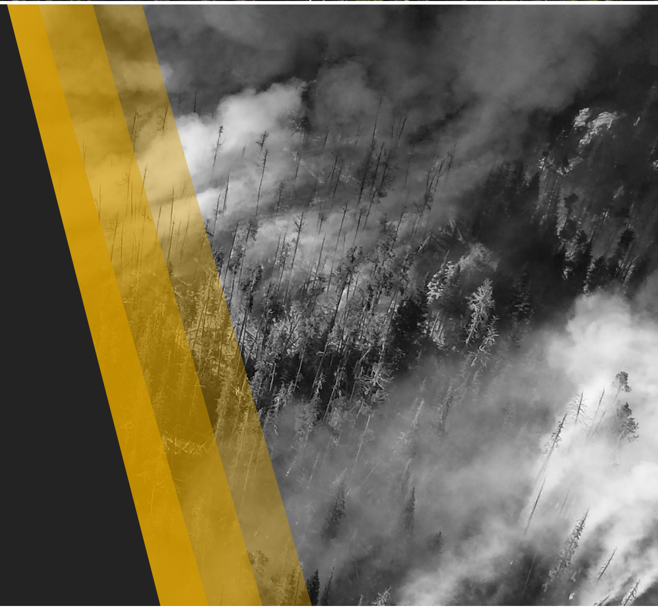
A spatial database of historical wildfire and timber harvest in the boreal Area of the Undertaking of Ontario: The methodological framework



Technical Report TR-37

Science and Research Branch

Ministry of Natural Resources and Forestry

Ontario

Science and Research Technical Report TR-37

A spatial database of historical wildfire and timber harvesting in the boreal Area of the Undertaking of Ontario

The methodological framework

M. Ouellette[1], T.K. Remmel[2], and A.H. Perera[1]

[1] Forest Landscape Ecology Program, Ontario Forest Research Institute, 1235 Queen Street East, Sault Ste. Marie, Ontario, P6A 2E5

[2] Department of Geography, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3

2020

Science and Research Branch

Ministry of Natural Resources and Forestry

Copies of this publication are available from info.mnrfscience@ontario.ca.

Cette publication hautement spécialisée, A Spatial Database of Historical Wildfire and Timber Harvesting in the Boreal Area of the Undertaking of Ontario: The Methodological Framework, n'est disponible qu'en anglais conformément au Règlement 671/92, selon lequel il n'est pas obligatoire de la traduire en vertu de la Loi sur les services en français. Pour obtenir des renseignements en français, veuillez communiquer avec le ministère des Richesses naturelles et des Forêts au info.mnrfscience@ontario.ca.

Cover photo: skeeze (pixabay.com/images/id-1161868/)

Some of the information in this document may not be compatible with assistive technologies. If you need any of the information in an alternate format, please contact info.mnrfscience@ontario.ca.

Cite this report as:

Ouellette, M., T.K. Remmel and A.H. Perera. 2020. A spatial database of historical wildfire and timber harvesting in the boreal Area of the Undertaking of Ontario: The methodological framework. Ontario Ministry of Natural Resources and Forestry, Science and Research Branch, Peterborough, ON. Science and Research Technical Report TR-37. 31 p. + appendix.

Abstract

We produce a spatial database of boreal forest fire and harvesting disturbances for Ontario's Area of the Undertaking. Disturbances are mapped annually for the period 1972–2018; our workflow allows annual updates to be assessed and incorporated on an ongoing basis. The approach identifies training areas for fire and harvesting either within the scene of interest or moving outward in space and/or time according to a hierarchy of priorities defined *a priori*. An unsupervised ISOData classification on the training data then defines a series of classification clusters that are modelled by a smooth polynomial function to identify a local minima at the point along the classification clusters that best distinguishes disturbances from non-disturbances. This threshold is applied to the scene to map disturbances. This process is conducted independently on individual scenes of Landsat MSS, TM, ETM+, or OLI imagery to produce point maps, with each point representing 1.44 ha. The identification of disturbances is also cross-referenced back through time to ensure that the earliest date for each disturbance is that recorded in the database. Finally, we intersect the identified disturbance mapping with other harvesting and fire databases to assess the accuracy of our product using an ensemble method. An ensemble confidence value is attached to each mapped location in the database. Our methods are documented, based on scientifically defensible and repeatable approaches. We avoid providing disturbance boundaries as these are generally scale dependent, susceptible to transitional or ecotonal characteristics, and will vary with respect to individual needs.

Résumé

Une base de données spatiales sur les feux de forêts et les récoltes de bois dans le passé dans la région boréale du Secteur d'exploitation forestière de l'Ontario : le cadre méthodologique

Nous produisons une base de données spatiales sur les feux dans les forêts boréales et les perturbations dans les récoltes dans le Secteur d'exploitation forestière de l'Ontario. Les perturbations sont cartographiées annuellement pour la période allant de 1972 à 2018; le déroulement de notre travail fait en sorte que des mises à jour annuelles puissent être évaluées et intégrées de façon continuelle. La démarche consiste à établir des zones de formation sur les feux et les récoltes soit sur la scène d'intérêt, ou en se déplaçant dans l'espace et (ou) dans le temps selon une hiérarchie de priorités définies *a priori*. Une classification non supervisée ISOData sur les données de formation définit par la suite une série de regroupements de classification qui sont modélisés selon une fonction polynomiale souple afin d'établir des minimums locaux à un point le long des regroupements de classification qui permet de mieux distinguer les perturbations des non-perturbations. Ce seuil est appliqué à la scène afin de produire une cartographie de perturbation. Ce processus est mené de façon indépendante sur chacune des scènes d'image Landsat MSS, TM, ETM+ ou OLI en vue de produire des cartes de points, où chaque point représente 1,44 ha. La concordance des perturbations relevées est également vérifiée dans le temps afin de s'assurer que la date la plus ancienne de chaque perturbation est consignée dans la base de données. Enfin, nous effectuons un croisement de la cartographie de perturbation relevée avec d'autres bases de données sur les récoltes et les feux en vue d'évaluer l'exactitude de notre produit selon une méthode d'ensemble. Une valeur de confiance d'ensemble est associée à chaque emplacement cartographié dans la base de

données. Nos méthodes sont consignées, selon des approches scientifiquement justifiables et reproductibles. Nous évitons d'établir des limites de perturbation, car celles-ci dépendent généralement de l'échelle, sont susceptibles aux caractéristiques transitoires ou écotonales et diffèreront en fonction des besoins individuels.

Acknowledgements

This work was possible due to the extensive insights, discussions, reviews, comments, and contributions at working meetings by many MNRF employees (in alphabetical order): Graham Atkinson, Jeff Bowman, Den Boychuk, Mike Brienesse, Lisa Buse, Stephen Casselman, Maureen Kershaw, Derek Landry, Rob Luik, Rob Mackereth, Stephen Mayor, Jim McLaughlin, Ed Mick, Dave Morris, Maara Packalen, Bill Parker, Sharon Reed, Geordie Robere-McGugan, Ian Sinclair, Ian Smyth, and Larry Watkins. Similarly, we are thankful for the contributions by the following individuals external to the MNRF (in alphabetical order): Joe Bennett (Carleton University), Yan Boucher (Ministère des Forêts, de la Faune et des Parcs), Carissa Brown (Memorial University), Domgmei Chen (Queens University), Michael Drescher (University of Waterloo), Erik Emilson (CFS), Marie-Josée Fortin (University of Toronto), Benoit Hamel (CFS), Patrick James (Université de Montréal), Brian Kielstra (CFS), Shawn Leroux (Memorial University), Stephen Murphy (University of Waterloo), Andrew Trant (University of Waterloo), Lisa Venier (CFS), Marc-André Villard (Université de Montréal), Peter Vogt (European Commission Joint Research Centre), Yolanda Wiersma (Memorial University), and Mike Wulder (CFS).

Contents

Background

Boreal forest Area of the Undertaking of Ontario

Ontario’s boreal forest is a very large and heterogeneous area. Not only is it the most extensive (about 45 million ha and representing 50% of the province[1]) of all vegetation types in Ontario, but it also includes at least 7 ecoregions due to its geo-climatic variability. Almost all of this boreal forest area, nearly 30 million ha, is in Ontario’s managed forest landscape known as the AOU, or the Area of the Undertaking (figures 1 and 2). Administratively in Ontario, the Ministry of Natural Resources and Forestry (MNRF) is the custodian of the AOU, a vast public landscape that is about 87% publicly owned, with responsibilities for governance of its use: conservation (e.g., parks and protected areas), timber harvesting (e.g., by issuing forest licences), recreation (e.g., by issuing hunting and fishing permits and monitoring by enforcement), and disaster protection (e.g., by managing wildfires and floods).

Figure 1. Map of Ontario showing, in grey shading, the boreal forest portion of Area of the Undertaking (AOU) as modified for our study.

[1] www.ontario.ca/page/annual-report-forest-management-2013-2014

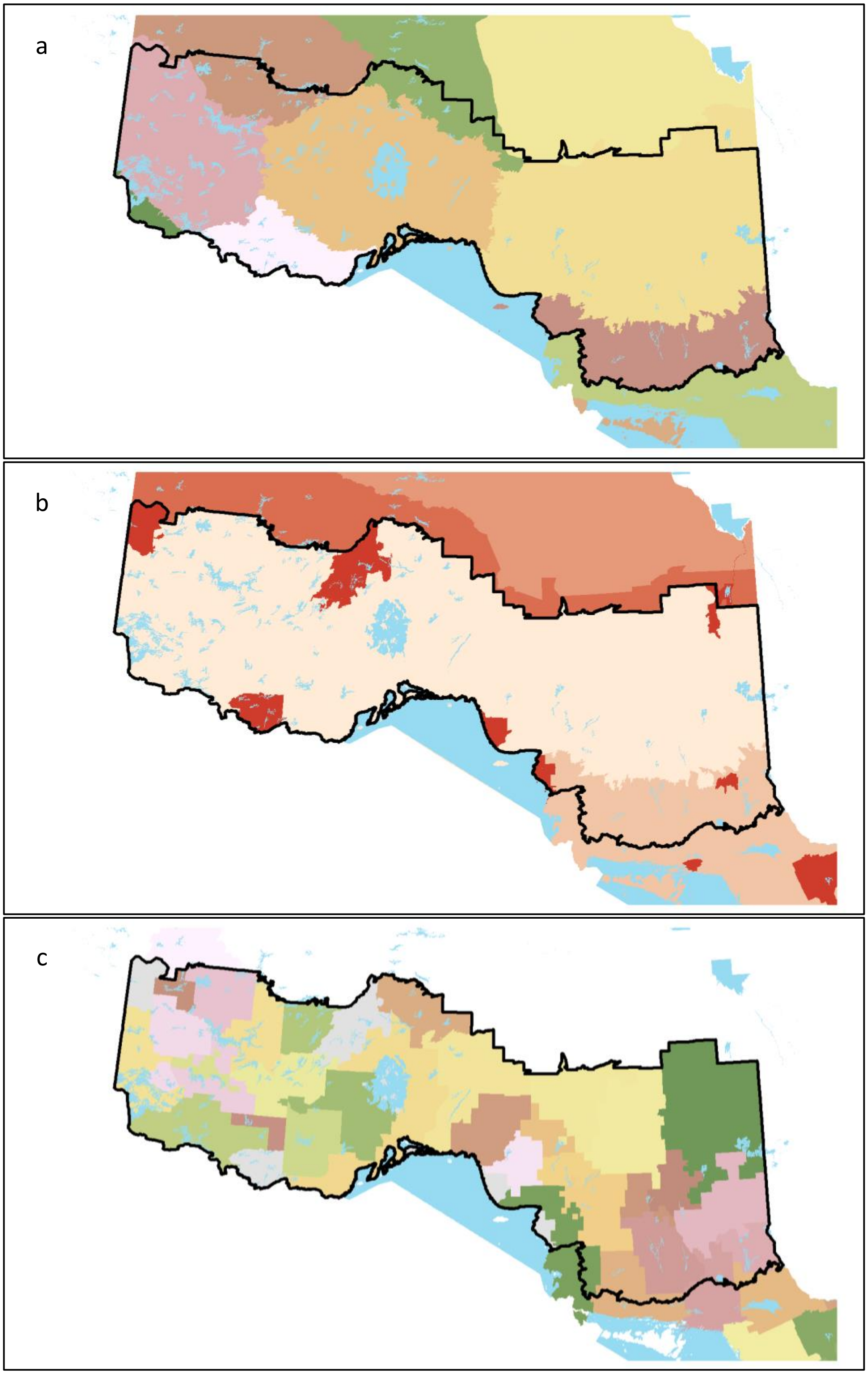


Figure 2. Demonstrating the complexity of Ontario's managed boreal forest region in light of multiple and competing jurisdictions and states: (a) eco-regions, (b) fire management zones, and (c) forest management units. Each combination is subject to different fire pressure, suppression effort, and management approaches.

