Project Report

Farmland Monitoring in Mohall and Grano Region

ZHENNI JIANG

42383142

L2B GEOB 373

Abstract:

This project aim at evaluating the value of 2002 Landsat 7 ETM+ in monitoring the status of farmlands in terms of the distribution of crops, the richness of crops, and abundance of different types of crops in the north-western part of North Dakota. Additionally, observe and identify other objects within the research region that may display relations with or influence on the farmlands. After operating supervised and unsupervised classifications, output maps and graphs demonstrate that there are mainly 4 types of crops: A, B, C and D, among which C shows relative stable and high abundance. Some fields have two or more types of crops, and some field has nothing but ground. Also, all types of crops preferably to be planted near the water bodies so that sufficient irrigation can be provided for vegetation.

Introduction:

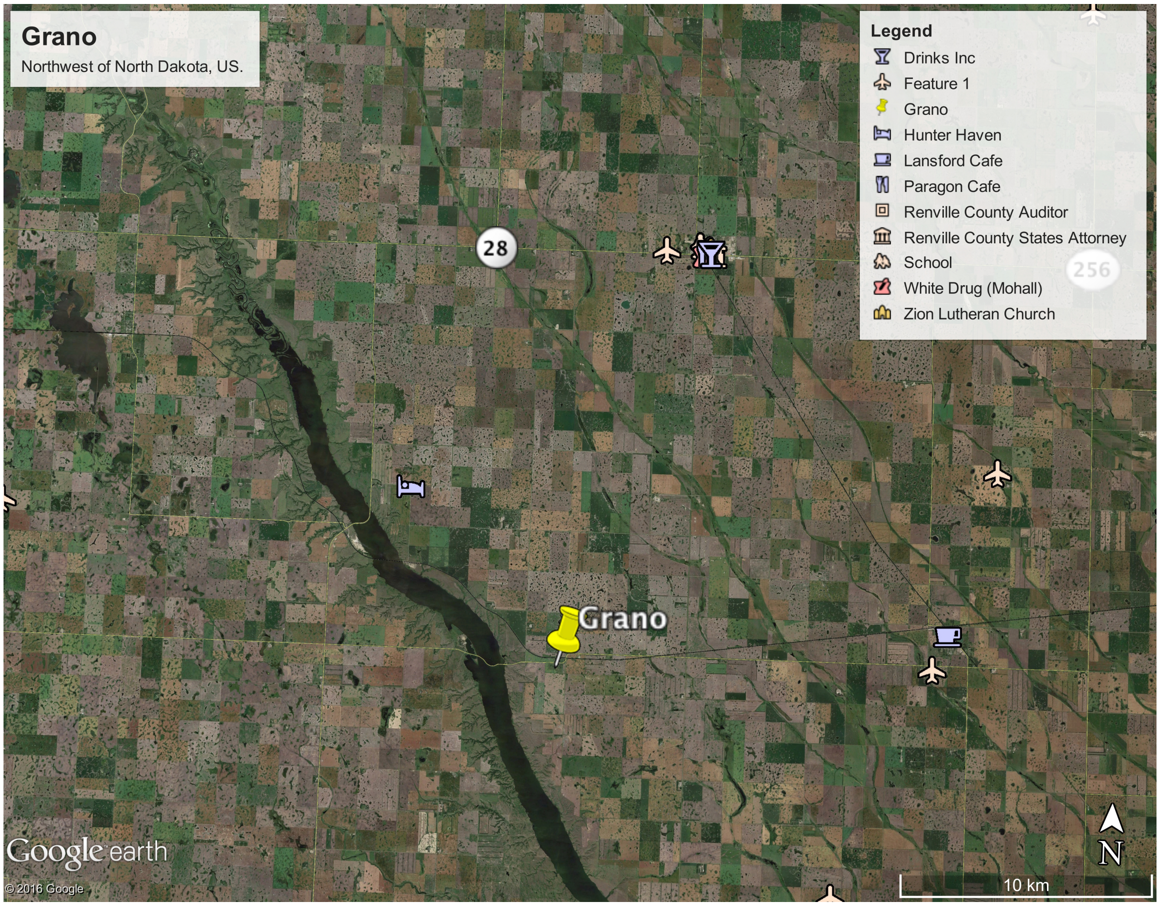

Figure 1. This map is retrieved from the Google Earth 19/8/2013 imagery.

The area covered by the image is located in the northwest of the North Dakota state with large scales of agricultural activities as we can observe this phenomenon from the image. In addition, there's a northwest-southeast water body, namely Lake Darling, appeared in the image with several other smaller water body features flowing in the same direction on the east next to it. With the assistance of Google Earth (as shown in Figure 1), the image delimited-region has several airports and a school in Mohall with some relaxation infrastructures such as cafes, restaurants and church. Additionally, there are some linear features that represent roads and railways that across the landscape and across the lake. There's also a wildlife protection refuge in the area to the west of the Lake Darling.

