Huronia Land Conservancy Mapping Project 2024

Includes:

Final Report

By: Benjamin Krawec

Completed for: Huronia Land Conservancy

Supervising Professor: Autumn Watkinson

Trent Community Research Centre Project Coordinator: Matthew Walmsley

Course Code: ERSC 4840H

Course Name: Community-Based Resource Mgmt

Completion Date: April 21, 2025

Project ID: 6269

Introduction

The Huronia Land Conservancy (HLC) was incorporated in 2012 with the goal of protecting and conserving Natural and Cultural Heritage Areas from Barrie to the Penetanguishene Peninsula. While much of North Tiny is currently protected by Simcoe County Forests and Awenda Provincial park, large areas of ecological significance are currently unprotected. The HLC aims to address this shortfall.

The primary objectives of this project were:
Delineate the presence, coverage and general height classes of stands of mature forest within the Region of Interest through the generation of a Crown Delineation Model (CDM).

The creation of an information package which can assist the Huronia Land Conservancy with recruiting the public to assist in conserving the North Tiny Forest. This information package consists primarily of mapping products supported by a concise power point presentation.

The creation of a package of GIS layers and datasets which can assist the HLC and allied organizations with identifying key pieces of land to conserve; ie, habitat for species at risk, preservation of ecosystem services associated with mature forest, preservation of trees with particularly valuable genetics for future restoration projects.

The secondary objective of this project was to generate a georeferenced contact list for landowners within the HLC study area which can allow the HLC to directly contact the owners of land deemed to be of high conservation value. At the time of this writing the time and resources were unavailable to complete this objective.

Unique Ecological features of the North Tiny Forest at a glance

During my site visit with Peter it became clear that one of the things which makes the North Tiny Forest truly special is the sheer age and continuity of it. Being able to highlight areas of particularly ancient and ecologically valuable forest is key to illustrating what it is that makes this place worth protecting.

North Tiny Forest is home to:
15 Areas of National Scientific Interest
Approximately 7.5 square kilometres classified as "Significant Groundwater Recharge Areas"
59 reported Species At Risk

Amphibians	1
Birds	20
Bryophytes	1
Fishes	1
Insects	7
Mammals	3
Reptiles_Turtles	11
Vascular Plants	15

The Ontario Ministry of the Environment, Conservation and Parks directly refers to the northwestern section of the study area as critical habitat in its Recovery Strategy for the Red-headed Woodpecker (Melanerpes erythrocephalus) in Ontario.

Included Data

With the exception of particularly large files such as LiDAR point clouds and aerial imagery, all data products, maps, geopackages and datasets which were created in the course of this projects or used in its creation are included in the .zip file, 'NorthTinyDataPackage.' A description of this zip file's contents follows.

Maps

A series of maps generated from the CDM outputs, basemaps and public geospatial data have been included. These depict a general overview of where areas of high conservation value lie in relation to the boundaries of Simcoe County Forests and Awenda Provincial Park.

Geopackages

The following geopackages are included as the final deliverable. Their utility in the project as well as their original data sources and the processes by which they were derived are explained in the following section. The majority of these files were generated using QGIS 3.38.2-Grenoble but they should be compatible with any desktop GIS software.

NorthTinySoil.gpkg

Contains information related to soil or bedrock geology.

Bedrock geology   -   Explanatory factor in soil chemistry and plant community structure. Clipped from 'sgu_polybedrock_polygon.shp'

Soil polygons -   Explanatory factor in plant community composition. Predictor of nesting habitat for herpetofauna. Clipped from 'SOIL_SURVEY_COMPLEX_ESRI_EN' within the LIO Composite File Geodatabase.

NorthTinyBasemaps.gpkg

Contains information useful for generating topographic maps.

Roads -   Line vector. Predictor of habitat fragmentation and road mortality. Clipped from 'ORN_SEGMENT_WITH_ADDRESS_ESRI_EN' within the LIO Composite File Geodatabase.

Watercourse   -   Line vector. Predictor of habitat or habitat connectivity for particular species. Clipped from 'OHN_WATERCOURSE.shp'

Contours   -   Line vector. Aid in navigation, visualization of topography. Derived from LiDAR derived Digital Terrain Model using the 'contour' function in QGIS.

Waterbody   -   Polygon vector. Predictor of habitat or habitat connectivity for particular species. Clipped from 'OHN_WATERBODY.shp.'

Buildings   -   Polygon vector. Predictor of anthropogenic habitat. Predictor of habitat fragmentation or disruption. Derived from LiDAR data using the "Classify Buildings' function in ArcGIS Pro.

NorthTinyCanopyModel.gpkg

Contains models of the tree canopy within the study area.