The boreal portion of the AOU, and the remote areas north of it, are not heavily populated, with only 780,140 inhabitants or 0.9 people per square kilometre.[2, 3] The region is sparsely developed by settlements except for a few urban areas and sporadic townships and farms, so the forest cover is not subject to urbanization and industrialization pressures. However, this region is subject to periodic broad scale disturbances caused by many natural (non-anthropogenic) and anthropogenic agents of change. According to the State of Ontario's Natural Resources – Forests 2016 report (OMNRF 2016), wildfire is the most significant (by the magnitude of frequency, extent, and intensity) natural disturbance, followed by windthrow and floods that cause sporadic and infrequent disruptions of forest cover in relatively small extents. Of the biotic disturbance agents, insect defoliation dominates forest disturbances in extent, though it is infrequent compared to more frequent disturbance by beavers that dam streams and flood considerable area of forest cover. Of anthropogenic disturbances in the boreal forest portion of the AOU, the most prominent is timber harvesting where large extents of forest cover are removed, following the related expansion of the network of access roads. Development of mining activities and hydroelectric reservoirs along with transmission corridors are relatively less common in the boreal forest portion of the AOU except for few isolated locations.

History of wildfire and timber harvesting disturbances

Extensive and frequent wildfire disturbances are an intrinsic component of boreal forest landscapes worldwide. They typically originate from lightning strikes during storms and cause varying degrees of damage to forest cover, ranging from locations of complete incineration to those of partial scorching of vegetation matter (de Groot et al. 2013). These events are ecologically important in many ways, for example, to periodically release nutrients trapped in dead biomass and seeds from serotinous cones, initiate the plant succession process, and create critical habitat for pyrophillic plant and animal species (Nappi and Drapeau 2009). Boreal wildfires also create a shifting spatial mosaic of forest ecosystems that is important to sustain the biodiversity and abiotic fluxes at many ecological scales (Goldammer and Furyaev 1996).

These wildfires have been increased, both in number and area, by people, whether initiated accidentally or purposely, since the boreal forest has been inhabited (Campos-Ruiz et al. 2018). The proportion of human-ignited forest fires has increased over time near settlements, recreational areas, and transportation corridors. Fire suppression activities focus on areas of settlement and development based on defined provincial fire management zones (Figure 2) that are heavily managed to limit effects on people, dwellings, and other infrastructure. While the effectiveness of forest fire management by suppression, in terms of changing the forest fire regime and balancing costs with benefits, is hard to quantify and likely unknown fully (Cumming 2005, Rijal et al. 2018), it can be argued that extinguishing fire ignitions in the intensive fire management zone represents most of the success. Here we refer to all forest fires that altered

---

[2] 2016 census – Statistics Canada: en.wikipedia.org/wiki/Northern_Ontario

[3] Statistics Canada – www12.statcan.gc.ca/

forest cover, regardless of their source of ignition and the level of suppression activities, as *wildfires*.

Timber harvesting activities in the boreal forest must always have been linked with human inhabitation, with collection of firewood and building material occurring even before European settlement. Organized and extensive activities of timber harvesting in the boreal forest had commenced at least a century ago (Lambert and Pross 1967). According to the reports by ministers in charge of crown (public) forest, the income from timber harvest has been considerable since early 1900s and fostered large scale expansion of the timber industry in Ontario (Epp 2000). In time, timber harvesting activities have expanded northwards and westwards from central Ontario, and various policies have been developed to guide timber harvesting operations to ensure the long-term health of Crown forests, broadly guided by the Crown Forest Sustainability Act (1994, S.O. 1994, c.25) and the Environmental Assessment Act (and specifically, Declaration Order MNR-75). In this context, forest management plans for individual management units must be submitted and approved by the MNRF to account for access roads and water crossings, harvesting plans, and renewal strategies, and ensure the ongoing maintenance of forests.

The harvesting intensity (i.e., how much timber to harvest) and the size and spatial patterns of harvest blocks have also changed with time, guided by periodic revisions to forest management policies (OMNRF 2017). Along with expansions in timber harvesting, the access road network has proliferated, especially with the mechanization of timber extraction and transport. Many primary, secondary, and tertiary harvest roads have been created, originating from public transport corridors, to facilitate expanding timber harvesting areas and most of the tertiary roads have historically not been decommissioned. Timber harvesting practices in the boreal forest of Ontario faced intense public scrutiny in the 1980s, resulting in almost a decade-long process of environmental assessment hearings (Euler and Epp 2000). Subsequent revisions to forest legislation, policies, and management guidelines have drastically changed the spatial and temporal patterns of timber harvesting in the boreal portion of the AOU, including the construction and decommissioning of harvest roads (e.g., OMNR 2012).

Extant information about forest landscape disturbances

Present knowledge about the wildfire and timber harvesting disturbances in a portion of boreal forest in Ontario’s AOU is derived primarily from data that is gathered, compiled, and reported by the two major land management agencies, the MNRF (OMNRF 2016) and Natural Resources Canada (NRCAN 2018). This information is supplemented by databases compiled periodically by ad hoc developers (such as forest companies, ENGOs, and individual researchers). Several major limitations are associated with the collective body of information thus assembled and reported. First, and most importantly, no generalized standards are in place for the assembly or portrayal of annual historical disturbance information, with respect to the data sources, spatial resolutions, projections, or definitions. In fact, almost all information sources do not specify methodological details or assumptions made during disturbance detection or the spatial portrayal of the resulting disturbance information. Second, many spatial and temporal gaps exist in the coverage of historical disturbances in the boreal portion of the AOU, resulting in unevenness in disturbance mapping through space and time that, when coupled with varied methods of production, lead to products that are difficult to use, compare, or analyze consistently. Third, the extant information is generally presented deterministically and without estimates of error or variability in their accuracy that would typically accompany geographical information in its metadata or as uncertainty layers. Collectively, the fragmented coverage of data, database inconsistencies, and the absence of methodological details make the reliability and defensibility of the present body of information weak, which has led to frequent disputes among land management agencies, the forest industry, and ENGOs.

Uses of forest landscape disturbance information

The demand for reliable and defensible information of Ontario's forest landscape disturbances is not only broad but is growing. First, the MNRF, as the custodian of Ontario's forest landscapes, has a legal requirement to report annually on the *state of the forest* (e.g., OMNRF 2016). Forest disturbances are a crucial element in these reports, dating back over a century (e.g., *Report of the Minister of Lands, Forests and Mines of the Province of Ontario*, 1919). These reports provide the Ontario public with an authoritative assessment of the forest resources and form the basis for narratives of the trends in forest disturbances. Furthermore, these reports are a source for scaling up forest disturbance information to broader levels, e.g., for reports at national levels, such as The State of Canada's Forests annual reports that are issued by Natural Resources Canada each year and used in global statements of forest assessments (FAO 2015). These national and global reports in turn have been a source for several recent studies and reports of national and global trends in boreal forest landscapes, for example, carbon dynamics and biodiversity patterns (Chaste et al. 2018). In addition, conservation organizations that focus on the boreal forest biome, at provincial, national, or global scales, periodically produce reports of the present state of the boreal forest and its future, as well as ecological consequences of disturbance, based on information about forest landscape disturbances (e.g., International Boreal Conservation Science Panel[4]).

Second, forest disturbance history is a vital source of information for strategic as well as operational land use planning. The demand for spatially and temporally accurate past wildfire and timber harvesting information for applied uses has grown recently, especially with the advent and popular use of spatially explicit simulation models and planning decision support systems (Perera and Cui 2010, Shinneman et al. 2012). Scenario simulations of future land management possibilities and optimizations for timber and other resource extractions, as well as for designing plans for conservation efforts, now require reliable spatial data with high fidelity on past changes to forest composition and structure. Accurate disturbance history also is an integral component in the forest inventories that constitute the foundation for all forest management planning activities (Twery 2004). The ability to have a confidence surface based on an ensemble of disturbance data sets, updated annually, will benefit the continuous inventory initiative going forward.

Third, and perhaps the most diverse and exacting requirement for forest landscape disturbance information comes from the research arena. Accurate and spatially and temporally explicit data on past wildfires and timber harvesting are a main input for many scientific studies, for current and emerging topics. The array of such topics from the disciplines of ecology (Grondin et al.

---

[4] www.borealscience.org/panel

2014), socio-economics (Brecka et al. 2018), and policy development (Bergeron et al. 2004) is vast, therefore here we list only few as broad examples that focus on determining changes in boreal forest landscape disturbance patterns, and their ecological and socio-economic consequences. These examples are shifts in disturbance regime spatial patterns, temporal trends, intensities, and scale; broad scale shifts in species composition, community assemblies, spatial patterns and temporal dynamics; cross-scale spatial and temporal cumulative effects on ecological processes such as abiotic flow, elemental flux, species migration, and gene flow; ecological system resilience and adaptive capacity at multiple scales (e.g., thresholds, recovery, heterogeneity, self-organization) in relation to future context changes and new stressors; changes in regulating, provisioning, and using socio-cultural ecosystem services; and effectiveness and effects of past and present land management policies. We remind the reader that these are only few broad scale topical groups, and they each hierarchically include numerous finer scale research questions.

Furthermore, many responsibilities (including reporting requirements and activities) of the MNRF use information about past wildfire and timber harvesting disturbances in Ontario's boreal portion of the AOU. Applications such as compiling annual state of the forest reporting, supporting the integrated monitoring framework, updating the forest resources inventory, cataloguing wildfire information, and examining the effectiveness of forest management guides will benefit from improvements made to available information.

We anticipate that many developer and user communities will be involved with assembling, analyzing, and applying this spatial database of historical wildfire and timber harvesting in the boreal portion of the AOU of Ontario. That unified vison is illustrated in Figure 3, focusing on the broadest of research and management questions: What are the ecological, socio-economic, and policy consequences of 50+ years of accumulating wildfire and timber harvesting disturbances in the boreal portion of the AOU of Ontario?

Goal and objectives

This database development initiative is predicated on the recognition of many weaknesses in the extant data of historical wildfire and timber harvesting information, and the broadening demand for improvements in reliability and defensibility of that information. This initiative was in the context of many direct applications in the MNRF that demand improved defensibility in information used, and emerging research needs to study patterns and consequences of accumulating forest disturbances in the boreal portion of the AOU that require reliable spatial and temporal information.

Our broad goal was to compile a spatiotemporally comprehensive data set that can be used to detect past disturbances and is explicit in assumptions, definitions, methods, biases, and errors. By detailing the methodological stages of derivation and assembly, we emphasized repeatability of all steps in the process. The output is updateable and includes a spatially explicit confidence surface, where a location's 3-year disturbance history is considered to avoid identifying disturbances in consecutive years at any location. This forms a better assessment of error propagation and alignment with extant data.

The primary objective of this report is to comprehensively describe the first phase of this project: detecting and mapping wildfires and timber harvesting in the boreal portion of the AOU of Ontario during the period 1972–2019. Specifically, we describe the definitions, assumptions, primary data sources, rationale for selecting the disturbance-detection methods, all steps in data processing and assembly, spatial and temporal resolutions, as well as sources of errors in compiled information and quantitative estimates of their magnitude. We also include a case study area where the resulting data layers are presented for illustrative purposes. Finally, we describe our intentions to expand the disturbance information temporally: backcasting before 1972 and adding future disturbances annually (to be done by the Forest Landscape Ecology Program at the Ontario Forest Research Institute).

