Fitting Brown Dwarf Distributions in the Milky Way: Vertical Scale Heights of the Thin Disk

Alasdair Koplick (10729216)
Project partner: Ewan Scopes

13 May 2024

ABSTRACT

The aim of this project was to fit scale heights of brown dwarf distributions in the Milky Way, making use of the catalogue of 1580 brown dwarfs obtained last semester. This involved generating model distributions of brown dwarfs based off previous works; ignoring the halo and thick disk, and expressing the thin disk as a double exponential. The completeness of our selection chain from last semester was then determined using a fake catalogue of dwarfs and applied to the model distribution, to account for any biases. The model distribution was then fit to the data from last semester to find the scale height parameter of the thin disk for the M, L and T dwarf types, as well as an overall value. The results obtained were an overall scale height of $H_z = 490 \pm 64$ pc, and for individual types scale heights of $H_{zM} = 420 \pm 53$ pc, $H_{zL} = 440 \pm 84$ pc and $H_{zT} = 950 \pm 185$ pc. These results are roughly in agreement with other papers. The completeness, fittings and comparisons to other papers then allowed further investigation and improvement of the selection chain developed last semester, as well as exploration of an alternate method for classifying dwarf subtypes based on colour colour space, resulting in new scale heights $H_z = 360 \pm 38$ pc, $H_{zM} = 330 \pm 39$ pc, $H_{zL} = 330 \pm 41$ pc and $H_{zT} = 500 \pm 74$ pc. These results agree well with other papers. The dependence of scale height on spectral type was determined to be inconclusive.

1 INTRODUCTION

Brown dwarfs are objects similar to stars, with a key difference in that they fail to acquire enough mass to produce the pressure needed to ignite fusion (the hydrogen burning limit being $\sim 0.064M_\odot - 0.087M_\odot$, from Auddy et al. (2016)) during the gravitational collapse in their formational stages. Due to their abnormal formation compared to other stars, and lack of presence in media and fiction since first being theorised in by Kumar (1962) and confirmed by Rebolo et al. (1995), it would be easy to think they have minimal impact on the Milky Way as a whole. However, brown dwarfs makeup around 50% of the stars in our local neighbourhood, as found by Aganze et al. (2022a) and discussed in Koplick (2024). Indeed, within the Milky Way Mužić et al. (2017) estimates there are $0.25 - 1 \times 10^{11}$ brown dwarfs, a significant fraction of its stellar mass of $\sim 6.5 \times 10^{10} M_\odot$ as given by McMillan (2011). The inability of brown dwarfs to fuse hydrogen also gives them two key attributes compared to other stars. Firstly, brown dwarfs occupy spectral types M, L T and Y, or the "latest" known spectral types (where later refers to dimmer stars), and all subtypes in-between (eg: M4,5,6 etc...). Note that in this report we will focus on M, L and T as the faintness of Y types makes them challenging to detect in surveys prior to JWST, see Burgasser et al. (2024). Secondly, since they simply cool in space after formation, as studied by Baraffe et al. (2003), Burrows et al. (1997) and others, they have very long lifespans. This combined with their ubiquity within the Milky Way make their distribution and ages a useful tracer of galactic formation and helpful for determining the timescales of various galactic processes, as discussed by Burgasser (2008b).

However, while brown dwarfs can be incredibly useful to study, they also pose contamination problems in projects such as Euclid (see Racca et al. (2016)) and other deep field surveys, as discussed in Caballero et al. (2008). Since most deep field surveys are conducted at near infra-red (NIR) wavelengths, and brown dwarf spectra peak around $0.5 - 2.5\mu m$ (see Mclean et al. (2003) for an early example of typical brown dwarf spectra), there is often a serious challenge associated with removing these foreground contaminant sources as they are easily confused with desired objects. This is worsened by the fact that brown dwarf spectra can look very similar to mid-redshift ($\sim Z = 6$) galaxies, as shown in Fig 2 of Koplick (2024). Indeed, it took until JWST spectroscopy by Burgasser et al. (2024) to show that some objects previously thought to be active galactic nuclei were infact more likely to be brown dwarfs. In some cases, such as galaxy surveys, the desired sources have angular extent while the brown dwarfs can be treated as point sources allowing them to be differentiated and removed. Indeed Bowler et al. (2016) suggests Euclid should be able to resolve galaxies to help prevent dwarf contamination. However, in many scenarios (such as Barnett et al. (2019)) this is not applicable, and also adds the requirement of resolving the desired sources, which is not always practical. As such, understanding brown dwarf distributions can help in any general case, as this allows a statistical approach to removing them.

Despite brown dwarfs being both contaminants and useful objects to study in their own right, the scale height of their distribution in the Milky Way is still not well known. Some recent values include $H_z \sim 300$ pc from Jurić et al. (2008), $H_z \sim 175 \pm 100$ pc from Aganze et al. (2022b) and $H_z \sim 290 \pm 50$ pc from Ryan et al. (2011) (Table 5 in Section 5.1 gives a more detailed list of results). In this report we aim to further constrain the scale height of brown dwarf populations. Part of the reason for these discrepancies is the different galactic components used in modelling, with Jurić et al. (2008) modelling the halo, thin and thick disks compared to Ryan et al. (2011) using just the thin disk. These different components of galactic structure were originally proposed by Oort (1926), who suggested a halo and disk based on the motions of stars with respect to the sun. Later the

© 0000 The Authors

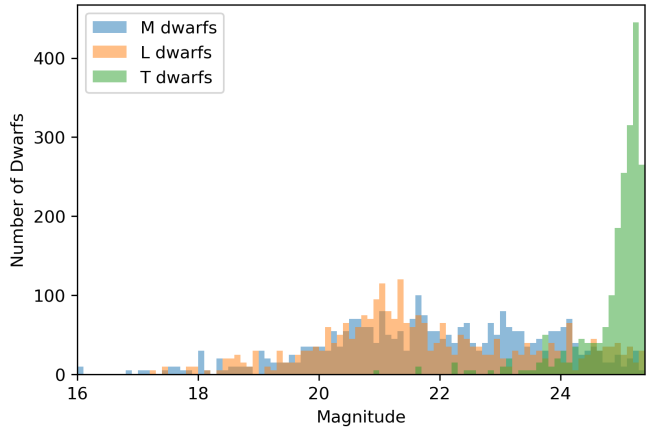

Figure 1. The data on brown dwarf counts by magnitude from the results obtained last semester, Koplick (2024). The different spectral types M, L and T are distinguished in by color. The Y axis is analogous to the surface density of dwarfs as this has been normalised to a 1deg^2 survey area.

distinction of the thin and thick disks was proposed by Gilmore & Reid (1983).

Last semester, the focus was on finding brown dwarfs within the COSMOS survey using data from the VISTA telescope (see Sutherland et al. (2015)), as discussed in Koplick (2024). Crucially COSMOS gave us a larger dataset than previous papers such as Aganze et al. (2022a) (focusing on nearby brown dwarfs) and Ryan et al. (2011) (focusing on brown dwarfs outside the local neighbourhood). Brown dwarfs were identified using a selection chain. First brown dwarf spectra models were carefully selected to use for spectral energy distribution (SED) fitting, and the SONORA from Marley et al. (2021) and LOWZ from Meisner et al. (2021) were chosen. The COSMOS catalogue was then trimmed to remove objects too faint ($J_{mag} > 25.3$) and a ks band colour space cut was applied to remove galaxies. Eazypy (or just "eazy" from Brammer (2021))was then used for SED fitting, and objects fit better with the brown dwarf models than the provided galaxy models were selected. This left a catalogue of 1580 brown dwarfs in Koplick (2024), the distribution of which depending on magnitude can be seen in Fig 1.

This semester, we fitted scale heights to this data to attempt to better constrain the distribution of brown dwarfs in the Milky Way. This starts by closely following Caballero et al. (2008), with the modeling of brown dwarf number densities in the Milky Way as a thin disk. Indeed, modeling brown dwarf distributions as a thin disk specifically, and fitting the scale height based on cumulative counts across different magnitudes has previously been well explored by Holwerda et al. (2023) and Pirzkal et al. (2009), to name a few (but not with the same amount of data that we have obtained from COSMOS). Using these models, counts for stars at different magnitudes will be generated, ie: a model distribution. These counts are then adjusted for any bias in our selection chain (or "completeness") through the use of another "fake" distribution consisting of only brown dwarfs. Finally we are able to fit our model distribution to our results from the previous semester (see Fig 1), giving us the best fitting scale heights for our thin disk model. We then investigate making changes to our selection chain motivated by the completeness of our selection chain, and discrepancies in results and methods compared to other papers. Also of interest is the dependence of scale height on spectral type, as this is an active area of discussion, with Ryan et al. (2011) suggesting that brown dwarfs should slowly depart from the scale height spectral type dependence followed by main sequence stars.

This report will firstly discuss the theory behind the thin disk model, the assumptions made and why they are valid, and derive the integrals to be evaluated for the dwarf number density. A background on completeness and why it is considered will then be included, before looking at the methods used to evaluate the integrals and determine the completeness of our data. The completeness graphs obtained will then be shown, and following that, the fits to the data and scale heights will be shown. Some of our scale heights are seen to be physically unreasonable and the data-set is found to contain contaminants (non-dwarf objects), after the removal of which we will arrive at our first set of results. The validity of these results and completenesses will then be discussed and compared to similar papers, resulting in an investigation of our selection chain. A secondary method of identifying the subtypes of dwarfs in our dataset will be tested, and the results of that compared to the prior results. Finally the dependance on scale height and physical implications of our results will briefly be discussed.

2 THEORY

In order to fit parameters to the data in Fig 1 and determine scale heights, the expected distribution of brown dwarfs in the Milky Way must be modelled, as mentioned above. For the case of modelling brown dwarfs, the thick disk and halo components of the Milky Way can be ignored, leaving only the thin disk. This assumption can safely be made for two reasons, as outlined in Caballero et al. (2008). Firstly, since brown dwarfs do not fuse hydrogen but slowly cool in the void of space, older brown dwarfs are incredibly faint (brown dwarf cooling models are discussed by Baraffe et al. (2003), for example). As the thick disk is expected to be composed of much older stars, $\gtrsim 10$ Gyr as outlined by Fuhrmann (1998), it is expected that any brown dwarfs in the thick disk will be too dim to detect. Hence, for modelling the brown dwarfs expected to be seen by a telescope in a survey area, they can be neglected. Secondly, as shown by Pirzkal et al. (2005), the halo component makes a negligible contribution to the overall number density of brown dwarfs. Caballero et al. (2008) estimates the error on these assumptions to be within $\lesssim 10\%$, which is similar to the error on other constants used later ($R_\odot$ and $Z_\odot$ in Equation 2). Hence, the distribution of brown dwarfs is safely modelled as a thin disk.

2.1 The Thin Disk Model

We now aim to arrive at an integral that can be evaluated to give the number density of brown dwarfs seen within a survey area, to produce a model distribution of brown dwarfs. This section closely follows the same derivation from Caballero et al. (2008). Expressing the number density n of stars in a thin disk as a double integral, from Bahcall & Soneira (1980), it can be written that:

$$n(r,z) = n_0 e^{-\frac{r}{H_r}} e^{-\frac{z}{H_z}} \tag{1}$$

where r and z refer to the two coordinate vectors labelled in Fig 2, H_r and H_z are the scale heights for the two exponentials, and n_0 is some initial distribution to scale the exponentials. Now, a correction is needed, as these exponentials are centered at the galactic centre, but the data is taken from telescopes on Earth (such as VISTA, whose capabilities are outlined in Sutherland et al. (2015)), a distance $R_\odot$ and height $Z_\odot$ from the galactic centre as also shown in Fig 2. This allows the exponential to be written as:

$$n(l,b,d) = n_0 e^{-\frac{R(l,b,d)-R_\odot}{H_r}} e^{-\frac{|Z_\odot - dsin(b)|}{H_z}} \tag{2}$$

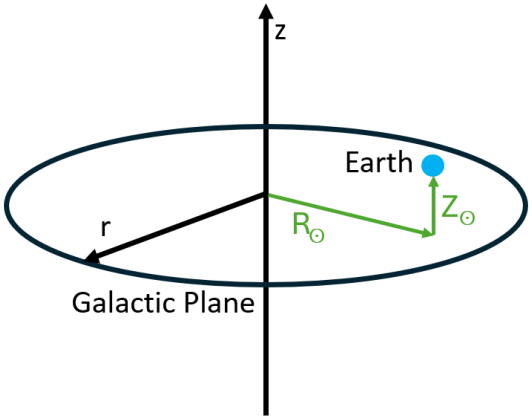

Figure 2. Diagram showing the Milky Way as a disk with the r and z coordinates used in Equation 1, as well as the distance from the galactic centre to Earth (blue), showing $R_\odot$ and $Z_\odot$.