The raster imagery is obtained in 2002 by Landsat 7 ETM+ sensor with a resolution of 30 m. From the metadata, we can know that its referencing system is UTM-14N, the map unit is meter with a unit distance of 1. The whole project area is of 1,564.20 km square/156,420 hectares. The sensor that is used to obtain the imagery has the characteristics of eight-band, multispectral scanning radiometer, high-resolution, Instantaneous Field of View (IFOV) of 15 m for band-8, spectral range of 0.45-12.50 μm, temporal resolution of 16 days, and a swath of 183 km (NASA, 2016).

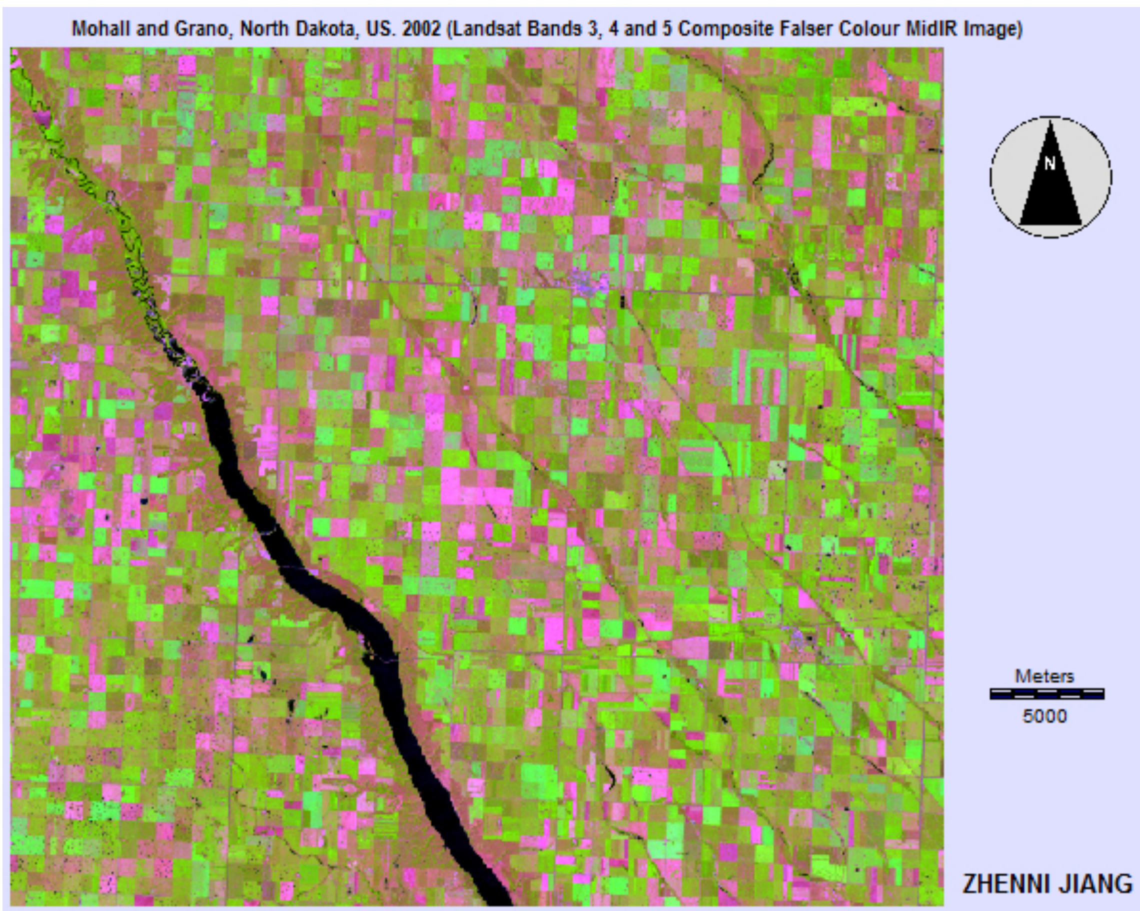

Figure 2. This map shows the Mohall and Grano, North Dakota, US region, the image is taken at 2002. The composite image is made by combining Landsat7 ETM+ Bands 3, 4 and 5.

