I have used Landsat imagery for over 25 years and am currently using the Landsat imagery being distributed through the USGS EROS Data Center.

Over the past year I have had some issues that I'd like to talk about with you …

My problem is like this:

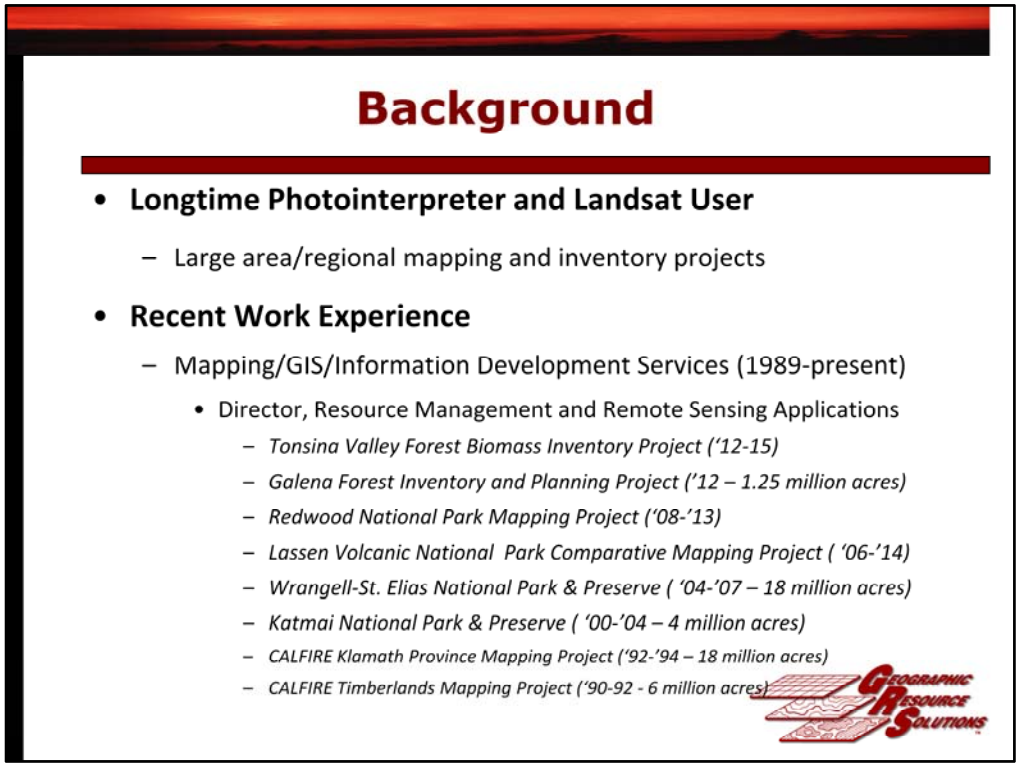
Imagine you really like ice cream and a new ice cream store opened in your town that advertised FREE ICE CREAM, as much as you could eat.

You said "WOW that's great" and went to the ice cream store to get some.

But when you got there and surveyed the menu you found that all the flavors of the free ice cream contained **NUTS – Lots of NUTS** … and you don't like **NUTS**!

Turned out it was not so great a deal after all.

That's sort of the problem I have with this free imagery.[Next]



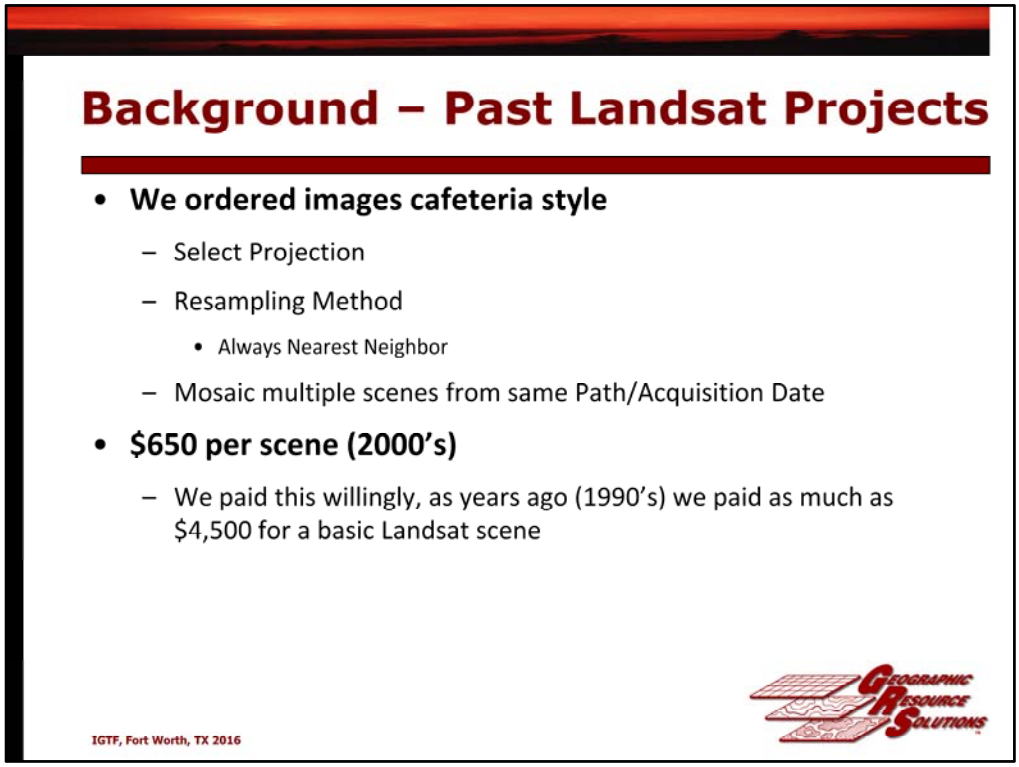
Back in 1974 I was quite fortunate to land a job with one of the major forest industries in the redwood region in northern California, a company noted for it's progressive approach to managing their timberlands.

My position was a new one and it encompassed some very interesting responsibilities that included managing the forest inventory, developing and applying growth and harvest simulation models, and developing GIS capabilities.

I worked in the field and in the office, and as part of my inventory management responsibilities I performed vegetation typing using stereo photography, wax pencils, and a zoom transfer scope.

After 15 years of this work, I left to Co-found Geographic Resource Solutions, where I have now worked the past 27 years. I started performing Image Classification Services and using Landsat imagery for large area mapping projects in the early 1990's when we landed the California Timberlands Mapping (6MM acres) and Klamath Province Mapping (18MM acres) projects for the California Dept. of Forestry and Fire Protection.

Since successfully completing those projects I have completed many other projects, primarily in California, the Pacific northwest, and Alaska covering over 50 million acres and including 4 National Parks. We currently have a Blanket Purchase Agreement with the BLM in Alaska to provide all Forest Inventory and Mapping services for the next 5 years.

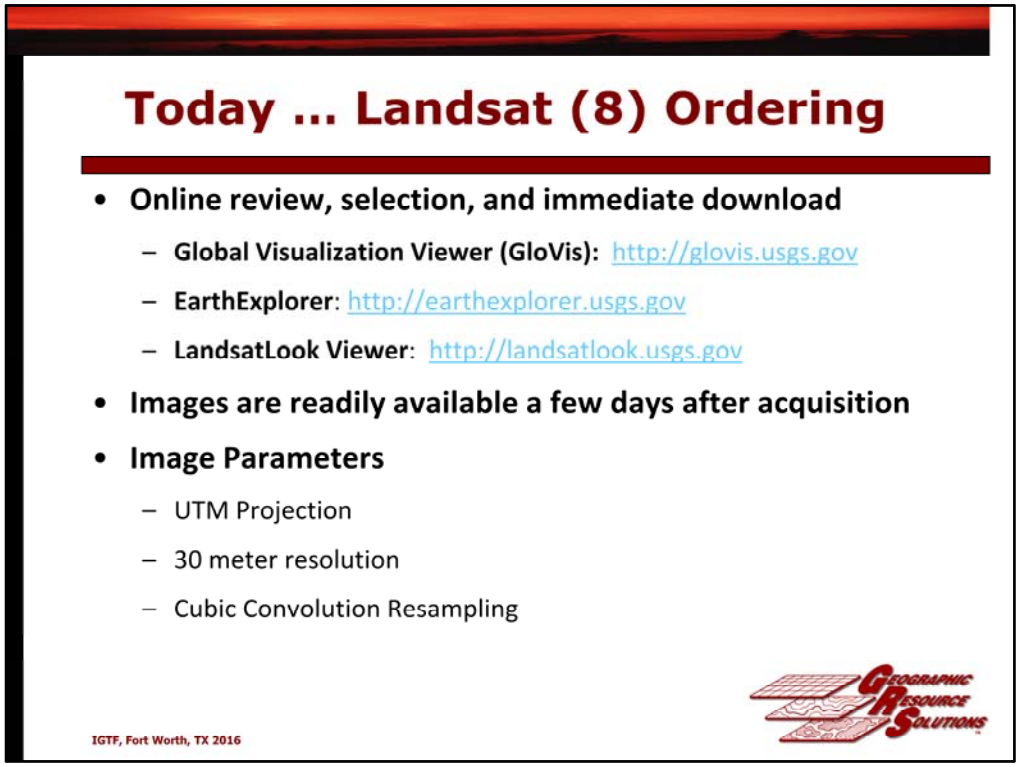


For all of these past projects that were based on Landsat imagery …

We ordered the imagery cafeteria style specifying the parameters of the imagery we needed.

We paid for the imagery … and it wasn't a problem, as we got the imagery we needed.

Now it is **ALL FREE!**



For our most recent efforts in Alaska, Colorado, and Northern California we have acquired the free Landsat 8 imagery at the GloVis website.

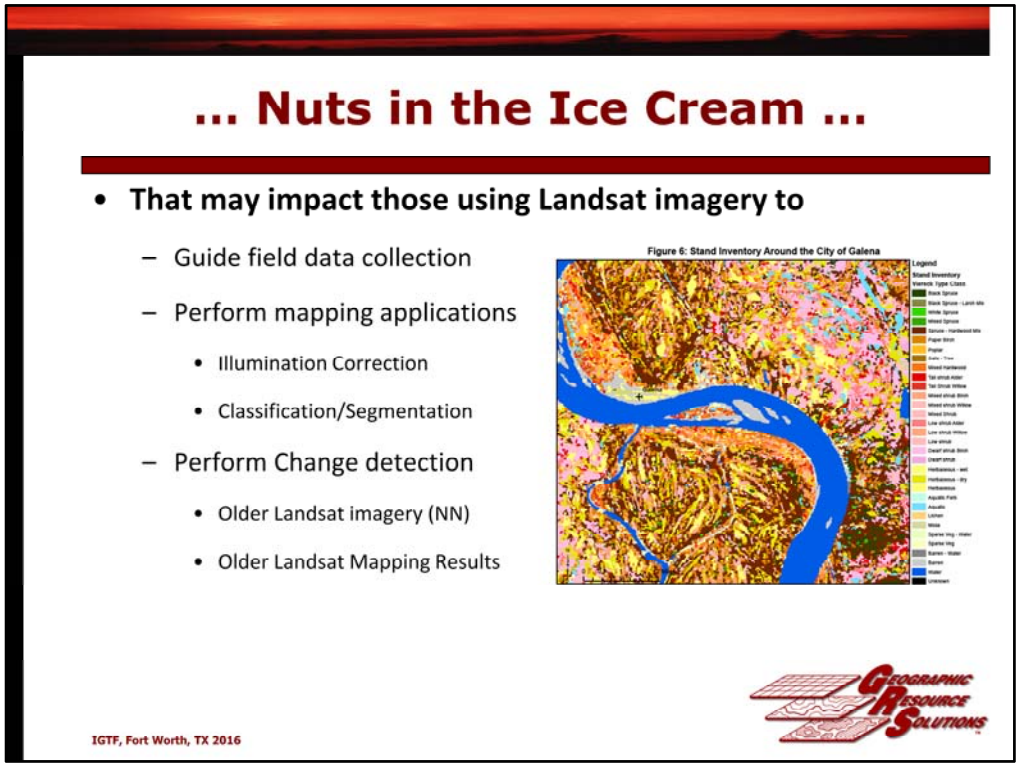
It was really easy to find, review, select, and download.

But then we found NUTs in our ice cream! Lots and Lots of Nuts!

