P1.3

USING A REFLECTIVITY "SCALE FILTER"
TO IMPROVE PERFORMANCE OF THE
STORM CELL IDENTIFICATION AND TRACKING (SCIT) ALGORITHM

Gregory J. Stumpf[1,2,*], Walter D. Zittel[3], Arthur Witt[2], Travis M. Smith[1,2],
Valerie K. McCoy[1,2], Shannon Dulin[1,2]

[1]Cooperative Institute for Mesoscale Meteorology Studies, University Of Oklahoma,
Norman, Oklahoma.
[2]NOAA/National Severe Storms Laboratory, Norman, Oklahoma.
[3]NOAA/NEXRAD Radar Operations Center, Norman, Oklahoma.

1. Introduction

The National Severe Storms Laboratory (NSSL) Storm Identification and Tracking Algorithm (SCIT; Johnson et al 1998) was implemented into the WSR-88D system and has been in operational use for about six years. During this time, it has become apparent that the algorithm faces shortcomings in its time association abilities. Most of these problems can be traced back into the horizontal and vertical association techniques in the algorithm, and are basically inherent in any algorithm which uses heuristic rules and centroid to centroid associations. A study has shown that given approximately 100 storm events ranging from 10-60 minute lifetimes, and from 5 locations nationwide, the present time association technique did not work about 25% of the time (Witt and McCot 2002). In other words, for a given storm event, for one out of every four volume scans, the storm ID is reassigned and all time-trends are restarted.

The NSSL partnered with the NEXRAD Radar Operations Center (ROC) to study various methods to improve SCIT time association. The first attempt to improve the algorithm included modifying source code to some of the heuristic rules in the vertical association technique (the "vertical-merge technique"; Witt and McCoy 2002). This improved the time association success rate to 84%. However, a less-costly alternative is subsequently tested and is reported here.

Many WSR-88D algorithms, including the SCIT algorithm, operate using native polar radar data. For the reflectivity moment, the resolution is 1° in azimuth and 1 km in range. The spatial resolution of polar data decreases with increasing range due to broadening of the beam. At near ranges, single storm cells can be represented by many reflectivity peaks, each being potential candidates for storm cell detection (Fig. 1). At far ranges, this problem is reduced.

We developed an initial "scale filter" technique that is designed to smooth out the fine-resolution reflectivity peaks at close ranges to the radar while retaining the peaks at farther ranges. The purpose of the filter is to reduce the number of reflectivity peaks in storms at near-

1

range. Figure 2 shows an example of the same data in Figure 1 with the scale filter applied. The single supercell storm nearest the radar was incorrectly characterized by 4 separate cell detections prior to the filter technique. The filtering removed the additional reflectivity peaks and resulted in only one storm cell detection.

