PennWest-Clarion

ASSESSMENT OF A DECADE OF URBAN EXPANSION IN MOON & MARSHALL TOWNSHIPS

Yasser Ayad
Professor, GIS
yayad@pennwest.edu

Abstract

In the state of Pennsylvania, Allegheny County has the second largest population after Philadelphia. For some time, the county's population has witnessed a steady and continuous decline. Not until the last 10 years that its population started to show some increase. According to the latest Census data of 2020, Allegheny County witnessed a 2.2% increase from 2010. Moon Township was one of the largest communities that had a significant population increase during this period (12.6%). On the other hand, Marshall Township, located in the north and shares a border with Butler county's Cranberry township, has almost doubled its population (45.8%) during the same period. In this study, we take a deeper look at the spatial changes that occurred during the last 10 years. Assessment of vegetation cover in general and tree cover in specific was conducted. However, despite the significant population growth of those areas, the results show a general increase in vegetation cover in both townships. The methods used in the current study adopted modified vegetation indices that were calculated from high resolution multispectral aerial images, as well as the extraction of tree cover from LiDAR point cloud data. The results show potentials and challenges of those methods. Finally, it is suggested that more structured accuracy assessment to be conducted on the presented results, and method refinements to include other hybrid techniques are to be investigated.

Contents

Introduction

In the state of Pennsylvania, Allegheny County has the second largest population after Philadelphia. For some time, the county's population has witnessed a steady and continuous decline. Not until the last 10 years that its population started to show some increase. According to the latest Census data of 2020, Allegheny County has witnessed a 2.2% increase from 2010. Moon Township, located west of the City of Pittsburgh and north of the Pittsburgh International Airport was one of the largest communities that had a significant population increase between 2010 and 2020 (12.6%). On the other hand, Marshall Township, which is located in the northern part of the County and just south of the expanding Cranberry Township in Butler County, has almost doubled its population (45.8%) in the same period (Roberts et al., 2021). Any population change would necessarily impact the overall socio-economic and environmental fabrics of the county.

Figure showing the location of the two municipalities under investigation in this study: Moon township and Marshall township.

This study investigates the population growth in those two municipalities and assesses the change in vegetation cover due to the likely urban expansion that occurred during the period under investigation using available high-resolution aerial imagery and Light Detection And Ranging (LiDAR) datasets. The following table summarizes the data used in this study and their sources:

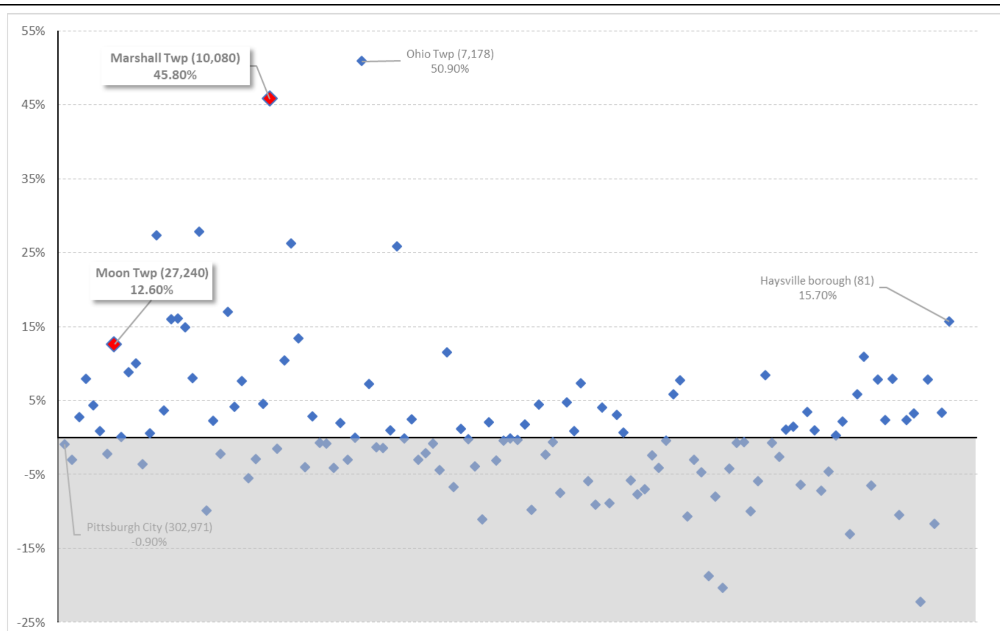
DATA TYPE	SOURCE	NOTES/DESCRIPTION
NON-SPATIAL		
Census Blocks (2010 & 2020) Census Population P1 (2010 & 2020) Census Municipality Boundaries Census County Boundaries	US Census Bureau	https://www.census.gov/
SPATIAL		
NAIP 2010	PASDA-USDA	- Spatial resolution for NAIP 2010: 1 meter
NAIP 2019	PASDA-USDA	- Spatial resolution for NAIP 2019: 0.6 meter
LiDAR LAS files 2006	PASDA-PAMAP	- LAS 2006 point spacing: 4.716 ft - 5.442 ft
LiDAR LAS files 2020	PASDA-USGS	- LAS 2020 point spacing: 1.203 ft - 1.77 ft
		- Data was acquired from: https://www.pasda.psu.edu/

It is important to note that there was no perfect match between the years where census data were available and those of spatial nature. For example, the LiDAR LAS files were first available in Pennsylvania starting 2005 (PAMAP program) and depending on the county location, LiDAR data collection ended in 2007. Similarly, but at a narrower range, the NAIP aerial imagery were available in multi-band in 2010 and 2019 with different spatial resolutions.

Population Growth

Marshall and Moon Townships were due to their population size and their percentage population growth during the past 10 years (Roberts et al., 2021). A scatter plot was created based on the data compiled by Roberts et al. (2021). The plot shows the percent change of population between 2010 and 2020 census. The distribution of all municipalities in Allegheny County were sorted by their population count in 2020. Therefore, Pittsburgh City, with the highest population in the group (302,971), falls to the far left of the graph and has a population decline of -0.90%, while Haysville Borough, with the lowest population (81), falls to the far right of the graph with a population growth of 15.70%.

