P53

IMPROVEMENTS TO CLUSTER IDENTIFCATION AND TRACKING IN A CIRCULATION DETECTION ALGORITHM

Madison Burnett[1], Travis Smith[2,3], Valliappa Lakshmanan[2,3]

[1]University of Missouri, Columbia, Missouri.
[2]Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma.
[3]NOAA/National Severe Storms Laboratory, Norman, Oklahoma.

1. Introduction

Both the TDA (Mitchell et al 1998) and MDA (Stumpf et al 1998) were developed and used with WSR-88D data beginning in the early 1990s to help forecasters learn to recognize signatures of circulations far enough in advance to issue tornado warnings with ample lead times. With the development of new radar technologies, it has become necessary to develop a new rotation detection algorithm for forecasters. This study makes improvements to the algorithm now in development and analyzes slightly altered versions of this algorithm to determine which performs best in a set of case studies. The Kmeans (Lakshmanan et al 2003) algorithm is examined to determine the best adaptable parameter settings for it to work with azimuthal shear data, rather than reflectivity fields or satellite IR temperatures that previous studies have used.

The algorithm identifies and tracks circulation “clusters” in maximum azimuthal shear fields. The last several positions of these circulation clusters through time are connected and shown as the “past tracks”. While these tracks should be predominantly linear if the algorithm accurately monitors and tracks these circulations, the tracks initially appeared erratic. This proved especially true at the smallest size scale, which study deals with exclusively. The number of clusters identified can also be problematic. Identifying too many clusters increases the false alarm rate, while too few could led to undetected mesocyclone-scale and tornado-scale circulations.

Three different size scales are used to identify and track both smaller tornado-scale and larger mesocyclone-scale circulations. The three main parameters of the algorithm initially altered in hopes of improving the algorithm were the filter types, size thresholds, and data ranges. Once parts of the algorithm were more closely examined through this process, it became apparent that other changes could be made as well.

This work attempts to make the number of circulations identified more accurate and improve the linearity of the cluster past tracks by altering the filter types, size thresholds, data ranges and by changing the tracking algorithm itself. The algorithm is tested on 11 different tornadic case studies (Table 1). Once improvements are made, different versions of the algorithm are compared in three different categories using a statistical program called w2scoretrack. The goal of this project is to develop a set of adaptable parameters for the algorithm that accurately identify and track the circulations.

2. Method

a. Creating the maximum azimuthal shear fields

Due to the noisiness of the velocity field, a more organized field was desired for the circulation tracking and identification in this algorithm. By taking the least linear squares derivative (LLSD) (Smith and Elmore 2004) of the velocity field, an azimuthal shear field was created. This field is smoother and more ideal for tracking. The field is then converted from polar to Cartesian coordinates. Finally the maximum azimuthal shear values in the elevation scans between layers (0-3 km in this study) are incorporated into a two-dimensional field known as the maximum azimuthal shear field. Clusters are identified on the azimuthal shear field using an enhanced watershed transformation (Lakshmanan et al 2009).

b. Improving the algorithm

The algorithm was evaluated using case studies in the WDSS-II graphical user interface. The 11 cases chosen were tornadic, used WSR-88D "super-resolution" data, and occurred between May 2007 and February 2009. Several hours of radar data from each case study is used.

First, the algorithm was processed on the Greensburg, Kansas tornado case from the Dodge City WSR-88D (KDDC). This case was chosen randomly from the supplied case studies and while it is an example of a larger, extremely strong circulation, it also contained many smaller, less intense circulations.

The algorithm's performance was evaluated and several changes were made in hopes of improving its cluster identification and tracking. The first of these changes was a time/space correction on the azimuthal shear field that takes storm motion into account. This helped to eliminate "doubles" or "shadows" that occur when the same circulation is shown in the data at two different locations, and is an artifact of storm motion. These appeared in the field when not all of the elevation scans in the specified layer taken into account in a time-step were updated. Before this change, the lowest scans received new data while the highest scans were still using the older data, producing a "shadow" or "double" in the maximum azimuthal shear field.