In order to aid the feature identification, a colour-composite-image of bands 3, 4 and 5 is created. The resultant is a MidIR false colour view as showed in Figure 2. By assigning the band 3 for blue, band 4 for green, band 5 for red, and using linear stretch with saturation points, 24-bit composite with original values and stretched saturation points, 1% saturated rules, a band345-composite image is generated. The reason of using the MidIR image is that it is the most informative composite in this dataset due to the fact that the inclusion of two infrared bands (NIR & MIR) in the image sharpens and contrasts the image more than 234 NIR composite or 123 true colour composite, which provides a better presentation of and highlights the different vegetation types and moisture content according to its spectral reflectance and brightness in infrared section. Thereby, using the 345-band combination is more beneficial for identifying and distinguishing features.

Following, I will introduce the methods that I used and analyze the results that were derived from operating the IDRISI.

Analysis:

To identify the abundance and location of different crops as well as other land covers in the region, the supervised and unsupervised classification is implemented.

Unsupervised classification:

First, using all of the bands from 1 to 7 to run a CLUSTER function. Then, choosing the “fine generalization” while retaining all clusters to create the cluster image. Thereafter, producing a histogram based on the map. Due to the over hundreds of clusters, further classification is needed. Following the classification rule of the relevant number of classes, reclassify is applied (1,738,000*1%=17380) and it generates a histogram with 26 classes. Still, there are many irrelevant classes that are of the less importance for the classification. Therefore, according to the diagram pattern, the more appropriate rule is set up as 3.5% elimination (1,738,000*3.5%=608,300). Thereby 19 classes are left out of the map. Thus, only 7 classes left. Then run the CLUSTER again using the same condition but restrict the class to 7. The classes then are assigned the names of Crop A, B, C and D, water, bare ground and building based on true colour image and Google Earth assistant as the figure shown below.

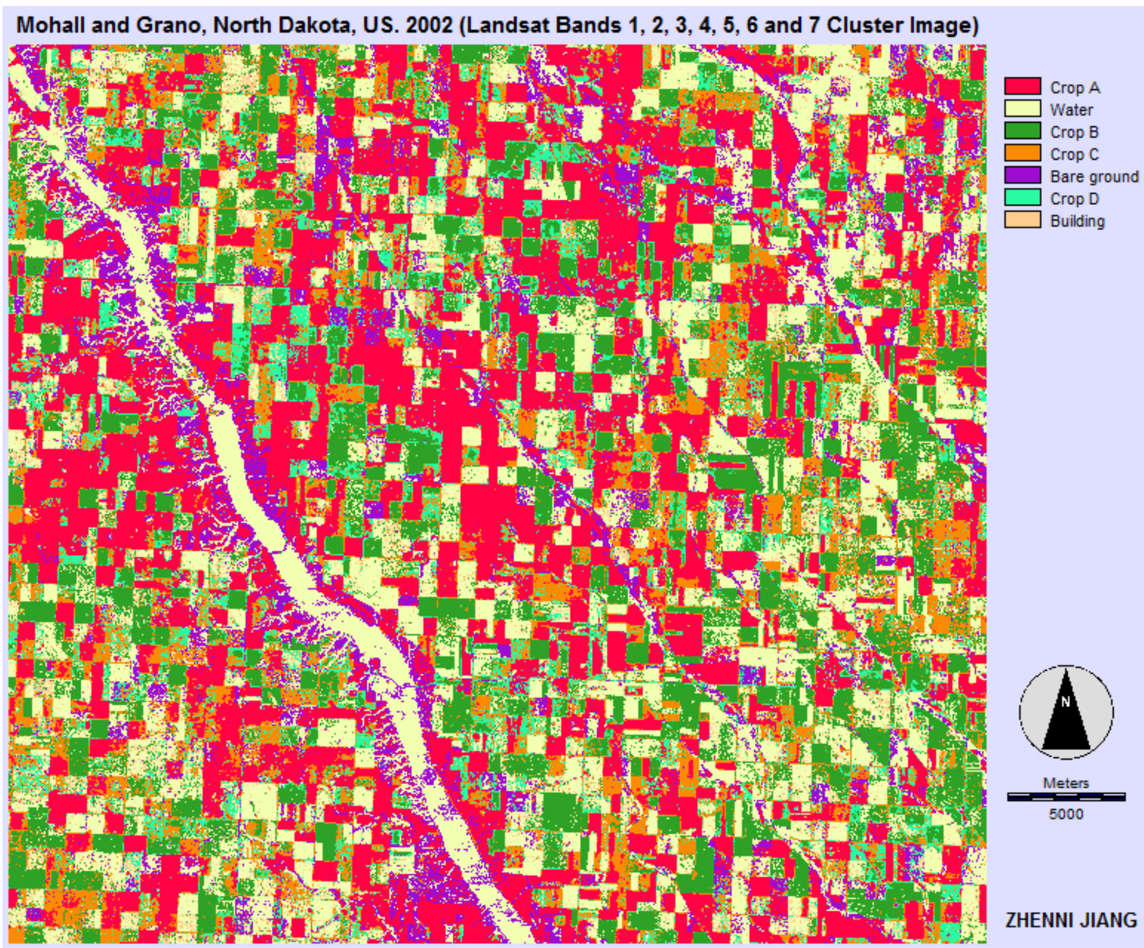

Figure 3. This map shows the re-clustered map based on the new classified numbers of category.