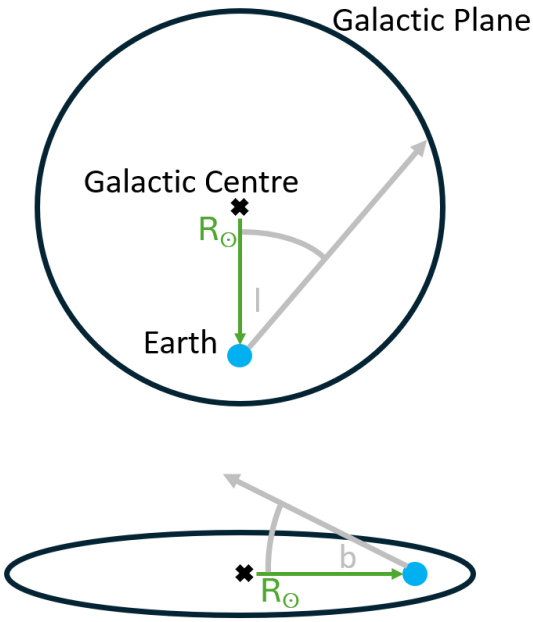

Figure 3. Diagram showing how the galactic coordinate system is defined from top-down and side-on views, with galactic longitude denoted as l and galactic latitude as b.

where we have also combined our z and r coordinate vectors into a single $R(l, b, d)$, the distance between the object being observed and the galactic centre. This is expressed in galactic longitude l and latitude b using the galactic coordinate system shown in Fig 3, and d, the distance between the brown dwarf and our sun. As such all other terms are now constants that can be found in papers such as Chen et al. (2001). The distance $R(l, b, d)$ can then be rewritten as:

$$R(l, b, d) = \sqrt{R_\odot^2 + d^2 cos(b)^2 - 2R_\odot d cos(b) cos(l)} \tag{3}$$

based on the assumption from Bahcall & Soneira (1980) that R is measured along the galactic plane. The further assumption of $d \ll R_\odot$ allows us to write:

$$R(l, b, d) \approx R_\odot - d cos(b) cos(l). \tag{4}$$

This can be safely assumed as the coordinates of the COSMOS survey we are simulating are point away from the galactic plane, so the distance from any dwarf to the galactic centre will be much greater than its distance to our sun. Combining Equations 2 and 4, yields;

$$n(l, b, d) = n_0 e^{\frac{d cos(b) cos(l)}{H_r}} e^{-\frac{\pm Z_\odot \pm d sin(b)}{H_z}}, \tag{5}$$

from which terms are combined to be written as:

$$n(l, b, d) = n_0 e^{\frac{\mp Z_\odot}{H_z}} e^{-\frac{d}{d_B(l,b)}}, \tag{6}$$

where;

$$\frac{1}{d_B(l, b)} = -\frac{cos(b) cos(l)}{H_r} \pm \frac{sin(b)}{H_z}. \tag{7}$$

This assists with integration later, as since the variable d will be integrated over, it is beneficial to separate it from l, b. The $\pm$ signs come from the absolute in Equation 2 and correspond to cases of looking above or below the galactic plane. This is discussed in detail in Section 2.2.1 when considering the physicality of the integral. Finally, the first half of Equation 6 can be written as a single term;

$$n_{A\mp} = n_0 e^{\frac{\mp Z_\odot}{H_z}}, \tag{8}$$

to end up with the integrand:

$$n(l, b, d) = n_{A\mp} e^{-\frac{d}{d_B(l,b)}} = n(d). \tag{9}$$

Note that the final equality in Equation 9 is due to the fact that our coordinates l, b will be fixed prior to integrating over the distance d.

2.2 Integration to Find Total Brown Dwarf Counts

We are now left with an equation that can be integrated over a volume to find the number of brown dwarfs inside that volume:

$$N = \int_V n(z)\, dV \tag{10}$$

where the variable d from Equation 9 has been re-written as z for clarity later. Assuming that our survey area is small, which is valid as the COSMOS field used to obtain data last semester in Koplick (2024) has an area of 2 deg^2, see Euclid Collaboration et al. (2022), we can assume that our volume to integrate over is a cone. This results in a line of sight integral:

$$N = \pi \left(\frac{b}{h}\right)^2 \int_{d1}^{d2} z^2 n(z)\, dz. \tag{11}$$

where z is the distance along line of sight, or height of our cone. The extra factor of $\pi(\frac{b}{h})^2 z^2$ comes from the integration of the cone over the other two axis, shown in Caballero et al. (2008). Here b is the base of the cone (survey area) and h is the height (distance from our sun to the dwarf). Approximating the cone as a square based pyramid, and considering the small angle approximation and the equation for angular size:

$$\theta = \frac{arclength}{radius}, \tag{12}$$

it can easily be seen that;

$$\pi \left(\frac{b}{h}\right)^2 = S. \tag{13}$$

Where S is the area of the survey in steradians (rad^2). This allows us to write our full integral from Equations 9, 11 and 13:

$$N = S n_{A\mp} \int_{d1}^{d2} e^{-\frac{z}{d_B(l,b)}}\, dz \tag{14}$$

2.2.1 Physical Scenarios & Discontinuities

Physically, this integral represents taking a cone of sky and counting the number of dwarfs found within it. It is therefore important to consider the different manners in which the galactic plane (the $z = 0$ plane of the Milky Way as shown in Fig 2) is intersected by our cone, as shown in Fig 4. This is because we expect the density of stars to exponentially increase in the direction of the galactic plane, and exponentially decrease away from it. For example, in the case of looking through the disk, the integral will first increase until it hits the disk, at which point it will start decreasing. In this case the integral would need to be split in half as it would be discontinuous at the galactic plane. This is seen in case I in Fig 4. There are three distinct cases found, as shown in Fig 4, where the two scenarios in case III are equivalent since the integral simply decreases in both. Other cases, such as the three cases identical to those in Fig 4 but with $d1$ and $d2$ switched, can be ignored as the way in which the integral is evaluated means $d1 < d2$ so they will never occur.

Within our integral, these cases are handled by the $\pm$ and $\mp$ signs in Equations 7 and 8, meaning that care must be taken to chose the correct normalisation factor $n_{A\mp}$ such that the integral overall either increases or decreases as required. Or in the discontinuous case I, such that it first increases then decreases. Defining the integral for ease of notation as

$$\int_{d1}^{d2} e^{-\frac{z}{d_B(l,b)}}\, dz \equiv I(d_B\pm, d1, d2) \tag{15}$$

it can then be seen that from evaluating Equation 15 that I is:

$$I(d_B\pm, d1, d2) = \left[(-1)e^{-\frac{z}{d_{B\pm}}}(d_{B\pm}z^2 + 2d_{B\pm}z + 2d_{B\pm}^3)\right]_{d1}^{d2}. \tag{16}$$

Here, the $\pm$ on the d_B corresponds to the positive and negative cases in Equation 7. This allows us to write the three cases for the number of brown dwarfs N_i where i corresponds to the different cases in Fig 4;

$$N_I = S\left[n_{A+}I(d_{B+}, d1, Z_\odot) + n_{A-}I(d_{B-}, Z_\odot, d2)\right], \tag{17}$$

$$N_{II} = S\left[n_{A+}I(d_{B+}, d1, d2)\right] \tag{18}$$

$$N_{III} = S\left[n_{A-}I(d_{B-}, d1, d2)\right]. \tag{19}$$

Now, the discontinuity in case I can be seen in Equation 17, where it must be evaluated in two halves using the distance to the galactic plane $Z_\odot$. Equations 17, 18 and 19 can then be used to calculate the expected numbers of dwarfs (by type) at different magnitudes, similar to Fig 1.

2.3 Completeness of The Data

One thing that must also be considered when generating a model distribution are any biases in the selection chain used to select brown dwarfs from the COSMOS survey. This is done by determining the completeness of the selection chain. The completeness can be defined as the fraction of successfully identified brown dwarfs after the selection chain, out of the total number of brown dwarfs put into the selection chain, similarly to Brown (2009). Successful identification requires that a given brown dwarf in the COSMOS survey is both identified correctly by the selection chain as a brown dwarf, and also not accidentally eliminated. As a result, the completeness of our selection chain depends on the cuts to the data made at each step

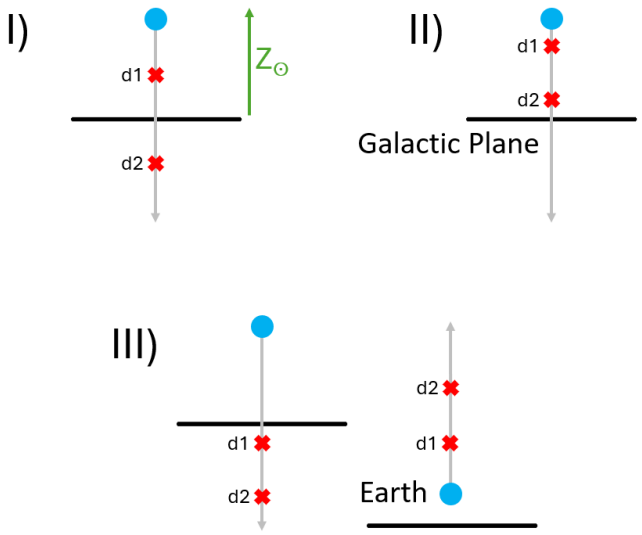

Figure 4. Diagram showing the different physical scenarios depending on what angle and distances are being evaluated. These result in different integrals to be evaluated. The two diagrams in case III are equivalent and have been included for clarity.

within the chain. Whilst these cuts aim to remove galaxies and retain brown dwarfs, at each step some brown dwarfs that appear similar to galaxies in these cuts will lost. For this reason, our final dataset will be biased against some types of brown dwarfs. The completeness then informs us in what way the selection chain is biased, allowing the distribution of dwarfs produced last semester to be accurately modeled. Each step of the selection chain is now explained along with how it is expected it to affect the completeness. After this a brief summary of the dependencies at each step given in Table 1:

1. Objects fainter than $J_{mag} = 25.3$ were discarded. This is the 5σ limit on the VISTA telescope (from Sutherland et al. (2015)), hence the limit on what we can observe. This is similar to how Guieu et al. (2006) treated the limit on completeness. A cutoff in completeness beyond 25.3 magnitudes is therefore expected as we have no objects beyond this.
2. The COSMOS data was plotted in Y-J vs H-Ks colour colour space, as this separates stars and galaxies into a stellar locus and galactic locus. Objects beyond where the density of the stellar locus fell below 5σ (as the density followed a Gaussian) were assumed not part of the stellar locus and removed. See Fig 7 in Koplick (2024). We will refer to this as the "Ks band color space cut". This roughly biases against dwarfs that are "redder" or emit at longer wavelengths. As a result the latest and dimmest dwarf types will be removed, reducing the completeness at high magnitudes.
3. After SED fitting in eazy (Brammer (2021)) against model brown dwarf SEDs, only objects with a $\chi^2_{dwarf} < \chi^2_{gal}$ were kept (ie: objects fit better by model dwarf SEDs than the provided galaxy templates). This biases against dwarfs with particularly noisy SEDs, expected to be the T dwarfs as they are the dimmest. Hence, the completeness is expected to possibly be reduced at high magnitudes.
4. Objects with $\chi^2_{dwarf} > \chi^2_{max}$ were removed, with χ^2_{max} being the maximum expected χ^2 for a brown dwarf (determined from our model dwarf SEDs). This means dwarf types poorly represented our model dwarf SEDs set will be biased against as they could be fit worse than χ^2_{max}.