Next, a 30° conical search radius was incorporated into the algorithm to help provide a better first guess approximation as to where clusters would be located in future time steps. Previously, the algorithm just looked to the nearest cluster without accounting for storm motion. The cluster identification was also made a function of not only the distance between clusters, but also a function of cluster peak values and sizes. The hope of this was that cluster identifications would remain assigned to the same circulations from time step to time step since the cluster's size and peak values would not fluctuate much between these steps.

A nearest neighbor method was then implemented when converting the data to Cartesian coordinates. A weighted combination method using the nearest range-gate to each grid point center was used instead of the maximum value within a radius from the grid point center. The filter, size thresholds, and data ranges were also changed from the initial version to the improved version. The algorithm went from using a median filter initially, to an erosion/dilation filter. The filter takes every pixel touching a background pixel and changes it into a background pixel. Then the filter dilates, or makes every background pixel that is touching an object pixel into an object pixel. The rest of these parameter changes can be seen in Table 2.

Once all these changes had been made, the case was processed on the 'New Hampshire tornado' case from the Boston, Massachusetts WSR-88D (KBOX) on 24 July 2008. Since this case had much lower maximum azimuthal shear values and was smaller in size, it was chosen to test the changes made based on the Greensburg case.

c. Creating and scoring different tracking and identification methods

After it was determined that the algorithm performed well qualitatively in both the Greensburg and New Hampshire cases, this improved algorithm was processed on all 11 case studies. Afterwards, small changes were made to this improved algorithm's filter and thresholds to create four additional, slightly altered versions of the algorithm. The performance of each of the five total versions was then scored in three categories using a technique developed by Lakshmanan et. al (2009). The parameters used for these versions are shown in Table 3.

To provide temporal continuity, only clusters lasting longer than 300 seconds (longer than one full volume scan) are used from each of the 11 cases. Since the tracks eliminated by this condition are very short, they have very little error and bias the overall results in each case. By eliminating them, values in every category increased.

The versions were scored on median duration, position error, and mismatch error. Median duration per cluster represents the median lifetime per cluster measured in seconds. Larger values in this category would indicate more continuous, long-lived circulations as opposed to short-lived circulations that oscillate above and below the data or size thresholds.

Position error measures the position error in relation to a straight line between cluster locations over time. It essentially measures how linear the tracks remain through time. Small values in this category would be indicative of somewhat linear paths and few cluster mismatch errors over time. Mismatch errors refer to the amount of maximum azimuthal shear value fluctuation per cluster over time. When mismatches occur, the maximum azimuthal shear value for a specific cluster identification will change when the particular identification number is assigned to a cluster with different characteristics. The fluctuation in these values over time can be a tell-tale sign of a cluster mismatch. Since mismatches cause higher values, a good algorithm would want to have smaller values in this field.

3. Results

While only qualitative comparisons are made between the pre-study version of the algorithm and the post-study version, it appears the changes made in this study significantly improve the algorithm. Figure 3 helps to illustrate this fact. This particular time step helps to illustrate how the path(s) have become more linear, mismatches have been reduced, and it appears that the more intense clusters are being properly identified. After qualitatively reviewing the performance of both the pre-study version of the algorithm and the post-study version, it does appear that there are significant improvements.

When examining the performances of the four additional, slightly altered versions of the post-study algorithm with the scoring algorithm, the results are much more quantitative. Figures 4, 5, and 6 indicate a great deal of overlap between the different versions, especially in terms of confidence intervals. In terms of position error (or linearity error), all five scored versions seem to have a median value very near 1 km and their confidence intervals almost completely overlap. An ideal version of the algorithm would have very low values in this category. Thus, in terms of position error there is no clear “best” version.

Figure 5 shows the relationship performance of the different versions in terms of cluster durations. Unlike the other two statistically evaluated categories, ideal duration values would be rather high. This would indicate that clusters are being accurately tracked throughout time. The SizeIncrease version does appear to have a longer median duration, but this is probably because since it increases the minimum size threshold, it also helps to filter out some of the shorter-lived clusters. This can be seen in Figure 7, which shows how many clusters are used by each of the versions. It can be seen that the SizeIncrease version uses many fewer clusters that some of the other methods. Since this is the case, it is difficult to count this as a “win” for the SizeIncrease method. Strong circulations of smaller sizes could slip through without being detected using this particular version.

The numbers of cluster identification mismatches are shown in Figure 6. While there is still a larger overlap between confidence intervals, the Median0.05 version does appear to have a lower median value that the other versions.

