USDA-FSA
Aerial Photography Field Office
Fugro EarthData Yazoo County, Mississippi
GeoSAR Evaluation Report

Prepared by Geospatial Services Branch Service Center Support Section

December 17, 2008

Table of Contents

Introduction ……..... 1

Deliverables ……..... 1

Acquisition …….. 2

CLU Maintenance ……... 2

Horizontal Accuracy …… 4

Classification of GeoSAR Data for Crop Indentification …… 5

Using the Unsupervised Classification in Erdas Imagine …….. 5

Subtracting the P band from the X band …… 10

Comparing X and P band DEMs to the NED ……. 11

Conclusion …….. 15

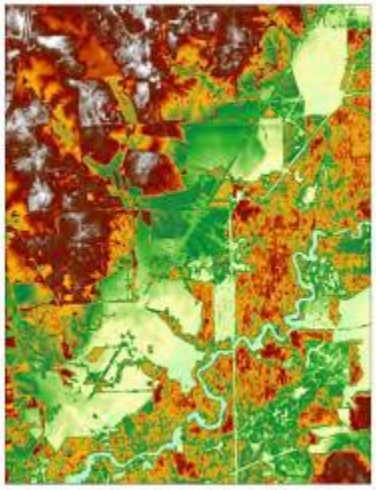
Appendix Documents
Working with GEOSAR Data in ERDAS Imagine
Final Report for USDA J07-0020

Introduction

The GeoSAR data evaluated was collected by Fugro EarthData, under contract with the USDA Farm Service Agency's Aerial Photography Field Office. GeoSAR is a system of data capture and is a type of Airborne Interferometric Synthetic Aperture Radar (IFSAR). It was collected in Yazoo County, Mississippi on August 29 - 31 2007. The purpose of this collection was to demonstrate the utility of Airborne IFSAR data to replace optical imagery data collected in the National Agricultural Imagery Program (NAIP). The GeoSAR data was evaluated on elevation quality versus the NED, all weather acquisition capabilities, its potential as a CLU creation source, semi-automated crop identification and any other general applications.

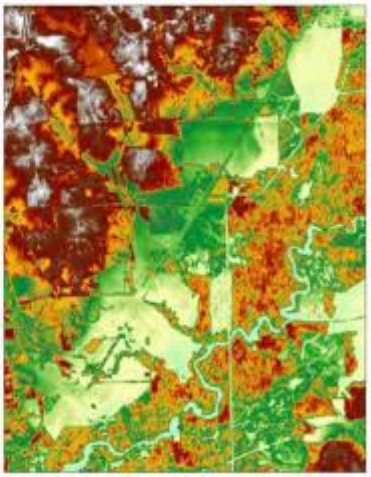
Deliverables

X band Magnitude 3 meter



X band DEM 3 meter

P band Magnitude 5 meter

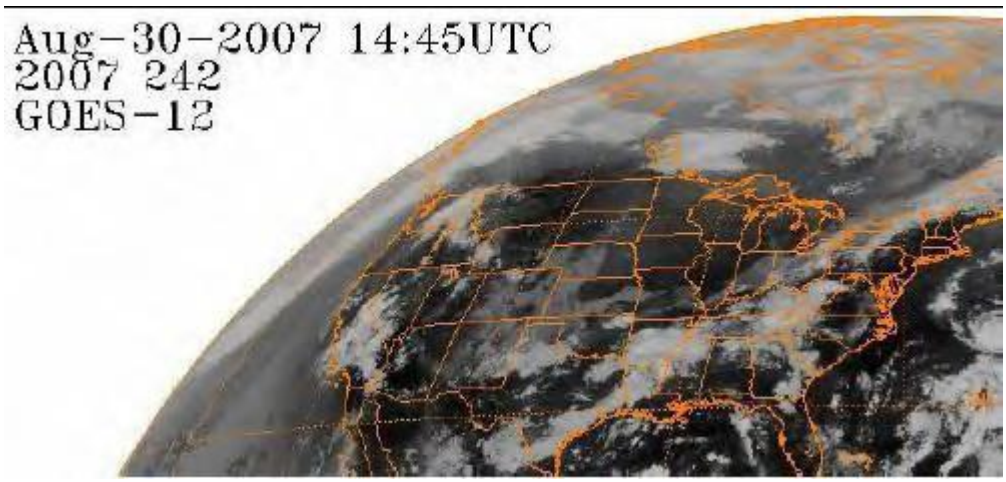


P band DEM 5 meter

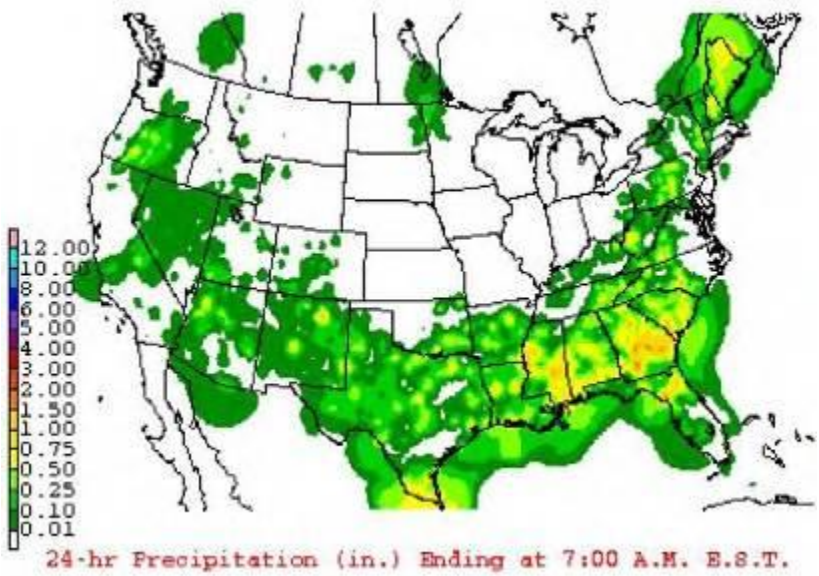
The deliverables came in three different geographic subsets – quads, swaths and counties. The data consisted of 32 bit Geotiffs and 8 bit MrSIDs, most analysis was processed using the 32 bit data. Magnitude refers to the raw return value relative to surface type.

The DEMs were created from these magnitude values and converted to equal elevation. No FGDC metadata was supplied for this data. Future deliveries should include this data.

Acquisition



GOES Weather Satellite



Precipitation Weather Map

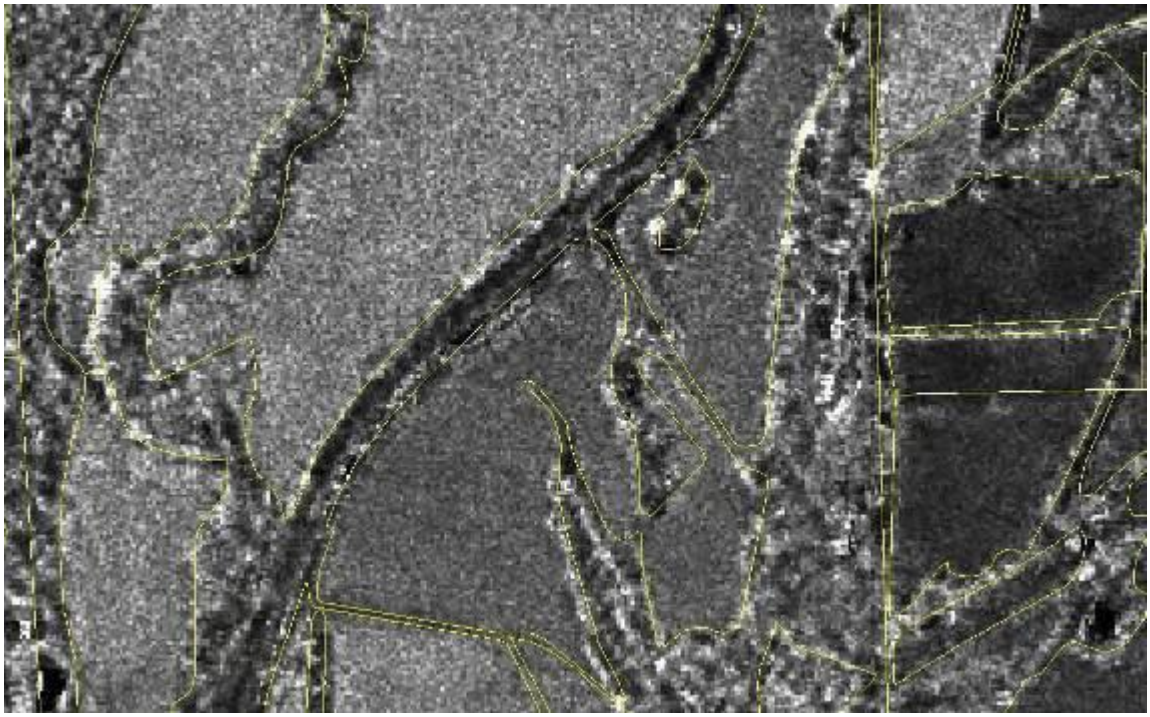
The GEOSAR data was collected over 3 days of partly cloudy weather conditions and low light. The quality and usability was excellent considering the collection conditions. The ability to be collected at any time of day during most any conditions would have definite benefits to the user.

CLU Maintenance

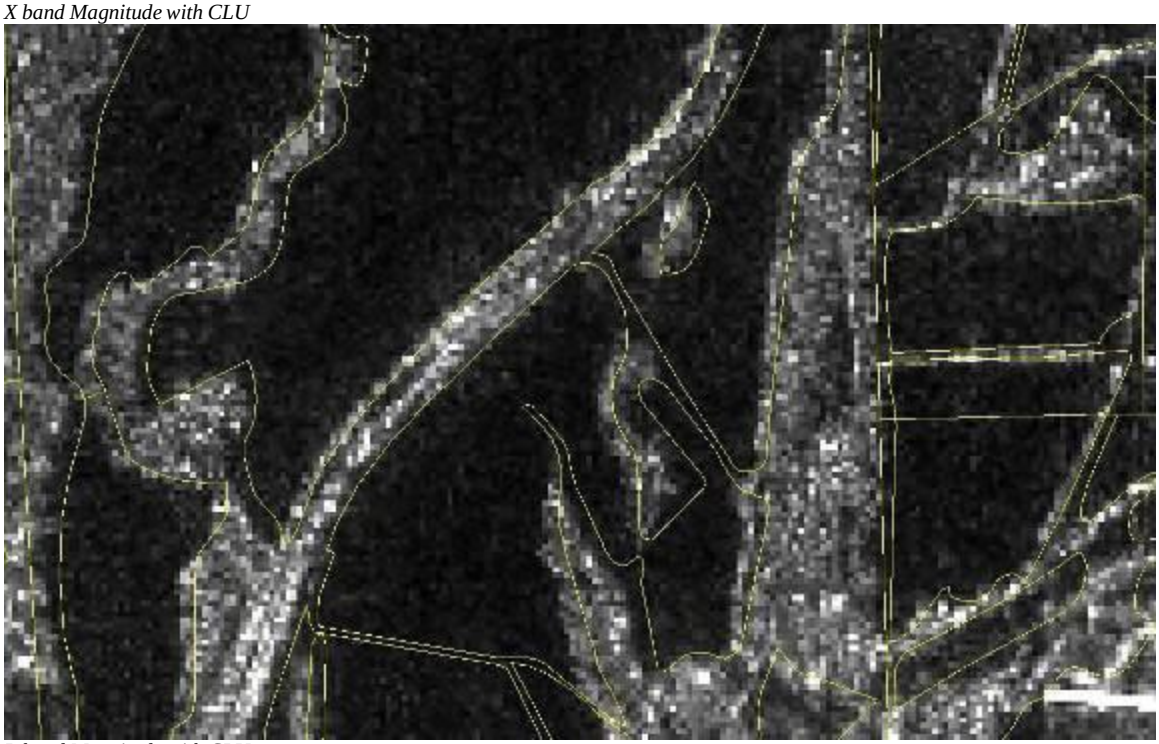
Analysis for the applicability of creating, editing and maintaining CLU with X and P band magnitude datasets yielded inconsistent results. The coarseness (3 and 5 meter) of the data makes determining discrete field boundaries difficult. If 2 CLU polygons

bordering each other contain a different type of vegetation it is possible to see the boundary between the areas and would allow for editing or creating CLU. However, if the vegetation type is the same, it is often impossible to see the small buffer area between fields which is needed to draw CLU. This would lead to creating large CLU where there should be more then one CLU to cover the area. The GEOSAR data could be used to supplement the imagery used to edit CLU, but it would be especially difficult to use GeoSAR without the imagery. Without experience with the GEOSAR data, it would be extremely difficult to edit CLU or identify basic features.

2007 NAIP with CLU



X band Magnitude with CLU

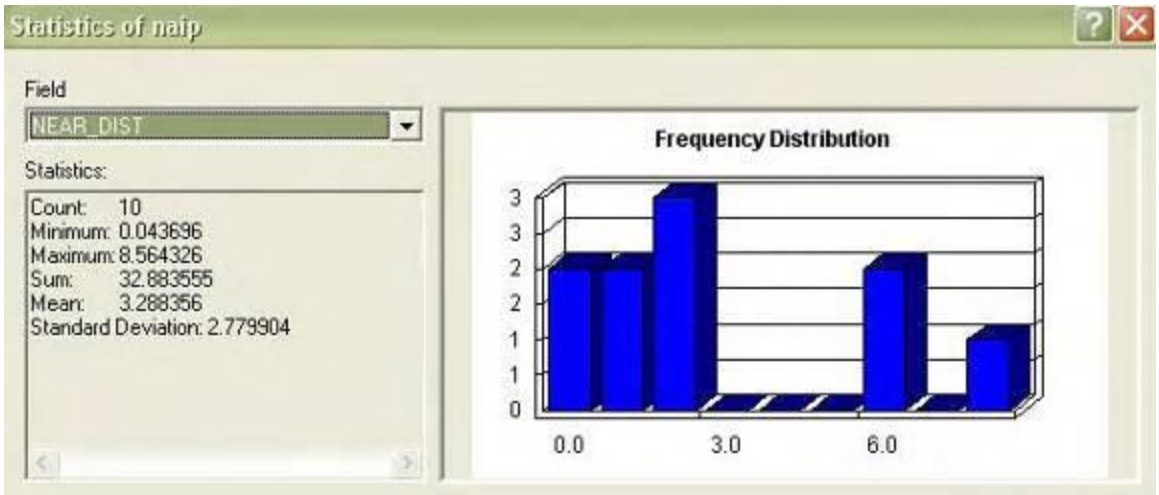


P band Magnitude with CLU

Horizontal Accuracy

The GEOSAR data horizontal accuracy was checked against the 2007 NAIP. This was a relative accuracy inspection process. Ten points were selected that were easily identifiable on both the NAIP image and the GEOSAR image. The mean offset was just over 3 meters with a maximum of 8 meters of difference. These offsets are a bit large,

but considering that the GEOSAR data resolutions are 3 and 5 meters respectively, the offsets are somewhat equivalent to a 1 or 2 meter offset of NAIP, well within tolerance. Based on this, test data derived from the GEOSAR collection would overlay correctly on NAIP and vice versa.



Classification of GeoSAR Data for Crop Identification

The classification procedures produced patterns which could be seen approximating the outlines of fields, as represented by the CLU polygons. The analysis used the same process with four different tiles, in order to compare the results accurately. This was done while bearing in mind that there could be many potential variables when running the processes, and many different tests which could be conducted.

Using the Unsupervised Classification in ERDAS Imagine

In the first tile examined, the pattern was surprisingly clear and similar to the Fugro EarthData report. The imagery was compared to CLU files with *crop* data collected through the crop maintenance tool included in the attribute table. This crop description (in the field DESC_CROP) assigns a [Mississippi] crop to the CLU polygons. It does not allow for fields which may contain more than one crop. Fields in Yazoo County contained both harvested and un-harvested crops at the time of data collection, and this was obviously not recorded in the attribute table. The upland cotton polygons, shown with magenta outlines in the image below, as well as sweet potatoes (aqua lines) stood out very clearly as classes 5 and 6 in the unsupervised Classification. These classes are represented in dark green in the image below. The neighborhood classification also clearly displayed areas of lower elevation, such as water and road surfaces. The hypothesis going forward was that other tiles would exhibit the same patterns after the same processes were run

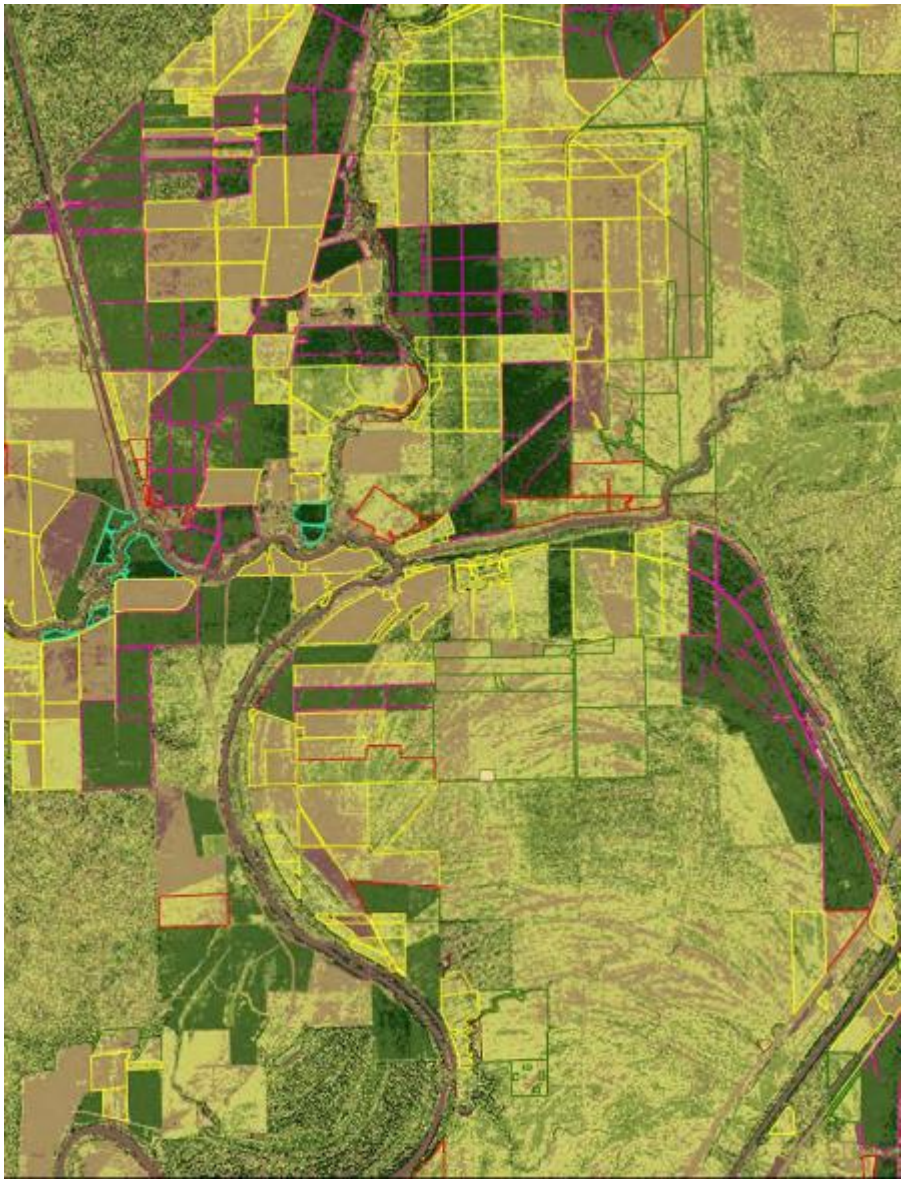


Figure 1: Tile G-6 showed upland cotton (in magenta) and sweet potatoes (aqua) very clearly, as darker green was used to symbolize classes 5 and 6. Corn (yellow) and soybeans (red) had displayed more variety in the classification.

Tile H-5 displayed the same pattern as tile G-6. Cotton appeared easily, but pixel classes for soybeans and corn were too mixed to show any pattern. Tile F-2 showed a somewhat different pattern. Cotton did not appear as clearly as it had in the previous two tiles examined. The unsupervised classification was run again with only four possible classes, and this produced an image with cotton displaying more prominently.

Tile G-2 was chosen for a series of more intensive tests because it had a higher number of CLU polygons than the previous three tiles. Most of the tiles, such as G-6, pictured above (figure 1), did not have a full range of CLU polygons throughout. In tile G-2, there is a greater coverage of CLU; figure 2, below, shows CLU for G-2 above the NAIP image.

Cotton (in magenta) again stands out, since most of the vegetation appears green on the NAIP image. Some cotton fields in the northeastern quadrant of the image appeared brown; this was explained by the fact that the NAIP imagery in this flight line was acquired on September 15, 2007. The adjacent NAIP flight lines were acquired in August, before the GEOSAR project was flown. The soybeans, outlined in red, appear in different stages of growing/harvest, as does the corn, outlined in yellow. The green CLU polygons are CRP areas, and the reddish brown polygons are fallow.

After the unsupervised classification was run on the tile, the neighborhood process was also run. The upland cotton areas are again visible as darker green tones, symbolizing pixel classes 5 and 6. The soybean polygons are mixed between darker and lighter greens, while the corn polygons are often brown, which might indicate a lower elevation. This would be consistent with the seemingly bare fields seen in the NAIP image.

At this point, it is important to note that the image used for the classification was the X band, a 32 bit tiff, displaying the magnitude. The pixel values are without units, measuring the strength of the return signal. As a result, the values may not be indicative of vegetation height, especially since they may not bounce back from the exact top of the plant, just as the P band may not bounce back from ground surface of the image. Another factor to consider is that the ground sample distance of the X band pixels is 3 meters, while the resolution of the P band pixels is 5 meters. As a result, pixel boundaries for the two layers would not coincide, and any results would be somewhat coarse when compared to the NAIP imagery. The classification process would, by its nature, resample the data to create a new file with 5 meter GSD.

.