NorthTinyCDM   -   Crown Delineation Model. Polygon vector. Attribute '_max' represents the maximum value for the Canopy Height Model within each feature. Derived from Ontario Classified Point Cloud – 2022LHURONGEORGIANBAY via process delineated in 'Methods' section. Predictor of habitat, habitat connectivity and ecosystem services or recreational value. On the ground validation for height data suggests an error rate of about ten percent. Some of this error is related to operator error while using the clinometer. While ten percent difference between predicted and measured values is too large for granular analysis, at the landscape scale this is useful for predicting general height classes. Canopy width and crown delineation validation has been conducted qualitatively and deemed good enough to be useful but has yet to have its % error measured quantitatively.

CHM -   Raster image representing tree heights across the study area. Derived from Ontario Classified Point Cloud – 2022LHURONGEORGIANBAY via process delineated in 'Methods' section. Potentially useful for canopy analysis.

CHM_Smoothed   -   CHM after undergoing a gaussian smoothing process in R Studio to improve canopy delineation. Potentially useful for improving the CDM. The smoothing process is described in more detail under methods.

NorthTinySpecialZones.gpkg

Contains polygon vectors depicting the locations of areas noted as ecologically or culturally valuable.

SignificantGroundwaterRecharge   -   Areas in which the rate of groundwater recharge is estimated to be 15 percent higher than the rest of the watershed. Clipped from 'CA_INFO_GIS_V_SWP_SGRA' found on Simcoe County's open GIS portal.

ANSI -   Areas designated as being of National Scientific Interest. Clipped from 'ANSI_ESRI_EN' within the LIO Composite File Geodatabase.

Wetlands   -   Designated wetlands. Clipped from 'Wetlands.shp' from Ontario Geohub.

NorthTinyBoundaries.gpkg

Depicts boundaries of areas protected by public ownership or the bounds of the study area.

SimcoeCtyFor -   Manually derived from maps and images found on Simcoe County website.

Awenda Park Boundary   -   Clipped from 'PROV_PARK_ADMIN_ZONE_ESRI_EN' within the LIO Composite File Geodatabase.

NorthTinyROI   -   Manually generated via verbal and written instructions from the host.

NorthTinyUtil.gpkg

ExpandedROI -   Derived via expanding NorthTinyROI. Used to clip vectors and rasters to a useable size while still providing a little context from outside the study area.

ValidationPoints   -   Point vector. Derived from data collected from the field by myself and the host, Peter Stubbins. Recorded species, location data, height (m), canopy width (m), diameter (cm), elevation, horizontal and vertical accuracy (m)

Linesplitter   -   Line vector. Used for cutting the CHM into manageable chunks to process for the Canopy Delineation Model.

ValidationWindows   -   Polygon vector used to perform zonal statistics on the CHM in order to compare measured tree heights with values from the CHM. Derived from ValidationPoints via the 'Rectangles, Ovals, Diamonds' feature in QGIS. Polygon size is derived from canopy width and horizontal accuracy.

ValidationPoints_x_CHM_zonalstats_unsmoothed -   Dataset. Compares measured vs model-predicted values. Derived from 'Zonal Statistics' function in QGIS.

ValidationPoints_x_CHM_Smoothed_zonalstats   -   Dataset. Compares measured vs model-predicted values from the smoothed Canopy Height Model. Used for trouble shooting potential issues with inaccuracy. Derived from 'Zonal Statistics' function in QGIS.

Other data

SARObservations_LocationStripped.csv   -   NHIC provided dataset with sensitive information such as names of observers, personal contact information and location data stripped out. Includes habitat descriptions and other useful data.

SAR_Observation_Count.csv   -   Count of observations per species in the study area.

CHM_Smoothing_Crown_Delineation.Rmd   -   R notebook containing source code for the CHM smoothing and crown delineation process as well as detailed notes which can guide others who want to improve on the current model or generate their own CDM.

CDM_Workflow.png -   Image depicting a detailed workflow beginning with classifying LiDAR data and ending with a CDM.

Methods

Data sources
High resolution satellite imagery is available from the South Central Ontario Orthophotography Project (SCOOP 2018) and is available for use under the Open Government License. Used in basemaps and for validating the Canopy Height Model as well as crown delineation output. Due to the size of this file it was not included in the package, however it is available for download from Ontario Geohub.

LiDAR data as well as LiDAR derived Digital Terrain Model and Digital Surface Model from the Ontario Classified Point Cloud available on GeoHub. This dataset came with the ground, water and noise layers already classified. . Used for generating the Canopy Height Model as well as deriving information about habitat and environmental conditions.

Land Information Ontario (LIO) Warehouse Open Data Products (Composite File Geodatabase) contains all publicly available products provided by Land Information Ontario. Contains information regarding wetlands, ANSI sites, waterbodies, provincial park boundaries.