Taking into account all three categories evaluated by the statistical software, it does not appear that there is a definite “best” method for circulation detection and tracking. There was a great deal of confidence interval overlap between all the methods and the only category with a possibly significant performance was the Median0.05’s performance in terms of a lower mismatch error.

In an attempt to gain some kind of quantitative results comparing the pre-study algorithm to the post-study algorithm, the statistical software was processed on the Greensburg, Kansas once using the pre-study version of the algorithm and once with the post-study version of the algorithm. The results of this comparison are shown in Figures 8, 9, and 10.

As seen in Figure 8, the median linearity/position error is roughly 0.5 km lower in the post-study version of the algorithm than in the pre-study version. The confidence intervals still overlap to a certain degree, but these results are encouraging. In terms of cluster duration (Figure 9), the median post-study version is actually reduced. The reasons for this are unknown and it should be noted that the confidence intervals in this particular category are incredibly large and overlapping. Median mismatch error (Figure 10) appears to have stayed nearly the same between the two versions, although it the confidence interval has become narrower.

4. Conclusion

Since the scoring categories can work against each other in some cases, it is hard to score well in all three categories. As an example, an older cluster identification that becomes associated with a new cluster would have increased duration, but also increased mismatch error. The “best” method of cluster tracking and identification could not score the worst in any category and would ideally rank highly in all three.

The statistical significance of the results in all three categories is uncertain because of larger confidence interval overlap. The position error comparison is very inconclusive due to similar median values and largely overlapping confidence intervals. The SizeIncrease method appears to have the longest duration, but since it tracks fewer clusters than most of the other methods, there is doubt to whether it would identify smaller, possibly tornadic circulations. The Median_0.005 method has a lower median value than the other methods in mismatch error, but the confidence intervals still overlap greatly. This overlap makes it very difficult to identify a best method for tracking and identification of circulation clusters.

There is much more confidence in the qualitative improvement from the initial version to the improved version of the algorithm processed on the Greensburg case. Track linearity appears better as well as accurate cluster identifications. The quantitative results of the scoring algorithm suggest support to these observations. Linearity error lessens in the improved version while mismatch error seems to have decreased slightly as well. Duration also decreased, most likely due to a decreased minimum size threshold in the improved version.

While nothing concrete arose from the pre-study version versus post-study version of the algorithm in terms of statistical categories based on the 05-06 May 2007 case, it did show possible signs of improvement in linearity error. Had this analysis been processed on all 11 cases, the results would be more substantial.

Overall there was no best method for cluster identification and tracking determined from this study, but all methods performed well. There was significant qualitative improvement from the algorithm's initial version and some suggested quantitative improvement as well.

5. Acknowledgements

This research was funded and performed under an appointment to the National Oceanic and Atmospheric Administration, Ernest F. Hollings Undergraduate Scholarship Program. We would also like to thank Kiel Ortega for case study preparation and advice.

6. References

Lakshmanan, Valliappa, Brett Roberts, Travis Smith, 2009: An objective method of evaluating and devising storm tracking algorithms, *Wea. Forecasting*, In press.

V. Lakshmanan, K. Hondl, and R. Rabin, ``An efficient, general-purpose technique for identifying storm cells in geospatial images,'' *J. Ocean. Atmos.*

Tech., vol. 26, no. 3, pp. 523-37, 2009.

V. Lakshmanan, R. Rabin, and V. DeBrunner, ``Multiscale storm identification and forecast,'' *J. Atm. Res.*, vol. 67, pp. 367-380, July 2003.

Mitchell, E.D., S.V. Vasiloff, G.J. Stumpf, A. Witt, M.D. Eilts, J.T. Johnson, and K.W. Thomas, 1998: The National Severe Storms Laboratory Tornado Detection Algorithm. *Wea. Forecasting*, **13**, 352–366.

Smith, T.M. and K. L. Elmore, 2004: The use of radial velocity derivatives to diagnose rotation and divergence. Preprints, 11th Conf. on Aviation, Range, and Aerospace, Hyannis, MA, Amer. Meteor. Soc., P5.6 - CD preprints.

