Pairwise Change Analysis Techniques for Quantitative Geographic Data:

A Survey of Techniques Introduced in the United Nations Institute for Training and Research (UNITAR) Workbook

Andrew Munsch
GEOG 5565
December 14, 2007

Introduction

This project attempts to apply the techniques introduced in the Change and Time-Series analysis UNITAR workbook to examine suburban development in the Twin Cities Metropolitan Area, specifically in the area surrounding the Lafayette Freeway in northern Dakota County from 1986 to 2002.

The Lafayette Freeway, now marked as US Highway 52, was constructed in the mid 1970's to provide a freeway link between downtown St. Paul and the suburbs in northeastern Dakota County (West St. Paul, South St. Paul, and Inver Grove Heights, see the map, right, or an overview of the study area). However, the freeway only extended as far south as State Highway 110, an east-west highway, which was converted into Interstate 494 in the mid 1980's. After I-494 was completed in 1986, the southern end of the freeway was extended southward piece by piece until 1994, when it was finally linked with the main 4-lane highway connecting the Twin Cities and Rochester (US 52/ Highway 55), giving St. Paul's its first continuous expressway connection to southeastern Minnesota. The new freeway replaced the old routes into downtown St. Paul via Robert Street and Concord Boulevard.

Study Area Map

With the given changes to the transportation network in northern Dakota County, the question arises as to how development in the region has changed over time since 1986, and if raster data of surface permeability can be used to analyze this change and/or detect significant trends. The UNITAR workbook on Change and Time Series analysis provides a framework of techniques and exercises to introduce methods to analyze change in a variable between two images using the IDRISI Andes software with raster data. The primary goal of the following investigation is therefore to examine the usefulness of the techniques introduced in the UNITAR workbook by modifying them for use in a local suburban context with impervious surface data.

The following investigation covers the first four exercises in the workbook, which examine pairwise comparison techniques for quantitative data. The exercises were conducted using both the data provided with the workbook and impervious surface data for our study area in northeastern Dakota County. The impervious surface data of the Twin Cities metro area, which is quantitative in nature, is analogous to the NDVI data (an index measuring vegetation density) used in the UNITAR exercises. The impervious surface data was developed by the Met Council in conjunction with the Remote Sensing Lab at the University of Minnesota. The local data has a horizontal resolution of 30 meters. In contrast to the UNITAR data, the local data to be used will have a uniform scale, and all exercises will be completed for the same geographic study area. Images for the years 2002 and 1986 were used. The 1986 image represents the area before the construction of the south end of the Lafayette Freeway, while the 2002 image represents the area after the completion of the freeway in 1994, allowing time for any subsequent development to take place.

UNITAR Workbook Exercises

Exercise 1: Simple Differencing

The UNITAR workbook starts out with the most basic of comparison techniques: simple differencing. Essentially, the difference of the value of each pixel is taken between the two images in question. The resulting image shows the raw difference in value for each pixel. The workbook uses Normalized Difference Vegetation Index (NDVI) images of the African continent for 1987 and 1988 for the first three exercises. For our local example, we will use the impervious surface index images of the study area for 1986 and 2002.

The exercise first asks for a subjective visual comparison of the two images of Africa (see figures 1 and 2 below). Overall, the 1988 image appears to have more green (high NDVI value) areas, although there are a few areas that appear to have a smaller NDVI value, clustered in east-central Africa.
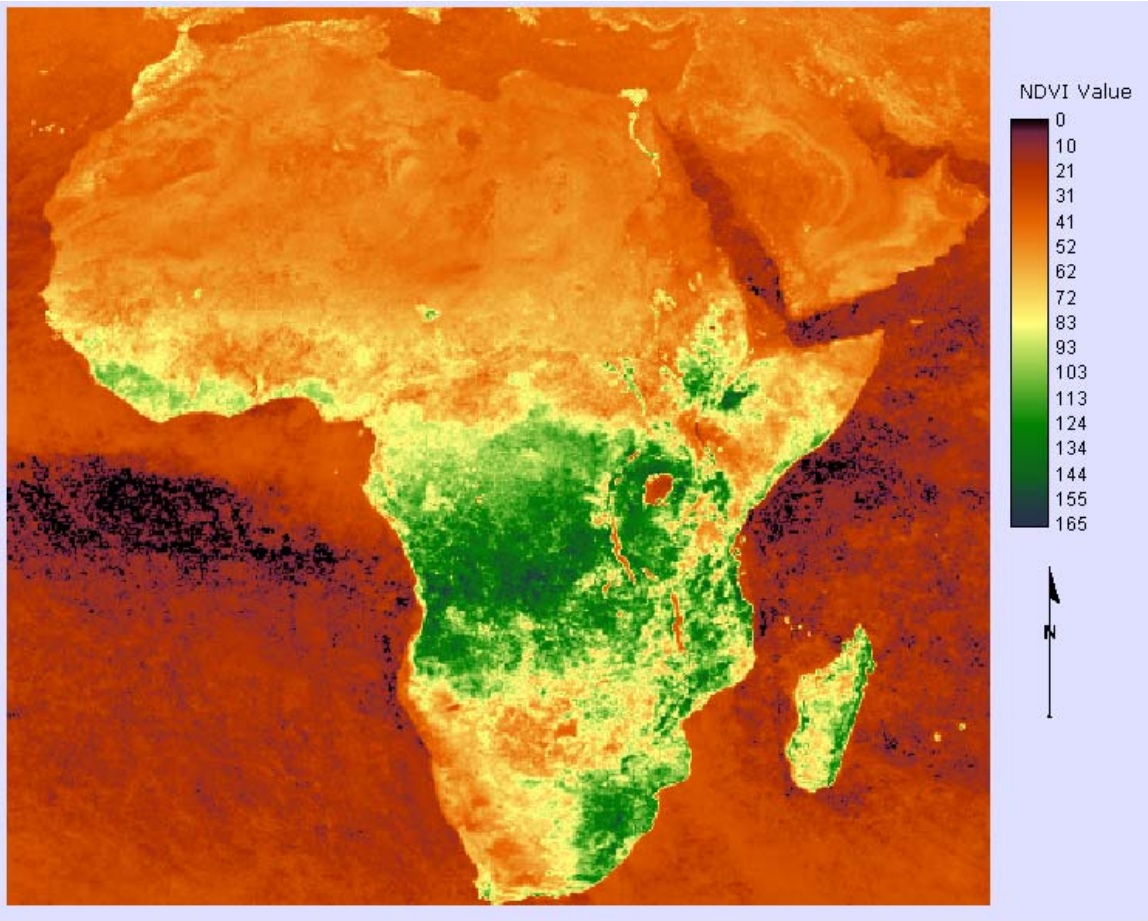
NDVI Value

0
10
21
31
41
52
62
72
83
93
103
113
124
134
144
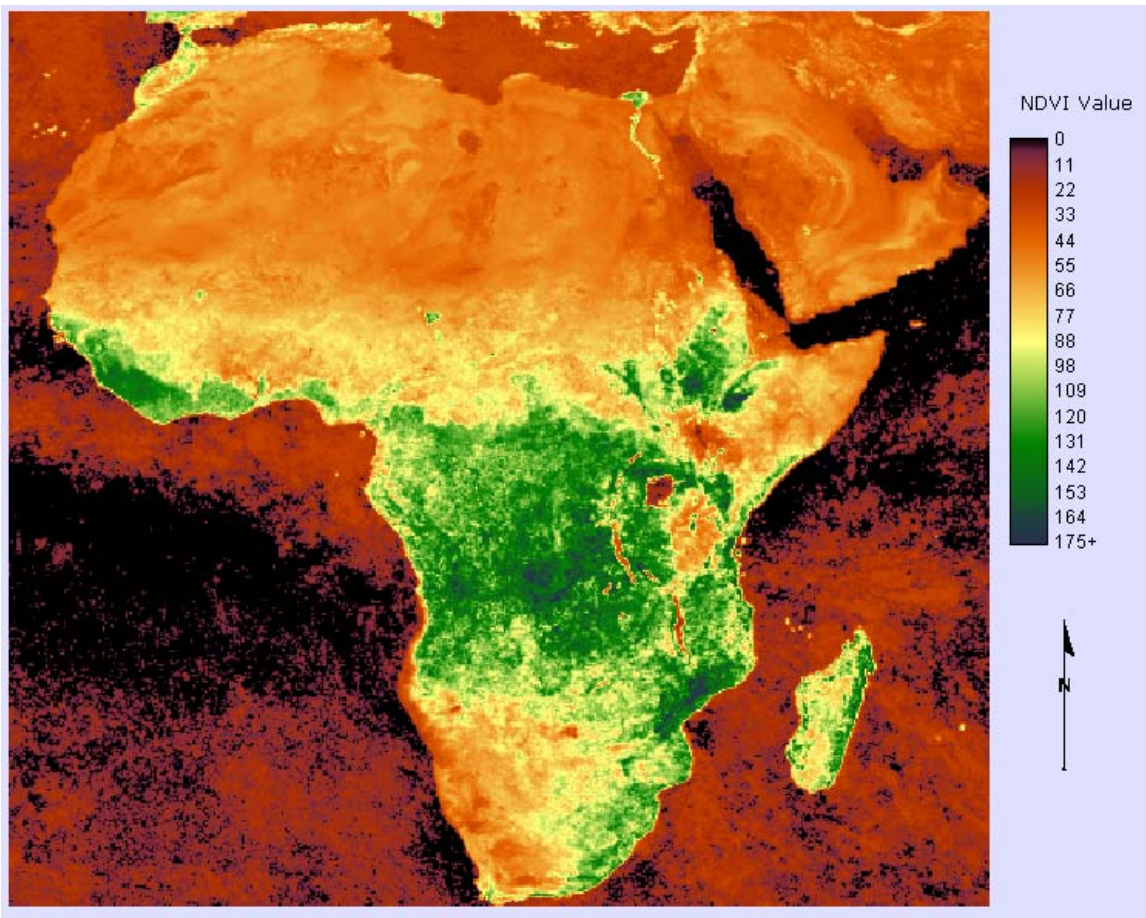155
165