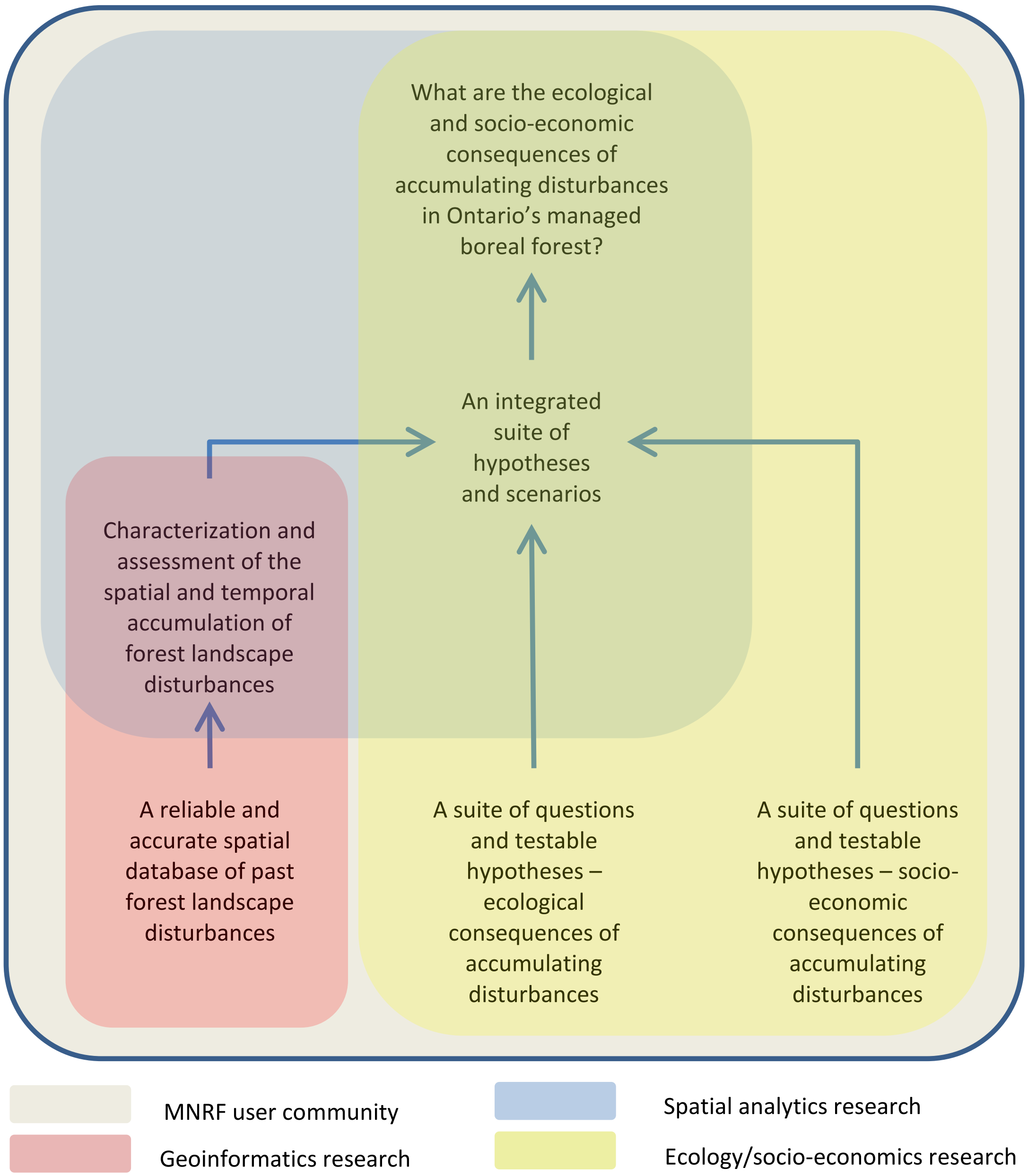


Figure 3. Broad user groups with example questions and hypotheses that each can investigate using disturbance mapping for the boreal forest portion of the AOU in Ontario.

We briefly describe our plans for detecting and mapping several broad scale anthropogenic disturbances that have accumulated during the same period: proliferation of access road networks and the expansion in infrastructural features. We include appendices with details about data (Table A1.1) and a list of acronyms used in this report (Table A1.2). For a list of Landsat scenes that were used for this study and those that were excluded during processing, please contact the authors.

Methods

General approach

We use methods to fulfill the premise of producing a complete boreal disturbance data set for the boreal forest portion of the AOU in Ontario for the years 1972 to present. We outline the modular workflow for producing the data, such that all decisions are clearly documented and, if necessary, could be adjusted to produce different versions of the database. This latter point is pivotal, as it provides a clear way to reproduce the disturbance mapping when new science, understanding, thresholds, or decision-making criteria become available; these adjustments can then be implemented consistently across the area of study.

The geographic extent and scope of this mapping and attribution exercise is limited to the boreal forest portion of the AOU in Ontario. We did not include the Far North or regions directly adjacent (e.g., Whitefeather management unit) to the northern boundary of the AOU since they are not currently subject to harvesting pressure (nor have they been) and we have not included the Great Lakes St. Lawrence southern portion of the AOU due to it not containing boreal forest and thus have not formed the basis for reporting or prior mapping efforts. These reasons make any ensemble confidence assessment difficult, inconsistent, and often impossible due to the lack of comparative data sets.

We split our methods into 3 distinct phases (Figure 4) that divide our timeline into discrete subsets. Each phase represents specific processing differences and requirements. Phase I (1972–1983) covers the period from the onset of Landsat Multispectral Scanner System (MSS) availability until a somewhat arbitrary date set by Hermosilla et al. (2016), which represents a point at which data was already being made available. Phase II is split into 2 parts: II-A (1984–2015) to identify the data made available by the mentioned study, and II-B (2016–present) to identify the most recent data. The final phase, Phase III, is beyond the scope of this report and represents future work of backcast modelling based on known disturbance recovery trajectories (Kennedy et al. 2010) to predict disturbances in the period before the availability of Landsat imagery (from about 1967–1972), relying on detected disturbances in 1972 and projecting into the past to identify disturbance dates.

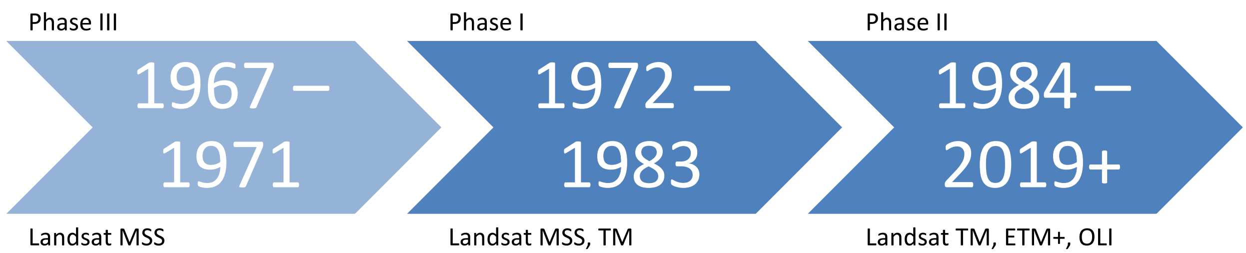


Figure 4. Timeline of data availability and temporal phases for processing data to develop a database of disturbance history for the Area of the Undertaking in Ontario, Canada. Phase III (1967–1971) is identified as future work and is beyond the scope of this report as it requires backcast modelling to map disturbances that occurred before the availability of commercial satellite imagery (e.g., Hermosilla et al. 2016). Phase II incorporates data produced by the Canadian Forest Service, a project led by Dr. Mike Wulder (e.g., White et al. 2017). Phase I is the preliminary focus of this report as it required the production of a new workflow and reporting process (Smyth 2020). MSS=Multispectral Scanner System, TM=Thematic Mapper, ETM+=Enhanced Thematic Mapper, OLI=Operational Land Imager.

We endeavoured to map fire and harvested areas among the many possible forest changes. The generic workflow is depicted in Figure 5, outlining the main processing steps along with the software and programming environments in which each was implemented. The layers produced with our methods rely primarily on spectral results and the ability to identify signatures associated with fire or harvesting disturbances relative to undisturbed locations. Since no boundary is delineated, vertex spacing consistency is not an issue and resulting mapped points each represents an area of 1.44 ha. This spatial resolution is the result of an aggregation of an appropriate number of pixel neighbours and is assigned an ensemble assessment of confidence.

The approach that we developed and implemented allows errors and uncertainty to be assessed at each mapped point. We strived to retain and attribute all instances where errors and uncertainty may arise. While we detail these throughout the methods section, we bring attention to locations where Landsat image scenes overlap and, because scenes are individually classified, the same locations may be classified more than once at these locations. Agreements and discrepancies at such locations are also tracked and reported. The benefit of the presented method is that attribution allows these various aspects to be unpacked and independently assessed as desired. Ultimately, users can elect to select, repackage, reassess, or use any combination of points to form clusters representing areas of disturbance, depending on their needs, biases, or to test and compare explicit assumptions.

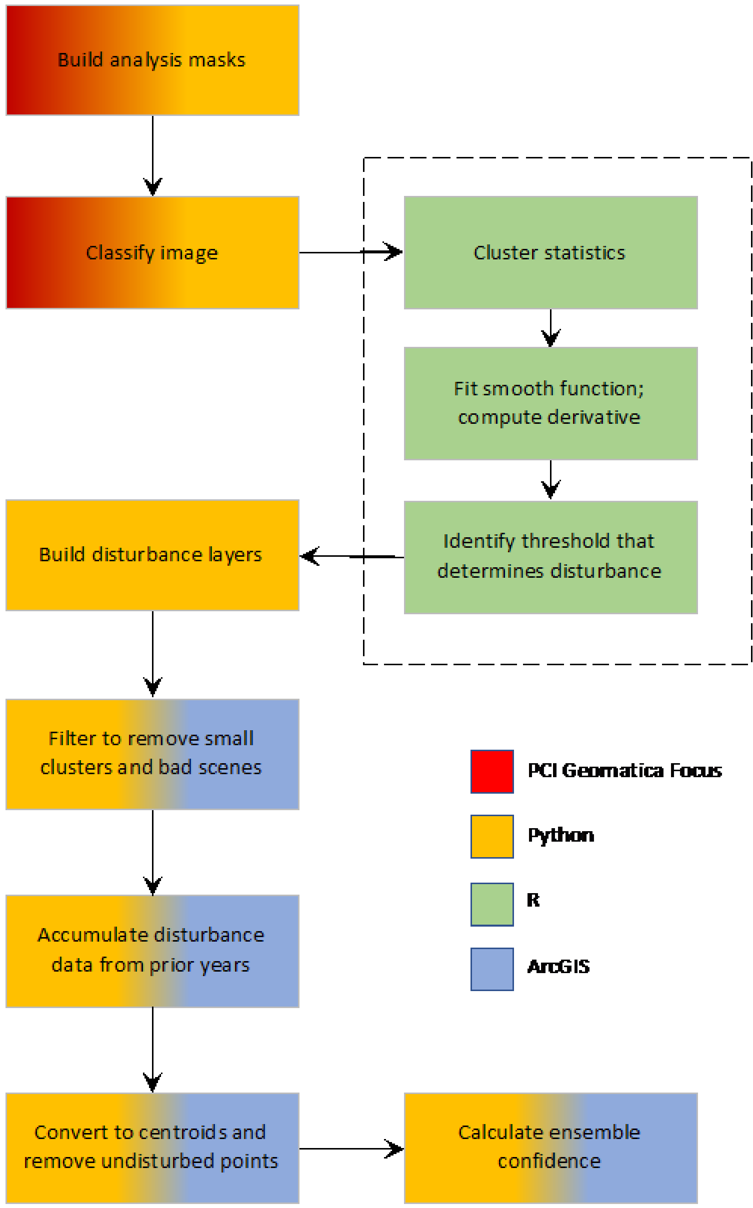


Figure 5. Generic workflow identifying the computing and programming environment used to implement disturbance database development.