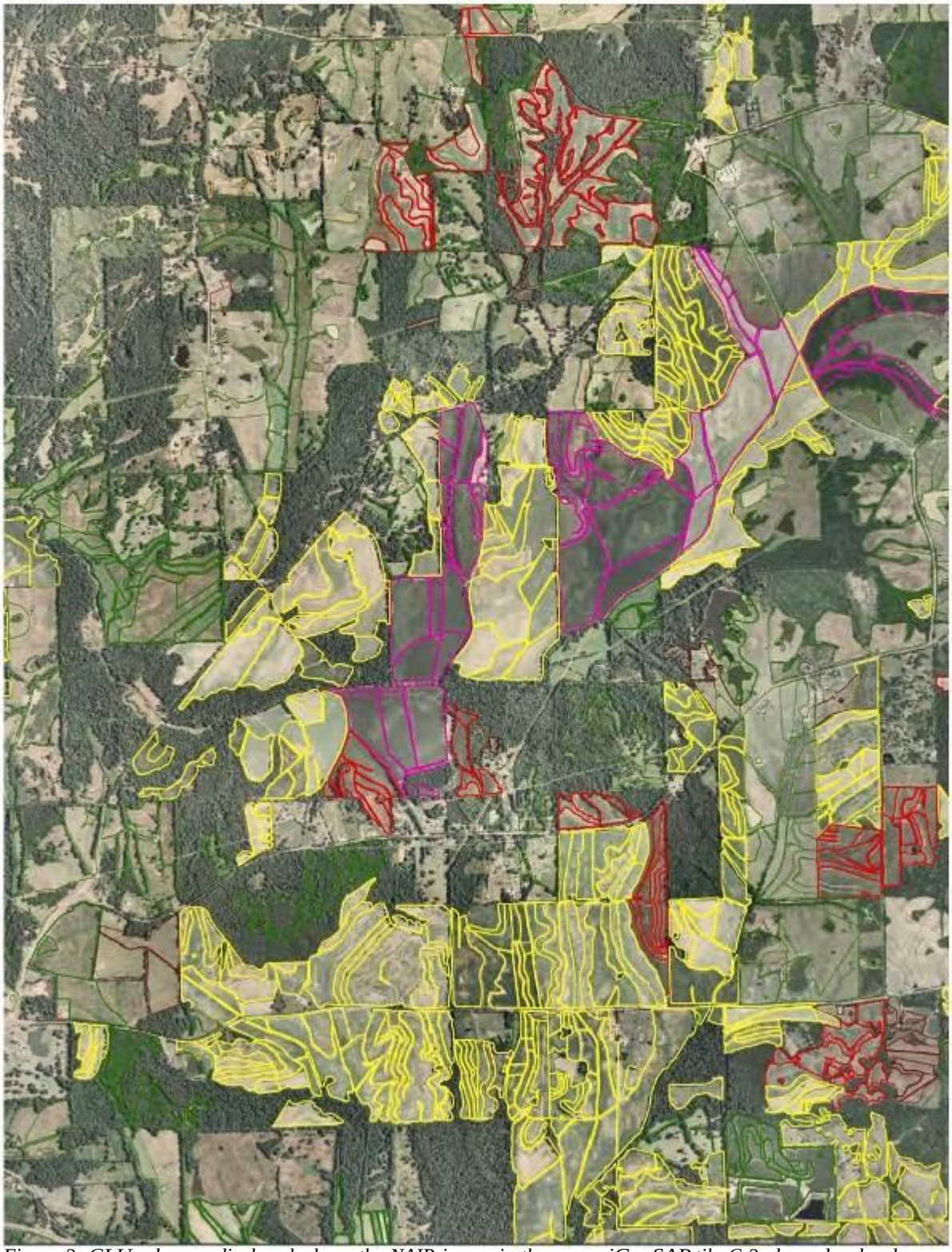


Figure 2: CLUpolygons displayed above the NAIP image in the area ojGeoSAR tile G-2 showed upland cotton (in magenta) clearly as compared to the other crops. Corn polygons (yellow) appear to be bare fields, cmd soybeans (red) display more variety in the greenish tone.

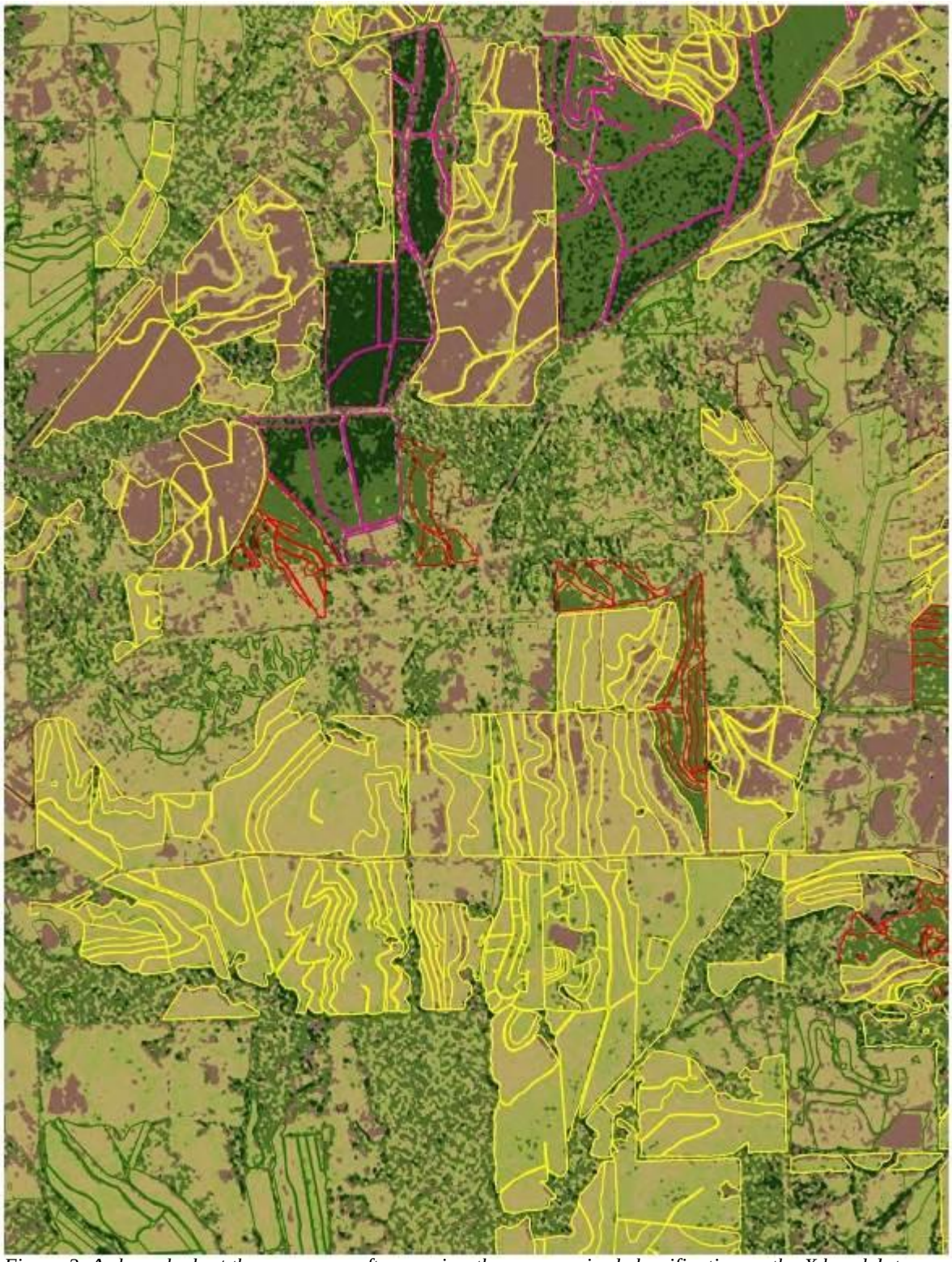


Figure 3: A closer look at the same area qfter running the unsupervised classification on the X bcmd data The uplcmd cotton polygons (magenta) stcmd out with darker green pixels used to symbolize classes 5 and 6, indicating a higher rejlectcmce value. The soybecmfields (red) are somewhat green, while the cornfield (yellow) display as lighter greens, tcm cmd dark brown, indicating lower rejlectcmce values.

After creating the classified and symbolized X band image, an image of only the dark green polygons (class 6) was extracted. A visual analysis was performed next, marking polygons by the perceived presence of upland cotton, based on the prevalence of pixels from this class.

The results were (by polygon):

CLU	APFO classification:			
Contains:	Cotton	Not Cotton	% Cotton	% Not Cotton
Cotton	112	34	73.2	22.22
Not Cotton	7		4.58	

And by acres:

CLU	APFO classification:			
Contains:	Cotton	Not Cotton	% Cotton	% Not Cotton
Cotton	1578	165.7	87.57	9.2
Not Cotton	58.24		3.23	

Acreages were taken from the CALC_ACRES field in the attribute table.

The study by Fugro found an 84% correlation between CLU reported and their classified map; an 86% correlation between map data and reference data, a 76% correlation between CLU reported crop data and (low flying aerial) photo-interpreted crop type, and an 88% correlation between photo-interpreted crop type and CART classified crop label. (CART being the automated Classification and Regression Tree analysis used by Fugro.) Their unit of comparison appears to be the individual polygons. Findings on one tile, comparing both polygons (fields) and acres, were 73% and 88%, respectively. The 73% comparison is slightly lower than their correlation for 76% for the same test, a comparison of CLU reported data and photo-interpreted crop type, while the APFO assessment by acreage is very close to their CART results. This would seem to show that even with coarse pixel resolution, inadequate software tools, and a manual assessment, the GEOSAR data could potentially be as useful as the imagery itself in some situations.

An important part of the result is the observation that while most cotton polygons contained predominantly class 6 pixels, the converse is not true. Polygons containing class 6 pixels were not necessarily cotton.

Subtracting the P band from the X band

Other tests were run using the method of subtracting band P from band X, since "theoretically" this should produce a model of vegetation height. Obviously, the reality of this data is that the returns are not clearly the "tops" of vegetation for the X band, or the ground surface for the P band. The tests were done with both the Mag data and the DEM data. With the Mag data, the resulting image was processed with the unsupervised classification in ERDAS Imagine, as described above. This process produced some strikingly clear patterns, which are described in greater detail in the appendix.

The DEM image produced patterns which closely resembled the natural color NAIP image. Because it lacked the tonal variety of NAIP, it would probably be less useful.

Because the pixels have a 3 meter GSD in the X band, and a 5 meter GSD in the P band, the pixel boundaries do not coincide. A new image made by subtracting one band from another would involve resampling, and GSD would be 5 meters. The resolution is much coarser than NAIP imagery, and as a result, the clarity of details would be greatly diminished.

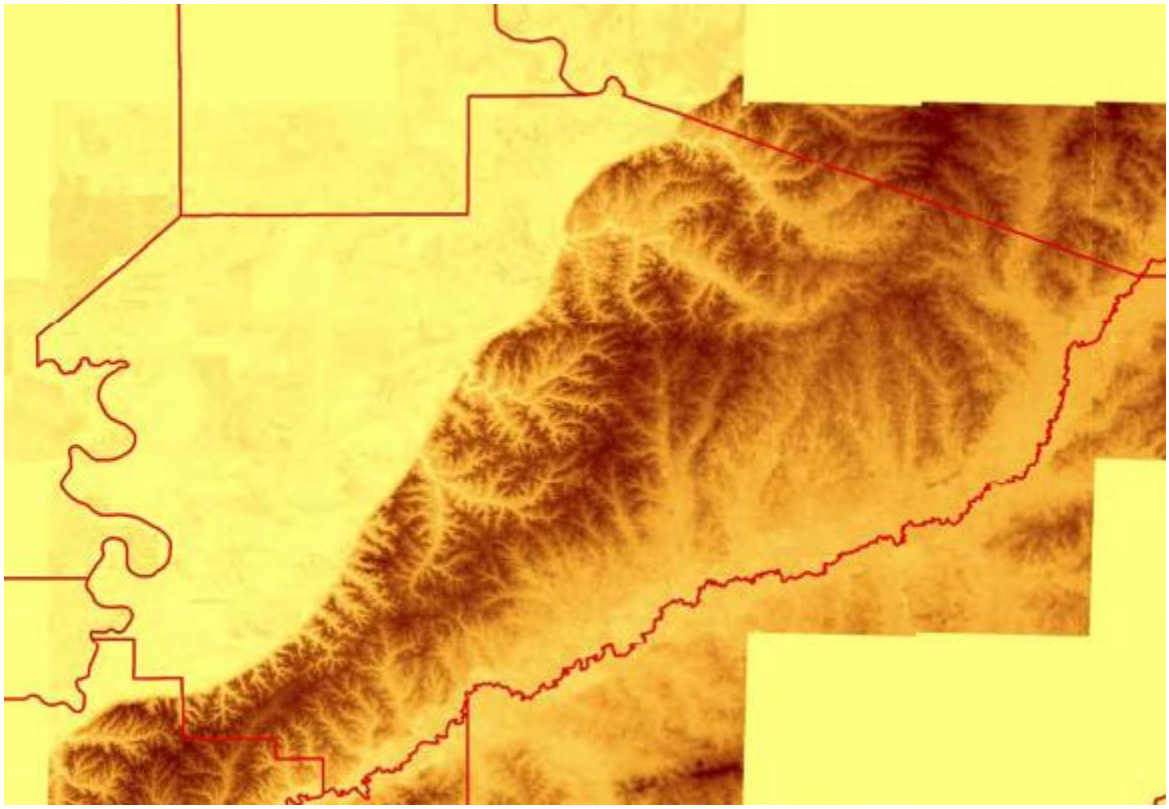
The GeoSAR data from Fugro displayed patterns coinciding with the crops as recorded in the CLU polygons with 578 crop attribute data. The tests at APFO were run with minimal image analysis tools in ArcGIS and ERDAS Imagine. More sophisticated software and greater operator education would produce more useful results. However, for the general user in a county or state office, it is doubtful that GeoSAR alone, at the present time, would be as useful as the basic natural color or color infrared imagery provided by the NAIP program.

Comparing X and P band DEMs to the NED

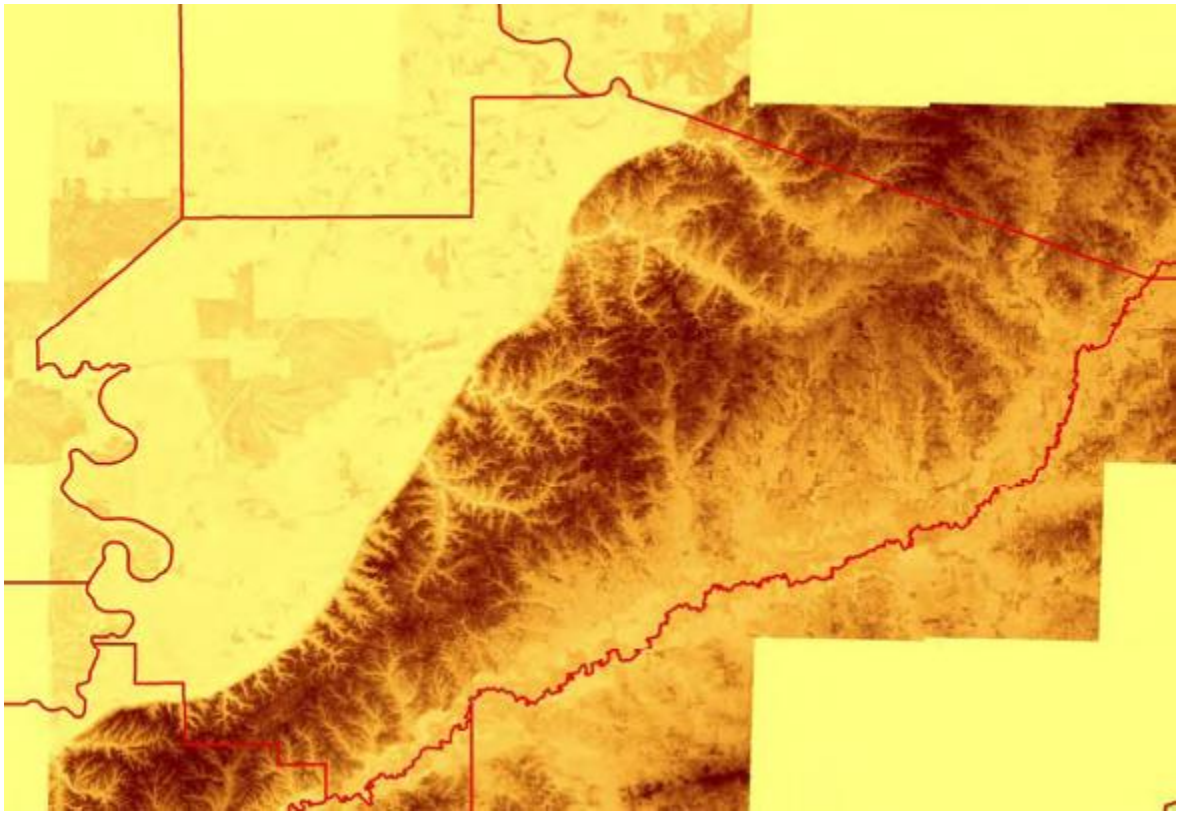
Both ArcScene and ArcMap were used to compare the P-band and X-band DEMs to the NED. NED data was downloaded for Yazoo County, MS from the USGS website, The National Map Seamless Server. For comparison purposes, only one quarter of the county NED data was used. The values for this section of the NED ranged from 19 to 122 meters. The Yazoo x-band DEM values ranged from 0 to 92 meters, and the p-band DEM values ranged from 0 to 91 meters. There is no frame of reference to compare the x and p band DEMs to the NED.

For the sake of comparison, the value of 19 was added to the p-band DEM to adjust it closer to the lowest value of the NED in this area. The next step was to subtract each band from the NED to calculate the difference. Using the Minus tool in ArcToolbox, each DEM layer was subtracted from the NED layer. The resulting values were divided into 5 meter increments to show the variation in values. It is impossible to compare these DEMs to the NED without knowing the real Geoidal heights of the x and p-band DEMs. Also note the horizontal datum of the x and p-band DEMs is WGS-84, and the datum of the NED is NAD83. It is also difficult to compare 3 and 5 meter DEMs to the NED which has a 10 meter resolution. Future comparisons should include resampling to a consistent resolution.

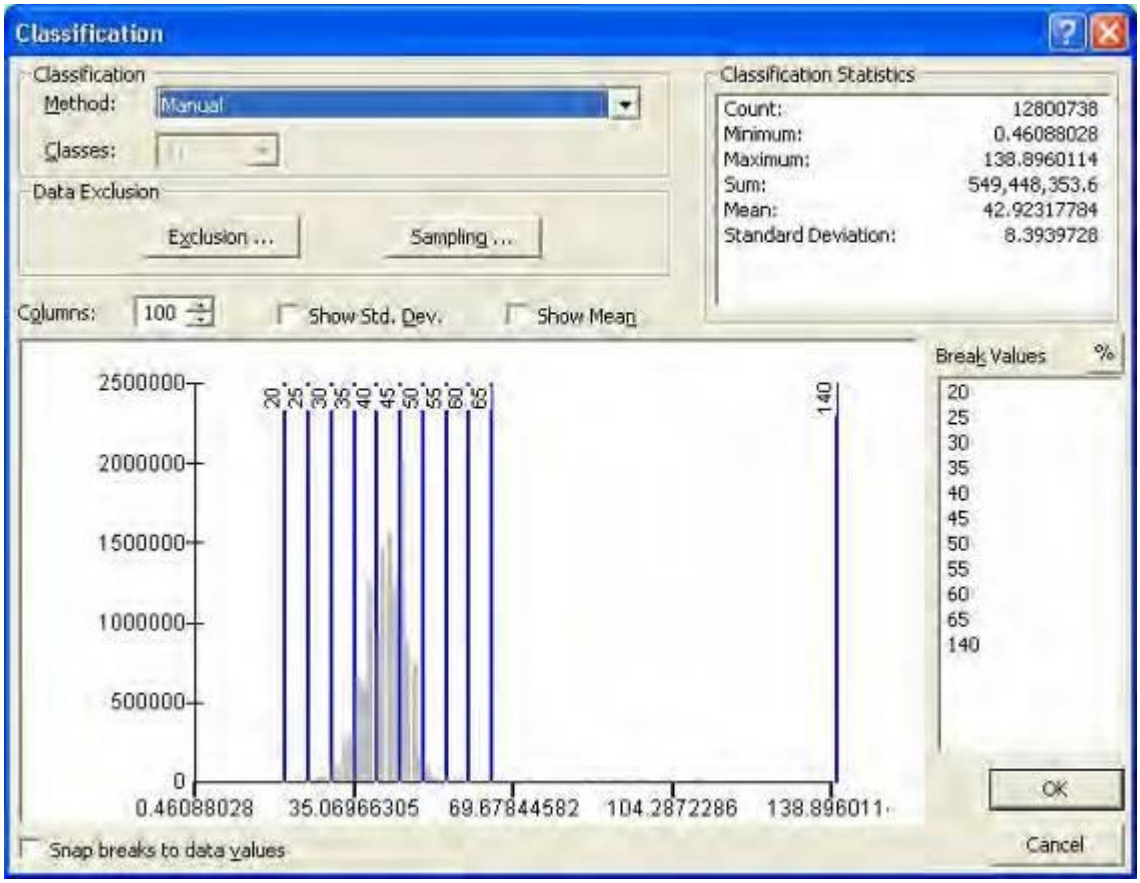
During the collection of the GeoSAR data, Fugro EarthData placed 2 X-band corner reflectors in the field that were GPS controlled. The heights in the GeoSAR DEMs at these locations were within 2m of the GPS height (on average). Further study of the DEMs should include vertical absolute accuracy control points which could be cross referenced with the DEMs to determine the average offset from true ground.



