Assessing the errors generated from classification of remotely sensed data using spatial autocorrelation

Edward Park
Department of Geography and the Environment
University of Texas at Austin
geo.edpark@utexas.edu

Abstract

In this paper, a method of analyzing the pattern of error when classification was done from remotely sensed data by using spatial autocorrelation analysis will be introduced. Various sites were picked (water, tree, grass, sand, and urban region) and corresponding reference data were supplied for comparison after classification. Classified images were compared to the reference data to assign white color (0) to the pixels that agree and grey to black color (1, 2, 3, 4; depending on the degree of disagreement) to the pixels that disagree. Thus black and white images (difference image) were produced and spatial autocorrelation was performed within grey and black pixels in difference images. Several methods of classification were applied including maximum likelihood, ISODATA and minimum distance to find out the most suitable classification after measuring spatial autocorrelations of difference images.

- *Some of the important **keywords** are in **bold** case.*

1. Introduction

Remote sensing might be an area that has been developed fastest along with advancement of other scientific technologies. Constant improvements of 4 dimensions of resolution (spatial, spectral, temporal, radiometric resolution) and quality of product increased the accuracy of measurement as well as convenience of using remote sensing data by making other ancillary data unnecessary which were indispensable in the past. Thus, remote sensing application field is stretching out quickly; however, there is an intrinsic problem that could not be solved even field of remote sensing progresses in a great scale which is also an inherent limitation that raster has. Every image is a grid based raster format so, pattern of error is exhibited which is different from the real and this could not be solved fundamentally no matter how much high the resolutions are. Since remote sensor calculates and defines brightness value (BV) of each pixel by weighted mean, '**smoothed**' representation is shown in remotely sensed imagery whether it is continuous landscape or discrete object. To overcome this phenomenon, various trials are being made and giving an edge effect by filtering of pixels is one of the methods used most commonly.[1] Another

[1] Spatial frequency in remotely sensed imagery may be enhanced or subdued by low-pass filtering or high-pass

commonly used method is a sub-pixel unmixing (spectral mixture analysis) which divides each proportion for various features mixed in a pixel. However there are some cumbersomeness in this method that analyst should have knowledge about the corresponding region, every feature and endmember characteristic in order to do this. Also through this method, it is possible to figure out each feature's proportion within a pixel; however, analyst still has to make an educational guess for the location of features[2]. Finally pixel based analysis like this sub-pixel unmixing is quite limited to small region which is most of the time not a proper method for classification of land cover type. Thus efforts to get over these limitations of resolutions should focus more on **assessing accuracy** along with various image enhancements.[3]

Smoothing effect of spatially varying landscapes appears on image and this effect creates the **mutuality** (dependency) between adjacent pixels. This is called in another word spatial autocorrelation and this mutuality between pixels is clearly exhibited at the boundary of classes when compared to real shaped feature that is delineated correctly. Spatial autocorrelation is the tendency for like things to occur near one another in geographic space (Slocum et al., 2009). The concept of spatial autocorrelation could sound general, conceptual and subjective. Because visual interpretation could be different from person to person and pattern discovered could not be a significant pattern across region only to be a chance factor in some cases. Thus there have been strong motivations for producing objective measure of spatially autocorrelated pattern.

In this study, SAC will be applied to see how accuracy of land cover type identification changes when various classification methods were used over a certain area. Congalton (1988) mentioned that the aspect of remote sensing most affected by spatial autocorrelation is in the analysis of classification error. Land cover type identification using remotely sensed image has been a major practice in this field therefore, there has been a lot of efforts finding out proper algorithm to improve the accuracy of land cover classification (Lu, et al., 2003). Starting from Campbell (1981), endeavors to measure the error of spatial pattern from remotely sensed data have been made. Verbyla et al. (1995) performed the accuracy assessment after classification using reference grid, and Henebry (1993) and Bruzzone (2000) have done analysis on multi-temporal change detection using spatial dependency while Chou (1991) and Qi et al. (1996) studied the changing pattern of spatial autocorrelation depending on spatial resolution. As getting toward recently, there were more researches applying spatial autocorrelation to see the changing pattern in certain regions in the world than developing methodologies. Elobaid et al. (2006) analyzed the changing pattern of tree's diameter in Malaysian forest using spatial autocorrelation and Lu et al. (2003) compared the results of changing pattern of spatial autocorrelation depending on various classification methods in Amazon basin. Other than these, there have been numerous studies applying the measure of spatial dependency to land cover change or classification whether directly or indirectly. However, this is the first time to compare the results of various classification methods to the hand digitized reference image and apply spatial autocorrelation to the quantified degree of difference.[4] I set my study area where there are both continuous fields and discrete objects to extend the applicable realm compared to Congalton (1988) who relatively focused on

---

filtering respectively.

[2] Especially in my research looking at the spatial distribution of error pattern induced by resolution, location of feature is also important as its proportion within pixel.

[3] If it is impossible to eradicate the error anyway, the best way is to understand the error completely and know the limitations of it.

[4] Congalton (1988) only used maximum likelihood method which is used most commonly and binary decision (0 or 1) which only tells agreement or disagreement.

continuous landscape only. Through this study, it is expected to find out the proper classifier in multi-dimensions of land cover type.

2. METHOD

2.1 Description of Study Area and Data Sources

Pattern of spatial autocorrelation was applied to the Cypress Creek Arm region of Lake Travis in Austin, Texas. There was a malfunction of scan line correction in ETM+; however, as seen in Figure 3 the study area is at the satellite path along nadir which is not affected by the malfunction. I chose the datum of NAD 1983 and Universal Transverse Mercator (UTM) zone 14 for the projection which the standard line exactly goes through the study region providing the highest accuracy of representation.

