original version of MARCS, an opacity distribution function (ODF) was used, due to computing constraints. This method comprises of choosing a wavelength region and computing a fraction of lines that have an absorption larger than some quantity l. A program would then create a function that represented the line absorption in this wavelength region based on that fraction of lines [Gustaffson et al., 1975]. Today, other groups use a similar method called "correlated k distributions", converting the lines in a frequency interval into *k distributions* of absorption strengths [Gustaffson et al., 1991]. In the version of MARCS we use today, the opacity input consists of lists of opacity sampling files for each molecule, that need to be calculated beforehand. For this calculation, a partition function for the molecule is needed. This is a critical step in modeling atmospheres, as there are multiple ways to calculate partition functions that the absorption coefficients are derived from, and line list completeness (in terms of both energies and line number) needs to be considered. A comparison of methods for calculating the total internal partition function can be found in Popovas and Jørgensen [2016]. The authors argue that a standard coherent set of partition functions is crucial in making comparisons of atmospheric models and that at the very least, it should be clear which method is used, as it affects the whole physical structure. MARCS has the option of calculating plane-parallel or spherical symmetric models and assumes LTE. Chemical equilibrium from minimization of Gibbs free energy is calculated in the GGChem code during MARCS atmosphere computing. MARCS uses a source function that is dependent on monochromatic scattering and absorption coefficients σ_λ and κ_λ [Gustafsson et al., 2008]:

$$S_\lambda = \frac{\kappa_\lambda}{\kappa_\lambda + \sigma_\lambda} B_\lambda(T) + \frac{\sigma_\lambda}{\kappa_\lambda + \sigma_\lambda} J_\lambda, \tag{19}$$

where $B_\lambda(T)$ is on the form of equation (10). MARCS uses a mixing length approach to convection, adding a convective flux to the energy equilibirum. The numerical methods MARCS uses are detailed also in Gustafsson et al. [2008]. Juncher et al. [2017] show how clouds can be implemented in MARCS via alternating execution of MARCS, and a cloud modeling program by Christiane Helling, called DRIFT (since renamed *static weather*). They model atmospheres for ultra-cool dwarfs ($T = 2000 - 3000\,\mathrm{K}, log(g) = 4.5$) and a giant gas-planet WASP 19-b.

2.7 Partition Function

2.7.1 Outline

The total internal partition function is defined in the context of thermodynamical equilibrium and determines the distribution of atoms and molecules in well-defined energetic states and temperatures. It is central to much of astrophysics. Since it is a statistical sum of Boltzmann factors (energy terms), it relies totally on how energetic states are defined. When physicists use different or inconsistent methods for assigning energetic states, this, in turn, affects the entire thermodynamics of their model. How to obtain a more accurate method of calculating partition functions is the subject of much research.

To increase the accuracy and reproducibility of astrophysical research, a standard set of Q_{tot} should ideally be used. This is not a problem that is solved as of yet, let alone reached a consensus about. Until then, clarity regarding the treatment of this problem is important when publishing results. Popovas and Jørgensen [2016] give a good discussion of the problems with the typical semi-classical approach to this problem. This outline of the typical approach is based on their work.

2.7.2 Energies

The Born-Oppenheimer approximation states that the motions of nuclei and electrons can be separated, due to the large difference in mass. By making use of the Born-Oppenheimer approximation, one can assume independence of vibrational and rotational energies, and likewise between vibrational and electronic [Zielinski et al., Updated 2021]. This allows for the calculation of independent contributions for the total internal partition sum, based on each energetic regime. The translational contribution is simple to calculate and depends on particle number N, temperature N, pressure P and particle mass m:

$$Q_{tr} = \frac{Nk_BT}{\lambda^3 P}, \qquad \lambda = \sqrt{\frac{2\pi\hbar^2}{mk_BT}}, \tag{20}$$

where h, k_B are natural constants. For vibrational energies, the harmonic oscillator potential can be used to describe a diatomic molecule sufficiently well when close to the potential minimum (minimum internuclear distance $R = R_e$.) A better approximation is the Morse potential:

$$U(R) = E_D\left[1 - e^{-a(R-R_e)}\right]^2, \qquad \Delta E(v) = h\omega\left[1 - \frac{h\omega}{2E_D}(v+1)\right], \tag{21}$$

where E_D is the binding energy of the potential well. With this potential, energy separations ΔE decrease with vibrational quantum number v. This Morse potential is an anharmonic oscillator. Given in term values $G(v) = E(v)/hc$:

$$\text{H.O.: } G(v) = \omega_e\left(v + \frac{1}{2}\right) \qquad \text{Anh.O: } G(v) = \omega_e\left(v + \frac{1}{2}\right) - \omega_e\chi_e\left(v + \frac{1}{2}\right)^2, \tag{22}$$

$$\text{where in case of the Morse potential:} \quad \omega_e = \frac{\omega}{c}, \quad \omega_e\chi_e = \frac{h\omega^2}{4cE_D}, \quad \omega = a\sqrt{\frac{2E_D}{\mu}}. \tag{23}$$

In (23), ω_e is the wave number, $\omega_e\chi_e$ anharmonicity term, and ω vib. frequency. In the harmonic approximation, we can sum Boltzmann terms and get an approximate Q_v:

$$Q_v \approx \sum_{v=0}^{\infty} e^{-hc\omega_e(v+\frac{1}{2})/k_BT} = \left(1 - e^{-hc\omega_e/k_BT}\right)^{-1}. \tag{24}$$

For rotation, the most simple description is the rigid rotator:

$$E(J) = \frac{1}{2}I\omega^2 = \frac{J^2}{2I}, \quad \text{with sq. length of } J \text{ in QM context: } E(J) = \frac{J(J+1)\hbar^2}{2I}. \tag{25}$$

Here, J is the total angular momentum, and I is the moment of interia. (25) would give again give an analytically calculable partition function, but it is not very physically appropriate. Using instead the non-rigid rotor allows for a centrifugal force acting on the atoms, increasing R. This causes an increase in the moment of inertia, thereby lowering E_{rot}. In term values, the distortion D_e yields

$$F(J) = B_e J(J+1) - D_e J^2(J+1)^2, \qquad B_e = \frac{h^2}{8\pi^2 cI}, \tag{26}$$

where B_e is the main molecular constant. There are $2J+1$ orientations available for the rotational axis. Additionally, one needs to either consider electron spin degeneracy or add a symmetry factor σ, equal to 2 for homonuclear molecules and 1 for heteronuclear molecules. For RRA with σ, the Q is computed similarly to the H.O.:

$$Q_{rot} = \frac{k_B T}{\sigma hc B_e}. \tag{27}$$

Interaction between vibration and rotation can now be considered by adjusting the vibrational levels with a molecular constant B_v, using the rotation-vibration interaction constant α_e:

$$B_v = B_e - \alpha_e \left(v + \frac{1}{2} \right). \tag{28}$$

The energies of excited electrons are high, and contribute little to the total partition function. For some molecules, where there are degenerate electron states, the electronic partition function used is

$$Q_{el} = \sum \tilde{\omega_e} e^{-T_e/K_b T}, \qquad \tilde{\omega_e} = (2 - \delta_{0,\Lambda})(2S+1) \quad \text{(elec. stat. weight)}. \tag{29}$$

In (29), $\delta_{0,\Lambda}$ is the Kronecker delta (equal to 1 at $\Lambda = 0$, and 0 otherwise).

2.7.3 Usage and Cutoffs

There are problems in the semi-classical approach for high energy levels. There need to be cutoffs for v and J which represents a limit where the description of molecules as isolated atoms, connected by rotating springs, breaks down. Where to place these cutoffs is unclear. One can stop at the dissociation energy, or for the Morse potential, at the point $v_{max} = \frac{2E_d - h\omega}{\omega}$, where the energy separation between states, ΔE has reduced to zero. Equivalently, for the non-rigid rotor (eq. 26): $J_{max} = \frac{\sqrt{2B_e} - 2\sqrt{D_e}}{2\sqrt{D_e}}$. For both, however, more terms still ought to be added to the approximation to account for higher states. Since the semi-classical approach breaks down at the highest level, there is not much left, other than to use experimental values. This is very problematic, as assumptions are needed for non-measured levels, while measured levels have to be categorized in terms of some insufficient semi-classical theory, and even then, one would observe also dissociation and recombination energies. The incorrect assignment has usually been ignored, and experimental values fitted and extrapolated. Popovas and Jørgensen [2016] list 8 cases of increasing complexity that have been used in literature to compute Q_{int}, ranging from using simply the HOA and RRA, to dropping both HOA and RRA, including molecular constants that vary with level, dissociation, and finally abandoning molecular constants, as they are rooted in classical mechanics.

3 Method

3.1 Methods Outline

Line lists were collected from publicly available sources online and converted into a format that is suitable for an opacity sampling script. Observed medium-resolution spectra of red dwarfs were downloaded from the SpeX Prism Library, expected to showcase dominating absorption from H_2O and TiO in different wavelength regions. A red giant was also selected, to show line list performance in a broader wavelength span. From the OS files, model atmospheres were computed with fairly standard parameters, and from those, synthetic spectra were generated, containing 27 opacity sources. These synthetic spectra are compared to the observed ones. The model structures and single-molecule contribution spectra, in different cases, are also examined.

3.2 OS Absorption Coefficients

MARCS obtains absorption coefficients from line lists by way of *opacity sampling.* The opacity sampling files consist of statistic absorption coefficients, constructed from a line list and a partition function, by random sampling of wavenumbers. The opacity (mass absorption coefficient at wavelength λ, μ_λ) is obtained when the absorption coefficients from the OS file are multiplied by the partial pressure of a molecule in the atmospheric model. The opacities are used in MARCS to compute self-consistent radiative transfer, energy balance, and $T - P_g$ profiles. The OS files are specified for a set of different temperatures, making it possible to interpolate between these, for use at any temperature in the atmosphere model. MARCS and the accompanying spectrum computation program SYNTOS uses OS files, computed with OS-script OS_EXOMOL, written by Jørgensen in December 2020. The program samples equidistant points from a scrambled selection of lines, and is designed to have an equal contribution to opacity according to energies, causing it to sample more, where the contribution to a computed spectrum's resolution is greatest. After sampling random wavenumbers, contributions from the line list in that range are summed up (making it possible for several lines to contribute), and in the case of molecular OS files, the integrated line strength (s_0) is distributed onto a temperature-dependent Gauss-profile. MARCS can then interpolate in temperature between the statistically selected points. Only a temperature-dependant Gauss function for broadening is used, omitting the Voigt profile for molecules, since we assume that crowding of lines with broad Doppler wings renders the full Voigt profile unnecessary. This allows for using the same molecular OS file for any model, while the atomic OS's have to correspond to the parameters of the atmospheric model. The integrated absorption coefficients presented in Figure 5 are smoothed plots of the opacity sampling absorption coefficients taken at 2700 K.

3.3 Synthetic Spectra

The spectrum program SYNTOS, written by U. G. Jørgensen, computes a synthetic spectrum with any chosen number of OS-absorption coefficients and a given model atmosphere from MARCS. The resolution of the spectra is generally 1/10 to 1/20 of the resolution of the OS files. The script can also compute named filter magnitudes and colors, and fold the output with an instrumental profile. It produces up to 9 spectra simultaneously, each with their specified combination of molecular OS files, along with a continuum and a Planck function. First, wavenumbers are calculated, and then an absorption coefficient value from the nearest OS-point is assigned to the wavenumber. Then, interpolation between temperatures listed in the OS file is performed to match the model temperature and adjust the value thereby (opacity at λ is a function of T). The program uses opacity data for 27 molecules, as these are the ones that, currently, there are both partial pressures readouts from the model file, and listed OS files. 27 molecules should be more than sufficient for the temperatures we are working with [Gustafsson et al., 2008].

3.4 Collection and Conversion of Line Lists

3.4.1 OS Parameters

OS_EXOMOL computes opacities at 12 different temperatures between 500-4000 K, with a Doppler broadening of 3 km/s and $Z/Z_{\odot} = 1$. The wavenumber interval increases in steps from 5-500 after reaching certain wavelengths, between 400 cm^{-1} and 80000 cm^{-1}, equivalent to 24 μm - 1200 Å (note that the maximum value wave number in the water lists is $\approx 41,500$ in POKAZATEL.) The molecular OS files are in cm^2/mol. This results in an opacity of cm^2/g when multiplied by partial pressure in MARCS. For all lists, the partition function provided by the creator of the list, available together with the lists, was used for OS calculation. All opacity sources other than water lists BT2, POKAZATEL and HITEMP2010 were included and specified in the version of MARCS provided by Jørgensen.

3.4.2 SCAN

The current SCAN list of 100 million lines is available by anonymous ftp. The line list is in the format (ν_{ij} [cm^{-1}], s_0 [cm/molecules], A_{ij}].) The OS file for SCAN contains 72,492 wavenumbers.

3.4.3 BT2

The BT2 line lists can be downloaded from The ExoMol website by searching "Data" and "by molecule". The data is presented in *levels* files and *transitions* files. The *levels* files contain 221,097 energy levels, their energies and accompanying quantum numbers. The *transitions* files contain level references for each transition and the Einstein coefficient. The partition function is listed along with these files. Since the OS_EXOMOL script uses gf-values for line strength, equation (16) was used for obtaining for f_{ij}, and thereby $gf = g_i f_{ij}$. The final line list contains 505,806,255 lines in the format (ν_{ij} [cm^{-1}], gf [cm/molecule], A_{ij}, ISO), where "ISO" is a isotope reference number ("1" throughout). The list covers a spectral range of 0-30,000 cm^{-1}. The computed OS file contains 86,336 wavenumbers.