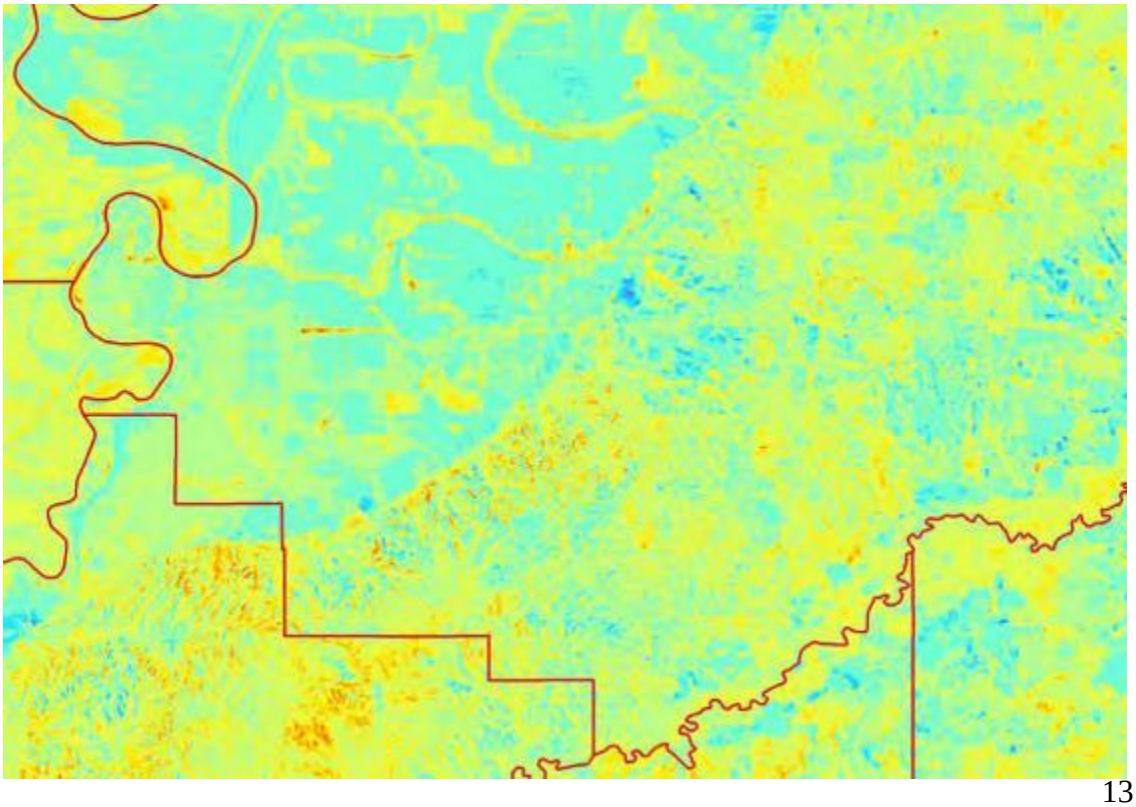
P-band DEM mosaic of Yazoo County, MS

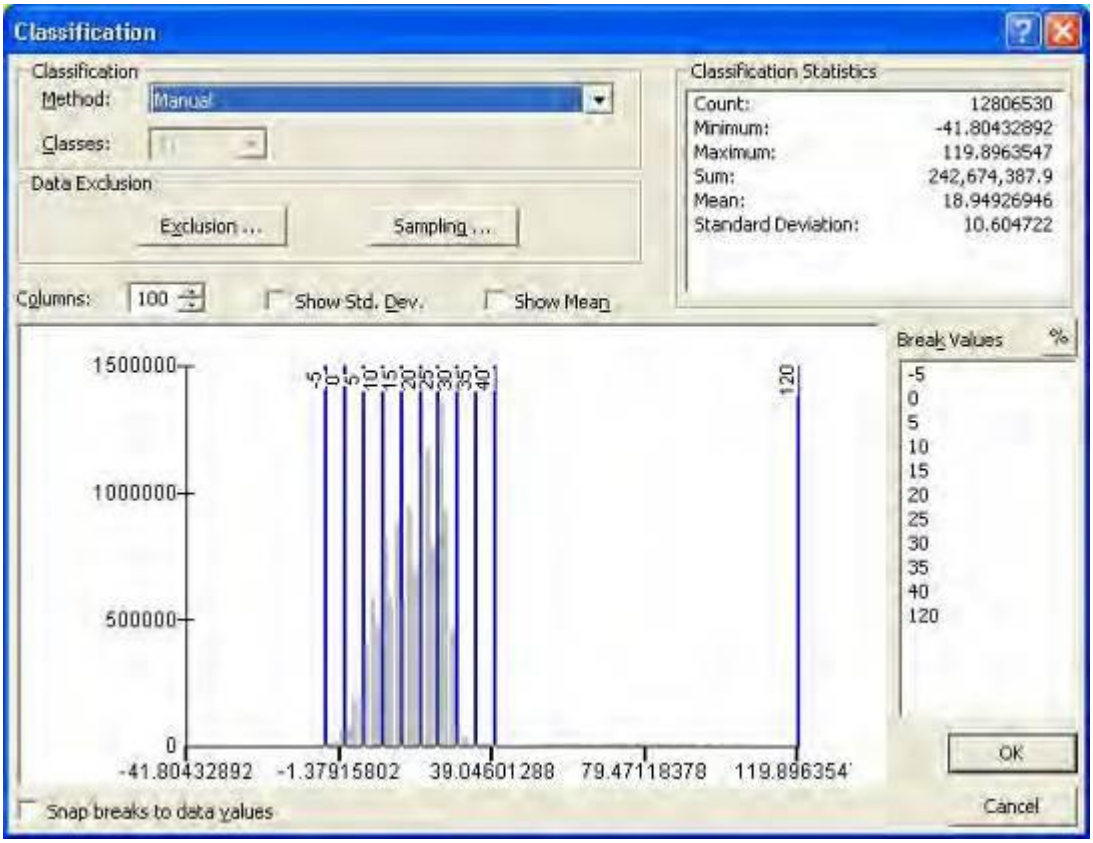


X-band DEM mosaic of Yazoo County, MS

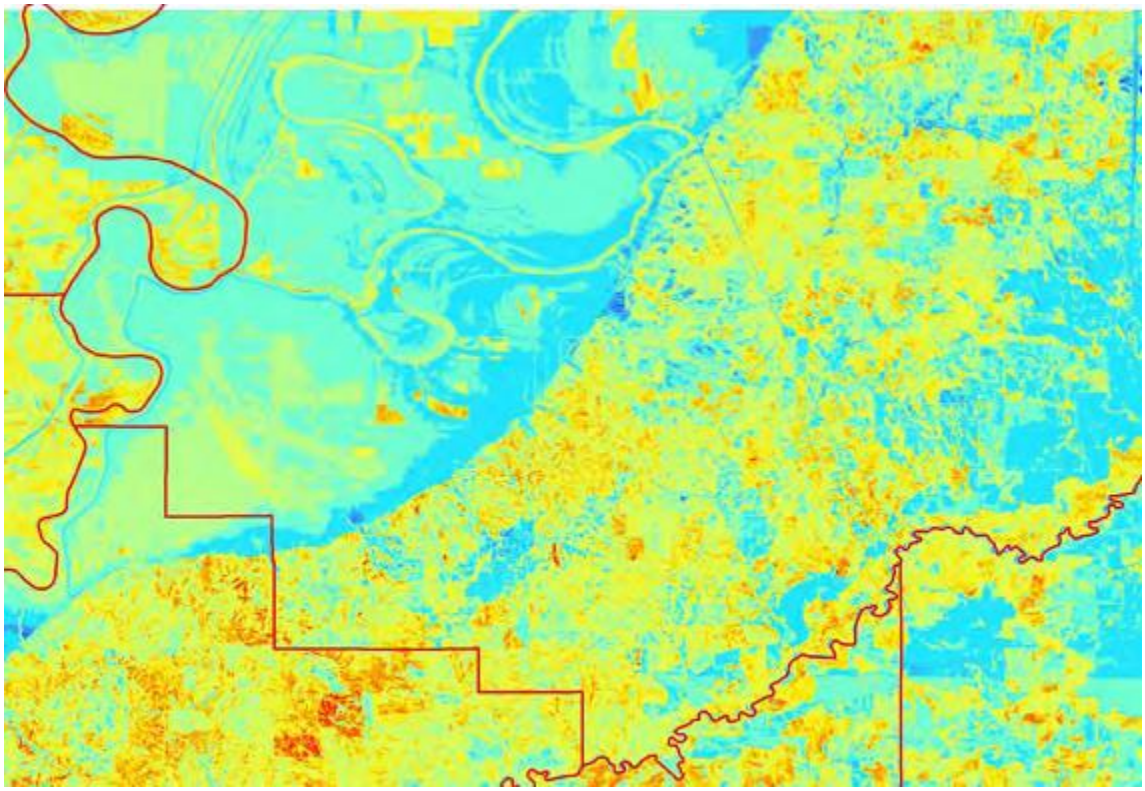


NED minus P-band distribution of values above and resulting classification below with a 19m adjustment.





NED minus X-band distribution of values above and resulting classification below with no adjustment.



Conclusion

The provided GeoSAR data falls in-line with other FSA datasets quality standards. No data quality or functional problems were found within the testing parameters performed at APFO. These parameters consisted of analysis tests which compared GeoSAR to NAIP imagery and GeoSAR DEMs to the NED. In the future, if more GeoSAR data was acquired a preliminary inspection process should be created for initial quality assurance and usability testing. This process would ensure that the data was usable in any analysis similar to what was preformed on this particular GeoSAR submittal or any future analysis which may be carried out on it or subsequent submittals.

Collected under harsh conditions and within a short time frame, after a year of precise planning, the GeoSAR data performed excellently. The accuracy matches up to the 2007 NAIP imagery collected during the same time frame and is a very rich dataset from which a great deal of data can be extracted. The sheer amount and types of data which can be extracted from both the magnitude and DEM returns is the data's greatest strength (and weakness). While a lot of data can be extracted, it requires a different set of skills and training to use proficiently. The data is very different from typical imagery or raster data which is dealt with on a regular basis and there is little documentation, literature, or training available for GeoSAR/IFSAR technologies. This makes understanding and extracting data from GeoSAR very time intensive and tedious.

Current FSA programs and job scopes would benefit little from this data at present. It has great potential for crop identification if exact study parameters are in effect before collection, but on a large scale is probably not more accurate or cost effective compared to current techniques at the present time. In visual inspections of GeoSAR, compared to NAIP, it is also harder for the untrained eye to identify field boundaries, building types and crop types; these functions are a core everyday task for FSA employees nationwide. Perhaps its greatest potential asset to FSA is the DEMs which can be derived. The DEM could enhance the ortho production process, which is critical for the vertical accuracy of NAIP. The GeoSAR data will be offered to other USDA agencies and other federal agencies for additional analysis. It is quite likely that these agencies will have far greater potential uses for, and expertise with, GeoSAR, which will provide future opportunities for GeoSAR collection and study for them and FSA alike.

Appendix Documents

Working with GEOSAR Data in ERDAS Imagine

Introduction.

This portion of the GEOSAR test uses ERDAS Imagine, in conjunction with ArcGIS. The ERDAS license at APFO does not include the Classifier module; as a result analysis was severely limited. The license does include an *Unsupervised Classification* function within the *Data Preparation* menu, as well as a *Neighborhood* function within the *GIS Analysis* menu, located under the Image Interpreter menu.

In spite of the limitations, the data produced patterns which could be seen approximating the outlines of fields, as represented by the CLU polygons. The analysis used the same process with four different tiles, in order to compare the results accurately. This was done while bearing in mind that there could be many potential variables when running the processes, and many different tests which could be conducted.

The data, when viewed in ArcGIS, could be symbolized as *Stretched*, where a color ramp is applied to the range of values in each pixel, or as *Classified*, in which the data is broken into classes using one of seven possible methods. The data breaks can be determined either by the software or by the user. The X band is the “first return” of the data, while the P band penetrated the vegetation to presumably reflect off the ground. Since the X band would reflect the vegetation surface, and would possibly model the crops in the field at the time of acquisition, I chose to look at this band when searching for data patterns. I chose to look at the 32 bit Mag tiff images.

Using the Unsupervised Classification in ERDAS Imagine

Tile G-6 was the first one examined. The X band image was used in an *Unsupervised Classification* in Imagine, with 6 classes and 10 iterations specified, stopping when the program reached .950 convergence. After the process had completed, the output image was used in a neighborhood classification, which would convert some random pixels to the dominant class value, giving the impression of a more cohesive polygon. The neighborhood definition was 7 x 7 pixels. The classes would range from 1 (lowest magnitude or height) through 6 (highest magnitude or height), with 0 being unclassified pixels. In all images analyzed, no pixels were classed as 0.

In the first tile examined, the pattern was surprisingly clear, and similar to the report and APFO examination in ArcGIS. The imagery was compared to CLU files with the *578* data included in the attribute table. This crop description (in the field DESC_CROP) assigns a [Mississippi] crop to the CLU polygons. It does not allow for fields which may contain more than one crop. Fields in Yazoo County

contained both harvested and un-harvested crops at the time of data collection, and this was obviously not recorded in the attribute table. The upland cotton polygons, shown with magenta outlines in the image below, as well as sweet potatoes (aqua lines) stood out very clearly as classes 5 and 6 in the Unsupervised Classification. These classes are represented in dark green in the image below. The neighborhood classification also clearly displayed areas of lower elevation, such as water and road surfaces. My hypothesis going forward was that other tiles would exhibit the same patterns after the same processes were run

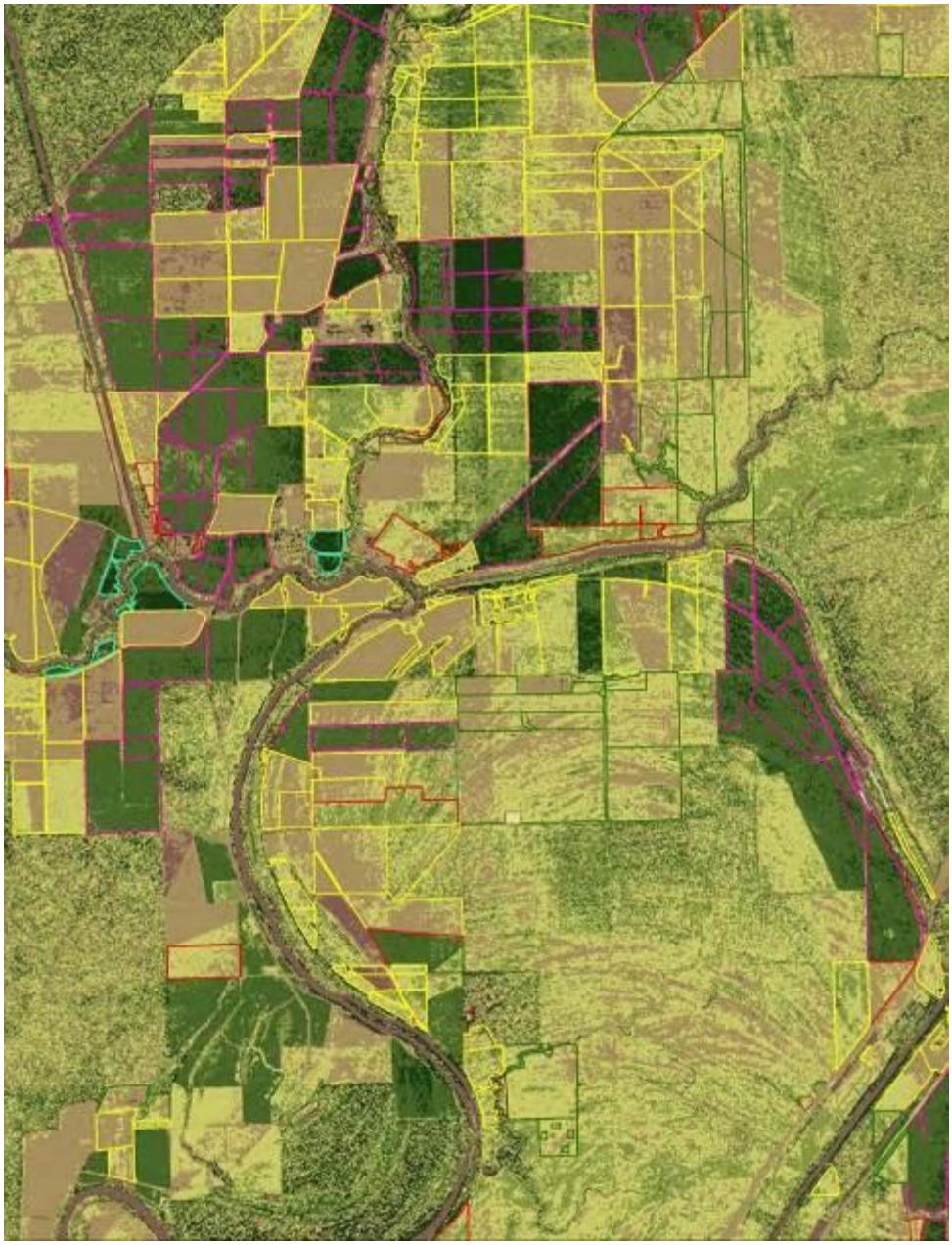


Figure 1: Tile G-6 showed upland cotton (in magenta) and sweet potatoes (aqua) very clearly, as darker green was used to symbolize classes 5 and 6. Corn (yellow) and soybeans (red) had displayed more variety in the classification.

Tile H-5 displayed the same pattern as tile G-6. Cotton appeared easily, but pixel classes for soybeans and corn were too mixed to show any pattern. Tile F-2 showed a somewhat different pattern. Cotton did not appear as clearly as it had in the previous two tiles examined. The unsupervised classification was run again with only four possible classes, and this produced an image with cotton displaying more prominently. This example underscored the fact that an arbitrary decision when selecting the process parameters will yield arbitrary results. A supervised classification, with a representative signature file, would probably yield a more valid result.

Tile G-2 was chosen for a series of more intensive tests because it had a higher number of CLU polygons than the previous three tiles. Most of the tiles, such as G-6, pictured above (figure 1), did not have a full range of CLU polygons throughout. In tile G-2, there is a greater coverage of CLU; figure 2, below, shows CLU for G-2 above the NAIP image. Cotton (in magenta) again stands out, since most of the vegetation appears green on the NAIP image. Some cotton fields in the northeastern quadrant of the image appeared brown; this was explained by the fact that the NAIP imagery in this flight line was acquired on September 15, 2007. The adjacent NAIP flight lines were acquired in August, before the GEOSAR project was flown. The soybeans, outlined in red, appear in different stages of growing/harvest, as does the corn, outlined in yellow. The green CLU polygons are CRP areas, and the reddish brown polygons are fallow.

After the unsupervised classification was run on the tile, the neighborhood process was also run. As with the previous tiles, the classification resulted in 6 classes of pixel values, and the neighborhood definition was 7 x 7 pixels. The resulting image is shown in a close-up view in Figure 3, below. The color scheme used for symbolization was the same one used for tile G-6. The upland cotton areas are again visible as darker green tones, symbolizing pixel classes 5 and 6. The soybean polygons are mixed between darker and lighter greens, while the corn polygons are often brown, which might indicate a lower elevation. This would be consistent with the seemingly bare fields seen in the NAIP image.

At this point, it is important to note that the image used for the classification was the X band, a 32 bit tiff, displaying the magnitude. The pixel values are without units, measuring the strength of the return signal. As a result, the values may not be indicative of vegetation height, especially since they may not bounce back from the exact top of the plant, just as the P band may not bounce back from ground surface of the image. Another factor to consider is that the ground sample distance of the X band pixels is 3 meters, while the resolution of the P band pixels is 5 meters. As a result, pixel boundaries for the two layers would not coincide, and any results would be somewhat coarse when compared to the NAIP imagery.

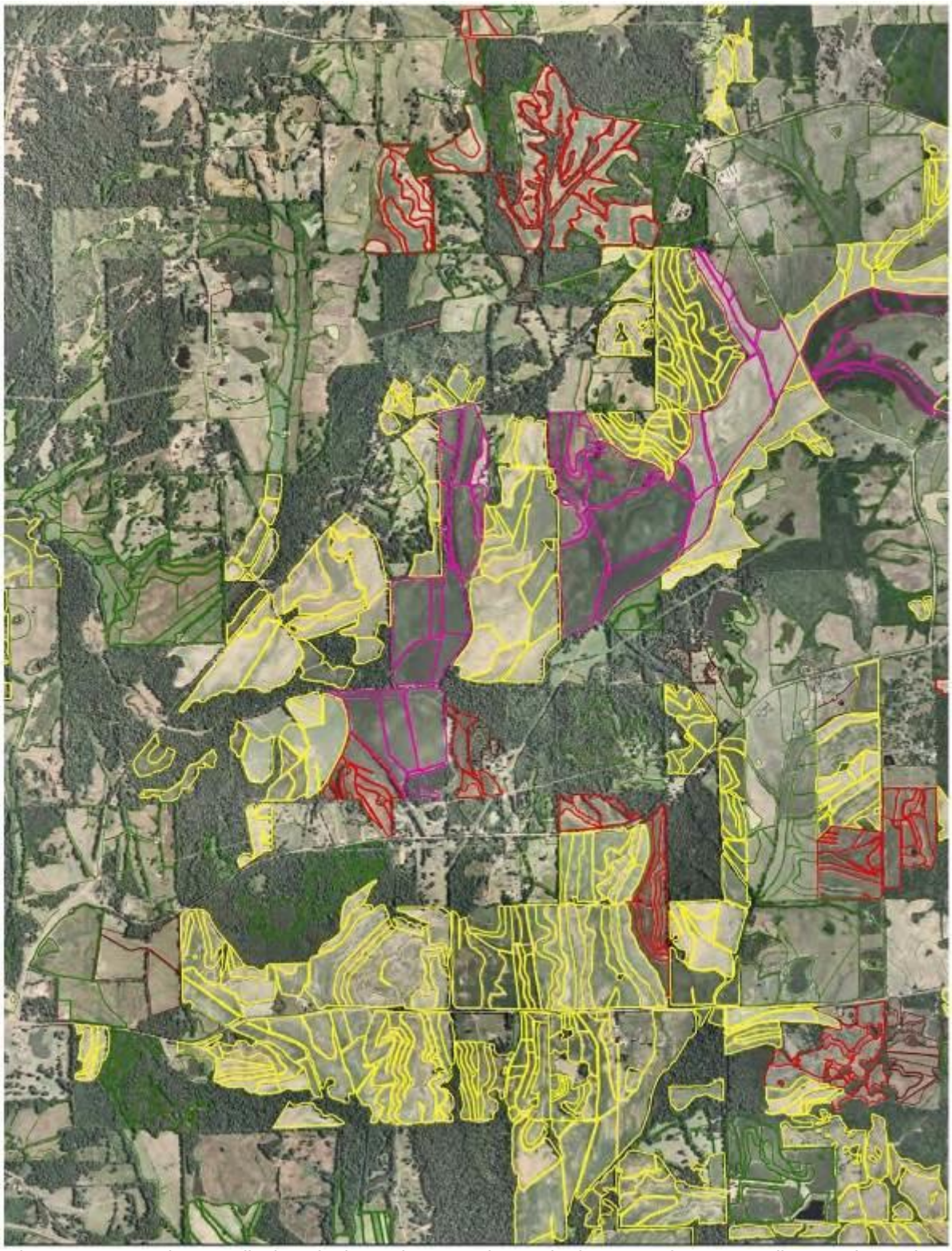


Figure 2: CLU polygons displayed above the NAIP image in the area of GeoSAR tile G-2 showed upland cotton (in magenta) clearly as compared to the other crops. Corn polygons (yellow) appear to be bare fields, and soybeans (red) display more variety in the greenish tone.

