

NSIDC Special Report 3

Intercomparison of Resampling Methods for SMMR Pathfinder in EASE-Grid

Kenneth W. Knowles, 1997

This document compares several methods for resampling SMMR Pathfinder data to the EASE-Grid format, including:

- [Drop in the bucket](#)
- [Nearest neighbor](#)
- [Bilinear](#)
- [Cubic convolution](#)
- [Sin\(x\)/x](#)
- [Inverse distance squared](#)
- [Backus-Gilbert](#)



TABLE OF CONTENTS

1	ABSTRACT.....	2
2	COMPARISON CRITERIA.....	2
3	RESAMPLING METHODS.....	2
3.1	Drop in the bucket.....	2
3.2	Nearest neighbor.....	3
3.3	Bilinear.....	3
3.4	Cubic convolution.....	4
3.5	Sin(x)/x.....	5
3.6	Inverse distance squared.....	6
3.7	Backus-Gilbert.....	7
4	SAMPLE IMAGES.....	7
4.1	Full resolution, full hemisphere.....	7
4.2	10 times expanded resolution for selected areas.....	11
5	PLOTS AND INTERPOLATIONS.....	26
5.1	T _b Histograms.....	26
5.2	T _b Gradients.....	36
5.3	Daily differences and histograms.....	36
5.4	Comparison to Tessellation.....	48
5.5	Problems with the data.....	53
6	GLOSSARY.....	54
7	REFERENCES.....	56

1 ABSTRACT

Several methods for resampling SMMR Pathfinder data to the EASE-Grid format are compared. The following methods are described and their properties are discussed: [drop in the bucket](#), [nearest neighbor](#), [bilinear](#), [cubic convolution](#), [sin\(x\)/x](#), [inverse distance squared](#), [Backus-Gilbert](#). Where appropriate, each method is simulated by a 2-D transfer function and the frequency response is derived and is used to describe the expected properties of the method. Three methods, including drop in the bucket, nearest neighbor, and inverse distance squared, are tested on actual data. Sample images at 25 km and 2.5 km resolution are presented. Each full grid is then summarized by histograms of brightness temperature values and gradients, and these are compared to the original input data histograms. For each method, sample data are compared day to day. Each method is compared to a tessellation over a small area for a single orbit. Finally, problems with the data discovered during these tests are outlined.

2 COMPARISON CRITERIA

Assume that the original signal is continuous and band limited, and that the original signal can be completely reconstructed from the swath data. The goal then is to determine an interpolation method that will also allow us to reconstruct the original signal. There is no way for us to directly compare the results of any resampled data to the original signal nor even to the swath data. The best we can do is to evaluate the interpolation methods according to their characteristics. The best method I believe is the one that does the least smoothing to the data while adding as little noise as possible.

Of course, there are problems with these assumptions. The "original" original signal was the 2-D brightness temperature signal of the earth. This was sampled by the back and forth scanning of the SMMR instrument. That time continuous signal was then band limited along scan; however, it was not band limited across scan. Any across scan signal with spatial frequency greater than 1 cycle per 50km would then be aliased back into the sampled data as noise. A further complication with these assumptions is that the "original" sampled data were themselves resampled along scan so that all channels were coincident with the 37GHz sample locations. The effects of this as to smoothing the data are not studied here.

3 RESAMPLING METHODS

3.1 Drop in the bucket

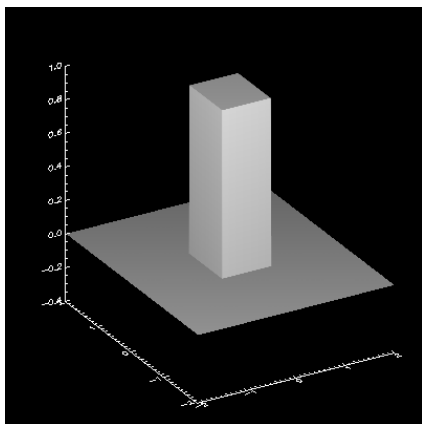
All data samples that fall within a grid cell are averaged together. This method cannot be simulated by a transfer function and so no frequency response can be derived. Due to the averaging involved

with this method, peaks in the histogram of Tb values will shift if the input histogram is not symmetrical. It will also decrease the occurrence of high gradients in the output. The output will appear to be smoothed and will have a different mean value. Since no spatial interpolation is done, there will also be geolocation errors.

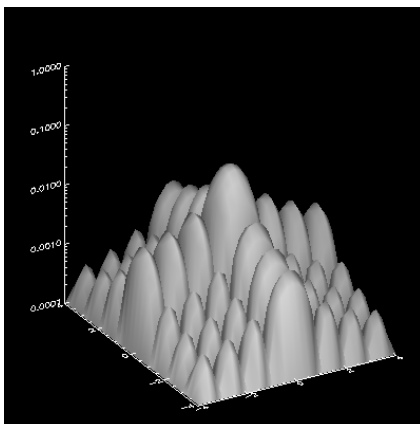
3.2 Nearest neighbor

Each grid cell is assigned the value of the nearest data sample. The transfer function is a square wave. The frequency response shows a lot of energy outside of the nyquist interval. This would indicate a very noisy output image. Because the input data are not averaged in any way, the output histogram of values should closely match the input. However, a noisy transfer will tend to increase the occurrence of high gradients in the output. The output image will appear speckled and will have a high incidence of geolocation errors.

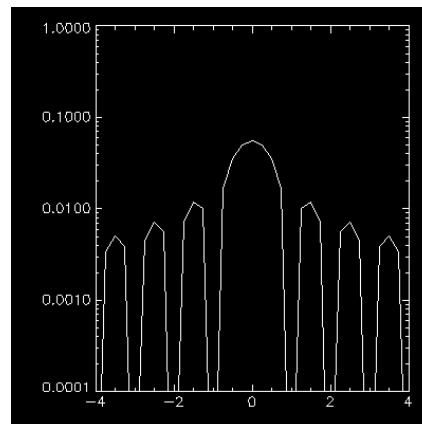
Transfer Function



Frequency Response



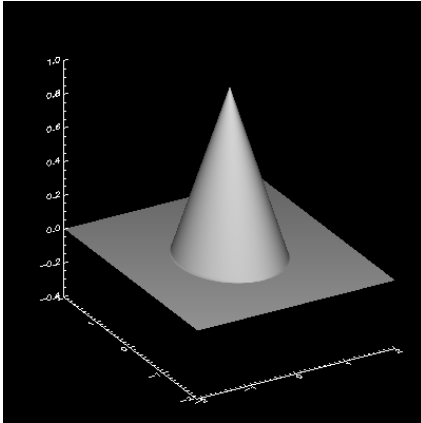
Profile at y=0



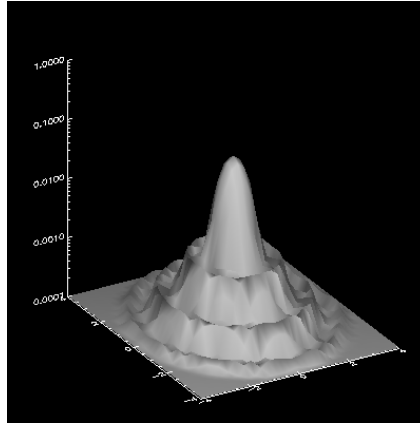
3.3 Bilinear

A grid cell is located in the sample data coordinate system at (u, v) which is in the middle of a quadrilateral of 4 data samples, (i, j) , $(i+1, j)$, $(i, j+1)$, $(i+1, j+1)$, where $i = \text{int}(u)$ and $j = \text{int}(v)$. First, the top two data samples are interpolated and the bottom two data samples are interpolated, then the results of these two interpolations are interpolated. This can also be expressed as a weighted average with weights as follows: $w(i, j) = (1 - |u - i|)(1 - |v - j|)$. The transfer function is a squared off cone.

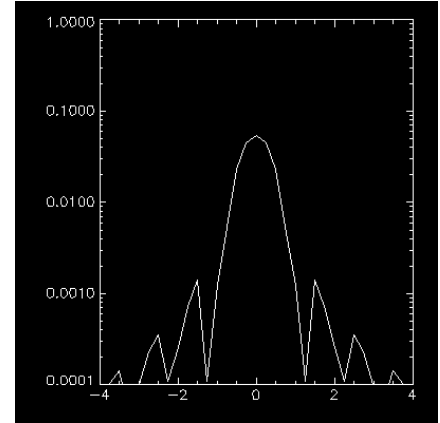
Transfer Function



Frequency Response

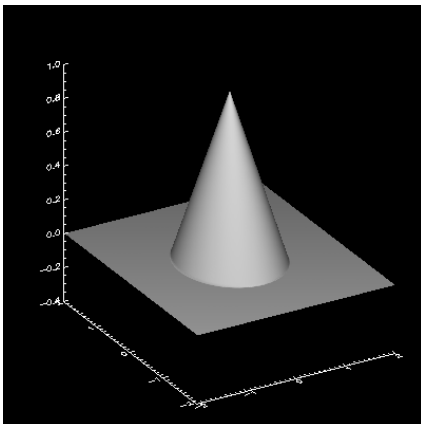


Profile at y=0

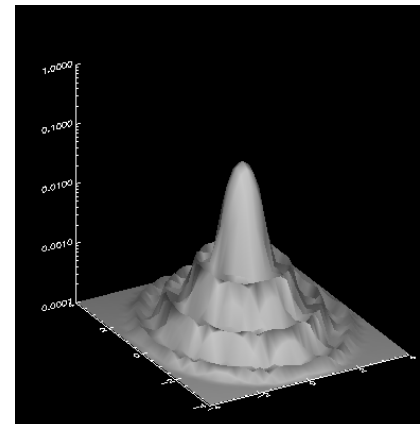


If the data coordinates of the grid cell are not known, then linear interpolation can be approximated by using weights: $w(i,j) = (1-d)*(1-d)$, where d is the distance from the grid cell center to the sample (i,j) . The transfer function for this method is a cone. The frequency response for this function has relatively small side lobes, but a fat center lobe which expands beyond nyquist. Because of the averaging this method could shift the histogram of Tb values and decrease the occurrence of high gradients. The output will be noticeably smoothed.

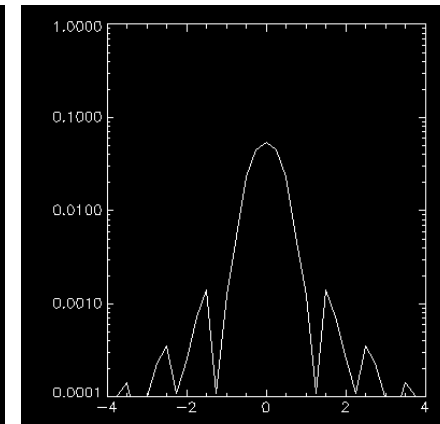
Transfer Function



Frequency Response



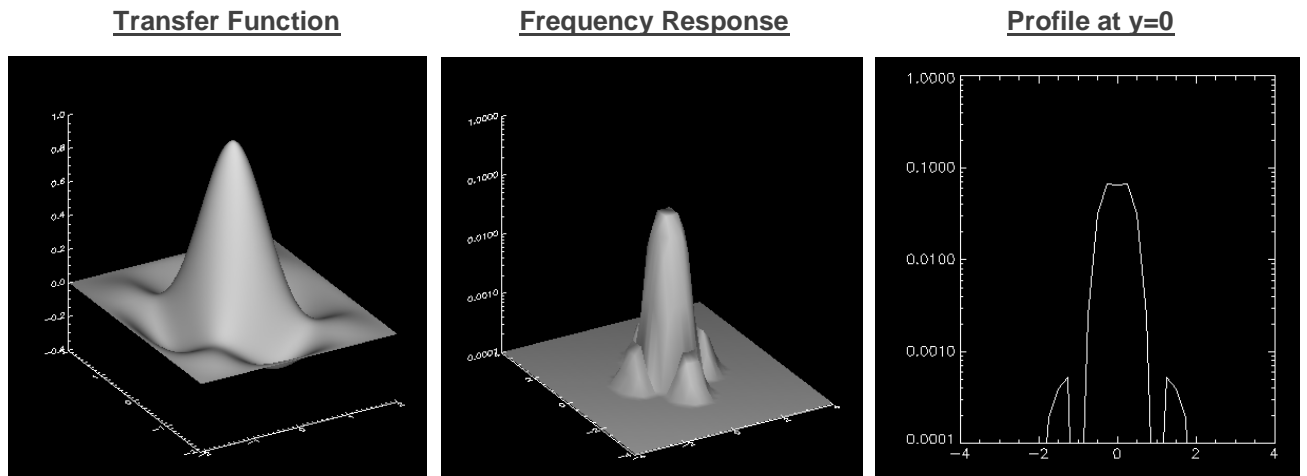
Profile at y=0



3.4 Cubic convolution

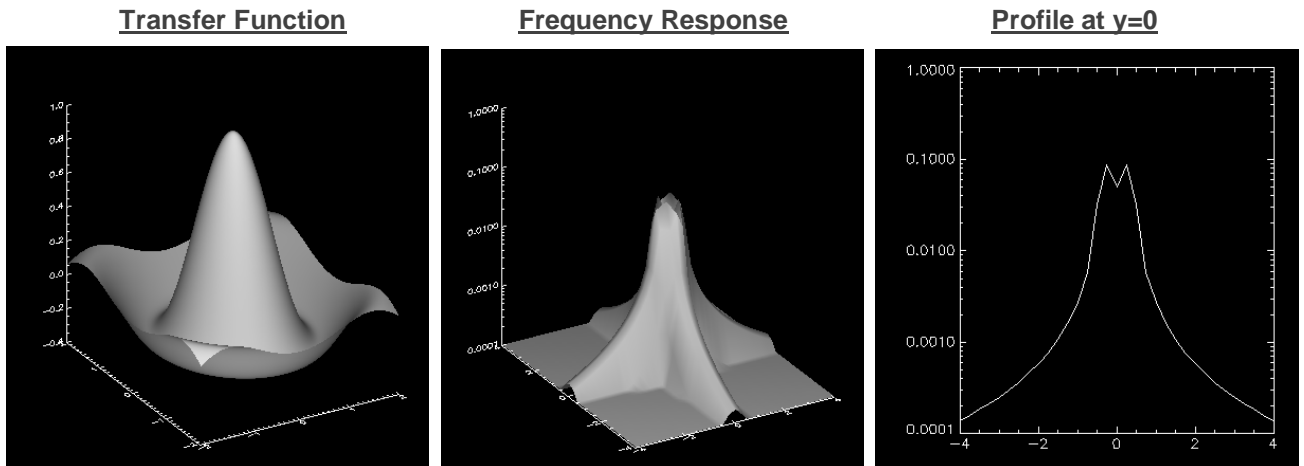
Cubic convolution is similar to bilinear, but requires 16 surrounding samples and uses cubic splines rather than linear interpolation. If the grid cell is at column u in the data, samples at columns $i-1$, i , $i+1$, and $i+2$ are used. Similarly for rows substituting v for u and j for i . Four interpolations are done across each row, then the four results are interpolated. This can also be expressed as a weighted sum: $w(x,y) = C(x)*C(y)$, for $x = (i-1, i, i+1, i+2)$, $y = (j-1, j, j+1, j+2)$, where $C(i-1) = p*(1-p)^2$, $C(i) = 1 - 2*p^2 + p^3$, $C(i+1) = p*(1+p-p^2)$, $C(i+2) = (1-p)*p^2$, $p = u - i$, and similarly for j ,

let $p = v - j$. The transfer function is a squared off splash. The frequency response shows a narrow center, flanked by very small side-lobes. Because the transfer function has negative and positive values this method should sharpen the image while not adding significant noise. The histograms of Tb values and gradients for the output should match the input. This technique requires a backwards navigation to determine sample data coordinates for grid cell centers. This is not a simple transformation when relying solely on the data within the SMMR Pathfinder files.



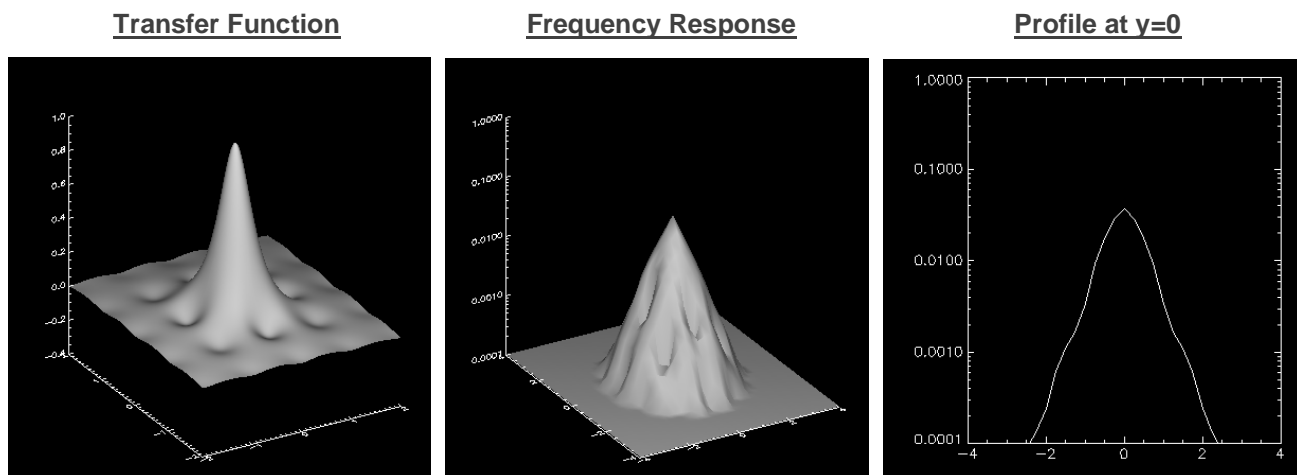
3.5 Sin(x)/x

Sine x over x interpolation is a weighted average over any size kernel with weights solely based on the distance of a sample from the grid cell center. The weight is: $w(d) = \sin(2\pi d)/(2\pi d)$. The transfer function is the classic damped wave rotated about the z-axis. The signal is limited and non-periodic so the expected pair of spikes in the frequency response get smeared out a bit.

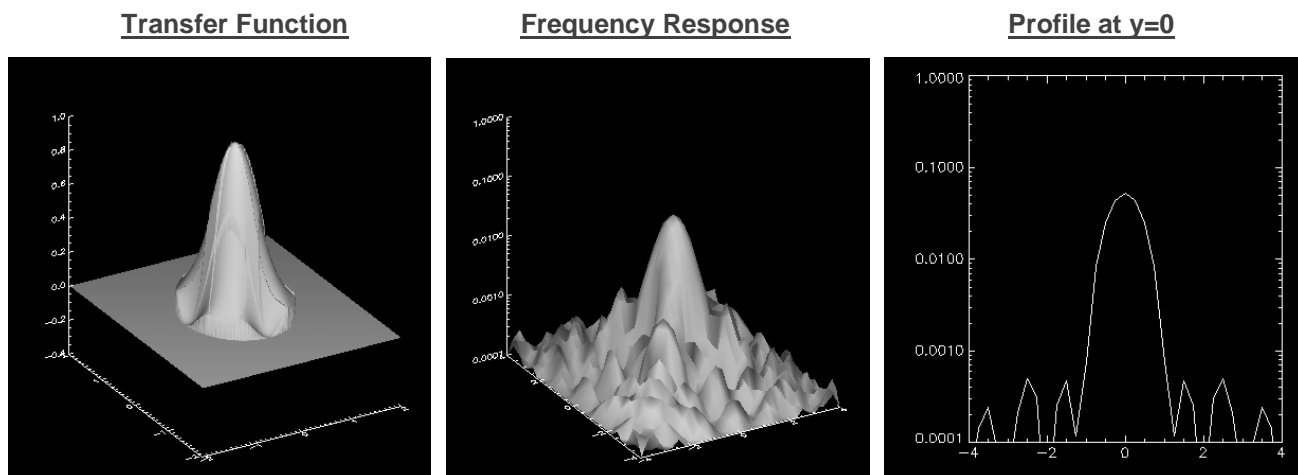


3.6 Inverse distance squared

Inverse distance squared is another weighted average based on the distance of a sample from the grid cell center. The weight is: $1/d^2$. The transfer function is a dimpled spike and the frequency response is a single fat lobe extending well beyond nyquist.



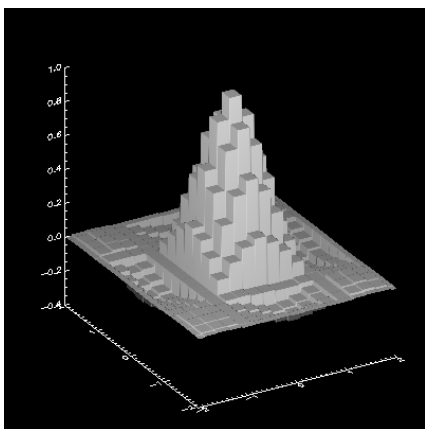
Limiting the maximum distance to 1 sample interval (25km) shrinks both the transfer function and the frequency response's center lobe. This is a forward resampling technique that is easily applied, and a good alternative to backwards navigation methods when the relationship of the grid cell location to the sample coordinate space is difficult to back out (as is the case with the SMMR Pathfinder data).



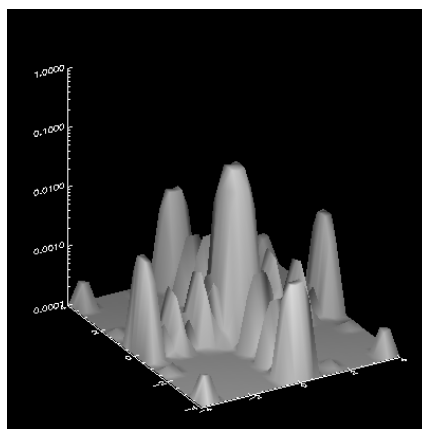
3.7 Backus-Gilbert

The SSMR swath data were originally sampled at different intervals for each channel. The Pathfinder data have been interpolated along scan to provide coincident measurements for all channels at the original 37GHz locations. Because of this, the Backus-Gilbert optimal interpolation technique, used by the SSM/I EASE-Grid resampling method, is not appropriate for the SSMR Pathfinder data. However, it is included here for comparison. The ideal transfer function would be the antenna gain pattern. However, backing out the weights for any given location within the sensor field of view is too computationally intensive for routine data processing, so a 16x16 sub-sampled approximation is used. The coefficients for the weighted sum are pre-computed at 1/4 sample and 1/4 scan intervals and the nearest-neighbor to the actual grid cell center is used. Because of this combination the transfer function appears blocky and the frequency response has some side lobe spikes outside of nyquist.

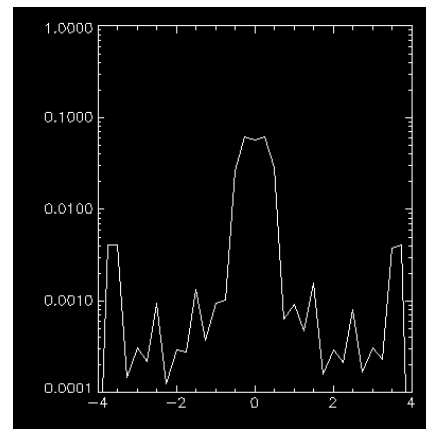
Transfer Function



Frequency Response



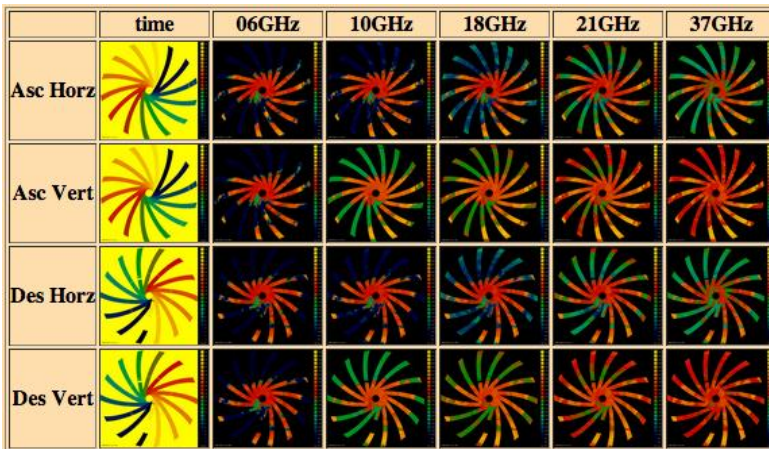
Profile at y=0



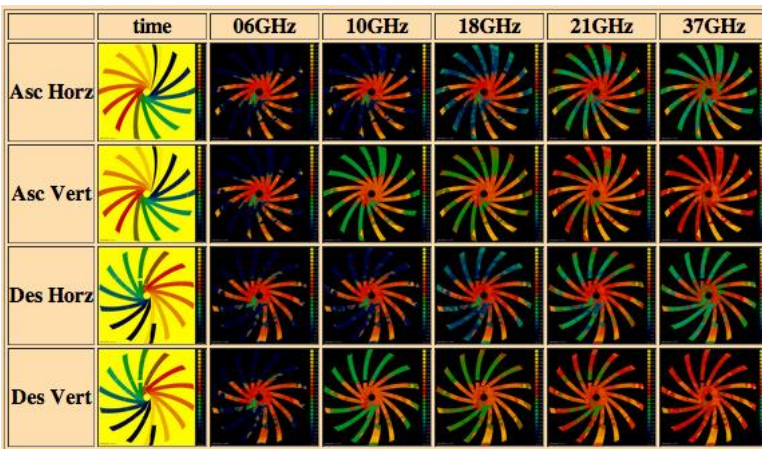
4 SAMPLE IMAGES

4.1 Full resolution, full hemisphere

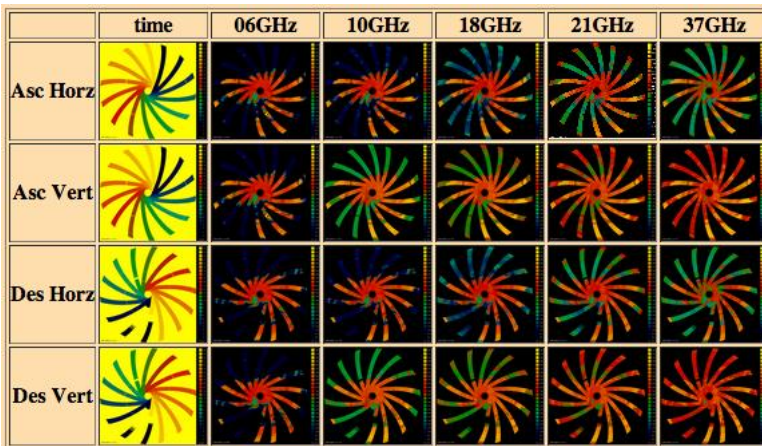
Three interpolation methods were selected for processing of full hemisphere grids. The sample data are all from the Northern hemisphere, for six consecutive days (three "on" periods) chosen arbitrarily (i.e. these were the first data available to me).



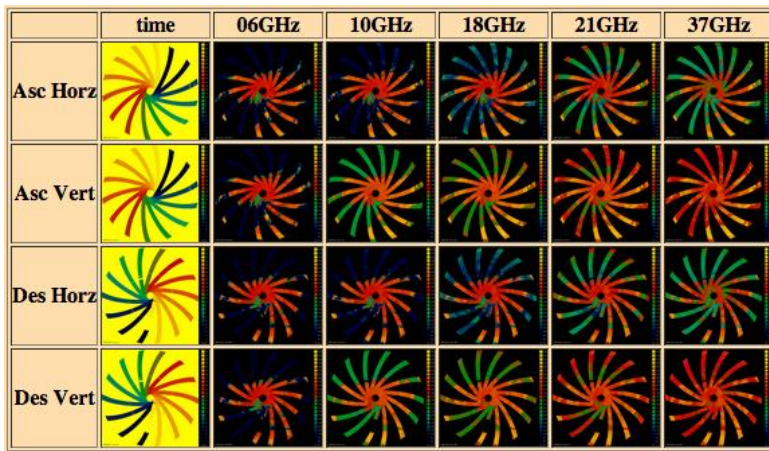
Nearest neighbor, north 25km 1981 day 81 (March 22).



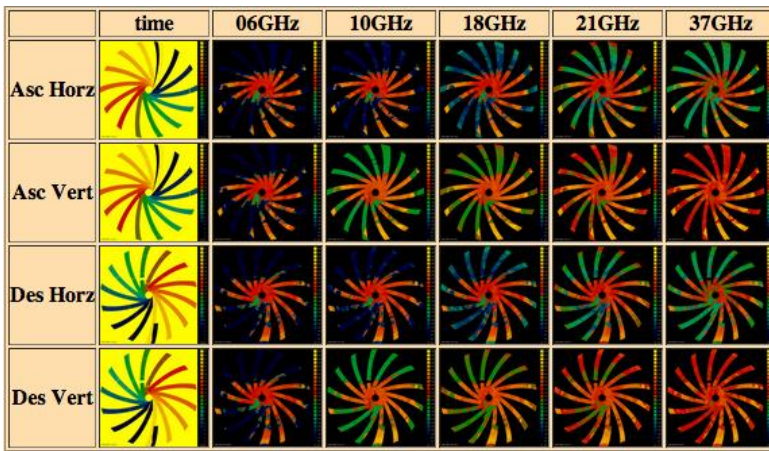
Nearest neighbor, north 25km 1981 day 83 (March 24).



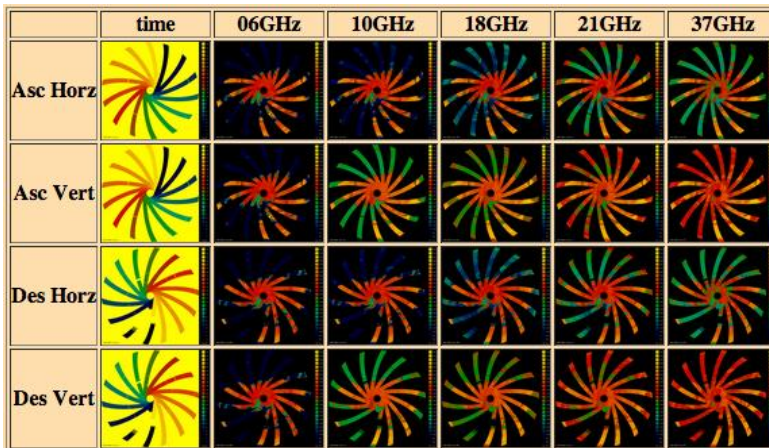
Nearest neighbor, north 25km 1981 day 85 (March 26).



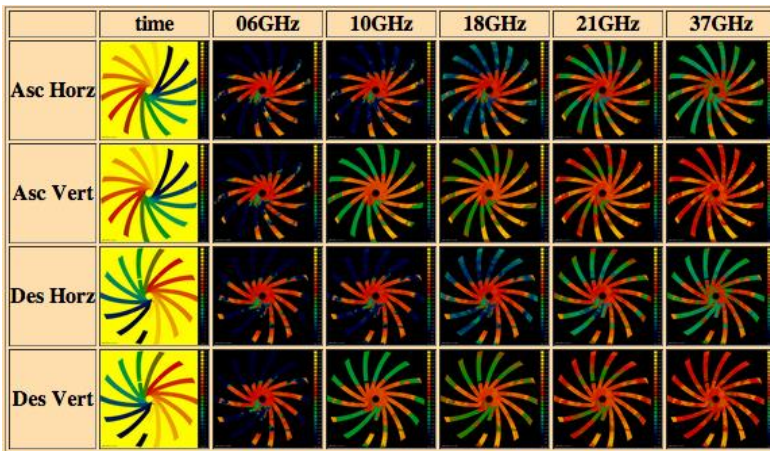
Drop in the bucket, north 25km 1981 day 81 (March 22).



