ForestNow Forest Health Monitoring Using AI and Satellite Images in the Romeo Malette Forest

Project: 10B-2018 - Forest Health Monitoring from Satellite Capture and Machine Learning

A project funded through the Knowledge Transfer and Tool Development program administered by the Forestry Futures Trust Committee

Global Surface Intelligence Ltd.
125 Princes Street, Edinburgh
Scotland, UK, EH2 4AD

Registration No. SC43901

https://www.surfaceintelligence.com

Peter Young
Global Surface Intelligence
125 Princes Street
EDINBURGH
EH2 4AD
Scotland

December 30th, 2020

Ms. Shelley Vescio
Programs Coordinator
Forestry Futures Trust Committee
Suite 2003 - 1294 Balmoral Street
Thunder Bay, ON
P7B 5Z5

Dear Ms. Vescio

We are pleased to submit the final report for the project 10B-2018 *"Forest Health Monitoring from Satellite Capture and Machine Learning"*. Please find the enclosed document and final progress report.

As mentioned, we will be in touch with the Ministry of Natural Resources and Forestry to present the results of this project directly.

We would like to thank the Forestry Futures Trust Committee to allowing GSI to participate in the Knowledge Transfer and Tool Development program.

Please feel to reach us if you have any questions or comments.

Sincerely,

Peter Young

Interim CEO

Executive Summary

In this project, GSI attempted to determine if it was possible to use artificial intelligence to analyse satellite imagery to monitor the health of the forest by detecting areas of concern with the end goal of reducing or eliminating the Ministry of Natural Resources and Forestry (OMNRF)'s current annual aerial survey method. GSI has used its models powered by its patented machine learning algorithms successfully in many forestry applications across the globe over the years. By adapting these algorithms and using reference data from the OMNRF, we turned our focus to detecting areas of tree damage caused by either biotic or abiotic factors (excluding fire). We applied three separate methods to detect this damage using:

1. LiDAR Reference Data: Train with height data to identify areas of significant structure change which is more typical of abiotic damage caused from several weather events.
2. Change Detection Across Time Periods: Adapting our fire detection algorithms, to compare a pre-damage period to a post-damage period and comparing for differences outside of normal expected changes (e.g. leaf-on vs leaf-off)
3. OMNRF Aerial Damage Survey Data: Using the OMNRF annual aerial damage surveys to train our models we attempted to predict in areas where we reserved the survey data and/or in the same area the year after.

Overall, the results were less than satisfactory; however, Method #1 did show promising results by clearly mapping the outline of annual harvest which we see as a direct resemblance to damage such as windthrow. As a result, it is strongly plausible to expect this method would be successful; however, the OMNRF damage reference data did not contain any visually recognizable abiotic damage (using 20 cm resolution imagery) to validate this hypothesis.

The other two methods did not predict any discernible patterns of damage when compared to the OMNRF damage reference data. For Method #3, we believe that the generalized polygons identifying damage are too general in the sense that they likely include both damaged and un-damaged trees within; therefore, our models are likely predicting the presence of both conditions resulting in a more random scatter.

GSI suggests that if the OMNRF has further interest in developing an automated method to detect areas of potential damage, they should focus on improving the data quality and location precision of the damage data. GSI would invite the opportunity to continue working with the OMNRF on such improvements.

Overall objectives

The original goal of this project was to determine if it was possible to detect forested areas that may be affected by some pathogen (biotic or abiotic) through satellite-based remote sensing technology combined with artificial intelligence models. Ideally, forest pathogens or damage can be detected early enough to allow for mitigation measures to prevent further damage (e.g. harvest) or better, possible treatment to prevent further decline in health (e.g. aerial spray) and help prevent further forest timber losses.

The project used three different methodologies each to be applied in a two-stepped approach. We first attempted to detect forest damage in a yearly manner, then if successful, we would move onto a finer time scale of a within-year approach.

To train our ForestNow platform, we used reference data provided by the Ministry of Natural Resources and Forestry (OMNRF) which included LiDAR, aerial damage surveys, and historic forest inventory data. This reference data is used for both training the ForestNow machine learning platform and validation of results.

This project timeline spanned across two time periods. The Year-1 portion covered the years from 2015-2018 and the Year-2 portion added the year 2019.

Deliverables

Deliverables are in raster format. The raster format is pixel-based (10m resolution) which is an excellent format for displaying results and are flexible as it could allow forestry analyst to auto-delineate areas of concern based on specific thresholds of damage. If such a threshold is specified, GSI has the capability to perform this auto-delineation to provide maps of damage areas.

Area of Interest

The area of interest (AOI) used in this project is the Romeo Malette Forest (Figure 1).