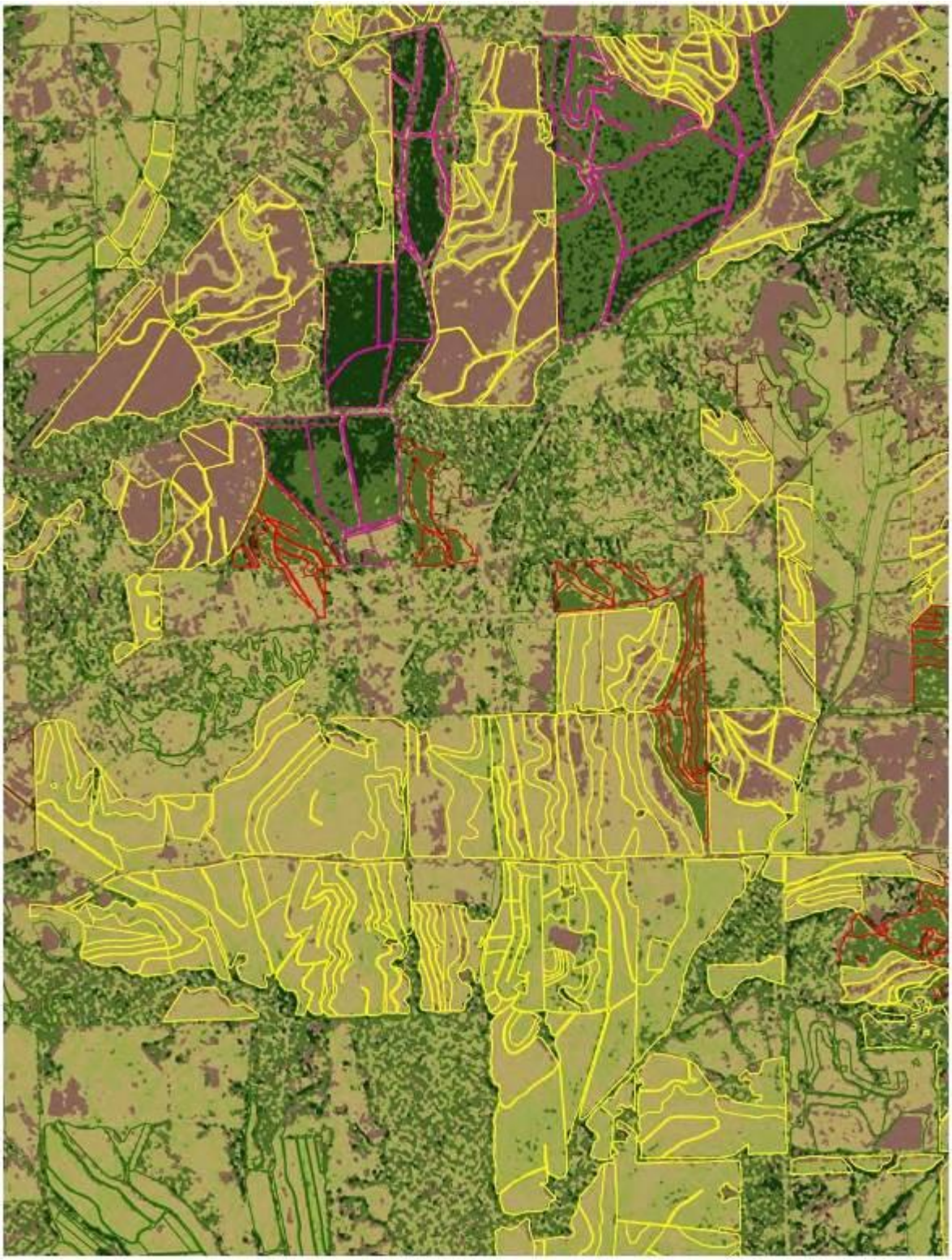


Figure 3: A closer look at the same area after running the unsupervised classification on the X band data. The upland cotton polygons (magenta) stand out with darker green pixels used to symbolize classes 5 and 6, indicating a higher reflectance value. The soybean fields (red) are somewhat green, while the corn field (yellow) display as lighter greens, tan and dark brown, indicating lower reflectance values.

After creating the classified and symbolized X band image, an image of only the dark green polygons (class 6) was extracted, using the *Extract by Attributes* tool in ArcGIS. I then did a visual analysis, marking polygons by the perceived presence of upland cotton, based on the prevalence of pixels from this class. The results were (by polygon):

CLU	Classed by me as:			
Contains:	Cotton	Not Cotton	% Cotton	% Not Cotton
Cotton	112	34	73.2	22.22
Not Cotton	7		4.58	

And by acres:

CLU	Classed by me as:			
Contains:	Cotton	Not Cotton	% Cotton	% Not Cotton
Cotton	1578	165.7	87.57	9.2
Not Cotton	58.24		3.23	

Acreages were taken from the CALC_ACRES field in the attribute table.

The study by Fugro found an 84% correlation between CLU reported and their classified map; an 86% correlation between map data and reference data, a 76% correlation between CLU reported crop data and (low flying aerial) photo-interpreted crop type, and an 88% correlation between photo-interpreted crop type and CART classified crop label. (CART being the automated Classification and Regression Tree analysis used by Fugro.) Their unit of comparison appears to be the individual polygons. My findings on one tile, comparing both polygons (fields) and acres, were 73% and 88%, respectively. The 73% comparison is slightly lower than their correlation for 76% for the same test, a comparison of CLU reported data and photo-interpreted crop type, while my assessment by acreage is very close to their CART results. This would seem to show that even with coarse pixel resolution, inadequate software tools, and a manual assessment, the GEOSAR data could potentially be as useful as the imagery itself in some situations.

An important part of the result is the observation that while most cotton polygons contained predominantly class 6 pixels, the converse is not true. Polygons containing class 6 pixels were not necessarily cotton.

Subtracting the P band from the X band

My next test was to create an image for tile G-2 by subtracting band P from band X, since theoretically this should produce a model of vegetation height. A flaw in this logic exists because the magnitude band is not a measure of vegetation height; the units recorded are of signal strength rather than elevation. However,

for experiment's sake I proceeded with the analysis. The new image was created in ArcGIS using the Minus tool. It was processed in ERDAS using the same steps as the previous images. A close-up of the image, showing the same polygons as Figure 3, gives a strikingly different result when using the same color symbolization as the previous images.

Figure 4: After subtracting P band values from X band values, and running an unsupervised classification on the resulting image, the upland cotton polygons (magenta) are immediately obvious as class 6 (the highest values).The soybean filed (red) are also a darker green,indicating higher values, and the corn (yellow) displays with more green (higher values) than in the X band image, or as interpreted from the NAIP image.

In this image, the upland cotton polygons (outlined in magenta) were immediately obvious in dark green, symbolizing class 6, the greatest difference in magnitude, and presumably greatest distance between ground and plant height. Comparing this image with the NAIP imagery shows the same magenta polygons with a richer green tone than other fields, implying the presence of growing vegetation. It was a bit surprising to see how closely the dark green pixels matched the polygon boundaries.

The soybean polygons, outlined in red, were also a darker green in the X – P image; they seemed to have a higher percentage of class 5 pixels along with class 6. This again is consistent with the NAIP image. The soybean polygons are green, but with a lighter tone than the cotton, implying vegetation which is not as well developed as the cotton plants.

The other results for the image were more mysterious. The CRP polygons, outlined in green, show a variety of land covers in the NAIP image. In the X – P image, the CRP polygons are depicted with class 1 – 3 pixels, implying smaller vegetation heights. This symbolization is seen in wooded areas, and pixel values in the raw X – P image for these areas were sometimes surprisingly low negative numbers. Lighter (lower classification) pixels are seen in between the dark green cotton fields; on the NAIP imagery trees are visible between the fields. This result seems counter-intuitive, since trees should be “higher” and therefore darker (a higher classification score). It would take more study and understanding of this imagery to determine why this pattern is emerging.

Also unexpected were the results for the corn polygons. These were all now a definitive green, depicting mixtures of classes 4, 5 and 6. The pixel patterns don’t match the patterns of green vegetation in the image, and it was especially surprising to see green patterns in areas where the NAIP image appears to show bare ground.

A visual photo-interpretation, based on the X – P image was done for all three crops (cotton, soybeans, and corn) analyzed by Fugro. For cotton, the CLU polygons were compared to pixels in class 6, and the results recorded by the number of polygons and by acreage.

	Contained Cotton	Partial Cotton	No Cotton
Polygons	132; 91%	8; 5.5%	5; 3.5%
Acres	1699.96; 97.5%	28.85; 1.65%	14.74; 0.85%

Acreage data came from the CALC_ACRES field in the attribute table.

A comparison of cotton CLU polygons with pixels from classes 4 and 5 found eight polygons (out of 145) which were predominantly represented by those two classes.

This result was significantly higher than the results found by Fugro in analyzing cotton results.

Results for soybeans in class 6 showed:

	Contained Soybeans	Partial Soybeans	No Soybeans
Polygons	70; 31.67%	40; 18.1%	111; 50.23%
Acres	667.6; 35.42%	334.96; 17.77%	882.3; 46.81%

Results for soybeans in classes 4 and 5 together showed:

	Soybeans	Partial Soy	Minimal Soy	No Soy
Polygons	131; 61.5%	52; 24.4%	29; 13.6%	1; 0.5%
Acres	1133.2; 60.8%	427.5; 22.9%	301.4; 16.2%	2.03; 0.1%

Soybeans appear to respond with a mixture of pixel values from classes 4, 5 and 6. A signature file containing representative combinations of pixels depicting soybeans could probably be used in a supervised classification to more clearly identify fields.

Another tile, F3, was chosen to repeat the sequence of analysis steps, and compare them with two of Fugro's field points, which fell on CLUs within the tile's area.

Figure 5: Two of Fugro's field points are displayed against the NAIP image. How would an analyst interpret the crops planted near these points?

Figure 5, above, details the two field points on the NAIP image. The top point lies in a field of green vegetation. From previous tiles, I would guess that it might be upland cotton. The second point lies in a bare field with some crop apparently still in the field to the northwest and west of the point. The green tone appears similar to the cotton, but previous tiles, and Fugro's report, have made me suspect corn.

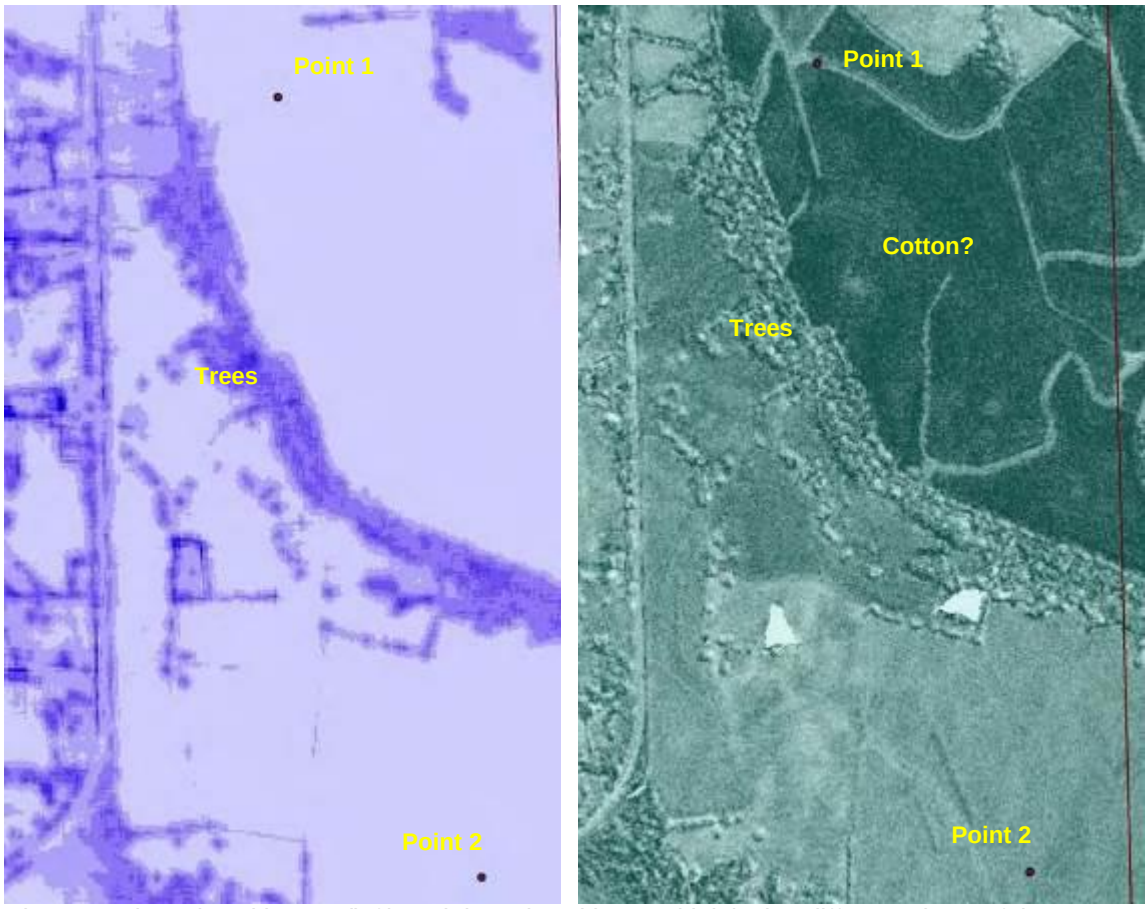


Figure 6: The P band image (left) and the X band image (right) give different views of the same area. The road and the trees are the clearest points of reference.

Figure 6 compares the P and X band magnitude images for the area. In the P band image, displaying the second return, or (theoretically) the ground surface, the roads have a stronger signal, implying that they are higher than the fields. The trees also displayed a darker color; this is an enigma noted in other tiles as well. The X band image displayed a darker tone in the fields near the top point, reinforcing the theory that it may be cotton. The brightest areas, two small polygons near the center of the image, do not appear clearly on the NAIP or P band image, and may be water. The trees display the shapes one would expect from canopies, but the tone is lighter than that of the fields immediately to the north. This is the other part of the enigma – the X band, in the trees, does not appear to be an indicator of vegetation height. The field containing the lower point shows a uniformity of tone lacking in the NAIP image. The GEOSAR would have been flown 3 – 6 days after the NAIP, and it is possible that the entire field

was harvested between the two image acquisition dates. The field adjacent to the road appears bare in the NAIP image, but is a darker tone than the fields immediately to its east.

An image made by subtracting the P band from the X band displays the same patterns seen in other tiles. The fields around the top point, suspected to be cotton, are the same darker tone seen in other tiles. The fields around the second point appear darker than would be expected, since these fields appear to be bare in the NAIP image. Again, the biggest surprise is in the trees – these display tones which are much lighter than the fields to the north.

Figure 7: The X band minus P band image, before the data was classified.

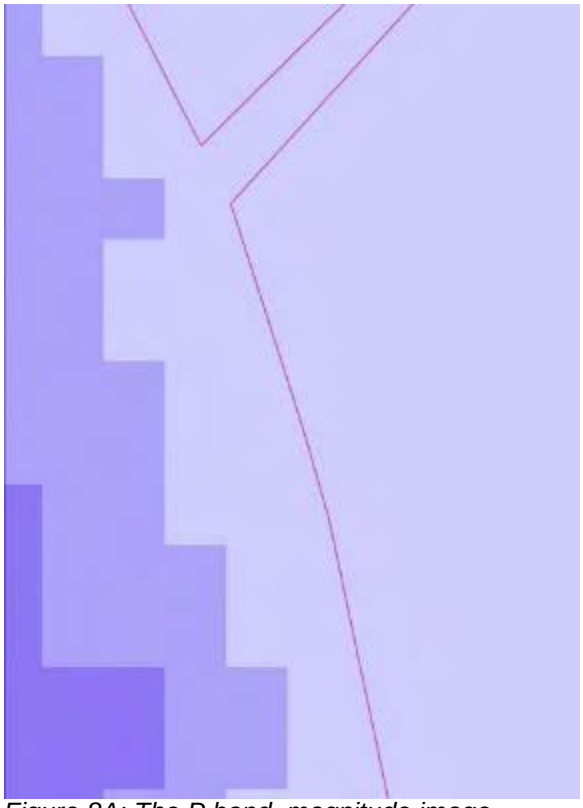


Figure 8A: The P band, magnitude image

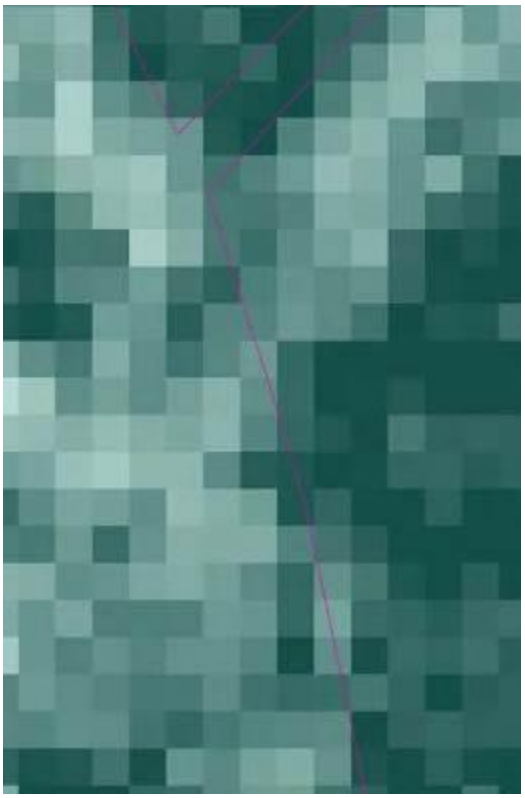


Figure 8B: The X Band, magnitude image

Figure 8C: X minus P image, magnitude

A digression will examine the imagery at the pixel level. The first point to consider is that the two images have different Ground Sample Distances (GSD), meaning that the pixels will not be aligned. Figure 8A shows the P band, with its 5 meter pixel resolution. The magenta line is the CLU boundary for the upland cotton polygons. The P band is noticeably lower to the left of the line. Figure 8B displays the X band, with 3 meter pixel resolution. In this image, the darker tones are to the right of the CLU boundary. When selecting the same pixel location on both images, the field area gave a value of 42 in a dark pixel to the right of the line in the X band, and 15 for the corresponding

location on the P band. A light toned pixel on the X band gave a reading of 13, while the same spot on the P band read 30.As a result, the pixels in Figure 8C display a lighter tone (lower values) in the area east of the polygon boundary. For some reason, the magnitude imagery records a lower value for the P band than for the X band in wooded locations. A greater understanding of the magnitude signals, and what information they really capture, would greatly enhance the usability of this imagery.

Figure 9: the X minus P image after Unsupervised Classification

Figure 9 displays the X minus P image after the Unsupervised Classification has been run. This displays the same pattern shown before: the area near the top point shows the same dark color consistent with cotton, while the fields near the bottom point are a mixture of classes 5 and 6, perhaps indicating corn. The

wooded areas exhibit the lighter yellow and brown tones indicated classes with lower values. The same image, after the neighborhood tool was used, presents a more generalized pattern; perhaps this step is not as useful as I had previously thought, since it gives a coarser view of the data. Perhaps a 3x 3 pixel neighborhood, rather than 7 x 7, would have given a better picture. This tool was used in an attempt to create areas of colors which were consistent. This would be important is an attempt were made to convert the raster to a vector file.

Figure 10: The X minus P image after the Neighborhood function has been run. This shows less detail in the field areas.

Solving the question of field identification, the CLU boundaries are displayed above the X – P classified image. As expected, the dark green fields near the top

point were cotton, and the fields around the bottom point were corn. However, much more study and experience would be needed before recommending this imagery as an easy tool to use in crop classification.

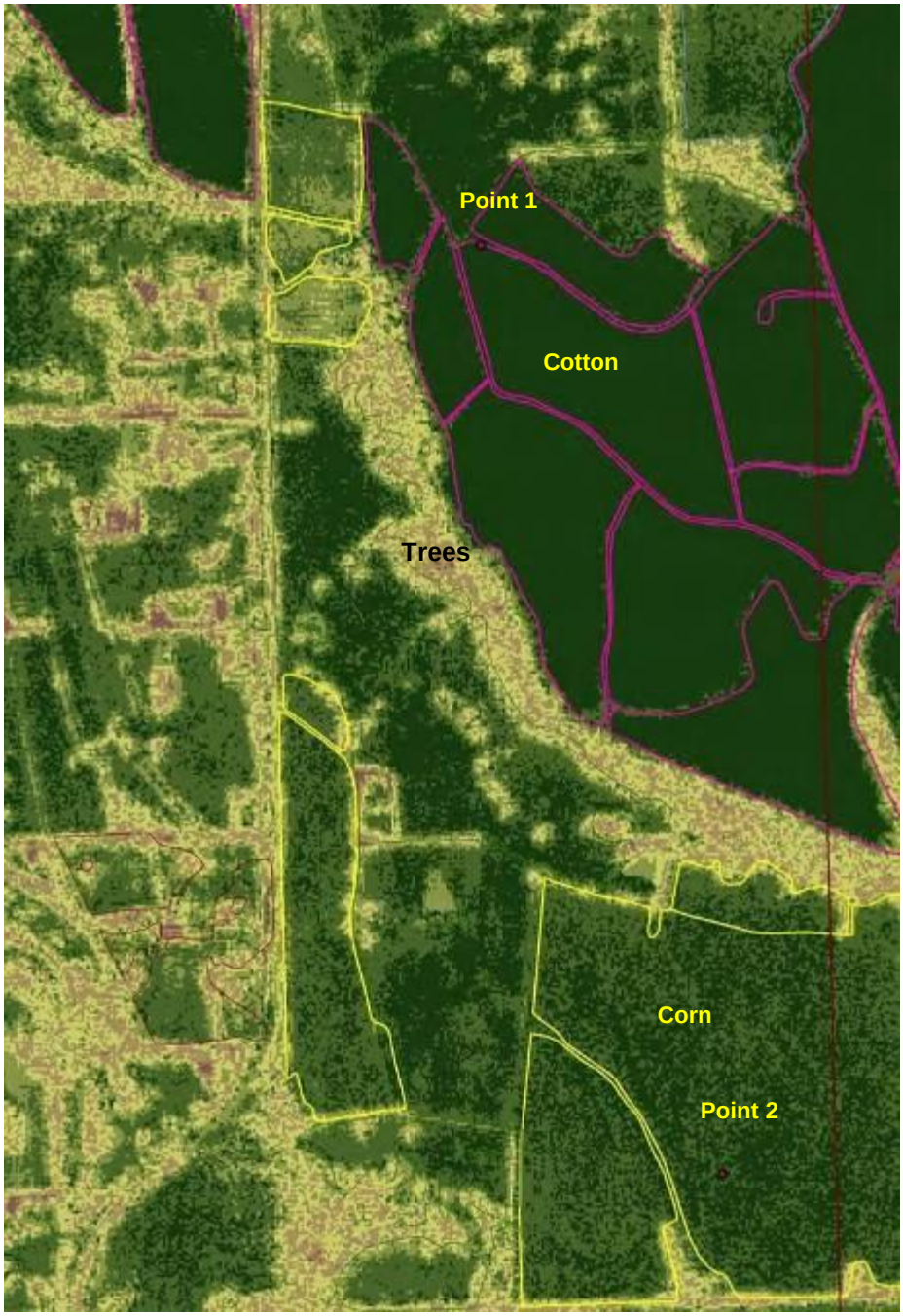


Figure 11: The classified image with CLU borders.