Figure 1: NDVI Image of Africa - 1987

Figure 2: NDVI Image of Africa - 1988

In a visual comparison of the images of our study area (see figures 3 and 4 below), the northern areas of both images appear very similar, showing the dense impervious surface density of the urban environments developed well before 1986 in West and South St. Paul. Evident in both images is the South Robert Street commercial strip in the northwest part of the image, running north and south, as well as the South St. Paul stockyards in the northeastern edge of the study area, developed for heavy industrial uses. However, the southern area of the 2002 image has quite a bit more non-zero areas registered than the 1986 image, especially in the area around and to the south of the Lafayette freeway, which seems to indicate a southward push of the developed area.
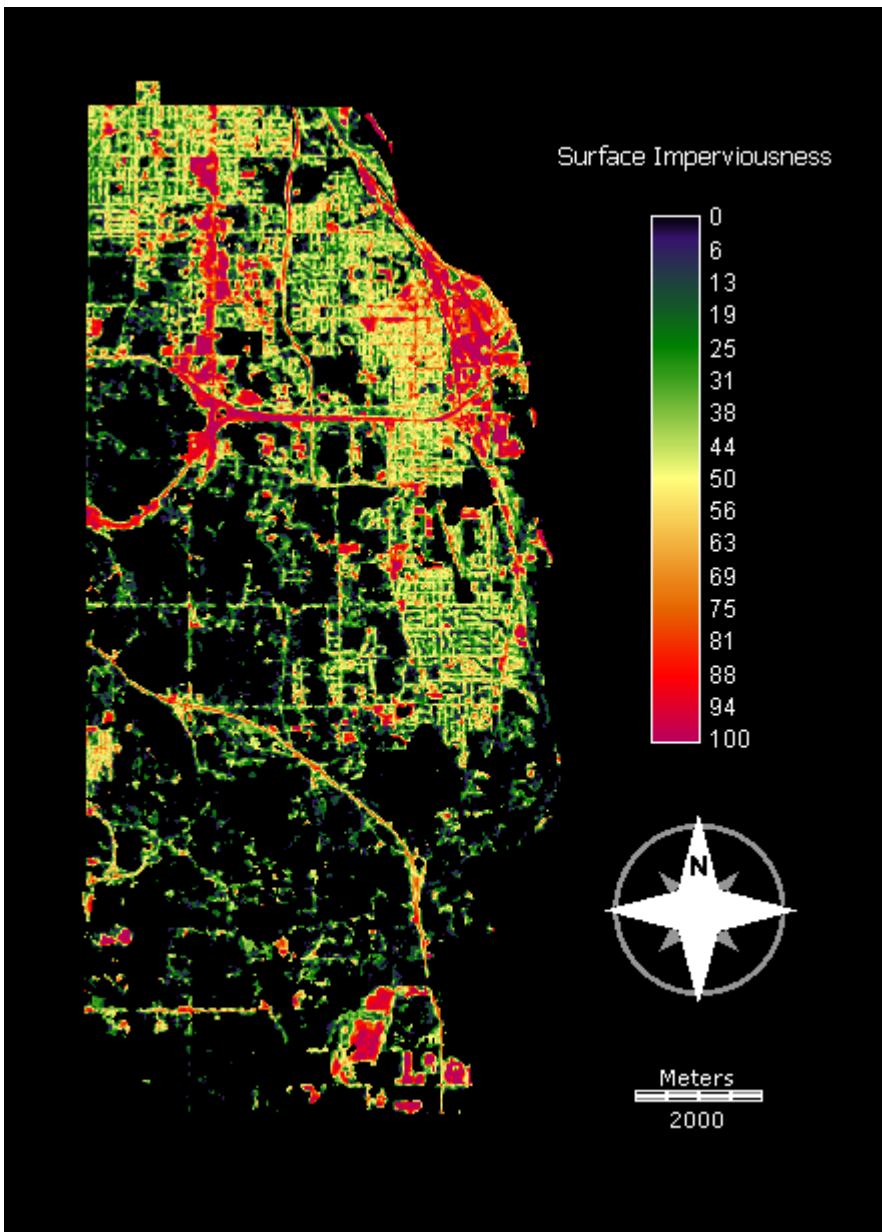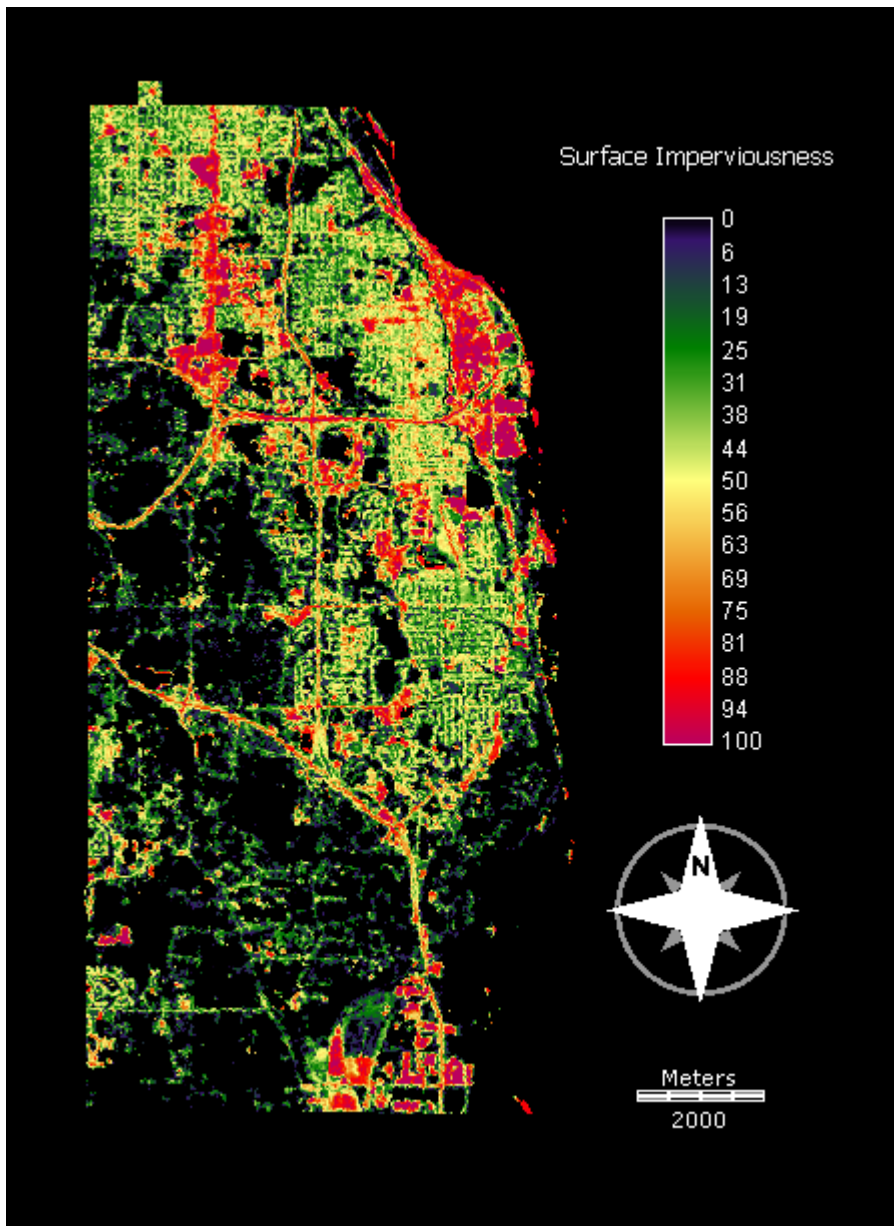
Figure 3: Surface Imperviousness - 1986

Figure 4: Surface Imperviousness - 2002

The next step in the exercise is to perform a difference operation on the two images using the image calculator in IDRISI using the expression:

[difference image] = [image time 2} – [image time 1].

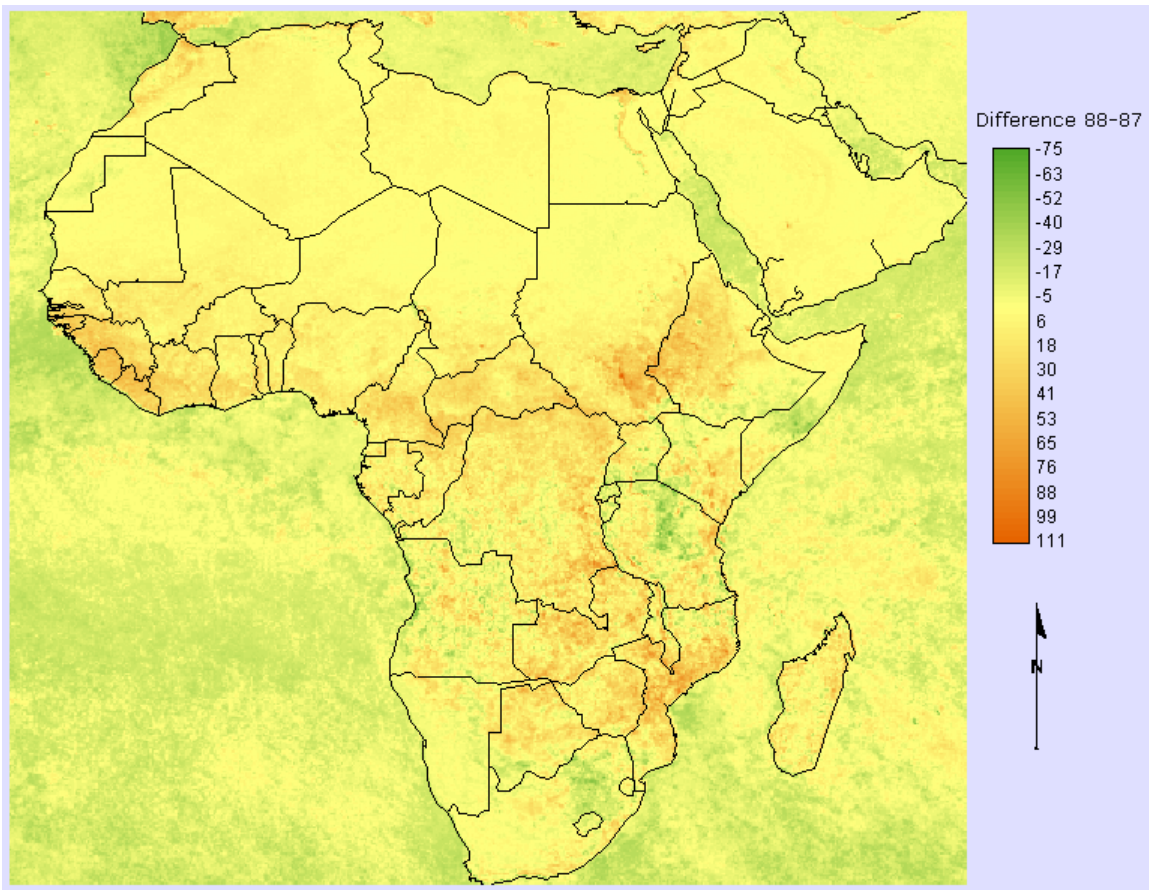The results are shown below for Africa (figure 4) and for our study area (figure 5).

Figure 5: Difference in NDVI Value, 1988-1987

For the Africa difference image, deeper shades of orange represent areas where the NDVI value has increased, while green areas express a decrease. There seems to be a widespread overall increase in a band across north central Africa, extending southward into south central Africa, with more intense increases occurring at the ends of the bands. Large areas showing a decrease appear mainly in three clusters, one in southern Somalia, another in eastern South Africa, and another in north central Tanzania.

Figure 6: Difference in Surface Imperviousness Value, 2002-1986
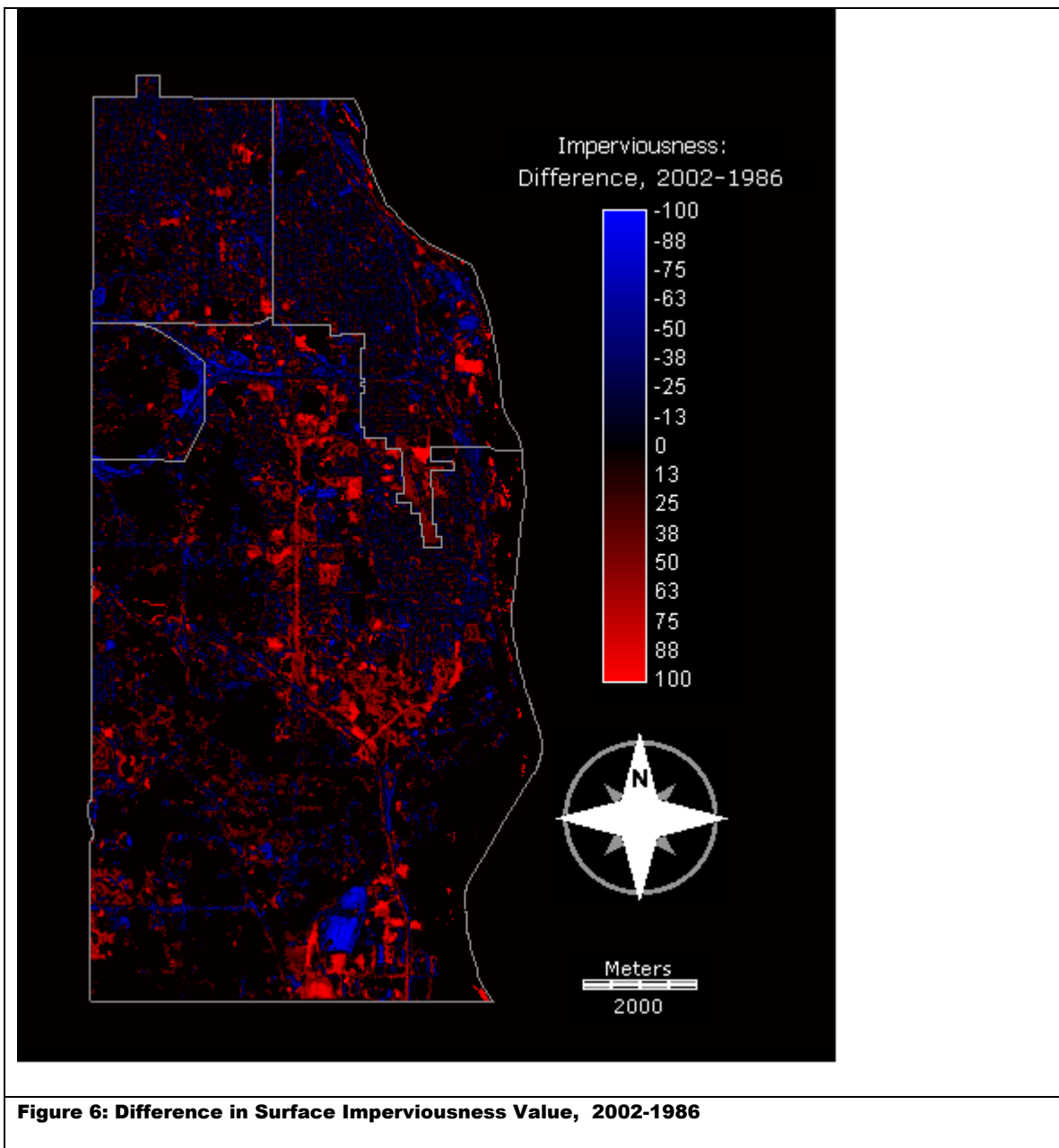
For our study area, the overall trend seems to be an increase in impervious surface value,

represented by areas shaded red in Figure 6 (above). This is especially true in an area of

north-central to southeastern Inver Grove Heights in the vicinity of the new freeway, and

especially at its southern end and around the Concord Boulevard corridor, which was

already noticeable by simple image to image comparison. Other areas of increase include

a region corresponding to the South St. Paul stockyards, , and the area around the South

St. Paul airport (in the southern peninsula-like extension of South St. Paul's city limits). Areas of decreased imperviousness, shaded blue, seem to be clustered along the 494 corridor in northern Inver Grove Heights and Sunfish Lake, as well as a rather conspicuous cluster in south central Inver Grove Heights. The former area might be explained by the fact that 494 was completed in 1986, and the area around the construction zone might have shown up as a barren or more impervious surface when the data was taken. Landscaping and plant growth in the years following the Interstate's completion could explain the decrease in apparent imperviousness. The latter area in south-central Inver Grove Heights corresponds to an area occupied by a large solid waste landfill. It might be expected that this area, which is constantly in flux due to dumping activity, would show large changes in reflectance to remote sensing devices, especially if the area of the landfill was covered with soil and grass that would end of being classified as a natural porous landscape. There also appears to be a large amount of low-level negative difference in the northern parts of the image, corresponding to areas of older residential districts. Perhaps natural tree growth could account for the widespread apparent decrease in surface imperviousness in these areas.

