Identification of Forest Road Construction Year from Historical Remotely Sensed Data by Integrating Change Detection and Geometric Connectivity

By

Kirsten Rachelle Noltie

A thesis submitted to the Graduate Program in Geography and Planning

in conformity with the requirements for the

Degree of Master of Science

Queen's University

Kingston, Ontario, Canada

October, 2023

Abstract

Current and highly accurate forest disturbance information is critical for monitoring and managing forested ecosystems. The types, rates, and responses to past disturbances allow for improved predictions for how forests will respond to current and future disturbances. Forest roads, a short-term forest stand replacing disturbance, are associated with ecosystem degradation. Determining the age of different elements of forest road networks can enable long term monitoring and assist with sustainable forestry management. To determine the age of construction of forest roads within Ontario's managed forests, a semi-automatic road extraction and Post-Classification Change Detection (PCCD) approach was developed with Best Available Pixel (BAP) composites of Landsat imagery. A study site was selected near Savant Lake, Ontario for the years between 1974 and 2019. The BAP composite approach uses sensor, day of year (DOY), distance to cloud or cloud shadow, opacity, and pixel brightness scores to create yearly composites which are free of clouds or anomalies and optimized for the intended analysis. An unsupervised classifier, Jenks Optimization, was used to differentiate road and non-road pixels from an image masked with current known road areas. This method aims to reduce the average deviation of each value from the mean of the class and increase the deviation of each class from the mean of other classes. PCCD was completed with the pre- and post-1984 periods resulting in overall accuracies of 80% and 68%, respectively. Four logic rules, which leverage the topological relations between road segments and address road gaps due to occlusions or sensor anomalies, were applied twice throughout the process. The logic rules increased the overall accuracy up to 3% for each annual binary road network and up to 3% for the final year of construction road network. Understanding historic types, rates, trends, and patterns of ecosystem change across large areas with high spatial detail is essential for effective forest monitoring and

sustainable management. Time series analyses based on optimized annual BAP composites and post-processing logic rules can provide information to support these endeavors.

***Keywords*—** Forest road, logging, forestry, Best Available Pixel (BAP) composite, pixel based composites, Landsat imagery, road extraction, change detection, classification, unsupervised classification, logic rules, post-processing, remote sensing, geographic information systems (GIS)

Acknowledgements

I would like to extend my sincere thanks to everyone who has supported the completion of this thesis. This research would not have been completed without the continued support of my supervisor, Dr. Dongmei Chen. Dr. Chen provided key guidance and the technical expertise that was crucial to my success. I am very grateful for the opportunities that she has extended to me throughout my journey at Queen's University. The comments, advice, and opportunities provided by Dr. Tarmo Remmel, from York University, and Marc Oullette, from the Ontario Ministry of Natural Resource and Forestry, were also vital to the successful completion of these manuscripts.

I would like to thank all the faculty and staff of the Queen's University Department of Geography and Planning. The work of the faculty, through their teaching, their commitment to their own high-quality research, and advice for graduate school provided essential inspiration, knowledge, and insights. To the staff, thank you for your continued efforts to ensure that all my course work, enrollment, funding, and teaching assistant documentation was always in order. We all know that nothing in the department would run successfully without your efforts!

I further wish to extend my gratitude to the Geography Cohort of 2018, who continue to inspire me with their commitment to their research, their careers, and their wonderful zest for life and adventure. I am so happy I was able to work alongside such amazing people and researchers. I look forward to seeing the amazing things you will continue to accomplish.

This work would not have been completed without the fierce support of Barb Lotan. I will be eternally grateful for her efforts and commitment to change. Her support has allowed me to continue my journey as a woman in STEM.

I wish to present my gratitude to all my friends and family. With special acknowledgement to Patricia, Brant, and Leigh-Anne Noltie as well as Jessica Lecours for their love, emotional support, and understanding.

And finally, to Edgar and Lucy Noltie. Their support, companionship, and editing efforts are always appreciated. So grateful to have you in my life for this journey.

Land Acknowledgement

Queen's University is situated on the traditional territory of the Anishinaabe and Haudenosaunee peoples. The study site for this research is on the traditional territory of Anishinaabeg encompassing Treaties #3, #5, and #9, as well as unceded territory. I am grateful to be able to live, research and play as an uninvited guest on these lands. It is important to acknowledge the significance of these territories for the Indigenous peoples who lived, and continue to live, upon them. The practices and spirituality of these Indigenous peoples are tied to the land and will continue to develop in relationship to the territory and all its inhabitants.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
AOT	Aerosol Optical Thickness
AVHRR	Advanced Very-High-Resolution Radiometer
BAP	Best Available Pixel
BEAST	Bayesian Estimator of Abrupt Change, Seasonal change, and Trend
BT	Band Test
C-V	Chan-Vese
C2C	Composite2Change
CART	Classification and Regression Trees
CFSA	Crown Forest Sustainability Act, 1994, S.O. 1994, c. 25, 1994
CON	Conventional Logging
COSSIM	Cosine Similarity Compositing
CPD	Change Point Detection and Analysis
CVA	Change Vector Analysis
DEM	Digital Elevation Model
DN	Digital Number
DOY	Day of Year
DOYSD	Day of Year Standard Deviation
EBP	Excluded Brightness Parameter
EROS	Earth Resources Observation and Science Center
ESPA	EROS Center Science Processing Architecture
ETM+	Enhanced Thematic Mapper Plus
ETOPO5	Earth Topography Five Minute Grid
FIM	Forest Information Manual
FMPM	Forest Management Planning Manual
FOSM	Forest Operations and Silviculture Manual
GARAC	Global and Local Regional Active Contour
GCP	Ground Control Point
GIS	Geographic Information System
GPS	Global Positioning System
GVF	Goodness of Variance Fit
ha	Hectare
HWBP	Heavily Weighted Brightness Parameter
IBP	Included Brightness Parameter
iid	Independent and Identically Distributed
ISODATA	Iterative Self-Organizing Data Analysis
KIA	Kappa Index of Agreement
km	Kilometer
K-L	Karhunen-Loeve
LaSRC	Land Surface Reflectance Code
LBF	Local Binary Fitting
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LiDAR	Light Detection and Ranging
LTMD	Long-Term Management Direction

LTS	Landsat Time Series
LULC	Land Use Land Cover
m	Metre
µm	Micrometer
MA	Map Algebra
MAX-NDVI	Maximum NDVI
MAX-RNB	Maximum RNB
MED-NIR	Median NIR
MERIS	Medium Resolution Imaging Spectrometer
MLC	Maximum Likelihood Classification
MNDMNRF	Ministry of Northern Development, Mines, Natural Resources and Forestry
MNRF	Ontario Ministry of Natural Resources and Forestry
MODIS	Moderate Resolution Image Spectroradiometer
MSS	Multispectral Scanner
nm	Nanometer
NASA	National Aeronautics and Space Administration
NBR	Normalized Burn Ratio
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
NLCD	National Land Cover Database
OLI	Operational Land Imager
OPFSF	Ontario Policy Framework for Sustainable Forests
ORN	Ontario Road Network
PAC	Phenology Adaptive Compositing
PC	Principal Component
PCA	Principal Component Analysis
PCCD	Post-Classification Change Detection
QA	Quality Assurance
RF	Random Forest
RIL	Reduced Impact Logging
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar
SDAM	Square Deviation of Array Mean
SDCM	Square Deviation of Class Mean
SLC	Scan Line Corrector
SPOT	Satellite Pour l'Observation de la Terre
SVM	Support Vector Machine
SWIR	Shortwave Infrared
TAR	Tape Archives
TIRS	Thermal Infrared Sensor
TM	Thematic Mapper
TOA	Top of Atmosphere
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
VCT	Vegetation Change Tracker
VHR	Very High Resolution

WELD	Web-Enabled Landsat Data
WGS	World Geodetic System
WPS	Weighted Parametric Scoring
WRS	World Reference System

Chapter 1: Introduction

1.1 Problem Statement and Contributions

The Canadian portion of the boreal forest experiences significant anthropogenic disturbances. In 2010, an estimated polygonal area of 24 million hectares (ha), 60% of which can be attributed to cut blocks (areas authorized for timber harvest), were identified. Approximately 600 000 kilometers (kms) of linear features related to resource extraction were mapped, with 80% attributed to roads and seismic exploration lines (Pasher et al., 2013). Thirty one percent of Canada's boreal forest is accessible by road. In 1987, the Ontario portion of the boreal forest had 33 000 km of logging roads, with an annual increase of approximately 1 700 kms of new roads (Wildlands League et al., 2001). Between 1990 and 2015, Canada's forested area decreased 0.3% (3.483 million km^2 to 3.471 million km^2) with 6% of the decrease due to forestry roads (Statistics Canada, 2018). Hosaka et al. (2014) note that between 6% to 25% of logged area is damaged from forest road networks. Forest roads have a wide range of environmental and social implications for the forests in which they exist, including soil degradation, carbon emissions, hydrological changes, habitat fragmentation, increased forest fire risk, biological invasions, and deforestation (Mann et al., 2019). As a result, it is essential to include the accurate detection and year of construction attribution for forest roads in forest monitoring approaches.

Previous forest road monitoring approaches have relied on aerial photography or high or moderate-resolution satellite imagery to digitize the linear disturbances; however, these methods have many limitations (Z. Chen et al., 2014). Other forest disturbance monitoring approaches often do not include a specific forest road class or are limited to the mid-1980s to 2010s (Hermosilla et al., 2015b; C. Huang et al., 2010; Kennedy et al., 2010; L. Liu et al., 2013). A semi-automatic road extraction and year of construction assignment algorithm using the Landsat

sensor series is proposed in this thesis to enhance the available forest road data for decision-making and planning within the Ontario boreal forest. There are four key objectives for this approach: (1) increase the temporal span of forest road monitoring to include both the underutilized Multispectral Scanner (MSS) sensor (1972 to 1974) and recent roads with the Operational Land Imager (OLI) sensor (2013 to present); (2) develop optimized pixel-based composites using the Best Available Pixel (BAP) protocol to enhance forest road extraction; (3) determine the feasibility of an unsupervised classification and change detection approach for attribution of forest road year of construction; and (4) enhance forest road extraction accuracy through the use of four topology based logic rules. These objectives were achieved through the development of a semi-automated road extraction and year of construction attribution algorithm for the Landsat MSS, Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and OLI imagery.

1.2 Thesis Organization

This thesis is organized into five chapters. Chapter 1 introduces forest roads and their monitoring. This chapter also describes the research objectives and contributions. Chapter 2 provides a review of the literature, including clearly defining forest roads and their characteristics. Characteristics of roads in medium resolution satellite imagery and the difficulties in their accurate detection are then discussed. Following, a description of the Landsat sensors and considerations for analyses using the series of sensors is provided. Finally, a review of prominent approaches for road extraction and change detection is described. Chapter 3 first reviews the BAP compositing approach first implemented by White et al., (2014). Next, it extends the BAP compositing approach to 1974 and beyond 2013 through the inclusion of the MSS and OLI sensors. The BAP protocol is then optimized for road extraction in the Ontario

boreal forest. Chapter 4 outlines the proposed algorithm, including road extraction, change detection, and post-processing using the four logic rules. The accuracy of the proposed algorithm is assessed, specifically examining the influence of the binary and network level implementation of the logic rules. Chapter 5 provides a summary of this thesis, contributions, and the directions for future research.

1.3 References

Chen, Z., Jefferies, B., Adlakha, P., Salehi, B., & Power, D. (2014). Monitoring linear disturbance footprint based on dense time series Landsat imagery. *Canadian Journal of Remote Sensing, 40*, 348–361. https://doi.org/10.1080/07038992.2014.987375

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment, 170*, 121–132. https://doi.org/10.1016/j.rse.2015.09.004

Hosaka, T., Niino, M., Kon, M., Ochi, T., Yamada, T., Fletcher, C., & Okuda, T. (2014). Effects of logging road networks on the ecological functions of dung beetles in Peninsular Malaysia. *Forest Ecology and Management, 326*, 18–24. https://doi.org/10.1016/j.foreco.2014.04.004

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment, 114*(1), 183–198. https://doi.org/10.1016/j.rse.2009.08.017

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. *Remote Sensing of Environment, 114*, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Liu, L., Tang, H., Caccetta, P., Lehmann, E. A., Hu, Y., & Wu, X. (2013). Mapping afforestation and deforestation from 1974 to 2012 using Landsat time-series stacks in Yulin District, a key region of the Three-North Shelter region, China. *Environmental Monitoring and Assessment, 185*(12), 9949–9965. https://doi.org/10.1007/s10661-013-3304-2

Mann, D., Agrawal, G., & Joshi, P. K. (2019). Spatio-temporal forest cover dynamics along road networks in the Central Himalaya. *Ecological Engineering, 127*, 383–393. https://doi.org/10.1016/j.ecoleng.2018.12.020

Pasher, J., Seed, E., & Duffe, J. (2013). Development of boreal ecosystem anthropogenic disturbance layers for Canada based on 2008 to 2010 Landsat imagery. *Canadian Journal of Remote Sensing*, *39*(1), 42–58. https://doi.org/10.5589/m13-007

Statistics Canada. (2018). *Human activity and the environment: Forests in Canada*. https://www150.statcan.gc.ca/n1/en/pub/16-201-x/16-201-x2018001-eng.pdf?st=ycHBl_tT

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., Coops, N. C., Hall, R. J., Hostert, P., Dyk, A., & Guindon, L. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing, 40*(3), 192–212. https://doi.org/10.1080/07038992.2014.945827

Wildlands League, Henschel, C., & McEarchern, G. (2001). *Good boreal forestry: Protecting the forest's critical characteristics*. http://wildlandsleague.org/attachments/Good%20Boreal%20Forestry%20FS4.pdf

Chapter 2: Literature Review

Roads are an essential anthropogenic feature required for many socio-economic functions. Despite their importance for human connectivity, roads are considered one of the most significant interferences into forested landscapes and are recognized for their negative impact on these regions (Mann et al., 2019). Forest roads are not uniform features and have varying functions, construction practices, appearances, impacts, related legislation, and monitoring requirements. Forest road monitoring is increasingly relying on remote sensing to cover large spatial expanses to identify and date forest disturbances (Bilyk et al., 2021). This chapter reviews the breath of literature on forest roads and their effective monitoring. Section 1 describes forest roads, their classification and construction within Ontario, and related legislation. Section 2 covers the development of road models in remote sensing imagery and the challenges in effectively and efficiently extracting them for use in decision making. Section 3 describes the Landsat Mission and why this sensor is the most appropriate for this road extraction context. Section 4 provides an overview of Landsat Sensors. Finally, Section 5 and 6 discuss automated road extraction and change detection approaches, respectively. These topics provide key background information for the proposed approach for road extraction and year attribution for forest roads in Ontario.

2.1 Ontario's Forests and Forest Disturbances

In Canada, predominately forested ecosystems cover approximately 65% of the country, or approximately 650 million hectares (ha) (Wulder et al., 2011). Globally, Canadian forests represent 10% of all forested areas and 28% of the boreal zone (Natural Resources Canada, 2021; White et al., 2014). These ecosystems encompass varying landcovers including trees, shrubs, wetlands, and lakes. The boreal forest is the largest forested region in Ontario, containing

two thirds of Ontario's forests or 50 million ha (Ontario Ministry of Natural Resources and Forestry, 2019; Wulder et al., 2011). Due to the vast expanse of this region and impact of various anthropogenic and natural forest disturbances, remote sensing is increasingly utilized (Win et al., 2012)

Natural and anthropogenic disturbances can occur as both non-stand and stand replacing forest disturbances. Non-stand replacing disturbances are often subtle changes to forest environments that occur gradually over a longer period (Goodbody et al., 2020; Hermosilla et al., 2015b). These changes are related to drought, stress, insect infestation, or disease. They provide insights on factors influencing forest growth, health, and broader underlying forest dynamics. Non-stand replacing changes are often not documented as the magnitude of change is low, and they may not result in a change to land cover class (Ahmed et al., 2017; Gómez et al., 2015).

In contrast, stand replacing disturbances result in a complete change in land cover class. This type of disturbance can result in long- or short-term changes in land cover. For example, the development of roads, urbanization, industrial activity, harvesting, deforestation, and wildfire are all stand replacing disturbances. Urbanization, industrial activities, paved road development and deforestation all result in long-term change. Harvesting, wildfires, and unpaved forest roads are short-term stand replacing disturbances that will return to a vegetation state (Hermosilla et al., 2015b). Quick onset disturbances (i.e., wildfires) will influence forest structure, function, and composition whereas slower while continuous changes (i.e., insect infestation) have longer term and less obvious impacts.

These disturbances have varying spatial impacts. Nationally, the boreal forest is mostly intact; however, wildfires have significant effects on northern regions whereas anthropogenic disturbances are more frequent in highly populated areas or regions with high forest productivity

(Hermosilla et al., 2015b). Nationally, fires impacted around 40.6 million ha between 1985 and 2010. In the same twenty-five-year period, only 16.7 million ha were subject to harvesting (Hermosilla et al., 2019). This translates to approximately 2 million ha impacted by wildfire, 1 million ha impacted by harvesting activities, and 45 00 ha impacted by deforestation annually (Hermosilla et al., 2015b).

The monitoring of various forest disturbances is essential for meeting a variety of forest management needs, including carbon accounting, biodiversity assessment, habitat fragmentation analyses, wildlife management, and climate change mitigation approaches. The cumulative effects of these disturbances can be assessed at regional and national scales (Z. Chen et al., 2014; Gómez et al., 2015; Hermosilla et al., 2019) Determining the dates of forest road construction, a short-term stand replacing disturbance, serves as the focus of this work.

2.1.1 Forest Roads Defined

Logging or forest roads are a specific type of linear forest disturbance. The terms logging or forest road are often used colloquially without further definition; however, forest roads may be defined in a variety of ways. Ontario forest road specifications will be defined in Chapter 2 Section 1.3. Within broader research, logging roads may include paved or unpaved roads built to allow wheeled vehicles to transport logs from the forest. These roads can be broadly divided into primary access roads and secondary roads. Primary roads serve to permanently access the forest concession. The secondary roads branch off the primary roads and are usually dead-end roads built for a limited temporal period. Their lifespan may range from a few months to a few years but are often resultant from a single timber harvesting operation (Kleinschroth et al., 2016; Kleinschroth & Healey, 2017). These definitions usually exclude roads built primarily for public use (Kleinschroth et al., 2016).

Depending on the study, skid trails may be included or excluded within the road definition. Skid trails are used to extract logs from their felling site using heavy machinery (Kleinschroth & Healey, 2017; Minn et al., 2022). Kleinschroth, Gourlet-Fleury, et al. (2016) note that skid trails are often designed for tracked vehicles and are therefore narrower than primary and secondary roads. This design prevents skid trails from having a continuous opening in the canopy which would allow them to be effectively monitored from remotely sensed imagery. In this application, only the primary, branch, and operational roads defined by the Ontario Ministry of Natural Resources and Forestry (MNRF) are considered. Due to the spatial resolution of the sensors selected as well as inability to penetrate the tree canopy for the extent of the study period, skid trail extraction was not within the scope of this study.

2.1.2 Forest Roads and Ontario Forestry Legislation

Currently, 44% of Crown Land (land owned by the Ontario provincial government) forests are managed forests (Ontario Ministry of Natural Resources and Forestry, 2020). Forest access roads on Crown Land are currently managed under the Ontario Policy Framework for Sustainable Forests (OPFSF). This framework provides a broad direction for forest policies with a focus on sustainability. The policies in the OPSF can be implemented through the Crown Forest Sustainability Act, 1994, S.O. 1994, c. 25 (CFSA). This act enables legislation for the regulation of forest planning, monitoring, forestry operations, licensing, trust funds, processing facilities, remediation, and enforcement of these policies (Ontario Ministry of Natural Resources and Forestry, 2020). The CFSA requires the provision of four manuals to guide and support the management of various aspects of Ontario forestry. These manuals include: (1) The Forest Management Planning Manual (FMPM); (2) the Forest Information Manual (FIM); (3) the Forest Operations and Silviculture Manual (FOSM); and (4) the Scaling Manual (SM) (Ontario Ministry of Natural Resources and Forestry, 2020).

The FMPM provides the broad requirements for the management of Ontario's forest systems, including a description of planning, implementation, monitoring, reporting specifications, and documentation cycles. The FMPM for each management unit is prepared by an interdisciplinary team every ten years and is valid for a ten-year period. This document provides the long-term management direction (LTMD) of the forestry unit and the planning operations of forest access (including forest access roads), harvest, renewal, and maintenance. For each FMPM, the team must confirm and update the existing roads and water crossings within their management unit. The Ministry currently denotes three forest road classes: primary, operational, and branch roads, each of which will be described below.

2.1.3 Ontario Forest Road Classes, Construction, and Decommissioning

Roads within Ontario forests are not uniform features. These roads range from unpaved forest trails, gravel industrial roads, and wide paved highways with medians (Hermosilla et al., 2015b). The MNRF defines three hierarchical road types: (1) primary roads; (2) branch roads; and (3) operational roads. Primary roads act as the main road system and provide access to a management unit. These roads are usually permanent, but do not always have consistent use for forestry management purposes. When considering the construction of a new primary road or extension to a primary road, potential 1 kilometer (km) wide corridors must be evaluated. A variety of factors are considered in the planning of new primary roads. These roads must be aligned with the proposed harvest distribution for at least four planning cycles (40 years). Considerations include how the physical environment, proposed engineering, and safety concerns act as constraints or opportunities, particularly when considering the development of other regional resources. There must be confirmation that the roadway aligns with existing planning initiatives in the region (Ontario's Crown Land Use Policy Atlas, conservation interests, provincial park and lake management plans and stewardship agreements). The results

of consultation with known interested parties, organizations, as well as First Nation and Métis communities are also incorporated into road planning. The maintenance and protection of important ecological features are identified for the proposed corridors. The proposed use of the roads, which includes public access, use restrictions, maintenance requirements, decommissioning specifications, and transfer provisions, are reviewed. Finally, the relative cost of road corridor alternatives, including construction, use management, maintenance, transfer, and decommissioning, are evaluated (Ontario Ministry of Natural Resources and Forestry, 2020).

Branch roads are roads, other than primary roads, which extend from primary or other branch roads. These roads provide access to and through areas of operation within a management unit. Whenever a new road, that has not already been classified as a primary or operational road, is required to provide access between, though, or to separate areas of operation, this road will be considered a branch road. Branch roads require the consideration of 1 km corridors and have similar planning requirements as primary roads (Ontario Ministry of Natural Resources and Forestry, 2020)

Operational roads are roads within designated operational road boundaries, other than primary or branch roads, which provide short-term access for harvest, maintenance, or renewal processes. These roads are generally not maintained after they are no longer used for a specific harvesting operation and are often decommissioned. Areas where operational roads can be constructed within a ten-year period will be identified through operational road boundaries. As opposed to road corridors, operational road boundaries allow some flexibility in road location in relation to local terrain limitations or other factors (Ontario Ministry of Natural Resources and Forestry, 2020). A variety of conditions may be applied within these operation zones including the following: no road construction permitted, specified methods of construction or use of

structures, construction time constraints, marking or flagging of locations prior to construction, and notification to First Nations and Métis communities in advance of construction. For each final operational road completed, the location and conditions will be reported back to the ministry.

When considering decommissioning roads, there are two major types: physical and natural abandonment. For physical abandonment, there is a deliberate act by the forest management team to make the road unusable by vehicular traffic. Physical abandonment may include processes to reduce the environmental impact of ceasing to maintain the roadway. Natural abandonment occurs when no physical changes are made to the road, but road maintenance has ceased. Specific steps should be taken to reduce erosion and to carefully remove structures that cannot be left intact (Ontario Ministry of Natural Resources, 1990).

For each of these road types within a road network, the stakeholders must provide a use management strategy. This strategy includes maintenance, monitoring, access, or restrictions, when and how the roads will be transferred to the management of the MNRF and plans for decommissioning.

For existing roads, the FMPM planning team must confirm the following: (1) road or road network identifier, (2) road class, (3) geographic location, (4) assigned responsibility for the road and water crossing, and (5) the associated management strategy. In addition to confirmation of current forest access, historic forest condition information is also essential for LTMD development. Historic forest information allows stakeholders to understand long-term forest dynamics as well as the effects of past forest policy and management (Ontario Ministry of Natural Resources and Forestry, 2020). Although the historical development of forest access is included in the summary requirements, often the specific date of historic road access

construction is not known, particularly for operational roads. Increased information about roadways will help the planners understand how previous human intervention and natural processes relate to current forest conditions within each management unit.

2.2 Roads in Remotely Sensed Imagery

2.2.1 Road Characteristics

There are several linear forest disturbances visible in remotely sensed imagery such as seismic lines, roads, pipelines, and powerline rights of way. These disturbances are common features within resource extraction intensive regions. These anthropogenic disturbances can be confused with natural linear features including streams, ravines, and wide creeks (Bocking et al., 2017). It should be noted that some man-made linear disturbances may not be visible in remotely sensed imagery, as they occur under the tree canopy. These disturbances include underground pipelines or trenches and embankments that serve to demarcate land boundaries (Affek et al., 2017).

Forest access roads in Ontario follow very strict construction and maintenance requirements that dictate their appearance in the natural environment. The way that roads are conceptualized within a remotely sensed image differs from their complexity on the ground. One of the key problems in road extraction is how to describe road features, as this directly relates to the choice of methodology. In general, most researchers agree that roads are bright, elongated geometric features (linear) that have slowly changing grey values. However, depending on the resolution of the image, spectral range and sensitivity, scale of the object relative to the image resolution, image complexity, and noise interference from buildings, shadows and other road obstructions, more detailed road features should be considered (J. Chen et al., 2014; Y. Li et al., 2009; J. Wang, 1993; W. Wang et al., 2016).

A review of the literature indicates that there are various road characteristics that can be considered in remotely sensed imagery, including: width, area, length to width ratio, texture, spectral properties, brightness, and topology. All these factors are related to the functional purpose the road exhibits as well as environmental and anthropogenic constraints (H. Wang et al., 2016). Depending on the functional purpose of a given roadway, roads may have a variety of widths. However, these widths will not suddenly vary. Instead, roads will often intersect with roads with contrasting widths (Bae et al., 2015; Cardim et al., 2018; Y. Li et al., 2009; Soni et al., 2018; W. Wang et al., 2016). A road is a continuous feature and road systems will usually have a larger area than other manmade structures. This attribute allows erroneously identified segments with small areas to be ruled out as belonging to road classes (Shi et al., 2014). In remotely sensed images, roads are generally narrow and longer than other artificial objects. This high length to width ratio allows roads to be conceptualized as linear geometric features with slow variation in curvature (Besbes & Benazza-Benyahia, 2014; Cardim et al., 2018; Sghaier & Lepage, 2016; Shi et al., 2014). The road texture relates to the spatial distribution of pixel intensity values within a given region; this is a measure of homogeneity (W. Wang et al., 2016). The spectral characteristics of roads will vary depending on the type of image. In panchromatic images, roads are generally lighter or darker than the background; however, in some area of the electromagnetic spectrum, the grey levels of main roads do not vary significantly from the background (Dal Poz et al., 2012; Geman & Jedynak, 1996). The spectral complexity of roads is also increased due to linear features varying spectrally and spatially along their extent. The intensity of the contrast between a road segment and the background may slowly vary across time and space (Wang, 1993). This variation often makes it difficult to extract roads from imagery using only local spectral information (Kahraman et al., 2018). When considering the

topological features of a road network, roads will include intersections. Typical road junctions include “T” intersections, “Y” intersections, “+” intersections as well as “O” roundabouts (Cardim et al., 2018; Qiaoping & Couloigner, 2004; W. Wang et al., 2016). These interactions are often referred to as junctions or nodes. Roads are not suddenly interrupted but may end with a road termination (Qiaoping & Couloigner, 2004; W. Wang et al., 2016).

The complexity of the background, or surrounding area also strongly impacts linear feature delineation and extraction. Image analysis in urban areas will see more complexity in the background and may experience more building shadows. This complexity may result in greater radiometric heterogeneity along the roadway than in rural areas (L. Liu & Lim, 2016; Soni et al., 2018; H. Wang et al., 2016). The combination of these features as well as their functional purpose relate to the general road classification within research applications. Road definitions vary significantly as standards for road construction vary heavily both geographically and temporally.

Many road extraction and delineation algorithms have been proposed in the literature; there are no feature extraction approaches that are successful for all road types (Z. Chen et al., 2014; W. Wang et al., 2016). The performance of a given methodology is strongly dependent on the type of road to be detected. What produces strong results for an urban six lane paved highway may not be successful for a gravel country road in a rural area (Y. Li et al., 2009; L. Liu & Lim, 2016; Ünsalan & Boyer, 2011). Most methods focus on a specific type of road with a very limited set of characteristics under tight conditions (Cardim et al., 2018). In addition to the type of road, the type of imagery being used will have a strong impact on results (Q. Zhang et al., 2002).

2.2.2 Challenges in Road Delineation in Remotely Sensed Imagery

Although roads, a stand replacing disturbance, are well suited for monitoring through remotely sensed imagery, there are multiple barriers to their successful extraction. These include the sensor characteristics, background complexity, road occlusion, and the limits of photo interpretation. In addition to the variation in road types, each sensor aboard a remote sensing satellite has its own unique characteristics, including image size, spatial resolution, acquisition geometry, and spectral resolution (Cardim et al., 2018; Kahraman et al., 2018). Each of these unique sensors are also impacted differently by weather conditions, changes in light, and responses of earth surface features. Analysts must be careful to consider not only the target road class characteristics but how the sensor will impact the extraction of the structural road model (Kahraman et al., 2018; Soni et al., 2018).

Some of the largest issues relating to roads in remotely sensed imagery are the background complexity, shadows, and occlusions. Many researchers rely on the fact that road pixels should not vary significantly in grey value from other road pixels, but instead should have a stark contrast to the background (L. Liu & Lim, 2016). However, if the researcher relies on the spectral homogeneity of the road, other objects that are spectrally similar may be misclassified as road segments (R. Li & Cao, 2018). Such areas may include bare soil, bedrock, or areas of new construction (R. Li & Cao, 2018). Depending on the type of imagery, this homogeneity may result in the extracted roads being wider than the ground truth roads or unrelated objects being classified as road (T. Zhou et al., 2019). In contrast, shadows, terrain differences and occlusions from cars, trees or buildings reduce the spectral homogeneity (G. Cheng et al., 2017; R. Li & Cao, 2018; Woodcock et al., 2008). As the spatial continuity is interrupted, there will be gaps in the digital road extraction (J. Wang et al., 1992; T. Zhou et al., 2019). Thes gaps may happen for the entire road segment or during centerline extraction and result in the disruption to the road

topology (L. Liu & Lim, 2016). It should also be noted that due to the limits of photointerpretation, some roads may never be delineated or extracted. For example, roads that are not visible through the canopy will not be accessible through multispectral or hyperspectral imagery (W. B. Cohen et al., 2010).

2.3 The Landsat Mission

Due to the breadth, temporal span, repetitive data acquisition, and image quality, satellite remote sensing is an essential data source for generating the dates of construction for forest road segments (Bodart et al., 2011; Chander et al., 2009). The Landsat satellites (1-5, 7-9) provide the longest continuous series of satellite imagery globally (Chander et al., 2009; Woodcock et al., 2008). First launched in 1972, these medium spatial resolution sensors are heavily utilized for monitoring earth surface changes for decision making. Over the past five decades, technological advances have enabled the amount and quality of data collected by the Landsat program to steadily increase (Chander et al., 2009).

Landsat imagery was selected for this application due to its spatial and temporal resolutions. Previous research on forest roads often utilized aerial photography or high or moderate-resolution satellite imagery to digitize the linear disturbances. There are significant limitations to this approach. High-resolution imagery and aerial photography often have inconsistent acquisition and small spatial coverage (Z. Chen et al., 2014). The limited acquisition range of this type of imagery often prevents the analyst from accurately determining the age of the road segments. Extending the interpretation to large areas is often prohibitively difficult due to the high processing complexity and logistical constraints of high-resolution imagery. Overall, it can be concluded that high resolution imagery and aerial photography are not suitable for this application (Z. Chen et al., 2014).

Coarse spatial resolution imagery is also not appropriate for delineating linear disturbances. Sensors like Advanced Very-High-Resolution Radiometer (AVHRR), Moderate Resolution Image Spectroradiometer (MODIS), and Medium Resolution Imaging Spectrometer (MERIS) with resolutions ranging between 250 m to 1 km are often used for land use monitoring at global or continental scales (Bodart et al., 2011; Thompson et al., 2015). Although this type of imagery is often low cost with short revisit periods, the coarse resolution only allows for the detection of larger scale forest patterns. Local or regional patterns are often lost due to the pixel sizes (Bodart et al., 2011). Larger pixel sizes subsume moderate sized landscape features, reduce the variation between pixels, and limit the ability to map finer patterns or delineate detailed land cover classes (Thompson et al., 2015). Processes like logging, deforestation, urban sprawl, and land abandonment all occur at spatial scales that cannot be addressed by coarse resolution imagery (Griffiths et al., 2013).

To overcome the issues with high-resolution and coarse resolution imagery as well as aerial photography, moderate spatial resolution imagery (60 to 30 m) can be used to map and monitor forest changes at regional scales (Bodart et al., 2011; Griffiths et al., 2014). This pixel size prevents the mapping of finer scale earth surface features; however, landscape level patterns can be captured while avoiding the computational complexity associated with hyper spectral and spatial sensors (Young et al., 2017). The use of moderate resolution imagery for manual interpretation often leads to inconsistencies due to interpretation variation (Z. Chen et al., 2014). The use of automated approaches can significantly reduce this variation.

The characteristics of the Landsat sensors allow for the monitoring of natural disturbances and anthropogenic activity over time and space. Although this sensor is suitable for forest road analysis, there are numerous challenges, particularly when working with large

geographic extents. The relatively small swath and sensor-specific-field-of view of the Landsat sensors requires the use of multiple Landsat footprints for full regional coverage (Griffiths et al., 2013, 2014). Additionally, frequent cloud cover and phenological and radiometric impacts from the time of acquisition require specific workflows for processing this data for use over large areas (Griffiths et al., 2014). Free and open access to the full Landsat archive in addition to technological advances has enabled scientists to maximize the information available within this imagery (Bullock et al., 2020; Hermosilla et al., 2015a; Woodcock et al., 2008). Advanced and automated pre-processing methods, improved processing algorithms, and enhanced data storage allows for significant image processing (Griffiths et al., 2014). One such method is Best Available Pixel (BAP) compositing discussed in Chapter 3.

The length of the Landsat imagery archive makes it increasingly beneficial for long-term time series-based change analyses of earth surface features. Not only are there five decades of global imagery available, but the sensors also have a low revisit period. Landsat sensors aboard satellites 1-3 have an 18-day revisit period whereas Landsat sensors aboard satellites 4-5, 7-9 have a 16-day revisit period. These temporal characteristics facilitate the analysis of landscapes through time at a variety of scales (Young et al., 2017). Researchers have utilized Landsat imagery for forest monitoring since the start of the Landsat program (Heller, 1975). Prior to the opening of the Landsat Archive in 2008, researchers predominately focused on time intervals between three and ten years using only one or two scenes. Studies with annual time intervals usually utilized a single Landsat scene. Applications that utilized greater spatial extents were limited to coarser time intervals due to computational complexity (W. B. Cohen et al., 2010; Hermosilla et al., 2019). Although data and processing limitations necessitated the monitoring of forest ecosystems at a coarse temporal frequency, forest processes are not well characterized

through bi-temporal style methods. Using only a few time points may capture the focal forest loss or fragmentation but often does not address regrowth or concurrent forest changes (Hermosilla et al., 2019). Forest disturbances and their underlying mechanisms vary across time and space at multiple scales. Forest spectral responses must be interpreted spatially and temporally to develop accurate conclusions (Gómez et al., 2015; Hermosilla et al., 2019). Following the opening of the USGS Landsat archive, continuously monitoring large geographic areas at fine time intervals is becoming the norm. These approaches are often referred to as dense Landsat Time Series (LTS) analyses (W. B. Cohen et al., 2010; Hermosilla et al., 2019).

Despite the use of LTS datasets, many time series studies have not used the full duration of the Landsat archive. Instead, researchers have used annual data starting with the Thematic Mapper (TM), followed by the Enhanced Thematic Mapper Plus (ETM+) and recently the Operational Land Imager (OLI) sensors. These approaches have not leveraged the twelve years of data captured by the Multispectral Scanner (MSS) system from 1972 to 1984 (Savage et al., 2019). This analysis will use almost the entire duration of the Landsat program (1974 to 2019) to determine accurate forest road dates. Each of the sensors included in this analysis will be discussed in detail in the section below.

2.4 The Landsat Sensors

This semi-automatic approach uses imagery from the Landsat sensors onboard Landsat satellites 1-5, 7 and 8. These satellites can be broadly characterized in three different groups: (1) Landsat 1-5: MSS Sensor, (2) Landsat 4-5: TM Sensor and 7: ETM+ Sensor, and (3) Landsat 8: OLI Sensor.

2.4.1 Landsat 1-5: MSS Sensor

The MSS sensors completed continuous data collection from July 1972 to October 1992 (Helder et al., 2012). Landsat 1-5 MSS sensors all have four reflective bands at 79 m spectral resolution (resampled to 60 m). The spectral range is 0.6 to 1.1 micrometers (µms). The MSS sensor on Landsat 3 also included a fifth thermal band (Young et al., 2017). The naming convention shifts between sensors with Landsat 1-3 having reflective bands 4 through 7 and Landsat 4-5 having bands 1 through 4. Landsat 1-3 had an 18-day revisit interval and used World Reference System (WRS)-1. Landsat 4 and 5 had a 16-day revisit interval and used WRS-2 (Earth Resources Observation and Science Centre, 2018). Please see Table 1 for the specific band designations.

Table 1: MSS and TM band designations (United States Geological Survey, n.d.)

LANDSAT MSS SENSORS				**LANDSAT TM SENSORS**		
Landsat 1-3 Bands	**Landsat 4-5 Bands**	**Wavelength (µm)**	**Resolution (m)**	**Landsat 4-5 Bands**	**Wavelength (µm)**	**Resolution (m)**
Band 4 - Green	Band 1 - Green	0.5-0.6	60	Band 1 - Blue	0.45-0.52	30
Band 5 - Red	Band 2 - Red	0.6-0.7	60	Band 2 - Green	0.52-0.60	30
Band 6 - Near Infrared (NIR)	Band 3 - Near Infrared (NIR)	0.7-0.8	60	Band 3 – Red	0.63-0.69	30
Band 7 - Near Infrared (NIR)	Band 4 - Near Infrared (NIR)	0.8-1.1	60	Band 4 - Near Infrared (NIR)	0.76-0.90	30
				Band 5 - Near Infrared (NIR)	1.55-1.75	30
				Band 6 - Thermal	10.40-12.50	120 (30)
				Band 7 - Shortwave	2.08-2.35	30

Infrared (SWIR)		

Table 2: ETM+ and OLI band designations (United States Geological Survey, n.d.)

LANDSAT ETM + SENSOR			**LANDSAT OLI SENSOR**		
LANDSAT 7 BANDS	**Wavelength (µm)**	**Resolution (m)**	**Landsat 8-9 Bands**	**Wavelength (µm)**	**Resolution (m)**
Band 1 - Blue	0.45-0.52	30	Band 1 - Coastal aerosol	0.43-0.45	30
Band 2 - Green	0.52-0.60	30	Band 2 - Blue	0.45-0.51	30
Band 3 - Red	0.63-0.69	30	Band 3 - Green	0.53-0.59	30
Band 4 - Near Infrared (NIR)	0.77-0.90	30	Band 4 - Red	0.64-0.67	30
Band 5 - Shortwave Infrared-1 (SWIR-1)	1.55-1.75	30	Band 5 – Near Infrared (NIR)	0.85-0.88	30
Band 6 - Thermal	10.40-12.50	60 (30)	Band 6 - Shortwave Infrared-1 (SWIR-1)	1.57-1.65	30
Band 7 - Shortwave Infrared-2 (SWIR-2)	2.09-2.35	30	Band 7 - Shortwave Infrared-2 (SWIR-2)	2.11-2.29	30
Band 8 - Panchromatic	0.52-.90	15	Band 8 - Panchromatic	0.50-0.68	15
			Band 9 - Cirrus	1.36-1.38	30
			Band 10 - Thermal Infrared-1	10.6-11.19	100

(TIRS-1)		
Band 11 - Thermal Infrared-2 (TIRS-2)	11.50-12.51	100

When considering the Landsat archive in its entirety, the imagery before 1984 (MSS data) is noisier than post-1984 datasets. The MSS data has significant weaknesses compared to the imagery collected from newer sensors, including fewer bands, relatively coarser spatial resolution, narrow spectral resolution, and low radiometric resolution (Savage et al., 2018). There are also significant data gaps; however, the imagery between 1972 and 1984 contains important data that can be incorporated into time series analyses with reasonable processing steps (Braaten et al., 2015; Savage et al., 2018).

2.4.2 Landsat 4,5 and 7: TM and ETM+ Sensors

The next series of satellites, Landsat 4, 5, and 7, include either the TM or ETM+ sensors. Landsat TM has a spectral range of 0.45 to 12.5 µm across seven bands, a spatial resolution of 30 m (120 m for the thermal band), a revisit period of 16 days, 8-bit radiometric resolution, and utilizes WRS-2 (Neigh et al., 2013). Please see Table 1 for the specific band designations.

Landsat 6 failed to reach orbit on October 5th, 1993; therefore, no data was collected from the Enhanced Thematic Mapper. The ETM+ sensor features a spectral range of 0.45 to 12.50 µm, a spatial resolution ranging from 15 m (panchromatic) to 60 m (thermal), a revisit period of 16 days and continues to use the WRS-2. Please see Table 2 for the specific band designations. The Landsat 7 products are delivered with 8-bit radiometric resolution. After the ETM+ Scan Line Corrector (SLC) failed in June 2003, imagery has been acquired and delivered with significant data gaps (Landsat Missions, n.d.-a).

2.4.3 Landsat 8: OLI

The OLI sensor includes nine bands with a spectral range of 0.43 to 2.29 µm, 16-bit pixel depth, with a spatial resolution of 30 m for the reflective bands and 15 m for the panchromatic band. Please see Table 2 for the specific band designations. Like the TM and ETM+ sensors, there is a 16-day revisit interval and data are collected using the WRS-2 (Landsat Missions, n.d.-b). Landsat 9, with OLI-2 and TIRS-2 onboard, was launched in 2021 (Neigh et al., n.d.). All data collected by this sensor is outside of the study period and therefore not currently included in the algorithm.

2.5 Change Detection Approaches

Most change detection methods are dependent on spectral differences. Reflectance or intensity differences are attributed to changes in the major LULC type within a pixel or object as opposed to interference from the atmosphere, illumination, viewing angle or other system variations (Afify, 2011; Radke et al., 2005; Sundaresan et al., 2007; Q. Zhang et al., 2002). The general process for change detection involves identifying the location of change, classifying the change, quantifying the types of change, and assessing the accuracy of the change through a variety of statistics (Amin & Shivakumar, 2017; Hussain et al., 2013; Radke et al., 2005). Basic change detection methods take the image sequence (two images) and generate a binary image, referred to as a change mask, which identifies changes to the last image. Change detection can be expressed in the following generic rule presented in Equation 1.

$$B(x,y) = \begin{cases} 1 \\ 0 \end{cases} \quad [1]$$

Where 1 represents a significant change in pixel brightness, *B*, at location x, y

And 0 otherwise (Radke et al., 2005).

Change detection has been investigated by researchers for multiple decades, with many approaches being proposed to meet the needs of different applications (Minu & Shetty, 2015). Change detection approaches are often divided into two major categories: pre-classification change detection (the analysis of a transformed image of the same area on at least two different dates) and PCCD (the analysis of at least two pre-existing classifications of the same area on different dates) (Hussain et al., 2013; Peiman, 2011). The level of change information derived from change detection approaches can also be divided into two main categories: binary changes (change or no change) and detailed "from-to" change classes (post-classification comparison for multiple classes) (Hussain et al., 2013). Depending on the application, the choices from either category above can be used successfully. Researchers often test a variety of change detection techniques to determine which method provides the best quantitative results in each application (Minu & Shetty, 2015).

Map algebra (MA) (image differencing and image rationing), principal component analysis (PCA), change point analysis (CPA), and PCCD are common change detection methods. These methods are all pixel-based approaches. Regional level analyses using medium resolution imagery usually employ a pixel-based approach as opposed to object-based analyses (Hussain et al., 2013). It should be noted that change detection methods are not applied in isolation. Often multiple methods, referred to as hybrid approaches, are applied. Hybrid approaches are often used to locate as well as identify the type of change if change attribution is not intrinsic to a given method (Tewkesbury et al., 2015).

Map Algebra for Change Detection (MA)

There are multiple forms of MA for change detection. The most common are image differencing and image ratioing. Image differencing is a very popular approach among

researchers as it is simple to implement and interpret (Minu & Shetty, 2015; Sundaresan et al., 2007; Tewkesbury et al., 2015). To employ this algorithm, one image is subtracted from another image. These images are of the same spatial area, but from different dates (Amin & Shivakumar, 2017; Sundaresan et al., 2007). The images must be co-registered or spatially aligned (Rosin & Ioannidis, 2003). The difference of the intensity values for each pixel or feature are stored in a third image. These intensity differences should reflect the type and magnitude of change present on the ground (Tewkesbury et al., 2015). An analyst will select a threshold value to separate meaningful change from noise to create a binary change image (Radke et al., 2005; Sundaresan et al., 2007). Image ratioing is a similar process, however, the bands are divided instead of subtracted (Afify, 2011). This process may better enhance the small differences between the spectral features on the earth's surface (Yang et al., 2019). Within the category of MA for change detection, image differencing is more popular as comparative studies often indicate it produces superior results (Q. Zhang et al., 2002).

There are multiple forestry applications which employ a MA approach to change detection. Win et al. (2012) analyzed the impact of selection logging on natural teak forests in Myanmar using image differencing on two SPOT-5 images. Their work indicated that logging roads impacted canopy change more than tree felling. Borrelli et al. (2014) completed an analysis of forest change in Italy using Landsat imagery between 2002 and 2011 using a Normalized Difference Vegetation Index (NDVI) pixel-oriented image differencing technique. Their approach achieved an overall accuracy average of 0.997, with a Kappa Index of Agreement (KIA) of 0.77. These studies indicate the suitability of MA approaches for moderate resolution forest change detection applications.

Principal Component Analysis (PCA)

Another form of change detection is PCA, a linear transformation technique. The technique may also be referred to as eigenvector, Hotelling, or Karhunen-Loeve (K-L) transformation (Koutsias et al., 2009). Despite the multiple names, these transformations are all based on eigenvalue decomposition of the covariance (or correlation) matrix of the image bands under investigation. This process converts the correlated spectral bands into a set of linearly uncorrelated bands, referred to as principal components (PCs). These PCs are usually visualized as greyscale images representing the component score (Dadon et al., 2019; Rodarmel & Shan, 2002). In addition to change detection, PCA has multiple applications including reducing data dimensionality, image compression and image enhancement (Fung & LeDrew, 1987; Koutsias et al., 2009).

PCA operates on the assumption that multi-, hyper-, and ultra-spectral image bands are highly correlated. The imagery bands are often numerically and visually similar. This correlation arises due to the interrelationship between multiple factors. Earth surface features within a series of bands may have similar spectral signatures. For example, within the visible portion of the electromagnetic spectrum, vegetation has a low reflectance. Additionally, there may be some areas of sensor band overlap, although usually quite small (Schowengerdt, 2007).

There are two predominant ways to use the PCA approach for remote sensing change detection. The first method is to conduct a PCA on multiple image stacks, then use the PC image in other change detection frameworks (image ratioing or differencing) (Deng et al., 2008). This approach assumes that the derived PCs contain similar spectral information, and the area of change is a relatively small proportion of the image (Fung & LeDrew, 1987). The next method takes multiple image bands, collected at two different times, and combines them. The combined bands are then transformed into PCs. There will be high correlation between two images in

unchanged areas and lower correlation in areas with change (Deng et al., 2008). Generally, when PCA is applied to a multidate integrated set of bands, minor components will contain most of the change information whereas higher order PCs will have unchanged information. In practice, the amount of change information contained in each PC varies between images (Fung & LeDrew, 1987; P. Gong, 1993; Koutsias et al., 2009). Although an analysis of the PCs can indicate where a change has taken place, this method cannot provide change attribution or determine the specific change that has occurred (Fung, 1992). Researchers often employ other change detection approaches like post-classification change detection, visual inspection; and classification methods to determine "from-to" class information (Deng et al., 2008; Fung, 1992; Radke et al., 2005).

