Comparing Positions of Molecular Clouds and Masers in the Milky Way

Sara Duval

ABSTRACT

In my project I compare Thomas Rice's dendrogram catalog of molecular gas in the Milky Way with data from the Bar and Spiral Structure Legacy (BeSSeL) Survey. Rice, a member of Harvard University's class of 2012, created this catalog as his senior thesis. He used the dendrogram technique to segment CO in the Milky Way into individual clouds. The dendrogram creates a tree diagram, in which the leaves are peaks in CO intensity. The BeSSeL Survey found parallaxes for methanol maser sources using the Very Large Baseline Array (VLBA). Because these sources are associated with regions of massive star formation in the spiral arms of the galaxy, their locations in the Milky Way yield information about the star forming regions in the spiral arms. They should correspond to peaks in CO, which were identified as clouds in the dendrogram. By matching points between the catalogs based on their locations and by plotting BeSSeL points directly on the dendrogram, I show correspondence between the two data sets. A parallax distance could be assigned to those clouds that have an associated maser, and these parallax distances could change what we know about the molecular clouds. I also show that the dendrogram technique is a viable automated technique for segmenting CO into clouds and assigning masers to those clouds.

1. Introduction

The Milky Way is our home galaxy, yet our knowledge of its structure is limited. The structure of the Milky Way is important not only because we would like to learn about our home galaxy. Knowledge of the structure affects the study of star formation. The locations of the spiral arms affect our knowledge about the distances to star forming regions in the

Galaxy, which in turn affects our measure of their sizes, pressure, and other properties. The difficulty in creating a model of the Milky Way stems from the fact that we are located in the plane of the Galaxy. It is impossible for us to get a top down view of the Galaxy, so we must rely on other methods of studying its structure. Studying the spiral structure of the Milky Way requires studying sources that are found in the spiral arms. Models of the galaxy have been created based on observations of gas, dust, and both young and old stars (Steiman-Cameron 2010). Common to all of these is the need to measure a distance for each source in order to determine its 3D location. Distances can be measured in many ways. For nearby sources, spectrophotometric distances can be found, but because of dust extinction farther sources usually depend on kinematic distances (Steiman-Cameron 2010). These distance measurements add difficulty to creating a model of the galaxy because kinematic distances have uncertainty due to their dependence on a rotational model and the distance ambiguity for sources interior to the Sun. Locations of spiral arms have also been found through electron density mapping, magnetic field mapping, and by assuming a spiral shape between tangent point directions of the arms (Steiman-Cameron 2010).

Due to the difficulty in studying the spiral structure of our Galaxy, there is not an agreed upon model. In summarizing the various models of the Galaxy, Steiman-Cameron (2010) stated that about 100 models have been proposed and that they differ not only in the location of arms but in the number of arms. Figure 1 shows just a few of the many models. Most models have arms that are roughly in the same positions, with either four or two main arms (Steiman-Cameron 2010). The Georgelin & Georgelin (1976) model includes 4 spiral arms and was created through a combination of radio observations of HII regions and optical observations of young stars. In contrast the Churchwell et al. (2009) model includes two main arms and two much less prominent arms. Their model is based on stellar overdensities on lines of sight tangent to spiral arms due to the increased path length through the arms (Churchwell et al. 2009). The lack of agreement between models makes the study of the spiral structure of the Milky Way an important research topic.

Part of contributing to the creation of models of our Galaxy involves finding 3D locations of sources found in the spiral arms. Thomas Rice (2012) created a catalog of molecular clouds in our Galaxy, assigning 3D locations to clouds in an effort to help trace the spiral arms and learn about the star forming regions of the Galaxy. He used the dendrogram technique to separate the Dame et al. (2001) CO survey into individual clouds. This method segments clouds hierarchically, creating a tree diagram in which the leaves are molecular clouds, branching off of larger and less intense clouds, branching off of even larger features (Rice 2012). Figure 2 is a diagram that shows how the dendrogram works. Rice then assigned kinematic distances to clouds using their velocities and the rotation curve of the Galaxy. His kinematic distances are subject to uncertainty due to uncertainty in the rotation curve as well as distance ambiguity in the inner galaxy. Distance ambiguity is caused by the fact that there are two possible distances for each velocity in the inner galaxy. Rice resolved this ambiguity by using the relationship between velocity dispersion and size of the cloud. Each distance leads to a different size for the cloud, and if one distance leads to a size that is inconsistent with the velocity dispersion then that distance can be eliminated (Rice 2012). Rice found 313 gravitationally bound clouds for which the

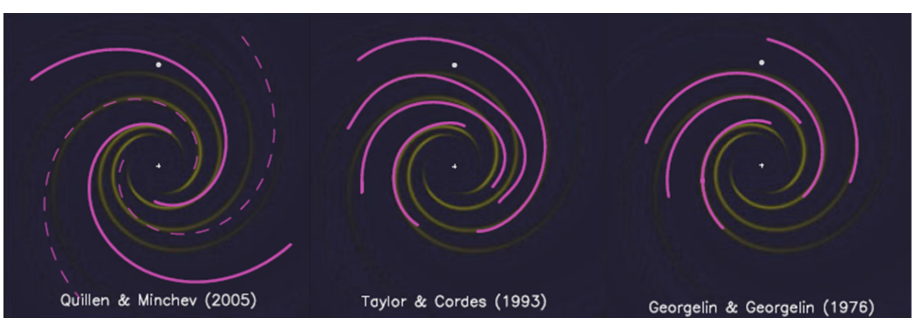

Fig. 1.— Three different models of the spiral arms of the Milky Way. (image taken from Steiman-Cameron, 2010)

disambiguation was successful.

In my project, I compared Rice's dendrogram catalog to data from the Bar and Spiral Structure Legacy (BeSSeL) Survey. This survey is measuring trigonometric parallaxes of methanol masers using the National Radio Astronomy Observatory's Very Long Baseline Array (VLBA) (Reid et al. 2009). The VLBA is an interferometer consisting of ten 25m radio telescopes spanning more than 5000 miles. The survey has measured parallax distances to 102 masers that are associated with star forming regions. In my project I looked at the consistency of Rice's 3D locations of molecular clouds with BeSSeL's more accurate positions of masers in order to contribute to the goal of assigning more accurate distances to the molecular clouds. Using parallax distances from BeSSeL would allow more accurate estimates of the properties of molecular clouds.