EarthExplorer: http://earthexplorer.usgs.gov – allows geographical searches of data held in the USGS archives

Global Visualization Viewer (GloVis): http://glovis.usgs.gov – a browse-based viewer for USGS Landsat Archive data sets

LandsatLook Viewer: http://landsatlook.usgs.gov - a prototype tool that allows rapid online viewing and access to the USGS Landsat archive



What were those NUTS that we were finding everywhere ?

They were digital artifacts that had been created because the imagery had been terrain corrected using the Cubic Convolution resampling algorithm!

These artifacts can be a real problem to those that use imagery to:

guide field data collection efforts;

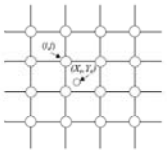
perform mapping applications; or

perform change detection, especially if one is processing new imagery relative to older previously NN'd imagery.

I decided to bring this issue up here because I think it is really important. So far, I haven't been successful in getting anyone at EROS Data Center to recognize that it is causing problems for people like me who process the imagery the way that we do … and I do not think I am the only one who does want the **nutty** imagery.

Some Simple Definitions

- **Cubic Convolution**
  - Assigns the value of a pixel based on the weighted average of a 4x4 grid of the nearest 16 pixels.
  - New pixel values may be **calculated** and assigned that may not exist in original 4x4 grid of values.
  - Likely does not preserve original values.



- **Nearest Neighbor**
  - Assigns the value of the pixel whose center is nearest to the center of the output pixel.
  - It represents a **transfer** of original values to the new raster accompanied by possible slight shifts in spatial accuracy.
  - Preserves original values.

IGTF, Fort Worth, TX 2016

So what's the big deal …

The CC approach assigns a value to the resampled pixel location based on a weighted averaging of the values of the nearest 16 pixels in the original grid.

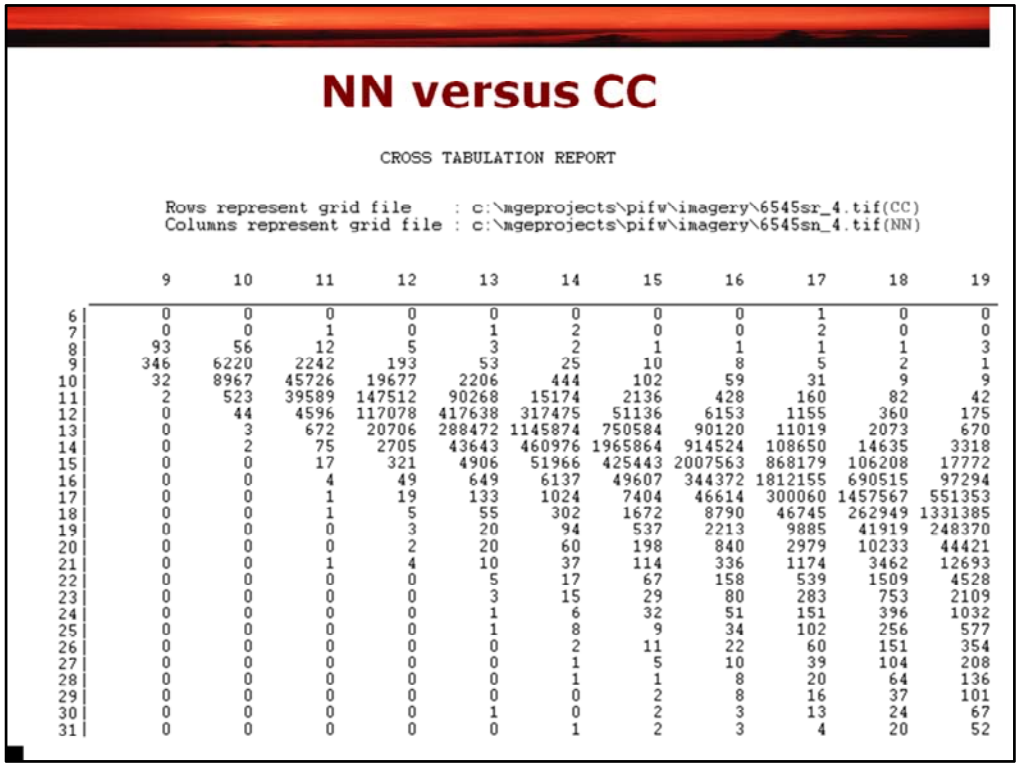
The CC approach does not preserve the original pixel value unless all surrounding (16 nearest) average that same value.

The literature quite clearly points out that this algorithm should not be used with generalized or classified data values. … Okay

The NN approach is unique in that it is the only resampling method that does not generate interpolated or averaged new values –

it transfers values based on the nearest pixel center point.

This approach should always be used with classified data and it must be used if it is important to preserve ORIGINAL data values.



NN versus CC

CROSS TABULATION REPORT

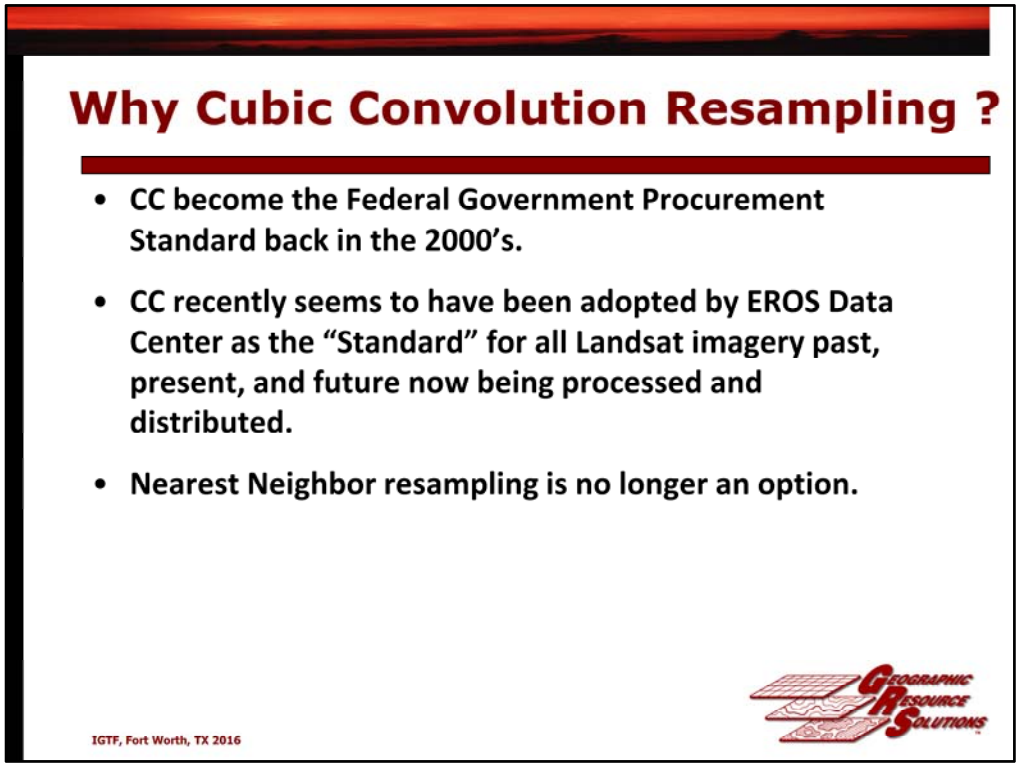
Rows represent grid file : c:\mgeprojects\pifw\imagery\6545sr_4.tif(CC)
Columns represent grid file : c:\mgeprojects\pifw\imagery\6545sn_4.tif(NN)

	9	10	11	12	13	14	15	16	17	18	19
6	0	0	0	0	0	0	0	0	1	0	0
7	0	0	1	0	1	2	0	0	2	0	0
8	93	56	12	5	3	2	1	1	1	1	3
9	346	6220	2242	193	53	25	10	8	5	2	1
10	32	8967	45726	19677	2206	444	102	59	31	9	9
11	2	523	39589	147512	90268	15174	2136	428	160	82	42
12	0	44	4596	117078	417638	317475	51136	6153	1155	360	175
13	0	3	672	20706	288472	1145874	750584	90120	11019	2073	670
14	0	2	75	2705	43643	460976	1965864	914524	108650	14635	3318
15	0	0	17	321	4906	51966	425443	2007563	868179	106208	17772
16	0	0	4	49	649	6137	49607	344372	1812155	690515	97294
17	0	0	1	19	133	1024	7404	46614	300060	1457567	551353
18	0	0	1	5	55	302	1672	8790	46745	262949	1331385
19	0	0	0	3	20	94	537	2213	9885	41919	248370
20	0	0	0	2	20	60	198	840	2979	10233	44421
21	0	0	1	4	10	37	114	336	1174	3462	12693
22	0	0	0	0	5	17	67	158	539	1509	4528
23	0	0	0	0	3	15	29	80	283	753	2109
24	0	0	0	0	1	6	32	51	151	396	1032
25	0	0	0	0	1	8	9	34	102	256	577
26	0	0	0	0	0	2	11	22	60	151	354
27	0	0	0	0	0	1	5	10	39	104	208
28	0	0	0	0	0	1	1	8	20	64	136
29	0	0	0	0	0	0	2	8	16	37	101
30	0	0	0	0	1	0	2	3	13	24	67
31	0	0	0	0	0	1	2	3	4	20	52

I first saw what I thought were significant image pixel value differences when I was ordering imagery for Hawaii 15 years ago and was by mistake delivered a CC resampled image.

I requested the NN resampled image and later compared band 4 using a cross-tabulation process.

This report indicated that there were many, many changes of the original values … and that most original values (horizontal rows) had been changed to another value.



But how did we end up with Cubic Convolution being our only resampling option ?

This is what I think happened:

Back in the early 2000's … the Federal Procurement Standard was adopted including Cubic Convolution resampling. I've heard stories about how this happened,

but there's no time for that now.

More recently, the CC algorithm seems to have been adopted by EROS Data Center as the "Standard" for all Landsat imagery past, present, and future now being processed and distributed.

The Nearest Neighbor resampling approach is no longer an option.

Their Customer Service reps have said to find a value Added Contractor listed on their website to do this work.

We have tried to do this …

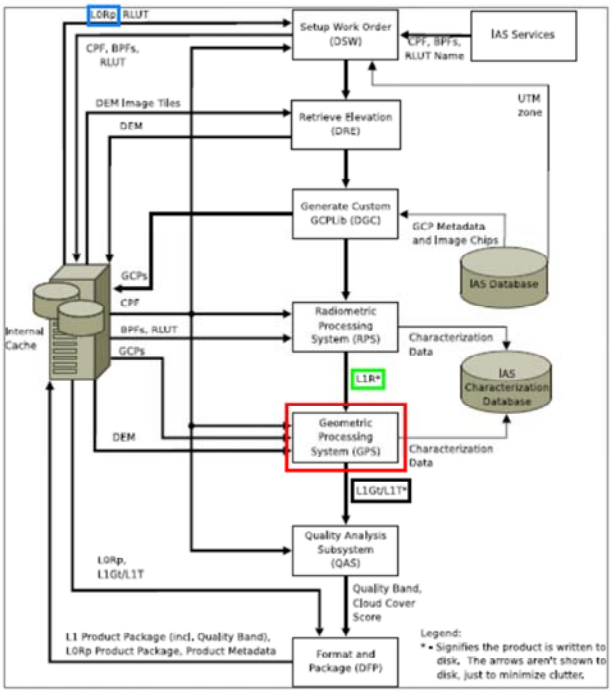


Figure 4-1. LPGS Standard Product Data Flow

- 44 -

LSDS-1574
Version 1.0

We did not find one vendor with this capability and we thought it important to find someone who could do what EROS Data Center used to do, for the sake of continuity over time.

We did find one vendor who said they would try to develop that capability.