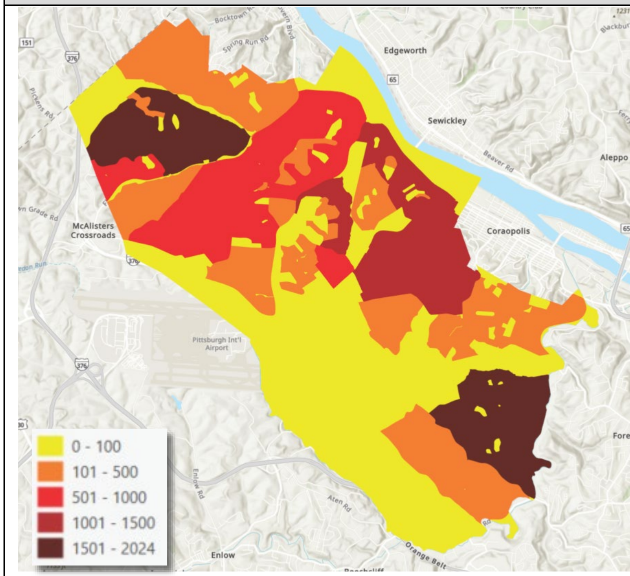
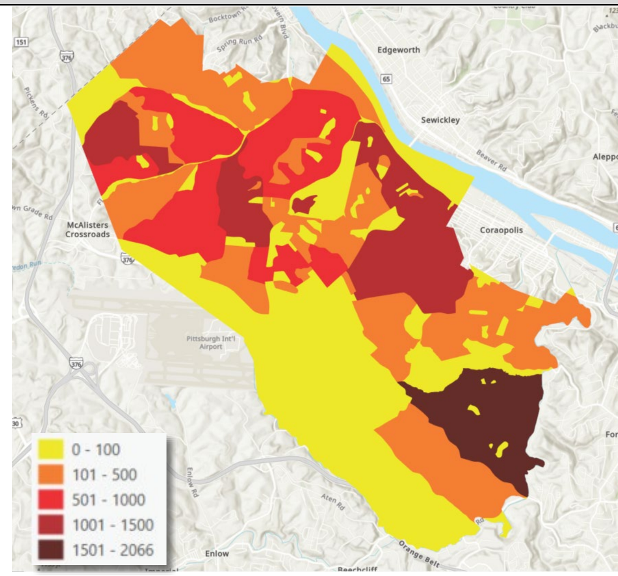
The population count for Moon Township in 2020 was 27,240 with 12.60% growth from 2010, and Marshall Township population was 10,080 in 2020 with a growth of 45.80%. Making them among the top townships in terms of population size and growth during the period under investigation.



Scatter plot showing the percent population change by municipality from 2010 to 2020 in Allegheny County. The data are sorted by population count where highest population counts in 2020 are located in the far left of the graph while the lowest are located in the far right. Moon Township and Marshall Township were selected for the current study and are highlighted in red.

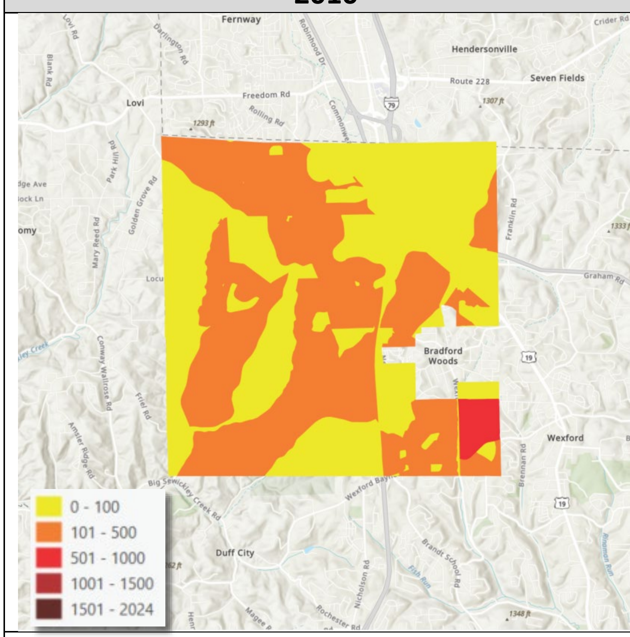
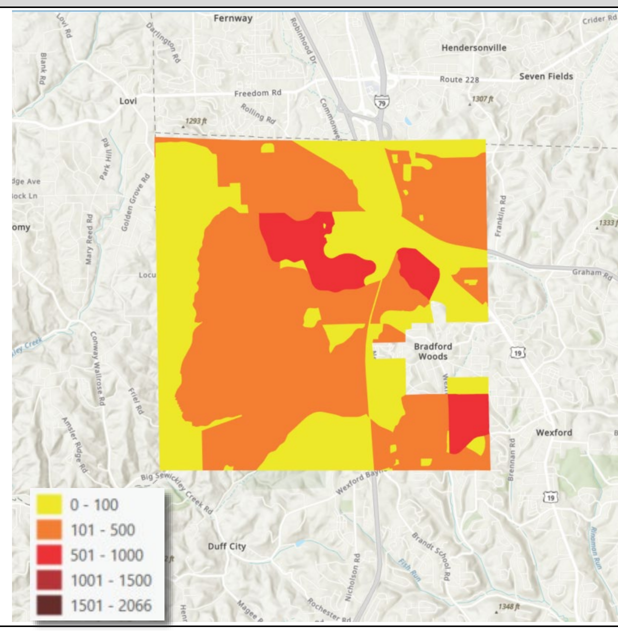
Even though some of the census block boundaries have changed between 2010 and 2020, the overall distribution of population change in both municipalities still demonstrates some growth patterns. The following maps were developed using census block Tiger line files and census block-level population data (P1) downloaded from the Census Bureau website (https://www.census.gov/).

MOON TOWNSHIP POPULATION CHANGE (2010-2020)

2010	2020
	

Moon township had more population counts than Marshall in both years. Some population changes are shown in the maps. Most notable the area to the south as well as the northwest borders of the township has witnessed the most changes in the municipality. It is important to note that some census boundary lines have also changed between the two dates which might have resulted in the aggregation or splitting of the values which might have affected class distribution on those maps.

MARSHALL TOWNSHIP POPULATION CHANGE (2010-2020)

2010	2020
	

Not as populated as Moon township, nevertheless, Marshall township witnessed some population change during the period between 2010 and 2020, especially at the center where areas that fell in the lowest class break (0-100) and the following one (101-500) have increased enough to be moved up one level each and be classified at a higher class (101-500 and 501-1000 respectively).

Detecting Spatial Change

Detecting changes in the spatial configuration would require an initial classification of temporal data. Accurate and equivalent classification for compared dates is essential in order to calculate reliable changes that would be useful in future decision making and planning processes.

Aerial images are becoming readily available at date intervals that would help compare the extent and magnitude of spatial changes. However, with the advent of high and very high-resolution imagery acquired using sensors that are mounted on either aircrafts of Unmanned Aircraft Systems (UAS), many technical challenges emerged. Among those are the spectral and spatial complexity those images are capable of recording. For example, shadows and cloud covers are becoming among the most challenging features to detect and correct. The main problem with shadows, whether caused by tree canopy, structures, or clouds, is the reduction or the total loss of information in an image (Zhang et al., 2015). Different methods have been developed to reduce the effects of shadows in images. Among those are image segmentation and using modified vegetation indices to detect the areas affected by shadows and therefore enabling either the isolation of those areas from analyses or identifying them for further correction procedures (Arévalo et al., 2008; Kwatra et al., 2012; Shahtahmassebi et al., 2013).