Another approach is through ISOCLUST. Running an ISOCLUST function, and then set the desired classes to 7.

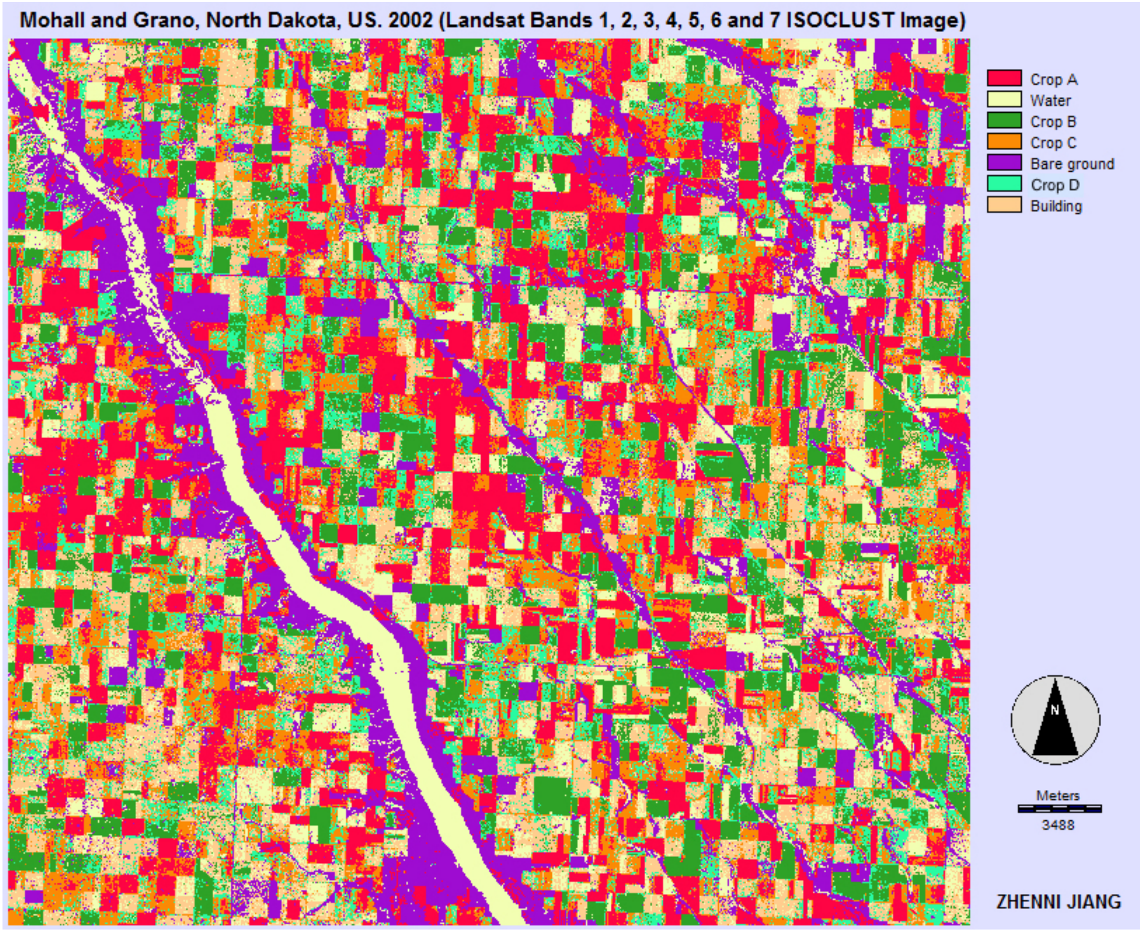

Figure 4. A map shows the ISOCLUST Analysis with 7 classes by using Landsat 7 ETM+ Bands 1-7.

The final product map derived from operating CLUSTER analysis visually shows that the Crop A occupies more areas than that are depicted in ISOCLUST image. Through operating the Area function in TerrSet, evidence of such justification is

CLUSTER Category	Hectares	Legend
1	44922.780000	Crop A
2	33627.330000	Water
3	25180.380000	Crop B
4	21270.060000	Crop C
5	13271.400000	Bare ground
6	12794.220000	Crop D
7	5353.830000	Building

Figure 5. This chart shows the CLUSTER analysis statistical result of each land cover area.

