Jacob Andrew Christensen
Project #2: “The H-R Diagram”
Physics 228, Winter 2022

Introduction

The Hertzsprung-Russel (H-R) Diagram is a powerful classification tool which allows us to compare a star’s color to its magnitude, or temperature to brightness. Based solely on a star’s position on the H-R Diagram, an astronomer can deduce its mass and composition, which reveals important information such as evolutionary phase (where the star came from and where it will be going) and internal structure. But just how accurate is the theoretical H-R diagram? In an effort to demonstrate how an H-R diagram works and the validity of such predictions, I will be conducting an analysis on stars catalogued by the Yale Trigonometric Parallax Dataset. This extensive dataset provides us with a star’s measured parallax and corresponding errors, which, along with parameters including apparent magnitude and color, will allow us to generate structures to be compared to predictions. Additionally, I will investigate where different stars lie on the graph based on parameters we measure besides the ones explicitly shown, giving a glimpse into the wide range of uses the H-R diagram has. Finally, I will demonstrate how errors in measurement change a star’s position on the diagram, important for judging appropriate error tolerance and further establishing the relationships which define an H-R diagram.

Procedure/Data

A total of 6220 stars were retrieved from the Yale dataset, a large enough sample to clearly reveal patterns to extrapolate from. Each star that was selected had its designation, parallax (arcseconds), apparent magnitude (v filter), B-V color index, and parallax uncertainty (milliarcseconds) recorded. I imported these values into Microsoft Excel to utilize its plotting tools, from which I created several different H-R diagrams (figures 1-6). Besides color, doing this required an absolute magnitude for each star, which I calculated from distance given by parallax and apparent magnitude. An example star is given in Table 1, along with the formulation for distance and absolute magnitude derived from the given parameters. Each diagram was constructed on the same scale to allow for proper cross-examination. Note that the vertical axes are flipped per H-R diagram tradition, so that the brightest and bluest stars appear in the upper left-hand corner of each graph.

Analysis

To begin my analysis, I created a baseline H-R diagram with every star in the sample displayed; I will discuss how this graph compares to the theoretical one in features such as shape. Then, I will analyze the effects of error in parallax measurements by creating three H-R diagrams, each sensitive to a different error threshold. To support my visual inspections, I make calculations for the expected positional shifts based on a given error. Finally, I reserve two sections for demonstrating how distance and apparent brightness isolate certain stars on the H-R diagram and offer interpretations for the behavior of each. It is important in any scientific analysis to include details that offer context. Thus, every H-R diagram includes a larger, red point that represents our Sun.

Analysis 1

Figure 1 displays the results of graphing every single star's absolute magnitude against B-V color. I marked each region according to the theoretical H-R diagram and my own intuition. Immediately we recognize the iconic shape of the H-R diagram. The main sequence is decently well-defined, with groupings of stars both above and below which I designate as white dwarfs and giant stars. Something to note is the width of the main sequence—it is much wider than the theoretical curve. This seems to cause the main sequence to bleed into the giant stars. Looking at the giants section itself, we can see a semblance of the red-giant branch. Notice how the grouping itself tapers off towards the top, an indication of where the supergiants reside, but still is not entirely clear in this manner. Also, when looking further up the main sequence, we do not have a well-defined horizontal branch from the giant/supergiant sections, a possible result of sampling bias, or inaccuracy in measurements. The white dwarf group, in contrast to the giants, is clear and well-defined, sitting exactly where I would expect it to be. It would be hard not to notice the outlying points on this graph, with one point as far as 4.25 B-V value. We do not see such stars on the "perfect" H-R diagram; either the ideal graph only considers the most abundant stars without mentioning oddities, or these points are spurious. Finally, I considered the Sun, the largest and red point. As expected, it lies within the main sequence and near its middle, next to the branch to the giants section.

Analysis 2

To explore the effects of parallax error, I made three separate H-R diagrams, one for stars with parallax error smaller than 50 mas (Figure 2), another with error smaller than 100 mas (Figure 3), and a final graph with error larger than 100 mas (Figure 4). Examining each after the other, we notice that the width of the main sequence increases with increasing error values. This is consistent with the idea of noise in data, where the variability in values increases as more error is introduced. We know that the absolute magnitude that we plot on the H-R diagram depends on the distance we measure to a star, which is proportional to its parallax angle. So, more variability caused by error in parallax angle means more variability in absolute magnitude, which shifts our stars upwards or downwards on the graph, causing the width to increase. Something I found interesting is that the more outlandish points only appear with the higher error values, including the outliers far to the right of the main sequence. For instance, the point at 4.25 B-V value has an error of 170 mas in its parallax measurement, a very high value. This is good evidence for such points being spurious data. Also, the white dwarf grouping all but vanishes in the >100 mas graph, suggesting that the measurements done on these stars were very accurate.