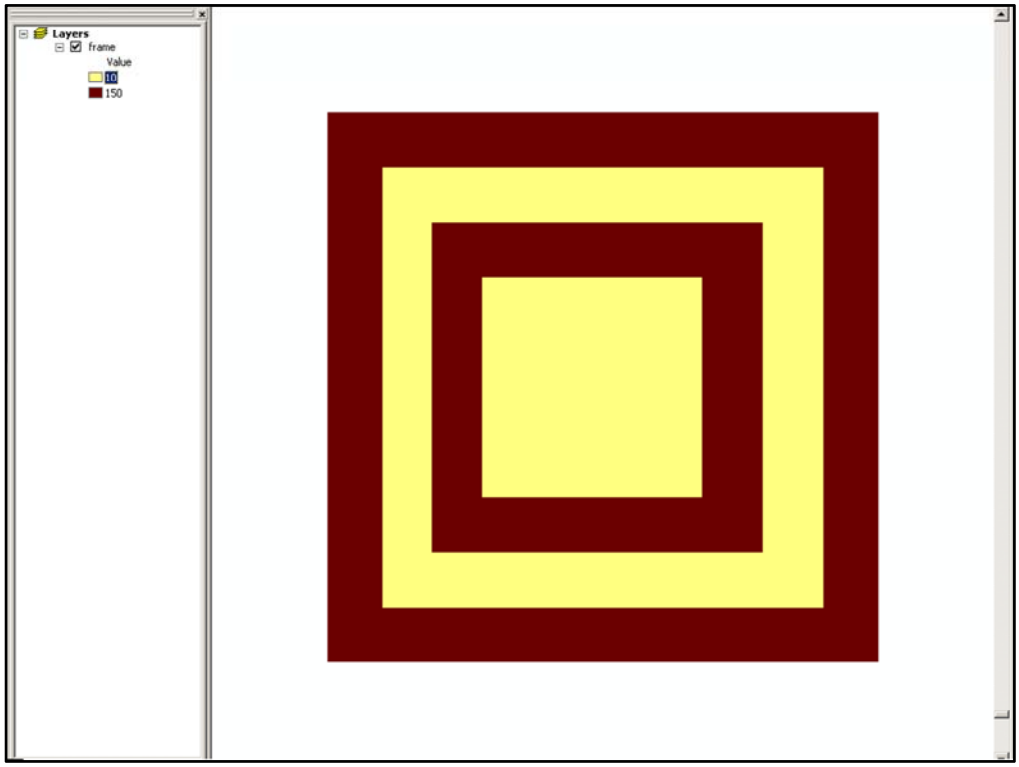
But then, rather than provide the L1R radiometrically corrected image EROS Data Center is said to save before the terrain correction process is applied, EROS Data Center delivered the raw Level 0 image data that has never had any radiometric corrections. They would not deliver the L1R.

The contractor I found who said he would try to repeat the process failed miserably.

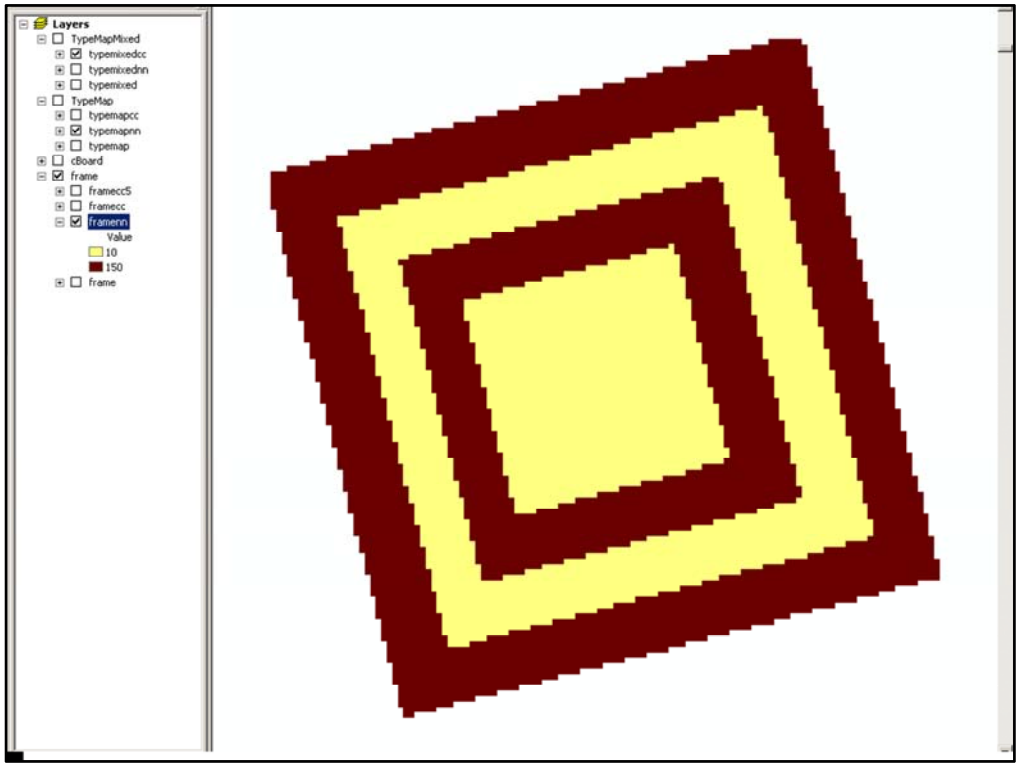
And EROS Data Center still say they won't provide the Level 1R imagery, even though this diagram in their L8 handbook indicates they save these files.

So what am I to do?

So Let's look at how differently these resampling algorithms operate. Let's resample a simple two color set of frames.

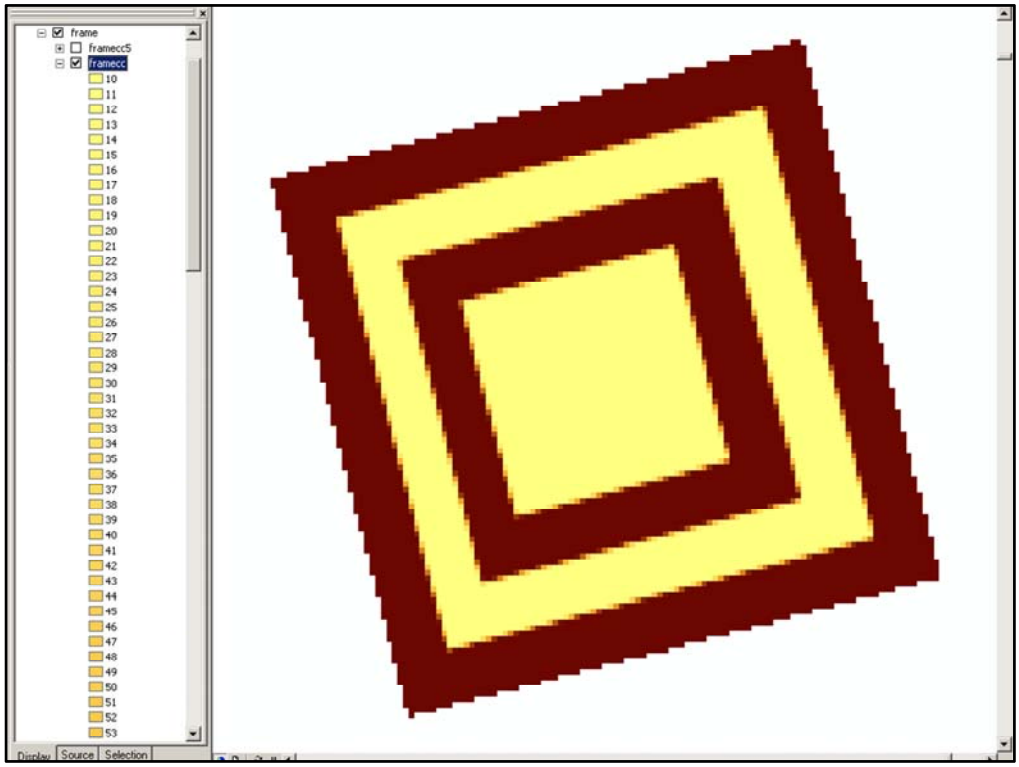


Here is a simple raster data set representing frames comprised of two different pixel values …



The original pixel values projected with NN resampling exhibit some minor spatial displacement – resulting in the jagged edges you see.

This is a major complaint about this process – the jagged edges create a lousy picture … but the original data/spectral content has been preserved.



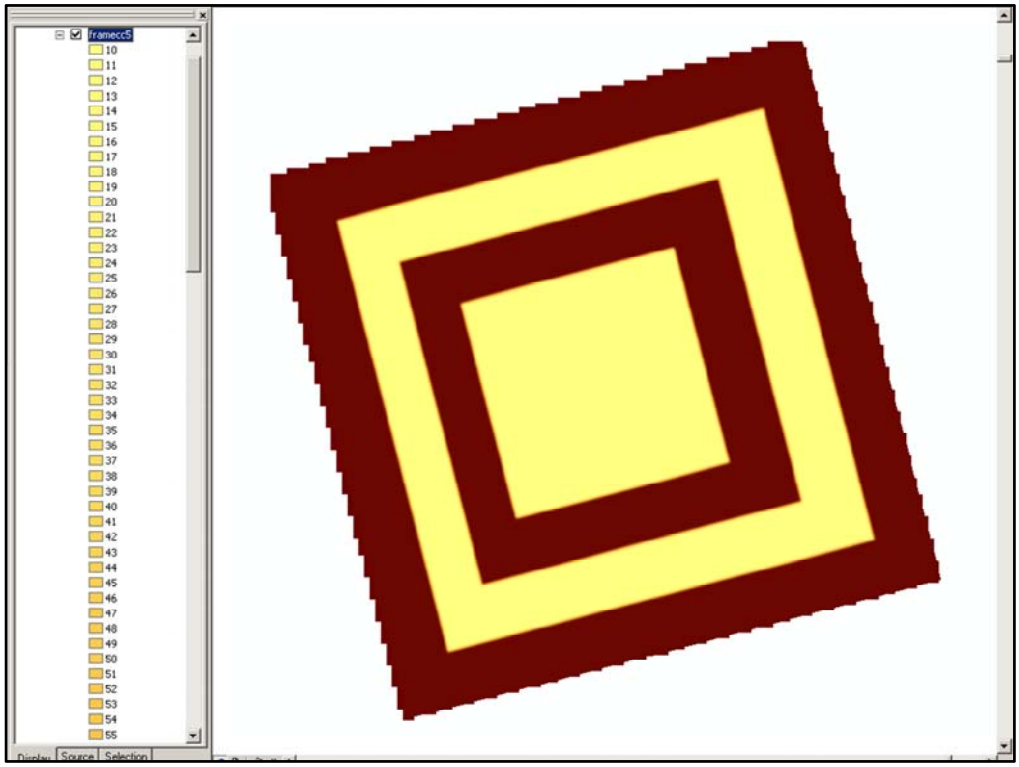
If we look at the CC'd version of the image, we see less jaggedness along the interior edges, as some fuzziness has been created.

The fuzziness is due to some of the averaged pixel values along the edges. There are now many new pixel values that have been created in addition to than the original pixel values.

But the Image has the **Same mean as the Original and NN'd images!**

The application of the CC is an attempt to better define the edges … The apparently smoother transition of pixel differences is what makes this resampling algorithm most useful.

CC can be used to make a "prettier, more representative picture" of the original, especially if we reduce the pixel size to 5m during the CC resampling, as shown in this …



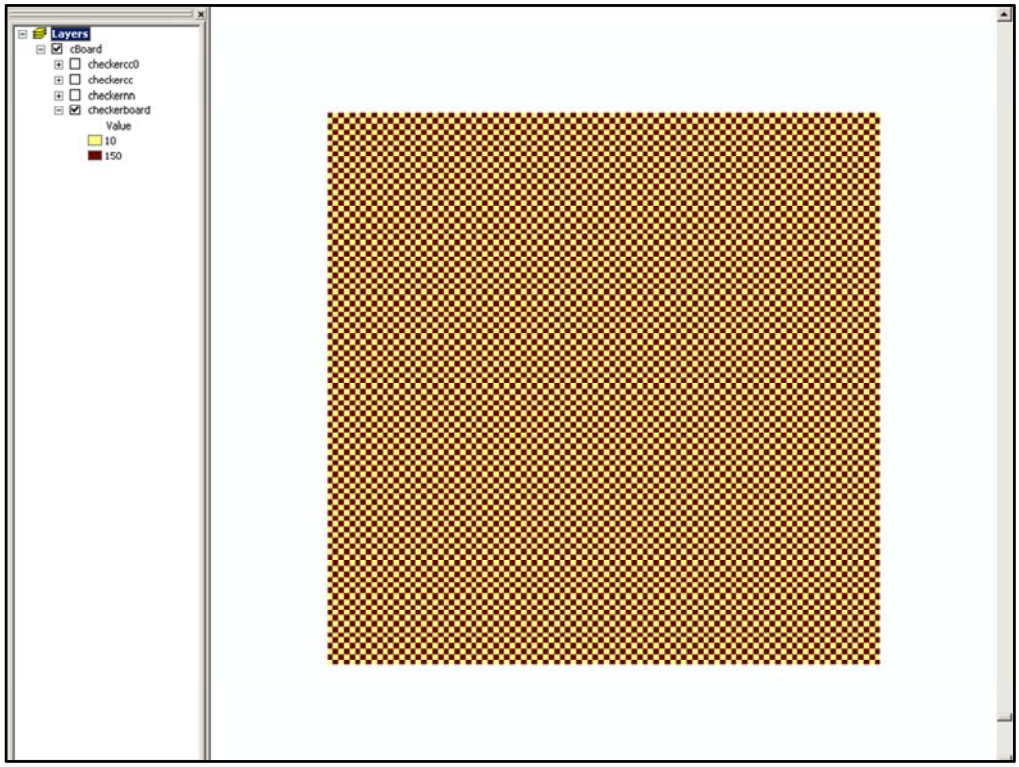
… next slide.