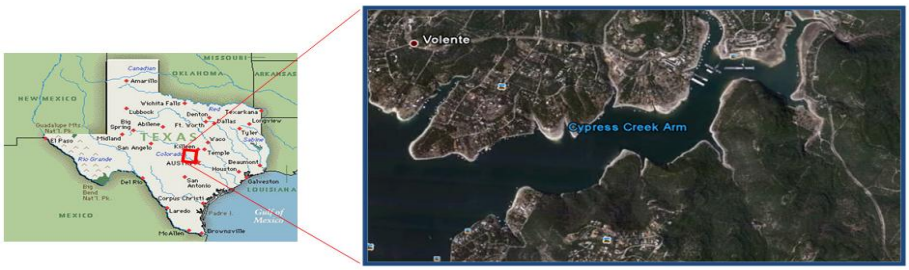


Figure 1. Study Area: Cypress Creek Arm from Google Earth **(**approximately 2km*1.5km). Extent: 30.433 (N), 30.415 (S), -97.877 (E), -97.906 (W).

Table 1. Imagery Data Sources

	Image Date	Scene ID	Path	Row	Spatial Resolution (m)	Cloud (%)	Bands Used
Satellite Image	Feb-1. 2009	LE70270392009032EDC 00	27	39	30	10	RGB, NIR
Aerial Photo	Feb-10. 2009	*TOP0809_50cm_3097_3 3_4_CIR_10032009*			0.5	0	GB, NIR
	Feb-10. 2009	*TOP0809_50cm_3097_3 3_4_NC_1032009*			0.5	0	GB, NIR
	Feb-3. 2009	*MANSFIELD_DAM-SEB1~4 (mrsid)*			0.15	0	RGB

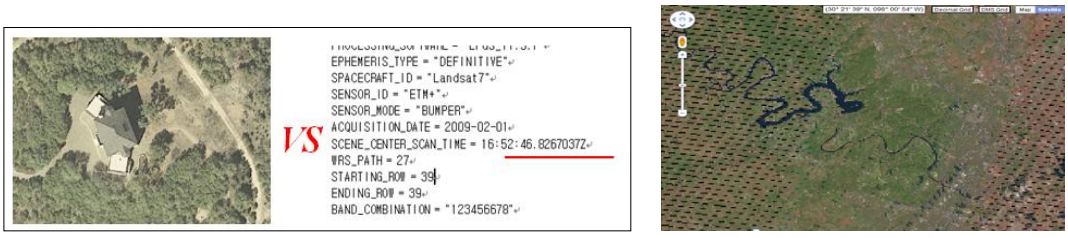


Figure 2 (left). Aerial photo seems to be taken in the afternoon around 3 (left) while the Landsat image (right screen shot) was taken at 4:52 pm (underlined red).

Figure 3 (right). My study area lies under the nadir of whisk broom sensor of Landsat 7 which is not affected by the malfunction of scan line correction. I have learned that scan line moves toward east as going from band 1 to 8.

2.2 Data Preprocessing

After deciding which data to use and proving relevance of those data, data preprocessing was started in earnest by digitizing. Since already made land cover products usually does not match the date with Landsat imageries and even I found some relevant land cover products that is in a close date with Landsat imageries, defined classes were not relevant for this type of analysis. In this case, I considered reclassifying them; however, there were difficulties finding proper reference data. Therefore I decided to make my own reference data. Digitizing was done in ArcGIS and reference vector image was rasterized into 10 meter grid cells using maximum area condition because it is thought to be the most appropriate for land cover mixed with continuous field and discrete objects. I maintained 1:300 scales when digitizing to apply same precision across the image. The threshold for digitizing was 7.5 meter and this is an arbitrary number I set because the BV from remote sensor is calculated based weighted mean and my assumption was that it would be classified correctly when at least it has width and height longer than 7.5 meter especially for urban structures which was the trickiest part. As hand digitizing takes a lot of time, I could not set larger region than this at this time. However, for my future research I will be applying this method to a wider and variety range of area.

TBC image was resampled into 10 meter using bilinear resampling method. The reason for choosing bilinear method is because it seemed to be the one most appropriate. Since there was no need to maintain pixel value as before classification, nearest neighbor method was not needed. Cubic convolution smoothes quite much and diminishes the contrast between objects which is necessary for classification so it is not considered to be proper as preprocessing before classifying. For these reasons it bilinear resampling seemed to the most proper method in this type of landscape where discrete object and continuous field coexist. After unifying the resolution for every image, I clipped out the study area using ERDAS Imagine AOI (area of interest), so that every image have 279x198 pixels.

2.3 Classification

Classification starts by defining each class scheme. In this study (for both reference and TBC images) classes were defined relatively generally (water, tree, grass, urban (house, paved road), sand) and these classes are assigned from 1 to 5 based on the order of spectral values (data type: short integer): water = 1, tree = 2, grass = 3, urban = 4 and sand = 5.

Classification is broadly divided into supervised and unsupervised classification. Supervised classification is used with other references or field data when analyst knows the identities and location of classes. Then sample training sites from TBC images are picked to form a character signature to be used to classify pixels. To maximize the differentiation between features, combination of visible and near infrared band was used. There are hundreds of classification methods (decision rule) those are being used; however in this study, only several most commonly used types of classification offered by ERDAS Imagine and ArcGIS will be evaluated. In unsupervised classification, computer automatically creates signature of feature based on spectral value and classify pixels (usually ISODATA clustering), therefore analyst only need to define number of classes. Admittedly, we can guess that there should be a certain type of classifer suitable for particular land cover which will be more precise than others and this kind of study was first performed by Congalton (1988) and he proved the change of classification precision depending on the land cover type. However, Congalton only used the maximum likelihood[5] and did not compare between classification methods.

[5] There must have been limitations at that time to do most of the works manually.

Table 2. Classification Methods

1	Parametric Maximum Likelihood (MLH)	6	Parallelepiped Minimum Distance (MND)
2	Parametric Mahalonobis (MHN)	7	Feature Space Maximum Likelihood (MLH)
3	Parametric Minimum Distance (MND)	8	Feature Space Minimum Distance (MND)
4	Parallelepiped Maximum Likelihood (MLH)	9	ISODATA 5 classes
5	Parallelepiped Mahalonobis (MHN)	10	Mixed

In this study, 8 supervised and 2 unsupervised classifications methods were used and these are organized in table 2. Three parametric classifications were performed using different decision rules (Maximum likelihood, Mahalonobis, Minimum distance) which assume normal distribution of each class. One of the Non-parametric classifications I used was the classifications using parallelepiped decision rule. The reason of using parallelepiped decision rule is that it is good for broad classification since the separabilities between classes were very high enough.[6] Another non parametric classification rule was using feature space which normally gives a great insight of pixel distribution across image. However, feature space using multiple bands (more than 3, 4 bands in this case: RGB and NIR) could not be represented on screen. Choosing correct training sites greatly affects the result of supervised classification. Hence it is important to confirm that the average brightness value of each training sites actually agree with the corresponding class. For this, I spatially linked reference and TBA images in ERDAS to choose training sites by visually inspecting.

