Marion, IA Case Study: Change Detection of Tree Canopy Before and After the August 2020 Midwest Derecho Using Vegetation Indices with Corroboration from Lidar-Derived Canopy Estimates

Michael Aitchison

University of Iowa

Introduction to Environmental Remote Sensing

Professor Linderman

December 17, 2021

Table of Contents

Introduction....3
Data Sources....4
Methodology....6
Results....8
Conclusions....8
References....9
Appendix A: Examples of *Trees From LiDAR Tool and Sample Data* Point Outputs, Compared to NAIP Imagery....11
Appendix B: Histogram of Short- and Long-term Image Differences, NDWI and SAVI....12
Appendix C: Short-term Highlighted Change Map, 1 s.d. Change Threshold....13
Appendix D: Long-term Highlighted Change Map, 1 s.d. Change Threshold....14
Appendix E: Short-term Highlighted Change Map, 2 s.d. Change Threshold....15
Appendix F: Long-term Highlighted Change Map, 2 s.d. Change Threshold....16
Appendix G: Side-by-side Comparison, Short-term Highlighted Decreasing Change at a 2 s.d. Change Threshold (top) vs. Overlap with Identified Tree Canopy Locations (bottom)....17
Appendix H: Side-by-side Comparison, Long-term Highlighted Decreasing Change at a 2 s.d. Change Threshold (top) vs. Overlap with Identified Tree Canopy Locations (bottom)....18

Introduction

On August 10 and 11 of 2020, a derecho (a storm with strong straight-line winds comparable to a hurricane) struck the midwestern United States. While the effects of the storm were felt throughout Nebraska, Iowa, and Illinois, eastern Iowa was among the hardest-hit regions, with the Cedar Rapids metropolitan area suffering from particularly severe damage resulting from gusts up to a reported 140 mph. Tens of thousands of Iowans were without electrical power for days after the event, and hundreds of thousands of trees throughout the metropolitan area were either severely damaged or felled by the strong winds. This project aims to visualize the spatial distribution of tree canopy damage in Marion, IA (a major part of the Cedar Rapids metropolitan area) in both the short and long terms through pre-classification change detection procedures applied to remotely sensed multispectral data. Methods such as these have long been employed by urban foresters and other interested stakeholders to estimate urban forest extents and changes (Walton et al., 2008). This change will be corroborated with lidar data, high-resolution aerial true-color imagery, and ancillary datasets used to confirm the spatial distribution of tree canopy cover within the study region.

Change detection is the measure of change in land cover or land use between two or more instances in time. The purpose of such a procedure is typically to evaluate the magnitude and/or direction of change in biophysical features (e.g. physical extent of rainforest ecoystem or urban sprawl) resulting from some process being studied (e.g. deforestation to make room for agricultural land or influx of new population from immigration or changing birth/death rates). In this case, the change detection is being run on the biophysical characteristics of tree canopy coverage resulting from the 2020 derecho. There are two overall types of change detection procedures: pre-classification and post-classification.

In post-classification change detection, images from each instant in time are classified using the same land cover scheme into a number of discrete land categories. The two (or more) classifications are compared to highlight areas which stayed consistent and areas which changed land covers. This method has the disadvantage of potentially greater rates of inaccuracy as classification errors compound between classified images, but has the advantage of discerning qualitative change. This method was not applied in this example because the goal is simply to measure the change in vegetation cover (specifically tree canopy cover) over time. While it is possible that a post-classification procedure would reveal differences in land cover classes between tree canopy and non-tree vegetation, a pre-classification procedure on one or more vegetation indices verified through other means accomplishes a similar goal, without the aforementioned reduction in land cover class accuracy that is probable in a post-classification procedure.

In pre-classification procedures such as image differencing, a simple difference of pixel values is calculated between corresponding bands or indices from different images, in order to measure the change from one instant in time to the next. This method has a tendency to be more accurate than post-classification change, as errors do not tend to compound between classified images. However, qualitative aspects of the land change cover (i.e. which land cover classes tended to transition over time, and into which other types) cannot be determined; in the case of a vegetation index, it can only be concluded that the value measured by the index (e.g. biomass) has increased or decreased. Because land cover classes within the realm of vegetation can’t be easily discerned through pre-classification change detection (e.g. differentiating tree canopy coverage, grass coverage, or crop coverage), it was necessary in this case to verify the relationship between tree canopy spatial distribution and vegetation index change through other means.

Vegetation indices are calculated using spectral reflectance values estimated from both the red visible EM spectrum and the near infrared EM spectrum, used to measure the biomass or vegetative vigor within the study area. There are multiple variations upon the basic vegetation index model to choose from. The normalized difference water index (or NDWI) is one common choice. While not a "vegetation" index per se, it is useful in measuring changes in plant water content, which can then be compared against another vegetation index to provide additional context; it is also said to be less sensitive to atmospheric scattering effects than other indices such as the normalized difference vegetation index, or NDVI (*NDWI: Normalized difference water index*, 2011). It is calculated by subtracting the reflectance value in the short-wave infrared spectrum from the reflectance value in the near infrared spectrum, then dividing by the sum of both reflectance values (to normalize the index to a value between -1 and 1).

The soil-adjusted vegetation index (SAVI) is very similar to the most common vegetation index (NDVI) in that it is calculated by subtracting the reflectance value in the red visible spectrum from the reflectance value in the near infrared spectrum, then dividing by the sum of both reflectance values (to normalize the index to a value between -1 and 1); however, it also incorporates the use of a soil calibration factor (L) to minimize soil background influences resulting from first order soil-plant spectral interactions. A value of 0.5 for L has been found to minimize the soil brightness variations and significantly reduce the need for any additional calibration based on different soil types, and so L has here been assigned a value of 0.5. In general, the use of SAVI "can significantly reduce the error in estimation of vegetation coverage caused by soil brightness as compared with NDVI when percentage of vegetation cover is low", which is often the case in large portions of urban landscapes (Wang & Xu, 2009).