Based on the literature, there are few change detection applications of PCA for forest disturbance monitoring. One such study was conducted by Lasaponara (2006) who evaluated the impact of forest fires and desiccation on Sicilia Island between 1999 and 2002. Fung and LeDrew (1987) also explored some forest disturbances in their PCA change detection study of south-central Ontario, Canada. Newer publications rely on PCA as an essential pre-processing step before applying a different change detection approach (Celik, 2009; Dharani & Sreenivasulu, 2021; P. Gong, 1993; Koutsias et al., 2009; Rodarmel & Shan, 2002)

Change Vector Analysis (CVA)

Change Vector Analysis (CVA) is another popular method for unsupervised change detection. This approach uses pixel-wise radiometric comparison and represents change as a vector in multi-dimensional spectral space (Du et al., 2020; Polykretis et al., 2020). After image pre-processing, including geometric registration and radiometric correction, CVA is applied to two images from different dates (Du et al., 2020). A change vector includes the angle of change

(vector direction) and the magnitude of the change between two dates (He et al., 2011; Tewkesbury et al., 2015).

There are multiple examples of CVA applications for forest disturbance monitoring. Vorovencii (2017) used a CVA approach to assess desertification risk in south-west Romania between 1984 and 2011 using Landsat imagery. The highest overall accuracy and KIA achieved within their study were 87.67% and 0.8150, respectively. Marinelli et al. (2022) applied CVA to Light Detection and Ranging (LiDAR) data of a coniferous forest stand on Vancouver Island. This approach enabled multiple types of forest disturbances, including growth, regeneration, and harvest, to be identified between 2004 and 2008 with 98.07% overall accuracy. Xiao et al. (2021) successfully used a direction-dominated CVA approach to analyze forest changes on multiple Sentinel-2 image pairs in Huangshan and Chengde, China. These studies indicate that CVA can successfully extract forest disturbance information using a variety of sensors, data types and algorithm variations.

Change Point Detection and Analysis (CPD)

Another method for determining change between multiple remotely sensed images is Change Point Detection and Analysis (CPD). The goal of early work was to find shifts in the mean of independent and identically distributed (iid) Gaussian variables for the purposes of quality control in industrial settings (Truong et al., 2018). CPD seeks to identify changes in the model of a signal or time series dataset (Truong et al., 2018). Changes can occur within these datasets due to external events acting on the system or internal dynamics (Aminikhanghahi & Cook, 2017). A change or breakpoint represents a change in the pixel's distribution through time (Hermosilla et al., 2015a). Metrics derived from the change event, as well as the periods pre- and post-change can be used to attribute the type of change event. For example, change magnitude

and duration could be key properties in determining the type of change that has occurred (Hermosilla et al., 2015b).

CPD analyses can be an asset for forest disturbance monitoring. One such example is the Bayesian Estimator of Abrupt change, Seasonal change, and Trend known as "BEAST" (Zhao et al., 2019). This method is an ensemble time series decomposition approach for assessing non-linear ecosystem change across multiple timescales for moderate scale resolution sensors. The authors tested this approach on multiple cases studies including 495 Landsat images in the Southern United States, which resulted in an 82.3% producer's accuracy and 73.2% user's accuracy for disturbance and recovery events. When analyzed separately, disturbance events were detected with much greater accuracy. Another study included part of the Shawnee State Forest in Ohio, United States, which experienced an ice storm and fire within the study period (2002-2014). Using the BEAST approach on MODIS data resulted in a R^2 of 0.91 in a leave-out cross validation accuracy assessment (Zhao et al., 2019). Other researchers are continuing to use this methodology, like Higginbottom and Symeonakis (2020), who applied the BEAST approach to a Global Inventory Monitoring and Modelling System third generation NDVI dataset from the AVHRR sensor between 1982 and 2015 over the African Savannah. Other researchers continue to investigate forest change, like Ukoha et al. (2021), who used Bayesian CPD between 2000 and 2020 in Akure Forest Reserve, Ondo State in Southwestern, Nigeria. Their overall accuracy ranged between 0.77 and 0.84. These examples indicate the applicability of CPD for forest monitoring applications.

Post-Classification Change Detection (PCCD)

PCCD, or delta classification, is the final change detection method that will be outlined. This method detects change through comparing previous classified images from multiple dates (W. Zhou et al., 2008). It is important that the classification scheme for all target images is

compatible (Q. Zhang et al., 2002). PCCD can be conceptualized as the overlaying of spatially coincident thematic maps from multiple dates (Tewkesbury et al., 2015). This popular approach for change detection is often applied to moderate and coarse resolution remotely sensed imagery. Researchers continue to propose methods for applying this approach to higher resolution imagery (Radke et al., 2005; W. Zhou et al., 2008). Not only are areas of change detected, but this method also provides "from-to" change information (Minu & Shetty, 2015; W. Zhou et al., 2008).

Using a post-classification approach reduces the impacts of using images from multiple dates and sensors. Each image is classified into the relevant land cover classes based on its own characteristics and data distribution. Image normalization is not required when using this approach, as any data defects can be contained within a singular classification (Jakubauskas et al., 1990; Tewkesbury et al., 2015). Image level classification can often account for issues that may arise when using imagery acquired in different seasons or with different phenology, soil moisture and atmospheric conditions (Jakubauskas et al., 1990; Peiman, 2011; Q. Zhang et al., 2002; W. Zhou et al., 2008). As a result, PCCD often requires fewer pre-processing steps than other methods like raster arithmetic and PCA (Afify, 2011). Q. Zhang et al. (2002) also notes that post-classification can be less impacted by registration errors and allows analysts to use a priori information to develop classification schemes.

Despite the many benefits of PCCD, there are multiple limitations of this change detection approach. Post-classification approaches are most heavily criticized for their ability to compound errors from the original classified images (Singh, 1989; Tewkesbury et al., 2015; Q. Zhang et al., 2002). PCCD can be preceded by either supervised or unsupervised classification algorithms, which have a variety of characteristics impacting their accuracy (Hussain et al., 2013). Additionally, researchers note that this method may require more steps prior to

implementation, as it requires the map production and classification of multiple different images. These maps may be generated from different data sources and algorithms, which may result in inconsistencies or inaccuracies in the final change detection product (Tewkesbury et al., 2015). For strong results, the maps must have near perfect spatial alignment, the same legend and thematic meaning of the classes, and high classification accuracy (Colditz et al., 2012). Spatial misregistration issues are particularly impactful at the edges or boundaries of classified land covers (W. Zhou et al., 2008). W. Zhou et al. (2008) notes that although post-processing refinements can be made, commission errors caused by spatial inaccuracies were still significant for multiple classes within their study. Despite these limitations, PCCD is still used due to its ease of use and ability to compare products from different satellites (Colditz et al., 2012).

Despite its simplicity, PCCD is still a common approach for change detection with authors like Giri et al. (2008), Halder et al. (2021), and Mansaray et al. (2016). Giri et al. (2008) examined the impact of tsunami-affected mangrove forests in southeastern Asia between 1975 and 2005. Landsat time-series data was classified using an Iterative Self-Organizing Data Analysis (ISODATA) clustering approach before PCCD was conducted. Halder et al. (2021) conducted similar work, examining disturbances to the Sundarbans mangrove forest of India between 1990 and 2019. Maximum likelihood classification (MLC) was conducted on Landsat imagery prior to post-classification comparison. The overall accuracies of the classified images ranged between 72% and 85.9%. PCCD continues to be prominent in more general LULC as well, like in Mansaray et al. (2016), who assessed the deforestation and urban expansion in Freetown, Sierra Leone from 1986 to 2015. Their study had an overall accuracy of 87.3% for the classified data for which they had reference data. These studies indicate the continued applicability of PCCD for forest disturbance studies.

2.6 Road Extraction Approaches

Researchers have proposed many road extraction algorithms. These include road tracking, mathematical morphology, image processing including active contour (snake), and level set, dynamic programming, as well as region or classification-based methods (Mena, 2003; Shi et al., 2014; W. Wang et al., 2016). These approaches are also frequently used in combination. Image analysts determine which type of road extraction method they will use based on a variety of application properties: image type (spectral, spatial, and temporal resolution), available training and validation data, input features, and required outputs. There is no method that will produce quality results in all road extraction applications. Often, researchers use their expert and application specific knowledge as opposed to the collective knowledge of the field due to the breadth of approaches (Khatami et al., 2016; Soni et al., 2018). To evaluate the accuracy of road extraction methods, they must be applied to the same dataset to account for the wide range of contextual factors (Cardim et al., 2018). Each of these approaches will be discussed below. To give a broad overview of the various approaches, the algorithms presented are not limited to medium resolution imagery but span a variety of imagery types.

Road Tracking Based Road Extraction

Road finding and tracking approaches are very prevalent in the early literature related to road extraction. These methods generate road networks from a set of seed points or points indicating the start of road segments. These seed points can either be selected manually or automatically (Mena, 2003). Three major road finding algorithm classes are present in the literature: edge linkers, correlation trackers and region-based followers. Edge linkers use edge operators to determine the magnitude and angle properties in the direction of the expected road away from the seed point. Then, the edges are linked based on criteria defined by the user. Correlation trackers assume that there is some pattern or texture intrinsic to the road surface; this

method is often used in combination with edge linkers. For example, multiple lane roads will have a bright line down them and there are usually tire wear mark patterns within a lane. Region based methods assume that there is a consistent intensity in the road region in comparison to the background; this method is often used in combination with correlation trackers (Barzohar et al., 1996; McKeown & Denlinger, 1988).

The work of Nevatia and Babu (1980) was pivotal to road tracking approaches. Although their work was intended for general linear feature extraction, there are many applications within road extraction research. Their approach derived edge magnitude and direction through convolution of the image using edge masks. The edges were thinned (only local maxima were retained) then linked with other edges based on proximity and orientation. The edges were linked by approximating the location of piecewise linear segments. McKeown and Denlinger (1988) built on this approach by developing a multilevel approach for road extraction that enables multiple low level processors' outputs to be evaluated by higher level analysis methods. This approach addressed the issue that most early road trackers struggled to recover after an extraction failure. This failure was often compounded as most methods could not automatically detect when there was a runtime failure or deviation from the road surface. McKeown and Denlinger (1988) added additional sources of knowledge to combat single source failures. Vosselman and De Knecht (1995) introduced template matching to road detection methodologies. An operator initializes the algorithm by generating two seed points that define a short road segment. Cross section profiles are generated between these two points, with the average becoming the template for the profile matching process. This profile is used to predict the next road segment adjacent to the current road segment. The properties are then matched, and the road parameters are updated. This process continues until a stopping threshold is reached. This approach uses least squares

profile matching with a Kalman filter. Rejected profile matches often indicate road intersections, which suggests width properties could enhance this type of approach. Since these early works in road tracking and finding, researchers have extended the methods to include tracking frameworks for automatic extraction under the supervision of a human operator to address tracking failures (J. Zhou et al., 2006), online learning methods to generate optimal templates (X. Zhou & Mahmassani, 2007), and multiple cooperative tracking windows (J. Cheng et al., 2012). Mathematical morphology was another method that emerged early in the road extraction discourse.

Mathematical Morphology

Mathematical morphology is an approach used in image processing, pattern recognition, computer vision and other applications. In the 1980s, researchers first applied mathematical morphology to road extraction problems (Wang et al., 2016). This approach is based on set theory developed by Matheron (1974) and Serra (1982)
(Mena, 2003; Soni et al., 2018). Set theory is often used in image processing for geometrical shapes. Common mathematical morphology operations include union, intersection, dilation, erosion, and other associated operations (Mena, 2003). Mathematical morphology has multiple benefits over road finding methods including less computational cost, lack of reliance on seed points for initialization, and decreased complexity (Soni et al., 2018). Despite these benefits, this set of methods often depends on a fixed shape structure element, which can prevent the algorithms from being flexible enough to manage road curves and rectilinear structures (Kahraman et al., 2018; Shi et al., 2014). Researchers have continued to propose methods to address this limitation, however, have struggled to find a balanced solution (Valero et al., 2010).

One of the early examples of mathematical morphology methods is the work by C. Zhang et al. (1999). This study included the segmentation of 1 m resolution imagery. Trivial opening

was then applied to reduce noise in the image. This method allows many of the objects with similar spectral characteristics to the road surface to be removed based on their geometric qualities. Researchers continue to apply these methods to road extraction problems including Valero et al. (2010), Shi et al. (2014) and Bakhtiari et al. (2017). Valero et al. (2010) address the fixed structuring element limitation of mathematical morphology. For each pixel, a morphological profile is generated from a granulometric analysis using path openings and closings. Linear pixel information is extracted from the profiles. The road network is then detected through classifying the extracted vectors. This approach outperforms other methods based on rotating rectangular elements, as the morphological operators are flexible enough to include rectilinear and moderately curved elements. Shi et al. (2014) built on this approach by starting their analysis using path openings and closings to generate morphological profiles. These profiles are used in a spectral-spatial binary support vector machine (SVM) classification (road/non-road). Local Geary's C is used to measure the local homogeneity of the pixel values. The results of the classification and homogeneity analysis are combined based on fusion rules. In contrast to Shi et al. (2014) and C. Zhang et al. (1999), Bakhtiari et al. (2017) use mathematical morphology as a secondary method. Their approach begins with a rough outline of the road surface being generated through the Canny operator. Adjacent segments are merged using the Full Lambda merging method. The image is then classified using the SVM classifier. Then, the morphological operators open and erode to remove non-road pixels/noise and fill gaps within the road. The inclusion of the operators served to increase the classification accuracy. Although researchers continue to address the difficulties associated with structuring elements, it may still be difficult to get high accuracy with only mathematical morphology approaches (W. Wang et

al., 2016). As a result, multimethod approaches like Bakhtiari et al. (2017) continue to be prevalent.

Image Processing Based Road Extraction

The active contour model and geometric active contour model are two other road extraction methods. These are better known as the snake and level set methods, respectively (W. Wang et al., 2016). Active contour methods are based on the optimization of energy functions and curves (Soni et al., 2018). The snake method was first proposed by Kass and Witkin (1988). A snake is an energy-minimizing spline algorithm that is guided by external constraints and image forces that guide it towards lines and edges. The curve or contour deforms from an arbitrary initial position towards the target feature and creates a best fit through minimizing an energy function (Amini, 2009; L. D. Cohen, 1991). Due to the visual appearance of slithering by the contour, active contour models are often referred to as "snakes" (Kass & Witkin, 1988).

Snakes are less sensitive to initialization points, noise, and are not as constrained by low-level sequential decision making compared to other early extraction methods (Kass & Witkin, 1988; Nakaguro et al., 2011; Zafiropoulos & Schenk, 1998). Despite these benefits, the lack of topological flexibility in the classical snake model is a major limitation. To address limitations, researchers including L. D. Cohen (1991), Jeon et al. (2002), Nakaguro et al. (2011), Saati and Amini (2017), and Zafiropoulos and Schenk (1998) have built on the work of Kass and Witkin (1988).

L. D. Cohen (1991) adds another force to increase the dynamic nature of snakes. Zafiropoulos and Schenk (1998) propose two methods for incorporating the colour component of remotely sensed imagery into the snake method. Nakaguro et al. (2011) address some of the topological limitations of snakes. Building on the works of L. D. Cohen, (1991) (balloon term), Xu and Prince (1998) (gradient vector flow), and Rochery et al. (2006) (quadratic active

contours), Nakaguro et al. (2011) use multiple quadratic snakes. These snakes can split, merge, and disappear under specific conditions, enabling the creation of topological flexible snakes without a reduction in simplicity and efficiency. Jeon et al. (2002) and Saati and Amini (2017) incorporate snakes into their hybrid approaches.

The level set method was proposed by Osher and Sethian (1988). This method uses an implicit energy function within a high number of dimensions in a fixed rectangular grid to model a region's contours. These approaches can handle a complex topology (splitting, merging, or developing holes) and incorporate a priori domain knowledge (Cai et al., 2019; Chan & Vese, 2001; Keaton & Brokish, 2002; Y. Liu et al., 2014).

Researchers have developed a variety of level set methods to address application specific limitations of the classical level set model. These include the Chan-Vese (C-V), local binary fitting (LBF), and global and local regional active contour (GARAC) approaches. Chan and Vese (2001) address the inability to identify weak object boundaries through basing the stopping criteria of the curve on the Mumford-Shah segmentation approach. The researchers extend this approach in Vese and Chan (2002) to represent multiple regions as piecewise constants. C. Li et al. (2007) developed the LBF method to segment images with inhomogeneous regions. H. Wang et al. (2016) developed the GARAC method, which implements a two-stage segmentation approach using local and global active contours. Researchers continue to improve on these approaches and apply them to road-based applications including Abdollahi et al. (2019), Keaton and Brokish (2002) and Rajeswari et al. (2011)

Region Based Road Extraction: Supervised Classification

Region based methods extract the initial road network based on image classification. In remote sensing, image classification refers to the division of an image into non-overlapping regions. Contiguous pixels often share similar attributes (intensity, colour, shape, texture, etc.).

Commonly, image classification algorithms are applied to the spectral data of an image, grouping the image into spectrally homogenous units. These regions usually correspond to an object or type of land cover (Balaji, 2014). In traditional remote sensing classification, it is assumed that each pixel is pure (only contains one land cover type) and is therefore labeled as a single land use. However, a pixel may contain more than one LULC. As a result, fuzzy classification and spectral mixture analyses techniques were developed to address mixed pixels and uncertain class boundaries (R. Li & Cao, 2018).

Although classification provides strong results for many applications, there are cases outside of their limits. Classification works under the assumption that there is a consistent association between the spectral response and earth surface features and phenomena. In reality, there is a complex relationship between scene complexity, scale, sensor, and earth surface features that prevent all sources of spatial variation from being modeled or predicted (Belgiu & Drăgu, 2016).

Classification methods can be divided into two main methods: supervised and unsupervised classification (Soni et al., 2018). Often, supervised classifiers provide more accurate results; however, supervised methods have increased complexity and time burdens (Zaidi et al., 2017). Each of these methods will be discussed in detail below. Hybrid classification, a combination of supervised and unsupervised methods, can also be utilized (K. Liu et al., 2011).

Supervised classifiers rely on an image analyst to collect representative training samples (R. Li & Cao, 2018). LULC classes are defined in advance by an analyst and the spectral properties are learned from training samples (Ghosh et al., 2011; X. Gong et al., 2019; Keuchel et al., 2003). Training samples are pixels that are identified with the desired spectral range for

each class. The training samples are used to train the classifier. The target pixels within the image are labeled according to the classifier's decision rules (R. Li & Cao, 2018; Roy et al., 2010; Soni et al., 2018). Supervised classifiers are highly dependent on quality of the training data provided by the analyst (Belgiu & Drăgu, 2016; Büschenfeld & Ostermann, 2012; X. Gong et al., 2019; Roy et al., 2010).

Efficient and effective supervised classifiers will possess the following properties: reduce the Hughes phenomenon (number of variables is larger than the number of samples), manage non-linear variables as well as imbalanced training samples, noisy samples, and unlabeled data, and reduce computation time (Belgiu & Drăgu, 2016). Supervised classifiers are frequently used as they are often more robust than their unsupervised counterparts (Belgiu & Drăgu, 2016). Many supervised classification algorithms have been developed, including but not limited to the following: SVM, MLC and Random Forest (RF). Each of these algorithms will be briefly described below.

SVM is a non-parametric supervised classification method developed in the framework of statistical learning theory by Vapnik and Chervonenkis (1971) (later extended by Vapnik (1999). This algorithm does not model or assume the distribution of the dataset; instead, it separates the different classes by searching for optimal boundaries between classes (Belgiu & Drăgu, 2016; Keuchel et al., 2003; Pal, 2005). The optimal boundary is achieved through repetition, data conversion, and mapping multi-dimensional data into higher dimensional space used for classification. The original SVM algorithm was designed as a binary classifier, but researchers have designed strategies to adopt this method to multiclass studies (Bakhtiari et al., 2017). SVM is a strong algorithm for heterogeneous datasets where pure training pixels are challenging to acquire (Amin & Shivakumar, 2017). This approach also has strong accuracy on

high dimensionality data, even when trained with a small number of samples (Han et al., 2018; Jozdani et al., 2019; Keuchel et al., 2003). The SVM classifier has been applied to road extraction by multiple researchers (Song & Civco, 2004; Ziems et al., 2012) as well as other LULC applications (Han et al., 2018; Otukei & Blaschke, 2010; Sanchez-Hernandez et al., 2007).

MLC, based on Bayes classification, is the most common statistical approach for classifying LULC and other remote sensing images (Soni et al., 2018; H. Wang et al., 2016; Zaidi et al., 2017). This classification algorithm is a parametric method, meaning that class data is presumed to be distributed according to a defined probability model (usually Gaussian distribution) (Ali et al., 2018; Keuchel et al., 2003). Each data point is assigned a class for which its values are most likely (highest a posteriori probability) (Keuchel et al., 2003). MLC considers within class variance-covariance distributions and is the strongest known parametric classifier (Otukei & Blaschke, 2010). Although this method is very popular, there are multiple drawbacks to this approach. Using a unimodal distribution for each class is highly limiting, as multimodal distributions are more commonly encountered in real datasets. MLC is also based only on spectral information as opposed to incorporating other attributes, which can lead to noisy results (Ali et al., 2018; Belgiu & Drăgu, 2016; Keuchel et al., 2003). MLC has been used in many LULC classification applications (Ali et al., 2018; Mosammam et al., 2017; Otukei & Blaschke, 2010).

RF is the final supervised classifier discussed here. RF is a machine learning ensemble classifier, which are approaches that are based on several supervised classifiers that are trained using bagging, boosting, or a variation of these approaches (Belgiu & Drăgu, 2016). Ensemble classifiers have been reported to have increased accuracy in comparison to the individual

classifiers making up their ensemble (Pal, 2005). Developed by Breiman (2001), the RF classifier uses a set of Classification and Regression Trees (CARTs) to make predictions. RF addresses the limitation of overfitting using a single CART (Belgiu & Drăgu, 2016; Georganos & Kalogirou, 2022; Gibson et al., 2020; Loozen et al., 2020). The forest (multiple trees) grows to a user determined number of trees. The final classification is taken through calculating the arithmetic mean of all the class assignment probability produced by all trees within the forest. New unlabeled data is then evaluated against each of the previously generated trees in the ensemble. Each tree generates a class membership for the unlabeled data. Majority voting is used to select the final class (Belgiu & Drăgu, 2016; Gibson et al., 2020; Griffiths et al., 2014; Loozen et al., 2020; Pal, 2005).

RF and other ensemble algorithms improve accuracy as the error of a single classification is diminished through using majority voting on multiple classifications (Gibson et al., 2020). The RF classifier is also less sensitive to poor quality training samples and overfitting in comparison to other machine learning algorithms. Overfitting is avoided as RF randomly samples the predictor space (Loozen et al., 2020). As a non-parametric approach, RF can also accommodate non-normal datasets and non-linear relationships (Loozen et al., 2020; Thompson et al., 2015). Despite the many benefits of RF classifiers, it should be noted that classification accuracy decreases when the classifier is trained on different study sites. This approach is not highly transferrable (Belgiu & Drăgu, 2016). RF classifiers have been applied to LULC classification (Colditz et al., 2012; Haas & Ban, 2018) as well as mapping boreal forest habitats and tree canopy (Karlson et al., 2015; Räsänen et al., 2013).

Region Based Road Extraction: Unsupervised Classification

In contrast to supervised classifiers, unsupervised classifiers do not use any training data or prior knowledge from the analyst (Han et al., 2018; R. Li & Cao, 2018; Richards, 2013).

Instead, the remotely sensed image pixels are classified based on natural groupings or clustering of the spectral values (R. Li & Cao, 2018; Verma et al., 2017). These groups will have homogeneity within the classes defined for one or more variables or (Khan, 2012). Unsupervised classification is also known as clustering classification or point group analysis (Bo et al., 2009; Han et al., 2018). Although unsupervised classification methods have a lower standard of a priori knowledge, there are multiple limitations. The accuracy of these classifiers is generally lower than supervised classifiers (Hussain et al., 2013; H. Wang et al., 2016). The user may struggle to label classifier generated clusters. The number of clusters also influences the outcome of the classification result; therefore, the selection of an appropriate number of clusters is paramount (Hussain et al., 2013). The k-means, ISODATA, Jenks Optimization, and Mean Shift Classifier methods will be discussed below.

The k-means algorithm was first developed by MacQueen (1967), although the version developed by Lloyd (1982) is the most common current application (Khan, 2012). The algorithm is also referred to as migrating means or iterative optimization (Richards, 2013). K-means generates optimal clusters by minimizing the intraclass cluster variance. The added variances of each cluster should explain the total variance within the pixel values (Wattelez et al., 2022). This approach requires a priori knowledge of the number of clusters and an initial seeding value for the center of the cluster in the feature space. The k-means approach is highly sensitive to the selection of the seeding value, with its selection strongly impacting the cluster assignment (Khan, 2012). Inappropriate initialization parameters may result in empty clusters, increased number of iterations, and increased change of the means being stuck in local minimums. Researchers have developed multiple approaches for better estimating initial clusters, including Kaufman's initialization (Xia et al., 2017). The initial number of clusters is often chosen higher

than required with the understanding that spectrally similar clusters can be merged after the algorithm has been applied (Richards, 2013). ISODATA builds on k-means by introducing checks on the clusters formed (Richards, 2013; Verma et al., 2017). These checks relate to the number of pixels assigned to each cluster as well as the shape of the cluster within the spectral space.

Advantages of k-means and ISODATA are that they are conceptually simple, easy to implement, and effective (Leichtle et al., 2017; Xia et al., 2017). In addition to the sensitivity to the initialization values, one disadvantage of these algorithms is that for each iteration, the pixel must be compared to the cluster centers. As a result, k-means and ISODATA are not computationally efficient, particularly for larger datasets (Richards, 2013). Despite these limitations, k-means and ISODATA are still heavily used within the literature for LULC applications (Markogianni & Dimitriou, 2016; Naikoo et al., 2020; Patra et al., 2018)

Another clustering method that can be used for unsupervised classification is Jenks Optimization. This method was developed by Jenks (1967), as described in Jenks and Caspall (1971). Jenks built upon the work of Fisher (1958). Jenks Optimization is most used in thematic map classification but also has applications in geographic research (Fraile et al., 2016). This method seeks to classify data based on the statistically significant breaks or naturally occurring gaps within the dataset. This method aims to reduce the average deviation of each value from the mean of the class and increase the deviation of each class from the mean of other classes (Gui et al., 2022). Thus, Jenks Optimization reduces intraclass variation and maximizes interclass variation (Amirruddin et al., 2020; de Smith et al., 2018; Fraile et al., 2016; Hou et al., 2022). Classification results are based on iterative computations of goodness-of-fit tests (goodness of variance fit, GVF) to indicate class performance (Fraile et al., 2016). Values are moved from the

class with the largest deviation from the class mean until the intraclass deviation reaches the lowest value (Jiang, 2013; Zongfan et al., 2022).

The Jenks Optimization can be applied as follows. The analyst selects the number of classes, k, for the one-dimensional attribute, x, to the classified. A set of k-1 random values in the range of [min{x}, max{x}] are generated. These values are the initial class boundaries. The mean of each initial class and GVF are calculated. Each class value is then systematically reassigned to adjacent classes to improve the GVF. This process iterates until the GVF reaches a certain threshold (optimal interclass and intraclass variance). The process can be repeated and the GVF values can be compared (de Smith et al., 2018; do Carvalhal Monteiro et al., 2018). GVF values range from 0 to 1, where 1 is optimal and 0 indicates minimal performance (Amirruddin et al., 2020). GVF can be obtained by dividing the difference between the square deviation of array mean (SDAM) and the square deviation of class mean (SDCM) by SDAM. These statistics are presented in Equation 2, Equation 3, and Equation 4.

$$GVF = \frac{SDAM - SDCM}{SDAM} \text{ [2]}$$

$$SDAM = \sum (x_i - X)^2 \text{[3]}$$

$$SDCM = \sum \sum (x_i - Z)^2 \text{[4]}$$

Where X is the dataset mean and Z is the average class value (Fraile et al., 2016; Gui et al., 2022). Generally, classification results where the GVF is greater than 0.7 are deemed acceptable (Hou et al., 2022). Although this approach can only be applied to one dimensional dataset, there are multiple advantages to its use. The results are replicable, efficient, conceptually simple and there are no modifiable parameters requiring expert a priori information for generating replicable results (Khan, 2012). In contrast to k-means and ISODATA, Jenks Optimization calculates the

maximum distance between cluster centers in addition to minimum distances between values and their assigned cluster centers. This approach can result in more natural clusters (Khan, 2012).

The final unsupervised classification method described here is mean shift, a nonparametric clustering approach for estimating density gradients (Bo et al., 2009). Mean shift was developed in Fukunaga and Hostetler (1975) for pattern recognition but was built upon by Y. Cheng (1995) for cluster analysis on imagery. There have been continual improvements on this approach since the 1970s (Vlachopoulos et al., 2020). Mean shift uses a generalized multivariate kernel approach, Parzen windows. This approach estimates the local maxima of the modes of underlying distribution of the input feature space. This method converges with an adaptive gradient ascent method (Vlachopoulos et al., 2020). Unlike the other approaches listed, mean shift does not require the number of clusters to be determined a priori. Instead, the number of clusters is determined automatically in the algorithm (Bo et al., 2009; Vlachopoulos et al., 2020). The mean shift clustering approach has been applied to many LULC applications (Friedman et al., 2013; Vlachopoulos et al., 2020).

2.7 Conclusion

In this chapter, an overview of literature related to forest road extraction from Landsat imagery was presented to support the development of a semi-automatic algorithm. The proposed approach for forest road extraction and year of construction assignment incorporates knowledge from Ontario Road legislation and construction practices, road conceptualization in remotely sensed imagery, sensor characteristics, and best practices in automated road detection and change detection research. Chapter 3 discusses the preparation and optimization of Landsat imagery BAP Composites for the input into the road extraction and change detection algorithm.

2.8 References

Abdollahi, A., Pradhan, B., & Shukla, N. (2019). Extraction of road features from UAV using a novel level set segmentation approach. *International Journal of Urban Sciences, 23*(3), 391–405. https://doi.org/10.1080/12265934.2019.1596040

Affek, A. N., Zachwatowicz, M., Sosnowska, A., Gerlée, A., & Kiszka, K. (2017). Impacts of modern mechanised skidding on the natural and cultural heritage of the Polish Carpathian Mountains. *Forest Ecology and Management, 405*, 391–403. https://doi.org/10.1016/j.foreco.2017.09.047

Afify, H. A. (2011). Evaluation of change detection techniques for monitoring land-cover changes: A case study in new Burg El-Arab area. *Alexandria Engineering Journal, 50*, 187–195. https://doi.org/10.1016/j.aej.2011.06.001

Ahmed, O. S., Wulder, M. A., White, J. C., Hermosilla, T., Coops, N. C., & Franklin, S. E. (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. https://doi.org/10.1080/2150704X.2016.1233371

Ali, M. Z., Qazi, W., & Aslam, N. (2018). A comparative study of ALOS-2 PALSAR and Landsat-8 imagery for land cover classification using maximum likelihood classifier. *Egyptian Journal of Remote Sensing and Space Science, 21*, S29–S35. https://doi.org/10.1016/j.ejrs.2018.03.003

Amin, K., & Shivakumar, B. R. (2017, November 21). Change detection using image differencing: A study over area surrounding Kumta, India. *Proceedings of the 2017 2nd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2017*. https://doi.org/10.1109/ICECCT.2017.8117851

Amini, J. (2009). Road extraction from satellite images using a fuzzy-snake model. *Cartographic Journal*, *46*(2), 164–172. https://doi.org/10.1179/000870409X459923

Aminikhanghahi, S., & Cook, D. J. (2017). A survey of methods for time series change point detection. *Knowledge and Information Systems*, *51*, 339–367. https://doi.org/10.1007/s10115-016-0987-z

Amirruddin, A. D., Muharam, F. M., Ismail, M. H., Ismail, M. F., Tan, N. P., & Karam, D. S. (2020). Hyperspectral remote sensing for assessment of chlorophyll sufficiency levels in mature oil palm (Elaeis guineensis) based on frond numbers: Analysis of decision tree and random forest. *Computers and Electronics in Agriculture*, *169*. https://doi.org/10.1016/j.compag.2020.105221

Bae, Y., Lee, W. H., Choi, Y., Jeon, Y. W., & Ra, J. B. (2015). Automatic road extraction from remote sensing images based on a normalized second derivative map. *IEEE Geoscience and Remote Sensing Letters*, *12*(9), 1858–1862. https://doi.org/10.1109/LGRS.2015.2431268

Bakhtiari, H. R. R., Abdollahi, A., & Rezaeian, H. (2017). Semi automatic road extraction from digital images. *Egyptian Journal of Remote Sensing and Space Science*, *20*(1), 117–123. https://doi.org/10.1016/j.ejrs.2017.03.001

Balaji, T. (2014). Robust and realistic classification of massive gray level thresholding in remote sensing images. *International Journal of Computer Science and Engineering*.

Barzohar, M., Cooper, D. B., & Member, S. (1996). Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*(7). https://doi.org/10.1109/34.506793

Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Besbes, O., & Benazza-Benyahia, A. (2014). Joint road network extraction from a set of high resolution satellite images. *2014 22nd European Signal Processing Conference (EUSIPCO)*.

Bilyk, A., Pulkki, R., Shahi, C., & Larocque, G. R. (2021). Development of the Ontario Forest Resources Inventory: A historical review. *Canadian Journal of Forest Research*, *51*, 198–209. https://doi.org/10.1139/cjfr-2020-0234

Bo, S., Ding, L., Li, H., Di, F., & Zhu, C. (2009). Mean shift-based clustering analysis of multispectral remote sensing imagery. *International Journal of Remote Sensing*, *30*(4), 817–827. https://doi.org/10.1080/01431160802395193

Bocking, E., Cooper, D. J., & Price, J. (2017). Using tree ring analysis to determine impacts of a road on a boreal peatland. *Forest Ecology and Management*, *404*, 24–30. https://doi.org/10.1016/j.foreco.2017.08.007

Bodart, C., Eva, H., Beuchle, R., Raši, R., Simonetti, D., Stibig, H. J., Brink, A., Lindquist, E., & Achard, F. (2011). Pre-processing of a sample of multi-scene and multi-date Landsat imagery used to monitor forest cover changes over the tropics. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*, 555–563. https://doi.org/10.1016/j.isprsjprs.2011.03.003

Borrelli, P., Modugno, S., Panagos, P., Marchetti, M., Schütt, B., & Montanarella, L. (2014). Detection of harvested forest areas in Italy using Landsat imagery. *Applied Geography*, *48*, 102–111. https://doi.org/10.1016/j.apgeog.2014.01.005

Braaten, J. D., Cohen, W. B., & Yang, Z. (2015). Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems. *Remote Sensing of Environment*, *169*, 128–138. https://doi.org/10.1016/j.rse.2015.08.006

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Bullock, E. L., Woodcock, C. E., & Olofsson, P. (2020). Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment, 238*. https://doi.org/10.1016/j.rse.2018.11.011

Büschenfeld, T., & Ostermann, J. (2012). Automatic refinement of training data for classification of satellite imagery. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *I–7*, 117–122.

Cai, Q., Liu, H., Qian, Y., Zhou, S., Duan, X., & Yang, Y. H. (2019). Saliency-guided level set model for automatic object segmentation. *Pattern Recognition*, *93*, 147–163. https://doi.org/10.1016/j.patcog.2019.04.019

Cardim, G. P., da Silva, E. A., Dias, M. A., Bravo, I., & Gardel, A. (2018). Statistical evaluation and analysis of road extraction methodologies using a unique dataset from remote sensing. *Remote Sensing*, *10*(620). https://doi.org/10.3390/rs10040620

Celik, T. (2009). Unsupervised change detection in satellite images using principal component analysis and k-means clustering. *IEEE Geoscience and Remote Sensing Letters*, *6*(4), 772–776. https://doi.org/10.1109/LGRS.2009.2025059

Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, *10*(2). https://doi.org/10.1109/83.902291

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, *113*, 893–903. https://doi.org/10.1016/j.rse.2009.01.007

Chen, J., Feng, H., Pan, K., Xu, Z., & Li, Q. (2014). An optimization method for registration and mosaicking of remote sensing images. *Optik*, *125*, 697–703. https://doi.org/10.1016/j.ijleo.2013.07.034

Chen, Z., Jefferies, B., Adlakha, P., Salehi, B., & Power, D. (2014). Monitoring linear disturbance footprint based on dense time series Landsat imagery. *Canadian Journal of Remote Sensing*, *40*, 348–361. https://doi.org/10.1080/07038992.2014.987375

Cheng, G., Wang, Y., Xu, S., Wang, H., Xiang, S., & Pan, C. (2017). Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(6), 3322–3337. https://doi.org/10.1109/TGRS.2017.2669341

Cheng, J., Ding, W., Ku, X., & Sun, J. (2012). Road extraction from high-resolution SAR images via automatic local detecting and human-guided global tracking. *International Journal of Antennas and Propagation*, *2012*. https://doi.org/10.1155/2012/989823

Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *17*(8).

Cohen, L. D. (1991). Note: On active contour models and balloons. *CVGIP: Image Understanding*, *53*(2), 211–218.

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync - Tools for calibration and validation. *Remote Sensing of Environment*, *114*, 2911–2924. https://doi.org/10.1016/j.rse.2010.07.010

Colditz, R. R., Acosta-Velázquez, J., Díaz Gallegos, J. R., De Vázquez Lule, A., Rodríguez-Zúñiga, M. T., Maeda, P., Cruz López, M. I., & Ressl, R. (2012). Potential effects in multi-resolution post-classification change detection. *International Journal of Remote Sensing*, *33*(20), 6426–6445. https://doi.org/10.1080/01431161.2012.688148

Crown Forest Sustainability Act, 1994, S.O. 1994, c. 25. Retrieved April 16, 2023, from https://www.ontario.ca/laws/statute/94c25

Dadon, A., Mandelmilch, M., Ben-Dor, E., & Sheffer, E. (2019). Sequential PCA-based classification of Mediterranean forest plants using airborne hyperspectral remote sensing. *Remote Sensing*, *11*(23). https://doi.org/10.3390/rs11232800

Dal Poz, A. P., Gallis, R. A. B., Da Silva, J. F. C., & Martins, É. F. O. (2012). Object-space road extraction in rural areas using stereoscopic aerial images. *IEEE Geoscience and Remote Sensing Letters*, *9*(4), 654–658. https://doi.org/10.1109/LGRS.2011.2177438

de Smith, M. J., Goodchild, M. F., & Longley, p. A. (2018). *Geospatial analysis: A comprehensive guide to principles, techniques and software tools* (6th Edition). Winchelsea Press.

Deng, J. S., Wang, K., Deng, Y. H., & Qi, G. J. (2008). PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, *29*(16), 4823–4838. https://doi.org/10.1080/01431160801950162

Dharani, M., & Sreenivasulu, G. (2021). Land use and land cover change detection by using principal component analysis and morphological operations in remote sensing applications. *International Journal of Computers and Applications*, *43*(5), 462–471. https://doi.org/10.1080/1206212X.2019.1578068

do Carvalhal Monteiro, R. L., Pereira, V., & Costa, H. G. (2018). A multicriteria approach to the human development index classification. *Social Indicators Research*, *136*(2), 417–438. https://doi.org/10.1007/s11205-017-1556-x

Du, P., Wang, X., Chen, D., Liu, S., Lin, C., & Meng, Y. (2020). An improved change detection approach using tri-temporal logic-verified change vector analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, *161*, 278–293. https://doi.org/10.1016/j.isprsjprs.2020.01.026

Earth Resources Observation and Science Centre. (2018). *USGS EROS Archive - Landsat Archives - Landsat 1-5 Multispectral Scanner (MSS) Level-1 Data Products*. https://www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-archives-landsat-4-5-thematic-mapper-tm-level-1-data

Fisher, W. D. (1958). On grouping for maximum homogeneity. *Journal of the American Statistical Association, 53*(284), 789–798. https://doi.org/10.1080/01621459.1958.10501479

Fraile, A., Larrodé, E., Alberto Magreñán, & Sicilia, J. A. (2016). Decision model for siting transport and logistic facilities in urban environments: A methodological approach. *Journal of Computational and Applied Mathematics*, *291*, 478–487. https://doi.org/10.1016/j.cam.2014.12.012

Friedman, L., Netanyahu, N. S., & Shoshany, M. (2013). Mean shift-based clustering of remotely sensed data with agricultural and land-cover applications. *International Journal of Remote Sensing*, *34*(17), 6037–6053. https://doi.org/10.1080/01431161.2013.793866

Fukunaga, K., & Hostetler, L. D. (1975). The estimation of the gradient of a density function, with applications in pattern-recognition. *IERE Transactions on Information Theory*, *1*.

Fung, T. (1992). Land use and land cover change detection with Landsat MSS and SPOT HRV data in Hong Kong. *Geocarto International*, *7*(3), 33–40. https://doi.org/10.1080/10106049209354378

Fung, T., & LeDrew, E. (1987). Application of principal components analysis to change detection. *Photogrammetric Engineering and Remote Sensing*, *53*(12), 1649–1658.

Geman, D., & Jedynak, B. (1996). An active testing model for tracking roads in satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*(1). https://doi.org/10.1109/34.476006

Georganos, S., & Kalogirou, S. (2022). A forest of forests: A spatially weighted and computationally efficient formulation of geographical random forests. *ISPRS International Journal of Geo-Information*, *11*(9). https://doi.org/10.3390/ijgi11090471

Ghosh, A., Mishra, N. S., & Ghosh, S. (2011). Fuzzy clustering algorithms for unsupervised change detection in remote sensing images. *Information Sciences*, *181*, 699–715. https://doi.org/10.1016/j.ins.2010.10.016

Gibson, R., Danaher, T., Hehir, W., & Collins, L. (2020). A remote sensing approach to mapping fire severity in south-eastern Australia using Sentinel 2 and Random Forest. *Remote Sensing of Environment*, *240*. https://doi.org/10.1016/j.rse.2020.111702

Giri, C., Zhu, Z., Tieszen, L. L., Singh, A., Gillette, S., & Kelmelis, J. A. (2008). Mangrove forest distributions and dynamics (1975-2005) of the tsunami-affected region of Asia. *Journal of Biogeography*, *35*(3), 519–528. https://doi.org/10.1111/j.1365-2699.2007.01806.x

Gómez, C., White, J. C., Wulder, M. A., & Alejandro, P. (2015). Integrated object-based spatiotemporal characterization of forest change from an annual time series of Landsat image composites. *Canadian Journal of Remote Sensing*, *41*, 271–292. https://doi.org/10.1080/07038992.2015.1089162

Gong, P. (1993). Change detection using principal component analysis and fuzzy set theory. *Canadian Journal of Remote Sensing*, *19*(1), 22–29. https://doi.org/10.1080/07038992.1993.10855147

Gong, X., Shen, L., & Lu, T. (2019). Refining training samples using median absolute deviation for supervised classification of remote sensing images. *Journal of the Indian Society of Remote Sensing*, *47*(4), 647–659. https://doi.org/10.1007/s12524-018-0887-7

Goodbody, T. R. H., Tompalski, P., Coops, N. C., White, J. C., Wulder, M. A., & Sanelli, M. (2020). Uncovering spatial and ecological variability in gap size frequency distributions in the Canadian boreal forest. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-62878-z

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., & Hostert, P. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment*, *151*, 72–88. https://doi.org/10.1016/j.rse.2013.04.022

Griffiths, P., Van Der Linden, S., Kuemmerle, T., & Hostert, P. (2013). Erratum: A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(5), 2088–2101. https://doi.org/10.1109/JSTARS.2012.2228167

Gui, R., Song, W., Pu, X., Lu, Y., Liu, C., & Chen, L. (2022). A river channel extraction method based on a digital elevation model retrieved from satellite imagery. *Water*, *14*(15). https://doi.org/10.3390/w14152387

Haas, J., & Ban, Y. (2018). Urban land cover and ecosystem service changes based on Sentinel-2A MSI and Landsat TM data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(2), 485–497. https://doi.org/10.1109/JSTARS.2017.2786468

Halder, S., Samanta, K., Das, S., & Pathak, D. (2021). Monitoring the inter-decade spatial–temporal dynamics of the Sundarban mangrove forest of India from 1990 to 2019. *Regional Studies in Marine Science*, *44*. https://doi.org/10.1016/j.rsma.2021.101718

Han, X., Pan, J., & Devlin, A. T. (2018). Remote sensing study of wetlands in the Pearl River Delta during 1995–2015 with the support vector machine method. *Frontiers of Earth Science*, *12*(3), 521–531. https://doi.org/10.1007/s11707-017-0672-x

He, C., Wei, A., Shi, P., Zhang, Q., & Zhao, Y. (2011). Detecting land-use/land-cover change in rural-urban fringe areas using extended change-vector analysis. *International Journal of Applied Earth Observation and Geoinformation*, *13*, 572–585. https://doi.org/10.1016/j.jag.2011.03.002

Helder, D. L., Karki, S., Bhatt, R., Micijevic, E., Aaron, D., & Jasinski, B. (2012). Radiometric calibration of the Landsat MSS sensor series. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(6), 2380–2399. https://doi.org/10.1109/TGRS.2011.2171351

Heller, R. C. (1975). *Evaluation of ERTS-1 data for forest and rangeland survey*. https://www.fs.usda.gov/research/treesearch/29965

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015a). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sensing of Environment*, *158*, 220–234. https://doi.org/10.1016/j.rse.2014.11.005

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015b). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment*, *170*, 121–132. https://doi.org/10.1016/j.rse.2015.09.004

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Pickell, P. D., & Bolton, D. K. (2019). Impact of time on interpretations of forest fragmentation: Three-decades of fragmentation dynamics over Canada. *Remote Sensing of Environment*, *222*, 65–77. https://doi.org/10.1016/j.rse.2018.12.027

Higginbottom, T. P., & Symeonakis, E. (2020). Identifying ecosystem function shifts in Africa using breakpoint analysis of long-term NDVI and RUE data. *Remote Sensing*, *12*(11). https://doi.org/10.3390/rs12111894

Hou, C., Xie, Y., & Zhang, Z. (2022). An improved convolutional neural network based indoor localization by using Jenks natural breaks algorithm. *China Communications*, *19*(4), 291–301. https://doi.org/10.23919/JCC.2022.04.021

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *80*, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

Jakubauskas, M. E., Lulla, K. P., & Mausel, P. E. (1990). Assessment of vegetation change in a fire-altered forest landscape. *Photogrammetric Engineering and Remote Sensing*, *56*, 371–377.

Jenks, G. F. (1967). The data model concept in statistical mapping. *International Yearbook of Cartography*, *7*, 186–190.