NHIC has provided a geodatabase containing species at risk reports for the study area

Crown delineation
After much trial and error, a process was developed for identifying areas which are likely to have old-growth characteristics. In order to do this a model of the tree canopy had to be created in which individual tree crowns were modelled and segmented.
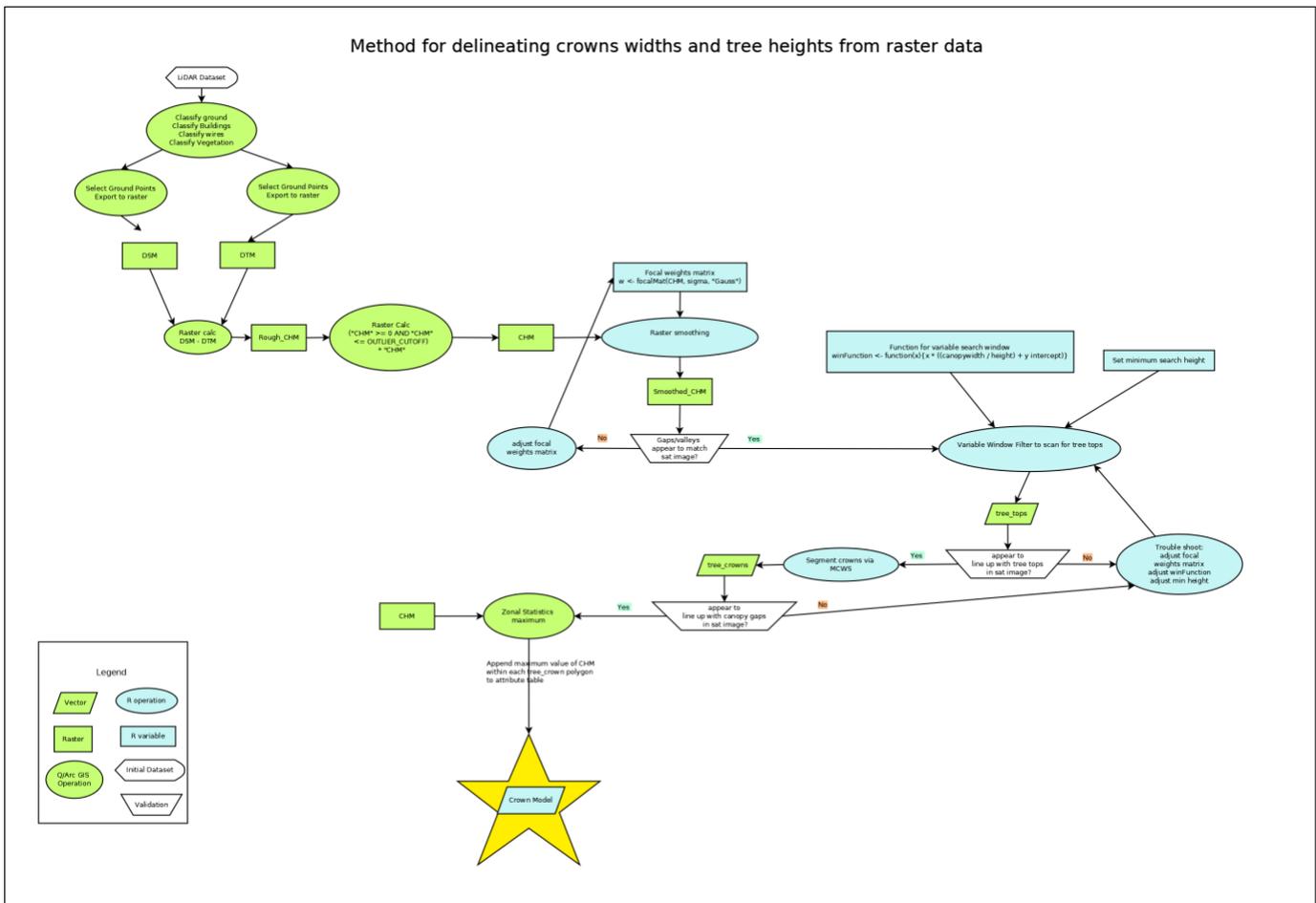
Figure 1: Workflow for generating a CDM, beginning with LiDAR data. A higher resolution version of this image is available in the included .zip file.

First, buildings, wires, and vegetation needed to be classified from the LiDAR point cloud; in that particular order. This process is described in more explicit detail in an attached document, 'BenjaminKrawec_Report_ERSC3160_Revised_04_04.pdf,' a report from a previous CBR I conducted which focused mainly on LiDAR classification. Once this process is complete, two raster files need to be exported from the classified dataset; a Digital Surface Model, which displays the height in meters above sea level of every recorded point which represents everything above the ground, such as trees, shrubs and buildings; and a Digital Terrain Model, which displays the height in meters above sea level of every recorded point which represents the ground or water. In this case we adjust our selection query to exclude power-lines, low vegetation and buildings.