To detect changes in vegetation land cover (and ultimately tree canopy damage), Sentinel 2 multispectral satellite system data was used to visualize the change in vegetation within the study area. A pre-classification image differencing procedure was performed four times in total: two short-term differences were derived using the NDWI and SAVI, and two long-term differences were likewise derived using the same indices. Significant change within these image differences (modeled as both 1 standard deviation and 2 standard deviations beyond the average difference value for each temporal pair) were then classified, aggregated in raster format, and highlighted. The correlation between detected vegetation index change (particularly decreasing values in both vegetation indices) and existing tree canopy coverage was further supported through the use of lidar data, high-resolution aerial true-color NAIP imagery, and planimetric polygon spatial data (outlining building and pavement footprints in the Linn County, IA area). In prior research, tree canopy estimates from lidar data and vegetation indices have been found to have a statistically significant positive correlation, and proven useful in estimating urban tree inventories in situations where direct ground data collection is problematic due to private property rights (Klobucar et al., 2021).

Data Sources

Sentinel 2 Level-1C, Marion, IA
- 13-band Imagery
- Spatial reference: UTM Zone 15N, WGS 84
- Resolution: 10 – 60 meters (varies by band)
- Downloaded from: USGS Global Visualization Viewer (https://glovis.usgs.gov/)
- Date of Imagery: August 6, 2020

Sentinel 2 Level-1C, Marion, IA

13-band Imagery
Spatial reference: UTM Zone 15N, WGS 84
Resolution: 10 – 60 meters (varies by band)
Downloaded from: USGS Global Visualization Viewer (https://glovis.usgs.gov/)
Date of Imagery: August 16, 2020

Sentinel 2 Level-1C, Marion, IA

13-band Imagery
Spatial reference: UTM Zone 15N, WGS 84
Resolution: 10 – 60 meters (varies by band)
Downloaded from: USGS Global Visualization Viewer (https://glovis.usgs.gov/)
Date of Imagery: August 6, 2021

Polygon boundaries of Marion, IA

Extracted from *City Limit Boundaries for the State of Iowa* [geojson]
Spatial reference: WGS 84
Temporal extent: Published August 31, 2020; last updated November 17, 2021
Source: Iowa Department of Transportation
Downloaded from: https://data.iowadot.gov/datasets/IowaDOT::city-4/explore

The Iowa LiDAR Project - LAS File, Marion, IA

Tile ID: 06124650, 06124652, 06124654, 06124656, 06124658, 06124660, 06144650, 06144652, 06144654, 06144656, 06144658, 06144660, 06164650, 06164652, 06164654, 06164656, 06164658, 06164660, 06184650, 06184652, 06184654, 06184656, 06184658, 06184660, 06204650, 06204652, 06204654, 06204656, 06204658, 06204660, 06224650, 06224652, 06224654
Projection: UTM Zone 15N, NAD 83
Downloaded from: Iowa Lidar Mapping Project (http://www.geotree.uni.edu/lidar/)
Date Created: May 2009

Planimetric Polygon [Linn County, Iowa planimetric polygons located outside the City of Cedar Rapids boundaries]

Spatial reference: WGS 84
Temporal extent: Published August 18, 2021
Source: Linn County, Iowa GIS
Downloaded from: https://opendata-linncounty-gis.opendata.arcgis.com/datasets/linncounty-gis::planimetric-polygon/about

2019 National Agriculture Imagery Program (NAIP) 3-Band Imagery, Johnson County, IA

True Color (Red, Green, and Blue)
Projection: UTM Zone 15N, NAD 83
Resolution: 1 meter
Downloaded from: Iowa Geospatial Data (https://geodata.iowa.gov/pages/imagery)

These datasets were chosen based on several criteria. First, Sentinel 2 multispectral data was chosen for the study due to its relatively high spatial resolution compared to readily available alternatives (10 to 20 meters for the bands used in this exercise, compared to 30 for most Landsat bands). The date for the first image (August 6, 2020) was chosen due to its timing just four days before the derecho hit. The latter date (August 16, 2020) was similarly chosen due to its close timing after the derecho event (6 days). Keeping the selected imagery within the same month time span avoided the problem associated with seasonal changes in vegetation reflectance profiles. This was also the justification for the timing of the last set of Sentinel 2 imagery (August 6, 2021), which was chosen to be one full year after the first image from before the derecho, allowing for the comparison of dates from the same seasonal range on the anniversary of the event.

Data availability was also a factor, as cloud cover over the area of interest prevented other candidate imagery from being chosen. Were cloud cover not an obstacle, it might have been preferable to use imagery from somewhat later after the derecho event (e.g. August 28, 2020), as this would have allowed slightly more time to pass between the derecho and the image date, which might have shown more significant changes as dead vegetation showed altered reflectance profiles and was removed. However, later imagery also ran the risk of introducing significant variation from seasonal influences, rather than solely the change of interest from the impact of the storm.

The *City Limit Boundaries for the State of Iowa* GeoJSON file was downloaded and the polygon representing the city of Marion, IA was extracted to define the area of interest used to subset the original Sentinel 2 imagery sets to only the area within Marion's city limits. This ensured that the analysis was confined to the study area, and that extraneous data did not influence variables (such as the mean and standard deviation of detected vegetation index change).

The selected lidar data was chosen simply for its availability as the most recent freely accessible lidar dataset for the study area. While there are some limitations associated with lidar data collected nearly 11 years before the derecho event, it should be noted that the slow rate at which urban forests typically change (outside of major events like the derecho) means that the overall spatial distribution of trees within Marion, IA is likely to be quite similar from 2009 to 2020. The 3-band NAIP imagery chosen for analysis alongside the lidar data was selected from 2009 as well, in order to most closely match the lidar dataset. It would have been preferable to obtain 4-band NAIP imagery (including the near infrared band), but that was not available until 2010, with leaf-on 4-band NAIP imagery only available in 2011. Finally, the planimetric data was also selected based on availability. In truth, a set of data from 2009 (the same time frame as the lidar and NAIP imagery) would have been preferable to more directly line up with structures in existence at the time, but this was not available at the time of writing.