In this study, segmentation didn't lead to a visually accurate classification of neither of the NAIP images. Vegetation Indices (VI) on the other hand showed some promise as they started detecting some information that were previously masked by shadows. Three Vegetation Indices were tested in this study, namely: Normalized Difference Vegetation Index (NDVI), Triangular Vegetation Index (TVI), and Modified Soil Adjusted Vegetation Index (MSAVI). All three were presented by (Zhang et al., 2015) which concluded that a higher level of detection of information within shadowed areas for TVI and MSAVI is possible.

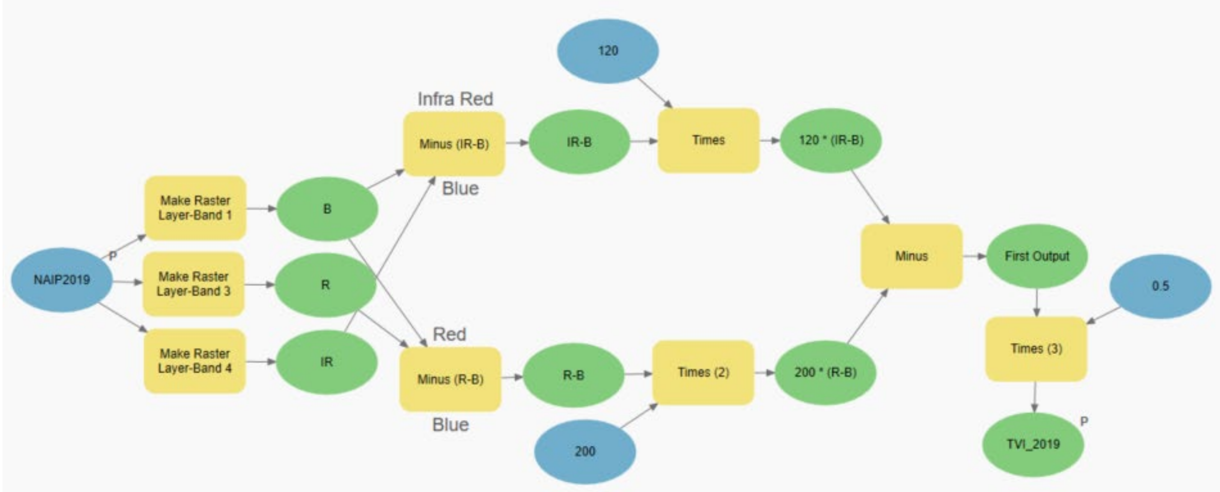
Calculating Vegetation Indices

NDVI was applied for both dates 2010 and 2019, which seemed to extract most of the vegetation from both imageries. A reclassification was carried out of the outputs in an attempt to isolate the shadowed areas. For example, for 2010, pixels with values of less than 115 seemed to include shadowed areas, while for 2019 the pixel value was 100. But it seems that the NDVI was not the best solution in either case.

In addition, TVI was calculated for both dates which, as described by Zhang et al. (2015), would help differentiate between illuminated and shadowed vegetation based on the following equation:

$$\boldsymbol{TVI = 0.5 \times \{120 \times (IR - B) - 200 \times (R - B)\}}$$

Where IR is the infrared band, R is the red band and B is the blue band of each of the NAIP dates. Each of the IR, R, & B bands was extracted from the composite image. The following model was created to calculate TVI:



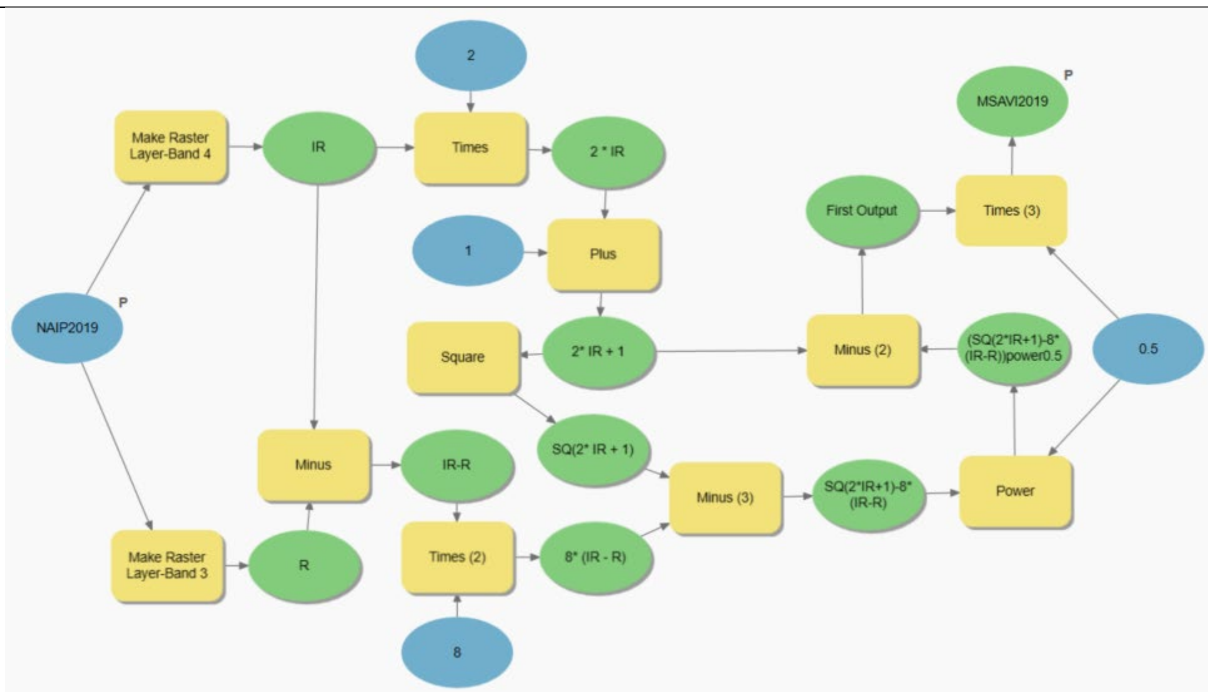
The process of calculating the TVI involved the extraction of individual spectral bands (IR, R, and B) and running the given equation at different steps on each of the given dates.

The resulting TVI values ranged from -18,180 to 11,560 for the 2010 NAIP and from -18,800 to 13,060 for the 2019 NAIP, which showed a great amount of variation in pixel values. This could present a higher level of flexibility in reclassification and extraction of shaded areas.