From these two files the raster calculator in QGIS was used to subtract the values of the Digital Surface Model from the Digital Terrain model. The product of this operation is the Canopy Height Model (CHM), which represents the height of all vegetation in the study area.
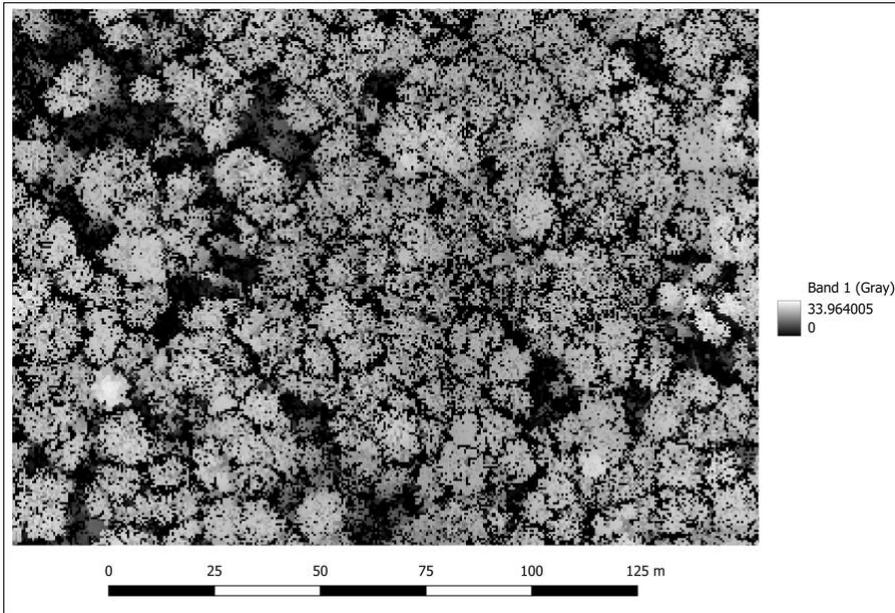
Band 1 (Gray)
33.964005
0

0   25   50   75   100   125 m

Figure 2: Sample of the generated CHM. This level of fine detail makes it difficult for the computer to differentiate between individual trees

Post processing is necessary in order to fill gaps and simplify the CHM as well as remove outliers. Outliers were set at 0 meters and 33 meters. The high cutoff was chosen after speaking to Tim Tully, the Park Naturalist with Awenda Park.

Smoothing was somewhat more complex. This process as well as the final canopy delineation are explained in greater detail within the attached R notebook, CHM_Smoothing_Crown_Delineation.Rmd. Several different smoothing operations were performed in R using operations found in the 'raster' and 'terra' libraries.
The method which produced the best results for tree crown delineation was a variable window focal mean operation in which the window size is determined by a Gaussian filter. Essentially this means that the edges of features within the image were blurred and blended together, and the radius of the blurring was determined by a function which examined the value of the pixel it was centred on.
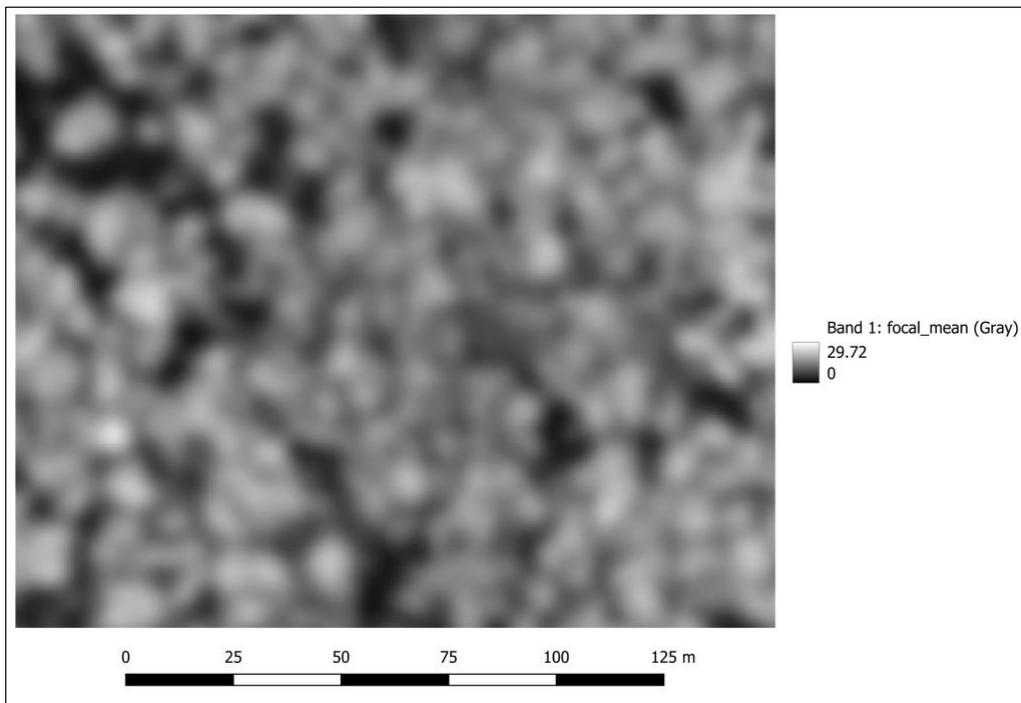
Figure 3: The same CHM after being smoothed