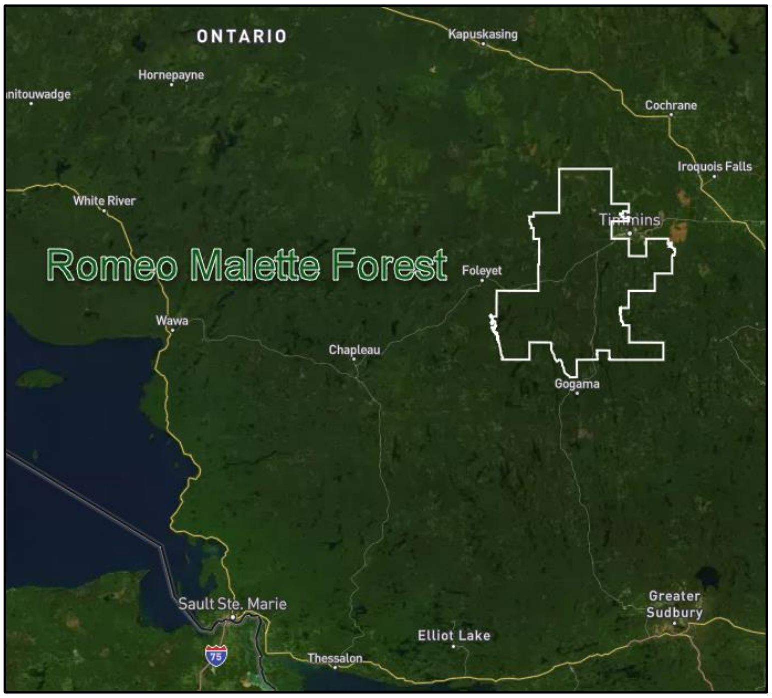

Figure 1. Map showing the area of interest (AOI); the Romeo Malette Forest (RMF).

Romeo Malette Forest, Ontario

The Romeo Malette Forest (RMF) was chosen as suggested by the OMNRF since there were several concurrent trials occurring on that forest with the acquisition of new single-photon LiDAR, and there is currently a concerning spruce budworm outbreak occurring in the northwest portion of this forest. This area was also used as part of two other project carried out by GSI.

Target Audience and Benefits

All forest stakeholders can benefit from monitoring forests for damage since early detection allows for more treatment options. Such treatments could lead to reducing spread and/or salvage harvesting timber before it decays beyond minimum merchantable quality specifications. As such, the following stakeholders will no doubt have an interest in forest health monitoring data:

- Government of Ontario – Ministry of Natural Resources and Forestry (OMNRF)
- Forest Industry Companies
- Government of Canada – Food Inspection Agency

Model Data

Satellite Images

GSI used both reflectance (multi-spectral bands) and synthetic aperture radar (SAR) from various publicly available satellite images (e.g. Sentinel 1 and 2, Modis, Landsat, etc). Reflectance data provides a much broader range of useful data; however, since it cannot penetrate through clouds, its frequency of clear usable scenes can be limiting. GSI has internal processes to reduce the impact caused by cloud cover from supplementing the model with the use of partially clear scenes. SAR on the other hand, can penetrate through cloud (therefore more frequently reliable); although, the data bands it provides do not allow for an in-depth analysis compared to reflectance bands.

GSI continuously ingest images throughout the calendar year and as a result, it can detect changes occurring through the year which may be unique to a specific species being affected by a pathogen for example.

This method is like our species identification process which was explored in another Knowledge Transfer and Tool Development project. Please refer to the following document for more process and result details: http://www.forestryfutures.ca/upload/464883/documents/69DB663DDC7DB0E8.pdf

Training Data

The primary source for the training data was provided by the OMNRF; specific data used was as follows:

- Lidar Point Cloud: Captured in 2018 using Single Photon (~25 points/m2).
- Lidar-Based CHM model (0.5m x 0.5m)
- OMNRF Annual Health Aerial Surveys: A polygon shapefile identifying areas of damage by specific biotic or abiotic factor.

Methodology

Approaches

GSI applied three methods to assess the feasibility for identifying biotic or abiotic damage to the forest; though, excluding fire as per the OMNRF's guidance. Each of these methods are further described below. The general approach for each method was to first determine if it was possible to positively identify a year-to-year detection of damage. If possible, then we would explore whether it was possible to detect damage within-year (excluding Method 1).

GSI uses a regression-based approach to generate our raster layer predictions which involves training with examples of known conditions or measurement such as aerial survey data at specified locations or tree attributes derived from LiDAR survey data. The training data is then fed into the system along with satellite imagery for the same locations within an acceptable timeframe inline with the training data.

This regression approach is more effective and flexible compared to an alternate approach which is a classification-based analysis. Regression provides a precise quantitative approach where it predicts a continuous range of data for every pixel. For example, in this project, the range of values for each pixel is a gradient from 0-100% which represents the probability that the pixel has a presence of tree damage or in the case of tree heights, a predicted height in meters.

In the original proposal we planned on analysing multiple years which included 2015; however, due to the absence of images, the analysis of 2016 was not possible. Since this project involves comparing one year to the next, the first possible predicted outcomes will be for 2017 as a result.

Method 1 - Annualized height estimates using LiDAR

This method consists of using differences in height estimate between years to find areas that have been disturbed. This method is more suited to find disturbances such as blowdown or ice damage which was the focus of this project. This method could also be used to identify burnt or harvest areas; however, those two sources of disturbances were explicitly excluded as per OMNRF's directive.