3.4.4 POKAZATEL

The POKAZATEL line list is available through ExoMol and can be obtained in the same way as BT2, along with its partition function. It is listed on the website as the recommended line list for $^1H_2{}^{16}O$. The line list was collected and converted in the same way as for BT2, with the exception that it was done separately for six subdivisions of the list, calculating OS files for each, before merging them back together. This was done due to the sheer number of lines in the POKAZATEL list. The final line list covers a spectral range of 0-41,140 cm^{-1}, and contains 5,569,365,927 lines. The computed OS file contains 92,682 wavenumbers.

3.4.5 HITEMP

The HITEMP2010 water list and partition function are available through HITRANonline. The list contains 114,241,164 transitions in the spectral range of 0-30,000 cm^{-1}. The format of HITEMP2010 is a 160-character length fixed-width format, displaying many properties of each transition, and including information about line shape. This format requires reading in Fortran or using an equivalent subroutine. This was done with a Fortran reading package in a Python script used for the subsequent conversion. A definition of all the units and formats are given in the presenting paper, Rothman et al. [2010]. The downloaded version lists intensity at a reference temperature of 296 K, requiring one to calculate back to the temperature needed, as described in the article. Since the quantities used for conversion of BT2 and POKAZATEL are listed, eq. (16) was used again, and the line list was put in the same format as the others. The OS file we use contains 86,141 wavelengths. HITRAN includes lineshape parameters, which can be used if no opacity function with line shape is used.

3.5 SpeX Prism Data

The SpeX Prism Library (SPL) [Burgasser, 2014] contains low-resolution NIR spectra of cool dwarfs, obtained with the SpeX spectrograph on the 3m NASA Infrared Telescope Facility on Hawaii. The spectra used have a resolution of 120 and span 0.65-2.55 μm. The resolution is sufficient for comparison with our models, and they fit our interests in the NIR spectra of cool dwarfs. The three cool dwarfs chosen from the SPL are part of the *Two Miron All-Sky Survey* 2MASS [Skrutskie et al., 2006 February].

3.6 Star Selection

As cloud formation is not included in the version of MARCS used for the line list comparison, the stars chosen should not be too far below the limit for cloud formation $T_{eff} < 2700$ K. Four different stars around this limit are chosen, excluding also warmer stars, as water is the most important contributor to opacity at the temperature of late M dwarfs ($\approx$ 2500 K), dominating the opacity in the infrared [Barber et al., 2006]. Atoms were deemed not necessary to include in the atmosphere models, because of the low temperature. The warmest of the cool dwarfs chosen, 2MASSJ00583814-1747311 and 2MASS J17364839+0220426, were both originally published in Burgasser et al. [2014]. The first was originally observed on 18 Sep 2003 and has been assigned NIR spectral class M6. Its temperature is assumed to be $\approx$ 2900 K. The second was originally observed on May 23 2003 and was assigned NIR spectral class M8. It is assumed to be $\approx$ 2700 K. The coldest object, 2MASSJ00163761+3448368, was originally observed on Sep 9, 2005, and originally published in Kirkpatrick et al. [2010]. Its assigned NIR spectral type is M8.5, and it is assumed to be $\approx$ 2500 K. The resolution of all three spectra is 120. The fourth object is a cool giant, SV Pegasus [Aringer et al., 2002], taken by the *Short Wavelength Spectrometer* aboard the *Infrared Space Observatory* (ISO-SWS) [de Graauw et al., Nov 1996]. The SWS covers the range 2.38-45.2 μm with spectral resolution between 1000-2000. Astronomical database Simbad lists the spectral type of SV Peg as M7 D. It was previously fitted to a temperature of 2900 K and surface gravity $\log g = 0.0$ using the SCAN line lists by Jørgensen et al. [2001].

3.7 Execution and Parameters

The MARCS atmosphere models contain OS files for 53 molecules, many of them from ExoMol. Gustafsson et al. [2008] lists sources for continuous absorption from atoms and ions, as well as molecular absorption before the addition of 30 new molecules by Rune Kjærsgaard. A table of references for these 30 molecules is found in section A.1. An atmosphere model is made for each spectrum needed to fit the observed spectra (varying T_{eff} and $\log g$). Several other atmosphere models are made to study the effect of each of these parameters on the atmosphere structure and temperature, and how they change the contribution to opacity from certain molecules. Every time, the option to begin the computation from a similar model is chosen, and the model is set to converge when a temperature change of less than 1.0 K occurs from an iteration, while a constraint is put on the maximum number of iterations to 30. All models have solar elemental abundances and range from wavenumbers 400 - 80000 cm^{-1} (1200 Å - 25 μm). The output includes a temperature-pressure structure, contributions to pressure and flux from many molecules, ions, and atoms, absorption and scattering coefficients, radiation pressure, turbulence- and electron pressures, and thermodynamic properties. MARCS is set to compute plane-parallel atmosphere models. Temperature-pressure structures from converged models were extracted with a SHELL-program and plotted in Python. SYNTOS was set to average the spectrum over 20 OS-points, and it does not have the spectrum folded with an instrument profile. MARCS, along with SYNTOS, OS_EXOMOL and line list converting scripts, some provided by Rune Kjærsgaard, and plotting programs in IDL, were all run on the UCPH HPC-computers via remote SSH connection. Most are optimized to utilize the multicore capabilities of the computing center. This required becoming familiar with each program, its input, naming practice, and language, along with C SHELL for navigating the HPC computers. All data and code used in this project are, as of writing, available on the UCPH HPC network.

4 Results

4.1 Comparison of OS Plots

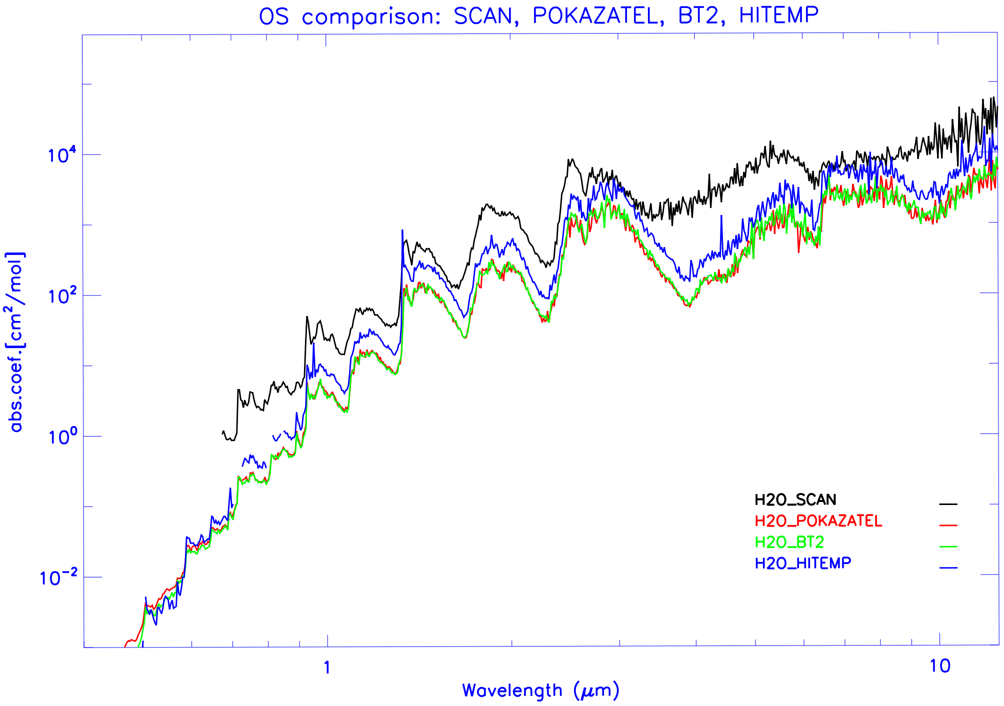

Figure 5: Each of the four opacity sampling absorption coefficients, plotted by the wavelength at 2700 K. BT2 and POKAZATEL coincide through much of the spectral range, as expected from Polyansky et al. [2018]. SCAN (black) shows no absorption shortward of 0.66μm, due to its cutoff at 15000 cm^{-1}

It is clear from Figure 5, how similar the POKAZATEL and BT2 OS files are. The HITEMP OS follows these two very closely, although shifted by a small factor. SCAN, on the other hand, differs from the others in a lot of places, along with having higher overall absorption. The plot also shows the lower energy cutoff at SCAN, resulting in the absorption coefficient beginning around 0.66 μm (the equivalent of $\nu = 15000\text{cm}^{-1}$). There are small gaps in the low wavelength region of the HITEMP OS, caused by missing data in the line list for the preferred conversion method. This does not cause any problem in the use of the OS.

4.2 Effect of Temperature

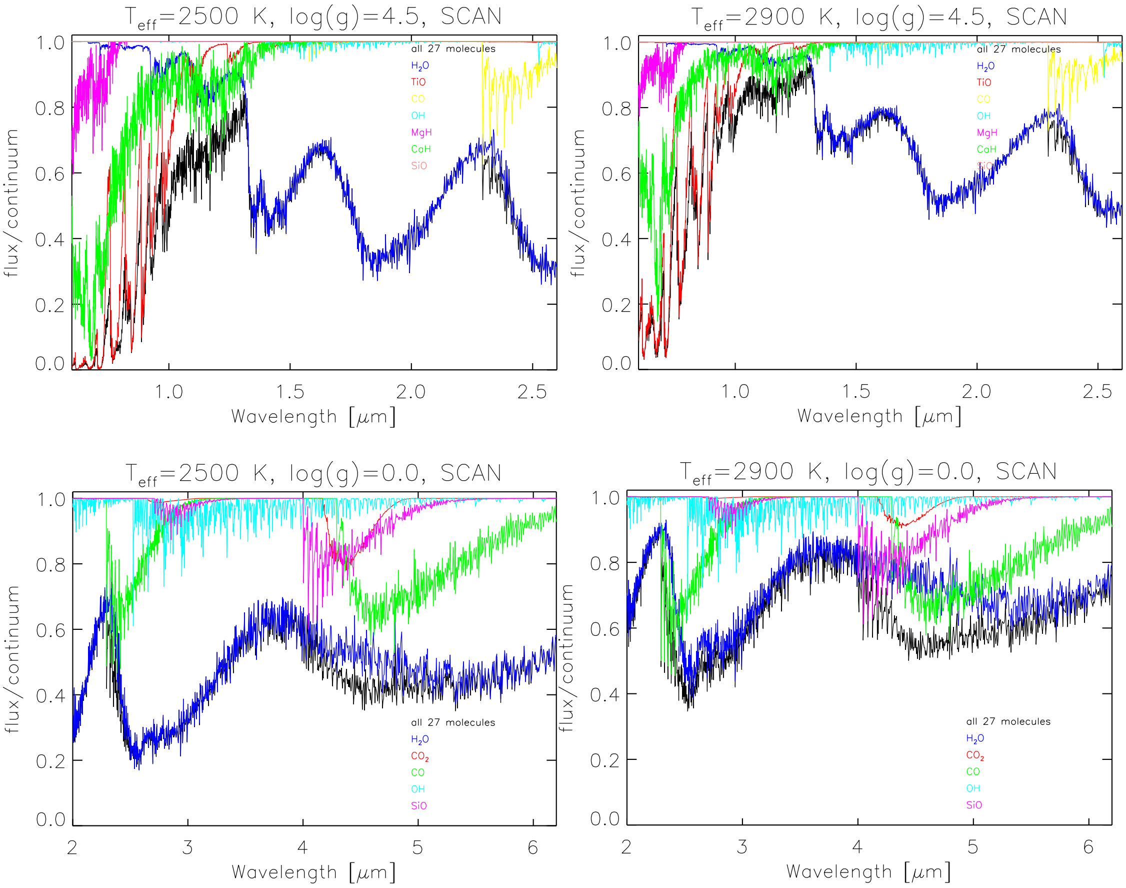

Figure 6: Four plots of synthetic spectra using different parameters. The top spectra contain 7 single-molecule spectra to show their effect on the total absorption in our fits, and the bottom plots show only 5 single-molecule spectra, as fewer molecules make a difference in this region. The top two are taken in the wavelength region for our dwarf fits at $\log g = 4.5$, and the bottom in the region relevant to the giant fit at $\log g = 0.0$.

Figure 6 is a comparison of synthetic spectra at the extremes of our fitting temperatures. Above, for $\log g = 4.5$, in 0.65-2.55 μm, the fitting region for the dwarfs, and below, $\log g = 0.0$ in 2.0-6.2 μm, the fitting region for of giant. The plots show the opacity of the molecules, which have the largest effect on the shape of the combined spectrum. The top plots show that the spectra spanning 0.65-2.55 μm, as expected, are dominated by TiO (red), shortward of 1.3 μm, and water (blue), longward of 1.3 μm. The combined spectrum (black) shows much more absorption of TiO (red) at 2500 K, strongly affecting the shape of the spectrum shortward of 1.3 μm. At 2500 K, there is also a shifting downward of the absorption of water (blue), with more pronounced absorption features around 1.4 μm and 1.8-1.9 μm, compared to the 2900 K plot. The bottom two plots show that increasing the temperature decreases the total absorption as a result of less absorption from water in 2.0-6.2 μm. Every other molecule is largely unaffected in comparison. At 2500 K, most of the partial pressure comes from H_2O, CO, and H_2. H_2 only absorbs very little at very narrow intervals because of its homonuclear structure. CO dominates partial pressure in the deep layers, while almost absent in upper layers. Despite its large partial pressure contribution, CO does not contribute much to absorption. For TiO, the case is the opposite. While it has a small partial pressure, it has very strong absorption per molecule and many electronic transitions in the visual range [Bernath, 2019].