Table 1. Table listing the different parameters that each step of our selection chain, and hence our completeness, depends on.

Selection Step	Dependencies
$J_{mag} < 25.3$ cut	J band magnitude
K_s 5σ cut	Dwarf type
$\chi^2_{dwarf} < \chi^2_{Gal}$	Dwarf Type, Noise
$\chi^2_{dwarf} < \chi^2_{max}$	Representation in model set
Type Identification	Representation in model set

5. The subtype of each dwarf is then identified, by taking the temperature of the best fitting model SED as the temperature of the dwarf. These temperatures are then matched to a subtype using a conversion table, for example Mamajek (2022). This will bias against types that are poorly represented in the model set at certain magnitudes, resulting in dips in the completeness. This will also result in baising towards certain subtypes.

The completeness function C can hence be expressed as:

$$C(\text{type},\ J_{mag},\ \text{noise},\ \text{representation}) = \frac{N_{ident}}{N_{tot}}, \tag{20}$$

where N_{ident} is the number of successfully identified brown dwarfs (as discussed above), and N_{tot} the total number of brown dwarfs.

As a side note, completeness and contamination (the existence of non-dwarf objects in our final catalogue of brown dwarfs) do not depend on each other. However, naively, having a high completeness tends to be a result of a lenient selection chain allowing all dwarfs through, which results in a higher contamination as more non-dwarfs will also be let through. Meanwhile a stricter selection chain would result in lower contamination, but also a lower completeness as more dwarfs would be incorrectly eliminated. As a result, there is a trade-off between contamination and completeness. Our aim here is to maximise completeness and minimize contamination.

3 METHOD

In this section, the method for evaluating the integrals found in Section 2.3 will be presented, including choosing which one to evaluate, and obtaining constants for the calculation. Then, the method used to find the completeness will be discussed. This will ultimately result in the distribution that is fitted to the data from last semester (Fig 1).

3.1 Calculating the Number Densities of Brown Dwarfs

Several things must be considered before evaluating the number density integrals in Equations 17, 18 and 19. Firstly the coordinates of the survey that is being modelled are needed. These can be found from EuclidConsortium (2023), with the data shown in Table 2. The data obtained last semester came from the COSMOS-Wide survey (referred to in this report as just COSMOS) so our galactic coordinates are $(236.8, 42.1)$. This corresponds to looking up and away from the centre of the Milky Way, which means we will be evaluating case III in Fig 4, or Equation 19. The next consideration are the constants $R_\odot$ and $Z_\odot$. These have been estimated using several different methods over the past 30 years, such as fitting the luminosity of red clump stars (Stanek & Garnavich (1998)) and star counts (Humphreys & Larsen (1995)), with some results outlined in Table 3. Values from Chen et al. (2001) were chosen, as they used a thin disk model and star counts (similarly to us). Furthermore their value of 27 ± 4 pc is within the range found by Joshi (2007), who performed statistical analysis of several different methods from previous papers to find a range of values for $Z_\odot$ of $13 - 28$ pc. Initial values for H_z and H_r were also taken as 400 and 2250 pc respectively based on Chen et al. (2001) and other papers. Following that, the normalisation factors n_0 in Equation 8 were acquired from Table 2 in Barnett et al. (2019). Finally, before calculating the integral, the limits $d1$ and $d2$ in Equation 19 need to be calculated. This was done by using Table 3 in Caballero et al. (2008), which gives the estimated absolute magnitude of every dwarf subtype, and the distance modulus equation:

Table 2. The coordinates and areas of the Euclid Auxiliary Fields that will be used to calibrate the Euclid Satellite. This includes the COSMOS-Wide field used to produce the brown dwarf catalogue last semester that is being modelled.

Field	RA:Dec (deg)	Galactic coordinates (l, b)	Area (deg^2)
CANDELS/AEGIS	214.827:+52.82	(96.438, 59.5521)	1.0
GOODS-NORTH	189.250:+62.25	(125.847, 54.7965)	0.5
COSMOS-Wide	150.119:+02.21	(236.818, 42.1239)	2.0
VVDS-Deep	36.500:-04.50	(172.006, -58.0557)	0.5
CDFS (inside EDFF)	53.117:-27.81	(223.576, -54.4374)	0.5
SXDS	34.500:-05.00	(169.759, -59.7518)	2.0

Table 3. Brief comparison of different values for $R_\odot$ and $Z_\odot$, obtained using different methods. Not all of these papers give values for both, indicated by –.

Paper	Method	$R_\odot$ (kpc)	$Z_\odot$ (pc)
Chen et al. (2001)	Thin disk	8.6 ± 0.2	27 ± 4
Stanek & Garnavich (1998)	Luminosity fits	8.2 ± 0.2	–
Humphreys & Larsen (1995)	Star counts	–	20.5 ± 3.5
Joshi (2007)	Various	–	$13 - 28$

$$m - M = -5 + 5\log d, \tag{21}$$

where M is the absolute magnitude, m the apparent magnitude and d the distance (integral limit). This is performed twice for each magnitude bin seen in Fig 1 to calculated both $d1$ (the lower limit) and $d2$ (the upper limit). The apparent magnitudes m are simply the magnitude of that bin for $d1$, and the magnitude of that bin + the bin size for $d2$. This is then repeated for every subtype of dwarf.

The integral in Equation 19 can then be calculated by evaluating it at each limit and taking the difference, in each magnitude bin. This gives us the number of brown dwarfs of each subtype in each magnitude bin that are within our cone of survey area. This is shown in Fig 5. Hence, this is the distribution of dwarfs expected when looking at the COSMOS survey. However, it does not look similar to the data from last semester (Fig 1), as it has not yet been adjusted for the biases of the selection chain, ie: the completeness.

3.2 Finding the Completeness

The generated distribution of dwarfs in Fig 5 now needs to be adjusted for completeness. The aim is to have a weighting that can be applied to every magnitude bin, for every dwarf subtype, using Equation 20. This weighting represents how well these dwarfs pass

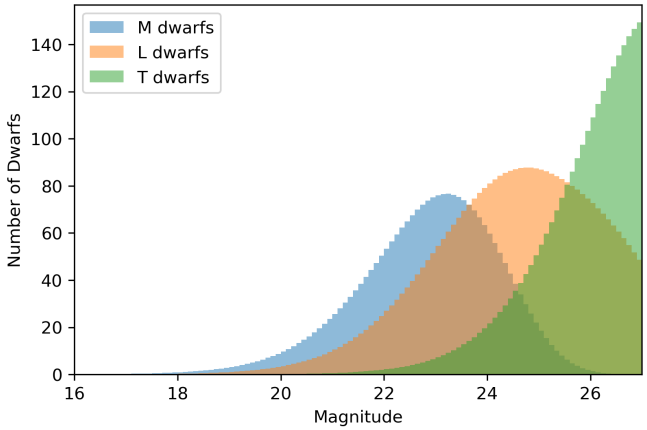

Figure 5. Model distribution of brown dwarfs generated using the thin disk model, coordinates for the COSMOS survey, and values for $R_\odot$ and $Z_\odot$ from Chen et al. (2001). Axis and colors are the same as Fig 1. Only types M8 and above as our selection chain is not sensitive to earlier spectral types. Subtypes have been grouped together into overall types M, L and T.

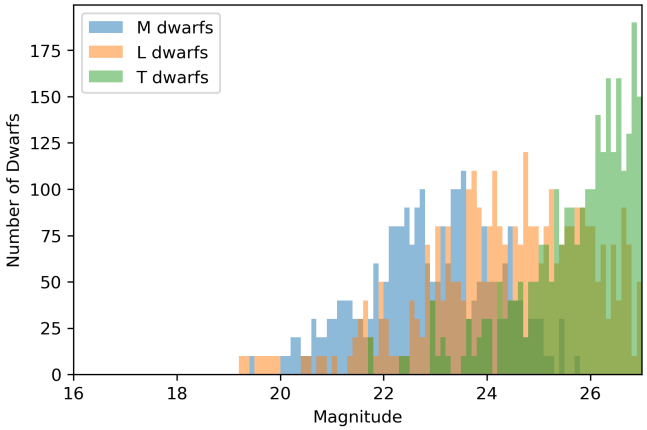

Figure 6. The generated distribution of dwarfs after randomly drawing to ensure there is a discrete number of dwarfs in each bin. Axis and colors are the same as Fig 1.

the selection chain (see Section 2.3). In order to find the fraction in Equation 20, a "fake catalogue" consisting of only brown dwarfs will be generated, modified for noise and other effects encountered when observing, and then fed through the selection chain from last semester. A fake catalogue of only dwarfs is necessary as it ensures the number of dwarfs fed into the selection chain is known. This allows the number of dwarfs that are successfully identified to be compared to the total, for each subtype and in each magnitude bin. Hence the completeness C as a function of magnitude can be found per subtype.

To generate the fake catalogue, the model distribution of brown dwarfs shown in Fig 5 must be made discrete, as Equation 19 evaluates to non discrete numbers especially at bright magnitudes where small numbers of dwarfs (such as 0.01) are expected. To do this dwarfs are drawn at random with a weighted probability to create a new distribution. To determine the weighted probability, Fig 5 is first used to find how many dwarfs of each subtype are in each magnitude bin (denoted as N_i where i is the dwarf type M7, M8 etc...). Also important is the total number of dwarfs N_{tot} within each magnitude bin. Now, within each magnitude bin, dwarfs are drawn randomly with weighted probabilities until N_{tot} dwarfs have been drawn. The weighted probability for a given subtype is the number of times it appears in the magnitude bin, normalised by N_{tot}. For example, the M8 dwarfs would have a weighting W of:

$$W = \frac{N_{M8}}{N_{tot}}. \tag{22}$$

This results in a randomly drawn, discrete distribution of dwarfs, shown in Fig 6.

Spectra are then needed for each of these dwarfs, so that the distribution can be turned into a catalogue and the SED fitting in the selection chain can be performed. Hence, a "characteristic spectra" of each dwarf subtype is needed. Since the spectral types of brown dwarfs are well defined by the SpeX standards from the SpeX PRISM Library, by Burgasser (2008a), this is a good place to take characteristic spectra from. However, some of the SpeX standard spectra appeared to be duplicated types (eg: 2 L0s and no L4), so a new selection from all available spectra in the library was made to ensure one high quality spectra of each type was present. For the M and L dwarfs one of each optical spectral type was chosen, while for the T

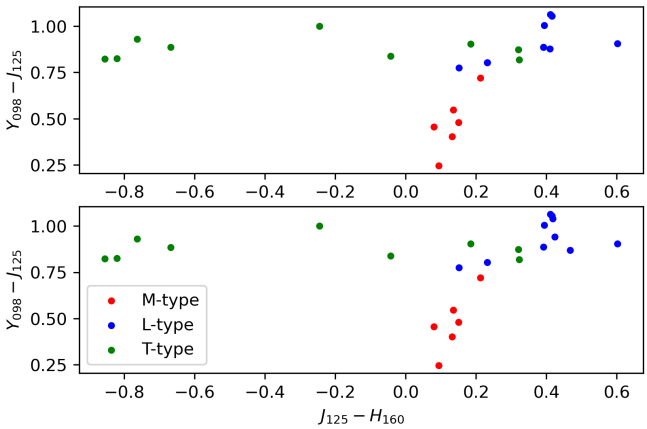

Figure 7. Comparison of the SpeX Standards containing duplicates (top) and the new selection of SpeX spectra used here for the characteristic spectra (bottom), in J-H Y-J color color space. Minimal differences can be seen in the L types.

dwarfs one of each NIR spectral type was chosen. This was following advice from private communication with A. Burgasser, as the NIR spectral types better represent the T dwarfs. This made a minimal difference overall, as can be seen in Fig 7 where a comparison of our new selection and the old SpeX standards is shown, although some of the L dwarfs are slightly better spaced in colour colour space. Now, when a specific subtype of dwarf is randomly drawn, its corresponding characteristic spectra is also selected. This results in a catalogue of brown dwarf spectra that follows the distribution in Fig 6.