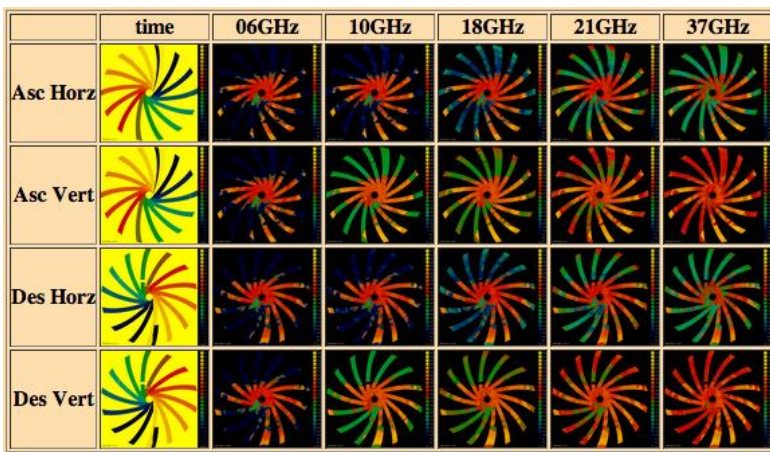
Drop in the bucket, north 25km 1981 day 83 (March 24).



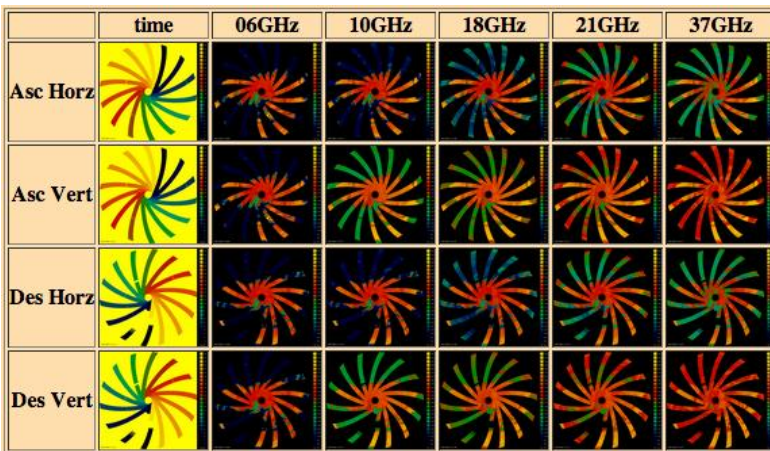
Drop in the bucket, north 25km 1981 day 85 (March 26).



Inverse distance squared, north 25km 1981 day 81 (March 22).



Inverse distance squared, north 25km 1981 day 83 (March 24).



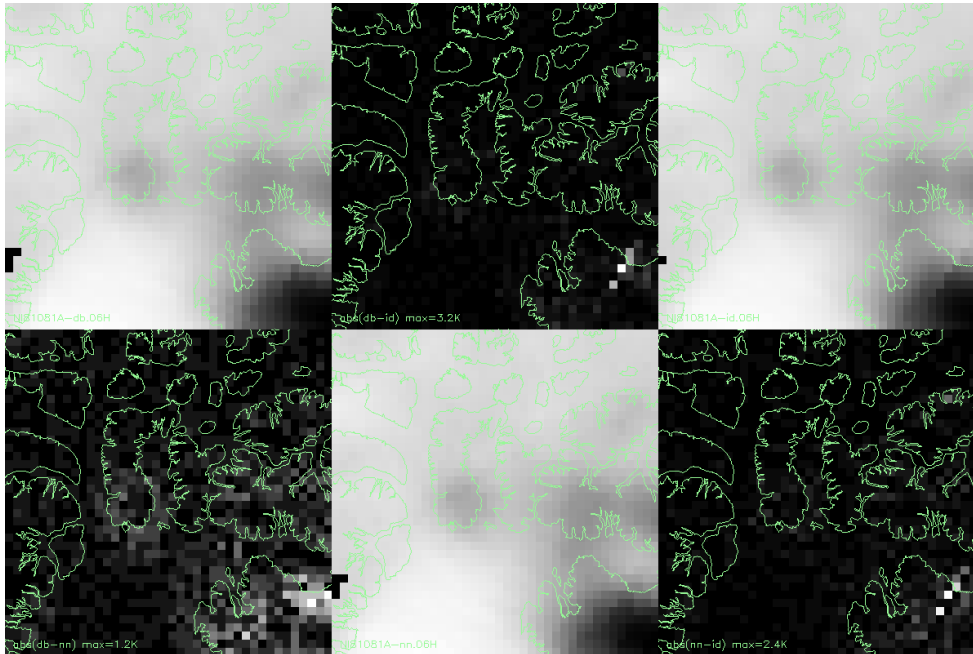
Inverse distance squared, north 25km 1981 day 85 (March 26).

4.2 10 times expanded resolution for selected areas

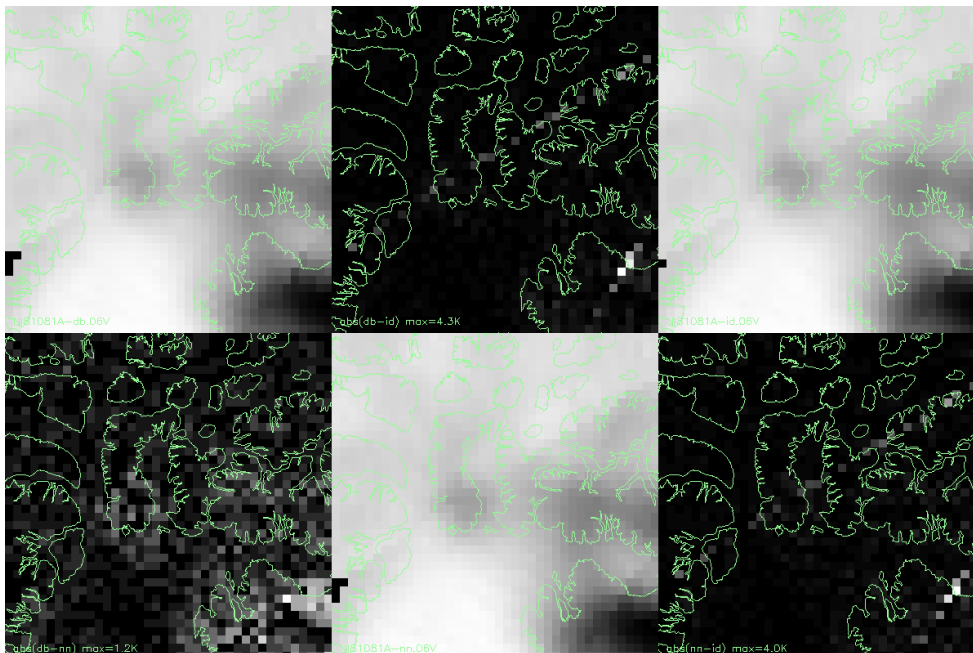
These images zoom in to visually assess differences. See below for mosaic images.