Stumpf, G.J., A. Witt, E.D. Mitchell, P.L. Spencer, J.T. Johnson, M.D. Eilts, K.W. Thomas, and D.W. Burgess, 1998: The National Severe Storms Laboratory Mesocyclone Detection Algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 304–326.

Figures and Tables

Case Study Date	WSR-88D Radar	Description
04-05 May 2007	KDDC	Greensburg, KS EF-5
18-19 August 2007	KTLX	Tornadic cells associated with Hurricane Erin
07 January 2008	KMKX	Few tornadic supercells
05-06 February 2008	KNQA	Super Tuesday Outbreak
07-08 April 2008	KTLX	Supercells with no tornado reports
22 May 2008	KFTG	Windsor, CO EF-3
22-23 May 2008	KDDC	Several strong supercell tornadoes
25-26 May 2008	KDVN	Parkersburg, IA EF-5
12-13 June 2008	KICT	Few supercell tornadoes
24 July 2008	KGYX	New Hampshire tornado
10-11 February 2009	KTLX	Supercell tornadoes

Table 1: The dates, radars, and brief descriptions of the 11 cases used for this study are shown above.

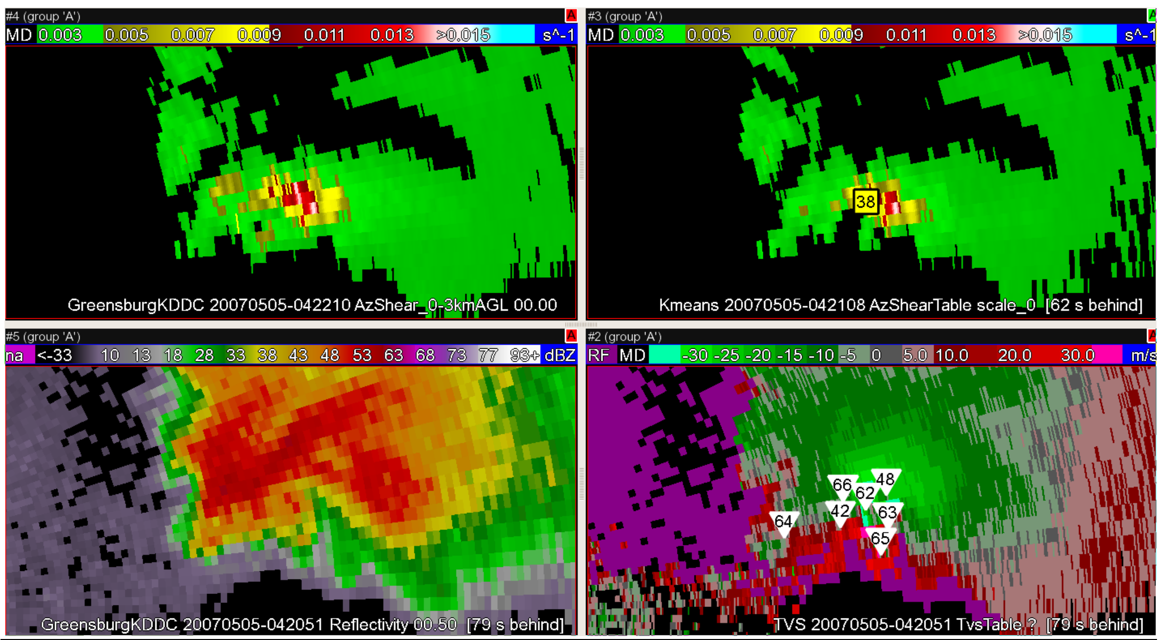

Figure 1: (Bottom left): The reflectivity image associated with the supercell that spawned the Greensburg, Kansas tornado on 06 May 2007 at 0420Z. (Bottom right): The white triangles denoted Tornado Vortex Signatures that have been registered in the velocity data associated with the reflectivity field in the bottom left. These TVS signatures were used to detect and track circulations in the Tornado Detection Algorithm. Note that seven circulations are detected in this velocity couplet alone. These signatures often register in areas of clear air as well. This example also shows how much range folding and general messiness can appear velocity fields. (Top left): The maximum azimuthal shear below 3 km field associated with the reflectivity image in the bottom left. Note how the field looks cleaner and there is a more defined area of rotation (located in the red colors). (Top right): The algorithm has been applied to the field shown in the top left. The yellow box designates the identified cluster. This algorithm identifies one well-defined circulation as opposed to the seven circulations registered by the TDA in the velocity field.