Despite the massive difference in scale between our study area and the Africa data used by the UNITAR workbook, the simple difference technique appears to be quite effective at bringing areas of raw positive and negative difference in a variable between two images into focus in comparison to subjectively comparing the two images.

Exercise 2: Thresholding

In the first exercise, we were easily able to find areas of positive and negative difference in a variable between two images. The second exercise in the UNITAR workbook seeks to find a way to determine a threshold, above or below which the

difference value represents real change instead of just normal variability. The UNITAR

workbook bases the determination of this threshold on the fact that values above or below

two or more standard deviations of the mean of a set of values that follow a normal

distribution can be seen as showing real change, as we can expect 95% of values to be

within 2 standard deviations, and 99% of values to be within three standard deviations.

The UNITAR workbook therefore sets the initial thresholds for change using the

following formula:

Upper/Lower Threshold = (Mean Change) +/- 3*(Standard Deviation).

This all assumes that our data follow a normal distribution, which they of course do not.

Using the HISTO module of IDRISI, we can determine the mean and standard deviations

of both the Africa NDVI difference image, and our study area difference image from

exercise 1. These values are given in table 1 below:

Table 1: HISTO Results

Data	Mean	Std. Deviation	+/- Threshold Value (3σ)
Africa NDVI	-1.62	14.315	41.325/-44.565
Local Impervious Surface Data	3.094	23.53	73.684/-67.496

According to the values in table 1, differences above 41.325 or below -44.565 show real

change with the Africa NDVI data according to the criteria put forth by the workbook,

while changes above 73.684 or below -67.496 represent real change of impervious

surface data for our study area.

Using the RECLASS module in IDRISI, we can isolate these areas from the

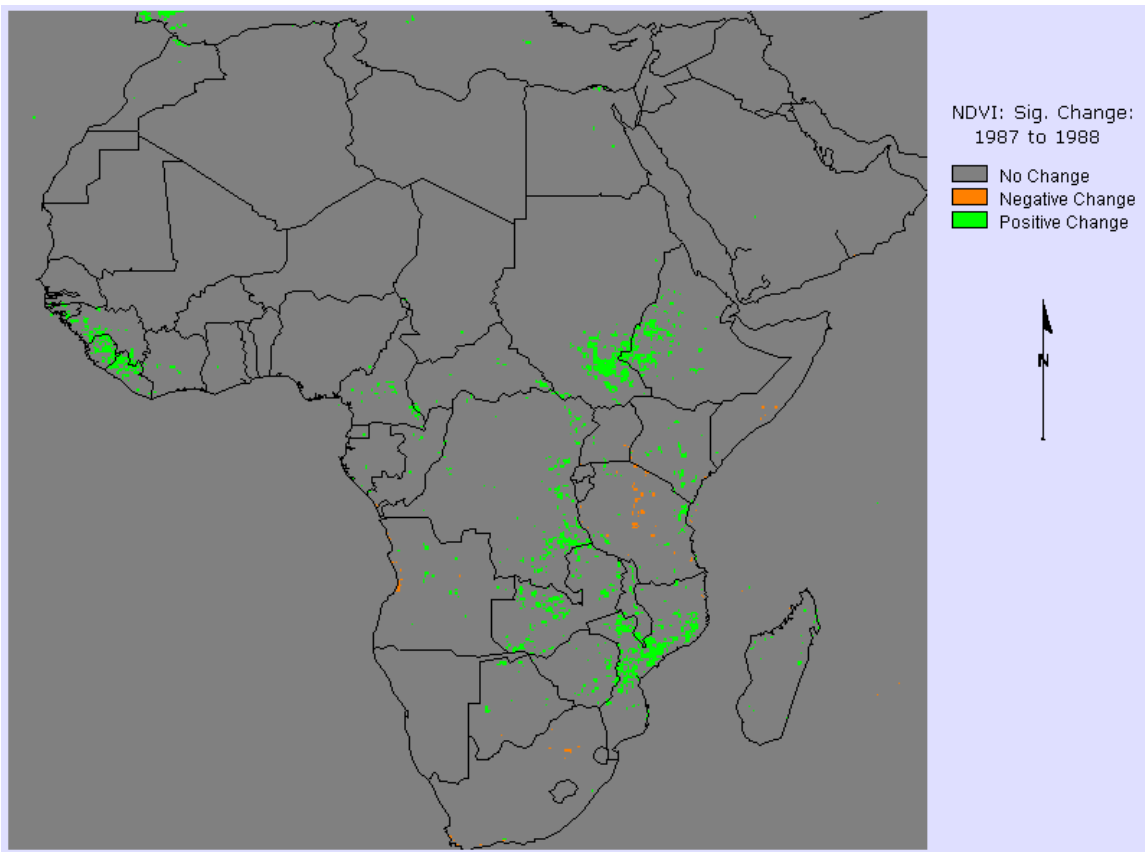difference images created for exercise 1. The results are given below:

Figure 7: Significant Change in NVDI 1987 to 1988

The resulting image of Africa (Figure 7, above) still shows a large number of areas where significant positive change has occurred, mainly in the areas that were already evident in the simple difference image. Although widespread increase was evident in the simple difference image, the change image brings focus to several clusters of positive change. These include areas in Mozambique, the southern Sudan and southeastern Ethiopia, and the western costal states (Liberia, Sierra Leone, Guinea). Areas of significant negative change appear to be quite limited in comparison to the areas of negative difference in the simple difference map, although they do appear in the locations expected. These areas are limited to a smattering in central Tanzania, one small area in northern South Africa, and another area on the west coast in Angola that was not as evident on the simple difference map.
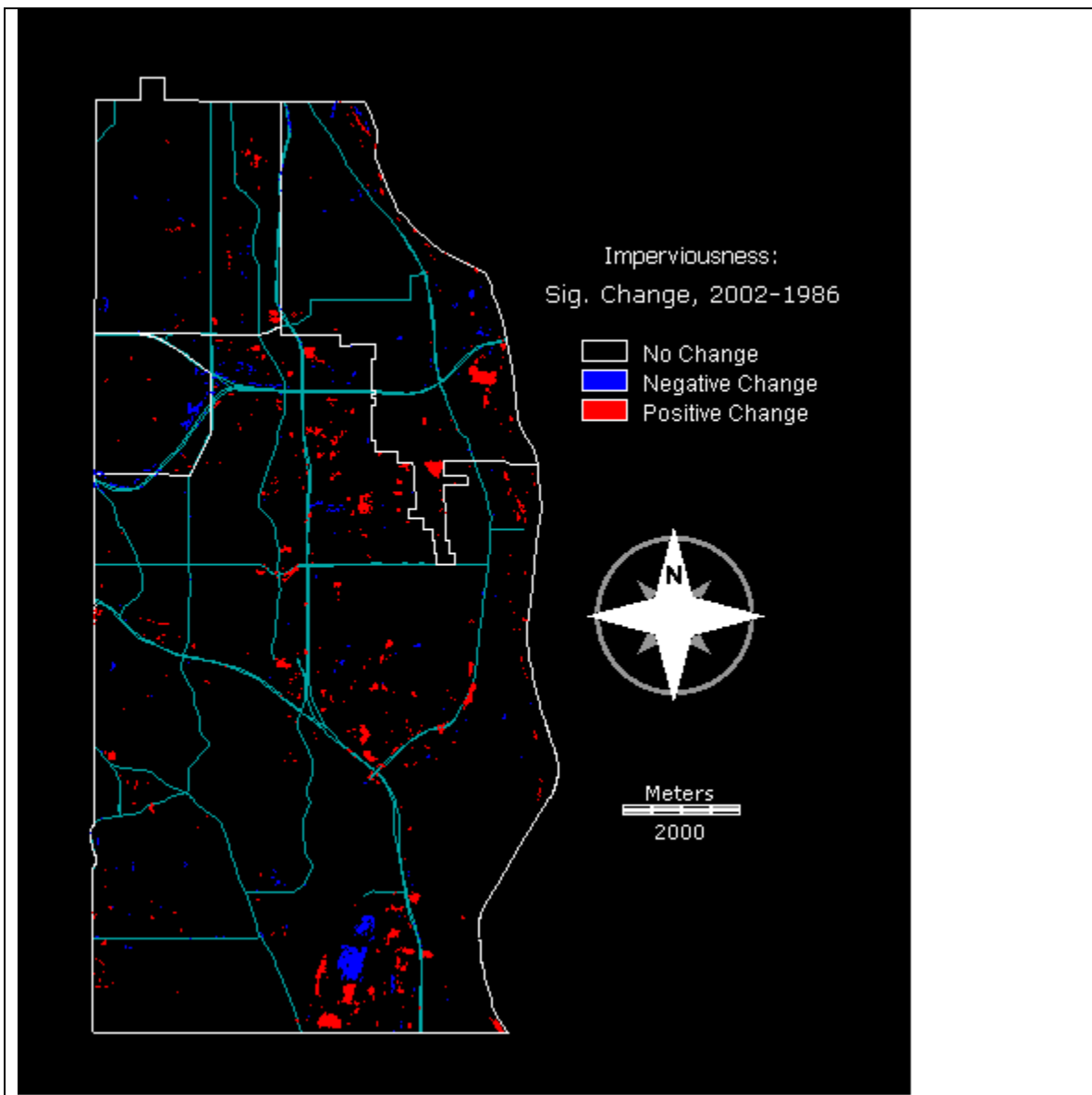
Figure 8: Significant Change in Impervious Surface Value, 1986 to 2002

Figure 8 (above) shows the areas of significant change in surface imperviousness for our study area (major highways have been added for clarity). The areas of positive and negative change occur mostly in the areas already pointed out in the simple difference image. However, much of the scattered low-level negative difference in the northern part of the map has been removed, leaving only limited focused areas around the map. The negative change areas in the western 494 corridor and at the landfill are again evident. Significant positive change seems more focused in the areas immediately to the southeast

of the interchange of 494 with the Lafayette Freeway in northern Inver Grove Heights. These areas were harder to see in the wider area of positive difference seen in the simple difference map.

Overall, the technique of thresholding seems to be a decent tool for focusing in the most significant areas of change in a given spatial variable. However, the UNITAR workbook's focus is on finding a statistical value that determines "real change" from natural variability. The technique of using 3 standard deviations on either side of the mean seems rather arbitrary, with its theoretical back-up based on abstract statistical theory, rather than on real-world knowledge of what constitutes real change in a given geographic context. This is harder to see with continental scale data, such as the Africa data given in the workbook example, where changes in vegetation could mean different things depending on geographic location within the continent. Only the most obvious changes apparent from small-scale, low resolution satellite images would be of interest here, so perhaps the arbitrary technique put forth here is adequate with small-scale data. However, at the higher resolution data used for our study area, we have a better knowledge at what constitutes a change from a non-impervious to an impervious surface, namely suburban development. In this case, ground-truth could be used to better determine a threshold, perhaps finding different thresholds for different types of development.

Exercise 3: Regression

For the third exercise, the UNITAR workbook takes a step back to assess the possibility that the two images used for comparison could differ, not because of actual changes on the surface, but because of changes to the remote sensing device, or in other factors that affect how the given variable is measured. Simple bivariate regression is put

forth as a way of determining if one of the two images being compared exhibits a systematic "gain" or "offset" from the other. Simply put, the relationship between the two images is reduced to a simple linear equation, where the later of the two images is used as the dependent variable, and the earlier image is given as the independent variable. Once this relationship is known, it can be applied to the time 1 image to correct for any differences in the mean and in variation between the two images. The methods from exercises 1 and 2 to determine difference and significant change can then be conducted as usual.