Figure 1. Original (unfiltered) reflectivity field. Overlaid are SCIT cell detections based on using the unfiltered field.
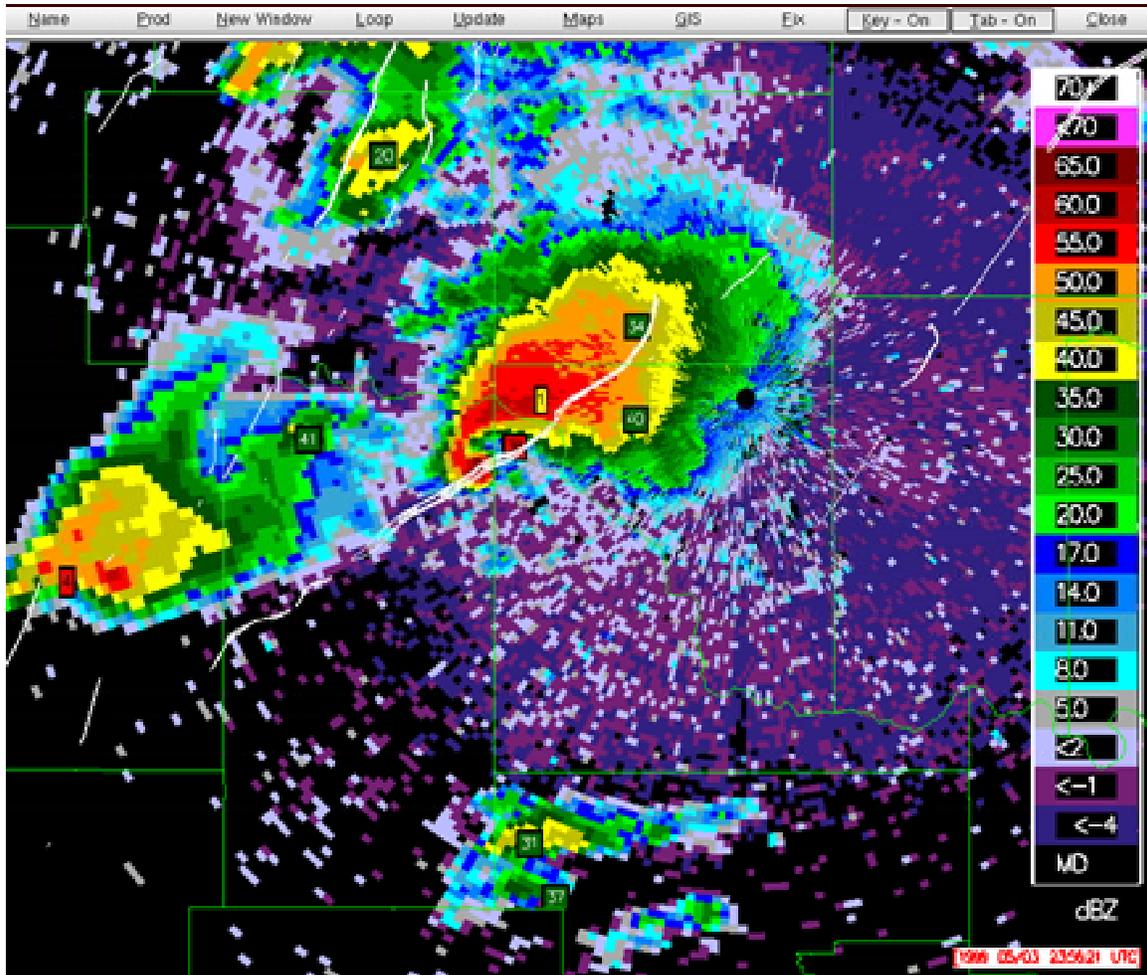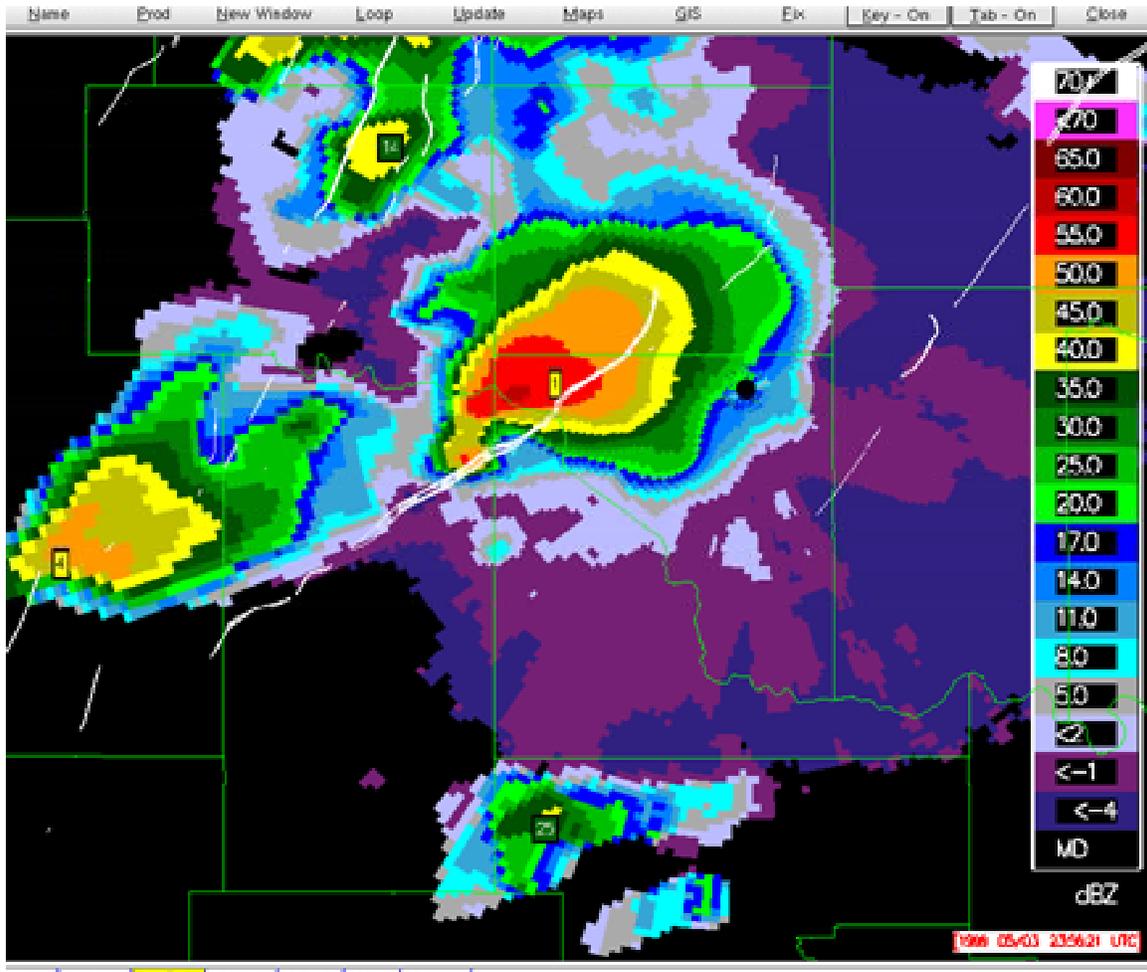
Figure 2. Reflectivity field with the scale filtering applied. Overlaid are SCIT cell detections based on using the scale filtered field.

One limitation of the filtering process is that reflectivity amplitudes are damped, especially at near ranges. This has two undesired effects. First, small and isolated (but valid) cells can be "filtered away", such that their size and intensity no longer meet the thresholds used to define cell detections. The second effect causes diagnosed strength attributes of the cell detections such as maximum reflectivity and Vertically Integrated Liquid (VIL) to be reduced.

2. Data

The same data sets used to test the "vertical-merge technique" are used to test the filter technique. Storms cells were identified from 5 different events

collected nationwide. These events represent a variety of storm types, including supercells, multi-cell clusters with light steering winds, hurricane outer rain bands, and squall lines, and are listed in Table 1. The data were derived from Level-II base radar data, and processed using the NSSL Warning Decision Support System – Integrated Information (WDSSII; Hondl 2003).

Table 1. List of storm events, locations, dates, and number of cells used in the analyses.

| Site	Radar ID	Storm Type	Date	# Volume Scans	# of Cells
Norman, OK	KTLX	supercells	050399	29	4
Eglin AFB	KEVX	hurricane bands	100495	15	15
Phoenix, AZ	KIWA	multicell	081596	18	31
St. Louis, MO	KLSX	squall line	041594	16	35
Detroit, MI	KDTX	supercells	062296	20	10
TOTAL				**98**	**95**

3.   Method

The technique uses a simple range-adjusting kernel such that more data points are smoothed at closer ranges to the radar. The kernel is centered on each data bin and its size is held "constant", so the number of azimuth bins within the kernel increases with decreasing range. The number of range bins within the kernel always remains constant with range. Figure 3 shows an example of a 7 km kernel at two different ranges. At the near range (on the right), 7 km represents an azimuthal extent of 24 degrees (24 radials), and thus 168 data values are used to compute the filtered value at the kernel's center. At the farther range (on the left), 7km represents an azimuthal extent of only 8 radials, and only 56 data values are used to compute the filtered value. Data values contained within each kernel are then sorted, and the value representing some point along the distribution of values in the bin is used to represent the new data value representing the kernel's center bin. Both the kernel size (e.g., 5km, 7km, etc.) and distribution percentage [i.e., 50% (median), 80%, etc.] are adaptable. A functional description of the source code is included in Appendix A.