Furthermore, the MSAVI was calculated based on the following formula:

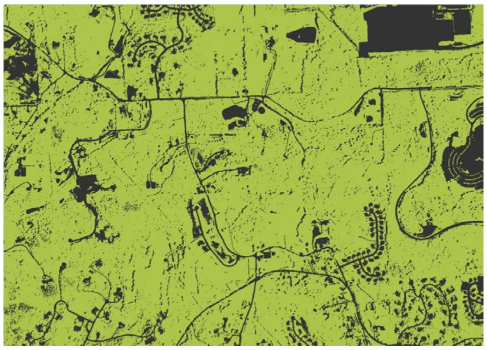
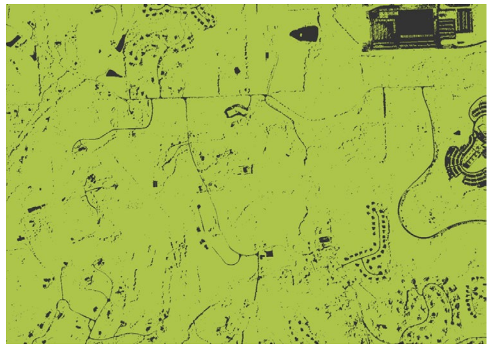
$$\boldsymbol{MSAVI = 0.5 \times \left\{(2IR + 1) - \left[(2IR + 1)^2 - 8 \times (IR - R)\right]^{0.5}\right\}}$$

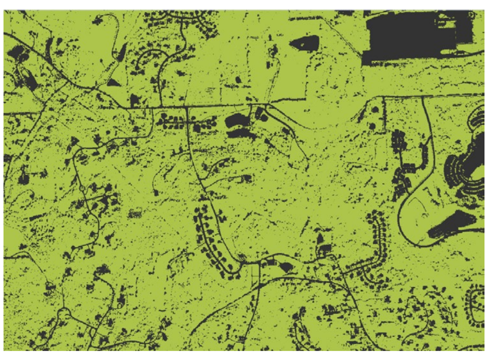
Where IR is the infrared band, and R is the red band of each of the NAIP dates. Each of the bands was extracted from the composite image. The following model was created to calculate MSAVI:



The process of calculating the MSAVI involved the extraction of individual spectral bands (IR and R) and running the given equation at different steps for each of the given dates.

The resulting MSAVI values ranged from -4.216 to 0.567 for the 2010 NAIP and from -19.184 to 0.746 for the 2019 NAIP. A comparison between the calculated vegetation indices (NDVI, TVI and MSAVI) is shown in the following figures:

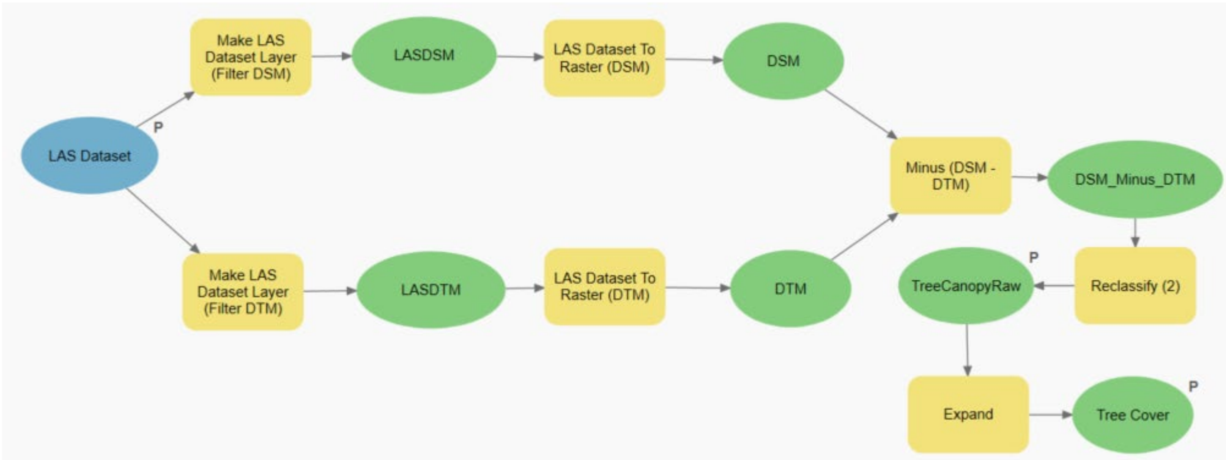
Vegetation Indices Comparison for 2010		
		
NDVI 2010-Values varied between 31 and 140, vegetation was identified between 115 and 140, with a large overlap between vegetated and no-vegetated surfaces.	TVI 2010-Values varied between -18,180 and 11,560, vegetation was identified between the values -800 and 11,560. Large overlap was observed between vegetated surfaces with structures and man-made features.	MSAVI 2010-Values varied between -4.216 and 0.567, vegetation was identified between values 0.031 and 0.567. MSAVI provided a clearer distinction between vegetation and non-vegetation features for this date. This result was adopted for further change detection.

Vegetation Indices Comparison for 2019		
		
2019 NDVI-Vegetation is identified in green, values vary between 2 and 161, vegetation has higher index value starting at about 100, any value below 100 would cover a mix between vegetation and non-vegetation surfaces.	2019 TVI-Vegetation is identified in Green, values vary between -18,800 and 13,060, vegetation was identified between values -2,999 and 13,060. TVI was capable to separate vegetated surface whether shadowed or not. This result was adopted for further change detection	2019 MSAVI-Values vary between -19.184 and 0.746. Vegetation was identified between values of -0.5 and 0.746.

Finally, the datasets from both dates were reclassified to either 1 for non-vegetation or 2 for vegetated land cover in order to facilitate later change detection calculations.

Tree Cover Extraction

Using the downloaded LiDAR LAS files, LAS datasets were created for each of the available dates. In this study, a tree canopy extraction model was used to extract the tree cover. Hansen and Fuglsang (2014) describes a method that calculates the difference between the Digital Surface Model (DSM) and the Digital Terrain Model (DTM) in order to extract the tree canopy (tree locations and canopy heights). This method was adopted and applied according to the following model:



Tree cover calculation using the difference between DSM and DTM calculated from LAS datasets of 2006 and 2020. Tree heights were omitted for the purpose of the current study, tree canopy with heights was also saved for possible future use.

Although the adopted model was designed to extract the tree canopy using a given LAS dataset, LAS files differed in specifications from year to year. In the current study, the point spacing of the 2006 PAMAP dataset varied between 4.716 ft and 5.442 ft, while it varied between 1.203 ft and 1.77 ft for the 2020 USGS dataset. So, after testing a 5, 7 and 10 ft pixel size in the rasterization procedure, 7 ft pixel size was adopted and proved to produce a better overall representation. In addition, the 2006 dataset required expanding the result by 1 pixel towards the end of the process in order to avoid unnecessary speckles and voids in the output. Visual testing was carried out in order to confirm those methods.