However there are fine references, it is always possible to miss something when analyst only sees the sites through photographs. 5 classes are fairly big and general divisions and even though they look distinct from each other some features are not clear just by looking at the photograph. Classes of water[7], tree, urban and sand are quite clear but a lot of time indistinct feature was grass. Grass is normally differentiated with urban which is upper adjacent class most of the time; however, sometimes confused with adjacent lower class tree depending on its type and occasionally looks similar to sand especially when it is wet. Therefore I did a brief field trip to the most accessible couple of grass sites to see and confirm that there actually is grass field. After confirming the correspondence of grass between reference image and actual field, I compared 11 other training sites I picked by visual inspection for grass with the actually visited sample sites in figure 4 which is located at 30°24'55.26"N 97 °53'09.85"W. For this comparison I first calculated the average of correlation coefficients of RGB values between the sample site and other 11 sites. The correlation coefficients ranged from 0.56296 to 0.98411 and average of them was 0.80019 which is quite high enough to assume that the changing pattern of each of RGB based on the sample training site is similar. The average value of RGB between the sample site and other training sites is compared to see how much their degree of spectral range agree to each other. Average RGB value for the sample site was 895 while the counterpart of other training sites was 700.21 which is quite lower than the sample one. Since the lower RGB average will set the classification standard darker than actual, my concern was that if there would be any overlap with its lower adjacent class (decreasing the separability between the tree class). So I looked at the average values for tree and it turned out to be 420 which is still greatly

Signature Name		1	2	3	4	5
Water Training Sites	1	0	552.987	638.343	748.287	1032.68
Tree Training Sites	2	552.987	0	189.327	589.33	1233.16
Grass Training Sites	3	638.343	189.327	0	5.40997	20.1096
Urban Traning Sites	4	748.287	589.33	5.40997	0	5.2317
Sand Traning Sites	5	1032.68	1233.16	20.1096	5.2317	0

6 separability of cell array from ERDAS Imagine.

7 Sometimes I got confused between shade and water however, it is easy to figure out when zoomed out from 300:1 scale and use backup reference image for comparison.

separable from grass class[8]. Through this process the definition of grass class which was a little bit confused initially became clear and accurate.

Figure 4. Typical grass site (location: 30°24'55.26"N 97 °53'09.85"W, mid-lower part of image).

Figure 5. From left: water, tree, urban.

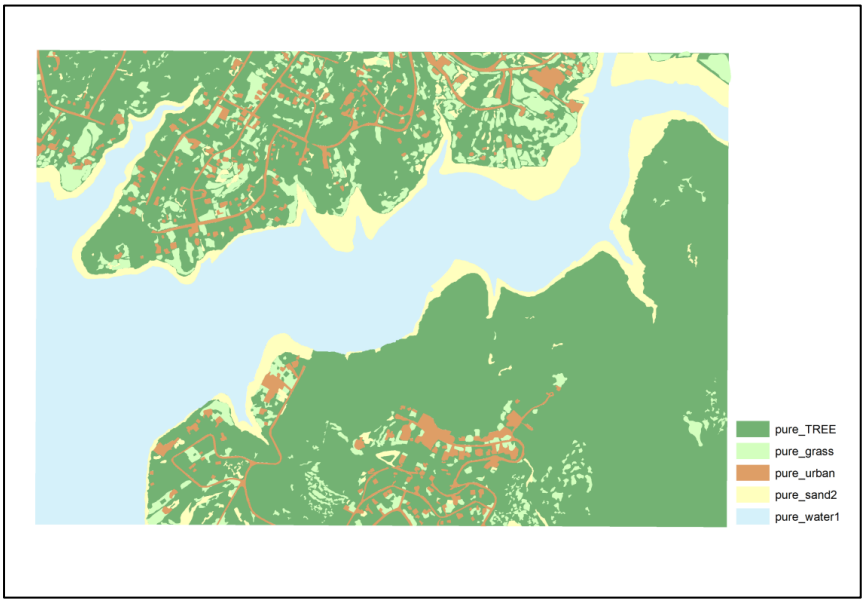


Figure 6. Final reference map after classified.

[8] Only looking at the average doesn't justify the high separability; however, along with high correlation between every training sites makes it relevant to assume as correctly defined classes.

	Signature Name	Color	Red	Green	Blue	Value	Order	Count
▶	Water Training Sites		0.156	0.312	0.413	13	13	355
	Tree Training Sites		0.550	0.368	0.342	14	26	409
	Grass Training Sites		0.766	0.757	0.688	15	39	228
	Urban Traning Sites		0.800	0.973	0.926	9	48	101
	Sand Traning Sites		0.892	1.000	1.000	17	61	246

Figure 7. Training sites for supervised classification from ERDAS Imagine signature editor.

ISODATA (Iterative Self-Organizing Data Analysis) clustering was used for unsupervised classification which doesn't require training sets and automatically group pixels based on spectral similarity. I did two classifications using ISODATA clustering: first one was producing 5 classes and assigning class values (1~5) to each of the corresponding classes and for the second one, it created 30 classes automatically and I reclassified them into 5 classes and used maximum likelihood method for land cover classification. For the later one, class signature was created by automatic computer process and I determined the classification method which is not normal. Figure 8 below is TBC images after classification.