GSI used the 2018 LiDAR data to derive a Canopy Height Model (CHM) and we used that 2018 training to predict heights for 2016, 2017, and 2018 using images from each of those years. To avoid an over-fitting to 2018 reference data, we use a 50:50 split method which is training on one half of the data (using an east-west split) to train the model then predict on the other half. We repeat doing the reverse split and combine both predicted halves together to produce an independently derived height model for 2018 to create a result that avoids predicting and training on the same pixel.

With a height model for each year, GSI compared between years to attempt to detect areas of significant disturbance inline with the areas identified by OMNR aerial survey.

Method 2 – Detect annual damage from year-to-year change detection

GSI extrapolated from a previously developed methodology used in detect forest wildfire damage, where we attempted to modify algorithms to detect changes from one year to the next. This method does not need training data to initialize, rather it based on a comparison of phenology differences from one year to the next while trying to identify areas falling outside a "normal" expected range of variation. This method on its own will not detect the source of the change; however, its results could be used to highlight areas where further resources should be deployed to investigation further.

Method 3 – Using OMNRF aerial survey data to predict next year's damage

GSI used the OMNRF annual health aerial survey data of a given year to train with, and then apply that training to satellite images of the next year to identify potential new areas of damage. Using the health data of that next year, we compared the effectiveness of our prediction against this real data.

We also attempted to predict the next year's damage by sub-dividing the year into quarters to see if detection is also noticeable at that time scale.

Using the OMNRF annual health aerial survey data, GSI converts this data from a damage/no-damage dichotomous raster coverage for each source of damage. Then using the regression approach, GSI predicts a gradient of damage (0-100%) in a wall-to-wall fashion for every 10-meter pixel of the AOI. The result is a damage probability layer for each source of damage.

Validation

Validation of results is a crucial step in assessing the accuracy of the results and the most important factor is the independence of the validation.

The annual health data provided by the OMRNF is in polygon shape which is collected via human observation from a fixed-wing flight flown a grid pattern from which data is recorded at a relatively high speed. GSI sees two potential concerns with this data that became more obvious as we progressed through this project:

1) The areas outlined as damaged are likely generalized to include damage and non damaged locations. We conjecture that there is a high likelihood that the polygons include nondamaged trees and thus translates to presence of damage information with low precision for specifically identifying damaged sites.

2) There is a clear north-south linear pattern that is visually noticeable from the mapped patterns of observed damage (Figure 2). We conjecture this is likely the result of the flight grid and suggest there may some "blind spots" in the recorded data.

Figure 2. A sample area of the recorded 2018 forest health aerial survey results showing areas of recorded damage. A noticeable north-south linear pattern is obvious.

We assume the data to be 100% correct and use it for model training and cross-validation. The impact of this initial assumption is two-fold: 1) model training with this data will likely include both false-positives and false-negatives, which would lead to a less precise training, and 2) cross-validating with this data could thus result in what appears to be false-positives in the survey data "blind spots" mentioned previously.

The GSI machine learning algorithms are very robust and can still predict reasonable result despite having imprecise data; however, due to the nature of the reference data, the validation for this project must be largely based on a subjective visual observation which we will present visuals to support our observations and conclusions.

Results

Between-Year and Within-Year Damage Detection

We tested both approaches of attempting to detect damage across years and within-year. The between-year approach yielded some mix results. However, there were some compelling results for some of the years where the outcomes were predicted which is further detailed in each of the below sections outlining the Methods separately.

Once we ran the between year approach, we then identified areas showing strong correspondence between predicted and observed damage. We then tested the within-year approach for areas identified for change between years. We completed a total of 76 different iterations of within-year assessments. the results here did provide any notable change in vegetation state within year that we could identify. These within-year results are not documented further because we were unable to extract anything valuable for change detection.

Method 1

Method 1 uses LiDAR as its raw training data. This application yielded some very good results in line with the original intent of this part of the project. This method was primarily to detect windthrow and/or ice damage. However, the reference data provided by the OMNRF for the RMF was imprecise as mentioned above and were thus inadequate for GSI to use as validation for this method. We also carried out a visual inspection tree damage in 2018 with the 2018 high-resolution (20-centimeter) provided by the OMNRF. Figures 4 and 5 below are some examples of OMNRF disturbance polygons identifying wind and ice damage. They both showed the disturbance was visually unidentifiable or undetectable with the between year approach.

Figure 4. Example showing an area identified as having wind damage from the annual aerial survey overlaid on a 20-centimeter 2018 image where no noticeable damage is identifiable.

Figure 5. Example showing an area identified as having ice damage from the annual aerial survey overlaid on a 20-centimeter 2018 image where no noticeable damage is identifiable.

As a result, we focused our attention on analysing harvested areas as a surrogate for this type of damage instead since those areas have a clear height differential before and after. Results for this test were excellent with a clear differentiation between years and comparison to annual update polygon shapes as shown in the series of images in Figure 6. The bottom right image shows minimal differences between GSI's predicted results versus the OMNRF's reference data.

An important thing to highlight about this method is that GSI can predict heights in years where there is no LiDAR reference data available. The LiDAR data available was from the year 2018; however, GSI predicted heights for 2016, 2017, and 2018 with enough accuracy to clearly outline where harvest has occurred across years. This is highlighted by the results of the implied harvest for the year 2017 which is derived from the height difference between GSI's predictions of the 2016 and 2017 years (Figure 6, image #2).