4.2.1 Baffin Bay

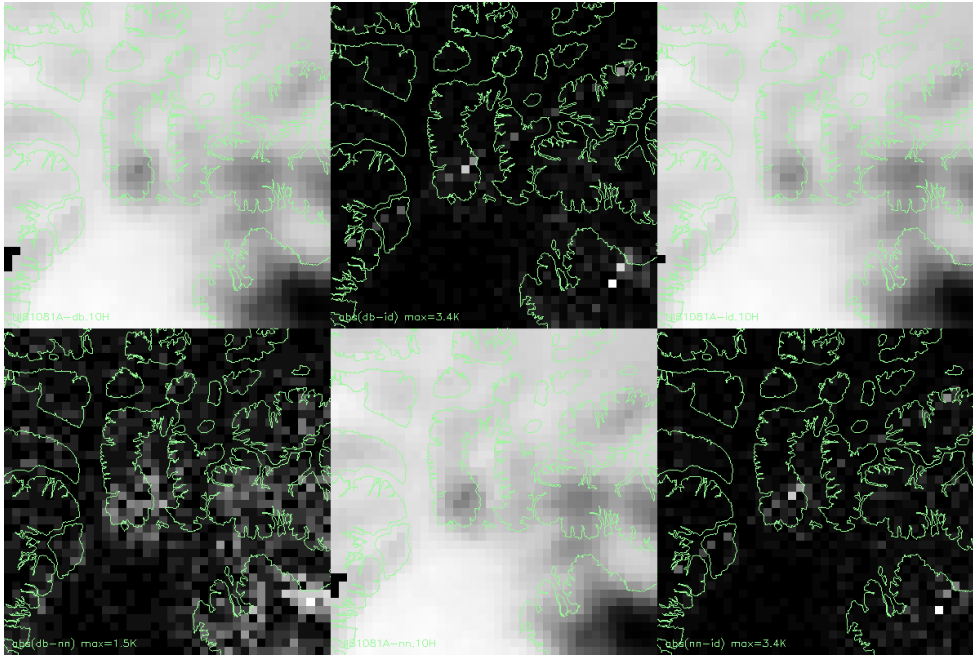
NI81081A.06H



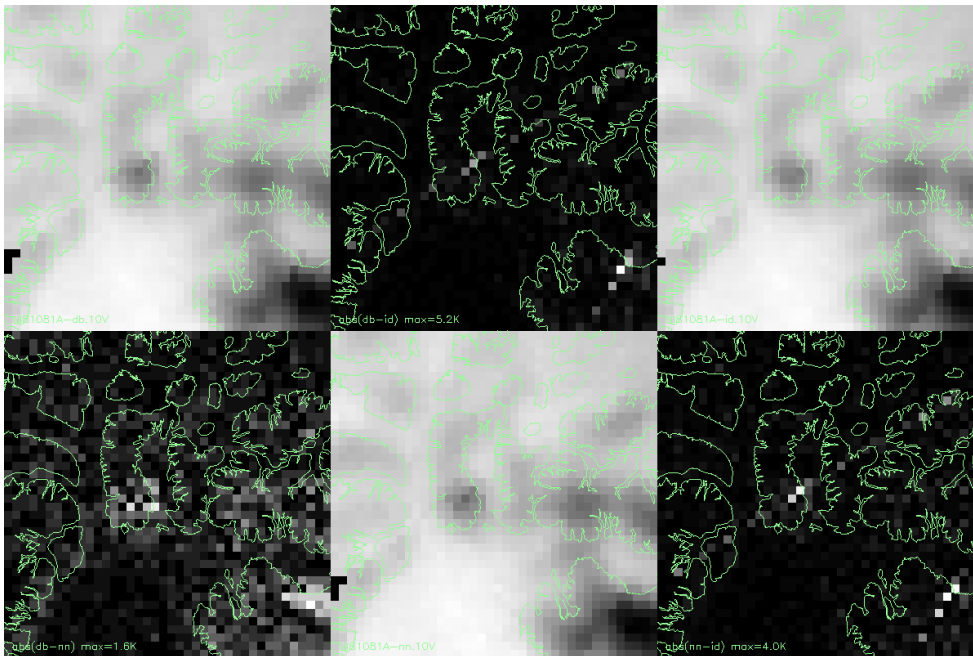
NI81081A.06V



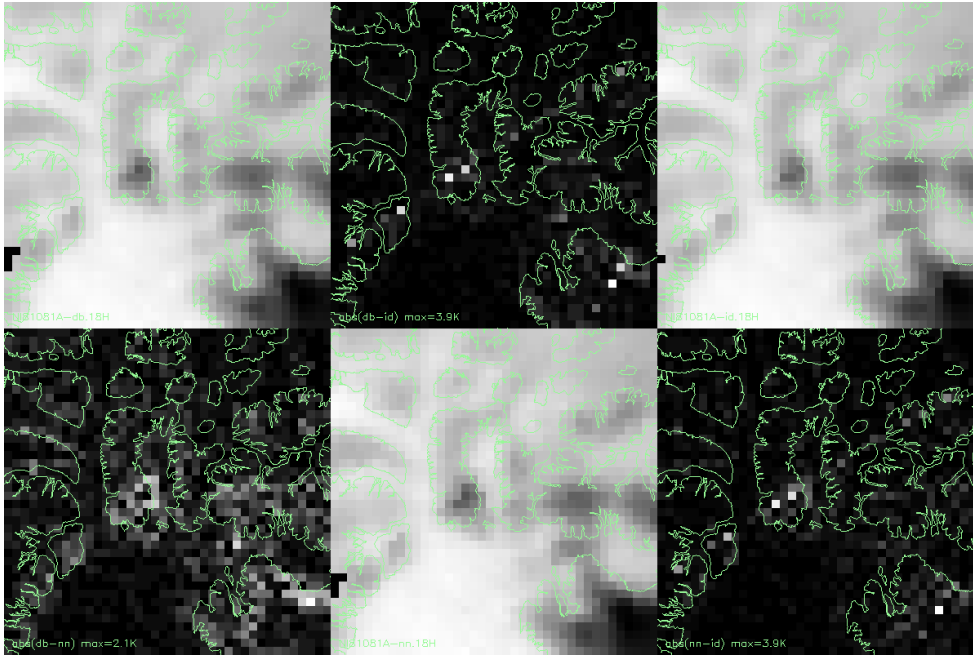
NI81081A.10H



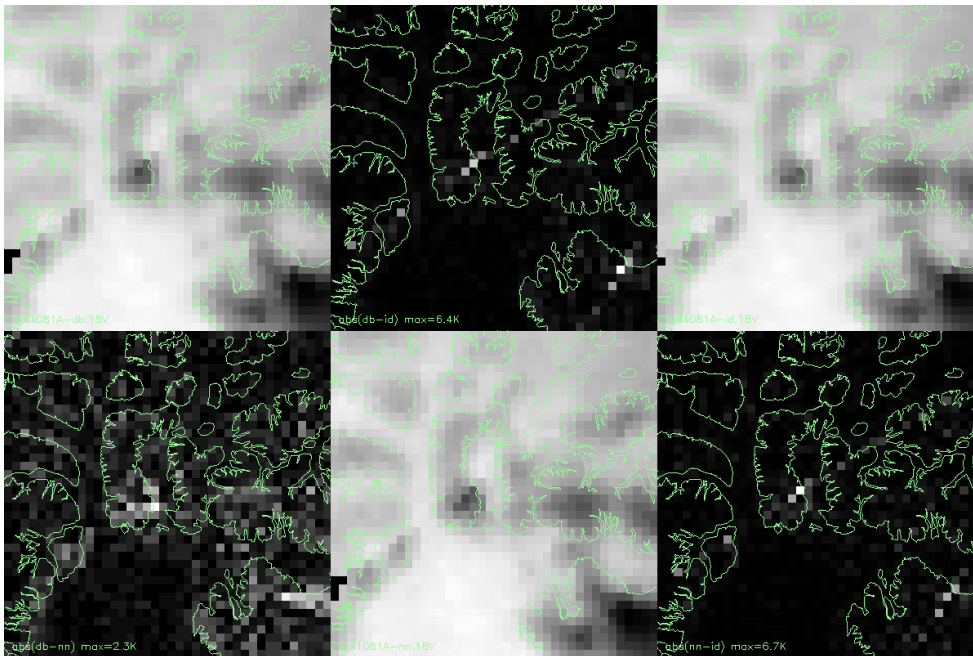
NI81081A.10V



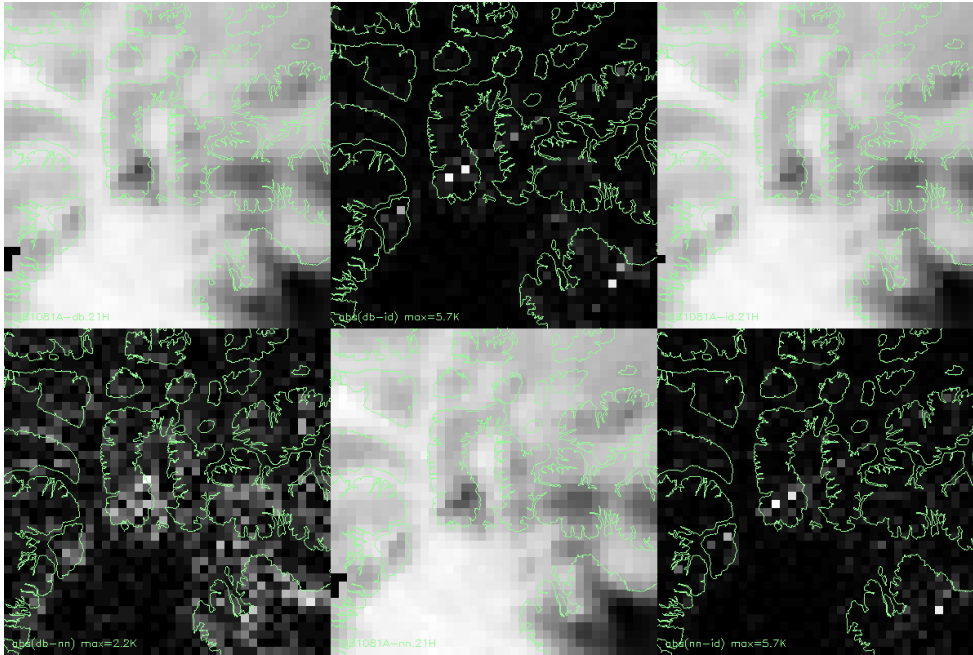
NI81081A.18H



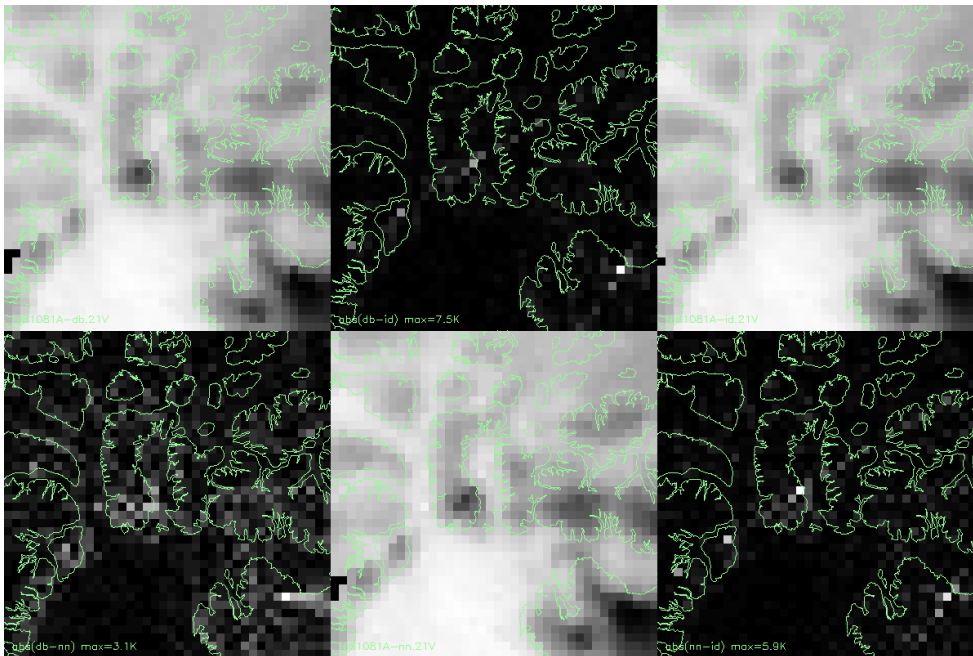
NI81081A.18V



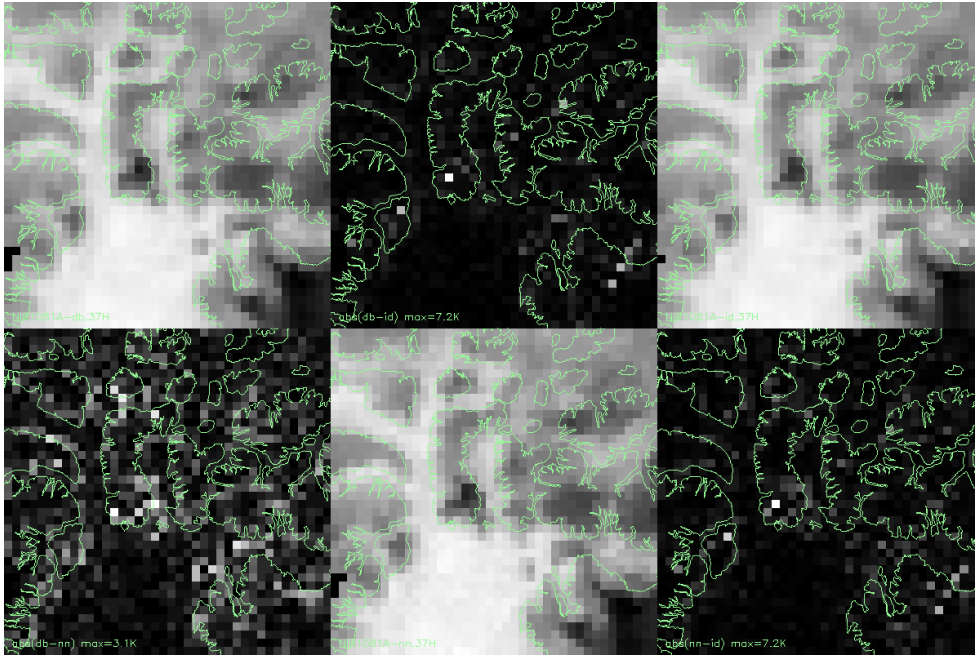
NI81081A.21H



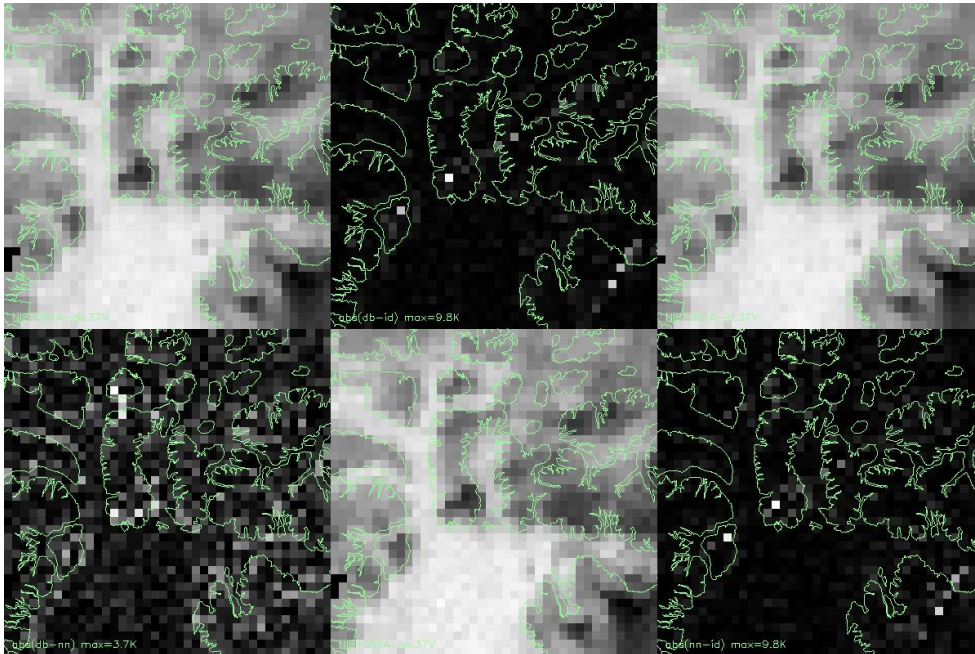
NI81081A.21V



NI81081A.37H

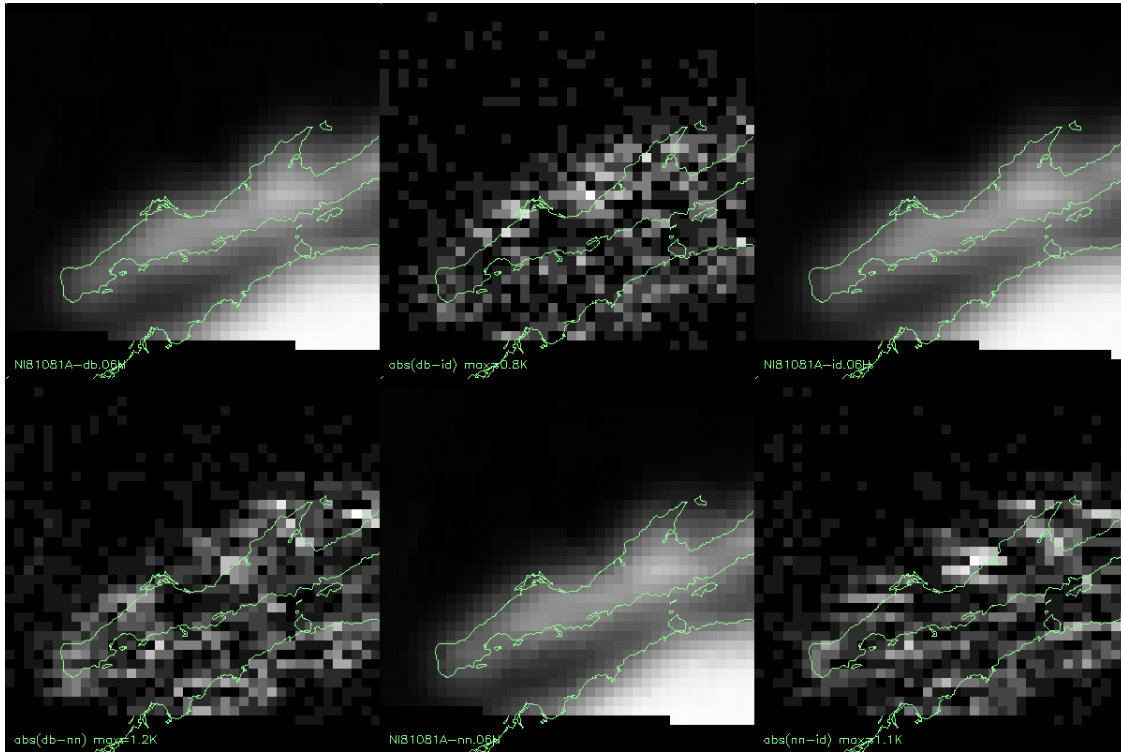


NI81081A.37V

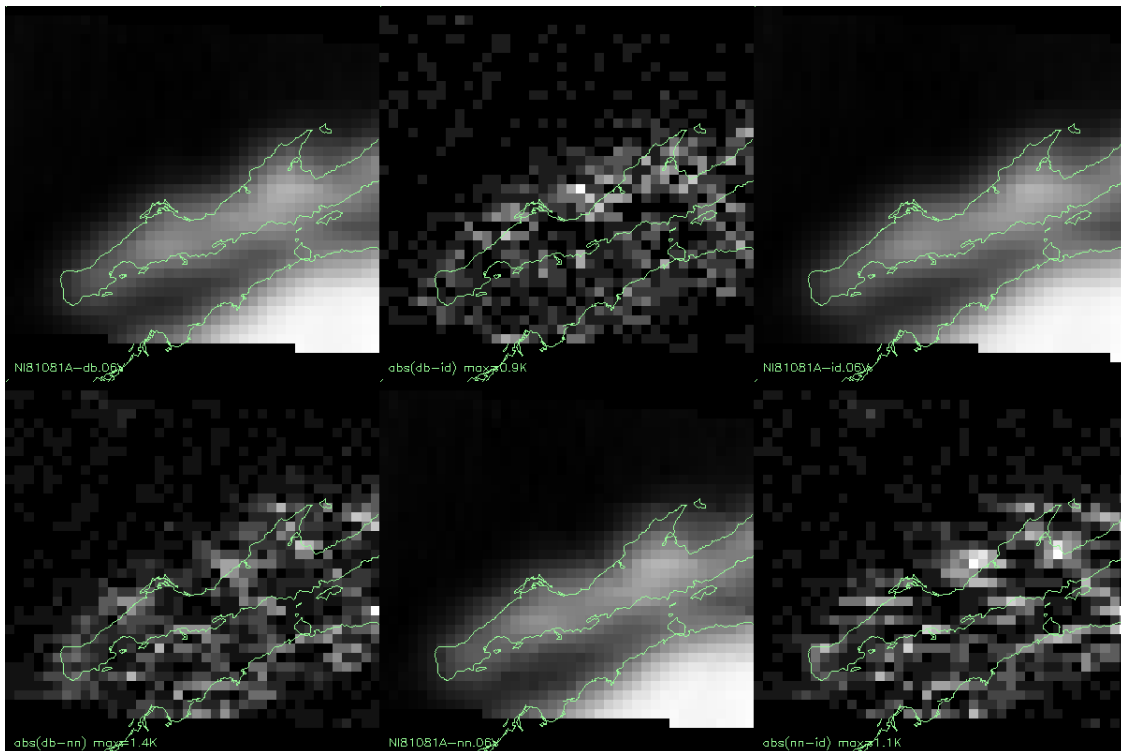


4.2.2 Gulf of California

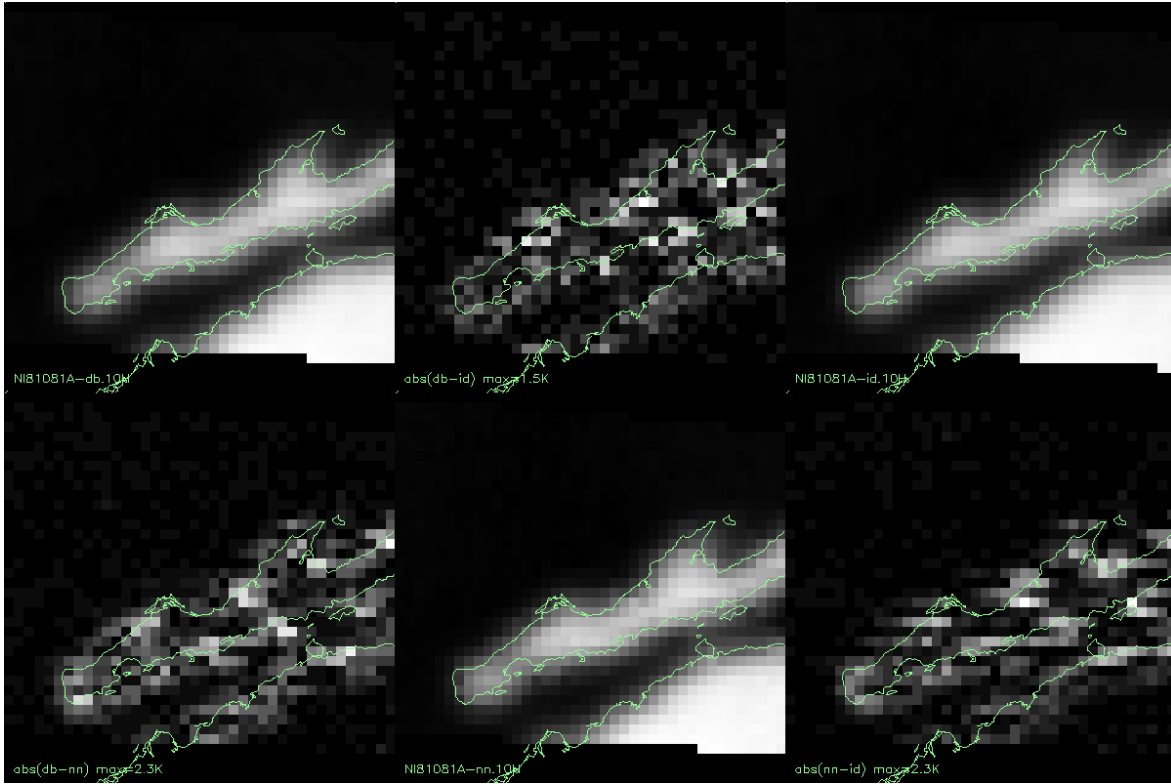
NI81081A.06H



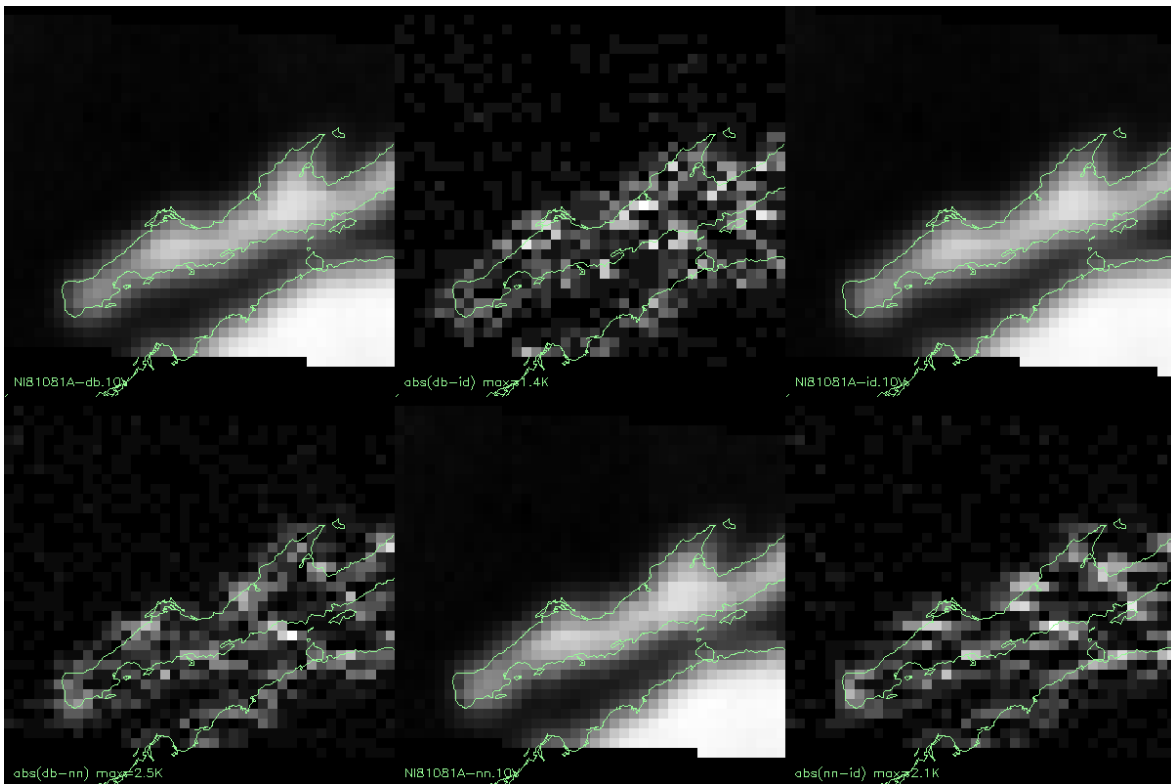
NI81081A.06V



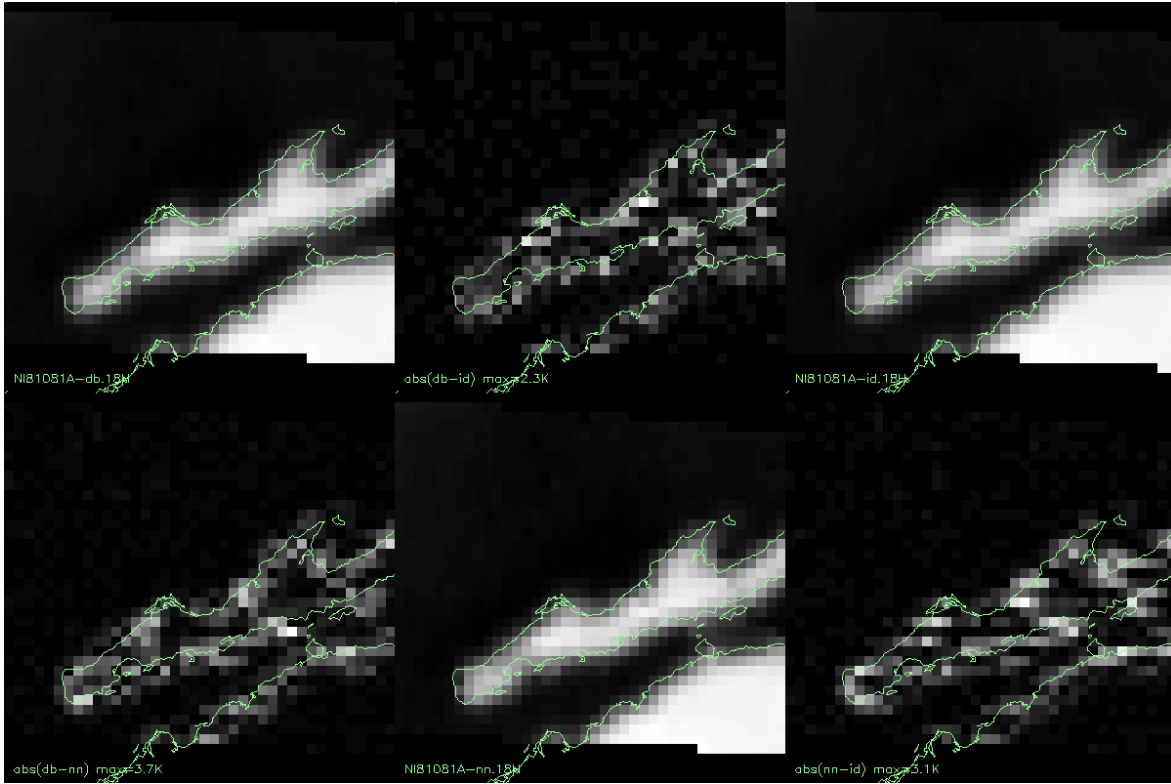
NI81081A.10H



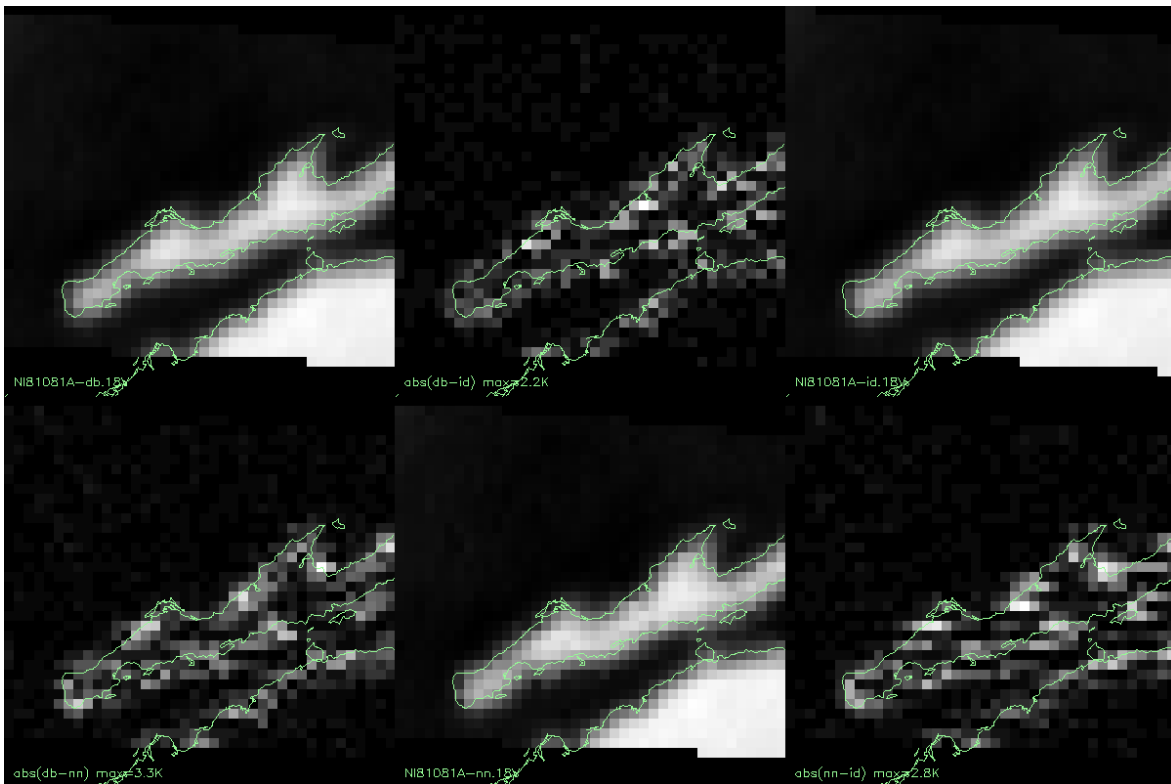
NI81081A.10V



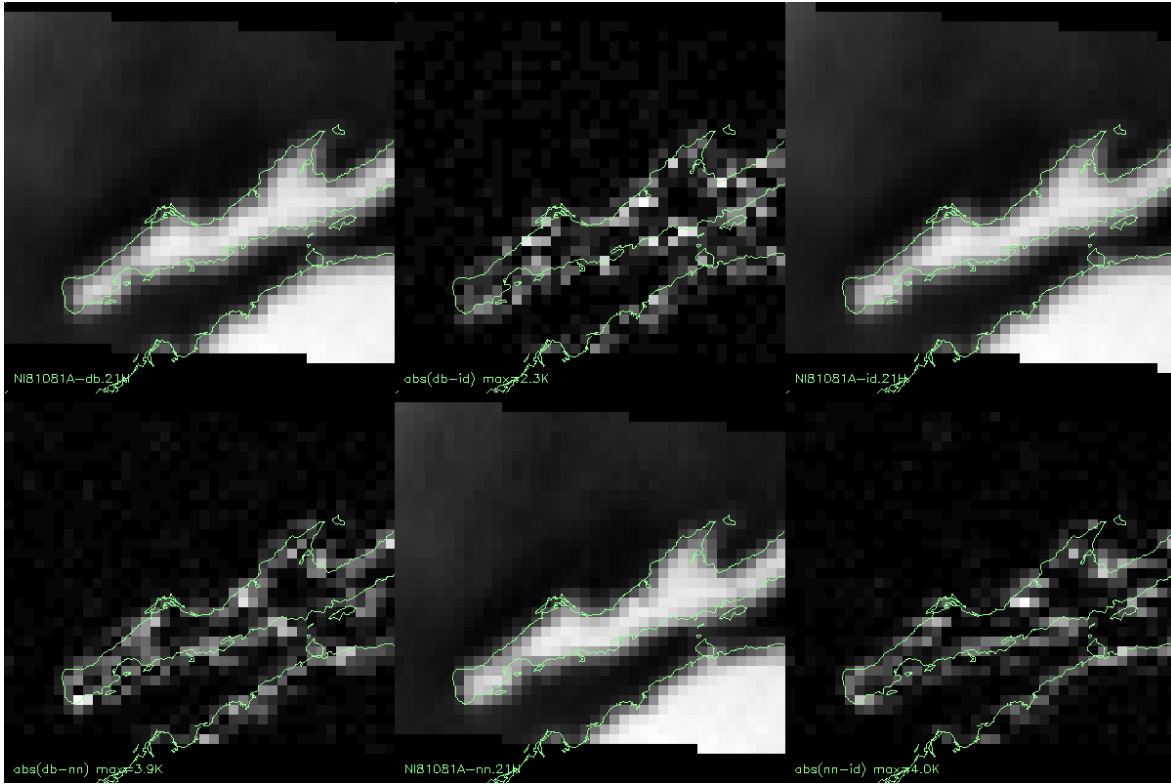
NI81081A.18H



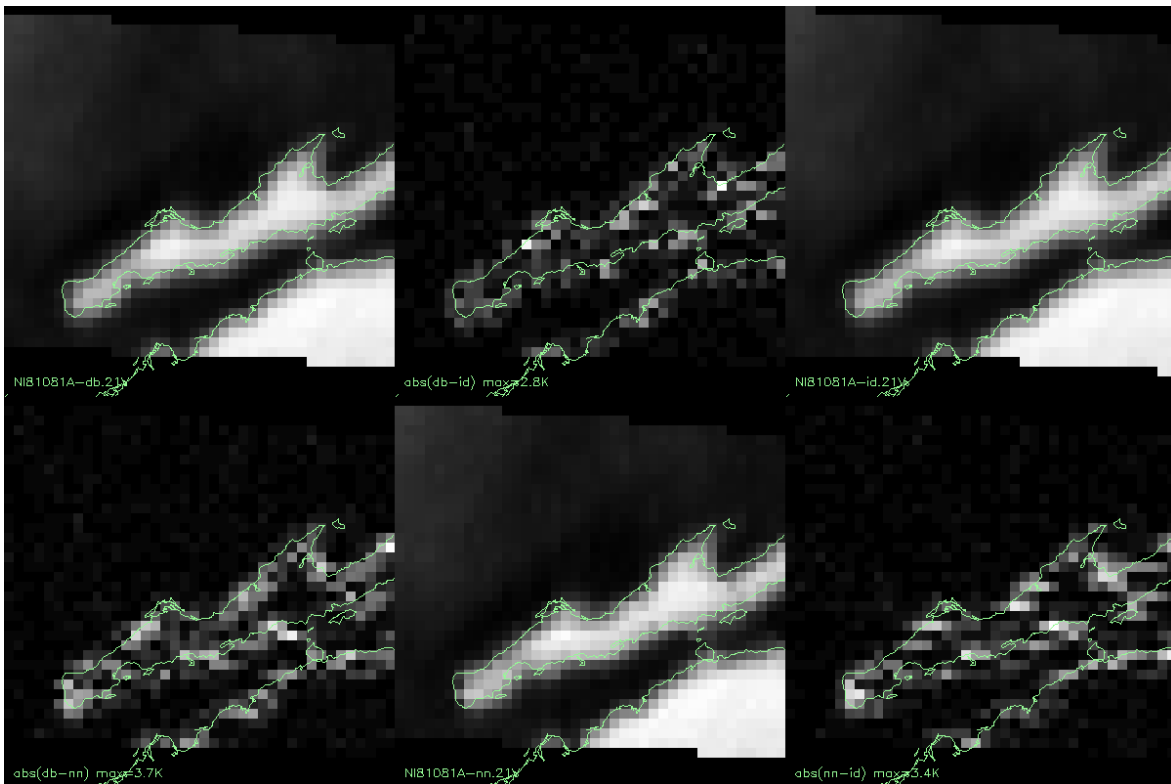
NI81081A.18V



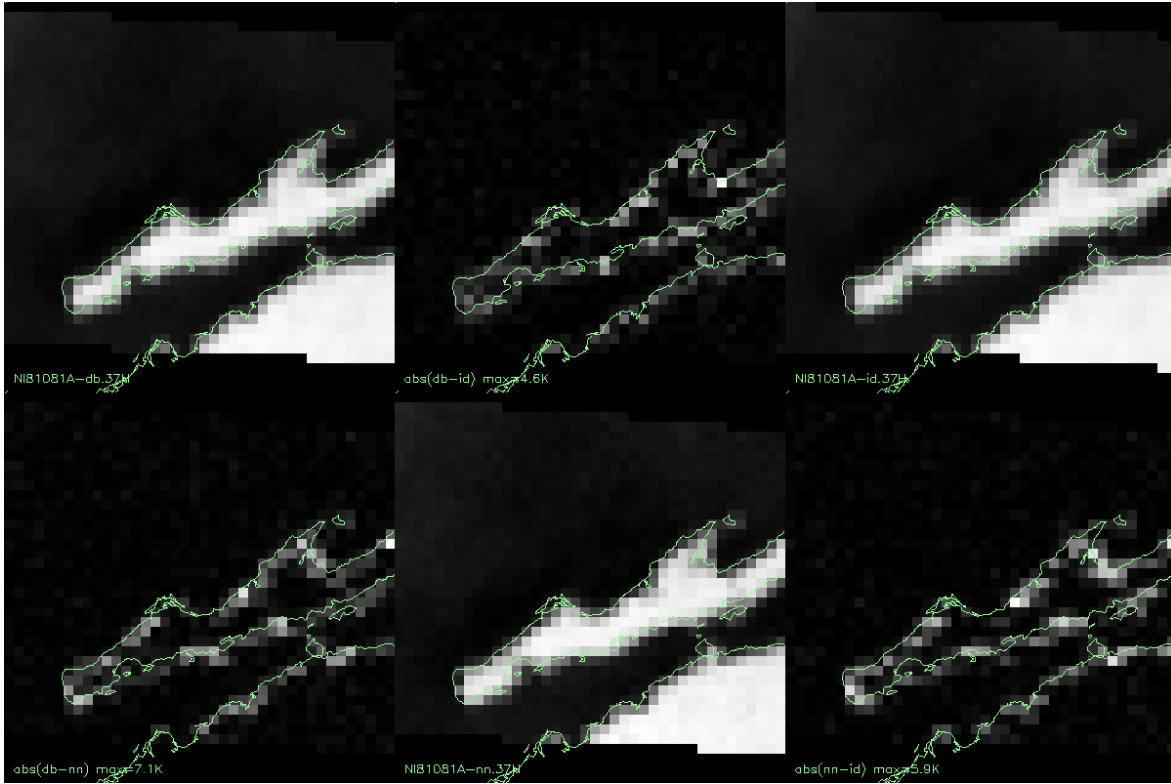
NI81081A.21H



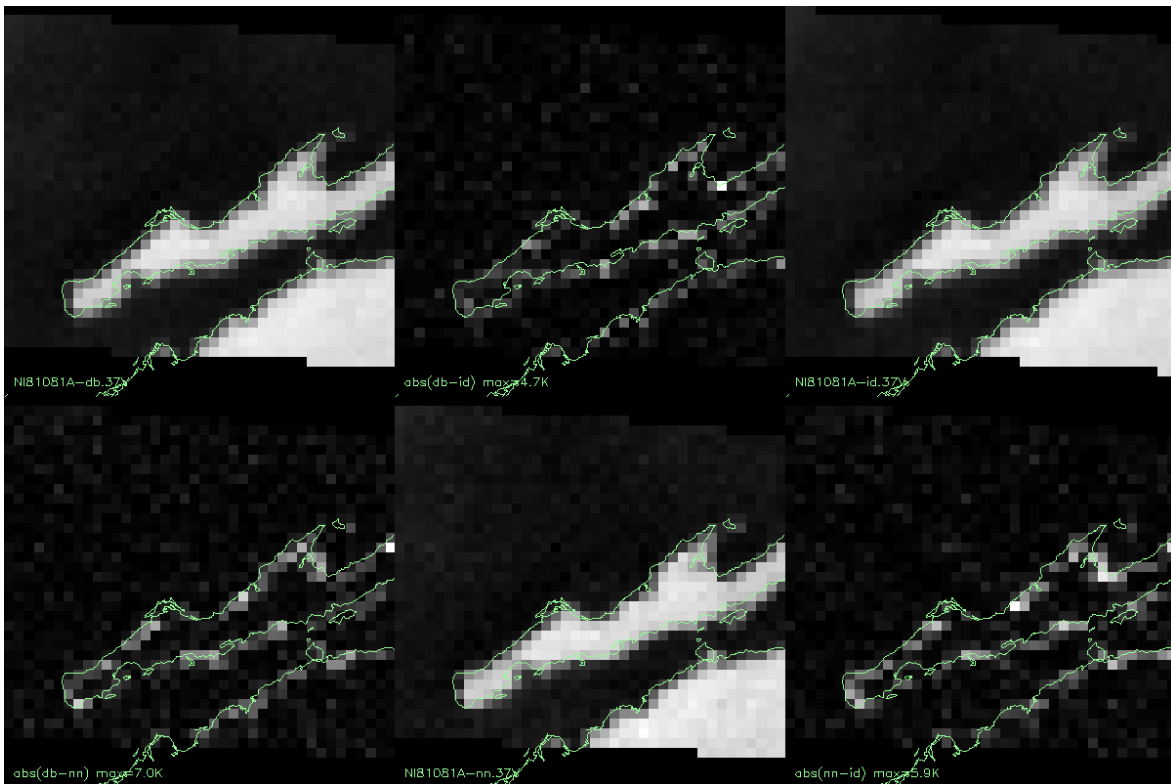
NI81081A.21V



NI81081A.37H

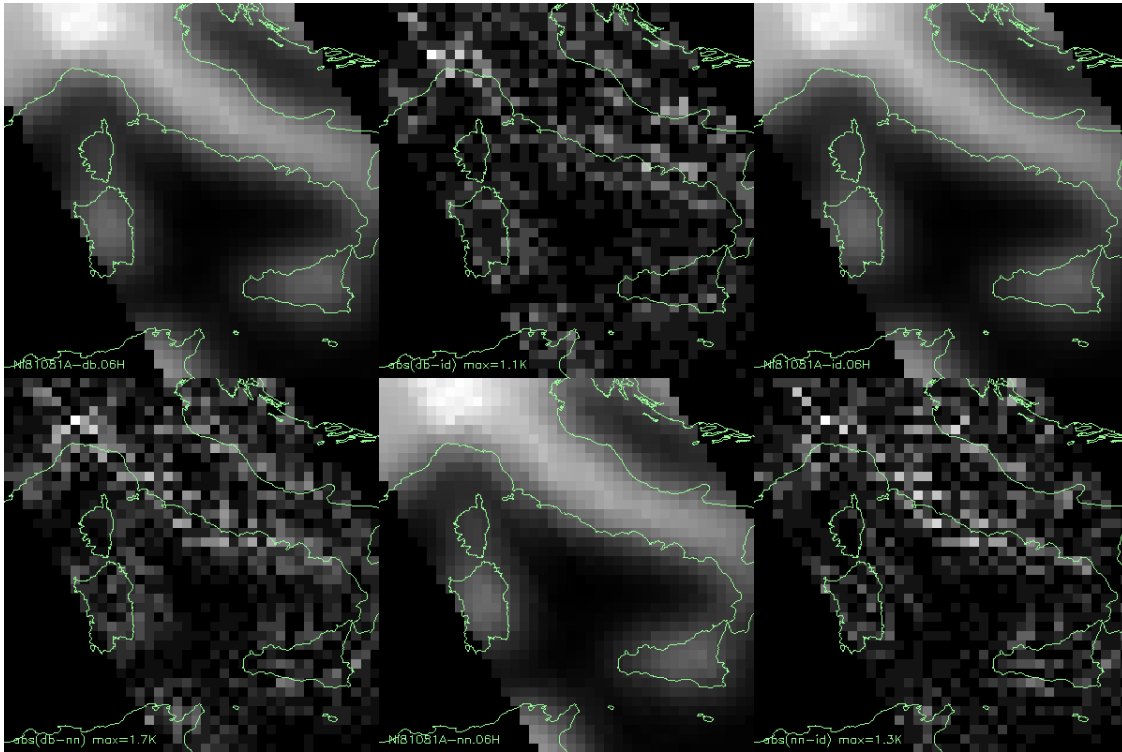


NI81081A.37V

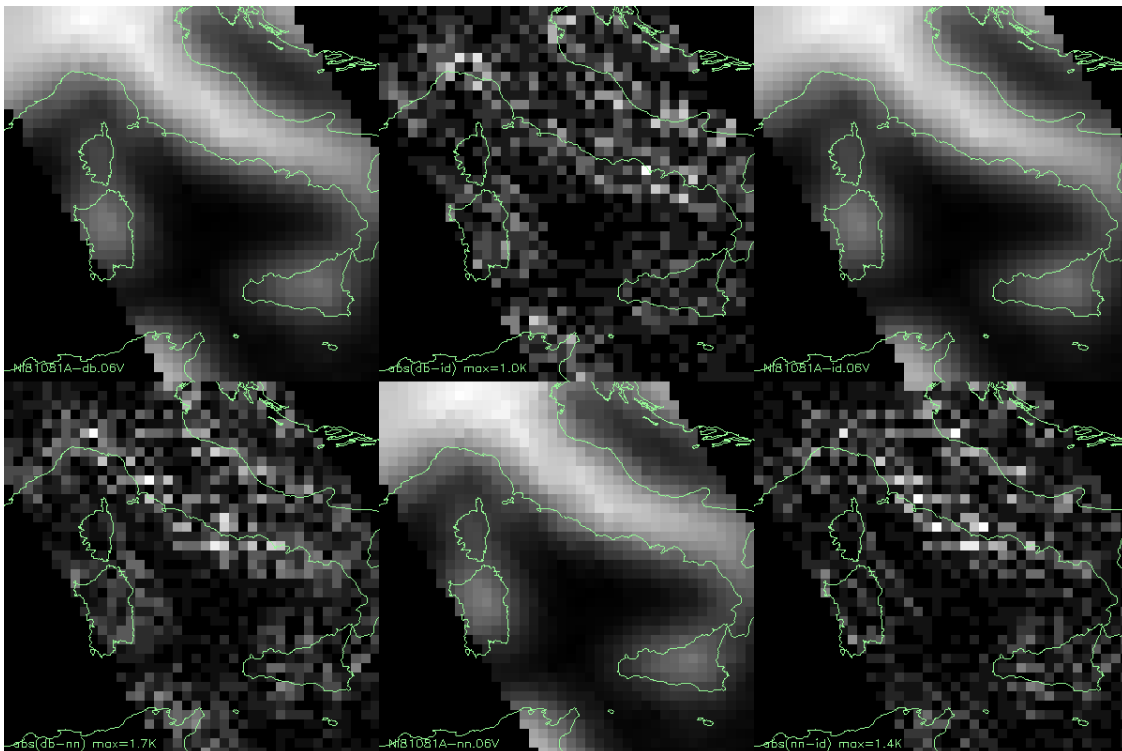


4.2.3 Tyrrhenian Sea