ISOCLUST Category	Hectares	Legend
1	30755.340000	Crop A
2	15519.780000	Water
3	19131.750000	Crop B
4	25450.110000	Crop C
5	24673.680000	Bare ground
6	17347.140000	Crop D
7	23542.200000	Building

Figure 6. This chart shows the ISOCLUST analysis statistical result of each land cover area.

provided by charts 5 and 6. In addition, the purple area is comparative obvious in ISOCLUST image, which represents Bare ground, statistically supported by the charts. Additionally, there are several straight lines presented as purple, which stand for roads, railways and bridges that contribute to the transportation convenience between regions and connect both sides of the lake. Also, the water in CLUSTER image appears some noises while ISOCLLUST image appears to be purer and more homogenous, yet more square-shape-field are categorized as water land cover, which may due to the large quantities of ponds spread over the landscape. Crop B shows higher occupancy in CLUSTER image while Crop C and D show over 4,000 ha higher distribution in ISOCLUST image. The hugest distinguish between the two analysis is Building features, which shows almost 5 times difference gap. Upon the additional information offered by the Google Earth, the ISOCLUST displays relative better classification than CLUSTER process in that it successfully delineate the large area of bare ground features around the lake and the other stream alike features as well as fewer noises. However, CLUSTER shows higher efficiency as well as less building features, which appears to be more appropriate.

Hinge on both images along with the Boolean images, we can notice that the distribution of Crop A has substantial distribution in northern part as well as neighbouring areas of the lake, Crop B mostly concentrates in the eastern part of the region, Crop C has quite discrete distribution yet a few parts appear as large clusters, and Crop D mainly concentrates near the lake. Overall, Crop A has the highest productions in this region with a generally square-field-shape production.

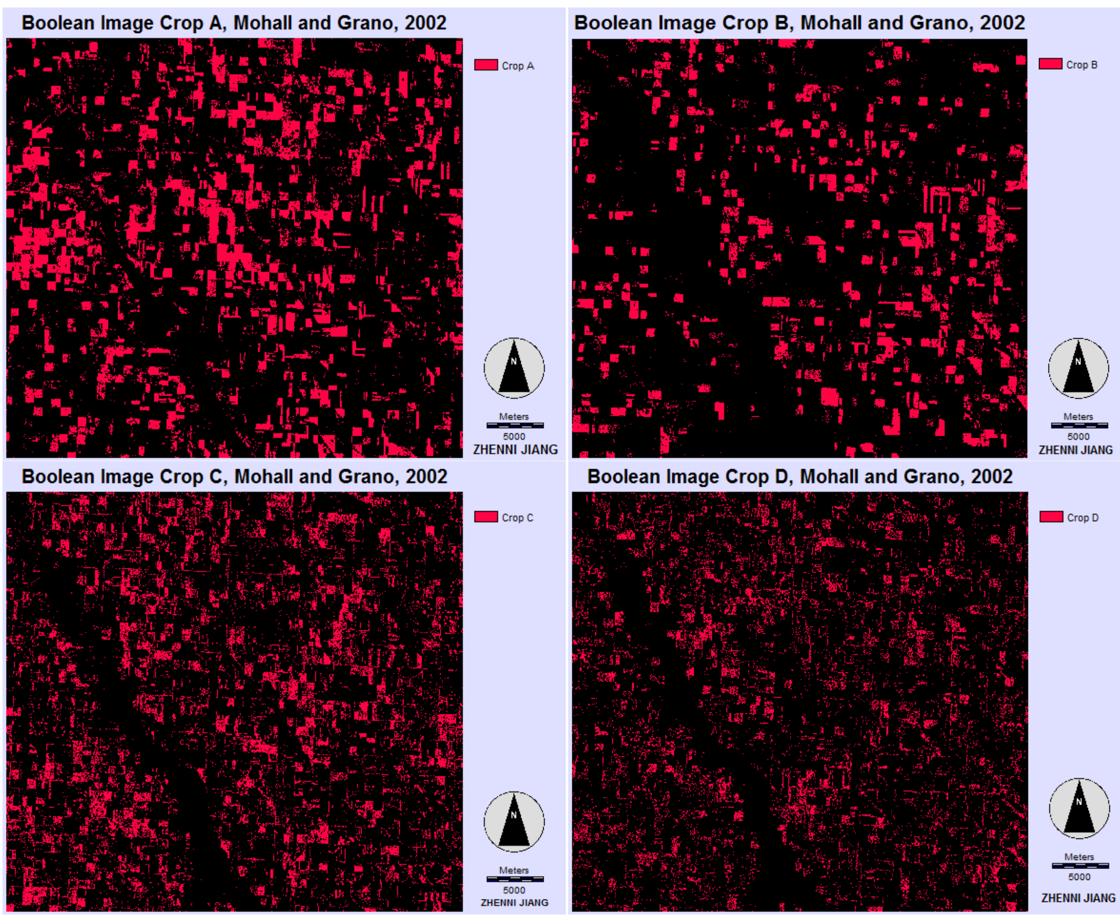

Figure 7, 8, 9 and 10. The images above show the distribution shape, area and location of the Crop A, B, C and D.