The first task was to optimize the filter parameters that give best time association with least impact to number of detections. With the optimal set of parameters, the next step is to mitigate any reflectivity amplitude damping caused by the filtering technique. The first mitigation technique involves the development of a range-dependent bias correction applied to the filtered data. If that technique proves unsuccessful, then a second mitigation technique will be tested. That technique involves using the filtered data for cell detection, and the original unfiltered data for cell diagnosis. These techniques will all be compared to the "vertical merge" technique developed and tested by Witt and McCoy (2002).

4

Figure 3. Filter kernels (white boxes with cross-hairs) for near range (on right) and far range (on left). Comparison of two kernels is provided in the text.
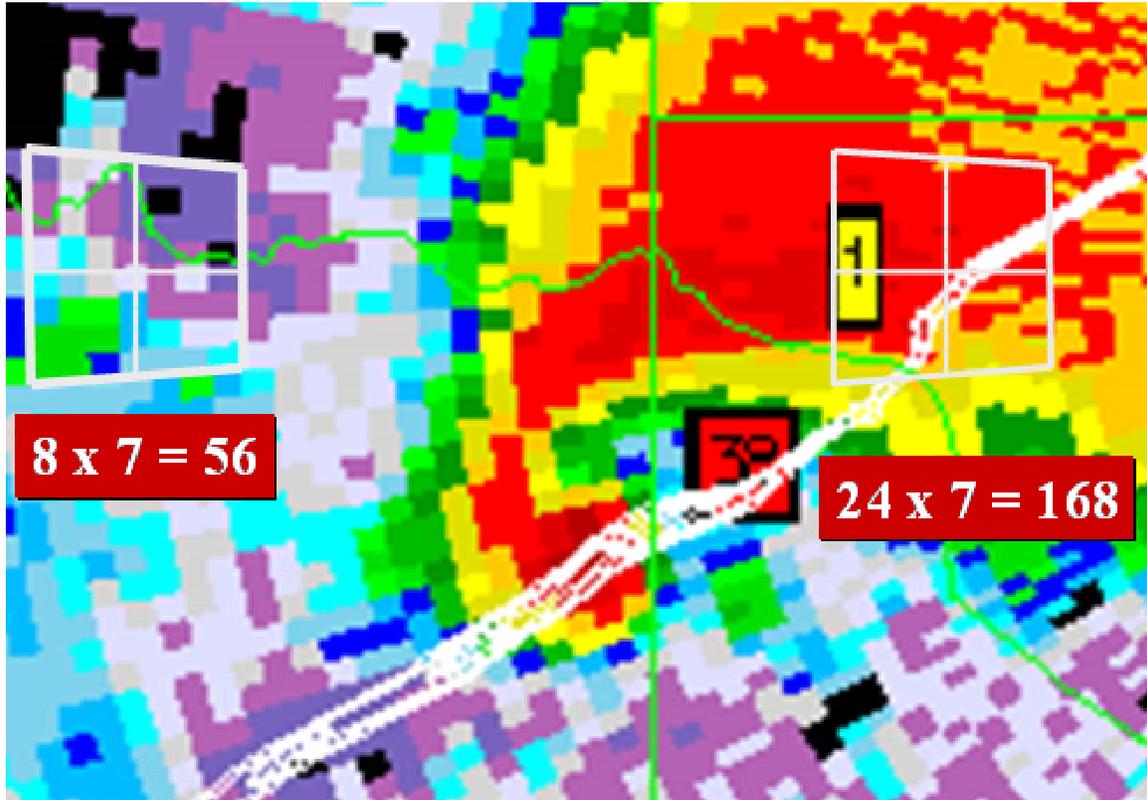
4. Results

a. *Optimization of filter parameters*

The first task was to run the scale filter with a nine different combinations of filter parameters to determine the optimal parameter that maximizes time association success rate

POD = [# TA success / (#TA success + #TA failures)]   (1)

A time association success was defined as a correct time association between matched cell detections on two consecutive volume scans (the cell ID remains unchanged across the two volume scans). A time association failure was defined as a missed time association between matched detections on two consecutive volume scans (the cell ID changes across two volume scans). Correct cell matching was determined using human-truthed cell matching across volume scans. Where there were gaps in the radar data that were larger than one volume scan (> 6 minutes), or when storm cells were not detected, time association could not be preformed and thus was not expected. Thus, these gaps are not considered as time association failures.

The nine parameter combinations included following values: Kernel size of 5km, 7km and 9km; distribution percentage of 50% (median filter), 70%, and 90%. The results of the nine runs, plus results using no filters and the

results using the vertical merge technique are shown in Table 2.

The various filter parameters made little variation in the time association success rate, as shown in the table below. The 70% distribution and 7 km kernel size had the greatest success. In every case, the success rate continued to greatly outperform both the baseline SCIT (no filter) and the vertical merge technique developed and tested by Witt and McCoy (2002).

Table 2. Time association success rates (Eq. 1) for the nine combinations of scale filter parameters, the baseline SCIT (unfiltered data) and the vertical merge technique.

Distribution (%)	50	50	50	70	70	70	90	90	90	*Baseline*	*Vertical*
Kernel Size (km)	5	7	9	5	7	9	5	7	9	*(no filter)*	*Merge*
TA Success Rate	93.9	96	95.4	95.3	97.1	95.9	95.4	96.6	93.2	77.2	83.9

b. *Reduction in small and isolated cell detections due to amplitude damping*

An unwanted side-effect of the amplitude damping is the reduction of small and isolated (but valid) cells which can be "filtered away", such that their size and intensity no longer meet the thresholds used to define cell detections. Originally, we were to determine the number of cells deleted by this effect. Multiple detections of the same cell are intended to be removed by this process, so they were not included in the statistics.