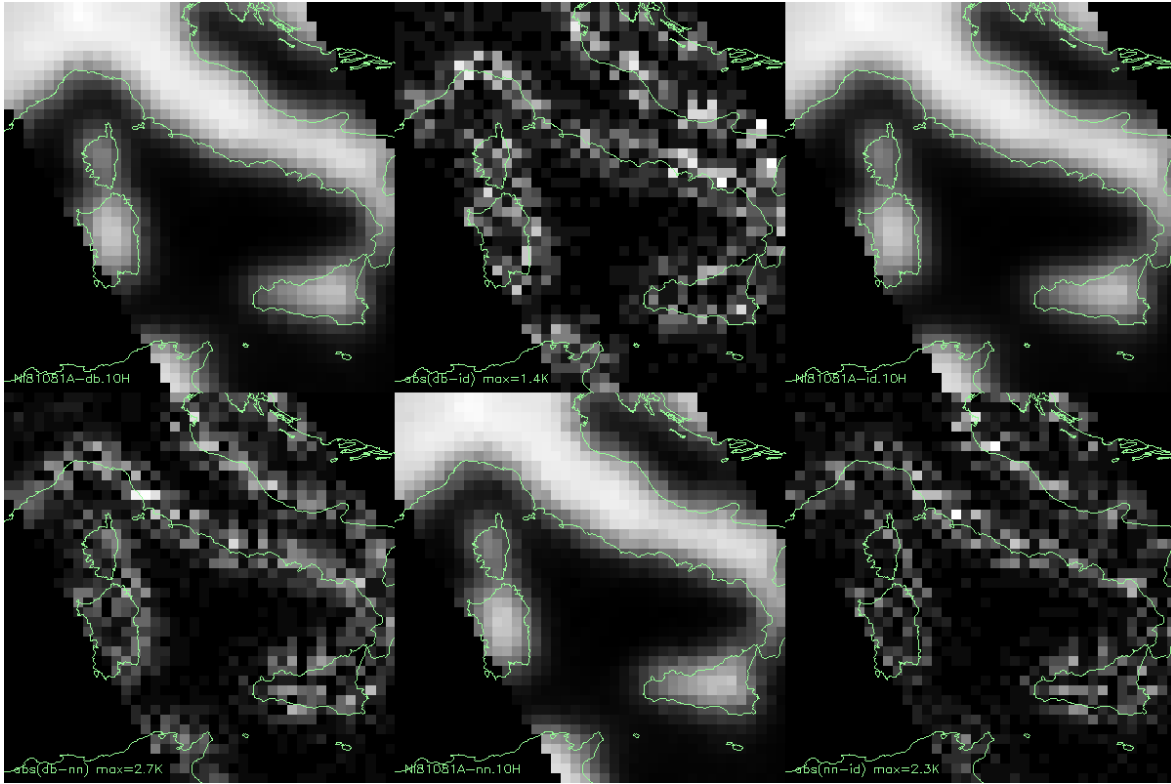
NI81081A.06H



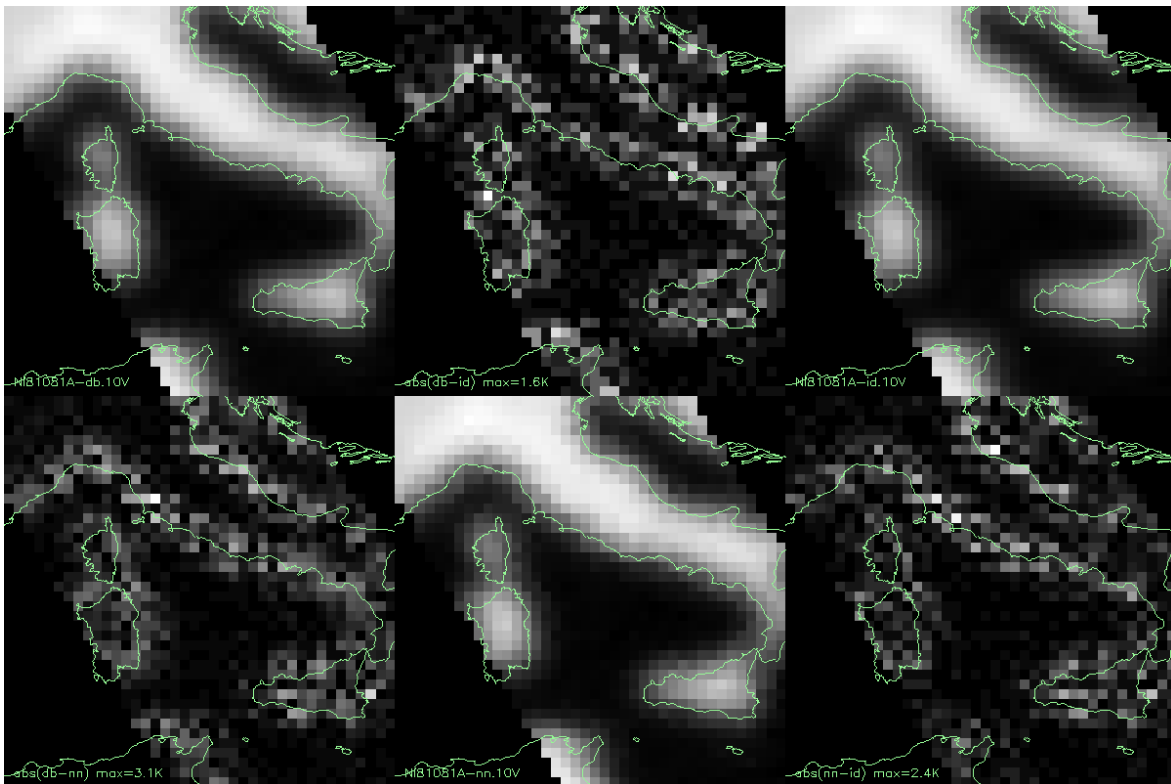
NI81081A.06V



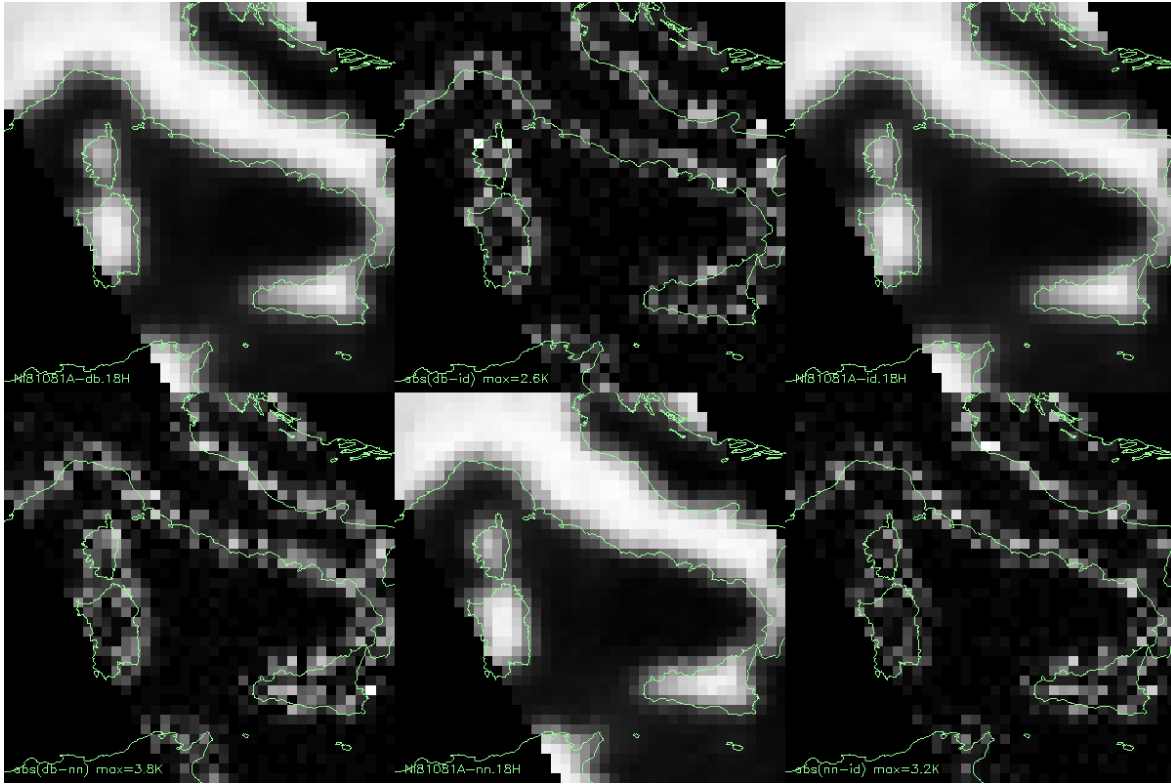
NI81081A.10H



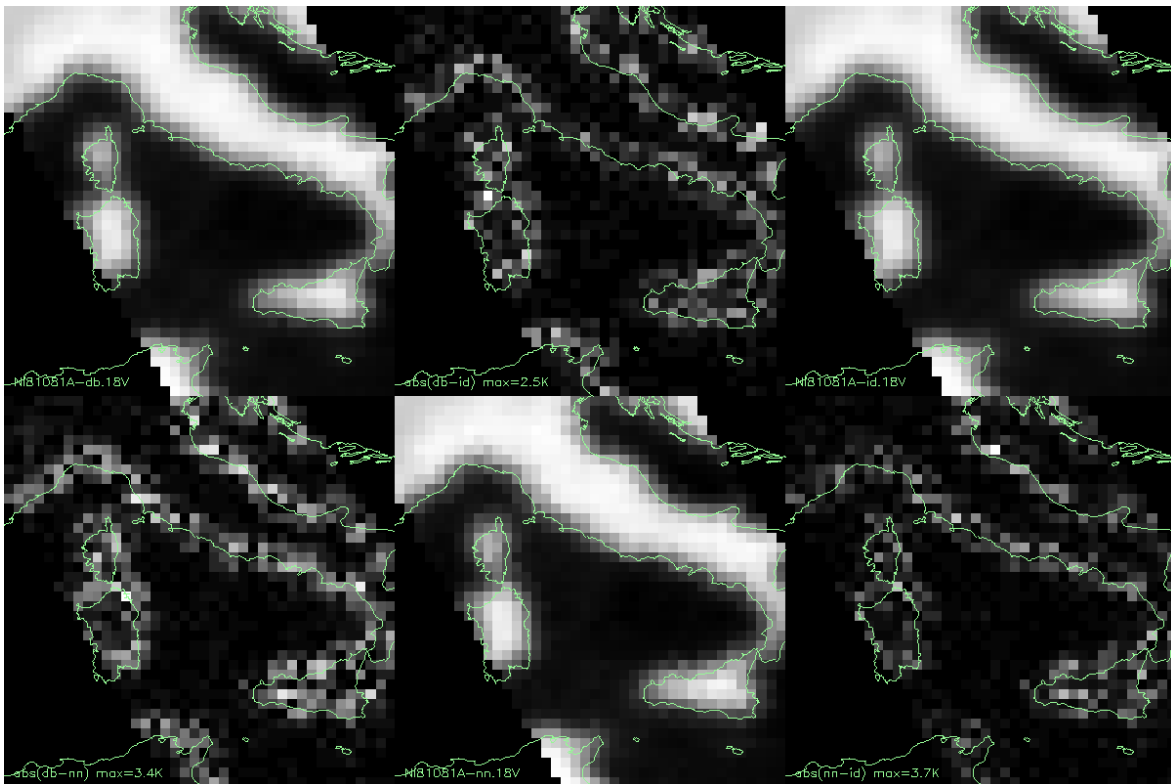
NI81081A.10V



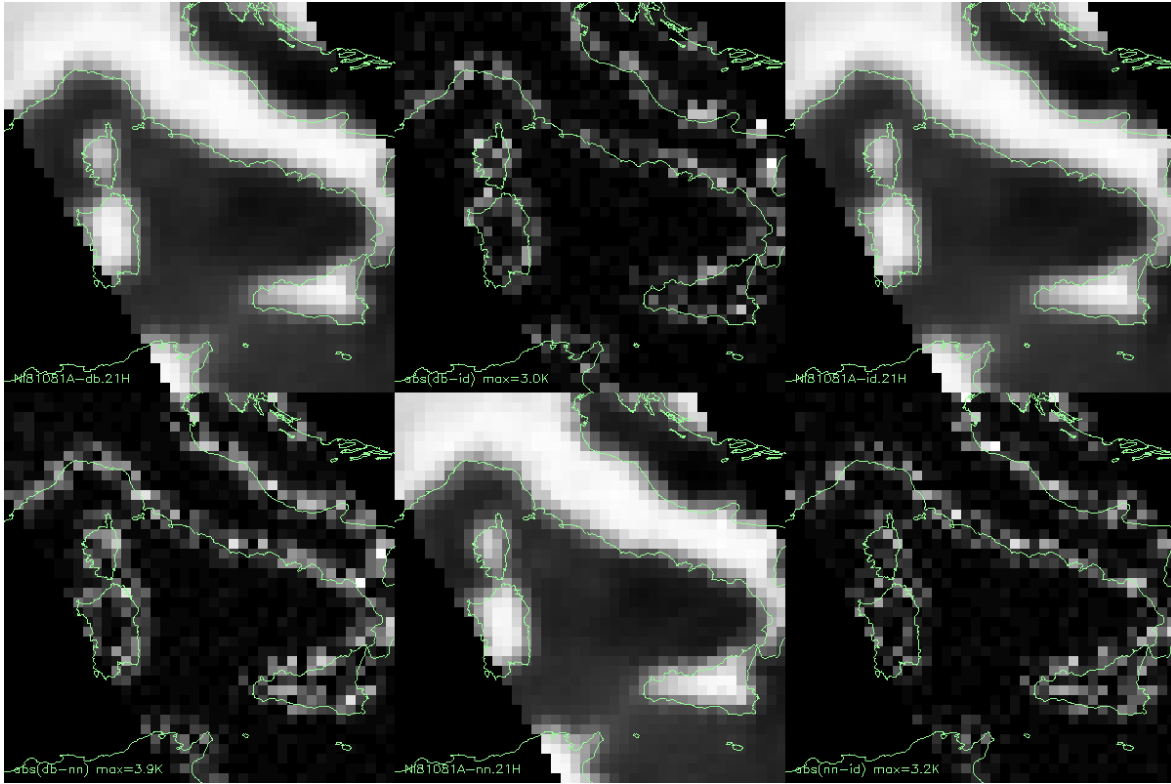
NI81081A.18H



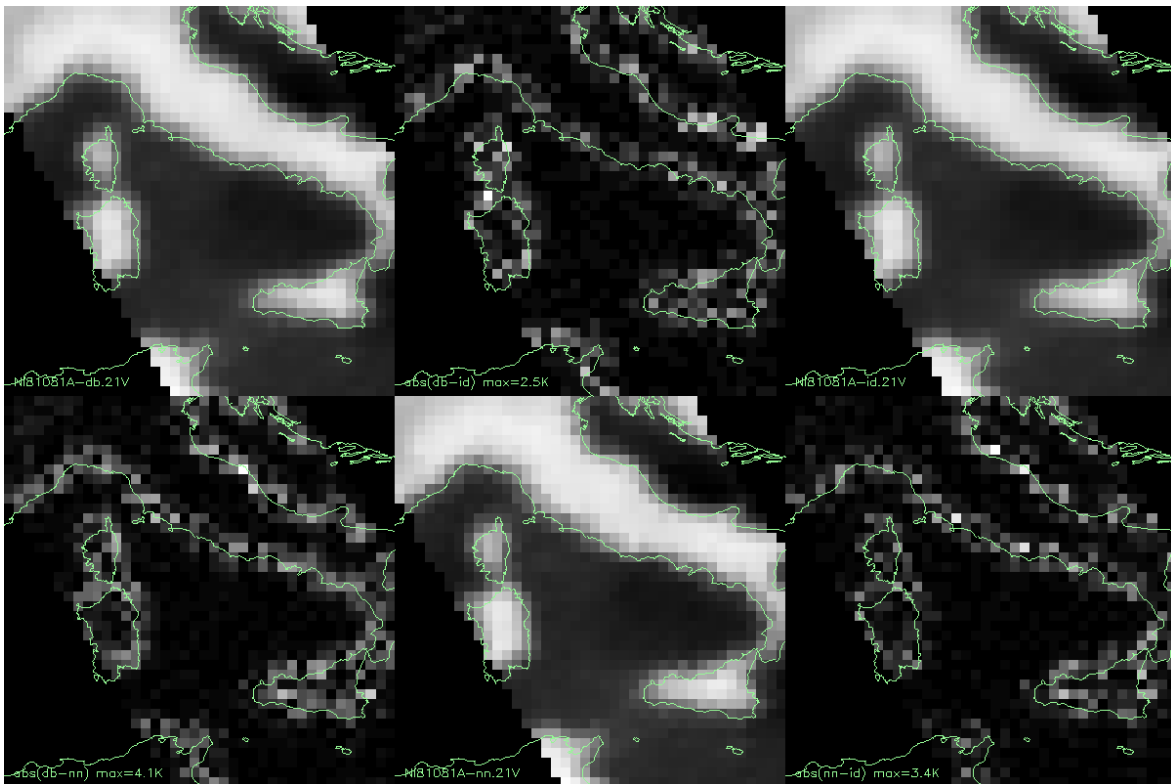
NI81081A.18V



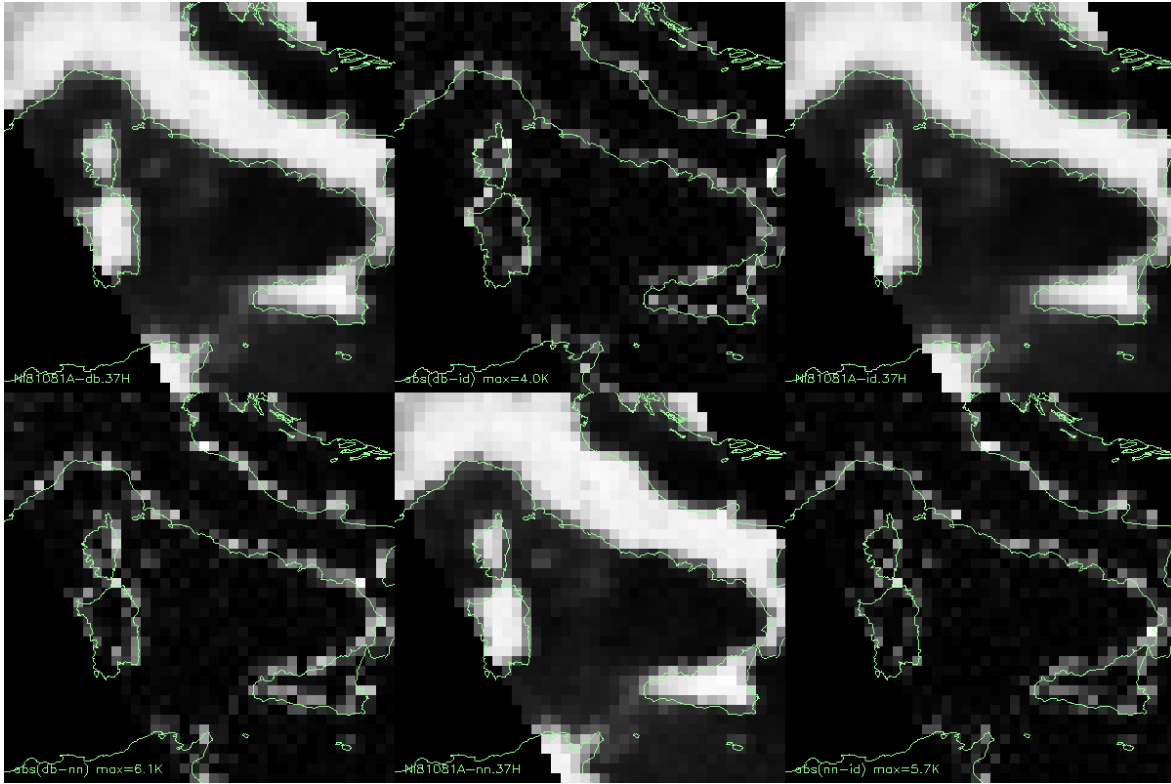
NI81081A.21H



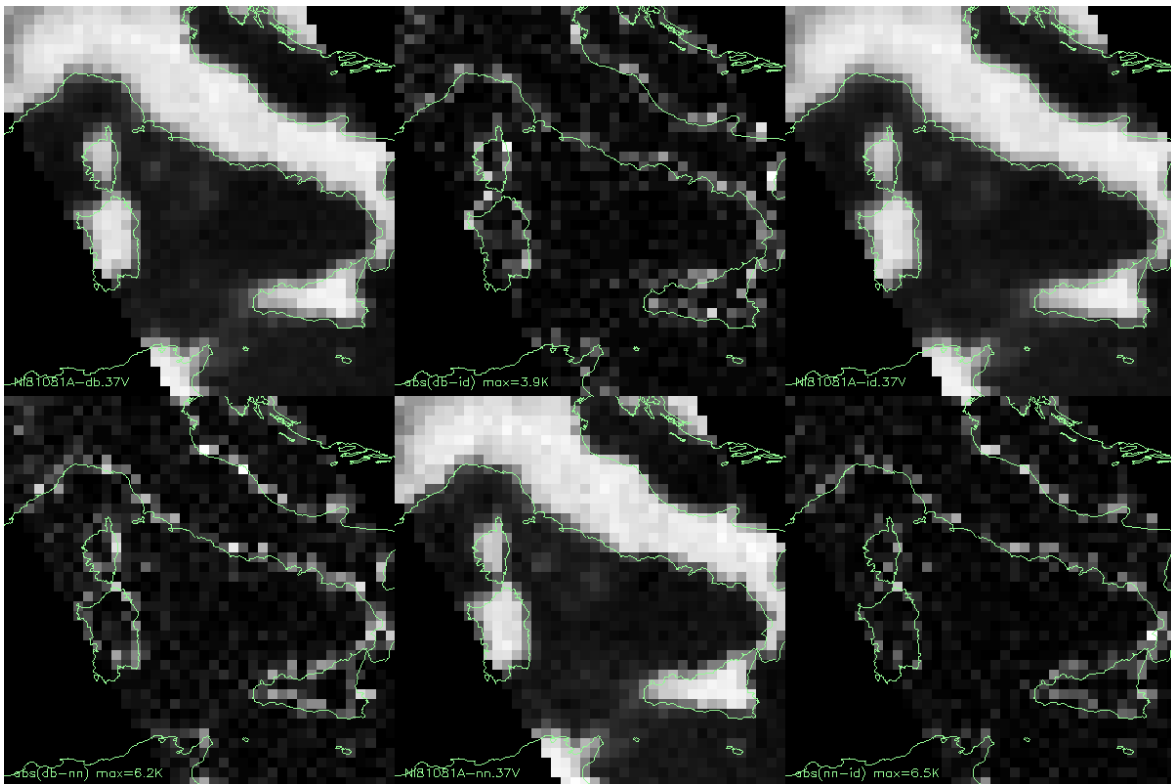
NI81081A.21V



NI81081A.37H



NI81081A.37V

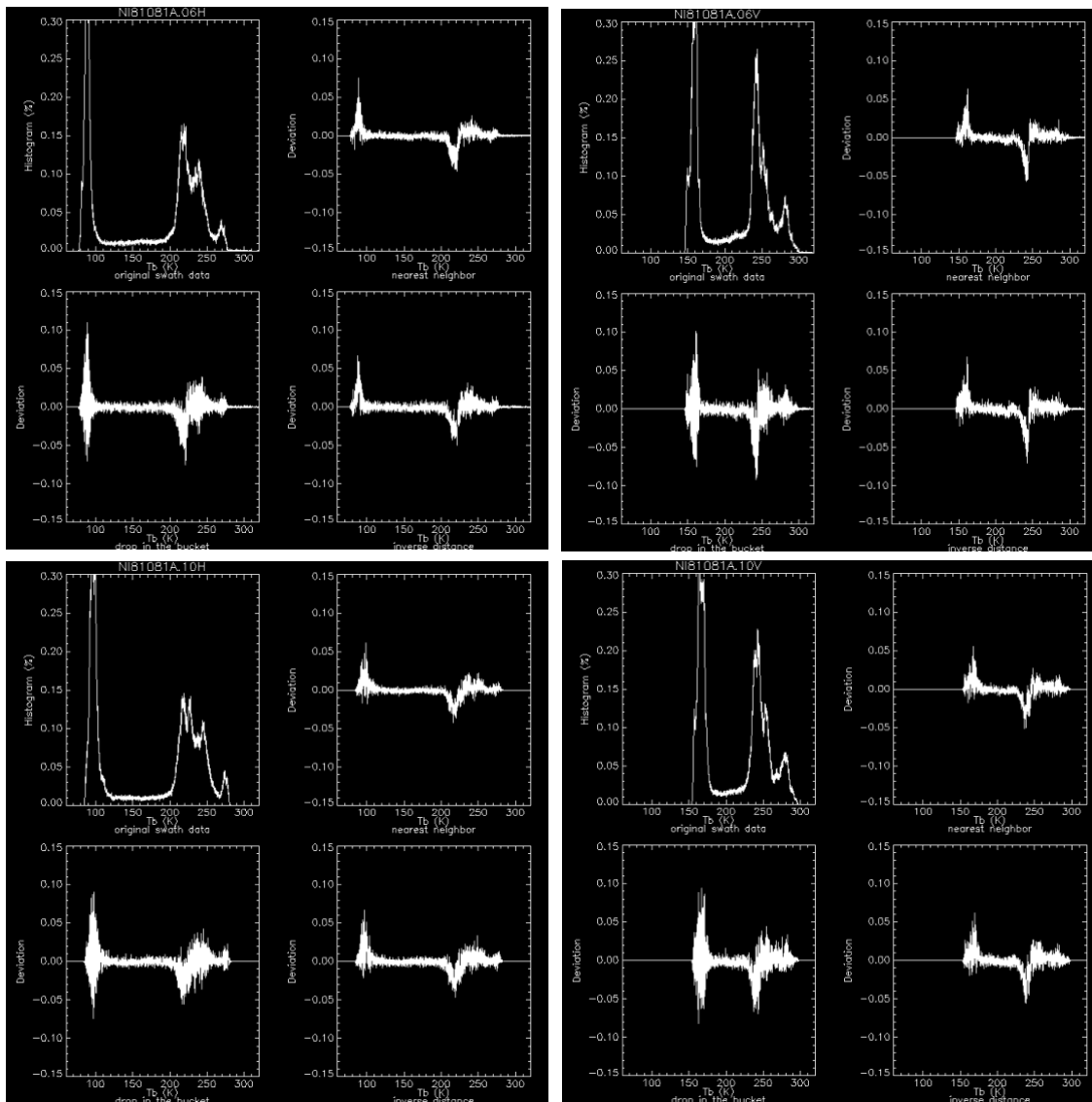


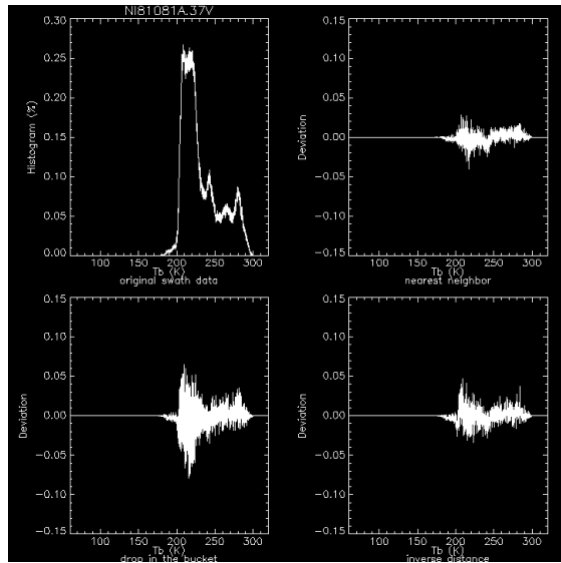
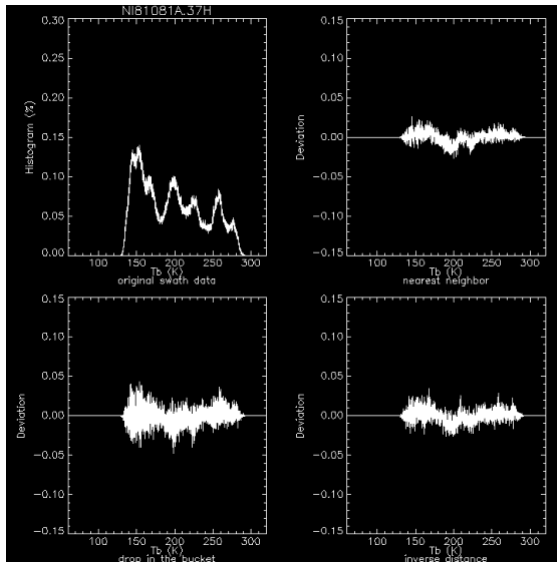
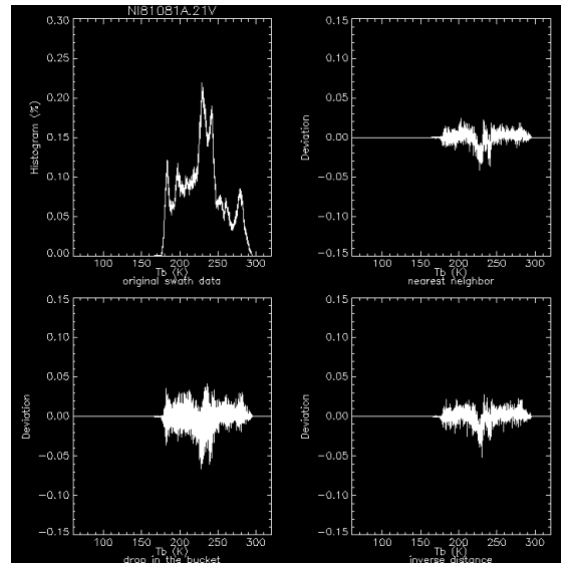
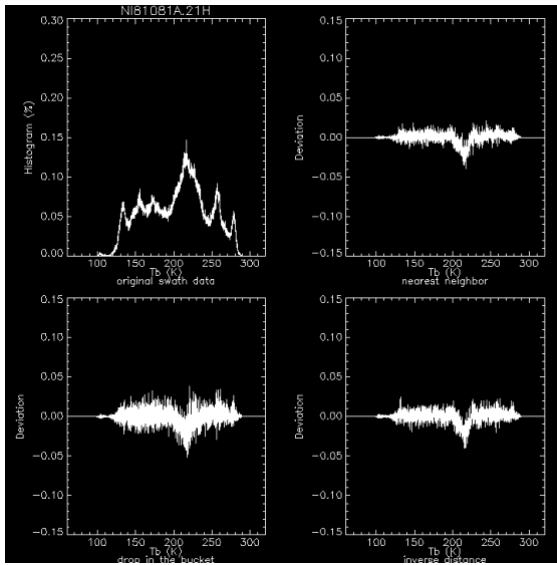
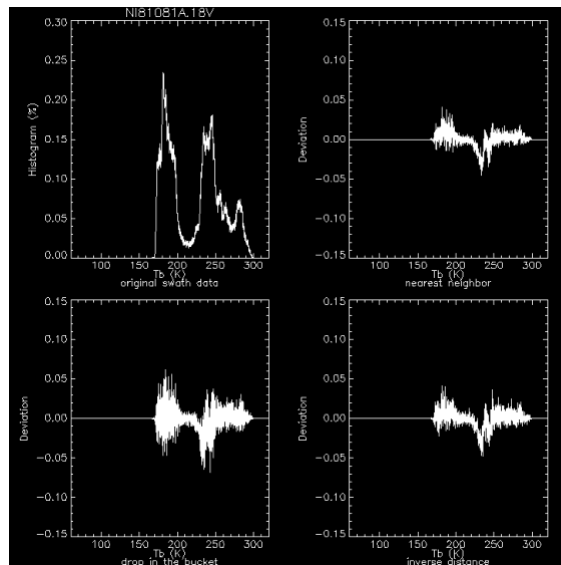
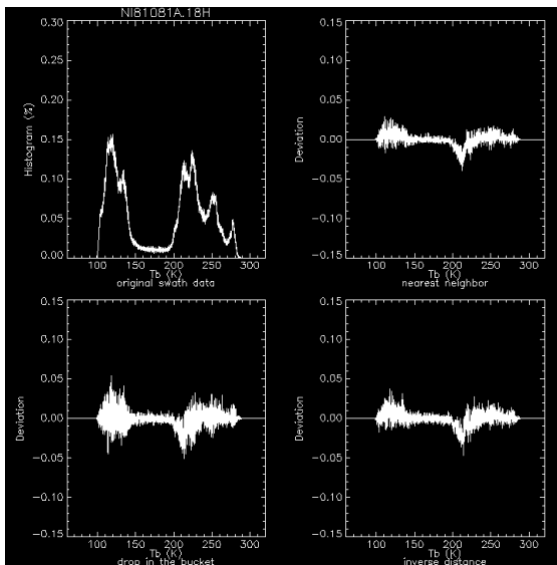
5 PLOTS AND INTERPOLATIONS

5.1 T_b Histograms

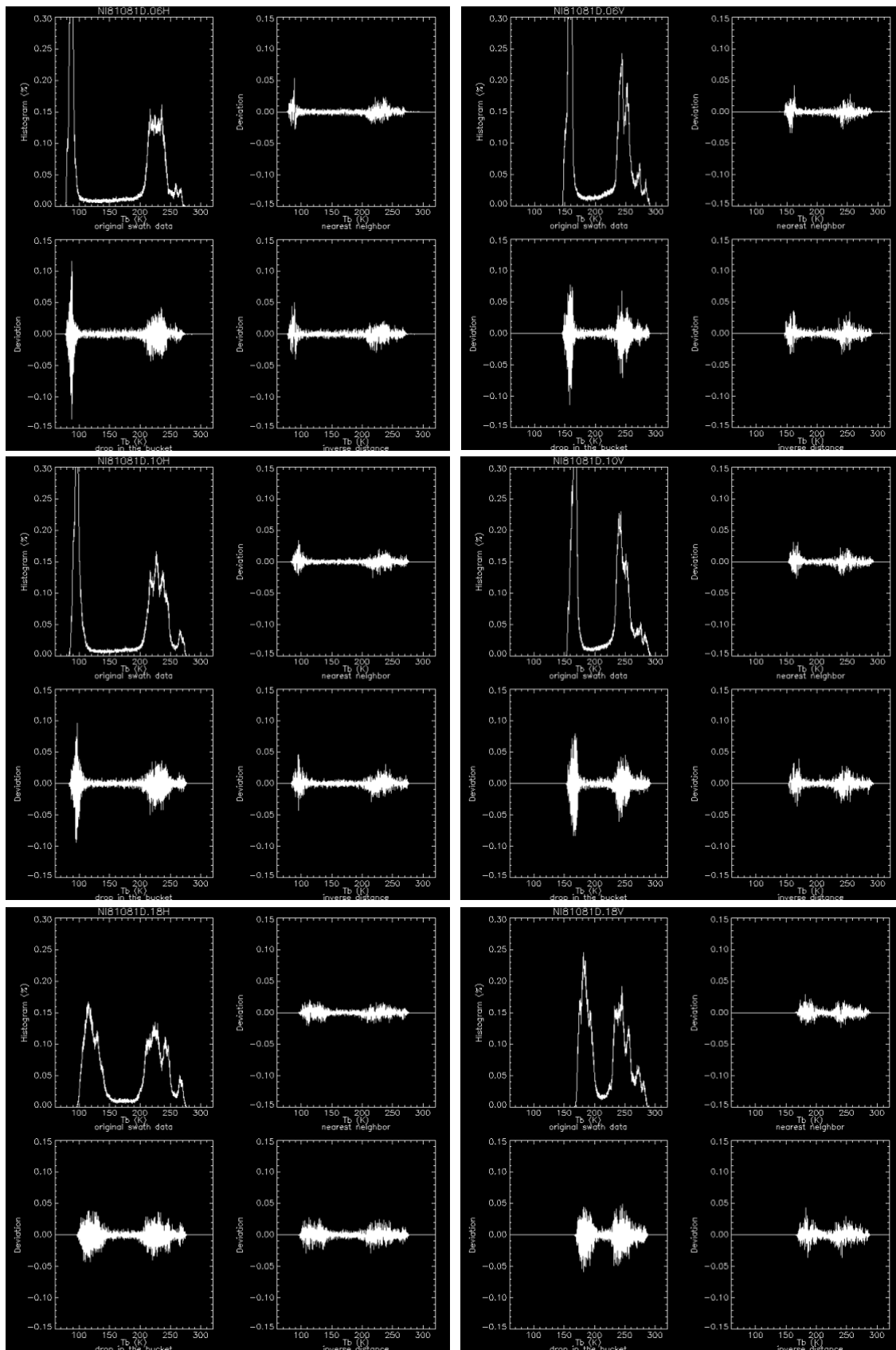
Each histogram plot shows the histogram of brightness temperatures for the original input data (all data that fall within the grid and satisfy the local time selection criteria). The other plots show where the histogram of the resampled data deviates from the original histogram. A shift in the peak values will be shown as a bulge in the difference plot.