To further investigate the effects of parallax error, I made some calculations of expected magnitude shifts based on an uncertainty percentage of 10 percent. These calculations can be found as a supplementary attachment labeled “Displacement Calculation”. I found that a deviation which causes a 10 percent overshoot of the actual parallax will cause an increase of 0.207 magnitude (a decrease in brightness), and an undershoot of 10 percent will cause a decrease of 0.228 (an increase in brightness). Does this error effect a star’s B-V color, too? An error in parallax means an error in distance. But B-V color does not depend on distance as it describes the color of a star. It is normalized in this sense, giving the same answer whether you use the difference in apparent or absolute magnitudes. So even though a difference in distance changes absolute magnitude, taking the difference between the two removes any consideration of such, and the star remains in its same position on the H-R diagram B-V axis.

Analysis 3

It is interesting to study the stars which appear closest to our Sun. This area is referred to as the “solar neighborhood” and constitutes a good random sample of stars. To reflect this in an H-R diagram, I sorted the stars by distance and only graphed the 250 closest ones to the Sun (Figure 5). From this graph we see that most stars that are closest to the Sun lie on the main sequence, just like it does, with the Sun being one of the brighter stars. It appears that essentially every single giant star has been removed from the graph. Also, the graph is missing many of the extremely luminous stars. Note that we do still have some white dwarfs, however. If we consider this a good representation of the Milky Way star population as a whole, the most abundant star population would be the brown dwarfs and red, dim stars. Thus, the Sun is within the more abundant star groups, but not within the *most* abundant population. Does this mean we see mostly brown dwarfs when looking at the night sky? It is possible that, according to the average night-time sky observer, the most abundant star populations are actually the large and luminous stars. This simple error is due to how much easier it is to see brighter stars than dimmer ones, especially with the naked eye.