Methodology

Each Sentinel 2 image was first aggregated from individual JPEG 2000 files of each band (from each image date) into multispectral IMAGINE Image files using ERDAS Imagine's "Layer Stack" tool. A subset of each stacked image was then created using the Marion, IA boundaries polygon file (which was first projected into WGS84 UTM 15N to match the imagery reference systems). Each subset multispectral image was changed from unsigned 16-bit top-of-atmosphere (TOA) reflectance values (the format of each image upon download) to values between 0 and 1. This was accomplished by converting each 16-bit value to a float value and dividing by 10,000. Vegetation indices (NDWI and SAVI) were then calculated for each of the reflectance-converted stacked images according to the following formulas (where ρ is a reflectance value for a given pixel in a specified EM band):

$$NDWI = \frac{\rho_{nir} - \rho_{swir}}{\rho_{nir} + \rho_{swir}}$$

$$SAVI = \frac{(1+L)(\rho_{nir} - \rho_{red})}{\rho_{nir} + \rho_{red} + L} \quad where\, L = 0.5$$

Band 8a and 12 reflectances were used for the NDWI formula for the NIR and SWIR values, respectively; band 8 and 4 reflectances were used for the SAVI formula for the NIR and visible red values, respectively. A pre-classification image differencing procedure was used to depict change between time periods, and pixels with a high amount of change in either direction were classified using the ArcGIS Pro Reclassify tool. For NDWI changes, significantly negative amounts of change were classified as an integer value of 1, while significantly positive amounts of change were classified as an integer value of 2; for SAVI changes, significantly negative amounts of change were classified as an integer value of 10, while significantly positive amounts of change were classified as an integer value of 20.

These highlighted change rasters were then added together using the raster calculator, and the resulting values classified according to the combined directions of change of each vegetation index and highlighted by color. Specifically, a final output raster integer value of 11 (meaning a pixel where both NDWI and SAVI had changed significantly negatively) was highlighted in red, while an integer value of 22 (meaning a pixel where both NDWI and SAVI had changed significantly positively) was highlighted in green. Pixels where SAVI had changed significantly negatively while NDWI change significantly positively were highlighted in blue, while the inverse case was highlighted in purple; however, given the significant correlation between NDWI and SAVI, there were relatively few examples of these classifications in the output raster. All other pixel values were left unhighlighted within the final raster output to emphasize the locations of significant change. This was done for each combination of “before” and “after” imagery: from August 6, 2020 to August 16, 2020; and from August 6, 2020 to August 6, 2021 (calculating the change over 10 days as well as the change over a full year). Each case was also modeled using both a +/-1 standard deviation change threshold and a +/-2 standard deviation change threshold.

To check that these paired changes in NDWI and SAVI correlate with tree canopy rather than some other form of vegetation or land cover, an analysis using lidar data, NAIP 3-band imagery, and planimetric polygons was used to establish the overall locations of trees within the Marion, IA study area. For the sake of the analysis, it is assumed that the spatial distribution of urban trees in the study area is generally similar (though likely not identical) to the spatial distribution of trees on August 6, 2020 (the date of the Sentinel 2 imagery used as a “before-derecho” sample). An ArcGIS Pro toolbox – authored by Arthur Crawford, a Senior Content Product Engineer at Esri – was used to streamline this process (Crawford, 2015). In short, the tool works by using NAIP 4-band or 3-band imagery to compute a value similar to the NDVI to identify areas of vegetation. These locations of vegetation are then compared to known areas over a certain threshold above the estimated bare earth elevation at that location (derived from the lidar data), as well as ancillary data in the form of a polygonal layer of known building/pavement footprints. Identifying pixels with detected vegetation where lidar has indicated features significantly above ground level and not within a 2-meter buffer distance from a building footprint allows the tool to infer that this is likely the location of a tree. From there, the tool estimates the point location of the tree trunk as the point of local maxima where the tree has been identified, and estimates the average radius of the tree canopy around the trunk based on that trunk location estimate and surrounding estimated trunk locations.

It should be noted that this procedure is not a foolproof way to determine the exact location of tree trunks or the exact radius of tree canopies. In particular, prior literature indicates that the use of very high resolution (VHR) lidar to identify trees in a similar manner has a tendency to generate “false negative errors concentrated on small trees and false positive errors in private gardens" (Ardila et al., 2012). However, for the purposes of spatially mapping the overall geographical distribution of tree canopy coverage throughout an urban area it is quite effective. This fact is illustrated through some example outputs in Appendix A. Following the identification of tree trunk points and radius attribute information, a buffer was computed using the radius attribute field of each point to estimate the overall spatial extent of tree canopy from the NAIP imagery and lidar datasets. Lastly, to visualize the aproximate degree of spatial overlap between estimated tree canopy and derived vegetation index highlighted change, areas of the +/-2 standard deviation short- and long-term bivariate highlighted change maps containing decreases in both indices (NDWI and SAVI) were exported and converted to polygons, then selected by location based on whether they intersected with the previously derived tree point radius buffers and made into a new layer and map.

Results

The results of each change detection procedures can be found in Appendices C through F. The highlighted areas of significant positive or negative change have been superimposed over a basic greyscale basemap for improved reference to the Marion, IA community. Both the short term and long term highlighted change maps indicate substantial correlation between NDWI and SAVI, with very few pixels revealing one index increasing while the other one decreases. At a change threshold level of one standard deviation, a significant amount of change both positive and negative can be seen in both time frames, with decreasing change regions tending to be located towards the southwest of the study area while increasing change regions tending to be located towards the northeast. At a higher change detection threshold of two standard deviations it is evident that there is a great deal more decreasing change than increasing change over both time frames, with the greatest relative difference between increasing and decreasing change in the short-term highlighted change map.

Comparing areas of significant decreasing vegetation indices change in both time frames to the locations identified as tree canopy through the lidar and NAIP data, it becomes apparent that there is a great deal of overlap between significant change and tree canopy as of 2009 (Appendices G and H). However, the amount of similarity between highlighted decreasing change and identified 2009 tree canopy varies based on time frame. There is more overlap in the short-term highlighted change map than in the long-term highlighted change map (where one or two sizable regions of detected change do not appear to overlap with identified 2009 tree canopy).

Conclusions

As can be seen in Appendix B, each of the histograms for each change dataset has a negative mean value along with a significant negative skew, indicating that there is more negative change (decreasing vegetation indices) than positive change (increasing vegetation indices), particularly in the short-term change map. This intuitively makes sense – the derecho would naturally cause a greater than normal amount of decrease in vegetative vigor due to storm damage, particularly to trees. This trend is also reflected in the spatial distribution of change at a two standard deviation change threshold.