The first step in this process is to use the REGRESS module in IDRISI to find the linear equation that represents the relationship between the two images. The results of this operation for both the Africa data and our study area are given in Figures 9 and 10 below.
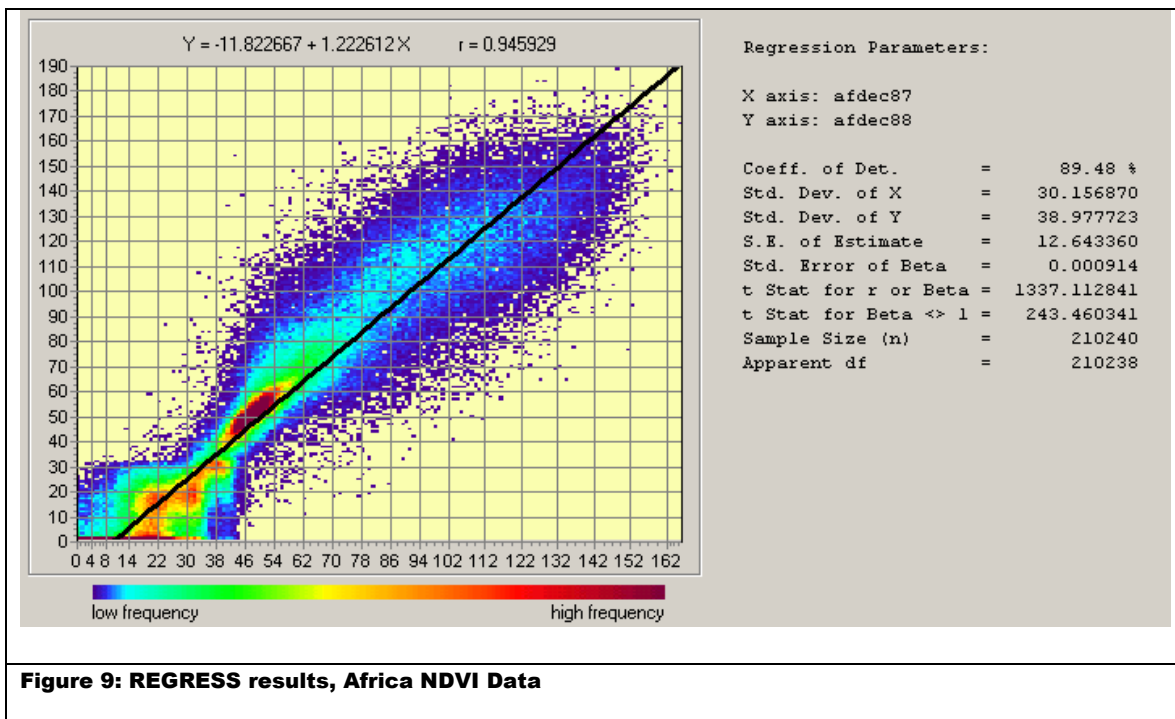

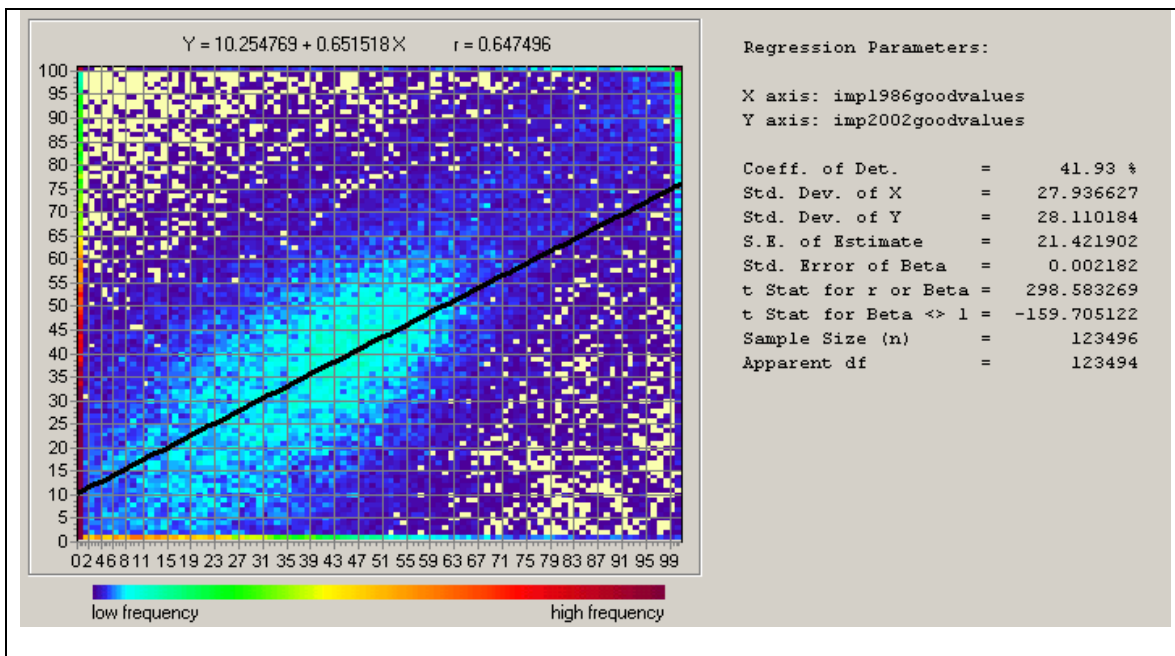**Figure 9: REGRESS results, Africa NDVI Data**

Figure 10: REGRESS results, Study Area Impervious Surface Data

Note how different the results of the REGRESS module were for our study area data in comparison the Africa data. The coefficient of determination for the Africa data is 89.48%, and the t statistic is a very high 1337.113. Questions of spatial autocorrelation and overestimation of degrees of freedom aside, the UNITAR workbook concludes that these statistics are enough to show that there is definitely a gain between the two NDVI images, and they proceed with the rest of the procedure accordingly. However, the relationship between our impervious surface images is not nearly as strong. The coefficient of determination is only 41.93%, and the slope is much lower at .651. Our t statistic is also much lower at 298.58, and we have fewer degrees of freedom. This all seems to point out that our data does not show a definite gain, at least in comparison to the Africa data. However, we will play devils advocate and go along with the procedure anyway.

The second step in this process is to apply the regression equation from the REGRESS module using the image calculator in IDRISI to the time 1 image to create an adjusted image. See figures 11 and 12 below for the results.
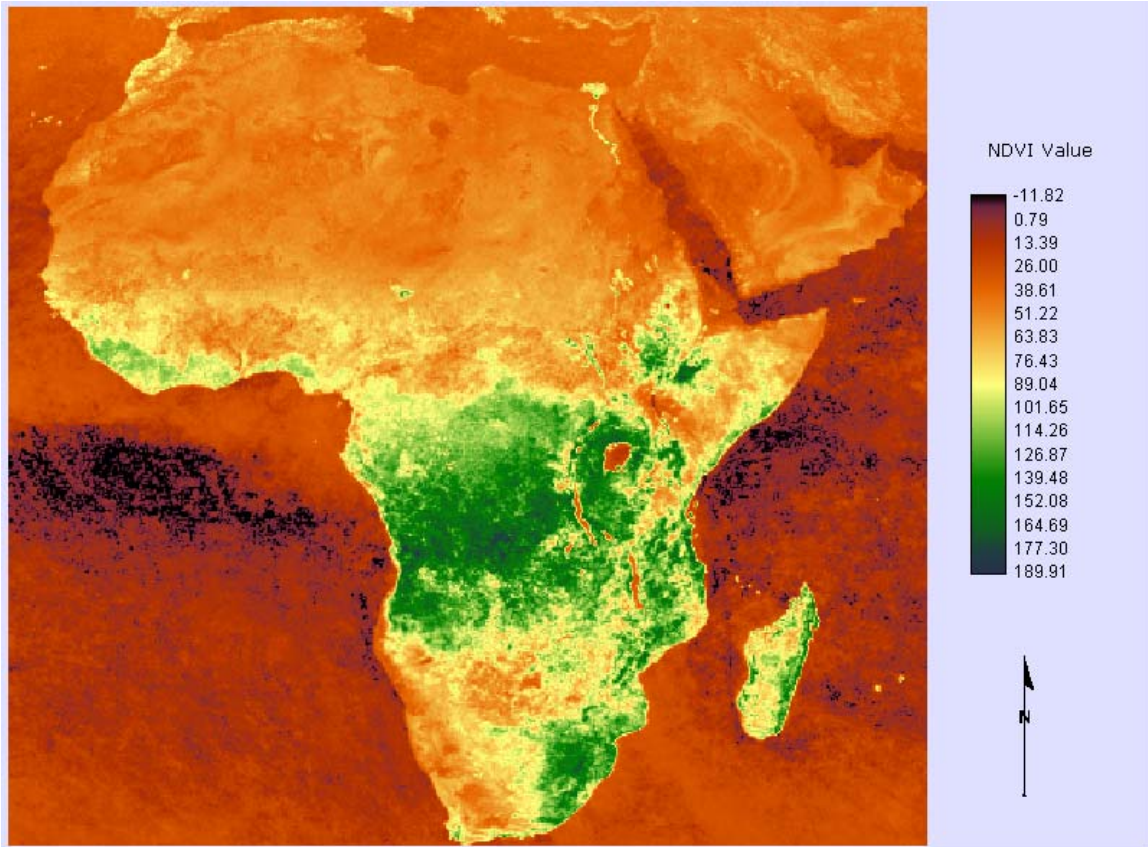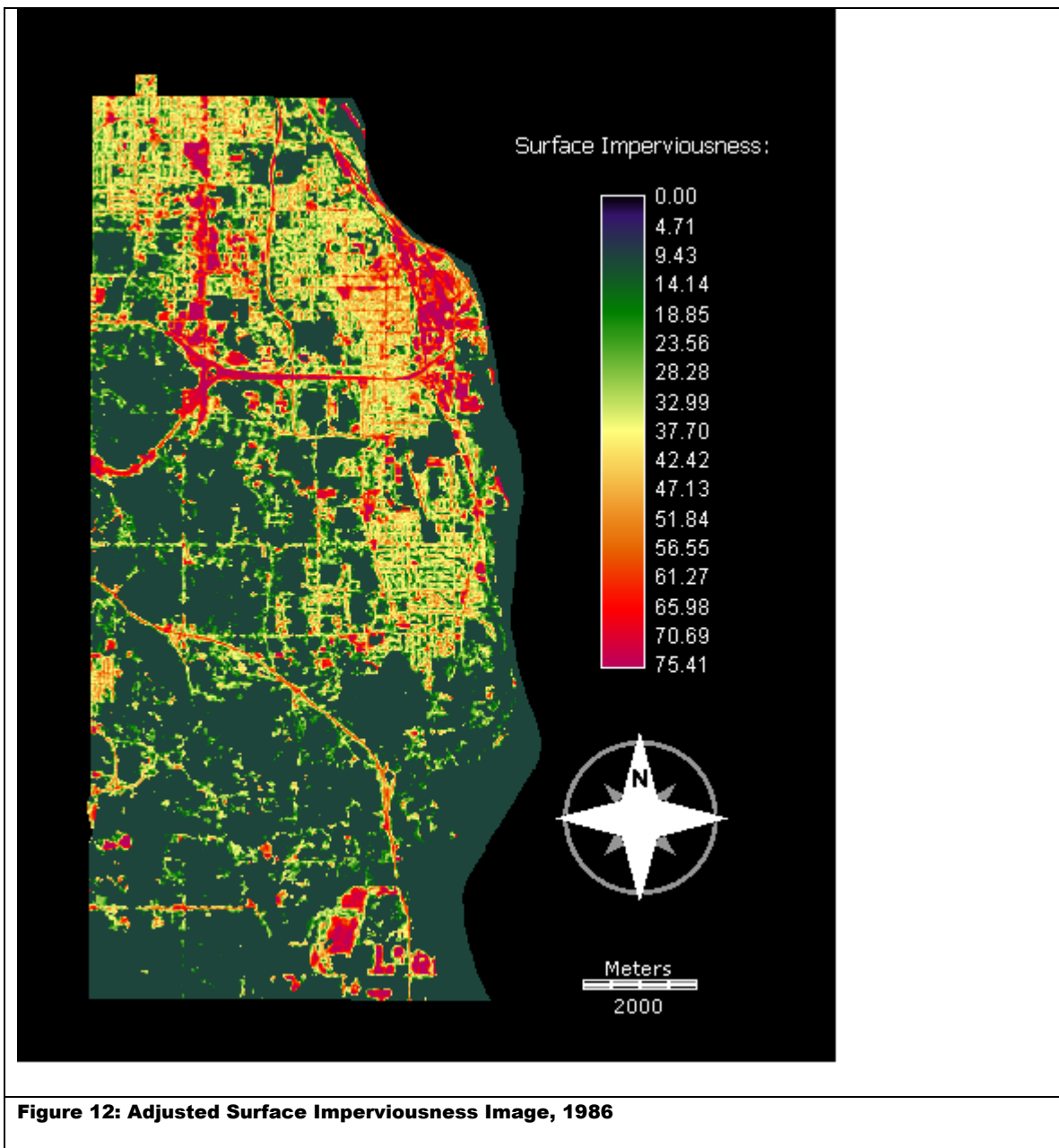

Figure 11: Adjusted 1987 NDVI Image

Figure 12: Adjusted Surface Imperviousness Image, 1986

The adjusted 1987 NDVI image of Africa (Figure 11) appears to be very similar in appearance to the unadjusted image (Figure 1). However, our adjusted 1986 study area (Figure 12) image appears very different from the original (Figure 3). What were once zero imperviousness values have now been offset to 10.254769, the intercept of our regression equation. This would make more sense if the 2002 image had some sort of

overall offset for the background values, but this is not the case. Again, it seems that these procedures do not make sense for the data we are using for our study area