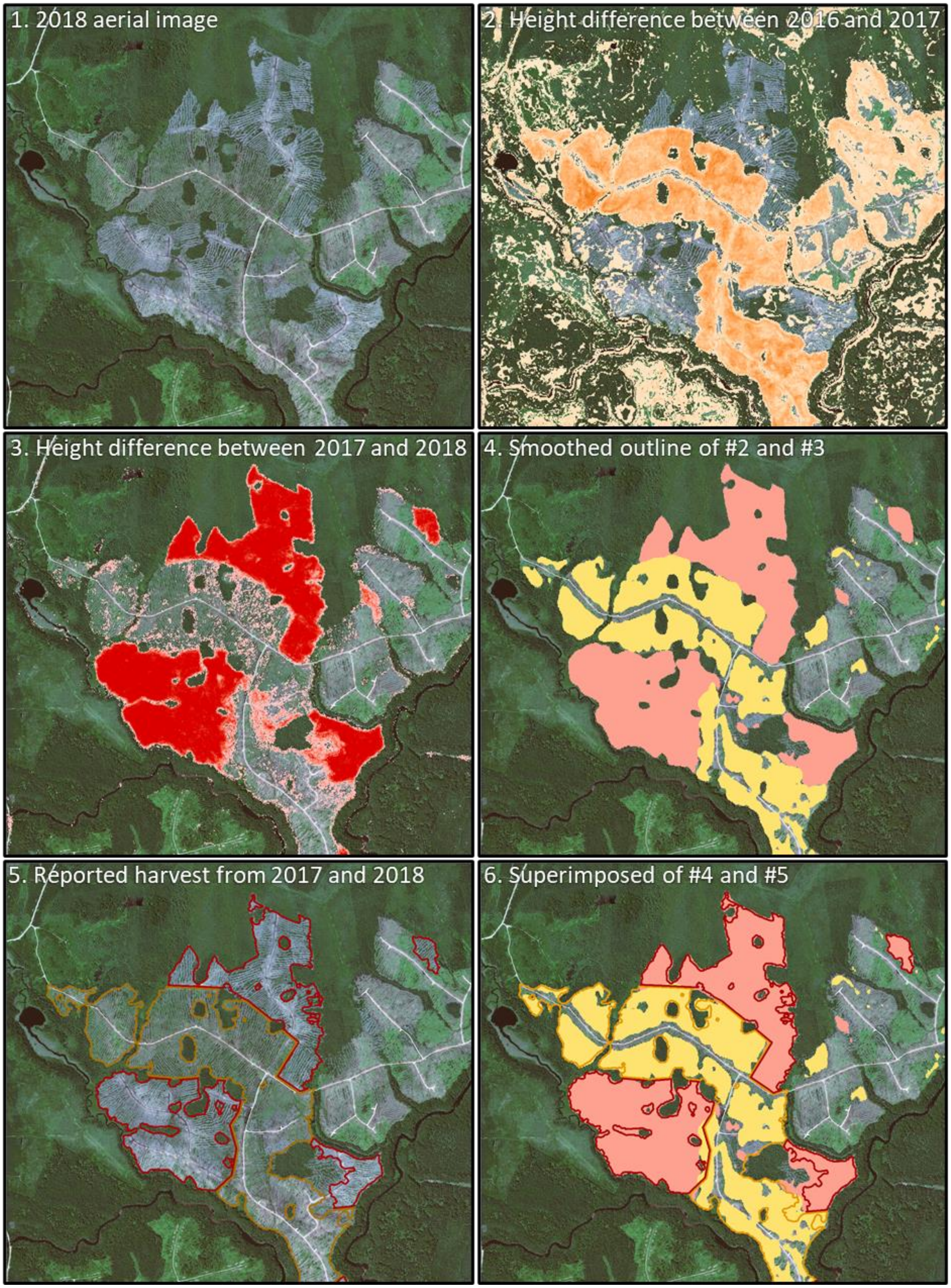

Figure 6. Shows a progression of steps using GSI's method of training with 2018 LiDAR and comparison to annually reported harvest polygons (from top to bottom and left to right). 1) A 20-centimeter resolution aerial image taken in 2018. 2) Gradient of height difference between 2016 and 2017 satellite images. 3) Gradient of height difference between 2017 and 2018 satellite images 4) Polygon shapes of harvested

areas by applying a threshold and a smoothing filter to #2 and #3. 5) The outline of the annual report harvest polygon shapes. 6) An overlay of #5 overtop #4 to show the comparison of the auto-delineation with reported harvest.

Method 2

GSI tested changed detection in a comparison methodology where no training data is used. This method compares one time period to another and tries to detect any significant phenology changes between the two periods. This method was adapted from one of our existing models where the algorithms clearly identified the pre- and post-change from wildfires; however, in the case of this project, we were not able to detect changes with any discernable pattern that was close to the areas of damage provided by the OMRNF. We were able to detect harvested areas where the change is much more drastic; however, we found that our Method 1 was still more effective than this one.