In conclusion, the unsupervised classification method has the general routine of "pixel clustering→classifying→assigning names to classes→validation". To some extent, it reduces the human error occurrence even without background knowledge about the area. However, it also has limitations, such as mismatches between clusters, decision-making difficulty of the right amount of classes, and limited control over results.

Supervised:

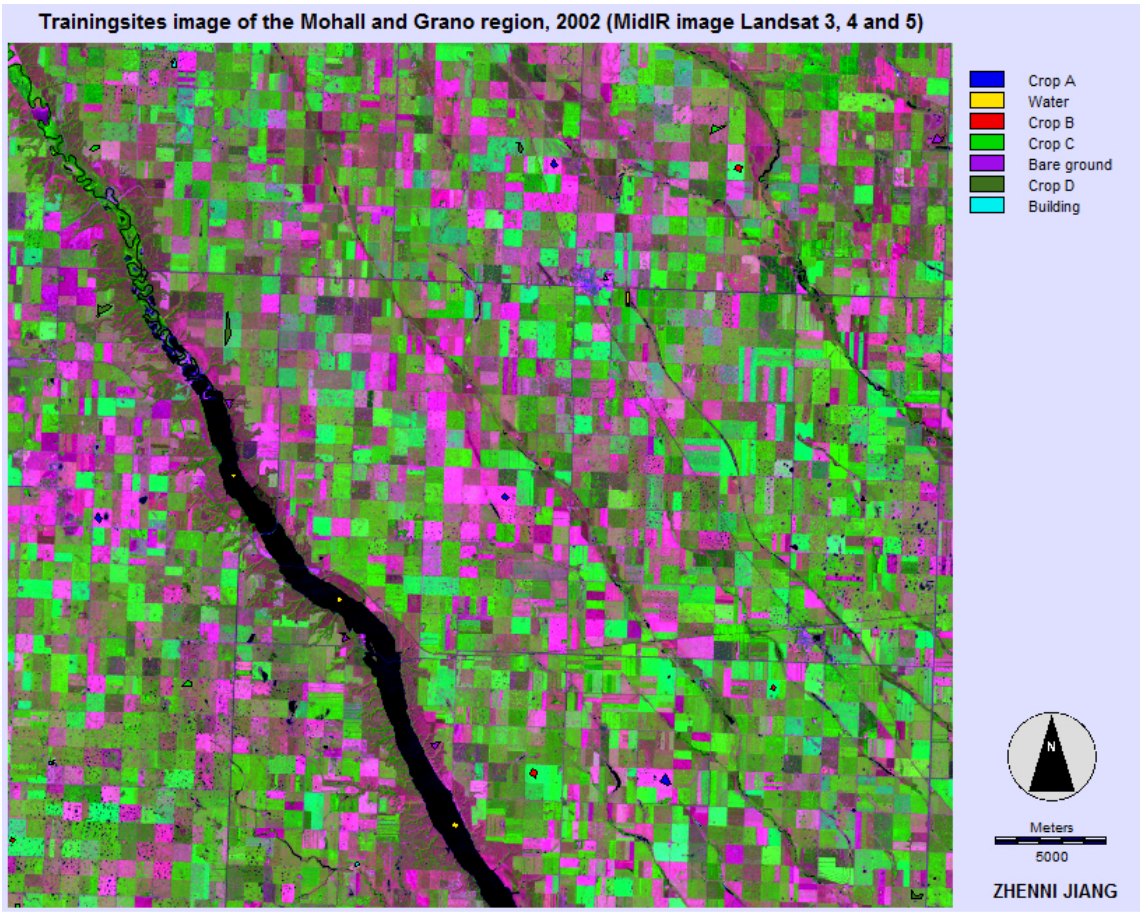

Figure 11. This is a training site map with different colours of polygons showing the different class training sites.