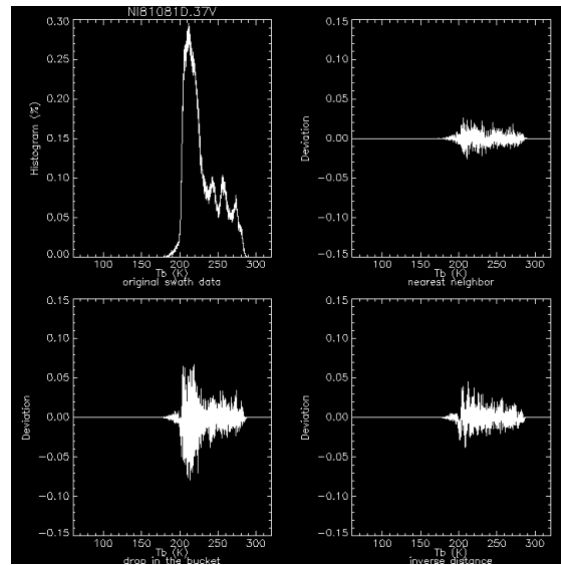
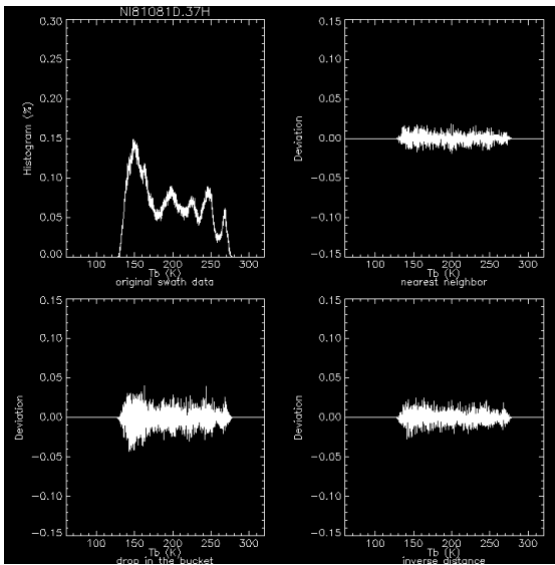
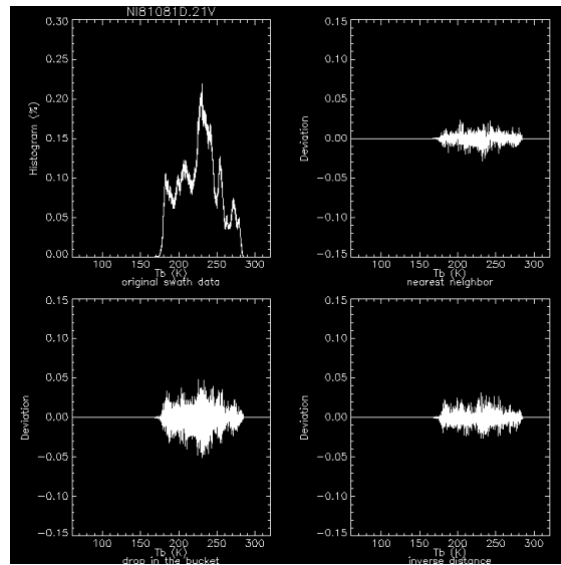
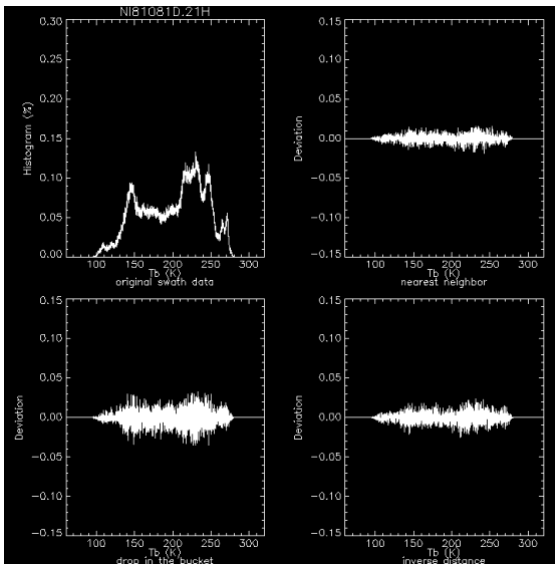
Northern Hemisphere, March 22, 1981, Ascending



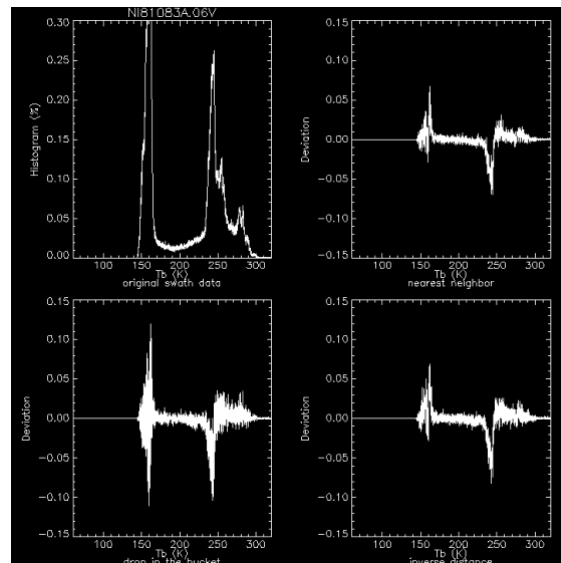
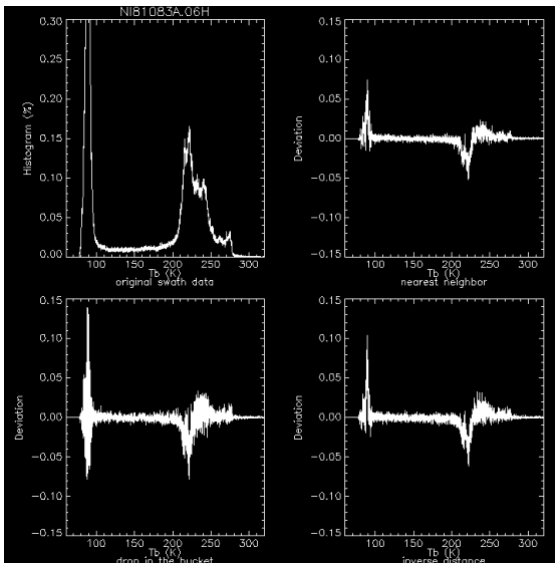


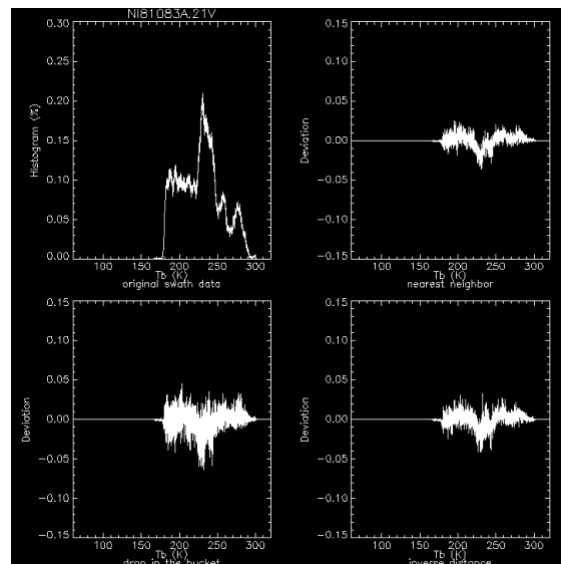
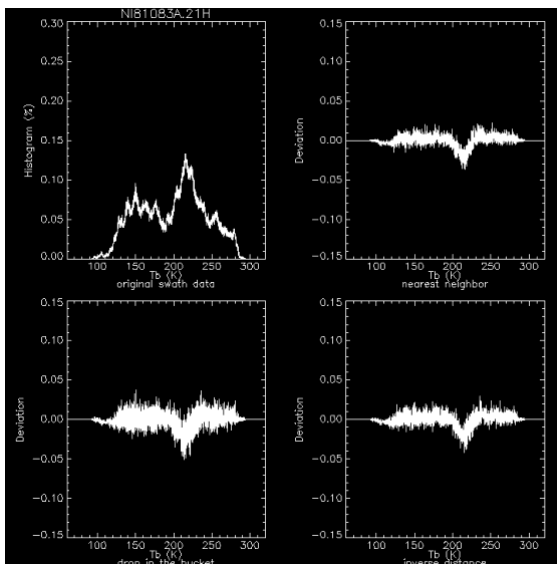
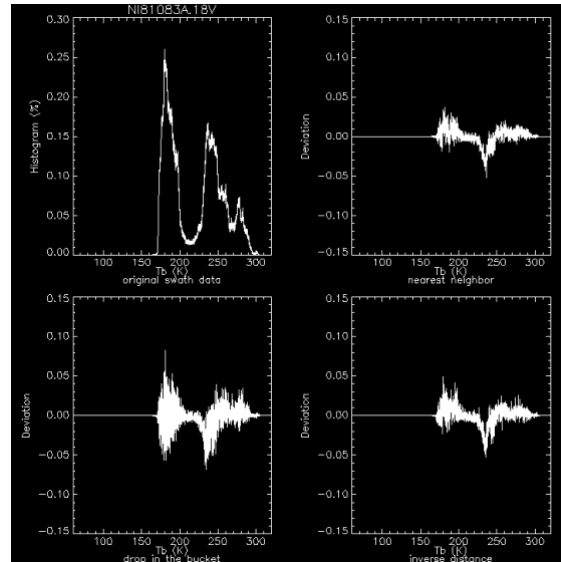
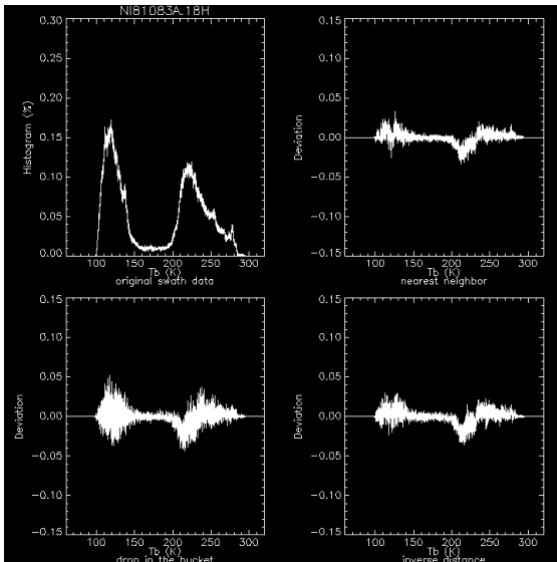
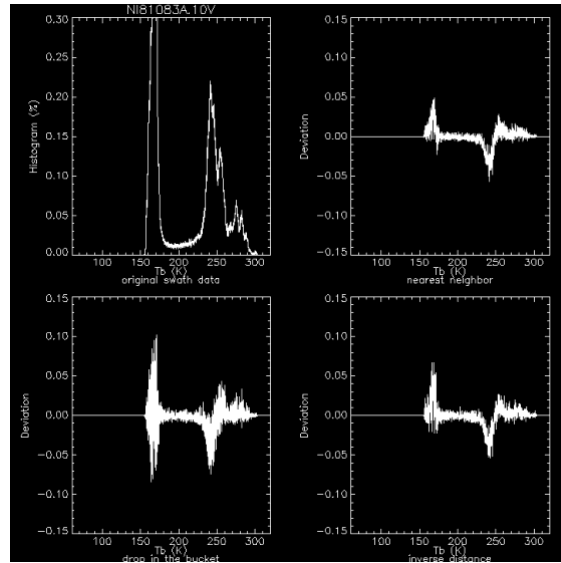
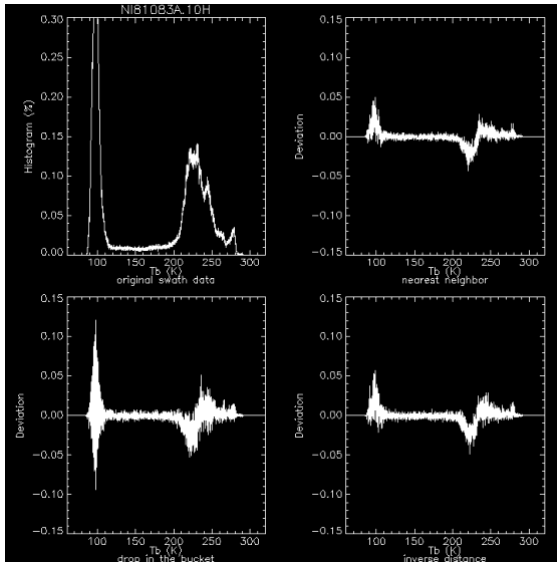
Northern Hemisphere, March 22, 1981, Descending

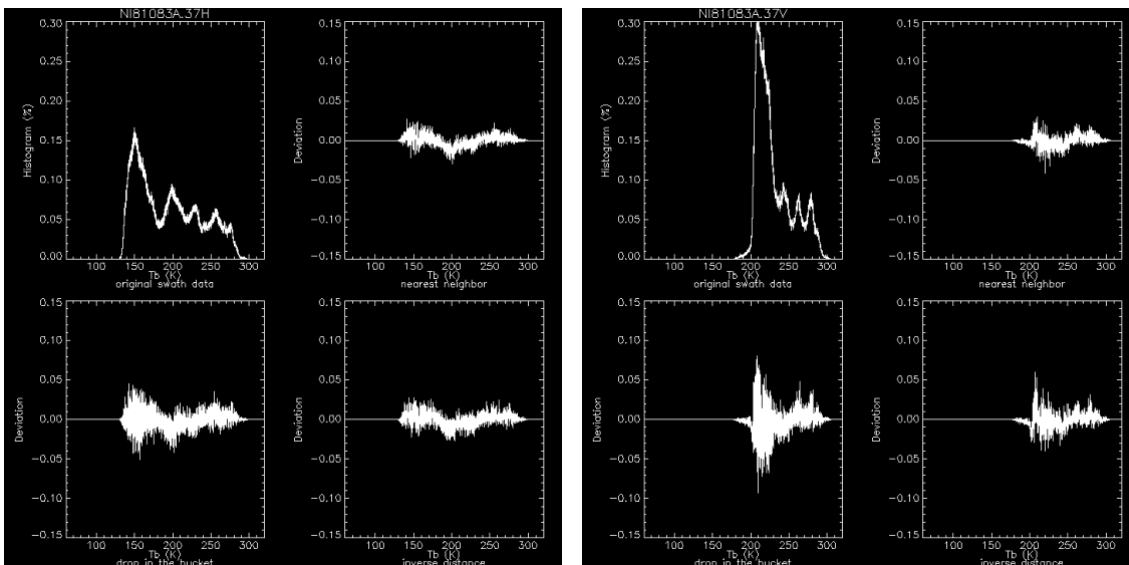




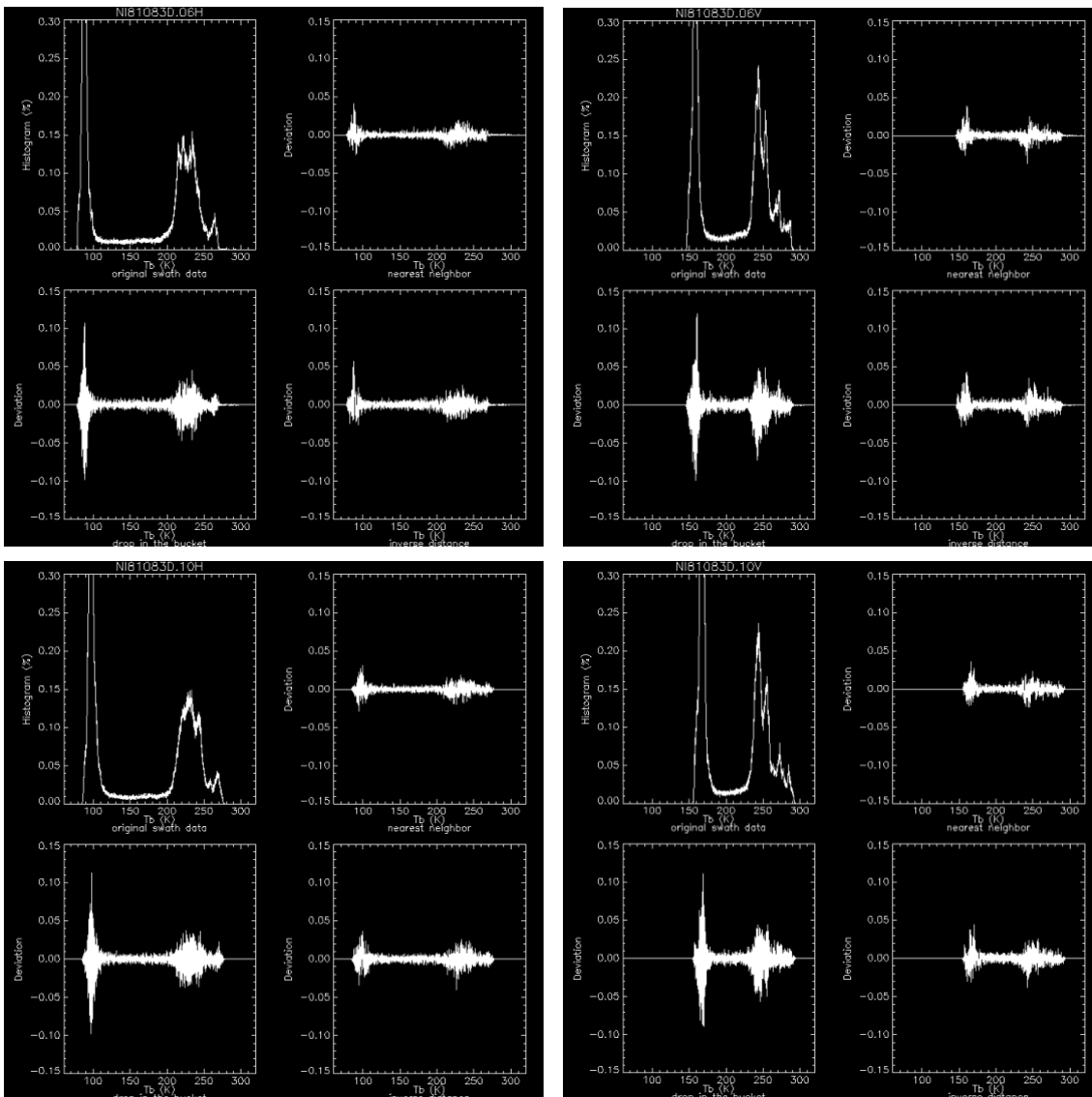
Northern Hemisphere, March 24, 1981, Ascending

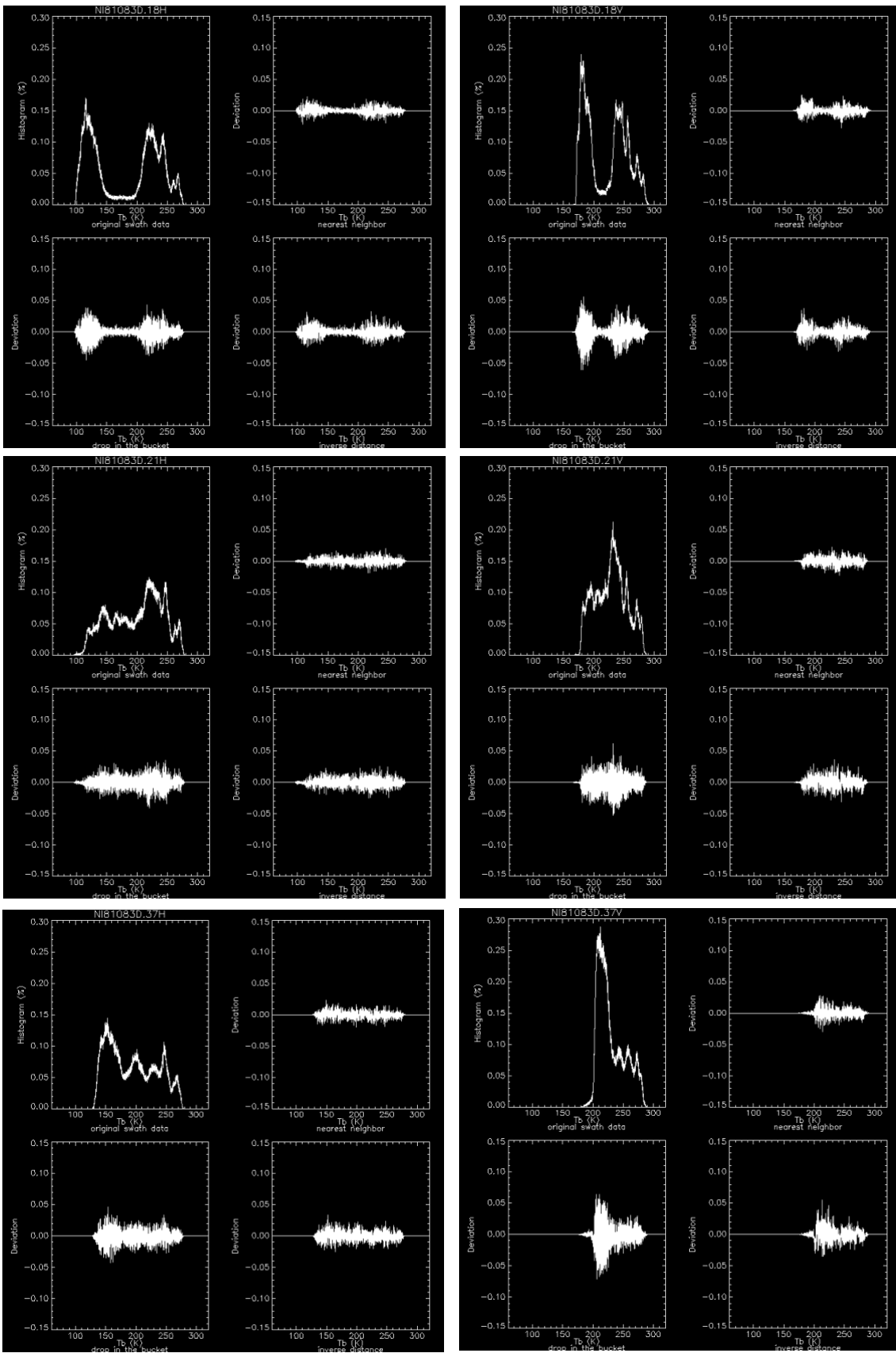




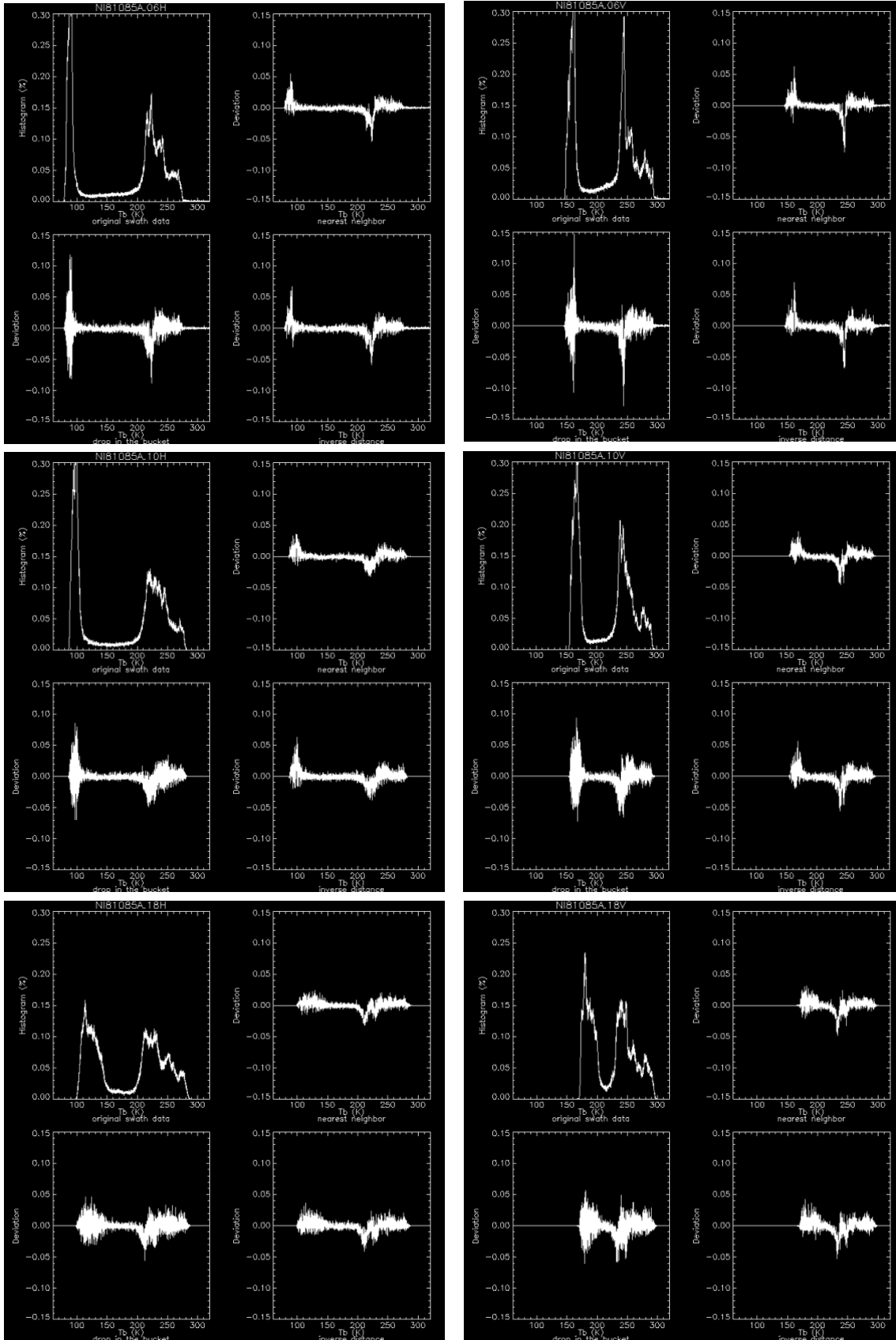


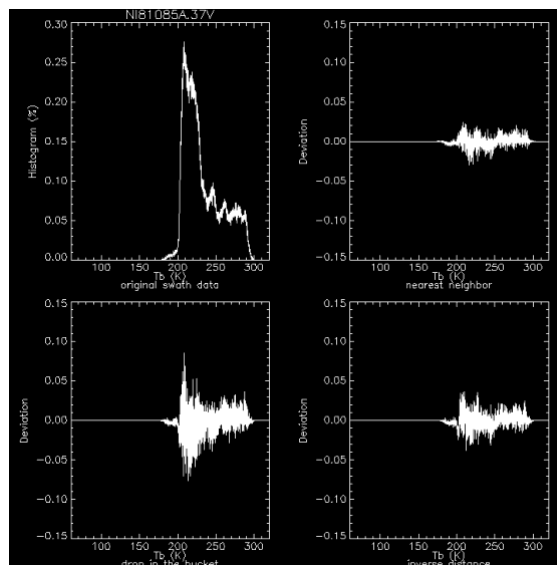
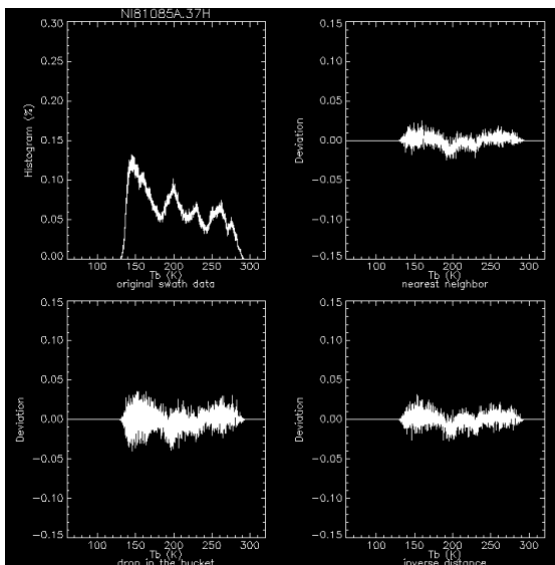
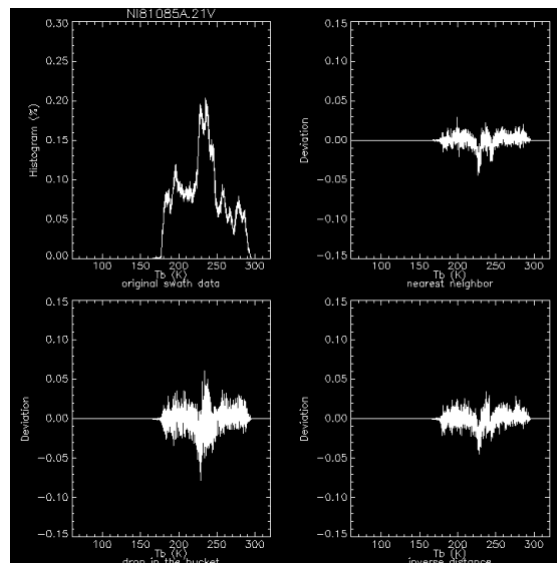
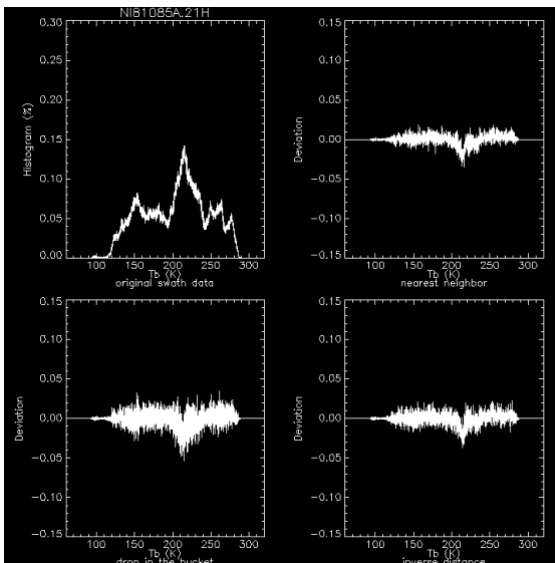
Northern Hemisphere, March 24, 1981, Descending



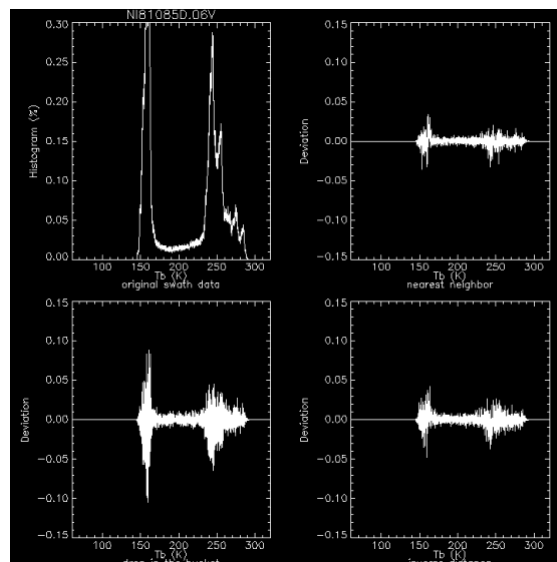
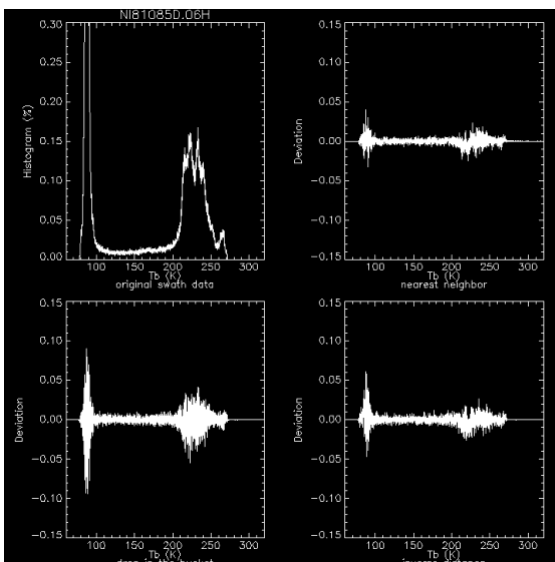