2. Methods

The first step in my project was to compare the two data sets visually to see whether they lined up at all. I did so using WorldWide Telescope. WorldWide Telescope is a Microsoft program that displays astronomical objects both on the sky and in 3D. By displaying Rice's molecular cloud catalog and the BeSSeL masers, I was able to compare them in the sky view and in a 3D view. I found that the two data sets have points that appear to line up very well, as well as some points that do not appear to line up at all. Figure 3 shows WorldWide Telescope's 3D model of the Milky Way with the BeSSeL points plotted in pink and molecular clouds plotted in red. It also shows a magnified view of one region of the sky in which the data sets line up well. After observing the data sets in WorldWide Telescope I decided that they agreed well enough to compare them further.

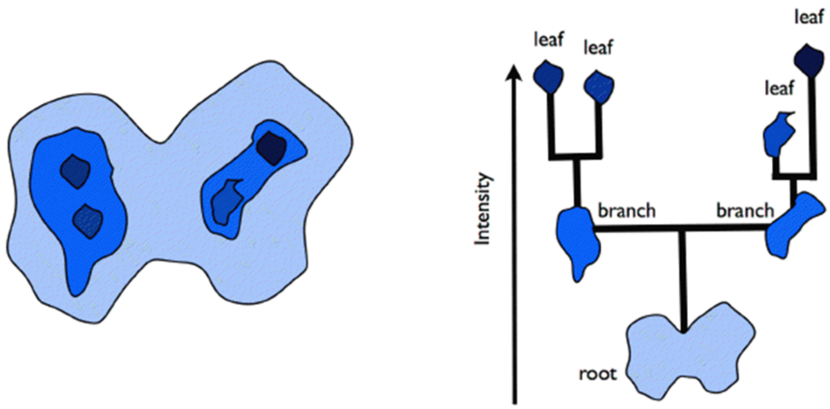

Fig. 2.— Image showing how a dendrogram is created from a contour map in intensity. Peaks in intensity become leaves that branch off of the less intense features. (image taken from Rice, 2012)

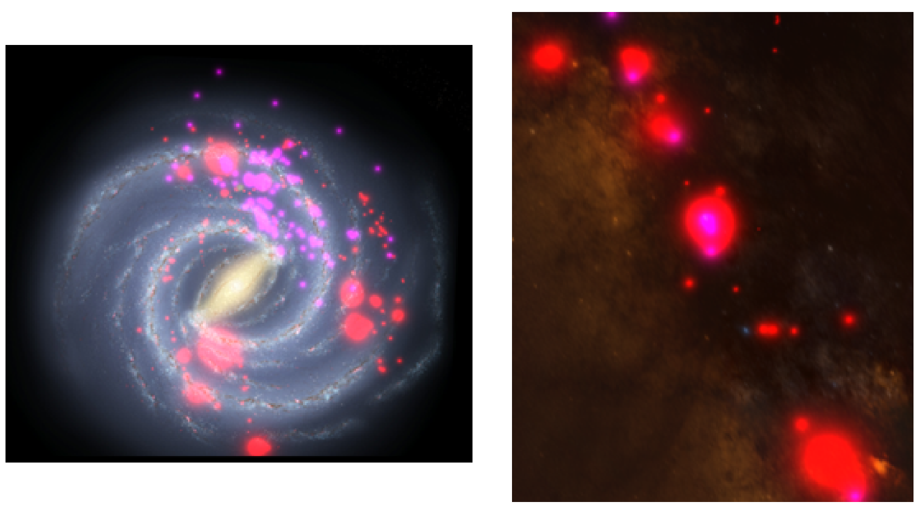

Fig. 3.— Left: A top down view of the Milky Way from WorldWide Telescope. Right: A region of the sky as viewed in WWT. Pink points are sources from BeSSeL and red points are molecular clouds from the dendrogram catalog. Sizes of the points are not to scale.

For the next phase in my project, I used Topcat to compare the two data sets. Topcat is a program that can match data across catalogs, both by sky coordinates and by 3D location. Figure 4 shows a plot of both data sets generated by Topcat overlayed on an image of the Milky Way. It is apparent in the plot that both data sets trace the spiral arms of the galaxy, but their agreement with each other is less apparent. I matched the catalogs in Topcat to see if I would get many matches or just a few. I found that half of the BeSSeL sources are within 5 degrees of a cloud from the dendrogram catalog. Topcat is not able to match sources within different distances for each point based on their sizes or errors, so I used 5 degrees as the angular distance for matching because it is the maximum angular extent of the molecular clouds. I also used Topcat to match sources in 3D. As the distance within which to match data points I used the average error in the BeSSeL data, 500 pc. I found 34 matches within 500 pc. I found enough matches in Topcat to make me want to try a more detailed way of looking at the correspondence between the two data sets.

Next, I plotted the BeSSeL points on Rice's dendrogram using an IDL program written by Christopher Beaumont. The program used the longitude, latitude, and velocity of the BeSSeL points to place them on the corresponding dendrogram feature. Its outputs are a plot showing the dendrogram diagram with the points overlayed, shown in figure 6, and a table of dendrogram feature number and x and y coordinates on the plot for each BeSSeL point. In figure 6 it is difficult to see the structure of the underlying dendrogram. This structure can be seen more clearly when zoomed in, as in figure 7. From these outputs, I was able to determine for each BeSSeL point the tree level (top, middle, or bottom) on which it is located. Masers at the top of the tree correspond to peaks in CO intensity, masers in the middle correspond to less intense and larger CO features, and masers near the bottom are not obviously associated with a particular CO cloud. Some examples of masers at the top of the dendrogram are points 10, 44, and 82 on Figure 5. Examples of points in the middle include 54 and 72, and examples of points on the bottom are 39 and

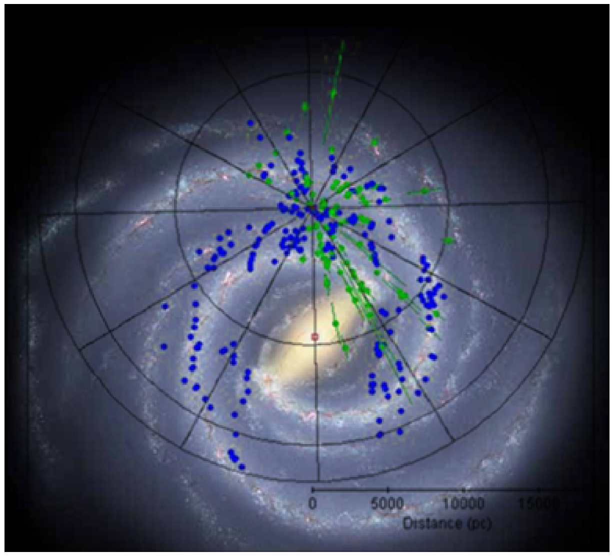

Fig. 4.— A top down plot of both catalogs created in Topcat. Green points are masers from BeSSeL and blue points are molecular clouds from the dendrogram catalog. The plot is overlayed on Robert Hurt's depiction of the Milky Way.