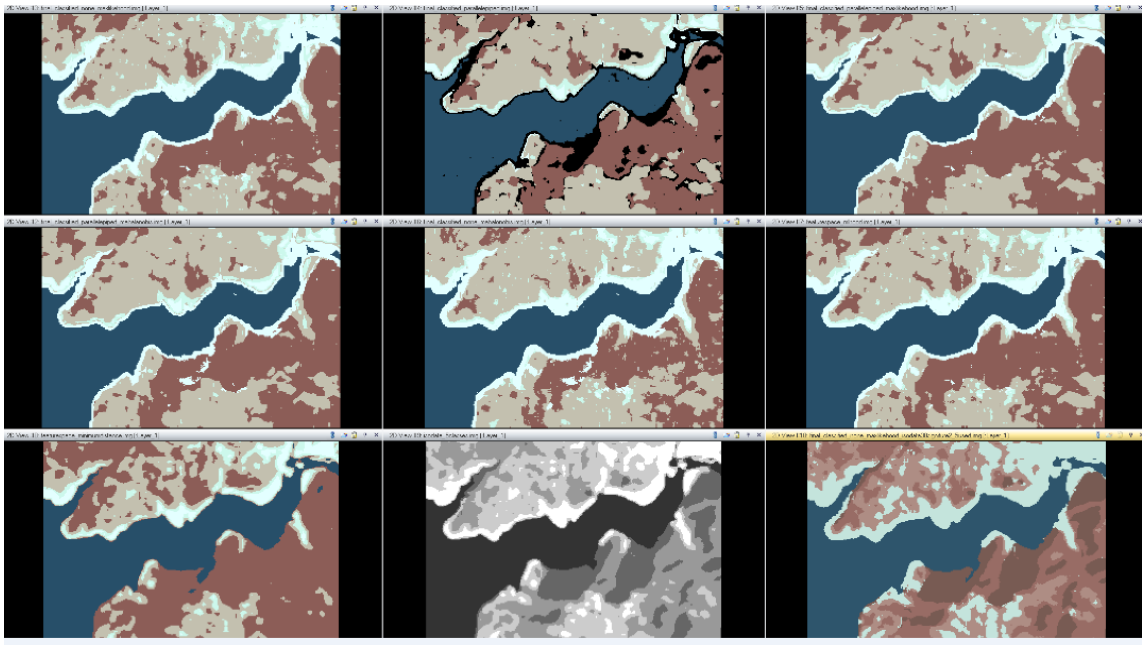


Figure 8. from upper-left clockwise: none-parametric MLH, parallelepiped MND, parallelepiped MLH, feature space MLH, mixed, ISODATA 5 classes, feature space MND, parallelepiped MHN, none-parametric MHN (center). Only 9 maps were included since I could not find any good layout to put 10 maps in a rectangular frame.

2.4 Difference Image

After creating 1 to 5 scaled reference image and TBC images, next step was to make difference image which is the last part of data preprocessing before actually starting the analysis. Difference image is a raster image that is made of the **difference value** between reference image and TBC images for each grid. Table 3 below shows the possible values of difference image. Different values of 0, 1, 2, 3, and 4 depending on the degree of disagreement is assigned to each of the pixels in difference image in a white (0) to grey color scale (1, 2, 3, 4). Color scheme in table 3 is assigned to the corresponding pixel. For example, if C_2 (tree) is classified as C_4 (urban) then 2 (grey 40%) will be assigned to the corresponding pixel in difference image. This is performed using raster calculator and all the negative

numbers are converted to absolute numbers.[9] Figure 9 is difference images created.

Table 3. Difference Image Table

ETM+ / Reference	C_1	C_2	C_3	C_4	C_5
C_1	0	1	2	3	4
C_2	1	0	1	2	3
C_3	2	1	0	1	2
C_4	3	2	1	0	1
C_5	4	3	2	1	0