4.3 Effect of Surface Gravity

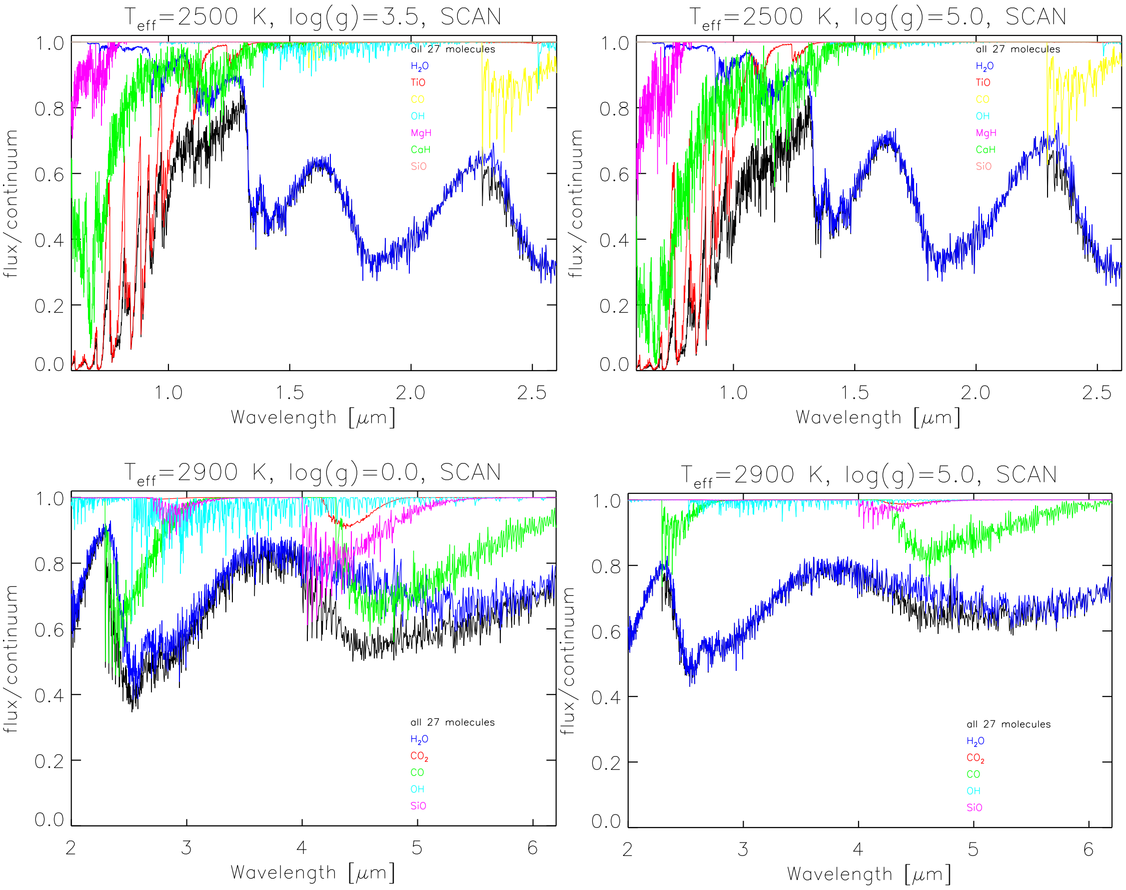

Figure 7: Four plots of spectra were computed with different surface gravity. The plots contain separate single-molecule spectra for each molecule that has an important contribution in this region at the given temperature. The top two is a comparison between $\log g = 3.5 - 5.0$ at constant temperature 2500 K, to show the effect of those molecules in the context of our dwarf fits. The bottom two are at a warmer constant temperature of 2900 K and show a more dramatic change in surface gravity from 0.0 to 5.0 in the 2.0-6.2 μm giant fitting region.

On Figure 7 is shown the effect on spectra when substantially changing the surface gravity parameter, $\log g$. The top two spectra are computed at 2500 K, with surface gravities of

$\log g$ = 3.5 on the left, and $\log g$ = 5.0 on the right. The slope of the TiO-dominated part is slightly steeper at $\log g$ = 5.0, while the two peaks (meaning less absorption) at 1.7 μm and 2.3 μm are taller. The effect seen directly on these spectra from raising the surface gravity is then similar to that of lowering the temperature, but significantly less powerful. On the bottom, the same comparison is made in 2.0-6.2 μm, and with a more extreme difference. Here, increasing the surface gravity lowers the absorption of all species shown, resulting in a lower total absorption, and a smoother spectrum. The spectrum seems more affected in the short wavelengths.

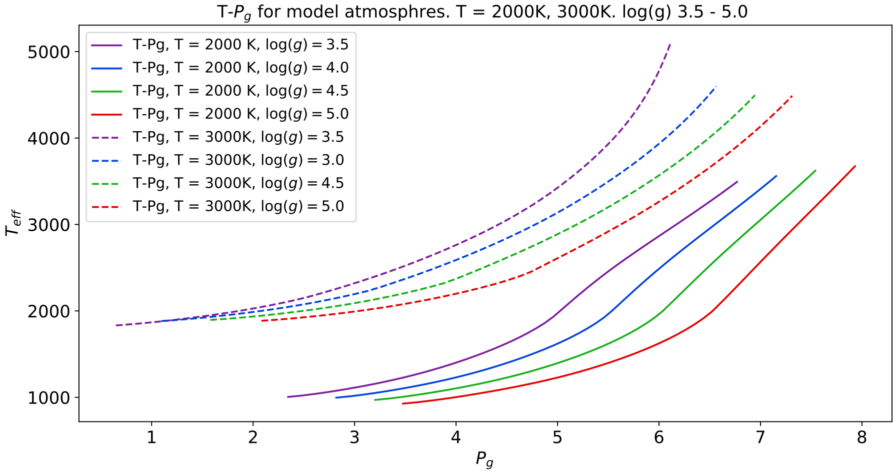

Figure 8: Temperature-pressure profiles of MARCS atmospheres with varying surface gravity. The upper dotted lines are calculated at 3000 K, and the filled lines below at 2000 K.

The plot in Figure 8 shows temperature-pressure profiles calculated by MARCS with the SCAN line list. In the bottom half, are shown the profiles of 4 models with surface gravities spanning $\log g = 3.5 - 5.0$ at 2000 K, and above, the same for 3000 K. The plot shows that the shape of the structure remains quite similar at these surface gravities, although shifted to lower pressures when there is lower surface gravity. For the 3000 K models, outward heat transport seems to be more efficient at low surface gravity, possibly due to more efficient convection. For the 2000 K models, the structures exhibit a small "kink", corresponding to the convection beginning in that layer of the atmosphere.

4.4 Line List Structure Comparison

in Figure 9, a comparison of model atmosphere $T - P_G$ structures for each different line list at T = 2900 K is shown at two different values for surface gravity. The plots exclude the very deepest part to more easily show the differences in the upper atmosphere. The excluded parts of the structures have very similar shapes to each other. On the left, models of surface gravity $\log g$ = 4.5 are plotted in solid, and on the right, $\log g$ = 0.0 models are plotted in dotted lines. For a lower surface gravity atmosphere, high pressures are never reached, while the outer layers expand to lower temperatures, and the inner layers (not shown at this temperature range) have a very large temperature gradient. The plot shows how similar the $\log g$ = 0.0 models are, while there are more noticeable differences in the shapes of the higher surface gravity structures. BT2 (green) and POKAZATEL (black) almost completely overlap, which is expected when examining the similarity of the OS absorption coefficient shown in Figure 5. SCAN has the most unique shape here, with cooler, more expanded upper layers and lower minimum pressure.

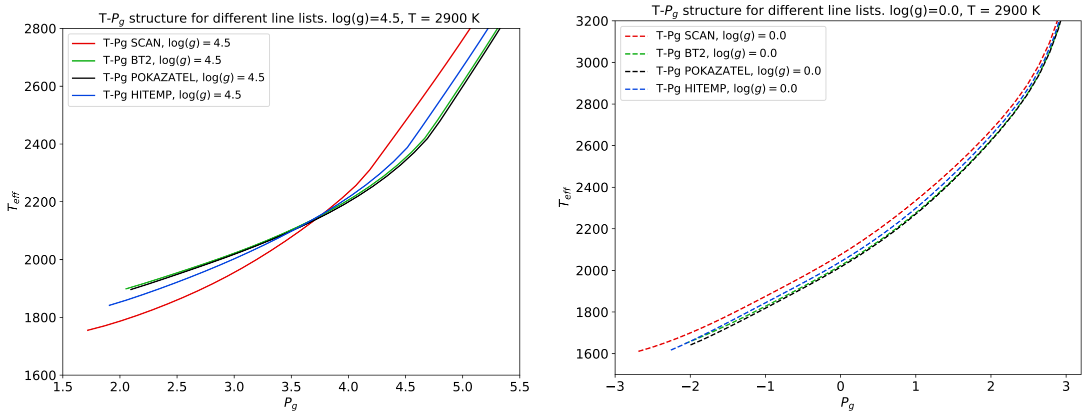

Figure 9: Temperature-pressure profiles of MARCS atmospheres calculated with different line lists for both high and very low surface gravity. The models on the left are calculated with $\log g = 4.5$ The models on the right with dotted lines are calculated with $\log g = 0.0$. Both plots are limited to lower temperatures to reveal differences in the upper atmosphere.

4.5 Stellar Fits

The synthetic spectra all have different multiplicative scaling to match the observed spectra. A method of simply normalizing the observed spectrum to the synthetic in the point closest to the continuum was attempted, but since the different line lists behaved differently at such points, simply multiplying the total flux of a spectrum and fitting "by hand" was easier, and allowed for better alignment. Many of the presented dwarf plots feature a large difference in fit quality between the 0.65-1.3 μm region and the 1.3-2.55 μm region. For this reason, scaling was adjusted to fit as well as possible in the 1.3-2.55 μm-region. This can be seen more clearly on plots limited to this region in the appendix, section A.2. The SV Peg fits all exhibit the same overestimated absorption in the long-wavelength region of the synthetic spectrum. For this reason, they are manually scaled to the 2.0-3.8 μm region, and plots limited to this region are shown in the appendix, section A.3. The wildly varying part of the red lines at each end of the spectral range signifies where the observed spectrum ends, and should not be taken as data that is being analyzed. In all following synthetic-observed spectral fits, the observed spectrum is plotted in red, and the synthetic in black. For an easier direct comparison of line lists for each star, see appendix section A.4. No goodness-of-fit parameters are included, due to the manual scaling involved, and the primary feature of the plots being the difference in behavior longward and shortward of 1.3 μm. The spectra labelled as "all" include the following 27 molecules: C_2, CaH, CH, CN, CO, CO_2, FeH, H_2O, HCN, MgH, OH, SiO, TiH, TiO, CrH, VO, ZrO, H_2, AlH, CS, HCl, NaCl, NH_3, PH_3, SiS, CH_4 and SH.

4.5.1 SCAN

in Figure 10, the best fit for each star is shown using the SCAN OS. The top left fit is of the M8.5 star. Top right is the M8, bottom left is the M6 dwarf, and finally, on the bottom right, is the giant, SV Peg. Apart from an underestimation of the absorption around 1.4-1.65 μm, the synthetic spectrum fits the whole spectrum reasonably well in each case. On the bottom right is the plot of the giant, SV Peg. The "tail" longward of 4.2 μm displays a small excess of absorption in the synthetic spectrum. SCAN's energy cutoff of 15000 cm^{-1} (0.66 μm), is high enough to not cause any problems since the observed spectrum spans 0.65-2.55 μm. The observed spectra are averaged over every 2 points. Before scaling, the synthetic spectra were normalized to 1 at their flux maximum. The top left M8.5 synthetic spectrum does not fit as well as the others, with a lack of absorption around 1.7-2.0 μm, and an "angled" shape in the top of the TiO-dominated region 1.0-1.3 μm.

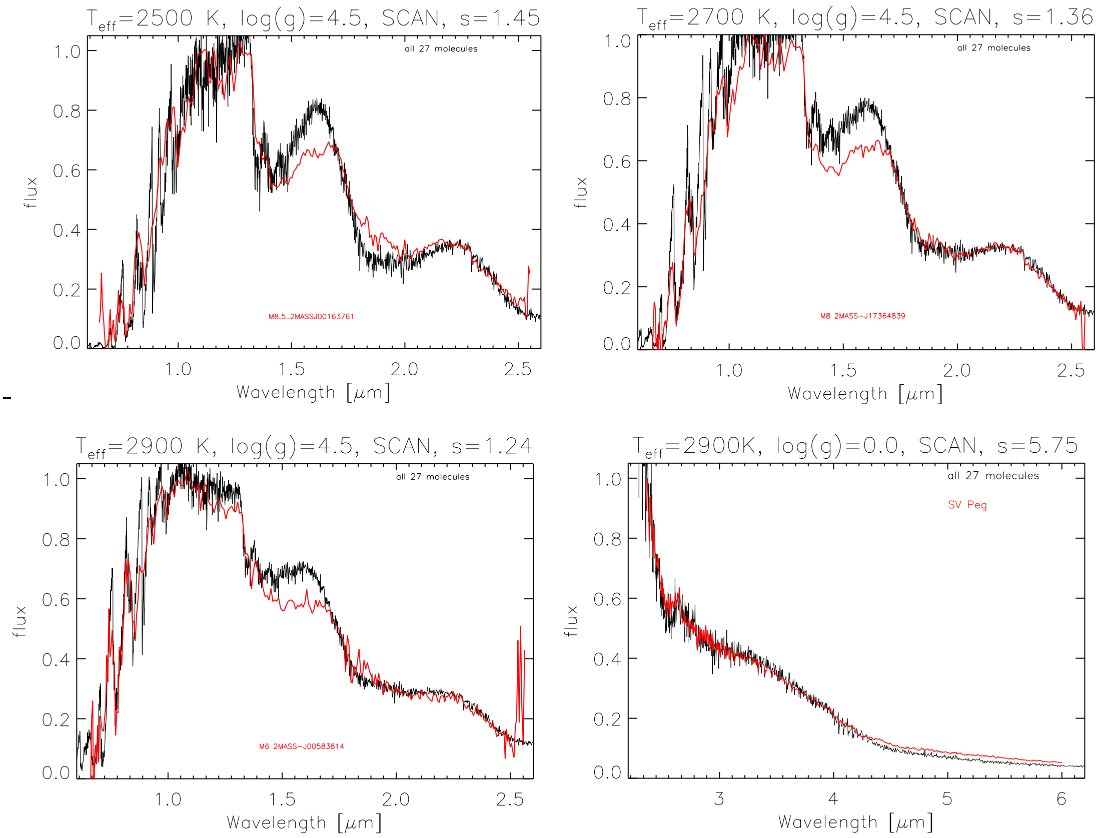

Figure 10: Observed spectra fitted with synthetic spectra. Computed with MARCS and SYNTOS with the SCAN line list. Top left; 2MASS J00163771 (M8.5). Top right: 2MASS J17364839 (M8). Bottom left: 2MASS J00583814 (M6). Bottom right: SV Peg (M7) giant. The synthetic spectra are normalized at their flux maximum and manually scaled.

