Land Use Land Cover mapping through time (1984-2020) using CCDC Algorithm for Toronto and Region Conservation Authority

Shakeeb Tahir, Supervisors: Dr. Mitchell Bonney and Dr. Yuhong He shakeeb.tahir@mail.utoronto.ca

Department of Geography, Geomatics and Environment, University of Toronto Mississauga

Introduction

Problem:

Local governments and conservation authorities need accurate spatial data quantifying land use land cover (LULC) change through time to meet their management needs. In the Toronto area, the Toronto and Region Conservation Authority (TRCA) is interested in developing an automated method for mapping LULC across their lands, which can potentially (given consistent data and methodology) be used to detect changes over time. The LULC map time series would provide essential information for a wide variety of TRCA projects, including Natural Heritage System/Water Resources System development, habitat restoration initiatives, monitoring programs, and watershed planning.

Objective:

We aim to use an automatic method, Continuous Change Detection and Classification (CCDC) algorithm [1] to classify LULC for the TRCA over years. CCDC is proposed in this study because the algorithm makes use of satellite images - Landsat time-series pixels – by creating continuous (e.g., seasonal) spectral curves that can classify LULC at any point in time and may better separate spectrally similar classes that often fool single-date methods (e.g., coniferous vs. deciduous forest).

Methodology

STUDY AREA: TRCA Watersheds
DATA: Landsat Images and TRCA Ecological Land Classification Data
METHODS:

1. A stratified random sample of 1000 training points was created using the TRCA provided Ecological Land Classification Data [2]
2. The Ecological Land Classification shapefile's 25 classes were grouped and simplified to 16 classes.
3. The training points were then verified over images by visual interpretation in order to ensure that they were accurate. Inaccurate points were removed from the sample.
4. CCDC algorithm was performed based on training points to classify LULC via a web app [3]. The classified maps were interpreted and training points were added/removed based on classes which were over/under classified.
5. Random proportional sample points were then created and visually interpreted for accurate assessment of the CCDC classified map from 2019
6. Accuracy assessment was done for the regular sample points, paired-down classes, and their "fuzzy" classifications where if a main or secondary visual interpretation is correct, the reference is considered correct.

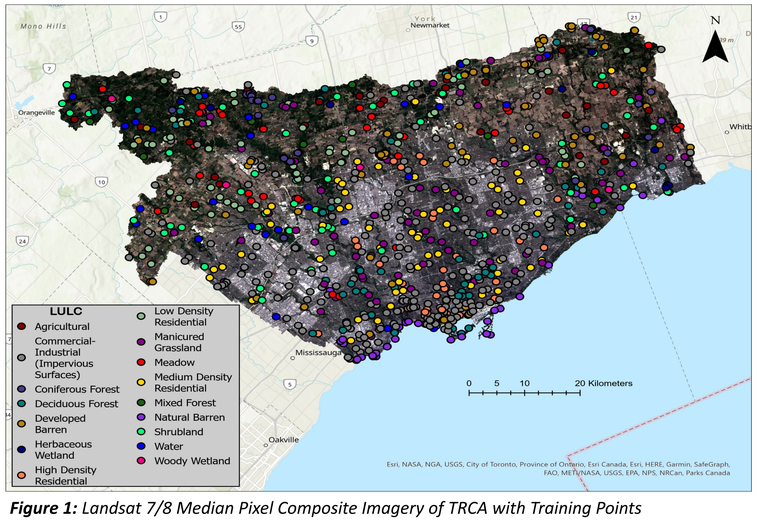

Figure 1: *Landsat 7/8 Median Pixel Composite Imagery of TRCA with Training Points*

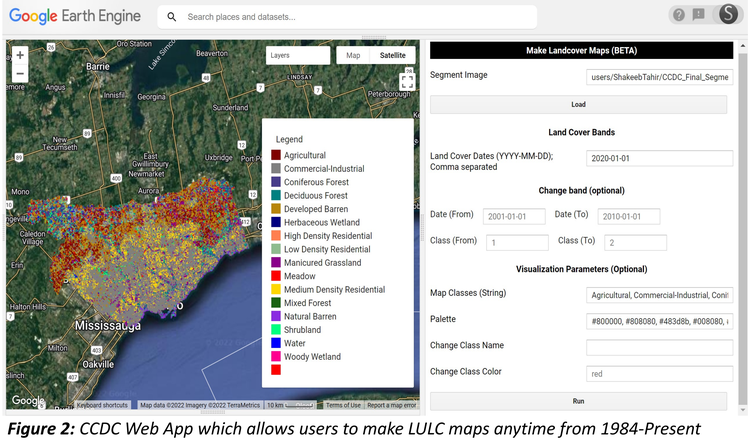

Figure 2: *CCDC Web App which allows users to make LULC maps anytime from 1984-Present*