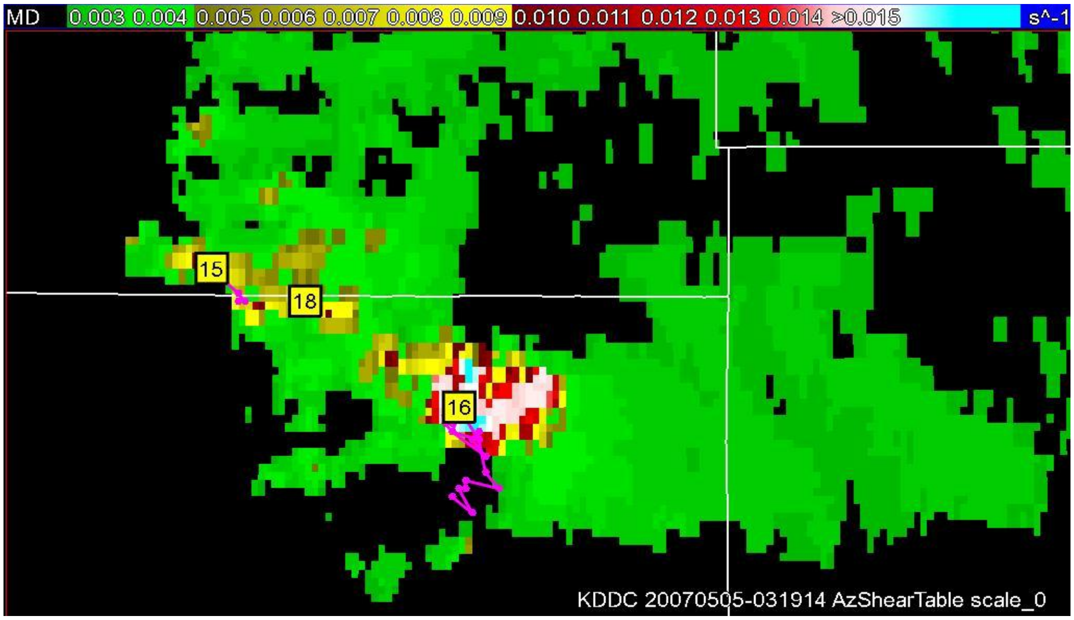

Figure 2: The above image is the maximum azimuthal shear field associated with the supercell that spawned the Greensburg, Kansas tornado of 05-06 May 2007. This particular time step is at 0319Z on 06 May 2007. The yellow, red, and pink pixels indicate areas of high azimuthal shear values in the 0-3 km layer. The yellow numbered boxes are identified circulation clusters and the pink lines are the tracks that these clusters have taken through time. Ideally these lines would be somewhat linear.

Algorithm Version	**Filter**	**Minimum Data Threshold**	**Maximum Data Threshold**	**Minimum Size Threshold**
Initial (pre-study)	5 x 5 pixel median	0.005 s^{-1}	0.02 s^{-1}	30 pixels
Improved (post-study)	3 x 3 pixel erosion/dilation	0.006 s^{-1}	0.02 s^{-1}	25 pixels

Table 2: The parameters used in this study are shown above for both the pre-study version of the algorithm and the improved version. The values of these parameters apply only to the smallest size scale, which was the focus of this study.