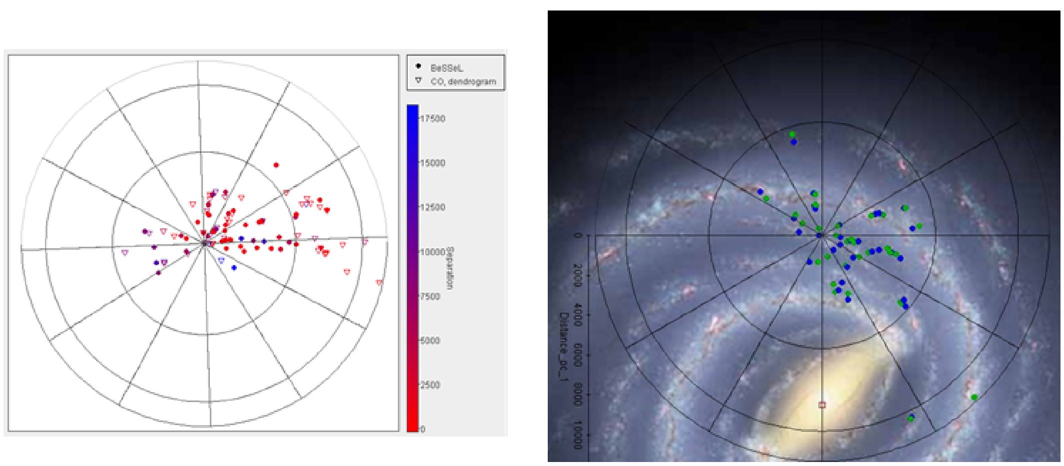

Fig. 5.— Left: A top down plot of the sources that were matched within 5 degrees. The color scale denotes 3D separation in parsecs, with blue being larger separation. Right: A top down plot of the sources that were matched within 500 pc, overlaid on Hurt's image of the galaxy. Green points are masers from BeSSeL and blue points are molecular clouds from the dendrogram catalog.

86. I also found for each BeSSeL point the number of other points that are located on the same dendrogram feature and whether or not the other points agree on distance. This information can show how well the two data sets agree. For example, points near the top of the tree and points on the same feature that agree on distance would indicate a good agreement between the two data sets. Points near the bottom of the tree or multiple points on the same feature with different distances would indicate poor agreement between the data sets, at least for those points. It is possible that there could be agreement for some points and disagreement for others.

I then looked at Thomas Dame's information about each BeSSeL source. He has assigned each BeSSeL point to a molecular cloud by eye. He has done so by going through velocity slices in CO for each point and finding where the maser best lines up with a molecular cloud. For many of the sources he was able to find a clear association between the maser and the CO. Figure 8 shows his image of the maser and CO for a BeSSeL point with a clear association, while Figure 9 shows an image for a point without a clear association to a molecular cloud. Dame was able to give each maser a distance based on the kinematic distance of the corresponding molecular cloud. From Dame's information, I was able to add to the information learned from the dendrogram plot whether or not each BeSSeL point had a clear association and the corresponding cloud distance. I was able to create a table that included for each maser the tree level, the number of other masers that share its location on the dendrogram, whether or not other masers on the same location agree in distance, and whether or not there is a clear association according to Dame. Table 1 summarizes these results for some of the BeSSeL points. I also was able to compare three distance measurements: the parallax distance from BeSSeL, the kinematic distance of the corresponding dendrogram feature, and the kinematic distance of the corresponding cloud as assigned by Dame. An important note is that kinematic distances depend on the rotation curve used; different rotation curves will lead to different distances for the same velocity.