The third step in this procedure is to create simple difference images using the same procedure as in Exercise 1. The resulting images are given in figures 13 and 14 below.
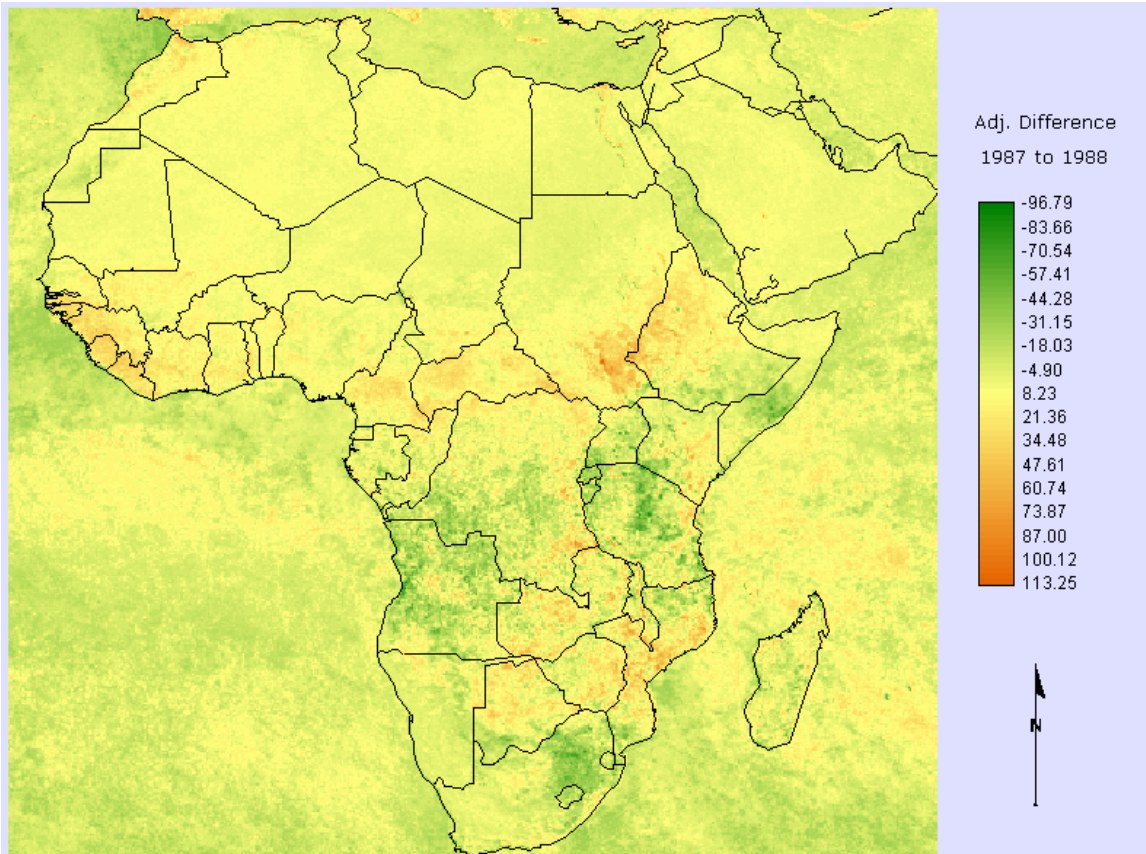

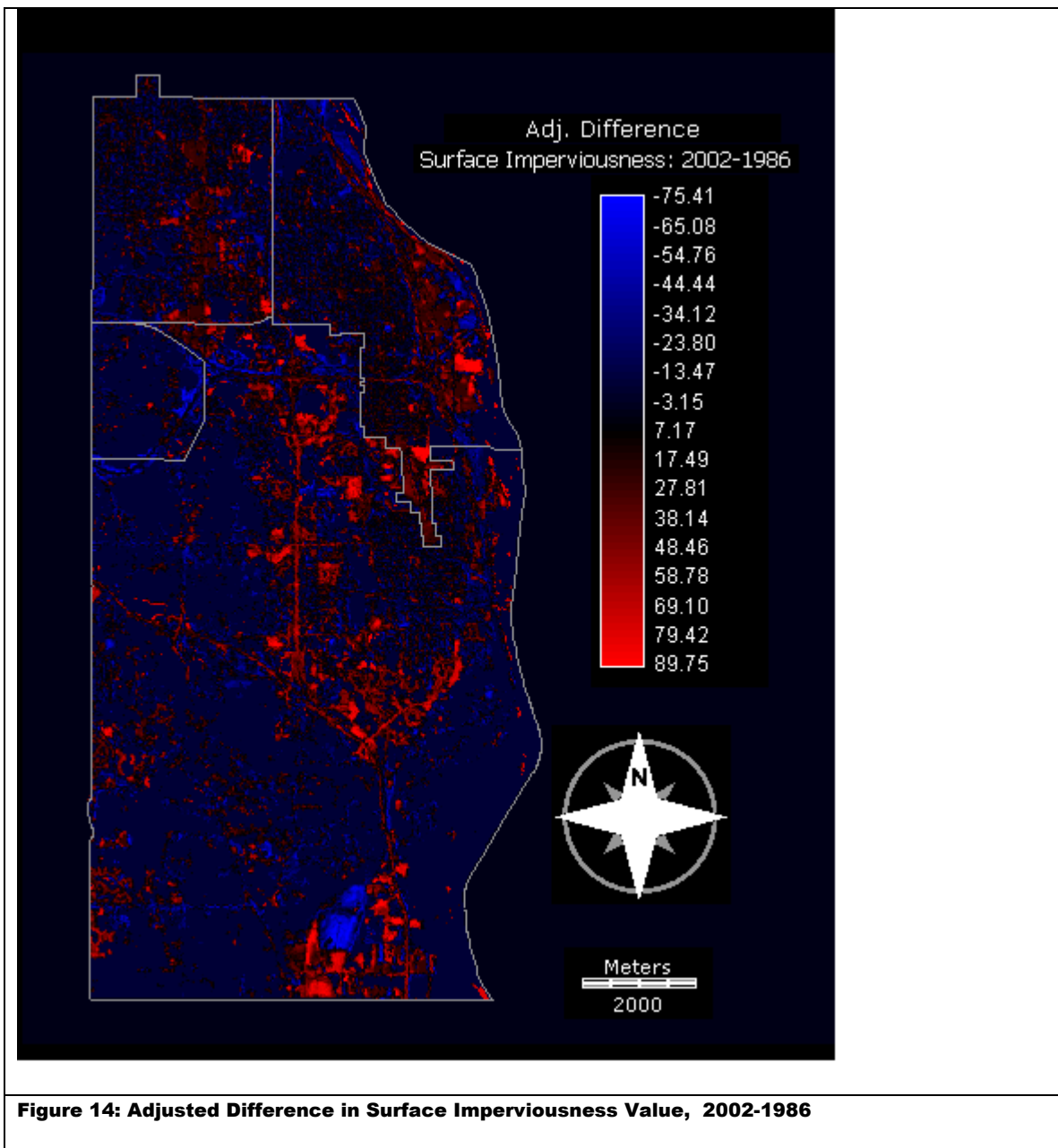**Figure 13: Adjusted Difference in NDVI Value, 1988-1987**

Figure 14: Adjusted Difference in Surface Imperviousness Value, 2002-1986

The adjusted difference image of Africa (Figure 13) shows a great deal more negative NDVI change in comparison to the unadjusted difference image (Figure 5). The adjusted study area image (Figure 14) also shows a great deal more negative change. However, much of this is false, introduced by the artificial background offset as seen in figure 12.

Table 2: Comparison of HISTO Results: Adjusted vs. Unadjusted Difference Images

	Africa NDVI	Africa NDVI (Adjusted)	Study Area Impervious Surface Value	Study Area Impervious Surface Value, Adjusted

Mean	-1.62	0	3.094	0
Std. Deviation	14.315	12.656	23.53	21.424
+/- Threshold Value (3σ)	41.325/-44.565	37.968/-37.968	73.684/-67.496	64.272/-64.272

From the results of HISTO given in Table 2, we can see that the mean of both adjusted

difference images was adjusted to 0 (a slight increase for the Africa data, a larger

decrease for our study area data). Also, the standard deviation was reduced for both

images.

The final step is to create significant change images using the new threshold

values given in Table 2 for the adjusted difference images. The results are given in
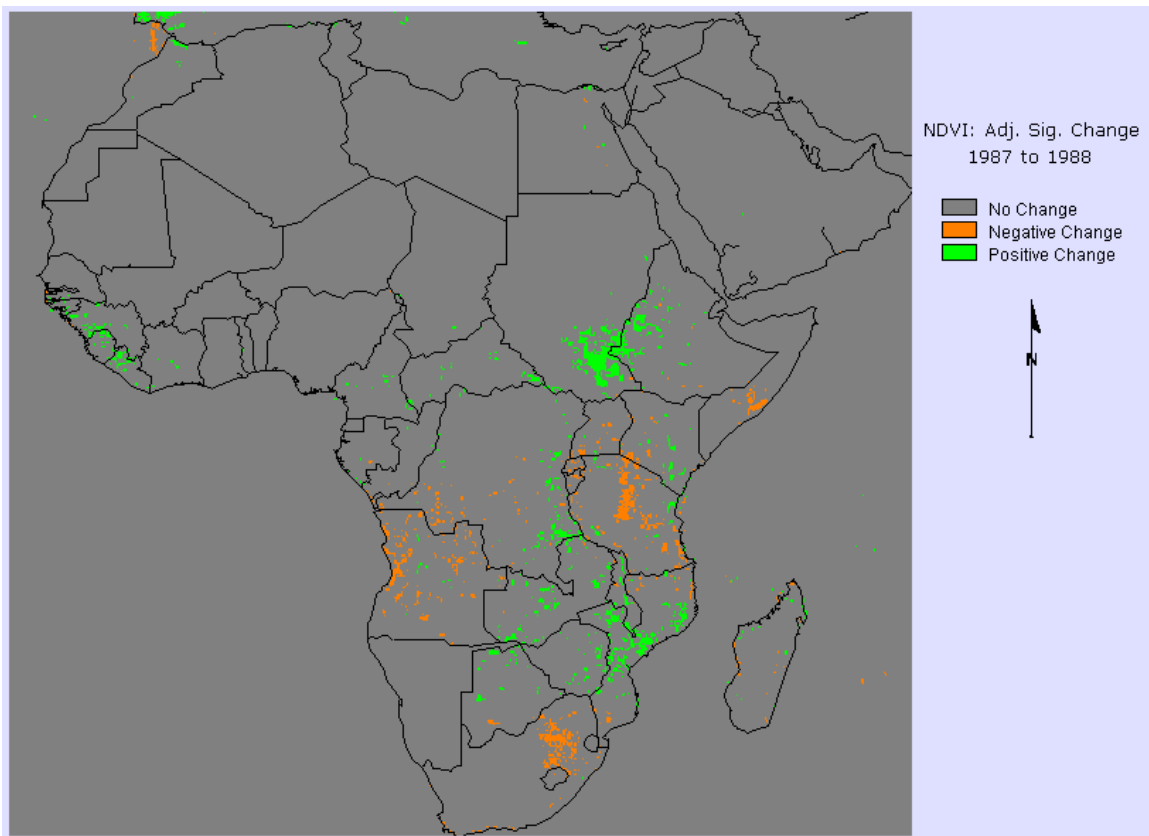
figures 15 and 16 below:

Figure 15: Adjusted Significant Change in NVDI 1987 to 1988

The adjusted change map of Africa (figure 15) shows a great deal more significant

negative NDVI changes than in Figure 7, mostly in the areas already noted, but with a

much greater extent. If we make the assumption that the sensor used to detect the vegetation index was indeed operating differently at the two data points, than the regression adjustment has indeed allowed us to get a more balanced picture of both significant positive and negative change.
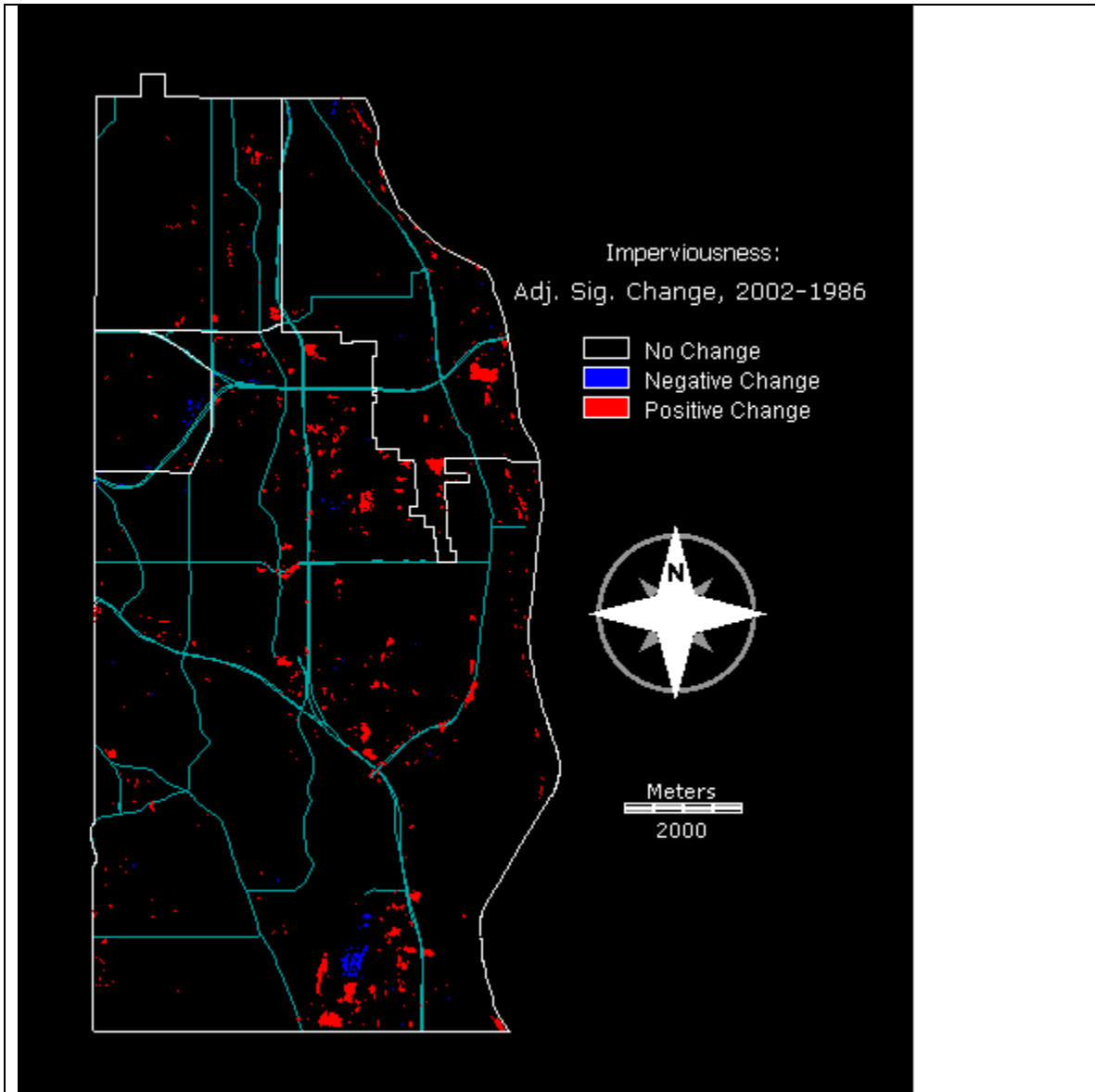