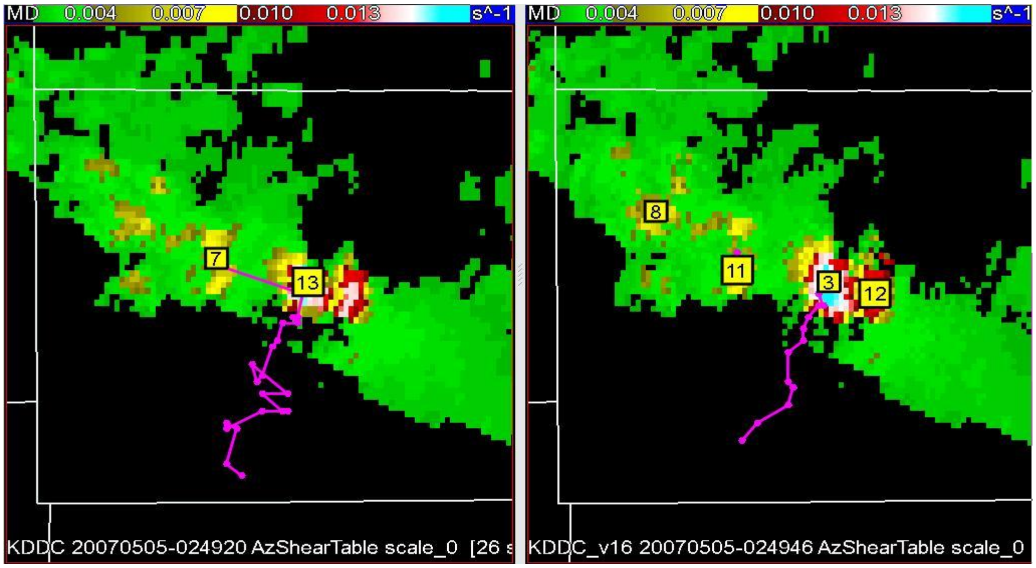

Figure 3: (Left): This image shows the pre-study version of the algorithm processed on the maximum azimuthal shear below 3 km field associated with the supercell that spawned the Greensburg, Kansas tornado on 05-06 May 2007. This particular time step is at 0249Z on 06 May 2007 and shows two identified circulation clusters, labeled 7 and 13. (Right): This image shows the closest possible time step (26 seconds behind the image on the left) using the post-study version of the algorithm on the improved background fields. The four yellow boxes numbered 8, 11, 3, and 12 refer to the identified circulation clusters.

	Filter	**Minimum Data Threshold**	**Minimum Size Threshold**
Improved	3 x 3 pixel erosion/dilation filter	0.006 s^{-1}	25 pixels (~ 6 km^2)
5 x 5	5 x 5 pixel erosion/dilation filter	0.006 s^{-1}	25 pixels (~ 6 km^2)
Median	3 x 3 pixel median filter	0.006 s^{-1}	25 pixels (~ 6 km^2)
Median 0.005	3 x 3 pixel median filter	0.005 s^{-1}	25 pixels (~ 6 km^2)
Size Increase	3 x 3 pixel erosion/dilation filter	0.006 s^{-1}	40 pixels (~ 10 km^2)

Table 3: The parameters of the improved version of the algorithm (grey row) and of the fourth slightly altered algorithms are shown above (white rows).

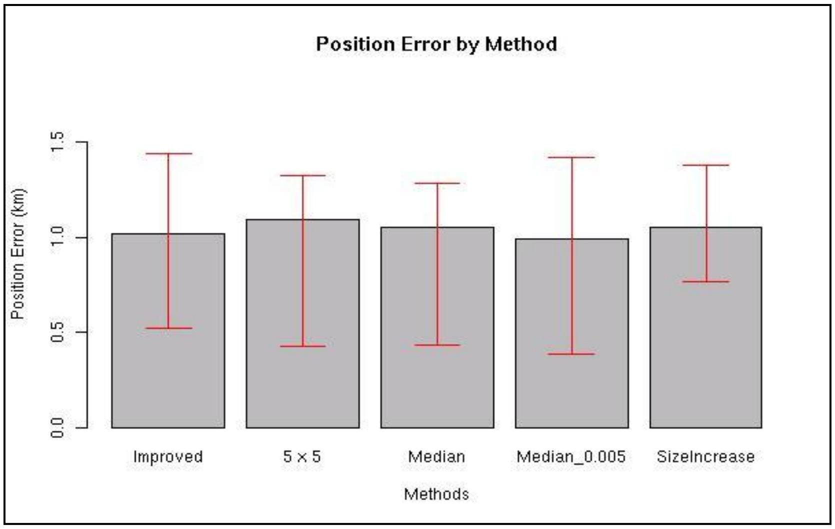

Figure 4: The above graph shows how the performance of the original post-study improved version of the algorithm with the four slightly altered versions of the improved version in terms of cluster position error. The grey columns show the median values in each category and the red whiskers enclose the 25th and 75th percentiles. The data is the collected from the algorithms' performances in all 11 cases.