Tree tops were identified using a Variable size Search Window (VSW). A VSW identifies the highest points in a raster by comparing them to their neighbours. The size of the neighbourhood being searched (the window) is derived by a function from the height of the point being examined. The idea is that the search window is roughly equal to the canopy width of a tree, if you can figure out *generally* what the ratio of width to height is for a given species then you can create a reasonably accurate search window. The proportion for canopy width to height (0.144 + 0.44) was initially derived by performing a linear regression on data obtained from the Tallo dataset (Jucker et al., 2022) and then adjusted by randomly selecting points within the study area and comparing the height reported in the CHM to the measured width of the canopy as depicted in high resolution satellite imagery.

The VSW for Tree top identification process in R was created using the code

```
function(x){x * 0.144 + 0.44}.
```

Where x is equal to the value of the raster (height) at the location being examined.
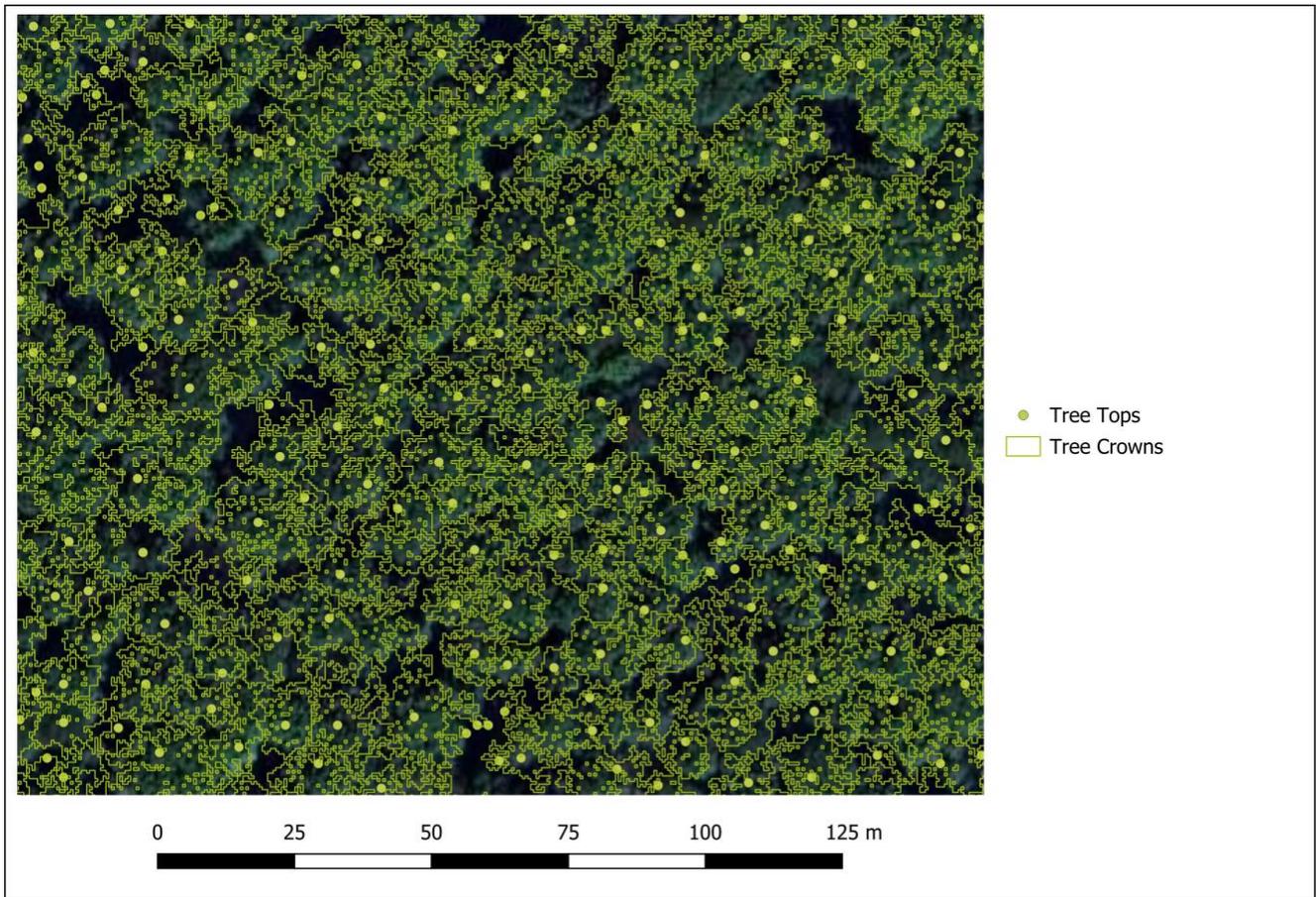
Figure 4: First attempt at running the R script for canopy delineation prior to smoothing
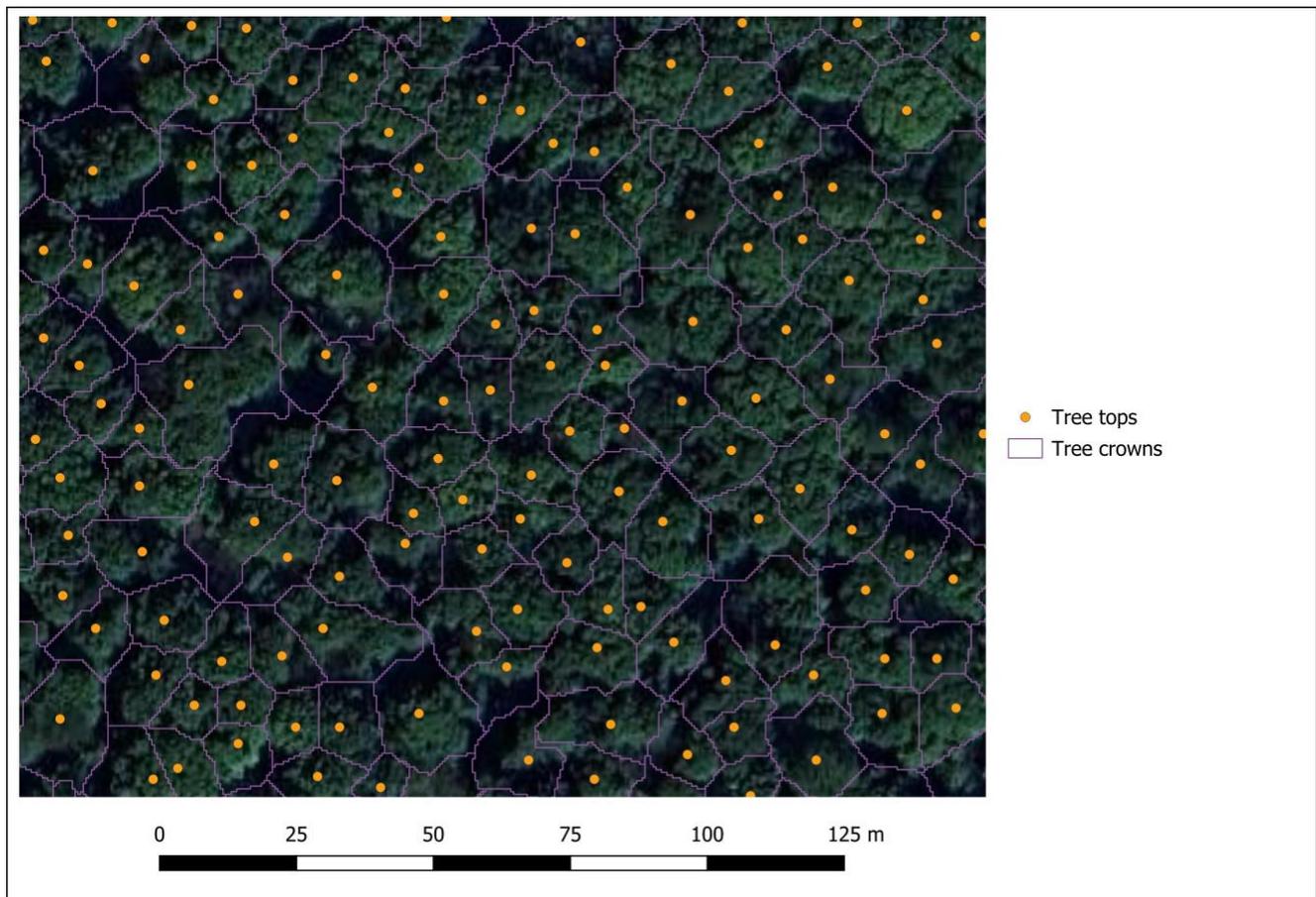
Figure 5: Crown delineation output post-smoothing.