Here I have changed the output pixel size from 30 meters to 5 meters.

We see the benefits of CC, as it creates much sharper image boundaries – just as were shown in the original.

This further illustrates how the CC approach is most useful in recreating the sharp edges – but at what cost ?

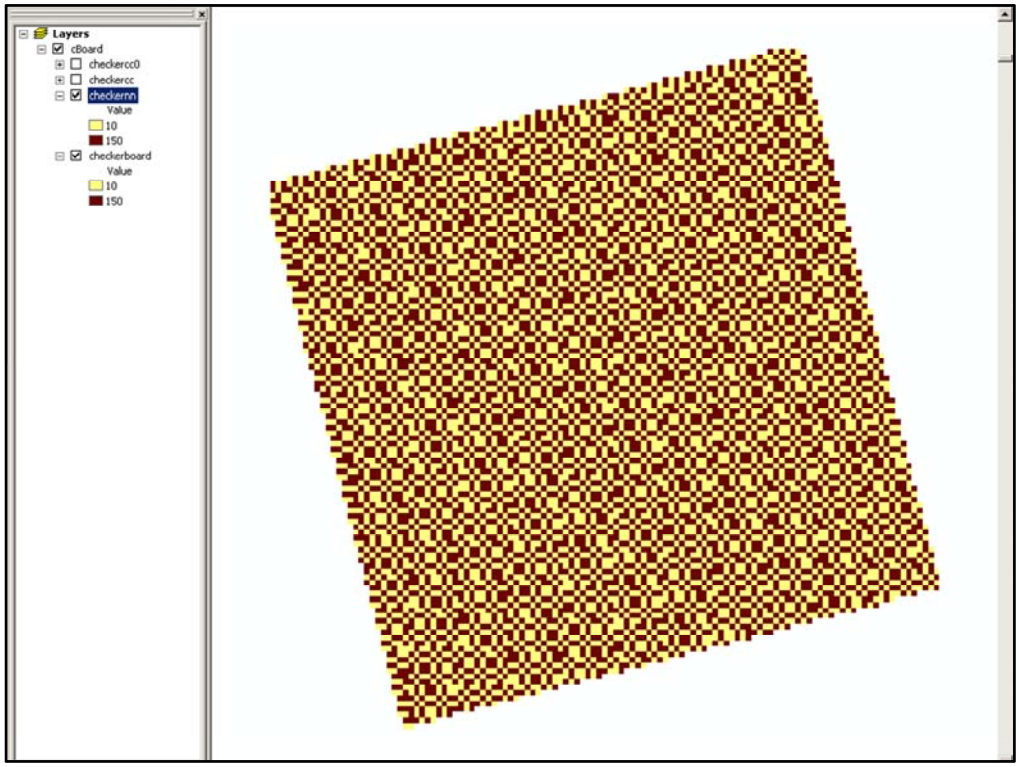
Let’s go on and look at another very different example …



… a 100x100 checkerboard pattern which represents a bimodal distribution of the values of 10 and 150

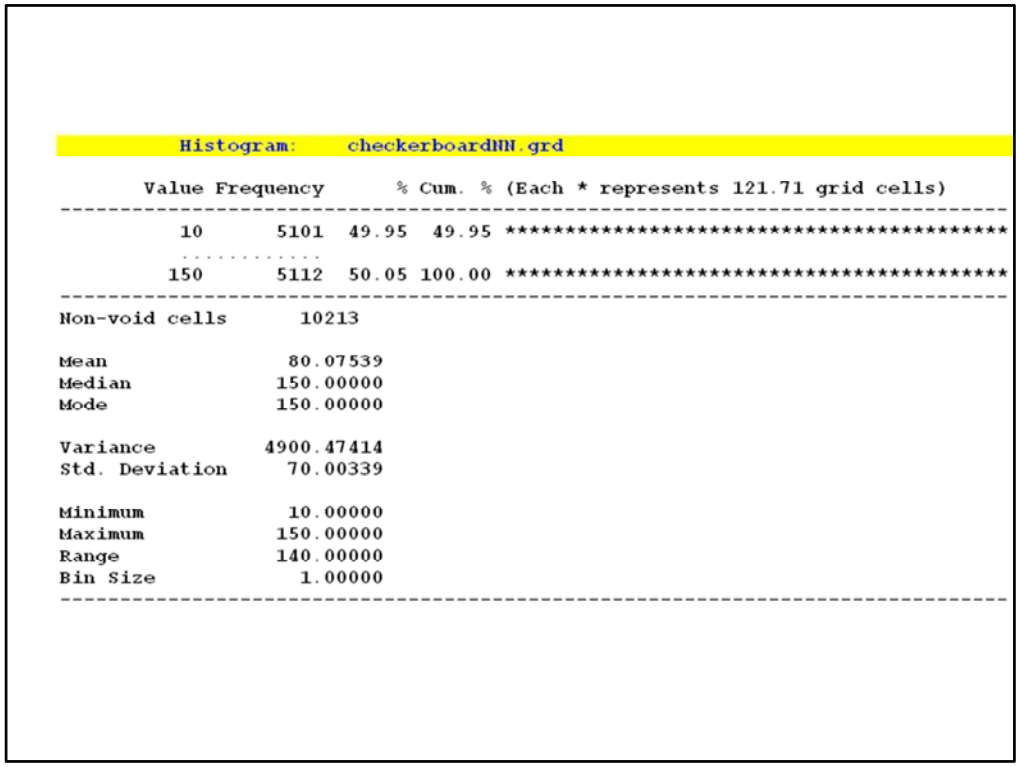
The average pixel value of this image is 80.

There are the same number of pixels of each value.



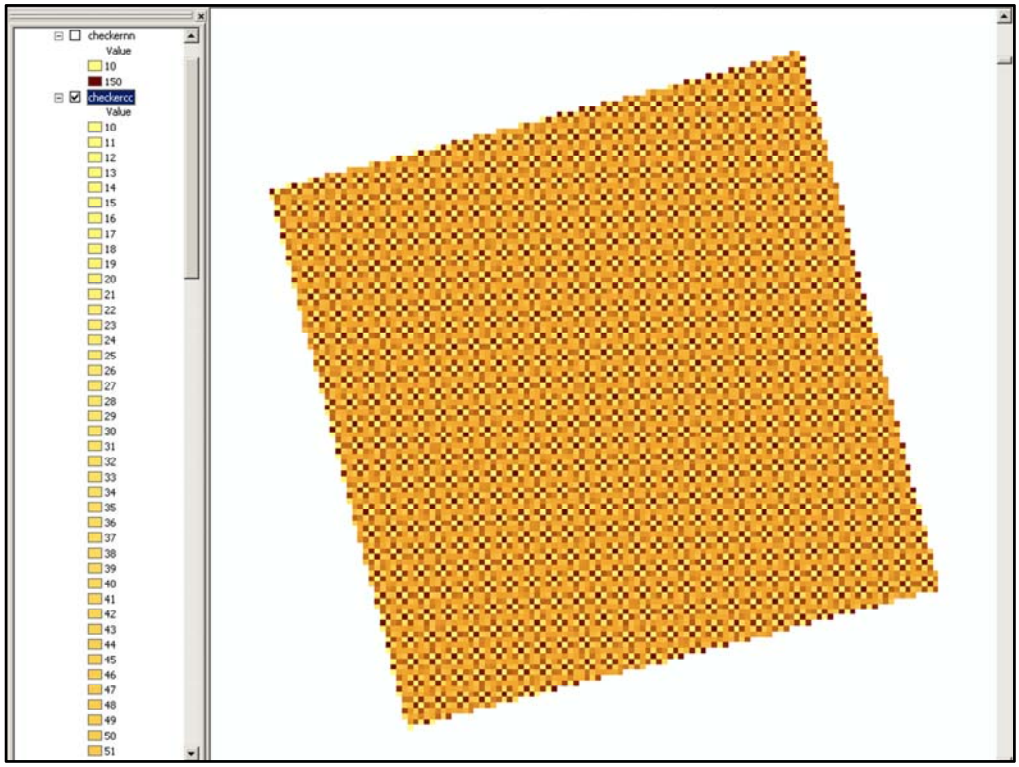
The results of the NN resampling shows some minor positional shifting of values – by as much as ±21m – or half the (diagonal) size of a pixel.

There is some obvious Spatial Distortion, evidenced by some new clumpiness that has been introduced, **[next]**



```
          Histogram:     checkerboardNN.grd

       Value Frequency      % Cum. % (Each * represents 121.71 grid cells)
-------------------------------------------------------------------------------
          10      5101  49.95  49.95 ******************************************
           .............
         150      5112  50.05 100.00 ******************************************
-------------------------------------------------------------------------------
Non-void cells      10213

Mean                80.07539
Median             150.00000
Mode               150.00000

Variance          4900.47414
Std. Deviation      70.00339

Minimum            10.00000
Maximum           150.00000
Range             140.00000
Bin Size            1.00000
-------------------------------------------------------------------------------
```

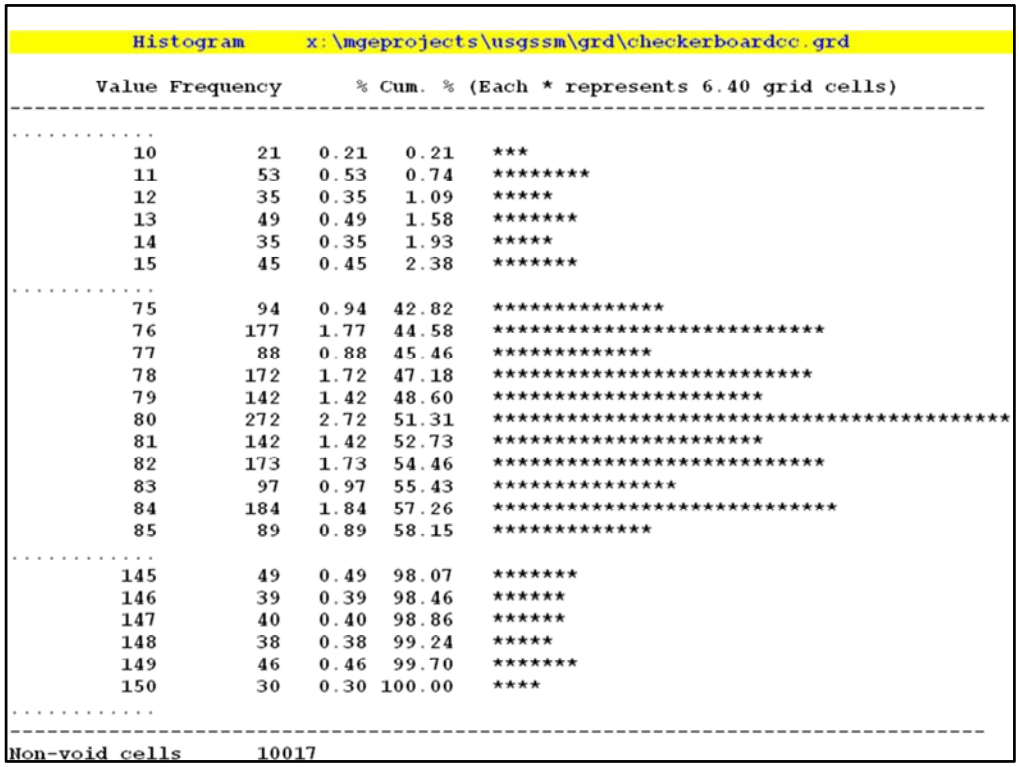
but all the original values are maintained and the average values (mean, mode, and median) and the variance value of this image are maintained.