However, the amount of increasing highlighted change (green areas) in the map at a change detection threshold of one standard deviation is at first glance surprising, particularly in the short term. This could be related to a number of factors. As previously noted, increasing highlighted change is distributed towards the northeast portion of Marion; a visual examination of the imagery indicates that this area has much fewer trees than the southwestern portion of the city, and more grass and cropland. It could be that this difference in vegetation cover type caused an apparent increase in vegetation between the dates in question. It might also be possible that downed trees / tree limbs were either blown by the storm into this area, or that downed trees / tree limbs were removed and hauled off to other parts of town. This might cause them to appear as “increased vegetation” in the vegetation indices. Increased vegetation content in the long-term maps (both one and two standard deviation thresholds) makes more sense, as more time has passed allowing for replanting efforts to increase vegetation content.

It also makes intuitive sense that similarities between the identified 2009 tree canopy cover and the highlighted decreasing vegetation indices from the change maps would be greater in the short term than in the long term. It stands to reason that tree damage from the derecho would make up the majority of the significant negative vegetation index change immediately after the event. It also stands to reason that there would be less correlation between detected change and 2009 canopy cover in the long term, given that more time has passed and therefore more opportunities for vegetation change unrelated to either trees or the derecho have occurred. One additional interesting result to note is that the change in vegetation indices in the short term has relatively overlap with the change in vegetation indices in the long term. This may be related to a difference between trees which were immediately killed vs. trees which merely sustained damage; it’s possible that trees which were initially damaged took longer to show that damage in the form of being removed or losing significant amounts of their crowns.

On the whole, there is substantial evidence that using SAVI and NDWI image differencing procedures is a reasonably effective means to visualize change in tree canopy coverage following severe weather damage. This was confirmed by the substantial degree of overlap between the highlighted change and the derived tree canopy coverage from 2009. However, there are limits to this method – the lidar and NAIP imagery used to confirm the locations of tree canopy is over a decade older than the derecho event being studied here. With improved access to more recent data from just before the event, the location of trees prior to the storm would be known with more confidence. Other methods such as object-oriented classification might also have improved results in distinguishing tree canopy cover from other vegetation types.

References

Ardila, J. P., Bijker, W., Tolpekin, V. A., & Stein, A. (2012). Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images. International Journal of Applied Earth Observation and Geoinformation, 15, 57–69. https://doi.org/10.1016/j.jag.2011.06.005

Crawford, A. (2015) *Trees From LiDAR Tool and Sample Data* [GIS Tool]. Downloaded from: https://hub.arcgis.com/content/d280e1c023084992b98088d9f14d3252/about. Accessed: 2021 Dec 12.

Klobucar, B., Sang, N., & Randrup, T. B. (2021). Comparing ground and remotely sensed measurements of urban tree canopy in private residential property. Trees, Forests and People, 5. https://doi.org/10.1016/j.tfp.2021.100114

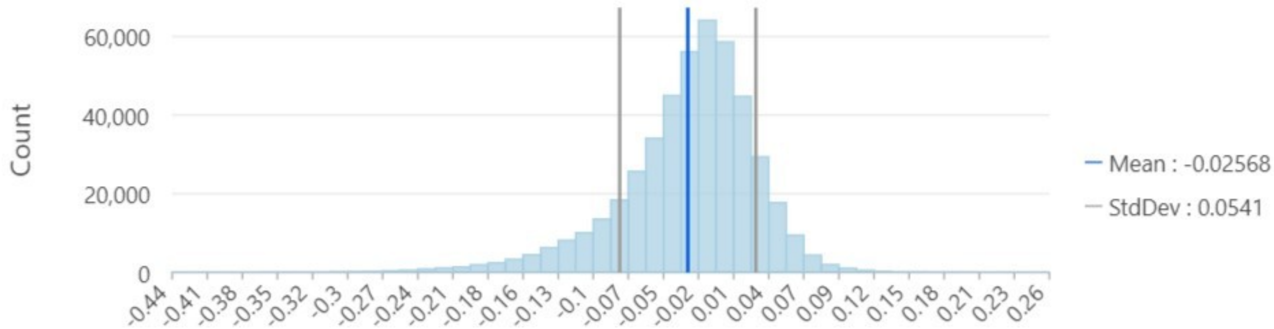
NDWI: Normalized difference water index. (2011). JRC European Commission. https://edo.jrc.ec.europa.eu/documents/factsheets/factsheet_ndwi.pdf

Walton, J., Nowak, D., & Greenfield, E. (2008). Assessing urban forest canopy cover using airborne or satellite imagery. Arboriculture & Urban Forestry, 34(6), 334–340. https://www.researchgate.net/publication/255635468_Assessing_Urban_Forest_Canopy_Cover_Using_Airborne_or_Satellite_Imagery

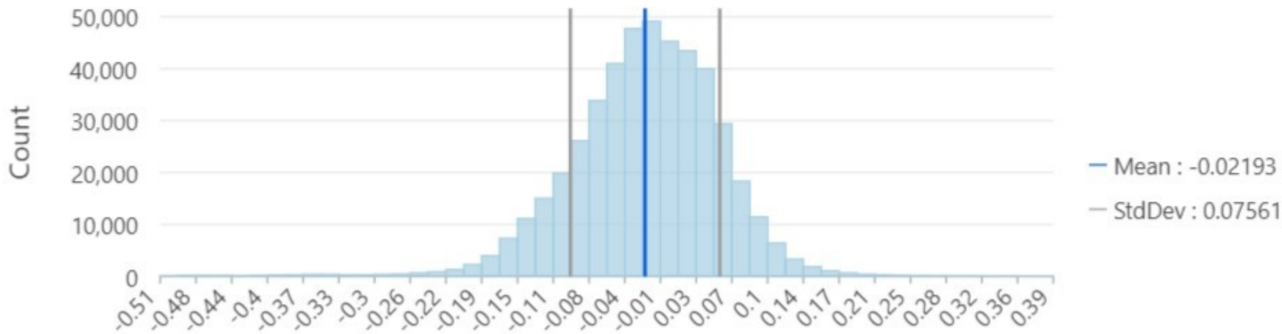
Wang, F., & Xu, Y. J. (2009). Comparison of remote sensing change detection techniques for assessing hurricane damage to forests. Environmental Monitoring and Assessment, 162(1-4), 311–326. https://doi.org/10.1007/s10661-009-0798-8