After verifying that the output lines up with what was seen on the satellite imagery well enough for our purposes, a join-by-location operation was performed in QGIS in order to generate a polygon layer in which individual tree polygons include a value for height.

Validation and Results

The model successfully predicted the locations of tracts of woods which were showing characteristics of old growth.

An output of predicted locations for trees >21 m tall was exported to a geoPDF. Myself and Peter went out to visit three of these areas and recorded the height (m), canopy spread (m) and diameter at breast height (cm), as well as species, x and y coordinates and elevation. Vertical and horizontal accuracy for each location was also recorded.

This data was then compared to the CDM output as well as the CHM. Locations of validation data points were loaded into QGIS and converted to a polygon shapefile, where features were represented as circles whose diameters were the recorded canopy width plus double the value of the recorded horizontal accuracy. Zonal statistics were calculated for the smoothed CHM within the validation

windows, reporting the minimum, maximum, minority and majority values of the CHM. Overall, there was a 25% disagreement between CHM reported maximum heights and clinometer measured values.

Overall there was a 25 percent disagreement between CDM reported values and clinometer measured values, with clinometer measured values being, on average, 25 percent higher.

After going back to the drawing board, I determined that the large disagreement was related to the Gaussian filter applied during the smoothing process. Using the polygons delineated in the CDM I used QGIS to run a zonal statistic analysis on the non-smoothed CHM and replaced the existing height value with the maximum value reported by zonal statistics.
When the heights for the CDM were determined using the CHM which had not been smoothed, disagreement between the CDM and measured values was reduced to ten percent.

Canopy_Delin_test
Crown heights (m)
- 8 - 10.8
- 10.8 - 13.1
- 13.1 - 15.3
- 15.3 - 17.6
- 17.6 - 19.8
- 19.8 - 21.9
- 21.9 - 24.3
- 24.3 - 29.7

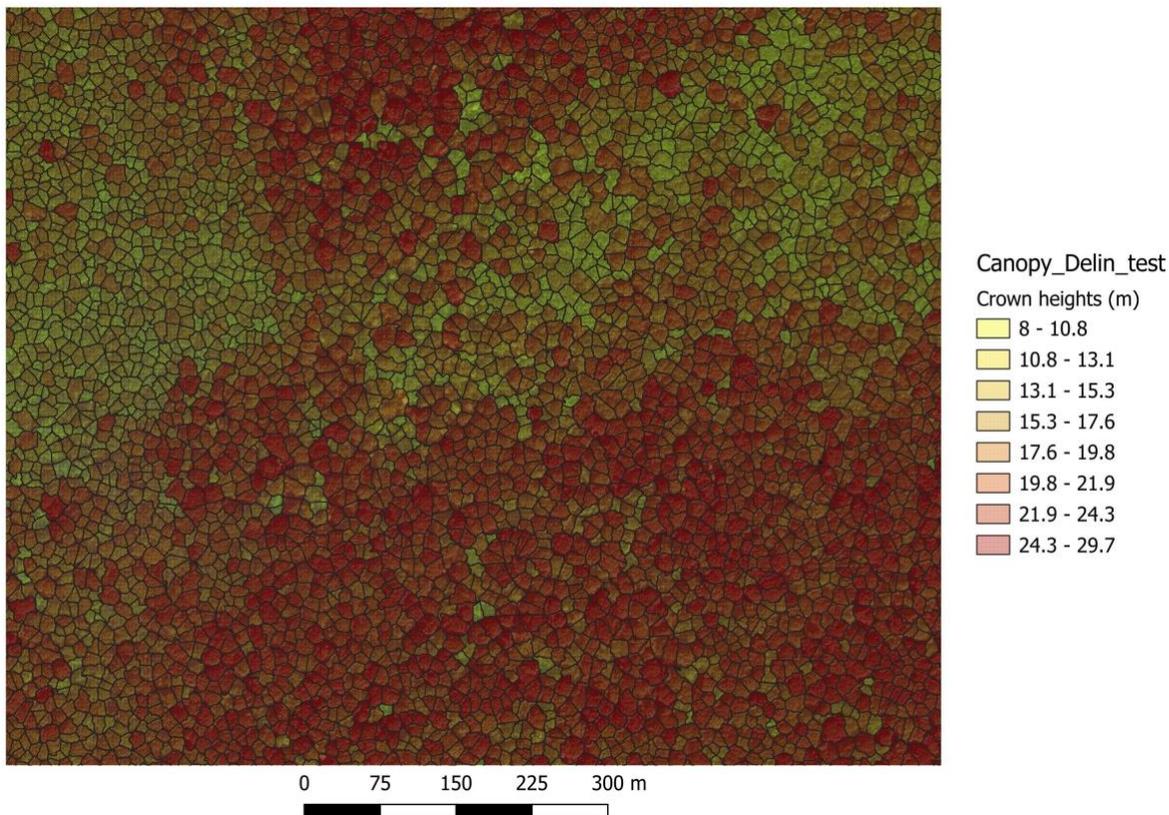0   75   150   225   300 m

Figure 6: Delineated crown layer with height data. This layer is now able to have search terms applied to it in order for trees to be highlighted according to characteristics such as height or crown diameter.