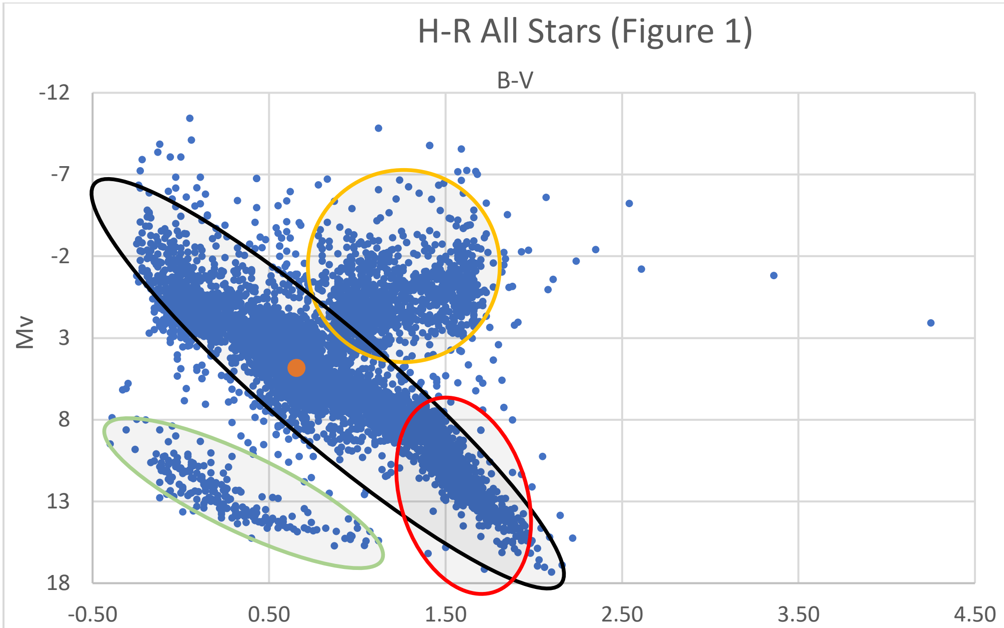
Analysis 4

As my final analysis, I created an H-R diagram for the brightest 250 stars in our night sky, based on apparent magnitudes (Figure 6). To do this, I sorted stars from smallest apparent magnitude to largest (brightest to dimmest) and graphed only the first 250 in the list. This graph is almost an opposite to the closest 250 graph. This time, the entire middle and bottom end of the main sequence has been removed—only the giants and very luminous main sequence stars remain. What is interesting here is that, even though we sorted the data based on apparent magnitude, it was only stars with a large absolute brightness that remained. I initially predicted that the distribution of this graph would include more star groups, by merit of distance and how it affects the light we receive from a star. It appears that stars which are truly brighter than others still win out against distance and its effect on apparent magnitude. For this reason, doing any sort of apparent magnitude-limited survey will preferentially select stars whose true brightness tend to be within this limit as well. This is not to say that an H-R diagram created from apparent magnitudes would encode as much information; distance still is a factor, which would cause many stars to be grouped together that otherwise would not be if distances were normalized. Also, we would be unable to find parameters such as a star's luminosity from such a graph. So, the graphs of closest 250 and brightest 250 are seemingly opposites because less abundant luminous stars are still the most apparently bright by a large amount. This means that an H-R diagram of the most apparently bright stars does not constitute a good representation of the general populace of stars (compared to the solar neighborhood H-R diagram).

Conclusion

My findings only highlight a few of the ways we can use the powerful tool that is the H-R diagram. I found that, first and foremost, the theoretical H-R diagram does a satisfactory job of representing real data, considering the error and admittedly small sample size I worked with. I addressed the discrepancies between my graph and the ideal one both by analysis and by determining the impact that error in parallax has on data. I concluded that an error as small as 10 percent can shift magnitude values within half a magnitude's range, but B-V color remained the same. Then, I demonstrated how the H-R diagram can be used to study the populations of stars, comparing the nearest and brightest stars, finding that the stars we see the most are not necessarily the most abundant stars in the Milky Way.

To conclude, I will extend this analysis I have done here to other galaxies. We now know that the most abundant stars are the small and less luminous ones, and that the ones we tend to see are the more luminous ones. So, when observing a galaxy with all its light integrated together, which kinds of stars will we see from this spectrum? Even though red, small stars make up the majority of stars by a long shot, we will still only see the most luminous stars, evident by which stars show up on the H-R diagram as the brightest. This is just yet another way to apply the H-R diagram to our understanding of what is really out there, beyond our world. It can help us compare what we see to what is true about stars.