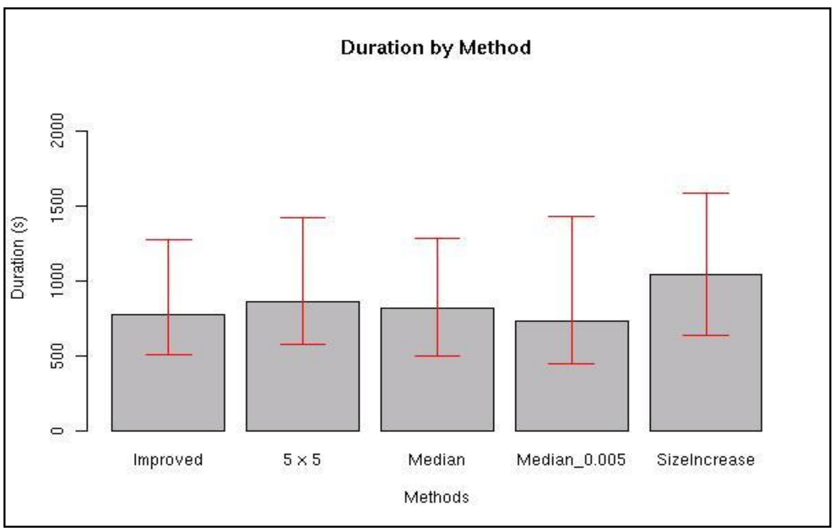

Figure 5: The above graph shows how the performance of the original post-study improved version of the algorithm with the four slightly altered versions of the improved version in terms of cluster duration. The grey columns show the median values in each category and the red whiskers enclose the 25th and 75th percentiles. The data is the collected from the algorithms' performances in all 11 cases.

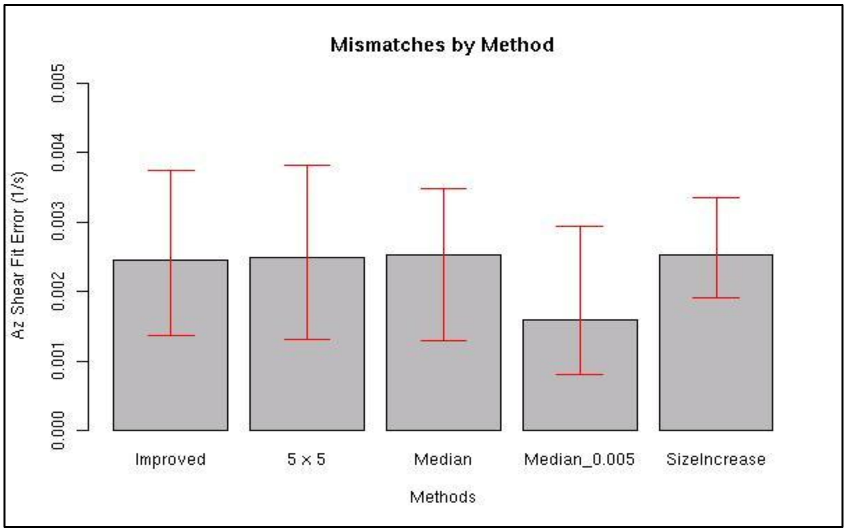

Figure 6: The above graph shows how the performance of the original post-study improved version of the algorithm with the four slightly altered versions of the improved version in terms of mismatch error. The grey columns show the median values in each category and the red whiskers enclose the 25th and 75th percentiles. The data is the collected from the algorithms' performances in all 11 cases.