4.5.2 BT2

Figure 11 shows four fits of the same four stellar objects, in the same order, using the BT2 OS. All four spectra fit very well in the region longward of 1.3 μm, though the M8.5 fit again has too much flux in a small peak around 1.6 μm. Shortward of 1.3 μm, there is a large overestimation of absorption in all three dwarf spectra. This region corresponds roughly to transitions between 7000-15400 cm^{-1}, while BT2 has a much larger cutoff, and is complete up to 20000 cm^{-1}. Scaling the spectrum to fit better with the long-wavelength region did not particularly magnify this problem, as there is no "in-between" good fit. The SV Peg synthetic spectrum on the bottom right fits overall quite well with the observed spectrum, though not as well as the SCAN fit, showing slightly too little absorption around 2.5 μm. There is again excessive absorption at the tail in the long-wavelength region. The best-fitting spectra in the long-wavelength region are the M8 and M6 fits, as with SCAN.

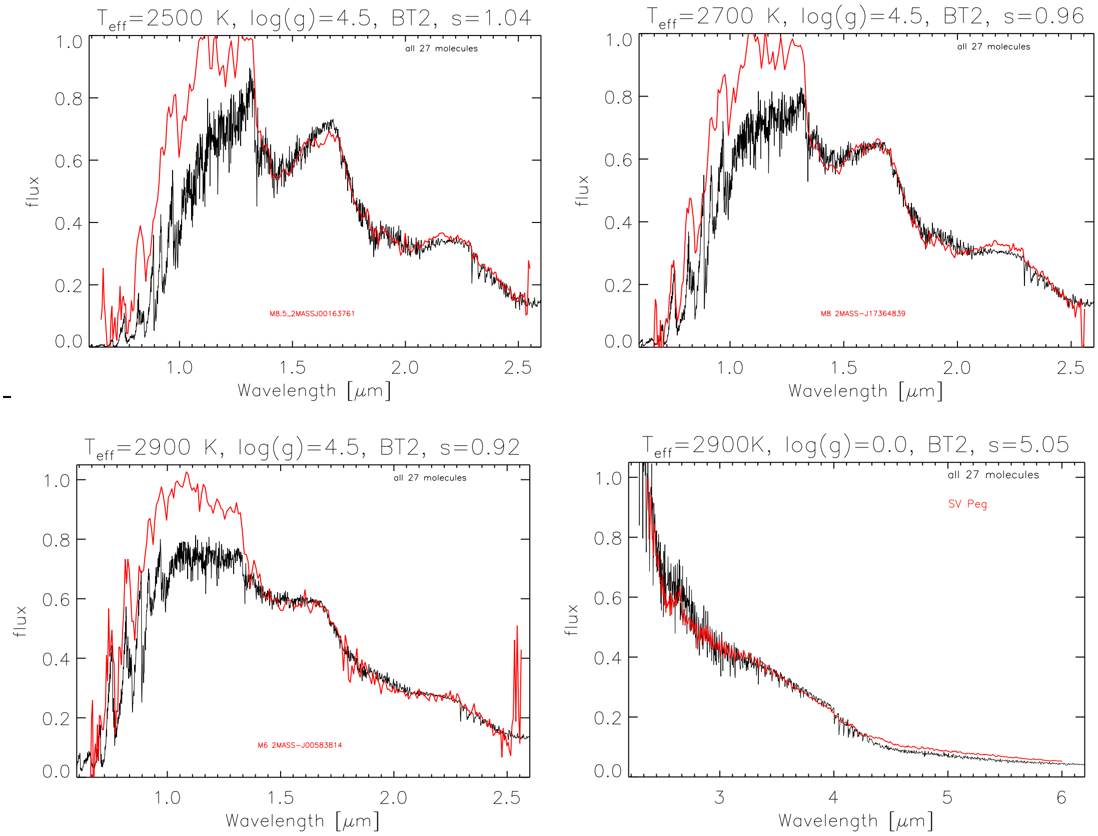

Figure 11: Observed spectra fitted with synthetic spectra. Computed with MARCS and SYNTOS with the BT2 line list. Top left; 2MASS J00163771 (M8.5). Top right: 2MASS J17364839 (M8). Bottom left: 2MASS J00583814 (M6). Bottom right: SV Peg (M7) giant. The synthetic spectra are normalized at their flux maximum and manually scaled.

4.5.3 POKAZATEL

Figure 12 shows the four fits done in the same way as previous ones, with the POKAZATEL OS. The synthetic spectra fit the observed very well in the long-wavelength region but show the same problems in the short-wavelength region as BT2. Overall, these plots bear extreme similarity to those of BT2, as was the case with OS plots (fig. 5) and atmospheric structures (fig. 9). If any differences can be pointed out, these spectra may be a very slightly worse fit in the long-wavelength region. The M8 and M6 fits are the best of the dwarf fits in this region, as with the two previous lines.

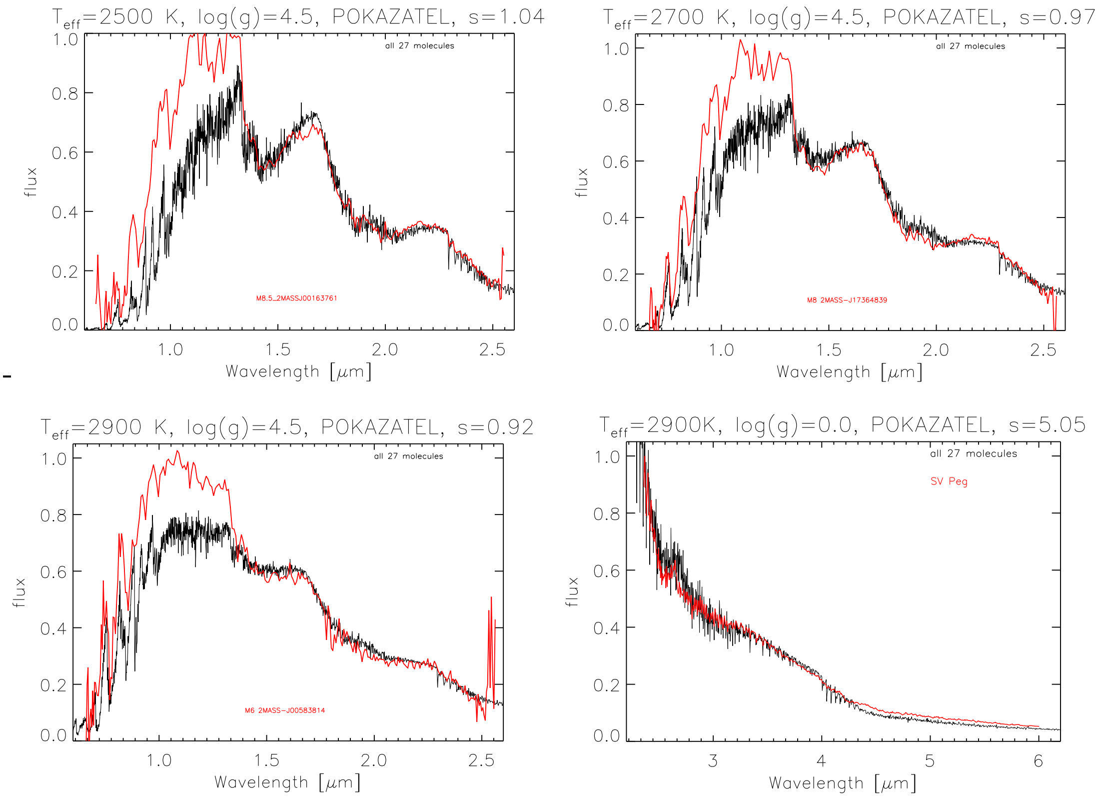

Figure 12: Observed spectra fitted with synthetic spectra. Computed with MARCS and SYNTOS with the POKAZATEL line list. Top left; 2MASS J00163771 (M8.5). Top right: 2MASS J17364839 (M8). Bottom left: 2MASS J00583814 (M6). Bottom right: SV Peg (M7) giant. The synthetic spectra are normalized at their flux maximum and manually scaled.

4.5.4 HITEMP

in Figure 13 are shown the stellar fits with the HITEMP OS. All four fit remarkably well in the region longward of 1.3 μm, with the exception of a sharp peak at 1.6 μm in the M8.5 fit. The short-wavelength region does not show as dramatic excess of absorption as BT2 and POKAZATEL. The fit of SV Peg (bottom right) is quite similar to that of BT2 and POKAZATEL, but the spectrum fits slightly better in the short-wavelength region, namely around the small peak at ≈ 2.6 μm. Like the case with the other line lists, the HITEMP spectra fit best with SV Peg and the M8 and M6 spectra longward of 1.3 μm.

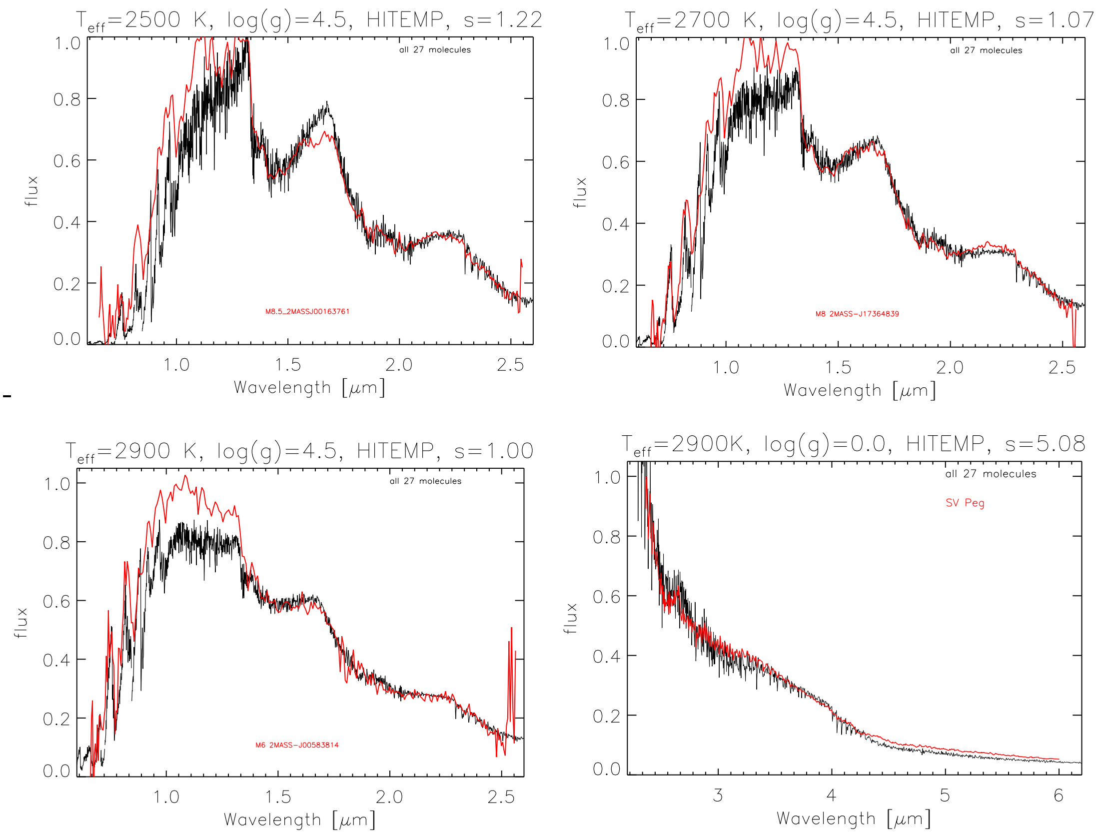

Figure 13: Observed spectra fitted with synthetic spectra. Computed with MARCS and SYNTOS with the HITEMP line list. Top left; 2MASS J00163771 (M8.5). Top right: 2MASS J17364839 (M8). Bottom left: 2MASS J00583814 (M6). Bottom right: SV Peg (M7) giant. The synthetic spectra are normalized at their flux maximum and manually scaled.

4.5.5 Water-only Spectra

Figure 14 shows synthetic spectra at 2700 K, $\log g = 4.5$, each one calculated with different line list OS files. They are scaled to match the M8 observed spectra, and compared to its spectrum, to show discrepancies in overall absorption. These spectra show how crucially important water absorption is to this spectral region, as almost the whole spectrum is shaped by water. They show that the gap at the flux peak around 1.6 μm simply follows from the water line lists (although slightly affected by CaH and OH as seen on fig. 6). They also show, most importantly, that in order to scale the synthetic spectra so that they fit one of the observed spectra in the long-wavelength region, they get quite different shapes in the short-wavelength region. Here it is very clear, why SCAN (shown top left), does not portray the excess absorption that the others do, and why HITEMP does so to a lesser degree than BT2 and POKAZATEL. Making such a plot in the 2.0-6.2 μm region is not necessary, since the synthetic spectra are largely similar for the SV Peg fits.