Figure 7: Depiction of a split tree error, in which the CDM interpreted a single tree as multiple trees. Possibly due to issues in the smoothing process

Figure 8: Depiction of error related to tile edge. The CHM was too large to process all at once, so it was split into pieces. Where the pieces touch one another tree canopies have been split between one tile and the other. Solutions to this are theoretically possible but are currently beyond the scope of this project

Figure 9: Depiction of error related to multiple trees being read as a single tree. It was noted that this error was most common among shorter canopy heights.

Figures 7 through 9 depict errors in crown delineation. These errors seem to cluster around trees of similar size category. They interfere little with analysis related to percent coverage of a particular size class or proximity of habitat features to particular size classes but they do have implications for operations which depend on the ability to count individual trees, such as density. Therefore it is not recommended to apply the model as it currently is towards density measurements or to apply statistical analysis to the canopy width value attached to the CDM.

<u>Conclusion</u>

The initial research question, "What is it that makes the North Tiny Forest so special?" does not have a short answer, but hopefully the assembled spatial data has pushed us closer towards having clear responses to that question which are grounded in thoughtful analysis and real-world observation. This project succeeded in generating a workable canopy delineation, putting together a data package which can be of immense use to staff or volunteers who are inclined to use it, as well as an information package which the HLC can use as part of their advocacy in the region.

Interacting with members and associates of the HLC has been educational and enjoyable and I foresee myself volunteering my time towards their projects. If any Trent Students or HLC staff or volunteers wish to contact me with questions related to this project they can feel free to reach out to me at benkrawec@gmail.com.

References

Https://zenodo.org/records/6637599. (n.d.). [Dataset]. https://doi.org/10.5281/zenodo.6637599
Jucker, T., Fischer, F. J., Chave, J., Coomes, D. A., Caspersen, J., Ali, A., Loubota Panzou, G. J., Feldpausch, T. R., Falster, D., Usoltsev, V. A., Adu-Bredu, S., Alves, L. F., Aminpour, M., Angoboy, I. B., Anten, N. P. R., Antin, C., Askari, Y., Muñoz, R., Ayyappan, N., … Zavala, M. A. (2022). Tallo: A global tree allometry and crown architecture database. Global Change Biology, 28(17), 5254–5268. https://doi.org/10.1111/gcb.16302

LAKE SIMCOE REGION CONSERVATION AUTHORITY. (2017). Significant Groundwater Recharge Area (SGRA) (Dataset No. CA_INFO_GIS_V_SWP_SGRA) [Shapefile]. https://simcoe.ca/explore/geographical-information-systems-maps/

Land Information Ontario. (2021a). OHN_WATERBODY [Dataset]. https://geohub.lio.gov.on.ca/

Land Information Ontario. (2021b). OHN_WATERCOURSE [Dataset]. https://geohub.lio.gov.on.ca/

Ontario Geological Survey, Ministry of Northern Development and Mines. (2003). Sgu_polybedrock_polygon [Dataset]. https://geohub.lio.gov.on.ca/

Ontario Ministry of Natural Resources and Forestry. (n.d.). Ontario Classified Point Cloud—Lidar Derived—2022LHURONGEORGIANBAY [LiDAR point cloud]. Retrieved October 15, 2024, from https://data.ontario.ca/dataset/ontario-classified-point-cloud-lidar-derived

Ontario Ministry of Natural Resources and Forestry. (2019). Land Information Ontario (LIO) Warehouse Open Data Products (Composite File Geodatabase) [Dataset]. https://geohub.lio.gov.on.ca/documents/10685ba12bcc48f1a45525fd8d67e1ba/about
Ontario Ministry of Natural Resources and Forestry - Provincial Mapping Unit. (n.d.-a). Non_Sensitive—WETLAND [Dataset]. https://geohub.lio.gov.on.ca/

Ontario Ministry of Natural Resources and Forestry - Provincial Mapping Unit. (n.d.-b). Wetlands [Dataset]. Retrieved September 15, 2024, from https://geohub.lio.gov.on.ca/datasets/mnrf::wetlands/about

Simcoe County. (n.d.). Forests—Simcoe County. https://simcoe.ca/explore/forests/