Finally, the spectra are integrated through the VISTA filters (shown in Koplick (2024) Fig 5), to give SEDs that can be fitted in eazy. Noise is then applied to each filter integrated flux, by drawing a value from a gaussian distribution with a standard deviation equal to the 1σ detection depth of that filter, to offset the filter-integrated flux with. This is to simulate noise seen when observing that could push dwarfs from one subtype to another, eg: an T2 may appear to be a T3 due to noise. The 5σ depths were obtained from Table 1 in Adams et al. (2023). We now have a catalogue of "fake dwarfs" that follow the distribution in Fig 6, accounting for observational effects.

3.3 Completeness Graphs

After generating a catalogue of fake dwarfs, it can now be run through our selection chain, and the number of dwarfs that are successfully identified can be compared to the total number put in. Hence the completeness C in Equation 20 can be found.

One problem that immediately emerges is the lack of certain dwarf types at some magnitudes. For example, even when generating huge catalogues (with $> 30,000$ objects), L and T dwarfs still rarely appear at bright magnitudes, from $16 - 21$. This is easily seen in the distribution in Fig 6, as these dwarfs simply do not exist at bright magnitudes. This results in the completeness fluctuating between 0 and 1 at bright magnitudes for L and T dwarfs, as either a single dwarf is drawn, or none are, in which case the fraction in Equation 20 evaluates to 0. Hence the completeness is only either 0 or 1. This also means that in some brighter bins where there may be 100% completeness (or $C = 1$), there is actually 0 as no dwarfs are drawn in this bin in the first place. This is corrected against for two reasons; firstly, the completeness function should be a smooth curve, and secondly, we expect dwarfs at brighter magnitudes to have very high completenesses (from Section 2.3 only the dimmer stars are expected to be biased against in steps 3 and 4 of our selection chain). To correct this we require that > 10 objects are in a magnitude bin before calculating the completeness, otherwise the completeness is simply set to 1. The exception to this is if there are 0 objects in every bin, in which case the completeness is set to 0, as the > 10 requirement would set it to 1 everywhere, which is wrong. This also gives a minimum resolution of $< 10\%$, the same as the error on the assumptions made in the thin disk model (as found by Caballero et al. (2008)), and similar to the error on the constants $R_\odot$ and $Z_\odot$. To further assist with increasing the number of later type dwarfs at brighter magnitudes, the size of the magnitude bins were increased brighter magnitudes, shown in Fig 8. This results in jagged graphs, and as mentioned above, a smooth function is expected for C. Hence, a Savitzky-Golay filter was used to smooth the completeness (Savitzky & Golay (1964) and Virtanen et al. (2020)).

Once these problems have been addressed, the completeness can be plotted. The completeness of a typical dwarf subtype (in this case T3) at each stage in the selection chain is shown in Fig 9 to give an idea of the typical shapes. It can be seen that the $J_{mag} < 25.3$, and Ks band color space cuts perform as expected, keeping all objects ($C = 1$) expect for the dimmest objects ($C < 1$). The "Eazy Applied" graph corresponds to steps 3, 4 and 5 in the selection chain and the dip comes from T3 dwarfs being poorly represented by the model dwarf SEDs at those magnitudes. The completeness graphs as a function of magnitude for each subtype are now shown in Fig 10. Values larger than 1 in Fig 10 occur due to the subtype identification based on the best fitting model SED at the end of our selection chain (step 5). For example, when the fake catalogue is generated there is an initial total number of L6 dwarfs (N_{tot}), but at the end of the selection chain other dwarfs have been incorrectly identified as L6. This results in a larger final number of "successfully identified" L6 dwarfs (N_{ident}), hence Equation 20 evaluates as > 1. As the completeness is a weighting that describes the bias of the selection chain, a completeness larger than 1 means a dwarf is biased for, and a completeness less than 1 means a dwarf is biased against by the selection chain. The validity of these completenesses will later be discussed in Section 5.2.

3.4 Fitting The Model to The Data

Once the completeness graphs in Fig 10 have been obtained, the model distribution of dwarfs in Fig 5 can be weighted based on the completenesses. This produces the model distribution in Fig 11. This distribution can now finally be fit to the data obtained last semester in Fig 1, as now the biases of the selection chain and observational effects have been accounted for. The data was fitted using the curve_fit routine from scipy, from Vugrin et al. (2007) and Virtanen et al. (2020), a non-linear least sqaures regression algorithm. A one variable fit varying the scale height was performed for each dwarf type (M, L and T) independently and then an overall fit of the whole distribution was also performed.

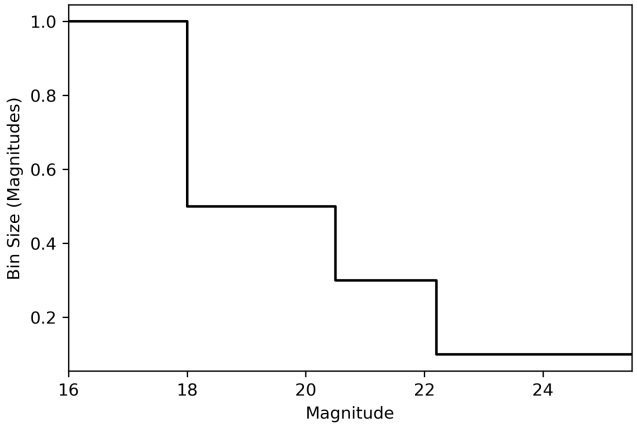

Figure 8. Figure showing how the size of bins in magnitudes (Y axis) changes across magnitudes (X axis) to force larger numbers of dwarfs at brighter magnitudes.

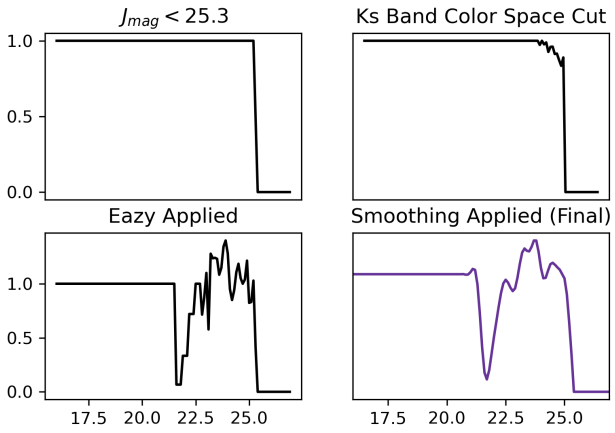

Figure 9. How the completeness of an average dwarf subtype (T3 in this case) is affected by the different stages of the selection chain as described in Section 2.3, left to right top to bottom. The Y axis is completeness (C) and the X axis magnitude. Note that stages 3, 4 and 5 have been combined into the third "Eazy Applied" graph. The "Smoothing Applied" graph is colored to demonstrate that it is these "Final" curves that are plotted in Fig 10.

4 RESULTS

The results from fitting the model distribution in Fig 11 to the data in Fig 1 are now shown in Fig 12. Scale heights for the different dwarf types (M, L and T) have been fit individually, along with an overall fit made by summing the distributions. This yields a scale height of

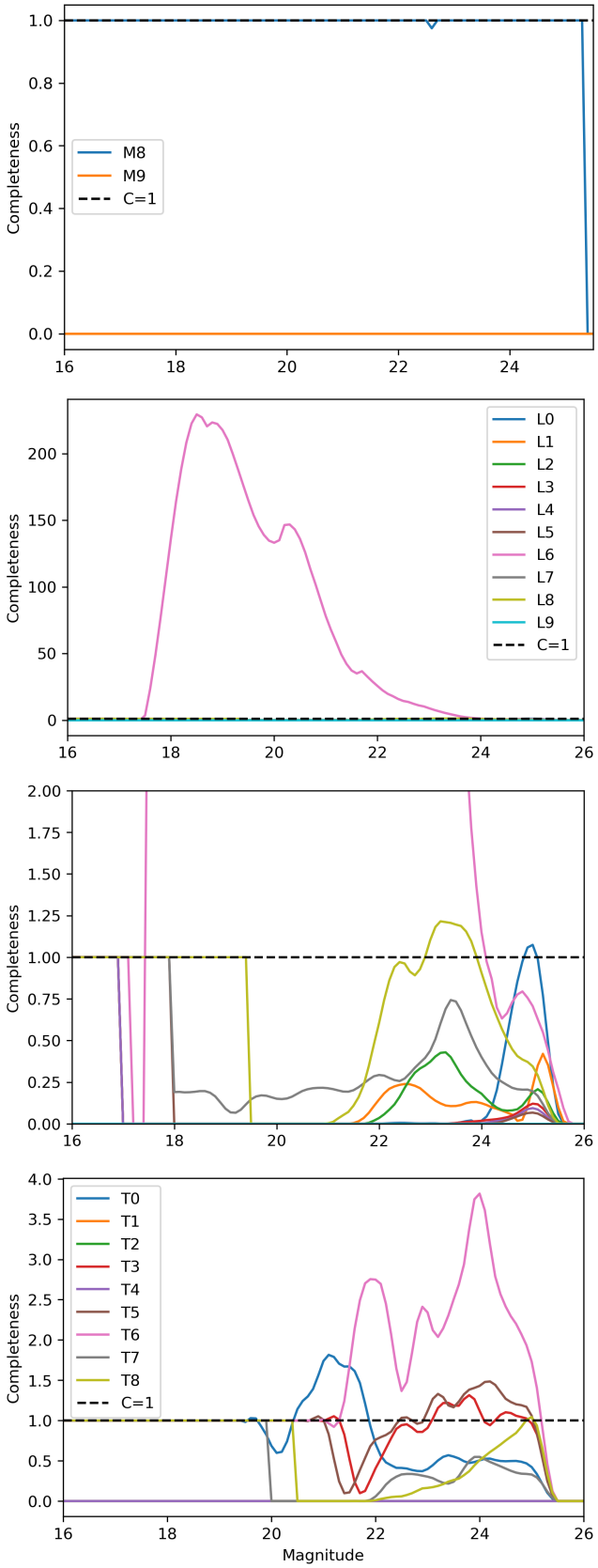

Figure 10. Completeness graphs for different dwarf subtypes, in magnitude vs completeness (C). The top image shows the M dwarfs, the middle two the L dwarfs (where the third image is zoomed in so subtypes other than L6 can be seen), and the bottom image shows the T dwarfs.

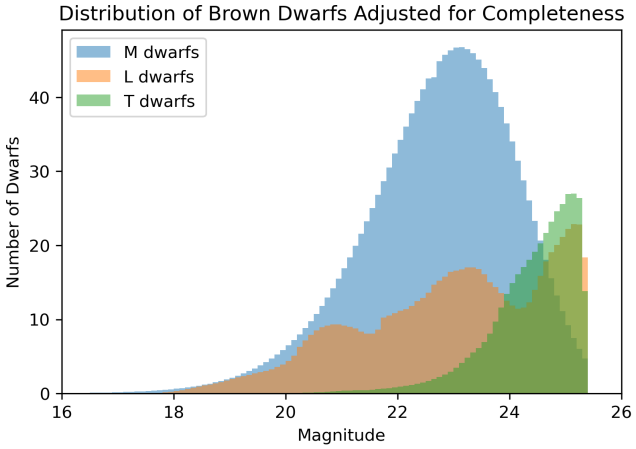

Figure 11. The distribution of dwarfs after adjusting for completeness. Axis and colors are the same as previous distribution graphs. The significant decrease in number from Fig 5 is due to dwarfs being biased against by the completeness. The $J_{mag} < 25.3$ cut from the selection chain can also clearly be seen.