The scope of the analysis was broadened since more analysis beside just weak and small cells was needed. Differences in the kernel size result in minor differences in the number of detected cells. Besides looking at just small and isolated cells, all "distinct" cells were included, even if there was some "contact" with another cell. Additionally, many new detections that occurred as a result of the various filters, including false detections (caused mainly by residual ground clutter), were considered.

For each of the 9 filters (in each section), the number of cells (either deleted or added, relative to the no-filter run), the average maximum reflectivity and VIL for all the cells, and the number of cells having non-zero Probability of Severe Hail (POSH; Witt et al 1998) values are shown in Table 3. The results indicate that the distribution percentage value is more dominant than the kernel size. And, as would be expected, the lower the distribution percentage value (corresponding to more smoothing), the greater the number of cell deletions. Increasing kernel size also leads to more cell deletions. The reverse generally applies to the new detections, valid or false, although there's less of an effect from the kernel size.

As for the character of the cells being added or deleted, there's not much difference in the average maximum

6

reflectivity and VIL for the deletions, whereas for valid additions, the 70% and 90% distribution percentage filters had lower values than the 50% filters, due to detecting more very weak cells.

To get a measure of whether or not any of these cells were really significant, we kept track of the number of cells with non-zero POSH. There were several of these among the deletions, but none among the valid additions. Although some of these cells with non-zero POSH were relatively small in size, others were quite large, including all the non-zero POSH cells. For these large cells that were not detected, it is likely that the smoothing of the reflectivity field resulted in 2D components whose centroids were outside of the search radius for 3D association.

Lastly, concerning the false detections, these came from three sources: residual ground clutter (all from the KIWA case), 2nd-trip high-dBZ ring (on the KLSX 4/15/94 case), and a few instances of reflectivity peaks in the anvil of an MCS.

We recommend avoiding using any of the 90% filters, due to the large number of added false detections. The 70% filters offer a "middle-of-the-road" approach. The 50% filters have the advantage of little or no false detections, and the additional valid detections don't include very weak cells (i.e., cells that are not operationally relevant).

Therefore, based on these findings, we recommended using the 50% distribution percentage and 7 km filter parameters.

Table 3.   For each of the 9 filter parameter combinations, the number of cells deleted (top), the number of valid cells added (middle) and the number of false cells added (bottom). Included are the average reflectivity (dBZ) and VIL, and the number of cells meeting non-zero POSH thresholds.

Cells Deleted					
	# of	*Average*	*Average*	*POSH*	*POSH*
Filter	*cells*	*dBZ*	*VIL*	*10-40%*	*50+%*
50-5	120	42.2	5.2	2	1
50-7	140	41.6	4	0	0
50-9	179	41.7	4.3	1	0
70-5	30	41.8	4.6	0	0
70-7	54	42.2	4.9	0	0
70-9	78	41.6	4.4	0	0
90-5	15	41	3.6	0	0
90-7	24	43.2	5.8	0	0
90-9	37	43.2	5.9	0	0

Valid Cells Added

Filter	# of cells	Average dBZ	Average VIL	POSH 10-40%	POSH 50+%
50-5	10	41.8	5.8	0	0
50-7	35	43.4	7.1	0	0
50-9	30	44.5	8.1	0	0
70-5	48	38.1	2.4	0	0
70-7	36	39.6	3.1	0	0
70-9	61	40.8	4.4	0	0
90-5	185	38.6	2.6	0	0
90-7	158	39	2.8	0	0
90-9	141	39	2.8	0	0

False Cells Added

Filter	# of cells	Average dBZ	Average VIL	POSH 10-40%	POSH 50+%
50-5	1	38	1	0	0
50-7	0				
50-9	0				
70-5	8	41.8	1.1	0	0
70-7	5	42.6	1.2	0	0
70-9	3	41	1.3	0	0
90-5	51	46.3	1.4	1	0
90-7	54	45.9	1.2	1	0
90-9	46	45.3	1.1	0	0

c. *Range-dependent bias correction to adjust for amplitude damping*

To address the second effect of the amplitude damping, a range-dependent bias correction

$$\text{Bias} = \text{value of unfiltered data} - \text{value of filtered data} \qquad (2)$$

was calculated and tested to recover the damped values. We chose to use a bias correction versus a ratio percent correction, because dBZ is measured on an open-ended log scale (- infinity to + infinity). In order to use a percent correction, dBZ would have to be converted back to Z, the ratios computed and corrected, and Z converted back to dBZ.

Instead of doing filtered versus non-filtered reflectivity comparisons using 2D features as originally planned, we chose to use the actual polar gridded data for each elevation angle. The latter data are much less sparse, and should give better range-dependent statistics on

the amplitude damping of reflectivity due to the filter process.

The entire 5 case data set was run and a statistical analysis comparing each sample volume from each elevation scan between the filtered (using the 50% and 7 km thresholds) and unfiltered data was performed. The biases (Eq. 2) of every sample volume whose original

reflectivity was 30 dBZ or greater were recorded. Figure 4 depicts the upper quartile (75th percentage) of the bias for each set of sample volumes divided into 5 km range bins (e.g., all samples from 0-5 km, 5-10 km, 10-15 km, etc.). As expected, the bias is always positive (unfiltered data have higher reflectivities than filtered data), and the bias decreases with increasing range.