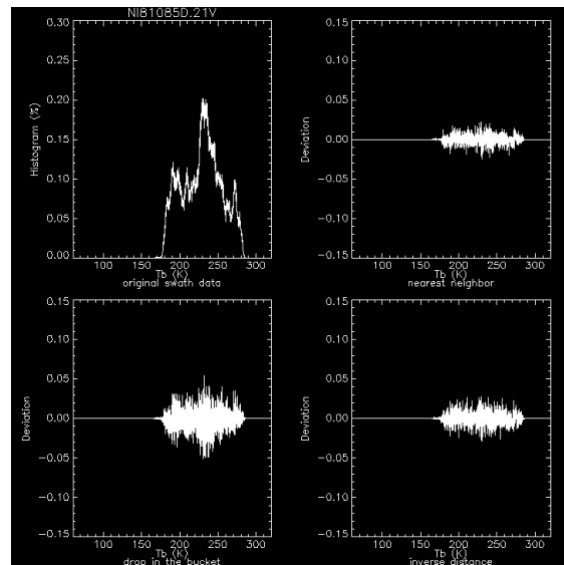
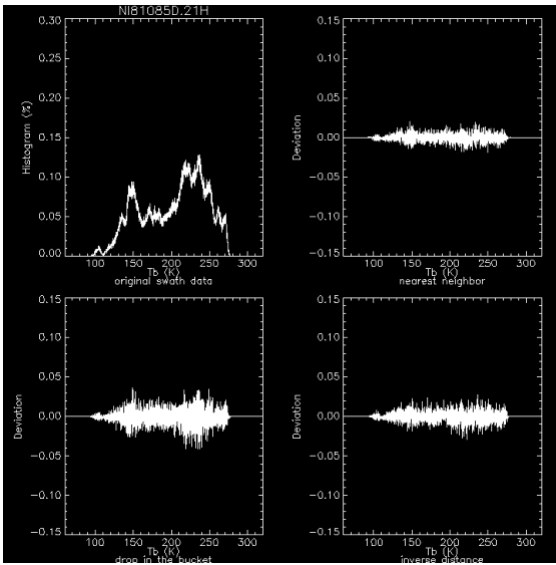
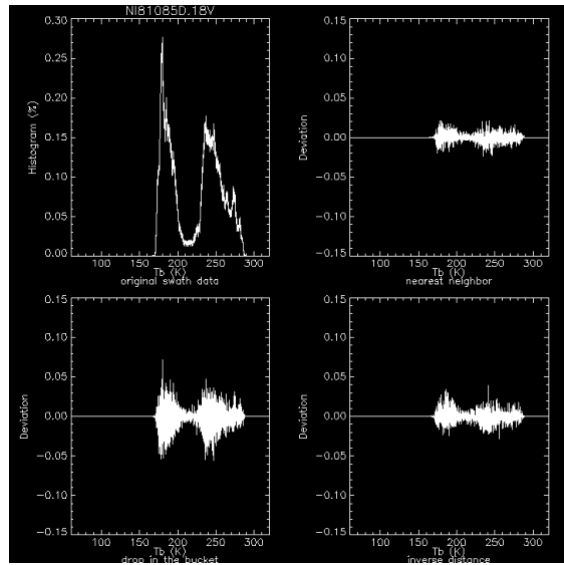
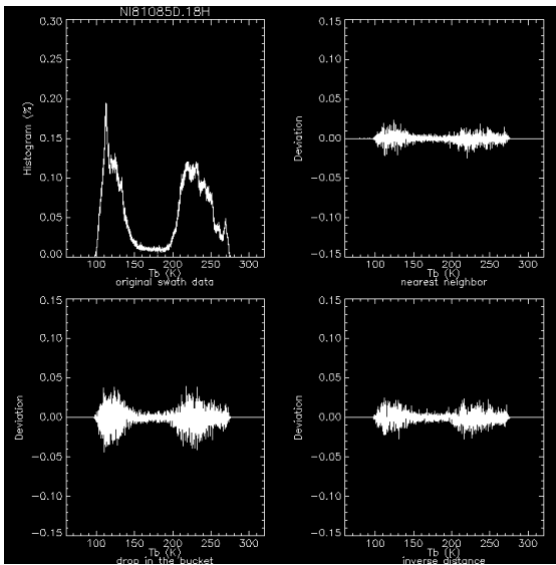
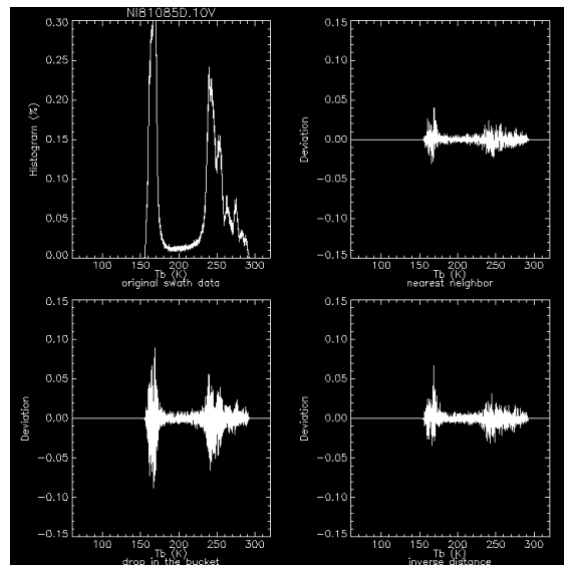
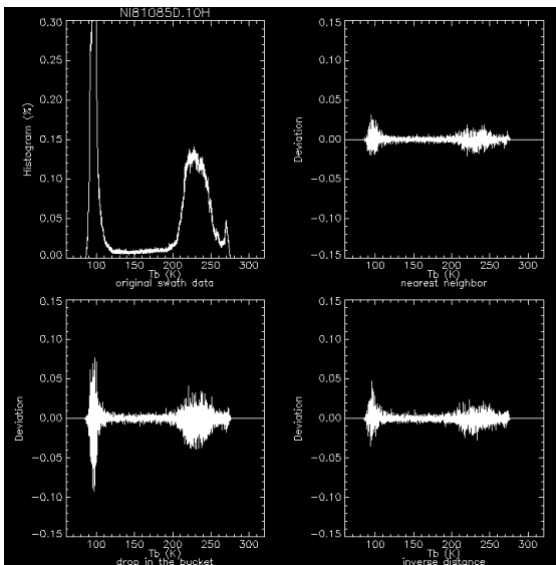
Northern Hemisphere, March 26, 1981, Ascending

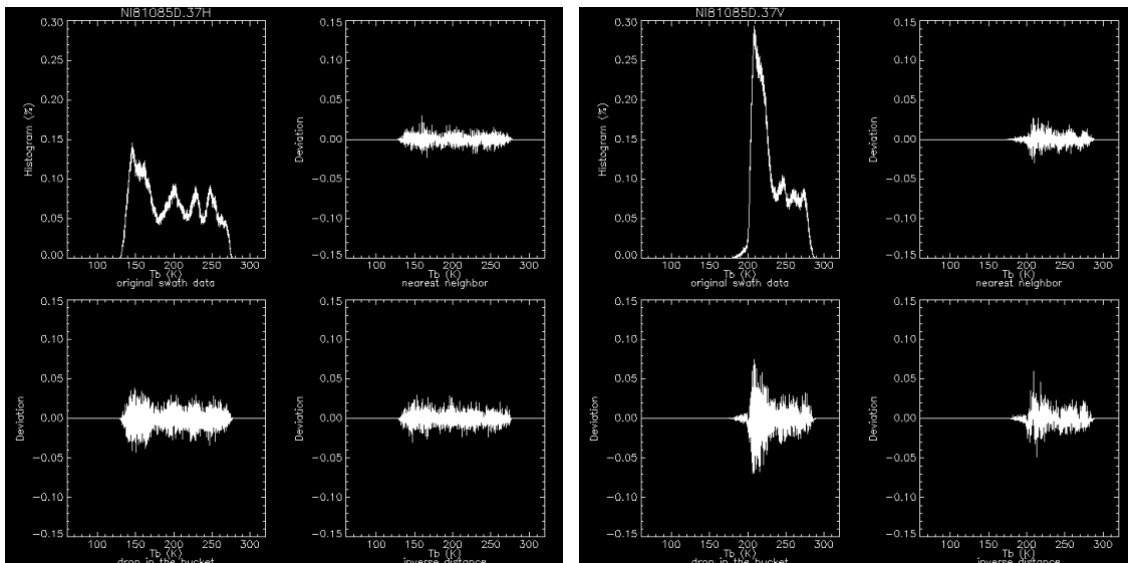




Northern Hemisphere, March 26, 1981, Descending



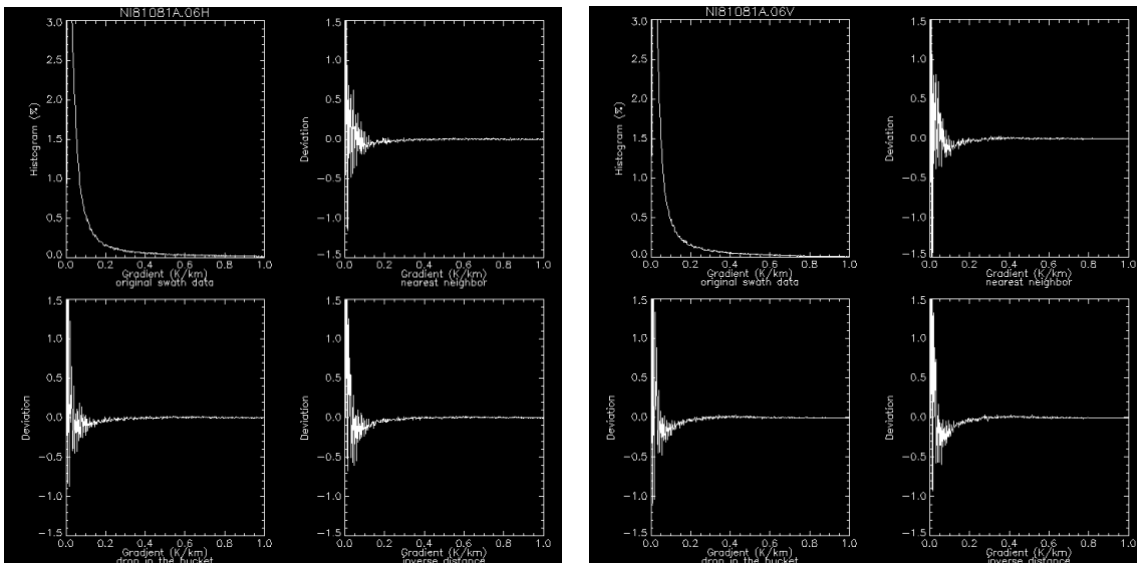


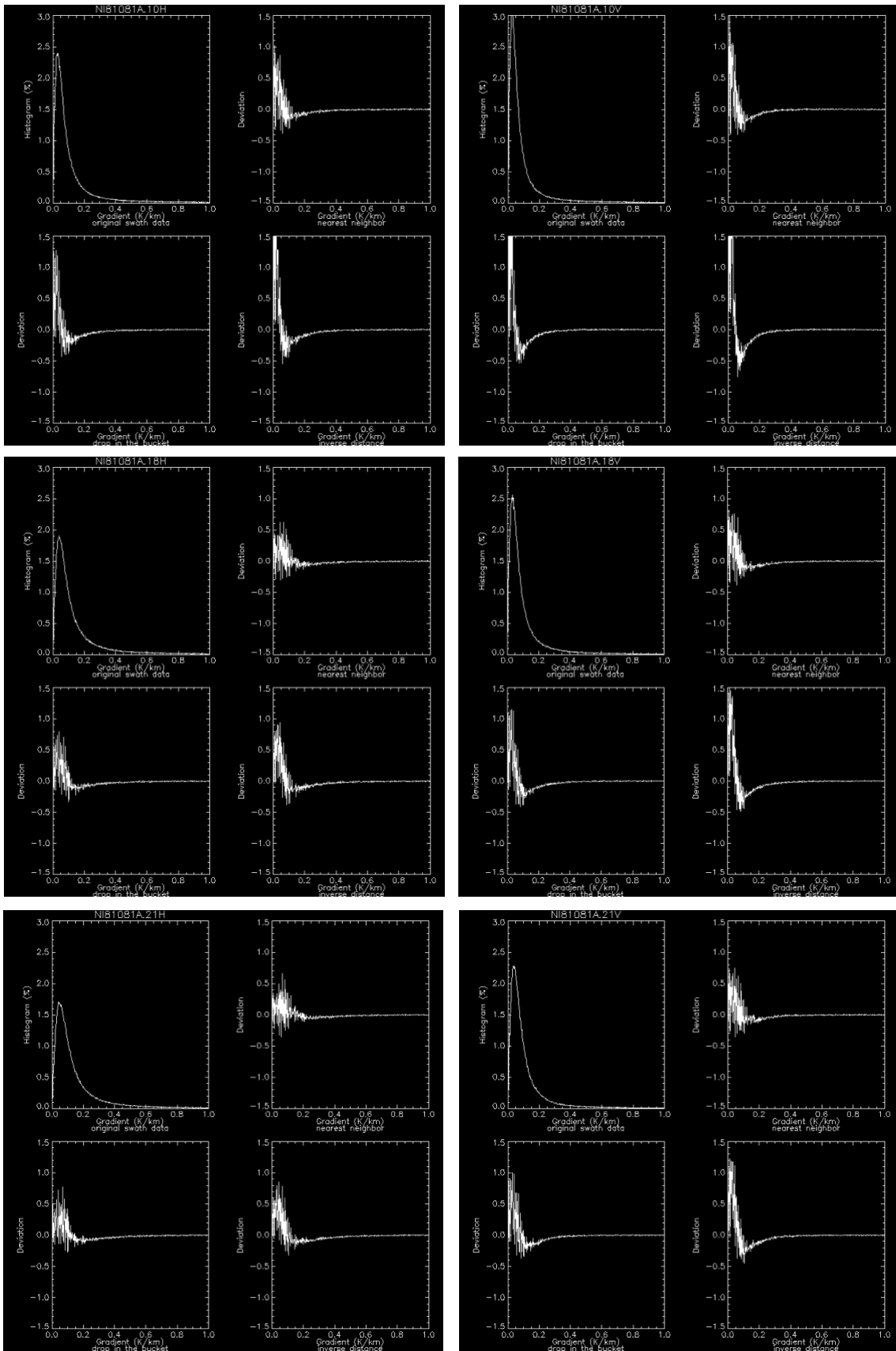


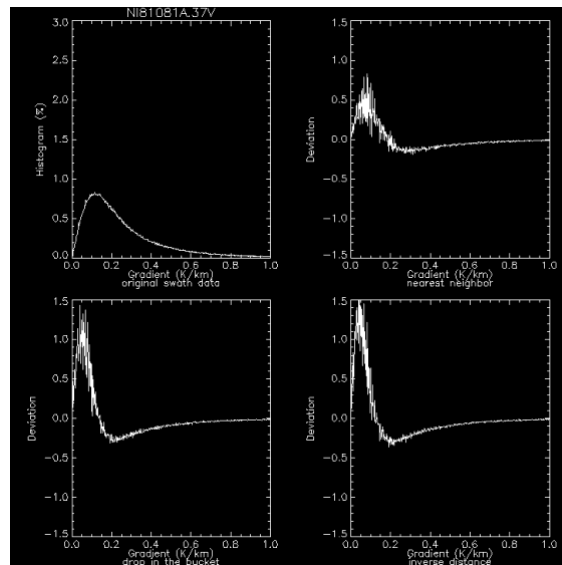
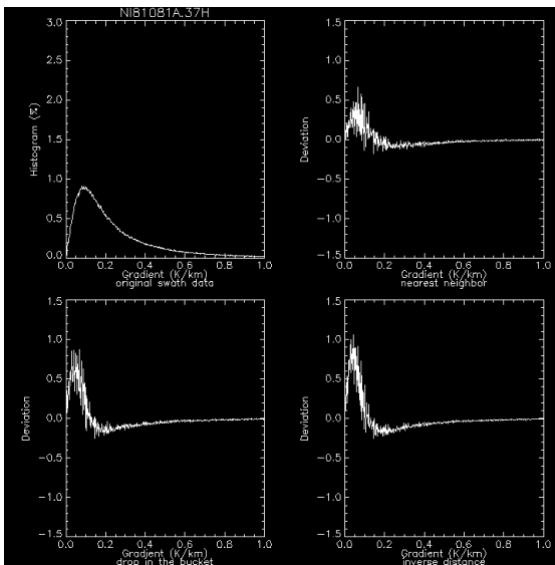
5.2 T_b Gradients

Each gradient plot shows the histogram of brightness temperature gradients for the original input data (all data that fall within the grid and satisfy the local time selection criteria). The other plots show where the histogram of gradients in the resampled data deviates from the original. An increase in low gradients indicates smoothing. An increase in high gradients indicates noise.

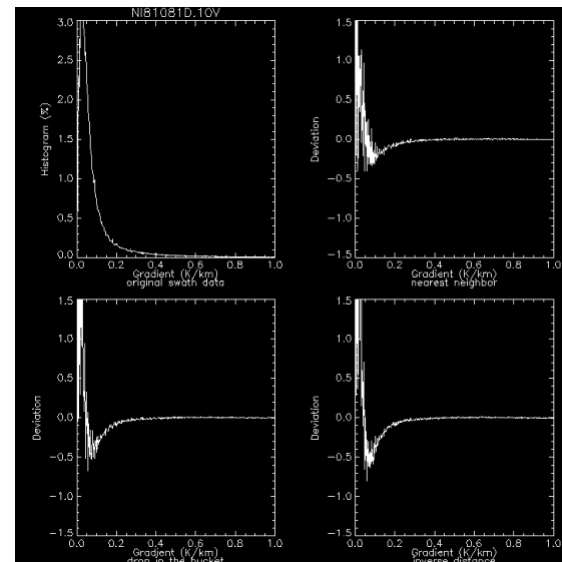
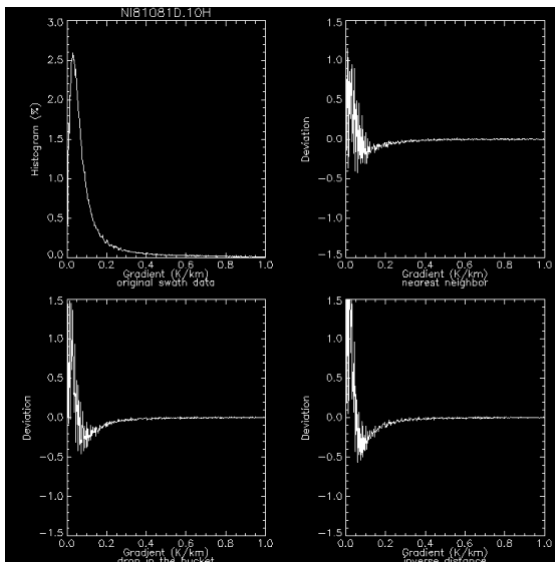
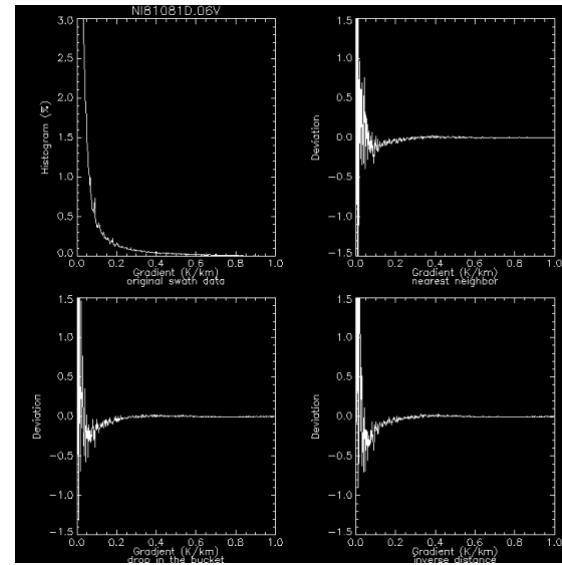
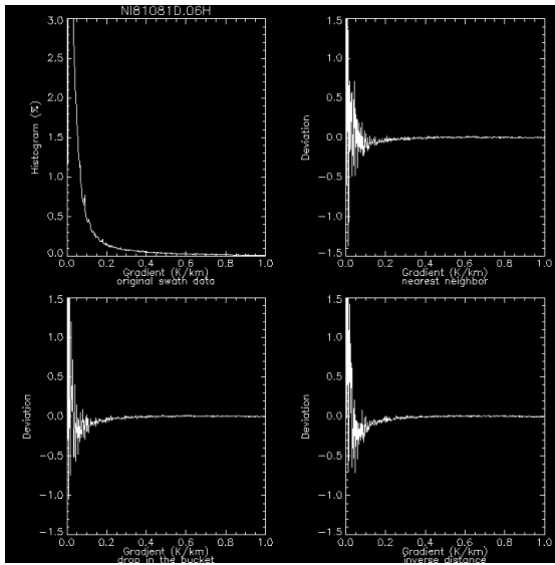
Northern Hemisphere, March 22, 1981, Ascending

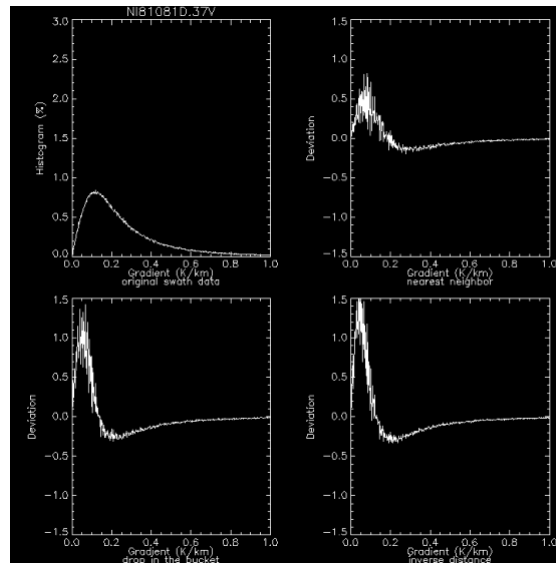
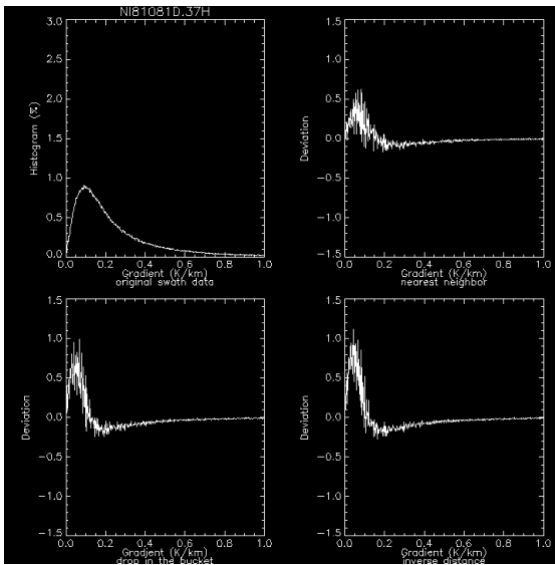
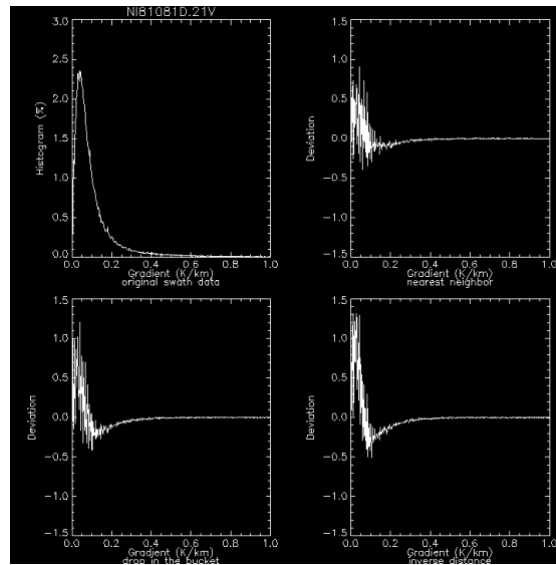
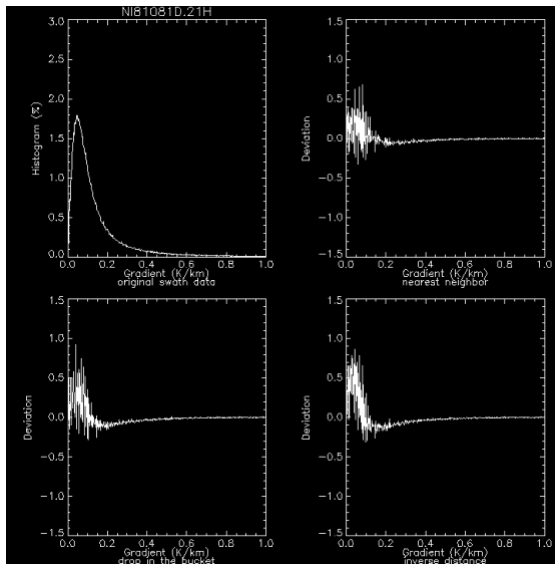
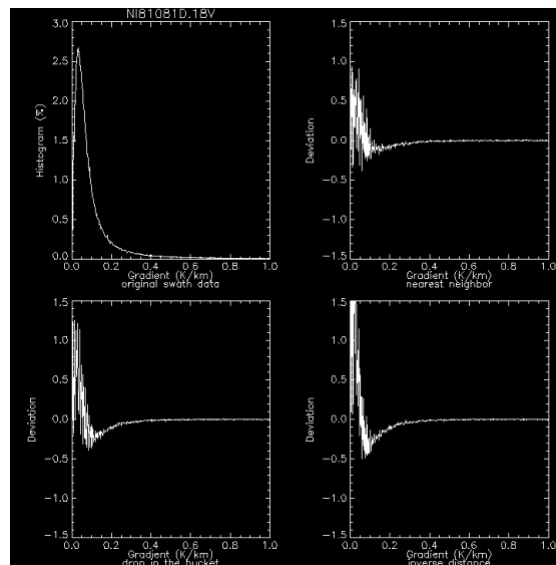
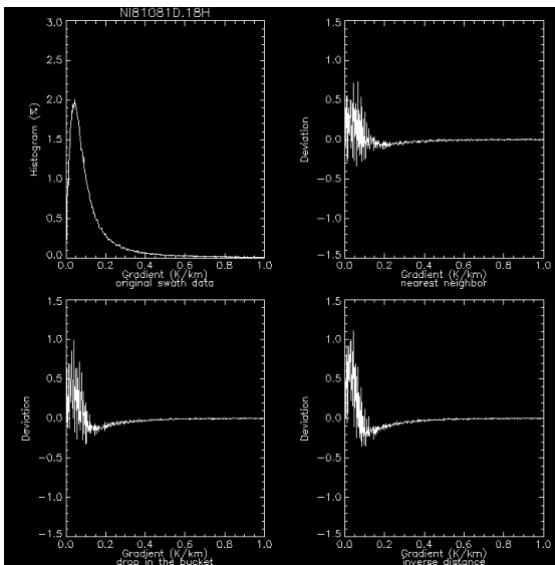




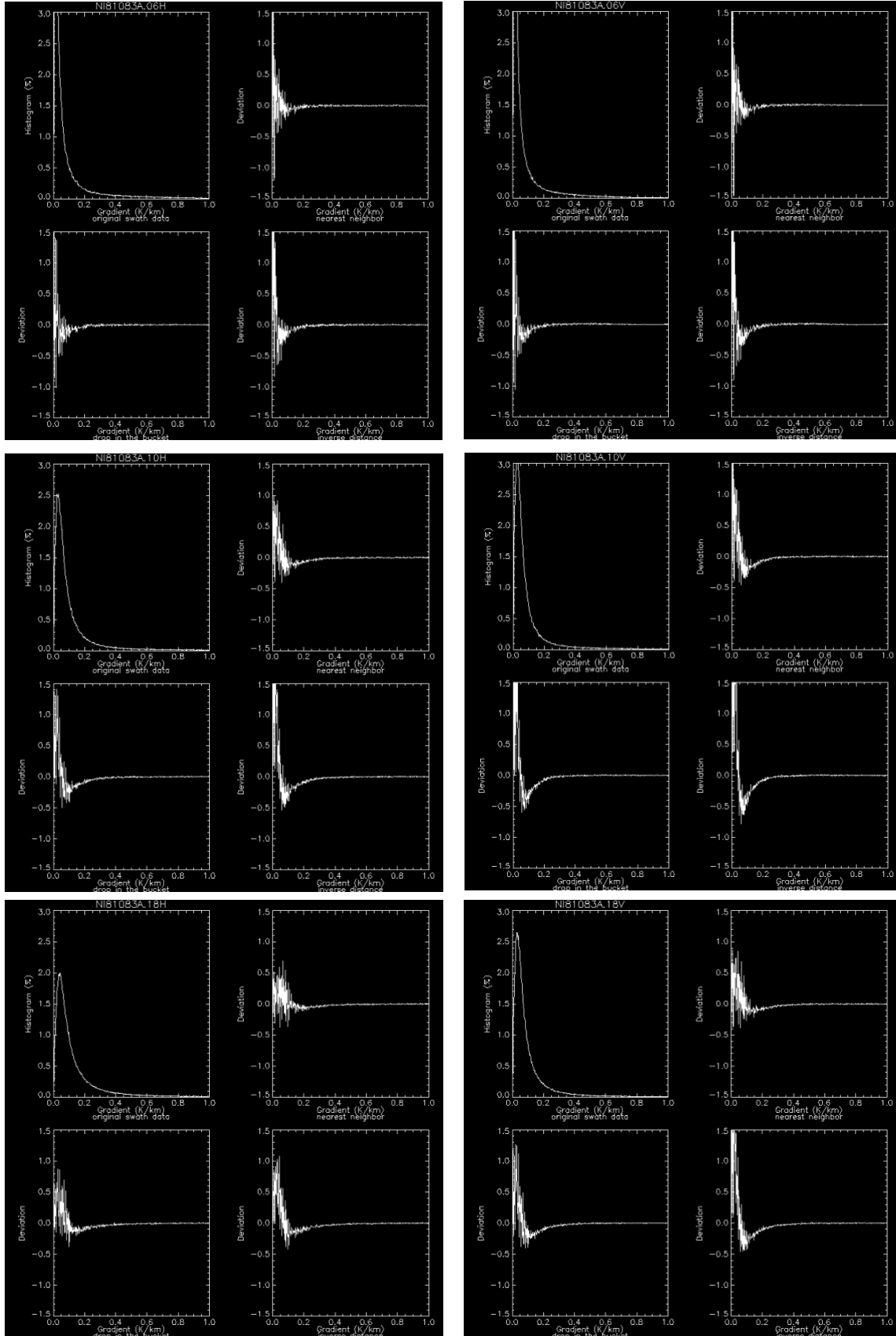


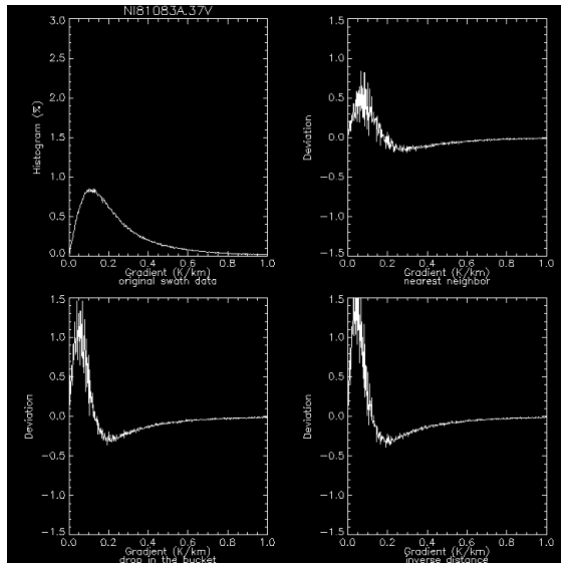
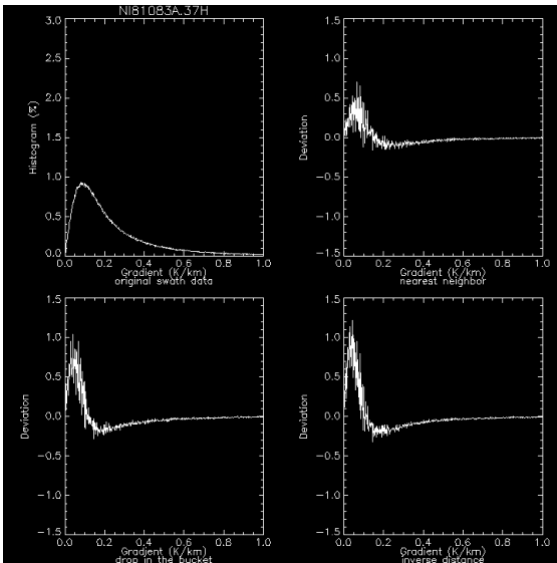
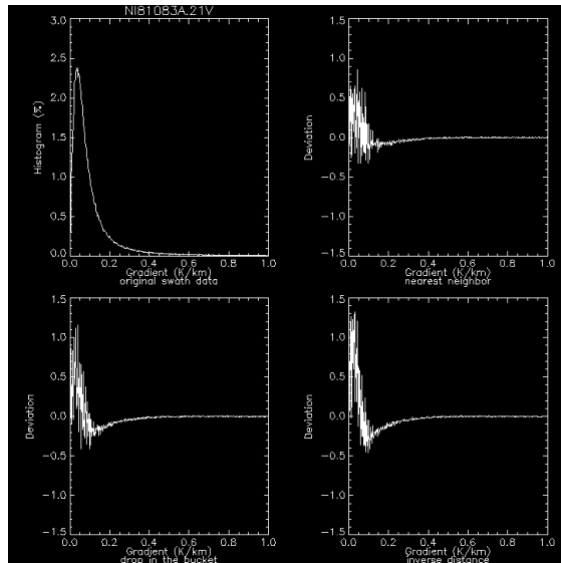
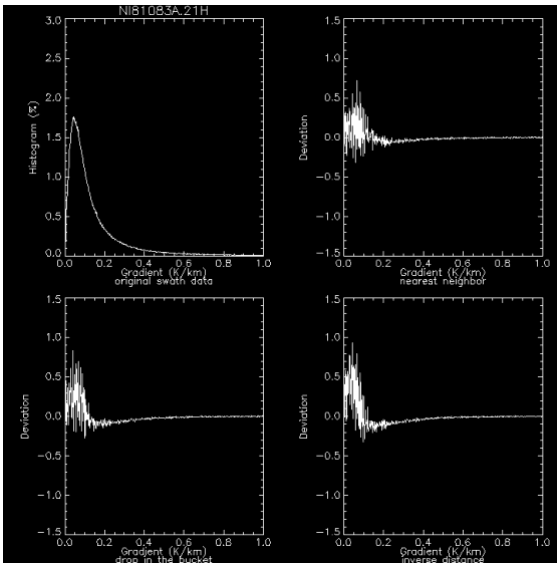
Northern Hemisphere, March 22, 1981, Descending