Output disturbance layers are provided not as deterministic and discrete entities (i.e., polygons) but as points to which attributes relating to the disturbance type, timing, and confidence are attached. This approach achieves several interconnected desired characteristics: (1) manageable file sizes (2) an ability to focus on identifying disturbance locations rather than attempt to delineate complex boundaries, and, most importantly, (3) we can attribute each point with measures of confidence based on an ensemble assessment of agreement. Each location in a disturbance can have its own level of confidence and thereby produce opportunities for assessing spatially varying confidence. Locations exist in the data layers if they are identified as disturbed by our methods or if identified as a disturbance by any other known data source; the confidence attribute, however, will reflect the ensemble assessment across all available data sets.

Phase I: 1972–1983

Landsat Multispectral Scanner System data assembly and preparation

Select valid Landsat scenes

The initial step in this process was to identify and download relevant Landsat MSS scenes using the USGS Earth Explorer web portal (earthexplorer.usgs.gov). All images that met the minimum of L1T quality, had ≤30% cloud cover, and were acquired between August and October (as close to 31 October as possible) for time the interval identified were selected as candidate scenes for download (n > 3000, about 40-65 annually, depending on the year). The August to October period was selected since it is the closest time for which we can obtain snow-free imagery at the termination of the fire season. Given that the years observed straddled the change in how scenes are spatially referenced and identified, all scenes from 1972–1982 (Landsat 1-3) were identified using the Worldwide Reference System (WRS)1 path/row system and scenes from 1982–1983 (Landsat 4–5) were identified using the WRS2 path/row system[5] (Figure 6).

The scene selection process was partially manual, but downloading was automated using the Bulk Downloader tool. Some scenes were clearly miscoded in terms of cloud cover and easily deleted by manual decision making at this stage, particularly if most of a scene was covered with opaque clouds. Landsat Thematic Mapper (TM), Enhanced Thematic Mapper (ETM+), and Operational Land Imager (OLI) scene path-row identifiers are provided in Figure 7. A filtering of scenes to select the clearest candidates resulted in a final set of about 360 images for further pre-processing according to the workflow depicted in Figure 8 and Figure 9.[6]

[5] www.usgs.gov/land-resources/nli/landsat/landsat-shapefiles-and-kml-files

[6] Details about Multispectral Scanner System and Thematic Mapper scene availability based on search criteria (e.g., clouds, date) and scene from which classification training signatures were obtained are available from the authors.

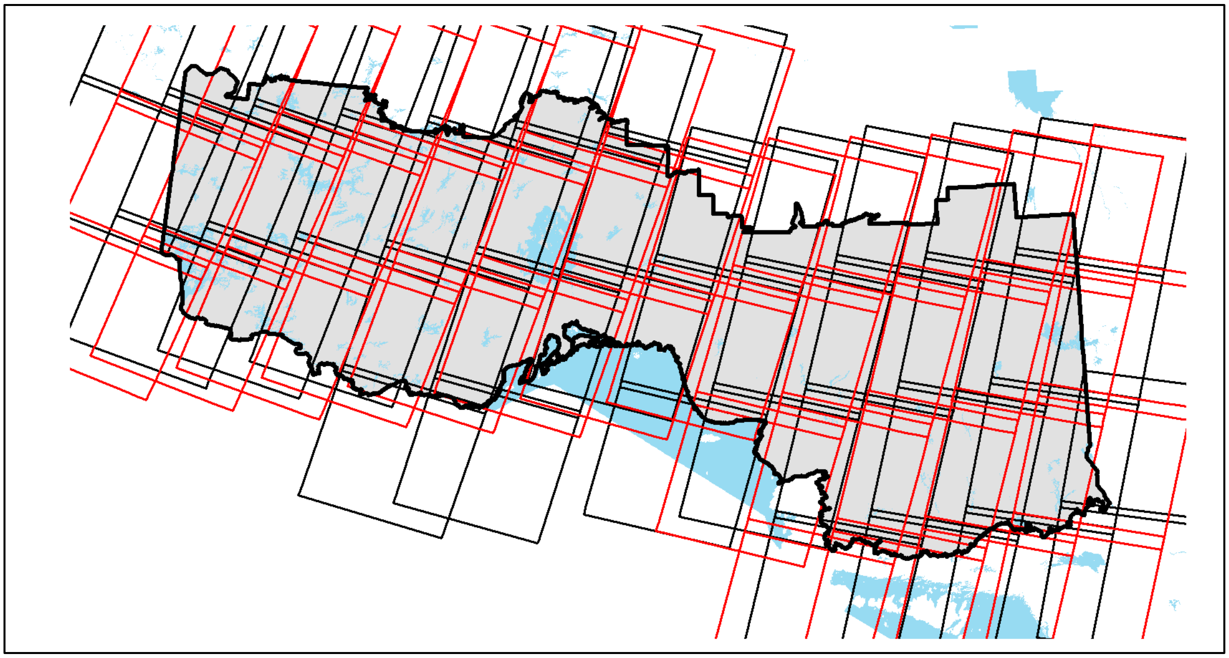


Figure 6. World Reference System (WRS)1 and WRS2 boundaries for identifying Landsat 1–5 scenes (red outlines represent Multispectral Scanner System scene boundaries and black outlines represent Thematic Mapper, Enhanced Thematic Mapper+, and Operational Land Imager) for the Area of the Undertaking in Ontario, Canada (as shown in Figure 1).

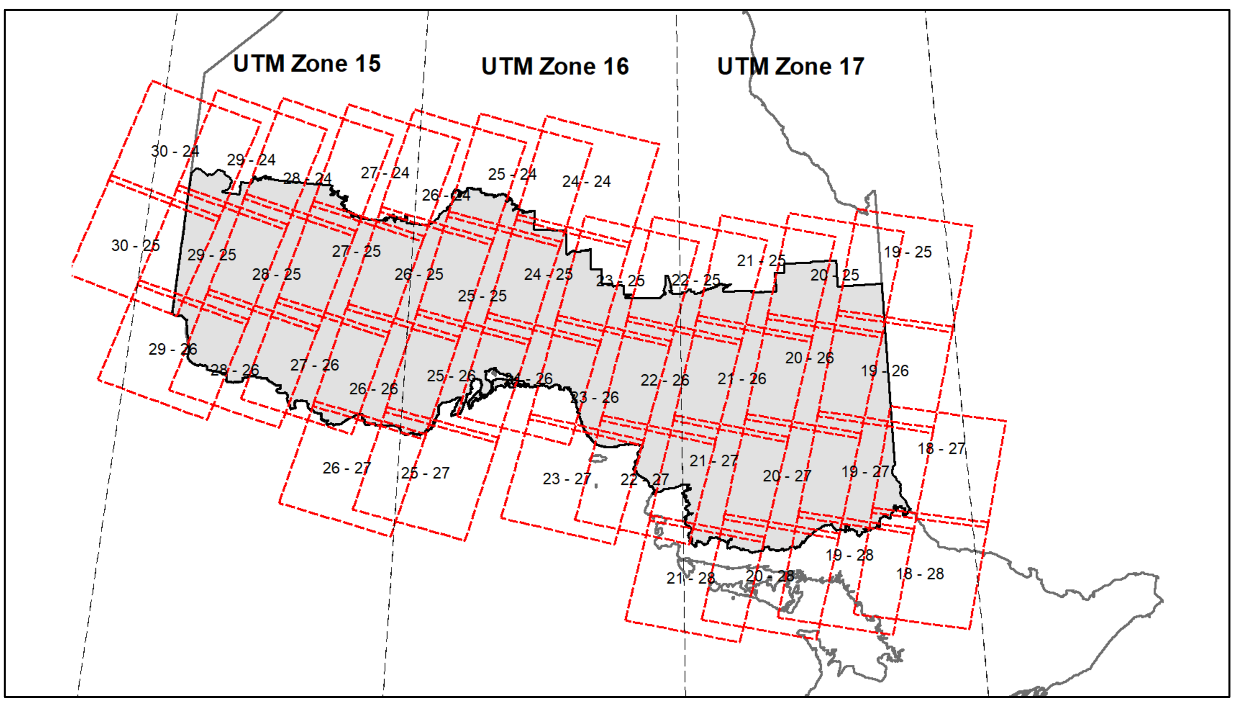


Figure 7. Positions of Landsat Thematic Mapper scenes (numbered boxes) in the World Reference System path-row construct overlain on Ontario's Area of the Undertaking with Universal Transverse Mercator (UTM) zone boundaries identified.

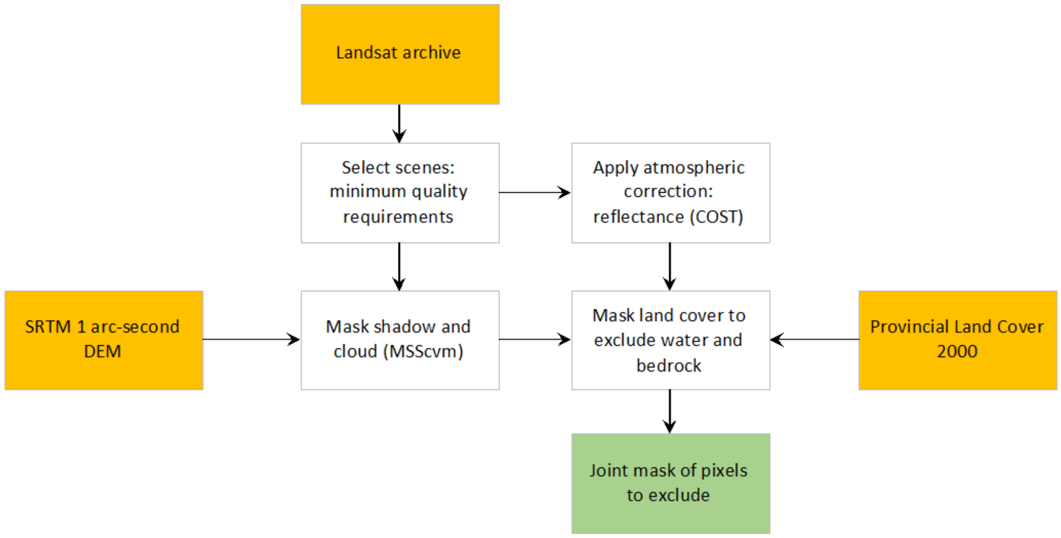


Figure 8. Pre-processing workflow (1972–1983) applied to Landsat Multispectral Scanner System (MSS) data. COST=Cosine of the Solar Zenith Angle, SRTM=Shuttle Radar Topography Mission, DEM=digital elevation model, MSScvm=Multispectral Scanner System clear-view-mask. Orange boxes are input data, white boxes are processes, and green box is the resulting data layer.

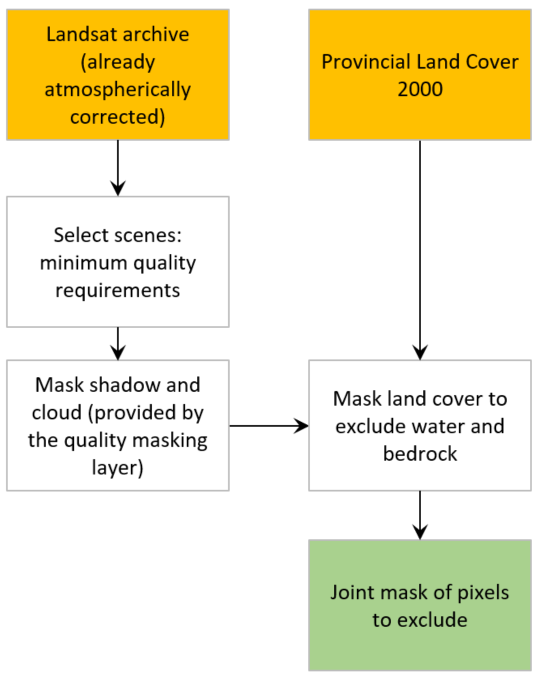


Figure 9. Pre-processing workflow (1984 onward) for Landsat Thematic Mapper, Enhanced Thematic Mapper, and Operational Land Imager data. For these data, the atmospheric correction has already been applied and the cloud/shadow masking is based on the supplied quality layer (Provincial Land Cover 2000[7]; Spectranalysis Inc. 2004). Orange boxes are input data, white boxes are processes, and green box is the output data layer.