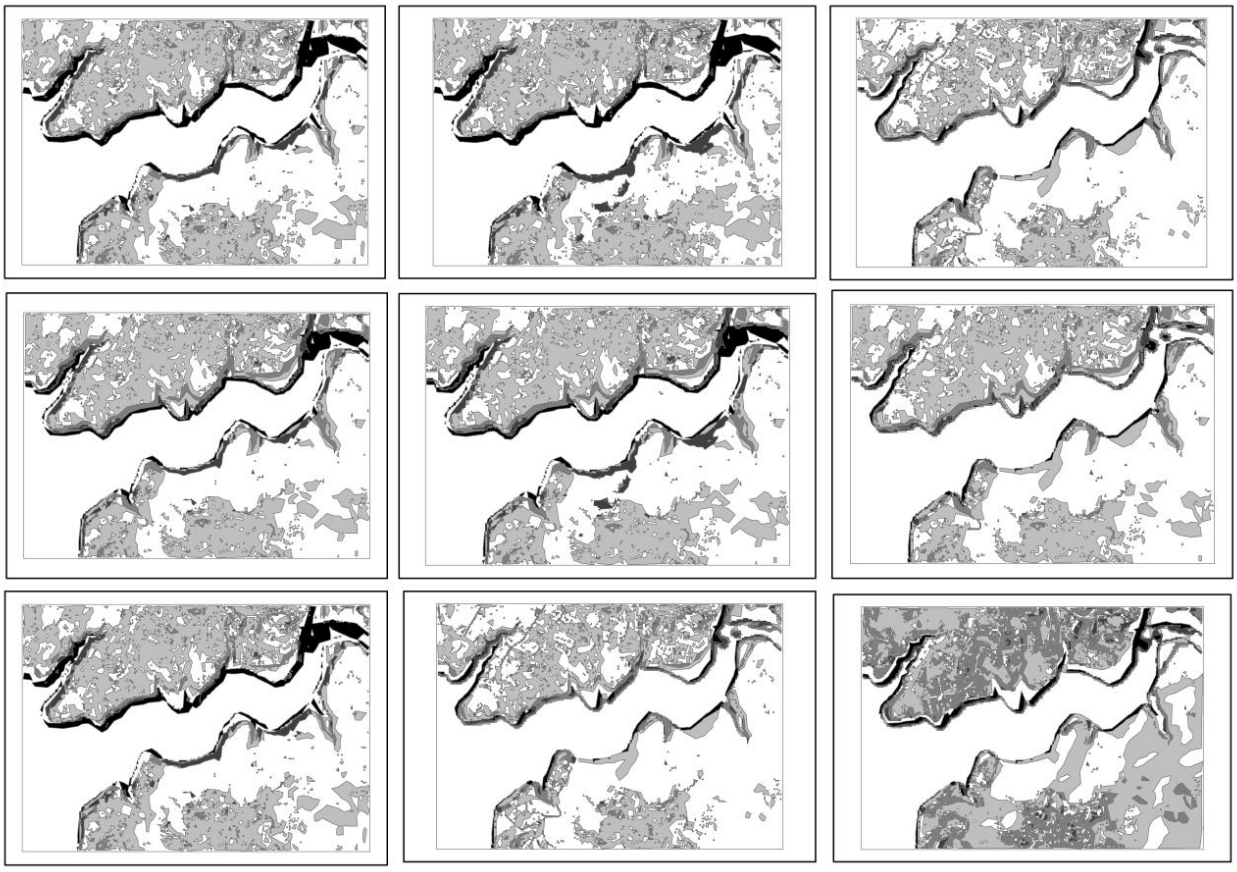


Figure 9. Difference images. From upper-left clockwise: none-parametric MLH, parallelepiped MND, parallelepiped MLH, feature space MLH, mixed, ISODATA 5 classes, feature space MND, parallelepiped MHN, none-parametric MHN (center).

[9] Since raw values of difference image ranged from -4 to 4, I transformed -1 to 1, -2 to 2 and so on.

2.5 Spatial Analysis

Spatial autocorrelation is the 1st rule of geography that 'everything is related to everything else, but near things are more related than distant things' said by Waldo Tobler in 1970. As the name describes it is the measure of how variable correlates with itself through the space therefore if there is any systematic pattern in spatial distribution between variables, it is considered to be spatially autocorrelated (whether positively or negatively). This is what makes geography meaningful that if everything is distributed in a random pattern, there is no need to do any kinds of spatial analysis. As I mentioned in the beginning of this paper, there were strong necessities to measure this conceptual pattern numerically and the most commonly used indexes which are **Moran's I, General G and LISA** (Local Indicator of Spatial Association)[10].

Moran's I is basically a measure of distribution of density that takes in account of both feature's location and value at the same time. Using these two variables this index evaluates whether the pattern of density distribution is clustered, dispersed, or random. Positive Moran's I index indicates the positive spatial autocorrelation while negative means negative spatial autocorrelation. The Moran's I index ranges from -1 to 1 where it means dispersion as the value of index gets close to -1 while it mean clustering toward 1. Value of 0 indicates the random distribution. Moran's I function in ArcGIS also calculates the z-score and p-value to see the significance of the index. Analyst could set an arbitrary significant threshold and decide whether the hypothesis is acceptable or rejectable (ArcGIS Resource Center). Moran's I function was performed for every difference image and the index, z-score (also p-value) and distance were calculated (Table 5). There are many options and parameters to set in performing this function such as conceptualization of spatial relationship. Fixed distance, zone of indifference and polygon contiguity were three most relevant options in this case and I chose the zone of indifference to be used. Polygon contiguity makes sense that takes every adjacent polygon into account; however, it failed to differentiate well enough between features. Fixed distance performs the calculation only for features within certain distance but not for features outside the distance; while zone of indifference also calculates the features outside the certain distance based on inverse distance weights. Since the zone of indifference still counts the features outside the distance, though not much but it looked to be the most proper method of conceptualization especially when I did simulation using same difference image only changing this condition (Figure 11). Also it was possible to determine the existence of geographical effects among features to support the choice of zone of indifference. For example sand (C_5) is naturally distributed alongside the coastline (C_1) which means they are physically adjacent and characteristically dependent to each other. As seen in the figure 10, the magnified diagram of coastline, difference feature (value from 1 to 4) is created fiducially from standard value of 0 induced from the necessity of being adjacent; it is relevant to tell that there is a potential effect to each other.

Next step was to see whether the local spatial autocorrelation exists in difference image by using LISA. Unlike the global Moran's I which measures the spatial autocorrelation across the whole image, LISA calculates spatial autocorrelation index as well as z-score and p-value for every polygon in an image with the cluster and outlier type. The cluster and outlier type (COType in attribute table of newly created output feature class) assigns HH for cluster of high values, LL for cluster of low values, HL for the outliers which is a high value surrounded by low values and LH for outliers which is the a low value surrounded by high values. This new output feature class gives good visual interpretation of the density of

[10] For the calculation of indexes, refer to the ArcGIS Resource Center (http://help.arcgis.com/en/arcgisdesktop/10.0/).

distribution. Figure 11 shows the results of LISA performed for every difference image to see the local spatial autocorrelation pattern. I maintained the same condition for conceptualization of spatial relationship as Moran's I.

There is another method of measuring the spatial clustering of features based on its location and value which is using the General G statistic that calculates the concentrations of high or low values over the study area. Same as Moran's I this General G index gives information about whether the values are clustered or not and particular distance between each other affects the result in a great scale. Likewise, since z-score is an important indicator of evaluating the relevance of the General G index and this varies depending on the distance, it is the most significant job to find the highest z-score to reject the null hypothesis that is assuming random distribution. Because the highest clustering occurs at the highest z-score. Zone of indifference was used for the conceptualization of spatial relationship by the same reason as Moran's I, and after finding the initial distance band, General G tool is ArcGIS is used several times by difference distance bands to find the pick of z-score which corresponds to the point of maximum clustering. As mentioned above to look at the compactness of grouping distance was significant and to find out the optimal distance, there was a necessity to set the relatively high clustered region as a standard distance. I first looked through some of the region that has grouping by visual inspection and then zoomed in to see the number of features are in that region to operate the 'calculate distance band from neighbor' tool in Arc Toolbox and obtained the average value of 68.69 (meter). This value is the length of band measured to find 8 nearest neighbor features of relatively high clustered region within entire map that should be bracketed high and low to determine the value to be used in General G calculation. To include less clustered region in the calculation, I decided to start from distance of 30 meter and increase by 5 meter to see how the z-score and measure of General G statistic changes depending on the various distances. For the calculation on the difference image produced by none-parametric maximum likelihood classification, the highest z-score was 6.23235 when the distance was 40 meter. Positive (and highest) z-score which is statistically significant and the larger observed General G index than expected General G index indicate the pattern of clustering of high valued features in the study area (Getis, 1992; Allen, 2010). The optimal distances which provides the most significant clustering for every 10 classification and the corresponding highest z-score, observed and expected General G indexes are organized in the table 4 below.