The results from another KTTD showed that we could predict relatively stable results for tree attributes year-on-year for the same harvest blocks which suggest GSI can distinguish between various conditions effectively. The results from GSI's *"Project: 9B-2018 - Post-Harvest Surveys from Satellite Capture and Machine Learning"* showed that we could successfully predict tree species composition and attributes (i.e. height, density and stocking) with minimal variation between years. This "normalization" of the signal from satellite images across years for a given area is very important when analysing areas where no change is expected; however, finding changes in areas where the damage does not immediately kill and/or completely defoliate a tree is more complicated since the phenology change can be more temporary. For example, in the case of spruce budworm damage, the insect typically only eats the current year's foliage; therefore, the host tree species (spruce and fir) do not permanently change color until several years of heavy defoliation where the trees succumb, and all foliage turns brown.

However, we believe that the damage from the biotic and abiotic in this forest and/or time period were too subtle to detect. The variations from one year to the next remained within our models' thresholds; therefore, no areas of damage were detected. This phenomenon is likely due to significant differences across year's in the: 1) within seasonal phenology changes, 2) year-to-year differences in atmospheric differences, and 3) timing of images causing too much "noise" in the analysis.

Method 3

Predicting 2018 Damage

Given our limited successes using Methods 1 and 2, GSI took a stepwise approach where we started with scenarios where there was the best chance at success. This approach meant that we started with the years that had the most cloud-free satellite images throughout the growing season which is based on our previous experience on other projects. We only analyzed damage from the eastern spruce budworm and the large aspen tortrix as they were the only 2 with enough reference data. Only the spruce budworm damage prediction showed any sign of plausibility; therefore, the results are presented for this type of damage only.

The following were the stepped approach taken and observations of the results:

1. Based on our preliminary assessment, we initially started by using the 2017 OMNRF damage data to train from with 2017 images to predict damage in 2018. This year pairing had the most abundant satellite images in each year. Unfortunately, the results were disappointing with minimal visual correlation against the OMRNF damage reference data from 2018. The predicted result had no discernible pattern to the reference data. All four steps were completed; however, all gave similar results.
2. With the results of step 1 in hand, we then tested whether it was possible to train and predict accurately for data in the same year. We conjectured this approach would elicit a determination as to whether any correlation existed at with the OMNRF's reference data. However, same-year

analysis does require model over-fitting to be controlled. Overfitting occurs when training and predicting is completed on the same pixels as areas identified as damage and when no data is withheld for independent cross-validation. Without independent cross-validation the accuracy measure of a training model alone obfuscates accuracy where there is no refence training data. (see Figure 7). We ran our initial analysis on 2018 data as it also had the most cloud free sentinel images. Figure 8 depicts the mapped results which show improved visual matching compared with the earlier approach that trained in one year and predicted in another. With this said, there seeming appears to be too much random presence prediction scattered outside known presence locations. In the Year-2 component of the project, we then explored using the 2018 predicted damage data from the satellite imagery and based on the OMNRF damage data to train and predict potential damage using 2019 satellite imagery. the predicted data in 2019 was then compared to the 2019 OMNRF damage data as cross-validation. We tested two new methods but neither showed results did with any significant consistency with the 2019 reference data. The two attempts were:

a. Training with the 2018 OMNRF damage data in the original polygon form to predict results for 2019 (Figure 9). No visual correlation was deduced from this attempt.
b. Same scenario as a) with the difference of using the 2018 results of step #2 instead of the original reference data from the OMNRF. (Figure 10). Similarly, no visual correlation was deduced from this attempt.