First, set 4 training sites for each class with a size around 8 ha per training site. Then create a signature graph by using the SIGCOMP showing the differences among these training sites. As the graph above shows, water has the lowest spectral signature in general with the peak at Band 6 (TIR) and trough at the Band 5 (MIR). The Crop A has the in average highest spectral signature but relative low at Band 4 (NIR) comparing to the other three crops. Crop B fluctuates the most that hold its crest at Band 4, which totally opposite with Crop A. Crop B also retain the lowest point at Band 7. Crop C displays a relatively smooth and mild signature pattern whereas Crop D reaches the

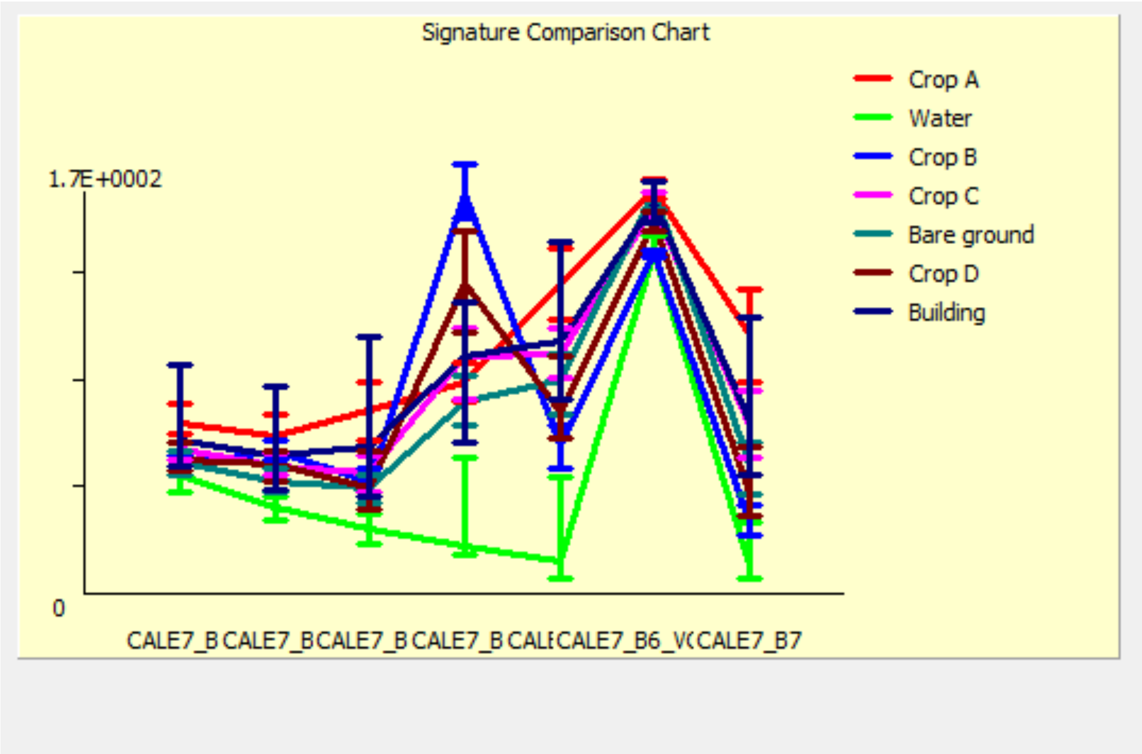

Figure 12. This is a signature graph showing the maximum, minimum and mean spectral value for each class.

lowest point for Band 3 (RED) among the four types of crop and gets its lowest at Band 7 (MIR). Nevertheless, all of these crops shows similar patterns compared to other features so that it's quite easy to separate them from other features especially water bodies. The minimum and maximum of the same land cover show general small gap, typically at the Band 6 has the smallest deviation from the mean value.

For the supervised classification, I choose the Maximum likelihood (MAXLIKE) hard classifier on account of its efficiency and accuracy while still evaluates both the variance and covariance of pixels from training sites. The resultant map as shown on the next page. It delineates lots of building class, which actually includes some other concrete features such as road, playground, concrete floor, and other infrastructures. Crop A shrinks to few blocks; black highlights the uncertain area that uses the 1% probability threshold. Crop C and Crop D shows a lot of discrete features, Crop B presents as blocks in the classified image. Water features mainly appear as the lake and several small areas in the northeast. Bare ground is still shown as purple that mainly encompasses water features.