If we look at the CC'd checkerboard image, we see a very different result in which the checkboard pattern has been replaced by a fuzzy systematically patterned image.

In this case, many, many pixel values have been altered.

If we look at the CC'd image values, we find that [next]



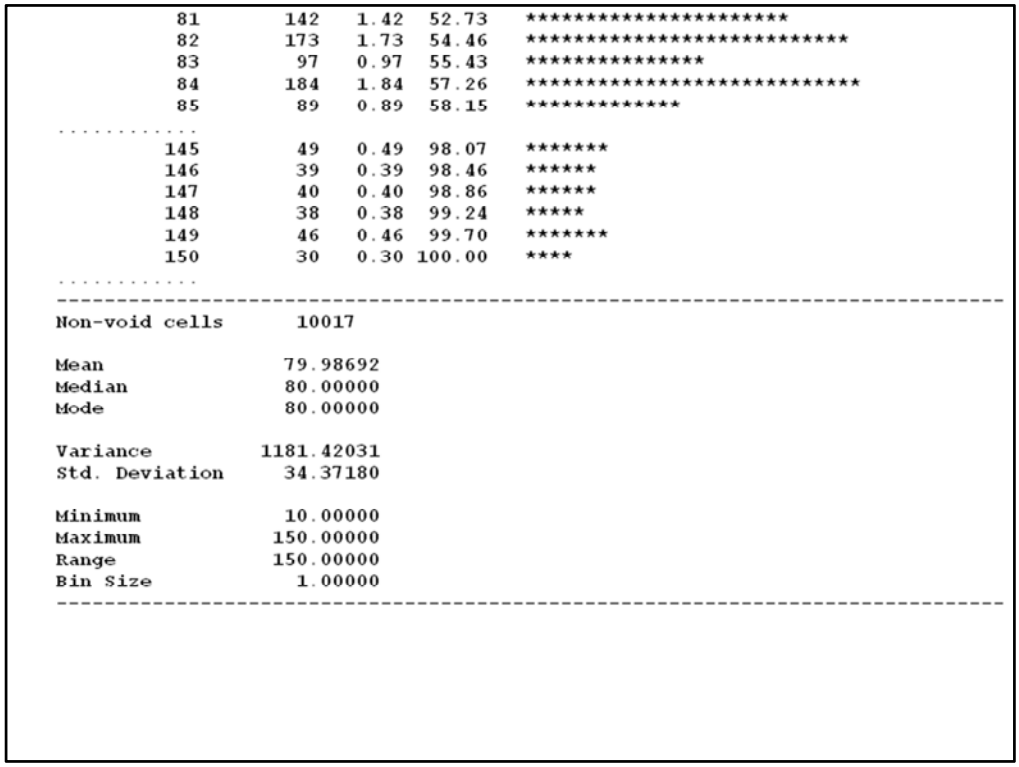
Histogram x:\mgeprojects\usgssm\grd\checkerboardcc.grd

Value	Frequency	%	Cum. %	(Each * represents 6.40 grid cells)
10	21	0.21	0.21	***
11	53	0.53	0.74	********
12	35	0.35	1.09	*****
13	49	0.49	1.58	*******
14	35	0.35	1.93	*****
15	45	0.45	2.38	*******
75	94	0.94	42.82	**************
76	177	1.77	44.58	***************************
77	88	0.88	45.46	*************
78	172	1.72	47.18	*************************
79	142	1.42	48.60	*********************
80	272	2.72	51.31	**
81	142	1.42	52.73	*********************
82	173	1.73	54.46	*************************
83	97	0.97	55.43	***************
84	184	1.84	57.26	****************************
85	89	0.89	58.15	*************
145	49	0.49	98.07	*******
146	39	0.39	98.46	******
147	40	0.40	98.86	******
148	38	0.38	99.24	*****
149	46	0.46	99.70	*******
150	30	0.30	100.00	****

Non-void cells 10017

the 10/150 bimodal distribution has been altered to form a **very flattened normal distribution**

that extends from 10 to 150 and now tends towards the average pixel value of 80 [next]

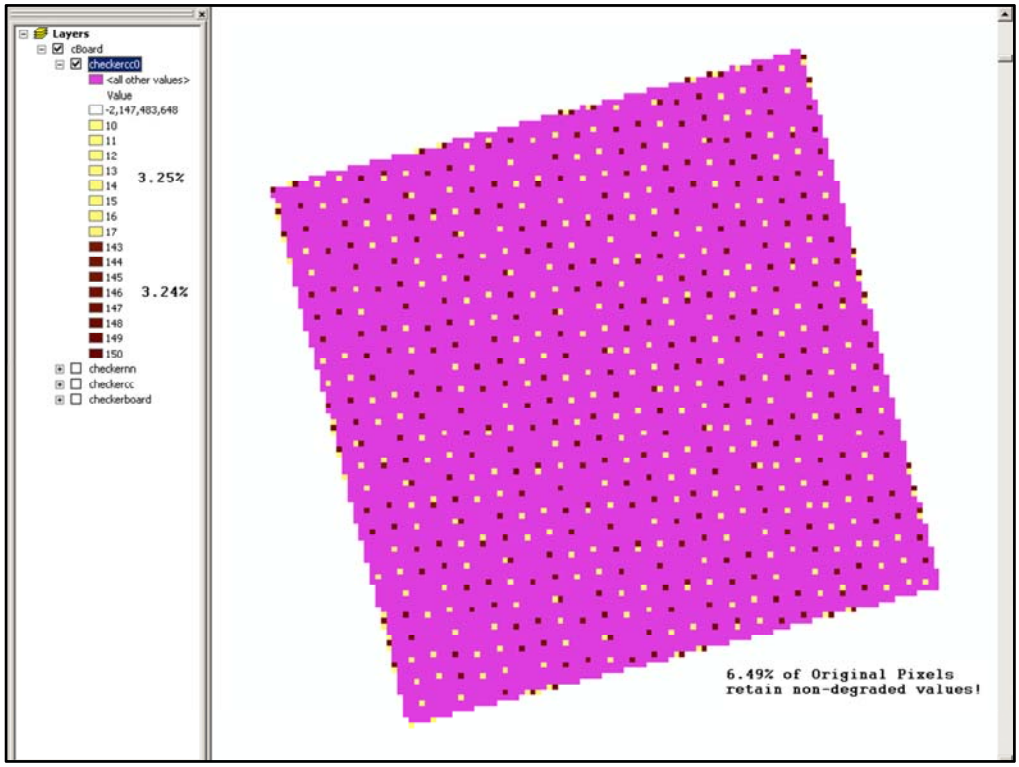


```
            81       142   1.42  52.73   **********************
            82       173   1.73  54.46   ***************************
            83        97   0.97  55.43   ***************
            84       184   1.84  57.26   ****************************
            85        89   0.89  58.15   *************
.............
           145        49   0.49  98.07   *******
           146        39   0.39  98.46   ******
           147        40   0.40  98.86   ******
           148        38   0.38  99.24   *****
           149        46   0.46  99.70   *******
           150        30   0.30 100.00   ****
.............
--------------------------------------------------------------------------------
Non-void cells      10017

Mean               79.98692
Median             80.00000
Mode               80.00000

Variance         1181.42031
Std. Deviation     34.37180

Minimum            10.00000
Maximum           150.00000
Range             150.00000
Bin Size            1.00000
--------------------------------------------------------------------------------
```

This image has the same mean of approximately 80 –

But now the resampled image has only 1/4th the variance … 1181 versus 4900.

Could this decreased variance be why some might consider the CC'd image a better representation of the original image ?



Ultimately many, many averaged values have been introduced into the resampled image.

Some of these new artifacts of the CC resampling algorithm are shown as purple in this example.

The purple values represent averaged pixel values that do not fall within [what I typically observe are] statistical tolerances of 2 standard deviations of mean training area pixel values if centered on the original pixel values of 10 and 150, AS SHOWN IN THE LEGEND.

In this case, the replacement of the original values occurs on a very large scale …

If I identify all of the pixels that are within the typical statistical limits of our spectral training data (10-17 and 143-150)

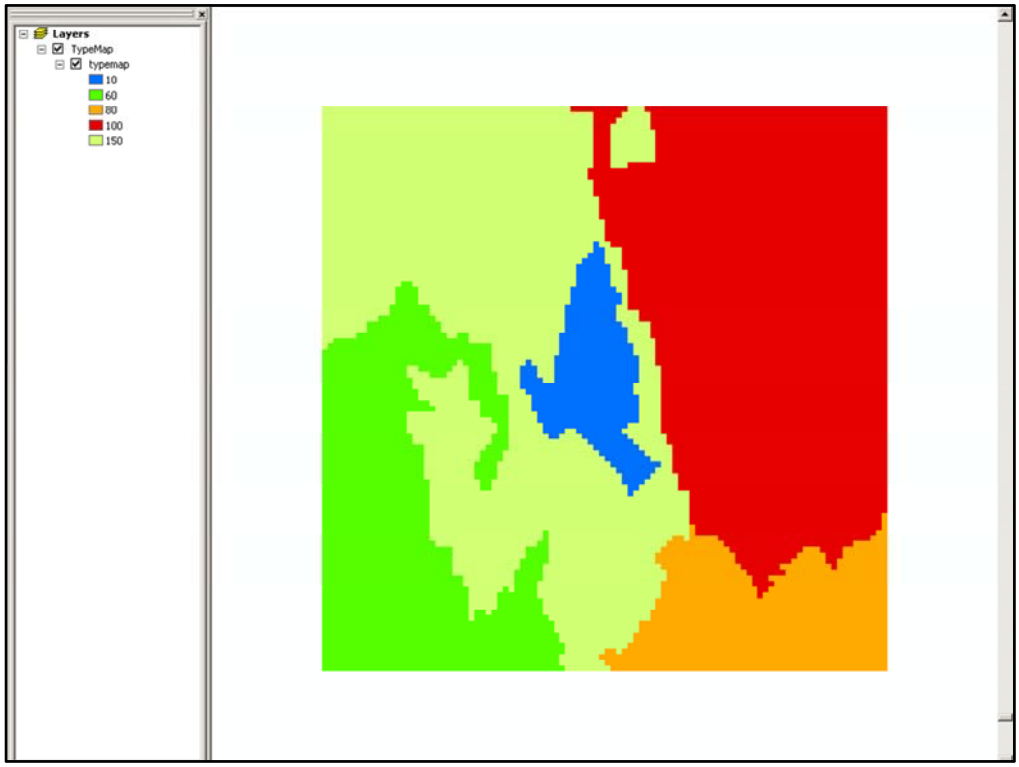
I find that only 6.49% of the original pixel values are now represented in the CC'd image.

In this CC'd image we do not see the results of SPATIAL DISTORTION, but rather STATISTICAL/SPECTRAL DISTORTION

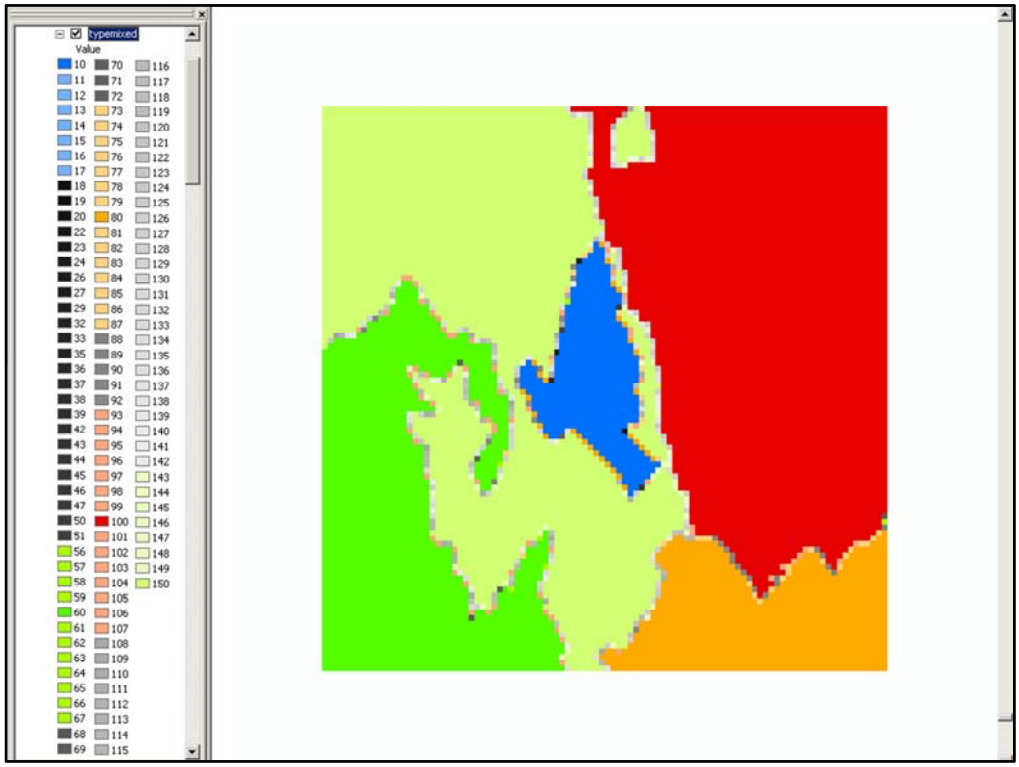
10 – 17 = 325 or 3.25%

143-150 = 324 or 3.24%

Total of 6.5% ~ the SAME original pixel value.