Jenks, G. F., & Caspall, F. C. (1971). Error on choroplethic maps: Definition, measurement, reduction. *Annals of the Association of American Geographers*, *61*(2), 217–244. http://www.jstor.org/stable/2562442

Jeon, B.-K., Jang, J.-H., & Hong, K.-S. (2002). Road detection in spaceborne SAR images using a genetic algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(1). https://doi.org/10.1109/36.981346

Jiang, B. (2013). Head/Tail breaks: A new classification scheme for data with a heavy-tailed distribution. *Professional Geographer*, *65*(3), 482–494. https://doi.org/10.1080/00330124.2012.700499

Jozdani, S. E., Johnson, B. A., & Chen, D. (2019). Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. *Remote Sensing*, *11*(14). https://doi.org/10.3390/rs11141713

Kahraman, I., Karas, I. R., & Akay, A. E. (2018). Road extraction techniques from remote sensing images: A review. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*(4/W9), 339–342. https://doi.org/10.5194/isprs-archives-XLII-4-W9-339-2018

Karlson, M., Ostwald, M., Reese, H., Sanou, J., Tankoano, B., & Mattsson, E. (2015). Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and Random Forest. *Remote Sensing*, *7*(8), 10017–10041. https://doi.org/10.3390/rs70810017

Kass, M., & Witkin, A. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 321–331. https://doi.org/10.1007/BF00133570

Keaton, T., & Brokish, J. (2002). A level set method for the extraction of roads from multispectral imagery. *Applied Imagery Pattern Recognition Workshop, 2002*, 141–147. https://doi.org/10.1109/AIPR.2002.1182268

Keuchel, J., Naumann, S., Heiler, M., & Siegmund, A. (2003). Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. *Remote Sensing of Environment*, *86*, 530–541. https://doi.org/10.1016/S0034-4257(03)00130-5

Khan, F. (2012). An initial seed selection algorithm for k-means clustering of georeferenced data to improve replicability of cluster assignments for mapping application. *Applied Soft Computing Journal*, *12*(11), 3698–3700. https://doi.org/10.1016/j.asoc.2012.07.021

Khatami, R., Mountrakis, G., & Stehman, S. V. (2016). A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sensing of Environment*, *177*, 89–100. https://doi.org/10.1016/j.rse.2016.02.028

Kleinschroth, F., Gourlet-Fleury, S., Gond, V., Sist, P., & Healey, J. R. (2016). Logging roads in tropical forests: Synthesis of literature written in French and English highlights environmental impact reduction through improved engineering. *Bois et Forets Des Tropiques*, *328*, 13–26. https://doi.org/10.19182/bft2016.328.a31299

Kleinschroth, F., & Healey, J. R. (2017). Impacts of logging roads on tropical forests. *Biotropica*, *49*(5), 620–635. https://doi.org/10.1111/btp.12462

Koutsias, N., Mallinis, G., & Karteris, M. (2009). A forward/backward principal component analysis of Landsat-7 ETM+ data to enhance the spectral signal of burnt surfaces. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(1), 37–46. https://doi.org/10.1016/j.isprsjprs.2008.06.004

Landsat Missions. (n.d.-a). *Landsat 7*. Retrieved April 15, 2023, from https://www.usgs.gov/landsat-missions/landsat-7

Landsat Missions. (n.d.-b). *Landsat 8*. Retrieved April 15, 2023, from https://www.usgs.gov/landsat-missions/landsat-8

Lasaponara, R. (2006). On the use of principal component analysis (PCA) for evaluating interannual vegetation anomalies from SPOT/VEGETATION NDVI temporal series. *Ecological Modelling*, *194*(4), 429–434. https://doi.org/10.1016/j.ecolmodel.2005.10.035

Leichtle, T., Geiß, C., Wurm, M., Lakes, T., & Taubenböck, H. (2017). Unsupervised change detection in VHR remote sensing imagery – An object-based clustering approach in a dynamic urban environment. *International Journal of Applied Earth Observation and Geoinformation*, *54*, 15–27. https://doi.org/10.1016/j.jag.2016.08.010

Li, C., Kao, C.-Y., Gore, J. C., & Ding, Z. (2007). Implicit active contours driven by local binary fitting energy. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1–7. https://doi.org/10.1109/CVPR.2007.383014

Li, R., & Cao, F. (2018). Road network extraction from high-resolution remote sensing image using homogenous property and shape feature. *Journal of the Indian Society of Remote Sensing*, *46*(1), 51–58. https://doi.org/10.1007/s12524-017-0678-6

Li, Y., Xu, L., & Piao, H. (2009). Semi-automatic road extraction from high-resolution remote sensing image: Review and prospects. *Proceedings - 2009 9th International Conference on Hybrid Intelligent Systems, HIS 2009*, *1*, 204–209. https://doi.org/10.1109/HIS.2009.317

Liu, K., Shi, W., & Zhang, H. (2011). A fuzzy topology-based maximum likelihood classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(1), 103–114. https://doi.org/10.1016/j.isprsjprs.2010.09.007

Liu, L., & Lim, S. (2016). A framework of road extraction from airborne LiDAR data and aerial imagery. *Journal of Spatial Science*, *61*(2), 263–281. https://doi.org/10.1080/14498596.2016.1147392

Liu, Y., Dai, Q., Liu, J. B., Liu, S. Bin, & Yang, J. (2014). Study of burn scar extraction automatically based on level set method using remote sensing data. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0087480

Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. https://doi.org/10.1109/TIT.1982.1056489

Loozen, Y., Rebel, K. T., de Jong, S. M., Lu, M., Ollinger, S. V., Wassen, M. J., & Karssenberg, D. (2020). Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the Random Forests method. *Remote Sensing of Environment, 247*. https://doi.org/10.1016/j.rse.2020.111933

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability: Vol. 1: Statistics* (pp. 281–297). University of California Press.

Mann, D., Agrawal, G., & Joshi, P. K. (2019). Spatio-temporal forest cover dynamics along road networks in the Central Himalaya. *Ecological Engineering*, *127*, 383–393. https://doi.org/10.1016/j.ecoleng.2018.12.020

Mansaray, L. R., Huang, J., & Kamara, A. A. (2016). Mapping deforestation and urban expansion in Freetown, Sierra Leone, from pre- to post-war economic recovery. *Environmental Monitoring and Assessment*, *188*(8). https://doi.org/10.1007/s10661-016-5469-y

Marinelli, D., Coops, N. C., Bolton, D. K., & Bruzzone, L. (2022). Forest change detection in LiDAR data based on polar change vector analysis. *IEEE Geoscience and Remote Sensing Letters, 19*. https://doi.org/10.1109/LGRS.2020.3022282

Markogianni, V., & Dimitriou, E. (2016). Landuse and NDVI change analysis of Sperchios river basin (Greece) with different spatial resolution sensor data by Landsat/MSS/TM and OLI. *Desalination and Water Treatment*, *57*, 29092–29103. https://doi.org/10.1080/19443994.2016.1188734

Matheron, G. (1974). *Random sets and integral geometry*. Wiley.

McKeown, D. M. Jr., & Denlinger, J. L. (1988). Cooperative methods for road tracking in aerial imagery. *Proceedings CVPR '88: The Computer Society Conference on Computer Vision and Pattern Recognition*, 662–672. https://doi.org/10.1109/CVPR.1988.196307

Mena, J. B. (2003). State of the art on automatic road extraction for GIS update: A novel classification. *Pattern Recognition Letters*, *24*, 3037–3058. https://doi.org/10.1016/S0167-8655(03)00164-8

Minn, S. T., Mizoue, N., & Ota, T. (2022). Evaluating ground disturbance at elephant skid trails, logging roads and log landings under the Myanmar selection system. *Journal of Forest Research, 27*(6), 409–418. https://doi.org/10.1080/13416979.2022.2067618

Minu, S., & Shetty, A. (2015). A comparative study of image change detection algorithms in MATLAB. *Aquatic Procedia*, *4*, 1366–1373. https://doi.org/10.1016/j.aqpro.2015.02.177

Mosammam, H. M., Nia, J. T., Khani, H., Teymouri, A., & Kazemi, M. (2017). Monitoring land use change and measuring urban sprawl based on its spatial forms: The case of Qom city. *Egyptian Journal of Remote Sensing and Space Science*, *20*(1), 103–116. https://doi.org/10.1016/j.ejrs.2016.08.002

Naikoo, M. W., Rihan, M., Ishtiaque, M., & Shahfahad. (2020). Analyses of Land Use Land Cover (LULC) change and built-up expansion in the suburb of a metropolitan city: Spatio-temporal analysis of Delhi NCR using Landsat datasets. *Journal of Urban Management*, *9*(3), 347–359. https://doi.org/10.1016/j.jum.2020.05.004

Nakaguro, Y., Makhanov, S. S., & Dailey, M. N. (2011). Numerical experiments with cooperating multiple quadratic snakes for road extraction. *International Journal of Geographical Information Science*, *25*(5), 765–783. https://doi.org/10.1080/13658816.2010.498377

Natural Resources Canada. (2021). *Boreal forest*. https://natural-resources.canada.ca/our-natural-resources/forests/sustainable-forest-management/boreal-forest/13071

Neigh, C., Taylor, M. P., & Rocchio, L. (n.d.). *Landsat 9*. Retrieved April 15, 2023, from https://landsat.gsfc.nasa.gov/satellites/landsat-9/

Neigh, C., Taylor, M. P., & Rocchio, L. (2013). *The thematic mapper*. NASA Landsat Science. https://landsat.gsfc.nasa.gov/article/the-thematic-mapper/

Nevatia, R., & Babu, K. R. (1980). Linear feature extraction and description. *Computer Graphics and Image Processing*, *13*, 257–269. https://doi.org/10.1016/0146-664X(80)90049-0

Ontario Ministry of Natural Resources. (1990). *Environmental guidelines for access roads and water crossings*. King's Printer for Ontario, 2012–23. https://www.ontario.ca/page/environmental-guidelines-access-roads-and-water-crossings

Ontario Ministry of Natural Resources and Forestry. (2019, March 3). *Forest management: Boreal landscapes*. King's Printer for Ontario, 2012–23. https://www.ontario.ca/page/forest-management-boreal-landscapes

Ontario Ministry of Natural Resources and Forestry. (2020). *Forest management planning manual*. https://files.ontario.ca/mnrf-forest-management-planning-manual-en-2021-04-09.pdf

Osher, S., & Sethian, J. A. (1988). Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics, 79*, 12–49.

Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, *12*(1 Suppl.). https://doi.org/10.1016/j.jag.2009.11.002

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*(1), 217–222. https://doi.org/10.1080/01431160412331269698

Patra, S., Sahoo, S., Mishra, P., & Mahapatra, S. C. (2018). Impacts of urbanization on land use / cover changes and its probable implications on local climate and groundwater level. *Journal of Urban Management*, *7*(2), 70–84. https://doi.org/10.1016/j.jum.2018.04.006

Peiman, R. (2011). Pre-classification and post-classification change-detection techniques to monitor land-cover and land-use change using multi-temporal Landsat imagery: A case study on Pisa Province in Italy. *International Journal of Remote Sensing*, *32*(15), 4365–4381. https://doi.org/10.1080/01431161.2010.486806

Polykretis, C., Grillakis, M. G., & Alexakis, D. D. (2020). Exploring the impact of various spectral indices on land cover change detection using change vector analysis: A case study of Crete Island, Greece. *Remote Sensing*, *12*(319). https://doi.org/10.3390/rs12020319

Qiaoping, Z., & Couloigner, I. (2004). Automatic road change detection and GIS updating from high spatial remotely-sensed imagery. *Geo-Spatial Information Science*, *7*(2), 89–95. https://doi.org/10.1007/BF02826642

Radke, R. J., Andra, S., Al-Kofahi, O., & Roysam, B. (2005). Image change detection algorithms: A systematic survey. *IEE Transactions on Image Processing*, *14*(3), 294–307. https://doi.org/10.1109/TIP.2004.838698

Rajeswari, M., Gurumurthy, K. S., Reddy, P. L., & Omkar Senthilnath J., S. N. (2011). Automatic road extraction based on normalized cuts and level set methods. *International Journal of Computer Applications*, *18*(7). https://doi.org/10.5120/2298-2988

Räsänen, A., Rusanen, A., Kuitunen, M., & Lensu, A. (2013). What makes segmentation good? A case study in boreal forest habitat mapping. *International Journal of Remote Sensing*, *34*(23), 8603–8627. https://doi.org/10.1080/01431161.2013.845318

Richards, J. A. (2013). Clustering and unsupervised classification. In *Remote Sensing Digital Image Analysis* (pp. 319–341). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30062-2_9

Rochery, M., Jermyn, I. H., & Zerubia, J. (2006). Higher order active contours. *International Journal of Computer Vision*, *69*(1), 27–42. https://doi.org/10.1007/s11263-006-6851-y

Rodarmel, C., & Shan, J. (2002). Principal component analysis for hyperspectral image classification. *Survey and Land Information Systems*, *62*(2), 115–000. https://www.researchgate.net/publication/265198128

Rosin, P. L. (2002). Thresholding for change detection. *Computer Vision and Image Understanding*, *86*, 79–95. https://doi.org/10.1006/cviu.2002.0960

Rosin, P. L., & Ioannidis, E. (2003). Evaluation of global image thresholding for change detection. *Pattern Recognition Letters*, *24*, 2345–2356. https://doi.org/10.1016/S0167-8655(03)00060-6

Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., Loveland, T. R., Vermote, E., & Zhang, C. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, *114*, 35–49. https://doi.org/10.1016/j.rse.2009.08.011

Saati, M., & Amini, J. (2017). Road network extraction from high-resolution SAR imagery based on the network snake model. *Photogrammetric Engineering and Remote Sensing*, *83*(3), 207–215. https://doi.org/10.14358/PERS.83.3.207

Sanchez-Hernandez, C., Boyd, D. S., & Foody, G. M. (2007). Mapping specific habitats from remotely sensed imagery: Support vector machine and support vector data description based classification of coastal saltmarsh habitats. *Ecological Informatics*, *2*(2), 83–88. https://doi.org/10.1016/j.ecoinf.2007.04.003

Savage, S. L., Lawrence, R. L., Squires, J. R., Holbrook, J. D., Olson, L. E., Braaten, J. D., & Cohen, W. B. (2018). Shifts in forest structure in northwest Montana from 1972 to 2015 using the Landsat archive from Multispectral Scanner to Operational Land Imager. *Forests*, *9*(157). https://doi.org/10.3390/f9040157

Schowengerdt, R. A. (2007). Chapter 5: Spectral transforms. In *Remote Sensing* (Third Edition). Elsevier Inc. https://doi.org/10.1016/B978-0-12-369407-2.50008-5

Serra, J. (1982). *Image analysis and mathematical morphology*. Academic Press.

Sghaier, M. O., & Lepage, R. (2016). Road extraction from very high resolution remote sensing optical images based on texture analysis and beamlet transform. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(5), 1946–1958. https://doi.org/10.1109/JSTARS.2015.2449296

Shi, W., Miao, Z., Wang, Q., & Zhang, H. (2014). Spectral-spatial classification and shape features for urban road centerline extraction. *IEEE Geoscience and Remote Sensing Letters*, *11*(4), 788–792. https://doi.org/10.1109/LGRS.2013.2279034

Singh, A. (1989). Review article: Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, *10*(6), 989–1003. https://doi.org/10.1080/01431168908903939

Song, M., & Civco, D. (2004). Road extraction using SVM and image segmentation. *Photogrammetric Engineering & Remote Sensing*, *12*, 1365–1371. https://doi.org/10.14358/PERS.70.12.1365

Soni, P. K., Rajpal, N., & Mehta, R. (2018). A comparison of road network extraction from High Resolution Images. *First International Conference on Secure Cyber Computing and Communications*, 525–531. https://doi.org/10.1109/ICSCCC.2018.8703276

Sundaresan, A., Varshney, P. K., & Arora, M. K. (2007). Robustness of change detection algorithms in the presence of registration errors. *Photogrammetric Engineering & Remote Sensing*, *73*(4), 375–383. https://doi.org/10.14358/PERS.73.4.375

Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., & Fisher, P. F. (2015). A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, *160*, 1–14. https://doi.org/10.1016/j.rse.2015.01.006

Thompson, S. D., Nelson, T. A., White, J. C., & Wulder, M. A. (2015). Mapping dominant tree species over large forested areas using Landsat Best-Available-Pixel image composites. *Canadian Journal of Remote Sensing, 41*, 203–218. https://doi.org/10.1080/07038992.2015.1065708

Truong, C., Oudre, L., & Vayatis, N. (2018). Selective review of offline change point detection methods. *Signal Processing, 167*. https://doi.org/10.1016/j.sigpro.2019.107299

Ukoha, P. A., Okonkwo, S. J., & Adewoye, A. R. (2021). Bayesian Change Point detection of vegetation cover dynamics of Akure forest reserve, Ondo State in Southwestern, Nigeria. *Journal of Applied Sciences and Environmental Management*, *25*(8), 1449–1452. https://doi.org/10.4314/jasem.v25i8.25

Ünsalan, C., & Boyer, K. L. (2011). Review on building and road detection. In *Multispectral Satellite Image Understanding* (pp. 139–144). Springer London. https://doi.org/10.1007/978-0-85729-667-2_11

Valero, S., Chanussot, J., Benediktsson, J. A., Talbot, H., & Waske, B. (2010). Advanced directional mathematical morphology for the detection of the road network in Very High Resolution remote sensing images. *Pattern Recognition Letters*, *31*(10), 1120–1127. https://doi.org/10.1016/j.patrec.2009.12.018

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks, 10*(5), 988–999. https://doi.org/10.1109/72.788640

Vapnik, V. N., & Chervonenkis, A. Y. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications*, *16*(2), 264–280. https://doi.org/10.1137/1116025

Verma, A. K., Garg, P. K., & Hari Prasad, K. S. (2017). Sugarcane crop identification from LISS IV data using ISODATA, MLC, and indices based decision tree approach. *Arabian Journal of Geosciences, 10*(16). https://doi.org/10.1007/s12517-016-2815-x

Vese, L. A., & Chan, T. F. (2002). A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, *50*(3), 271–293. https://doi.org/10.1023/A:1020874308076

Vlachopoulos, O., Leblon, B., Wang, J., Haddadi, A., LaRocque, A., & Patterson, G. (2020). Delineation of bare soil field areas from unmanned aircraft system imagery with the mean shift unsupervised clustering and the random forest supervised classification. In *Canadian Journal of Remote Sensing* (Vol. 46, Issue 4, pp. 489–500). Bellwether Publishing, Ltd. https://doi.org/10.1080/07038992.2020.1763789

Vorovencii, I. (2017). Applying the change vector analysis technique to assess the desertification risk in the south-west of Romania in the period 1984–2011. *Environmental Monitoring and Assessment, 189*(10). https://doi.org/10.1007/s10661-017-6234-6

Vosselman, G., & De Knecht, J. (1995). Automatic extraction of man-made objects from aerial and space images road tracing by profile matching and Kalman filtering. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, 265–274. https://doi.org/10.1007/978-3-0348-9242-1_25

Wang, H., Huang, T. Z., Xu, Z., & Wang, Y. (2016). A two-stage image segmentation via global and local region active contours. *Neurocomputing*, *205*, 130–140. https://doi.org/10.1016/j.neucom.2016.03.050

Wang, J. (1993). Linda — A system for automated linear feature detection and analysis. *Canadian Journal of Remote Sensing*, *19*(1), 9–21. https://doi.org/10.1080/07038992.1993.10855146

Wang, J., Treitz, P. M., & Howarth, P. J. (1992). Road network detection from SPOT imagery for updating geographical information systems in the rural-urban fringe. *International Journal of Geographical Information Systems*, *6*(2), 141–157. https://doi.org/10.1080/02693799208901901

Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., & Eklund, P. (2016). A review of road extraction from remote sensing images. *Journal of Traffic and Transportation Engineering (English Edition)*, *3*(3), 271–282. https://doi.org/10.1016/j.jtte.2016.05.005

Wattelez, G., Dupouy, C., & Juillot, F. (2022). Unsupervised optical classification of the seabed color in shallow oligotrophic waters from Sentinel-2 images: A case study in the Voh-Koné-Pouembout lagoon (New Caledonia). *Remote Sensing*, *14*(4). https://doi.org/10.3390/rs14040836

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., Coops, N. C., Hall, R. J., Hostert, P., Dyk, A., & Guindon, L. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, *40*(3), 192–212. https://doi.org/10.1080/07038992.2014.945827

Win, R. N., Suzuki, R., & Takeda, S. (2012). Remote sensing analysis of forest damage by selection logging in the Kabaung Reserved Forest, Bago Mountains, Myanmar. *Journal of Forest Research*, *17*(2), 121–128. https://doi.org/10.1007/s10310-011-0276-3

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., Vogelmann, J., Wulder, M. A., & Wynne, R. (2008). Free access to Landsat imagery. *Science*, *320*(5879), 1011. https://doi.org/10.1126/science.320.5879.1011a

Wulder, M. A., White, J. C., & Coops, N. C. (2011). Fragmentation regimes of Canada's forests. *Canadian Geographer*, *55*(3), 288–300. https://doi.org/10.1111/j.1541-0064.2010.00335.x

Xia, H., Karimi, H. A., & Meng, L. (2017). Parallel implementation of Kaufman's initialization for clustering large remote sensing images on clouds. *Computers, Environment and Urban Systems*, *61*, 153–162. https://doi.org/10.1016/j.compenvurbsys.2014.06.002

Xiao, P., Sheng, G., Zhang, X., Liu, H., & Guo, R. (2021). Direction-dominated change vector analysis for forest change detection. *International Journal of Applied Earth Observation and Geoinformation*, *103*. https://doi.org/10.1016/j.jag.2021.102492

Xu, C., & Prince, J. L. (1998). Generalized gradient vector flow external forces for active contours. *Signal Processing*, *71*, 131–139.

Yang, K., Pan, M., Luo, Y., Chen, K., Zhao, Y., & Zhou, X. (2019). A time-series analysis of urbanization-induced impervious surface area extent in the Dianchi Lake watershed from 1988–2017. *International Journal of Remote Sensing*, *40*(2), 573–592. https://doi.org/10.1080/01431161.2018.1516312

Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R., & Evangelista, P. H. (2017). A survival guide to Landsat preprocessing. *Ecology*, *98*(4), 920–932. https://doi.org/10.1002/ecy.1730

Zafiropoulos, P., & Schenk, T. (1998). Extraction of road structures with color energy models. *Electronic Imaging*. https://doi.org/10.1117/12.333792

Zaidi, S. M., Akbari, A., Samah, A. A., Kong, N. S., & Gisen, J. I. A. (2017). Landsat-5 time series analysis for land use/land cover change detection using NDVI and semi-supervised classification techniques. *Polish Journal of Environmental Studies*, *26*(6), 2833–2840. https://doi.org/10.15244/pjoes/68878

Zhang, C., Murai, S., & Baltsavias, E. P. (1999). Road network detection by mathematical morphology. *ISPRS Workshop 3D Geospatial Data Production: Meeting Application Requirements*, 185–200. https://doi.org/10.3929/ethz-a-004334280

Zhang, Q., Wang, J., Peng, X., Gong, P., & Shi, P. (2002). Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. *International Journal of Remote Sensing*, *23*(15), 3057–3078. https://doi.org/10.1080/01431160110104728

Zhao, K., Wulder, M. A., Hu, T., Bright, R., Wu, Q., Qin, H., Li, Y., Toman, E., Mallick, B., Zhang, X., & Brown, M. (2019). Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote Sensing of Environment*, *232*. https://doi.org/10.1016/j.rse.2019.04.034

Zhou, J., Bischof, W. F., & Caelli, T. (2006). Road tracking in aerial images based on human-computer interaction and Bayesian filtering. *ISPRS Journal of Photogrammetry and Remote Sensing*, *61*(2), 108–124. https://doi.org/10.1016/j.isprsjprs.2006.09.002

Zhou, T., Sun, C., & Fu, H. (2019). Road information extraction from high-resolution remote sensing images based on road reconstruction. *Remote Sensing*, *11*(79). https://doi.org/10.3390/rs11010079

Zhou, W., Troy, A., & Grove, M. (2008). Object-based land cover classification and change analysis in the Baltimore metropolitan area using multitemporal high resolution remote sensing data. *Sensors*, *8*, 1613–1636. https://doi.org/10.3390/s8031613

Zhou, X., & Mahmassani, H. S. (2007). A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework. *Transportation Research Part B: Methodological*, *41*(8), 823–840. https://doi.org/10.1016/j.trb.2007.02.004

Ziems, M., Breitkopf, U., Heipke, C., & Rottensteiner, F. (2012). Multiple-model based verification of road data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *1–3*. https://doi.org/10.5194/isprsannals-I-3-329-2012

Zongfan, B., Ling, H., Xuhai, J., Ming, L., Liangzhi, L., Huiqun, L., & Jiaxin, L. (2022). Spatiotemporal evolution of desertification based on integrated remote sensing indices in Duolun County, Inner Mongolia. *Ecological Informatics*, *70*. https://doi.org/10.1016/j.ecoinf.2022.101750

Chapter 3: Optimal Best Available Pixel Composite Approach for Forest Road Extraction

3.1 Introduction and Objectives

Government agencies, foresters, biologists, and environmental scientists all require easily accessible and inexpensive data to monitor forest changes under their management. Highly accurate and current land cover change information is critical for monitoring and managing forested ecosystems (Ahmed et al., 2017; Savage et al., 2018). The types and rates of these land cover changes, as well as the response of the ecosystem to past changes, allows for improved predictions for how forests will respond to current and future disturbances (Bodart et al., 2011). Mapping changes from the earliest time available can assist in developing forest management policies that can adequately address the accelerating forest disturbances associated with climate change (Savage et al., 2018). Through remotely sensed approaches, changes to forested ecosystems can be monitored. Satellite-based remote sensing can provide spatially extensive, consistent, and accessible datasets in various parts of the electromagnetic spectrum, making it suitable for monitoring forest health and mortality (Byer & Jin, 2017).

The boreal forest in Ontario has undergone a variety of forest disturbances across time and space. Stand-replacing forest disturbances are those which result in short- or long-term changes to the land cover (urbanization, road construction, industrial activity, harvesting or wildfires). Non-stand replacing disturbances are lower magnitude disturbances that do not result in a full change to land cover type (insect infestation, drought stress, or disease) (Ahmed et al., 2017; Hermosilla et al., 2015b). Forest roads are a type of stand replacing forest disturbance well suited for monitoring through remotely sensed techniques.

To extract and attribute year of construction to historic forest roads from satellite imagery, cloud and anomaly-free images from comparable sensors should be used. To meet this objective, Best Available Pixel (BAP) composites of Landsat imagery from 1974 to 2019 using Sensors 1-5, 7 and 8 were generated. BAP composites enable the most suitable pixel in a Landsat pixel stack to be selected and composited into an analysis ready image based on user-defined parameters (Beckschäfer, 2017; Corbane et al., 2020; Griffiths et al., 2014; Hermosilla et al., 2015a; White et al., 2014). This paper seeks to evaluate the impact of BAP composite parameters and image bands on road extraction outcomes using the semi-automated algorithm presented in Chapter 4. In addition, the BAP compositing approach has been extended to include the Multi Spectral Scanner (MSS) system and Operational Land Imager (OLI) imagery. The original BAP algorithm only included Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors. Due to the limitations of the MSS imagery (fewer bands, narrow spatial resolution, data gaps, poor cloud masking, coarse resolution, and sensor anomalies), previous studies often do not include the years 1972 to 1984 in their forest monitoring. To the author's knowledge, currently Francini et al. (2023) is the only other BAP compositing execution that has extended the algorithm to include the OLI sensor. The proposed approach differs slightly from Francini et al. (2023) in the management of missing opacity information. The BAP compositing MSS and OLI extensions allow for the full archive to be leveraged and for the addition of critical time series change analysis information. These results will enable optimized BAP imagery to be used for semi-automated forest road extraction procedures.

This chapter will provide an overview of BAP composites and previous applications of pixel-based composites. This summary will be followed by a detailed description of the methodology, including the study site, candidate images, image processing, BAP algorithm and

parameters, as well as accuracy assessment. Two BAP compositing analyses were conducted. The first examines how the addition of a brightness parameter impacts road extraction results. The second examines which Landsat band combinations produce the strongest road extraction results. The results of the BAP compositing testing will be presented followed by a discussion of the significance and concluding remarks.

3.2 Literature Review

3.2.1 Pixel Composites

Free and open access to the entire Landsat series of sensors (from 1972 to present) since 2008 has allowed for significant scientific innovation (Hermosilla et al., 2016; Thompson et al., 2015; Woodcock et al., 2008). Increased computing power and affordable data storage options have supported these advances (Corbane et al., 2020; Maxwell & Sylvester, 2012). The development of automated approaches to radiometric and geometric pre-processing have also created standardized image products to enhance research endeavors (Griffiths et al., 2013; Masek et al., 2006; Roy et al., 2010; Thompson et al., 2015). These scientific and logistical advances have facilitated the development of large-area, application specific, pixel-based compositing approaches. These compositing approaches use the best pixel in each pixel series from a set of candidate images acquired on multiple dates (Beckschäfer, 2017). In this case, best is defined as having the highest scores or strongest value for multiple user defined characteristics such as year, day of year (DOY), distance to clouds and shadows, spectral characteristics (maximum normalized difference vegetation index (NDVI) or median near infrared (MED-NIR) value), phenology, sun angle, and sensor type (Corbane et al., 2020; Hermosilla et al., 2016; Man et al., 2018; Thompson et al., 2015). These pixels, independent of their neighbors' characteristics, are then combined into a singular image that is ideally free of clouds and sensor anomalies as well as optimized for the intended analysis (Beckschäfer, 2017; Corbane et al., 2020; Griffiths et al.,

2014; Hermosilla et al., 2015a; White et al., 2014). Pixels that do not have sufficiently suitable spatial or temporal observations become areas of no data and can be addressed through a variety of protocols (Hermosilla et al., 2016).

3.2.2 Historical Progression of Remote Sensing Composites

Remote sensing composite approaches were initially developed for sensors with a wide swath and near daily global coverage but had high percentages of cloud cover. Combining these images would reduce view angle effects, aerosol contamination, cloud cover, other atmospheric impacts, and increase the usability of the entire dataset (Griffiths et al., 2014; White et al., 2014). Prior to the opening of the Landsat archive, compositing was predominantly used for lower spatial resolution imagery like Advanced Very-High-Resolution Radiometer (AVHRR) (Cihlar & Manak, 1994; Holben, 1986) and Moderate Resolution Image Spectroradiometer (MODIS) (Ju et al., 2010; Roy, 2000). When the Landsat data became available for free download, the United States Geological Survey (USGS) saw an annual increase in image downloads of 50 times (Young et al., 2017). In addition to being the longest historical repository of remotely sensed data, systematic collection practices have been integral to research innovation. The Landsat sensors have remained consistent through time due to rigorous quality assurance (QA) and calibration approaches by the Earth Resources Observation and Science Center (EROS) (Beckschäfer, 2017; Braaten et al., 2015). This standardization allows for rich Landsat Time Series (LTS) analyses, particularly for the TM, ETM+ and OLI sensors (Beckschäfer, 2017; Hermosilla et al., 2016). Many researchers have proposed methodologies for rule-based compositing for multiple Landsat images (Hermosilla et al., 2015a; Man et al., 2018; Thompson et al., 2015). These approaches enable Land Use Land Cover (LULC) change and monitoring applications at a variety of scales (regional to global) (Hermosilla et al., 2015a). The

development of annual cloud-free composites enables wall to wall multi-year time series that are essential for many research and science applications (White et al., 2014).

3.2.3 Types of Pixel-Based Composites

There are multiple types of pixel-based composites to meet the needs of different applications, including: (1) annual or single year, (2) multi-year and (3) proxy-value image composites. The most significant difference between these composites is how pixels are handled if there are not suitable observations within the target year (Hermosilla et al., 2015a). An annual composite is comprised of the best pixel for each location from a singular target year. In addition to application specific rules (sensor preference, clouds, atmospheric opacity, etc.), annual composites may be configured to model a specific time of year or phenological window (Target DOY ± 30 Days (within a singular year)) (Corbane et al., 2020; White et al., 2014). Multiple annual composites can be created for a singular year to create seasonal composites. These may be useful to classify a range of LULC types as some land covers are spectrally similar for some parts of the year but differ significantly in others (Bleyhl et al., 2017). If no pixels within the series satisfy the criteria, the pixel is not filled and is coded as "NoData," resulting in incomplete coverage (White et al., 2014). A multi-year image composite will use data from outside of the target year to fill pixels when no suitable observations are found within the target year (Target DOY ± 30 Days (± 1 or 2 Years)) (Corbane et al., 2020; Hermosilla et al., 2015a). Although this serves to fill the data gaps, this approach may confound series patterns and trends (Hermosilla et al., 2015a). Proxy composites were developed to overcome the limitations associated with multi-year composites. "NoData" values are assigned a synthetic value generated through examining the complete pixel time series and determining a value most spectrally similar in both time and space (Hermosilla et al., 2015a). If the pixel series is stable through time, the proxy value can be calculated through averaging (or another method) the entire series. If the pixel's value varies,

change events in the trajectory will be determined and the pixels within the pixel's spatial neighborhood may be incorporated. These proxy values may also be assigned to pixels with anomalous values (scores that exceed expected values or pixels flagged with haze or other atmospheric effects) (White et al., 2014).

3.2.4 Shift from Scene to Pixel-Based Composites: Advantages and Disadvantages

The recent shift to pixel-based compositing may signal the end of scene-based compositing approaches. Pixel-based approaches include multiple improvements over scene-based algorithms, which include (but are not limited to): analysts are not restricted to a few low cloud cover images but instead can access more seasonally optimal acquisitions (Griffiths et al., 2013); a reduction in excessive cloud coverage or data gaps (like the Landsat 7 Scan Line Corrector (SLC) failure) (Bleyhl et al., 2017; Griffiths et al., 2014); footprints artificially generated through path and row acquisition can be replaced with larger scale study sites (Griffiths et al., 2013); overcoming discontinuity in image archives (Bleyhl et al., 2017; Corbane et al., 2020); addressing limitations related to atmospheric and radiometric inconsistencies due to seasonal or sun angle differences (Corbane et al., 2020); entire imagery archives can be leveraged by using partial images (Griffiths et al., 2013); the incorporation of a wider breadth of spatial and temporal considerations as opposed to just temporal resolution and acquisition path and row (Hansen & Loveland, 2012; White et al., 2014); observation frequency can be increased through using pixels from imagery previously discarded due to high cloud coverage as well across track overlap (Griffiths et al., 2013); and fusion of data from multiple sensors can allow for enhanced broad area mapping and monitoring (Hansen & Loveland, 2012; White et al., 2014).

Despite these advantages, there are many challenges that remain for analysts utilizing a pixel-based composite methodology. When considering large complex study sites, like Canada

or Ontario's boreal forest, the optimal number of cloud-free observations may not be reached. The probability of cloud-free observations is determined by the combination of local weather conditions, acquisition schedule, and downlink activity (White et al., 2014). Cloud-free observations are more difficult to obtain in mountainous regions with increased cloud cover and have greater implications for crops that are temporally dynamic such as rice (Bleyhl et al., 2017; Francini et al., 2023; Griffiths et al., 2013; Man et al., 2018). This cloud cover, in addition to strict compositing rules, data availability, and continual haze and smoke, can result in data gaps or spectral anomalies. These data gaps can persist spatially and temporally through timeseries composites (Gómez et al., 2015). Gap infilling techniques can be applied, allowing for exact and estimate values within the pixel series, but add complexity to the algorithm. Fine spatial resolution imagery composites would be exceptionally time consuming to process and prohibitively computationally expensive due to complexities associated with many acquisition dates like varying illumination conditions and view angles (Thompson et al., 2015). As a result, imagery with large or moderate spatial resolution is used for mapping larger areas. The pixel sizes of this data can subsume smaller scale objects and earth surface features, which can reduce the contrast between pixels and limit types of applicable analyses for these composite products (Thompson et al., 2015). Finally, discrete changes in a growing season and other seasonal artifacts may make it difficult to attribute some LULC changes within a pixel series. This attribution error may be exacerbated by a larger observation acquisition date range (Francini et al., 2023). As a result, BAP composites may not be appropriate for all applications (Corbane et al., 2020; Gómez et al., 2015).

3.2.5 Remote Sensing Composites and Forestry Applications

Many researchers have used compositing approaches on a variety of study sites and phenomena to generate and analyze cloud and anomaly free images. Early pixel composting

approaches used singular rules to develop composites including Holben (1986), Roy et al. (2010), and Potapov et al. (2011, 2012). These methods select observations based on band statistics. Other researchers began to use multiple rule composting approaches such as Griffiths et al. (2013), or implemented spectral similarity like Flood (2013)'s Medoid approach. Most early approaches used either top of atmosphere (TOA) reflectance or surface reflectance imagery, with varying rules for cloud and shadow, image date, and target spectral indices. Until 2013, researchers developed custom cloud masking algorithms. The FMask algorithm (Zhu & Woodcock, 2012) and EROS systematic pre-processing was not adopted until 2013 (Flood, 2013; Griffiths et al., 2013). White et al. (2014) introduced the BAP compositing approach to meet the needs of forest ecosystem monitoring in Canada. Their work was applied to a Canadian national dataset in 2010 as well as annual composites of Saskatchewan and Newfoundland between 1998 and 2012. This approach was then applied by Hermosilla et al. (2015a, 2015b) who detected and characterized annual forest change between 1984 and 2012 using BAP composites generated from the White et al., (2014) BAP composite protocol. Their approach is referred to as the Composite2Change (C2C) and used normalized burn ratio (NBR) composites to complete breakpoint detection and change type attribution using object metrics and a random forest (RF) classifier. This work was extended further by White et al. (2017), who completed a nationwide characterization of forest disturbance and recovery between 1985 and 2010. Future works on pixel-based composting incorporated systematic pre-processing and composting approaches. Beckschäfer (2017) applied a BAP composite algorithm to 270 Landsat NDVI Thematic Mapper (TM) and ETM+ images of rubber plantations in China to conduct a time series analysis. This process enabled the age of the rubber trees grown between the years 1988 and 2015 to be determined with a root mean square error (RMSE) of 2.5 years. In 2017, Frantz et

al. introduced a phenology adaptive approach for pixel composites using Landsat and MODIS. This algorithm generates homogeneous and seamless multi-annual composites suitable for most seasons. Man et al. (2018) used Landsat 8 BAP composites in their land cover analysis of Hanoi, Vietnam, known as a perpetually cloudy region. Using BAP composites, the researchers were able to achieve 99.78% cloud-free images and enhance the classification accuracy by an average of 20.47% compared to non-composite imagery. Morresi et al. (2022) analyzed burn severity in the western Italian Alps during 2017 using phenology coherent (based on the work of Frantz et al. (2017)) reflectance composites from Sentinel-2 images. The classification accuracy increased approximately 5% using reflectance composites in comparison to uncalibrated bi-temporal indices. The proposed approach builds on the BAP approach by White et al., (2014) by extending their protocol to the MSS and TM sensors as well as optimizing the composite for forest road monitoring.

3.2.6 Designing Pixel Composites for Road Extraction

Creating a compositing methodology that covers different dates (and in some cases, different years) requires the consideration between inter- and intra-annual landcover changes, phenology effects, persistence of cloud coverage, topographic effects, data availability, and dynamism of landscape processes (Griffiths et al., 2013; White et al., 2014). Different information needs (mapping varying types of disturbances or estimating biophysical properties) also influence compositing decisions, such as the target date and application specific rules. Analysts must assess tradeoffs in the creation of their application specific composites. Overall, researchers aim to create phenology consistent composites with minimal data gaps that best reflect their target phenomenon (White et al., 2014). A general LULC mapping application may focus on the main period of photosynthetic activity and select a seasonal window to achieve a predominately leaf on state (Griffiths et al., 2013).

Francini et al. (2023) developed an assessment approach for pixel-based image composites. In their analysis, they highlight five key characteristics of strong image composites including (1) no data gaps, (2) no atmospheric interference (cloud or haze), (3) radiometrically consistent composite products, (4) proximity of pixel dates to target DOY and (5) spatial congruence of pixels. Limited observations because of clouds, haze, and related shadows can result in data gaps or pixels with no observations within the study period. The impact of undetected clouds, haze, or smoke can also limit the ability to detect earth surface changes. The number of days between the target DOY and acquisition pixel is important to ensure there is consistent phenology within the resultant composite and that accurate statistics can be developed. Limiting the number of observation dates with which pixels are selected is also essential to enhance spatial congruency of pixel acquisition dates, phenological phases, land covers, and atmospheric conditions. Using these five metrics, Francini et al. (2023) evaluated the BAP composite approach (White et al., 2014) as well as the Medoid pixel-based composites (Flood, 2013) in 2019 over Europe. Overall, it was found that the Medoid approach was more radiometrically consistent whereas the BAP approach was more temporally consistent. Medoid and BAP approaches both had reliable and consistent spectral values. The BAP approach selected observations from a smaller data range and a smaller day of year standard deviation (DOYSD) than Medoid in most cases. The smaller date range and DOYSD enables BAP composites to have reduced interference from atmospheric condition changes and enhances classification accuracy. The amount of noise was on average 3.7% greater for BAP composites than Medoid. These values are appropriate as Medoid focuses on radiometrically consistent surface reflectance data, while BAP composites have other criteria. Qiu et al. (2023) also evaluated multiple Landsat image compositing approaches. Their study evaluated ten

composting algorithms including MAX-NDVI (Holben, 1986), MAX-RNB (Qiu et al., 2023), MED-NIR (Potapov et al., 2011), WELD (Roy et al., 2010), BAP (White et al., 2014), PAC (Frantz et al., 2017), WPS (Griffiths et al., 2019), MEDOID (Flood, 2013), COSSIM (Nelson & Steinwand, 2015) and NLCD (Jin et al., 2023) (See Table 3 for the algorithm definition and a brief description). These composites were evaluated in six study sites with two images each in the categories of low, intermediate, and high cloud coverage in three compositing types: monthly, seasonal, and annual. Visual assessment indicates that BAP or NLCD would be most appropriate for seasonal composites and PAC or BAP is most suitable in annual composites. Visual assessment focused on the impact of salt-and-pepper effects, patches, and other image artifacts. Quantitative evaluation indicates that BAP performs second best in annual composite spectral performance, has the best spatial evaluation in all composite timeframes, has the lowest disagreement for annual land cover maps, and has strong annual land change performance compared to the other algorithms. The evaluations of the BAP approach by Qiu et al. (2023) and Francini et al. (2023) indicate the strong suitability of this composting algorithm for forest road monitoring applications.

Table 3: Qiu et al. (2023) pixel composite algorithms reviewed.

Algorithm	*Full Name*	*Data*	*Select Observations By:*	*Reference*
MAX-NDVI	Maximum NDVI	MODIS	Minimum normalized difference vegetation index	(Holben, 1986)
MAX-RNB	Maximum RNB	TM, ETM +, OLI	Maximum ratio of near infrared band to blue band	(Qiu et al., 2023)
MED-NIR	Median NIR	ETM +	Median near infrared band	(Potapov et al., 2011)
WELD	Web-Enabled Landsat Data	ETM+	Maximum normalized difference vegetation index or brightness temperature	(Roy et al., 2010)

BAP	Best Available Pixel	ETM+	Score of sensors, day of year, cloud distance and atmospheric opacity	(White et al., 2014)
PAC	Phenology Adaptive Composting	TM, ETM+, OLI	Weighted scores of day of year, cloud distance, haze optimized transformation, and spectral correlation	(Frantz et al., 2017)
WPS	Weighted Parametric Scoring	NASA HLS V1.2	Weighted scores of sensors, day of year, cloud distance, haze optimized transformation, and cloud coverage	(Griffiths et al., 2019)
MEDIOD	Medoid	TM, ETM+	Minimum Medoid to other data in spectral bands	(Flood, 2013)
COSSIM	Cosine Similarity Compositing	TM	Cosine similar to other data in spectral bands	(Nelson & Steinwand, 2015)
NLCD	National Land Cover Database	TM, ETM+, OLI/TIRS	Minimum distance to pseudo median data in spectral bands	(Jin et al., 2023)

The proposed approach was based on the work of Griffiths et al. (2013), White et al. (2014), and Hermosilla et al. (2015a), who evaluated pixel suitability with flexible parametric weighting schemes. An annual pixel composite was designed to enhance the ability to detect forest roads. Five scores are calculated for each pixel and include sensor score, DOY score, distance to cloud or cloud shadow score, opacity score, and pixel brightness score. The sensor and DOY scores are calculated at the image level (all pixels for that image receive the same score) whereas the remaining scores are calculated at the pixel level. Each of these scores are summed for each pixel and the pixel with the highest score (the “best” available pixel) was selected for use in the image composite (White et al., 2014). This approach will be discussed in further detail in the Section 3.6

3.3 Methodology

3.3.1 Study Site

The study site for this analysis is a 3 405 squared kilometer region of northern Ontario, Canada (Please see Figure 1). This site is within a single Landsat scene, World Reference System-1 (WRS-1) path 029 and row 025 as well as WRS-2 path 027 and row 025. This area is within the World Geodetic System (WGS) 1984 Universal Transverse Mercator (UTM) Zone 15N.

The boundary of the study site is 105 kilometers (kms) northeast of Dryden, Ontario, and 190 kms northwest of Thunder Bay, Ontario. There are two communities within the study area, Savant Lake as well as the southern part of Ojibway Nation of Saugeen Reserve. Savant Lake, within the Thunder Bay District (population 146 048 in 2016), is located along the Highway 599 Wilderness Corridor. The Canadian National Railway transcontinental main line runs through the Savant Lake community. Twenty kms northwest of Savant Lake is the Ojibway Nation of Saugeen, with an on-reserve population of 92 (2007) (Ojibway Nation of Saugeen, n.d.). Between Savant Lake and Ojibway Nation of Saugeen, Highway 516 branches off Highway 599 to the west.

The Ecological Land Classification system is used to describe ecosystem types. The Ontario Ministry of Natural Resources and Forestry (MNRF) has classified Ontario into ecoregions using this system based on ecological functions, composition, and structure. The study site is within Lake Nipigon or 3W Ecoregion. The climate in this ecoregion is moist and cold, with a short mean growing season of 161 to 182 days. The study site is within the Precambrian shield with granitic and gneissic bedrock (Crins et al., 2009). Over half of the ecoregion has thinly covered acidic bedrock with poor substrate quality. The majority of the landcover within the ecoregion is forested with 33.2% as mixed forest, 19.9% coniferous forest,

and 17.1% deciduous forest. Water and sparse forest account for 10.9% and 5.6% of the land cover, respectively. Forestry is the predominant land use activity within the ecoregion (Crins et al., 2009).

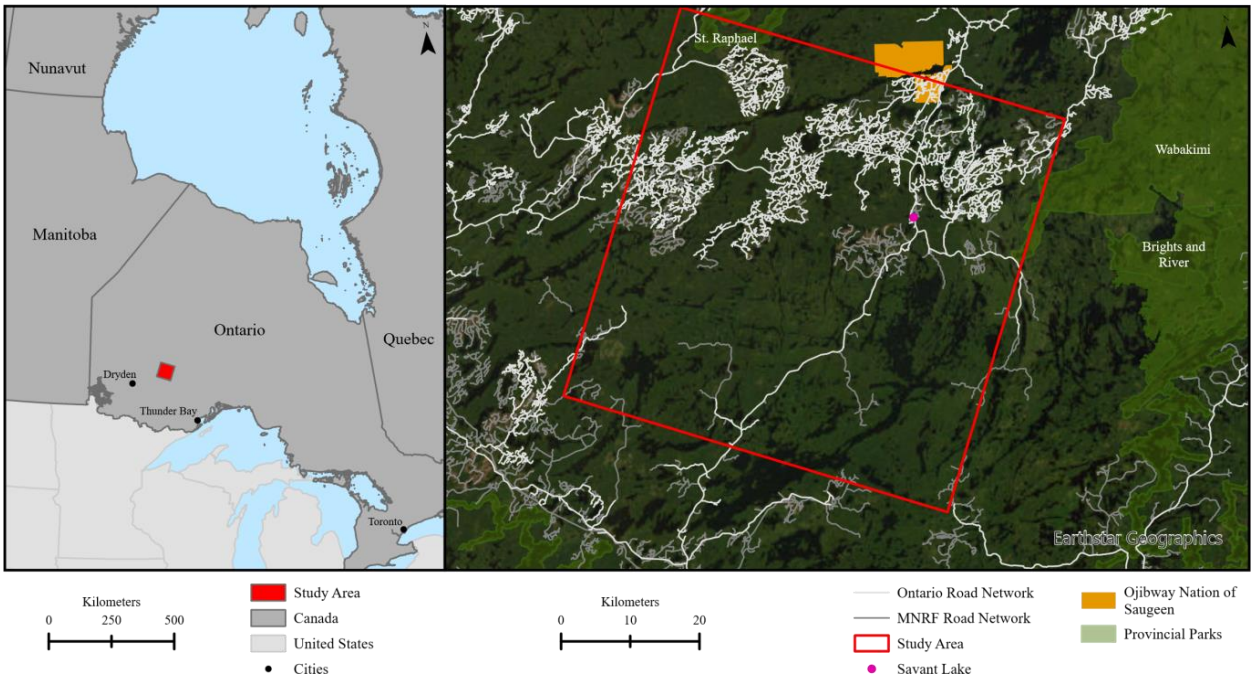

Figure 1: Map of Lake Savant study site (Land Information Ontario, 2012, 2015; Ontario Ministry of Natural Resources and Forestry, 2019b, 2019a; Statistics Canada, 2016; United States Census Bureau, 2018)

3.3.2 Candidate Landsat Images

Landsat imagery was selected for this application to balance the needs for data processing efficiency, inexpensive and easily accessible data, extensive temporal range, and spatial resolution. Please see Chapter 2 Sections 3 and 4 for further details.

To create BAP composites suitable for determining the year of construction for forest roads, MSS, TM, ETM+, and OLI images were acquired from the USGS EROS Science Processing Architecture (ESPA) interface (https://espa.cr.usgs.gov/). Application ready products with consistent radiometric and geometric properties, like those provided by the USGS, reduces pre-processing requirements, and permits the automation of the BAP compositing approach (Gómez et al., 2015).

Surface reflectance products were used between 1982 to 2019 (TM, ETM+ and OLI sensors). Surface reflectance products (as opposed to at-sensor radiance or TOA reflectance) were selected as this unit enhances the comparison between multiple images due to accounting for atmospheric effects (aerosol scattering and thin clouds) (Young et al., 2017). Currently, there are no surface reflectance products available for the MSS sensor. As a result, for 1974 to 1982 data (MSS), digital numbers (DNs) are used. MSS sensor recorded values are calibrated to radiance using the temporally varying sensor gains and offsets before being converted to 7-bit unsigned integers. DNs are used as opposed to radiance values as the conversion is linear and little enhancement is achieved compared to using the DN data. As post-classification change detection (PCCD) is used in this approach and the images are processed using sensor specific values, absolute correction is not strictly a necessity (Young et al., 2017). When the USGS develops a suitable atmospheric compensation algorithm and MSS surface reflectance products become available, the algorithm could be updated to incorporate this enhanced data (Young et al., 2017).