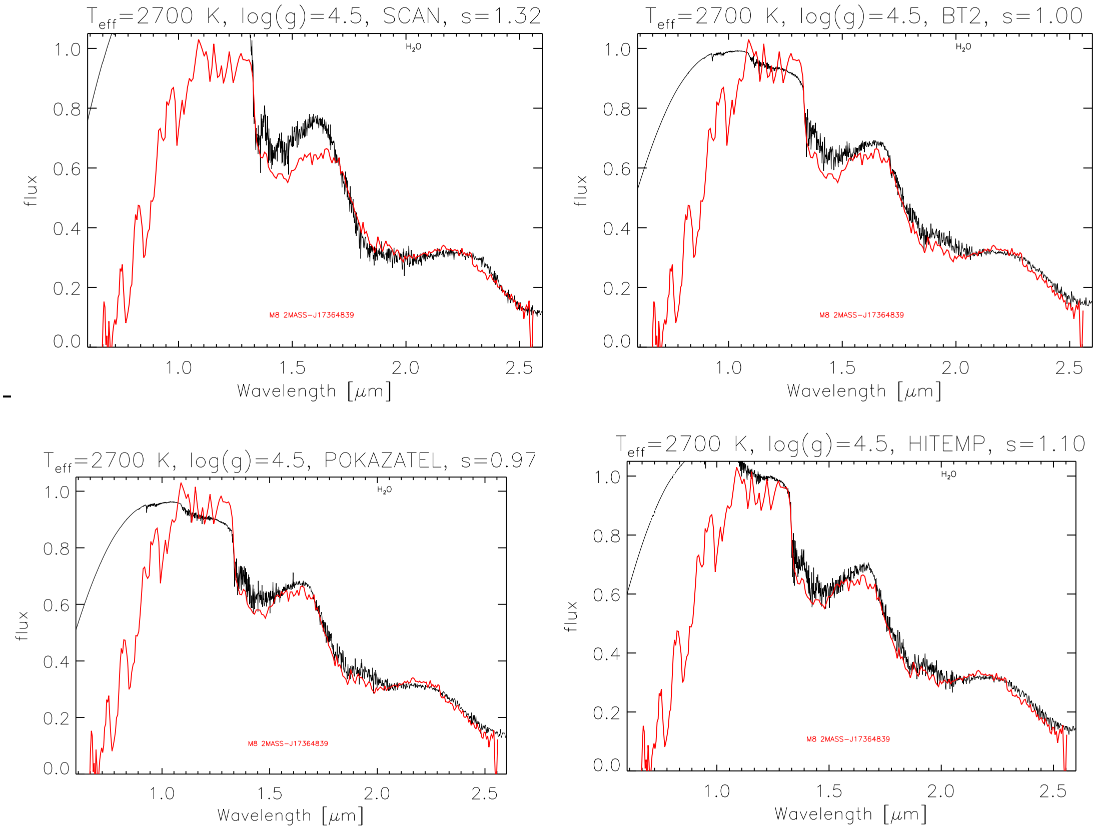

Figure 14: Four synthetic spectra with same parameters and different line lists OS files. The are scaled and compared to the observed M8 spectra of 2MASS J17364839. Top left: SCAN. Top right: BT2. Bottom left: POKAZATEL. Bottom right: HITEMP.

4.6 Results Summary

The OS coefficients in Figure 5 are very similar for BT2 and POKAZATEL, while HITEMP follows them in large wavelengths with a somewhat surprising factor stronger absorption. The SCAN OS plot is quite different from the others, while its lower cut-off is visible. This is expected, due to differences in methodology. Spectra in Figure 6 show that water absorption increases at lower temperature in both wavelength regions, lowering the total absorption, while the other molecules with significant contributions are largely unaffected. Spectra in Figure 7 show that the effects on the spectra of increasing the surface gravity are similar to slightly decreasing temperature in the short-wavelength region, while in the long-wavelength region, it lessens the contributions from other molecules. Plots of structure in Figure 8 show that computed atmosphere structures are quite similar for different surface gravities, while convection is more efficient for low surface gravity high temperature, causing more efficient heat loss from deep layers. The similarity of structures is expected, as the increased opacity causing radiation pressure is counterbalanced by the strong gravity as per hydrostatic equilibrium. As is the more efficient convection in low surface gravity. Line list structure comparison in Figure 9 shows similar structures, while SCAN has a more extended, cooled atmosphere, possibly due to a stronger representation of weak lines. The spectral fits in Figures 10, 11, 12, 13 show a good overall fit for SCAN, while BT2, HITEMP, and POKAZATEL show a very good fit longward of 1.3 μm, and a match of the overall shape of the short-wavelength region, while the short and long-wavelength regions requires different scaling to fit the observed spectra. The reason is seen clearly on water-only spectra in Figure 14, where the short-wavelength region have completely different shapes at their best-fit scaling. This is quite surprising, as these lists are large, published lists, that have been tested in other applications. A possible reason is the differences in line list construction methods.

5 Discussion

In this section, the underlying assumptions when fitting, the similarity of the line lists, and the 1.3 μm discrepancy will be discussed. After that, options for how to acknowledge and proceed around the currently unsolved problems in this subject are proposed. Finally, an outline of the prospects for atmospheric analysis is given.

5.1 Assumptions Regarding Fits

Since $\log g$ makes quite little impact on the spectra, as shown in Figure 7, whereas it has a larger effect on the atmospheric structure, as shown in Figure 8, it should not be treated as a fitting parameter for the spectra, when an approximate value can easily be guessed from information about the star, and the exact value within that range does not matter much for the model. For example, since it is known that SV Pegasus is a red giant, one must assume that its surface gravity is close to $\log g = 0.0$, and there is no gain from trying to obtain a more accurate value from fitting. For this reason, surface gravity has in some places been varied in steps of 0.5 at a time, to see if it made a beneficial difference within a reasonable range. The manual scaling method should not introduce any error or unfair comparison. In either case, the water-only spectra have been scaled in the same way, and display the same characteristics seen with the spectra containing all 27 molecules.

5.2 Similarity of BT2, POKAZATEL, and HITEMP

The synthetic spectra made with OS files computed from HITEMP, BT2, and POKAZATEL line lists, look qualitatively quite similar, especially the latter two. However, the plots examine a small region of wavelengths only, although larger than possible with a laboratory experiment. Recall that I have chosen red dwarfs to showcase the use of MARCS in cool exoplanet-like atmospheres while choosing a temperature range, where water dominates the absorption. The similarity of the line lists is also displayed on the OS-plot in Figure 5. There, BT2 and POKAZATEL coincide or lie very close, while HITEMP lies above them, which alludes to a possible error in scaling of the HITEMP line lists in conversion from its downloaded form, or perhaps a consequence of extrapolation from fitted values in the construction of HITEMP. We would expect HITEMP to coincide with BT2 at the intervals, where BT2 was used as the main water input.

Nevertheless, the water-only spectra for HITEMP show too little absorption in the short wavelength for HITEMP, so such an error does not critically affect my results.

5.3 Splitting the Spectrum

All the presented line lists would fit better if the plots were separated, shortward and longward of 1.3 μm, and the two halves were scaled differently. This is, however, exactly the point of using this kind of data for testing line list performance. This effect may be due to different handling of the weak lines, and the differing degree of empirical adjustment in the calculation when constructing line lists. This would explain why SCAN behaves so differently, being the only list to fit both regions somewhat well. In our case, the dwarfs provide more valuable insight, as water lines are more dominant in their higher atmospheric pressure. The comparison carried out in Rothman et al. [2010] was in the range of 1.3-2.3 μm. It is natural for these types of experiments to focus on a narrow spectral range, but perhaps, it should be standard to use spectral fitting as a supplementary test in that case. The occurrence of two regions, where one part of the spectrum must fit poorly, can not be the effect of a simple scaling error, as the error would behave differently for different wavenumbers while having two regimes that fitted well on their own.

5.4 Settling the Differences

I have discussed how the approach in constructing a line list can differ, and how these should be tested. There, several options were outlined, that are important to the final result but seem to vary among researchers. I will now go through these issues and discuss them.

5.4.1 Partition Sum

There is no clear way to best construct a complete partition sum. I have described the outline of how one usually tries in practice, but there is no concrete agreement because every method will have its flaws. One can not simply integrate all the lines in one's list, since these lines need quantum numbers. As discussed, the semi-classical quantum mechanics used is inconsistent with the high energy levels one needs to compute a complete water line list, and have it be capable of correctly modeling the atmosphere energy transport. The best one can do is for now is to provide clear information about the method used or the partition function adopted in one's work.

5.4.2 Adjusting to Data

When is adjusting line intensities to laboratory data the best option, and when does it affect the further use of the line lists negatively? If one tries to measure the many weak lines in a laboratory sensor, using a detector, one's data will be integrated over some timescale, and the absorption of the strongest lines contained in the data will be measured. Adjusting line position to data is harmless enough while adjusting intensities can shift the behavior of many weak lines not directly visible in the data, but which affect the atmosphere structure, when the list is used for modeling. If it were simpler to construct many line lists for a molecule like water quickly, this could be tested through varying the degree of manipulation and testing the resulting line list with somewhat standardized experimental data and stellar spectra.

5.4.3 Line Shape and Absorption Coefficient

What is the best approach to representing the absorption coefficient, and how does this information affect the line shape when creating a spectrum? In recent similar work, the correlated k method, described in section 2.6, is sometimes used. This method can potentially misrepresent the weak lines, and a switch to OS from ODF in MARCS was made because ODF did not perform as well. The benefit of using OS is that strong and weak lines are represented proportionally to the quantity of them and that the opacity per gram of stellar material can be added linearly. Of course, when a sampled wave number is assigned an absorption strength, this will be "wrong" for their neighbours, but when enough points are sampled randomly, it will produce a sample that is both representative of the line population and includes the desired effects of weak and strong lines in the model and spectrum shape.

5.4.4 Best Test for Line Lists

There is no clear agreement that comparing synthetic and observed spectra is the best way of testing line lists, as some prefer laboratory data. As discussed, the method of comparing with laboratory tests is useful, but it is not thorough, since it does not reflect the physics of a real celestial body, where the entire thermodynamics reacts to the energy transport. Comparing with stellar objects also has the bonus, that it improves our capabilities of reproducing stellar spectra of different temperatures.

5.5 Future Prospects for Atmospheric Analysis

Atmospheric modeling is a heavily researched subject, and though it is very complex, it is still quite simple compared to the physical and chemical processes that occur in and around a large celestial body. Major areas for improvement include sphericity of the models, better treatment of convection, solving chemical abundances without the assumption of LTE, the modeling of disequilibrium chemistry, and an entire treatment of radiative transfer in a planet with a host star. Disequilibrium may be a key feature in planets containing life, since organisms through their metabolic processes or otherwise, affect the composition of the atmosphere continuously. Sphericity, partly through a different distribution of incoming radiation, affects temperature structure. This is shown for supergiants in comparison with plane-parallel models in Gustafsson et al. [2008]. These are all very complicated issues, and will likely take a long time and the combined effort of many researchers to solve. Perhaps a more immediate improvement would be to have a consensus on the use of partition functions and opacity function methods, as well as having fits of stellar spectra be a standard test, while generally including as many opacity sources as possible at a wide temperature range [Gustafsson et al., 2008]. With the advent of GGchem, adding the ability to calculate equilibrium chemistry down to 100 K [Woitke et al., 2018], and cloud formation programs, MARCS will be able to model increasingly cool atmospheres. When the aforementioned problems are solved or approximated, MARCS will finally be able to model planetary atmospheres. Until then, the $T - P_g$ profile we obtain will not have the characteristic features discussed in section 2.1.

6 Conclusion

In this work, I have aimed to examine the behavior of modern available line lists for $^{1}H_2{}^{16}O$ in atmosphere modeling with MARCS and determine which are best suited for application in cool atmospheres. I have collected and converted line lists for $^{1}H_2{}^{16}O$: BT2, POKAZATEL, and HITEMP to the format used by OS_EXOMOL, and calculated opacity sampling files for each. I have used those opacity sampling files, along with that of SCAN, in MARCS model atmospheres, to examine the difference in temperature-pressure profile, and to compare their ability to reproduce the spectra of three late-type M-dwarfs and one M-giant. I have computed spectra with SYNTOS for different molecular species, to examine their effect on the spectra in different wavelength regions, and how their opacities are affected by temperature and surface gravity. My plots showed that surface gravity has an effect on the spectra that is equivalent to very slightly lowering the temperature. For our dwarf fitting spectral range, we see that our spectra are most affected by TiO and CaH in 0.6-1.3 μm, and dominated by water H_2O in 1.3-2.6 μm. For our giant spectral range, I showed that water dominates the entire spectrum, with a strong effect in low surface gravity of CO and SiO, being muted at higher surface gravity. I examine temperature-pressure profiles of models, calculated with each line list, and with different surface gravities, for the SCAN line list, and different inclusions of strong and weak lines. I show that convection is more efficient in low surface gravity models, and SCAN shows a generally more extended and cooled upper atmosphere compared to the others. I collect three M-dwarfs from the SpeX prism library and a spectrum from SWS of an M-giant, calculate models in MARCS with each line list for water and obtain spectra to fit with these cool stars. In my results, SCAN displays the capability of roughly fitting to the entirety of the spectral range, while the remaining line lists show similar results, displaying two regions on each side of 1.3 μm, unable to fit the stellar spectrum simultaneously while all three are able to fit the long-wavelength region very well. I discuss the reason for this problem and the possibility that it stems from risky use of fitting when constructing the line lists, and may have been avoided if prior tests were not constrained to laboratory experiments. I have also discussed the benefits of opacity sampling, and the unsolved problem of how to best calculate the internal partition sum and assign quantum numbers. In the future, it would be best for researchers to explain the details of their approach to this problem, or for the community around this type of research to reach a consensus, or standard, on this topic.

7 Acknowledgements

A big thank you to Diana Juncher for letting me continue the work started by her in her unpublished paper on the same subject, and for helping me reduce OS calculation times. Also to Rune Kjærsgaard, for providing scripts for reading and converting BT2 and POKAZATEL line lists. Lastly, like to thank my supervisor, Uffe Gråe Jørgensen for all his help, and for providing the bulk of the data and code needed for this work.