Let look at a hypothetical example that might represent different vegetation/landscape features that are homogeneous areas …



The reality is that we do have mixed pixels along the edges of the types …

because the image pixel boundaries do not align with the vegetation/landscape feature boundaries.

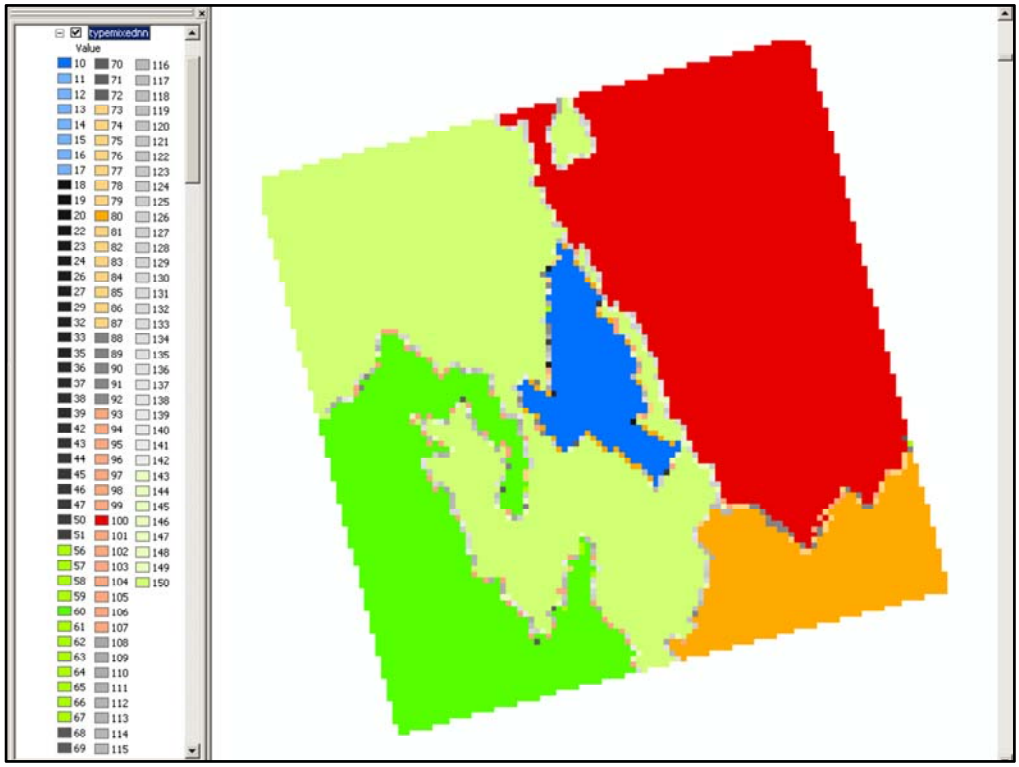
Here I have created a band of mixed pixels along the type edges that is typically one pixel wide, but is sometimes as much as two pixels wide.

The mixed shades of gray shown in the legend and map represent locations now having pixel values that fall outside the typical statistical ranges of ±2SDs of the mean values of the homogeneous types they are next to.

Some "mixed" pixels actually represent pixels from other types that are present elsewhere in the imagery and therefore represent CONFUSED pixel values – see ORANGE pixels along Lake boundary.

I wonder if the existence of such mixed pixels is why some may believe the CC algorithm doesn't really impact the image data as there are already mixed pixels – what are a few more ?

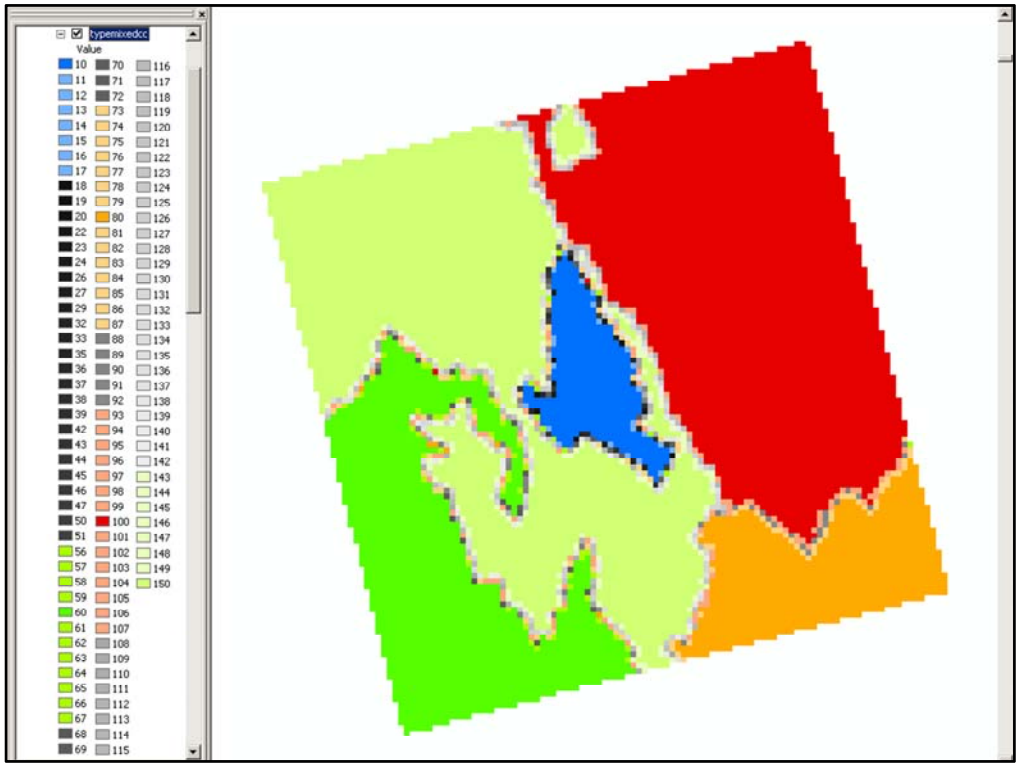
The NN image [next]



looks like this – it is rotated, but the values are preserved.

The mixed pixels still represent a band of pixels approximately one pixel wide.

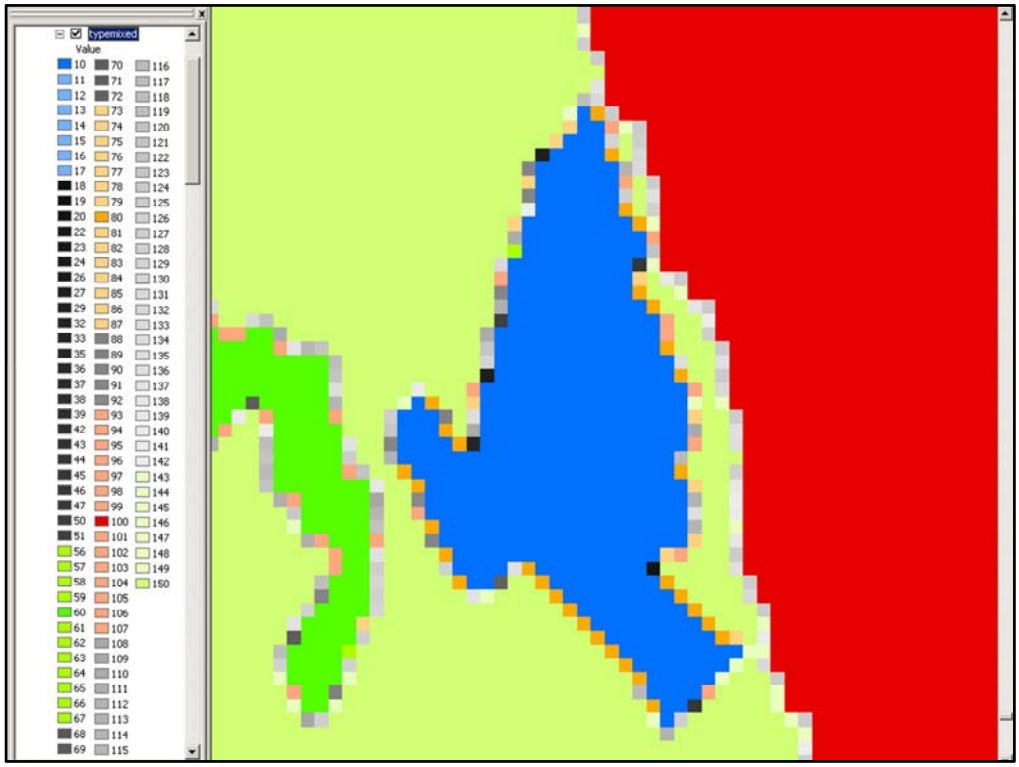
However, the CC'd image looks like **[next]**



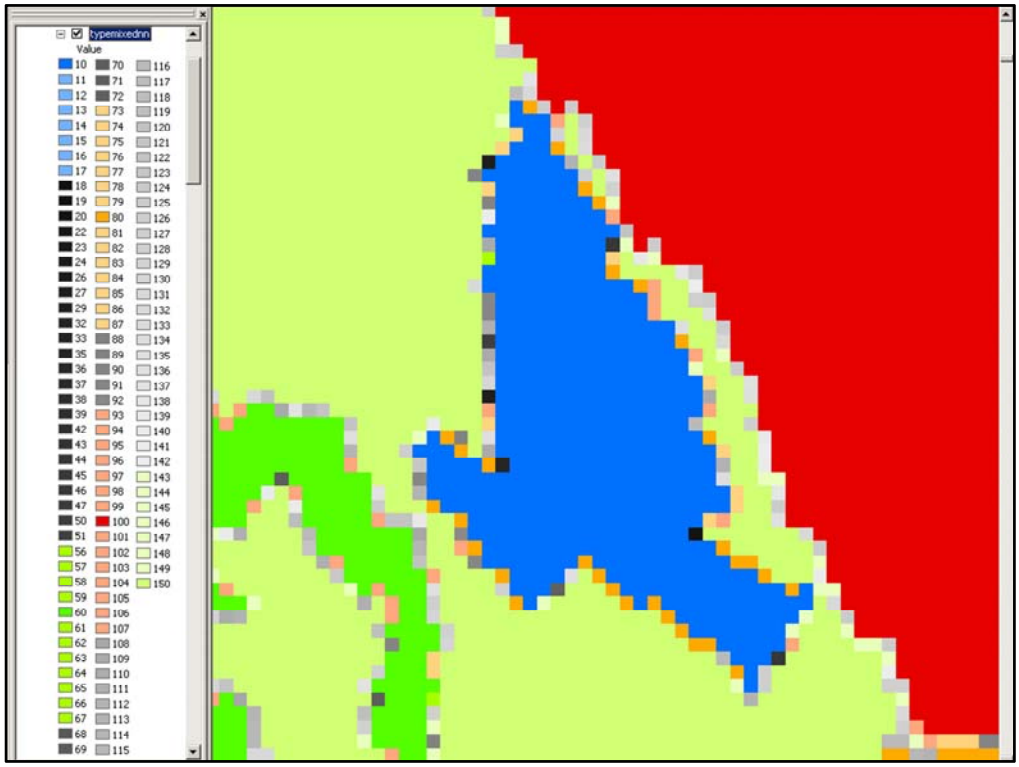
The CC's image looks like this …

The application of the CC has increased the number of mixed pixels by nearly 200% as there is now a band of mixed pixels along the edges that ranges from 2 to 4 pixels wide.