For DSM and DTM calculations, since there were no specific classes for vegetation (trees) nor buildings for neither 2006 nor 2020, the interpolation was carried out using information from the following table, those variables were adopted based on the visual observation of cloud point filters and return values for each date:

	LAS Dataset Filters			LAS Dataset Rasterization				
	Class Codes	**Return Values**	**Notes**	**Value Field**	**Interpolation Type**	**Cell Assignment**	**Void Fill Method**	**Sampling Type: Cell Size**
DSM	All but 2	- First of Many - Last of Many - 2nd Return	All Unflagged & Flagged points	Elevation	Binning	Maximum	None	7 ft
DTM	2	All return values	All Unflagged & Flagged points	Elevation	Binning	Average	None	7 ft

It is important to mention that some of ground points were included in some of the selected returns. Omitting those resulted in significant exclusion of tree points, therefore those points were elected to be included, keeping into consideration that the final result might include some misclassified pixels.

The difference between the produced DSM and DTM was calculated. The result at this stage contained tree cover and tree canopy elevations. For the purposes of the current study, a reclassification was carried out to the tree canopy to create a tree cover raster with no specific canopy elevation values (all values above 0 were assigned a class code of 1, all others were assigned 0).

Furthermore, to reduce the speckle effect, the classified areas from 2006 was increased by 1 pixel using the Expand function, the 2020 data did not need such refinement. A sample of the results are presented in the following figure:

2006	**2020**

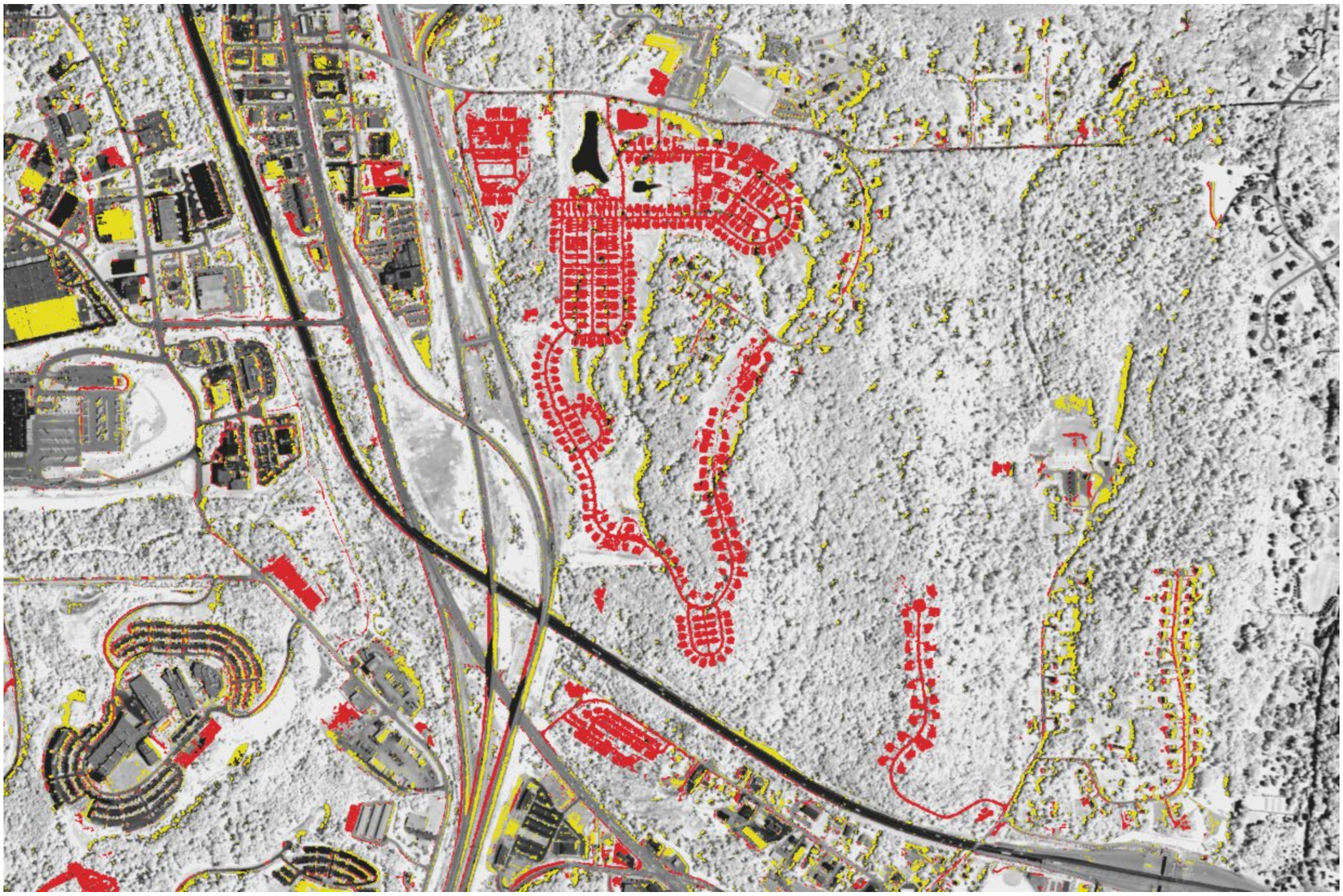
Result of the tree cover extraction from the LiDAR datasets overlaid on the corresponding NAIP imagery.

Change Detection

Comparison of the change in vegetated versus non-vegetated covers was carried out using the results from the vegetation indices method, while the comparison of the change in tree cover between the two given dates was done using the results from LiDAR data manipulation.