[7] www.sse.gov.on.ca/sites/MNR-PublicDocs/EN/CMID/Provincial%20Landcover%202000%20Edition%20Report%20-%2027%20Classes.pdf

Mask clouds and shadows and apply atmospheric correction

Due to the large number of scenes that needed to be processed, and our desire to automate processing to ensure reproducibility and reduce human error, we landed on an automated solution for masking clouds and shadows implemented in *R* (R Core Team 2019). We used the Landsat MSS clear-view-mask (MSScvm) tool by Braaten et al. (2015), which automated the cloud and shadow identification and thereby blocked areas from further processing. The tool is freely available via GitHub[8] and relies on the NASA derived 1 arc second Shuttle Radar Topography Mission digital elevation model (SRTM DEM)[9] for assessing topographic conditions of the landscape. The masked area of these identified cloud and shadow areas was further expanded by 1 pixel in all directions to capture edge effects and to be more cautious in building land cover masks (i.e., more likely to include cells in the mask than exclude them). Added to this masked area, we included all areas identified by the (Ontario) Provincial Landcover 2000 (Spectranalysis Inc. 2004) product as water (class = 1), shallow water (class = 2), and exposed bedrock (class = 5). Given the difference in spatial resolution between our mapping and the Provincial Landcover 2000 data, a nearest neighbour (NN) approach was implemented to identify cells requiring masking. The NN algorithm, as implemented in ArcGIS, resorts to selecting the lower right (southeast) corner cell, in the event of a tie in distances to nearest cells. A final mask of pixels to exclude from analysis was produced by combining the water, cloud, and shadow masks.

To further prepare the acquired MSS scenes for use and to allow for standardized comparisons among them, we atmospherically corrected each scene. While many atmospheric correction modules exist, we produced ground reflectance values using the "cosine of the solar zenith angle" (COST) method (Chavez Jr. 1996, Abdolrassoul and Turner 2007) available within TerrSet[10] software, a product of Clark Labs that integrates the IDRISI GIS Analysis and Image Processing tools. TerrSet software was used for reasons of licensing and ease of access, though other options exist. The inclusion of the COST atmospheric correction tools at no additional cost, and with results comparable to modelled atmospheric corrections (Chavez Jr. 1996), meant that we did not need to purchase the additional license for ATCOR (Richter and Schläpfer 2012), but we did need to move images among software environments. Furthermore, the COST algorithm is specifically designed for MSS imagery (which aligns perfectly with the requirements of Phase I), has been tested in North America, and has been shown to behave exceptionally well (Chavez Jr. 1996). The IDRISI raster images in floating point format (with reflectance values between 0.0 and 1.0) were multiplied by 10,000 and truncated to convert floating point values to integer outputs saved in *.tif* image format (but original precision is readily obtained by dividing values again by 10,000); all masked areas were forced to have no values.

Since the data product being compiled spans 3 UTM projection zones (15, 16, and 17) and many Landsat images cross one of these boundaries, special handling was required when processing scenes at the interfaces. For example, adjacent zones were at times processed together, selecting one or the other zone to represent all scenes and thereby reduce projection-induced

---

[8] github.com/jdbcode/MSScvm

[9] lpdaac.usgs.gov/products/srtmgl1v003/

[10] clarklabs.org/terrset/

edge effects. By keeping scenes intact, we maintained scene integrity while facilitating the proper processing of full scenes by COST and the further identification of pixels to be masked.

Detect fire and harvesting disturbance by image classification

The full workflow pertaining to image classification to identify disturbances is depicted in Figure 10. All imagery was imported to PCI Geomatica Focus proprietary format (.*pix*), maintaining the Landsat scene name for continuity and backwards compatibility. Once imported, each scene and year combination were processed in sequence to identify training areas. The locations of these training areas were guided by the MNRF fire and harvesting databases but also visual assessment. The goal was not to specify explicit cells representing the presence or absence of disturbances, but to select general areas that cover both disturbed and undisturbed pixels. The specifics of training area selection were subject to the presence and size of disturbances on scenes. The availability of fire training was also much more abundant in the western half of the boreal area of the AOU, as fires are much more prevalent there than in the east. We also preferred to use training local to a scene, rather than something more distant in space or time as identified below by our priority sequencing. Training for areas subjected to fire and harvesting were completed independently due to their spectral differences but also their prevalence on the landscape, which varied how training sites were identified. For any scene, we aimed to identify 6 to 10 training areas for harvesting disturbances and ≤4 training areas for fire disturbances.

The approach for identifying local training data when disturbances exist on a scene works well; however, this was not always the case and sometimes we had to obtain training data from another scene based on a logical sequence of decreasing priorities. The goal was to provide priority to imagery from as close in time and geography as possible to ensure consistent vegetation state, atmospheric condition, and thus representativeness of the reflectance values observed. The list below provides our sequence of decreasing priority for selecting training areas:

1 Same path + same day =
Select a scene northward or southward
2 Same path + closest day =
Select the same scene or 1 northward or southward
3 Neighbouring path + same row + closest day =
Select scenes westward or eastward
4 Neighbouring paths + different row + other path + closest day =
Select scenes diagonally from the west, east, north, or south
5 Same path OR row + different year =
Select same scene from a later year
6 Same path + closest day + other year =
Select scenes north or south from a later year
7 Neighbouring paths + same row + closest day + different year =
Select scenes to west or east from a later year
8 Neighbouring path, other row, other year =
Select scenes diagonally from the west, east, north, or south

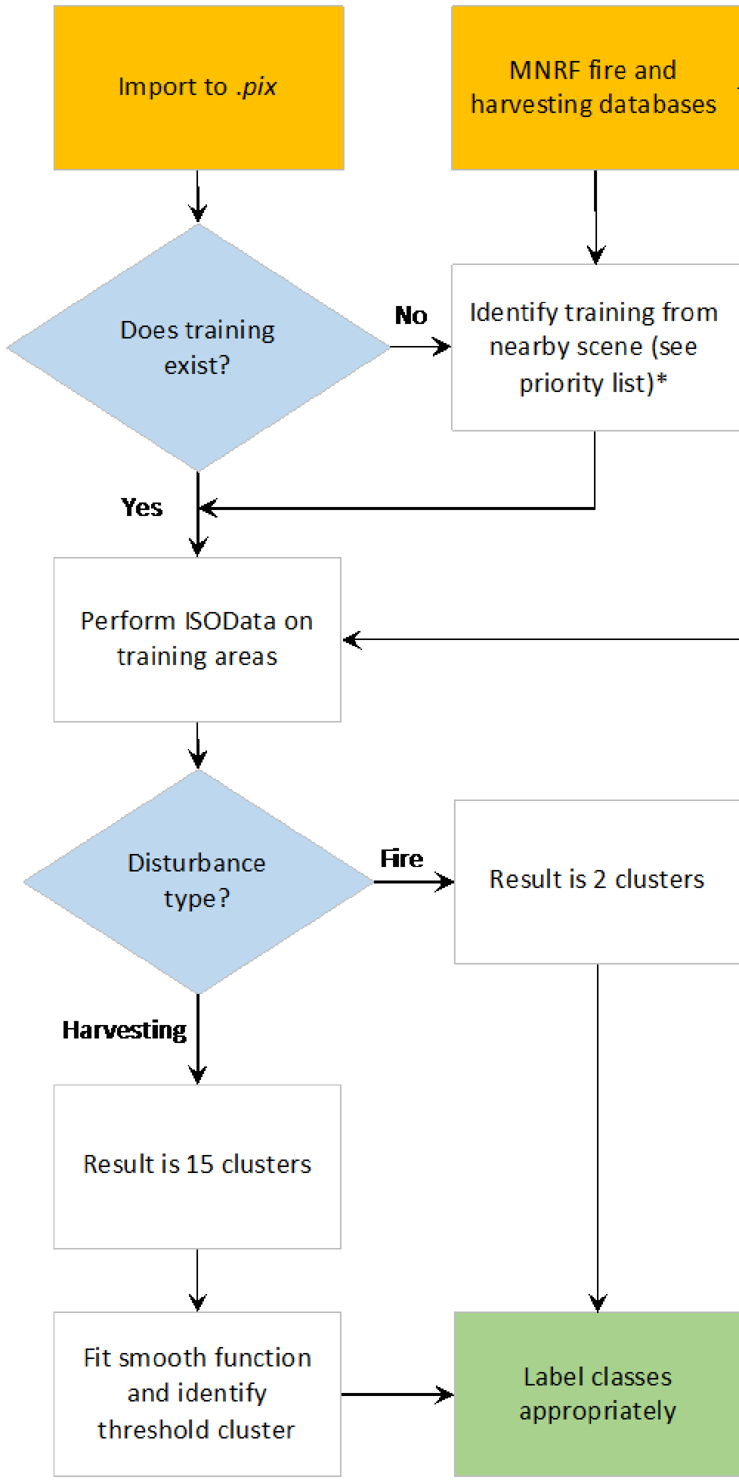


Figure 10. Workflow for image classification used in developing a disturbance database for the Area of the Undertaking in Ontario. Note (*): When no other harvest databases were available (i.e., 1972–1989, harvest training sites were selected manually by looking at patterns such as progressive cuts and verified by road accessibility and annual progression. This method was also applied in 2018–19 because the provincial harvest database had not yet been updated. In 2019, training for fire sites was based on the available ignition centroids as no other data was yet available. Orange shading indicates input data, white boxes are processes, blue diamonds are decisions, and the output data layer is shown in green.

Once training areas were identified, unsupervised ISOData (iterative self-organizing data analysis method) was run to produce clusters that would ultimately form scene-wide training signatures at the pixel level. For fire data, 2 clusters were produced; for harvesting data, 15 clusters were produced. The difference in the number of clusters stems from the quality of results obtained for fires but the spectrally more complex harvesting signatures. These targets remain flexible and analyses could be replicated with varying clusters to perform subsequent sensitivity analyses.

Harvesting clusters were then fit with a third-order polynomial smooth function in R to identify clusters that occur at points of flexure along this function (Figure 11). The identified points are used to split the classification training into harvested versus non-harvested training; where harvests were at the higher-valued clusters. Thresholds for splitting were selected by fitting a smooth polynomial function through the scatter of clusters versus cluster mean values and then computing the first derivative of this function. This information was then passed back to scripts (see Figure 5) that implemented the threshold selection by scanning for the minimum value, starting from the largest cluster identifier value. If the derivative had no minimum, then the average of all selected clusters was used instead. The cluster definitions and statistics were stored and used for subsequent maximum likelihood supervised classification that used the ISOData clusters as input training data. Classification output provided disturbance versus non-disturbance classes for all areas where pixels were not masked out due to shadow, cloud, water, or bedrock classes. The results are preliminary fire and harvesting disturbance maps.

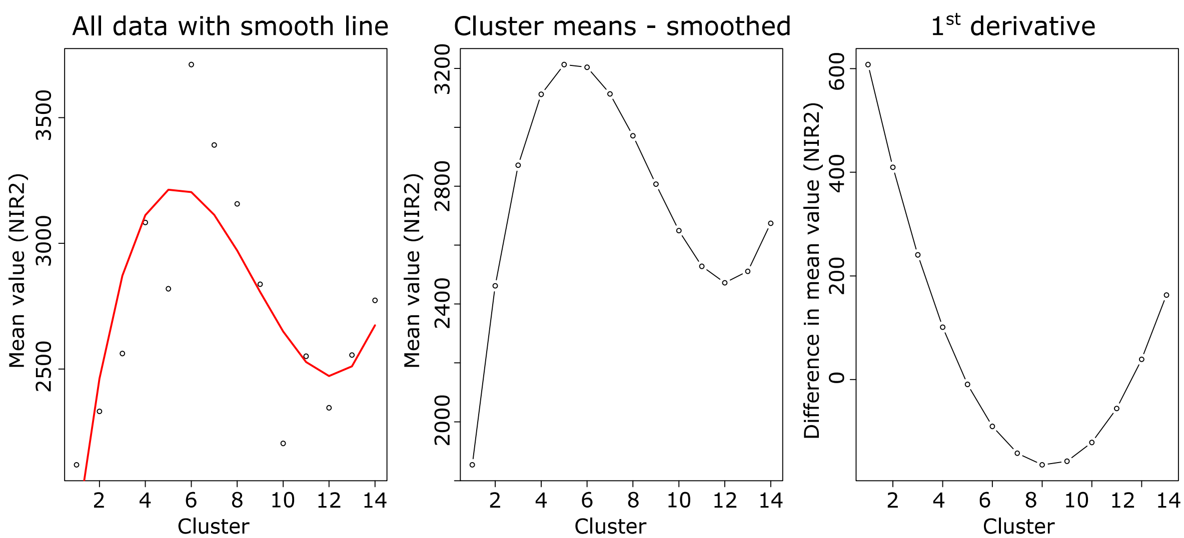


Figure 11. Example of ISOData (iterative self-organizing) cluster distributions for harvesting classification on a single image. The red line is a smooth function fit to the cluster means (on the NIR2 band).