Appendix A: Examples of *Trees From LiDAR Tool and Sample Data* Point Outputs, Compared to NAIP Imagery

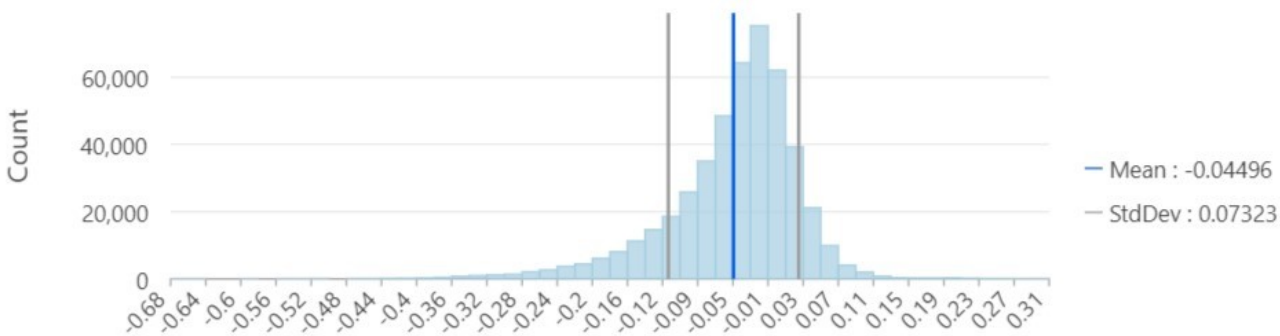
Appendix B: Histogram of Short- and Long-term Image Differences, NDWI and SAVI



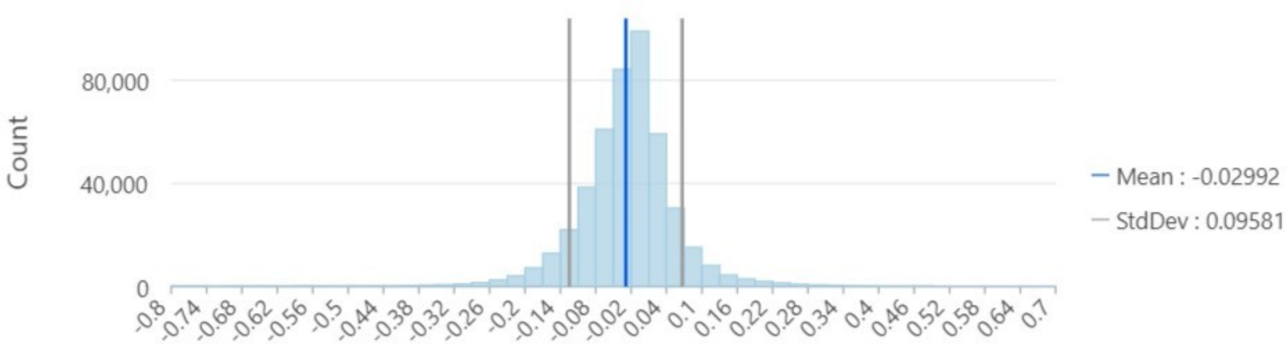
Short-term SAVI Image Difference Histogram