Results: Zoomed In CCDC Classification of TRCA Land throughout the years

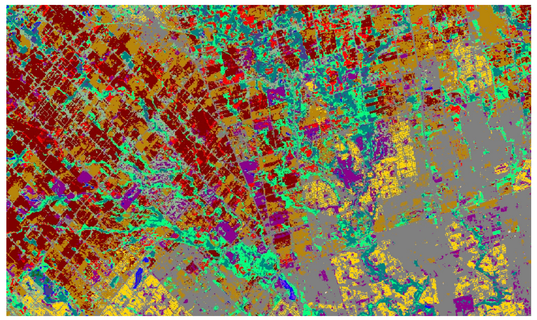

Figure 3: *1990 Zoomed-In LULC Classification*

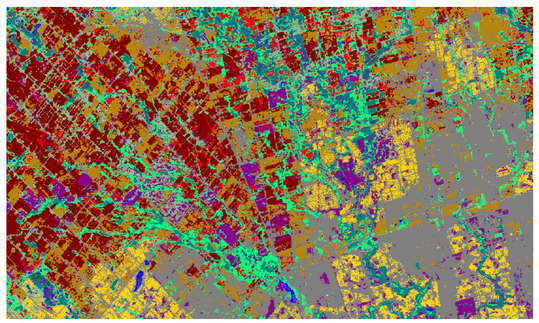

Figure 4: *2000 Zoomed-In LULC Classification*

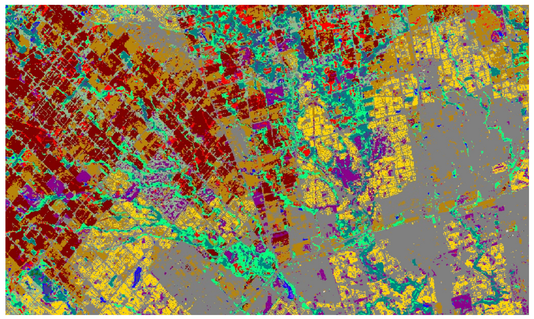

Figure 5: *2010 Zoomed-In LULC Classification*

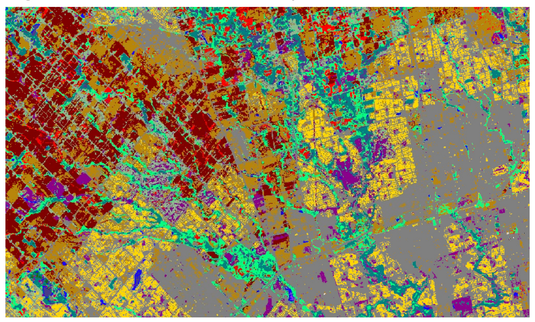

Figure 6: *2020 Zoomed-In LULC Classification*

Legend
- Agricultural
- Commercial-Industrial
- Coniferous Forest
- Deciduous Forest
- Developed Barren
- Herbaceous Wetland
- High Density Residential
- Low Density Residential
- Manicured Grassland
- Meadow
- Medium Density Residential
- Mixed Forest
- Natural Barren
- Shrubland
- Water
- Woody Wetland

Observations:

- CCDC algorithm allows us to create LULC maps from Landsat time series at any point of time between 1984 and now.
- For each pixel, the algorithm forms a continuous time-series curve of every available Landsat observation and uses its spectral and temporal information to label its LULC class for the pixel.
- Clear changes are visible through each decade from 1990 to 2020 when we zoom into the classified TRCA maps. Specifically, we can see the effects that urban sprawl has on many of the LULC classes such as agriculture, developed barren, and shrubland as they are replaced by classes such as medium-density residential or commercial-industrial.