Fugro provided pictures on the ground to accompany their field points. The two points used in this example look as expected from their position on the NAIP imagery. The top point, in the cotton fields, is shown in Figure 12; it gives an “as expected” view of the field, but the angle of the photo does not allow the viewer

to estimate the height of the plants. The bottom point, Figure 13, shows the bare corn field, as it appears in NAIP image. It is very puzzling; why does the X – P image show the cotton and corn plants as being “higher” than the trees?

Figure 12: The upland cotton plants near Point 1.

Figure 13: The bare corn field near Point 2.

Subtracting the P band OEM from the X band OEM

A final experiment was to create an X – P image from the OEM imagery. This should theoretically produce an image which is truly representative of the heights of vegetation.

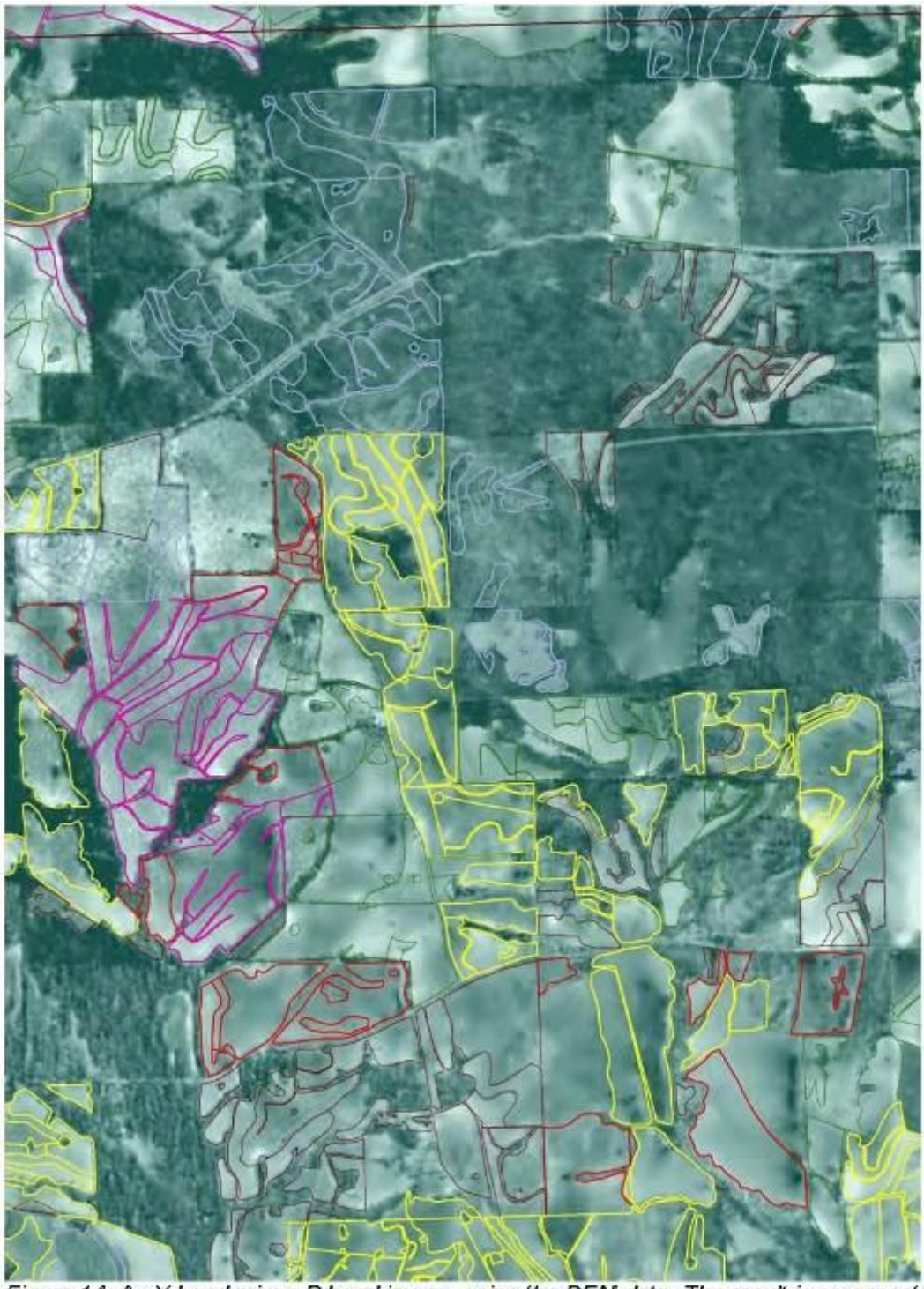


Figure 14: An X band minus P band image, using the DEM data. The result is as expected, with the trees showing elevation (darker green) and the field being lighter (lower elevation).

This hypothesis is correct; in comparing Figures 14 and 15, the X-P image of the OEM bands is close to what one would expect as a model of relative elevations in this area. It compares very closely to the NAIP image.

Figure 15: The NAIP imager with CLU boundaries overlaid.

In this experiment, the X minus P image gives the same sort of information as the NAIP – minus the tonal variations. This data would have value, but not as an indicator of crop type.

Summary

The GeoSAR data from Fugro displayed patterns coinciding with the crops as recorded in the CLU polygons with 578 crop attribute data. The tests at APFO were run with minimal image analysis tools in ArcGIS and ERDAS Imagine. More sophisticated software and greater operator education would produce more useful results. However, for the general user in a county or state office, it is doubtful that GeoSAR alone, at the present time, would be as useful as the basic natural color or color infrared imagery provided by the NAIP program.

Final Report for USDA J07-0020

GeoSAR Derived Automated Crop Cultivation Assessment USDA Yazoo County Mississippi

1 October, 2008

Prepared for

W. Geoffrey Gabbott, Contracting Officer
USDA FSA Aerial Photography Field Office
2222 W 2300 S Salt Lake City, Utah 84119
Phone 801-975-3500x207, Fax 801-975-3512
Email: geoffrey.gabbott@slc.usda.gov

Prepared by

Tom Carson
Chad Lopez
Jim Reis
Tim Leary
Steve Shaffer

Fugro EarthData, Inc.
7320 Executive Way
Frederick, MD 21704
301-948-8550
www.earthdata.com

The attached document contains proprietary and confidential information of Fugro EarthData.
The information may not, directly or indirectly, be displayed, provided, or transferred to any person or entity and may not be used, copied, or disclosed without the express written permission of Fugro EarthData.

TABLE OF CONTENTS

1 Introduction....1
2 Geosar technology overview....1
3 Description of Study area....2
4 Geosar data acquisition....3
4.1 Collection Plan....3
4.2 Local Weather at Time of Collect....5
5 Truth data collection....5
5.1 Ground Truth....5
5.2 Airborne Reconnaissance....6
6 Geosar data processing....7
6.1 Frequency Allocation....7
6.2 Profiling LIDAR....7
6.3 GeoSAR Production Flow....8
6.4 DATA PROCESSING STEPS....9
6.4.1 X-band and P-band Data Processing Production Steps....9
6.4.2 Wide Area Mosaicing Production Steps....10
6.4.3 Product Finishing and Packaging Production Steps....12
7 Comparison Between GeoSAR Imagery, NAIP Imagery and CLU Data....12
8 geodetic accuracy of geosar data....15
8.1 Vertical Accuracy....16
8.2 Horizontal Accuracy....17
8.2.1 Comparison to Corner Reflectors....17
8.2.1.1 Reflector 1....17
8.2.1.2 Reflector 2....20
8.2.2 Reflector Results....21
8.2.3 Comparison to Common Land Unit Data and NAIP Imagery....22
8.2.3.1 Accuracy of the CLU and NAIP....22
9 Crop and landcover classification experiments....23
9.1 Data Preparation and Training Site Collection....25
9.2 CART Analysis....26
9.3 Accuracy Assessment....27
9.4 Comparison of Classification results with CLU Reported Data....31
9.5 Classification Conclusions....33
10 Discussion and Conclusions....34
Appendix A - GeoStationary satellite images and precipitation maps....35

Appendix B Examples of data from crop field reports. 37

List of Tables

Table 1 Acquisition report for Yazoo data collects. Some acquisitions represent re-collects of the same line. The sorties are color coded. 4
Table 2 Comparison between XY position for Reflector 1 as measured by GPS and Image Mosaic. 17
Table 3 Comparison between XY position for Reflector 2 as measured by GPS and Image Mosaic. 20
Table 4 Measurement results comparing X and P Band mosaic images with CLU and NAIP data. 22
Table 5 List of datasets used. 25
Table 6 The number of CART analysis training sites and accuracy assessment sites for each crop type. 26
Table 7 Accuracy Assessment error matrix with corn subclasses using all four GeoSAR data products. 28
Table 8 Accuracy Assessment error matrix with single corn class using all four GeoSAR data products. 28
Table 9 Accuracy Assessment error matrix with corn subclasses but without the Σ_0 GeoSAR data product. 29
Table 10 Accuracy Assessment error matrix with single corn class but without the Σ_0 GeoSAR data product. 29
Table 11 Individual error matrix KAPPA analysis results. 30
Table 12 Comparative error matrix KAPPA analysis results. 30
Table 13 Comparison of crop identification by CART analysis with the data reported in the CLU dataset. 31
Table 14 Comparison of field reports with CLU data and CART classification. 32
Table 15 Comparison between fields identified by photo interpretation and the CLU labels. 32
Table 16 Comparison between fields identified by photo interpretation and the CART Classification labels. 33

List of Figures

Figure 1 GeoSAR IFSAR collection aircraft. 2
Figure 2 GeoSAR X-Band Image Mosaic of Yazoo County, Mississippi 3
Figure 3 The GeoSAR flight plan for Yazoo County, Mississippi. 5
Figure 4 ArcMap project showing GPS flight track points 7
Figure 5 High level flow diagram of GeoSAR production process. 9
Figure 6 Imagery Examples from 32090-H4 Quad. 14
Figure 7 Imagery Examples from 32090-F5 Quad. 15
Figure 8 Frequency histogram of X-Band DEM - lidar differences. 16
Figure 9 X-Band image with location of reflector 1 annotated. 18
Figure 10 Ground photo of the reflector 1. 19
Figure 11 Air photo with reflector 1 position annotated. 19
Figure 12 X-Band image with location of reflector 2 annotated. 20
Figure 13 Ground photo of the reflector 2. 21
Figure 14 Air photo with reflector 2 position annotated. 21
Figure 15 Frequency distribution of measured position differences. 23
Figure 16 Example of a portion of the See5 CART analysis classification rules. 27

1 INTRODUCTION

Under USDA contract AG-8447-C-06-0006, Fugro EarthData, Inc. (EarthData) conducted a pilot project to demonstrate the capabilities of the airborne GeoSAR Interferometric Synthetic Aperture Radar (IFSAR) to augment the USDA National Agricultural Photography Program (NAIP).

Under the scope of this contract, the EarthData GeoSAR system collected sufficient X-band and P-band interferometric SAR data using its standard imaging modes to fully map the county of Yazoo, Mississippi. GeoSAR collected a total of 16 lines of data in 5 sorties between August 29, 2007 and August 31, 2007. EarthData processed this data in its Frederick, Maryland facility to produce standard products, including 3-meter resolution X-band elevation models and orthorectified mosaics and 5-meter resolution P-band elevation models and orthorectified mosaics.

EarthData investigated how the unique characteristics of the X- and P-band SAR images can facilitate the collection of certain feature types, such as fence lines, that are not readily visible in standard NAIP imagery. EarthData also used the GeoSAR X- and P-band image data to determine whether the common crop types present in the data sets are separable within the feature space defined by the GeoSAR data. This allowed EarthData to determine that GeoSAR can be used as a tool to monitor compliance. The GeoSAR data was used in a CART regression tree analysis.

2 GEOSAR TECHNOLOGY OVERVIEW

GeoSAR is a dual-sided, dual-frequency, interferometric synthetic aperture radar (IFSAR) mapping system. The system is integrated onto a Gulfstream II business jet (Figure 1), and is wholly owned by Fugro EarthData Incorporated. It is capable of collecting data from 40,000 ft above ground level at an airspeed of over 400kts yielding a net collection rate of over 280 sq km per minute. The SAR operates at 2 frequencies simultaneously, X-Band, with a center frequency of 9700MHz, and P-Band, with a center frequency of 350MHZ. The system is designed to produce high accuracy digital elevation models (DEM) through the process of radar interferometry, as well as SAR orthophoto mosaics. GeoSAR is a commercialization of IFSAR technology developed by NASA Jet Propulsion Laboratory for wide area airborne mapping applications. EarthData has been commercially operating the GeoSAR aircraft since 2002 generating large cover area maps for the US National Oceanic and Atmospheric Administration in Southern California and for the National Geospatial-Intelligence Agency in South America. Image and DEM quality is excellent and is independent of cloud cover and sun illumination, yielding nearly all weather collection capability.

Figure 1 GeoSAR IFSAR collection aircraft. The P-Band antennas are contained in the tip-tanks on either side of the aircraft. The X-Band antennas are in the fairing visible near the fuselage on the underside of the aircraft. (Photo courtesy Fugro EarthData Inc.)

3 DESCRIPTION OF STUDY AREA

Yazoo County Mississippi is a predominantly rural county in west central Mississippi, with an area of 2,419 km^2 (934 mi^2). The major crops in the county, in terms of economic value, are corn, cotton, rice, wheat, hay, sorghum grain, soybeans and sweet potatoes. The major livestock commodities are cattle and catfish. The physiography of the county can be divided into two regions. To the west of the loess-capped bluff bisecting the county the landscape is part of the Yazoo-Mississippi Basin or delta. This terrain is level to very gently undulating near the Yazoo River and around abandoned and extinct river channels. The land cover in the delta consists of swampy forests, agricultural fields, and occasional catfish ponds. East of the bluff there are low, wooded Loess Bluffs or Brown Loam Hills. Among the hills nearest the bluff, valleys are often deep and steep-walled.

The two distinct physiographic regions are evident in the X-Band SAR image mosaic in Figure 2. The boundary between the Yazoo Delta region to the West and the Loess Hills to the East bisects the image from the NE to SW.

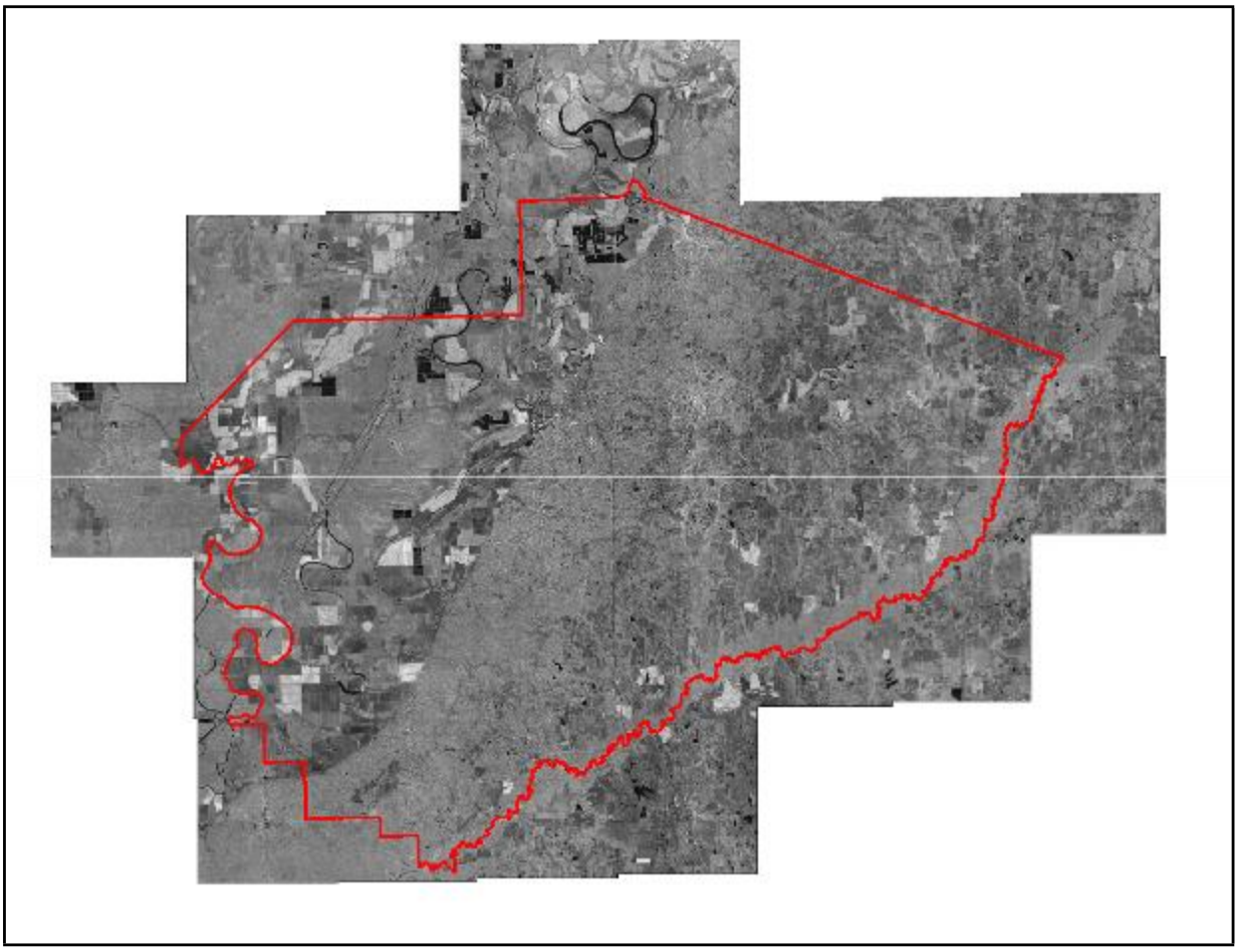


Figure 2 GeoSAR X-Band Image Mosaic of Yazoo County, Mississippi (red outline).

4 GEOSAR DATA ACQUISITION

4.1 Collection Plan

GeoSAR collected a total of 16 lines of data in 4 sorties between August 29, 2007 and August 31, 2007 (Table 1). The data was acquired both during the day and in the evening, demonstrating that the GeoSAR system can collect during time periods when optical sensors would not be able to operate. The GeoSAR also collected through cloud cover, as described in the next section. The lines were collected to cover the entire county with NS looking and EW looking data with at least 4 times redundant coverage. The average line length collected was 122 km. Figure 3 shows the basic flight plan, with the coverage redundancy color coded.

Table 1 Acquisition report for Yazoo data collects. Some acquisitions represent re-collects of the same line. The sorties are color coded.

USDA Acquisition Flight Report

Flight	Start	Stop	Elapsed	Line
20070829F1	5:01:53	5:04:43	0:02:50	YazooAll_02%001#010
20070829F1	5:17:17	5:27:58	0:10:41	YazooAll_02%001#009
20070829F1	5:33:35	5:44:36	0:11:01	YazooAll_02%001#008
20070829F1	5:49:43	6:01:30	0:11:47	YazooAll_02%001#007
20070829F1	6:06:39	6:17:50	0:11:11	YazooAll_02%001#006
20070829F1	6:22:31	6:34:32	0:12:01	YazooAll_02%001#005
20070829F1	6:39:44	6:50:54	0:11:10	YazooAll_02%001#004
20070829F1	6:55:47	7:07:01	0:11:14	YazooAll_02%001#003
20070829F1	7:14:25	7:24:32	0:10:07	YazooAll_02%001#016
20070829F1	7:29:03	7:39:28	0:10:25	YazooAll_02%001#015
20070830F1	11:06:07	11:15:16	0:09:09	YazooAll_02%001#017
20070830F1	11:20:35	11:30:44	0:10:09	YazooAll_02%001#016
20070830F1	11:35:36	11:46:13	0:10:37	YazooAll_02%001#015
20070830F1	11:50:42	12:01:45	0:11:03	YazooAll_02%001#014
20070830F1	12:05:56	12:16:48	0:10:52	YazooAll_02%001#013
20070830F1	12:32:36	12:43:19	0:10:43	YazooAll_02%001#012
20070830F1	12:48:04	12:54:51	0:06:47	YazooAll_02%001#011
20070830F1	13:10:57	13:21:01	0:10:04	YazooAll_02%001#011
20070830F1	13:33:23	13:43:06	0:09:43	YazooAll_02%001#001
20070830F1	13:47:40	13:56:16	0:08:36	YazooAll_02%001#002
20070830F2	18:51:48	19:01:57	0:10:09	YazooAll_02%001#010
20070830F2	19:09:14	19:20:58	0:11:44	YazooAll_02%001#007
20070830F2	19:25:15	19:36:43	0:11:28	YazooAll_02%001#006
20070830F2	19:43:36	19:54:43	0:11:07	YazooAll_02%001#003
20070831F1	16:59:11	17:08:57	0:09:46	YazooAll_02%001#010
20070831F1	17:18:24	17:30:31	0:12:07	YazooAll_02%001#005
20070831F1	17:34:45	17:46:08	0:11:23	YazooAll_02%001#004

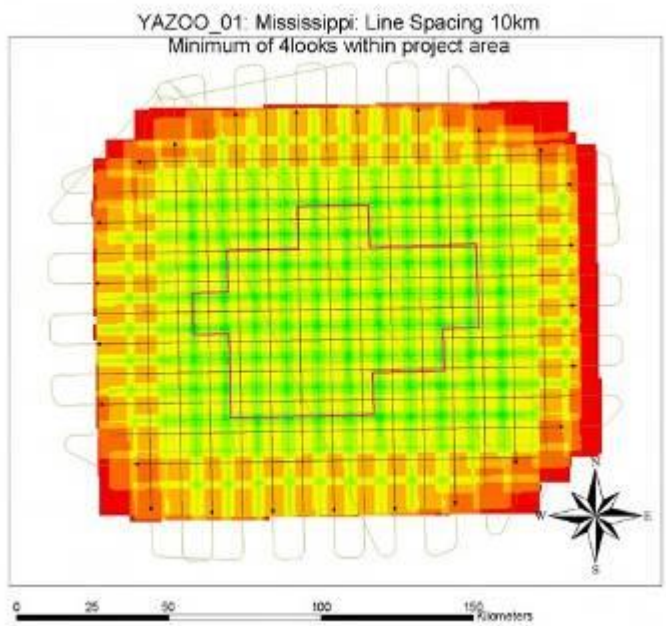


Figure 3 The GeoSAR flight plan for Yazoo County, Mississippi, showing the deliverable quad sheet layout. The figure is color coded by the number of looks over each point on the ground. Yellow indicates 4 looks, while green is more than 4 looks.