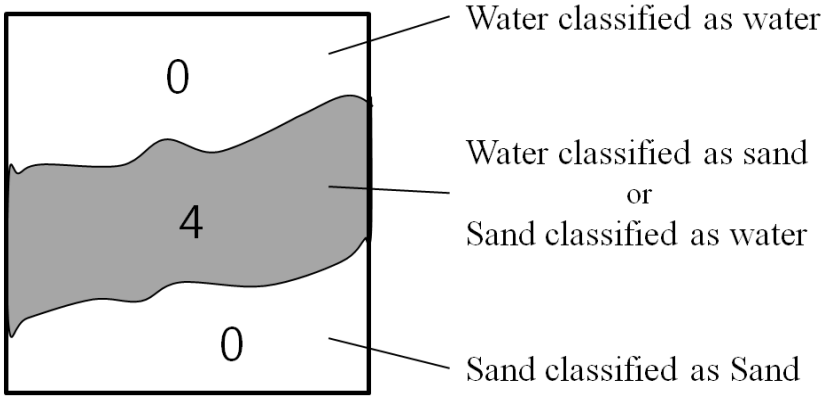


Figure 10. Magnified some boundary of coastline.

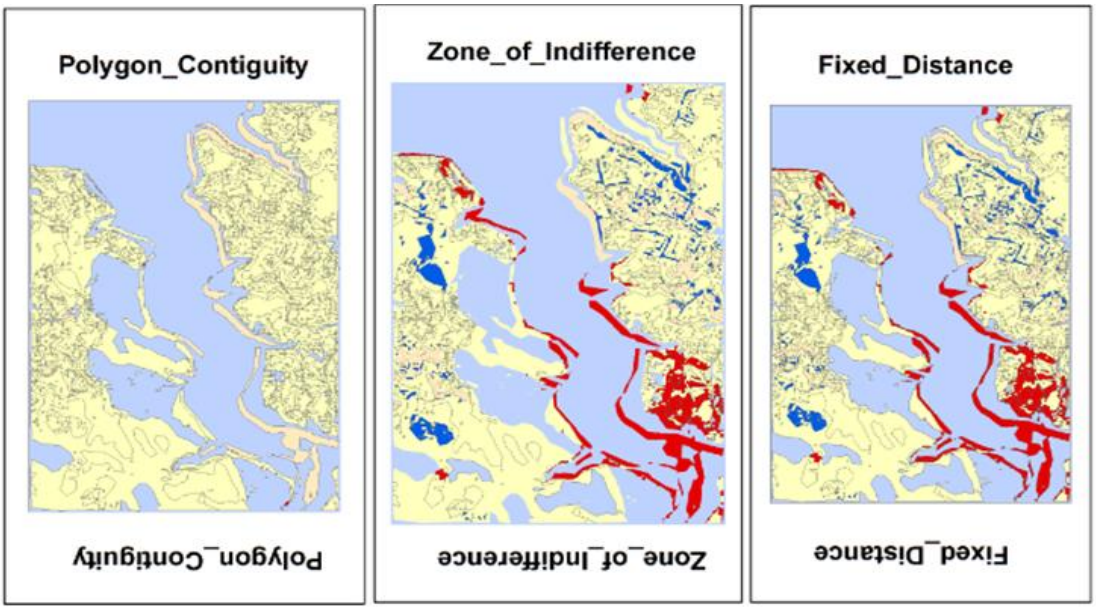


Figure 11. Simulation of using different conceptualization of spatial relationship (LISA).

RESULTS

Until now, using the most commonly used spatial autocorrelation indexes, I have looked at the classification induced spatial pattern of error and saw the clustering pattern of error. As seen in table 5, z-score (or p-value) for every classification is small enough to reject the null hypothesis which is assuming the random distribution to present the statistical significance. Hence the interpretation could be made that there are clustering patterns across the image entirely as it shows in the left adjacent Moran's I column. Moran's I is a relative measure and the significance based on z-score or p-value determines whether the index would be meaningful or not (Mitchell, 2005). Moran' I values range from 0.1087 to 0.174 and the minimum is when ISODATA classification is used while the maximum is when parallelepiped minimum distance classification is used. The average of them is 0.1485 and standard deviation is 0.0226 which is quite small that the difference of errors all-round is not big between classifications. Thus, global spatial autocorrelation tells about the general idea of distribution; however, it has certain limitation in representing the various complex features spread out in an unknown way (even though there is a certain spatial pattern). Therefore along with global spatial autocorrelation, looking at how the spatial aspect of error distribution of every feature is locally autocorrelated is meaningful and the result is shown in figure 12.

As mentioned in previous section, running LISA on ArcGIS creates a new feature class map looks like figure 12 and this provides an intuitive visual interpretation. Above all, this has a strong advantage of ability to look at both global error distribution pattern and autocorrelation of each feature at a time. As seen in figure 12, general pattern of LISA is similar for every classification method. The most obvious local autocorrelation pattern was exhibited at the boundary of water and sand which are two extremes in terms of spectral value. All of the LISA output images showed the high value's clustering along the water boundary (though some differences in degree exist). High valued features surrounded by low values (HL) were shown in a least amount that it wasn't shown in most of the classification method but only at the urban region in mixed classification. As mentioned above, characteristic signature was produced based on spectral reflectance automatically by computer for mixed classification, separability between classes should be smaller than the

training sites picked directly by analyst in supervised classification. [11] Like all other remote sensing applications, urban is the most complex and ambiguous region which is affected by spatial resolution in a greatest manner so it was first expected to have the most pattern of errors. On the other hand, there was a classification method which exhibited the low value's clustering in urban region which was the feature space maximum likelihood. This potentially means that the classification in this region has been done relatively correctly. Another classification method which had quite a lot of LL in urban region was the same maximum likelihood but using parallelepiped for decision rule. The most accurately classified feature over all difference images is water. It is easy to check in figure 9 that all water features are assigned by white color (0, agreement) and blue color (LH) for most of the LISA output images.