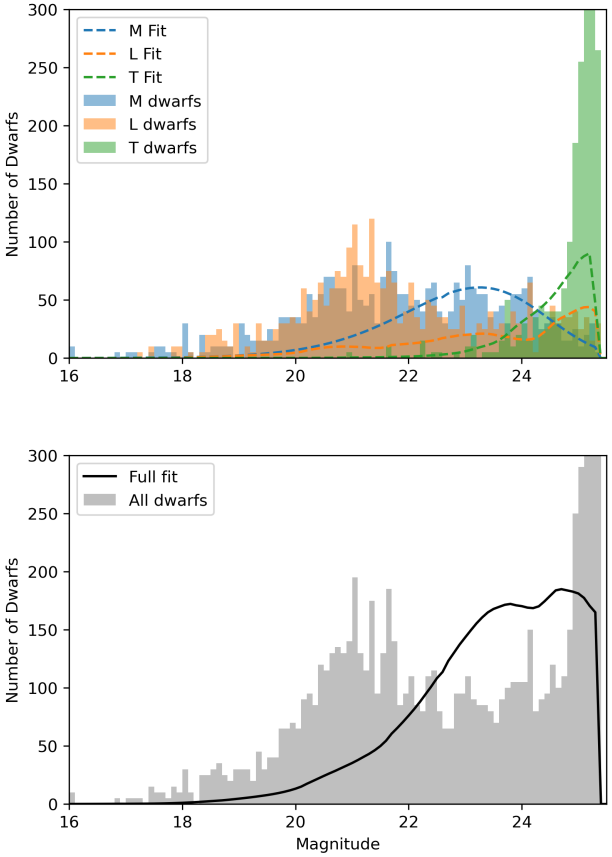

Figure 12. Top: individual fits (dashed lines) to data (solid bars) for scale heights of each dwarf type, distinguished by color. The large spike in T dwarfs at faint magnitudes can be seen. Axis and colors are the same as previous distribution graphs. Bottom: fit (black) to the data (grey) for all the distributions combined to give a single scale height. Axis are the same as previous distribution graphs. The features of Fig 11 can be seen in the fitted distributions (dashed and black lines).

$H_zM = 440 \pm 10$ pc, $H_zL = 480 \pm 36$ pc and $H_zT = 0.04 \pm 9 \times 10^9$ pc, as well as an overall scale height of $H_z = 610 \pm 16$ pc.

One thing that is immediately noticeable is the size of the fit for the T scale height. Indeed, the radius of the Milky Way is of the order $\sim$ 200 kpc, from Kochanek (1996), suggesting that our result of 40×10^3 kpc for the scale height of T dwarfs is highly unreasonable. This was immediately investigated. Initially, the data was fitted using a separate least squares fitting procedure that was self written, to check that the scipy fitting routine was not failing. This confirmed the large value of H_zT. Next the data from the previous semester was inspected, with the large H_zT value found to be due to the large (and largely unexpected) spike in T dwarfs at faint magnitudes, as seen in Fig 1 and 12. The most obvious suggestion is contamination of the dwarf catalogue with galaxies, or other faint non-dwarf objects. As a result, the selection chain was re-examined to determine why this large spike in T dwarfs was occurring, and to attempt to identify and remove any contamination.

4.1 Modification to Selection Chain to Prevent Contaimination

After inspecting the eazypy (Brammer (2021)) part of the selection chain (steps 3,4 and 5 as outlined in Section 2.3), it was determined that in certain cases all objects fed into eazypy were passing the χ^2 cuts. This is due to the eazypy failing to initialise templates and returning infinite χ^2_{gal} values for some faint objects. Since the $\chi^2_{dwarf} < \chi^2_{gal}$ condition is checked by requiring $\Delta\chi^2 = \chi^2_{dwarf} - \chi^2_{gal} < 0$, objects with an infinite χ^2_{gal} are always selected (as the galaxy fits are always worse than the dwarf fits). This results in many objects entering the catalogue that could either be dwarfs, galaxies, or other stars, resulting a large contamination of the catalogue at faint magnitudes; hence the large, faint T spike. The selection chain was modified to fix this by removing objects with $\chi^2_{gal} = \infty$. In addition to this step, to further prevent galaxy contamination, an extra check was added; requiring $\Delta\chi^2 < -10$. This prevents cases where an object is fit very similarly by both a galaxy template and a model dwarf SED. In these cases it maybe uncertain if the object truly is a dwarf, so excluding it is the safer bet to prevent contamination.

In order to check that these changes are actually reducing contamination, and not also removing dwarfs, the completenesses were compared before and after these steps. Since the completeness is a distribution of only dwarfs, if there is a minimal difference in completeness and a large number of objects removed from the data, then we can be certain that these steps are mostly removing non-dwarf objects. After these steps are taken, there is a $\sim$ 6% reduction in dwarf completeness compared to a $\sim$ 30% reduction of objects in the data. We can therefore be certain that of the around 200 objects removed from the data in this step, roughly $1-6/30 \approx 80\%$ should be contaminants. This is deemed acceptable, as other steps such as the $J_{mag} < 25.3$ cut have a much more significant effect on accidentally removing dwarfs than this ($\sim$ 50%), and there is no other way to easily remove these sources (recall the trade-off between completeness and contamination mentioned in Section 2.3). Therefore, both steps are added to the selection chain. The new distributions and fits after taking these steps are shown in Fig 13, and a summary of scale heights for these modifications is presented in Table 4. Indeed, these modifications produce better fits that are more consistent with other literature, especially reducing the value of H_zT to a much more physically reasonable number. This further suggests that there were in-fact non-dwarf objects contaminating the data as theorized, and that this contamination has successfully been reduced significantly.

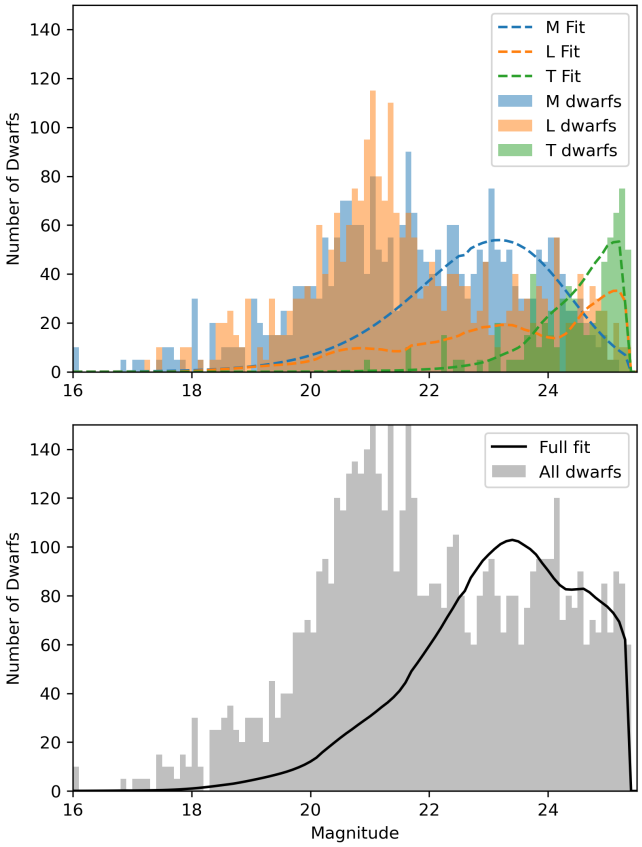

Figure 13. Top: fit for individual scale heights after excluding any values with $\chi^2_{gal} = \infty$ and $\Delta\chi^2 > -10$. The removal of the T spike at faint magnitudes can be clearly seen compared to the previous figure. Bottom: overall fit after applying the same constraints. This data is much better fit by the model distribution. The axis and colors are the same as Fig 12.

Table 4. Summary of the scale heights fitted before and after modifying the selection chain, where "both" refers to removing objects with $\chi^2_{gal} = \infty$ and $\Delta\chi^2 > -10$. "10% Err." refers to after an additional 10% error has been added. The T scale height value can be seen falling to a more physically reasonable value.

Method	H_zM (pc)	H_zL (pc)	H_zT (pc)	$H_zOverall$ (pc)
Original	440 ± 10	480 ± 36	$0.04 \pm 9 \times 10^9$	610 ± 16
No ∞	440 ± 10	480 ± 37	$0.09 \pm 4 \times 10^7$	530 ± 14
Both	420 ± 11	440 ± 40	950 ± 90	490 ± 15
10% Err.	420 ± 53	440 ± 84	950 ± 185	490 ± 64

4.2 Errors on Results

It is worth noting that the errors provided on the results in above are from the fitting routine only, and do not include any other sources of errors. As discussed in at the end of Section 2, there is a $\sim$ 10% error on the assumption of a thin disk only, and similar errors on the constants $R_\odot$ and $Z_\odot$. Therefore, a minimum error of 10% is also applied to our results, giving final scale heights of $H_zM = 420 \pm 53$ pc, $H_zL = 440 \pm 84$ pc, $H_zT = 950 \pm 185$ pc, and $H_z = 490 \pm 64$ pc. This is shown in Table 4, and is a error on the value obtained. The real error is likely larger.

Table 5. Comparison of scale height results from other papers. Thick disk and halo components refer to papers where the models used included additional thick disk and halo components ontop of the thin disk model we used. If no additional components are mentioned than a thin disk model was used. LN dwarfs refers to dwarfs found in the local neighbourhood, and typing by color refers to a different the method used to identify dwarf subtypes (discussed later).

Paper	Result	Notes
Jurić et al. (2008)	$H_{zM} \sim 300$ pc	thick disk component halo component
Bochanski et al. (2010)	$H_z = 300 \pm 15$ pc	thick disk component
Ryan et al. (2011)	$H_z = 290 \pm 40$ pc	typing by color
van Vledder et al. (2016)	$H_{zM} = 290 \pm 20$ pc	halo component
Rosell et al. (2019)	$H_{zL} \sim 450$ pc	very large sample
Sorahana et al. (2019)	$H_{zL} = 340 - 420$ pc	
Aganze et al. (2022b)	$H_{zM} = 250 \pm 50$ pc $H_{zL} = 150 \pm 50$ pc $H_{zT} = 175 \pm 100$ pc	only LN dwarfs
Holwerda et al. (2023)	$H_{zM} = 300 - 350$ pc $H_{zT} = 150 - 200$ pc	M4-M8 only T4-T8 only

5 DISCUSSION

5.1 Comparison to Literature

At a glance, the results from Table 4 after applying both additional constraints to our selection chain seem to agree with other papers shown in Table 5. Certainly, our value for H_{zL} falls within the values obtained by Rosell et al. (2019) and Sorahana et al. (2019). This is especially reassuring as both of these papers also used thin disk models, with Rosell et al. (2019) using a sample of dwarfs an order of magnitude larger than our own. Further similarities when comparing our model distribution (see Fig 5) to those produced by the Euclid team for simulations suggest that our method for producing these distributions is valid (via private communication with R.Bowler). Note this is most easily seen when M6 and M7 dwarfs are added to our distribution, which are not included in Fig 5.

On the other hand, our overall scale height, or H_z value, appears larger than both Ryan et al. (2011) and Bochanski et al. (2010), although not by an order of magnitude. One suggestion of where this difference could come from would be the additional thick disk component fit by Bochanski et al. (2010). However, the fact that Holwerda et al. (2023) obtains a similar value for H_{zM} to Jurić et al. (2008) (halo and thick disk) and van Vledder et al. (2016) (halo) while only using a thin disk model implies there is minimal effect when fitting additional galactic structure components. This is further suggested by Caballero et al. (2008) estimating the error on assuming only a thin disk model as $\sim$ 10% (as mentioned at the end of Section 2). This suggests that the discrepancy in results instead arises from differences in method, as Ryan et al. (2011) identify the subtype of dwarf using color color space, as opposed to the method used here of taking the temperature of the best fitting model dwarf SED.

Similarly, our H_{zM} and H_{zT} values are larger than other papers have found, especially the H_{zT} values. The more concerning disagreement is with Holwerda et al. (2023) rather than Aganze et al. (2022b). This is because Aganze et al. (2022b) has only 164 dwarfs from just the local neighbourhood, and hence uses a different method for estimating the scale height (luminosity functions and a Monte Carlo approach as outlined by Burgasser (2004)). Aganze et al. (2022b) also notes that their H_{zT} value is noticeably smaller than other, ground based studies. Meanwhile Holwerda et al. (2023) has 518 dwarfs from a JWST field survey, and also uses a thin disk model, so there should be minimal differences between the values obtained by them and us since the data and model is the same (assuming all field surveys are similar). This also suggests that it is the method, not the data and model, that is responsible for the discrepancy; Holwerda et al. (2023) uses a different method for identifying dwarfs from their data and also for identifying the subtype of dwarfs found.