4.2 Local Weather at Time of Collect

An important aspect of radar data collection is the ability to collect in overcast weather conditions, during periods where optical imaging sensors typically used in NAIP data collections would not be able to operate. The NOAA GOES-12 satellite images collected during August 29-31 are shown in Appendix A. These images record the fact that during the daylight hours of this time period optical sensors would have been largely weathered out from collecting data. The images show the typical cloudy weather patterns in the Southeastern United States during late summer. Appendix A also contains the NOAA archive 24-hour precipitation maps for that time period, showing that there were persistent cycles of precipitation in that same time period over the Yazoo Mississippi area.

5 TRUTH DATA COLLECTION

5.1 Ground Truth

A ground campaign was conducted in conjunction with the radar collection. A team of three EarthData scientists, accompanied by the local USDA extension officer, visited approximately 25 sites in Yazoo County from August 26 to August 29, 2007. At each site, data was collected on individual fields regarding the crop type, local soil roughness, row spacing, plant height, width, spacing, and predominant compass direction of rows. Approximately 220 ground-based photographs were taken on August 27 and August 29, 2007. The photographs assist in the evaluation of field and crop conditions. Soil cores were collected in 5 major soil types for later

laboratory analysis for bulk density, soil moisture and electrical conductivity. Examples of the ground truth data are tabulated in Appendix A.

Two trihedral X-band corner reflectors were placed in the study area, with their locations precisely measured through GPS. The corner reflectors will be used for absolute geometric and radiometric control of the radar data.

5.2 Airborne Reconnaissance

The team conducted an airborne reconnaissance for 2.5 hours of flight time in a light aircraft on August 26, 2007 for the acquisition of approximately 550 oblique, natural color images from an altitude of 150-200 meters. During the flight in which the aerial photos were taken, a GPS unit collected and stored flight track points throughout the entire flight. The flight track points were later downloaded and converted into a point shapefile with the time stamp of each point recorded in the attribute table. Each photo that was taken was also tagged with a time stamp that could be linked to a flight track point with a similar time. Each photo file name was manually inserted into the flight track point shapefile as an attribute of the specific point that matched its time. Each photo was attached to only one point. By using the hyperlink function in ArcMap, flight track points with an associated photo could be clicked on when displayed in ArcMap and the photo itself would be opened up in a viewer window (Figure 4). This allowed the photo to be compared directly with the imagery displayed in ArcMap (particularly the NAIP imagery) so that features in the photo could be located in the imagery and photo-identified.

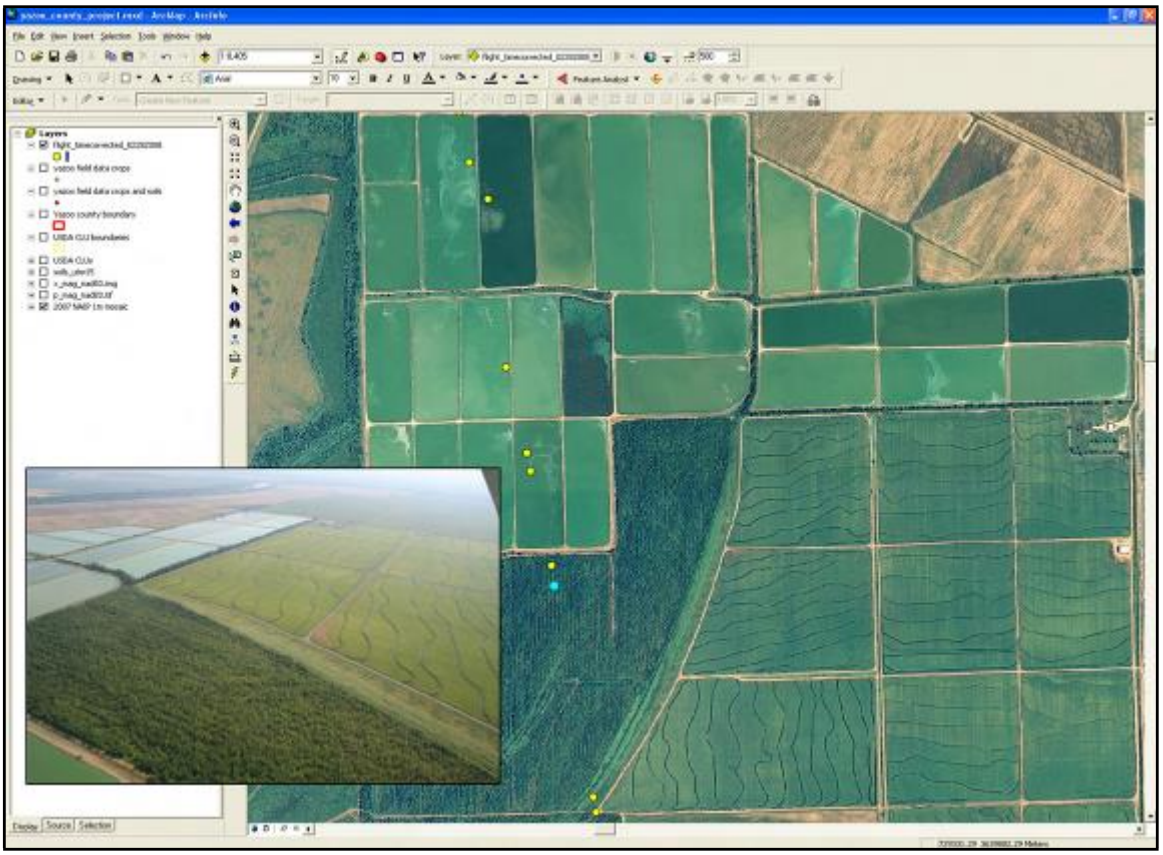


Figure 4 ArcMap project showing GPS flight track points on top of 2007 NAIP true color orthoimagery and aerial oblique photo (lower left) hyperlinked to highlighted point.

6 GEOSAR DATA PROCESSING

6.1 Frequency Allocation

Like all radar mapping system, GeoSAR emits broadband electromagnetic energy while acquiring data. Throughout the U.S. and its territories, GeoSAR is fully licensed by the U.S. Federal Communications Commission (FCC) for transmission as a primary user in X-band and as a secondary user for transmission in P-band. EarthData requires frequency permission for X-band transmission on 9630–9790 MHz and P-band transmission on 270–430MHz. GeoSAR is capable of arbitrarily notching the P-band waveform as required to prevent interference to primary users of the P-band frequencies utilized by GeoSAR. During the Yazoo Mississippi data collection, the GeoSAR P-band signal was notched to approximately 60% of its bandwidth. This notching reduced the image data quality of the P-band product.

6.2 Profiling LIDAR

A profiling LIDAR was added to the GeoSAR aircraft in 2004 to enhance the GeoSAR ability to collect high quality terrain data in the presence of dense foliage. The LIDAR is based upon the field-proven Leica ALS-40 LIDAR design. Leica, under contract to EarthData, modified a standard production instrument from a scanning

system with limited collection altitude to a profiling system with significant collection altitude. This profiling LIDAR is unique to the GeoSAR system and has been designed specifically to provide the following capabilities:

- Generate high density top surface and bottom surface terrain profiles (single line of points, versus an array image) from as altitudes up to as 40k ft MSL to support high quality radar DEM generation with minimal ground control
- At 30kft above terrain, the lidar spot size is 3m, the vertical accuracy less than 50cm, and the horizontal planimetric accuracy less than 3m.
- The LIDAR collects range and reflectance intensity for up to three returns (first, last, and middle) at a 10KHz pulse repetition rate (PRF). At a collection speed of 425kts (218m/s) this corresponds to a ground sample spacing of about 2.2cm between pulses.
- At 30kft above terrain, absolute height accuracy is less than 50cm and absolute horizontal accuracy is 3m or better

6.3 GeoSAR Production Flow

An overview of the overall GeoSAR process is shown in Figure 5. Data collection is highly automated. Before take-off, flight lines and radar operating parameters (such as gain settings, operating modes, and look angles) are loaded into an automatic radar controller (ARC) aboard the aircraft via an optical radar command disc. The ARC starts and stops the radar data collection as the plane passes preplanned pre-computed way-points. The plane flies in a straight line at a steady altitude over the mapping site, and the radar system automatically turns off during turns.

The GeoSAR system uses an Ashtech Z-12 GPS receiver used in post flight dGPS processing. The position of the P-band and X-band antennas as they move during flight are measured by an Antenna Positioning Measuring Unit (APMU that is mounted in a pod on the center of the fuselage. The lasers and tracking cameras inside this pod are combined with GPS data and inertial navigation measurements to precisely determine the orientation of the antennas for each transmitted radar pulse.

During a flight, which may last as long as 4.5 hrs, the raw radar returns are recorded on a high capacity digital direct recording system, storing up to 16TB worth of radar data.

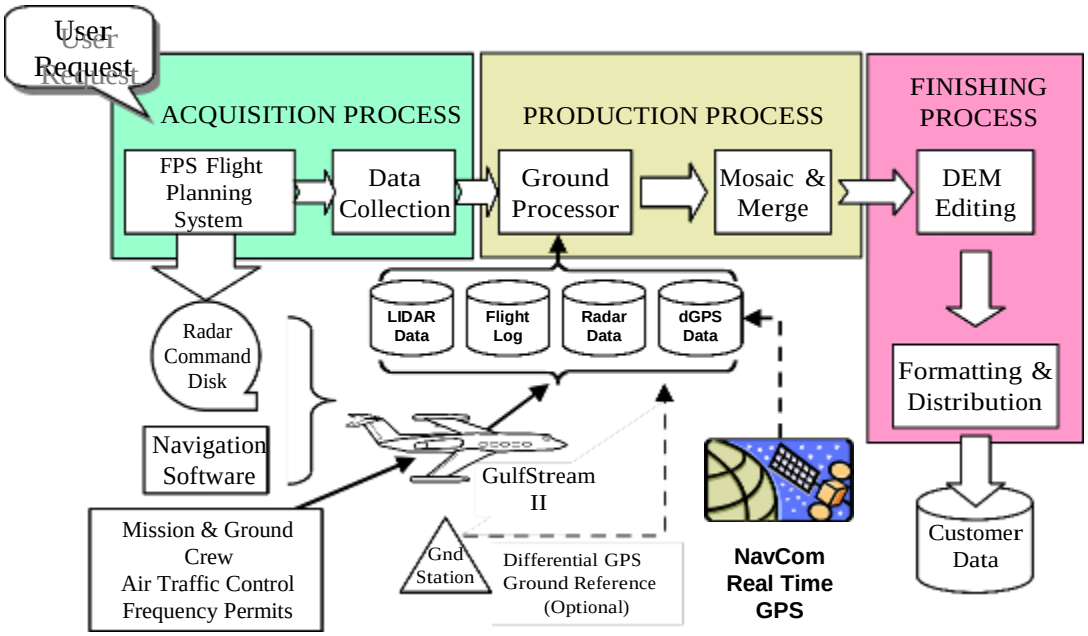


Figure 5 High level flow diagram of GeoSAR production process.

6.4 DATA PROCESSING STEPS

The processing center for GeoSAR data is located at the Fugro EarthData Incorporated (FEDI) facility in Frederick, Maryland, USA.

6.4.1 X-band and P-band Data Processing Production Steps

Seven major activities take place in this production step:

1. **Data Take Ingestion** is the process of incoming (EDI) reception and logging of the disk recorders and auxiliary data sent from the field into the production database. The base station GPS data is combined with the aircraft (real-time) GPS data to create a dGPS location for the aircraft relative to the base station ground control point. This data is entered into the production database.

2. **Motion Measurement Processing (MMP)** examines the auxiliary data for motion quality and prepares the parameters necessary for transferring the raw data off of the disks. MMP is the beginning step of the Ground Processor. The motion data is sent through a QA process to verify its quality. This process is run after each sortie, prior to the aircraft leaving the collection area.

   The MMP inputs the dGPS aircraft position location data and combines it with the auxiliary antenna motion data to generate a Time Varying Parameter (TVP) file for the data take, which is used by the X-band and P-band processors to motion compensate the raw data.

3. **Disk Transfer** is based upon data obtained from the MMP process, which identifies what data is located where on the disk drives. The Disk transfer process requires 2–3 times real time to complete

the transfer from the high-density aircraft disc modules into a format suitable for processing. Data transfer rates approach 200MBS.

4. **X-band Swath Processing** inputs operator or MMP specified parameters and outputs a co-registered X-band reflectance image and DEM at the specified ground sample distance. This is a computationally intense process, limited to about 10-quads/hour throughput. If slant plane products are required, then the slant plane data is also output from the processor.

5. **P-band Swath Processing** inputs operator or MMP specified parameters and X-band DEM and outputs a co-registered P-band reflectance image and DEM at the specified ground sample distance. This is a computationally intense process, limited to about 8-quads/hour throughput. If slant plane products are required, then the slant plane data is also output from the processor.

6. **Steps 2, 3, and 4**, while sequential, can be executing in parallel on different swaths. Up to a dozen different swaths may be in processing at single time.

7. **QC Swath** occurs when the radar processors have completed their reduction of the raw data into reflectance and DEM swaths. This is a manual operation where each swath is examined for processing anomalies, such as phase-unwrapped regions, ambiguity jumps, of noisy data. Approximately 20% of this data will be reprocessed with different input parameters to mitigate the anomalies.

8. **Ready to Mosaic** is the final QC check of the processed swaths to ensure that all the available data has been processed correctly and enough data is on hand to generate a composite mosaic for (a large portion of) the project area.

6.4.2 Wide Area Mosaicing Production Steps

Three major activities take place in this production step:

1. **Swath Processing** involves six iterative steps:

    a. Segmentation—chop the swaths into segments containing usable data, eliminating sub specification data (usually a result of severe motion artifacts due to turbulent weather during the data take).

    b. Swath De-Tilt—remove residual linear tilt in range for the swath (this is minimized by calibration, but might be required on wide swaths). LIDAR profiler data may be used in this step.

    c. Region Match Exclusion—mask out large water regions so they do not generate spurious inter-swath match points (this region is usually defined during the flight planning process).

d. Control Point Extraction—using ground control information, extract the position of the radar corner reflectors or photo identification points to a fraction of a pixel. LIDAR profiles are also matched against every intersecting swath.

e. Swath Point Matching—in the overlap regions of any two swaths, find all reflectance (and/or DEM) points which correspond to the same point on the ground. This process generates hundreds to thousands of match points per swath-pair.

f. Match Point Culling—severely cull the match points to retain only the very best, highly correlated points.

2. **Mosaic Affine Transformations** is the process of combining the match points (produced by item-f above) and the ground control points (item-d above) in a weighted least mean squared error affine transformation[1] for geographic registration to generate a composite mosaic of the region. Data from overlapping swaths are averaged in common areas and feathered at the boundaries. The same transformation parameters are applied for both the reflectance and height data. Null values in one swath are replaced by non-null values in overlapping swaths. Orthogonal cross-tie lines and LIDAR profiles provide powerful near range constraints to remove residual tilts. Ground control points remove residual systematic vertical bias.

3. **QC and Mosaic Statistics** are collected for the composite mosaic. Independent ground control and LIDAR profile data is used to estimate overall product quality. An important QC step is the creation of a height error map, which measures the standard deviation at each pixel in the output mosaic derived from the multiple independent looks. If the (3 to 7) separate swath measurements associated with a given pixel location are consistent, then the standard deviation of the height difference between them will be small. Height errors exceeding a project specific criteria, say 1m, must be remediated by further swath processing. This method for checking the DEM for internal consistency is one of the great benefits derived from GeoSAR's inherent dual sided, multilook collection strategy.

   Once the mosaic is declared completed by the mosaicing staff, independent ground control values are used to assess the overall quality of the final mosaic. If the result is not satisfactory, then the anomalies are noted and the composite mosaic is redone (usually) from the swath point matching or from the affine transformation step.

[1] This operation is similar to the bundle adjustment done in photogrammetry.

6.4.3 Product Finishing and Packaging Production Steps

Six major activities take place in this production step:

1. **Crop to Quads** chops the composite mosaic into units suitable for ingestion by the DEM and image editing workstation. The size chosen usually corresponds to the client's final mapping unit size.
2. **DEM Edit** is the manual process for removing residual radar artifacts from the DEM according to the client's specifications. Bald earth DEM editing is significantly more labor intense than is top surface editing. Although EarthData has developed many LIDAR-based "bulldozing" tools to facilitate tree and building removal from DEMs, the process still requires "mapping" judgment to arrive at a high quality product.
3. **Image Edit** is cosmetic smoothing or removal of radar artifacts to enhance the photo-like quality of the imagery. EarthData does not normally image edit the data unless the client specifically requests it because many end users feed the data into other analytical tools, which are adversely affected by de-speckling and spatial smoothing. EarthData has a full complement of standard image edit routines that can be applied to meet the client's needs.
4. **QC** is the final check on product data quality. Data failing to meet the delivery spec is sent back to production for remediation.
5. **Formatting** publishes the data onto client specified media and formats, and builds the metadata files according to the delivery specification.
6. **Delivery** is the process of shipping the data to the client and following up that the data is found to be satisfactory. CD and DVD are popular delivery formats. However, data sets from 2GB–150GB are best delivered by USB hard disk. Delivery media for larger data sets, e.g., raw phase history of complex slant plane data is on a case by case basis.

Remediation is a feedback loop where intermediate data is reprocessed at the point in the production chain capable of making it right. Remediation occurs internally within the GeoSAR factory as well. Under extreme circumstances, remediation may require collection and processing of additional lines. EarthData will not knowingly ship any substandard data.

7 COMPARISON BETWEEN GEOSAR IMAGERY, NAIP IMAGERY AND CLU DATA

FEDI received Common Land Unit (CLU) vector data and 2007 National Agriculture Imagery Program (NAIP) imagery as government furnished information. One of the tasks of this project was to determine the utility of the GeoSAR SAR imagery for the extraction of CLU vector data, and to determine the degree of registration of the SAR data with the CLU data. The CLU data is extracted from NAIP imagery, so it is anticipated that most of the features in the CLU data would be visible on the 2007 NAIP imagery. Figures 6 and 7 below show two

examples of conjugate areas of NAIP imagery and GeoSAR X-Band imagery both without and with CLU data superimposed in yellow. The CLU data was not used as control for the GeoSAR imagery, which was mosaiced without any ground control. So, the degree of registration between the CLU and SAR data is completely due to the native geodetic accuracy of the two products.

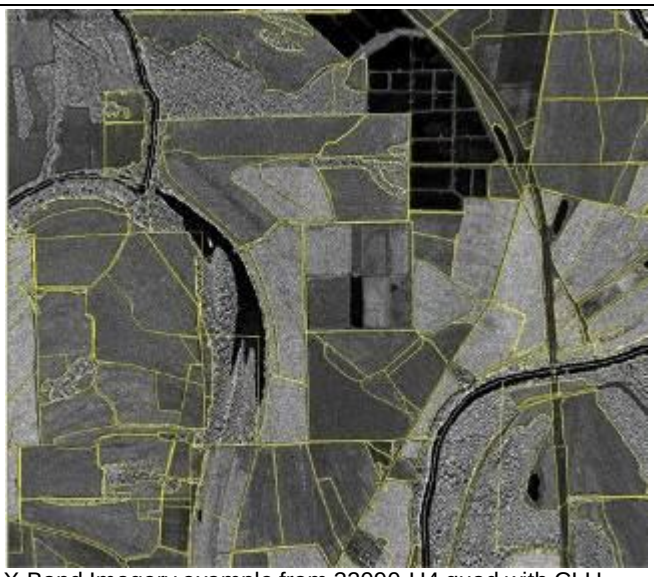
Figure 6, which is an example of data from the 32090-H4 quad, demonstrates a good correspondence between the X-Band SAR and the CLU vectors. Most of the vectors correspond to a linear feature that is easily recognizable on the SAR imagery. There are vectors that do not seem to have a correspondence with linear features or field boundaries. In most of these cases, however, the field boundary is not evident on the NAIP imagery either.

Figure 7 is an example of imagery and CLU data from the 32090-F5 quad. The imagery shows a similarly good correspondence between the CLU data boundaries and linear features from the GeoSAR data. This example shows some of the problems inherent in the NAIP and CLU data. The NAIP imagery has a cloud, which partially obscures the area with the cloud and its shadow. The CLU data vectors seem to have been drawn a little arbitrarily. In several cases, the vectors appear to be tracing the curving features associated with filled-in river meanders. These features do not seem to depict actual field boundaries in most cases. The polygons defined by these vectors also seem to encompass several crop types. This problem will evidence itself when attempting to evaluate classification results compared to CLU reporting data.

NAIP 2007 Imagery example from 32090-H4 quad.

NAIP 2007 Imagery example from 32090-H4 quad with CLU

X-Band Imagery example from 32090-H4 quad.

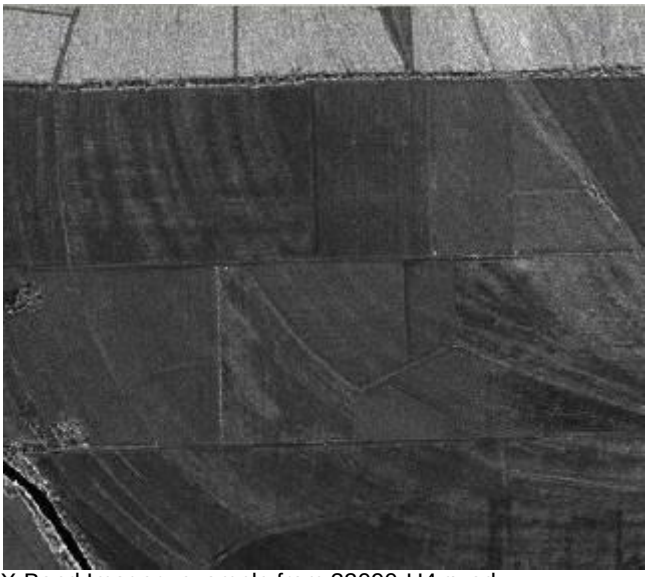


X-Band Imagery example from 32090-H4 quad with CLU

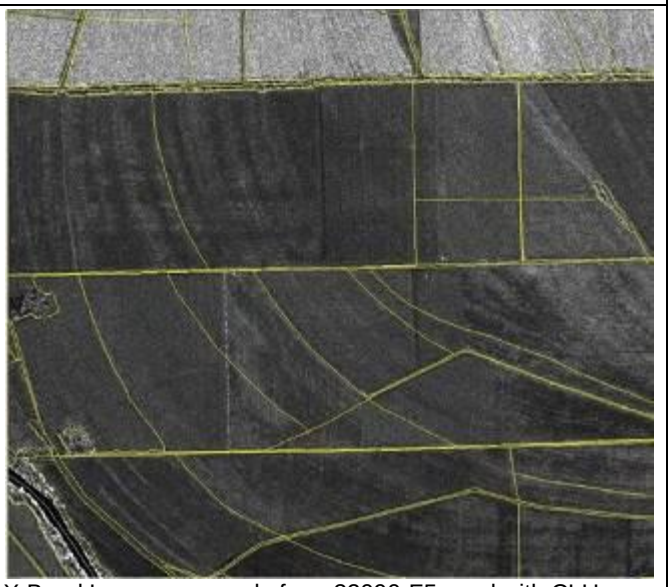
Figure 6 Imagery Examples from 32090-H4 Quad.