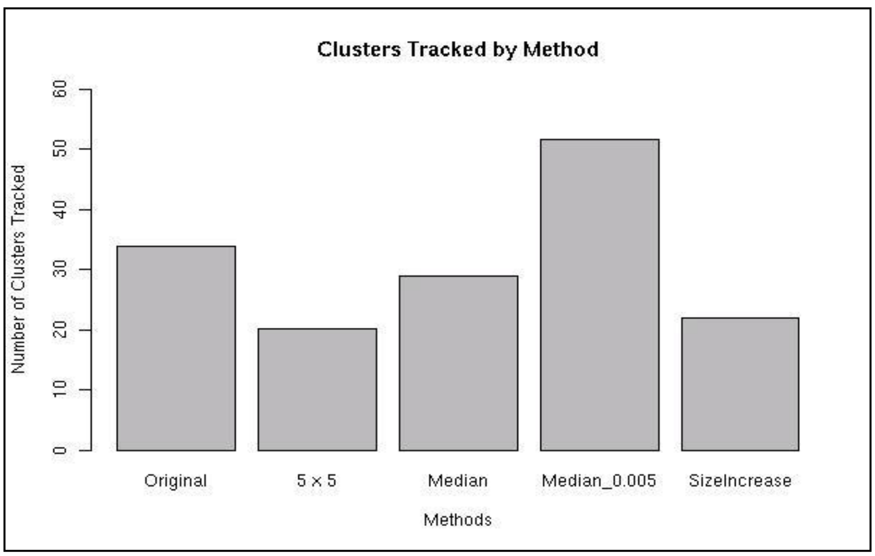

Figure 7: The above graph shows the number of clusters track by each method, or version per case.

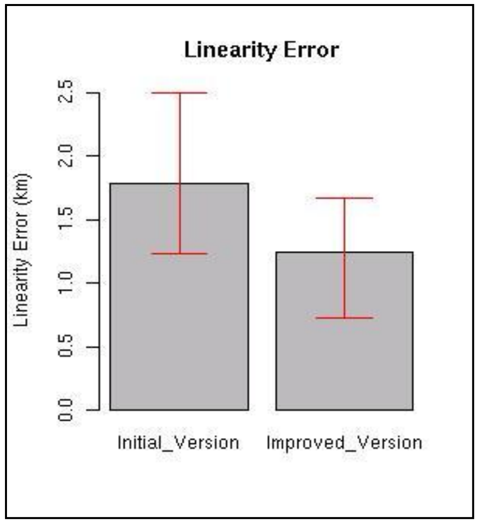

Figure 8: The graph above shows the relationship in terms of linearity/position error between the initial and improved versions of the algorithm in the 05-06 May 2007 case. The grey bars indicate the median values and the red brackets indicate 25th and 75th percentiles.

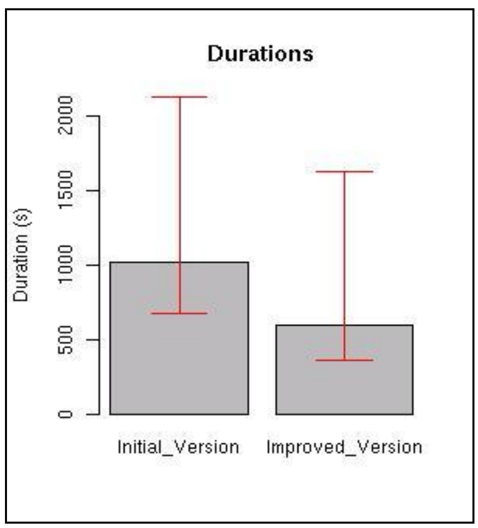

Figure 9: The graph above shows the relationship in terms of duration between the initial and improved versions of the algorithm in the 05-06 May 2007 case. The grey bars indicate the median values and the red brackets indicate 25th and 75th percentiles.

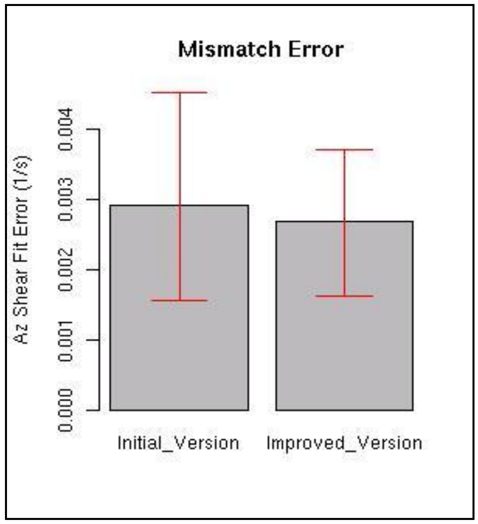

Figure 10: The graph above shows the relationship in terms of mismatch error between the initial and improved versions of the algorithm in the 05-06 May 2007 case. The grey bars indicate the median values and the red brackets indicate 25th and 75th percentiles.