Figure 4. Reflectivity bias correction as a function of range, as determined from the upper quartile (75th percentage; cyan curve) of the bias (unfiltered dBZ minus filtered dBZ) for 5 km range bins. Also shown are the median (50th percentage; purple) and the lower quartile (25th percentage; orange).
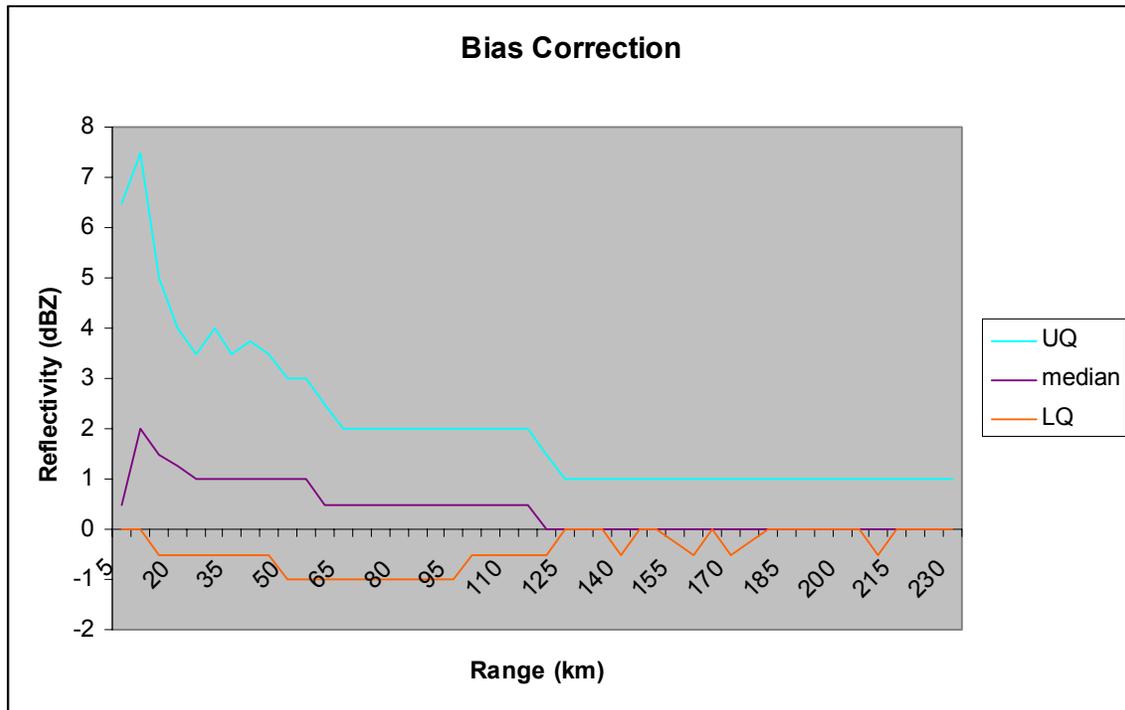
The upper quartile information (cyan curve in Figure 4) was used to develop a range-dependent bias correction function. The range-dependent bias curve was slightly smoothed and applied to the data filtered using the 50% and 7 km parameters. All 5 cases were run using the bias-corrected filtered data. The time association success rate (Eq. 1)

recomputed from the bias-corrected runs was 93.0%. This is a slight reduction from the non-corrected filtered data (96.0%).

The VIL trends for the bias-corrected runs were also tabulated. Shown in Figure 5 as the average VILs as a function of range (10 km range

increments) for all of the storm cells in the 5 cases combined. Included in Figure 5 are the values as determined from truth (human-interpreted cells, reflectivities, and calculated VILs), the baseline SCIT (unfiltered data), the vertical merge technique, the filtered data without bias correction, and the bias-corrected filtered data. In Figure 6, the percent difference (average ratios) from truth of the three runs (no filter, filter, and filter with bias correction) is also shown as a function of range. Note that the filter run without bias correction shows a smaller ratio at closer ranges. The average ratios for all ranges, all cases, and each of the three runs are shown in Table 4.

Figure 5. Average VIL as a function of Range (km) for "truth" data (magenta), baseline unfiltered data (yellow), vertical merge (blue), filtered data (cyan), and bias-corrected filtered data (purple).
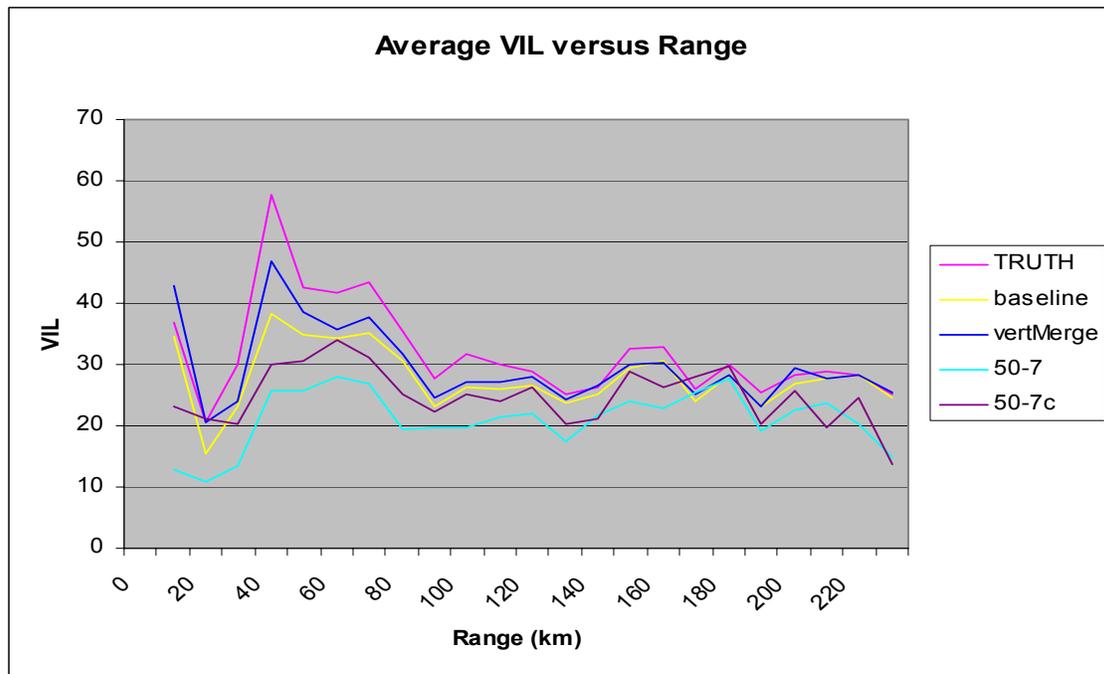
Figure 6. Percent difference (average ratios) from truth of the three runs (no filter, filter, and filter with bias correction) as a function of range. Baseline unfiltered data (yellow), vertical merge (blue), filtered data (cyan), and bias-corrected filtered data (purple).
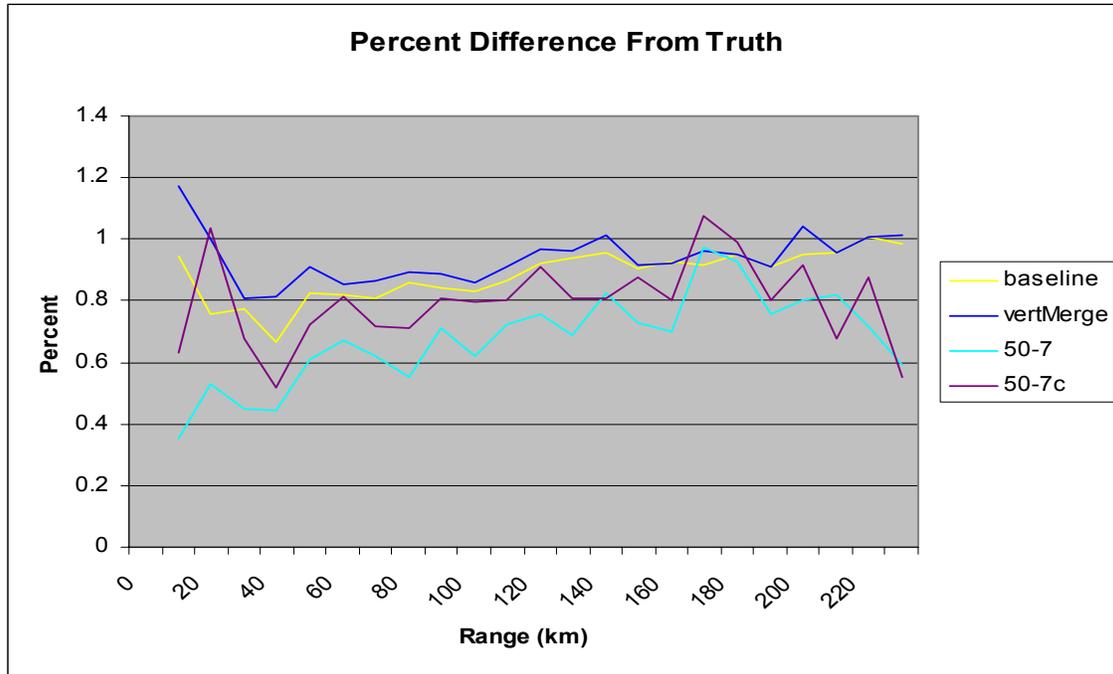