Given landscape variability, signature confusion does lead to commission errors that typically result in small clusters of pixels identified as disturbances. We spatially filtered the results to eliminate all clusters (8-cell contiguity rule) that are <4 ha (i.e., 11.1 MSS or 44.4 TM pixels). We then reprojected from UTM to Lambert Conformal Conic projection while resampling the disturbance layer to 1.44 ha spatial resolution (4 MSS cells) to align with standard reporting scales and to further reduce some of the spurious pixels identified as disturbances but deemed to introduce noise into the database. A further threshold for scene quality was applied that rejects the use of images where >75,000 ha is identified as harvest or >150,000 ha is identified as fire; these are typically noisy images and exceed normal expectations for these types of disturbance. The threshold values were selected by examining the size class distributions for historic disturbances and identifying areas at which large jumps occur and do not exclude large disturbances.

Since image scenes overlap along their margins and given that we assessed disturbances independently on each scene, we may have 1 to 4 scenes with separate disturbance classifications for the overlapping areas. Though this introduces non-uniform complexity (i.e., this overlap is constrained to scene margins, not the central area), we used this overlap to record agreement between or among the multiple classifications. Using the coding presented in Table 1, we indicated whether a location was identified as fire disturbed, harvesting disturbed, both fire and harvesting disturbed (uncertainty), or not disturbed. We completed this identical assessment on each overlapping area and then added the values obtained for all overlapping layers. For example, 10 + 10 + 10 + 1 = 31 would indicate a location identified as a harvest disturbance on 3 of 4 overlapping scenes, with 1 scene identifying it as a fire disturbance. Obviously, 40 and 4 are the ideal extremes (4 overlapping images identifying harvest or fire, respectively), but all possibilities were coded in this manner and retained as attributes to facilitate later interpretation if desired by a user. If a location was identified as fire disturbed or harvesting disturbed in the previous 2 years, then the 100s or 1000s column of this maximally 4-digit code was respectively was coded with a “1” rather than a “0”.

Table 1. Coding used for unique occurrences of disturbance classifications in the current year, in a 3-year window (that includes the current year), and further at multiple overlapping scene margins. The value can range from 1–4 to indicate up to 4 overlapping classifications at scene corners. Thus, a 1041 is a harvest location independently identified 4 times at a scene corner and as fire once and also identified as a harvest but not fire in the previous 2 years with a value of 1 (otherwise 0).

Case	DisturbanceCodeLandsat
Fire (current year)	1
Harvest (current year)	10
Fire (previous 2 years)	100
Harvest (previous 2 years)	1000

Data are provided on an annual basis, thus all disturbances tagged to a specific year were collected into independent layers and saved as points. The points represent the centroids of the original cells and were retained only where disturbances were identified (either by this method or by any of the overlain data products). By dropping the point locations where no disturbances were identified, file sizes were greatly reduced. All mapping conforms to a standardized AOU grid, such that points among years and UTM zones will always align properly.

Determine classification confidence through an ensemble approach

The assessment of accuracy in remote sensing classification is the best way to quantify the confidence in an interpreted product (Congalton 1991). However, for studies such as this one where classification was conducted for historical imagery, it is not possible to conduct field-based validation due to the time lag between dates represented by images and the present: since forest landscapes are dynamic the lag introduces uncertainty. In situations where a single classification, vote, or decision may carry uncertainty, it is possible to rely on ensemble voting to improve the accuracy of a decision (Bigdeli and Pahlavani 2016, Ko et al. 2016), where a greater number of votes for a specific decision indicates agreement among multiple sources.

We implemented an ensemble assessment of confidence in our disturbance labelling by overlaying other existing data products to identify agreements and disagreements. We recorded which products agree/do not in the attribute table, thus permitting each data product's correspondence to be assessed or weighted in whatever way a user wishes. The default computation of our reported confidence was to assume equal contribution to the weighting among all products available for a specific location. Our ensemble approach currently includes fire and harvesting data from several sources along with additional contextual layers (Table 2).

Given the attribution of ensemble confidence, a value that can change as additional data products are included either globally or locally in the confidence calculation, it is possible to map confidence surfaces for the boreal forest portion of the AOU in any given year. This approach provides a way to assess temporal consistency in the data, the thematic alignment with other products, and the spatial co-existence of disturbances mapped at any location. It will be possible to assess how close an identified disturbance is to the nearest identified location on other data products to assess spatial coincidence.

Table 2. A listing of databases used in the disturbance mapping, identifying their sources.

Database	Source
AOU	https://Derived product from ecoregions, parks, and forest management units
Ecoregions and ecodistricts	geohub.lio.gov.on.ca/datasets/ecodistrict
Forest management units	https://geohub.lio.gov.on.ca/datasets/forest-management-unit
Grid for centroids	Produced in-house for this project
Historical fire management zones	geohub.lio.gov.on.ca/datasets/historical-fire-management-zones
Natural Resources Canada fire and harvest database	https://opendata.nfis.org/mapserver/nfis-change_eng.html
Ontario Hydro Network (OHN) Waterbody	geohub.lio.gov.on.ca/datasets/mnrf::ontario-hydro-network-ohn-waterbody
Ontario Ministry of Natural Resources and Forestry fire database	https://geohub.lio.gov.on.ca/datasets/fire-disturbance-area
Ontario Ministry of Natural Resources and Forestry harvest database	Accessible via MNRF, forest analyst at the Policy Division
Province outline (modified to follow Great Lakes shorelines)	geohub.lio.gov.on.ca/datasets/province
Provincial satellite-derived disturbance mapping – Landsat time series: Satellite-based change detection for provincial monitoring	Accessible via MNRF (Smyth 2020)
SRTM 1 arc-second DEM	www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1-arc?qt-science_center_objects=0#qt-science_center_objects
UTM zone boundaries	hub.arcgis.com/datasets/esri::world-utm-grid
WRS1 and WRS2	www.usgs.gov/land-resources/nli/landsat/landsat-shapefiles-and-kml-files

Phase II: 1984–2019+

The period from 1984 onward has received considerable attention, primarily due to the availability of higher quality imagery (e.g., Landsat ETM+, OLI) and the increased number of available disturbance records, with its direct effect on current forest management practices. Given the existence of, and effort to produce data for this period, we simply incorporated those products rather than duplicate effort. While national products of land cover and disturbance are not new (Wulder et al. 2001, 2003), they generally rely on the interpretation of aerial photographs (Gillis 2001) or coarse spatial resolution satellite imagery (Wulder et al. 2010, Guindon et al. 2014), or the integration of remote sensing and photo interpreted products (Remmel et al. 2005). Recent research has relied on finer spatial resolution Landsat time-series analyses but tends to be more local (Schroeder et al. 2011, Ahmed et al. 2017).

This phase includes data from 2 sources: (1) a project lead by the Canadian Forest Service (CFS) in which the BAP (best available pixel) method for selecting values from within time series (Hermosilla et al. 2016) was used to produce national disturbance maps[11] on an annual time step (White et al. 2017), and (2) an internal MNRF effort to reproduce and automate the same methods (Smyth 2020). We do not implement these methods but rather include their supplied data and tie it to our location centroids. The impetus was to have continuity in the event that CFS stops producing annual updates. In this context, the addition of subsequent data years would be completed using this approach and compared with the methods presented in Phase I.

Phase III: 1967–1971 (future extension)

In July 1972, the ERTS-1 (Earth Resources Technology Satellite) began acquiring multispectral images for Earth observation purposes. This satellite was renamed Landsat-1 in 1975 and acquired over 300,000 images before it was decommissioned on January 6, 1978. This was the first satellite launched with the goal of studying and monitoring our planet and the first that provided such data commercially for research purposes. While Landsat-1 carried 2 types of sensors: (1) the return beam vidicon (RBV) and (2) the multispectral scanner system (MSS), the former failed in 1974 and thus the latter provided the greatest data resource and benefit. The MSS acquired imagery on 4 spectral channels (**Table 3**) that were replicated on later Landsat satellites for continuity.

[11] opendata.nfis.org/mapserver/nfis-change_eng.html

Table 3. Landsat-1 Multispectral Scanner System spectral channel descriptions (reprocessed Level-1 data product[12]) used in developing the disturbance database for Ontario's Area of the Undertaking.

Landsat-1 band*	Spectral range (μm)	Spatial resolution (m)
4	0.5 – 0.6	60
5	0.6 – 0.7	60
6	0.7 – 0.8	60
7	0.8 – 1.1	60

* Bands 1–3 were reserved for RBV: (1: blue-green; 2: yellow-red; 3: near IR).

Although satellite imagery does not exist for the period before July 1972, the legacy of prior disturbances may well be recorded on the available 1972 (and more recent) images. The legacies of such prior disturbances would appear as fire scars and stands at varying degrees of recovery and regrowth. If pixels in available imagery could be identified as belonging to a recovery state, and if the rate of recovery could be learned by examining the archive of available data, it may be possible to backcast predict the time of disturbance onset. This approach requires the use of Landsat time-series stacks to identify disturbance recoveries (Kennedy et al. 2010; Pflugmacher et al. 2012, 2014). As per Kennedy et al. (2012), the identification of disturbance and regrowth attributes would allow the identification of key points in the forest-disturbance-recovery trajectory. The extension to backcasting (historic projection) when disturbances occurred would rely on the identification of a point on the recovery limb and then, based on the recovery rate, use a regression relationship to identify the year of disturbance (Figure 12).

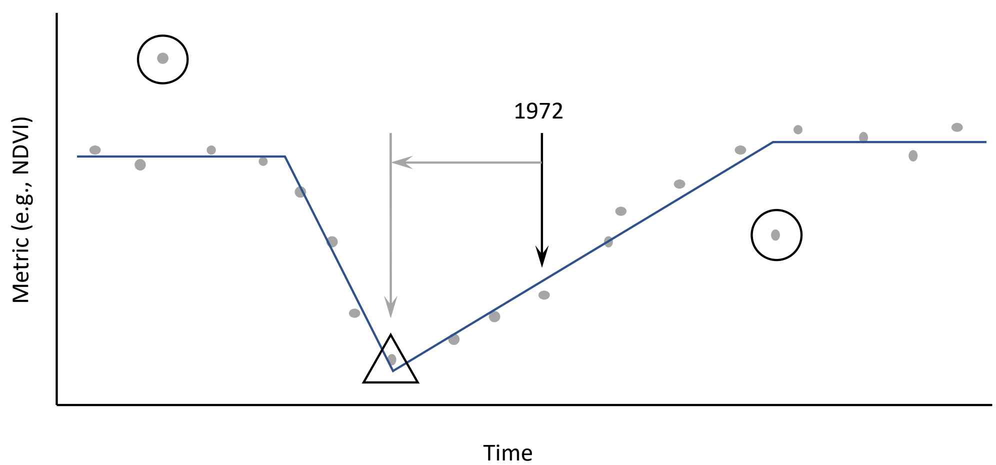


Figure 12. Theoretical disturbance recovery model for a single pixel's historic prediction of disturbance timing. Grey dots are Landsat measurements, circled dots are identified outliers, blue lines are regression relationships, and the triangle identifies the predicted disturbance timing given the 1972 (and more recent) data points. NDVI=normalized difference vegetation index

[12] www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-archives-landsat-1-5-multispectral-scanner-mss-level?qt-science_center_objects=0#qt-science_center_objects

While this backcasting approach is theoretically sound, initial exploration has shown that implementation is wrought with data quality complications. We propose that this extension be explored after the primary data set is complete and available, providing an opportunity to potentially extend the period for which disturbance mapping is available, including the period before satellite image availability.