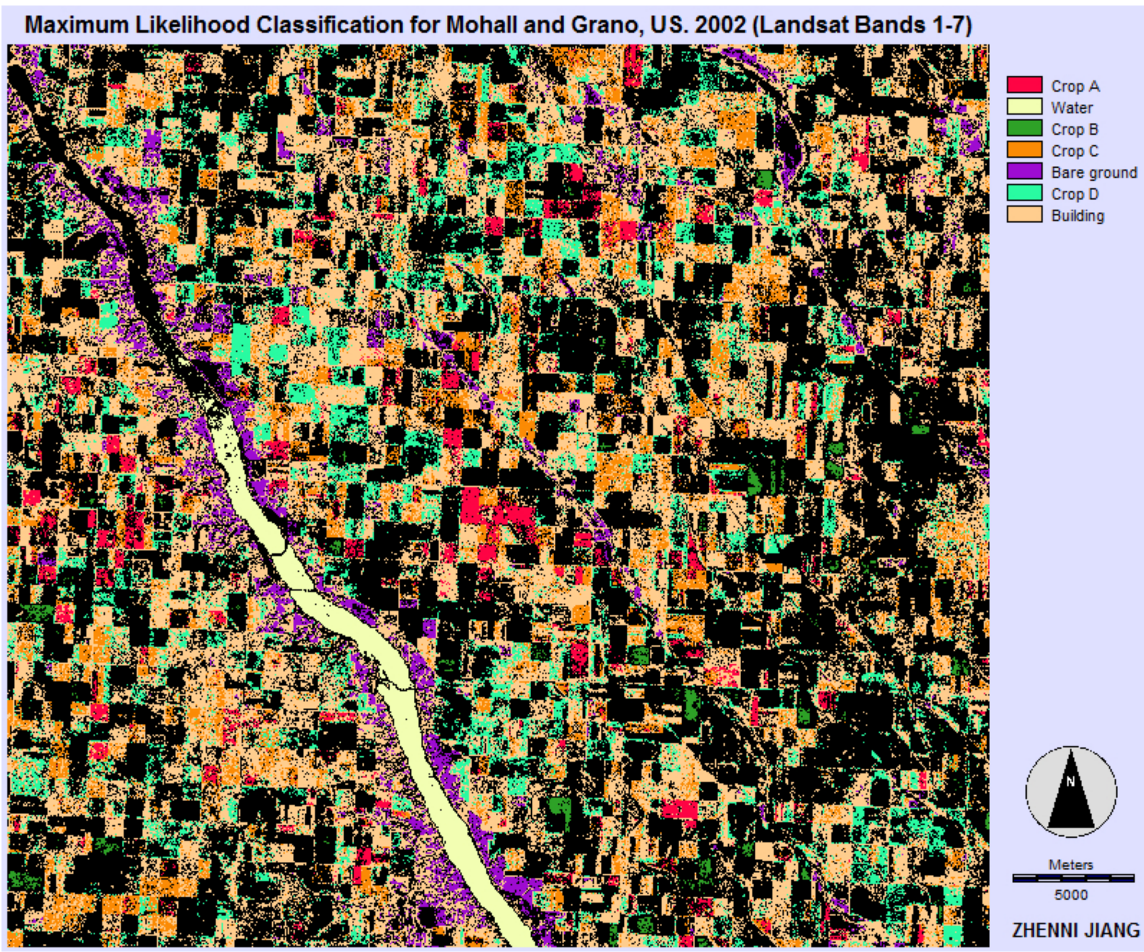

Figure 13. This map shows the area classified after MAXLIKE analysis.

The strengths of using the supervised classification are that analyst can choose the feature of its interest and select training sites based on observations and justifications. However, the weaknesses are that unique classes can be identified in other classes that increase the inaccuracy and uncertainty, also training sites selection procedure needs high level of accuracy and the training sites selected should possess high level of representative characteristic. Therefore, the more homogenous the training site is, the better result would be produced. Supervised classification would be really precise if the training sites are chosen appropriately. It also depends on the landscape appearance that may influence the efficiency and its difficulty degree.

Conclusion:

The ISOCLUST (unsupervised classification) is relative more successful than the CLUSTER. Because it well classified the Bare ground and water features whereas CLUSTER presents many more water features than the Google Earth shows. For crops, ISOCLUST has more homogenous field while CLUSTER presents too much noise. However, there are many disturbances appearing in the real world image, such as discarded land, plain floor and features that are hard to identify. Hence, the supervised image shows more a sincere answer to those unidentified areas (black). However, there are too many building features (too much inclusion in building features), which distract the view of crops and show too few Crop A. This may due to the mixture of crop A and Building features in one pixel that results in only showing Building rather than Crop A. Generally, Crop C has stable and abundant productions after analyzing all classifications. Crop A and B may both contain some identification problems that result in all three maps display dramatic differences. The Bare ground feature in the MAXLIKE image represents the best output among all the other classification methods – not too little or too much.

In the case of this project, I would suggest using the ISOCLUST unsupervised classification method to monitor the distribution and abundance of crops in the farmlands. However, for projects with sufficient budget and time, I would recommend using supervised classification to identify land cover and to produce an accuracy assessment to evaluate the result. To improve land cover identification, complex model should be built according to different regions and landscapes to increase the accuracy. In addition, accuracy assessment is necessary to validate the result.

Reference:

NASA. (2016). The Enhanced Thematic Mapper Plus. Retrieved from http://landsat.gsfc.nasa.gov/?p=3225