For the surface reflectance products (TM, ETM+ and OLI), Precision and Terrain Corrected (Level 1TP) datasets with a cloud cover below 70% were downloaded. Seventy percent was selected as the cloud cover threshold as when cloud cover exceeds this value, ground control points (GCPs) can become obscured; this reduces the efficacy of geometric corrections (White et al., 2014). For the TM and ETM+ datasets, data was pre-processed using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al., 2006; United States Geological Survey, 2020a; Wolfe et al., 2004). For the OLI datasets, pre-processing was applied prior to download through the Land Surface Reflectance Code (LaSRC) (United States

Geological Survey, 2020b; Vermote et al., 2018). These corrections enhance comparability between images and reduce post-processing at the application level.

Multiple types of corrections are applied through the LEDAPS and LaSRC including geometric and radiometric correction. Landsat Level-1 data products are precision registered and orthorectified using GCPs and the Earth topography five-minute grid (ETOPO5), a digital elevation model (DEM) with global coverage. Tier 1 products, as utilized in this study, are georegistered within a ≤12 metre (m) radial RMSE. This level of accuracy is appropriate for time series pixel level analyses (Young et al., 2017). Precise geometric registration is essential as image displacement will result in false change being detected within the image. Sub-pixel level geometric accuracy is recommended for most change detection applications. Increased registration accuracy is essential when using multiple sensors or different spatial resolutions (Hussain et al., 2013). In this approach, the extraction results are rejoined to the reference road network. This method functionally introduces a buffer factor which can reduce the impact of any misregistration. Other tier levels may not be appropriate for time series analyses as insufficient GCPs or elevation data can reduce the efficacy of precision or terrain corrections (Beckschäfer, 2017).

In addition to geometric corrections, absolute radiometric corrections are also completed prior to download. Radiometric corrections include a collection of pre-processing steps that correct sensor, solar, atmospheric, and some topographic effects. While DNs may be used for a single image analysis, absolute correction allows the values to be compared between scenes or images (Chander et al., 2009; Young et al., 2017). The LEDAPS and LaSRC protocols convert TOA reflectance to surface reflectance using algorithms by Masek et al. (2006) and Vermote et al. (2016) respectively. Only pre-processing steps necessary for the analysis are completed;

additional unnecessary processing increases the risk of introducing artifacts or errors into the dataset prior to analysis without adding any value (Young et al., 2017). The ability to accurately identify earth surface features and detect change over time is dependent on appropriately calibrated sensors – both in terms of accuracy and precision (Chander et al., 2009). It is essential that changes can be attributed to changes in radiance value as opposed to variation in solar illumination, atmospheric interference or viewing geometry (Afify, 2011); importance of consistent measures makes pre-processing an essential component of remote sensing analyses (Afify, 2011; Chander et al., 2009).

In addition to the corrected visible bands, multiple QA bands are required – pixel (cloud and cloud shadow) and radiometric saturation. These bands are included with visible bands in Tape Archives (TARs) downloaded from the USGS. Four QA products are used with this analysis: cloud and cloud shadow, opacity, radiometric saturation, and valid pixels. To generate cloud confidence, cloud shadow, as well as snow/ice flags, the CFMask is used. Opacity and aerosol information is provided as well. LEDAPS and LaSRC both provide radiometric QA bands, which indicate when specific bands were saturated during data collection, resulting in unusable results (United States Geological Survey, 2020b, 2020a). Valid pixel ranges for surface reflectance for each band are also provided in the documentation.

For the DN products (MSS), Terrain Corrected (Level 1TP) products with a cloud cover below 70% were downloaded when available**.** These images are radiometric, geometric, and precision corrected. The accuracy of these products is dependent on the GCPs and DEM used for these corrections. Despite the Level 1TP images being the best available MSS products, this level of pre-processing was not always available for the study site. When Level 1TP imagery was not available, Systematic Corrected (L1GS) products were used. This imagery is radiometrically

corrected as well as resampled for geometric correction. In the cases where L1GS imagery was used, the researcher visually validated that the imagery was aligned in the specific site that the image could be included within LTS analyses. In addition to the visible bands, the pixel quality band is also used. This band is included in the TAR downloaded from the USGS. The MSS QA band provides radiometric saturation, dropped pixel, and cloud flag information. No opacity information is provided (United States Geological Survey, 2018). Cloud masking for the MSS imagery is very difficult because if the pixel is very bright, it is difficult to discern whether it is snow, cloud, or sand. Later Landsat sensors have Shortwave Infrared (SWIR) bands and thermal bands that can assist with this differentiation (Braaten et al., 2015). Collection 2 MSS data is now available with enhanced cloud masking capabilities using the MSS cloud mask algorithm (Cubist model) (United States Geological Survey, 2020c). For future applications, the algorithm can be modified to include the specifications for this collection.

To test the effectiveness of the compositing approach, a subset of images was selected for testing. The algorithm is designed for the date range between 1974 and 2019. For this scene, images from 1972 and 1973 were omitted due to poor data quality at the start of the Landsat Mission. Images from 1982, 1986, and 1987 were omitted due to consistent cloud cover over the study site. There were no images with pixels suitable for composite. Forty-eight images were selected as the subset for this test from the years 1983, 1987, 1996, 2005, 2009, and 2019. Please see Figure 2 for the distribution of the acquisition DOY. Please see Figure 3 for the cloud coverage percentage for each image. The test years have the following number of quality images per year; 1983: 2, 1987: 6, 1996: 9, 2005: 7, 2009: 6, 2019: 14. The number of images per year is a product of agency goals, downlink activity, cloud coverage values as well as other QA metrics.

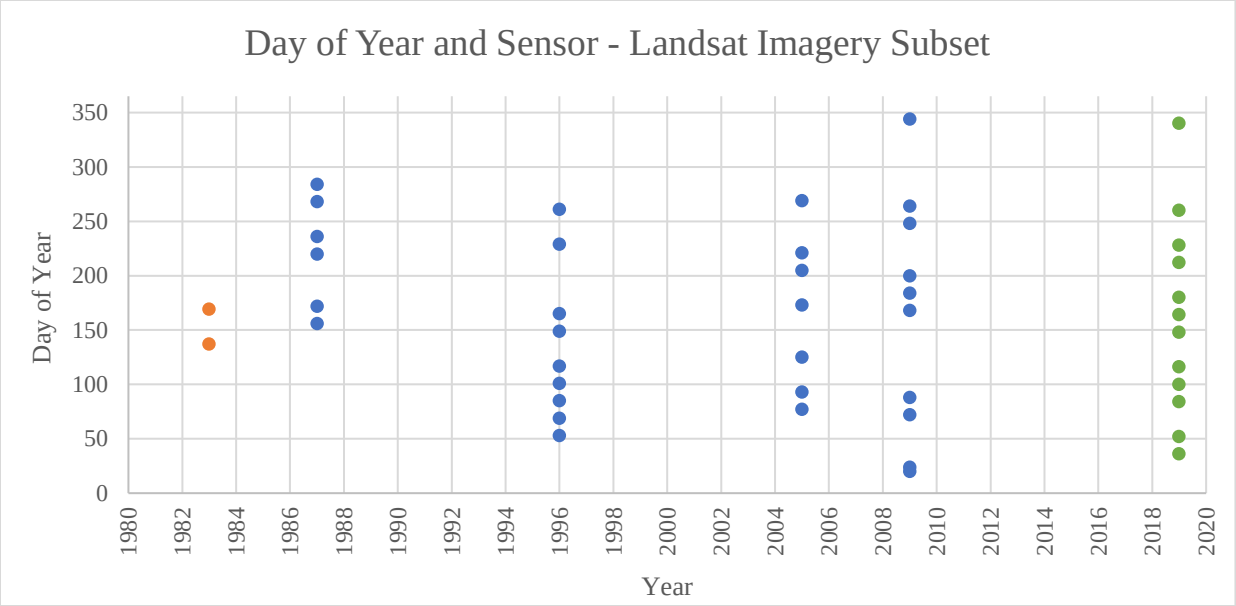

Figure 2: DOY score for Landsat image composite testing subset. Orange, blue, and green points represent MSS - Satellite 4, TM - Satellite 5 and OLI - Satellite 8, respectively.

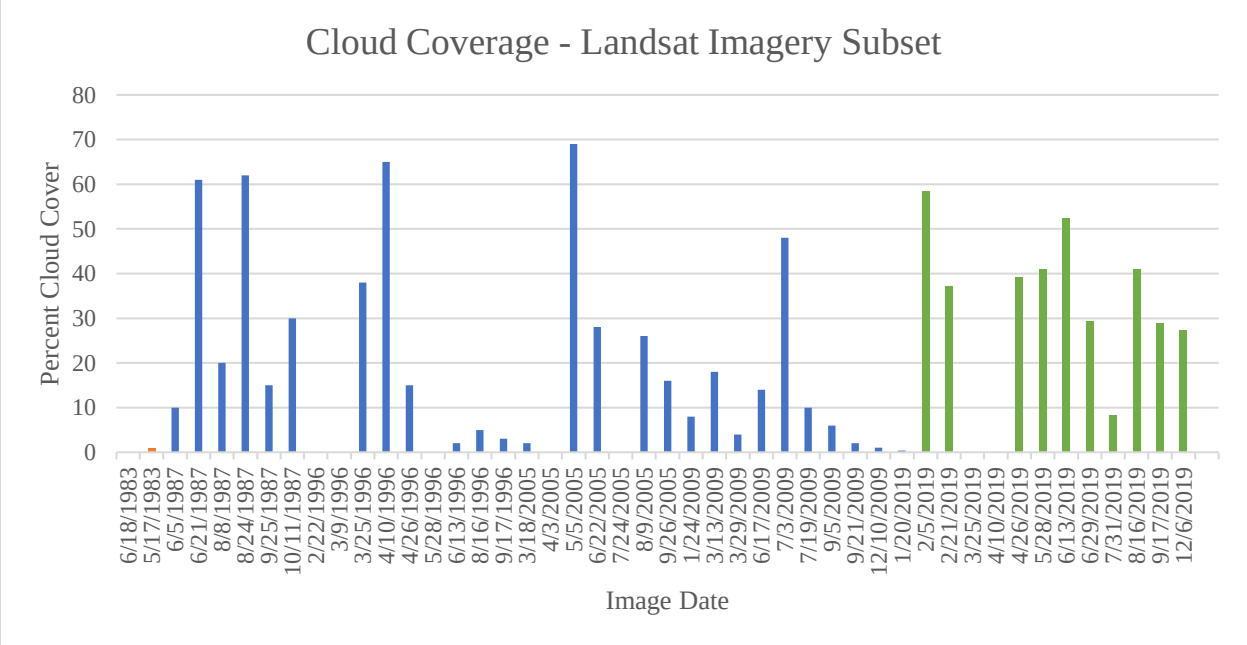

Figure 3: Cloud coverage for Landsat image composite testing subset. Orange, blue, and green points represent MSS - Satellite 4, TM - Satellite 5 and OLI - Satellite 8, respectively.

3.3.3 Reference and Ground Truth Data

To enhance road extraction, a mask of known roads in the study area is generated from reference data. The road mask is developed with two road networks: MNRF Road Segment and

Ontario Road Network (ORN) Segment with Address. These roads were accessed through Ontario GeoHub and Land Information Ontario (https://geohub.lio.gov.on.ca/).

To assess the results of the road extraction approach, ground truth data is required for accuracy assessment. Through collaboration with the MNRF, it was indicated that no accurate ground truth data for the historic forestry roads existed for this region. Ground truth data for historic forest road construction is unreliable across the province of Ontario. As a result, the author verified the existence of each road segment in the ground truth dataset manually from the Landsat satellite imagery. The entire road network within the study area was digitized as a vector based on reference road network for each target year. The analyst examined the road network in true and false colour composites as well as multiple individual bands, marking segments that existed in that year with a 1 and segments that did not exist with a 0 to create a binary classification (J. Wang, 1993). The vector road network was then converted into a raster. The imagery from multiple dates for six years was digitized: 1983 (MSS), 1987 (TM), 1996 (TM), 2005 (TM), 2009 (TM), and 2019 (OLI). These vector road networks were then converted to accuracy assessment points. For the images with 60 metre (m) resolution, there were 35 113 accuracy assessment points. For the images with 30 m resolution, there were 85 427 accuracy assessment points. These dates were selected as there were enough high-quality images to produce composites from at least two images. Road statistics for each reference road network and ground truth year are presented in Table 4. For a road to be identified, it had to be visible in the imagery by the analyst and be connected to the main network. Disconnected roads were more problematic in later years when uneven forest succession occurred on decommissioned roads. Any road segments not connected to the main network were not considered roads for this application.

Table 4: Road network lengths (kilometers)

DATASET	LENGTH OF ROAD (KM)
Full Study Site Network	2048.08
MNRF Study Site Network	1594.48
ORN Study Site Network	1520.31
1983 Ground Truth	868.57
1987 Ground Truth	955.82
1996 Ground Truth	1083.10
2005 Ground Truth	1294.44
2009 Ground Truth	1383.50
2019 Ground Truth	1512.40

3.3.4 Algorithm Overview

The proposed semi-automatic algorithm for forest road extraction using BAP composites was written in Python 3.7.0 (Python Software Foundation, 2001) using the arcpy site package for ArcGIS Pro 3.0.1 (Environmental Systems Research Institute, 2022). The algorithm can be divided into four main sections: (1) Pre-Processing, (2) BAP Compositing, (3) Road Extraction and Change Detection and (4) Logic Rule Execution. In this chapter (Chapter 3), the (1) Pre-processing and (2) BAP Compositing will be outlined. A methodology flow chart is provided in Figure 4. Two BAP compositing analyses were conducted. The first analysis seeks to find the optimal BAP parameters for road extraction. The second analysis shows which Landsat band combinations for the four sensors produce the strongest road extraction results. Results for two BAP composite analyses will be presented and discussed. The road extraction algorithm will be applied to derive these results, but the methodology will be discussed in depth in Chapter 4. The logic rules are not applied in the BAP composite testing presented in this chapter. For more information on the logic rules, please see Chapter 4.

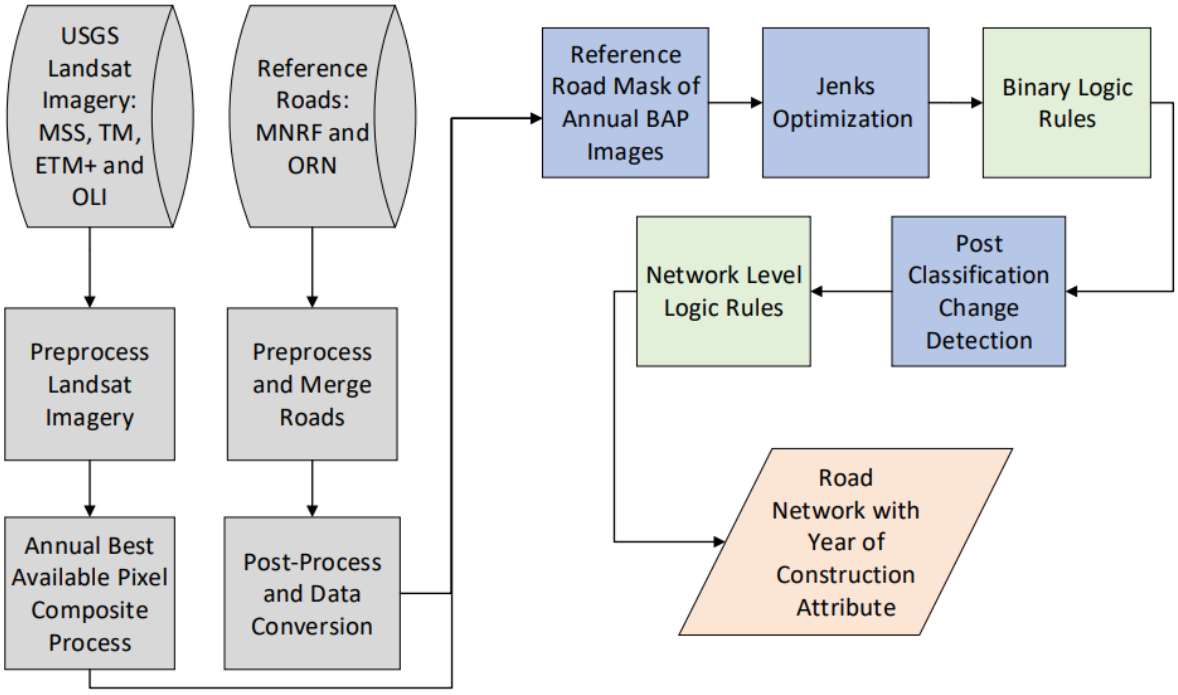

Figure 4: Methodology overview flowchart.

3.3.5 Landsat Imagery Pre-processing

Prior to BAP compositing, all images within the scene must be pre-processed. Where possible, the same pre-processing steps are applied to all input Landsat images. The Landsat sensors (MSS, TM/ETM+ and OLI) have different characteristics (band designations, available QA data, and formatting) that prevent uniform processing; the slight modifications made to account for these differences will be noted where applicable.

Images downloaded from the USGS are projected into WGS 1984 UTM Zone 15N. The provided projection is retained for this study. Downloaded images with below 70% cloud coverage had cloud, cloud shadow, water, snow, and ice classes masked where possible. If any algorithm encounters any imagery with greater than 70% cloud coverage, the images are automatically deleted from the input images and are not processed. The MSS quality assessment is not as sophisticated, so as a result, only clouds are masked in these images (see above for

discussion of MSS cloud masking limitations). It is essential that these pixels are masked, as the high reflectance values can often be misclassified as roads. Saturated pixels as well as invalid reflectance pixels are also masked as these pixels are unusable for any analysis. Opacity and aerosol quality information is also pre-processed for use in the BAP algorithm. Data is then clipped to the study site. If any images have complete cloud or high aerosol coverage, the image is dropped from the candidate images. After preprocessing, the percentage of area remaining in the study site is calculated for each image. Images that have an area of less than 50% of the study area are removed to reduce image noise.

3.3.6 Best Available Pixel Composite Approach

There are five variables that are used within the BAP composite testing. These include DOY, sensor, distance to cloud or cloud shadow, opacity score, and brightness score. Different combinations and weightings of these scores are tested to determine which BAP composite produces the strongest road extraction results. A description of the calculation of each of these scores is provided below.

Day of Year Score (DOY)

The DOY score is assigned to all pixels in an image relative to the target DOY assigned by the analyst. In this application, the target DOY assigned was September 1st (Day of Year 245). This day was selected to reduce road occlusions from leaves and snow cover. This date is a good balance between some defoliation of deciduous trees without snow cover. A DOY score was assigned using Equation 5 below from Griffiths et al. (2013):

$$Score_{DOY} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \quad [5]$$

Where μ denotes the mean and σ denotes the standard deviation of all input image DOYs. X_i is the DOY of the image being assessed. In this application, μ is set to the target DOY of 245. The

DOY score is scaled to a value between 0 and 1 by dividing by the maximum score for the entire dataset.

Sensor Score

The sensor score is generated to mitigate the impact of gaps associated with the Landsat 7 ETM+ SLC failure. Sensors are assigned scores based on Table 5.

Table 5: BAP Composite Sensor Scores (Griffiths et al., 2013)

Satellite	Sensor	Assigned Score
Landsat 1, 2, 3, 4 and 5	MSS	0.5
Landsat 4 and 5	TM	1
Landsat 7	ETM+	Before SLC Failure (May 30, 2003): 1
		After SLC Failure (May 31, 2003): 0.5
Landsat 8	OLI	1

Distance to Cloud or Cloud Shadow Score

Using the outputs from the cloud masking algorithms (CFMasks (TM, ETM+ and OLI) and MSS Cloud Masking), a distance to cloud or cloud shadow is assigned to each pixel. Pixels that are identified as clouds are assigned "NoData" and pixels that are greater than 50 pixels away from clouds are assigned a score of 1. In images with no clouds, all pixels are uniformly assigned a score of 1. Pixels that are not clouds and within 50 pixels away from cloud as well as cloud shadows are assigned a score between 1 and 0 using Equation 6 from Griffiths et al. (2013):

$$Score_{CloudDistance} = \frac{1}{1+e^{(-0.2\,(D_i, D_{req}) - (\frac{D_{req}-D_{min}}{2}))}} \quad [6]$$

Where D_i is the pixel's distance to another pixel identified as cloud or cloud shadow, D_{req} is the minimum required distance (50 pixels) and D_{min} is the minimum distance of the given pixel observations (0 pixels).

Opacity Score

Image haze or images with high opacity can reduce the quality of generated image composites. The opacity score is managed differently for each of the three sensor groups. For Landsat sensors TM and ETM+, LEADAPS use the dark dense vegetation method presented in Kaufman et al. (1997) to estimate aerosol optical thickness (AOT). The LEDAPS surface reflectance output includes an AOT map derived from the TM or ETM+ Blue band. AOT values less than 0.1 are clear, values between 0.1 and 0.3 are considered average and opacity values greater than 0.3 are hazy (United States Geological Survey, 2020a). Pixels with values less than 0.2 are assigned a value of 1 and pixels with an opacity greater than 0.3 are assigned "NoData." Pixels with a value greater than 0.2 and less than 0.3 are assigned a score using Equation 7 (Griffiths et al., 2013). The Landsat 8 OLI sensor does not have the same optical information. Instead, the LaSRC specifications provide pixel flags that indicate low, medium, and high-level aerosols (United States Geological Survey, 2020b). For this application, low-level aerosol values are assigned a 1, medium-level aerosols are assigned 0.5 and high-level aerosols are assigned "NoData." MSS imagery does not include any aerosol or opacity information in the QA data bands. As a result, all MSS data is assigned a pixel score of 1.

$$Score_{Opacity} = 1 - \frac{1}{1+e^{(-0.2\,(min(O_i, O_{max}))-(\frac{O_{max}-O_{min}}{2}))}} \quad [7]$$

Where O_i is the pixel's opacity value, O_{max} is the maximum opacity value (0.3) and O_{min} is the minimum opacity value (0.2).

Brightness Score

As forest road pixels are considerably brighter than the vegetation that surrounds them, preference is given to brighter pixels. This road property enables road pixels to be selected for compositing over occluded pixels that may prevent roads from being extracted. The brightness score is calculated using Equation 8:

$$Score_{Brightness} = \frac{BP_i}{BP_{max}} [8]$$

Where BP_i is the pixel brightness of the given pixel and BP_{max} is the maximum brightness value for the pre-process and clipped images for each year.

Final BAP Composite

The objective of annual image compositing is to populate a final composite image using the BAP from all the input images. After scores are assigned for the five variables for each image, the scores are summed according to Equation 9. All input images are then compared to determine the maximum score for a given pixel location. Once the highest score is determined, the pixel value associated with the highest score at that pixel location is used to populate the final image composite:

$$Final\ Score = (W * Score_{DOY}) + (W * Score_{Sensor}) + (W * Score_{CloudDistance}) + (W * Score_{Opacity}) + (W * Score_{Brightness}) [9]$$

Where *W* represents an assigned weight for each of the scores. The scores must add to 1 or the sum of the desired weights. $Score_{DOY}$, $Score_{Sensor}$, $Score_{CloudDistance}$, $Score_{Opacity}$ and $Score_{Brightness}$ represent the DOY score, sensor score, distance to cloud or cloud shadow score, opacity score and brightness score, respectively.

3.3.7 Road Extraction Algorithm

The road extraction process involves three steps: (1) pre-processing the reference road data, (2) generating a mask using the reference road data, and (3) extracting the roads using the Jenks Optimization algorithm. The ORN and MNRF road networks are merged to act as a mask of known roads for the yearly BAP composites. The Jenks Optimization algorithm is then used on the masked images. Pixels are classified into two classes: road and non-road. The raster results are not converted into polylines as the topology is not assessed. The logic rules are not

applied for this analysis. Please see Chapter 4 for a detailed description of the road extraction algorithm.

3.3.8 Accuracy Assessment

Results were evaluated using a quantitative accuracy assessment process. Broadly, this process involves collection of ground truth samples, generation of error matrices, and analysis of the results (Congalton, 1991; P. Gong & Mu, 2000; Q. Zhang et al., 2002). As noted above, six years were selected from the 45-year study period. The digitized road network for each of the selected years serves as the ground truth data. These raster cells were converted into 35113 and were 85427 points for the 30 m and 60 m resolution imagery, respectively. Error or confusion matrices describe the level of accuracy between accuracy assessment samples and the corresponding classified values for the same image (Han et al., 2018; Zaidi et al., 2017). Overall classification accuracy is assessed through user's accuracy, producer's accuracy, and the kappa coefficient (Zaidi et al., 2017). The confusion matrix is an extremely common approach to accuracy in remote sensing research (Zaidi et al., 2017). An automated approach was developed to generate accuracy assessment points and confusion matrices for each set of test results.

3.4 Results

Two series of tests were conducted to optimize the BAP compositing approach for road extraction. The first series of tests examined the parameters of the BAP composite protocol. Three approaches, (1) included brightness parameter (IBP), (2) excluded brightness parameter (EBP) and (3) heavily weighted brightness parameter (HWBP), were utilized. The three parameterizations were tested and compared to determine which resulted in the highest road extraction algorithm results through the use of a confusion matrix. The second test series determined which band combinations were the most appropriate for extracting roads from LTS imagery. Due to the spatial and temporal breadth of the dataset, reducing the number of bands

required for the analysis can significantly decrease processing time. By testing several single band combinations (one per sensor), the strongest bands could be incorporated into the automated algorithm. No multiple band combination results were included in the results as the accuracy did not increase significantly using multiple bands per sensor. The small increases in accuracy did not outweigh the increase in processing time and complexity. After completing this testing, the strongest band combination and parameterization were used to generate annual composites. The BAP composite processing and road extraction results are presented.

3.4.1 BAP Composite Parameter Testing Results

The BAP composite approach was developed for monitoring all forest disturbances. To enhance the extraction of forest roads, three parameterizations of the BAP composite scoring were completed. The first approach included a brightness parameter to prioritize the brightest pixels within a stack. The brightness score was evenly weighted with the other four parameters. The second approach did not include a brightness parameter. All parameters were equally weighted. The final approach included the brightness parameter, but had it weighed at 60% of the score. The remaining four parameters were 10% of the score each. The formulas for each of these calculations can be seen in Table 6.

Table 6: BAP composite score parameterizations

Parameterization	Formula
Included brightness parameter (IBP)	$Final\ Score = (0.2 * Score_{DOY}) + (0.2 * Score_{Sensor}) + (0.2 * Score_{CloudDistance}) + (0.2 * Score_{Opacity}) + (0.2 * Score_{Brightness})$ [10]
Excluded brightness parameter (EBP)	$Final\ Score = (0.25 * Score_{DOY}) + (0.25 * Score_{Sensor}) + (0.25 * Score_{CloudDistance}) + (0.25 * Score_{Opacity})$ [11]

Heavily weighted brightness parameter (HWBP)	$Final\ Score = (0.1 * Score_{DOY}) + (0.1 * Score_{Sensor}) + (0.1 * Score_{CloudDistance}) + (0.1 * Score_{Opacity}) + (0.6 * Score_{Brightness})$ [12]

After generating annual BAP composites according to each of the approaches, the forest roads were extracted. Confusion matrices were generated. Table 7 provides a summary of these results. The average overall accuracy and kappa coefficient do not vary significantly between approaches, with the highest average overall accuracy being 0.880 (IBP) and the other two approaches being within 0.002 (EBP and HWBP).

Table 7: Accuracy assessment for BAP composite parameterizations

Overall Accuracy by Composite							
Parameterization	**1983**	**1987**	**1996**	**2005**	**2009**	**2019**	**Average Overall Accuracy**
Included brightness parameter (IBP)	***0.939***	***0.933***	***0.873***	0.854	**0.848**	0.832	***0.880***
Excluded Brightness Parameter (EBP)	0.939	0.933	0.865	0.854	0.848	0.832	0.879
Heavily weighted brightness parameter (HWBP)	0.928	0.933	0.868	**0.855**	0.848	**0.837**	0.878
Kappa by BAP Composite							
Parameterization	**1983**	**1987**	**1996**	**2005**	**2009**	**2019**	**Average Kappa**
Included brightness parameter (IBP)	***0.877***	***0.866***	***0.744***	0.696	***0.656***	0.582	**0.737**
Excluded Brightness Parameter (EBP)	0.877	0.866	0.728	0.696	0.656	0.582	0.734
Heavily weighted brightness parameter (HWBP)	0.853	0.866	0.734	***0.698***	0.654	**0.600**	0.734

3.4.2 BAP Composite Band Combination Results

Six single band combination tests were completed. Each of the bands for the OLI, TM and ETM+ sensors were paired with the closest available band for the MSS sensor. The band

combinations are provided in Table 8. As the MSS sensors have fewer bands than the TM, ETM+, and OLI sensors, the closest comparable band was used in the band combinations. As a result, the MSS bands were included in multiple tests. The band combination tests were generated using the IBP BAP composite parameterization.

Table 8: Band combinations for BAP Testing. Please see Table 1 and 2 for Landsat sensor band designations.

SENSOR	BT-1 - RED/ SWIR-2	BT-2 - GREEN/ BLUE	BT-3 - GREEN/ GREEN	BT-4 - RED/ RED	BT-5 - NIR/ NIR	BT-6 - NIR/ SWIR-1
MSS - 123	5	4	4	5	6	7
MSS - 45	2	1	1	2	3	4
TM AND ETM+	7	1	2	3	4	5
OLI	7	2	3	4	5	6

A confusion matrix was generated for each year for the BAP composite band combinations. A summary of these results, the overall accuracy, and kappa coefficient values are presented in Table 9. For the years 1983, 1996, 2009, and 2019, the Red/SWIR-2 (Band Test-1 (BT-1)) band combination provides the highest overall accuracies. For 1983 (MSS sensor), the red band sees only a slight increase in overall accuracy over the green band (0.004), but significantly better accuracy than the NIR bands (0.255 and 0.283). For the years 1996 and 2009 (TM sensor), the red band had accuracy within 0.01; however, all other bands had an increase of 0.06 and 0.15. The year 2019 (OLI sensor) had a small increase in accuracy over the SWIR-1 band (0.02 increase), but an increase in accuracy of more than 0.3 in comparison to all other bands. In contrast, the years 1987 and 2005 have the highest accuracies in the green and red bands. These bands saw similar overall accuracy values to the SWIR-2, blue, green, and red bands, with only a 0.013 and 0.004 increase for 1987 and 2005, respectively. As most bands have the highest accuracy in the Red and SWIR-2 bands (and difference is small for years with better accuracy in other bands), these bands combined to generate an average overall accuracy

and kappa coefficient of 0.880 and 0.737. The full confusion matrix for BT-1 can be found in Section 4.3.

Table 9: Accuracy assessment for BAP composite band combinations

Overall Accuracy by Composite							
	1983	**1987**	**1996**	**2005**	**2009**	**2019**	**Average Overall Accuracy**
Red/SWIR-2 – BT-1	***0.939***	0.933	***0.873***	0.854	***0.848***	***0.832***	***0.880***
Green/Blue – BT-2	0.935	0.928	0.705	0.849	0.803	0.772	0.832
Green/Green – BT-3	0.935	**0.946**	0.777	0.857	0.842	0.799	0.859
Red/Red – BT-4	0.939	*0.944*	0.868	***0.858***	0.845	0.768	0.870
NIR/NIR – BT-5	0.684	0.718	0.736	0.707	0.696	0.736	0.713
NIR/SWIR-1 – BT-6	0.656	0.620	0.808	0.781	0.792	0.812	0.745
Kappa Coefficient by BAP Composite							
	1*983*	**1987**	**1996**	**2005**	**2009**	**2019**	**Average Kappa**
Red/SWIR-2- BT-1	***0.877***	0.866	***0.744***	0.696	***0.656***	***0.582***	***0.737***
Blue/Green – BT-2	0.867	0.856	0.393	0.675	0.519	0.428	0.623
Green/Green – BT-3	0.867	***0.891***	0.543	0.701	0.638	0.459	0.683
Red/Red – BT-4	0.877	0.887	0.736	***0.703***	0.643	0.423	0.711
NIR/NIR – BT-5	0.398	0.454	0.456	0.272	0.153	0.121	0.309
NIR/SWIR-1 – BT-6	0.343	0.806	0.607	0.483	0.464	0.436	0.523

3.4.3 Red and SWIR-2 Annual BAP Composite

Results were generated using the strongest BAP scoring parameterization and most suitable band for each sensor, IBP and BT-1-Red/SWIR-2. The BAP composite processing details, BAP composites and road extraction results are included below (Table 10). Based on the specifications for the BAP composites, not all images can be retained for pre-processing. Images with complete coverage from cloud cover, saturated pixels, snow, ice, or sensor anomalies cannot be included in the final composites. As the study site is less than an entire Landsat scene, the cloud cover information provided in the metadata does not necessarily reflect the amount of cloud cover within the specific study area. As a result, between 0 and 9 images were deemed unacceptable for compositing and removed from year input image stacks. The BT-1 combination, Red (MSS) and SWIR-2 (TM, ETM+ and OLI) bands, composting details are

provided in Table 9 to illustrate the range in annual composite results. Of the images remaining after pre-processing, all were used in their associated annual BAP composite. The images in composites ranged between 2 and 4, although the percentage of pixels from each image varied dramatically. In 1983, 98.37% of pixels come from one image whereas in 1987, the distribution between images was closer with 44.63% coming from one image and 55.37% from the other. DOY values ranged between 176 and 3 days from the target DOY (245).

Table 10: Processing details for BT-1 Red/SWIR-2 BAP composite

Year	Input Images	Images Retained After Pre-Processing	Images in Composite	Percentage from Each Image	DOY of Image (Target 245)
1983	2	2	2	1.53%; 98.37%	137; 69
1987	6	2	2	44.63%; 55.37	156; 222
1996	9	4	4	44.48%; 0.28%; 16.53%; 38.71%	149; 165; 229; 261
2005	7	2	2	80.11%; 19.89%	205; 221
2009	9	4	4	44.35%; 0.26%; 27.78%; 27.61%	168; 200; 248; 264
2019	13	3	3	45.33%; 40.06; 14.62%	180; 212; 260

BT-1 Red/SWIR-2 BAP composites for the six test years are presented in Figure 5. The maximum DN value for the MSS composite is 221. For the surface reflectance images, the maximum values range from 0.33615 and 0.7815. Each of the composites have less than 0.01% no data values after the compositing process. Based on a visual inspection of the final composites, there are no residual unmasked clouds; however, water is not completely masked in the MSS image due to the limited MSS QA band information. The impact of unmasked water is negligible as a mask of known roads is applied before any statistics are generated for the road

extraction algorithm. The roads extracted from the BT-1 annual composites are provided in Figure 6. For reference, the ground truth roads are provided in Figure 7. The confusion matrix for each annual composite is presented in Table 11.

Visual interpretation of the extracted roads indicates that there are many road segments identified that are disconnected from the main features. Road gaps are also more common in the later years. Based on the confusion matrices, the years 1983 and 1987 have the highest overall accuracies whereas 2009 and 2019 have the lowest overall accuracies. The user's and

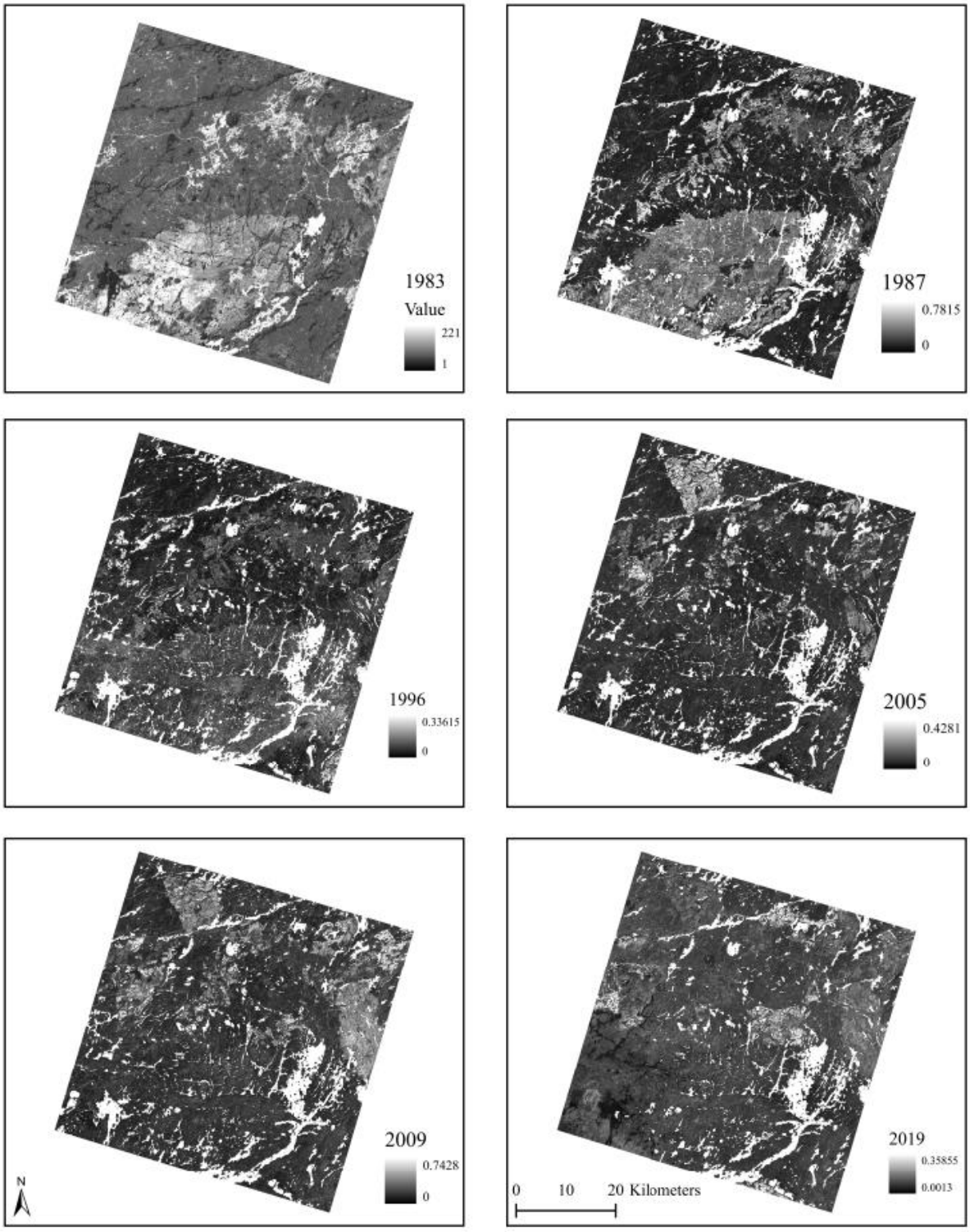

Figure 5: BAP composites Red and SWIR-2 Bands and IBP parameterization

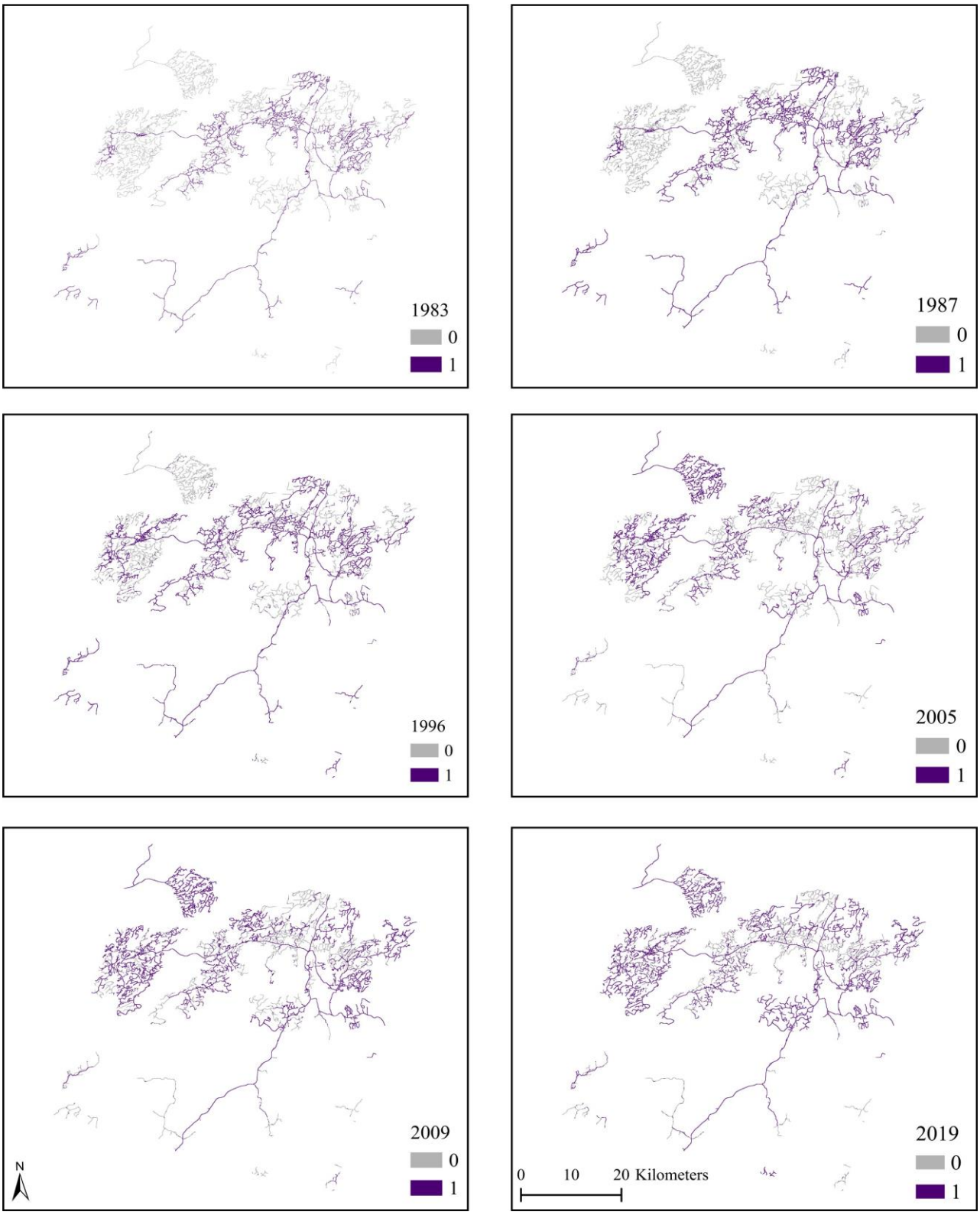

Figure 6: Roads extracted from BAP composites - Red (MSS Sensor) and SWIR-2 (TM, ETM+, and OLI Sensor) bands and IBP parameterization. One represents road and zero represents non-road pixels.

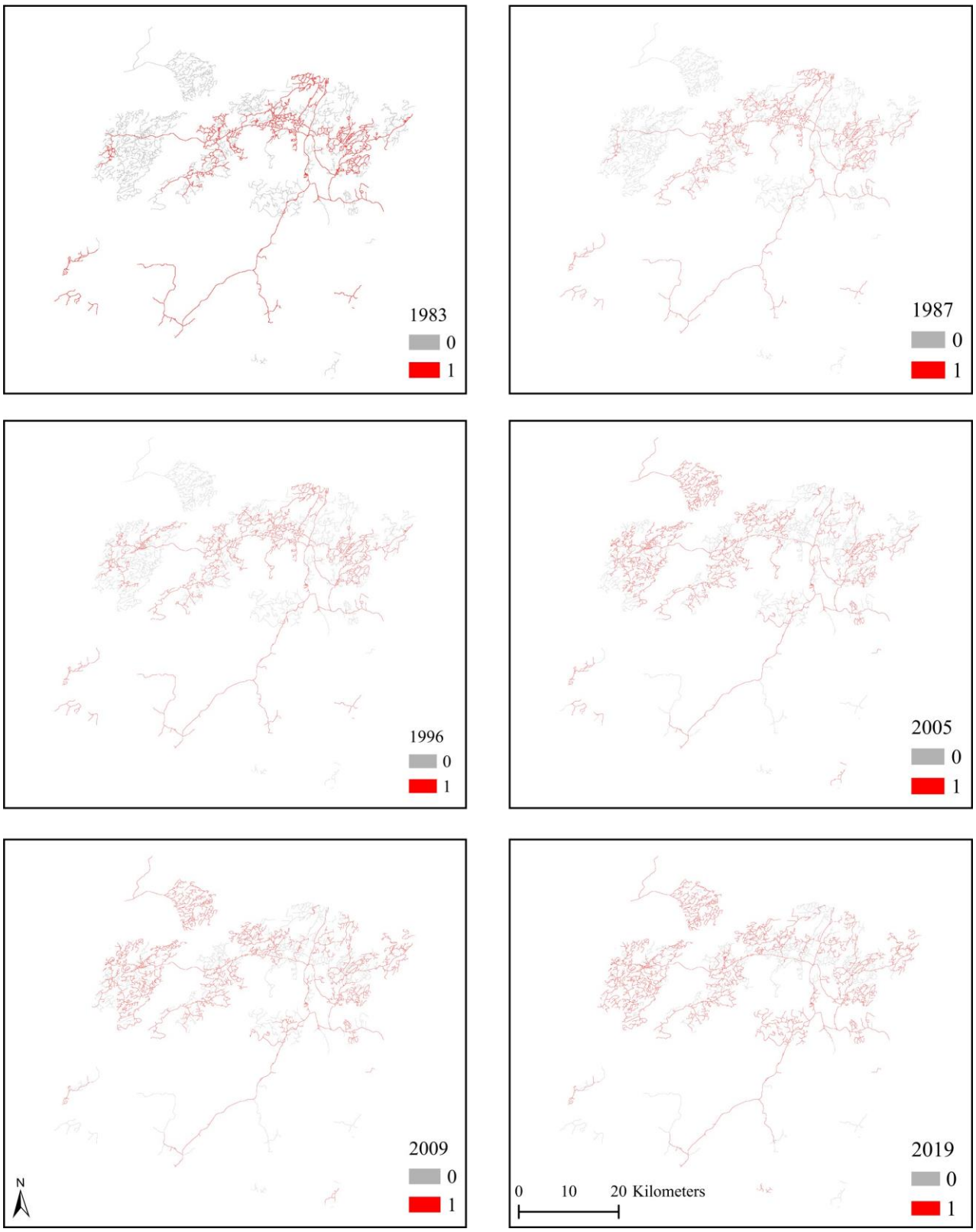

Figure 7: Ground truth roads digitized from true and false colour composites and individual bands. One represents road and zero represents non-road pixels.

Table 11: Accuracy assessment for road extracted from BAP composites - Red (MSS Sensor) and SWIR-2 (TM, ETM+, and OLI Sensor) bands and IBP parameterization.

1983					
Class Value	**Non-Road**	**Road**	**Total**	**User's Accuracy**	**Kappa**
Non-Road	18856	694	19550	0.964501	
Road	1434	14129	15563	0.907858	
Total	20290	14823	35113		
Producer's Accuracy	0.929325	0.953181		**0.939396**	
Kappa					**0.87661**
1987					
Class Value	**Non-Road**	**Road**	**Total**	**User's Accuracy**	**Kappa**
Non-Road	40490	553	41043	0.986526	
Road	5189	39195	44384	0.883089	
Total	45679	39748	85427		
Producer's Accuracy	0.886403	0.986087		**0.932785**	
Kappa					**0.86593**
1996					
Class Value	**Non-Road**	**Road**	**Total**	**User's Accuracy**	**Kappa**
Non-Road	32562	2979	35541	0.916181	
Road	7847	42039	49886	0.842701	
Total	40409	45018	85427		
Producer's Accuracy	0.805811	0.933826		**0.873272**	
Kappa					**0.74423**
2005					
Class Value	**Non-Road**	**Road**	**Total**	**User's Accuracy**	**Kappa**
Non-Road	27724	8690	36414	0.761356	
Road	3791	45222	49013	0.922653	
Total	31515	53912	85427		
Producer's Accuracy	0.879708	0.838811		**0.853899**	
Kappa					**0.69604**
2009					
Class Value	**Non-Road**	**Road**	**Total**	**User's Accuracy**	**Kappa**
Non-Road	21572	6721	28293	0.76245	
Road	6234	50900	57134	0.890888	
Total	27806	57621	85427		
Producer's Accuracy	0.775804	0.883358		**0.84835**	
Kappa					**0.65619**
2019					
Class Value	**Non-Road**	**Road**	**Total**	**User's Accuracy**	**Kappa**

Non-Road	16616	8651	25267	0.657617	
Road	5723	54437	60160	0.90487	
Total	22339	63088	85427		
Producer's Accuracy	0.743811	0.862874		**0.831739**	
Kappa					**0.582047**

producer's accuracies of road and non-road classes are above 0.8 between 1983 and 2005. Between 2005 and 2019, the non-road user's and producer's accuracies decrease to between 0.65 and 0.74 (excluding producer's accuracy of 2005). The non-road user's and producer's accuracies remain above 0.8. The lowest accuracy within the table is the 2019 non-road user's accuracy, indicating that there were high commission errors within this class. The kappa coefficients decrease as the years progress, ranging between 0.88 in 1983 and 0.52 in 2019.