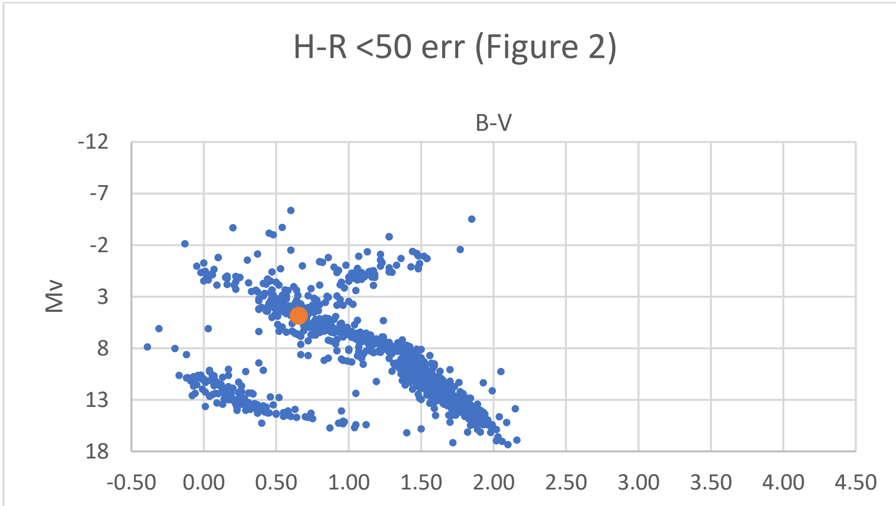


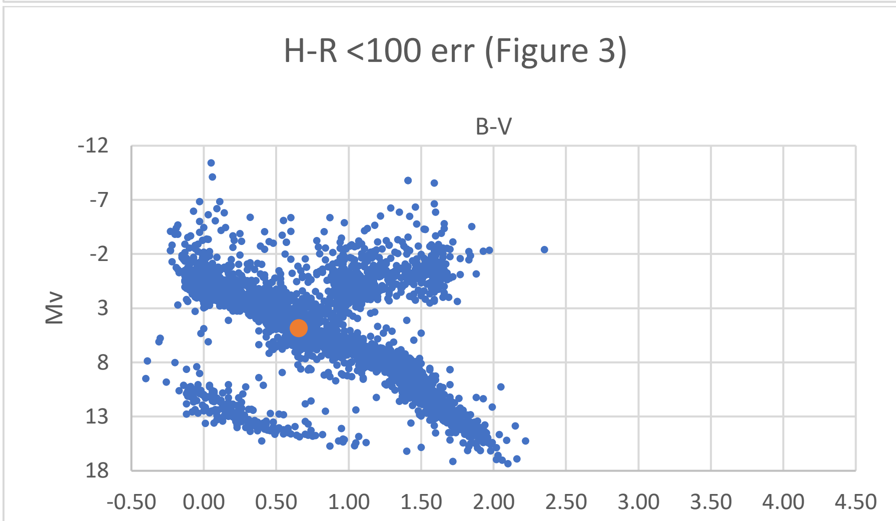
Black: Main Sequence
Orange: Giants
Green: White Dwarfs
Red: Red and Brown Dwarfs

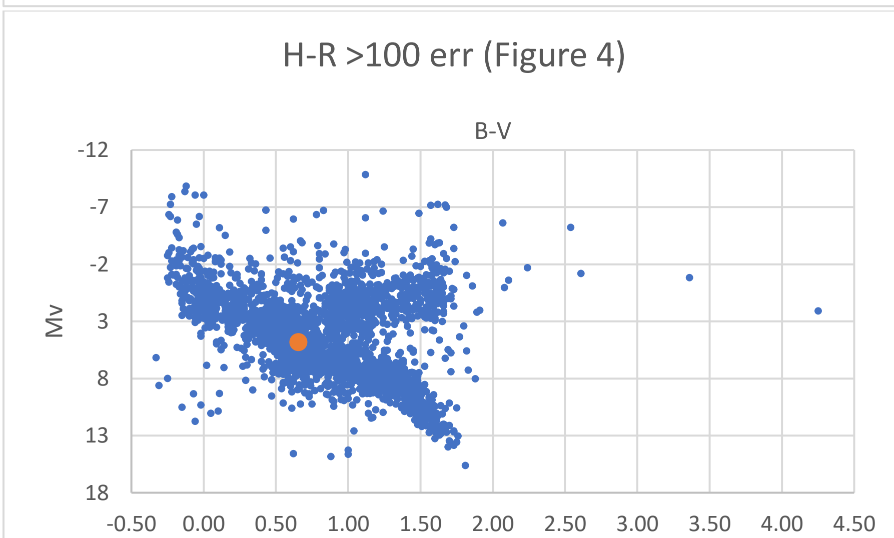
Table 1

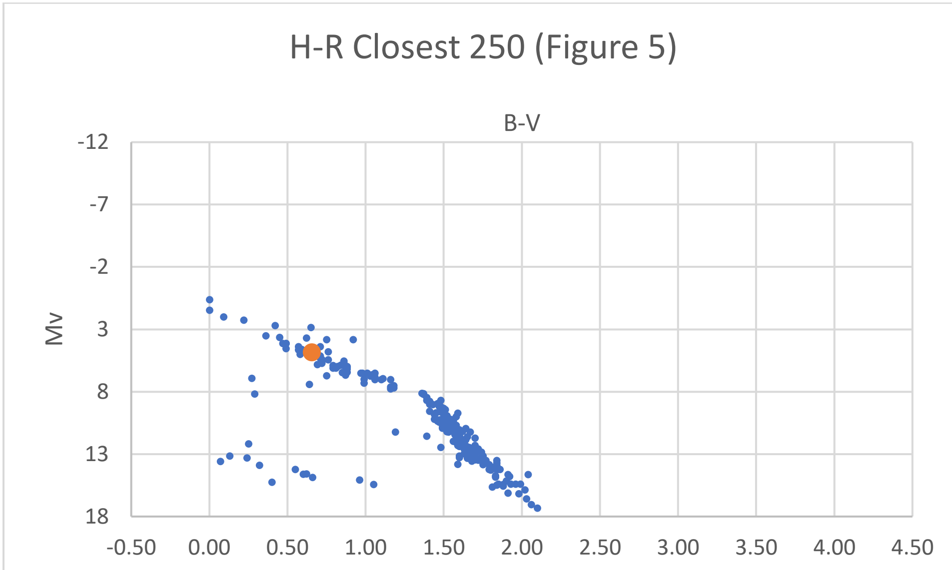
ID #	m_v	B-V	π"	err π (mas)	Mv
1577.00	-1.46	0.00	0.382	22	1.448042