For General G statistics, ranges of distance bands that provides the peak z-score were not showing any significant differences which ranged from 20 to 60 as seen in the table 4 and the corresponding z-score also were not much different from each others. This tells that the distribution of pattern of errors within difference images were not that different. All cases were statistically significant and had positive z-score. Also all of the observed General G values were higher than the expected General G values which meant there were pattern of clusterings of high valued features across the image. The result of General G is quite correlated with the LISA output image in figure 12. Thus results from three spatial autocorrelation indexes showed the positive autocorrelation of pattern of errors.

Table 5. Moran's I index for difference images.

	Images	Moran's Index	Z-score / P-value	Distance
1	Parametric MLH	0.1605	25.736 / 0.00	123.549
2	Parametric MHN	0.1603	24.475 / 0.00	107.365
3	Parametric MND	0.1207	27.033 / 0.00	144.849
4	Parallelepiped MLH	0.1620	33.308 / 0.00	212.797
5	Parallelepiped MHN	0.1594	32.987 / 0.00	212.797
6	Parallelepiped MND	0.1740	43.005 / 0.00	174.317
7	Feature Space MLH	0.1605	25.736 / 0.00	123.549
8	Feature Space MND	0.1207	27.033 / 0.00	144.849
9	ISODATA 5 classes	0.1087	30.502 / 0.00	171.853
10	Mixed	0.1578	27.452/0.00	146.456

[11] Signature of each class produced by computer tends to have a linear relationship between classes; therefore the separability should be lower than other classifications.

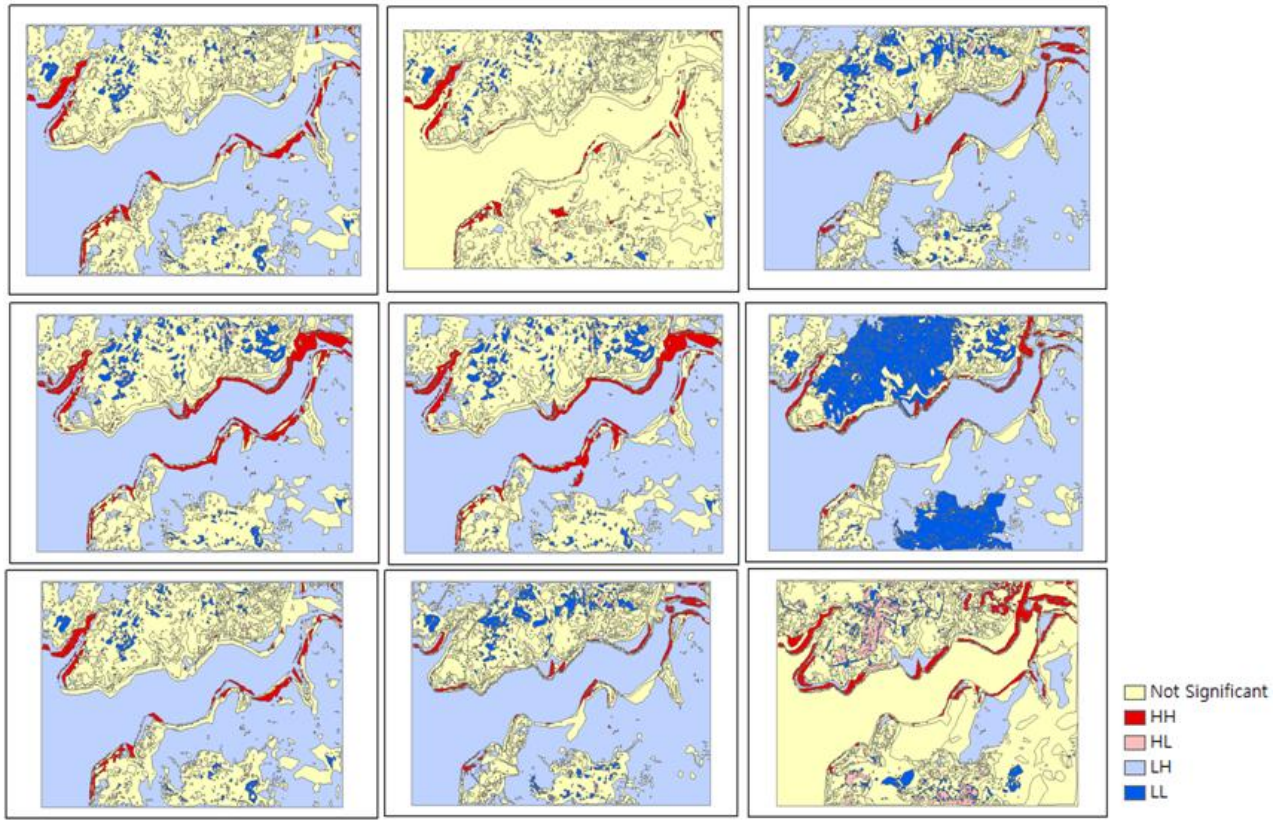
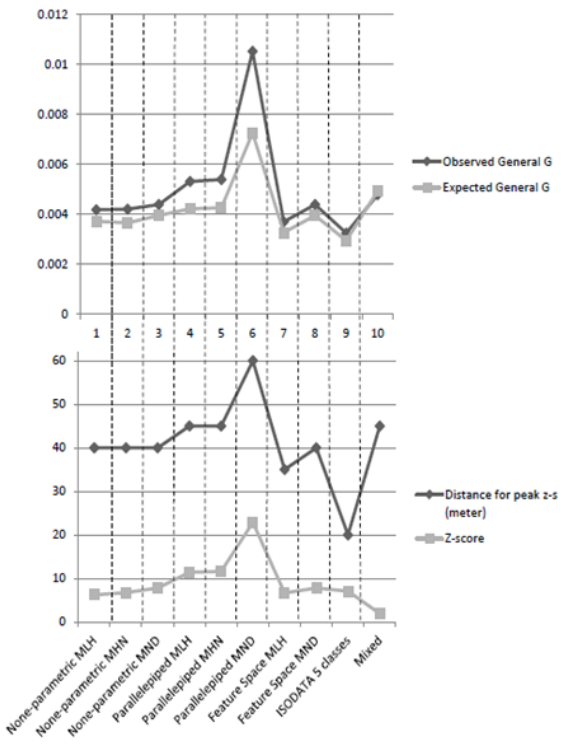