8 Bibliography

A. F. Al-Refaie et al. Exomol line lists – viii. a variationally computed line list for hot formaldehyde. *Monthly Notices of the Royal Astronomical Society*, 448:1704–1714, 2015. doi: 10.1093/mnras/stv091.

B. Aringer, F. Kerschbaum, and U. G. Jørgensen. H_2o in stellar athmospheres | ii. iso spectra of cool red giants hydrostatic models. *A%A*, 395(3):915–927, 2002. doi: 10.1051/0004-6361:20021313.

R. Barber, J. Tennyson, G. Harris, and R. N. Tolchenov. A high-accuracy computed water line list. *Monthly Notices of the Royal Astronomical Society*, 386:1087–1094, 2006. doi: 10.1111/j.1365-2966.2006.10184.x.

E. J. Barton et al. Exomol molecular line lists v: The ro-vibrational spectra of nacl and kcl. *Monthly Notices of the Royal Astronomical Society*, 442:1821–1829, 2014. doi: 10.1093/mnras/stt1105.

P. F. Bernath. Mollist: Molecular line lists, intensities and spectra. *University of S*, 2019. link: Elsevier.

D. M. Bittner and P. F. Bernath. Line lists for lif and licl in the x $1\sigma+$ ground state. *ApJSS*, 235:8, 2018. doi: 10.3847/1538-4365/aa9846.

S. Bryson et al. The occurrence of rocky habitable zone planets around solar-like stars from kepler data. *AJ*, 161(36):36, 2021. link: IOP, doi: 10.3847/1538-3881/abc418.

A. J. Burgasser. The spex prism library: 1000+ low-resolution, near-infrared spectra of ultracool m, l, t and y dwarfs. *ASI Conference Series*, 10:1–10, 2014. Link: SpeX Prism Library. arXiv: 1406.4887v1.

A. J. Burgasser et al. The 2mass wide-field t dwarf search. iii. seven new t dwarfs and other cool dwarf discoveries. *The Astronomical Journa*, 127:2856–2870, 2014. doi: 10.1086/383549.

D. C. Catling. *Atmospheric Evolution on Inhabited and Lifeless Worlds.* Cambridge University Press, 2017. doi: 10.1017/9781139020558.

A. Coppalle and P. Vervisch. Spectral emissivities of h_2o vapor at 2900 k in the 1-9-μm region. *JQSRT*, 35:121–125, 1986. doi: 10.1016/0022-4073(86)90108-1.

M. C. Cushing et al. The discovery of y dwarfs using data from the wide-field infrared survey explorer (wise). *ApJ*, 743(50):17, 2011. doi: 10.1088/0004-637X/743/1/50.

D. Darby-Lewis et al. Synthetic spectra of beh, bed and bet for modelling of emission from the jet plasma. *J. Phys. B: At. Mol. Opt. Phys*, 51, 2018. doi : 10.1088/1361-6455/aad6d0.

T. de Graauw et al. Observing with the iso short-wavelength spectrometer. *A&A*, 315: L49–L54, Nov 1996. Bibcode: 1996A&A...315L..49D.

D. J. Frohman, P. F. Bernath, and J. S. Brooke. Molecular line lists: The rovibrational spectra of naf and kf. *JQSRT*, 169:104–110, 2016. doi: 10.1016/j.jqsrt.2015.10.004.

D. F. Gray. *The Observation and Analysis of Stellar Photospheres (Third edition)*. Cambridge University Press, 2005.

B. Gustaffson, R. Bell, K. Eriksson, and Nordlund. A grid of model atmospheres for metal-deficient giant stars i. *A&A*, 42:407–432, 1975. ads bibcode: 1975A&A....42..407G.

B. Gustaffson, R. Bell, K. Eriksson, and Nordlund. A description of the correlated k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. *Journal of Geophysical Research*, 96(D5):9027–9063, 1991. doi: 10.1029/90JD01945.

B. Gustafsson, B. Edvardsson, K. Eriksson, U. G. Jørgensen, A. Nordlund, and B. Plez. A grid of marcs model atmospheres for late-type stars i. methods and general properties. *Astronomy & Astrophysics*, 486:951–970, 2008. doi: 10.1051/0004-6361:200809724, arXiv:0805.0554.

C. Helling. Exoplanet clouds. *Annual Review of Earth and Planetary Sciences*, 47: 583–606, 2019. doi: 10.1146/annurev-earth-053018-060401.

C. Helling et al. The mineral clouds on hd 209458b and hd 189733b. *Monthly Notices of the Royal Astronomical Society*, 460(1):855–883, 2016. doi: 10.1093/mnras/stw662.

B. Hesman, G. R. David, and D. A. Naylor. *The Abundance of Carbon Monoxide in Neptune's Atmosphere*. University of Saskatchewan, 2005. Link: Researchgate.

R. C. Hilborn. Erratum: Einstein coefficients, cross sections, f values, dipole moments, and all that. *American Journal of Physics*, 51(471), 1983. doi: 10.1119/1.13515.

S. Hou and P. F. Bernath. Line list for the ground state of caf. *JQSRT*, 210:44–51, 2017a. doi: 10.1016/j.jqsrt.2018.02.011.

S. Hou and P. F. Bernath. Line list for the mgf ground state. *JQSRT*, 203:511–516, 2017b. doi: 10.1016/j.jqsrt.2017.03.01.

H. R. Jones, Y. Pavlenko, S. Viti, and J. Tennyson. A comparison of water vapour line lists. *Cambridge workshop on cool stars, UC 2003*, 2003. Link.

D. Juncher, U. G. Jørgensen, and C. Helling. Self-consistent atmosphere modeling with cloud formation for low-mass stars and exoplanets. *Astronomy & Astrophysics*, 608 (A70):17, 2017. doi: 10.1051/0004-6361/201629977.

U. G. Jørgensen. in *Astrophysical Applications of Powerful New Databases,* ed. S.J. Adels & W. L. Wiese. ASP Conf. Ser. 78, 170, 1995.

U. G. Jørgensen and P. Jensen. The dipole moment surface and the vibrational transition moments of h2o. *Journal of Molecular Spectroscopy*, 161(1):219–242, 1993. doi: 10.1006/jmsp.1993.1228.

U. G. Jørgensen and H. R. Johnson. Radiative force on molecules and its possible role for mass loss in evolved agb stars. *A&A*, 265:168–176, 1992. doi: 10.1146/annurev-earth-053018-060401.

U. G. Jørgensen, P. Jensen, G. Sørensen, and B. Aringer. H2o in stellar atmospheres. *Astronomy & Astrophysics*, 372(1):249–259, 2001. doi: 10.1051/0004-6361:20010285 .

J. D. Kirkpatrick et al. Discoveries from a near-infrared proper motion survey using multi-epoch two micron all-sky survey data. *ApJS*, 190:100–146, 2010. doi: 10.1088/0067-0049/190/1/100.

J. D. Kirkpatrick et al. Further defining spectral type "y" and exploring the low-mass end of the field brown dwarf mass function. *ApJ*, 753(156):38, 2012. doi: 10.1088/0004-637X/753/2/156.

R. K. Kupparapu et al. Habitable zones around main-sequence stars: Dependence on planetary mass. *ApJL*, 787(2):L29, 2021. doi: 10.1088/2041-8205/787/2/L29.

G. Li et al. Direct fit of experimental ro-vibrational intensities to the dipole moment function: application to hcl. *JQSRT*, 112:1543–1550, 2011. doi: 10.1016/j.jqsrt.2011.03.014.

M. M. K. Liliana Mammino. *Frontiers in Computational Chemistry - Chapter 6 - Considering the Medium when Studying Biologically Active Molecules: Motivation, Options and Challenges*. Bentham Books, 2015. doi: 10.1016/B978-1-60805-865-5.50006-X, link to specific part retrieved feb 9 2022.

E. Manjavacas et al. New constraints on the formation and settling of dust in the atmospheres of young m and l dwarfs. *A&A*, 564(A55):21, 2014. doi: 10.1051/0004-6361/201323016 .

P. Matsyutenko, M. Grechko, T. R. Rizzo, O. V. Boyarkin, et al. State-selective spectroscopy of water up to its first dissociation limit. *J. Chem. Phys.*, 131(221105), 2009. doi: 10.1063/1.3273207.

D. Mihalas. *Stellar Atmospheres (Second edition)*. W. H. Freeman and Company, 1978.

NASA. *The Sun as an X-ray Source*. 2014. Link to NASA webpage, retrieved feb 10 2022.

NASA. *Strong Evidence For Coronal Heating Theory Presented at 2015 TESS Meeting*. 2015. Link to NASA webpage, retrieved feb 10 2022.

A. Owens et al. The rotation–vibration spectrum of methyl fluoride from first principles. *Physical Chemistry Chemical Physics*, 21:3496–3505, 2019. doi: 10.1039/C8CP01721B.

H. Partridge and D. W. Schwenke. The determination of an accurate isotope dependent potential energy surface for water from extensive *ab initio* calculations and experimental data. *J. Chem. Phys.*, 106(4618), 1997. doi: 10.1063/1.473987.

A. T. Patrascu et al. Exomol molecular line lists – ix. the spectrum of alo. *Monthly Notices of the Royal Astronomical Society*, 449:3613–3619, 2015. doi: doi.org/10.1093/mnras/stv507.

G. Paulose et al. Exomol molecular line lists – xii. line lists for eight isotopologues of cs. *Monthly Notices of the Royal Astronomical Society*, 454:1931–1939, 2015. doi: 10.1093/mnras/stv1543.

A. I. Pavlyuchko et al. Exomol molecular line lists – xi. the spectrum of nitric acid. *Monthly Notices of the Royal Astronomical Society*, 452:1702–1706, 2015. doi: 10.1093/mnras/stv1376.

O. L. Polyansky, A. A. Kyuberis, N. F. Zobov, J. Tennyson, S. N. Yurchenko, and L. Lodi. Exomol molecular line lists xxx: a complete high-accuracy line list for water. *Monthly Notices of the Royal Astronomical Society*, 480:2597–2608, 2018. doi: 10.1093/mnras/sty1877.

A. Popovas and U. G. Jørgensen. Partition functions i. improved partition functions and thermodynamic quantities for normal, equilibrium, and ortho and para molecular hydrogen. *Astronomy & Astrophysics*, 595(A130):23, 2016. doi: 10.1051/0004-6361/201527209.

L. Prajapat et al. Exomol molecular line lists – xxiii. spectra of po and ps. *Monthly Notices of the Royal Astronomical Society*, 472:3648–3658, 2017. doi : 10.1093/mnras/stx2229.

R. S. Ram et al. Einstein a-values and oscillator strengths of the a2–x2σ+ system of cp. *JQSRT*, 138:107–115, 2014. doi: 10.1016/j.jqsrt.2014.01.030.

T. Rivlin et al. Exomol molecular line lists – x. the spectrum of sodium hydride. *Monthly Notices of the Royal Astronomical Society*, 451:634–638, 2014. doi: 10.1093/mnras/stv979.

L. Rothman, I. Gordan, R. Barber, et al. Hitemp, the high-temperature molecular spectroscopic database. *JQSRT*, 111:2139–2150, 2010. doi: 10.1016/j.jqsrt.2010.05.001.

L. S. Rothman, C. P. Rinsland, A. Goldman, et al. The hitran molecular spectroscopic database and hawks (hitran atmospheric workstation): 1996 edition. *JQSRT*, 60(5): 665–710, 1998. doi: 10.1.1.560.3034 .

S. Shirin, O. L. Polyansky, N. F. Zobov, P. Barletta, and J. Tennyson. Spectroscopically determined potential energy surface of $h_2^{16}o$ up to 25000cm^{-1}. *J. Chem. Phys.*, 118: 2124–2129, 2003. doi: 10.1063/1.1532001.

A. J. Skemer, C. V. Morley, et al. The first spectrum of the coldest brown dwarf. *ApJL*, 826(2):L17 (5pp), 2016. doi: 10.3847/2041-8205/826/2/L17.

M. F. Skrutskie et al. The two micron all sky surver (2mass). *AJ*, 131(2):1163–1183, 2006 February. link: IOP.

C. Sousa-Silva et al. Exomol line lists – vii. the rotation–vibration spectrum of phosphine up to 1500 k. *Monthly Notices of the Royal Astronomical Society*, 446:2337–2347, 2014. doi: 10.1093/mnras/stu2246.

S. A. Tashkun. High temperature water line lists: comparison and validation. *Proc. of SPIE, 6580*, 2006. doi : 10.1117/12.724763.

J. Tennyon et al. Iupac critical evaluation of the rotational-vicrational spectra og water vapor, part iii: Energy levels and transition wavenumbers for $h_2^{16}o$. *JQSRT*, 117:29–58, March 2013. doi: 10.1016/j.jqsrt.2012.10.002.

D. S. Underwood et al. Exomol molecular line lists – xiv. the rotation–vibration spectrum of hot so2. *Monthly Notices of the Royal Astronomical Society*, 459:3890–3899, 2016. doi: 10.1093/mnras/stw849.

A. Upadhyay et al. Exomol line lists xxv: a hot line list for silicon sulphide, sis. *Monthly Notices of the Royal Astronomical Society*, 447:1520–1527, 2018. doi: 10.1093/mnras/sty998.

A. R. Whitehill. Vibronic origin of sulfur mass-independent isotope effect in photoexcitation of so2 and the implications to the early earth's atmosphere. *Proceedings of the National Academy of Sciences*, 110:17697–17702, 2013.

P. Woitke, C. Helling, et al. Equilibrium chemistry down to 100 k. *Astronomy & Astrophysics*, 604(A1), 2018. doi: 10.1051/0004-6361/201732193.

B. Yadin et al. Exomol line lists - i. the rovibrational spectrum of beh, mgh and cah in the x2+ state. *Monthly Notices of the Royal Astronomical Society*, 425:34–43, 2012. doi : 10.1111/j.1365- 2966.2012.21367.x.