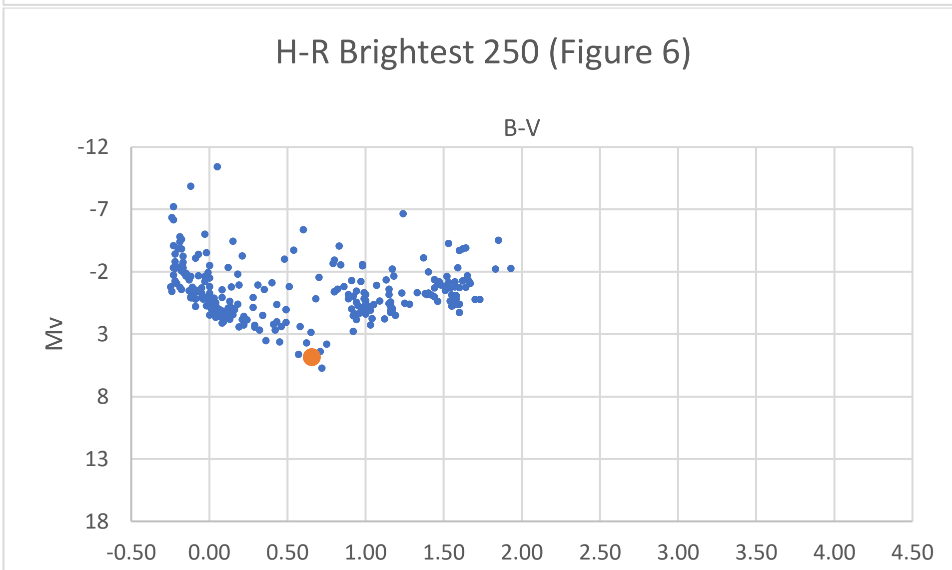
Equation: $M_v = m_v - 5 * \text{Log10}[1/(10 * \pi")]$


H-R <50 err (Figure 2)
B-V
Mv
-12
-7
-2
3
8
13
18
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50


H-R <100 err (Figure 3)
B-V
Mv
-12
-7
-2
3
8
13
18
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50


H-R >100 err (Figure 4)
B-V
Mv
-12
-7
-2
3
8
13
18
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50


H-R Closest 250 (Figure 5)
B-V
Mv
-12
-7
-2
3
8
13
18
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50


H-R Brightest 250 (Figure 6)
B-V
Mv
-12
-7
-2
3
8
13
18
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50

$\boxed{\text{Displacement Calculation}}$

$$M = m - 5\log\left(\frac{1/\pi''}{10}\right)$$

$$10\% = 100\% \times \frac{\delta\pi}{\pi}, \quad \frac{\delta\pi}{\pi} = 0.1$$

$$M = m - 5\log\left(\frac{1/(\pi'' \pm 0.1\pi'')}{10}\right)$$

$$M_{est} - M_{true} = -5\log\left(\frac{1/(\pi'' + 0.1\pi'')}{10}\right) + 5\log\left(\frac{1/\pi''}{10}\right)$$

overshoot

$$= 5\log\left(\frac{\left(\frac{1/\pi''}{10}\right)}{\frac{1/(\pi'' + 0.1\pi'')}{10}}\right)$$

$$= 5\log\left(\frac{1}{\pi'' \cancel{10}} \cdot \frac{1.1\pi'' \cancel{10}}{1}\right)$$

$$= 5\log(1.1) = \boxed{0.2069}$$

$$M_{est} - M_{true} = -5\log\left(\frac{1/(\pi'' - 0.1\pi'')}{10}\right) + 5\log\left(\frac{1/\pi''}{10}\right)$$

undershoot

$$= 5\log\left(\frac{1}{\cancel{\pi''}\cancel{10}} \cdot \frac{0.9\cancel{\pi''}\cancel{10}}{1}\right)$$

$$= 5\log(0.9) = \boxed{-0.228}$$

According to these differences, if a star's parallax is overshot by 10%, it'll find itself displaced downwards on the diagram by 0.2009 magnitude.

If a star is undershot by 10%, it'll move up on the diagram by 0.228 (technically down in value, though).