Data product

The final data product is an ever-evolving dynamic archive of boreal forest disturbances in Ontario. This product is provided under the Creative Commons license CC BY-NC-SA[13], allowing anyone to remix, adapt, and build on this work non-commercially, as long as they credit this work and license their new creations under identical terms. At the conclusion of each calendar year, the archive will be updated to include that year's disturbances and to update attribution based on a moving 3-year window of identified disturbances for each location. The archive is provided as an ArcGIS GeoDatabase called *BorealDisturbance* with one feature class (Feature Type: Simple, Geometry Type: Point, Coordinates have Z values: No, Coordinates have measures: No) for each calendar year of disturbances. Disturbance feature class layers are named, starting with *BD* (for boreal disturbance) followed by the year (e.g., *BD1972*). This initial release has 48 annual layers (1972–2019). Additionally, for context we provide the political boundary for the Province of Ontario (*Province*), the boundary of the Area of the Undertaking (*AOU*) portion of the boreal forest, and the reference grid from which mapping points were derived (*BaseRaster*). Common projection information is provided in Table 4. We are aware that access to this data in different formats is desired and are working towards providing options in a future release.

Table 4. Projection definition implemented for geographic data layers used in the development of the disturbance database for Ontario's Area of the Undertaking.

Parameter	Value
Projected coordinate system	NAD_1983_Lambert_Conformal_Conic
Projection	Lambert_Conformal_Conic
False_easting	0
False_northing	0
Central_meridian	-95
Standard_parallel_1	49
Standard_parallel_2	77
Latitude_of_origin	49
Linear unit	Metre
Geographic coordinate system	GCS_North_American_1983
Datum	D_North_American_1983
Prime meridian	Greenwich
Angular unit	Degree

[13] https://creativecommons.org/licenses/

The final attribute tables for feature classes include 13 distinct attributes as summarized in Table 5. The *Disturbed* attribute is a Boolean indicator of whether some form of disturbance is recorded at a point (1) or not (0). The *DisturbanceYear* attribute further identifies the earliest assessment of the year in which that disturbance occurred; integer values indicate the year and are constrained to the period of study (currently 1972–2019). These values constrain the disturbance point to a specific feature class layer but, when combined with data from other years, allow tracking through time. The *FireLandsat* and *HarvestLandsat* attributes encode whether our methods identify a location as being disturbed by fire or harvesting (1) or not (0), respectively.

The six attributes (*OtherFireAFFES*, *OtherHarvestMNRF*, OtherFireMNRFSRB, OtherHarvestMNRFSRB, *OtherFireNRCAN*, *OtherHarvestNRCAN*) — shown in grey in Table 5 — are used to track whether other disturbance products also identify fire or harvesting disturbances at our standardized mapping point locations. The products identified here are the Natural Resources Canada mapping of fire and harvesting and the fire and harvesting mapping as supplied by the MNRF. The MNRF harvesting data is from compiled from annual reporting data provided by individual forest management units; the data was obtained from a forest analyst at the Policy Division of the MNRF Crown Forests and Lands Policy Branch.

Table 5. Final attribute definitions used in the disturbance database for Ontario's Area of the Undertaking. AFFES = Aviation, Forest Fire and Emergency Services, SRB = Science and Research Branch, NRCan = Natural Resources Canada.

Attribute	Type	Range	Description
Disturbed	Boolean	0|1	Did a disturbance happen here?
DisturbanceYear	Integer	1972–2019	Year of disturbance, whatever type
FireLandsat	Boolean	0|1	Was it identified by us?
HarvestLandsat	Boolean	0|1	Was it identified by us?
OtherFireAFFES	Boolean	0|1	Was it in the AFFES database?
OtherHarvestMNRF	Boolean	0|1	Was it in the MNRF database?
OtherFireMNRFSRB	Boolean	0|1	Was it in the MNRF/SRB database?
OtherHarvestMNRFSRB	Boolean	0|1	Was it in the MNRF/SRB database?
OtherFireNRCAN	Boolean	0|1	Was it in the NRCAN database?
OtherHarvestNRCAN	Boolean	0|1	Was it in the NRCAN database?
HarvestPrevious	Boolean	0|1	Identified in previous 2-year window?
FirePrevious	Boolean	0|1	Identified in previous 2-year window?
DisturbanceCodeLandsat	Integer	0–4444	Landsat code for overlap areas
FireConfidence	Integer	0–100	Confidence (%)
HarvestConfidence	Integer	0–100	Confidence (%)

HarvestPrevious and *FirePrevious* identify whether a location was previously identified as subjected to either harvesting or fire in a 3-year window. This allows the truest year of disturbance to be identified, while providing confidence in having identified a disturbance at a specific location. The *DisturbanceCodeLandsat* synthesizes the combination of information pertaining to identified disturbances (harvests and fires) in a year, but also in the 3-year window preceding a specific year. Further, this code identifies the state of overlapping classifications along scene margins and is helpful for devising checks for classification consistency using ensemble approaches. The final 2 attributes, *FireConfidence* and *HarvestConfidence*, convey measures of classification confidence by assessing the ensemble decisions made about classification. Currently, the calculation relies on an equal weighting among data products and computes a simple weighted percentage of agreement (C) among the binary indicators (D_i) of disturbance by n data products:

$$C = 100 \sum_{i=1}^{n} \frac{D_i}{n}$$

The confidence attribute can logically be used to create customized output maps, such as the example provided in Figure 13. In this map, the disturbance point mapping is overlain on the original false-colour infrared representation of the 1989 Landsat TM image for context. On the image, the fire disturbance is clearly seen as the green tones. The red and tan points represent this feature quite well but the numerous light green points that represent low confidence extend well into undisturbed territory and result from highly simplistic fire mapping used to create the Aviation, Forest Fire and Emergency Services (AFFES) database. The high confidence points avoid false positives and avoid mapping residual vegetation that was unaffected within and adjacent to the fire footprint. The distinction between fire events, footprints, and residual vegetation are clearly articulated by Remmel and Perera (2009) and establish the basis for mapping both the internal and boundary complexity of boreal fires.

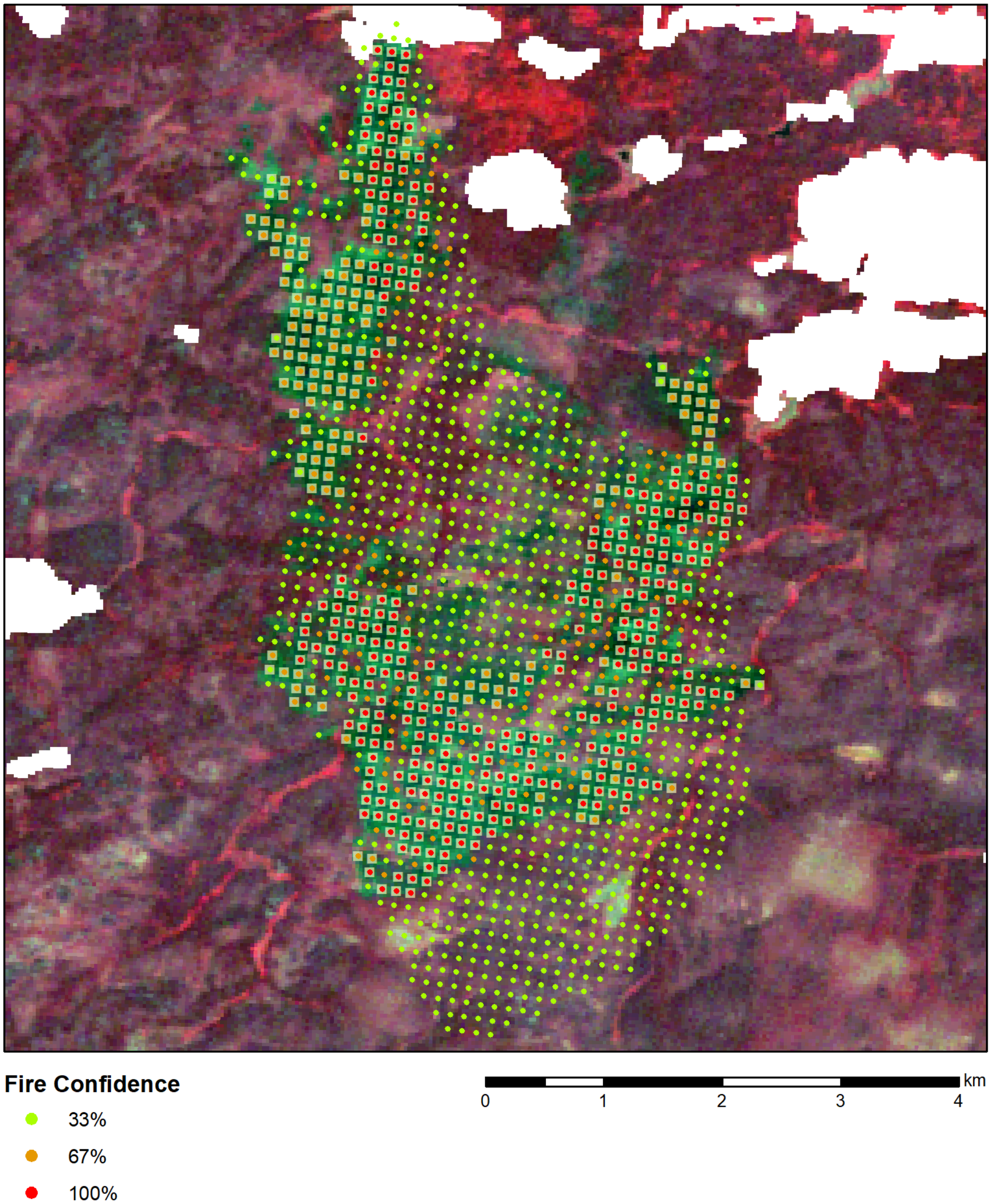


Figure 13. A subset of the data set (points) depicting a mapped fire disturbance with the original false-colour infrared Landsat TM scene rendered in the background for context. The confidence of each point is used to control its colour. The white areas are those that were masked out due to cloud, shadow, or waterbodies. Points printed on top of squares are locations mapped by the method described in this report to show that the intermediate confidence is due to our mapping in conjunction with Natural Resources Canada, rather than our omission and identification by other products.

Comments and future extensions

This data product is intended as a rigorous and comprehensive mapping of boreal forest fire and harvesting disturbances in Ontario's boreal forest portion of the AOU from 1972–2019. The availability of each new year's remote sensing imagery and our documented workflow (and collection of programming scripts) make it possible to continue adding annual data to this product. Hence, this is a dynamically growing archive that embodies both current and historical interest towards understanding and interpreting disturbance pressure in this vast area of natural resources.

Given the extent, remoteness, and historical aspect to this data, rigorous field validation is not possible, and thus we rely on computing the labelling confidence by an ensemble method that relies on additional disturbance mapping products. We provide all of this information as attributes tagged to the data. We also avoid the subtlety of mapping precise boundary locations (i.e., polygons) and rather focus on mapping regularly spaced points. This approach allows internal residual patches of undisturbed forest to be excluded from mapping, along with water, bedrock, and other non-burnable land cover types. Because of this, the computation of disturbance area and boundary perimeter lengths are not immediately possible. While it is possible to estimate area by assuming that each point represents an area of 1.44 ha, the additional consideration of confidence needs to be factored into any assessment (e.g., which points to include in any computation?).

While considerable consultation, thought, rigour, and science were applied in the production of this data product, with the goal of reproducibility, we wish to implement extensions, adjustments, and testing in future updates and releases. With the pending launch of Landsat 9 in November 2021, the availability of new image products (e.g., Sentinel, SPOT, lidar) will be inevitable.