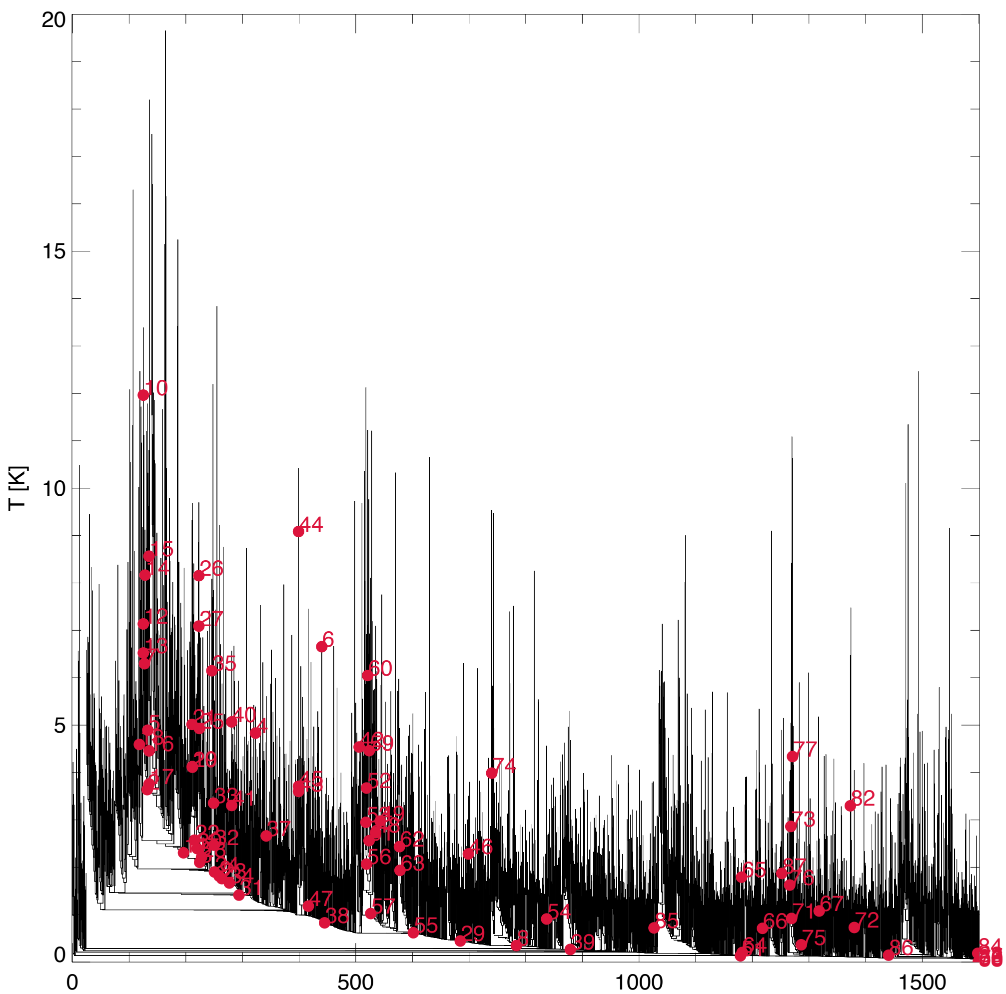

Fig. 6.— BeSSeL points plotted in red over the dendrogram from Rices catalog. This plot can be magnified to see more detail.

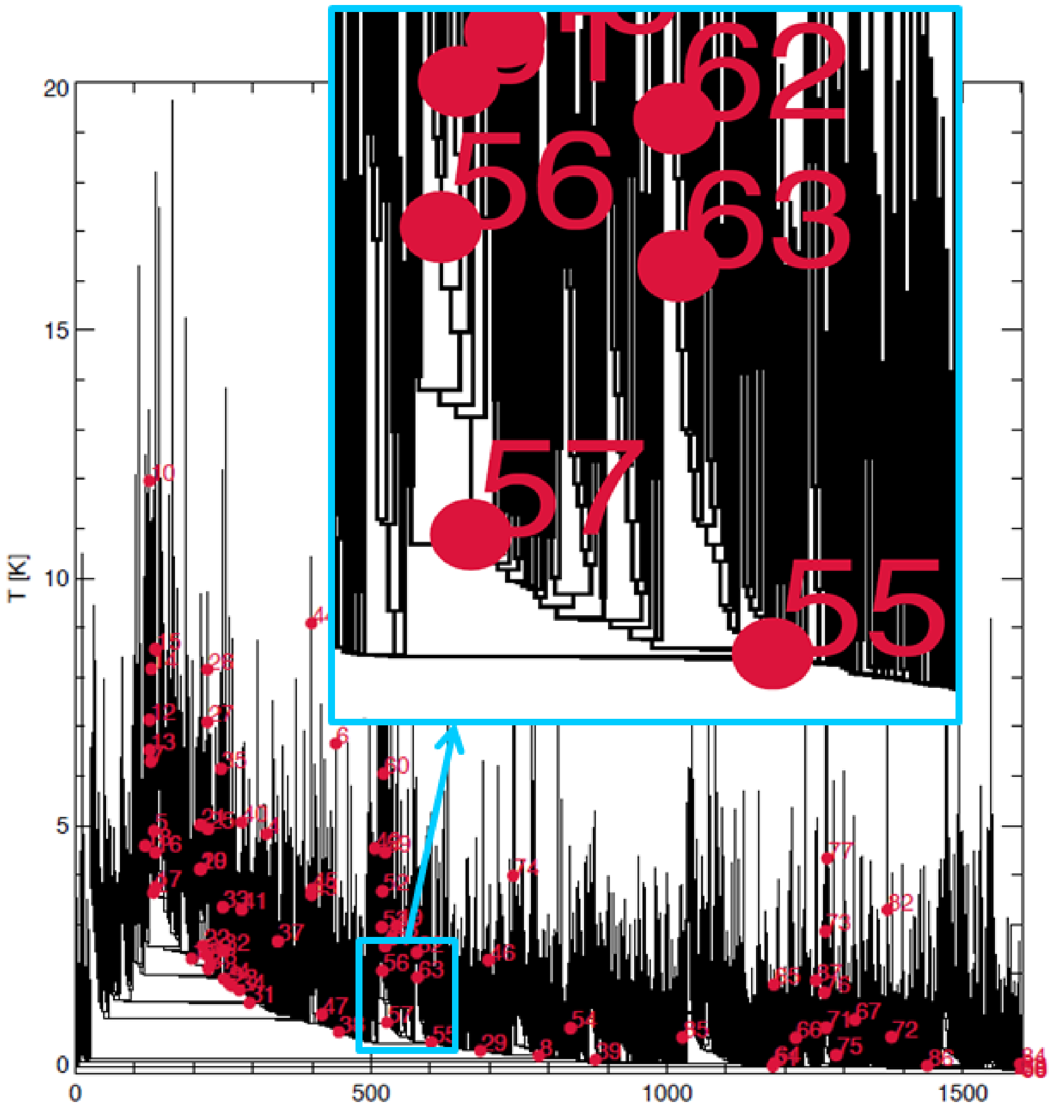

Fig. 7.— An image showing a zoomed in portion of the dendrogram to show the tree structure.

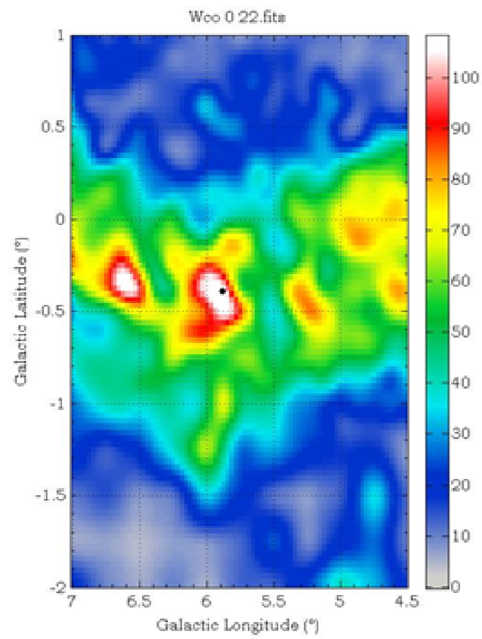

Fig. 8.— An image provided by Thomas Dame showing a clear association between a maser, shown as a black point, and CO.