Long-term SAVI Image Difference Histogram

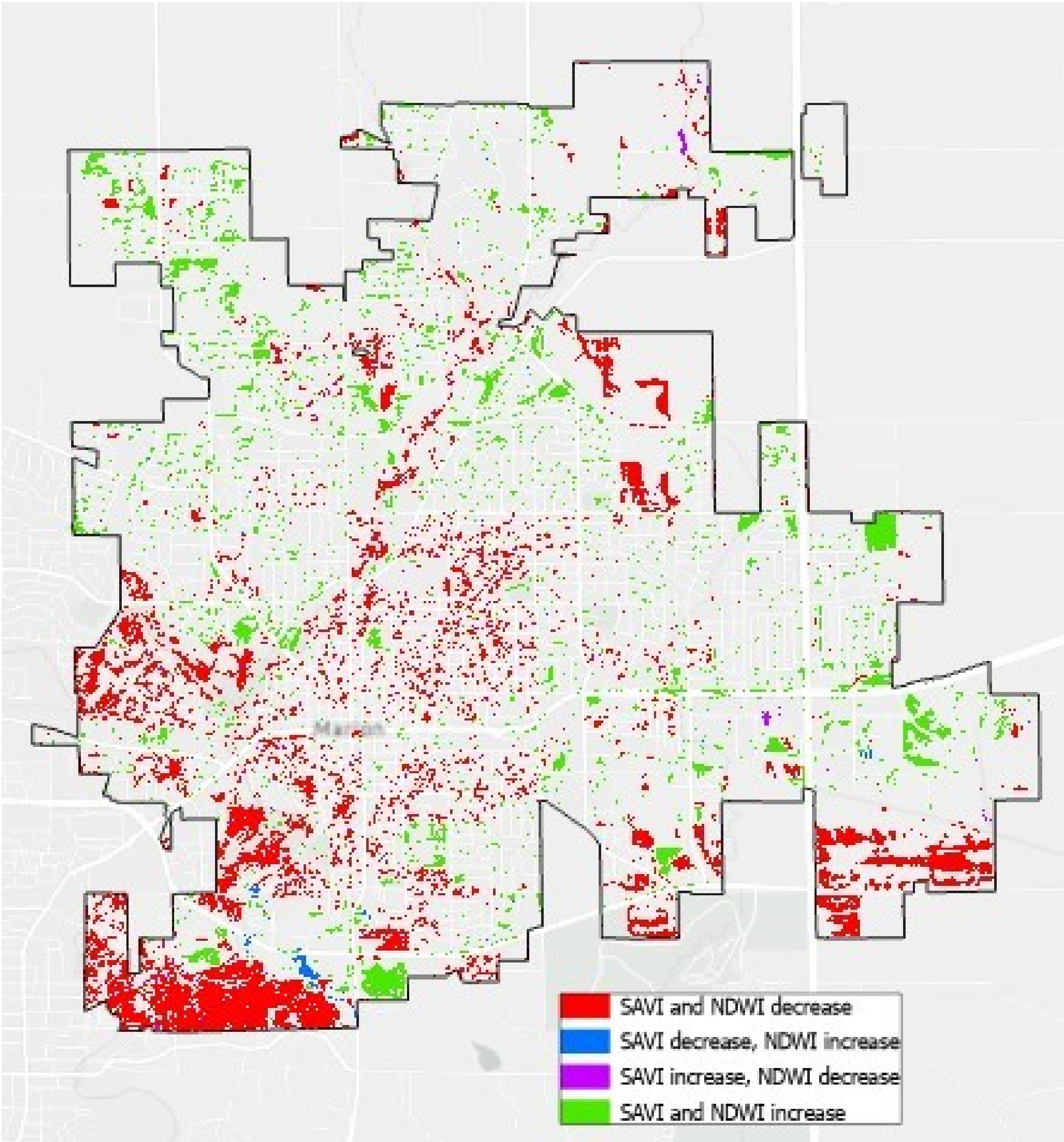


Short-term NDWI Image Difference Histogram

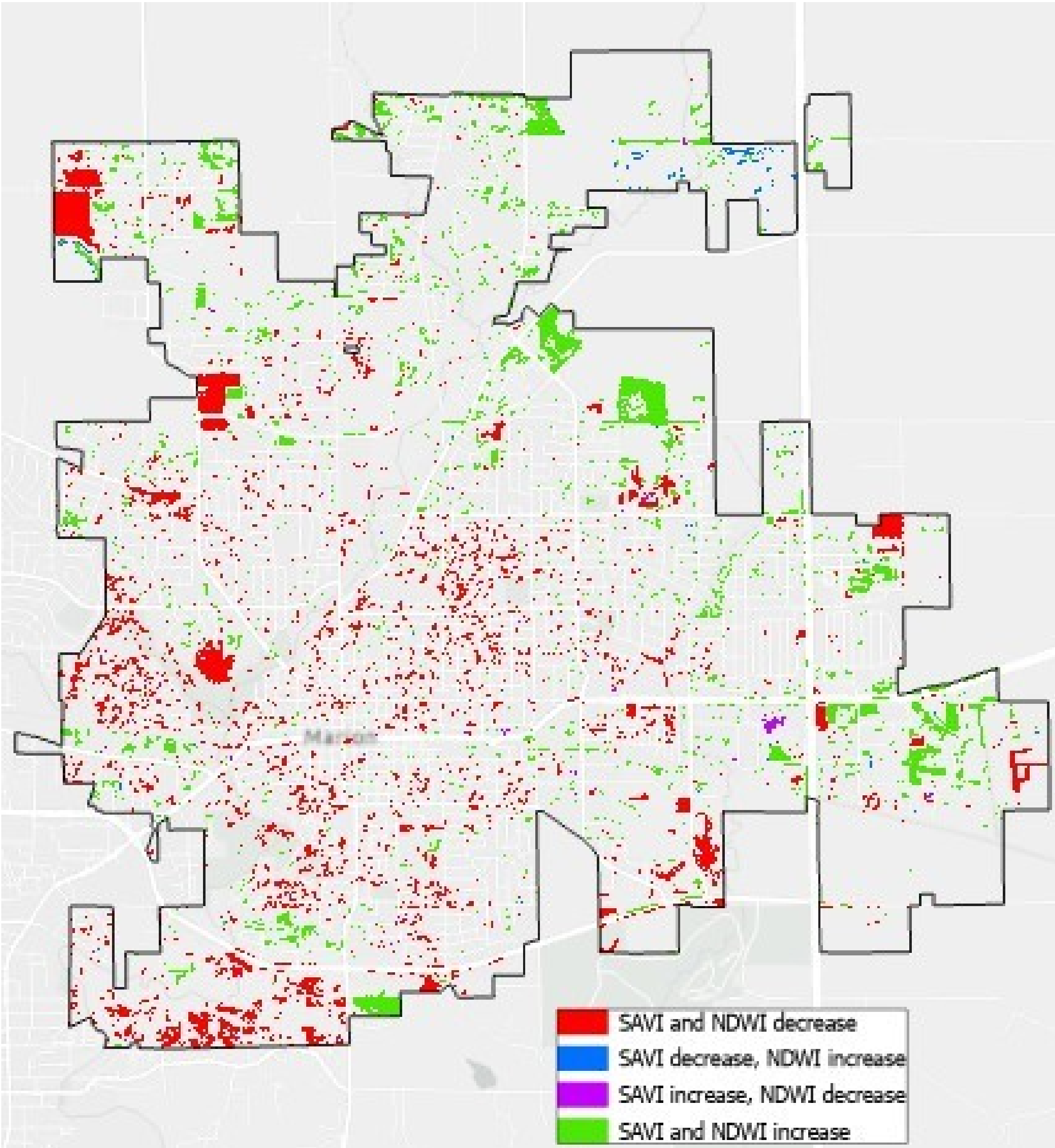


Long-term NDWI Image Difference Histogram