The desire is to further improve the reporting of data quality and diagnostics. For example, for instances where disturbed locations are not mapped (e.g., presence of cloud, excluded scene), rather than omit those points, we would prefer to include a *NoData* point with a descriptor as to why mapping was excluded.

In the continued spirit of providing quality metrics for these data, overlays with the provincial Forest Resources Inventory (FRI) are pending and we are actively seeking other products with which to compare these results. Similarly, we are considering reprocessing the MSS data using the best available pixel (BAP) approach (Hermosilla et al. 2016) to assess potential improvement to the detection of disturbances in otherwise relatively noisy data.

References

Ahmed, O.S., M.A. Wulder, J.C. White, T. Hermosilla, N.C. Coops and S. E. Franklin. 2017. Classification of annual non-stand replacing boreal forest change in Canada using Landsat time series: A case study in northern Ontario. Remote Sensing Letters 8(1): 29–37.

Abdolrassoul, S.M. and B.J. Turner. 2007. A comparison of four common atmospheric correction methods. Photogrammetric Engineering and Remote Sensing 73(4): 361–368.

Bergeron, Y., M. Flannigan, S. Gauthier, A. Leduc and P. Lefort. 2004. Past, current and future fire frequency in the Canadian boreal forest: Implications for sustainable forest management. Ambio 33(6): 356–360.

Bigdeli, B. and P. Pahlavani. 2016. High resolution multisensor fusion of SAR, optical and LiDAR data based on crisp vs. fuzzy and feature vs. decision ensemble systems. International Journal of Applied Earth Observation and Geoinformation 52: 126–136.

Braaten, J.D., W.B. Cohen and Z. Yang. 2015. Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems. Remote Sensing of Environment 169: 128–138.

Brecka, A.F.J., C. Shahi and H.Y.H. Chen. 2018. Climate change impacts on boreal forest timber supply. Forest Policy and Economics 92: 11–21.

Campos-Ruiz, R., M.-A. Parisien and M.D. Flannigan. 2018. Temporal patterns of wildfire activity in areas of contrasting human influence in the Canadian boreal forest. Forests 9(4): 159.

Chaste, E., M.P. Girardin, J.O. Kaplan, J. Portier, Y. Bergeron and C. Hély. 2018. The pyrogeography of eastern boreal Canada from 1901 to 2012 simulated with the LPJ-LMfire model. Biogeosciences 15(5): 1273–1292.

Chavez Jr., P.S. 1996. Image-based atmospheric corrections — revisited and improved. Photogrammetric Engineering and Remote Sensing 62(9): 1025–1036.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37(1): 35–46.

Cumming, S.G. 2005. Effective fire suppression in boreal forests. Canadian Journal of Forest Research 35(4): 772–786.

de Groot, W.J., A.S. Cantin, M.D. Flannigan, A.J. Soja, L.M. Gowman and A. Newbery. 2013. A comparison of Canadian and Russian boreal forest fire regimes. Forest Ecology and Management 294: 23–34.

Epp, E.A. 2000. Ontario forests and forest policy before the era of sustainable forestry. Pp. 237–275 *in* A.H. Perera, D. Euler, I. D. Thompson (eds.). Ecology of a Managed Terrestrial Landscape: Patterns and Processes of Forest Landscapes in Ontario. UBC Press, Vancouver, BC, in cooperation with the Ontario Ministry of Natural Resources.

Euler, D.L. and E.A. Epp. 2000. A new foundation for Ontario forest policy for the 21st century. Pp. 276–294 *in* A.H. Perera, D. Euler, I. D. Thompson (eds.). Ecology of a Managed Terrestrial Landscape: Patterns and Processes of Forest Landscapes in Ontario. UBC Press, Vancouver, BC, in cooperation with the Ontario Ministry of Natural Resources.

FAO. 2015. Global forest resources assessment 2015: How are the world's forest changing? Food and Agriculture Organization of the United Nations, Rome, IT.

Gillis, M.D. 2001. Canada's national forest inventory (responding to current information needs). Environmental Monitoring and Assessment 67(1/2): 121–129.

Goldammer, J.G. and V.V. Furyaev. 1996. Fire in Ecosystems of Boreal Eurasia. Kluwer Academic Publishers, Dordrecht, NL.

Grondin, P., S. Gauthier, D. Borcard, Y. Bergeron and J. Noël. 2014. A new approach to ecological land classification for the Canadian boreal forest that integrates disturbances. Landscape Ecology 29(1): 1–16.

Guindon, L., P.Y. Bernier, A. Beaudoin, D. Pouliot, P. Villemaire, R.J. Hall, R. Latifovic and R. St.-Amant. 2014. Annual mapping of large forest disturbances across Canada's forests using 250 m MODIS imagery from 2000 to 2011. Canadian Journal of Forest Research 44(12): 1545–1554.

Hermosilla, T., M.A. Wulder, J.C. White, N.C. Coops, G.W. Hobart and L.B. Campbell. 2016. Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring. International Journal of Digital Earth 9(11): 1035–1054.

Kennedy, R.E., Z. Yang, and W.B. Cohen. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. Remote Sensing of Environment 114(12): 2897–2910.

Kennedy, R.E., Z. Yang, W.B. Cohen, E. Pfaff, J. Braaten and P. Nelson. 2012. Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. Remote Sensing of Environment 122: 117–133.

Ko, C., G. Sohn, T. Remmel and J. Miller. 2016. Maximizing the diversity of ensemble random forests for tree genera classification using high density lidar data. Remote Sensing 8(8): 646.

Lambert, R.S. and P. Pross. 1967. Renewing Nature's Wealth: A Centennial History of the Public Management of Lands, Forests and Wildlife in Ontario, 1763–1967. Department of Lands and Forests, Toronto, ON.

Nappi, A. and P. Drapeau. 2009. Reproductive success of the black-backed woodpecker (*Picoides arcticus*) in burned boreal forests: Are burns source habitats? Biological Conservation 142(7): 1381–1391.

[NRCAN] Natural Resources Canada. 2018. The State of Canada's Forests. Annual Report 2018. Ottawa, ON. 84 p.

[OMNR] Ontario Ministry of Natural Resources. 2012. Annual report on forest management 2010/11. Ontario Ministry of Natural Resources, Toronto, ON. (p. 94).

[OMNRF] Ontario Ministry of Natural Resources and Forestry. 2017. Forest management planning manual. Ontario Ministry of Natural Resources and Forestry, Toronto, ON. (p. 462).

[OMNRF] Ontario Ministry of Natural Resources and Forestry. 2016. State of Ontario's natural resources – forests 2016. Ontario Ministry of Natural Resources and Forestry, Sault Ste. Marie, ON. (p. 41).

Perera, A.H. and W.B. Cui. 2010. Emulating natural disturbances as a forest management goal: Lessons from fire regime simulations. Forest Ecology and Management 259(7): 1328–1337.

Pflugmacher, D., W.B. Cohen and R.E. Kennedy. 2012. Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. Remote Sensing of Environment 122: 146–165.

Pflugmacher, D., W.B. Cohen, R.E. Kennedy and Z. Yang. 2014. Using Landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics. Remote Sensing of Environment 151: 124–137.

R Core Team. 2019. R: A language and environment for statistical computing (Version 3.6.1). R Foundation for Statistical Computing, Vienna, AT.

Remmel, T.K., F. Csillag, S. Mitchell and M.A. Wulder. 2005. Integration of forest inventory and satellite imagery: A Canadian status assessment and research issues. Forest Ecology and Management 207(3): 405–428.

Remmel, T.K. and A.H. Perera. 2009. Mapping natural phenomena: Boreal forest fires with non-discrete boundaries. Cartographica 44(4): 274–288.

Richter, R. and D. Schläpfer. 2012. Atmospheric/topographic correction for satellite imagery (ATCOR-2/3 user guide). German Aerospace Center, Wessling, DE. DLR report DLR-IB 565-02111. (p. 216).

Rijal, B., F. Raulier, D.L. Martell and S. Gauthier. 2018. The economic impact of fire management on timber production in the boreal forest region of Quebec, Canada. International Journal of Wildland Fire 27(12): 831–844.

Schroeder, T.A., M.A. Wulder, S.P. Healey and G.G. Moisen. 2011. Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. Remote Sensing of Environment 115(6): 1421–1433.

Shinneman, D.J., B.J. Palik and M.W. Cornett. 2012. Can landscape-level ecological restoration influence fire risk? A spatially explicit assessment of a northern temperate-southern boreal forest landscape. Forest Ecology and Management 274: 126–135.

Smyth, I. 2020 (in prep). *Provincial satellite derived disturbance mapping - LANDSAT time series*. Ontario Ministry of Natural Resources and Forestry, Science and Research Branch, Peterborough, ON. Science and Research Technical Manual TM-11.

Spectranalysis Inc. 2004. Introduction to the Ontario land cover data base, second edition (2000): Outline of production methodology and description of 27 land cover classes. Oakville, ON. 34 p.

Twery, M.J. 2004. Modelling in Forest Management. Pp. 295–305 *in* J. Wainwright and M. Mulligan (eds.), Environmental Modelling: Finding Simplicity in Complexity. John Wiley and Sons, Ltd., New York, NY.

White, J.C., M.A. Wulder, T. Hermosilla, N.C. Coops and G.W. Hobart. 2017. A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. Remote Sensing of Environment 194: 303–321.

Wilgress, A.T. 1919. Report of the Minister of Lands, Forests and Mines of the Province of Ontario, 1919-1922. Ryerson Press, Toronto, ON. (p. 176)

Wulder, M.A., J.A. Dechka, M.A. Gillis, J.E. Luther, R.J. Hall, A. Beaudoin and S.E. Franklin. 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. The Forestry Chronicle *79*(6): 1075–1083.

Wulder, M., M. Gillis, J. Luther and A. Dyk. 2001. A guide to the estimation of Canada's National Forest Inventory attributes from Landsat TM data. Natural Resources Canada, Canadian Forest Service, Ottawa, ON. Working paper. 84 p.

Wulder, M.A., J.C. White, M.D. Gillis, N. Walsworth, M.C. Hansen and P. Potapov. 2010. Multiscale satellite and spatial information and analysis framework in support of a large-area forest monitoring and inventory update. Environmental Monitoring and Assessment 170(1–4): 417–433.

Appendix

Table A1.1. Data availability for ensemble overlay confidence assessment in developing a disturbance database of Ontario's Area of the Undertaking. (NA=not available)

Date range		MNRF/AFFES	MNRF	MNRF/SRB		NRCAN	
	Landsat	Fire	Harvest	Fire	Harvest	Fire	Harvest
1972–1984	X	X	NA	NA	NA	NA	NA
1985–1989	X	X	NA	NA	NA	X	X
1990–2013	X	X	X	NA	NA	X	X
2014–2015	X	X	X	X	X	X	X
2016–2017	X	X	X	X	X	NA	NA
2018	X	X	NA	X	X	NA	NA
2019	X	X	NA	NA	NA	NA	NA

Table A1.2. List of acronyms that appear in this report.

Acronym	**Expanded form**
AFFES	Aviation, Forest Fire and Emergency Services
AOU	Area of Undertaking
BAP	Best available pixel
CFS	Canadian Forest Service
DEM	Digital elevation model
ENGO	Environmental non-government organization
ETM+	Enhanced Thematic Mapper Plus
GCS	Geographic Coordinate System
ISOData	Iterative self-organizing data
Lidar	Light detection and ranging
MNR	Ministry of Natural Resources
MNRF	Ministry of Natural Resources and Forestry
MSS	Multispectral Scanner System
NRCAN	Natural Resources Canada
NAD	North American Datum
NASA	National Aeronautics and Space Administration
OMNR	Ontario Ministry of Natural Resources
OMNRF	Ontario Ministry of Natural Resources and Forestry
OLI	Optical line imager
SPOT	Satellite Pour l'Observation de la Terre
SRB	Science and Research Branch
SRTM	Shuttle Radar Topography Mission
TM	Thematic Mapper
USGS	United States Geological Survey
UTM	Universal Transverse Mercator

(0k P.R. 20 05 25)
ISBN 978-1-4868-4433-3 (pdf)