Table 4. Average ratios of VIL values between "truth" and each of 3 different SCIT runs (reflectivity with no filter, filter, and bias-corrected filter). Also, average ratios of VIL values between "baseline" and the 2 filtered SCIT runs

Reflectivity data	Average ratio from truth	Average ratio from baseline
Baseline (no filter)	.864	-
Vertical Merge	.908	1.049
Filter	.672	.777
Bias-Corrected Filter	.796	.920

Since the average ratios were only 79.6% of truthed values, and only 92.0% of baseline values, we recommend that the original (unfiltered) reflectivity data are used for cell diagnosis, and the filtered data (with the 50% and 7 km parameters) is only used for cell detection.

Note that the vertical merge technique developed shows the best average VIL ratio from truth and improves upon the baseline. Recall from Table 2 that the time association success rate for the vertical merge technique is inferior to the filter techniques. Implications of these results will be discussed later.

Since the reflectivity bias correction was not deemed acceptable, and in the interest of time, the reduction in the number of detected small and isolated

(but valid) cells caused by the reduction of the maximum dBZ below SCIT algorithm thresholds was not determined.

d.   Use of original (unfiltered) reflectivity for diagnosis

The SCIT code was modified such that the filtered data are used for the detection of cells, and the original unfiltered data is used for cell diagnosis. The method is to determine the azimuthal and range extent of each 2D SCIT feature (on each elevation scan), and using the data from the original (unfiltered) reflectivity field, determine the maximum three-gate running average reflectivity of all data points within the 2D feature extent. These maximum reflectivity values are then used to calculate VIL, SHI, and other storm diagnostic parameters.

Because cell diagnostic information is used in the vertical and time association schemes (cell components are sorted by strength prior to association), there should be some differences in the overall time associations. The time association success rate (Eq. 1), recomputed 94.1%. This is a slight reduction from the non-corrected filtered data (96.0%) and a slight improvement from the bias-corrected filtered data (93.0%). Overall, this TA success rate is still quite higher than the baseline and vertical merge rates (Table 2).

The VIL trends for the final run were tabulated. Figures 7 and 8 repeat Figures 5 and 6, respectively, with the exclusion of the filter and bias-corrected filter runs, and the inclusion of the VIL ratios for the final run. Figure 9 is included to show the VIL ratios as percent difference of all four options from the baseline SCIT as a function of range. Again, note the range differences on the filter case without the bias correction, and how the final option is an improvement two the first two filter options at almost every range. Also, Table 5 repeats Table 4 wit the addition of the final run statistics.