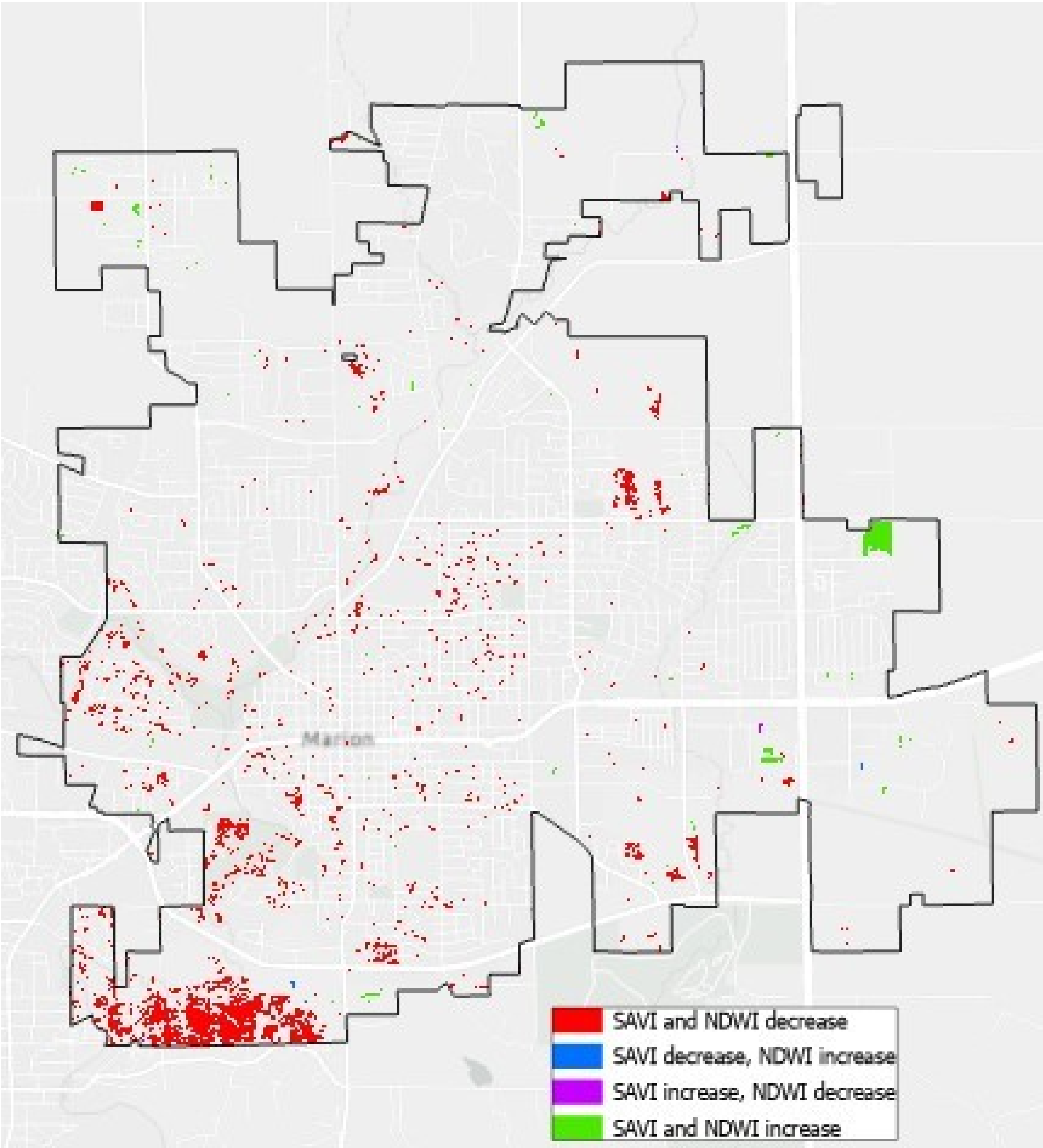
Appendix C: Short-term Highlighted Change Map, 1 s.d. Change Threshold



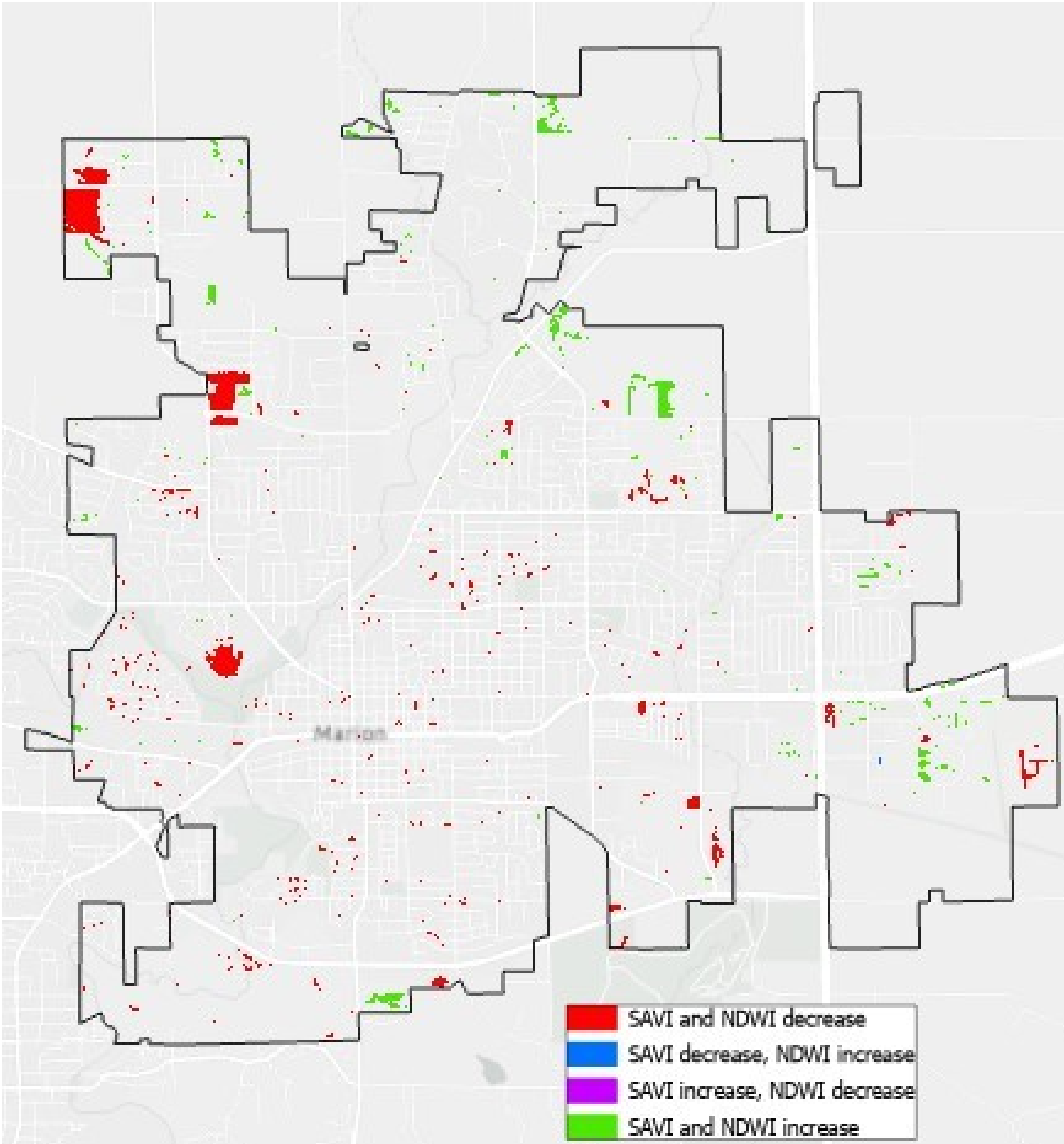
Appendix D: Long-term Highlighted Change Map, 1 s.d. Change Threshold