3.5 Discussion

The use of appropriately parameterized and optimized BAP composites can enhance the extraction of forest roads from historic Landsat imagery. Through the testing of the three BAP composite parameter calculations, it is evident that the pixel brightness parameter does not have a significant impact on the results of the BAP composites. The limited impact may be the result of the low number of images used for each composite. In some images, there is only one acceptable pixel within the stack based on the other parameters, particularly in images with heavy cloud or aerosol coverage. Additionally, selecting the brightness pixel within a stack may accidentally result in the prioritization of unmasked clouds or sensor anomalies. These erroneous pixels may balance out any of the gains achieved through prioritizing bright road pixels over darker forested pixels. The size of the forest road clearing, and tree type might also mean that the forest canopy does not create inconsistent occlusions. Additional testing for areas with higher percentages of deciduous trees, more candidate images, and skid road extraction may see more

significant differences in accuracies. The brightness score is included as it does have a minimal gain in accuracy without a significant increase in process time.

Selecting the best band for the composite is extremely important to reduce the amount of processing time while achieving strong results. Overall, the MSS sensor sees the highest accuracy in the red band whereas the TM, ETM+ and OLI sensors see the highest accuracies in the SWIR-2 band. Red is generally used to discriminate cultural features, but in this case, the higher wavelength range of SWIR-2 is generally more accurate for the later sensors. In 1987, the green band has the highest accuracy which indicates the influence of vegetation vigor may be of greater importance based on the DOY of the input images; however, the increase in accuracy is minimal. As a result, the red (MSS) and SWIR-2 (TM, ETM+ and OLI) bands are selected for use in the compositing algorithm.

The results of the road extraction on the final BAP composite algorithm (IBP and Red/SWIR-2) generated an average overall accuracy of 0.88 and kappa of 0.74. These results indicate the BAP composites are suitable for use within the road extraction algorithm. Despite these promising overall results, when examining the annual overall accuracy values, the accuracy declines over time. This decline could be the result of multiple factors. Phenological consistency is essential for BAP composites as each season will result in earth surface features having different spectral responses. Having multiple acquisition dates within the same composite may negatively impact the statistics generated for the extraction algorithm. In this case, the use of less images used within the composite does not account for all the differences in accuracy. In addition to the number of images, the average DOY for the composite may also be important for the extraction statistics. The average DOY increases over time, which may indicate the need for additional testing with an earlier target DOY. If there are not significant occlusions caused by

tree canopies, additional vegetation greenness may increase the difference between road and non-road classes. Finally, over time, the roads network becomes more complex. Roads that are abandoned or decommissioned change at different rates along their length. Road decommissioning may generate disconnected features as well as increase the number of segments that have mixed land covers. Additionally, the point at which they can be considered undriveable is also difficult to determine based on satellite imagery, as this would depend on multiple contextual factors on the ground. Masking known decommissioned roads or developing an accompanying decommissioned road extraction process might enhance these results. The issues with decommissioned and abandoned roads likely causes the higher number of false disconnected road pixels identified.

3.6 Conclusion

Through the testing of BAP composite parameters and Landsat band combinations, the optimal BAP composite parameterization was developed for road extraction between 1974 and 2019. Two series of tests, scoring parameters and band combinations, were conducted to determine this optimal composite implementation. The comparison of the parameter scoring results indicates that the use of the brightness parameter increases accuracy a nominal amount without increasing processing time. The comparison of the band combinations results shows that the selection of the most appropriate bands for composting, red (MSS) and SWIR-2 (TM, ETM+ and OLI) increased overall road extraction accuracy from 0.712 to 0.880. After reviewing the completed road extraction using the final BAP composite approach for the extended time frame of 1974 to 2019, the results for the early time period are acceptable; however, the decline in accuracy over time due to increased complexity from road decommissioning and inconsistent phenology is a limitation. Future work on the optimized BAP approach for road extraction will

focus on addressing improving phenology consistency in the compositing process as well as enhancing road pixel contrast from the background in more complex images.

3.7 References

Afify, H. A. (2011). Evaluation of change detection techniques for monitoring land-cover changes: A case study in new Burg El-Arab area. *Alexandria Engineering Journal*, *50*, 187–195. https://doi.org/10.1016/j.aej.2011.06.001

Ahmed, O. S., Wulder, M. A., White, J. C., Hermosilla, T., Coops, N. C., & Franklin, S. E. (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. https://doi.org/10.1080/2150704X.2016.1233371

Beckschäfer, P. (2017). Obtaining rubber plantation age information from very dense Landsat TM & ETM + time series data and pixel-based image compositing. *Remote Sensing of Environment*, *196*, 89–100. https://doi.org/10.1016/j.rse.2017.04.003

Bleyhl, B., Baumann, M., Griffiths, P., Heidelberg, A., Manvelyan, K., Radeloff, V. C., Zazanashvili, N., & Kuemmerle, T. (2017). Assessing landscape connectivity for large mammals in the Caucasus using Landsat 8 seasonal image composites. *Remote Sensing of Environment*, *193*, 193–203. https://doi.org/10.1016/j.rse.2017.03.001

Bodart, C., Eva, H., Beuchle, R., Raši, R., Simonetti, D., Stibig, H. J., Brink, A., Lindquist, E., & Achard, F. (2011). Pre-processing of a sample of multi-scene and multi-date Landsat imagery used to monitor forest cover changes over the tropics. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*, 555–563. https://doi.org/10.1016/j.isprsjprs.2011.03.003

Braaten, J. D., Cohen, W. B., & Yang, Z. (2015). Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems. *Remote Sensing of Environment*, *169*, 128–138. https://doi.org/10.1016/j.rse.2015.08.006

Byer, S., & Jin, Y. (2017). Detecting drought-induced tree mortality in Sierra Nevada forests with time series of satellite data. *Remote Sensing*, *9*(929). https://doi.org/10.3390/rs9090929

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, *113*, 893–903. https://doi.org/10.1016/j.rse.2009.01.007

Cihlar, J., & Manak, D. (1994). Evaluation of compositing algorithms for AVHRR data over land. *IEEE Transactions on Geoscience and Remote Sensing*, *32*(2), 427–437. https://doi.org/10.1109/36.295057

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Corbane, C., Politis, P., Kempeneers, P., Simonetti, D., Soille, P., Burger, A., Pesaresi, M., Sabo, F., Syrris, V., & Kemper, T. (2020). A global cloud free pixel- based image composite from Sentinel-2 data. *Data in Brief*, *31*. https://doi.org/10.1016/j.dib.2020.105737

Crins, W. J., Gray, P. A., Uhlig, P. W. C., & Wester, M. C. (2009). *Ecosystems of Ontario. Part 1, ecozones and ecoregions*. Ontario, Ministry of Natural Resources, Inventory, Monitoring and Assessment Section.

Environmental Systems Research Institute, Inc. (2022). *ArcGIS Pro* (3.0.1.).

Flood, N. (2013). Seasonal composite Landsat TM/ETM+ images using the medoid (a multi-dimensional median). *Remote Sensing*, *5*(12), 6481–6500. https://doi.org/10.3390/rs5126481

Francini, S., Hermosilla, T., Coops, N. C., Wulder, M. A., White, J. C., & Chirici, G. (2023). An assessment approach for pixel-based image composites. *ISPRS Journal of Photogrammetry and Remote Sensing*, *202*, 1–12. https://doi.org/10.1016/j.isprsjprs.2023.06.002

Frantz, D., Röder, A., Stellmes, M., & Hill, J. (2017). Phenology-adaptive pixel-based compositing using optical earth observation imagery. *Remote Sensing of Environment*, *190*, 331–347. https://doi.org/10.1016/j.rse.2017.01.002

Fung, T. (1992). Land use and land cover change detection with Landsat MSS and SPOT HRV data in Hong Kong. *Geocarto International*, *7*(3), 33–40. https://doi.org/10.1080/10106049209354378

Gómez, C., White, J. C., Wulder, M. A., & Alejandro, P. (2015). Integrated object-based spatiotemporal characterization of forest change from an annual time series of Landsat image composites. *Canadian Journal of Remote Sensing*, *41*, 271–292. https://doi.org/10.1080/07038992.2015.1089162

Gong, P., & Mu, L. (2000). Error detection through consistency checking. *Annals of GIS*, *6*(2), 188–193. https://doi.org/10.1080/10824000009480549

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., & Hostert, P. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment*, *151*, 72–88. https://doi.org/10.1016/j.rse.2013.04.022

Griffiths, P., Nendel, C., & Hostert, P. (2019). Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sensing of Environment*, *220*, 135–151. https://doi.org/10.1016/j.rse.2018.10.031

Griffiths, P., Van Der Linden, S., Kuemmerle, T., & Hostert, P. (2013). Erratum: A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of*

Selected Topics in Applied Earth Observations and Remote Sensing, *6*(5), 2088–2101. https://doi.org/10.1109/JSTARS.2012.2228167

Han, X., Pan, J., & Devlin, A. T. (2018). Remote sensing study of wetlands in the Pearl River Delta during 1995–2015 with the support vector machine method. *Frontiers of Earth Science*, *12*(3), 521–531. https://doi.org/10.1007/s11707-017-0672-x

Hansen, M. C., & Loveland, T. R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, *122*, 66–74. https://doi.org/10.1016/j.rse.2011.08.024

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015a). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sensing of Environment*, *158*, 220–234. https://doi.org/10.1016/j.rse.2014.11.005

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015b). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment*, *170*, 121–132. https://doi.org/10.1016/j.rse.2015.09.004

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., & Campbell, L. B. (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. https://doi.org/10.1080/17538947.2016.1187673

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, *7*(11), 1417–1434. https://doi.org/10.1080/01431168608948945

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *80*, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

Jin, S., Dewitz, J., Danielson, P., Granneman, B., Costello, C., Smith, K., & Zhu, Z. (2023). National land cover database 2019: A new strategy for creating clean leaf-on and leaf-off Landsat composite images. *Journal of Remote Sensing*, *3*. https://doi.org/10.34133/remotesensing.0022

Ju, J., Roy, D. P., Shuai, Y., & Schaaf, C. (2010). Development of an approach for generation of temporally complete daily nadir MODIS reflectance time series. *Remote Sensing of Environment*, *114*(1), 1–20. https://doi.org/10.1016/j.rse.2009.05.022

Land Information Ontario. (2012). *Indian reserve*. Ontario GeoHub. https://geohub.lio.gov.on.ca/datasets/lio::indian-reserve/about

Land Information Ontario. (2015). *Provincial park regulated*. Ontario GeoHub. https://geohub.lio.gov.on.ca/datasets/lio::provincial-park-regulated/about

Man, C. D., Nguyen, T. T., Bui, H. Q., Lasko, K., & Nguyen, T. N. T. (2018). Improvement of land-cover classification over frequently cloud-covered areas using Landsat 8 time-series composites and an ensemble of supervised classifiers. *International Journal of Remote Sensing*, *39*(4), 1243–1255. https://doi.org/10.1080/01431161.2017.1399477

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., Gao, F., Kutler, J., & Lim, T. K. (2006). A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters*, *3*(1), 68–72. https://doi.org/10.1109/LGRS.2005.857030

Maxwell, S. K., & Sylvester, K. M. (2012). Identification of "ever-cropped" land (1984-2010) using Landsat annual maximum NDVI image composites: Southwestern Kansas case study. *Remote Sensing of Environment*, *121*, 186–195. https://doi.org/10.1016/j.rse.2012.01.022

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/bm.2012.031

Morresi, D., Marzano, R., Lingua, E., Motta, R., & Garbarino, M. (2022). Mapping burn severity in the western Italian Alps through phenologically coherent reflectance composites derived from Sentinel-2 imagery. *Remote Sensing of Environment*, *269*. https://doi.org/10.1016/j.rse.2021.112800

Nelson, K. J., & Steinwand, D. (2015). A Landsat data tiling and compositing approach optimized for change detection in the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, *81*(7), 573–586. https://doi.org/10.14358/PERS.81.7.573

Ojibway Nation of Saugeen. (n.d.). *About the Ojibway Nation of Saugeen*. Retrieved April 16, 2023, from https://ojibwaynationofsaugeen.ca/index.php/home/about-us

Ontario Ministry of Natural Resources and Forestry. (2019a). *Land information Ontario data description: MNRF road segment*. https://www.sdc.gov.on.ca/sites/MNRF-PublicDocs/EN/CMID/MNR%20Road%20Segment%20-%20Data%20Description.pdf

Ontario Ministry of Natural Resources and Forestry. (2019b). *Land information Ontario data description: ORN road net element*. https://www.sdc.gov.on.ca/sites/MNRF-PublicDocs/EN/CMID/ORN%20Road%20Net%20Element%20-%20Data%20Description.pdf

Potapov, P., Turubanova, S., & Hansen, M. C. (2011). Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia. *Remote Sensing of Environment*, *115*(2), 548–561. https://doi.org/10.1016/j.rse.2010.10.001

Potapov, P. V., Turubanova, S. A., Hansen, M. C., Adusei, B., Broich, M., Altstatt, A., Mane, L., & Justice, C. O. (2012). Quantifying forest cover loss in Democratic Republic of the

Congo, 2000-2010, with Landsat ETM+ data. *Remote Sensing of Environment*, *122*, 106–116. https://doi.org/10.1016/j.rse.2011.08.027

Python Software Foundation. (2001). *The Python language reference*. https://docs.python.org/3/reference/

Qiu, S., Zhu, Z., Olofsson, P., Woodcock, C. E., & Jin, S. (2023). Evaluation of Landsat image compositing algorithms. *Remote Sensing of Environment*, *285*. https://doi.org/10.1016/j.rse.2022.113375

Roy, D. P. (2000). The impact of misregistration upon composited wide field of view satellite data and implications for change detection. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(4 II), 2017–2032. https://doi.org/10.1109/36.851783

Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., Loveland, T. R., Vermote, E., & Zhang, C. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, *114*, 35–49. https://doi.org/10.1016/j.rse.2009.08.011

Savage, S. L., Lawrence, R. L., Squires, J. R., Holbrook, J. D., Olson, L. E., Braaten, J. D., & Cohen, W. B. (2018). Shifts in forest structure in northwest Montana from 1972 to 2015 using the Landsat archive from Multispectral Scanner to Operational Land Imager. *Forests*, *9*(157). https://doi.org/10.3390/f9040157

Statistics Canada. (2016, November 16). *Provinces/Territories, cartographic boundary file - 2016 census*. https://open.canada.ca/data/en/dataset/a883eb14-0c0e-45c4-b8c4-b54c4a819edb

Thompson, S. D., Nelson, T. A., White, J. C., & Wulder, M. A. (2015). Mapping Dominant Tree Species over Large Forested Areas Using Landsat Best-Available-Pixel Image Composites. *Canadian Journal of Remote Sensing*, *41*, 203–218. https://doi.org/10.1080/07038992.2015.1065708

United States Census Bureau. (2018). *Cartographic boundary files - Shapefile*. https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html

United States Geological Survey. (2018). *LSDS-286 Landsat Multispectral Scanner (MSS) Level 1 (L1) Data Format Control Book (DFCB)*. https://earth.esa.int/eogateway/documents/20142/37627/LSDS-286_LandsatMSS-Level1_DFCB-v11.pdf

United States Geological Survey. (2020a). *Landsat 4-7 Collection 1 (C1) surface reflectance (LEDAPS) product guide*. https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1370_L4-7_C1-SurfaceReflectance-LEDAPS_ProductGuide-v3.pdf

United States Geological Survey. (2020b). *Landsat 8 Collection 1 (C1) Land Surface Reflectance Code (LaSRC) product guide*. https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1368_L8_C1-LandSurfaceReflectanceCode-LASRC_ProductGuide-v3.pdf

United States Geological Survey. (2020c). *Landsat Multispectral Scanner (MSS) Collection 2 (C2) Level 1 (L1) Data Format Control Book (DFCB)*. https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1416_LandsatMSS-C2-L1-DFCB-v3.pdf

Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, *185*, 46–56. https://doi.org/10.1016/j.rse.2016.04.008

Vermote, E., Roger, J. C., Franch, B., & Skakun, S. (2018). LaSRC (Land Surface Reflectance Code): Overview, application and validation using MODIS, VIIRS, LANDSAT and Sentinel 2 data's. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 8173–8176. https://doi.org/10.1109/IGARSS.2018.8517622

Wang, J. (1993). Linda — A system for automated linear feature detection and analysis. *Canadian Journal of Remote Sensing*, *19*(1), 9–21. https://doi.org/10.1080/07038992.1993.10855146

White, J. C., Wulder, M. A., Hermosilla, T., Coops, N. C., & Hobart, G. W. (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. https://doi.org/10.1016/j.rse.2017.03.035

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., Coops, N. C., Hall, R. J., Hostert, P., Dyk, A., & Guindon, L. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, *40*(3), 192–212. https://doi.org/10.1080/07038992.2014.945827

Wolfe, R., Masek, J., Saleous, N., Code, F. H., & Goddard, N. (2004). *LEDAPS: Mapping North American disturbance from the Landsat record*.

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., Vogelmann, J., Wulder, M. A., & Wynne, R. (2008). Free access to Landsat imagery. *Science*, *320*(5879), 1011. https://doi.org/10.1126/science.320.5879.1011a

Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R., & Evangelista, P. H. (2017). A survival guide to Landsat preprocessing. *Ecology*, *98*(4), 920–932. https://doi.org/10.1002/ecy.1730

Zaidi, S. M., Akbari, A., Samah, A. A., Kong, N. S., & Gisen, J. I. A. (2017). Landsat-5 time series analysis for land use/land cover change detection using NDVI and semi-supervised

classification techniques. *Polish Journal of Environmental Studies*, *26*(6), 2833–2840. https://doi.org/10.15244/pjoes/68878

Zhang, Q., Wang, J., Peng, X., Gong, P., & Shi, P. (2002). Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. *International Journal of Remote Sensing*, *23*(15), 3057–3078. https://doi.org/10.1080/01431160110104728

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94. https://doi.org/10.1016/j.rse.2011.10.028

Chapter 4: Forest Road Extraction, Year of Construction Assignment, and Logic Rule Application

4.1 Introduction and Objectives

Forest ecosystems provide resources and services that are essential for all living beings, including timber and non-timber forestry production, water resource protection, a range of habitats, biodiversity, and recreational features (Griffiths et al., 2014). Additionally, the role of forests in global climate trends, surface radiation regulation, and carbon sequestration have increasing importance as anthropogenic climate change intensifies (J. Chen et al., 2010; Griffiths et al., 2014; L. Liu et al., 2013). Forest resources and ecosystem services are dependent on forest type, condition, species composition, and management activities (Griffiths et al., 2014).

Canada's boreal forest is relatively intact but has undergone a variety of forest disturbances through time. In northern ecosystems, wildfires are the predominate concern, whereas anthropogenic activity significantly impacts settled areas and regions where forest productivity is high (Hermosilla et al., 2015b). Stand replacing forest disturbances, including urbanization, road construction, industrial activity, harvesting, and wildfires, may be short- or long-term in modification of land cover. Non-stand replacing forest disturbances are of lower magnitude and usually manifest as a variation in land cover as opposed to a full land cover change (Ahmed et al., 2017; Hermosilla et al., 2015b). Of the 300 million hectares (ha) of forested areas in the Canadian boreal zone, approximately 2 million ha are impacted by wildfires, 2 million ha are affected by harvesting related activities and 45, 000 ha are deforested annually (Hermosilla et al., 2015b).

Ontario, one of the provinces of Canada, has immense forest resources to manage. Ontario is 107 520 421 ha with 66% of the total land area including forest resources. Forests account for 56 094 000 ha whereas treed wetlands account for the remaining 14 972 000 ha. The Ministry of Northern Development, Mines, Natural Resources and Forestry (MNDMNRF) estimates that Ontario has 85 billion trees and more than 7.1 billion cubic meters of tree volume. With such a vast area and number of resources, maintaining accurate monitoring of land disturbance datasets is a challenge (Bilyk et al., 2021). The lengthy Landsat data archive allows researchers to detect stand and non-stand replacing changes through the examination of spectral trends in a spatially, temporally, and categorically systematic manner. These changes are particularly important in large, forested regions like Ontario, which have managed and unmanaged forest regions with a variety of stewardship organizations (Hermosilla et al., 2015a, 2015b). Understanding historic types, rates, trends, and patterns of ecosystem change across large areas with high spatial detail is essential for effective forest monitoring and sustainable management (Ahmed et al., 2017; Griffiths et al., 2014; Hermosilla et al., 2015b). Time series analyses based on annual remotely sensed imagery composites can provide rich information to support these endeavors (White et al., 2014).

Forest access roads are one type of disturbance well suited for monitoring through remotely sensed imagery. Commercial logging operations are initiated with the construction of roads to access forest stands, haul logs, and transport resources to processing facilities (Bowman et al., 2010; Mihell & Hunt, 2011). Forest access roads are generally considered a long-term forest stand replacing change as vegetation is removed for the construction of roads, highways, and access corridors. Due to the long linear shape of roadways, they can be separated from other anthropogenic land use changes like industrial development and urbanization (Hermosilla et al.,

2015b). Road delineation and extraction from remotely sensed images is both a challenging and important research endeavor. Road information is crucial for transportation, traffic management, planning, global positioning system (GPS) navigation, and map updating (Kahraman et al., 2018; Qiaoping & Couloigner, 2004; J. Wang, 1993). However, manual interpretation of imagery is resource and time intensive. As a result, a variety of automatic and semi-automatic methods have been developed (Bae et al., 2015; Bakhtiari et al., 2017; G. Cheng et al., 2017). A semi-automatic algorithm is proposed to extract forest roads and attribute year of construction for historic forest roads in Ontario's portion of the boreal forest between 1974 and 2019.

There are many approaches to identifying forest disturbances including Hermosilla et al. (2015b), C. Huang et al. (2010), Kennedy et al. (2010), and L. Liu et al. (2013). C. Huang et al. (2010) developed an automated algorithm called vegetation change tracker (VCT) which identified disturbance and recovery events using Landsat Time Series (LTS) stacks. Their approach first identifies persistent or changed forest, water, or non-forest pixels. Then, disturbance type, year of occurrence, and magnitude information are identified. Forest disturbance classes in their study sites included urbanization, logging and harvest, ice and wind damaged, fire, and fuel treatment. Although the overall accuracy and kappa coefficient ranged between 0.77 to 0.85 and 0.43 to 0.76, respectively, for their test locations, there is no specific forest road class. Kennedy et al. (2010) developed LandTrendr which identifies short- and long-term disturbance and recovery events in LTS stacks. In their analysis of 388 forested plots, disturbances identified included both harvest and fire as well as insects and pathogens. Although their disturbance identification was strong, there was not a focus on forest roads as a disturbance type. Similarly, L. Liu et al. (2013) developed an algorithm using LTS stacks to identify afforestation and deforestation in Yulin District, China between 1974 and 2012. Although their

methodology resulted in an overall accuracy of 89.1 % and kappa coefficient of 0.858, there is no specific forest road data available. Hermosilla et al. (2015b) developed a methodology to identify stand and non-stand replacing forest changes using LTS data between 1984 and 2012. Their approach uses breakpoint analysis to detect and date changes and a Random Forest (RF) classifier to attribute change type. Their change types include fire, harvesting, road and non-stand replacing changes. These changes are detected and dated with an accuracy 91.1%; however, roads were the most challenging to attribute with less than an 86% user accuracy. This work seeks to enhance forest disturbance monitoring techniques by extending forest road extraction and year of construction attribution to the years 1974 and 2019. This extension includes the MSS (Multispectral Scanner) and Operational Land Imager (OLI) sensors. In addition, a new semi-automatic forest road extraction technique has been proposed with the inclusion of four logic rules to enhance extraction and change detection accuracy.

The proposed approach uses the Jenks Optimization clustering technique to complete an unsupervised classification of Best Available Pixel (BAP) composite Landsat imagery between 1974 and 2019. A reference road mask, developed from known forest roads, is then applied, followed by the roads being converted to a vector format. Four logic rules are used to enhance the vector road network before post-classification change detection (PCCD) is conducted to determine the year of construction for each of the road segments within the network. The logic rules are then applied again to further refine the connectivity and year of construction attributes. The results from this algorithm can be used to update records of unknown years of road construction and provide critical information to forestry managers and stakeholders.

This chapter will provide a broad overview of relevant road extraction, change detection, and post-processing approaches. This overview will be followed by a detailed description of the

methodology of the algorithm including the study site, candidate images, image processing, and accuracy assessment. The results of the algorithm will be presented followed by a discussion of the significance and concluding remarks.

4.2 Literature Review

4.2.1 Forest Roads in Ontario's Boreal Forest

There are many linear disturbances within forested ecosystems (forest roads, seismic lines, roads, pipelines, and powerline rights of way). Forest or logging roads are a type of linear disturbance broadly defined as paved or unpaved roads used to allow wheeled vehicles to transport logs and equipment to and from the forest (Bocking et al., 2017). Forest access roads on Ontario Crown Land are managed under the Ontario Policy Framework for Sustainable Forests (OPFSF) (Ontario Ministry of Natural Resources and Forestry, 2020). Under the Public Lands Act, R.S.O. 1990, c. P.43, it states that forest roads (including the associated bridges, culverts, water crossing and related rights of way) should be reasonably traversable using a four-wheel drive truck. These roads exclude the King's Highway and any secondary highway or industrial route defined under the Public Transportation and Highway Improvement Act, R.S.O. 1990, c. P.50 (Ontario Ministry of Natural Resources and Forestry, 2020). The Ontario Ministry of Natural Resources and Forestry (MNRF) currently denote three forest road classes: primary (access to management unit), branch roads (branch off primary roads or other branch roads to provide access to management units), and operational roads (transportation for short term harvest, maintenance, or renewal purposes) (Ontario Ministry of Natural Resources and Forestry, 2020). In the Ontario context, skid trails are excluded from the road definition and official road network. These trails are used to drag logs from their felling site. As the trails are often designed for tracked vehicles (as opposed to wheeled vehicles), they are often narrower than other forest roads. The width of skid trails often prevents the development of a continuous canopy opening

(Kleinschroth, Gourlet-Fleury, et al., 2016; Kleinschroth & Healey, 2017; Minn et al., 2022). The small width makes it difficult to effectively monitor skid trails from moderate resolution remotely sensed imagery. As a result, only primary, branch, and operational roads as defined by the MNRF are analyzed within this study.

4.2.2 Roads in Remotely Sensed Imagery

Forest access roads in Ontario follow strict construction and maintenance requirements that determine their appearance in the natural environment. However, the conceptualization of roads in remotely sensed imagery differs from roads on the ground. In general, roads in remotely sensed images are characterized as bright elongated geometric features (linear) with slowly changing pixel values (J. Chen et al., 2014; Y. Li et al., 2009; J. Wang, 1993; W. Wang et al., 2016). The intensity of contrast between a road segment and the background may slowly vary across space and time. This change is influenced by the spectral range and sensitivity of the imagery used, scale of the object relative to image resolution, image complexity, and noise interference from buildings, shadows, or other road obstructions, as well as road materials. When considering the topological features of a road network, roads will include intersections. Typical road junctions include “T” intersections, “Y” intersections, “+” intersections as well as “O” roundabouts (Cardim et al., 2018; Qiaoping & Couloigner, 2004; W. Wang et al., 2016). These interactions are often referred to as junctions or nodes. Roads are not suddenly interrupted but may end with a road termination (Qiaoping & Couloigner, 2004; W. Wang et al., 2016).

As a result of the challenges associated with extracting roads from remotely sensed imagery (Cardim et al., 2018; G. Cheng et al., 2017; W. B. Cohen et al., 2010; Kahraman et al., 2018; R. Li & Cao, 2018; L. Liu & Lim, 2016; T. Zhou et al., 2019), roads can be divided into two classes: salient and non-salient roads. Salient roads are roads within an image that have a distinctive and consistent appearance. These roads are consistently not impacted by shadows or

occlusions from trees or buildings (Özkaya, 2012). Salient roads will not vary significantly in their spectral and spatial characteristics along their length (J. Wang et al., 1992). In contrast, non-salient roads are more difficult to detect as there is greater heterogeneity because of occlusion or shadow from other earth surface features. Generally, this heterogeneity will lead to poorer or incomplete detection or delineation (Özkaya, 2012; Soni et al., 2018). The methodology propose din this thesis was designed to capture salient roads through road extraction but also address non-salient roads through the four post-processing logic rules. These logic rules have been adapted for both binary road segments (road and non-road) as well as the final road network with year of construction attributes.

Researchers have proposed many road network extraction algorithms since the advent of remotely sensed imagery. However, there are no approaches that are suitable for all types of roads (Z. Chen et al., 2014; W. Wang et al., 2016). The performance of each methodology is strongly dependent on the type of road to be extracted (Y. Li et al., 2009; L. Liu & Lim, 2016; Ünsalan & Boyer, 2011). As a result, most methods focus on a specific road and road model conceptualization (Cardim et al., 2018). This methodology is designed specifically for the Ontario Forest Road framework.

4.2.3 Road Extraction

There are many approaches for road extraction presented in the literature. These algorithms include road tracking, mathematical morphology, image processing (snake and level set), dynamic programming, and classification. The first road extraction algorithms proposed were road finding and tracking approaches (J. Cheng et al., 2012; McKeown & Denlinger, 1988; Nevatia & Babu, 1980; Vosselman & De Knecht, 1995; J. Zhou et al., 2006; X. Zhou & Mahmassani, 2007). These approaches generally use seed points, and the roads are traced sequentially through the image using local pixel information (Dal Poz et al., 2006; Mena, 2003).

Edge linkers (edge operations and rule-based edge linking), correlation trackers (pattern based and often paired with edge linkers) and region-based followers (assumes consistent feature intensity and paired with correlation trackers) are common types of road finding and tracking approaches (Barzohar et al., 1996; McKeown & Denlinger, 1988). These approaches are criticized within the literature for computational cost and complexity as well as reliance on seed points for initialization (Soni et al., 2018).

Mathematical morphology is another approach that was adopted early in the road extraction literature (Bakhtiari et al., 2017; Besbes & Benazza-Benyahia, 2014; Kahraman et al., 2018; Shi et al., 2014; Valero et al., 2010; H. Wang et al., 2016; J. Wang, 1993; C. Zhang et al., 1999). This method is based on set theory developed by Matheron (1974) and Serra (1982) (Mena, 2003; Soni et al., 2018). Set theory is often used in image processing for geometrical shapes and these approaches are often combined with other approaches or used to enhance road features during post-processing (Bakhtiari et al., 2017; J. Wang, 1993). Although mathematical morphology techniques do not rely on seed points and have reduced computational cost, they often rely on fixed shape structure elements. This reliance often prevents the algorithm from being flexible enough to manage road curves (Kahraman et al., 2018; Shi et al., 2014).

Snake (Amini, 2009; J. Cohen, 1960; Jeon et al., 2002; Kass & Witkin, 1988; Nakaguro et al., 2011; Saati & Amini, 2017; Soni et al., 2018; Xu & Prince, 1998; Zafiropoulos & Schenk, 1998) and level set approaches are also common within the road extraction literature (Abdollahi et al., 2019; Cai et al., 2019; Keaton & Brokish, 2002; Y. Liu et al., 2014; Osher & Sethian, 1988; Rajeswari et al., 2011; Vese & Chan, 2002; H. Wang et al., 2016). These techniques are based on the optimization of energy functions and curves (Soni et al., 2018). A snake is an energy-minimizing spline algorithm that is guided by external constraints and image forces that

guide it towards lines and edges. The curve or contour deforms from an arbitrary initial position (user-defined or automatic) towards the target feature and creates a best fit through minimizing an energy function (Amini, 2009; L. D. Cohen, 1991). Level set uses an implicit energy function within a high number of dimensions in a fixed rectangular grid to model a region's contours. Curves are expressed as the level set of higher-dimensional continuous surfaces. Solving the curve function involves partial differential equations of zero level set (Y. Liu et al., 2014). Snake and level set methods overcome the limitations of early sequential road extraction approaches which are highly susceptible to propagating errors. They are not sensitive to initialization points and noise; however, there is no topological flexibility in the classical snake model. The level set method addresses the lack of topological flexibility but does not incorporate external forces of geometric constraint and results in image noise sensitivity (Kass & Witkin, 1988; Nakaguro et al., 2011; Zafiropoulos & Schenk, 1998). Although researchers have continued to build on these approaches, there are still limitations related to the extraction of weak boundaries and heterogeneous segments (Abdollahi et al., 2019; Cai et al., 2019; Vese & Chan, 2002).

Region based methods extract the road network based on image classification. Image classification refers to the division of an image into non-overlapping regions based on similar attributes (intensity, colour, shape, texture, etc.) (Balaji, 2014). There are two main types of classification: supervised and unsupervised. Supervised methods include Support Vector Machine (SVM) (Han et al., 2018; Otukei & Blaschke, 2010; Sanchez-Hernandez et al., 2007; Song & Civco, 2004; Ziems et al., 2012), Maximum Likelihood Classification (MLC) (Ali et al., 2018; Mosammam et al., 2017; Otukei & Blaschke, 2010), and Random Fores (RF) (Colditz et al., 2012; Haas & Ban, 2018; Karlson et al., 2015; Räsänen et al., 2013). Common unsupervised classifiers are k-means and its variant Iterative Self-Organizing Data Analysis (ISODATA)

(Markogianni & Dimitriou, 2016; Naikoo et al., 2020; Patra et al., 2018), Jenks Optimization (Hou et al., 2022; Jenks, 1967; Jenks & Caspall, 1971; Sadeghfam et al., 2016) and Mean Shift Classifier (Friedman et al., 2013; Vlachopoulos et al., 2020). Region-based approaches assume that there is a consistent relationship between spectral response within the imagery and the earth surface feature or phenomenon. Despite this assumption, there is often a complex relationship between scene complexity, scale, sensor, and earth surface features which can reduce the accuracy and applicability of these methods (Belgiu & Drăgu, 2016). These methods can be paired with post-processing approaches to address these limitations.

4.2.4 Change Detection

To determine the year of road construction, after the roads are extracted, change detection analysis will be employed. Change detection is a tool for determining the extent and type of land use land cover (LULC) change through the comparison of images acquired in the same area through one or more dates (Colditz et al., 2012; Fung, 1992; Minu & Shetty, 2015). In other words, it is "the process of identifying differences in the state of an object or phenomenon by observing it at different times" (A. Singh, 1989, p. 989). Remotely sensed imagery has regular coverage, a wide selection of spatial and spectral resolutions, digital formats suitable for quick computation, short collection intervals, and consistent image quality for assessing LULC (Afify, 2011; Hussain et al., 2013). These characteristics make this data an efficient and cost-effective method for the inventory and monitoring of the type and extent of earth surface changes (Afify, 2011).

Map algebra (MA) (image differencing and image rationing), principal component analysis (PCA), change vector analysis (CVA), change point detection and analysis (CPD), and PCCD are common change detection methods. MA approaches use mathematical operators (subtraction and division) on images that are spatially aligned. The resultant image displays the

type and magnitude of the change on the ground. A threshold value is used to separate meaningful change from image noise (Afify, 2011; Minu & Shetty, 2015; Sundaresan et al., 2007; Tewkesbury et al., 2015). This method is sensitive to sensor noise, variation in scene illumination, is often global, and does not consider local consistency or spatial autocorrelation (Hussain et al., 2013; Radke et al., 2005; Rosin, 2002).

PCA transformations are all based on eigenvalue decomposition of the covariance (or correlation) matrix of the image bands under investigation. This process converts the correlated spectral bands into a set of linearly uncorrelated bands, referred to as principal components (PCs) (Dadon et al., 2019; Koutsias et al., 2009; Rodarmel & Shan, 2002). In change detection applications, higher order PCs will have unchanged information whereas lower order PCs will have change information images (Deng et al., 2008; Fung & LeDrew, 1987; P. Gong, 1993; Koutsias et al., 2009). PCA change detection does not provide any from-to class change information and is heavily scene dependent (Dadon et al., 2019; Fung, 1992; Koutsias et al., 2009), making it hard to apply to large scale time series datasets.

CVA uses pixel-wise radiometric comparison and represents change as a vector in multi-dimensional spectral space (Du et al., 2020; Polykretis et al., 2020). Although CVA provides magnitude information as well as from-to change information, this technique is sensitive to the impact of varying atmospheric conditions, solar angle, soil moisture, and phenology (He et al., 2011). Different scenes, applications, and times of years result in varying magnitude for different change types (He et al., 2011; Xiao et al., 2021). These differences may increase computational complexity when applied to large scale time series datasets.

CPD seeks to identify changes in the model of a signal or time series dataset. It can be assumed that the whole non-stationary signal can be broken into multiple stationary segments

where the breakpoints represent change points of varying magnitude (Truong et al., 2018). Although CPD provides rich information and the ability to calculate a variety of additional statistics, the effectiveness declines as time series noise increases and change magnitude decreases. Additionally, change points that happen near the start or end of a time series may be subject to the margin effect and will not be interpreted as change points (Rahman et al., 2013; van den Burg & Williams, 2020).

PCCD enables each image to be classified into the relevant land cover classes based on its own characteristics and data distribution before change is evaluated between the images. Using a post-classification approach provides from-to change information and reduces the impacts of using images from multiple dates and sensors (Jakubauskas et al., 1990; Tewkesbury et al., 2015). This technique can often account for issues that may arise when using imagery acquired in different seasons or with different phenology, soil moisture, and atmospheric conditions. Despite the benefits of PCCD, it is criticized for compounding errors from the original classified images (A. Singh, 1989; Tewkesbury et al., 2015; Q. Zhang et al., 2002). PCCD can be preceded by either supervised or unsupervised classification algorithms, which have a variety of characteristics which impact their accuracy (Hussain et al., 2013). Although additional errors may be introduced, for a large LTS dataset, this method may strongly reduce computational complexity and cost associated with sensor and scene variations. Due to the use of multiple sensors over a 45-year period, PCCD was selected for this algorithm.

4.2.5 Post-Processing and Logic Rules

After the completion of road extraction and change detection, non-salient roads can remain unidentified and disconnected from the main road network because of occlusions. In addition, over or under detection can occur due to image quality, quantity, or extraction effectiveness. These missing, erroneous, or disconnected features reduce the integrity of the road

network (Fan et al., 2020). Road network shape and connectivity are defining components deemed essential for decision making related to the use of extracted networks. By using the geometric, radiation or pixel brightness, topological, and contextual features of road segments and networks, researchers can post-process their extracted roads to overcome the effect of non-salient road features. In early road research, mathematical morphology was a common post-processing approach to manage non-salient linear features. For example, L. W. Li et al. (2005) used morphological operators to refine binary road centerlines (road and non-road). Although this approach can reduce the impact of small road gaps, when there are larger distances between disconnected road segments, significant road shape deformation can be introduced. Other approaches leverage further linear feature properties to enhance gap filling procedures. Arrighi and Soille (1999) join disconnected contour lines with a propagation function to determine the extremities of the disconnected features, the Euclidean distance between the two extremities, and the differences between their directions to bridge long road gaps. In 2012, Samet and Hancer use the local and geometric properties of broken contour lines to automatically bridge gaps and enhance connectivity. Fan et al. (2020) detect breakpoints from a thinned road skeleton and removes short, isolated segments through mathematical morphology operations. The breakpoints are then clustered and joined through a fitted and filled polynomial curve. This approach can enhance connectivity, but as road complexity increases, the algorithm struggles with multiple matches for road breakpoints, resulting in unaddressed gaps. In addition to these approaches, knowledge based post-processing approaches are also commonly applied to initial extracted road networks. Researchers will use properties from their context specific road conceptualization and develop rules to modify the road network. J. Wang (1993) removes all disconnected road features that are less than a user defined length. Chaudhuri et al. (2012) used road length, number

of intersections, and proximity to neighboring segments to remove disconnected features and small non-road linear features. Knowledge based rules are also common in other remote sensing approaches. For example, H. Liu and Zhou (2004) developed six rules for post-classification comparison of the urban fringe of Beijing, China. These rules are “IF/THEN” statements that address inconsistencies in change type. Savage et al. (2018) employ eight rules to address temporal inconsistencies through local, ecosystem, and successional knowledge. Due to the presence of non-salient roads, logic rules are used as a post-processing step in this algorithm (Please see Section 4.3.8).

4.3 Methodology

4.3.1 Study Site

The study site for this analysis is a 3 405 squared kilometer region of northern Ontario, Canada within the Landsat scenes World Reference System-1 (WRS-1) path 029 and row 025 as well as WRS-2 path 027 and row 025 (Figure 8). This location is northeast of Dryden and northwest of Thunder Bay, Ontario. The town of Savant Lake and southern part of Ojibway Nation of Saugeen Reserve are contained within the study boundary. The MNRF has classified this region as part of the Lake Nipigon or 3W Ecoregion. For more information on the study site, please see Chapter 3 Section 3.1.

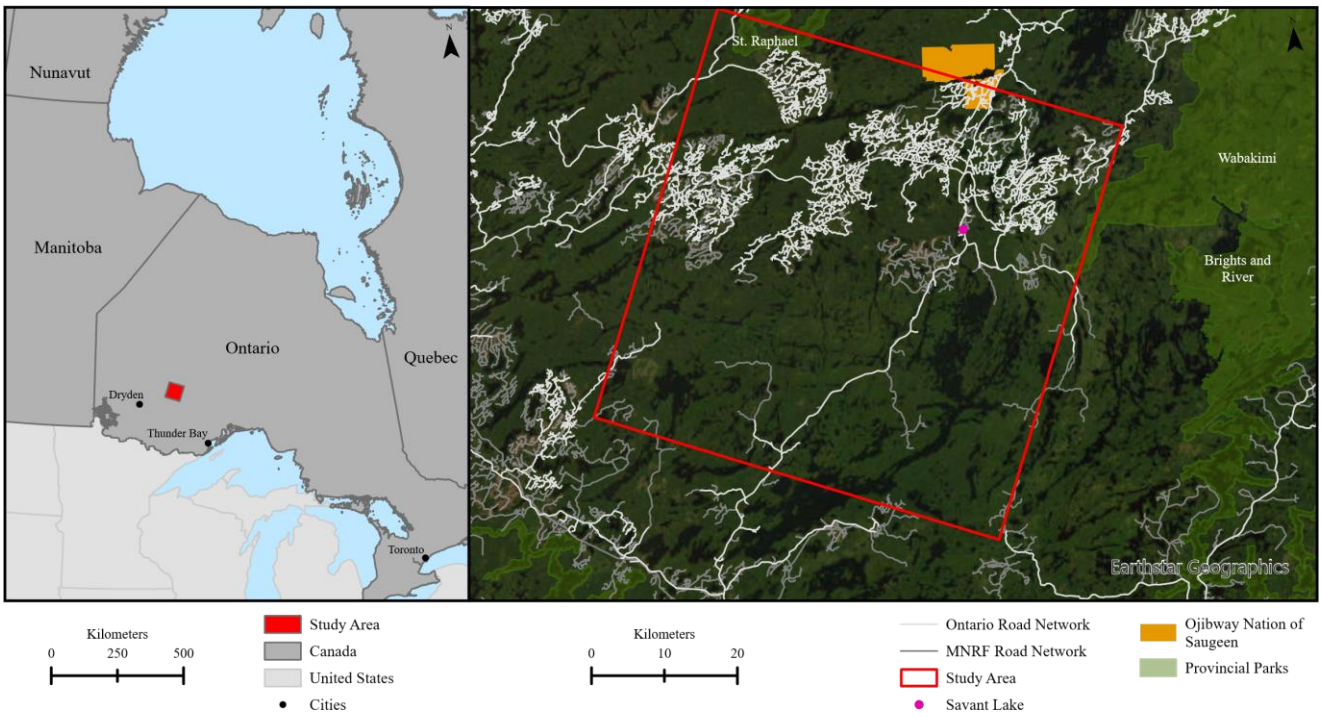

Figure 8: Map of Lake Savant study site (Land Information Ontario, 2012, 2015; Ontario Ministry of Natural Resources and Forestry, 2019b, 2019a; Statistics Canada, 2016; United States Census Bureau, 2018)

4.3.2 Candidate Images

Two subsets of images were selected from images downloaded from the USGS. These images were Landsat Level 1 Tier 1 Data products. Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and OLI images use surface reflectance with images being processed through the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (TM and ETM+) or Land Surface Reflectance Code (LaSRC) (OLI) programs. MSS data is pre-processed through the USGS, but no surface reflectance data is currently available. As a result, Digital Numbers (DNs) are used. For more information on the selected imagery, please see Chapter 3 Section 3.2. The first subset is used to test annual road level analyses (pre-change detection) whereas the second subset is used to test full network level analyses (post-change detection).

The first subset, also used in Chapter 3, includes 48 images to create BAP composites for the years 1983, 1987, 1996, 2005, 2009, and 2019. Please see Figure 9 for the distribution of

acquisition day of year (DOY) and sensors for the subset. The test years have the following number of quality images per year: 1983: 2, 1987: 6, 1996: 9, 2005: 7, 2009: 6, 2019: 14.

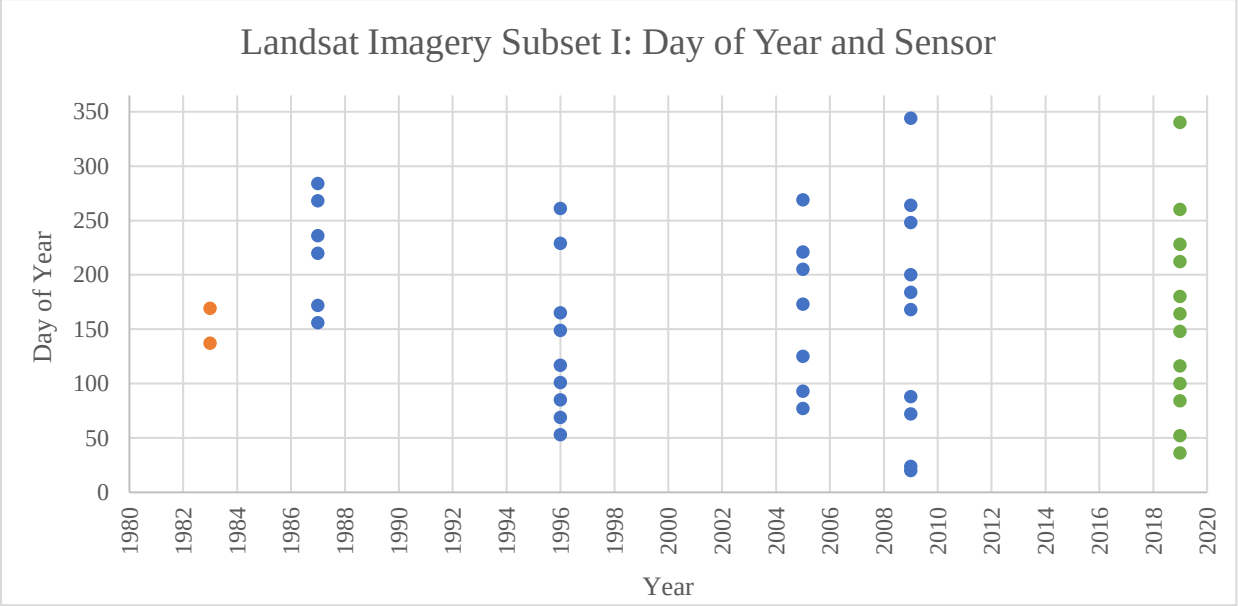

Figure 9: Day of year for Landsat image composite testing subset. Orange, blue, and green points represent MSS - Satellite 4, TM – Satellite 5 and OLI – Satellite 8, respectively.

For the second subset, there are 364 images used to generate BAP composites for the dates between 1974 and 2019. The following years are excluded: 1972, 1973, 1982, 1986 and 1987. Images from 1972 and 1973 were omitted due to poor data quality at the start of the Landsat Mission. Images from 1982, 1986 and 1987 were omitted due to consistent occlusions over the study site. There are fewer MSS images as clouds are often not fully masked due to insufficient cloud masking algorithms. This incomplete masking often leaves halos in the image and can lead to compounding errors during the compositing process. In the years the TM, ETM+ and OLI sensors are used, a greater number of input images is advantageous; for the MSS data, including very few images with very low cloud coverage is critical. Please see Figure 10 for the distribution of acquisition day of year and sensors for the subset. Please see Figure 11 for the number of images per year. Figure 11 includes the number of images in the input dataset as well

as the total number of images considered the final BAP composites after the algorithm removed unsuitable imagery.