Figure 7. Average VIL as a function of Range (km) for "truth" data (magenta), baseline unfiltered data (yellow), vertical merge (blue), and the final run using filtered data for detection and unfiltered data for diagnosis (red).
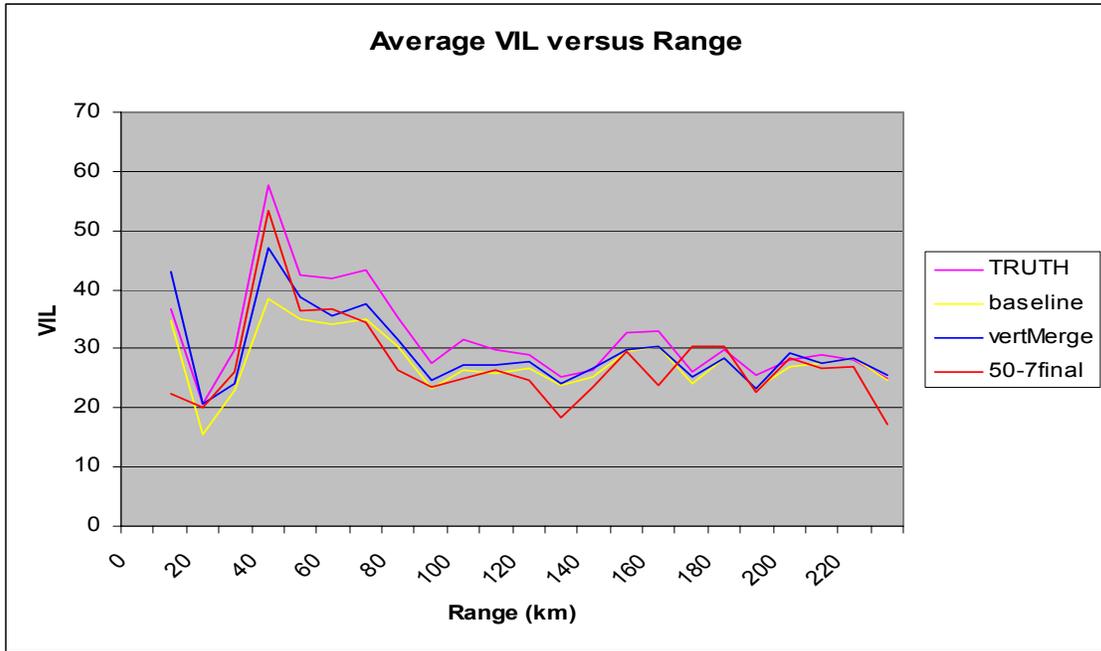

Figure 8. Percent difference (average ratios) from truth of the three runs (no filter, filter, and filter with bias correction) as a function of range. Baseline unfiltered data (yellow), vertical merge (blue), and the final run using filtered data for detection and unfiltered data for diagnosis (red).
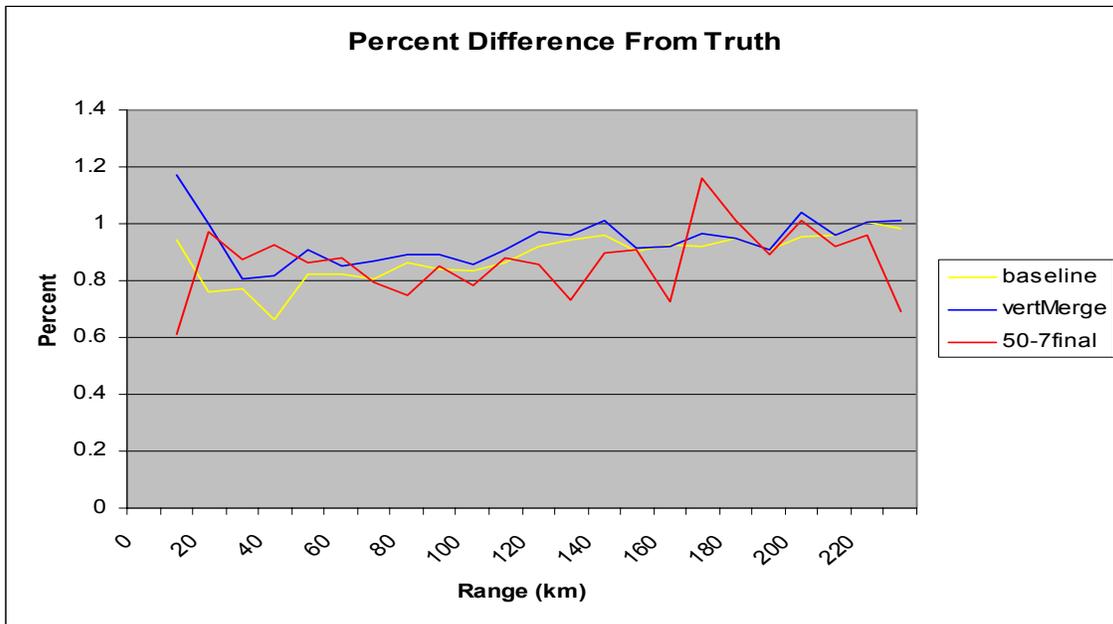
Figure 9. Percent difference (average ratios) from baseline of the three runs (no filter, filter, and filter with bias correction) as a function of range. Baseline unfiltered data (yellow), vertical merge (blue), filtered data (cyan), bias-corrected filtered data (purple), and the final run using filtered data for detection and unfiltered data for diagnosis (red).
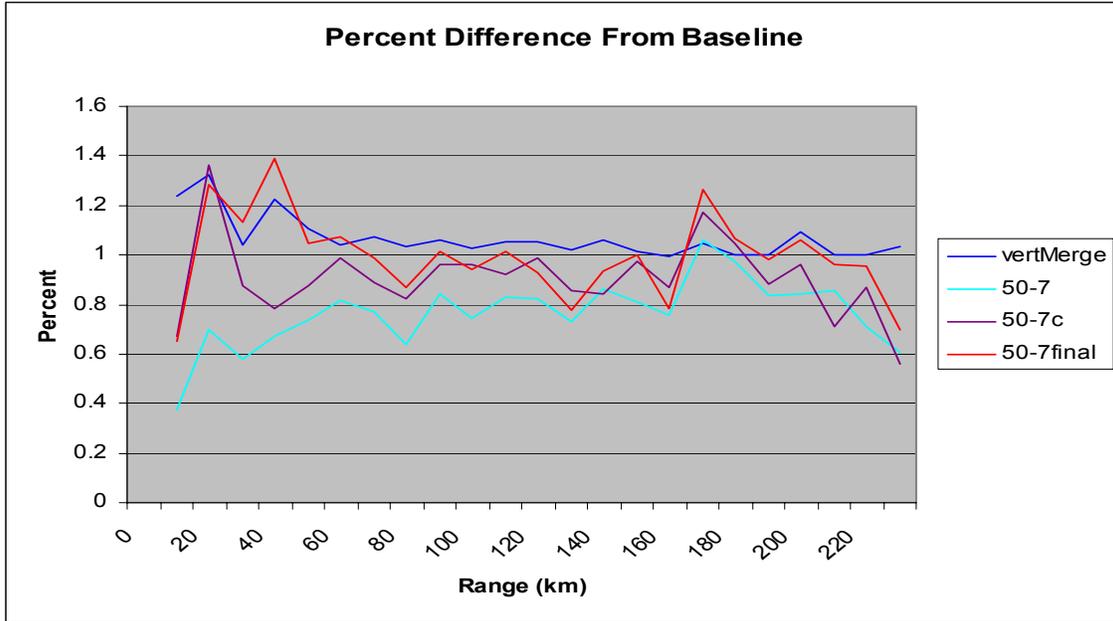