Figure 16: Adjusted Significant Change in Impervious Surface Value, 1986 to 2002

In contrast, the adjusted change image for our study area's value is questionable, as the assumption used for the Africa data is probably false in this case. It is possible that the impervious surface value images were created in a way that already adjusted for sensor

differences (a lack of metadata prevents us from learning about the process used to generate these images). In this case, the regression adjustment would be superfluous. The resulting image shows very little in the way of significant negative change. Although this might be expected in the context of a quickly developing suburban area, in this case it might not represent reality. It appears that the adjustment has just served to further restrict which areas of positive change are considered significant, all of which occur in the same areas noted earlier.

Exercise 4: Image Ratioing

The fourth exercise focuses on the use of image ratioing rather than image differencing in order to detect large relative changes instead of large absolute changes. In other words, changes from a value to zero to 10 could be seen as more significant than a change from 90 to 100. For this exercise, the data has been switched to larger scale images of the Rosso area of Mauritania in northwestern Africa at about the same scale as the images used for our study area. An obvious issue with ratioing is how zeroes in a given set of data are to be dealt with, as dividing by zero is usually ground for error in a mathematical operation. This will be especially important to consider for our impervious surface data, which is made up mostly of pixels with a value of 0. Luckily, this is the first issue the workbook covers in this exercise.

The NDVI values of the Mauritania data used in the workbook include negative and zero values, caused by the arid nature of the region. To deal with these values, the first step in exercise 4 is to use the RECLASS module with the 1977 and 1979 Mauritania NDVI images to give all negative and zero value data points a new value of .01, which essentially has the meaning of "no vegetation" (see figures 17 and 18 below for the adjusted NDVI images of Rosso, Mauritania).
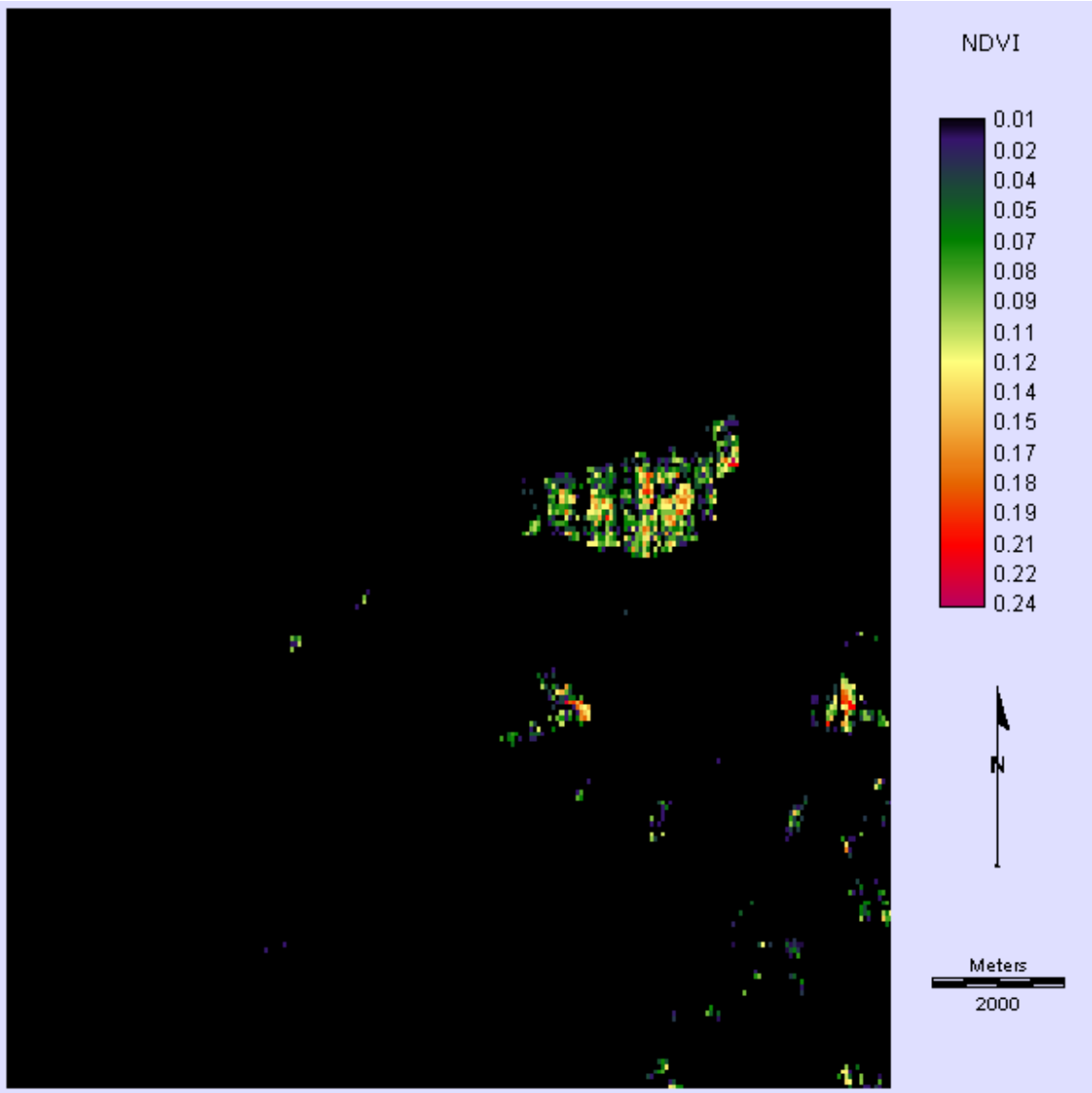
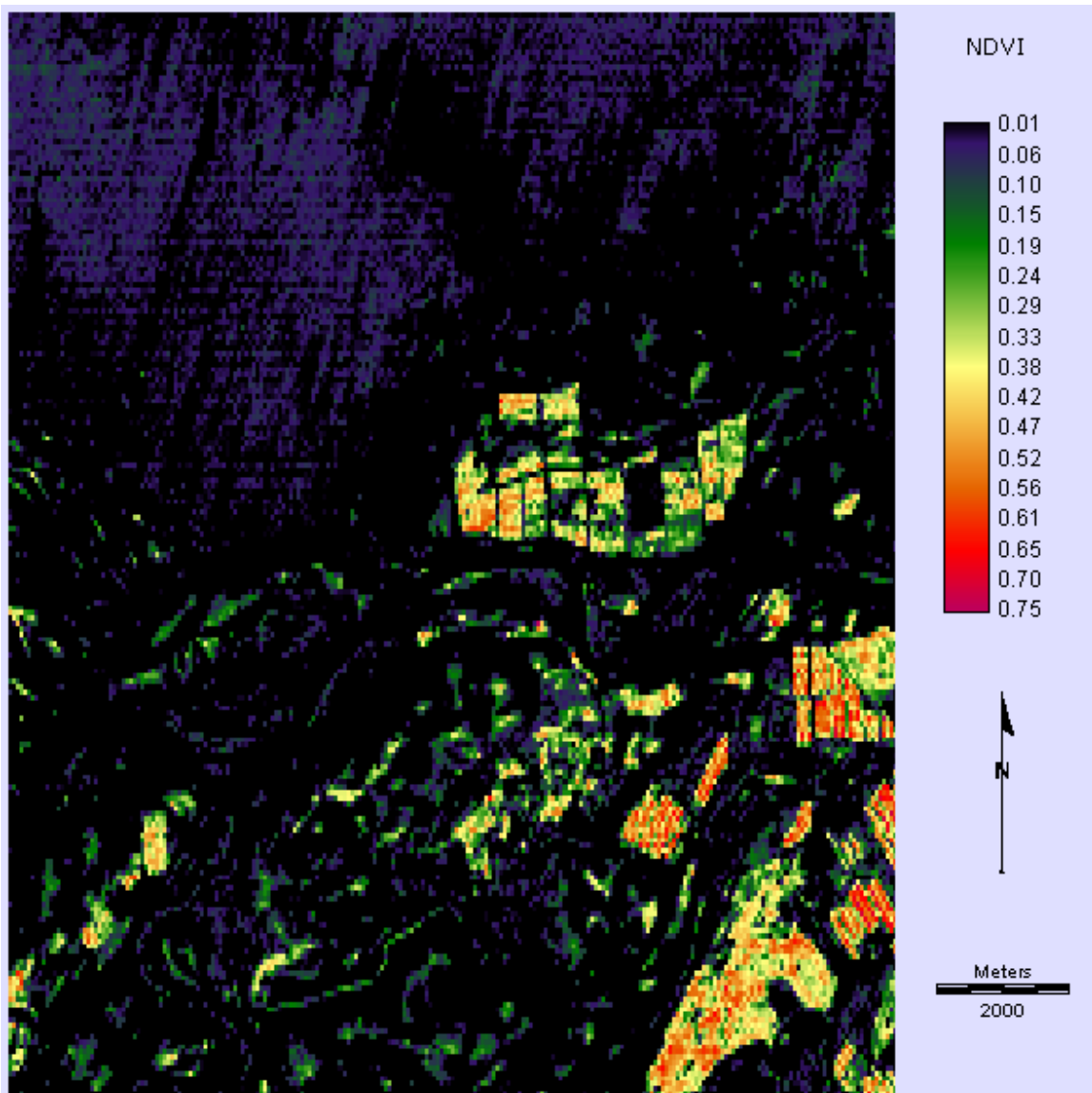**Figure 17: Adjusted NDVI Image of Rosso, Mauritania - 1977**

Figure 18: Adjusted NDVI Image of Rosso, Mauritania - 1979

However, for our data, .01 would be an inappropriate value to replace all of our

zero value data points. The imperviousness values in our data are all integers between 1

and 100. If a location has a value of 0 in 1986, and a value of 100 in the 2002 image, than

the resulting adjusted ratio value would be 10,000 (100 / .01)! By using 1 instead of .01

for our zero values, the resulting ratio image will have values ranging from 0 to 100, the

same scale already in use by our data. If a data point has a value of 1 in 1986, and again

in 2002, the ratio will be 1, i.e. no change (no figures are given of the adjusted 1986 and

2002 study area images, as they appear nearly identical to figures 3 and 4 above).