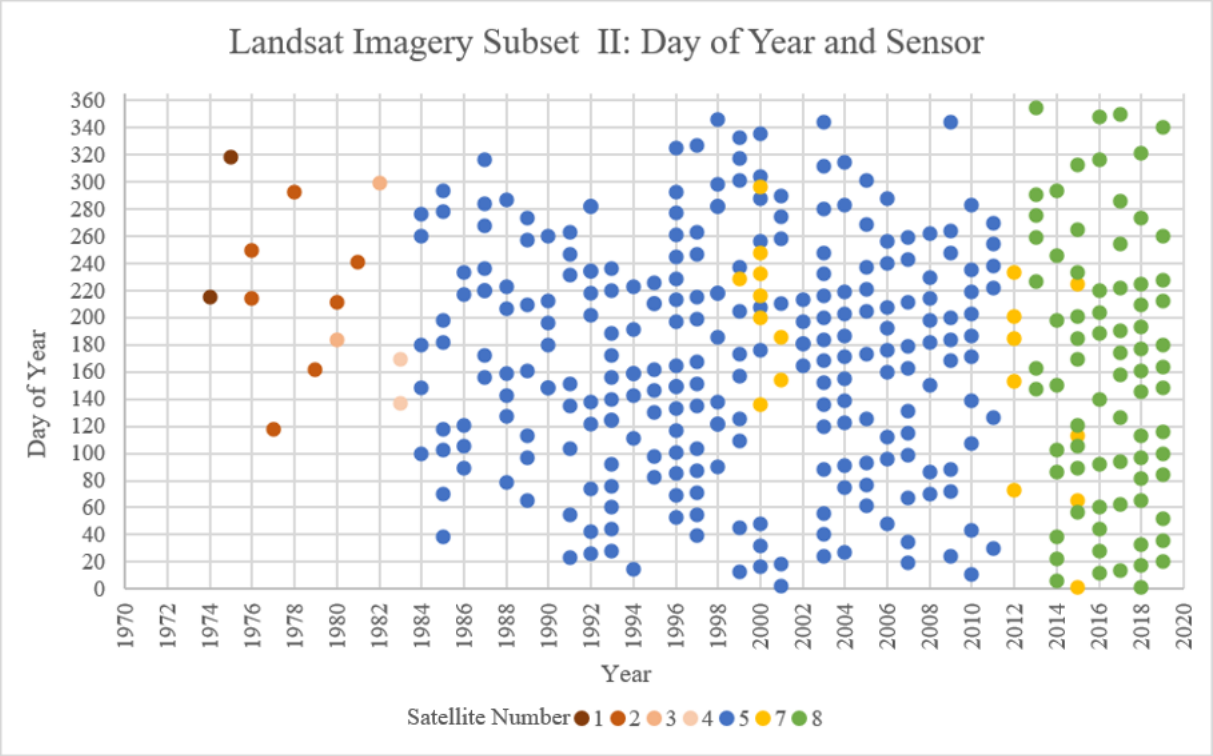

Figure 10: Day of year for Landsat image composite testing subset. Orange, blue, yellow, and green points represent MSS – Satellite1-5, TM – Satellite 5, ETM+ – Satellite 7, and OLI – Satellite 8, respectively.

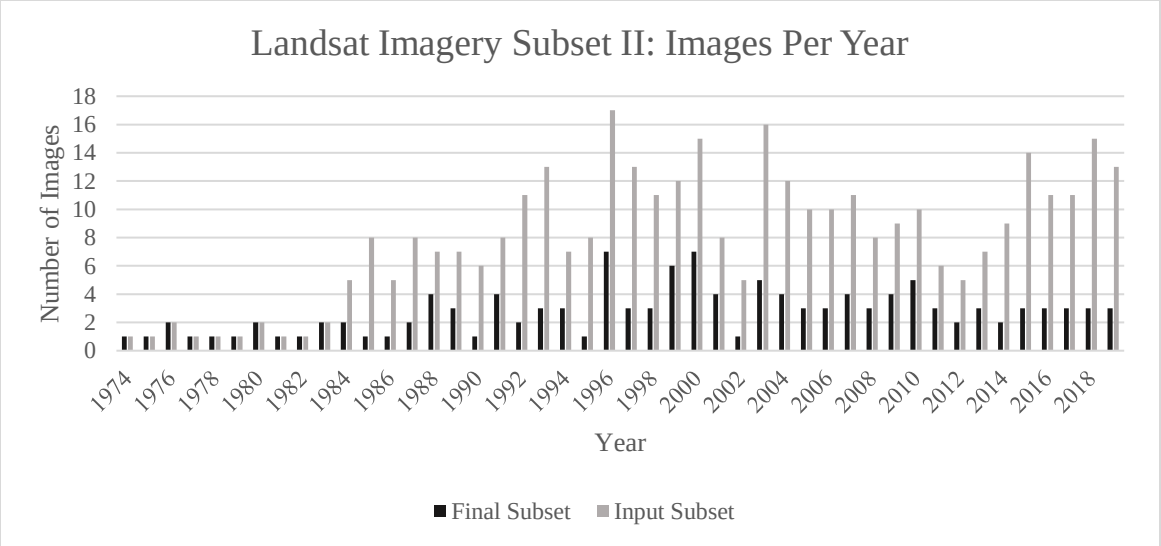

Figure 11: Number of images per year for Landsat Imagery Subset II. The input subset represents all images input into the algorithm whereas the final subset values represent all images retained after all poor-quality images were removed during the compositing process.

4.3.3 Reference and Ground Truth Data

To enhance road extraction, a mask of known roads in the study area is generated from reference data. After the construction date of each road segment is known, these segments can be used as seed points to find the location of previously undocumented roads in future research. The road mask is developed with two road networks: MNRF Road Segment and Ontario Road Network (ORN) Segment with Address. These roads were accessed through Ontario GeoHub and Land Information Ontario (https://geohub.lio.gov.on.ca/). The MNRF Road Segments are all roads under the jurisdiction of the MNRF. This includes roads that cross land owned by the MNRF, roads that are the responsibility of the MNRF to maintain, or if the road is maintained by another agency through agreement, but responsibility of the road will be returned to the MNRF if the agreement is terminated. All line segments in the dataset represent a single continuous object. To exist within this road network, roads must have existed for a sufficient period of time and have supported the use of a conventional four-wheel street legal vehicle. Specifically excluded from this dataset are paths for off-road vehicles (all-terrain vehicle trails) and temporary roads for forest or forest fire management (skidder trails and rehabilitated fire trails). When the road can no longer be traversed by a four-wheel street legal vehicle, the status of the segment is changed to decommissioned, but the feature is not removed from the database (Ontario Ministry of Natural Resources, 2019). Only the spatial information from this file was used. No attributes related to year of construction or decommission date were used for this study as these attributes are not available for all segments and the accuracy of some records is questionable (as noted by the owners of the dataset). The intended use of this algorithm is to update these records.

ORN Segment with Address is a database which includes over 250 000 kilometers (kms) of roads including municipal roads, provincial highways, recreational roads, and some resource roads. The ORN is the official reference source for roads data for the Government of Canada

(Ontario Ministry of Natural Resources and Forestry, 2019c). This dataset is derived from the ORN Road Net Element dataset which is a linear referencing system database (Ontario Ministry of Natural Resources and Forestry, 2019b).

To assess the results of the road extraction approach, ground truth data is required. Through collaboration with the MNRF, it was indicated that no accurate ground truth data for the historic forestry roads existed for this region. This is a common challenge for studies looking to validate the results of LTS detection of forest change over large areas (W. B. Cohen et al., 2010). To generate accuracy assessment points for the year of construction, 660 random points and 2150 points were created within the mask of known roads for the pre-1984 and post-1984 time periods, respectively. This division reflects the slightly different processing procedures and spatial resolution between the MSS and later sensors. The author manually confirmed the year of construction using the input Landsat images for the algorithm. Ideally, each year would have 50 samples (Congalton, 1991). Generally, the points for each year are quite clustered as forestry operations for that year are centered around a specific region. In some cases, points were added to better reflect this clustering. Some years, there are few or no roads built; as a result, these years have no or few points. For a road to be identified, it had to be visible in the imagery by the analyst and be connected to the main network. The year of construction was the first year the road was identified within the periods of 1972 to 1984 or 1985 to 2019, reflecting the two accuracy datasets.

4.3.4 Algorithm Overview

The proposed semi-automatic algorithm for forest road extraction and change detection using BAP composites was written in Python 3.7.0 (Python Software Foundation, 2001) using the arcpy site package for ArcGIS Pro 3.0 (Environmental Systems Research Institute, 2022). The algorithm can be divided into four main sections: (1) Pre-Processing, (2) BAP Compositing,

(3) Road Extraction and Change Detection and (4) Logic Rule Execution (binary and change year attribution). In this chapter (Chapter 4), the road extraction, change detection and logic rule execution will be discussed. The BAP compositing results as well as BAP approach testing are presented in Chapter 3.

4.3.5 Data Pre-processing and BAP Composites

Landsat imagery from the Landsat MSS, TM, ETM+ and OLI sensors are pre-processed for use in the BAP composite approach. The red band was used for the MSS data whereas the Shortwave Infrared-2 (SWIR-2) band was used for the TM, ETM+, and OLI sensors. These images are radiometrically and geometrically corrected projected surface reflectance (TM, ETM+ and OLI) and DN (MSS) products. Images where the full scene has cloud coverage over 70% are removed prior to pre-processing due to potential limitations to Ground Control Points (GCPs). Retained images have cloud, cloud shadow, water, snow, and ice classes masked where available. Saturated pixels and invalid pixels are also masked as these pixels cannot be used in further analysis. Opacity and aerosol quality information is pre-processed for use in the generation of BAP composite scores. These pre-processing steps ensure that any change information derived from the images can be attributed to true change on the ground instead of atmospheric, geometric, or sensor differences between multiple images (Afify, 2011). The data is then clipped to the study site defined by the user to reduce processing time. For further information on the imagery pre-processing, please see Chapter 3 Section 3.5.

BAP composites allow for a singular image with the best pixel in each location as defined by user specifications to be composed from multiple input images. In this application, five scores are generated and weighted to determine which pixel in each pixel stack should be used in the final annual BAP composite. This approach is based on the work of Griffiths et al. (2013), Hermosilla et al. (2015a), and White et al. (2014), and adjusted to reflect the needs of

road extraction in the Ontario boreal forest. In Chapter 3, multiple variable and weighting combinations for this flexible parametric weighting scheme were assessed. The approach that generated the best overall accuracy was the following:

$$Final\ Score = (0.2 * Score_{DOY}) + (0.2 * Score_{Sensor}) + (0.2 * Score_{CloudDistance}) + (0.2 * Score_{Opacity}) + (0.2 * Score_{Brightness}) \quad [1]$$

Where $Score_{DOY}$ prioritizes pixels close to the ideal date of acquisition (September 1 or day 245), $Score_{Sensor}$ prioritizes pixels with sensors that generate the fewest anomalies, $Score_{CloudDistance}$ prioritizes pixels further away from cloud and cloud shadows, $Score_{Opacity}$ prioritizes pixels with low Aerosol Optical Thickness (AOT) and $Score_{Brightness}$ prioritizes brighter pixels. More detailed descriptions and formulas for each of these scores can be found in Chapter 3 Section 3.6. The pixel with the highest score is then used in the final pixel composite generated for each year. The forest road network is extracted from these BAP composite images.

4.3.6 Road Extraction

The road extraction process can be broadly classified into three main steps: pre-processing the reference road data, generating a mask using the reference roads, and road extraction using Jenks Optimization.

Reference Road Generation

The two input reference road sets, MNRF and ORN Road Segments, have a variety of inconsistencies that must be addressed prior to being used as a mask. There are many roads that exist within both datasets, however, there are topological and spatial differences between the duplicated segments. These mismatched duplications must be addressed in a manner that best preserves the true location, shape, and topology of the road segments. As the MNRF segments include more forest roads with enhanced detail, the non-arterial roads as defined in the MNRF dataset are erased from the ORN dataset. A 25 metre (m) buffer is applied to the non-arterial

roads to account for the spatial variation of the digitized roads between the ORN and MNRF road networks. This approach results in major roads being retained in the ORN dataset and all roads (predominantly forest roads) being retained in the MNRF road network. These two road networks are then merged. Any roads within 30 m of each other are integrated into a singular line segment with common vertices. The road segments that are disconnected from other segments are then extended. Roads within 75 m of another road feature are extended to connect to that feature. The roads are then converted to multipart feature classes, where all line segments are considered one complex spatial entity with one feature identification number. This function enables the road segments to be manipulated individually as opposed to manipulating the full network. Any spatially coincident line segments, or segments within 30 m, are removed. This step addresses any duplicate lines being created when the major roads in each layer are merged. These processes of integrating, extending, and removing duplicate road segments do slightly alter the spatial location and topological relationships between segments. However, this process reduces duplicate road segments side by side which would have greater influence on the road connectivity overall. The slight change in spatial location is addressed by using a buffer around the line segments when creating the road extraction mask.

The road network is then buffered by 60 m so that the small linear features and some surrounding land cover can be used in calculating statistics for the road extraction. Sixty meters was selected as the buffer value to reflect the size of one or two pixels, for the MSS or TM, ETM+ and OLI sensors. Two raster road masks are then generated: (1) the first road mask is the MSS road mask which has a spatial resolution of 60 m, (2) the second is the road mask for the TM, ETM+, and OLI data which has a spatial resolution of 30 m. These masks are then used to prepare the BAP composites for road extraction.

BAP Composite Road Masking

To reduce the impact of incomplete cloud masking and the significant amount of exposed bedrock within the study area, the known road mask is used to extract a subset of the study area within 60 m of the reference roads. The mask limits the analysis to roads that are known to have existed in the study site. Roads previously unrecorded are outside the scope of this study. A convolution function is applied to each BAP composite in the pre-1984 dataset. Accuracy was not enhanced when the convolution function was applied to the post-1984 dataset based on the completion of an accuracy assessment matrix. Convolution enables the enhancement of imagery and the removal of spurious features within the imagery. A smooth arithmetic mean filter is used on each image. This low pass filter smooths the data by reducing data variation and noise. This filter was selected over Sobel, Laplacian, or other edge filters as it produces the best results for this dataset through a series of testing. The use of filtering algorithms in road extraction algorithms is extensively discussed in the literature and are commonly applied before line detection (J. Wang et al., 1992). The road masks are then used to extract the area around the reference roads from the convoluted BAP composites for each year.

Road Extraction using Jenks Optimization

The Jenks Optimization algorithm is then implemented through arcpy. Data is classified based on statistically significant breaks. The aim is to reduce the average deviation of each value from the mean of the class and increase the deviation of each class from the mean of other classes (Gui et al., 2022). This approach results in the reduction of intraclass variation and maximizes interclass variation (Amirruddin et al., 2020; de Smith et al., 2018; Fraile et al., 2016; Hou et al., 2022). Classification results are based on iterative computations of goodness-of-fit tests (goodness of variance fit, GVF) to indicate class performance (Fraile et al., 2016).

The algorithm is applied to each image with four class breaks. The values of the breaks are iteratively calculated for each year. Through empirical testing, four classes provided the best extraction results. These four classes can be broadly divided into the following (from lowest to highest brightness value): (1) non-road or background pixel, (2) mixed-pixel (forest-road and occlusion), (3) unpaved forest road, and (4) paved road. These four classes were then reclassified to create two binary classes: non-road pixels (original class 1) and road pixels (original classes 2 to 4). The reclassified layers were then smoothed using focal statistics using a 3 x 3 rectangular neighborhood. The maximum value was selected within this neighborhood as it is easier to remove disconnected features than to fill gaps in areas with complex topology. The raster roads are then converted to vector roads. To convert the roads, the raster roads were first converted to polygons with each road segment being assigned a binary road status (non-road or road). The reference road layer polyline was then split based on this polygon layer, had the road and non-road attributes assigned, then merged to make the final polyline layer. This approach was taken as converting the raster directly to a polyline would result in zigzag lines requiring smoothing. This smoothing process would remove the original topological relations that are leveraged in the logic rules applied. The issue of raster to vector conceptualization of roads and their subsequent conversions is addressed throughout the road literature (Mena, 2003; J. Wang et al., 1992). The splitting and rejoining of the road and non-road pixels and vector segments also reduce some of the road spurs that would be generated if the network was joined as a homogenous network. The binary logic rules as described in Section 3.8 are then applied three times. As the logic rules only process the segments with spatially coincident vertices to the target segment, each consecutive running of the logic rules will enable newly generated road to be evaluated against adjacent segments.

4.3.7 Change Detection

PCCD was used to attribute the year of construction to the road network (Jakubauskas et al., 1990; Tewkesbury et al., 2015). The roads were classified into road or non-road segments using the Jenks Optimization as described in Section 3.6. From-to-change information was generated for each year pair between 1974 and 2019. For the pre-1984 dataset, the change regions were mosaicked together starting in 1975. The first year, 1974, was mosaicked last as the main road network was continually detected as a change region throughout the period between 1974 and 1984. The areas of change within the road network were then converted from a raster to a vector for the logic rules to be applied.

The post-1984 dataset followed a similar process. From-to-change information was generated; however, due to the presence of excessive noise, change areas that had less than a 2500 m perimeter (approximately 1200 m linear length) were removed. The change regions were then mosaicked together starting in 2019. For the logic rules to be applied, the pre-1984 and post-1984 road networks were merged to provide more year of construction information. If there were any duplicate change areas, the post-1984 roads were retained. The final road segments were smoothed and the pre-1984 roads were removed before the accuracy assessment.

4.3.8 Logic Rules: Binary and Yearly

Four logic rules were introduced to reduce the impact of non-salient roads during the road extraction process. These logic rules leverage the topological relations between the road segments. The logic rules, described below, are applied to the binary road segments (road or non-road) or the full road network with assigned year of construction. The logic rules are presented in Table 12 and Figure 12.

Table 12: Logic rule description for binary and yearly approaches

Rule Number	Intersection Type	Binary Rule Description	Yearly Rule Description
1: Disconnected Features	All	The road segments must have connectivity with the main network	The road segments must have connectivity with the main network
2: Road Gaps	Line or T Intersection	If the target road segment with year A is between two road segments assigned year B, the target road segment should be assigned the year B	If the target road segment is assigned non-road but is between two road segments assigned road, the target road segment should be assigned as road
3: Unidentified Roads	Line or T Intersection	If the target road segment is between two unidentified road segments, the target road segment should also be assigned unidentified	If the target road segment is assigned as road, but is between two road segments assigned non-road, the target road should be assigned non-road
4: Dead-End Inconsistencies	Dead End	If a road segment assigned year A and the target road branching off that segment has been assigned an earlier year of construction, the target road segment should be assigned the year A	If a road section branches off the target road segment assigned as non-road, the target segment should be a road

Logic rules 2, 3, and 4 are applied first. These rules are the most likely to add road segments to the network, even in cases where the road segment is not connected to the main network. These detached road fragments can be built upon as the first application of the logic rules, allowing other connected road fragments to also be connected without adding any type of penalty. Then, logic rule 1 can be applied to remove any disconnected features that were erroneously added to during the first logic rules phase. To execute the first phase of the logic rules, a table with spatial information about the road segments is generated. For each road segment, the road segments with spatially coincident vertices are identified.

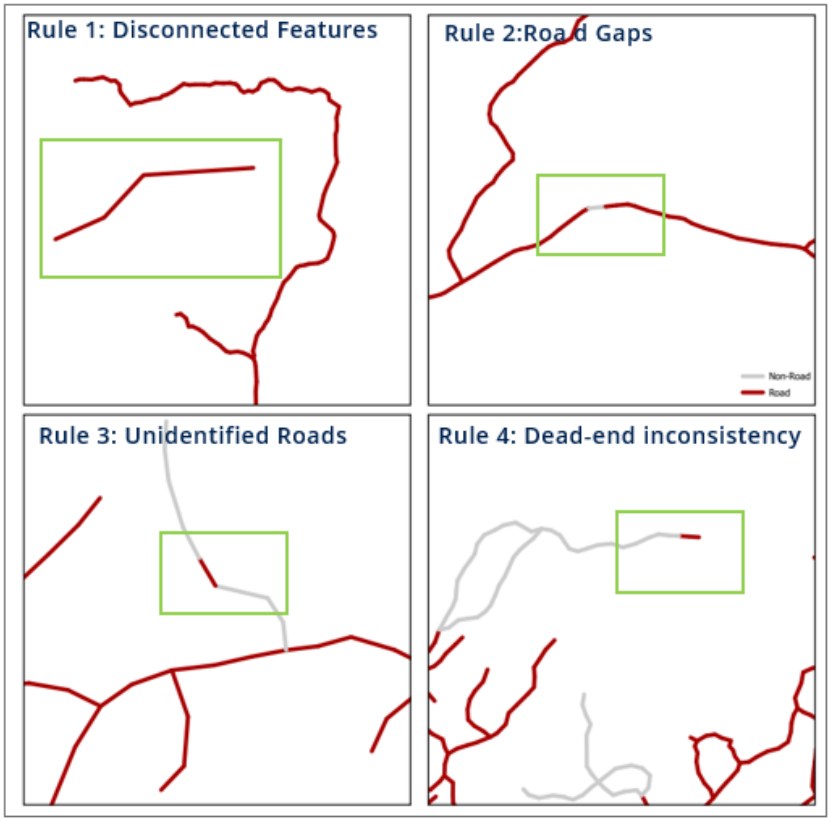

Figure 12: Logic rule road segment examples.

The road segment length and binary road value (road or non-road) or assigned year of construction are also included. The length of the road segments is used to ensure that the road segments are significant enough to be used to inform the road status of surrounding segments. For example, if the road segment is a small fragment of 10 m, the status of that road should not influence an adjacent road segment with a length of 700 m. In this application, the adjacent segments must be larger than the sum of the adjacent segments multiplied by 0.5. Multiple static distance lengths as well as percentages of the sum of the adjacent segments were tested; however, 50% of the sum of the segments was found to produce the highest accuracy.

The algorithm iteratively cycles through each road segment within the network for the application of logic rules 2, 3 and 4. For each target segment, it will identify the number of

elements with spatially coincident vertices. If there is only one spatially coincident point, the road is a dead end, and logic rule 4 will be applied. If there are two spatially coincident points on only one end of the segment, the road is functionally a dead end, but the two adjacent segments must have the road status for logic rule 4 to be applied. If the road has segments with spatially coincident vertices on both ends, the road is a line, and logic rule 2 and 3 can be applied. Again, the adjacent road segments must have the same road status for the year of construction to apply. If the roads were to have two different years, there is no appropriate logic to assign a year of construction as either one of those years or any year between those years. These steps are also applied to road and non-road status as well. If the road has more than 2 spatially coincident vertices, the logic rules are not applied. At intersections with 2 or more roads, a singular logic rule cannot be applied as the road could have been added at any stage of building the other three or more roads.

After the binary or yearly logic rules 2, 3 and 4 are applied, logic rule 1 is then used to remove small road fragments disconnected from the main network. An artificial study site boundary is generated for use in this algorithm. The road segments will not end at the boundary of the study area. In the event multiple adjacent study areas are processed and the resulting road networks are connected, the road network must retain topological relations with the road segments outside of the study site. To ensure exterior connectivity, any points at which reference roads intersect with the boundary are identified. All identified roads (either with a status of road or assigned a year of construction) are then buffered by 100 m. This buffering also accomplishes the reduction of small gaps between road segments that are too topologically complex for logic rules 2, 3 and 4. This buffer is converted to a single part feature before a spatial join is completed with exterior connectivity points. All road segments are retained during the spatial join. The

length of each of the road segments is calculated. Any road that is connected to the network based on the exterior connectivity spatial join or is larger than 20 m is retained. Ideally, all roads disconnected from the main network would be removed; however, as there are some road occlusions that cannot be addressed by the logic rules, large road segments that are disconnected are maintained. It is likely these roads should be connected to the main network. This process does add some 50 m road spurs on the edge of long segments with limited roads branching off. However, despite these spurs, accuracy is still increased through the application of the logic rules.

4.3.9 Accuracy Assessment

Results were evaluated using a quantitative accuracy assessment process. Broadly, this process involves sampling, collection of ground truth samples, generation of error matrices, and analysis of the results (P. Gong & Mu, 2000; Q. Zhang et al., 2002). Error or confusion matrices describe the level of accuracy between accuracy assessment samples and the corresponding classified values for the same image (Han et al., 2018; Zaidi et al., 2017). Overall classification accuracy is assessed through user's accuracy, producer's accuracy, and the kappa coefficient (Zaidi et al., 2017). These statistics allow for the assessment of the commission and omission errors and error bounds (Hermosilla et al., 2016). The confusion matrix is an extremely common approach to accuracy in remote sensing research (Zaidi et al., 2017). An automated approach was developed to generate accuracy assessment points and confusion matrices for each set of test results.

4.4 Results

Two tests were conducted to enhance the accuracy of road extraction and forest road disturbance year assignment. The first set of results evaluates the impact of the binary logic rules. These rules are applied to each year of extracted roads. The next set of results evaluates the

impact of the network level logic rules on the entire road network, with associated assigned years of construction, after the change detection process. This test was conducted on the pre- and post-1984 datasets separately to account for the differing sensor characteristics.

4.4.1 Binary Logic Rule Application

Table 13 displays the results of an accuracy assessment for three configurations of the binary road network for the years 1983, 1987, 1996, 2009, and 2019. These include (1) the road extraction without the application of the binary logic rules (results from Chapter 3 Section 3.4), (2) the road extraction with logic rules 2, 3 and 4 applied and (3) the road extraction with all logic rules applied. For all years excluding 1987, road extraction with all logic rules applied has the highest overall accuracy. The years 1983, 1987, and 2005 saw only a slight increase of 0.30%, 0.30%, and 0.9%, respectively whereas 1996, 2009, and 2019 see larger increases of 3.20%, 1.70% and 1%. In 1987, the highest overall accuracy is achieved with only logic rules 2, 3, and 4 applied. The highest kappa coefficient also occurs when only logic rules 2, 3, and 4 are applied. This result is strongly influenced by the kappa values of 1983 and 1987. The road network results of the two logic rule applications can be found in Figures 13 and 14. For the road network results for the road extraction with no logic rules applied, see Chapter 3 Section 3.4 – Figure 6. The ground truth road networks for each of the subset years can be found in Chapter 3 Section 3.4 – Figure 7.

Table 13: Accuracy assessment for three binary rule applications: (1) Without logic rules, (2) With logic rules 2,3 and 4, and (3) With all logic rules

Overall Accuracy by Logic Rule Application							
	1983	1987	1996	2005	2009	2019	Average Overall Accuracy
(1) No Logic Rules	0.939	0.933	0.873	0.854	0.848	0.832	0.880
(2) Logic Rules 2, 3 and 4	0.908	**0.954**	0.904	0.850	0.854	0.812	0.881
(3) Logic Rules 1, 2, 3, 4	**0.942**	0.936	**0.905**	**0.863**	**0.865**	**0.842**	**0.892**
Kappa by Logic Rule Application							
	1983	1987	1996	2005	2009	2019	Average Kappa
(1) No Logic Rules	0.877	0.866	0.744	0.696	0.656	0.582	0.737
(2) Logic Rules 2, 3 and 4	**0.955**	**0.908**	0.808	0.698	0.685	0.583	**0.773**
(3) Logic Rules 1, 2, 3, 4	0.882	0.873	**0.808**	**0.716**	**0.696**	**0.622**	0.766

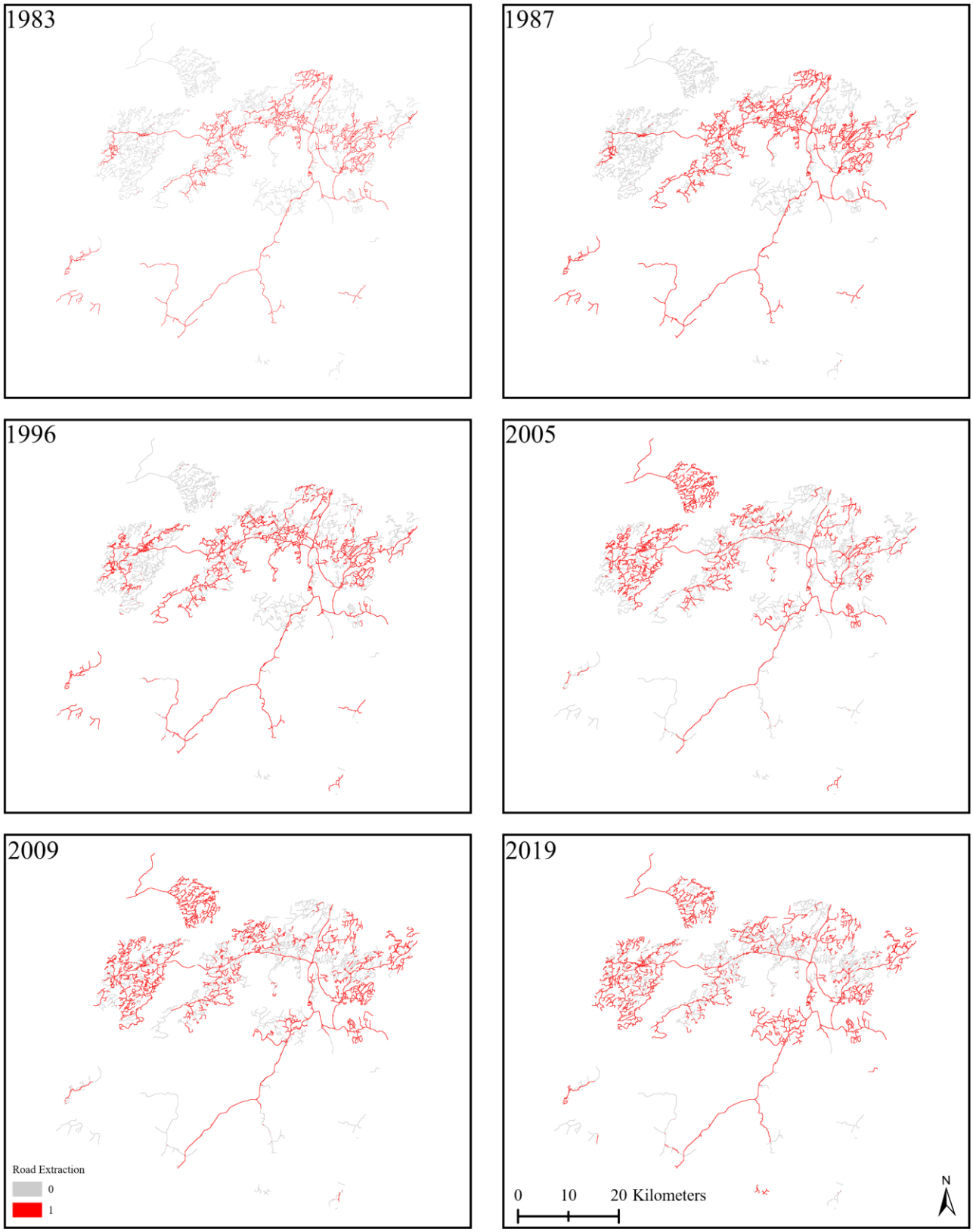

Figure 13: Roads extracted from Red and SWIR-2 annual BAP composite with logic rules 2, 3 and 4 applied.

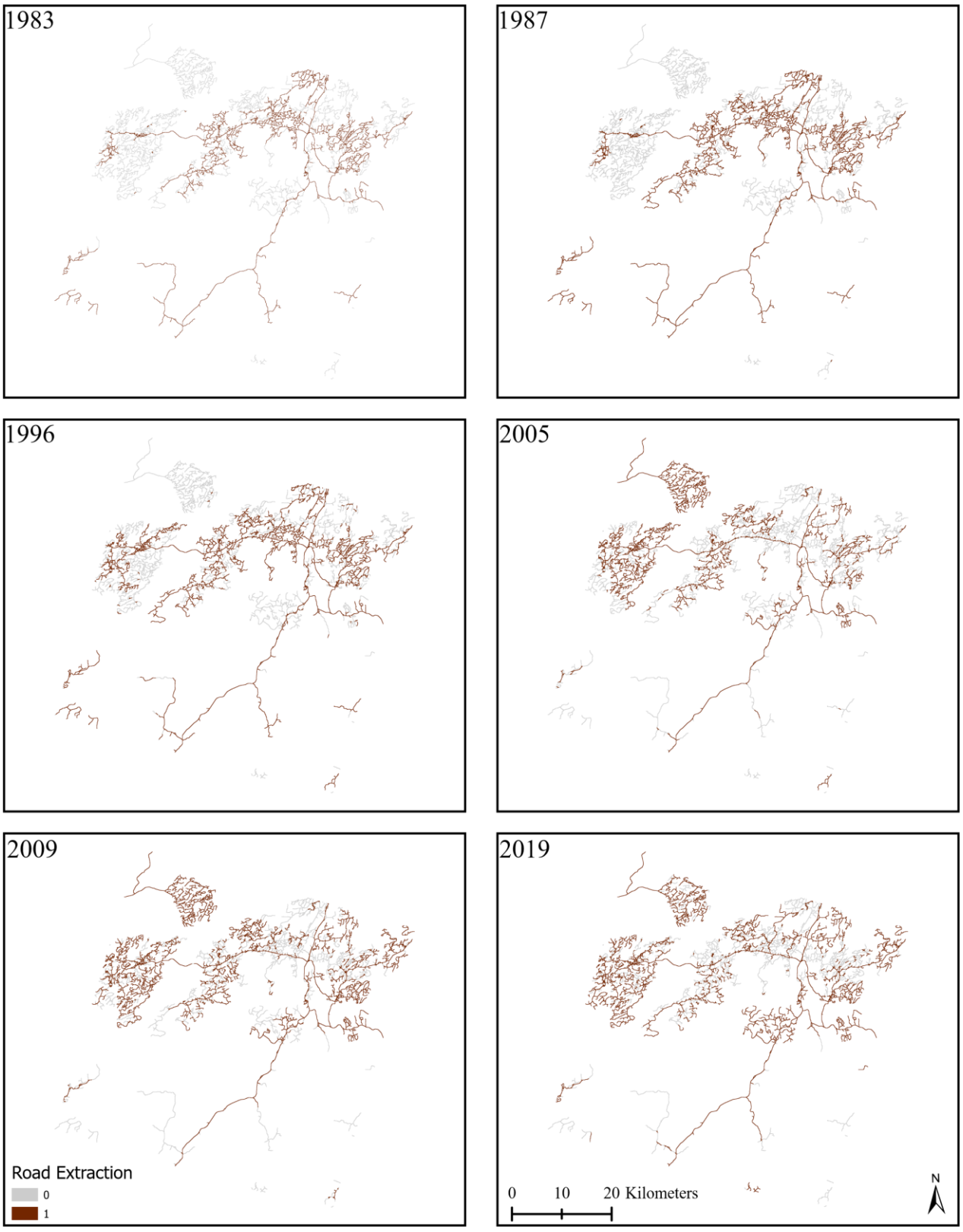

Figure 14: Roads extracted from Red and SWIR-2 annual BAP composite with logic rules 1, 2, 3 and 4 applied

4.4.2 Network Level Logic Rule Application

The four logic rules were applied to each of the road networks pre- and post-1984 to enhance their connectivity and year of construction attribution. Table 14 displays the accuracy assessment for the pre-1984 roads without the application of the binary or network level logic rules. The user's accuracy ranges between 0.482 and 0.964 whereas the producer's accuracy ranges between 0.528 and 0.985. The pre-1984 road extraction network had an overall accuracy of 0.770 and a kappa of 0.744.

Table 14: Accuracy assessment for extracted roads constructed between 1974 and 1984 without the application of logic rules.

Pre-1984 Road Network									
Year	**Non-Road**	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**	**1980**	**1981**
Producer's Accuracy	0.848	0.985	0.557	0.600	0.820	0.579	0.528	0.771	0.936
User's Accuracy	0.964	0.736	0.630	0.825	0.482	0.688	0.933	0.860	0.830
Year	**1983**	**1984**					**Overall Accuracy**		**0.770**
Producer's Accuracy	0.938	0.837							
User's Accuracy	0.726	0.911					**Kappa**		**0.744**

Table 15 presents the accuracy assessment for roads constructed after 1984 without the application of the binary or network level logic rules. The user's accuracy ranges between 0 (for years there are few data points or no data points due to lack of road construction) and 1. The producer's accuracy ranges between 0 and 0.969. The post-1984 road extraction network had an overall accuracy of 0.681 and a kappa of 0.667.

Table 15: Accuracy assessment for extracted roads constructed between 1984 and 2019 without the application of logic rules.

Post-1984 Road Network								
Year	**1985**	**1987**	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**
Producer's Accuracy	0.000	0.000	0.000	0.000	0.380	0.808	0.469	0.527
User's	0.000	0.000	0.000	0.000	0.864	0.756	0.652	0.569

Accuracy								
Year	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Producer's Accuracy	0.163	0.556	0.633	0.793	0.629	0.744	0.806	0.915
User's Accuracy	0.412	0.833	1.000	0.657	0.973	0.873	0.800	0.962
Year	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Producer's Accuracy	0.816	0.773	0.202	0.824	0.780	0.661	0.687	0.724
User's Accuracy	0.959	0.603	1.000	0.969	0.681	0.884	0.885	0.977
Year	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Producer's Accuracy	0.782	0.909	0.892	0.788	0.750	0.953	0.541	0.867
User's Accuracy	0.843	0.476	0.805	0.695	0.162	0.451	0.736	0.241
Year	**2019**	**Non-Road**						***0.681***
Producer's Accuracy	0.000	0.000				**Overall Accuracy**		
User's Accuracy	0.000	0.000				**Kappa**		***0.667***

Table 16 and 17 provide the accuracy assessment for the road networks with the four logic rules applied. Figure 15 includes the final road network with year of construction attribution for both the pre- and post-1984 time periods. The results for the pre-1984 road network show an increase in overall accuracy of 0.3. The accuracy increases from 0.770 to 0.800. The kappa sees a similar increase from 0.744 to 0.777.

Table 16: Accuracy assessment for extracted roads constructed between 1974 and 1984 with the application of logic rules.

Pre-1984 Road - Logic Rules Applied									
Year	**Non-Road**	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**	**1980**	**1981**
Producer's Accuracy	0.992	0.985	0.557	0.636	0.820	0.579	0.547	0.750	0.936
User's Accuracy	0.969	0.753	0.667	0.814	0.586	0.688	0.935	0.857	0.815
Year	**1983**	**1984**					**Overall Accuracy**		**0.800**
Producer's Accuracy	0.938	0.837							

User's Accuracy	0.750	0.911		**Kappa**	**0.777**

In contrast, the results for the post-1984 road network do not increase with the application of the four logic rules. Instead, the overall accuracy and kappa remain the same. Multiple individual years increase or decrease in accuracy based on the logic rules, but as these changes ultimately balance out, they do not enhance overall accuracy.

Table 17: Accuracy assessment for extracted roads constructed between 1984 and 2019 with the application of logic rules.

Post-1984 Roads - Logic Rules Applied								
Year	**1985**	**1987**	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**
Producer's Accuracy	0.000	0.000	0.000	0.000	0.380	0.808	0.469	0.509
User's Accuracy	0.000	0.000	0.000	0.000	0.864	0.766	0.682	0.571
Year	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Producer's Accuracy	0.163	0.556	0.633	0.841	0.629	0.744	0.806	0.915
User's Accuracy	0.467	0.833	1.000	0.683	0.986	0.873	0.806	0.962
Year	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Producer's Accuracy	0.816	0.773	0.202	0.824	0.780	0.661	0.687	0.690
User's Accuracy	0.959	0.582	1.000	0.969	0.696	0.884	0.868	0.976
Year	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Producer's Accuracy	0.782	0.909	0.892	0.788	0.750	0.953	0.541	0.867
User's Accuracy	0.843	0.476	0.805	0.695	0.200	0.436	0.736	0.245
Year	**2019**	**Non-Road**				**Overall Accuracy**		**0.681**
Producer's Accuracy	0.000	0.000						
User's Accuracy	0.000	0.000				**Kappa**		**0.668**

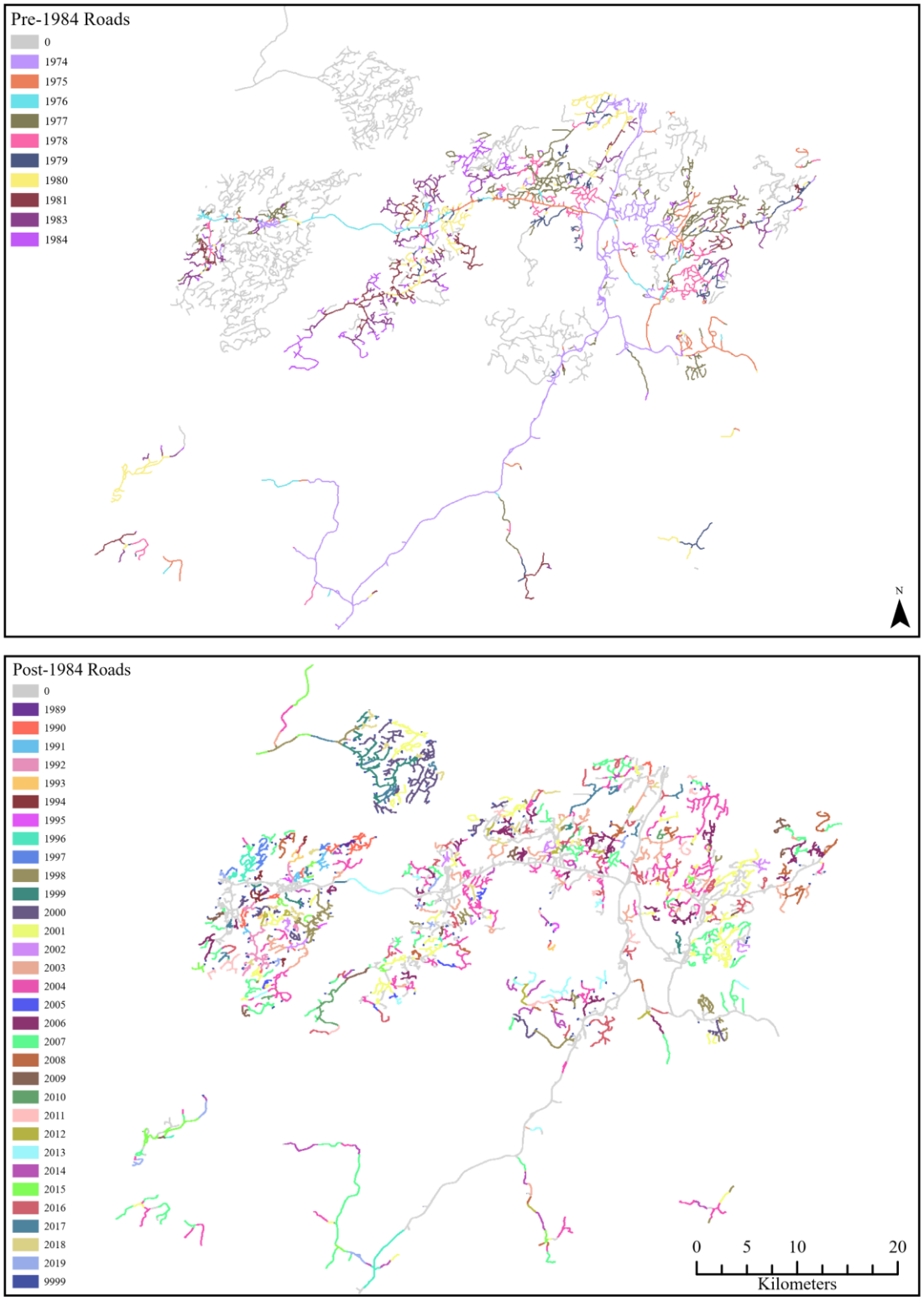

Figure 15: Final road network for roads constructed between 1974 and 1984 (upper) and roads construction between 1985 and 2019 (lower) with the application of logic rules.

4.5 Discussion

The results from the series of tests using the proposed semi-automatic algorithm for forest road extraction and year of construction attribution indicate that the proposed approach is suitable for the MSS imagery; however, the TM, ETM+, and OLI portion of the algorithm has multiple limiting factors that prevent successful implementation. The road extraction, without the enhancement of the binary logic rules, had an average overall accuracy of 0.892; however, the overall accuracy decreases as the year increases at the individual image level. The earlier years of 1983, 1987, and 1996 have results over 0.90 while the years 2005, 2009, and 2019 have accuracies around 0.85. As noted in Chapter 3, this may be the result of a combination of factors including phenological consistency of the BAP composites and the increase in road complexity as more roads are decommissioned or recommissioned over time. When all four binary logic rules are applied, there is no consistent temporal trend in the increase in accuracy. This indicates that the binary logic rules can be successfully applied to the extracted roads from all the Landsat sensors. The increase in accuracy is likely dependent on the level of noise in the image and the topological complexity of the erroneous road segments.

The results of the change detection and year attribution show varying results. The pre-1984 roads have an accuracy of 0.77. As the MSS image quality is generally poor, this value was anticipated. For example, the year 1977, which has a low user's accuracy of 0.482, has a significant amount of unmasked cloud and haze. This haze is brighter than the background pixels and will negatively impact the unsupervised classification. The haze increases the erroneous road segments identified. However, the post-1984 roads, which have higher quality input images, had a lower than anticipated overall accuracy of 0.68. Through examination of the image stack, this accuracy is likely the result of the propagation of classification errors, missing pixels/data gaps,

and the type of roads extracted. An example of classification error can be seen in the 2015 image, which has a user's accuracy of 0.162. This is the result of unmasked smoke that covers a large portion of the study area. The smoke results in the identification of erroneous road segments. As the road network is developed through mosaicking the images from latest to earliest age, these erroneous segments will be propagated through the image; a limitation common in PCCD change detection approaches. Any other single year anomalies in the pixel stack will also create multiple instances of change for a given road segment, reducing the accuracy of the entire classification. Similarly, the 2007 image is missing pixels within the composite due to persistent cloud cover. During the change detection process, it appears that the missing road segments are constructed the following year as the pixels are no longer missing. The number and type of road segments to be extracted may also influence the differences in accuracies between the MSS and TM, ETM+ and OLI imagery. Most road segments in the study area have an initial construction year prior to 1984. In the pre-1984 period, the roads that exist or are constructed include non-forest roads as well as a higher proportion of primary and branch roads. These roads are easier to identify as they are typically wider, have construction materials that appear brighter, and are more frequently travelled which reduces plant growth prior to decommissioning. In the post-1984 study period, there are fewer primary and branch roads constructed. The operational roads will be harder to detect due to their surface type and may have more complex intersections with other operational roads. There are also fewer roads constructed per year. In some years, only a single operational road was built. Smaller road segments to extract and classes with limited ground truth data will also negatively impact the classification results.

The results of the network level logic rules also present differing results for the pre-and post-1984 datasets. The pre-1984 dataset sees a 0.03 increase in overall accuracy with the implementation of all 4 logic rules. The post-1984 dataset sees minimal changes and no increase in accuracy. The lack of response to the logic rules in the post-1984 dataset may be the result of the initial poor classification. As there is propagated noise and erroneously dated segments, adjacent years of construction to a target segment may not be related, preventing the application of the logic rules. In addition, the roads constructed in the post-1984 dataset are much more topologically complex than the pre-1984 dataset. The spatial relationship between long expanses of highways or primary and branch roads will be simpler than shorter operational roads with increased connectivity to surrounding segments. As noted above, knowledge rules are limited to three road intersections with specific requirements for the years of adjacent segments. These specifications are less likely to occur with more complex topology and more transient road segments. Additionally, by 2019, the road network is complete, so there are fewer erroneous road segments that can be removed through logic rule 1. The amount and type of construction in the pre-and post-1984 study period may change regionally, so future work should consider extending the testing regions in addition to adding measures to increase initial post-1984 extraction accuracy.

4.6 Conclusion

Logging roads are associated with a variety of types of forest degradation. The forest clearing for road construction makes forest roads one of the most pervasive disturbances related to selective logging (Kleinschroth et al., 2015). Forest fragmentation, increased forest fire risk, edge effects, spread of invasive species, changes to the microclimate, increased siltation of aquatic habitats, decreased tree growth rates and reduced habitat suitability near the road surface

are concerning impacts of forest road construction (Kleinschroth et al., 2015; Mann et al., 2019; Sherba et al., 2014). Forest disturbance monitoring using remotely sensed imaged can provide information on forest road networks to allow stakeholders to study and address these concerns. The inclusion of the MSS imagery allows for historic road conditions to be incorporated into these efforts. The proposed semi-automatic approach for forest road extraction and change attribution provided an overall accuracy of 0.800 and 0.681 for the MSS and TM, ETM+, and OLI sensors, respectively. Through the development of binary logic rules, the road extraction was enhanced by up to 3% for images from all sensors. The network level logic rules enhanced the MSS road network by 3% while the TM, ETM+, and OLI derived network did not see significant enhancement. The low accuracy and lack of enhancement from the logic rules can be attributed to increased road complexity, decreased road construction, and data gaps in the BAP composites. Future work will focus on creating proxy annual BAP images to reduce data gaps as well as the inclusion of temporal logic rules to reduce temporal outliers within a pixel stack and reduce false change detection. Expanding the algorithm to include other forest road statuses may also help to address issues related to road complexity. Including road decommissioning and forest succession information may provide additional information with which to remove erroneous road segments.