Vegetation Gain/Loss

A change raster was created to indicate categorical differences from and to vegetation and non-vegetation land covers. The result was filtered to display only the changed pixels. The following figure shows a sample image of the change in land cover between 2010 and 2019:



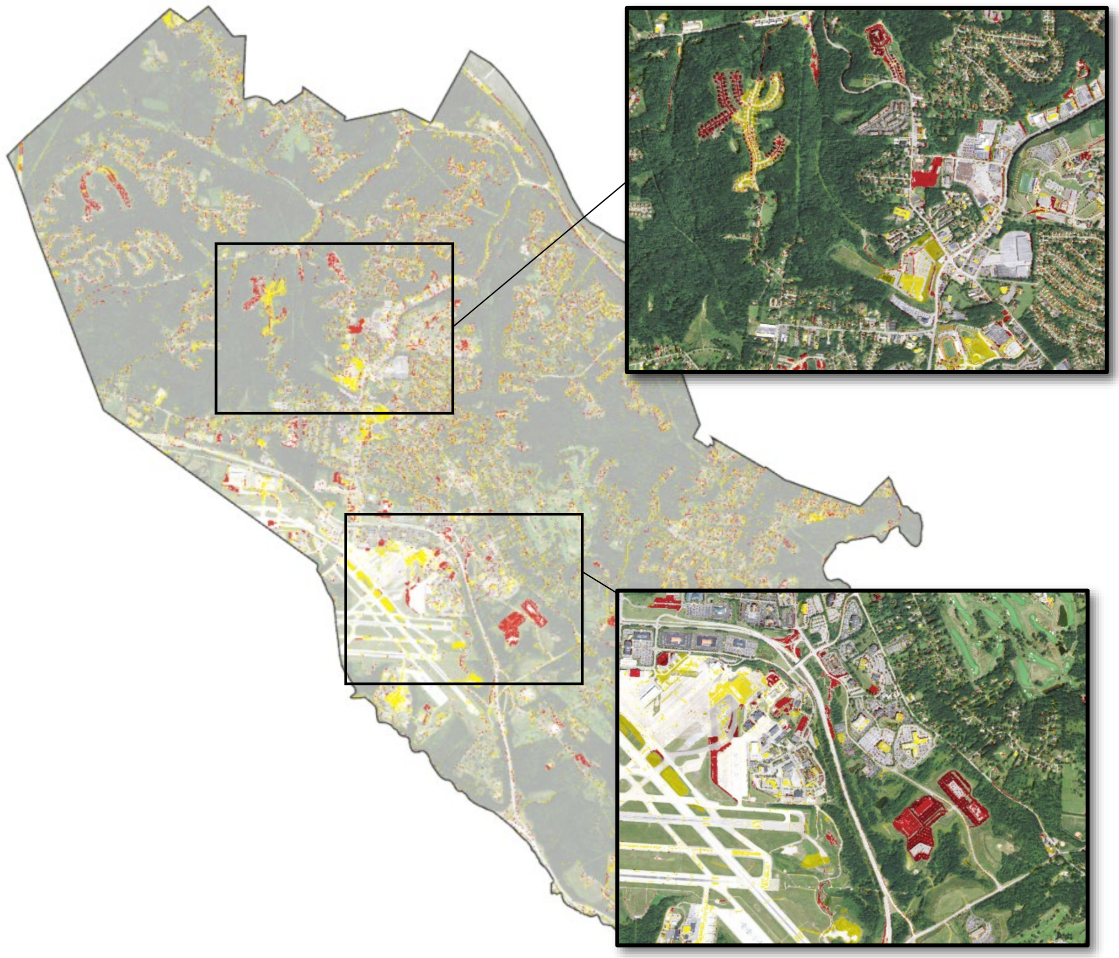
Difference image between 2010 and 2019, located at the north east of Marshall township, which shows the change from vegetation to non-vegetation land cover in red, and from non-vegetation to vegetation in yellow. It also shows that many pixels have been misclassified as a change to vegetation due to inaccuracies with the identification of some buildings and shadowed areas (top left of the image shows some buildings that never actually changed between the two dates, but the suggested method flagged them as changed nevertheless). On the other hand, the change to non-vegetation (red) was very close and fewer pixels were mis-classified in this category.

Overall, for both municipalities, there was a 3.39% (858.41 acres) change to non-vegetation and 4.67% (1,183.66 acres) of change to vegetation cover. The following table summarizes the type and amount of changes that occurred between 2010 and 2019 in both Moon township and Marshall township:

Class	Description	Count	Area (Acres)	Percent Change
1->2	Changed to Vegetation	4,790,17.00	1,183.66	4.67%
2->1	Changed to Non-Vegetation	3,473,859.00	858.41	3.39%
No Change	No Change	94,290,831.00	23,299.68	91.94%
Total		**102,554,817.00**	**25,341.75**	**100%**

The following figure shows the changes that took place between 2010 and 2019 for both Moon and Marshall Townships. Changes from vegetation to non-vegetation is depicted in red while the change to vegetated cover is depicted in yellow.

VEGETATION GAIN/LOSS BETWEEN 2010 AND 2019 FOR MOON TOWNSHIP



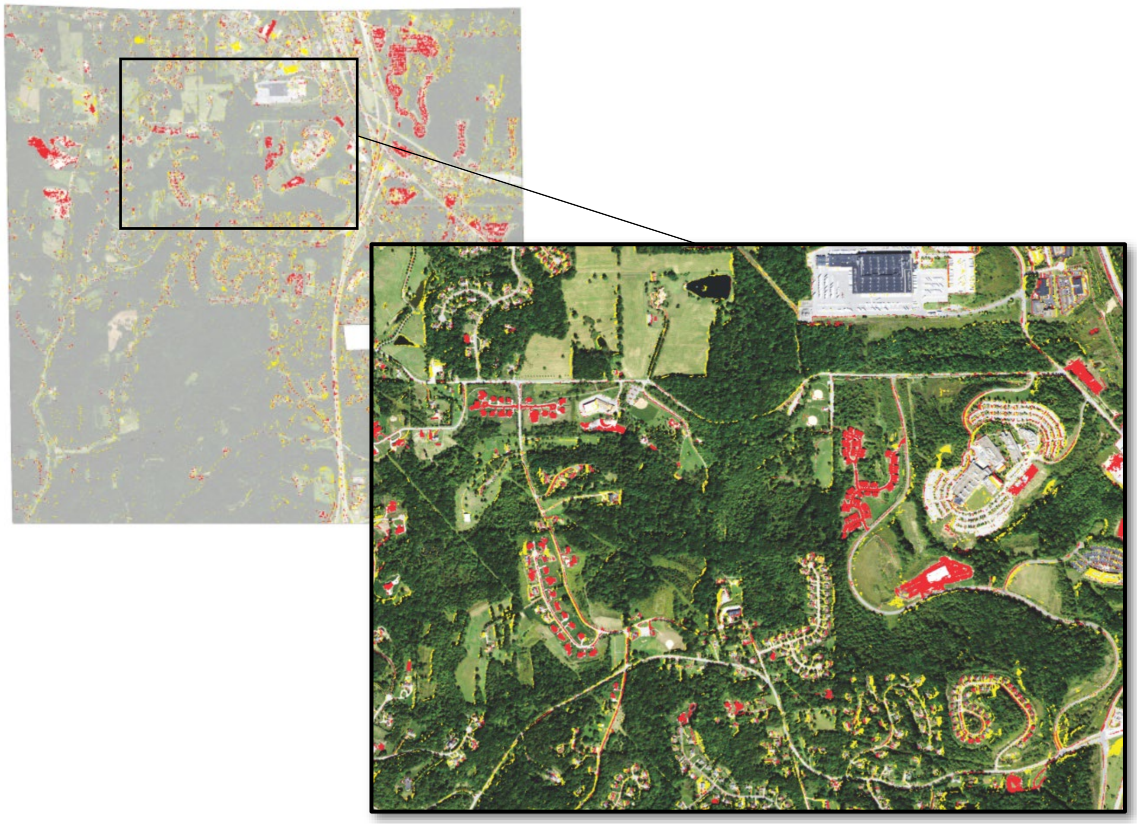
The adopted TVI and MSAVI methods showed an overall promising result with few exceptions. In Moon township, for example, as shown in the southern part of the change map above, areas marked in red were converted to non-vegetated cover between the two dates: note the large red areas to the south east of the enlarged frame. On the other hand, some areas were misclassified as gain of vegetation: Note the yellow patch of a residential neighborhood located in the north west part of the top enlarged frame.