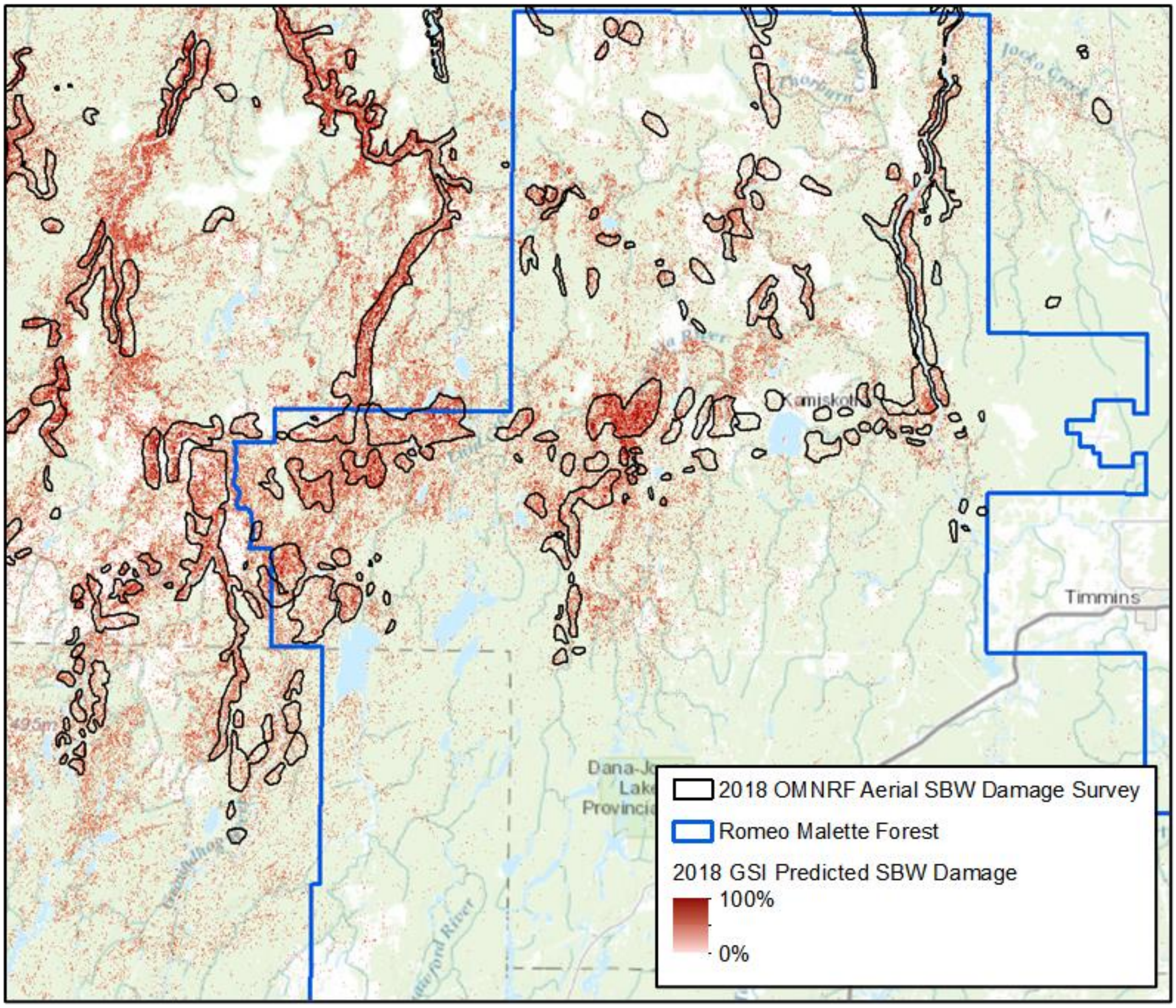

Figure 7. Shows the results of GSI's 2018 predicted proportional damage by training the 2018 OMNRF damage data and predicting on 2018 satellite imagery overtop the same area. There is a clearly a visual correlation of the results compared to the OMNRF 2018 reference data which is a result of overfitting.