4.7 References

Abdollahi, A., Pradhan, B., & Shukla, N. (2019). Extraction of road features from UAV using a novel level set segmentation approach. *International Journal of Urban Sciences, 23*(3), 391–405. https://doi.org/10.1080/12265934.2019.1596040

Afify, H. A. (2011). Evaluation of change detection techniques for monitoring land-cover changes: A case study in new Burg El-Arab area. *Alexandria Engineering Journal*, *50*, 187–195. https://doi.org/10.1016/j.aej.2011.06.001

Ahmed, O. S., Wulder, M. A., White, J. C., Hermosilla, T., Coops, N. C., & Franklin, S. E. (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. https://doi.org/10.1080/2150704X.2016.1233371

Ali, M. Z., Qazi, W., & Aslam, N. (2018). A comparative study of ALOS-2 PALSAR and Landsat-8 imagery for land cover classification using maximum likelihood classifier. *Egyptian Journal of Remote Sensing and Space Science*, *21*, S29–S35. https://doi.org/10.1016/j.ejrs.2018.03.003

Amini, J. (2009). Road extraction from satellite images using a fuzzy-snake model. *Cartographic Journal*, *46*(2), 164–172. https://doi.org/10.1179/000870409X459923

Amirruddin, A. D., Muharam, F. M., Ismail, M. H., Ismail, M. F., Tan, N. P., & Karam, D. S. (2020). Hyperspectral remote sensing for assessment of chlorophyll sufficiency levels in mature oil palm (Elaeis guineensis) based on frond numbers: Analysis of decision tree and random forest. *Computers and Electronics in Agriculture, 169*. https://doi.org/10.1016/j.compag.2020.105221

Arrighi, P., & Soille, P. (1999). From scanned topographic maps to digital elevation models. *Proceedings of Geovision'99, International Symposium on Imaging Applications in Geology*. http://web.ukonline.co.uk/soille/arrighi-soille99.pdf

Bae, Y., Lee, W. H., Choi, Y., Jeon, Y. W., & Ra, J. B. (2015). Automatic road extraction from remote sensing images based on a normalized second derivative map. *IEEE Geoscience and Remote Sensing Letters*, *12*(9), 1858–1862. https://doi.org/10.1109/LGRS.2015.2431268

Bakhtiari, H. R. R., Abdollahi, A., & Rezaeian, H. (2017). Semi automatic road extraction from digital images. *Egyptian Journal of Remote Sensing and Space Science, 20*(1), 117–123. https://doi.org/10.1016/j.ejrs.2017.03.001

Balaji, T. (2014). Robust and realistic classification of massive gray level thresholding in remote sensing images. *International Journal of Computer Science and Engineering*.

Barzohar, M., Cooper, D. B., & Member, S. (1996). Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18*(7). https://doi.org/10.1109/34.506793

Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Besbes, O., & Benazza-Benyahia, A. (2014). Joint road network extraction from a set of high resolution satellite images. *2014 22nd European Signal Processing Conference (EUSIPCO)*.

Bilyk, A., Pulkki, R., Shahi, C., & Larocque, G. R. (2021). Development of the Ontario Forest Resources Inventory: A historical review. *Canadian Journal of Forest Research*, *51*, 198–209. https://doi.org/10.1139/cjfr-2020-0234

Bocking, E., Cooper, D. J., & Price, J. (2017). Using tree ring analysis to determine impacts of a road on a boreal peatland. *Forest Ecology and Management*, *404*, 24–30. https://doi.org/10.1016/j.foreco.2017.08.007

Bowman, J., Ray, J. C., Magoun, A. J., Johnson, D. S., & Dawson, N. F. (2010). Roads, logging and the large-mammal community of an eastern Canadian boreal forest. *Canadian Journal of Zoology*, *88*, 454–467. https://doi.org/10.1139Z10-019

Cai, Q., Liu, H., Qian, Y., Zhou, S., Duan, X., & Yang, Y. H. (2019). Saliency-guided level set model for automatic object segmentation. *Pattern Recognition*, *93*, 147–163. https://doi.org/10.1016/j.patcog.2019.04.019

Cardim, G. P., da Silva, E. A., Dias, M. A., Bravo, I., & Gardel, A. (2018). Statistical evaluation and analysis of road extraction methodologies using a unique dataset from remote sensing. *Remote Sensing*, *10*(620). https://doi.org/10.3390/rs10040620

Chaudhuri, D., Kushwaha, N. K., & Samal, A. (2012). Semi-automated road detection from high resolution satellite images by directional morphological enhancement and segmentation techniques. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(5), 1538–1544. https://doi.org/10.1109/JSTARS.2012.2199085

Chen, J., Colombo, S. J., Ter-Mikaelian, M. T., & Heath, L. S. (2010). Carbon budget of Ontario's managed forests and harvested wood products, 2001-2100. *Forest Ecology and Management*, *259*, 1385–1398. https://doi.org/10.1016/j.foreco.2010.01.007

Chen, J., Feng, H., Pan, K., Xu, Z., & Li, Q. (2014). An optimization method for registration and mosaicking of remote sensing images. *Optik*, *125*, 697–703. https://doi.org/10.1016/j.ijleo.2013.07.034

Chen, Z., Jefferies, B., Adlakha, P., Salehi, B., & Power, D. (2014). Monitoring linear disturbance footprint based on dense time series Landsat imagery. *Canadian Journal of Remote Sensing*, *40*, 348–361. https://doi.org/10.1080/07038992.2014.987375

Cheng, G., Wang, Y., Xu, S., Wang, H., Xiang, S., & Pan, C. (2017). Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network. *IEEE*

Transactions on Geoscience and Remote Sensing, *55*(6), 3322–3337. https://doi.org/10.1109/TGRS.2017.2669341

Cheng, J., Ding, W., Ku, X., & Sun, J. (2012). Road extraction from high-resolution SAR images via automatic local detecting and human-guided global tracking. *International Journal of Antennas and Propagation*, *2012*. https://doi.org/10.1155/2012/989823

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, L. D. (1991). Note: On active contour models and balloons. *CVGIP: Image Understanding*, *53*(2), 211–218.

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync - Tools for calibration and validation. *Remote Sensing of Environment*, *114*, 2911–2924. https://doi.org/10.1016/j.rse.2010.07.010

Colditz, R. R., Acosta-Velázquez, J., Díaz Gallegos, J. R., De Vázquez Lule, A., Rodríguez-Zúñiga, M. T., Maeda, P., Cruz López, M. I., & Ressl, R. (2012). Potential effects in multi-resolution post-classification change detection. *International Journal of Remote Sensing*, *33*(20), 6426–6445. https://doi.org/10.1080/01431161.2012.688148

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Dadon, A., Mandelmilch, M., Ben-Dor, E., & Sheffer, E. (2019). Sequential PCA-based classification of Mediterranean forest plants using airborne hyperspectral remote sensing. *Remote Sensing*, *11*(23). https://doi.org/10.3390/rs11232800

Dal Poz, A. P., Zanin, R. B., & do Vale, G. M. (2006). Automated extraction of road network from medium-and high-resolution images. *Pattern Recognition and Image Analysis*, *16*, 239–248. https://doi.org/10.1134/S1054661806020118

de Smith, M. J., Goodchild, M. F., & Longley, p. A. (2018). *Geospatial analysis: A comprehensive guide to principles, techniques and software tools* (6th Edition). Winchelsea Press.

Du, P., Wang, X., Chen, D., Liu, S., Lin, C., & Meng, Y. (2020). An improved change detection approach using tri-temporal logic-verified change vector analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, *161*, 278–293. https://doi.org/10.1016/j.isprsjprs.2020.01.026

Environmental Systems Research Institute, Inc. (2022). *ArcGIS Pro* (3.0.1.).

Fan, D. L., Wang, B., Chen, Z. L., & Wang, L. (2020). Research on broken road connection method after road extraction from high-resolution remote sensing image. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* -

ISPRS Archives, *42*(3/W10), 387–395. https://doi.org/10.5194/isprs-archives-XLII-3-W10-387-2020

Fraile, A., Larrodé, E., Alberto Magreñán, & Sicilia, J. A. (2016). Decision model for siting transport and logistic facilities in urban environments: A methodological approach. *Journal of Computational and Applied Mathematics*, *291*, 478–487. https://doi.org/10.1016/j.cam.2014.12.012

Friedman, L., Netanyahu, N. S., & Shoshany, M. (2013). Mean shift-based clustering of remotely sensed data with agricultural and land-cover applications. *International Journal of Remote Sensing*, *34*(17), 6037–6053. https://doi.org/10.1080/01431161.2013.793866

Fung, T. (1992). Land use and land cover change detection with Landsat MSS and SPOT HRV data in Hong Kong. *Geocarto International*, *7*(3), 33–40. https://doi.org/10.1080/10106049209354378

Fung, T., & LeDrew, E. (1987). Application of principal components analysis to change detection. *Photogrammetric Engineering and Remote Sensing*, *53*(12), 1649–1658.

Gong, P. (1993). Change detection using principal component analysis and fuzzy set theory. *Canadian Journal of Remote Sensing*, *19*(1), 22–29. https://doi.org/10.1080/07038992.1993.10855147

Gong, P., & Mu, L. (2000). Error detection through consistency checking. *Annals of GIS*, *6*(2), 188–193. https://doi.org/10.1080/10824000009480549

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., & Hostert, P. (2014). Forest disturbances, forest recovery, and changes in forest types across the carpathian ecoregion from 1985 to 2010 based on landsat image composites. *Remote Sensing of Environment*, *151*, 72–88. https://doi.org/10.1016/j.rse.2013.04.022

Griffiths, P., Van Der Linden, S., Kuemmerle, T., & Hostert, P. (2013). Erratum: A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(5), 2088–2101. https://doi.org/10.1109/JSTARS.2012.2228167

Gui, R., Song, W., Pu, X., Lu, Y., Liu, C., & Chen, L. (2022). A river channel extraction method based on a digital elevation model retrieved from satellite imagery. *Water*, *14*(15). https://doi.org/10.3390/w14152387

Haas, J., & Ban, Y. (2018). Urban land cover and ecosystem service changes based on Sentinel-2A MSI and Landsat TM data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(2), 485–497. https://doi.org/10.1109/JSTARS.2017.2786468

Han, X., Pan, J., & Devlin, A. T. (2018). Remote sensing study of wetlands in the Pearl River Delta during 1995–2015 with the support vector machine method. *Frontiers of Earth Science*, *12*(3), 521–531. https://doi.org/10.1007/s11707-017-0672-x

He, C., Wei, A., Shi, P., Zhang, Q., & Zhao, Y. (2011). Detecting land-use/land-cover change in rural-urban fringe areas using extended change-vector analysis. *International Journal of Applied Earth Observation and Geoinformation*, *13*, 572–585. https://doi.org/10.1016/j.jag.2011.03.002

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015a). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sensing of Environment*, *158*, 220–234. https://doi.org/10.1016/j.rse.2014.11.005

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015b). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment*, *170*, 121–132. https://doi.org/10.1016/j.rse.2015.09.004

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., & Campbell, L. B. (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. https://doi.org/10.1080/17538947.2016.1187673

Hou, C., Xie, Y., & Zhang, Z. (2022). An improved convolutional neural network based indoor localization by using Jenks natural breaks algorithm. *China Communications*, *19*(4), 291–301. https://doi.org/10.23919/JCC.2022.04.021

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, *114*(1), 183–198. https://doi.org/10.1016/j.rse.2009.08.017

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *80*, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

Jakubauskas, M. E., Lulla, K. P., & Mausel, P. E. (1990). Assessment of vegetation change in a fire-altered forest landscape. *Photogrammetric Engineering and Remote Sensing*, *56*, 371–377.

Jenks, G. F. (1967). The data model concept in statistical mapping. *International Yearbook of Cartography*, *7*, 186–190.

Jenks, G. F., & Caspall, F. C. (1971). Error on choroplethic maps: Definition, measurement, reduction. *Annals of the Association of American Geographers*, *61*(2), 217–244. http://www.jstor.org/stable/2562442

Jeon, B.-K., Jang, J.-H., & Hong, K.-S. (2002). Road detection in spaceborne SAR images using a genetic algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(1). https://doi.org/10.1109/36.981346

Jiang, H., Feng, M., Zhu, Y., Lu, N., Huang, J., & Xiao, T. (2014). An automated method for extracting rivers and lakes from Landsat imagery. *Remote Sensing*, *6*, 5067–5089. https://doi.org/10.3390/rs6065067

Kahraman, I., Karas, I. R., & Akay, A. E. (2018). Road extraction techniques from remote sensing images: A review. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*(4/W9), 339–342. https://doi.org/10.5194/isprs-archives-XLII-4-W9-339-2018

Kass, M., & Witkin, A. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 321–331. https://doi.org/10.1007/BF00133570

Keaton, T., & Brokish, J. (2002). A level set method for the extraction of roads from multispectral imagery. *Applied Imagery Pattern Recognition Workshop, 2002*, 141–147. https://doi.org/10.1109/AIPR.2002.1182268

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. *Remote Sensing of Environment*, *114*, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Khatami, R., Mountrakis, G., & Stehman, S. V. (2016). A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sensing of Environment*, *177*, 89–100. https://doi.org/10.1016/j.rse.2016.02.028

Kleinschroth, F., Gourlet-Fleury, S., Gond, V., Sist, P., & Healey, J. R. (2016). Logging roads in tropical forests: Synthesis of literature written in French and English highlights environmental impact reduction through improved engineering. *Bois et Forets Des Tropiques*, *328*, 13–26. https://doi.org/10.19182/bft2016.328.a31299

Kleinschroth, F., Gourlet-Fleury, S., Sist, P., Mortier, F., & Healey, J. R. (2015). Legacy of logging roads in the Congo Basin: How persistent are the scars in forest cover? *Ecosphere*, *6*(4). https://doi.org/10.1890/ES14-00488.1

Kleinschroth, F., & Healey, J. R. (2017). Impacts of logging roads on tropical forests. *Biotropica*, *49*(5), 620–635. https://doi.org/10.1111/btp.12462

Koutsias, N., Mallinis, G., & Karteris, M. (2009). A forward/backward principal component analysis of Landsat-7 ETM+ data to enhance the spectral signal of burnt surfaces. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(1), 37–46. https://doi.org/10.1016/j.isprsjprs.2008.06.004

Land Information Ontario. (2012). *Indian reserve*. Ontario GeoHub. https://geohub.lio.gov.on.ca/datasets/lio::indian-reserve/about

Land Information Ontario. (2015). *Provincial park regulated*. Ontario GeoHub. https://geohub.lio.gov.on.ca/datasets/lio::provincial-park-regulated/about

Li, L. W., Liu, J. P., & Yin, Z. W. (2005). High-resolution remote sensing image road extraction based on mathematical morphology. *Remote Sensing Information*, *05*, 9–11.

Li, R., & Cao, F. (2018). Road Network Extraction from High-Resolution Remote Sensing Image Using Homogenous Property and Shape Feature. *Journal of the Indian Society of Remote Sensing*, *46*(1), 51–58. https://doi.org/10.1007/s12524-017-0678-6

Li, Y., Xu, L., & Piao, H. (2009). Semi-automatic road extraction from high-resolution remote sensing image: Review and prospects. *Proceedings - 2009 9th International Conference on Hybrid Intelligent Systems, HIS 2009*, *1*, 204–209. https://doi.org/10.1109/HIS.2009.317

Liu, H., & Zhou, Q. (2004). Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison. *International Journal of Remote Sensing*, *25*(5), 1037–1050. https://doi.org/10.1080/0143116031000150004

Liu, L., & Lim, S. (2016). A framework of road extraction from airborne lidar data and aerial imagery. *Journal of Spatial Science*, *61*(2), 263–281. https://doi.org/10.1080/14498596.2016.1147392

Liu, L., Tang, H., Caccetta, P., Lehmann, E. A., Hu, Y., & Wu, X. (2013). Mapping afforestation and deforestation from 1974 to 2012 using Landsat time-series stacks in Yulin District, a key region of the Three-North Shelter region, China. *Environmental Monitoring and Assessment*, *185*(12), 9949–9965. https://doi.org/10.1007/s10661-013-3304-2

Liu, Y., Dai, Q., Liu, J. B., Liu, S. Bin, & Yang, J. (2014). Study of burn scar extraction automatically based on level set method using remote sensing data. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0087480

Mann, D., Agrawal, G., & Joshi, P. K. (2019). Spatio-temporal forest cover dynamics along road networks in the Central Himalaya. *Ecological Engineering*, *127*, 383–393. https://doi.org/10.1016/j.ecoleng.2018.12.020

Markogianni, V., & Dimitriou, E. (2016). Landuse and NDVI change analysis of Sperchios river basin (Greece) with different spatial resolution sensor data by Landsat/MSS/TM and OLI. *Desalination and Water Treatment*, *57*, 29092–29103. https://doi.org/10.1080/19443994.2016.1188734

Matheron, G. (1974). *Random sets and integral geometry*. Wiley.

Maxwell, S. K., & Sylvester, K. M. (2012). Identification of “ever-cropped” land (1984-2010) using Landsat annual maximum NDVI image composites: Southwestern Kansas case

study. *Remote Sensing of Environment*, *121*, 186–195. https://doi.org/10.1016/j.rse.2012.01.022

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/bm.2012.031

McKeown, D. M. Jr., & Denlinger, J. L. (1988). Cooperative methods for road tracking in aerial imagery. *Proceedings CVPR '88: The Computer Society Conference on Computer Vision and Pattern Recognition*, 662–672. https://doi.org/10.1109/CVPR.1988.196307

Mena, J. B. (2003). State of the art on automatic road extraction for GIS update: A novel classification. *Pattern Recognition Letters*, *24*, 3037–3058. https://doi.org/10.1016/S0167-8655(03)00164-8

Mihell, K., & Hunt, L. M. (2011). Understanding residents' desired approaches to manage forest access roads: A case from northeastern Ontario, Canada. *Canadian Journal of Forest Research*, *41*(9), 1808–1818. https://doi.org/10.1139/x11-100

Minn, S. T., Mizoue, N., & Ota, T. (2022). Evaluating ground disturbance at elephant skid trails, logging roads and log landings under the Myanmar selection system. *Journal of Forest Research*, *27*(6), 409–418. https://doi.org/10.1080/13416979.2022.2067618

Minu, S., & Shetty, A. (2015). A comparative study of image change detection algorithms in MATLAB. *Aquatic Procedia*, *4*, 1366–1373. https://doi.org/10.1016/j.aqpro.2015.02.177

Mosammam, H. M., Nia, J. T., Khani, H., Teymouri, A., & Kazemi, M. (2017). Monitoring land use change and measuring urban sprawl based on its spatial forms: The case of Qom city. *Egyptian Journal of Remote Sensing and Space Science*, *20*(1), 103–116. https://doi.org/10.1016/j.ejrs.2016.08.002

Naikoo, M. W., Rihan, M., Ishtiaque, M., & Shahfahad. (2020). Analyses of Land Use Land Cover (LULC) change and built-up expansion in the suburb of a metropolitan city: Spatio-temporal analysis of Delhi NCR using Landsat datasets. *Journal of Urban Management*, *9*(3), 347–359. https://doi.org/10.1016/j.jum.2020.05.004

Nakaguro, Y., Makhanov, S. S., & Dailey, M. N. (2011). Numerical experiments with cooperating multiple quadratic snakes for road extraction. *International Journal of Geographical Information Science*, *25*(5), 765–783. https://doi.org/10.1080/13658816.2010.498377

Nevatia, R., & Babu, K. R. (1980). Linear feature extraction and description. *Computer Graphics and Image Processing*, *13*, 257–269. https://doi.org/10.1016/0146-664X(80)90049-0

Ontario Ministry of Natural Resources and Forestry. (2019a). *Land information Ontario data description: MNRF road segment*. https://www.sdc.gov.on.ca/sites/MNRF-PublicDocs/EN/CMID/MNR%20Road%20Segment%20-%20Data%20Description.pdf

Ontario Ministry of Natural Resources and Forestry. (2019b). *Land information Ontario data description: ORN road net element*. https://www.sdc.gov.on.ca/sites/MNRF-PublicDocs/EN/CMID/ORN%20Road%20Net%20Element%20-%20Data%20Description.pdf

Ontario Ministry of Natural Resources and Forestry. (2019c). *Land information Ontario data description: ORN segment with address*. https://www.sdc.gov.on.ca/sites/MNRF-PublicDocs/EN/CMID/ORN%20Segment%20with%20Address%20-%20Data%20Description.pdf

Ontario Ministry of Natural Resources and Forestry. (2020). *Forest management planning manual*. https://files.ontario.ca/mnrf-forest-management-planning-manual-en-2021-04-09.pdf

Osher, S., & Sethian, J. A. (1988). Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, *79*, 12–49.

Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, *12*(1 Suppl.). https://doi.org/10.1016/j.jag.2009.11.002

Özkaya, M. (2012). Road extraction from high resolution satellite images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS Congress*. https://doi.org/10.5194/isprsarchives-XXXIX-B4-143-2012

Patra, S., Sahoo, S., Mishra, P., & Mahapatra, S. C. (2018). Impacts of urbanization on land use / cover changes and its probable implications on local climate and groundwater level. *Journal of Urban Management*, *7*(2), 70–84. https://doi.org/10.1016/j.jum.2018.04.006

Polykretis, C., Grillakis, M. G., & Alexakis, D. D. (2020). Exploring the impact of various spectral indices on land cover change detection using change vector analysis: A case study of Crete Island, Greece. *Remote Sensing*, *12*(319). https://doi.org/10.3390/rs12020319

Public Lands Act, R.S.O. 1990, c. P.43. Retrieved September 5, 2023, from https://www.ontario.ca/laws/statute/90p43

Public Transportation and Highway Improvement Act, R.S.O. 1990, c. P.50. Retrieved April 16, 2023, from ontario.ca/laws/statute/90p50

Python Software Foundation. (2001). *The Python language reference*. https://docs.python.org/3/reference/

Qiaoping, Z., & Couloigner, I. (2004). Automatic road change detection and GIS updating from high spatial remotely-sensed imagery. *Geo-Spatial Information Science*, *7*(2), 89–95. https://doi.org/10.1007/BF02826642

Radke, R. J., Andra, S., Al-Kofahi, O., & Roysam, B. (2005). Image change detection algorithms: A systematic survey. *IEE Transactions on Image Processing*, *14*(3), 294–307. https://doi.org/10.1109/TIP.2004.838698

Rahman, A. F., Dragoni, D., Didan, K., Barreto-Munoz, A., & Hutabarat, J. A. (2013). Detecting large scale conversion of mangroves to aquaculture with change point and mixed-pixel analyses of high-fidelity MODIS data. *Remote Sensing of Environment*, *130*, 96–107. https://doi.org/10.1016/j.rse.2012.11.014

Rajeswari, M., Gurumurthy, K. S., Reddy, P. L., & Omkar Senthilnath J., S. N. (2011). Automatic road extraction based on normalized cuts and level set methods. *International Journal of Computer Applications*, *18*(7). https://doi.org/10.5120/2298-2988

Rodarmel, C., & Shan, J. (2002). Principal component analysis for hyperspectral image classification. *Survey and Land Information Systems*, *62*(2), 115–000. https://www.researchgate.net/publication/265198128

Rosin, P. L. (2002). Thresholding for change detection. *Computer Vision and Image Understanding*, *86*, 79–95. https://doi.org/10.1006/cviu.2002.0960

Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., Loveland, T. R., Vermote, E., & Zhang, C. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, *114*, 35–49. https://doi.org/10.1016/j.rse.2009.08.011

Saati, M., & Amini, J. (2017). Road network extraction from high-resolution SAR imagery based on the network snake model. *Photogrammetric Engineering and Remote Sensing*, *83*(3), 207–215. https://doi.org/10.14358/PERS.83.3.207

Sadeghfam, S., Hassanzadeh, Y., Nadiri, A. A., & Khatibi, R. (2016). Mapping groundwater potential field using catastrophe fuzzy membership functions and Jenks optimization method: A case study of Maragheh-Bonab plain, Iran. *Environmental Earth Sciences*, *75*(7). https://doi.org/10.1007/s12665-015-5107-y

Samet, R., & Hancer, E. (2012). A new approach to the reconstruction of contour lines extracted from topographic maps. *Journal of Visual Communication and Image Representation*, *23*(4), 642–647. https://doi.org/10.1016/j.jvcir.2012.02.005

Sanchez-Hernandez, C., Boyd, D. S., & Foody, G. M. (2007). Mapping specific habitats from remotely sensed imagery: Support vector machine and support vector data description based classification of coastal saltmarsh habitats. *Ecological Informatics*, *2*(2), 83–88. https://doi.org/10.1016/j.ecoinf.2007.04.003

Savage, S. L., Lawrence, R. L., Squires, J. R., Holbrook, J. D., Olson, L. E., Braaten, J. D., & Cohen, W. B. (2018). Shifts in forest structure in northwest Montana from 1972 to 2015 using the Landsat archive from Multispectral Scanner to Operational Land Imager. *Forests*, *9*(157). https://doi.org/10.3390/f9040157

Serra, J. (1982). *Image analysis and mathematical morphology*. Academic Press.

Sherba, J., Blesius, L., & Davis, J. (2014). Object-based classification of abandoned logging roads under heavy canopy using LiDAR. *Remote Sensing*, *6*(5), 4043–4060. https://doi.org/10.3390/rs6054043

Shi, W., Miao, Z., Wang, Q., & Zhang, H. (2014). Spectral-spatial classification and shape features for urban road centerline extraction. *IEEE Geoscience and Remote Sensing Letters*, *11*(4), 788–792. https://doi.org/10.1109/LGRS.2013.2279034

Singh, A. (1989). Review article: Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, *10*(6), 989–1003. https://doi.org/10.1080/01431168908903939

Song, M., & Civco, D. (2004). Road extraction using SVM and image segmentation. *Photogrammetric Engineering & Remote Sensing*, *12*, 1365–1371. https://doi.org/10.14358/PERS.70.12.1365

Soni, P. K., Rajpal, N., & Mehta, R. (2018). A comparison of road network extraction from High Resolution Images. *First International Conference on Secure Cyber Computing and Communications*, 525–531. https://doi.org/10.1109/ICSCCC.2018.8703276

Statistics Canada. (2016, November 16). *Provinces/Territories, cartographic boundary file - 2016 census*. https://open.canada.ca/data/en/dataset/a883eb14-0c0e-45c4-b8c4-b54c4a819edb

Sundaresan, A., Varshney, P. K., & Arora, M. K. (2007). Robustness of change detection algorithms in the presence of registration errors. *Photogrammetric Engineering & Remote Sensing*, *73*(4), 375–383. https://doi.org/10.14358/PERS.73.4.375

Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., & Fisher, P. F. (2015). A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, *160*, 1–14. https://doi.org/10.1016/j.rse.2015.01.006

Truong, C., Oudre, L., & Vayatis, N. (2018). Selective review of offline change point detection methods. *Signal Processing*, *167*. https://doi.org/10.1016/j.sigpro.2019.107299

United States Census Bureau. (2018). *Cartographic boundary files - Shapefile*. https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html

Ünsalan, C., & Boyer, K. L. (2011). Review on Building and Road Detection. In *Multispectral Satellite Image Understanding* (pp. 139–144). Springer London. https://doi.org/10.1007/978-0-85729-667-2_11

Valero, S., Chanussot, J., Benediktsson, J. A., Talbot, H., & Waske, B. (2010). Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images. *Pattern Recognition Letters*, *31*(10), 1120–1127. https://doi.org/10.1016/j.patrec.2009.12.018

Vese, L. A., & Chan, T. F. (2002). A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, *50*(3), 271–293. https://doi.org/10.1023/A:1020874308076

Vlachopoulos, O., Leblon, B., Wang, J., Haddadi, A., LaRocque, A., & Patterson, G. (2020). Delineation of bare soil field areas from unmanned aircraft system imagery with the mean shift unsupervised clustering and the random forest supervised classification. In *Canadian Journal of Remote Sensing* (Vol. 46, Issue 4, pp. 489–500). Bellwether Publishing, Ltd. https://doi.org/10.1080/07038992.2020.1763789

Vosselman, G., & De Knecht, J. (1995). Automatic extraction of man-made objects from aerial and space images road tracing by profile matching and Kalman filtering. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, 265–274. https://doi.org/10.1007/978-3-0348-9242-1_25

Wang, H., Huang, T. Z., Xu, Z., & Wang, Y. (2016). A two-stage image segmentation via global and local region active contours. *Neurocomputing*, *205*, 130–140. https://doi.org/10.1016/j.neucom.2016.03.050

Wang, J. (1993). Linda — a system for automated linear feature detection and analysis analysis. *Canadian Journal of Remote Sensing*, *19*(1), 9–21. https://doi.org/10.1080/07038992.1993.10855146

Wang, J., Treitz, P. M., & Howarth, P. J. (1992). Road network detection from SPOT imagery for updating geographical information systems in the rural-urban fringe. *International Journal of Geographical Information Systems*, *6*(2), 141–157. https://doi.org/10.1080/02693799208901901

Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., & Eklund, P. (2016). A review of road extraction from remote sensing images. *Journal of Traffic and Transportation Engineering (English Edition)*, *3*(3), 271–282. https://doi.org/10.1016/j.jtte.2016.05.005

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., Coops, N. C., Hall, R. J., Hostert, P., Dyk, A., & Guindon, L. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, *40*(3), 192–212. https://doi.org/10.1080/07038992.2014.945827

Xiao, P., Sheng, G., Zhang, X., Liu, H., & Guo, R. (2021). Direction-dominated change vector analysis for forest change detection. *International Journal of Applied Earth Observation and Geoinformation*, *103*. https://doi.org/10.1016/j.jag.2021.102492

Xu, C., & Prince, J. L. (1998). Generalized gradient vector flow external forces for active contours. *Signal Processing*, *71*, 131–139.

Zafiropoulos, P., & Schenk, T. (1998). Extraction of road structures with color energy models. *Electronic Imaging*. https://doi.org/10.1117/12.333792

Zaidi, S. M., Akbari, A., Samah, A. A., Kong, N. S., & Gisen, J. I. A. (2017). Landsat-5 time series analysis for land use/land cover change detection using NDVI and semi-supervised classification techniques. *Polish Journal of Environmental Studies*, *26*(6), 2833–2840. https://doi.org/10.15244/pjoes/68878

Zhang, C., Murai, S., & Baltsavias, E. P. (1999). Road network detection by mathematical morphology. *ISPRS Workshop 3D Geospatial Data Production: Meeting Application Requirements*, 185–200. https://doi.org/10.3929/ethz-a-004334280

Zhang, Q., Wang, J., Peng, X., Gong, P., & Shi, P. (2002). Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. *International Journal of Remote Sensing*, *23*(15), 3057–3078. https://doi.org/10.1080/01431160110104728

Zhou, J., Bischof, W. F., & Caelli, T. (2006). Road tracking in aerial images based on human-computer interaction and Bayesian filtering. *ISPRS Journal of Photogrammetry and Remote Sensing*, *61*(2), 108–124. https://doi.org/10.1016/j.isprsjprs.2006.09.002

Zhou, T., Sun, C., & Fu, H. (2019). Road information extraction from high-resolution remote sensing images based on road reconstruction. *Remote Sensing*, *11*(79). https://doi.org/10.3390/rs11010079

Zhou, X., & Mahmassani, H. S. (2007). A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework. *Transportation Research Part B: Methodological*, *41*(8), 823–840. https://doi.org/10.1016/j.trb.2007.02.004

Ziems, M., Breitkopf, U., Heipke, C., & Rottensteiner, F. (2012). Multiple-model based verification of road data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *1–3*. https://doi.org/10.5194/isprsannals-I-3-329-2012

Chapter 5: Conclusions

This research aimed to develop a semi-automatic algorithm for extracting and assigning year of construction to forest roads in the Ontario portion of the boreal forest using historic Landsat Best Available Pixel (BAP) composites and four logic rules. The objectives for this thesis were as follows: (1) to increase the temporal span of forest road monitoring to the Multispectral Scanner (MSS) sensor (1972 to 1974), which has historically been excluded, as well as to the Operational Land Imager (OLI) sensor (2013 to present); (2) to determine which configurations of the BAP compositing approach will optimize forest road extraction; (3) to assign year of forest road construction using a semi-automated unsupervised classification and change detection approach; (4) to enhance year of road construction attribution using topological logic rules.

Based on a series of analyses, it can be concluded that the proposed approach successfully extracts roads with a year of construction attribute for the MSS imagery (pre-1984) using the red band and the addition of a brightness parameter in the BAP composite protocol. The overall accuracy for the pre-1984 dataset was 80%. However, there are multiple limitations which prevent the successful extraction and change detection of the Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and OLI datasets (post-1984). Although the BAP composite approach was optimized using the Shortwave Infrared-2 (SWIR-2) bands and the addition of a brightness parameter, data gaps and temporal anomalies prevented the obtainment of high accuracy using Post-Classification Change Detection (PCCD) for this time period. The results of the accuracy assessment for the binary logic rules indicate that the binary rules enhance accuracy for the imagery from all four sensors. Up to a 3% increase overall was achieved depending on image noise, fragmentation, and topological complexity of annual the network.

The results of the accuracy assessment for the network level logic rules indicate enhancement of results for the MSS imagery, a 3% increase in overall accuracy. The TM, ETM+, and OLI network level logic rule results indicate further refinement of the input BAP composites is required prior to the successful implementation of these rules. Based on these conclusions, algorithm modifications should be focused on the extension of the BAP protocol to annual proxy composites for the OLI sensor. In addition, the development of temporal logic rules for the pixel stacks may address landcover anomalies which may decrease erroneous road segment detection and image noise.

Future work can be undertaken to enhance semi-automatic forest road monitoring, including enhancements to the algorithm to enable full scene processing as well as the inclusion of steps to allow the successful integration of tiled results. The proposed scope for future work is the provinces of Ontario and Quebec. This proposed extension to the study site also requires testing in additional study sites to ensure generalizability of the results over a provincial scale. In addition, the generated network from the proposed algorithm can be used to assist in determining the location of roads previously unrecorded within the study site. These annual known roads can be used as seed points for further road detection and extraction to provide a robust historic dataset crucial for effective and efficient forest monitoring initiatives.

Previous remote sensing research on forest disturbances excluded forest roads or only leveraged a small portion of the available Landsat data using only the TM and ETM+ sensors, preventing forest road monitoring from extending back to the 1970s (Hermosilla et al., 2015; Huang et al., 2010; Kennedy et al., 2010; Liu et al., 2013). Many forest road analyses still rely on manual digitization or analysis (Chen et al., 2014). These methods are prohibitively time consuming and often unfeasible for the scales required for historic province- or state-level

monitoring initiatives. Forest roads are significant source of anthropogenic disturbances within forested landscapes, being cited as an element of logging operations that have the most significant environmental impact (Kleinschroth, Gourlet-Fleury, et al., 2016). The impacts of forest roads vary based on construction materials, use, and decommissioning practices. Negative implications of the construction of forest roads include edge effects (Bocking et al., 2017; Kleinschroth, Gourlet-Fleury, et al., 2016; Laurance et al., 2009), changes to soil (Eaton et al., 2021; Kleinschroth & Healey, 2017) and hydrological regimes (Grayson et al., 1993; Laurance et al., 2009), deforestation (Affek et al., 2017; Kleinschroth & Healey, 2017) as well as habitat fragmentation and destruction (Affek et al., 2017; Hosaka et al., 2014; Kleinschroth, Healey, et al., 2016; Kleinschroth & Healey, 2017; Laurance et al., 2006; Legros et al., 2017). Despite environmental concerns, forest roads are often an essential feature required for socio-economic function. To support sustainable forest management as well as enable continued research efforts into the impact of forest roads on the forested ecosystem, historic forest road information obtained in an efficient and effective manner is crucial.

This work extends semi-automatic forest road monitoring approaches to the underutilized MSS sensor as well as proposes an approach to reduce the impact of non-salient roads by up to 6%. This work may assist government agencies, foresters, as well as other key stakeholders to include historic forest road year of construction into their rehabilitation and degradation mitigation plans or use in forest level analyses like changes to annual carbon emissions and sequestration metrics.

5.1 References

Affek, A. N., Zachwatowicz, M., Sosnowska, A., Gerlée, A., & Kiszka, K. (2017). Impacts of modern mechanised skidding on the natural and cultural heritage of the Polish Carpathian Mountains. *Forest Ecology and Management*, *405*, 391–403. https://doi.org/10.1016/j.foreco.2017.09.047

Bocking, E., Cooper, D. J., & Price, J. (2017). Using tree ring analysis to determine impacts of a road on a boreal peatland. *Forest Ecology and Management*, *404*, 24–30. https://doi.org/10.1016/j.foreco.2017.08.007

Chen, Z., Jefferies, B., Adlakha, P., Salehi, B., & Power, D. (2014). Monitoring linear disturbance footprint based on dense time series Landsat imagery. *Canadian Journal of Remote Sensing*, *40*, 348–361. https://doi.org/10.1080/07038992.2014.987375

Eaton, W. D., McGee, K. M., Donnelly, R., Lemenze, A., Larimer, M., & Hajibabaei, M. (2021). The impacts of a logging road on the soil microbial communities, and carbon and nitrogen components in a Northern Zone Costa Rican forest. *Applied Soil Ecology*, *164*. https://doi.org/10.1016/j.apsoil.2021.103937

Grayson, R. B., Haydon, S. R., Jayasuriya, A., & Finlayson, B. L. (1993). Water quality in mountain ash forests-separating the impacts of roads from those of logging operations. *Journal of Hydrology*, *150*, 459–480. https://doi.org/10.1016/0022-1694(93)90121-O

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment*, *170*, 121–132. https://doi.org/10.1016/j.rse.2015.09.004

Hosaka, T., Niino, M., Kon, M., Ochi, T., Yamada, T., Fletcher, C., & Okuda, T. (2014). Effects of logging road networks on the ecological functions of dung beetles in Peninsular Malaysia. *Forest Ecology and Management*, *326*, 18–24. https://doi.org/10.1016/j.foreco.2014.04.004

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, *114*(1), 183–198. https://doi.org/10.1016/j.rse.2009.08.017

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. *Remote Sensing of Environment*, *114*, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Kleinschroth, F., Gourlet-Fleury, S., Gond, V., Sist, P., & Healey, J. R. (2016). Logging roads in tropical forests: Synthesis of literature written in French and English highlights environmental impact reduction through improved engineering. *Bois et Forets Des Tropiques*, *328*, 13–26. https://doi.org/10.19182/bft2016.328.a31299

Kleinschroth, F., & Healey, J. R. (2017). Impacts of logging roads on tropical forests. *Biotropica*, *49*(5), 620–635. https://doi.org/10.1111/btp.12462

Kleinschroth, F., Healey, J. R., & Gourlet-Fleury, S. (2016). Sparing forests in Central Africa: Re-use old logging roads to avoid creating new ones. *Frontiers in Ecology and the Environment*, *14*(1), 9–10. https://doi.org/10.1002/FEEKleinscrothletter.1

Laurance, W. F., Croes, B. M., Tchignoumba, L., Lahm, S. A., Alonso, A., Lee, M. E., Campbell, P., & Ondzeano, C. (2006). Impacts of roads and hunting on central African rainforest mammals. *Conservation Biology*, *20*(4), 1251–1261. https://doi.org/10.1111/j.1523-1739.2006.00420.x

Laurance, W. F., Goosem, M., & Laurance, S. G. W. (2009). Impacts of roads and linear clearings on tropical forests. *Trends in Ecology and Evolution*, *24*(12), 659–669. https://doi.org/10.1016/j.tree.2009.06.009

Legros, D. L., Steinberg, B., & Lesbarrères, D. (2017). Middle of the road: Enhanced habitat for salamanders on unused logging roads. *Wildlife Research*, *44*(1), 1–8. https://doi.org/10.1071/WR14239

Liu, L., Tang, H., Caccetta, P., Lehmann, E. A., Hu, Y., & Wu, X. (2013). Mapping afforestation and deforestation from 1974 to 2012 using Landsat time-series stacks in Yulin District, a key region of the Three-North Shelter region, China. *Environmental Monitoring and Assessment*, *185*(12), 9949–9965. https://doi.org/10.1007/s10661-013-3304-2

Appendix 1: Forest Road Impacts

Forest roads are an element of logging operations that have one of the greatest environmental impacts (Kleinschroth, Gourlet-Fleury, et al., 2016). Although best practices have been developed for construction, maintenance, and decommissioning, recommendations are often not implemented, regulated, or enforced (Kleinschroth & Healey, 2017). The following sections will describe the environmental and social impacts of forest roads. As the proposed algorithm can be applied internationally, both temperate and tropical forests are discussed.

Forest roads have a wide range of environmental and social implications for the forests in which they exist. Negative impacts include human encroachment, carbon emissions, soil degradation, hydrological changes, edge effects and habitat fragmentation, desiccation and forest fires, biological invasion, habitat and biological changes, increased hunting, deforestation, and cultural feature destruction. The impacts of logging roads, as documented below, are diverse spatially and ecologically, whether they are direct or indirect processes. The impacts of roads differ significantly based on the organism, habitat, landscape properties, geographic location, spatial scale, weather conditions and so forth (Affek et al., 2017). The life stage of the road, construction processes, and rehabilitation or decommissioning efforts also have a significant impact on the type and level of degradation (Girardin et al., 2022; Turski et al., 2023).

Roads as an Indicator of Human Encroachment

Roads are often the most visible indicator of human activity within a forest and are therefore often used to estimate or model the global extent of forest degradation. Road networks are generally the only detectable aspect of selective logging, especially from medium to high resolution satellite imagery (Kleinschroth, Gourlet-Fleury, et al., 2016). Forest roads may also facilitate illegal logging, hunting, and the conversion of forested areas into agriculture. For example, in Brazilian Amazonia, 95% of all deforestation and forest fires occur within 50 km of

highways or roads. Most illegal gold mining operations in Suriname occur near roads. In tropical Africa, the population distribution of multiple species decreases near roadways whereas bushmeat production increases (Laurance et al., 2009). Laurance et al. (2015) have noted researchers state that the "first cut is the deepest" meaning that roads are perceived as the conduit to bring new environmental problems and exploitation of forest resources into previously inaccessible old growth forests (Kleinschroth & Healey, 2017).

Pedological Impacts

The development of logging roads can remove the forest topsoil and expose subsoil. This removal and subsequent heavy truck traffic can result in increased soil compaction and decreased bulk density. The exposed subsoil can also be subjected to increased soil erosion due to heavy rain, increased water runoff, and sediment transport into water bodies (Affek et al., 2017; Eaton et al., 2021). Increased amounts of trace elements can also destabilize soil chemistry and weaken its suitability for healthy forest stands (Turski et al., 2023). Dust, heavy metals, ozone, and organic molecules have been found at elevated levels within 10 to 200 m of roadways (Laurance et al., 2009). Large emissions of trace elements, lead, and de-icing salt are the result of the deposition of chemicals from exhaust emissions in traffic heavy regions, leaded gasoline usage, and road maintenance, respectively (Laurance et al., 2009). The harmfulness of these deposits is dependent on concentration and specific form of pollutant (Turski et al., 2023). These processes result in reduced bulk density, soil respiration, nutrient retention, oxygen levels, decomposition rates, carbon and nitrogen cycling, soil formation, productivity, and soil microbial communities (Eaton et al., 2021; Kleinschroth & Healey, 2017). The soil damage produced by forest roads has serious long-term impacts on forest recovery. Studies indicate that soil damage along logging roads can hinder tropical forest recovery for 10 to 15 years (Eaton et al., 2021).

Hydrological Impacts

The forest road related erosion mentioned above also results in the degradation of water quality and stream flow through accumulating sediment and debris (Affek et al., 2017; Kleinschroth & Healey, 2017). Contaminants enter aquatic environments during periods of heavy rain and negatively impact aquatic species, generate eutrophication, and cause biomagnification of heavy metals within aquatic food chains (Grayson et al., 1993; Laurance et al., 2009). Road crossings over waterways are particularly detrimental as bridge structures often erode over time. Although an herbaceous layer can reduce sediment production, it takes over twenty years to reach sufficient root strength to prevent landslides on steep slopes (Kleinschroth & Healey, 2017). Bridges and culverts may also dam streams, create artificial ponds and flooding, alter hydrologic processes, and change connectivity to downstream regions (Bocking et al., 2017; Kleinschroth & Healey, 2017). These changes in the hydrologic regime can impact water table depth, tree growth and morality rates, plant species composition, peatland type, soil aeration, substrate temperature, nutrient availability, and oxygen level in root zones (Bocking et al., 2017). Down-gradient of culvert blockages, there may be changes to water flow which can cause smaller streams to fail and result in plant desiccation during dry seasons. These changes can further negatively impact aquatic ecosystems as well as make landslides more prevalent (Laurance et al., 2009).

Edge Effects

The construction of roads creates an increased number of edges and intrusions into forest interiors. These edges can deplete forest resources, increase light infiltration, change local temperature, increase tree transpiration rates, and increase exposure to wind, dust, and other debris from road surfaces (Bocking et al., 2017; Kleinschroth, Gourlet-Fleury, et al., 2016; Laurance et al., 2009). Desiccation contributes to the relationship between forest roads and fire

occurrences. This phenomenon may be exacerbated by tree debris accumulated along the roadside during construction. Road corridors also have the paradoxical effect of increasing rates of fire ignition due to human activities as well as serving as transportation routes to fight interior fires. Road-related fire risk has been shown to decrease over time (Kleinschroth & Healey, 2017). The intensity of edge effects is dependent on the characteristics of the linear clearing. Width is an important attribute, with narrow clearings (less than 20 m) being less vulnerable to wind and desiccation stress than wider road corridors. Studies have indicated that forested regions within 50 to 100 m of edges have greater fluctuations in climate variables (light, temperature, and humidity). These sites are often drier and hotter than forest interiors as well as have increased tree mortality and canopy gaps (Laurance et al., 2009). Additional variables which may influence the impact of forest clearings include orientation of the clearing in relation to solar angles and maintenance of the road. Clearings parallel to the track of the sun will experience greater levels of sunlight throughout the day in comparison to perpendicular clearings, which will not see midday sun. This alignment with sun angle will result in greater heat and desiccation. Treating the edges with herbicides, fires, or mechanically removing the foliage disturbs adjacent vegetation and prevents the natural edge closure, further contributing to edge effect intensity (Laurance et al., 2009).

Habitat and Species Impacts

There are multiple interrelated impacts on animals and plants within forested areas in the vicinity of forest roads. These impacts include biological invasions, habitat loss and fragmentation, modification to animal patterns, and hunting stress. Logging trucks and other machinery have been shown to act as the vectors for the dispersal of exotic grass, tree, non-local vertebrate, fungi, and ant species (Bocking et al., 2017; Carvalho et al., 2021; Kleinschroth, Gourlet-Fleury, et al., 2016; Kleinschroth & Healey, 2017; Laurance et al., 2009). There is a

significant decline in the quality of habitat due to the greater presence of humans, noise contamination, and logging activities from forest roads. This decline in habitat quality often makes the habitat inaccessible for large territorial carnivores, birds, and anthrophobic species (Affek et al., 2017; Hosaka et al., 2014; Kleinschroth, Healey, et al., 2016; Kleinschroth & Healey, 2017; Laurance et al., 2006; Legros et al., 2017). Forest roads have a vast impact on forest dynamics including changes to tree species composition, forest structure, and microclimate. These changes have both positive and negative effects on individual species. Examples include early successional habitats facilitating increased survival of certain species, herbivore food resources, movement of carnivores, habitat suitability for reproduction, as well as predator-prey interactions (Affek et al., 2017; Carvalho et al., 2021; Kleinschroth, Healey, et al., 2016; Kleinschroth & Healey, 2017; St-Pierre et al., 2021). Forest roads also increase hunting, trapping, and fishing pressure on many species (Kleinschroth & Healey, 2017; Laurance et al., 2006; St-Pierre et al., 2021).

Deforestation

All the above impacts of forest roads have been the result of changes within a modified forest environment; however, deforestation as an indirect result of forest roads is also problematic. Generally, logging occurs within 25 km of detectable forest roads. The probability of logged forests ultimately becoming deforested is four times higher than undisturbed forests. Logging roads open forest access for conversion to a multitude of land uses (Kleinschroth & Healey, 2017). This statistic may oversimplify the relationship between logging roads and deforestation as the ultimate land conversion is often related to official governmental land re-designation, poor land use planning, and providing incentives for settlement (Kleinschroth & Healey, 2017). Concerns related to government zoning of logging roads also include providing unwanted access to Indigenous communities, omission of funding for transportation networks for

remote communities to access required services, and lack of consideration of health impacts on local communities due to respiratory problems related to heavy machinery traffic (dust and emissions) (Affek et al., 2017; Kleinschroth & Healey, 2017). This deforestation, in addition to more dynamic land clearing patterns related to forestry, results in increased carbon emissions from destroyed biomass. A study in Indonesia in 2014 indicates that road construction accounted for 14% of all forestry related carbon emissions (Kleinschroth & Healey, 2017). This percentage would vary significantly depending on the type of logging and regional environmental factors.