The following table represents the vegetation cover changes from 2010 to 2019. It is estimated that Moon township lost about 3.35% (517.58 acres) of its vegetation and gained about 4.82% (745.23 acres) during this period.

MOON TOWNSHIP: VEGETATION COVER CHANGE FROM 2010 (MSAVI) TO 2019 (TVI)

Class	Description	Count	Area (Acres)	Percent Change
1->2	Changed to Vegetation	3,015,839.00	745.23	4.82%
2->1	Changed to Non-Vegetation	2,094,561.00	517.58	3.35%
No Change	No Change	57,478,578.00	14,203.21	91.83%
Total		**62,588,978.00**	**15,466.01**	**100%**

VEGETATION GAIN/LOSS BETWEEN 2010 AND 2019 FOR MARSHALL TOWNSHIP



The adopted TVI and MSAVI methods showed an overall promising result with few exceptions. In Marshall township, for example, as shown in the northern part of the change map above, areas marked in red were converted to non-vegetated cover between the two dates: note the large red areas to that mark the addition of several residential buildings as well as a large red patch, on the right, that indicates an addition of a large parking lot around an existing building.

In Marshall township, the total gain of vegetation was about 4.44% (438.4 acres), and total loss was about 3.45% (340.81 acres). The following table summarizes the type and amount of change of vegetation in Marshall township between 2010 and 2019:

MARSHALL TOWNSHIP: VEGETATION COVER CHANGE FROM 2010 (MSAVI) TO 2019 (TVI)

Class	Description	Count	Area (Acres)	Percent Change
1->2	Changed to Vegetation	1,774,133.00	438.40	4.44%
2->1	Changed to Non-Vegetation	1,379,221.00	340.81	3.45%
No Change	No Change	36,810,402.00	9,096.01	92.11%
Total		**39,963,756.00**	**9,875.22**	**100%**

Tree Cover Change

A change raster was created to indicate categorical differences in gain and loss of tree cover between 2006 and 2020. The results display only the areas that changed. The following figure shows a sample image of the change in tree cover between 2006 and 2020 located in the north east of Marshall township:

The tree cover that remains unchanged between 2006 and 2020 is depicted in green. Tree loss is shown in red while tree cover gain is shown in yellow. It is important to note that some ground pixels were misclassified as trees from the USGS 2020 LiDAR dataset.

It is estimated that there was a 24.39% (66,580.76 acres) increase in tree cover and about 2.14% (5,836.38 acres) loss of tree cover in both municipalities during the period between 2006 and 2020. The following table summarizes the overall change that occurred in both municipalities:

Class	Description	Count	Area (Acres)	Percent Change
0->1	Changed to Trees	5,498,854.00	66,580.76	24.39%
1->0	Changed to No Trees	482,022.00	5,836.38	2.14%
No Change	No Change	16,562,238.00	200,537.49	73.47%
Total		**22,543,114.00**	**272,954.62**	**100%**

The following figure shows the changes in tree cover that took place between 2006 and 2020 for Moon Township. Changes from trees to no trees is depicted in red while the change to trees is depicted in yellow.

VEGETATION GAIN/LOSS BETWEEN 2010 AND 2019 FOR MOON TOWNSHIP

The detection of the loss in vegetation cover (red) was mainly main accurate when visually compared between the two dates. Many pixels were misclassified as addition of trees (yellow) as shown here.

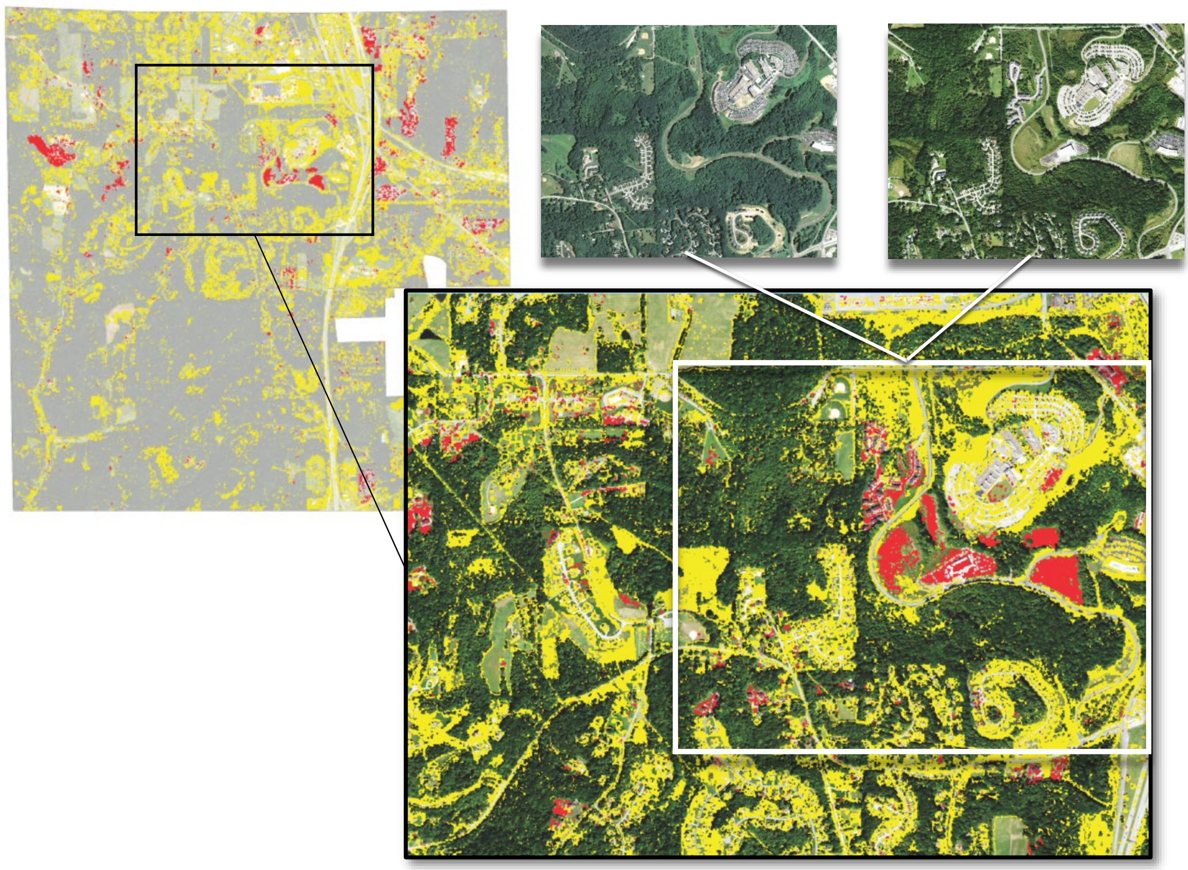
The following table represents the tree cover changes from 2006 to 2020. It is estimated that Moon township lost about 1.95% (3,248.18 acres) of its tree cover and gained about 25.23% (42,024.24 acres) during this period.