Figure 12. Local Indicator of Spatial Association (LISA) for every difference image. From upper-left: none-parametric MLH, parallelepiped MND, parallelepiped MLH, feature space MLH, mixed, ISODATA 5 classes, feature space MND, parallelepiped MHN, none-parametric MHN (center)

Table 4. General G statistics and plot

	Distance for peak z-score (meter)	Z-score	Observed General G	Expected General G
None-parametric MLH	40	6.2324	0.00418	0.003718
None-parametric MHN	40	6.7415	0.004204	0.003653
None-parametric MND	40	7.8224	0.004393	0.003952
Parallelepiped MLH	45	11.3821	0.005315	0.004218
Parallelepiped MHN	45	11.6502	0.005393	0.004253
Parallelepiped MND	60	22.8333	0.010532	0.007259
Feature Space MLH	35	6.653324	0.003698	0.003265
Feature Space MND	40	7.822443	0.004393	0.003952
ISODATA 5 classes	20	7.058271	0.003245	0.002943
Mixed	45	1.956224	0.004933	0.004773



CONCLUSIONS and LIMITATIONS

Although remote sensing is applied to variety of fields for its many strength, it possesses intrinsic limitations induced from various dimensions of resolutions. Hence, correctly understanding this drawback when studying and applying remote sensing is significant and this study showed one aspect of this by evaluating the error pattern when land cover was classified by using computerized process. Idea of using difference image was variously applied to many fields in remote sensing and other related fields and proven to be useful in understanding pattern of error as well through this study. It became possible to see the spatial pattern of error by applying spatial autocorrelation and also easy to be understood when visualized through difference image. Pattern of errors in image appears from various reasons that it may be affected by high variability if landscape, computer process, or characteristics of the remote sensor (Congalton, 1988). To reduce these errors scientists make effort developing more complicated algorithm and improve sensor's resolution; however, this increases the accuracy but cannot eradicate the error. Therefore as mentioned above, it is always important to know this unavoidable limitation of pattern of error.

By this time, we saw that the pattern of error appears for any kinds of classification method. Hence we should remind ourselves the incompleteness of computerized classification and become more critical about the process. In fact, computer and geospatial spatial softwares dealing with remote sensing or GIS have been developing fast, and accordingly users are given preference of doing complicated analysis in such a simple way. However, this phenomenon often times made analysts unnecessary to possess the thorough backgrounds of fundamental principles to perform analysis only increasing the dependency on computer unprecedentedly high. Because of this phenomenon, analysts often do not realize the incompleteness of human-made algorithms and computer's simplified (in user's perspective) automatic process in geospatial analysis and even they are conscious of this imperfection, most of the time people don't really have a chance to consider how much they are different from real world. Through this study precision was estimated by applying spatial autocorrelation to classified images using various methods, and with this it was possible to see the limits of human made classifications.

Reference

Allen, D. (2010) GIS tutorial 2: Spatial analysis workbook, ESRI Press, Redlands, CA.

Anselin, Luc. "Local Indicators of Spatial Association—LISA," *Geographical Analysis* 27(2): 93–115, 1995.

Bruzzone, L. (2000). Automatic analysis of the difference image for unsupervised change detection. IEEE transactions on geosciences and remote sensing, 38 (3).

Chou, Y. H. (1991). Map resolution and spatial autocorrelation. Geographical analysis, 23 (3).

Congalton, R. (1988). Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric engineering and remote sensing, 54 (4), 587-592.

Dormann, C., McPherson, J. M., Araujo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, D., Kuhn, I., Ohlemuller, R., Peres-Neto, P. R., Reineking, B., Schroder, B., Schurr, F., and Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography, 30, 609-628.

ESRI, Inc. (2010). *ArcGIS Resource Center*, ArcGIS v. 10, Redlands, CA.

Getis, A, and Ord, J. K. (1992) The analysis of spatial association by use of distance statistics. Geographical analysis, 24 (3).

Henebry, G. M. (1993). Detecting change in grasslands using measures of spatial dependence with Landsat TM data. Remote sensing of environment, 46, 223-234.

Jensen, J.R. (2005) *Introductory Digital Image Processing: A Remote Sensing Perspective*, 3rd ed., Prentice Hall, Upper Saddle River, NJ.

Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm? Ecology, 74 (6), 1659-1673.

Lu, D., Mousel, P., Batistella, M., and Moran. (2003). Comparison of land-cover classification methods in the Brazilian Amazon basin. American society for photogrammetry and remote sensing 2003 annual conference proceedings. Anchorage, Alaska.

Mitchell, Andy. *The ESRI Guide to GIS Analysis,* Volume 2. ESRI Press, 2005.

Monroe, D., Southworth, J., and Tucker, C. M. (2001). The dynamics of land-cover change in western Honduras: Spatial autocorrelation and temporal variation. AAEA annual meeting.

Moran, P. A. P. (1950). "Notes on Continuous Stochastic Phenomena". *Biometrika* 37 (1): 17–23

Overmars, K. P., de Koning, G. H., and Veldkamp, A. (2003). Spatial autocorrelation in multi-scale land use models. Ecological modeling, 164, 257-270.

Qi, Y., and Wu, J. (1996). Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. Landscape ecology, 11 (1), 39-49.

Rees, W. G. (2000). The accuracy of digital elevation models interpolated to higher resolutions. International journal of remote sensing, 21 (1), 7-20.

Russel, G. C. (1988). Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing, 54 (5), 587-592.

Slocum, T., McMaster, R., Kessler, F., and Howard, H. (2009) Thematic Cartography and Geovisualization, 3rd ed., Prentice Hall, Upper Saddle River, NJ.

Verbyla, D. L., and Hammond, T. O. (1995). Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. International journal of remote sensing, 16 (3), 581-587.