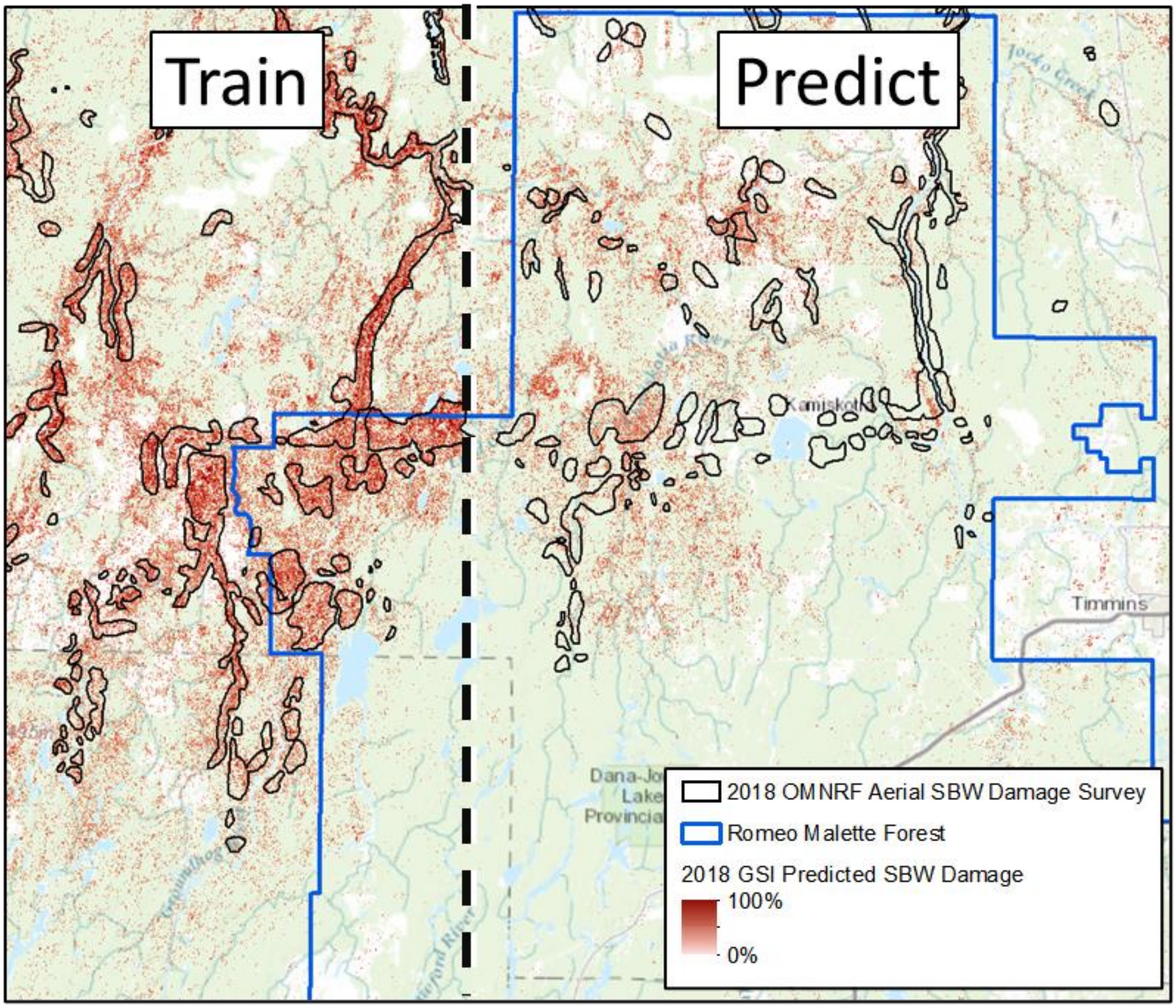

Figure 8. Shows the results of GSI's 2018 predicted proportional damage by training with the 2018 OMNRF damage data and predicting on 2018 satellite imagery where areas of training and predicting are separated for a more independent process. There is a clearly a visual correlation of the results compared to the OMNRF 2018 reference data on the left side caused by overfitting while the right side does show a weaker correlation; however, demonstrates some plausibility from an independent prediction.

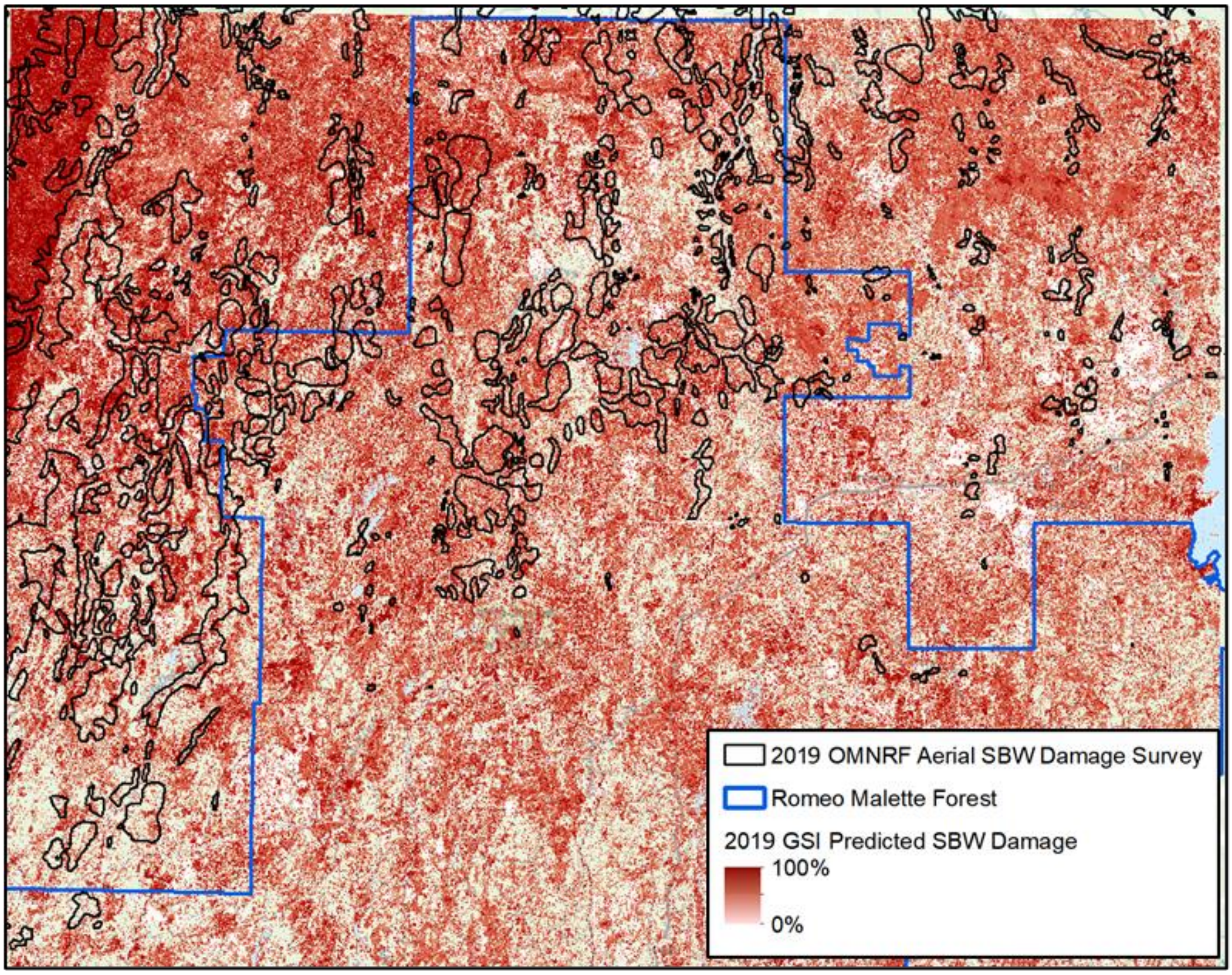

Figure 9. Shows the results of GSI's 2019 predicted proportional damage by training with the raw (polygon) 2018 OMNRF damage data on 2018 satellite imagery and predicting on 2019 images. No visual correlation between the results and the reference data can be deduced.