All of the above suggests that the selection chain and method of dwarf subtype identification should be investigated to ensure it is valid.

5.2 Completeness Validity

One way to investigate our selection chain and type identification is by looking at the completeness. Since the completeness is effectively accounting for the biases of our selection chain, which includes subtype identification, asking if our completeness graphs are reasonable can help find issues with our method. The completeness is also applied to the model distribution that is fit to our data (Fig 11), so changing the completeness will have a direct effect on our scale heights.

The biggest effect on the completeness graphs comes from the way in which the subtype of dwarf is identified; by using the temperature of the best fitting model dwarf SED and comparing that temperature to a table to determine the corresponding type (see Section 2.3). This can be understood by considering the massive spike in L6 dwarfs in Fig 10, other dwarfs that are not L6 get incorrectly identified as L6 dwarfs resulting in a much larger number of L6 dwarfs at the end of our selection chain than were put in at the start. This is referred to as "scatter", as non L6 dwarfs scatter into being L6s. To see which subtypes were being incorrectly assigned, and to what subtypes they were being wrongly assigned to, the heatmap in Fig 14 was produced. This compares the fraction of input dwarf subtypes to output dwarf subtypes. Since this is generated from the fake catalogue, it is known exactly which dwarf subtypes are being inputted, so gives an indication of how well the subtype identification step of our selection chain does. From inspecting this graph there is a roughly linear trend, which is good. A linear 1:1 input/output relation would mean dwarfs are being identified as the correct subtype, for example, 99.9% of M8 dwarfs are successfully identified as M8 dwarfs, and 36% of L6 dwarfs are correctly identified as L6. However, this is not a very strong trend. It can already be seen from Fig 10 that the subtypes L6 and T6 are biased for, ie: have a completeness larger than 1. But from Fig 14, many rows that are 0 all the way across can also be seen, such as M9, L9, T1, T2 and T4. This means that no dwarfs are ever identified as these subtypes and hence they have a completeness of 0, as after subtype identification no dwarfs of these subtypes exist. As a result these objects are heavily biased against, this is also shown in Fig 10 but is harder to see. There are also columns in Fig 14 that are 0 all the way up them, meaning that these dwarfs are removed during the SED fitting process done by eazy, as they are either better fit as galaxies or too poorly altogether (steps 3 and 4 in our selection chain as outlined in Section 2.3). This suggests that L4 and L5 dwarfs are not well represented in the model dwarf SEDs we are using to fit with.

The huge bias towards L6 dwarfs can be explained by considering the model dwarf SEDs, and temperature conversion table from

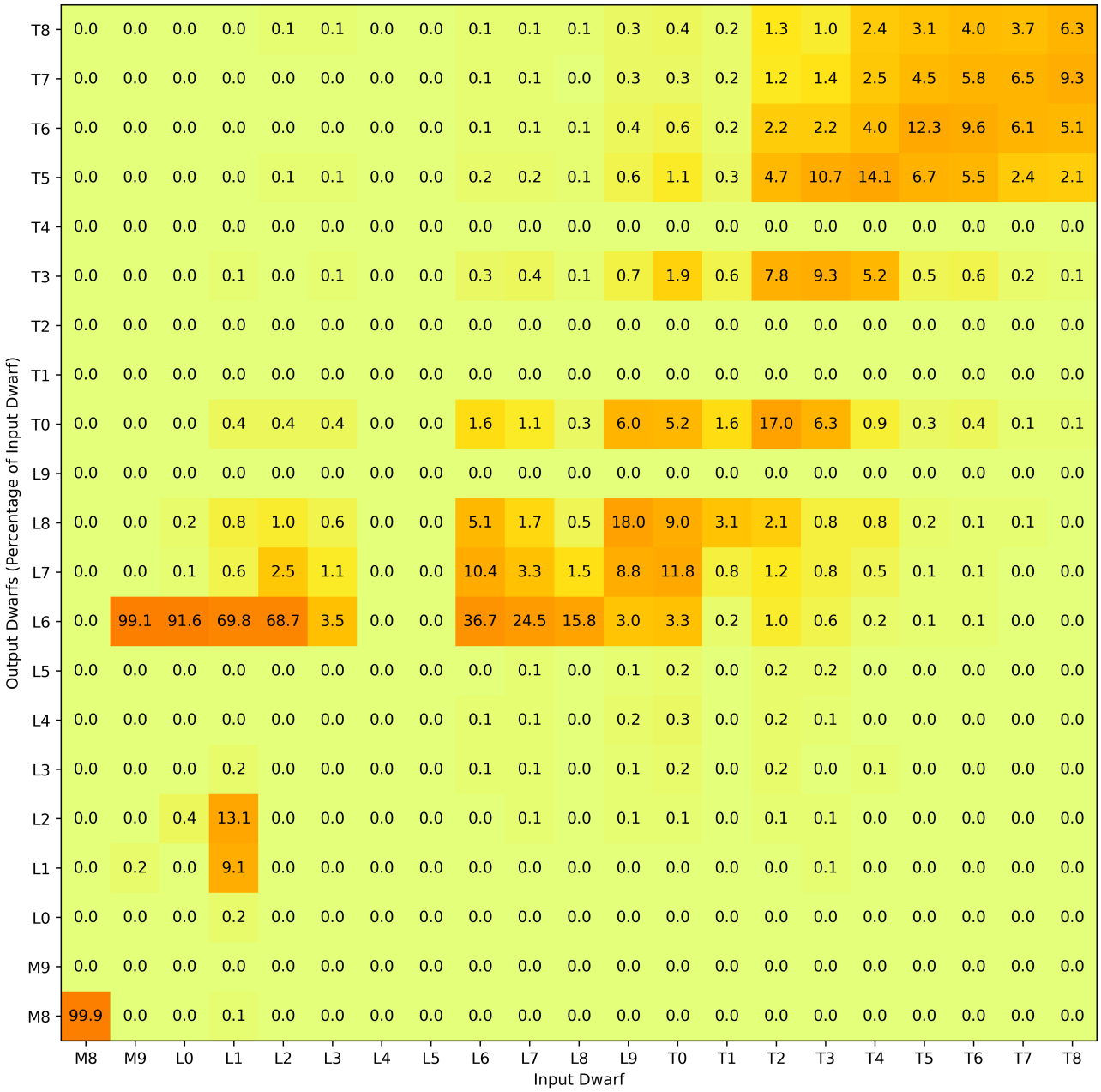

Figure 14. Input/output heatmap of dwarf subtypes and the fractions of input dwarfs that contribute to output dwarfs. The X axis is dwarf subtypes in and the Y axis is dwarf subtypes out. The numbers correspond to the percentage of input dwarf subtypes that contribute to that output subtype, eg; the 99.9 in the bottom left box means that 99.9% of M8 dwarfs are identified as M8 dwarfs, whilst the 99.1 in the M9 column means that 99.1% of M9 dwarfs are identified as L6 dwarfs.

Mamajek (2022) used. This is because the L6 temperature occurs at 1550 k, which is the peak temperature of the LOWZ model dwarf spectra (from Meisner et al. (2021)), one of the model dwarf spectra sets that are used to fit the data in eazypy (with the other being the SONORA models, from Marley et al. (2021)). Since the best fitting model is taken, this means any dwarf that is L6 or earlier and fit better by the LOWZ models than the SONORA models, ends up with a "best fit" temperature of 1550 k. This results in a pileup of L6 or 1550 k dwarfs, massively inflating the completeness. The bias for the T6s however, is less well explained. We expect the largest amount of scatter for the T dwarfs, as they have the smallest signal to noise ratio of all the dwarfs due to being the faintest. However it is uncertain why T6s specifically are preferred. One thing that was noticed was that the initial number of T6s (~ 5000) in the fake catalogue was much smaller than T5s (~ 10000) and T7s (~ 28000). Since scatter from closer subtypes is roughly stronger, as seen in Fig 14, this means the number of T dwarfs scattering into T6 would quickly exceed the initial number of T6s, resulting in a completeness greater than 1. This could be a random effect, where by chance when drawing randomly to create the model distribution in Fig 6, very few T6s were drawn initially. Alternatively, this could be an artifact of the maths behind our model. To test this, many random distributions could be drawn and the completenesses averaged, but this would require more time than was available.

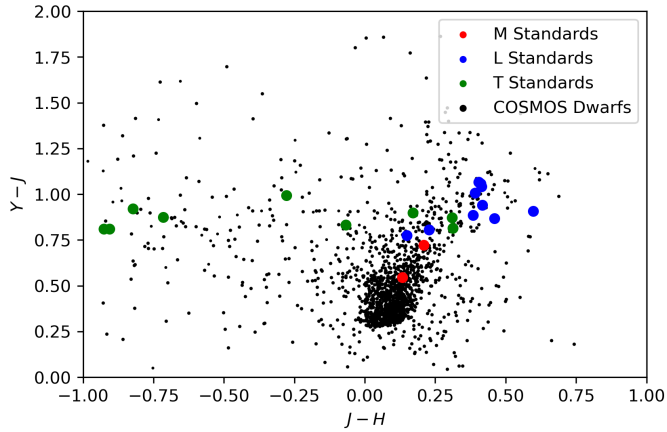

Figure 15. Plot of the SpeX standards and our 1580 dwarfs from the previous semester in J-H vs Y-J color color space. The type of these dwarfs was then determined as the type of the closest SpeX standard.

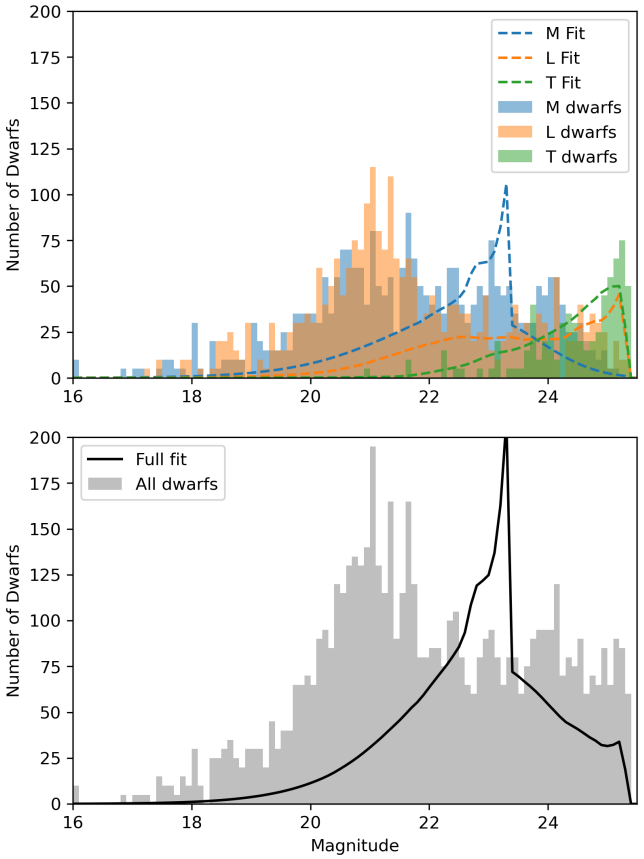

Figure 16. Fits of the data after performing identification of dwarf subtypes by color color space. The axis and colors are the same as Fig 12.