After we have our adjusted images, we are ready to start the ratioing process. In the exercise, this involves two steps. The first step is using an OVERLAY operation to produce a simple ratio image. The second step is to use a TRANSFORM operation to transform the resulting image into a logarithmic scale. The second step is necessary because the first image produced by OVERLAY will not be symmetric about zero, and will thus be difficult to interpret. For example, a large change from 25 to 50 between the 1986 and 2002 impervious surface value images would result in a ratio of 2. However, if the opposite were to occur, the ratio would be .5. In both cases, the change is a factor of 2. However, the scale does not allow us to see if the changes are negative or positive. Transforming to a logarithmic scale would make the doubling between 1986 and 2002 a positive value, while the halving would be a negative value, symmetric around zero, allowing us to compare positive and negative changes on a much more intuitive linear scale.
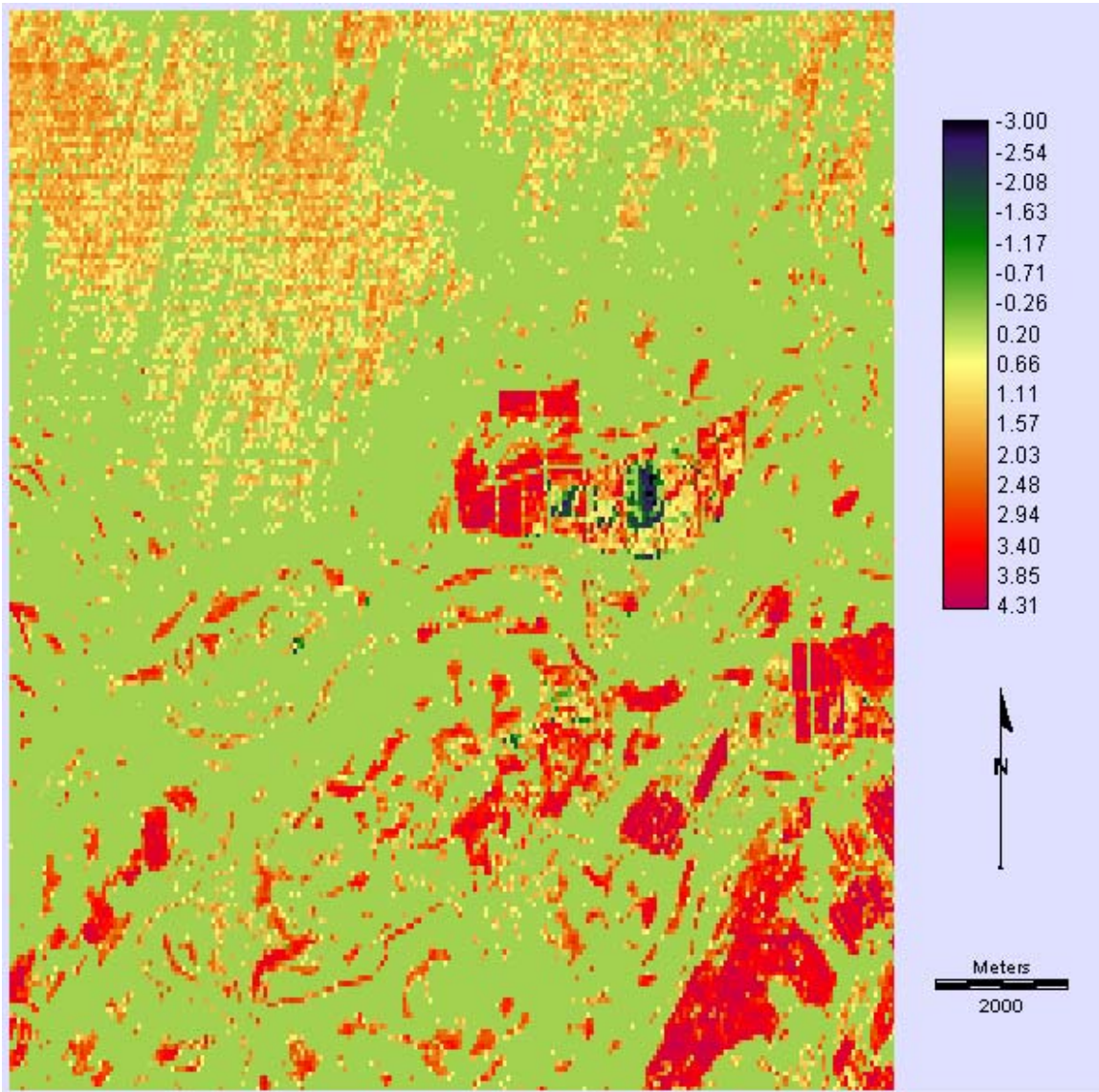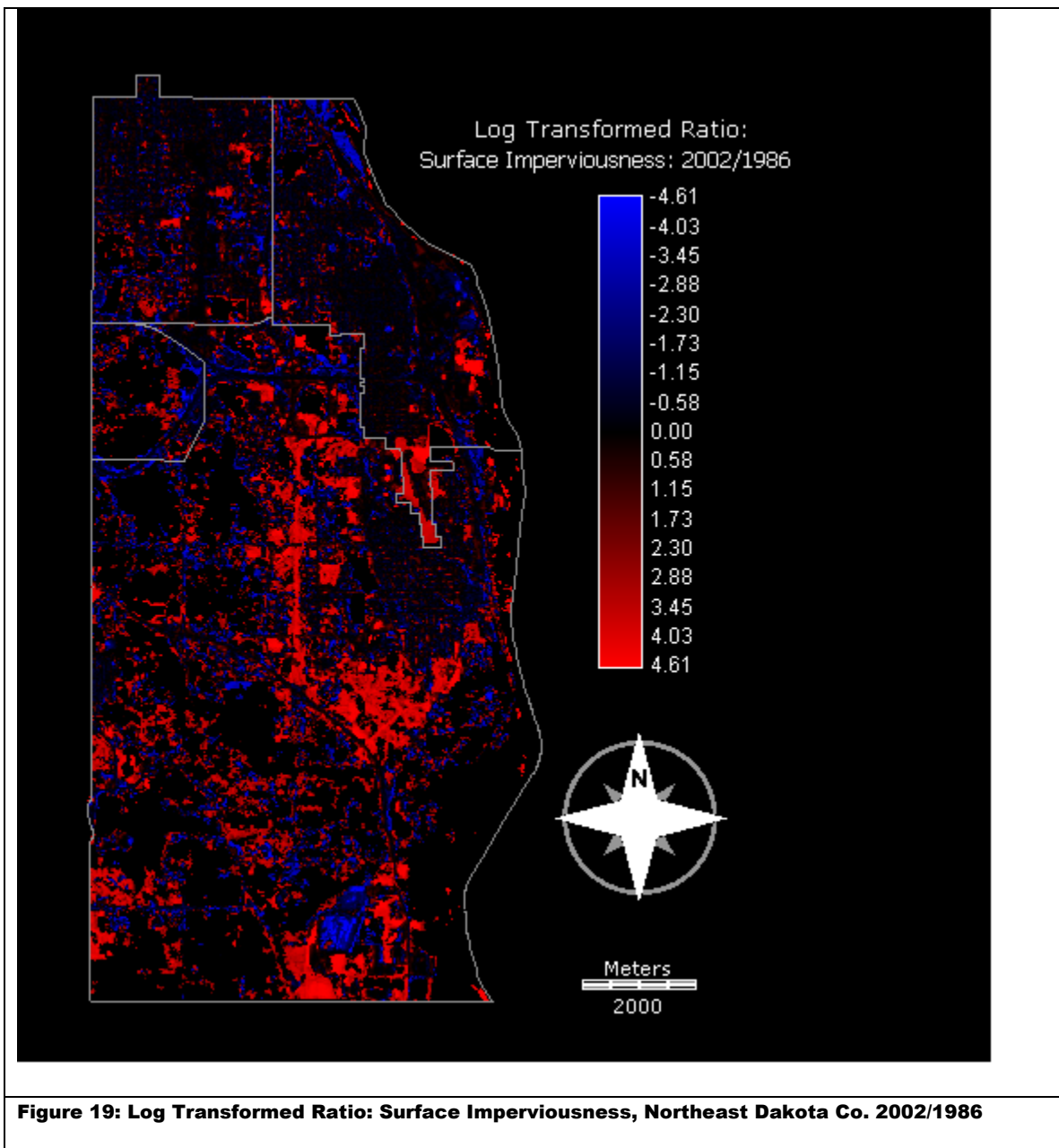
Figure 19: Logarithmic Transformed Ratio, NDVI - Rosso, Mauritania 1979/1977

Figure 19: Log Transformed Ratio: Surface Imperviousness, Northeast Dakota Co. 2002/1986

Figures 19 and 20 above show the final results of the OVERLAY and

TRANSFORM operations for the Mauritania and northeast Dakota County study area

respectively. The Mauritania image values range from about -3 to 4.31. The range of

values for our local study area is similar, ranging from -4.6 to 4.6. Positive values

indicate an increase in value over time, while the negative values indicate a decrease. In

comparison to our earlier difference maps for the northern Dakota County study area, the

log transformed ratio result seems much more cluttered, showing a great deal more of both positive and negative change. This should be expected, since the ratio technique is designed to show relative change between two images instead of absolute change, allowing areas of change not as prevalent to be brought to the forefront. Especially striking in Figure 19 is the area of increased imperviousness in around the new freeway corridor, which shows much more clearly on this image than in any of the earlier images.

The final step in the fourth exercise is to again use three standard deviations from the mean as the threshold for real change. The exercise also calls for the use of two standard deviations for comparison. Table 3 (below) gives the pertinent values.

Table 3: Statistics for Log Transformed Ratio Images

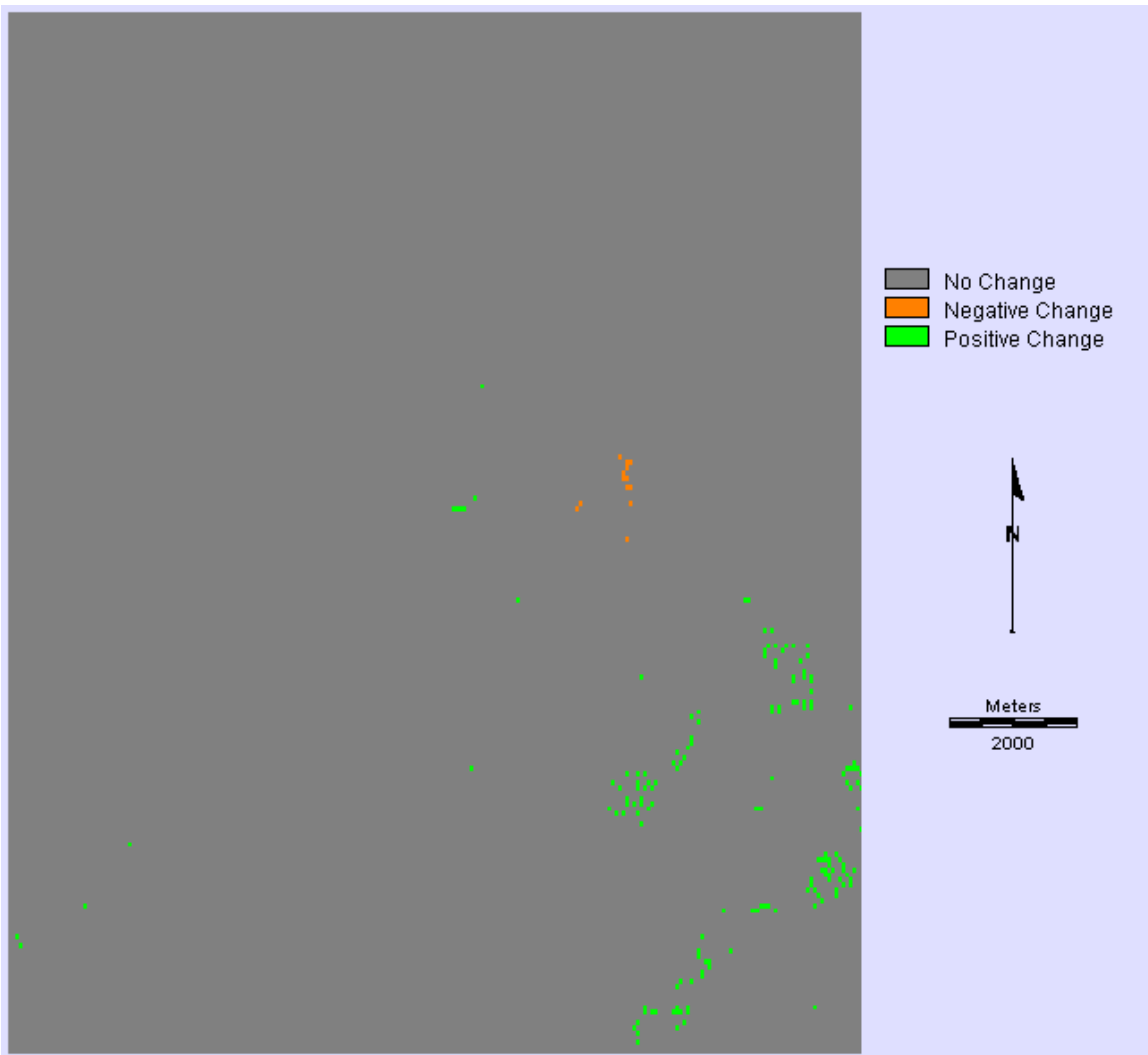| Data	Mean	Std. Deviation	+/- Threshold Value (3σ)	+/- Threshold Value (2σ)
Mauritania NDVI	.685	1.126	-2.693 / 4.063	-1.567 / 2.937
Local Impervious Surface Data	.316	1.511	-4.217 / 4.849	-2.706 / 3.338

Figure 20: Significant Change, from Logarithmic Ratio of NDVI (3σ) - Rosso, Mauritania 1979/1977