NAIP 2007 Imagery example from 32090-F5 quad.

NAIP 2007 Imagery example from 32090-F5 quad with CLU



X-Band Imagery example from 32090-H4 quad.



X-Band Imagery example from 32090-F5 quad with CLU

Figure 7 Imagery Examples from 32090-F5 Quad.

8 GEODETIC ACCURACY OF GEOSAR DATA

Because of the sparseness of the available ground control (with only 2 radar reflectors), the decision was made to process the mosaic product over Yazoo County with no control, and use the radar reflectors for check purpose. Also, while NAIP imagery and common land unit (CLU) data was available, it was decided not to include any control from these data sets in the mosaic. This way, the comparison with those data sets will represent a fully independent quality assessment. The only control that was used was applying the lidar points in the post processing of the elevation data. Here, a Z-bump was applied to the elevation models based on the mean height error compared to lidar profiler points.

8.1 Vertical Accuracy

As discussed in Chapter 6, tie points are found between the overlapping flight lines, and the match points are used as observations in an affine block adjustment program that tied the images together. This block adjustment is then used to produce a mosaic product of both the images and the elevation models. The affine solution for the X-Band block adjustment had a residual error of 93 cm. This was the average 3D residual error of the 132000 match points used to tie the 32 image swaths together, and represents a good estimate of the relative vertical error of the solution.

The resulting elevation model mosaics were compared to lidar data points. The lidar data is collected concurrently with the radar with a profiling lidar that operates on the GeoSAR platform. The lidar points are spaced approximately 5 cm apart, and have 20-30 cm absolute elevation accuracy. The lidar points were filtered using automated software so that only points that were in open terrain were used in the comparison.

The GeoSAR X-Band digital elevation model was compared with 695 lidar points in open terrain. The distribution of the point difference had a standard deviation of 92 cm, which tracks well with the affine solution statistics. The mean difference for the elevation points was 4.5 meters. Since the affine solution did not contain any control points, the mean difference was easily reduced to near zero with a controlled solution. The histogram of these points is shown below in Figure 8.

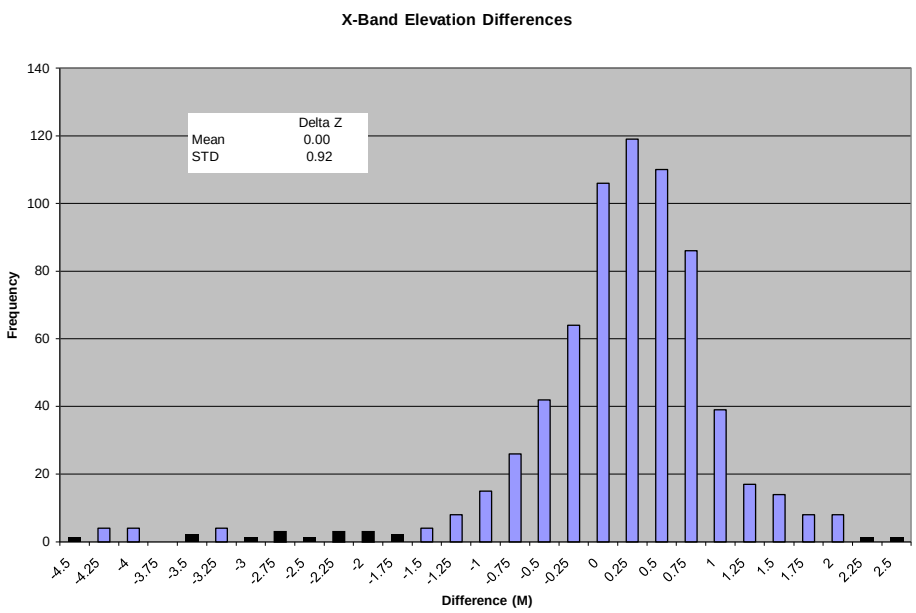


Figure 8 Frequency histogram of X-Band DEM - lidar differences.

8.2 Horizontal Accuracy

8.2.1 Comparison to Corner Reflectors

As part of the ground truth collection campaign during the GeoSAR collection, two corner reflectors were installed within the collection area. The corner reflectors (Figures 10 and 13), are designed to reflect the radar directly back to the antenna from all aspects. The reflectors are positioned to within 1 meter by GPS. The reflectors size makes them optimal for X-Band frequency, but they tend not to be visible at the P-Band frequency. For this reason, they did not appear on the P-Band imagery.

8.2.1.1 Reflector 1

Reflector 1 was placed in a large pasture (Figures 10 and 11) located in the sheet 32090-G3. The reflector position was measured through the use of a handheld GPS that was placed on the apex of the reflector and collected positions for about 30 minutes. The presumed horizontal accuracy of this point is about 1 meter. The position of the reflector point on the X-Band mosaic image was determined by measuring the center of the bright return (Figure 9) using ERDAS 9.0.

Table 2 Comparison between XY position for Reflector 1 as measured by GPS and Image Mosaic.

Reflector 1 Position		X E-W	Y N-S
GPS Measured Position		756767.17	3627537.12
X-Band Measured		756761.21	3627537.51
Xband – GPS (Meters)		-5.96	0.39

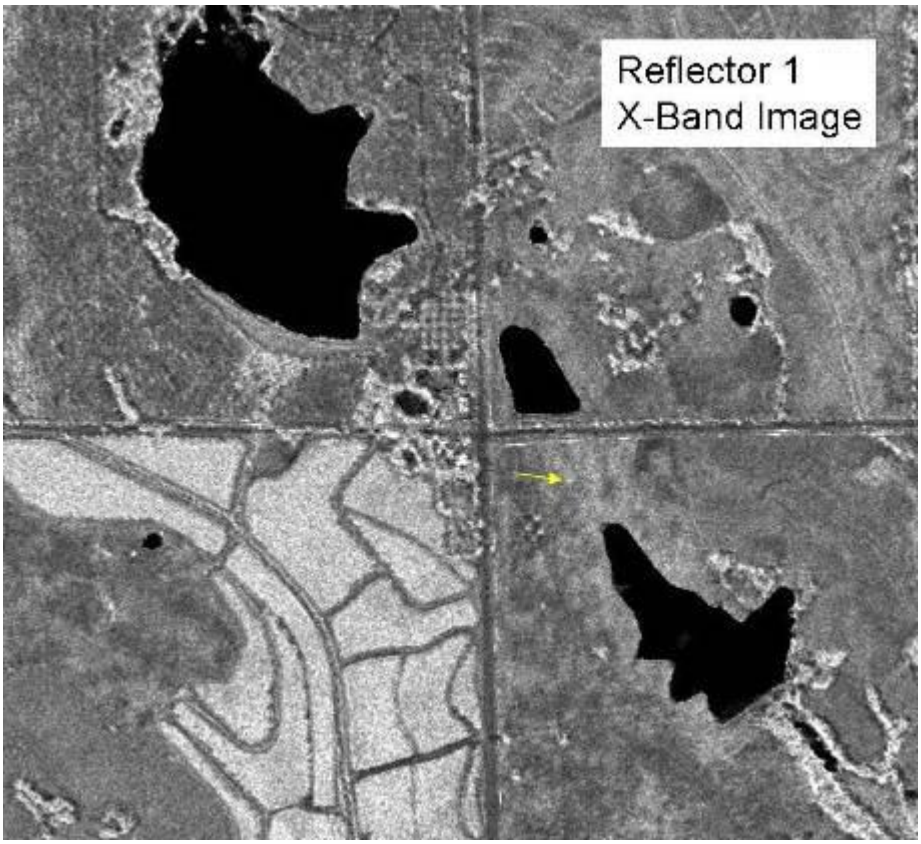


Figure 9 X-Band image with location of reflector 1 annotated.

Figure 10 Ground photo of the reflector 1.

Figure 11 Air photo with reflector 1 position annotated.

8.2.1.2 Reflector 2

Reflector 2 was placed on a grass covered peninsula in a man-made lake. The reflector position was measured in a manner similar to reflector 1 described above.

Table 3 Comparison between XY position for Reflector 2 as measured by GPS and Image Mosaic.

Reflector Positions		X E-W	Y N-S
GPS Measured Position		730734.46	3642288.98
X-Band Measured		730727.68	3642292.48
Xband – GPS (Meters)		-6.78	3.50

Figure 12 X-Band image with location of reflector 2 annotated.

Figure 13 Ground photo of the reflector 2.

Figure 14 Air photo with reflector 2 position annotated.

8.2.2 Reflector Results

Ideally, the reflector positions would have been carried in the block adjustment as control. Because there were only 2 reflectors available for this project, it was decided to use them as check points, and to process the data

without the use of ground control. The reflector results, while only consisting of 2 data points, show that the GeoSAR data likely exhibits a horizontal bias of approximately 6 meters.

8.2.3 Comparison to Common Land Unit Data and NAIP Imagery

The horizontal agreement between the GeoSAR X and P-Band radar image mosaics and the Common Land Unit (CLU) data and NAIP imagery provided by the USDA was measured. Common points appearing on all 4 data types were measured using the mensuration capabilities in ERDAS 9.0. All 4 data sources were displayed simultaneously in linked viewers. Points were measured manually and recorded in a spreadsheet for analysis. A total of 92 points were measured on 6 quad sheets (32090-E4, 32090-G4, 32090-H5, 32090-F4, 32090-H1, and 32090-G1).

8.2.3.1 Accuracy of the CLU and NAIP

The CLU vectors and 2-meter resolution NAIP imagery were received as government furnished information (GFI) by Fugro EarthData, and used without any modification. It was assumed that the CLU data were extracted from the NAIP imagery, and that the two data sources would have similar horizontal accuracy. According to USDA FSA material on the internet:

> The horizontal accuracy of NAIP has always been tied to Government furnished baseline imagery datasets, where 1 meter Ground Sample Distance (GSD) imagery has an accuracy that matches the reference baseline imagery dataset within 5 meters (90% confidence), and 2 meter GSD imagery matches the baseline imagery within 10 meters (90% confidence). This type of accuracy is defined as relative accuracy, and the deliverable is not tied to true ground, rather another imagery dataset, which was tied to ground within a certain confidence.

From this, it is assumed that the absolute and relative accuracy of the GFI is on the order of 10 meters. Table 4 and Figure 15 below shows the statistical results of the measurements. The results indicate that roughly the same bias shown in the reflector positions also is evident in comparing the CLU and NAIP data. The deltas between the SAR imagery and the CLU and NAIP imagery can be partially explained by operator misreading, the difference in resolution between the data sets. Overall, the differences are reasonable given the 10 meter uncertainty in the NAIP and CLU data.

Table 4 Measurement results comparing X and P Band mosaic images with CLU and NAIP data.

	Xband-Pband		Xband-CLU		Xband - NAIP		Pband-CLU		Pband-NAIP		CLU-NAIP	
	E-W	N-S	E-W	N-S	E-W	N-S	E-W	N-S	E-W	N-S	E-W	N-S
Mean	3.86	-1.47	-7.44	-5.97	-5.67	-3.03	-11.29	-4.49	-9.52	-1.56	1.77	2.94
STD	6.91	7.88	7.73	7.97	6.11	6.01	7.75	8.12	6.28	7.10	7.12	5.81

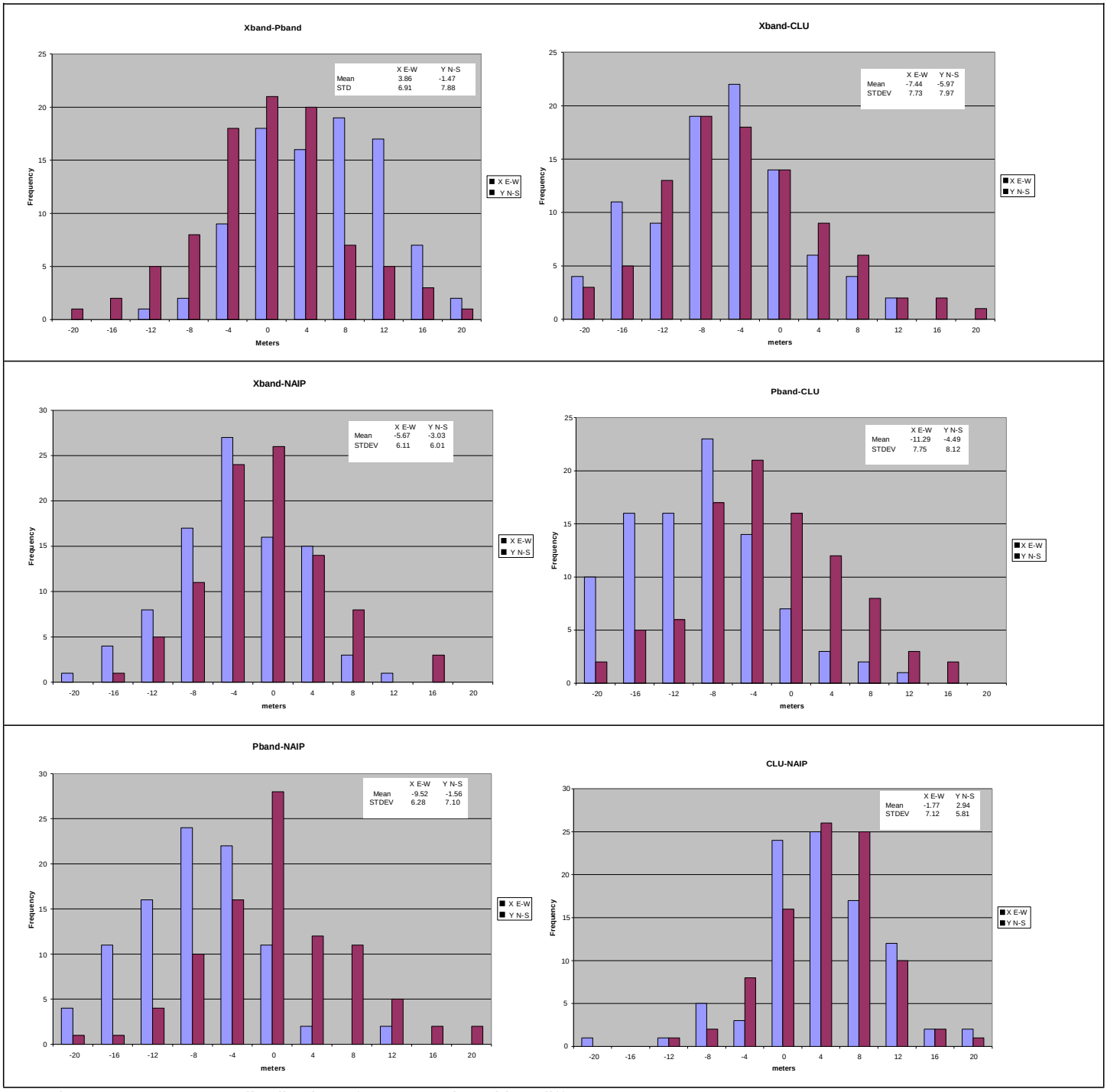


Figure 15 Frequency distribution of measured position differences.

9 CROP AND LANDCOVER CLASSIFICATION EXPERIMENTS

The purpose of this analysis is to test the applicability of GeoSAR data for mapping different crop types. Fugro EarthData, Inc. chose a semi-automated image classification methodology that has proven successful for creating land cover/land use and benthic habitat maps from high resolution digital orthophotography (Green, K.

and Lopez, C. 2007[2]). GeoSAR imagery has previously been used as an ancillary dataset in the analysis for land cover/land use mapping, and since GeoSAR produces high resolution digital imagery it was determined that this data would be an ideal candidate dataset for testing the same image classification techniques used for classifying multispectral imagery.

The techniques previously used for classifying land cover/land use and benthic habitats are a combination of object-oriented image analysis and Classification and Regression Tree (CART) analysis. Object-oriented image analysis is an extremely important technique in the classification of high resolution imagery. This analysis uses image objects or segments (i.e. groups of pixels delineated as polygons) for mapping instead of using individual pixels, allowing for the incorporation of shape and context into the creation of habitat data. While powerful in the classification of moderate resolution data, object-oriented classification is pivotal for automated classification of high resolution satellite or airborne multispectral imagery because of the mixture of both dark and illuminated pixels.

CART analysis is a statistical technique for image classification that builds tree diagrams for predicting variables from categorical and continuous data (Breiman et al., 1984[3]). It “mines” the sample data and builds rules which are if-then statements in hierarchical “trees” that condition the prediction of habitat classes. CART is powerful because

- it can accept both continuous and categorical data inputs,
- the results are easy to interpret,
- unlike a maximum likelihood classifier, no assumptions are required concerning the distributions of the independent variables,
- it identifies simple and complex relationships between variables that other techniques might not uncover, and
- it forces consistency and analytical rigor into the segment labeling process.

---

[2] Green, K. and Lopez, C. 2007. Using Object-Oriented classification of ADS40 Data to Map the Benthic Habitats of the State of Texas. Photogrammetric Engineering & Remote Sensing. Vol 73, No. 8: 861-865.

[3] Breiman, L., J.H. Friedman, R.A. Olshen, C.J. Stone, 1984. Classification and Regression Trees, Chapman & Hall, International Thomson Publishing, 358pp.

To run the CART analysis, two inputs are required. One input is the training sites which are the examples of each class that are being classified. This is the dependent variable (class). The second input is one or more datasets that are either raster or vector data which comprise the independent variables. The main independent variable input is generally the imagery used for the project, but can also include any other ancillary datasets such as other imagery or vector maps.

9.1 Data Preparation and Training Site Collection

The first step in the CART analysis was to collect the imagery used for classification. The main purpose of the project was to test GeoSAR data for crop classification. An advantage of the GeoSAR sensor is that multiple products are produced, all of which may contain valuable information for classifying different crop and cover types. The different GeoSAR products used in the CART analysis are listed in Table 5. Also available was 1m 3-band true color digital ortho imagery collected for the 2007 NAIP program which was used as a reference. All sets of imagery were mosaicked into single images covering the Yazoo county project area.

Table 5 List of datasets used.

Data	Description	Collection Date	Used in CART Analysis
GeoSAR P-MAG	P-band magnitude	August 29-31, 2007	Yes
GeoSAR X-MAG	X-band magnitude	August 29-31, 2007	Yes
GeoSAR X-DEM minus P-DEM height	Height difference between X and P DEMs	August 29-31, 2007	Yes
GeoSAR Σ_0	Radar imagery normalized for local incident angle.	August 29-31, 2007	Yes and No
True Color NAIP orthoimagery	Digital aerial imagery collected for NAIP	Mostly collected ~3 weeks before the GeoSAR data	No, reference data only

The next step was to collect training sites for input into the CART classifier. Training sites were collected in the form of polygons with each polygon being labeled as a single crop class. The classes used in the analysis were:

1. Corn
    a. Standing (uncut)
    b. Cut (harvested)
2. Cotton
3. Soybeans

Corn, cotton, and soybeans were the only crop types tested in this analysis because they appeared to be the most abundant types in the Yazoo project area and these were the only types we could reliably collect enough training data for based on the available field data. At the time of data collection, corn harvesting was occurring

in some fields and had already occurred in others, while other fields were un-harvested, so this class was split into two subclasses: “uncut” and “harvested”.

Rather than creating a new set of polygons from the imagery to classify, USDA Common Land Unit (CLU) polygons for Yazoo County were used in order to simplify the analysis. Only CLU polygons that were homogenous or nearly homogenous in cover type, based on photo interpretation using the GeoSAR and NAIP imagery, were used for training sites. There were many examples of CLU polygons that contain multiple cover types, and sites like these used as training sites for image classification can result in an analysis with high percentages of errors.

Training sites for each of the three crop types were collected based on either the field data collected or the aerial photos taken over the project area using a hand held digital camera. Since the field data was limited, the aerial photos were utilized extensively to identify more fields. By identifying what crops could be seen in a photo, and then locating the crop field’s position in the imagery, it was possible to reliably record that field’s location in a shapefile and attribute it with a crop type.

The total number of training sites for each class is summarized in Table 6. Once all of the sites were collected, a random number generator was used to randomly select a number of sites for each class to be used for accuracy assessment only (also summarized in Table 6). The remaining sites were used for the CART analysis. Because of the limited number of total sites for each class and the need to have 10 sites per class for the CART analysis, only a limited number of sites were available to use in the accuracy assessment. All of the CART training and accuracy assessment sites were intersected with the imagery using the Zonal Stats tool in ArcMap in order to obtain the mean and standard deviation values of each image for each polygon. A shape index was also calculated for each polygon using the formula:

(polygon perimeter * 0.25) / [2 * square root(polygon area * 3.1415)].

Table 6 The number of CART analysis training sites and accuracy assessment sites for each crop type.

Crop	CART Training Sites	Accuracy Assessment Sites
Cut Corn	10	10
Standing Corn	10	18
Cotton	10	14
Soybeans	10	11

9.2 CART Analysis

See5 statistical software was used to perform the CART analysis and accuracy assessment. For the analysis, the crop type label of each training site was the dependent variable and the means and standard deviations of each image band and the polygon shape index were the independent variables. The See5 parameters used for this analysis were a 10 trial adaptive boosting with default values of 25% pruning CF and a 2-case minimum for tree diagram branches. The CART analysis determined which of the independent variables were highly

correlated with each crop type and built a set of classification rules based on the independent variable values (Figure 16). See5 then used the classification rules it generated to classify each of the accuracy assessment polygons. Accuracy matrices summarizing each accuracy assessment polygon's actual label and classified label were built using Microsoft Excel.

```
----- Trial 1: -----

Decision tree:

xmag <= 26.8441:
:...xmag_std <= 4.601: corn cut (16.2/5.2)
:   xmag_std > 4.601: corn standing (7/1.6)
xmag > 26.8441:
:...xmag <= 32.1409: soybeans (9.7/1.8)
    xmag > 32.1409: cotton (7.2)

----- Trial 2: -----

Decision tree:

xmag > 26.8441:
:...pmag <= 19.5931: cotton (11.6/2.5)
:   pmag > 19.5931: soybeans (4.7)
xmag <= 26.8441:
:...xmag > 22.2034: soybeans (8.8/5)
    xmag <= 22.2034:
    :...x_minus_p_std <= 2.2541: corn cut (10.8/4.1)
        x_minus_p_std > 2.2541: corn standing (4.1)
```

Figure 16 Example of a portion of the See5 CART analysis classification rules in the form of a dichotomous tree diagram.