MOON TOWNSHIP TREE CANOPY CHANGE FROM 2006 TO 2020

Class	Description	Count	Area (Acres)	Percent Change
0->1	Changed to Trees	3,470,750.00	42,024.24	25.23%
1->0	Changed to No Trees	268,265.00	3,248.18	1.95%
No Change	No Change	10,019,469.00	121,316.89	72.82%
Total		**13,758,484.00**	**166,589.31**	**100%**

In Marshall township, the result trends were no different than Moon township. Same pattern of misclassified conversion to tree covers was excessive in the township. The following figure depicts this result.

TREE COVER CHANGE BETWEEN 2006 AND 2020 FOR MARSHALL TOWNSHIP



The detection of the loss in vegetation cover (red) was mainly main accurate when visually compared between the two dates. Many pixels were misclassified as addition of trees (yellow) as shown here.

In Marshall township, the calculated total gain of tree cover was about 23.09% (24,556.52 acres), and the total loss was about 2.43% (2,588.19 acres). The following table summarizes the type and amount of change of tree cover in Marshall township between 2006 and 2020:

MARSHALL TOWNSHIP TREE CANOPY CHANGE FROM 2006 TO 2020

Class	Description	Count	Area (Acres)	Percent Change
0->1	Changed to Trees	2,028,104.00	24,556.52	23.09%
1->0	Changed to No Trees	213,757.00	2,588.19	2.43%
No Change	No Change	6,542,769.00	79,220.60	74.48%
Total		**8,784,630.00**	**106,365.31**	**100%**

Conclusions & Recommendations

This study explored the use of high resolution National Agricultural Imagery Program NAIP datasets as well as LiDAR point cloud data (LAS) in detecting and assessing the change in vegetation features in general and tree cover in specific. The methods adopted in this endeavor produced promising results that could be further developed to produce more accurate results. Nevertheless, some challenges were present. For example, the presence of shadows in high resolution imagery could prevent the methods from producing accurate classification of land cover in general and vegetation in specific, especially when using low spectral resolution imagery similar to NAIP, which contains only 4 spectral bands. Furthermore, classified LiDAR datasets are designed to serve certain purposes and applications, but specific features such as structures, vegetation, and tree cover could be absent from those classifications and the user is left with aggregated classes. A deeper analysis of the LiDAR datasets, including further investigations of signal intensity and point density could be helpful in future refinements. Finally, the compatibility of the datasets through the years could also constitute a challenge in some cases. Absence of time synchronized datasets prevents analysts from achieving accurate change detection results especially in those areas that witness rapid spatial changes.

Overall, for the current study, the spatial change that the area witnessed during the 10 years period between 2010 and 2020 was mainly characterized by an overall increase in vegetation cover. In some areas, it was very clear that there was a conversion of the natural landscape into more man-made structures, but generally, this change was accompanied by the addition of tree cover and open areas of lawn around the buildings, which might have had a major contribution to the presented results.

The mixed detection of tree cover especially from the USGS LAS files of 2020 could be reduced if a hybrid method would be developed using the vegetation indices used earlier in this study and/or other methodologies that would help in better identifying tree cover. Those misclassified pixels of structures and other features than trees could be removed. On the other hand, any other misclassified pixels or areas covered by another natural vegetation cover could present a challenge with this method. Furthermore, including a combination of tools and processes could be beneficial. For example, including segmentation techniques applied to the Vegetation Indices might produce higher accuracy results and is yet to be tested.

For future work that relates to this study, a more rigorous accuracy assessment of the classification results is needed. Refinement of the methods could be achieved if quantifiable accuracy assessment figures were available. The methods could also be applied to sub municipality levels (census tracts, blocks, or block groups) in order to assess changes in smaller map units and link the spatial results to population data in a more direct way.

Finally, the study showed potentials and limitations of using specific datasets in studying landscape change. Looking at two townships in Allegheny county gave an indication that there is significant development in the area which doesn’t appear to be slowing down in the near future. Nonetheless, this does not seem having a negative impact on the vegetation cover in the area. The urban expansions continue to take place in several surrounding municipalities in Allegheny, Beaver and Butler counties and a more thorough and expansive study of the population growth in those areas is highly recommended.

References

Arévalo, V., González, J., & Ambrosio, G. (2008). Shadow detection in colour high-resolution satellite images. *International Journal of Remote Sensing*, *29*(7), 1945–1963.

Hansen, H. S., & Fuglsang, M. (2014). An Operational Web-Based Indicator System for Integrated Coastal Zone Management. *ISPRS International Journal of Geo-Information*, *3*(1), 326–344.

Hanssen, F., Barton, D. N., Venter, Z. S., Nowell, M. S., & Cimburova, Z. (2021). Utilizing LiDAR data to map tree canopy for urban ecosystem extent and condition accounts in Oslo. *Ecological Indicators*, *130*, 108007.

Kwatra, V., Han, M., & Dai, S. (2012). *Shadow Removal for Aerial Imagery by Information Theoretic Intrinsic Image Analysis*.

Roberts, G., Pyun, Y., & Stucka, M. (2021, August 16). *How many people live in Allegheny County after the 2020 Census count?* [News]. USA TODAY Network. https://www.timesonline.com/story/news/2021/08/16/gda-2020-census-population-pa-nbct-42003/48479281/

Shahtahmassebi, A., Yang, N., Wang, K., Moore, N., & Shen, Z. (2013). Review of shadow detection and de-shadowing methods in remote sensing. *Chinese Geographical Science*, *23*(4), 403–420.

Zhang, L., Sun, X., Wu, T., & Zhang, H. (2015). An Analysis of Shadow Effects on Spectral Vegetation Indexes Using a Ground-Based Imaging Spectrometer. *IEEE Geoscience and Remote Sensing*