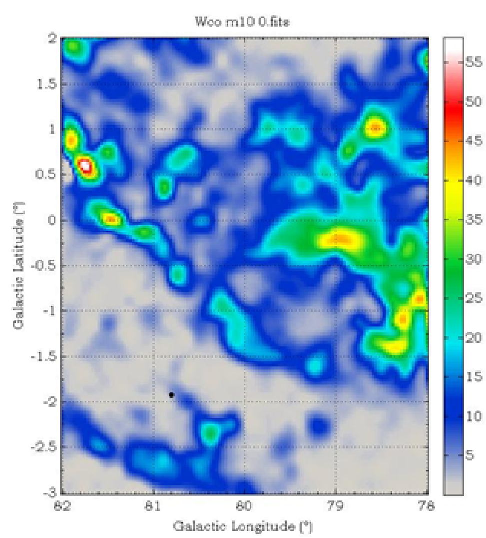

Fig. 9.— An image provided by Thomas Dame showing an unclear association between a maser, shown as a black point, and CO.

Rice and Dame used slightly different rotation curves, so this could explain some differences between the two kinematic distances.

3. Results

In my project, I found that the locations of BeSSeL sources show agreement with the locations of molecular clouds in the dendrogram. Changing the kinematic distances to parallax distances from BeSSeL would be a way to assign better distances to the molecular clouds and therefore better estimate the properties of those clouds. I also found that the dendrogram technique is effective in assigning masers to molecular clouds.

Through my work in Topcat, I found that half of the BeSSeL sources are within 5 degree of clouds from Rice's catalog and a third were within 500 pc. For both of these groups I created plots showing the distance according to Rice's catalog versus the distance according to the BeSSeL catalog. The points on the plot fall close to the line at which the two distances would be equal. For those that were matched based on sky coordinates, a few points showed large differences in the two distances. A possible reason for this could be errors in resolving the distance ambiguity. For example, for the point located at (1980, 15610) on figure 10, the near kinematic distance was chosen. The large parallax distance to the BeSSeL source indicates that the far kinematic distance should have been chosen instead. For points that were matched based on their 3D locations, the distances agree much better because those that were not close in distance were not matched. For these it is interesting to see whether Dame's cloud distances also agree well. Since these masers were found to be close to a molecular cloud in the dendrogram, it would make sense if Dame was able to assign them to a molecular cloud with a consistent distance. Figure 12 shows that Dame's cloud distances are similar to the BeSSeL distances for these sources.

Table 1. Information for some of the BeSSeL points.

Label on Plot	Galactic Long.	Galactic Lat.	Velocity (km/s)	Masers per Leaf	Ambiguous Distance	Tree Level	Clear Association
14	13.87	0.28	48	1	0	3	1
15	14.33	-0.64	22	1	0	3	1
16	14.63	-0.57	19	2	0	2	1
17	15.03	-0.67	22	2	0	2	1
18	16.58	-0.05	60	1	0	3	1
19	23	-0.41	80	2	0	3	1
20	23.43	-0.18	97	2	0	3	1
21	23.65	-0.12	83	1	0	3	1
22	23.7	-0.19	79	1	0	2	0
23	25.7	0.04	93	1	0	2	1
24	27.36	-0.16	92	1	0	2	1
25	28.86	0.06	100	1	0	3	no image
26	29.86	-0.04	100	2	1	3	1

Note. — For ambiguous distance, 1 means the distance is ambiguous, 0 means it is not. For tree level, 3 is top, 2 is middle, and 1 is bottom. For clear association, 1 is clear and 0 is unclear.