9.3 Accuracy Assessment

Tables 7-10 present the results of the accuracy assessment on the CART analysis mapping. These results are for a fully automated classification model (no manual editing done) that can used to create a full classified map. Tables 7 and 8 present the results of the analysis using all of the GeoSAR products listed in Table 5. The first 4 x 4 error matrix (Table 7) has corn broken out into the two subclasses "cut corn" (harvested) and "standing corn". When corn is broken out into the two subclasses, both the producer's and user's accuracies for each are quite low, indicating that there are substantial errors of omission and commission. However, the majority of the errors are between the two subclasses, so when they are combined into a single "corn" class (Table 8), the user's and producer's accuracies increase to 89% and 96%, respectively, which are extremely high accuracies. So while there is trouble distinguishing between cut corn and standing corn, the GeoSAR data are very accurate for identifying corn as a single class. Accuracies for cotton and soybeans are also quite high for a fully automated map. The overall map accuracy when cut corn and standing corn are single classes is 60%, which is good for an automated map, but increases to 85% when they are combined into a single class, which is at or near the overall map accuracy that is typically specified for a mapping project. Results such as this would mean that only

a small amount of manual editing would be required in order to achieve the typical 80-90% individual class accuracies and overall map accuracy required for a mapping project.

Table 7 Accuracy Assessment error matrix with corn subclasses using all four GeoSAR data products.

		Reference Data					
	4 X 4	Cut Corn	Standing Corn	Cotton	Soybeans	Total Map	User's Accuracy
Map Data	Cut Corn	4	8			12	0.33
	Standing Corn	5	8		1	14	0.57
	Cotton			12	2	14	0.86
	Soybeans	1	2	2	8	13	0.62
	Total Reference	10	18	14	11	53	
	Producer's Accuracy	0.40	0.44	0.86	0.73		0.60

Table 8 Accuracy Assessment error matrix with single corn class using all four GeoSAR data products.

		Reference Data				
	3 X 3	Corn	Cotton	Soybeans	Totals	User's Accuracy
Map Data	Corn	25		1	26	0.96
	Cotton		12	2	14	0.86
	Soybeans	3	2	8	13	0.62
	Totals	28	14	11	53	
	Producer's Accuracy	0.89	0.86	0.73		0.85

Tables 9 and 10 presents the results of a second CART analysis but with the GeoSAR Σ_0 product removed from the analysis. For reasons that are unknown, removing this dataset has the affect of increasing the individual accuracies for both cut corn and standing corn (4 x 4 error matrix in Table 9). However, the majority of the error is again between the two classes, so when they are combined into a single “corn” class, the accuracies for all three crop classes and the overall map accuracy (3 x 3 error matrix in Table 10) are similar to the first analysis presented in Table 7. While overall results for the two analyses are very similar, the ability to accurately map cut

corn and standing corn increases when the Σ_0 product is removed and only the GeoSAR P-Mag, X-Mag, and X-DEM minus P-DEM height layer are used in the analysis. This is a very interesting result that the use of a single data layer actually decreases accuracies for some classes and is worth exploring further to determine the reason why.

Table 9 Accuracy Assessment error matrix with corn subclasses but without the Σ_0 GeoSAR data product.

		Reference Data					
	4 X 4	Cut Corn	Standing Corn	Cotton	Soybeans	Total Map	User's Accuracy
Map Data	Cut Corn	6	5		1	12	0.50
	Standing Corn	4	12			16	0.75
	Cotton			12	2	14	0.86
	Soybeans		1	2	8	11	0.73
	Total Reference	10	18	14	11	53	
	Producer's Accuracy	0.60	0.67	0.86	0.73		0.72

Table 10 Accuracy Assessment error matrix with single corn class but without the Σ_0 GeoSAR data product.

		Reference Data				
	3 X 3	Corn	Cotton	Soybeans	Totals	User's Accuracy
Map Data	Corn	27		1	28	0.96
	Cotton		12	2	14	0.86
	Soybeans	1	2	8	11	0.73
	Totals	28	14	11	53	
	Producer's Accuracy	0.96	0.86	0.73		0.89

Tables 11 and 12 present KAPPA analysis results to determine whether or not the CART analyses are significantly different than by random chance and if there are pair-wise significant differences between error matrices(Congalton, Russell and Kass Green. 1999[4]). Table 11 summarizes the KAPPA results for all four error matrices. Each matrix is significantly better than random (indicating the classification is significantly better than classifying by random chance) as indicated by the Z statistics. Any Z statistic (absolute value) greater than 1.96 is significant at the 95% confidence lever. Table 12 summarizes the results of comparing each 4 X 4 error matrix and each 3 X 3 error matrix. The absolute values of the Z statistics for each test are not greater than 1.96 which indicates that error matrices 7 and 9 are not significantly different than each other and error matrices 3b and 4b are also not significantly different than each other. While matrices 8 and 10 do look very similar, there are differences in individual accuracy values between matrices 7 and 8. While they may look like they should be statistically different, the KAPPA analysis results indicate otherwise. This may be due to the low numbers of accuracy sites for each class used in the analysis because of the limited amount of field data collected. Typically a minimum of 30 accuracy sites would be used per class for analysis, so if more sites had been used, the analysis would have been more robust and therefore may have resulted in significant statistical differences. However, this should not dismiss the results of this analysis.

Table 11 Individual error matrix KAPPA analysis results.

Error Matrix	Overall Accuracy	KHAT	Z Statistic
Table 7	60%	0.47	5.236
Table 8	85%	0.76	9.831
Table 9	72%	0.62	7.452
Table 10	89%	0.81	11.766

Table 12 Comparative error matrix KAPPA analysis results.

Error Matrix	Total Number of Accuracy Sites	Z Statistic
7 vs. 9 (4 X 4 matrices)	53	-1.220
8 vs. 10 (3 X 3 matrices)	53	-0.553

[4] Discussion of KAPPA analysis found in: Congalton, Russell and Kass Green. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers. 137pp

9.4 Comparison of Classification results with CLU Reported Data

An important aspect of this study is to determine the utility of the GeoSAR data for monitoring crop compliance by the Farm Services Agency. To this end, after the CART classification analysis was completed, all CLU fields were classified. The actual utility of the CLU data was limited because the boundaries defined by the CLU often did not encompass a single field or single crop type, so a certain amount of confusion in the classification can be expected. Table 13 shows a comparison of the CLU fields as reported by FSA, and the fields as classified by CART analysis. The agreement is rather good for the Corn and Cotton crop types, with 88% and 87% agreement, respectively. The analysis, however, is remarkably different for the Soybean type. The agreement is only on the order of 35% for soybean.

Table 13 Comparison of crop identification by CART analysis with the data reported in the CLU dataset.

		Classified Map			Total CLU reported	% agreement CLU to Classified
		Corn	Soybeans	Cotton		
CLU Reported	Corn	**3256**	411	38	3705	0.88
	Soybeans	429	**315**	105	849	0.37
	Cotton	17	160	**947**	1124	0.84
	Other	7	6	2	15	0
	Total Class.	3709	892	1092	5693	
% agreement Classified to CLU		0.88	0.35	0.87		

This result leads to the question of why the soybean crop analysis was so poor compared to the others. The ground truth reports collected during the FEDI field work are summarized in Table 14. They show a good correlation between the field-identified corn and cotton crops with the CLU data and CART classified labels. For the soybeans, however, 2 of the 3 field reports disagree with the CLU labels. While this is admittedly a small statistic, it is a point of data that would agree with the classification analysis.

Table 14 Comparison of field reports with CLU data and CART classification.

Field Sites				CLU Reported Crop	Classified Crop Label
Site ID	Latitude	Longitude	Field Site Crop Type		
F15	32.9001	-90.4386	corn cut	corn	corn cut
F11	32.8992	-90.4423	corn cut	corn	corn cut
F3	32.8988	-90.4477	corn cut	corn	corn cut
A5	32.8533	-90.7037	corn cut	corn	corn cut
Y10	32.7253	-90.2525	corn cut	corn	corn cut
A2	32.8490	-90.7055	corn cut	corn	corn standing
Y3	32.8997	-90.4499	corn standing	corn	corn standing
Y4	32.7662	-90.4917	corn standing	corn	corn standing
Y7	32.6260	-90.3629	corn standing	corn	corn standing
Y9	32.7488	-90.5043	cotton	No Label	cotton
Y8	32.7411	-90.2569	cotton	cotton	cotton
B2	32.8581	-90.7041	cotton	cotton	cotton
B8	32.8644	-90.7006	cotton	cotton	cotton
B4	32.8605	-90.7005	cotton	cotton	cotton
Y2	32.6248	-90.3661	cotton	No Label	cotton
G10	32.6227	-90.3670	cotton	No Label	cotton
Y5	32.8993	-90.4499	fallow	corn	corn standing
Y6	32.8988	-90.4389	soybeans	soybeans	soybeans
Y12	32.7638	-90.5213	soybeans	corn	soybeans
Y1	32.6293	-90.3714	soybeans	wheat	soybeans

Table 15 below tabulates the comparison between fields identified by photo interpretation (PI) from the low level air photographs collected during the field survey, and the CLU data. These samples show that there is a 90% agreement between the photo interpretations of Soybeans and the CLU data. In other words, there were 20 fields identified as Soybeans in the CLU, and 18 of them were interpreted as soybeans by photo interpretation, the other 2 identified as corn.

Table 15 Comparison between fields identified by photo interpretation and the CLU labels.

	CLU Reported Crop								
Photointerpreted Crop Type	**Corn**	**Cotton**	**Soybeans**	**Wheat**	**Peanuts**	**Grass**	**No Label**	**Totals**	**% Agreement PI/CLU**
Corn	30		2	1		1	5	39	0.77
Cotton		13			1		3	17	0.76
Soybeans			18					18	1.00
Totals	30	13	20	1	1	1	8	74	
% Agreement CLU/PI	1.00	1.00	0.90	0.00	0.00	0.00	0.00		

There is also a good agreement between the photo interpreted field labels and the CART classification results. Table 16 below shows an 81% agreement for soybean fields.

Table 16 Comparison between fields identified by photo interpretation and the CART Classification labels.

	Classified Crop Label				% Agreement PI/Class
Photointerpreted Crop Type	**Corn Cut**	**Cotton**	**Soybeans**	**Totals**	
Corn	29		1	30	0.97
Cotton		15	2	17	0.88
Soybeans	3	2	13	18	0.72
totals	32	17	16	65	
% Agreement Class/PI	0.91	0.88	0.81		

This unfortunately means that the soybean classification result as compared to the bulk CLU data reporting is still ambiguous at this point. There will need to be a deeper analysis to determine the accuracy of the CLU data.

9.5 Classification Conclusions

The results of this analysis show that the ability to use the GeoSAR data products to map corn, cotton, and soybean crops is very promising. While the term “semi-automated” was used earlier to describe the methodology for the classification, this was actually a fully automated analysis because no manual editing was performed. However, for a complete image classification project, manual map editing would be a component in the process, hence the term “semi-automated.” The map accuracies achieved are very high for a fully automated analysis, and were this project to be carried out to the end by creating a classified map with class accuracies around the 85% individual level, only a relatively small amount of time would need to be used for manual editing to reach the accuracy specifications typical for an image classification project.

The classification accuracy results achieved for the 3 crop types analyzed using GeoSAR data were comparable to or better than results for similar crop types using MODIS (Doraiswamy, Stern and Akhmedov, 2007[5],)LANDSAT and SPOT (Viera, Mather, and Aplin, 2003[6]).

[5] Doraiswamy, Paul C., Stern, Alan J., and Akhmedov, Bakhyt. 2007. Crop Classification in U.S. Corn Belt Using MODIS Imagery. Presentation at the IGARSS 2007, Barcelona, Spain.

[6] Viera, Carlos; Mather, Paul; Aplin, Paul. 2003. Agricultural Crop Classification Using the Spectral-Temporal Response Surface. Anais XI SBSR, Belo Horizonte, Brazil, 2003.

Recommendations for future work mapping crops using GeoSAR data products include;

1. Collecting more field data and increasing the number of accuracy sites used for analysis.
2. Including more crop types in the analysis
3. Exploring the use of image segmentation software (such as Definiens Professional) to create new crop field polygons for classification. It was found that for this analysis, many USDA CLU polygons for Yazoo County were unusable because they encompass multiple habitats and crop types. Using such polygons in an object-oriented classification project can lead to a significant amount of error.
4. Exploring why the use of the GeoSAR Σ_0 product in the CART analysis introduced more error in the standing corn and cut corn subclasses.
5. Further analyze the CLU data to determine the reason for the low correlation between reported and classification results for the soybeans.

10 DISCUSSION AND CONCLUSIONS

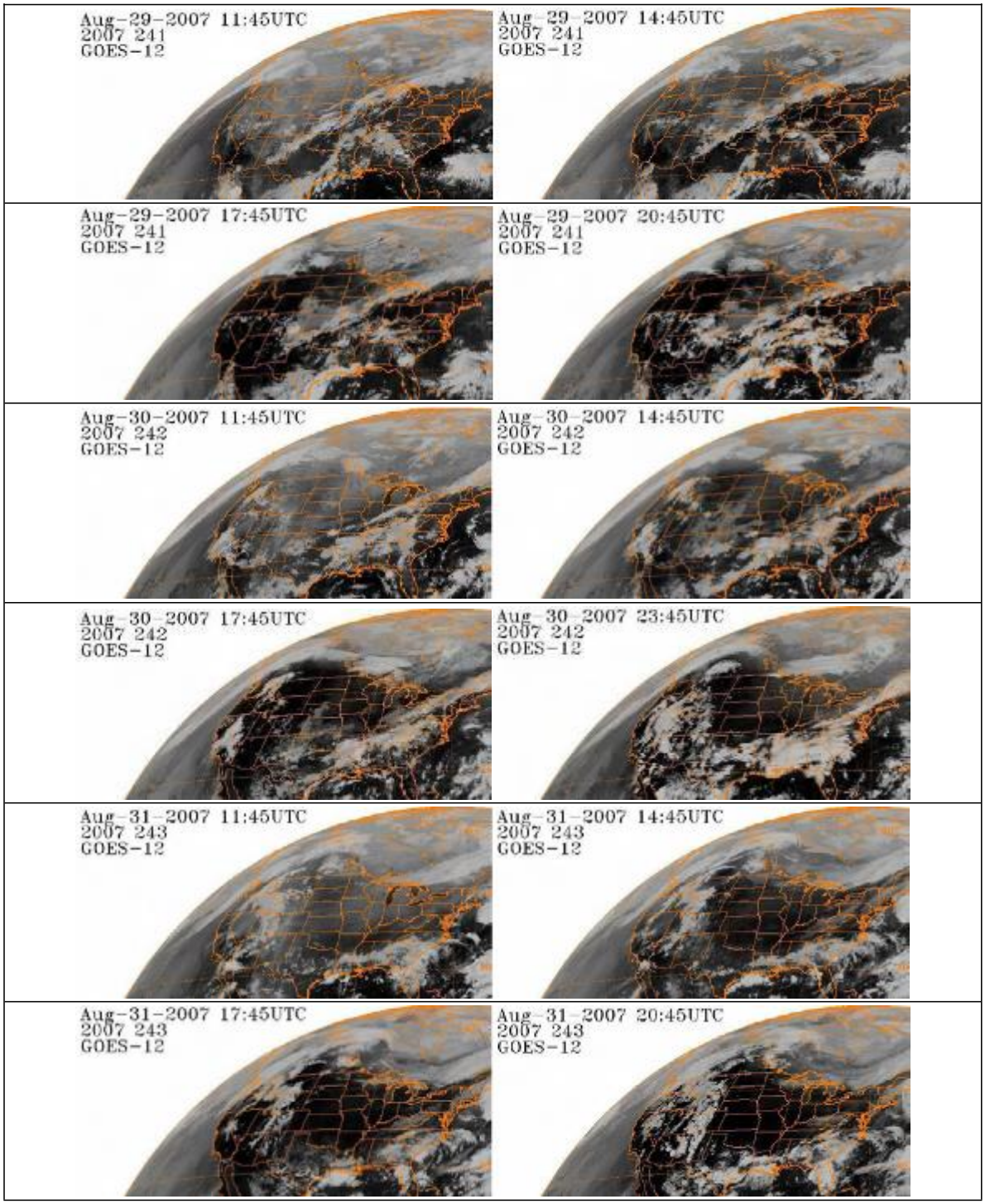
The results of this study demonstrated the utility of airborne IFSAR data collections for the augmentation of the NAIP imagery collection. It demonstrated that:

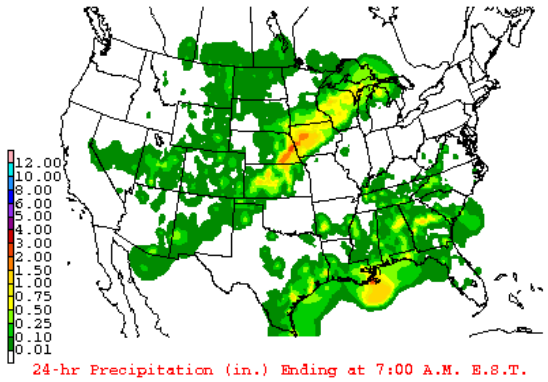
1. GeoSAR is fully capable of collecting data over a wide area in a short period of time;
2. GeoSAR can collect data at times and in weather conditions where a standard optical sensor would be unable to collect the data;
3. GeoSAR data can be used to produce precise digital orthophoto mosaics and digital elevation models;
4. The GeoSAR derived imagery products register well with USDA Common Land Unit (CLU) data;
5. The GeoSAR data is of sufficient quality to successfully extract CLU data vectors;
6. The GeoSAR data can be used to support fully automated crop classification at an accuracy equal to or superior to that of optical multispectral data.

This report is just the beginning of the analysis that can be made from this rich data set. Fugro EarthData Incorporated plans to continue analysis of this data to obtain maximum use from the data, and will be periodically publishing results.

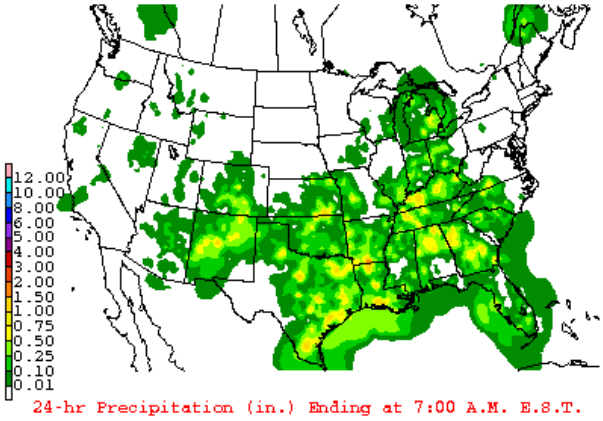
APPENDIX A - GEOSTATIONARY SATELLITE IMAGES AND PRECIPITATION MAPS

The images below were collected by the NOAA GOES-12 geostationary satellite during the period of GeoSAR data collection from August 29 to August 31, 2007. The images depict the typical cloud patterns over the Southeastern United States during the NAIP imaging season. The images also demonstrate that the GeoSAR system was collecting data in Yazoo County Mississippi while it was too cloudy to collect the data by standard optical means.

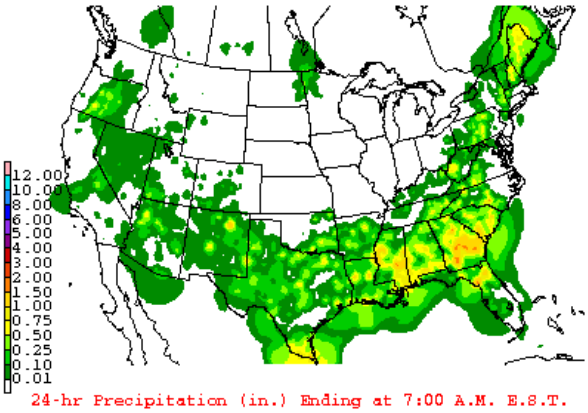




August 29, 2007



August 30, 2007



August 31, 2007

APPENDIX B EXAMPLES OF DATA FROM CROP FIELD REPORTS.

Latitude	Longitude	Crop	Date
32.89970000000	-90.44989000000	PG corn	8/27/2007
Row Direction	**Row Spacing (in.)**	**Plant Spacing (in.)**	**Crop Height**
265	38	4	11 (Ft.)
Stem Diameter (in.)	**Photo Range**		
1	2845-2848		

Latitude	Longitude	Crop	Date
32.89933000000	-90.44985000000	Corn Stubble	8/27/2007
Row Direction	**Row Spacing (in.)**	**Plant Spacing (in.)**	**Crop Height**
80	38	8	6-12 in.
Stem Diameter	**Photo Range**		
0	2849-2853		

Latitude	Longitude	Crop	Date
32.89878000000	-90.43888000000	PG soy	8/27/2007
Row Direction	**Row Spacing (in.)**	**Plant Spacing (in.)**	**Crop Height**
280	18	2	1-2 ft.
Stem Diameter	**Photo Range**		
0	2867-2870		

Latitude	Longitude	Crop	Date
32.74111100000	-90.25691700000	cotton	8/28/2007
Row Direction	**Row Spacing (in.)**	**Plant Spacing (in.)**	**Crop Height**
130	38	4-6	3 ft.
Stem Diameter	**Photo Range**		
0	2920-2930		

Latitude	Longitude	Crop	Date
32.74877800000	-90.50433300000	cotton-standing	8/26/2007
Row Direction	**Row Spacing (in.)**	**Plant Spacing (in.)**	**Crop Height**
85	38 double drill	8 double drill	3 ft.
Stem Diameter	**Photo Range**		
1	2821-2829		

Latitude	Longitude	Crop	Date
32.87308300000	-90.52025000000	rice-flooded	8/28/2007
Row Direction	**Row Spacing (in.)**	**Plant Spacing (in.)**	**Crop Height**
			3 ft.
Stem Diameter	**Photo Range**		
	2965-2980		

Latitude	Longitude	Crop	Date
32.76383300000	-90.52130600000	soy-row	8/26/2007
Row Direction	**Row Spacing (in.)**	**Plant Spacing (in.)**	**Crop Height**
180	38	4-6double drill	2 ft
Stem Diameter	**Photo Range**		
0	2837-2844		