L. Yorke et al. Exomol molecular line lists – vi. a high temperature line list for phosphorus nitride. *Monthly Notices of the Royal Astronomical Society*, 445:1383–1391, 2014. doi : 10.1093/mnras/stu1854.

M. Yousefi and P. F. Bernath. Line lists for alf and alcl in the x1 ground state. *ApJS*, 273:8, 2018. doi : 10.3847/ 1538- 4365/aacc6a.

S. N. Yurchenko, R. J. Barber, J. Tennyson, et al. A variationally computed line list for hot nh3. *Monthly Notices of the Royal Astronomical Society*, 413:1828–1834, 2013a. 10.1111/j.1365- 2966.2011.18261.x.

S. N. Yurchenko, J. Tennyson, et al. Vibrational transition moments of ch4 from first principles. *Journal of Molecule Spectroscopy*, 291:69–76, 2013b.

S. N. Yurchenko, W. Bond, et al. Exomol molecular line lists – xxvi: spectra of sh and ns. *Monthly Notices of the Royal Astronomical Society*, 478:270–282, 2018a. doi : 10.1093/mnras/sty939.

S. N. Yurchenko, H. Williams, et al. Exomol linelists xxviii: the rovibronic spectrum of alh. *Monthly Notices of the Royal Astronomical Society*, 479:1401–1411, 2018b. doi: 10.1093/mnras/sty1524.

T. J. Zielinski et al. *Quantum States of Atoms and Molecules.* Libretexts, Updated 2021. Link: Chapter 10.1, Libretexts.

A Appendix

A.1 Line List References for Molecules Included by Rune Kjærsgaard

Molecule	Reference
AlCl	Yousefi and Bernath [2018]
AlF	Yousefi and Bernath [2018]
AlH	Yurchenko et al. [2018b]
AlO	Patrascu et al. [2015]
BeH	Yadin et al. [2012], Darby-Lewis et al. [2018]
CaF	Hou and Bernath [2017a]
CH3F	Owens et al. [2019]
CH_4	Yurchenko et al. [2013b]
CP	Ram et al. [2014]
CS	Paulose et al. [2015]
H_2CO	Al-Refaie et al. [2015]
HCl	Li et al. [2011]
HNO_3	Pavlyuchko et al. [2015]
KCl	Barton et al. [2014]
KF	Frohman et al. [2016]
LiCl	Bittner and Bernath [2018]
LiF	Bittner and Bernath [2018]
MgF	Hou and Bernath [2017b]
NaCl	Barton et al. [2014]
NaF	Frohman et al. [2016]
NaH	Rivlin et al. [2014]
NH_3	Yurchenko et al. [2013a]
NS	Yurchenko et al. [2018a]
PH_3	Sousa-Silva et al. [2014]
PN	Yorke et al. [2014]
PO	Prajapat et al. [2017]
PS	Prajapat et al. [2017]
SH	Yurchenko et al. [2018a]
SiS	Upadhyay et al. [2018]
SO_2	Underwood et al. [2016], Whitehill [2013]

Table 1: List of references for the 30 molecules added to MARCS by Rune Kjærsgaard.

A.2 Dwarf Plots in 1.3-2.3 μm

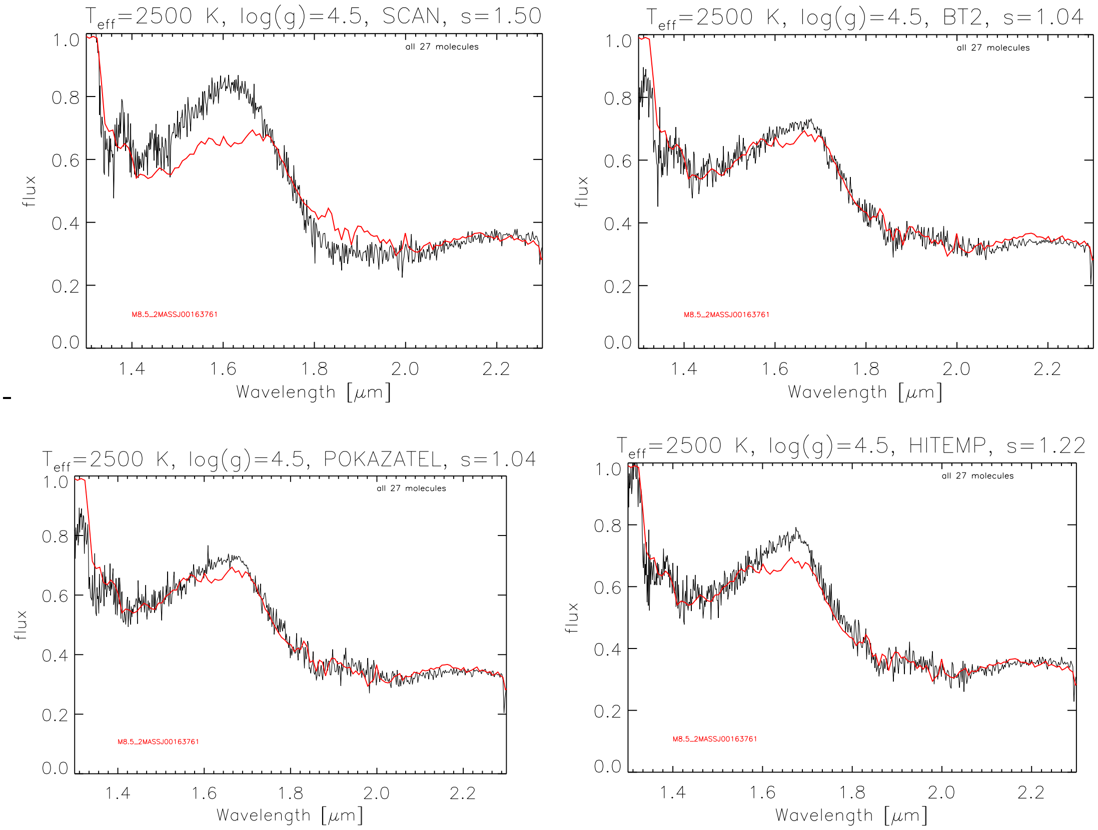

Figure 15: M8.5 dwarf plots limited to the 1.3-2.3 μm region, where manual scaling was chosen to fit this region.

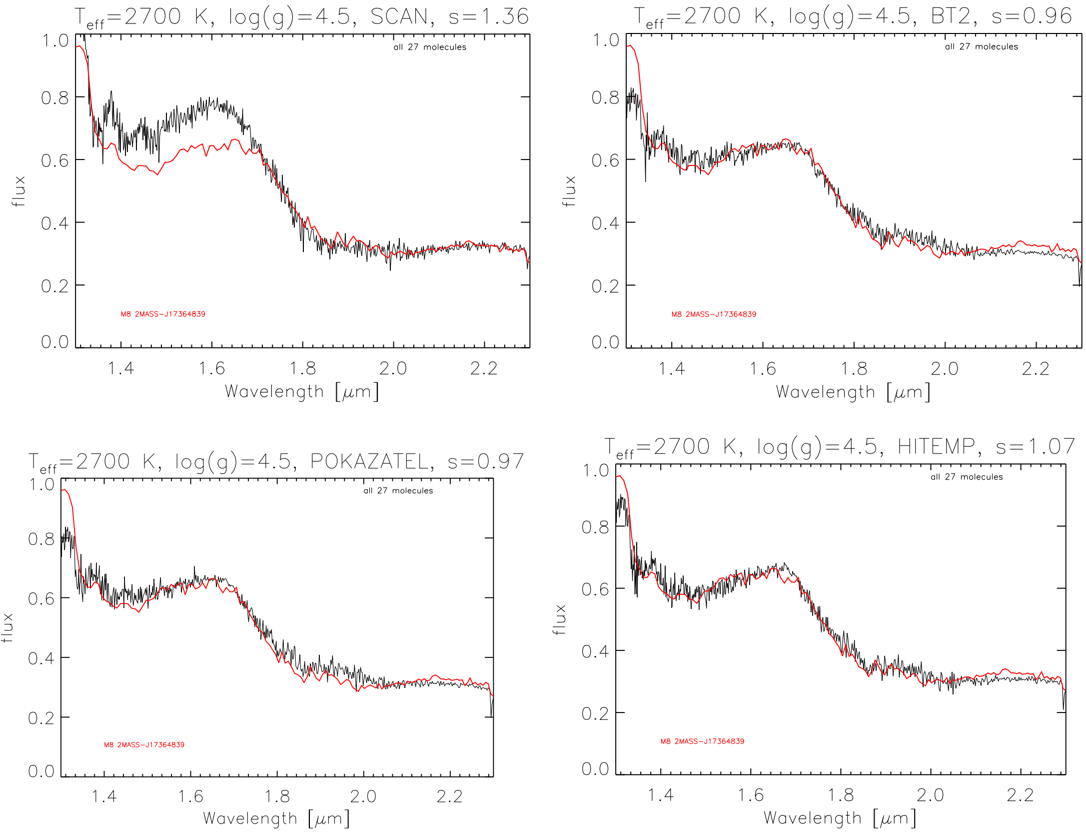

Figure 16: M8 dwarf plots limited to the 1.3-2.3 μm region, where manual scaling was chosen to fit this region.

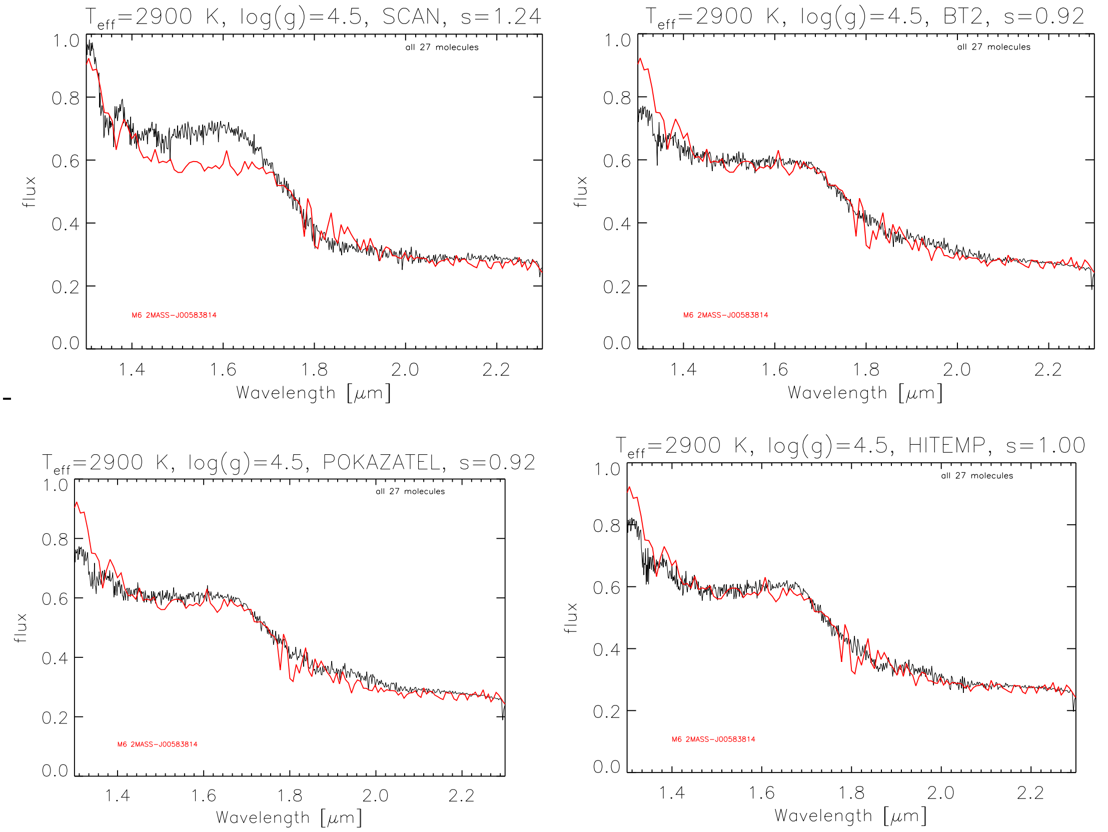

Figure 17: M6 dwarf plots limited to the 1.3-2.3 μm region, where manual scaling was chosen to fit this region.

A.3 SV Peg Plots in 2.2-3.8 μm

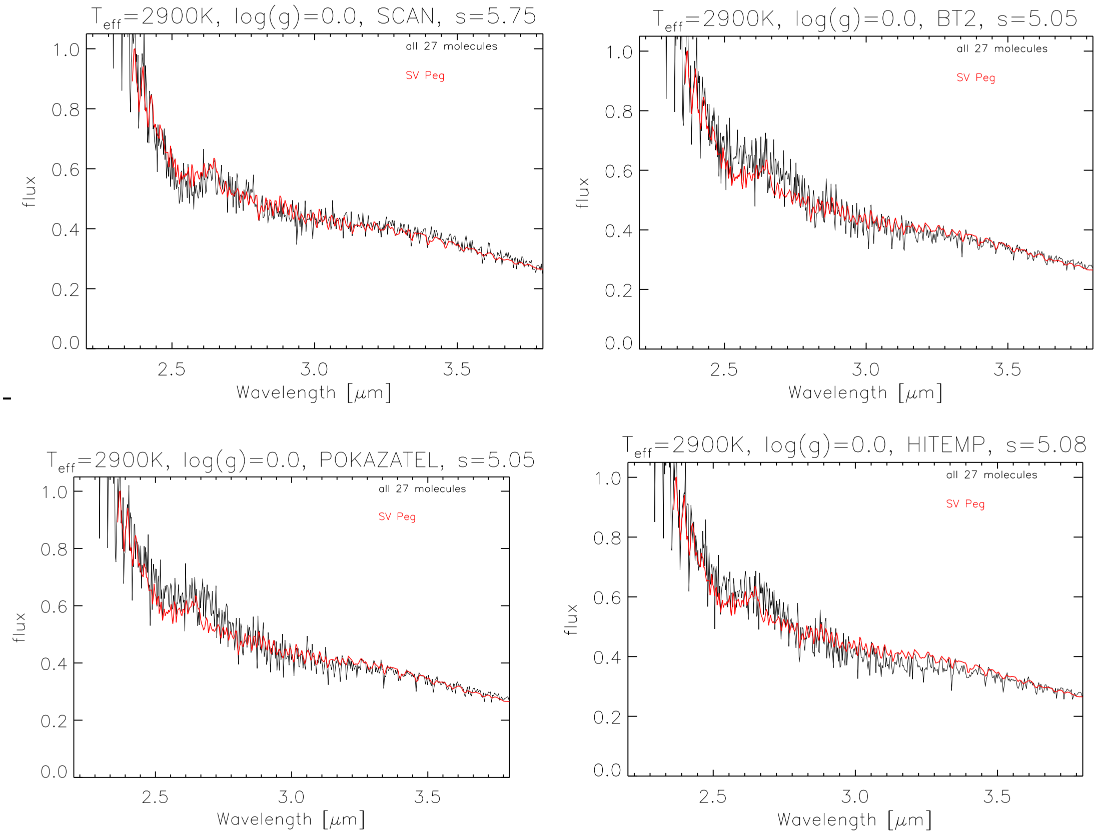

Figure 18: SV Peg plots limited to the 2.0-3.8 μm region, where manual scaling was chosen to fit this region.