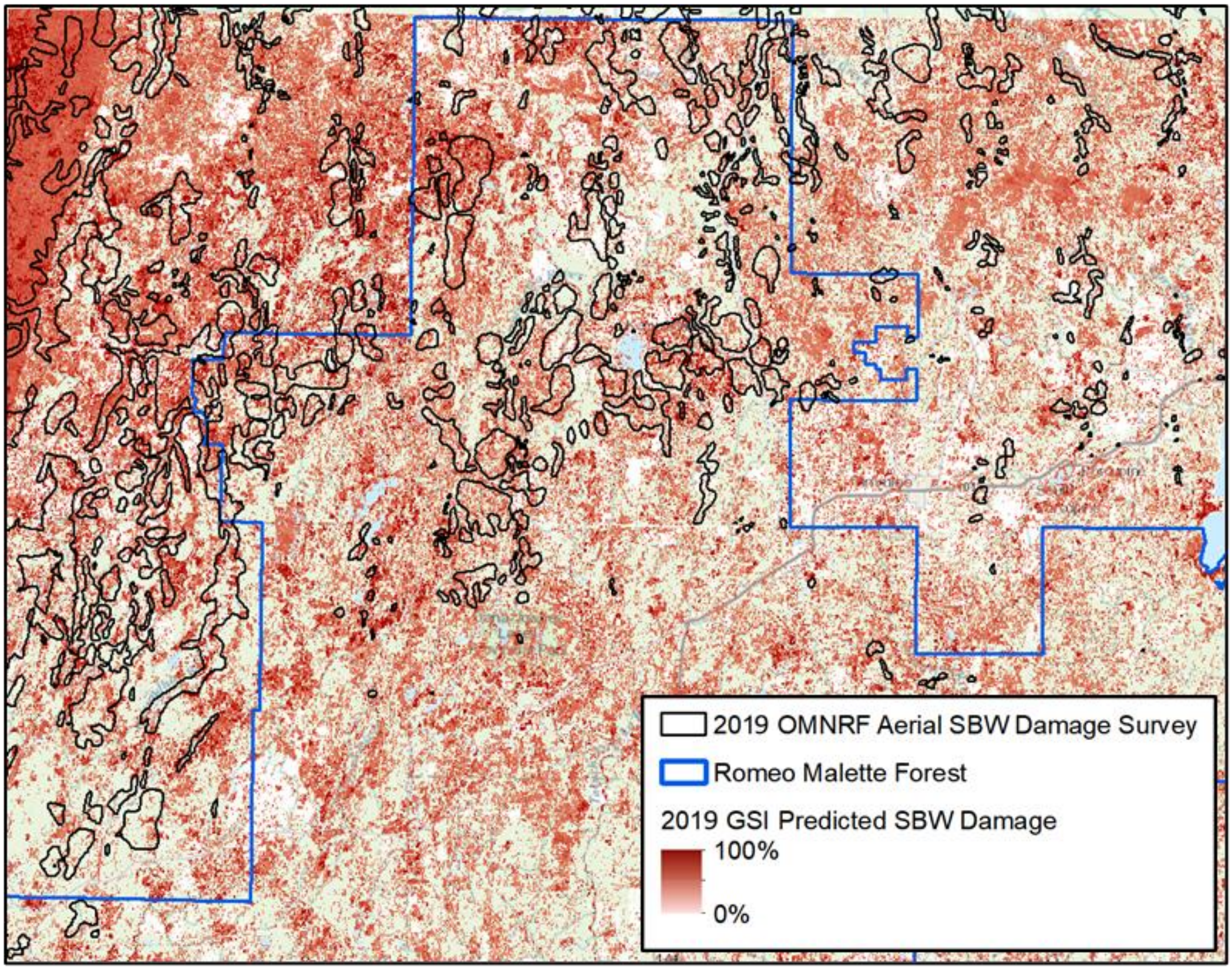

Figure 10. Shows the results of GSI's 2019 predicted proportional damage by training with pre-processed (raster) 2018 OMNRF damage data on 2018 satellite imagery and predicting on 2019 images. No visual correlation between the results and the reference data can be deduced.

Conclusion

We were able to show some consistency in tracking between disturbance linked to harvest which demonstrates plausibility for more significant damage events like windthrow. However, we were unsuccessful in using any of the OMNRF's damage data as source to train and predict with. We tested running models directly, in aggregate of polygons, training in 1 years and predicting the next, and training in the same year. In all cases we had low quality results. We believe the specific reason for this was because the ONMRF data polygons themselves likely have low precision and consist of amalgamation of locations that include presence and absence of each driver of forest / tree damage. GSI suggests that if the OMNRF had further interest in developing an automated or smart method to detect areas of potential damage- OMNRF should focus on improving the data quality and location precision of the damage data. GSI would invite the opportunity to continue working with the OMNRF on such improvements.