Our first result for the Mauritania data (Figure 20, above) using three standard deviations as the threshold shows that there are very few areas that meet our criteria for change, especially negative change. Loosening our thresholds to two standard deviations results in a dramatically different image (see figure 21 below), with a great deal more positive change. Negative change remains very small.
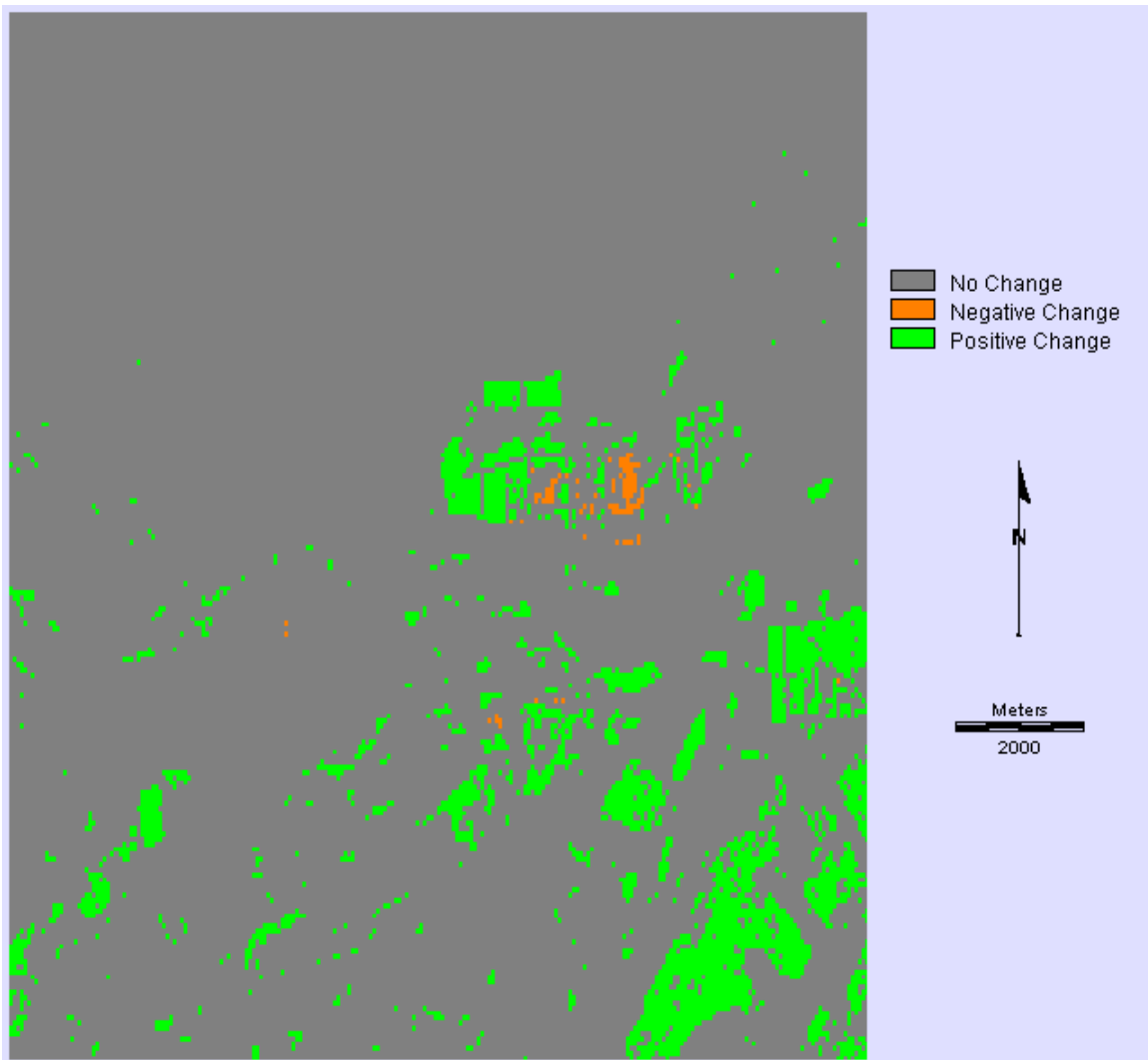
Figure 21: Significant Change, from Logarithmic Ratio of NDVI (2σ) - Rosso, Mauritania 1979/1977
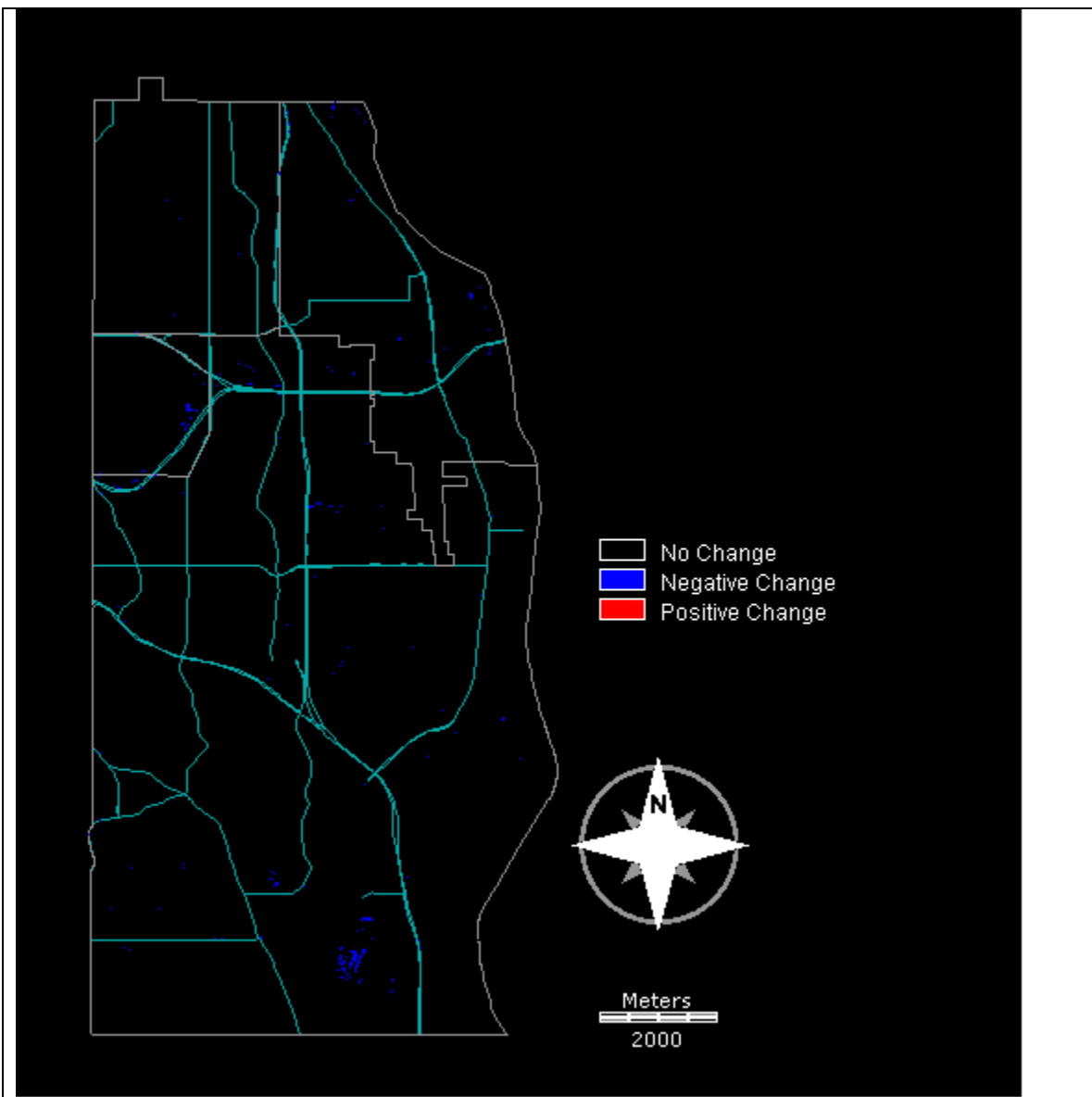
Figure 22: Significant Change, from Logarithmic Ratio: Surface Imperviousness (3σ), Northeast Dakota Co. 2002/1986

Our results using three standard deviations as a threshold for our local study area are even more dramatic (see Figure 22, above). Here, no positive change areas have met our criteria, and only a few areas of negative change have met our standard, again along the 494 corridor and at the landfill in southern Inver Grove Heights. This removes any data of interest to us, and also serves to show how the use of an arbitrary threshold can render our analysis effectively useless. Reducing the threshold to two standard deviations results in a very different image. This image shows more areas of change than any of the other

methods that we have tried in the first three exercises. Areas of positive change again

seem clustered along the southern end of the Lafayette Freeway, and around the south

end of Concord Boulevard. The areas of negative change, however, seem scattered, with

little pattern, other than to say that they seem prevalent in areas where not much positive
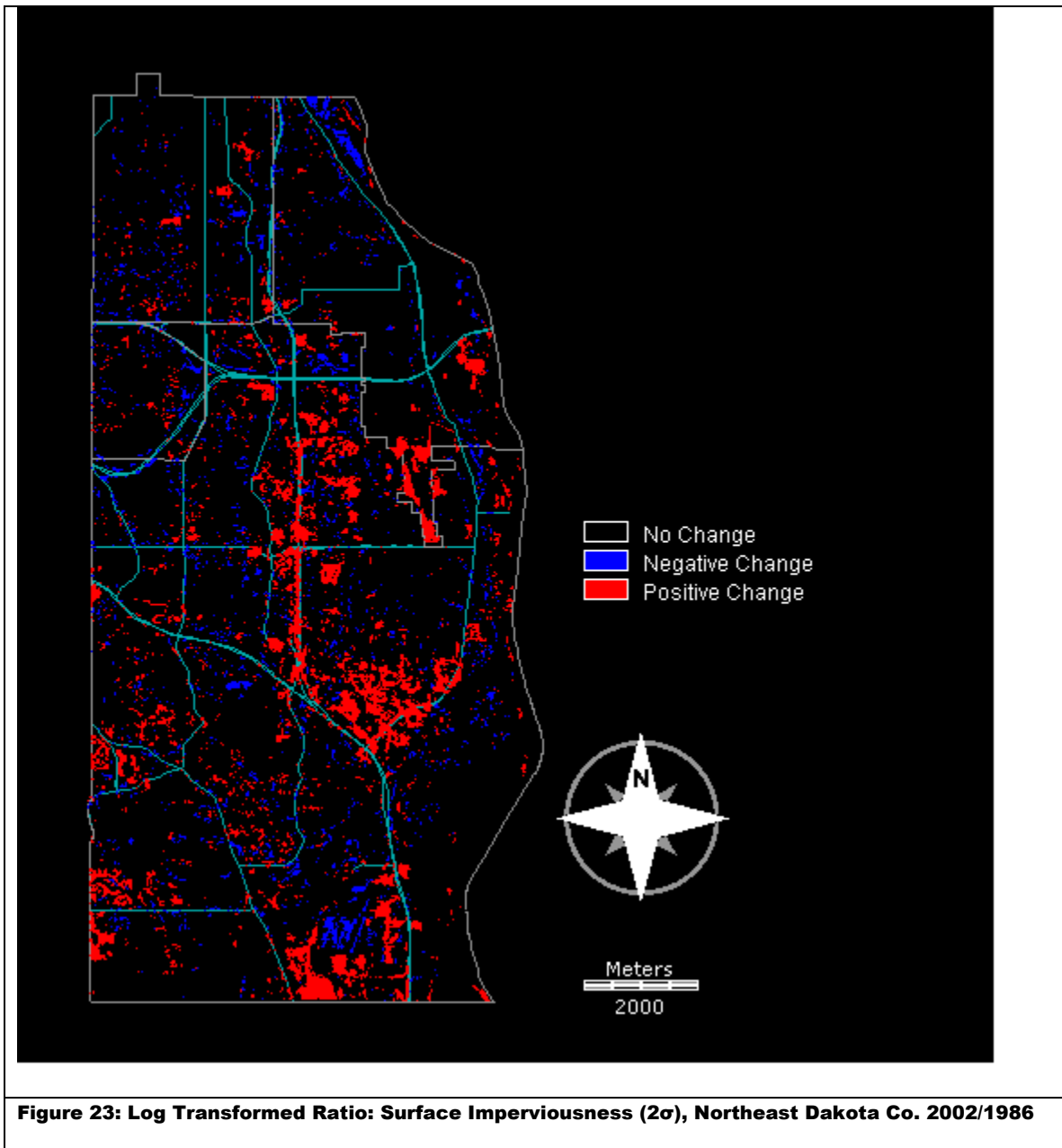
change is occurring.

Figure 23: Log Transformed Ratio: Surface Imperviousness (2σ), Northeast Dakota Co. 2002/1986

Conclusions:

Overall, the techniques introduced for pairwise image comparison seem useful when used with the pre-processed data provided with the UNITAR notebook for finding patterns of change. However, if these techniques are to be used in other contexts, much more knowledge of the data and local conditions on the ground would be required in order to determine better thresholds for detecting real change from normal local variability. This could be an entire avenue of investigation by itself.

For the local data used in the preceding exercises, the simple differencing technique in exercise 1 seemed the most useful, allowing us to at least see the general trends in surface imperviousness for our study area, and by proxy, development. The results of the other exercises, which put a lot of stock into the threshold values, could be seen as questionable because of the arbitrary use of an abstract statistical rule for their determination.

Bibliography:

Eastman, J. Ronald, Jean E. McKendry, and Michele A. Fulk. United Nations Institute for Training and Research Explorations in Geographic Information Systems Technology Volume 1: Change and Time Series Analysis. Clark Labs, 2007.

Minnesota Department of Transportation. Construction Project Logs, Dakota County. http://www.dot.state.mn.us/roadway/data/reports/projectlog.html