Let's look at this a little closer focusing on the 'lake' and surrounding area in this next example **[next]**



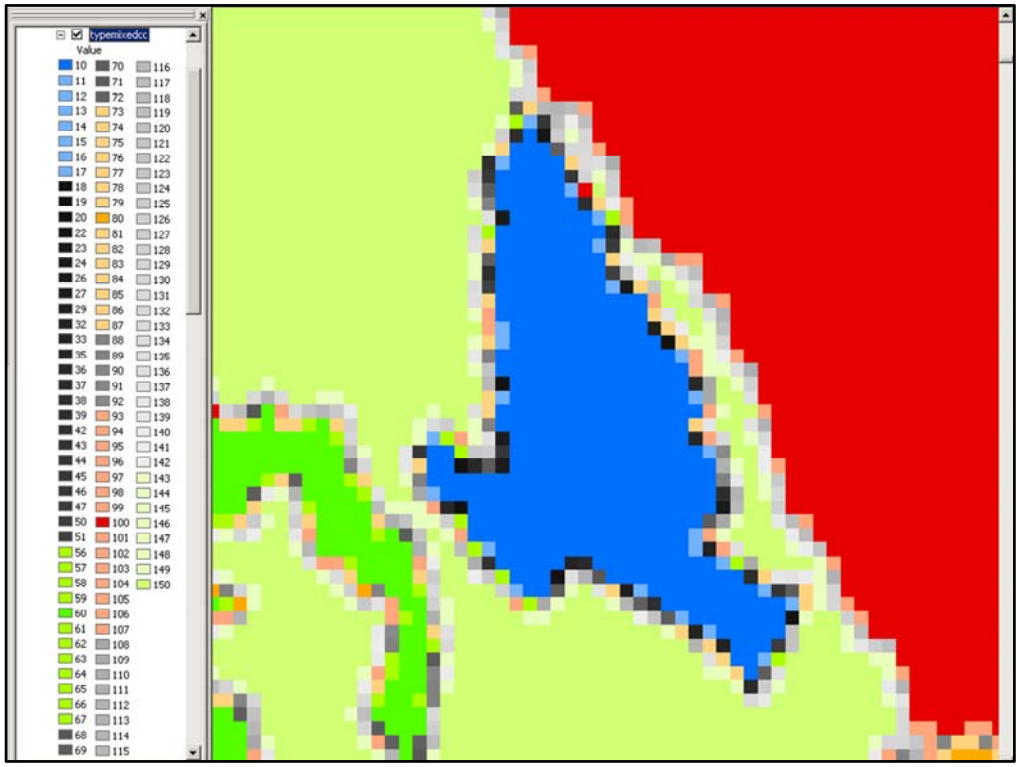
Here is the original type image having some mixed pixels along the area edges …



Here is the rotated NN image – note the preservation of the original pixel values. as well as the band of mixed pixels.

These is some slight shifting of locations, but the values are preserved.

[next]



However, the result of the CC algorithm is that there are now a lot more mixed pixels.

In addition, all of the original naturally occurring 'mixed' **pixel values have been altered.**

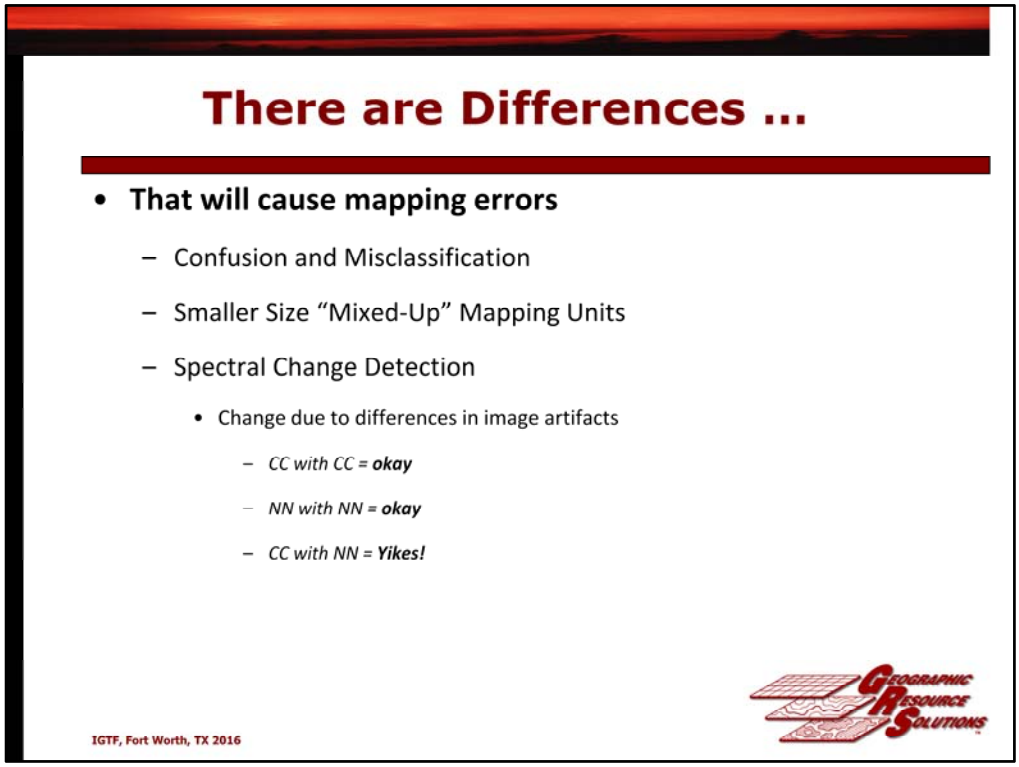
Three things have happened here …

1. **Many new "convoluted" values have been introduced that will likely be confused with vegetation/landscape type values from other parts of the project area that DO NOT actually occur in this particular area!**

2. **Small contiguous areas of mixed pixels have been created that will meet or exceed the minimum mapping unit size. Typical sizes of my past projects are 0.5 hectares which is about 1.2 acres or 6 pixels (see the north end of the lake). These convoluted areas are artifacts of the CC method that will lead to less accurate map data sets as they are large enough to form valid size polygons in the map data set.**

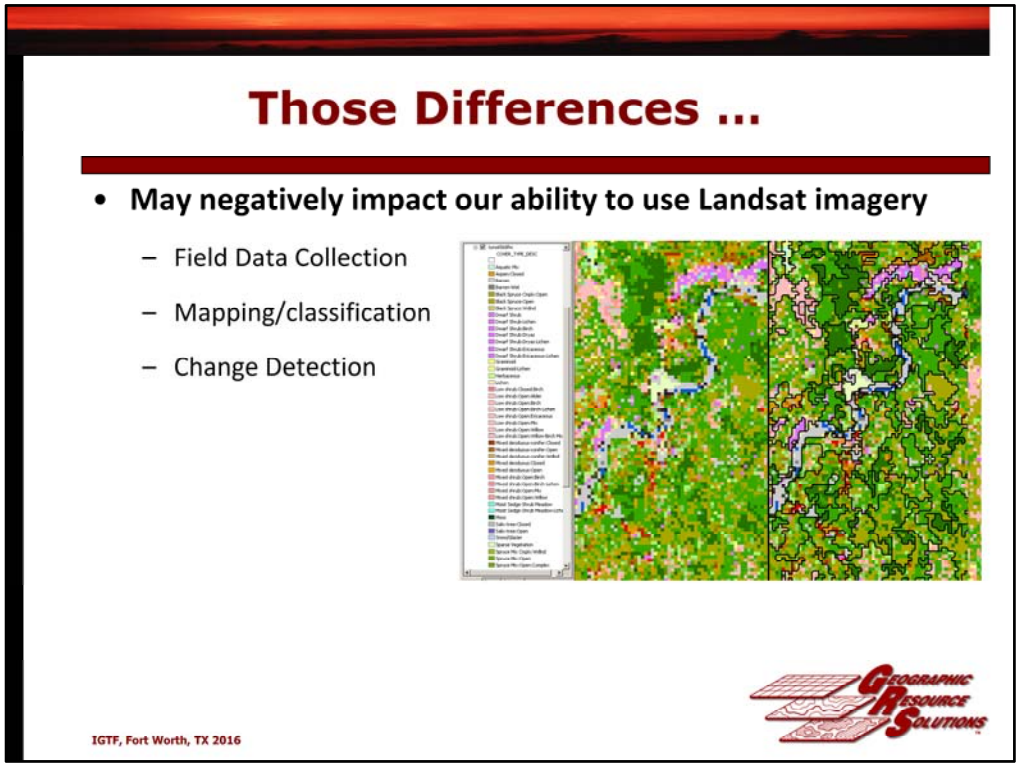
3. **There are now small "MIXED-UP" pixels in the output image. These are pixels that have replaced the original mixed pixels along the edges of type areas. Their values replace the original mixed pixel values.**

[See the top of the lake …]



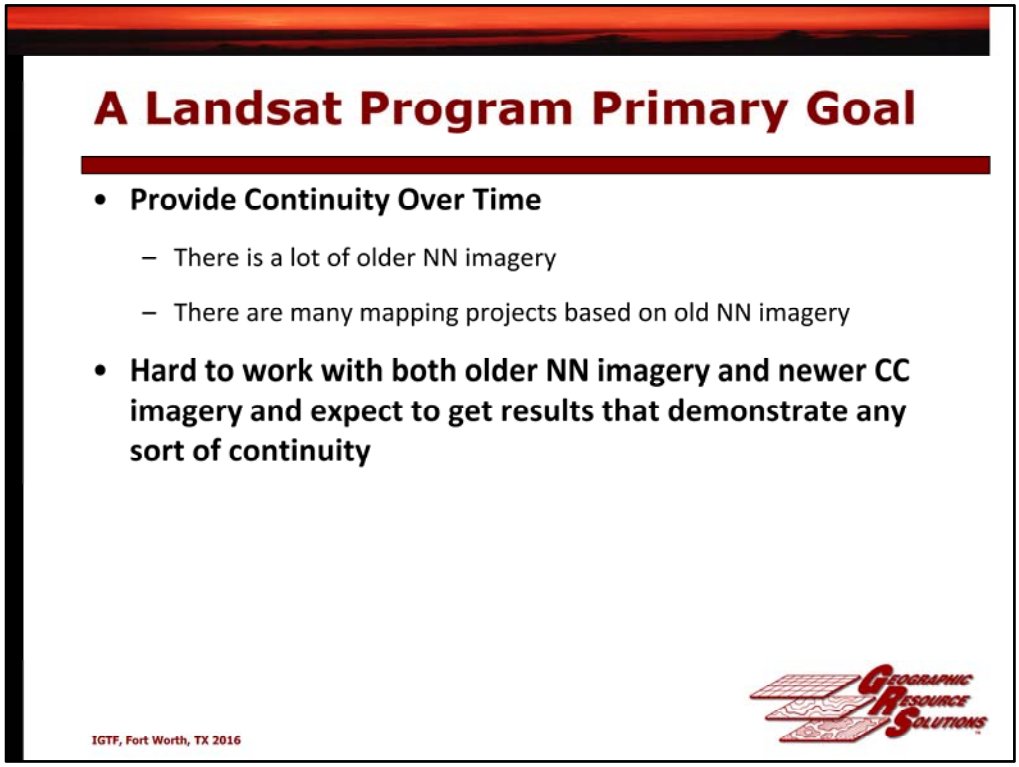
So there are some definite significant differences between the images developed using these two resampling algorithms.

I maintain that the CC'd imagery will cause problems relative to the NN'd imagery that include … [see slide]



Those Differences …

May negatively impact our ability to use Landsat imagery to perform analyses and develop accurate mapping data sets.



Now it just happens that one of the **Landsat Program's Primary Goals** is to provide continuity over time.

I believe the removal of the NN resampling option will make it difficult to work with both older NN imagery and the newer CC imagery and expect to get results that demonstrate continuity between the mapping and analytical products.

This past year a whole slew of new "Advanced" Landsat products have been introduced. There appears to be what I view as an attempt to produce a standardized or normalized "Surface Reflectance" image across paths and acquisition dates make a "pretty backdrop" of imagery for web and mobile applications.