The bias against the M9, L9, T1, T2 and T4 dwarfs can be understood by consider the spacing of the LOWZ and SONORA models in temperature, compared to the conversion table (Mamajek (2022)). At higher temperatures, both LOWZ and SONORA have temperature spacing of 1000 k between models, but the conversion table is more finely spaced than this, with T1, T2 and T4 being at 1240, 1220 and 1180 k respectively. The T3 dwarfs, however, are given a temperature of 1200k. Therefore, when the temperature of the best fit model is taken, being either 1100, 1200, 1300 k etc..., dwarfs can never be identified as T1, T2 or T4. They will always been identified as T3 if a temperature of 1200 k is assigned, or surrounding T subtypes if 1100 or 1300k is assigned. This can be seen in Fig 14, with T1, T2 and T4 dwarfs being preferentially assigned T3. They also scatter to T0 and T5, which occurs due to the noise added to the fake catalogue to account for observational effects. The same phenomena results in M9 dwarfs always being incorrectly identified as M8s, and L9s never being identified and scattering to close by subtypes.

All of the above suggests that the method of identifying dwarf subtypes by using the best fitting model dwarf SED is not ideal. Infact, most of the papers in Table 5 use various algorithms to identify dwarf subtype instead of SED fitting.

5.3 Subtype Identification using Color Space

Considering the drawbacks of subtype identification by SED fitting, it was decided to investigate identifying dwarf subtypes using color space, the method used by Ryan et al. (2011). The decision to investigate this method as opposed to any of the algorithms used by other papers was mainly made due to time constraints. Note that "color" refers to subtracting two filter-integrated fluxes from each other, such as J-H, ie: the filter-integrated H band flux subtracted from the filter-integrated J band flux. A color color space diagram can then be made by plotting this on each axis, such as the one in Fig 15, which is in J-H vs Y-J.

Ryan et al. (2011) Fig 3 and 4 shows the manner in which they have done this; boxes in color color space constrain objects as brown dwarfs of a certain type. However, here a more simple approach was opted for; simply plotting the SpeX standards, and the dwarfs to be typed, in color color space. The color space initially chosen was J-H vs Y-J as this is the same space used by Ryan et al. (2011). Then it is determined which SpeX standard each dwarf is geometrically closest to in this space, and the dwarf is given the subtype of that SpeX standard. Since the new set of SpeX standards we chose has one object of every subtype, every dwarf should be equally well represented. This is demonstrated in Fig 15. This was applied to the data from the previous semester to produce a new catalogue with subtypes identified by color. The fake catalogue used to generate completenesses was then also re-done to identify dwarf subtypes this way, resulting in data and a model distribution to fit to the data for this process. The data and fits for this are shown in Fig 16. This resulted in new scale height values of $H_{zM} = 330 \pm 49$ pc, $H_{zL} = 330 \pm 41$ pc, $H_{zT} = 500 \pm 74$ pc, $H_z = 360 \pm 48$ pc.

Looking once again to other papers, the decrease to all values obtained compared to Table 4 brings our results generally closer in line. Especially for the overall H_z value, which is now almost within the error on the value of $H_z = 290 \pm 50$ pc found by Ryan et al. (2011). It also brings us much closer to the $H_z = 300 \pm 15$ pc value found Bochanski et al. (2010). Our H_{zM} is also now in direct agreement with Holwerda et al. (2014) and Jurić et al. (2008), whilst being very close to the other papers in Table 5. However, this moves our H_{zL} value out of direct agreement with Rosell et al. (2019) and Sorahana et al. (2019) (although it is still extremely close to the latter), which is interesting as this subtype identification method produces a H_{zL} fit with much lower error than the SED fitting method. Likewise, despite a significant decrease of our H_{zT} value from 950 to 500 pc, it is still very far from the values obtained by Aganze et al. (2022b) and Holwerda et al. (2014). Further investigation could attempt to narrow this discrepancy down to a specific difference in method.

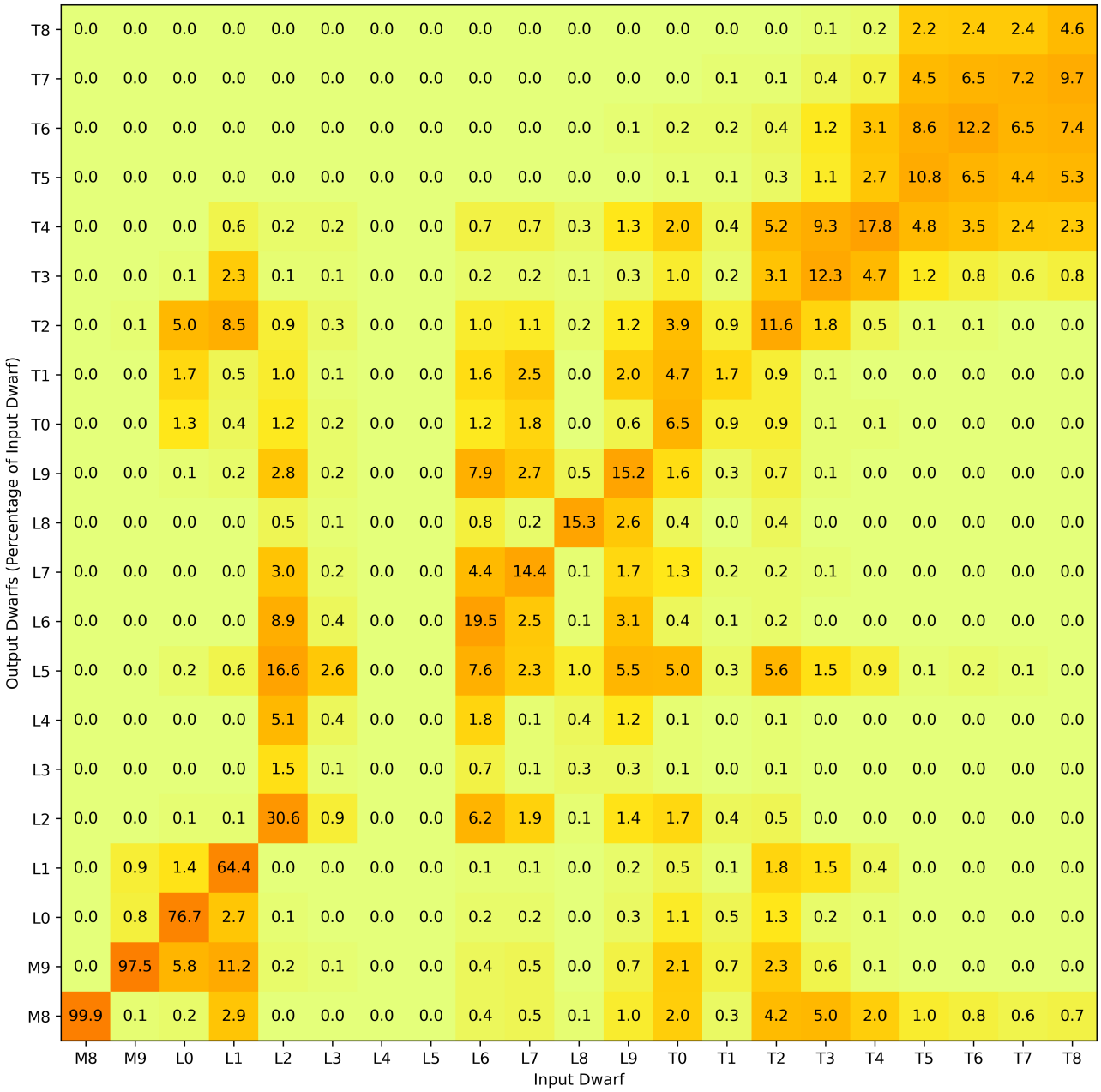

Figure 17. Input/output heatmap of dwarf subtypes for type identification in J-H, Y-J color color space. A much stronger linear trend can be seen compared to Fig 14. The axis are the same as Fig 14.

5.3.1 Comparison to SED fitting method

To compare this method to the SED fitting method, the input/output heatmap of subtypes can be looked at, the graph of which for subtype identification by color can be seen in Fig 17. Since the only change between this graph and Fig 14 is the different subtype identification method, comparing these two directly shows the impact of subtype identification by color. Some parts of Fig 14 and 17 are similar, such as the scatter at late T types, implying that it is indeed mostly caused by the noise in the data for these objects. The empty columns for L4 and L5 can still be seen too, as these are cases where eazy discards the objects as they are too poorly fit by the model dwarf SEDs, and changing how subtype identification is done will not fix poor representation in the LOWZ and SONORA model sets. However, the biggest change is the much stronger linear trend seen in Fig 17, suggesting that more types are correctly identified as what they are when using color color space. Also noticeable is the lack of empty rows, meaning that no subtypes end up not being assigned after subtype identification is gone. Both of these changes are a massive improvement, as the most ideal method of classification would have no subtypes missing and a perfectly linear input/output.

The completeness graphs themselves can also be directly compared (see Appendix A for the completeness graphs after subtype identification by color). Doing so also shows considerable improvements over the previous method, as expected from comparing the input/output heatmaps above. This includes the spike in the L6 completness being reduced from ~ 200 to ~ 2. Interestingly the spike in

the T6 dwarfs has disappeared, and been replaced by a larger spike in the T2 dwarfs. This further suggests that the T6 spike could have been caused by the random drawing of dwarfs as proposed in Section 5.2, as such a chance effect would likely not be repeatable. One interesting phenomena in the new completeness graphs is the spike in the M9 completeness to ~ 7. This can be understood from Fig 15 as of the two M type SpeX standards (in red), M9 is the lower down one, and can be seen to be closer to the vast majority of dwarfs. A large scatter of nearby types is therefore expected, and in Fig 17 it can be seen that the early L types contribute significantly to the M9s. There is also a spike in the L1 dwarf completeness at high magnitudes, although more time would be needed to fully investigate this.

Also briefly investigated was the use of different color-color spaces other than J-H vs Y-J, to determine if any were particularly better. Y-J vs H-Ks was looked at but seen to make a minimal difference other than having a larger H_{zT} value. This is included in Appendix A for a full comparison. Further investigation into other color spaces could be explored as an extension to this project, as the COSMOS catalogue provides 9 different observation bands in the NIR range.

5.4 Scale Height Dependence and Physical Implications of Results

The dependence of scale height on lateness of dwarf type is still a question up for discussion, as mentioned in the introduction. Since brown dwarfs cool through later spectral types as they age, as discussed by Baraffe et al. (2003) and Burrows et al. (1997), later type brown dwarf populations would be expected to consist of older dwarfs. Since older stellar populations are more dynamically relaxed, they would be expected to have an increased vertical scale height, see Ma et al. (2017) for simulation results. As a result, one naively expects that the scale heights of brown dwarfs will increase with lateness of type. However, some papers such as Ryan et al. (2017) argue that due to evolutionary effects, later types (L to early T) would have a reduced scale height. This is supported in data by Aganze et al. (2022b) and Holwerda et al. (2014), but would be in disagreement with Rosell et al. (2019). Looking at our results, this is consistent with our H_{zL} result both before and after applying typing by color color space, where instead of increasing our scale height flattens and is roughly similar to H_{zM}. However, the large scale H_{zT} value is not in agreement with this, as it is noticeably higher than both H_{zM} and H_{zL}. The tension between H_{zL} not being larger than H_{zM} but H_{zT} being much larger than both means that we are not able to conclusively say if our results agree that scale height should either increase with spectral type lateness, or if it decreases/flattens due to evolutionary effects (as suggested by Ryan et al. (2017)).

Physically, our results would suggest a thin disk consisting of M and L dwarfs with a less dense, much larger disk of T dwarfs. This could suggest that we have either thick disk or halo T dwarfs in our data. Since only a thin disk has been modelled, this would drive up the scale height of the T dwarfs. Alternatively it could be the result of further contamination of the data that has gone unnoticed.