Table 5. Average ratios of VIL values between "truth" and each of 4 different SCIT runs (reflectivity with no filter, filter, and bias-corrected filter, filter for detection and no filter for diagnosis). Also, average ratios of VIL values between "baseline" and the 2 filtered SCIT runs

Reflectivity data	Average ratio from truth	Average ratio from baseline
Baseline (no filter)	.864	-
Vertical Merge	.908	1.049
Filter	.672	.777
Bias-Corrected Filter	.796	.920
Filter Detection/ No Filter Diagnosis	.845	.976

The average VIL ratio for the final run was 84.5% of truth value. The average VIL ratio was 97.6% of the baseline SCIT value.

Note that the vertical merge technique developed still shows the best average VIL ratio from truth and the baseline as compared to each of the filter techniques. Recall from Table 2 that the time association success rate for the vertical merge technique is inferior to all of the filter techniques. Again, the implications of these results will be discussed later.

In the interest of time, the reduction in the number of detected small and isolated (but valid) cells caused by the reduction of the maximum dBZ below SCIT algorithm thresholds was not determined. It is assumed that the results would not differ much from those

of the filter run for the 50% and 7 km filter parameters.

5. Summary and recommendation for operational systems

The final option, using the filtered data for cell detection, and reverting back to the original unfiltered data for storm diagnosis, was quite successful. Figure 10 shows the final results of the original SCIT, the vertical merge technique tested prior to FY03, and the final filter option. The time association success rate was improved from 77.2% for the baseline SCIT using unfiltered data to 94.1% using the final technique. Furthermore, the average VIL values between the final technique and the baseline SCIT was nearly the same (only a 2.4% reduction).

Figure 10. Comparison of time association success rate (Eq. 1), and average VIL ratios from the baseline (unfiltered SCIT) and truth for the baseline SCIT (blue), vertical merge technique (red) and the final filter option (yellow).
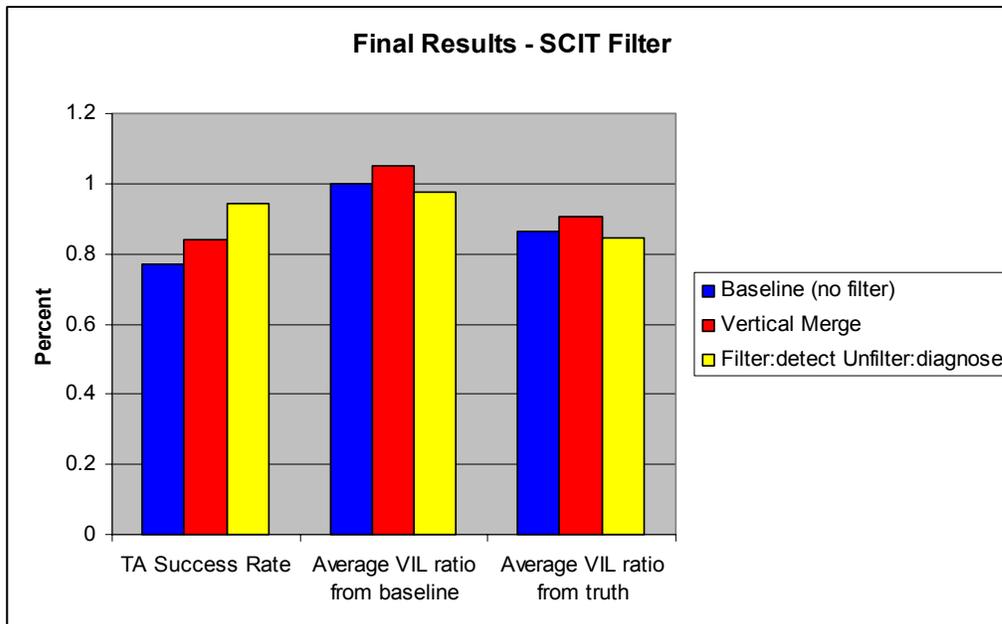
Considering that the vertical merge technique shows the best average VIL ratios as compared to truth, one recommendation for future work is be to combine the final filter technique with the vertical merge technique to maximize both time association success rate and to improve upon the VIL ratios as compared to truth. This option would be more expensive to implement, because it would involve numerous coding changes.

We recommended that this final technique, using the filtered data for detection, and the original unfiltered data for diagnosis, be implemented into the WSR-88D operational system as soon as possible. The new technique requires minimal coding changes with a large improvement in time association. For the baseline SCIT, storm tracks will be broken on average once every four volume scans (15-20 minutes). With the recommended technique, storm tracks should last on average over 9 volume scans, which is typically longer than the average storm lifetime.

6.   References

Hondl, K. D., 2003: Capabilities and components of the Warning Decision Support System – Integrated Information (WDSS-II). *Preprints, 19th Conf. On Interactive Info. Processing Systems*, Long Beach, California, Amer. Meteor. Soc., CD preprints.

Johnson, J. T., P. L. MacKeen, A. Witt, E. D. Mitchell, G. J. Stumpf, M. D. Eilts, and K. W. Thomas, 1998: The Storm Cell Identification and Tracking (SCIT) algorithm: An enhanced WSR-88D algorithm. *Wea. Forecasting*, **13**, 263-276.

Witt, A., and V. K. McCoy, 2002: Improvements to the WSR-88D Storm Cell Identification and Tracking algorithm. *Preprints, 21st Conf. On Severe Local Storms*, San Antonio, Texas, Amer. Meteor. Soc., 186-189.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286-303.