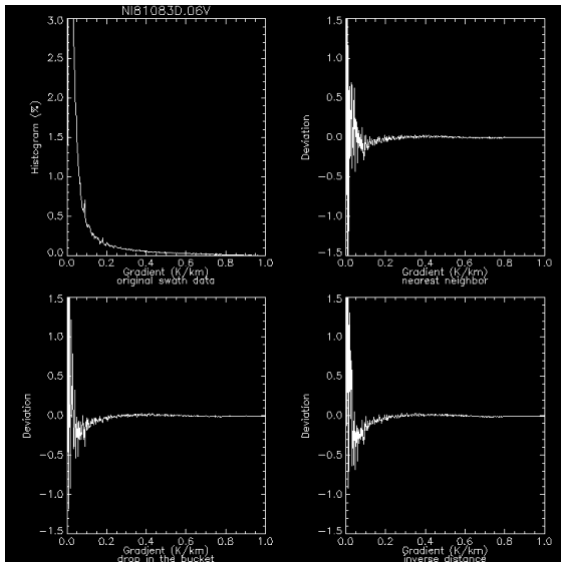
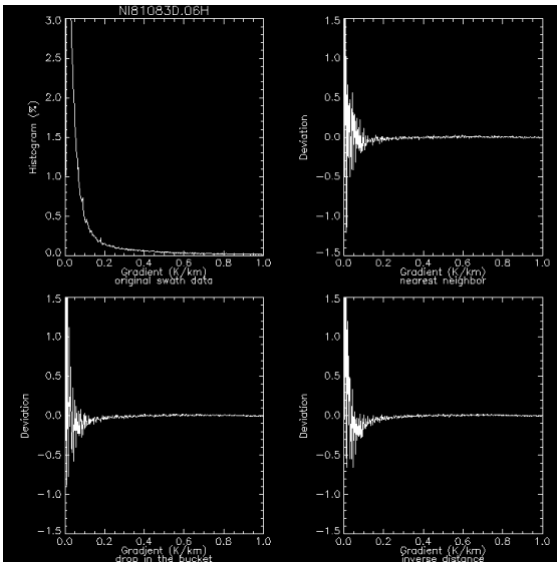


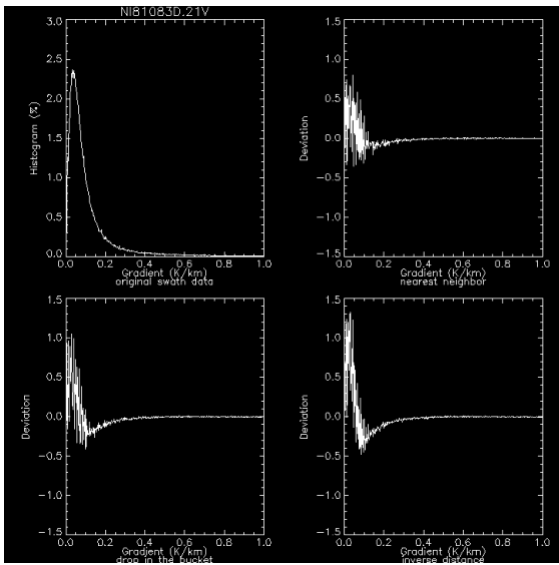
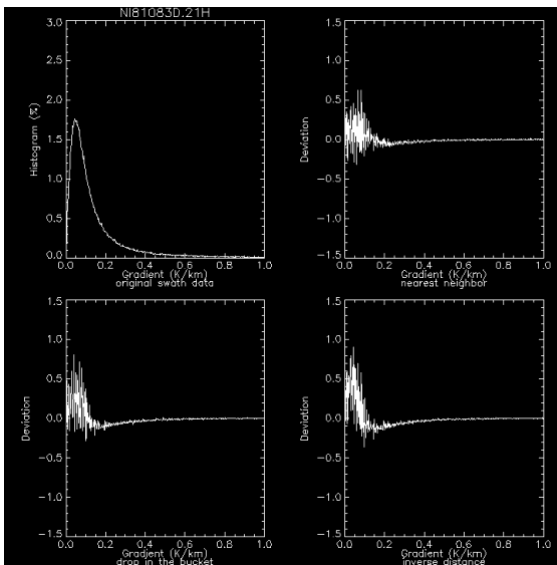
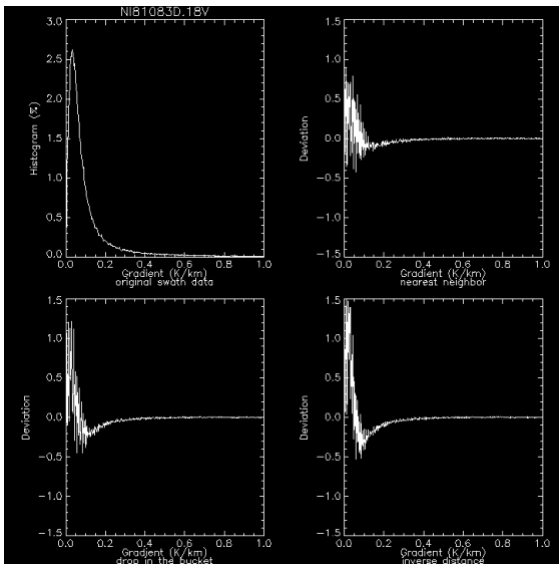
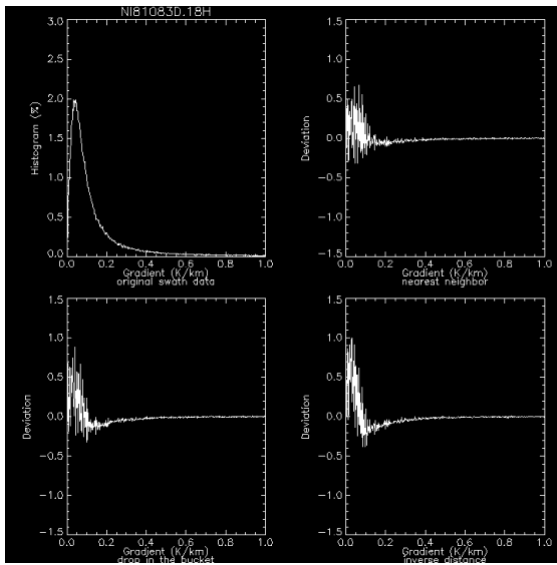
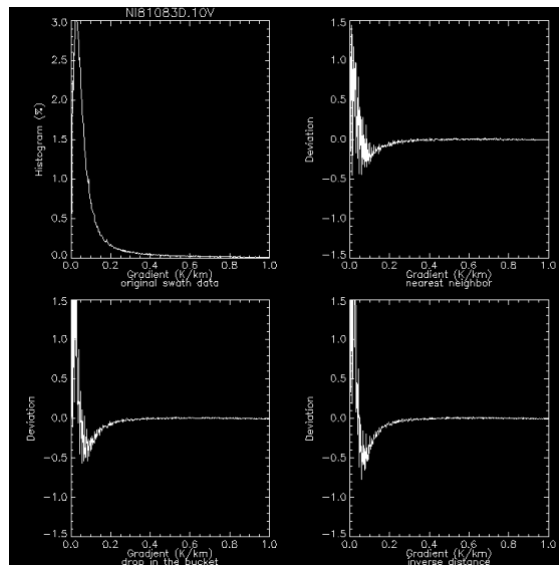
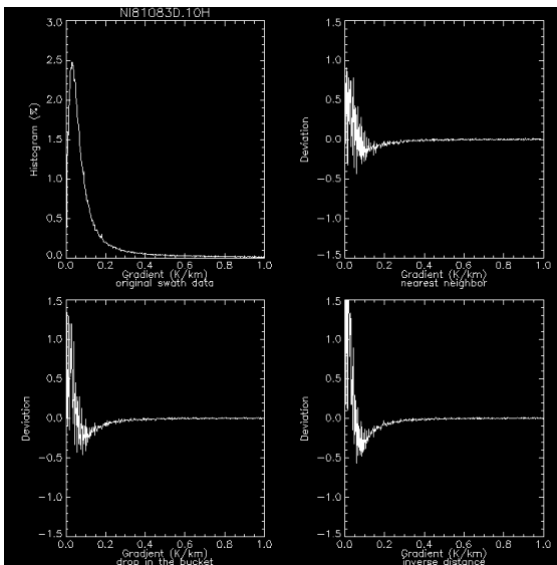
Northern Hemisphere, March 24, 1981, Ascending

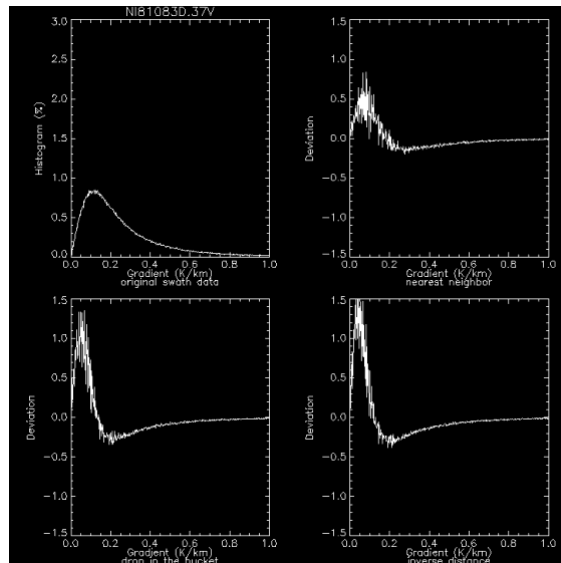
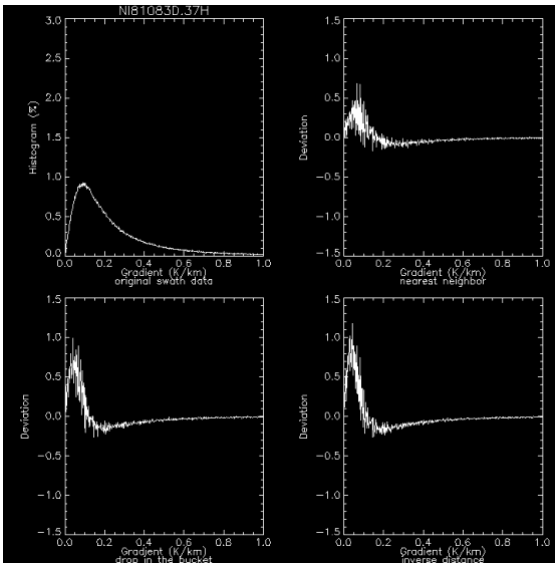




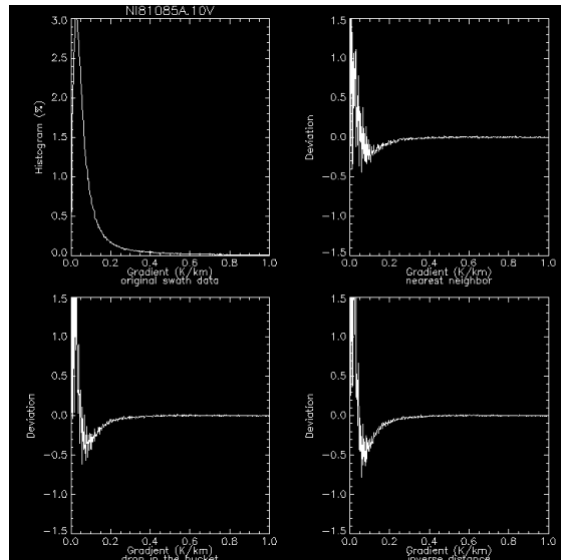
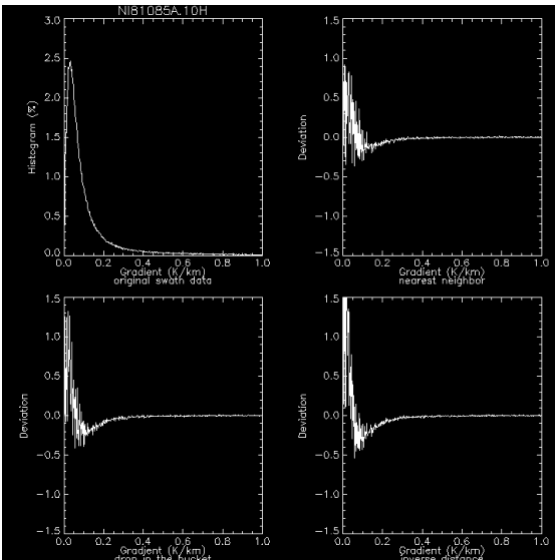
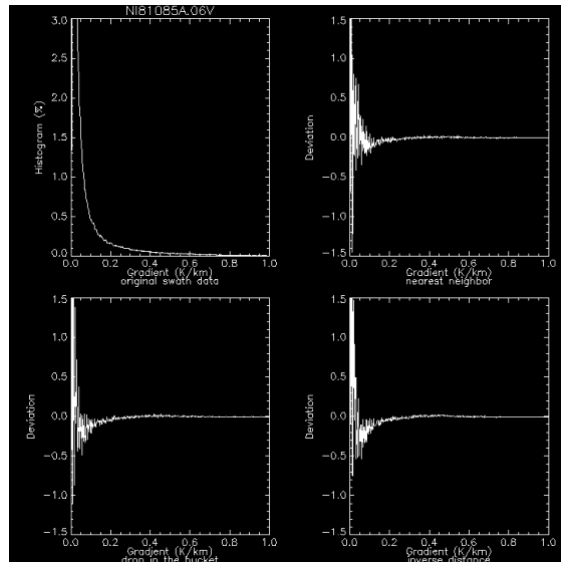
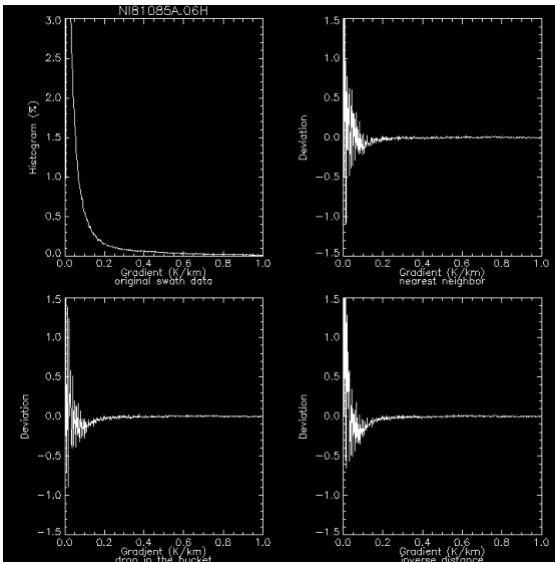
Northern Hemisphere, March 24, 1981, Descending

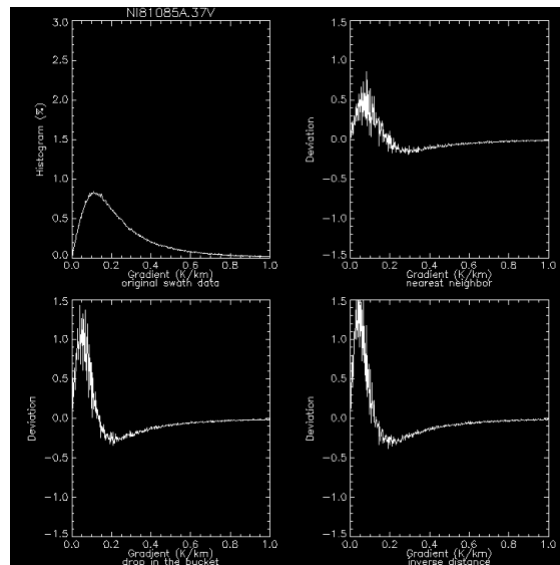
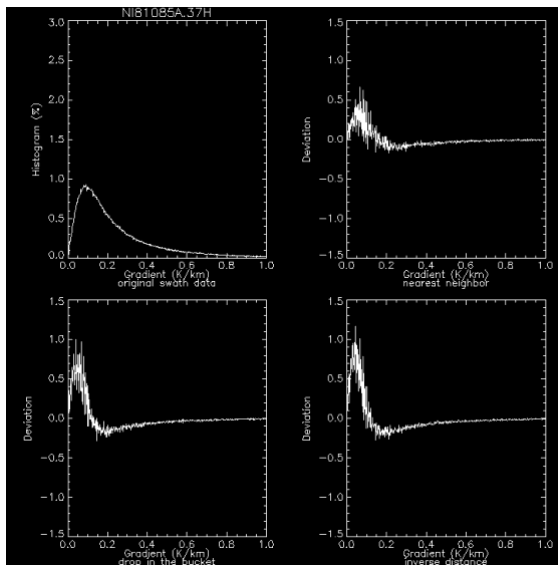
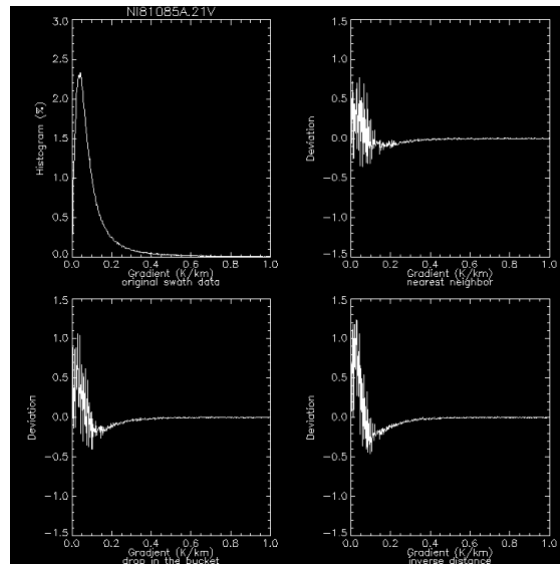
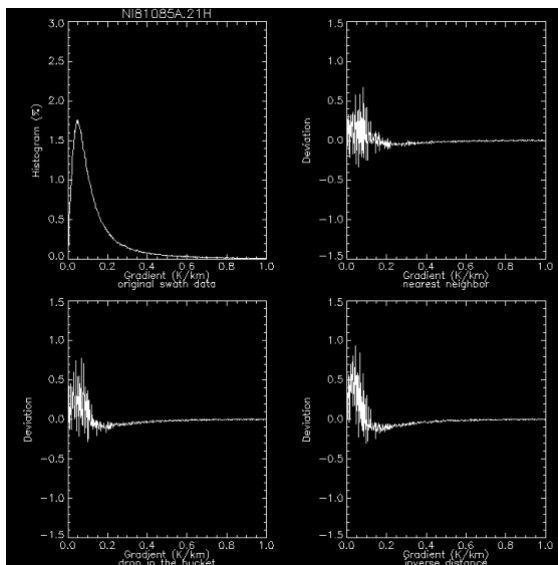
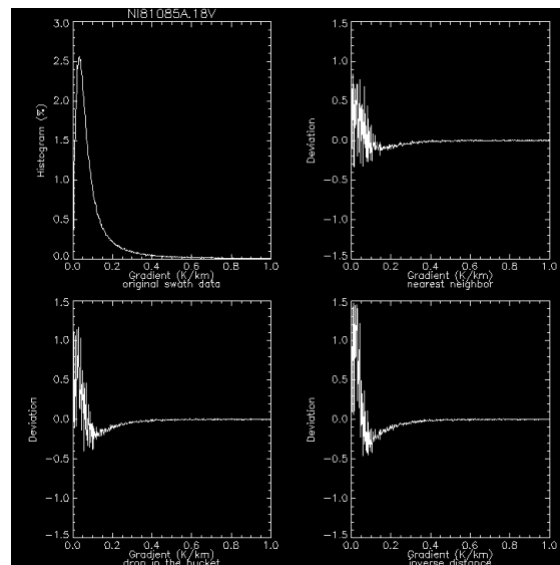
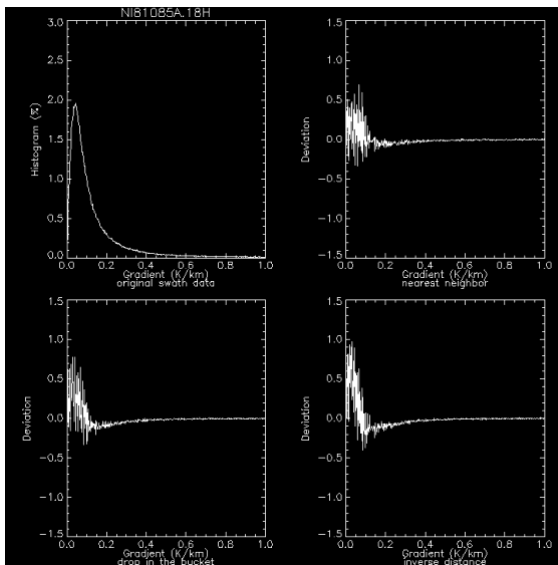




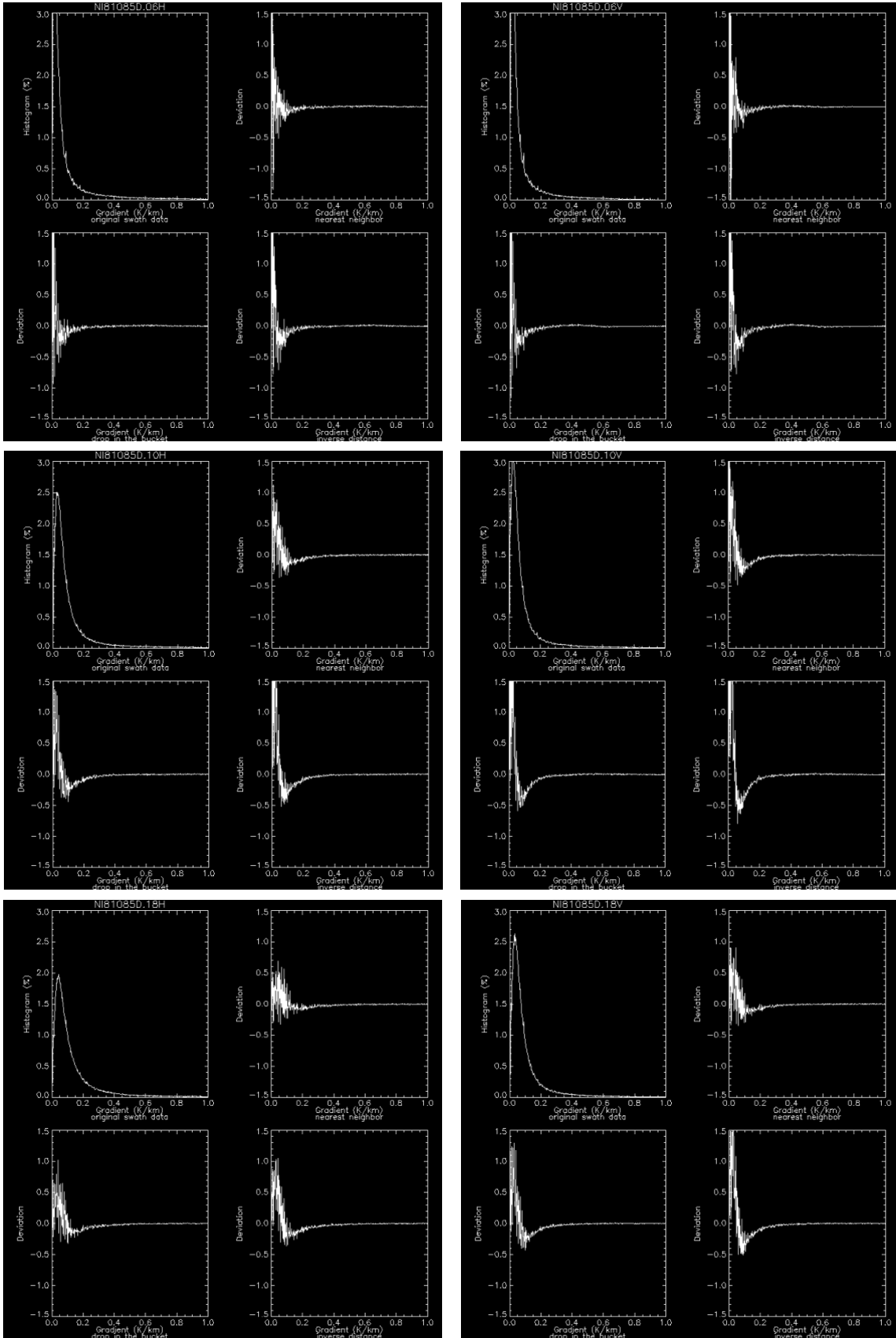


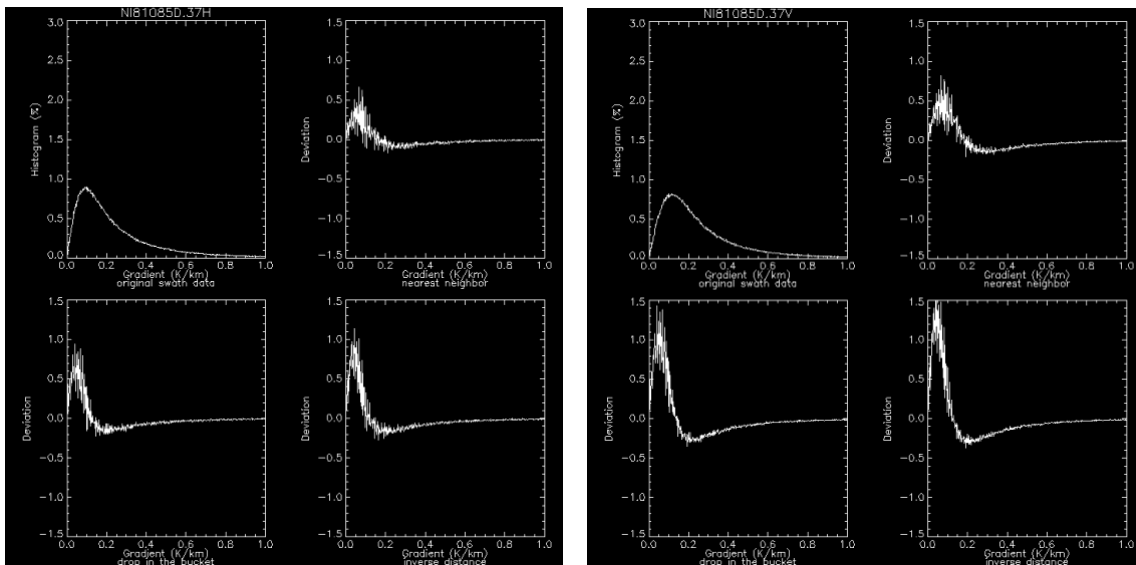
Northern Hemisphere, March 26, 1981, Ascending





Northern Hemisphere, March 26, 1981, Descending





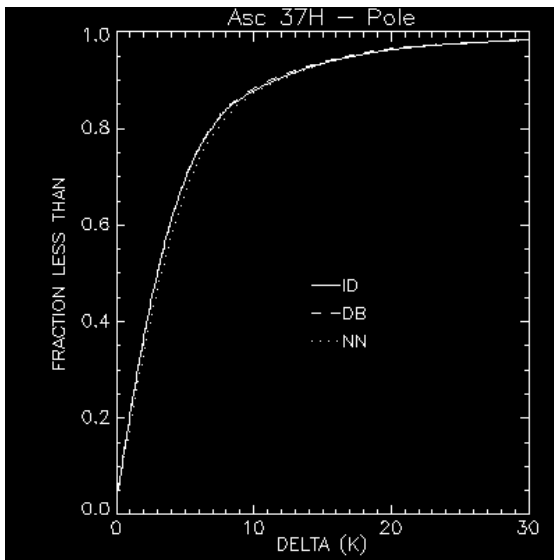
5.3 Daily differences and histograms

Permanent features should provide good references. These images examine day to day differences in images. Cumulative histogram plots of Tb differences are used to compare interpolation methods. Presumably the best method will have the lowest cumulative error.