6 CONCLUSION

Ultimately, using a catalogue of 1580 brown dwarfs from the COSMOS field survey (identified previously in Koplick (2024)) and a thin disk model, we find scale heights for each type of dwarf as $H_{zM} = 420 \pm 52$ pc, $H_{zL} = 440 \pm 84$ pc and $H_{zT} = 490 \pm 64$ pc, as well as an overall scale height for the entire population of $H_z = 490 \pm 64$ pc. These values are roughly in line with other papers shown in Table 4, but generally too large. After reviewing differences between our method and those used by other papers, the selection chain was changed so that identifying the subtype of dwarfs was done in color color space (similar to Ryan et al. (2011)). This resulted in new values of $H_{zM} = 330 \pm 39$ pc, $H_{zL} = 330 \pm 41$ pc, $H_{zT} = 500\pm = 74$ pc, $H_z = 360 \pm 48$ pc. These new values on the whole agree substantially better with the other papers listed in Table 4, and directly agree in many cases. We hope that these values help constrain the scale height of brown dwarf populations. From the above, we conclude that using the LOWZ and SONORA model dwarf SEDs to identfy dwarf subtypes is not a suitable method, and that many dwarfs are poorly represented by these models. This includes L4 and L5 dwarfs which are always fit too poorly by these models to be selected, as shown in Fig 14. Using more models in conjunction with these (such as the Cholla models in Karalidi et al. (2021), focusing on late type dwarfs) could greatly improve on this, as hopefully all dwarf types could be properly represented and a higher resolution in temperature space could be achieved. Future projects could also investigate different color color spaces to determine which bands are the most helpful for isolating and determining the subtype of brown dwarfs, or even analysis of the different methods used by different papers to further determine the discrepancies of scale height estimates within the last 20 years.

7 ACKNOWLEDGEMENTS

This research has benefited from the SpeX Prism Spectral Libraries, maintained by Adam Burgasser at http://www.browndwarfs.org/spexprism.

I would also like to express my appreciation once again to my project partner Ewan Scopes and supervisor Rebecca Bowler for their assistance and patience with me and my challenges.

REFERENCES

Adams N. J., Bowler R. A. A., Jarvis M. J., Varadaraj R. G., Häußler B., 2023, MNRAS, 523, 327
Aganze C., et al., 2022a, ApJ, 924, 114
Aganze C., et al., 2022b, The Astrophysical Journal, 934, 73
Auddy S., Basu S., Valluri S. R., 2016, Advances in Astronomy, 2016, 1–15
Bahcall J. N., Soneira R. M., 1980, ApJS, 44, 73
Baraffe I., Chabrier G., Barman T. S., Allard F., Hauschildt P. H., 2003, Astronomy & Astrophysics, 402, 701–712
Barnett R., et al., 2019, Astronomy & Astrophysics, 631, A85
Bochanski J. J., Hawley S. L., Covey K. R., West A. A., Reid I. N., Golimowski D. A., Ivezić , 2010, The Astronomical Journal, 139, 2679–2699
Bowler R. A. A., Dunlop J. S., McLure R. J., McLeod D. J., 2016, Monthly Notices of the Royal Astronomical Society, 466, 3612
Brammer G., 2021, eazy-py, doi:10.5281/zenodo.5012704, https://github.com/gbrammer/eazy-py
Brown R. A., 2009, The Astrophysical Journal, 699, 711–715
Burgasser A. J., 2004, The Astrophysical Journal Supplement Series, 155, 191–207
Burgasser A., 2008a, The SpeX Prism Spectral Libraries, https://cass.ucsd.edu/~ajb/browndwarfs/spexprism/index.html
Burgasser A. J., 2008b, Proceedings of the International Astronomical Union, 4, 317–326
Burgasser A. J., et al., 2024, UNCOVER: JWST Spectroscopy of Three Cold Brown Dwarfs at Kiloparsec-scale Distances (arXiv:2308.12107)
Burrows A., et al., 1997, ApJ, 491, 856
Caballero J. A., Burgasser A. J., Klement R., 2008, Astronomy & Astrophysics, 488, 181–190
Chen B., et al., 2001, ApJ, 553, 184

Euclid Collaboration et al., 2022, A&A, 662, A112
EuclidConsortium 2023, Euclid for Scientists: overview, https://www.euclid-ec.org/science/overview/%7D%7D
Fuhrmann K., 1998, A&A, 338, 161
Gilmore G., Reid N., 1983, Monthly Notices of the Royal Astronomical Society, 202, 1025
Guieu S., Dougados C., Monin J. L., Magnier E., Martín E. L., 2006, A&A, 446, 485
Holwerda B. W., et al., 2014, ApJ, 788, 77
Holwerda B. W., et al., 2023, Cosmic Evolution Early Release Science Survey (CEERS): Multi-classing Galactic Dwarf Stars in the deep JWST/NIRCam (arXiv:2309.05835)
Humphreys R. M., Larsen J. A., 1995, AJ, 110, 2183
Joshi Y. C., 2007, MNRAS, 378, 768
Jurić M., et al., 2008, ApJ, 673, 864
Karalidi T., Marley M., Fortney J. J., Morley C., Saumon D., Lupu R., Visscher C., Freedman R., 2021, The Astrophysical Journal, 923, 269
Kochanek C. S., 1996, The Astrophysical Journal, 457, 228
Koplick A., 2024, Finding Brown Dwarfs Within the COSMOS Survey, https://astro.twink.cafe
Kumar S. S., 1962, AJ, 67, 579
Ma X., Hopkins P. F., Wetzel A. R., Kirby E. N., Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Quataert E., 2017, Monthly Notices of the Royal Astronomical Society, 467, 2430
Mamajek E., 2022, A Modern Mean Dwarf Stellar Color and Effective Temperature Sequence, http://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txt
Marley M. S., et al., 2021, ApJ, 920, 85
McMillan P. J., 2011, Monthly Notices of the Royal Astronomical Society, 414, 2446
Mclean I., McGovern M., Burgasser A., Kirkpatrick J., Prato L., Kim S., 2003, The Astrophysical Journal, 596
Meisner A. M., et al., 2021, The Astrophysical Journal, 915, 120
Mužić K., Schödel R., Scholz A., Geers V. C., Jayawardhana R., Ascenso J., Cieza L. A., 2017, Monthly Notices of the Royal Astronomical Society, 471, 3699
Oort J. H., 1926, PhD thesis, -
Pirzkal N., et al., 2005, ApJ, 622, 319
Pirzkal N., et al., 2009, The Astrophysical Journal, 695, 1591–1603
Racca G. D., et al., 2016, in MacEwen H. A., Fazio G. G., Lystrup M., Batalha N., Siegler N., Tong E. C., eds, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. SPIE, doi:10.1117/12.2230762, http://dx.doi.org/10.1117/12.2230762
Rebolo R., Zapatero Osorio M. R., Martín E. L., 1995, Nature, 377, 129
Rosell A., et al., 2019, Monthly Notices of the Royal Astronomical Society, 489, 5301–5325
Ryan R. E., et al., 2011, The Astrophysical Journal, 739, 83
Ryan Russell E. J., et al., 2017, ApJ, 847, 53
Savitzky A., Golay M. J. E., 1964, Analytical Chemistry, 36, 1627
Sorahana S., Nakajima T., Matsuoka Y., 2019, The Astrophysical Journal, 870, 118
Stanek K. Z., Garnavich P. M., 1998, ApJ, 503, L131
Sutherland W., et al., 2015, Astronomy & Astrophysics, 575, A25
Virtanen P., et al., 2020, Nature Methods, 17, 261
Vugrin K., Swiler L., Roberts R., Stucky-Mack N., Sullivan S., 2007, Water Resources Research - WATER RESOUR RES, 43
van Vledder I., van der Vlugt D., Holwerda B. W., Kenworthy M. A., Bouwens R. J., Trenti M., 2016, Monthly Notices of the Royal Astronomical Society, 458, 425

APPENDIX A: ADDITIONAL GRAPHS

Additional graphs are included here for full comparisons to be made. Fig A1 shows completeness graphs for subtype identification in Y-J vs H-Ks space, and Fig A2 shows the fits for this.

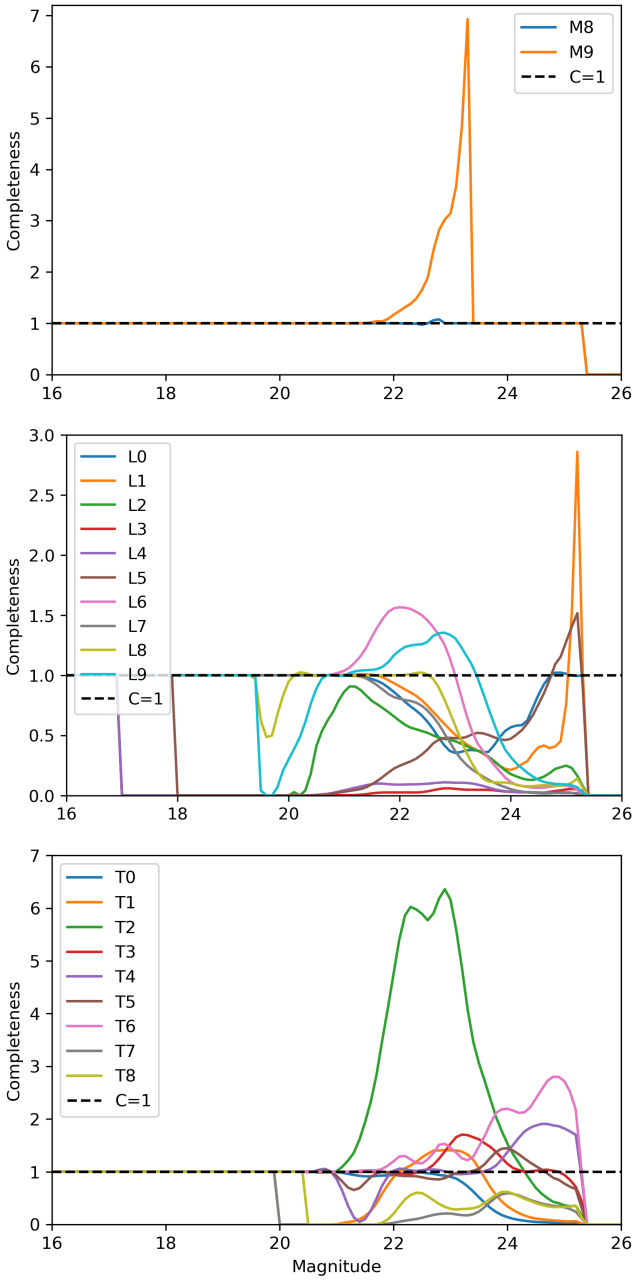

Figure A1. Completeness graphs for our selection chain when performing subtype identification in Y-J H-Ks color color space, in magnitude vs completeness (C), to be compared to Fig 10.

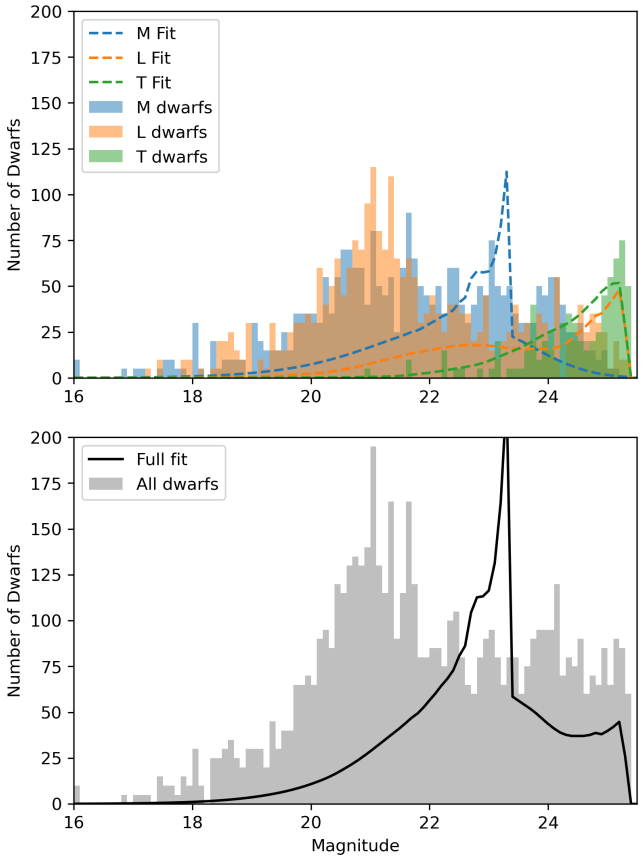

Figure A2. Fits of the data for Y-J vs H-Ks color color space typing, to be compared to Fig 16. The axis and colors are the same as 16.