A.4 Comparative Plots (full)

A.4.1 M8.5

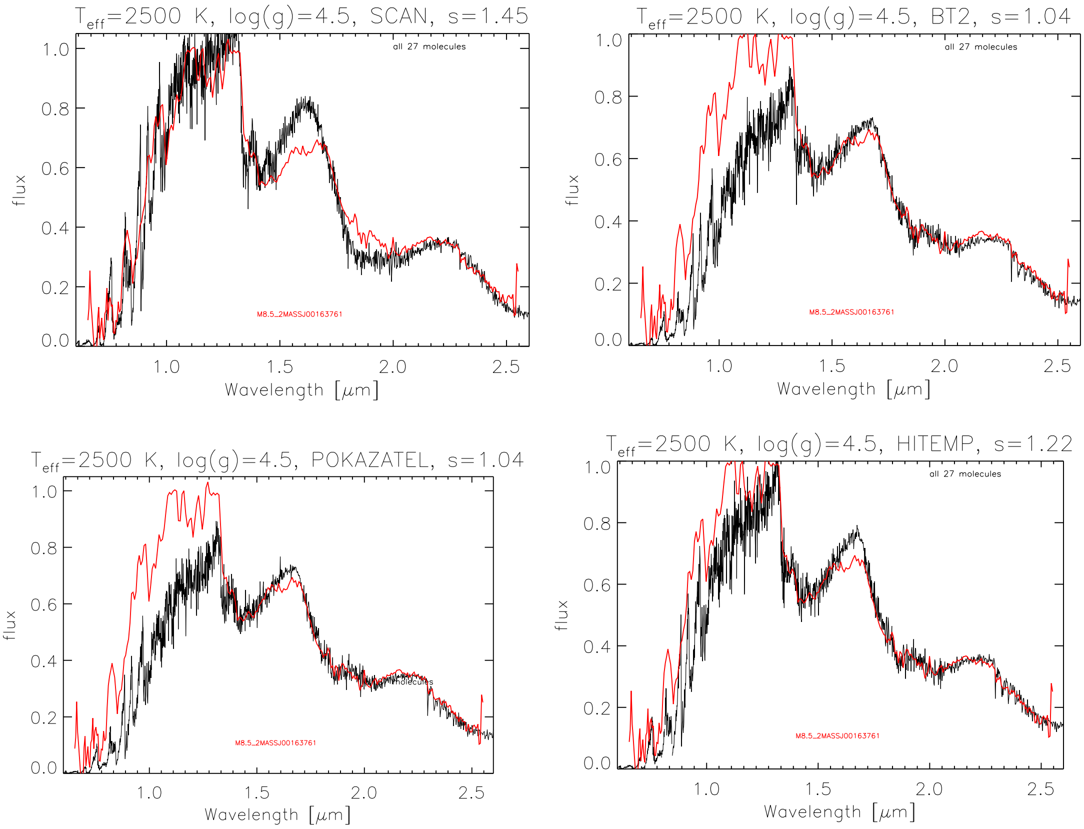

Figure 19: Comparative plot of the stellar fits for the M8.5 dwarf. Plots are the same as the stellar fits in 4.5, but line list performance is shown side-by-side.

A.4.2 M8

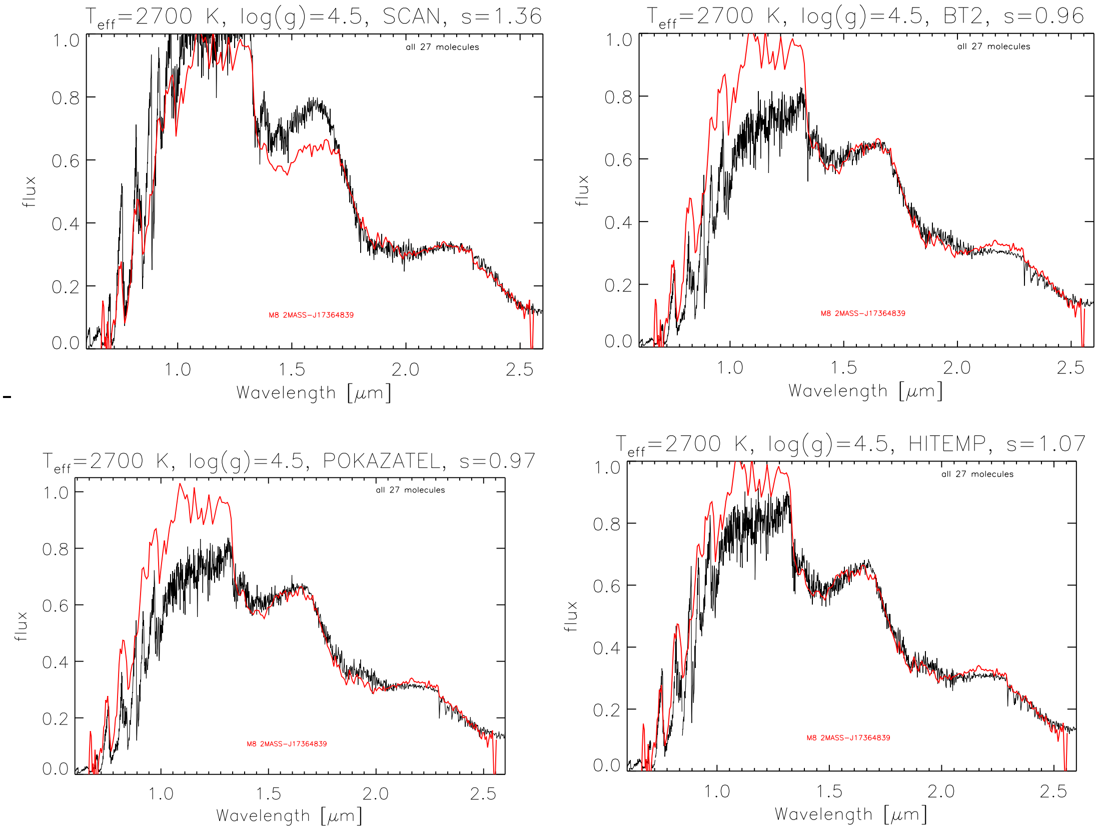

Figure 20: Comparative plot of the stellar fits for the M8 dwarf. Plots are the same as the stellar fits in 4.5, but line list performance is shown side-by-side.

A.4.3 M6

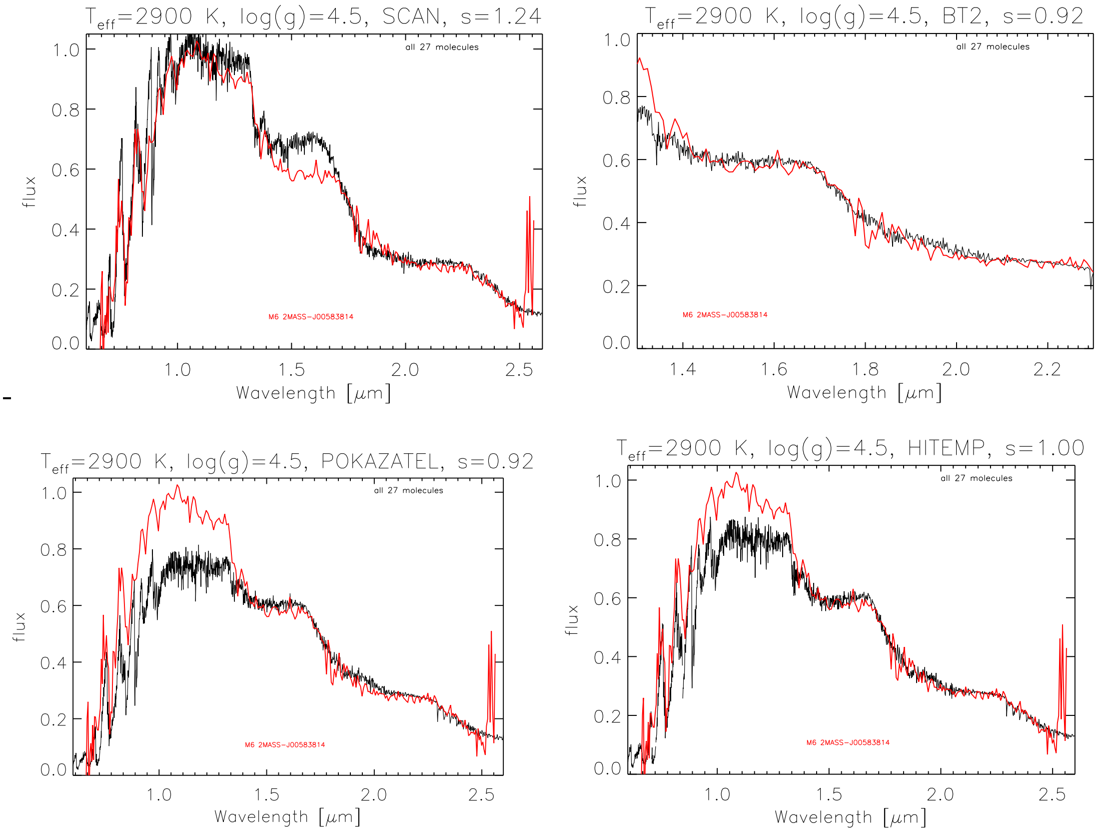

Figure 21: Comparative plot of the stellar fits for the M6 dwarf. Plots are the same as the stellar fits in 4.5, but line list performance is shown side-by-side.

A.4.4 SV Peg

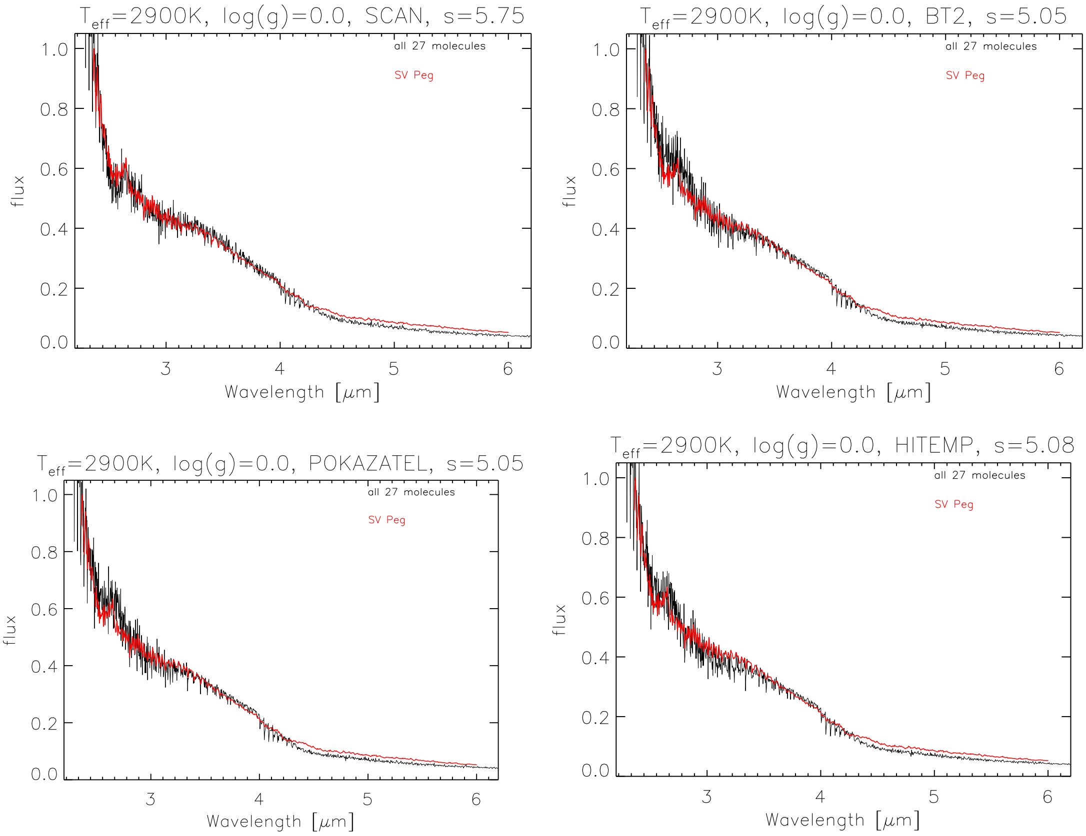

Figure 22: Comparative plot of the stellar fits for SV Peg. Plots are the same as the stellar fits in 4.5, but line list performance is shown side-by-side.