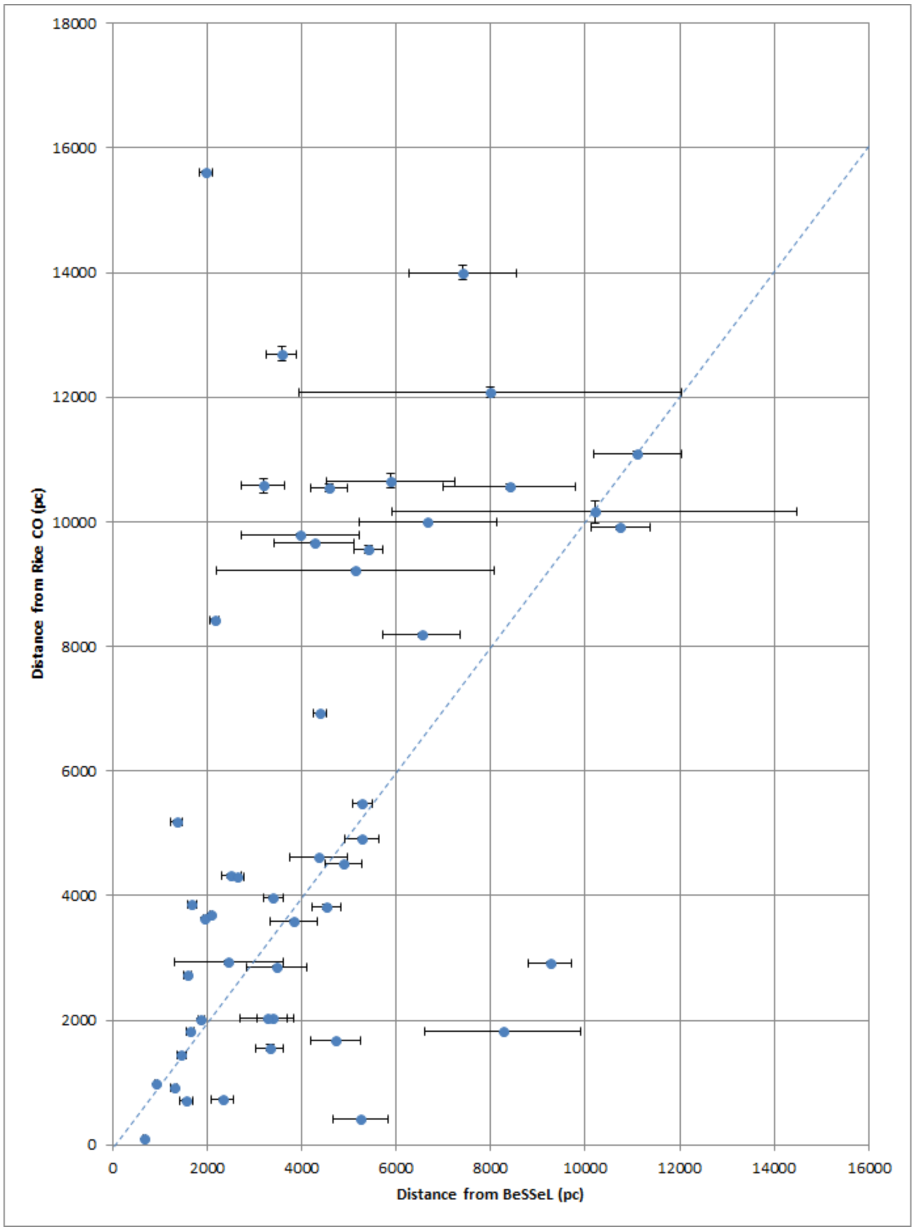

Fig. 10.— Distance comparison for sources matched within 5 degrees. The dashed line shows the line for which both distances are equal. Vertical error bars correspond to the size of the cloud in the dendrogram catalog, and horizontal error bars correspond to the error from BeSSeL.

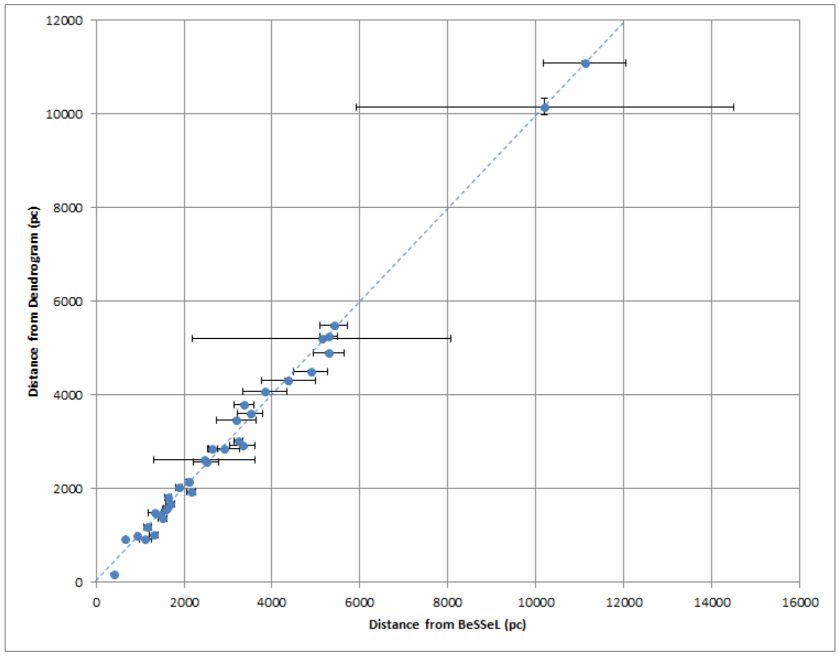

Fig. 11.— Distance comparison for sources matched within 500 pc. The dashed line shows the line for which both distances are equal. Vertical error bars correspond to the size of the cloud in the dendrogram catalog, and horizontal error bars correspond to the error from BeSSeL.

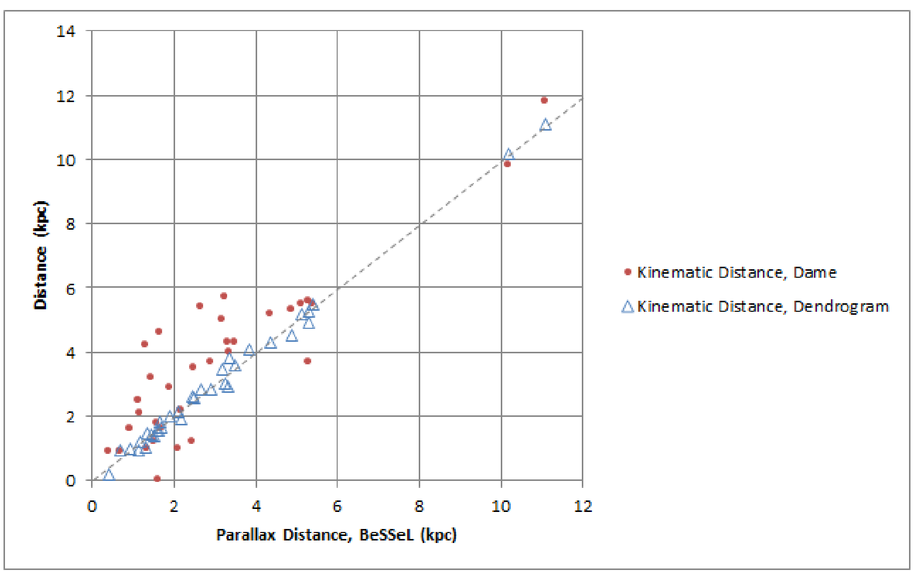

Fig. 12.— Distance comparison including the kinematic distance from the dendrogram and the kinematic distance from Dame.

The information listed in Table 1 allows for different ways in comparing the data. I found that 22 of the 102 BeSSeL points were not placed into the dendrogram. This is due to those objects being in areas of the sky that were not included in the dendrogram. I found that 10 of the 102 points are on the bottom of the dendrogram, 36 in the middle, and 34 at the top. Of those that were placed on the plot, 87% are at either the top or middle tree levels. Most of the points do not share a dendrogram feature with other points. For those that do, I found that 7 of the BeSSeL points have ambiguous distances, meaning that they do not agree with the distances of other points on the same dendrogram feature. Of those points with ambiguous distances, 5 were on the same feature at the bottom of the dendrogram. They were probably placed there because of a lack of clear association with any cloud. Two thirds of the points are at the top or middle tree level and have unambiguous distances. This is a good sign towards agreement of the two data sets. In my comparison to Dame's images of cloud association, I found that 69 of the 102 points have a clear association. Half of the points have a clear association according to Dame, an unambiguous distance, and are at the top or middle tree levels, and 20% of the points have a clear association according to Dame, an unambiguous distance, and are at the top tree level. This means that for at least half of the points, the data sets agree well. Figure 13 shows how many points fell into the categories of clear association, unambiguous distance, and tree level 2 or 3 (middle or top), and the various combinations of those categories.