Results: CCDC Overall 2020 Classification

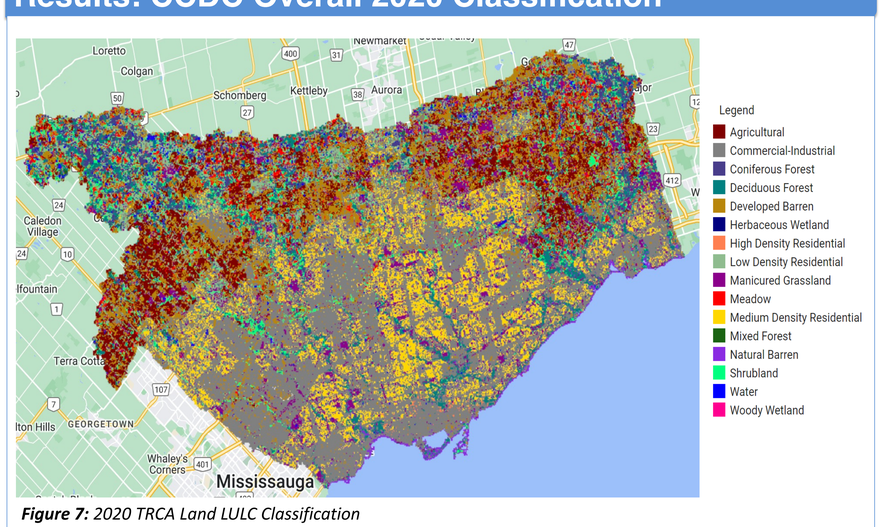

Figure 7: *2020 TRCA Land LULC Classification*

Discussion and Conclusion

- Commercial-Industrial is somewhat over-represented within the classification due to having more training points after the stratified sample was visually interpreted. As a result, classes such as High-Density Residential and Medium-Density Residential are less prevalent than they should be within the classification.
- Classification struggled somewhat with more of the natural LULC, often confusing classes such as Meadow and Shrubland. However, the continuous nature of the algorithm did help separate some spectrally similar classes such as Coniferous vs. Deciduous forest relatively well.
- Often, the random points created for the training data were near the edge of another class (such as a meadow point being near roads (commercial-industrial) or water points being near forests, which likely resulted in some pixels having mixed spectral values which would confuse the classification. In order to improve the training data, points could possibly be moved to more homogenous areas.
- Under-represented classes need more points in order to make sure they are represented within the classification.
- Grouping the classes when creating training data into even fewer classes (from 16 to 8-10) would increase overall accuracy as having many classes especially those which are spectrally similar makes it harder for the algorithm to distinguish between them.
- Overall, the results of the CCDC algorithm provide a good foundation for an initial framework of an automatic way to classify LULC over time.

Results: Accuracy Assessment

LULC	UA	PA
Agriculture	1.00	0.28
Com-Industrial	0.63	0.97
Coniferous Forest	1.00	0.61
Deciduous Forest	0.80	0.63
Developed Barren	0.12	0.80
Herbaceous Wetland	0.50	0.45
High Density Residential	0.63	0.5
Low Density Residential	0.83	0.83
Manicured Grassland	0.65	1.00
Meadow	0.82	0.53
Medium Density Residential	0.88	0.58
Mixed Forest	0.82	0.45
Natural Barren	0.36	0.8
Shrubland	0.45	1.00
Water	0.54	0.60
Woody Wetland	0.8	0.89
Overall Accuracy	~ 0.63	

Table 1: *16 Classes Fuzzy Classification*

Observations:
Lack of sample points seems to overestimate accuracy of some classes such as Woody Wetland. Classes such as Meadow and Shrubland were often confused with each other due to being spectrally similar. Water had lower than expected accuracy as it was over classified likely due to shadows being spectrally similar which could have confused the classification.

Accuracy Assessment Type	Overall Accuracy
16 Classes (Primary Only)	0.48
16 Classes (Fuzzy)	0.63
8 Classes (Primary Only)	0.58
8 Classes (Fuzzy)	0.71

Table 3: *Accuracy Assessment Types Summary*

LULC	UA	PA
Agriculture	1.00	0.28
Barren	0.20	0.90
Com-Industrial	0.63	0.97
Forest	0.95	0.61
Natural Vegetation	0.73	0.95
Residential	0.94	0.70
Water	0.54	0.60
Wetland	0.90	0.90
Overall Accuracy	~ 0.71	

Table 2: *8 Classes Fuzzy Classification*

References

[1] Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat Data. *Remote Sensing of Environment*, *144*, 152–171. https://doi.org/10.1016/j.rse.2014.01.011

[2] TRCA LandUse NaturalCover 2017. *TRCA Open Data Portal.* https://trcacamaps.opendata.arcgis.com/datasets/camaps::trca-landuse-naturalcover-2017/about

[3] CCDC LandCover App. *Google Earth Engine.* https://code.earthengine.google.com/?scriptPath=users%2Fparevalo_bu%2Fgee-ccdc-tools%3AAPPS%2Flandcover_app

Acknowledgements

- Toronto and Region Conservation Authority (TRCA)
- Jonathan Ruppert and Namrata Shrestha
- Research Opportunity Program (ROP)
- Professor Mitchell Bonney for feedback and support along with the professional interpretation of accuracy assessment points
- Professor Yuhong He's research group for comments and feedback