Unfortunately, imagery produced using the NN Resampling is not an ADVANCED option!

Higher Level Science Data Products

Surface Reflectance and other high level science data products can be ordered through the following pages:

USGS Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) On Demand Interface (https://espa.cr.usgs.gov/)

To begin the order, upload a text file (*.txt) listing one Landsat Level 1 or MODIS scene identifier (filename) on each line. Scene identifiers can be found in the search results on EarthExplorer (http://earthexplorer.usgs.gov) or GloVis (http://glovis.usgs.gov).

After uploading the scene list text file, a number of options can be selected, including:

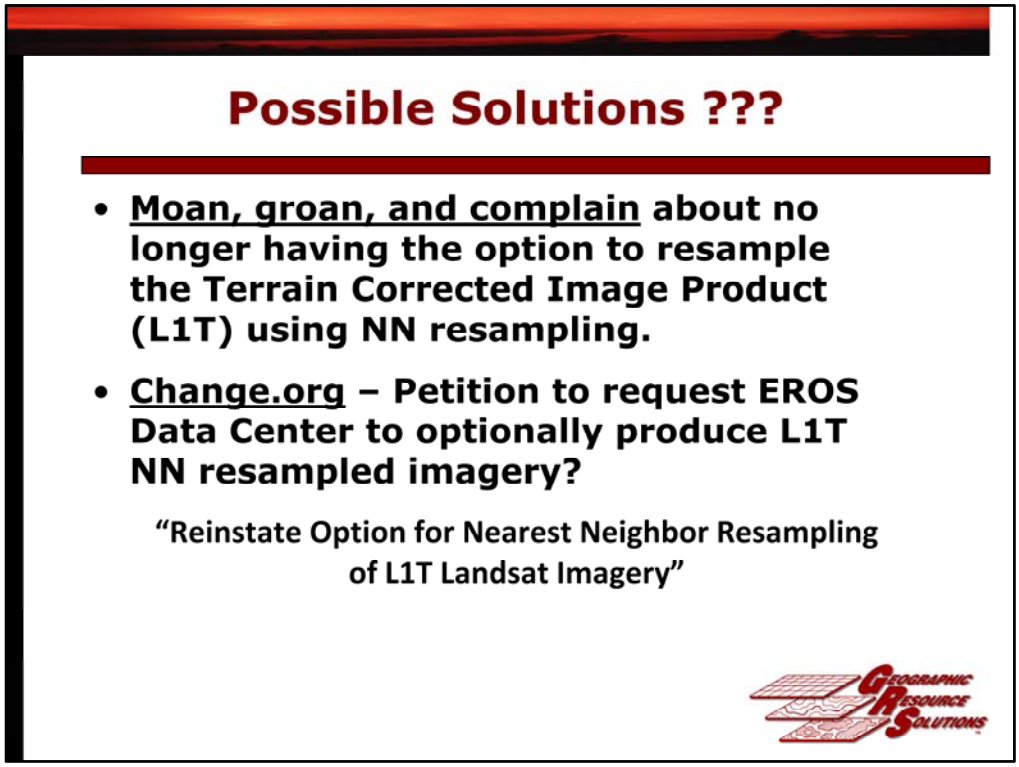
Source products (Original input Level 1 product or metadata)

Top of Atmosphere Reflectance, Surface Reflectance (SR), or Band 6 Brightness Temperature products

Surface Reflectance-based Spectral Indices (NDVI, NDMI, NBR, SAVI, EVI

Customizable output options: data format, reprojection, modifying the image extents, and pixel resizing

Intercomparison and Output Product Statistics Plotting



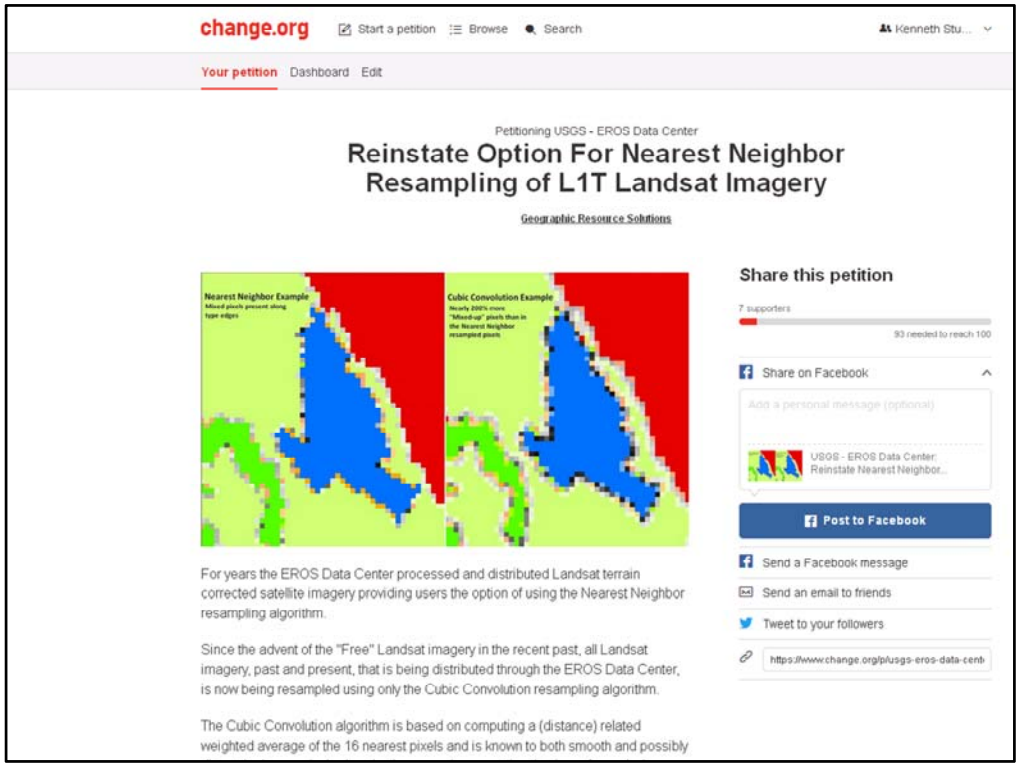
So what can we do ?

We'll, I decided to present this information and see what happened. I was hoping someone might tell me this was all unnecessary and that there was an easy way to order and acquire the NN resampled Landsat imagery, but that has not happened.

Instead I am finding that most users don't really understand the differences that result from these resampling algorithms.

In addition, they are fully trusting that the EROS Data Center Landsat/satellite image experts are providing the imagery that they need for their analyses and mapping efforts, as that was the initial impetus behind the Landsat program. Why would they change to being more concerned about generating a "pretty picture" than providing imagery to support our mapping and analysis applications ?

Finding no readily available solution I decided to be proactive and start a petition seeking reinstatement of the NN resampling methodology as an option.

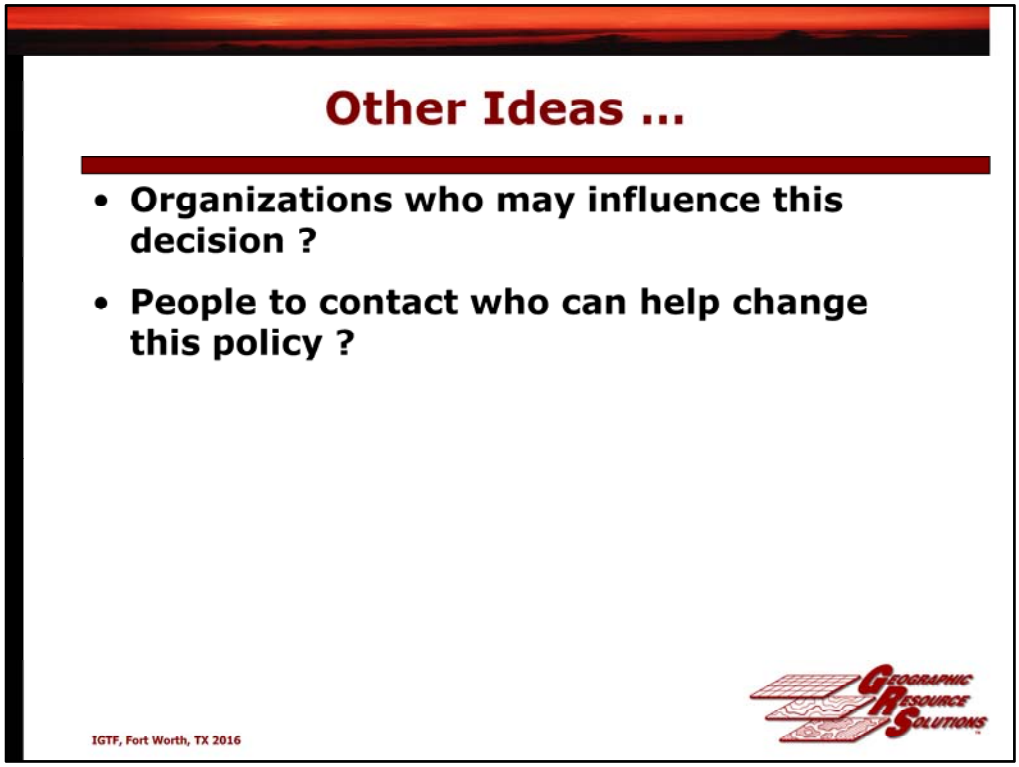


This morning I started a petition at Change.org to reinstate the NN resampling option.

I encourage any of you who need this type of Landsat imagery and support this resampling methodology to sign this petition.

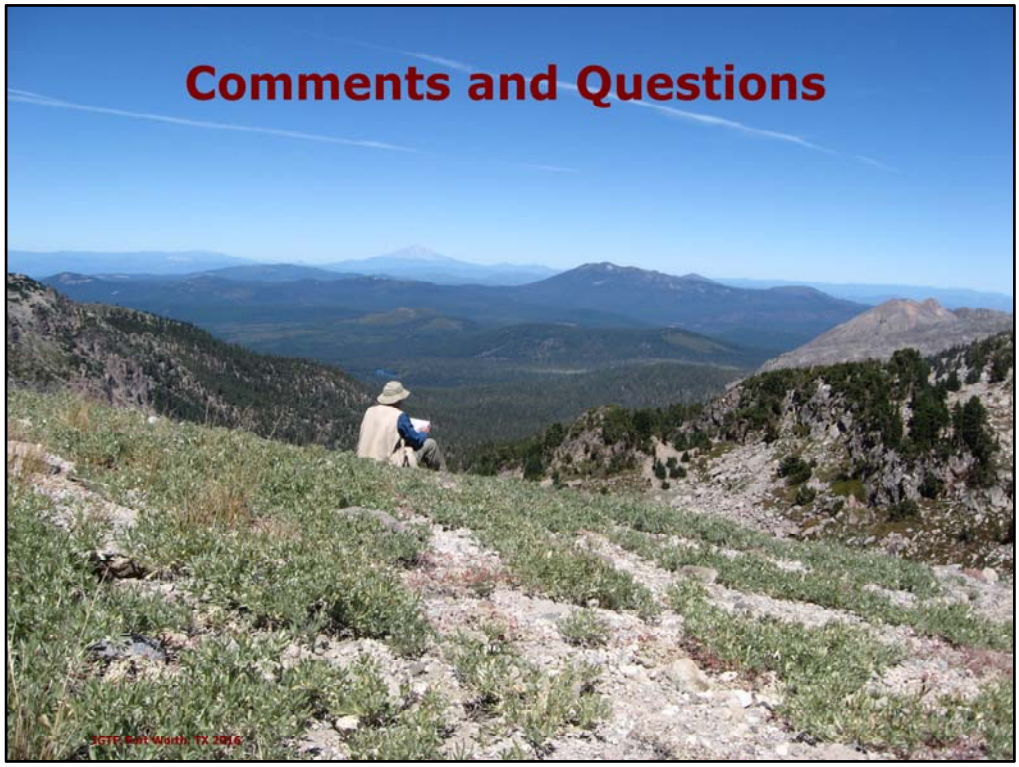
Link: https://www.change.org/p/usgs-eros-data-center-reinstate-nearest-neighbor-resampling-of-l1t-landsat-imagery

Just maybe we can convince the “powers that be” at USGS and the EROS Data Center to change their present policies and provide us with the option to have Terrain Correction performed using the NN resampling algorithm.



There may be other ways to effect change through other organizations and influential decision makers.

I am open to suggestion !



Thanks for you attention to this matter.