For those points that were at the top or middle of the dendrogram, I created a plot showing the kinematic distance from the dendrogram catalog versus the parallax distance from BeSSeL. Most of the points agree in distance. A linear fit to the plot yielded a line of slope .98 and $R^2 = .6$ after removing one outlier. This fit supports an agreement between the two data sets, but with a lot of scatter. The scatter shows that changing from the kinematic distance to the parallax distance for each point could significantly affect its distance and therefore our estimates of that cloud's properties. I found the rms of

Fig. 13.— A Venn diagram showing the number of maser that fell into each of these categories.

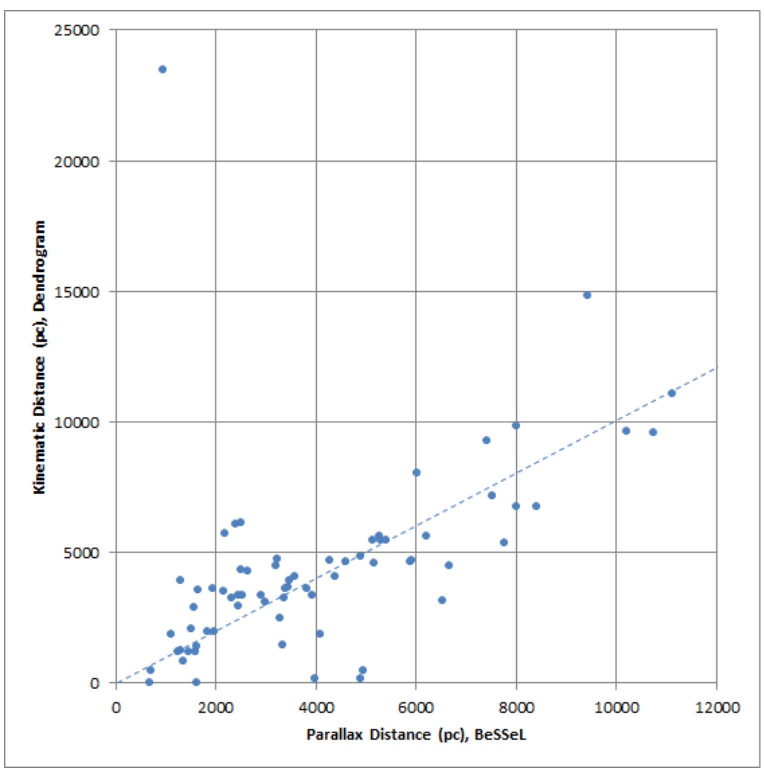

Fig. 14.— Distance comparison for sources at the top and the middle of the dendrogram.

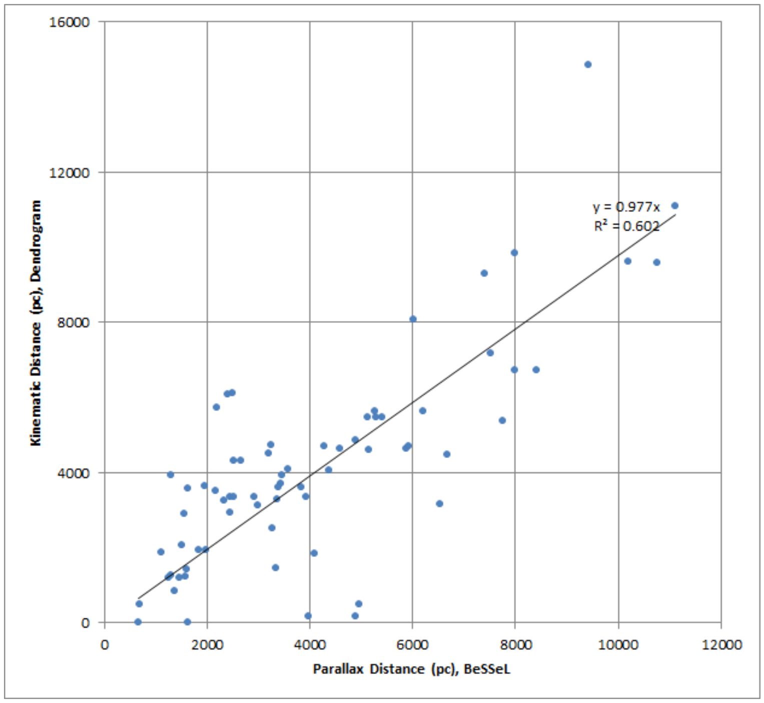

Fig. 15.— The plot from figure 14 with one outlier removed and a linear fit added.

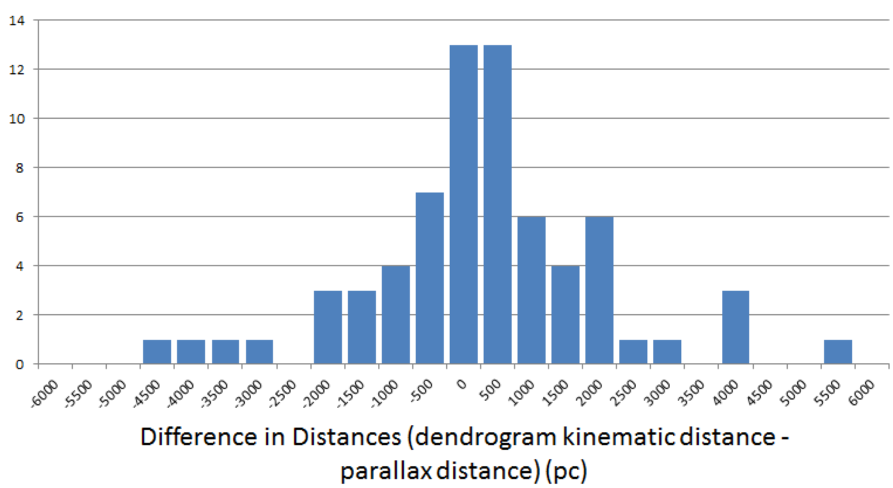

Fig. 16.— A histogram showing the residuals from the distance comparison.

the residuals of the points from the line corresponding to exact agreement to be 1746 pc (excluding the outlier). A histogram of the residuals is shown in figure 16. The histogram is symmetric about zero, which shows that there is not a systematic trend to the residuals.