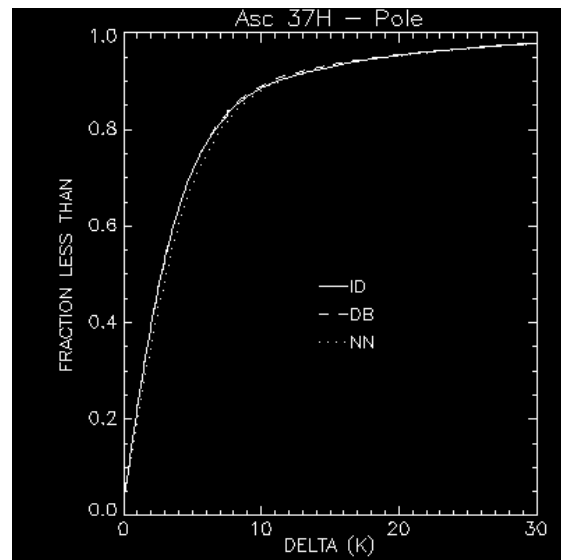
5.3.1 Full Resolution Daily Differences

Ascending Data – 37 GHz Horizontal

NI81081A vs. NI81083A

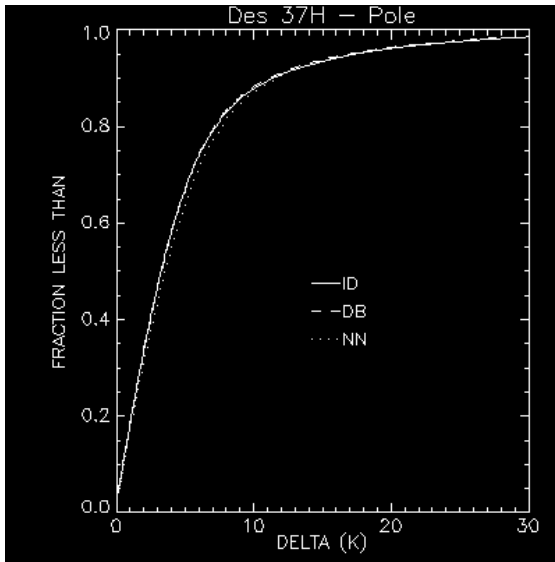


NI81083A vs. NI81085A

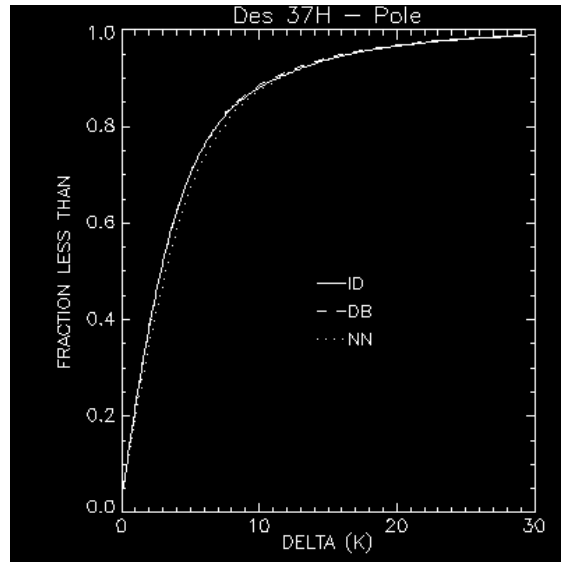


Descending Data – 37 GHz Horizontal

NI81081D vs. NI81083D



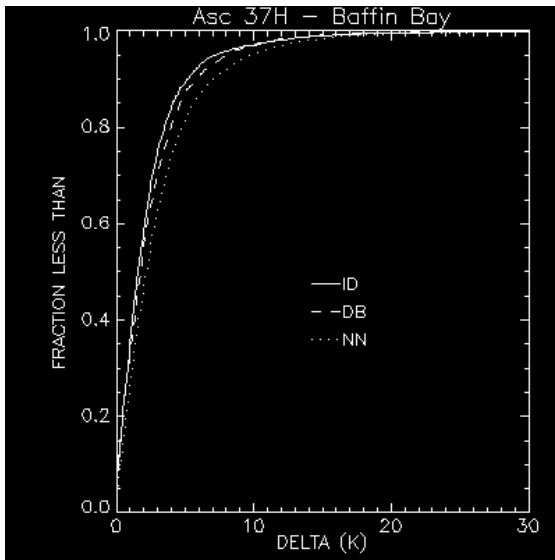
NI81083D vs. NI81085D



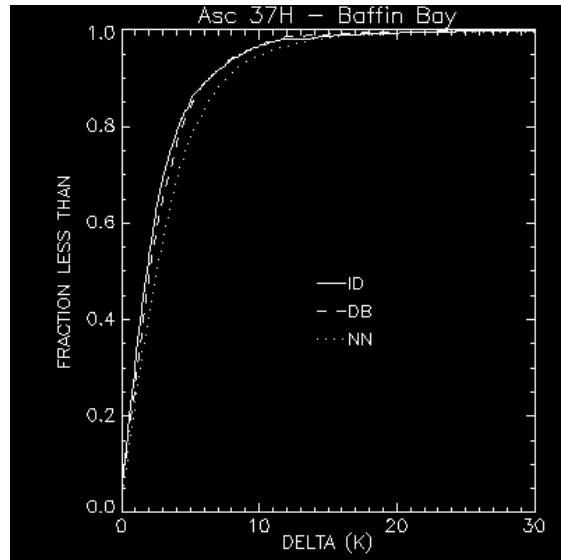
5.3.2 Baffin Bay Daily Differences

Ascending Data – 37 GHz Horizontal

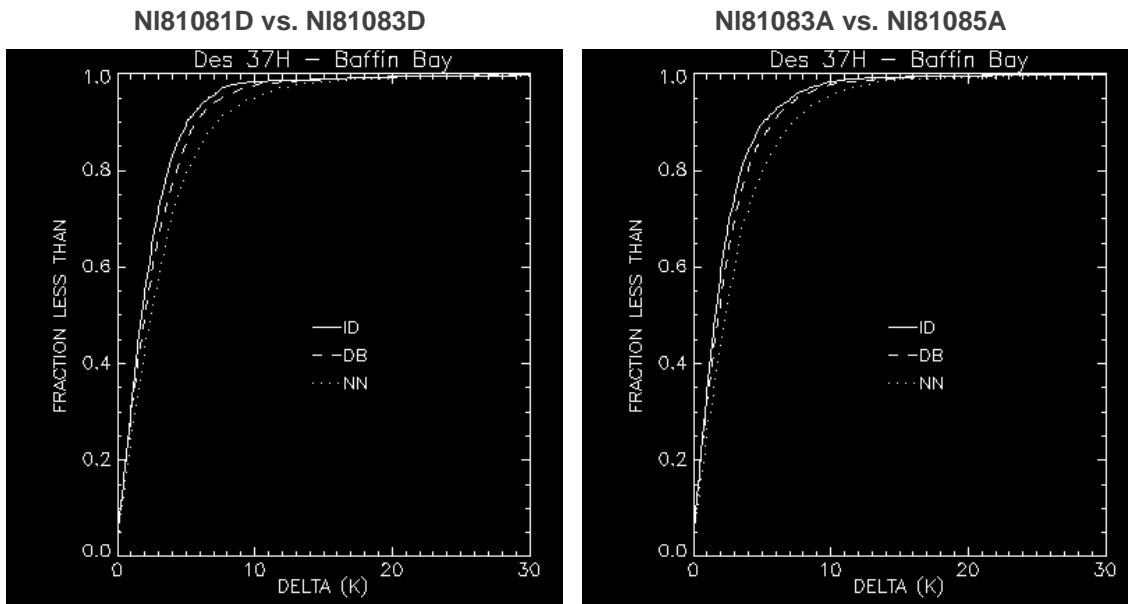
NI81081A vs. NI81083A



NI81083A vs. NI81085A



Descending Data – 37 GHz Horizontal

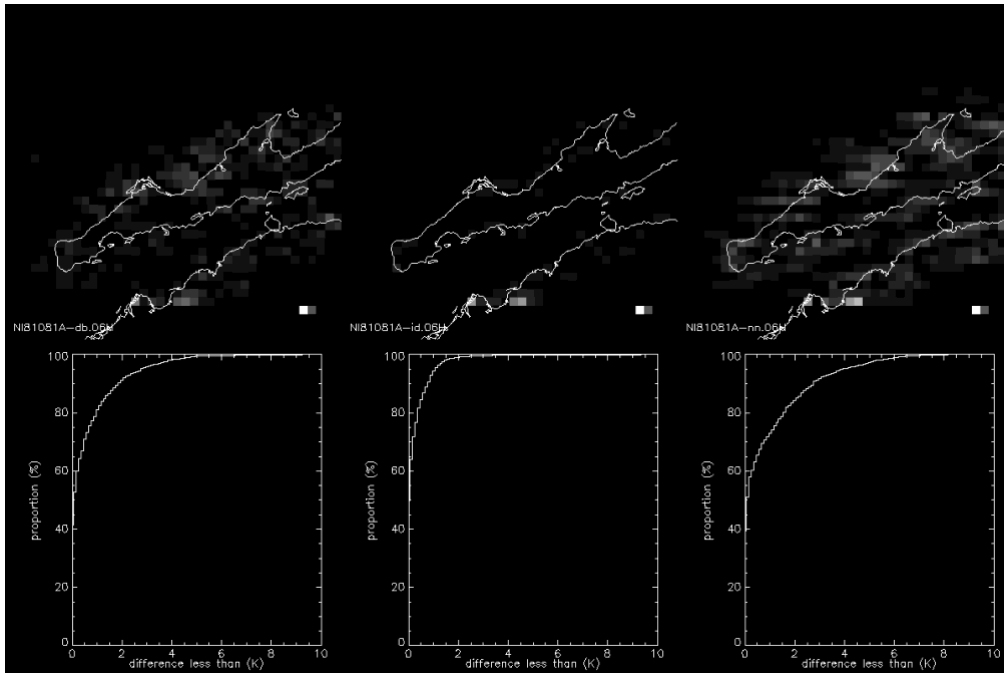


5.4 Comparison to Tessellation

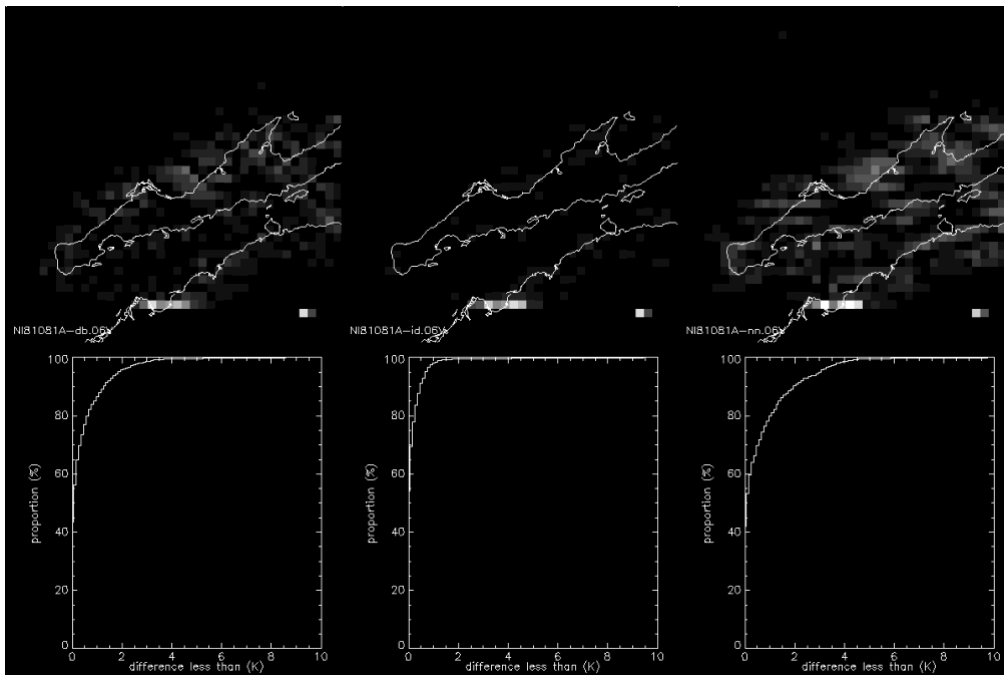
To tessellate is to cover with small tiles. Here the term specifically refers to an interpolation method. First, for each point, the set of points which are nearer to it than any other point is listed. This collection of sets of points is called the Delauney triangulation. The triangulation is then used to create a smooth surface over the original data. A set of quintic polynomials are used to fit a surface which is smooth and has a smooth first derivative. The values at the grid points are then taken from this surface. Tessellation is a good general solution to the problem of interpolating irregularly sampled data and can easily be adapted to interpolation on a sphere. It is, however, computationally expensive. Here the results of the test interpolation methods are compared to a tessellation over a small area for a single orbit.

5.4.1 Gulf of California

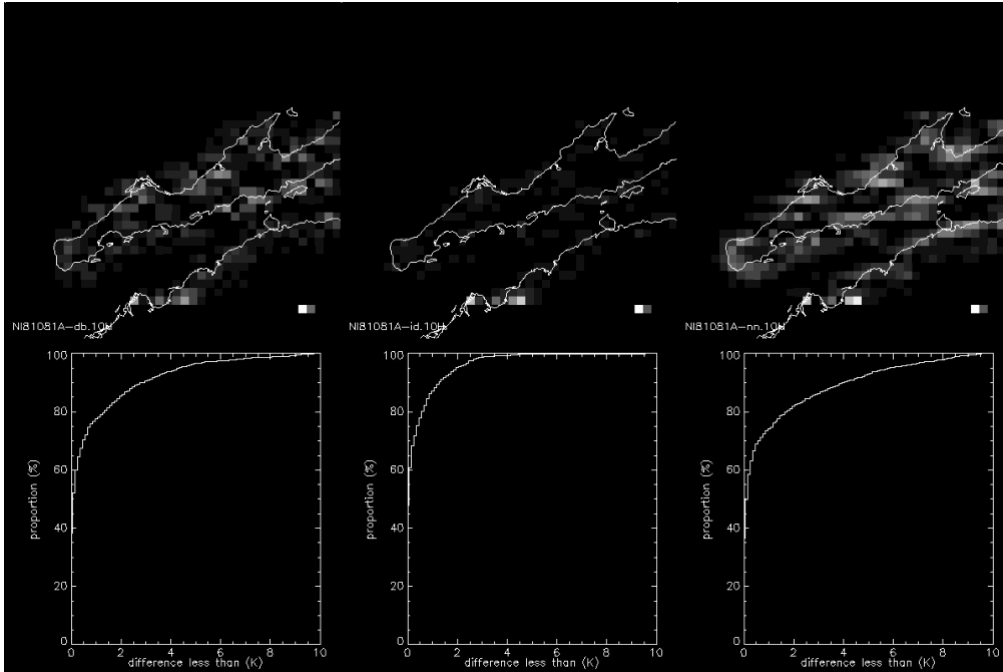
NI81081A.06H



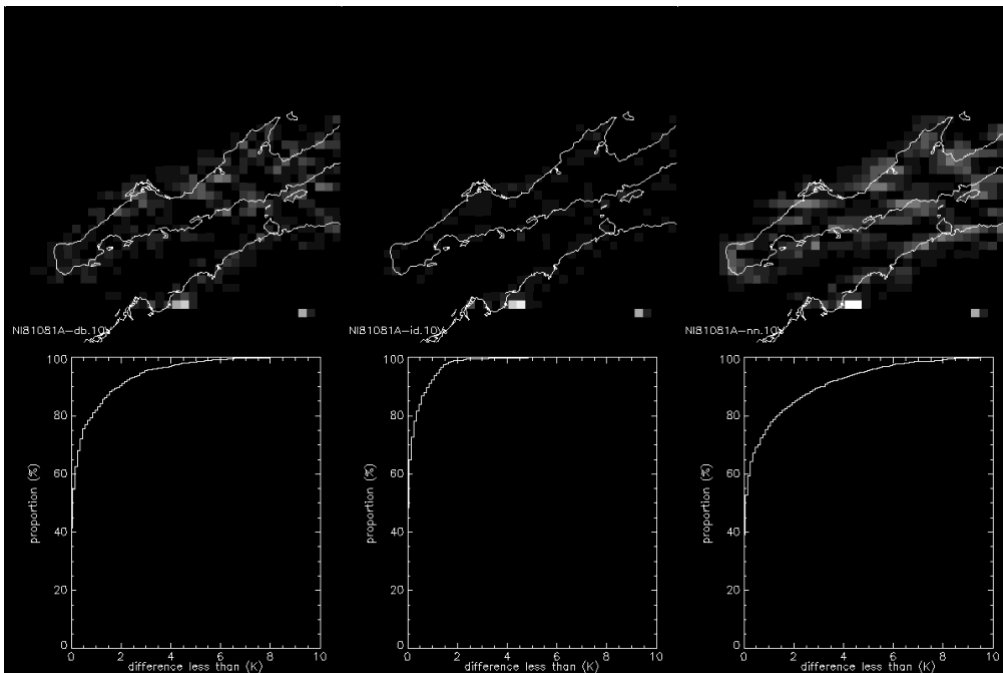
NI81081A.06V



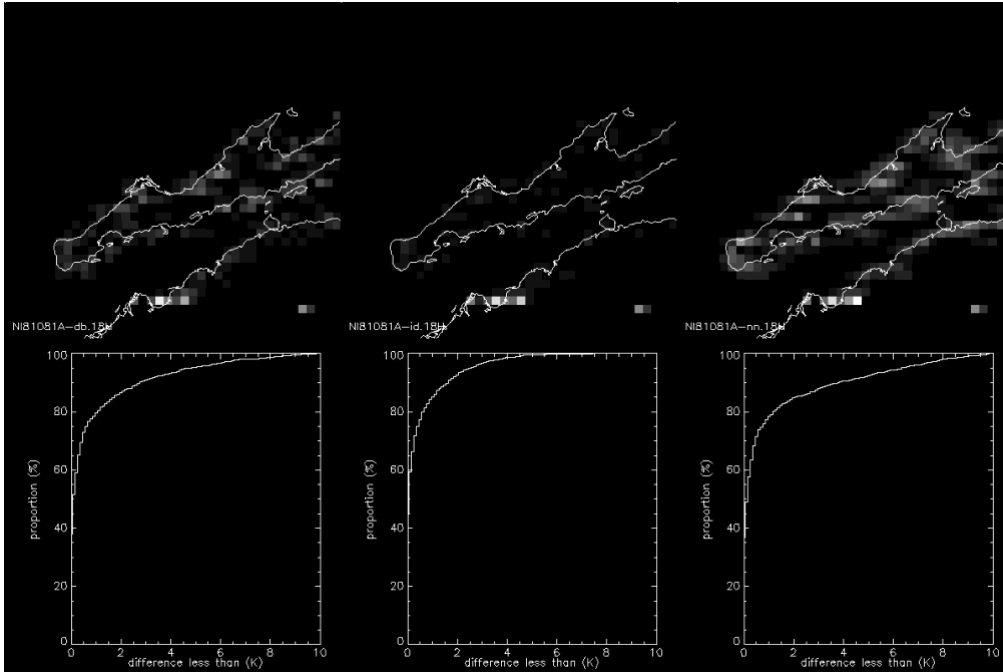
NI81081A.10H



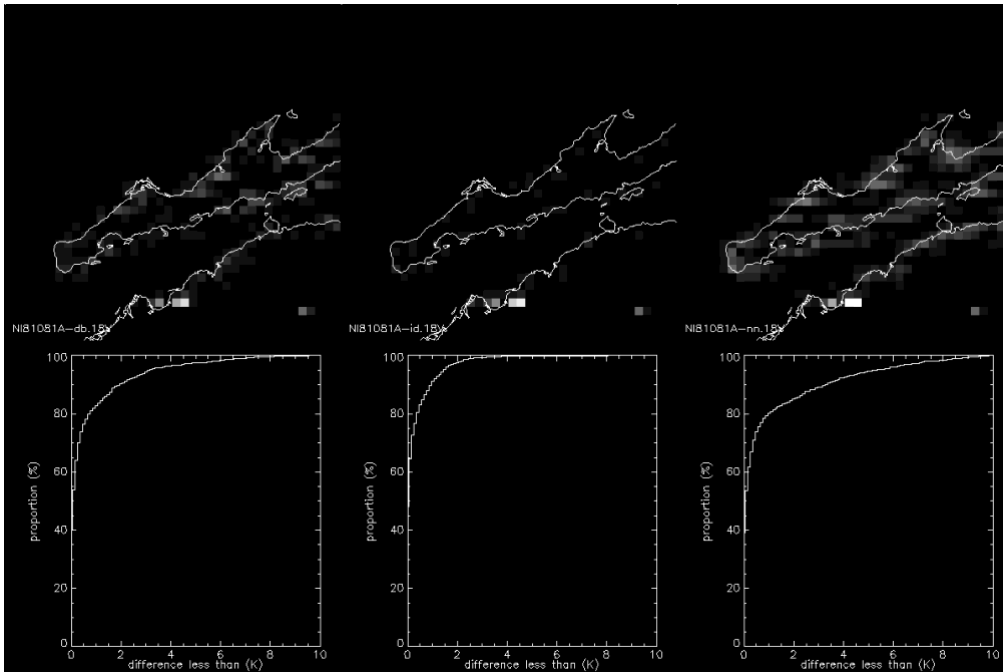
NI81081A.10V



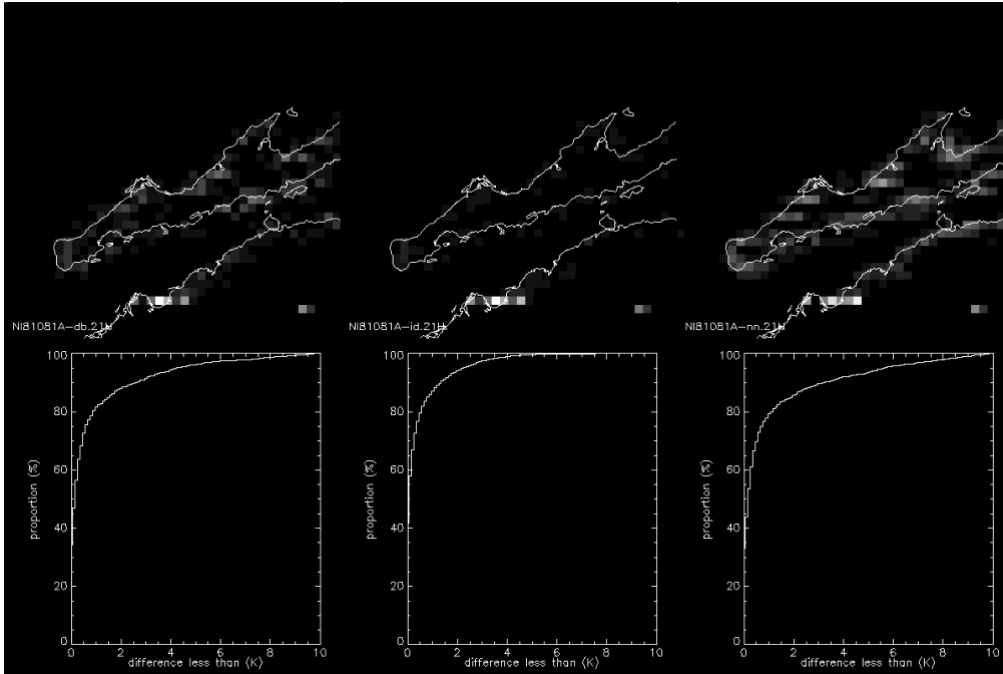
NI81081A.18H



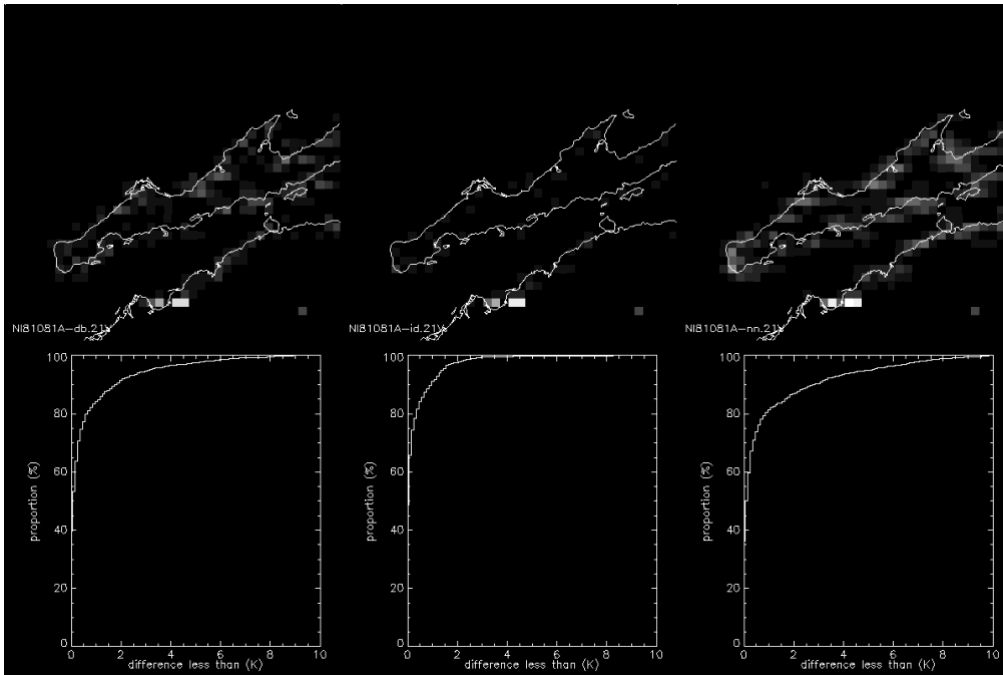
NI81081A.18V



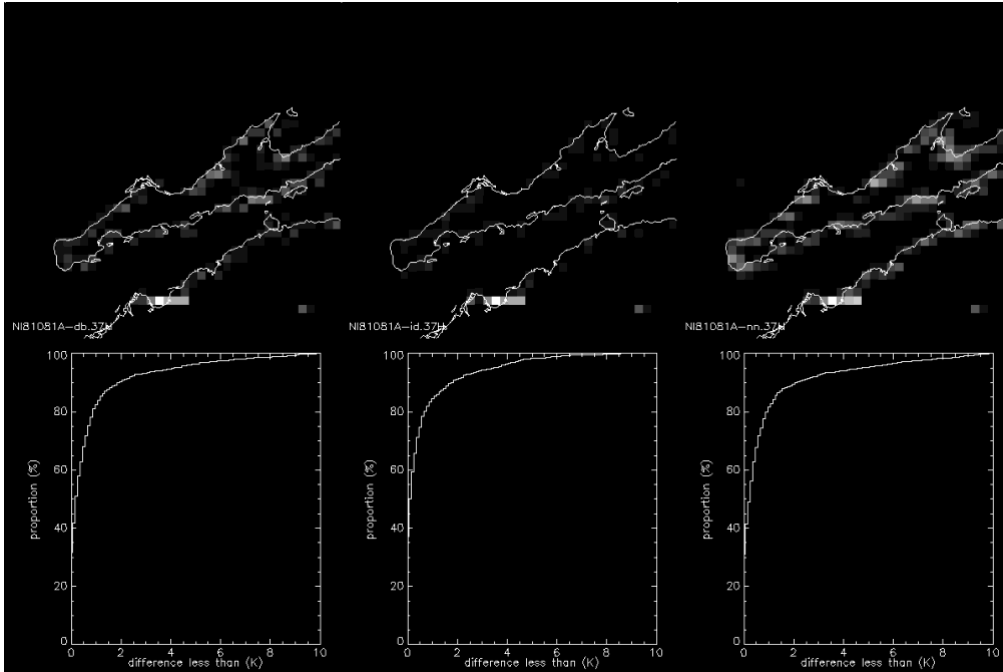
NI81081A.21H



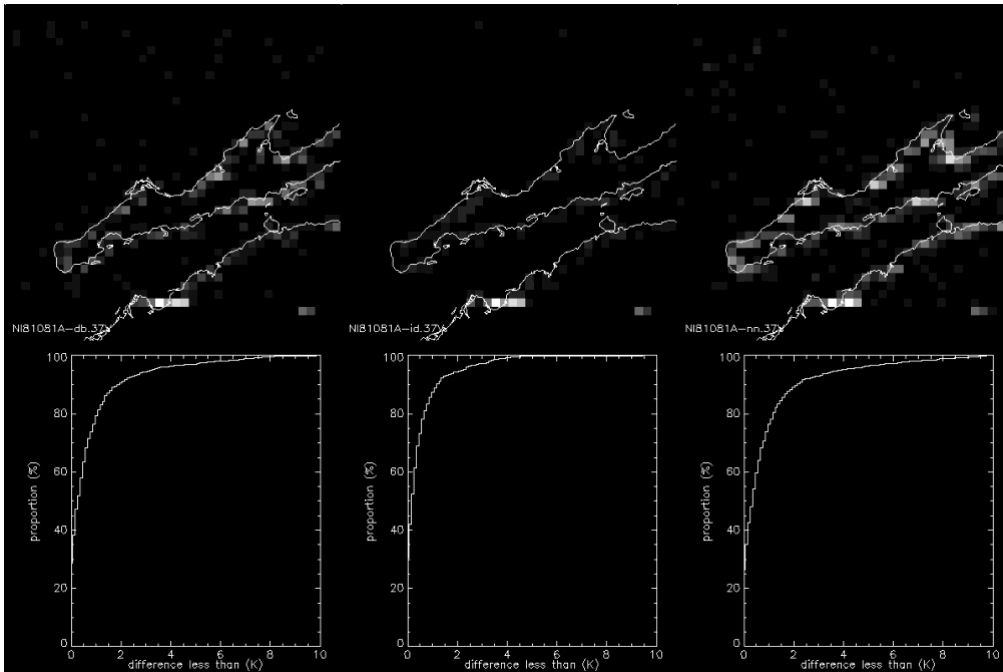
NI81081A.21V



NI81081A.37H



NI81081A.37V



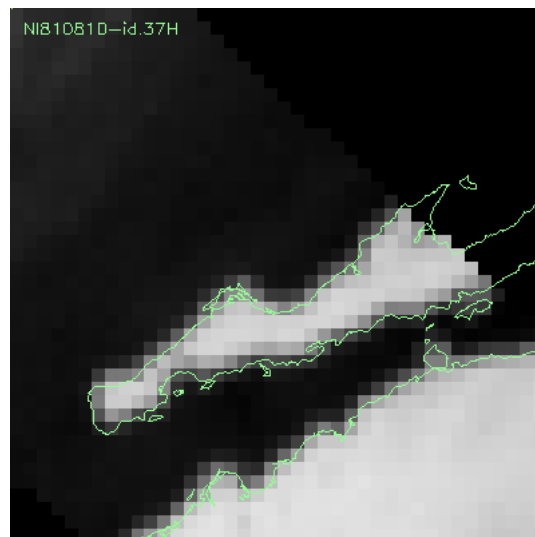
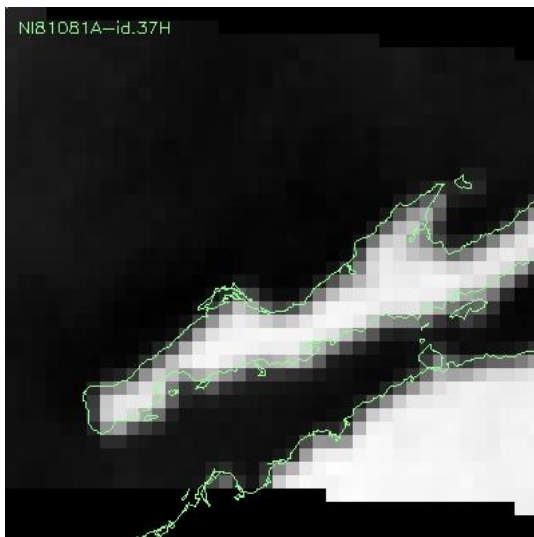
5.5 Problems with the data

While processing the data for these tests I noticed a few things. Since the sensor looks ahead of/behind??? the satellite I didn't want to use the asc/des flag in the header, but wanted to

discriminate based on the location of the sample at nadir. When the satellite is ascending toward the north pole, the latitudes should steadily increase and then decrease as the satellite descends. What I found however, is that the orbit wobbles, rapidly shifting back and forth between ascending (increasing latitude) and descending (decreasing latitude) for a few scans. I found this at the poles, but I suspect it wobbles everywhere in the orbit.

Another problem I noticed was that when I tried to run these same tests on the southern hemisphere, I found no ascending data below 40S for any of the sample days. I have not yet explored this further.

Finally, when comparing the daily differences, I had to compare ascending to ascending and descending to descending because there seems to be a systematic along track error in the geolocation. Take a look at this difference image.



6 GLOSSARY

Frequency Response

This plot graphically shows how the interpolation process responds to spatial frequencies in the original data. The horizontal axes show cycles per sampling interval and the vertical axis gives the relative strength of how those frequency components will be passed on to the interpolated signal. The ideal frequency response is a box of height 1.0 which spans the interval between $-1/2$ and $+1/2$ cycles per sample. Note that the frequency response plots have a logarithmic scale on the vertical axis.

Geolocation Errors

Errors in determining position on the Earth.

Geometric Mean

Square root of the sum of the squares.

Gradient

The gradient is the rate of change (i.e., magnitude of first derivative) at a point. Here it specifically refers to the rate at which the brightness temperatures change with respect to distance. For the swath data the gradient at a specific sample is approximated by taking the geometric mean of the T_b difference divided by distance for the adjacent samples along scan and across scan. For the gridded data it's the same except the adjacent column and row are used.

Noise

Not signal. Aliasing is a specific kind of noise where the energy from frequencies greater than the Nyquist critical frequency is spuriously reflected back into the sampled signal.

Nyquist Critical Frequency

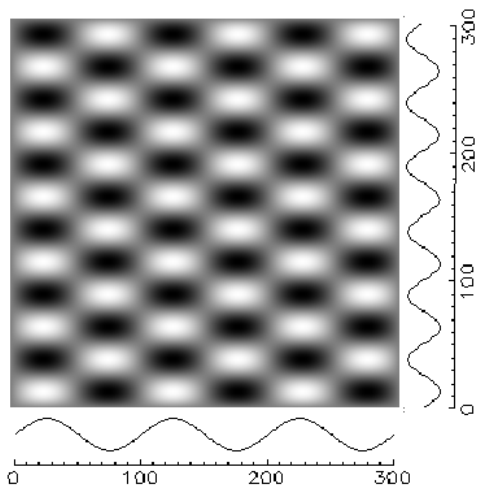
It takes two pixels to describe a cycle, so the critical frequency, $F_c = 1/2$ cycle per pixel. A digital signal cannot represent frequencies greater than F_c .

Smoothing

Smoothing eliminates high frequencies in the signal. If these high frequencies are noise, then this is a good thing. If they are signal, then smoothing is (probably) a bad thing. Averaging is a good example of smoothing. If the original data contains random errors (or, in a digital signal, aliasing) then averaging will eliminate the errors. If on the other hand the fluctuations in the sampled data are due to actual fluctuations in the measured signal, averaging will cause you to lose this information.

Spatial Frequency

We usually think of frequency as the rate of oscillation per time. Spatial frequency is the rate of oscillation per distance. Here's an example of an image that varies at the rate of 1 cycle per 100 samples (or .01 cycles per sample) in x and 1 cycle per 50 samples (or .02 cycles per sample) in y .



Transfer Function

The plot of the transfer function graphically shows the distribution of weights used to interpolate the original data. The point to be interpolated is at the center of the horizontal axes. The vertical scale gives the relative weight to be given to each original data point according to its (horizontal) spatial relationship to the interpolation point.

7 REFERENCES

Gonzalez, R. and P. Wintz. (1987). *Digital Image Processing*. Reading, MA: Addison-Wesley.

Teukolsky, et. al., *Numerical Recipes*.