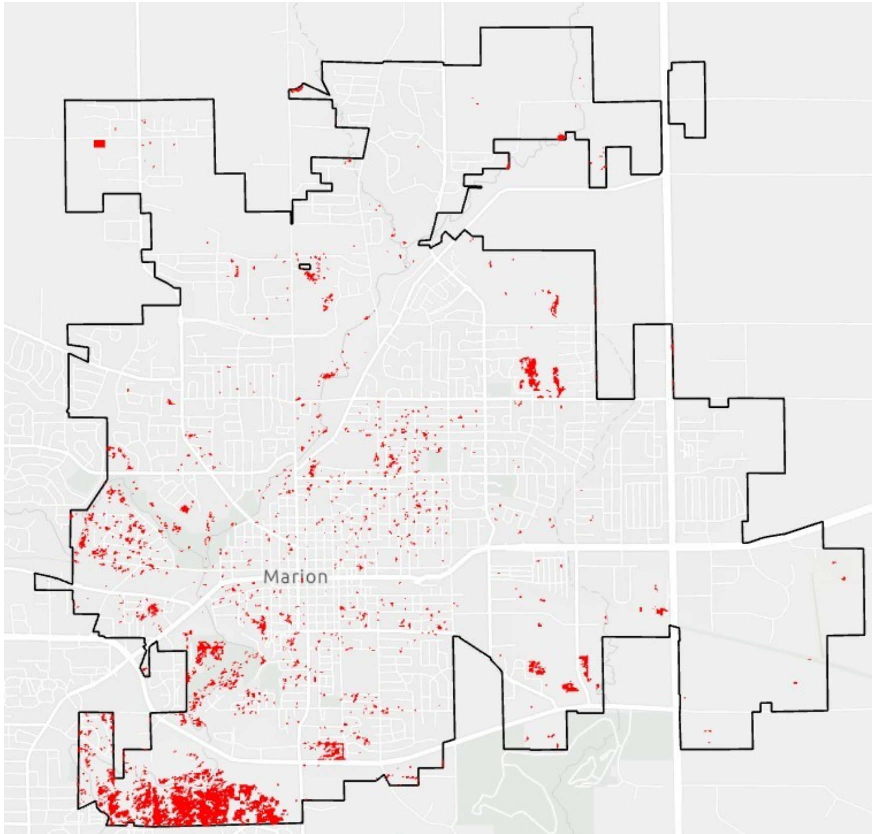
Appendix E: Short-term Highlighted Change Map, 2 s.d. Change Threshold

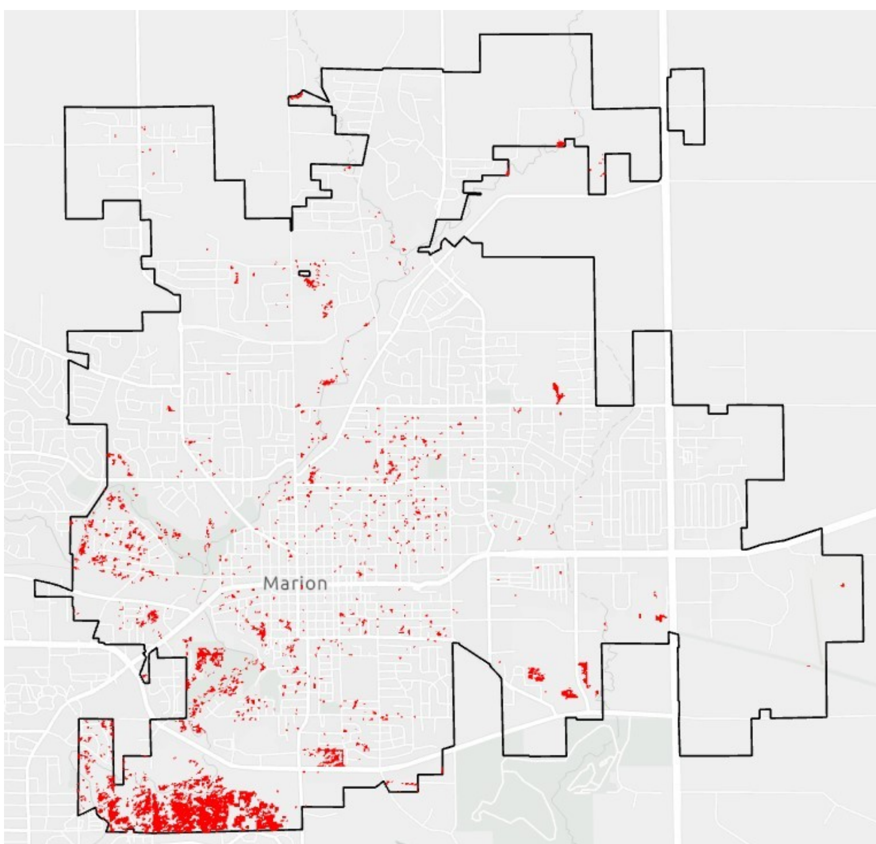


Appendix F: Long-term Highlighted Change Map, 2 s.d. Change Threshold



Appendix G: Side-by-side Comparison, Short-term Highlighted Decreasing Change at a 2 s.d. Change Threshold (top) vs. Overlap with Identified Tree Canopy Locations (bottom)





Appendix H: Side-by-side Comparison, Long-term Highlighted Decreasing Change at a 2 s.d. Change Threshold (top) vs. Overlap with Identified Tree Canopy Locations (bottom)