I then looked at the points that indicate disagreement: those at the bottom of the dendrogram and those that have an unclear association according to Dame. For the 10 points at the bottom of the dendrogram, all but 2 have a clear association according to Dame. This would suggest that the dendrogram was not able to assign these points to the correct cloud. However, in comparing Dame's cloud distances for these points to the parallax distances for these points, I found that the two distances are off by an average of 1.6 kpc. Most of Dame's cloud distances differ from the parallax distances by less than 1 kpc. The disagreement between the two distances shows that though Dame was able to assign a cloud to these points, it was with less accuracy than other points. This means that the points' locations at the bottom of the dendrogram could be due to a lack of obvious association with molecular clouds and not due to a problem with the dendrogram. Of the 9 points that have an unclear association according to Dames images, 4 are in the middle of the dendrogram, 2 at the bottom, and 3 are not included on the plot. Ignoring those not included on the plot, this means the majority of those with an unclear association are in the middle of the dendrogram, not at the bottom. Upon reexamining the images, the points in the middle of the dendrogram make sense because they are towards the edge of a cloud or in a less intense part of the cloud in the images. Figures 17 and 18 show the difference between an unclear association corresponding to a point in the middle of the dendrogram and an unclear association corresponding to a point at the bottom of the dendrogram. These results show that points with an unclear association according to Dame but found in the middle of the dendrogram do not indicate that they were assigned incorrectly in the dendrogram. Through comparison to Dame's information of each BeSSeL point, I was able to show that the dendrogram was able to assign masers to molecular clouds about as well

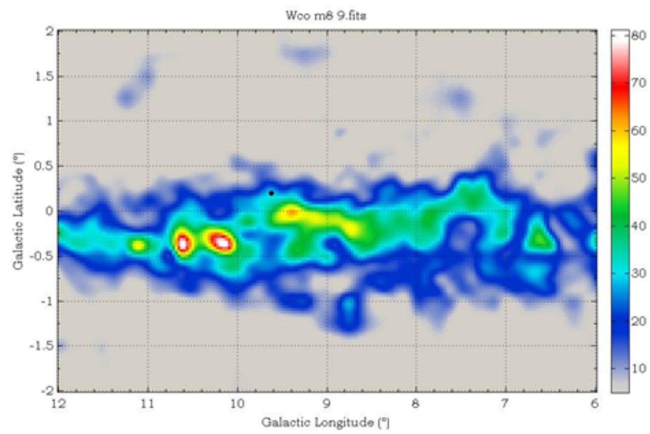

Fig. 17.— Image for a maser in the middle of the dendrogram. The maser is located on the edge of the cloud.

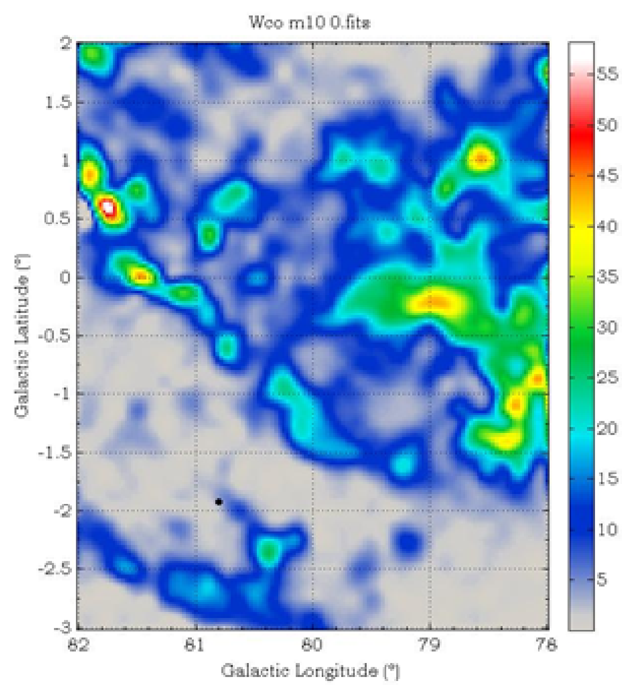

Fig. 18.— Image for a maser on the bottom of the dendrogram. The maser is not located in a cloud.

as Dame was able to do by eye.

4. Conclusion

By comparing data from the BeSSeL survey to Rice's dendrogram catalog, I was able to show that many of the masers are associated with molecular clouds. A comparison of the kinematic distances from the dendrogram and the parallax distances from BeSSeL shows that replacing the kinematic distances with the more accurate parallax distances could significantly affect our estimates of the properties of the molecular clouds. For example, changing to parallax distances could change our estimates of the masses of these clouds by an average factor of about 2.5. By comparing to Dame's information about the BeSSeL points, I was able to show that the dendrogram was able to assign masers to molecular clouds as well as Dame was able to do by eye. If the performance is equal, the dendrogram is a better option for this task because it is automated. It can do this for a large amount of data and calculate the properties of the molecular clouds at the same time.

5. Future Work

More information could be learned from comparing these two data sets. An important thing to do next would be to add the parallax distance of associated masers to the dendrogram catalog. Because parallax distances are more accurate than kinematic distances, this would be a way to refine the 3D locations of the molecular clouds identified in the dendrogram. This would improve our knowledge about the star forming regions in the Milky Way. More could also be learned by comparing to other data sets. The more data sets that are found to agree about the locations of star forming regions in the Galaxy, the more accurate a model of the Milky Way can be created. A long term goal for the future is to create a 3D model of the Galaxy with accurate locations of the spiral arms and regions of star formation.

REFERENCES

Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, PASP

Dame, T.M., Hartmann, D., & Thaddeus, P. 2001, ApJ

Georgelin, Y. M. & Georgelin, Y. P. 1976, A&A

Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2009, ApJ

Rice, T. S. 2012, "A Hierarchical Catalog of Molecular Clouds in the Milky Way"

Steiman-Cameron, T. Y. 2010, *Galaxies and Their Masks*, ed. D. L. Block et al.

This manuscript was prepared with the AAS LaTeX macros v5.2.