References

Affek, A. N., Zachwatowicz, M., Sosnowska, A., Gerlée, A., & Kiszka, K. (2017). Impacts of modern mechanised skidding on the natural and cultural heritage of the Polish Carpathian Mountains. *Forest Ecology and Management*, *405*, 391–403. https://doi.org/10.1016/j.foreco.2017.09.047

Bocking, E., Cooper, D. J., & Price, J. (2017). Using tree ring analysis to determine impacts of a road on a boreal peatland. *Forest Ecology and Management*, *404*, 24–30. https://doi.org/10.1016/j.foreco.2017.08.007

Carvalho, E. A. R., Nienow, S. S., Bonavigo, P. H., & Haugaasen, T. (2021). Mammal responses to reduced-impact logging in Amazonian forest concessions. *Forest Ecology and Management*, *496*. https://doi.org/10.1016/j.foreco.2021.119401

Eaton, W. D., McGee, K. M., Donnelly, R., Lemenze, A., Larimer, M., & Hajibabaei, M. (2021). The impacts of a logging road on the soil microbial communities, and carbon and nitrogen components in a Northern Zone Costa Rican forest. *Applied Soil Ecology*, *164*. https://doi.org/10.1016/j.apsoil.2021.103937

Girardin, P., Valeria, O., & Girard, F. (2022). Measuring spatial and temporal gravelled forest road degradation in the boreal forest. *Remote Sensing*, *14*(3), 457. https://doi.org/10.3390/rs14030457

Grayson, R. B., Haydon, S. R., Jayasuriya, A., & Finlayson, B. L. (1993). Water quality in mountain ash forests-separating the impacts of roads from those of logging operations. *Journal of Hydrology*, *150*, 459–480. https://doi.org/10.1016/0022-1694(93)90121-O

Hosaka, T., Niino, M., Kon, M., Ochi, T., Yamada, T., Fletcher, C., & Okuda, T. (2014). Effects of logging road networks on the ecological functions of dung beetles in Peninsular Malaysia. *Forest Ecology and Management*, *326*, 18–24. https://doi.org/10.1016/j.foreco.2014.04.004

Kleinschroth, F., Gourlet-Fleury, S., Gond, V., Sist, P., & Healey, J. R. (2016). Logging roads in tropical forests: Synthesis of literature written in French and English highlights environmental impact reduction through improved engineering. *Bois et Forets Des Tropiques*, *328*, 13–26. https://doi.org/10.19182/bft2016.328.a31299

Kleinschroth, F., & Healey, J. R. (2017). Impacts of logging roads on tropical forests. *Biotropica*, *49*(5), 620–635. https://doi.org/10.1111/btp.12462

Kleinschroth, F., Healey, J. R., & Gourlet-Fleury, S. (2016). Sparing forests in Central Africa: Re-use old logging roads to avoid creating new ones. *Frontiers in Ecology and the Environment*, *14*(1), 9–10. https://doi.org/10.1002/FEEKleinscrothletter.1

Laurance, W. F., Croes, B. M., Tchignoumba, L., Lahm, S. A., Alonso, A., Lee, M. E., Campbell, P., & Ondzeano, C. (2006). Impacts of roads and hunting on central African

rainforest mammals. *Conservation Biology*, *20*(4), 1251–1261. https://doi.org/10.1111/j.1523-1739.2006.00420.x

Laurance, W. F., Goosem, M., & Laurance, S. G. W. (2009). Impacts of roads and linear clearings on tropical forests. *Trends in Ecology and Evolution*, *24*(12), 659–669. https://doi.org/10.1016/j.tree.2009.06.009

Laurance, W. F., Peletier-Jellema, A., Geenen, B., Koster, H., Verweij, P., Van Dijck, P., Lovejoy, T. E., Schleicher, J., & Van Kuijk, M. (2015). Reducing the global environmental impacts of rapid infrastructure expansion. *Current Biology*, *25*(7), R259–R262. https://doi.org/10.1016/j.cub.2015.02.050

Legros, D. L., Steinberg, B., & Lesbarrères, D. (2017). Middle of the road: Enhanced habitat for salamanders on unused logging roads. *Wildlife Research*, *44*(1), 1–8. https://doi.org/10.1071/WR14239

St-Pierre, F., Drapeau, P., & St-Laurent, M. H. (2021). Drivers of vegetation regrowth on logging roads in the boreal forest: Implications for restoration of woodland caribou habitat. *Forest Ecology and Management*, *482*. https://doi.org/10.1016/j.foreco.2020.118846

Turski, M., Beker, C., & Czerniak, A. (2023). The impact of road investments on the forest environment—Case study: The impact of asphalt roads on the health condition and growth of trees. *Sustainability*, *15*(2). https://doi.org/10.3390/su15021307

Appendix 2: Mitigation of Forest Road Impacts

Forest roads are linear disturbances that can significantly impact the dynamics of the forest in which they are situated. These short- and long-term impacts have made the construction and maintenance of forest roads highly contentious. Decisions about road location, building, maintenance, and decommissioning are highly complex as there are many tradeoffs that must be considered by stakeholders. Traditional road planning has focused on economic and social factors. A shift to evaluating the ecological impacts of roads includes the more holistic appraisal of the interplay between road function, environmental factors, building regulations, and all living beings coexisting along the road corridor. A well-planned forest road is essential, regardless of its initial purpose (i.e., forest maintenance, resource extraction, recreation, or game control) (Demir, 2007). To achieve this kind of forest management, companies must be held accountable for careful management of the road networks within their concessions (Kleinschroth & Healey, 2017). Funding is often the most significant limitation for enacting mitigation strategies if human safety is not at play (Legros et al., 2017). There are many specific measures that can be implemented at varying scales and timeframes to reduce the impacts of forest roads. These measures provide a broad overview of mitigation strategies that may be applied globally. This paper does not seek to assess or comment on the specific strategies adopted within the study site. Instead, its purpose is to highlight the various options available to stakeholders globally. These include strategies at the planning, maintenance, and decommissioning phase of a forest road's lifespan.

Construction Practices

To mitigate the impact of forest roads, forestry companies and governmental organizations can engage in a variety of construction practices including minimizing the length and width of road corridors (taking into account solar angle and exposure to wind), adapt roads

to topographic features, ensure adequate drainage, protect slopes, limit stream and river crossings, optimize layouts, plan construction around large trees along the route, optimize the ratio between forest roads and skid trails, avoid collateral damage, prohibit the use of roads during wetter periods, and avoid the damming of streams and creeks (Carvalho et al., 2021; Kleinschroth, Gourlet-Fleury, et al., 2016; Laurance et al., 2009). Construction practices may also include the consideration of fuel consumption, which varies depending on road surfacing, slope gradient, vehicle tire type, and axel distribution as well as driver practices (Svenson & Fjeld, 2016). These practices indicate the importance of the shift from selective logging or conventional logging (CON) to reduced impact logging (RIL). Selective logging is a practice where only large trees of permissible commercial trees are selected and harvested. This practice occurs in approximately 20% of the world's forests (Minn et al., 2022). Selective timber harvest has the benefit of creating a diverse age structure and spatial pattern within forest stands, especially when compared to clear cutting. However, this method requires more frequent skidder access to the forest than other methods. Selective logging operations are so scattered that stands rarely remain cut free for any significant period. A semi-natural multiple age forest is achieved, but at the expense of significant human activity throughout the region (Affek et al., 2017). CON refers to untrained and unsupervised workers that work without using detailed management or planning documentation (Minn et al., 2022). Both CON and selective logging activities are often poorly planned, resulting in excessive road building activities (Laurance et al., 2009). RIL involves intense pre-harvest road and skid road planning and tree inventory, marking trees to be felled, skid road planning, and directional felling by trained workers. This method is designed to maximize efficiency while mitigating negative social and environmental impacts. Adopting RIL practices allows for less damage to adjacent vegetation and habitats as well as provides better

economic and social outcomes in comparison to CON or selective logging (Carvalho et al., 2021). Forest certification groups, such as the Forest Stewardship Council, provide market-based incentives for logging companies who abide by sustainable forest practices. These practices include prevention of long-term negative impacts on forests, such as poorly planned logging roads (Kleinschroth, Gourlet-Fleury, et al., 2016). Carefully planned forest road networks, such as networks that balance minimization of environmental damage with economic efficiency, have been shown to reduce forest degradation by 40% (Kleinschroth & Healey, 2017).

Maintenance Activities

After road construction, requiring environmentally conscious maintenance during operation can help reduce environmental degradation. These activities may include allowing secondary growth along road margins and embankments, mowing and brush cutting as opposed to the use of broad-scale herbicides or fire and maintaining a continuous canopy. These practices can lessen forest desiccation, reduce limiting factors on forest regeneration, lessen edge effects on microclimate, and reduce biological invasion (Laurance et al., 2009). Road degradation occurs unevenly across road networks because of road width, time since last maintenance, substrate type and compaction rate, traffic, travel speed, slope, and water accumulation or runoff (Girardin et al., 2022). Girardin et al. (2022) developed a road degradation model to anticipate when roads would require maintenance. Along with remote sensing analyses, this type of model can be integrated into forest management to plan the frequency of maintenance activities, budget effectively, and flag areas at the highest risk for environmental or safety risks if maintenance activities are not completed as scheduled. The maintenance of forest roads is a contentious issue within the literature, as maintenance activities, particularly if completed improperly, act as periodic landscape disturbances. These disturbances can often be as significant as road construction in terms of the type and extent of environmental impact. However, not maintaining

forest roads can also have significant negative impacts. The primary function of the road can be reduced in critical environment incidents like mass wasting events or flooding. These impacts can sometimes exceed the damages resulting from the construction phase (Demir, 2007).

Road Decommissioning and Abandonment

Road decommissioning and abandonment is another major element of concern for environmental stewardship. Like road maintenance, there is not a straightforward response to mitigating environmental impact. Over time, the road ecosystem matures and adapts to a variety of disturbances and ecological conditions. Despite mitigation practices, roads will always produce long-term legacies upon the landscape, of which the intensity is highly dependent on road type, use, and maintenance practices (Demir, 2007). Road abandonment and decommissioning allows successional processes to overtake the previously managed road corridor. The speed of these processes is very dependent on the type of road, geographic location, and ecological landscape factors (Demir, 2007). Forest regeneration after natural resource extraction is essential for sustainable forest management and biodiversity. The rate of reforestation and variation in species composition are dependent on the number of growing degree days, geographic location, proportion of conifers in adjacent forest stands, soil compaction, and proximity to wetlands (St-Pierre et al., 2021). The success of forest restoration can be measured structurally (height and stage of forest succession) as well as functionally (changes to predator and prey movement). As most logging roads are temporary, they are a great site for regeneration. However, reduced levels of regeneration have been noted along forest roads and skid trails due to unfavorable soil conditions (Kleinschroth, Gourlet-Fleury, et al., 2016). Researchers have noted that logging roads can add an extra 10 to 15 years to forest recovery following deforestation activities (Eaton et al., 2021). Rehabilitation techniques introduced can serve to accelerate the rate of succession to reach management goals after road abandonment

(Demir, 2007). Practices may include removing slash and debris piles after road decommissioning to help enhance the recovery rate of carbon and nitrogen cycles within soils (Eaton et al., 2021). Tree planting activities may also be employed (St-Pierre et al., 2021).

Although road decommissioning allows for forest regeneration, researchers argue that concentrating logging operations on existing road networks or reopening previously decommissioned logging roads have advantages over complete abandonment. Researchers have estimated that road construction accounts for 40% of selective logging operations. Reopening previously decommissioned or abandoned roads is less expensive and time consuming in comparison to the construction of a new road. This difference can be attributed to relatively lower amounts of biomass that need to be removed and the residual soil compaction. By reopening old roads, new forest areas do not need to be cleared in the vicinity of the former roads and commercial forestry activities can be concentrated and intensified in a previously disturbed area (Kleinschroth et al., 2016). There are some barriers to reopening logging roads, including poor documentation of the location of the previous roads or that more efficient routes were later identified (Kleinschroth et al., 2016). Despite these benefits, the reopening of roads that have been previously abandoned can have a significant impact on animal and plant populations that have just begun to recover. Poor management of managed forest resources in previously logged areas can result in unreliable regeneration of the most valuable and intensively harvested tree species, which may lead to further ecological complications. It may also continue to be more lucrative to open new areas where the highest value timber resources are located despite the rising transportation costs (Kleinschroth et al., 2016). How and when to decommission or reopen forest roads must be carefully evaluated and planned in the specific context of the road network.

Additional Strategies

Additionally, many stakeholders are calling for creating road-free and fully intact forest preserves (Laurance et al., 2009). Intact forest landscapes are defined as areas larger than 500 square kilometers with a minimum width of 10 kilometers that is outside of a 1 km buffer around any road or settlement. Forested areas that are not easily accessible by roads have the highest conservation value. Sparing high conservation value lands can strengthen other landscape fragmentation reduction strategies and enhance biodiversity (Kleinschroth et al., 2016).

One of the most significant issues with forest road environmental degradation mitigation strategies is that many regions focus solely on the road without evaluating indirect road consequences, such as hunting and secondary road expansion (Laurance et al., 2009). These oversights can include controlling access to currently used roads, constructing barriers, and removing stream crossings after decommissioning and prohibiting the transport of bushmeat in logging vehicles (Carvalho et al., 2021; Kleinschroth et al., 2016). Foresters may also consider using rail, when possible, as this controls access points and may strategically limit the encroachment into environmentally sensitive areas (Laurance et al., 2009). Governments should also regulate land use after the completion of logging activities and provide alternatives for incursions into frontier land (Kleinschroth et al., 2016). Remote sensing activities can help stakeholders better inventory, plan, and manage forested ecosystems including the documentation of forest roads.

References

Affek, A. N., Zachwatowicz, M., Sosnowska, A., Gerlée, A., & Kiszka, K. (2017). Impacts of modern mechanised skidding on the natural and cultural heritage of the Polish Carpathian Mountains. *Forest Ecology and Management*, *405*, 391–403. https://doi.org/10.1016/j.foreco.2017.09.047

Carvalho, E. A. R., Nienow, S. S., Bonavigo, P. H., & Haugaasen, T. (2021). Mammal responses to reduced-impact logging in Amazonian forest concessions. *Forest Ecology and Management*, *496*. https://doi.org/10.1016/j.foreco.2021.119401

Demir, M. (2007). Impacts, management and functional planning criterion of forest road network system in Turkey. *Transportation Research Part A: Policy and Practice*, *41*(1), 56–68. https://doi.org/10.1016/j.tra.2006.05.006

Eaton, W. D., McGee, K. M., Donnelly, R., Lemenze, A., Larimer, M., & Hajibabaei, M. (2021). The impacts of a logging road on the soil microbial communities, and carbon and nitrogen components in a Northern Zone Costa Rican forest. *Applied Soil Ecology*, *164*. https://doi.org/10.1016/j.apsoil.2021.103937

Girardin, P., Valeria, O., & Girard, F. (2022). Measuring spatial and temporal gravelled forest road degradation in the boreal forest. *Remote Sensing*, *14*(3), 457. https://doi.org/10.3390/rs14030457

Kleinschroth, F., Gourlet-Fleury, S., Gond, V., Sist, P., & Healey, J. R. (2016). Logging roads in tropical forests: Synthesis of literature written in French and English highlights environmental impact reduction through improved engineering. *Bois et Forets Des Tropiques*, *328*, 13–26. https://doi.org/10.19182/bft2016.328.a31299

Kleinschroth, F., & Healey, J. R. (2017). Impacts of logging roads on tropical forests. *Biotropica*, *49*(5), 620–635. https://doi.org/10.1111/btp.12462

Kleinschroth, F., Healey, J. R., & Gourlet-Fleury, S. (2016). Sparing forests in Central Africa: Re-use old logging roads to avoid creating new ones. *Frontiers in Ecology and the Environment*, *14*(1), 9–10. https://doi.org/10.1002/FEEKleinscrothletter.1

Laurance, W. F., Goosem, M., & Laurance, S. G. W. (2009). Impacts of roads and linear clearings on tropical forests. *Trends in Ecology and Evolution*, *24*(12), 659–669. https://doi.org/10.1016/j.tree.2009.06.009

Legros, D. L., Steinberg, B., & Lesbarrères, D. (2017). Middle of the road: Enhanced habitat for salamanders on unused logging roads. *Wildlife Research*, *44*(1), 1–8. https://doi.org/10.1071/WR14239

Minn, S. T., Mizoue, N., & Ota, T. (2022). Evaluating ground disturbance at elephant skid trails, logging roads and log landings under the Myanmar selection system. *Journal of Forest Research*, *27*(6), 409–418. https://doi.org/10.1080/13416979.2022.2067618

St-Pierre, F., Drapeau, P., & St-Laurent, M. H. (2021). Drivers of vegetation regrowth on logging roads in the boreal forest: Implications for restoration of woodland caribou habitat. *Forest Ecology and Management, 482*. https://doi.org/10.1016/j.foreco.2020.118846

Svenson, G., & Fjeld, D. (2016). The impact of road geometry and surface roughness on fuel consumption of logging trucks. *Scandinavian Journal of Forest Research, 31*(5), 526–536. https://doi.org/10.1080/02827581.2015.1092574

Appendix 3: Python Script

```
################################################################################
# Identification of Forest Road Construction Year from Historical Remotely
# Sensed Data by Integrating Change Detection
# and Geometric Connectivity

# Kirsten Noltie - 2023

################################################################################

# Import required modules and site packages
import arcpy
import os
import math
import numpy
import shutil
import time
from arcpy.sa import *
from datetime import date
from arcpy.geoprocessing import *
import statistics

# Enable files to be overwritten if true
arcpy.env.overwriteOutput = True

# Start timer for script
startTime = time.time()

################################################################################
# Input path from user. Required input data should include outer folder housing
# a subfolder for each raster within the scene to be preprocessed. Data should
# be in an unzipped tarball (from .tar.gz) file. Tarballs are the standard file
# format when downloading data from the USGS servers
inputPath = os.path.normpath(r"D:\NewSub")


# Reference road inputs. All input road files will be used (merged into a
# single file and automatically refined to reduce spatially near or coincident road segments
roadData =
["C:\RoadTest_ReferenceFiles\MNRRDSEG\MNRF_RoadSeg\MNRF_RoadSeg.gdb\
MNRF_ROAD_SEGMENT", "C:\RoadTest_ReferenceFiles\MNRRDSEG\ORNRoadSeg\
ORN_RoadSeg.gdb\OrnWithClass"]

# Input Band Numbers. Code can be modified to use any band within
# the input tarballs.
```

```
MSSE = "5"
MSSL = "2"
TM_ETMBand = "7"
OLIBand = "7"

# Preferred projection for output dataset
projection = arcpy.SpatialReference("WGS 1984 UTM Zone 15N")

# Area for raster subsetting/ analysis footprint - To be provided in
# raster and vector format
# Consider WRS changes when selecting footprint
clipRaster = r"C:\RoadTest_ReferenceFiles\extent_ras.tif"
clipPoly = r"C:\RoadTest_ReferenceFiles\ExtentRas.shp"

################################################################################
#Suffix for output BAP composites
name = "_BAPFinal_SWIR2.tif"

# Name of output files to store intermediate results
# Required for all functions
outerFile = "Results"
referenceRoad = "ReferenceRoads_0"
preprocessedImages = "PreprocessedImages_1"
BAPFolder = "BAP_2"
roadExtractSubFolder = "RoadExtract_3"
jenksOpti = "JenksOpt_04"
logicRules = "LogicRulesPrep_00"
diffFolder = "ChangeDetection_5"
finalFolder = "FinalResults_06"

bap_yearScore = "YearScore"
bap_sensorScore = "SensorScore"
bap_cloudScore = "CloudDistanceScore"
bap_opacityScore = "OpacityScore"
bap_brightnessScore = "BrightnessScore"
bap_preprocess = "BAPPreprocess"
bap_final = "FinalBAP"

yearScorePost = "Year_Score"
sensorScorePost = "Sensor_Score"
cloudScorePRe = "Cloud_Pre"
cloudScorePost = "Cloud_Score"
opacityScorePre = "Opacity_Pre"
opacityScorePost = "Opacity_Score"

intermediateResultFolder = [referenceRoad, preprocessedImages, BAPFolder,
```

```
                roadExtractSubFolder, jenksOpti, logicRules,
                diffFolder, finalFolder]

bapFolders = [bap_yearScore, bap_sensorScore, bap_cloudScore, bap_opacityScore,
        bap_brightnessScore, bap_preprocess,
        bap_final]

yearFolders = [yearScorePost]
sensorFolders = [sensorScorePost]
cloudScoreFolders = [cloudScorePRe, cloudScorePost]
opacityScoreFolders = [opacityScorePre, opacityScorePost]

# Base filepath for most geoprocessing and results
path = os.path.join(inputPath, outerFile)
pathbase = os.path.join(inputPath, outerFile)


################################################################################

# Create outer folder to hold all results and intermediate files
def createfolders(inpath, outfile):
   arcpy.CreateFolder_management(inpath, outfile)

   # Create folder for intermediate and final results
   for folder1 in intermediateResultFolder:
      arcpy.CreateFolder_management(path, folder1)

   for folder2 in bapFolders:
      arcpy.CreateFolder_management((os.path.join(path, BAPFolder)),
                        folder2)

   for folder3 in yearFolders:
      arcpy.CreateFolder_management((os.path.join(path, BAPFolder,
                              bap_yearScore)),
                        folder3)

   for folder4 in sensorFolders:
      arcpy.CreateFolder_management((os.path.join(path, BAPFolder,
                              bap_sensorScore)), folder4)

   for folder5 in cloudScoreFolders:
      arcpy.CreateFolder_management((os.path.join(path, BAPFolder,
                              bap_cloudScore)), folder5)

   for folder6 in opacityScoreFolders:
      arcpy.CreateFolder_management((os.path.join(path, BAPFolder,
```

```
                                    bap_opacityScore)), folder6)

    del folder6, folder5, folder4, folder3, folder2, folder1

# Create lists of all input images for preprocessing iteration
def listimages(inpath2):
    # Define lists for each sensor type for preprocessing
    LS8 = []
    LS457 = []
    LS12345 = []

    # List the files in the path provided by the user
    directory = os.listdir(inpath2)

    # Creates lists of the images for each sensor
    # MSS, ETM+, TM, and OLI/TIRS
    for folder in directory:
        if folder != outerFile:
            sensor = folder[1]
            if sensor == "C" or sensor == "O":
                LS8.append(folder)
            elif sensor == "E" or sensor == "T":
                LS457.append(folder)
            elif sensor == "M":
                LS12345.append(folder)
    print("Input Images:", "\n", "MSS: ", LS12345, "\n", "TM and ETM+: ",
          LS457, "\n", "OLI: ", LS8)
    return LS8, LS457, LS12345

# Remove rasters with cloud cover over 70% to reduce pre-processing errors
def cloudcover(listLS, sensorName):
    print("\n" + sensorName + " Files Removed due to Cloud Coverage "
                          "being Over 70%:")

    # Iterate through each list of folders provided by the user (recommended
    # for TM, ETM+ and OLI sensors). If the cloud coverage is too high in the study area
    #  the preprocessing will fail due to NoDATA rasters being produced
    for folder in listLS:
        # Changes workspace to each image folder provided by the user
        currentFolder = os.path.join(inputPath, folder)
        env.workspace = currentFolder
        rasterName = os.listdir(currentFolder)[0]

        # Parses the unique product identifier for each image
        productIdentifier = rasterName[:40]
        dateNum = productIdentifier[17:25]
```

```
        sensor = productIdentifier[:4]
        name_metadata = productIdentifier + "_MTL.txt"

        # Opens the metadata file to determine the cloud coverage
        fpath = os.path.join(inputPath, folder, name_metadata)
        f = open(fpath, "r")
        for x in f:
            if "CLOUD_COVER" in x and "LAND" not in x:
                percent = float(x[18:22])
        f.close()

        # If the percentage of cloud cover is over 80%, delete the folder
        # (is not processed) and notify the user
        if percent >= 70:
            print(productIdentifier)
            shutil.rmtree(os.path.join(inputPath, folder))
    print()

# Notify the user of the number of images retained for preprocessing
# Not all images will be used in the final BAP composite
def inpputyearnums(inpath3):
    # Outer folder provided by the user
    directory = os.listdir(inpath3)

    # Create empty dictionary and list for years
    yearDictionary = {}
    numlist = []

    # Lists all folders in the outer folder provided by the user
    for folder in directory:
        # Parse folders differently based on USGS naming conventions
        if folder != outerFile:
            if "_" in folder:
                yearInstance = folder[17:21]
            else:
                yearInstance = folder[10:14]
            numlist.append(yearInstance)

    # Determine the earliest and the latest year based on the input images
    fyear = min(numlist)
    lyear = max(numlist)

    # Creates list of the study year range. The last year must be one greater
    # because using the range alone results in the exclusion of the final year
    yearList = range(int(fyear), int(lyear) + 1)
```

```
    # Creates a dictionary with the value being an empty list to be filled
    # with the file names of the raster, the key is the year
    for yo in yearList:
        yearDictionary["L" + str(yo)] = []

    for i in numlist:
        imgYear = i
        for key in yearDictionary:
            yearID = "L" + imgYear
            if key == yearID:
                yearDictionary[key] = yearDictionary.setdefault(key, []) + [i]

    # Print years and image dates for user
    print("\n" , "Images per Year")
    for key in yearDictionary:
        print(key, len(yearDictionary.get(key)))

# Pre-process the imagery based on the specific sensor properties
# Each pre-processing strategy includes cloud, invalid pixel,
# unsaturated pixel, and snow/cloud shadow/water masking if available

# Pre-processes all items in the OLI Sensor List
def preprocessLS8(LS8list):
    # SQL Where Clauses for various pre-processing steps
    # Radsat Where Clause for Landsat 4-5 & 7 (ETM+ and TM) and
    # Landsat 8 (OLI/TIRS)
    whereClause2 = "Value = 32"
    # Valid Pixel Range for Landsat 4-5 & 7 (ETM+ and TM) and
    # Landsat 8 (OLI/TIRS)
    whereClause3 = "VALUE < 10000 And VALUE > 0"
    # Pixel QA Where Clause for Landsat 8  (OLI/TIRS)
    whereClause4 = "Value = 324 OR Value = 328 OR Value = 352 OR" \
                " Value = 388 OR Value =392 OR Value = 416 OR Value = " \
                "480 OR Value = 836 OR Value = 840 OR Value = 864 OR Value" \
                " = 880 OR Value = 900 OR Value =904 OR " \
                "Value = 928 OR Value =944 OR Value= 992 OR Value = 832 OR" \
                " Value = 336 OR Value = 368 OR Value = " \
                "400 OR Value = 432 OR Value = 848 OR Value = 880 OR Value" \
                " = 912 OR Value = 944 OR Value = 992  "

    # Cloud Where Clause for Cloud Score BAP Prep
    cloudEuclideanCO = "Value = 352 OR Value = 368 OR Value = 416 OR Value" \
                    " = 432 OR Value = 480 OR Value = 864 OR " \
                    "Value = 880 OR Value = 928 OR Value = 944 OR Value" \
                    " = 992 OR Value = 328 OR Value = 392 OR " \
                    "Value = 840 OR Value = 904 OR Value = 1350 "
```

```
inFalseRaster = 1

# Generate format for Extent Object to use in pre-processing
extentRaster = Raster(clipRaster)
topExtent = arcpy.GetRasterProperties_management(extentRaster,
                                    "TOP").getOutput(0)
leftExtent = arcpy.GetRasterProperties_management(extentRaster,
                                    "LEFT").getOutput(0)
rightExtent = arcpy.GetRasterProperties_management(extentRaster,
                                    "RIGHT").getOutput(0)
bottomExtent = arcpy.GetRasterProperties_management(extentRaster,
                                    "BOTTOM").getOutput(0)
extentFormat = str(leftExtent + " " + bottomExtent + " " +
            rightExtent + " " + topExtent)

# Pre-process each image in the Landsat 8 List
print("\n" + "The following OLI Images have been Preprocessed:")
for folder in LS8list:
   print(folder)
   # Changes workspace to each image folder provided by the user
   currentFolder = os.path.join(inputPath, folder)
   env.workspace = currentFolder
   rasterName = os.listdir(currentFolder)[0]

   # Parses the unique product identifier for each image
   productIdentifier = rasterName[:40]
   dateNum = productIdentifier[17:25]
   sensor = productIdentifier[:4]

   # Generate product identifier specific name files for input data
   name_pixel_qa = productIdentifier + "_pixel_qa.tif"
   name_radsat_qa = productIdentifier + "_radsat_qa.tif"
   name_band5 = productIdentifier + "_sr_band" + OLIBand + ".tif"
   aerosol_qa = productIdentifier + "_sr_aerosol.tif"
   rasterPrj = "memory\\" + "productIdentifier"

   # Generate file paths for saving intermediate results to disk
   subsetArea = os.path.join(path, preprocessedImages, sensor +
                     dateNum + ".tif")
   QAFile = os.path.join(path, BAPFolder, bap_opacityScore,
                  opacityScorePre, sensor + dateNum +
                  "_OpacRaw.tif")

   # Prepare Pixel QA Mask
   # Masks cloud, cloud shadow, water, snow, and cloud/water classes
   outInt1 = Int(name_pixel_qa)
```

```
        pixelQAMask = SetNull(outInt1, inFalseRaster, whereClause4)

        # Prepare Radsat Mask
        # Removes unsaturated pixels in input band
        outInt2 = Int(name_radsat_qa)
        radsatMask = SetNull(outInt2, inFalseRaster, whereClause2)

        # Prepare Input Surface Reflectance Band
        # Removes invalid pixel values (Greater than 1000 and less than 0)
        band5Valid = arcpy.sa.Con(name_band5, name_band5, 0, whereClause3)
        band5Valid = Raster(band5Valid) * 0.0001

        # Extracts the Pixel QA Mask from Band 5 Valid
        outExtractByMask1 = ExtractByMask(band5Valid, pixelQAMask)

        # Extracts the Radsat Mask
        outExtractByMask2 = ExtractByMask(outExtractByMask1, radsatMask)

        arcpy.ProjectRaster_management(outExtractByMask2, rasterPrj,
                                       projection, "NEAREST", "#",
                                       "#", "#", "#")

        # Clips the raster to the subset determined by the user
        # Essential step due to the change in WRS and satellite path changes
        arcpy.Clip_management(rasterPrj, extentFormat, subsetArea, "#",
                              "#", "NONE", "NO_MAINTAIN_EXTENT")

        # Create Cloud Layer for Cloud Score
        arcpy.BuildRasterAttributeTable_management(name_pixel_qa, "Overwrite")
        attExtract = ExtractByAttributes(name_pixel_qa, cloudEuclideanCO)
        attExtract = ExtractByMask(attExtract, clipRaster, "INSIDE")
        attExtract.save(os.path.join(path, BAPFolder, bap_cloudScore,
                                     cloudScorePRe, sensor
                                     + dateNum + "_CLRaw.tif"))

        # Copy the Opacity QA file for BAP Score processing
        OpacityS = ExtractByMask(aerosol_qa, clipRaster, "INSIDE")
        OpacityS.save(QAFile)

        # Delete variables and objects no longer in use
        del outExtractByMask1, outExtractByMask2, folder, outInt1, \
            outInt2, attExtract, name_band5

# Pre-processes images in the TM and ETM+ Sensor List
def preprocessLS457(LS457List):
    # SQL Where Clauses
```

```
# Cloud QA Where Clause for Landsat 4-5 & 7 (ETM+ and TM)
whereClause1 = "Value = 8 OR Value = 2 OR Value = 34 OR Value = 4 OR" \
            " Value = 12 OR Value = 20 OR Value = 36 OR " \
            "Value = 52 OR Value = 16 OR Value = 24 OR Value = 48" \
            " OR Value = 56 OR Value = 32 "
# Radsat Where Clause for Landsat 4-5 & 7 (ETM+ and TM) and Landsat 8
# (OLI/TIRS)
whereClause2 = "Value = 32"
# Valid Pixel Range for Landsat 4-5 & 7 (ETM+ and TM) and Landsat 8
# (OLI/TIRS)
whereClause3 = "VALUE < 10000 And VALUE > 0"
# Cloud Shadow BAP Prep File Where Clause
cloudEuclideanET = "Value = 72 OR Value = 136 OR Value = 96 OR" \
                " Value = 112 OR Value = 160 OR Value = 176 OR " \
                "Value = 224 "
inFalseRaster = 1

# Get extent of the raster to be used to subset the data
extentRaster = Raster(clipRaster)
topExtent = arcpy.GetRasterProperties_management(extentRaster,
                                        "TOP").getOutput(0)
leftExtent = arcpy.GetRasterProperties_management(extentRaster,
                                         "LEFT").getOutput(0)
rightExtent = arcpy.GetRasterProperties_management(extentRaster,
                                          "RIGHT").getOutput(0)
bottomExtent = arcpy.GetRasterProperties_management(extentRaster,
                                           "BOTTOM").getOutput(0)
extentFormat = str(leftExtent + " " + bottomExtent + " " +
                rightExtent + " " + topExtent)

print("\n" + "The following TM and ETM+ Images have been Preprocessed:")

# Pre-process each image in the Landsat 457 List
for folder in LS457List:
    print(folder)

    # Changes workspace to each image folder provided by the user
    currentFolder = os.path.join(inputPath, folder)
    env.workspace = currentFolder

    # Determines the unique product identifier for each image
    rasterName = os.listdir(currentFolder)[0]
    productIdentifier = rasterName[:40]
    dateNum = productIdentifier[17:25]
    sensor = productIdentifier[:4]
```

```
# Generates product identifier specific name files for input
# data depending on USGS formatting
if rasterName[7:9] == "SP":
   # Pixel QA Where Clause for Landsat 4-5 & 7 (ETM+ and TM)
   whereClause11 = "Value <> 5440"
   name_radsat_qa = rasterName[:40] + "_QA_RADSAT.TIF"
   name_cloud_qa = rasterName[:40] + "_SR_CLOUD_QA.TIF"
   name_pixel_qa = rasterName[:40] + "_QA_PIXEL.TIF"
   name_band5 = rasterName[:40] + "_SR_B" + TM_ETMBand + ".TIF"
   name_opacity = rasterName[:40] + "_SR_ATMOS_OPACITY.TIF"
   rasterPrj = "memory\\" + "productIdentifier"

else:
   # Generate product identifier specific file names for input data
   # Pixel QA Where Clause for Landsat 4-5 & 7 (ETM+ and TM)
   whereClause11 = "Value <> 66 AND Value <> 130"
   name_cloud_qa = productIdentifier + "_sr_cloud_qa.tif"
   name_radsat_qa = productIdentifier + "_radsat_qa.tif"
   name_pixel_qa = productIdentifier + "_pixel_qa.tif"
   name_band5 = productIdentifier + "_sr_band" + TM_ETMBand + ".tif"
   name_opacity = productIdentifier + "_sr_atmos_opacity.tif"
   rasterPrj = "memory\\" + "productIdentifier"

# Create file names for saving intermediate results
subsetArea = os.path.join(path, preprocessedImages, sensor +
                  dateNum + ".tif")

# Intermediate File paths
attextractpath = os.path.join(path, BAPFolder, bap_cloudScore,
                     cloudScorePRe, sensor + dateNum
                     + "_CLRaw.tif")
rawOpacPathPre = os.path.join(path, BAPFolder, bap_opacityScore,
                     opacityScorePre,
                     sensor + dateNum + "_OpacRaw.tif")

# Prepare Cloud Mask
# Mask cloud, cloud shadow, water, snow, and cloud/water classes
outInt1 = Int(name_cloud_qa)
cloudMask = SetNull(outInt1, inFalseRaster, whereClause1)

# Prepare Radsat Mask
# Removes unsaturated pixels in Band 5
outInt2 = Int(name_radsat_qa)
radsatMask = SetNull(outInt2, inFalseRaster, whereClause2)

# Prepare Band 5
```

```
# Removes invalid pixel values (Greater than 1000 and less than 0)
band5Valid = arcpy.sa.Con(name_band5, name_band5, 0, whereClause3)
band5Valid = Raster(band5Valid) * 0.0001

# Prepare Pixel QA
# Removes non-clear pixels
outInt4 = Int(name_pixel_qa)
pixelQAMask = SetNull(outInt4, inFalseRaster, whereClause11)

# Extract by Mask
# Completed for both the Radsat Mask, Pixel QA Mask and the Cloud
# Mask individually instead of merging the masks due to the impact of
# the NoData values on raster math
outExtractByMask1 = ExtractByMask(band5Valid, cloudMask)
outExtractByMask2 = ExtractByMask(outExtractByMask1, pixelQAMask)
outExtractByMask3 = ExtractByMask(outExtractByMask2, radsatMask)
arcpy.ProjectRaster_management(outExtractByMask3, rasterPrj, projection,
                        "NEAREST", "#", "#", "#", "#")

# Clips the raster to the subset determined by the user
# Essential step due to the change in WRS and satellite path changes
arcpy.Clip_management(rasterPrj, extentFormat, subsetArea, "#", "#",
                  "NONE", "NO_MAINTAIN_EXTENT")

# Create Cloud Layer for BAP Cloud Score
arcpy.BuildRasterAttributeTable_management(name_pixel_qa, "Overwrite")
attExtract = ExtractByAttributes(name_pixel_qa, cloudEuclideanET)
attExtract = ExtractByMask(attExtract, clipRaster, "INSIDE")
attExtract.save(attextractpath)

# Create Opacity Layer for BAP Opacity Score
rawOpacity = Raster(name_opacity)
arcpy.management.CalculateStatistics(rawOpacity)
if int(arcpy.GetRasterProperties_management(rawOpacity, "ALLNODATA").
            getOutput(0)) == 0:
   rawOpacityS = ExtractByMask(rawOpacity, clipRaster, "INSIDE")
else:
   rawOpacityS = CreateConstantRaster(0, "FLOAT", 30, extentFormat)
rawOpacityS = Float(Raster(rawOpacityS)) * 0.0010
arcpy.management.DefineProjection(rawOpacityS, projection)
rawOpacityS.save(rawOpacPathPre)

# Delete variables and objects no longer in use
del outExtractByMask1, outExtractByMask2, outExtractByMask3, folder, \
   outInt1, outInt2, outInt4, \
   attExtract, name_band5
```

```
# Pre-processes all items in the MSS Sensor list
def preprocessLS12345(LS12345List, inpath3):
    # SQL Where Clauses
    # QA Band Where Clause for Landsat 1-5 (MSS)
    whereClause9 = "Value = 2 OR Value = 34 OR Value = 76 OR Value = 124 OR " \
                   "Value = 112 OR Value = 116 " \
                   "OR Value = 120 OR Value = 124 OR Value = 40 OR Value = 36"
    # Cloud Shadow BAP Prep File Where Clause
    cloudEuclideanM = "Value = 124 OR Value = 112 OR Value = 116 OR Value = 120"
    inFalseRaster = 1

    # Get extent of the raster to be used to subset the data
    extentRaster = Raster(clipRaster)
    topExtent = arcpy.GetRasterProperties_management(extentRaster, "TOP")\
       .getOutput(0)
    leftExtent = arcpy.GetRasterProperties_management(extentRaster, "LEFT")\
       .getOutput(0)
    rightExtent = arcpy.GetRasterProperties_management(extentRaster, "RIGHT")\
       .getOutput(0)
    bottomExtent = arcpy.GetRasterProperties_management(extentRaster, "BOTTOM")\
       .getOutput(0)
    extentFormat = str(leftExtent + " " + bottomExtent + " " + rightExtent
                       + " " + topExtent)

    # Check to ensure that the metadata file is present for each MSS folder
    # Missing metadata files results in pre-processing errors
    directory = os.listdir(inpath3)
    for folder3 in directory:
       if folder3 != outerFile and folder3[1] == "M":
          metacheckpath = os.path.join(inpath3, folder3)
          metacheck = os.listdir(metacheckpath)
          substringMTL = "_MTL.txt"
          subMTL = any(substringMTL in string for string in metacheck)
          if subMTL is False:
             os.remove(metacheckpath)

    print("\n" + "The following MSS Images have been Preprocessed:")

    # Pre-process each image in the Landsat 12345 List
    for folder in LS12345List:
       print(folder)
       # Changes workspace to each image folder provided by the user
       currentFolder = os.path.join(inputPath, folder)
       env.workspace = currentFolder
```

```
# Determines the unique product identifier for each image
rasterName = os.listdir(currentFolder)[0]
productIdentifier = rasterName[:40]
dateNum = productIdentifier[17:25]
sensor = productIdentifier[:4]
satNum = productIdentifier[2:4]
rasterPrj = "memory\\" + "productIdentifier"
subsetArea = os.path.join(path, preprocessedImages, sensor +
                    dateNum + ".tif")

# Generate product identifier specific file names for input data
name_cloud_qa = productIdentifier + "_BQA.tif"

# Determine correct band based on MSS Satellite number (shift in band
# labelling after LS3)
if satNum == "01" or satNum == "02" or satNum == "03":
   name_band5 = productIdentifier + "_B" + MSSE + ".tif" # 5
elif satNum == "04" or satNum == "05":
   name_band5 = productIdentifier + "_B" + MSSL + ".tif" #2

# Determine TOA rescale values based on MTL file
fpath = Raster(os.path.join(inputPath, productIdentifier, name_band5))

# Prepare Pixel QA Mask
outInt1 = Int(name_cloud_qa)
OutRas = ExtractByAttributes(outInt1, whereClause9)

# Extract by Mask for Pixel QA Mask
outExtractByMask1 = ExtractByMask(fpath, OutRas, "OUTSIDE")
arcpy.ProjectRaster_management(outExtractByMask1, rasterPrj, projection,
                        "NEAREST", "#", "#", "#", "#")
arcpy.Clip_management(rasterPrj, extentFormat, subsetArea, "#", "#",
                  "NONE", "NO_MAINTAIN_EXTENT")
OutRas = SetNull(subsetArea, subsetArea, "Value = 65535 OR Value < 0")
OutRas.save(subsetArea)

# Create Cloud Layers for Cloud Score
OPPathPre = os.path.join(path, BAPFolder, bap_cloudScore, cloudScorePRe,
                    sensor + dateNum + "_CLRaw.tif")
OPPathPost = os.path.join(path, BAPFolder, bap_opacityScore,
                     opacityScorePre,
                     sensor + dateNum + "_ProxOpac.tif")
arcpy.BuildRasterAttributeTable_management(name_cloud_qa, "Overwrite")
attExtract = ExtractByAttributes(name_cloud_qa, cloudEuclideanM)
attExtract.save(OPPathPre)
attExtract.save(OPPathPost)
```

```
# Create a dictionary in the format of:
# yearDictionary = {"Lyear": ["imgID", "imgID", "imgID"], "Lyear":
# ["imgID", "imgID", "imgID"]...}
def moasicdictionary():
    # Changes workspace to the projected rasters folder
    currentFolder4 = os.path.join(path, preprocessedImages)
    env.workspace = currentFolder4

    noDataAppend = []

    # Create empty dictionary
    yearDictionary = {}

    # Lists all rasters in the Projected Raster Folder, reclassifies raster
    # to a value of 1, determines count
    # to calculate area, removes all rasters with less than 50% coverage to avoid
    # pixels amidst lots of cloud or snow cover
    rastListOrg = arcpy.ListRasters("*", "TIF")
    for r in rastListOrg:
        # Determines if the raster clipped to the subset include no data
        noData = arcpy.GetRasterProperties_management(r, "ALLNODATA")
        noDataR = noData.getOutput(0)
        # Appends the images with no data to a list to remove
        if noDataR == "1":
            noDataAppend.append(r)
        else:
            # If the image has valid pixels, calculate the percentage of the
            # study area covered
            table = os.path.join(path, logicRules, "Table.dbf")
            maxR = arcpy.GetRasterProperties_management(r, "MAXIMUM")\
                .getOutput(0)
            colR = arcpy.GetRasterProperties_management(r, "COLUMNCOUNT")\
                .getOutput(0)
            rowR = arcpy.GetRasterProperties_management(r, "ROWCOUNT")\
                .getOutput(0)
            outArea = Reclassify(r, "Value", RemapRange([[1, maxR, 1]]))
            outArea.save(os.path.join(path, logicRules, "area.tif"))
            arcpy.conversion.ExportTable(outArea, table)
            counter = 0
            for row in arcpy.da.SearchCursor(table, ["COUNT"]):
                if counter == 0:
                    numPixels = row[0]
            totalPixels = float(colR) * float(rowR)
            percentCov = (numPixels / totalPixels) * 100
            del outArea, maxR, colR, rowR, outArea, totalPixels, percentCov
```

```
    # Append images with less than 50% coverage to a list to remove
    if percentCov < 50:
        noDataAppend.append(r)

print("\n",
      'The following images have been removed because after preprocessing '
      'all pixels are NoData or have less than'
      ' 50% area coverage:')

for p in noDataAppend:
    arcpy.management.Delete(p)
    print(p)
print()

# List all remaining images and list input years
numlist = []
rastListOrg = arcpy.ListRasters("*", "TIF")
for r in rastListOrg:
    numlist.append(int(r[4:8]))

# Determine the earliest and the latest year based on the input images
fyear = min(numlist)
lyear = max(numlist)

# Creates list of the study year range. The last year must be one greater
# because using the range alone results in the exclusion of the final year
yearList = range(fyear, lyear + 1)

# Creates a dictionary with the value being an empty list to be filled
# with the file names of the raster, the key is the year
for yo in yearList:
    yearDictionary["L" + str(yo)] = []

# For each image in the image list, determines the year of the raster
# For each year that is available in the dictionary, creates a year
# ID for the image. If the year ID of the raster matches the key of the
# dictionary, assigns the image to the list. Creates an iterative solution
# to sort every
# raster file name into a list by year
for i in rastListOrg:
    imgYear = i[4:8]
    for key in yearDictionary:
        yearID = "L" + imgYear
        if key == yearID:
            yearDictionary[key] = yearDictionary.setdefault(key, []) + [i]
```

```
    print("Image dates for each year:")
    print(yearDictionary, "\n")

    print("*Revised* Number of Images Per Year:")
    for key, value in yearDictionary.items():
        print(key, len([item for item in value if item]))

    return yearDictionary, yearList

################################################################################
# Based on White et al. (2014) with modifications to process MSS data

# Day of Year Score
def doyscore(yeardictionary, yearlist):
    print("\n" + "BAP Image Processing:")
    # Determine Day of Year for Each Image
    DOYList = []
    for yp in yearlist:
        keyL = "L" + str(yp)
        for item in yeardictionary[keyL]:
            date_val = date(int(item[4:8]), int(item[8:10]), int(item[10:12]))
            day_of_year = date_val.strftime('%j')
            DOYList.append(int(day_of_year))
            del item

    # Target Day of Year (September 1)
    # Selected to achieve limited foliage prior to snow cover
    DOYMean = float(245)

    # Determine standard deviation of all the input days of the year
    DOYSTD = float(numpy.std(DOYList))

    # Append all DOY scores to a list to determine max DOY for normalization
    DOYScoreList = []
    for yl in yearlist:
        keyL = "L" + str(yl)
        for item in yeardictionary[keyL]:
            date_val = date(int(item[4:8]), int(item[8:10]), int(item[10:12]))
            day_of_year = float(date_val.strftime('%j'))
            lower = DOYSTD * math.sqrt(2.0 * math.pi)
            mult = math.e ** ((-1/2) * (((day_of_year - DOYMean)/DOYSTD)
                                ** 2.00))
            DOY_Score = (1/lower) * mult
            DOYScoreList.append(DOY_Score)
            del item
```

```
    # Maximum available score
    DOYScore_Max = max(DOYScoreList)

    # Iterate through the rasters using the year dictionary sublists
    for yq in yearlist:
        keyL = "L" + str(yq)
        for item in yeardictionary[keyL]:
            # Calculate the DOY for each raster
            date_val = date(int(item[4:8]), int(item[8:10]), int(item[10:12]))
            day_of_year = float(date_val.strftime('%j'))

            # Calculate DOY Score
            lower = DOYSTD * math.sqrt(2.0 * math.pi)
            mult = math.e ** ((-1/2) * (((day_of_year - DOYMean)/DOYSTD)
                               ** 2.00))
            DOY_Score = (1/lower) * mult

            # Normalizes the score between 0 and 1 based on the maximum and
            # minimum values
            DOY_Score_Scale = DOY_Score / DOYScore_Max

            # Create a constant raster for each input image with the assigned
            # DOY Score
            inputExtent = arcpy.sa.Raster(item).extent
            inputCellX = arcpy.GetRasterProperties_management(item,
                                                "CELLSIZEX").\
                getOutput(0)
            DOYRaster_Name = str(item[:12]) + "_DOY.tif"
            DOY_Raster = CreateConstantRaster(DOY_Score_Scale, "FLOAT",
                                      inputCellX, inputExtent)
            DOY_Raster.save(os.path.join(path, BAPFolder, bap_yearScore,
                                    yearScorePost, DOYRaster_Name))
            arcpy.DefineProjection_management(os.path.join(path, BAPFolder,
                                        bap_yearScore, yearScorePost,
                                        DOYRaster_Name), projection)

            # Delete variables and objects no longer in use
            del item, DOY_Raster, inputExtent, inputCellX, date_val, \
                day_of_year, DOY_Score_Scale

    print("DOY Score Complete")

# Scan Line Corrector Failure/ Sensor Score
def sensorscore(yeardictionary2, yearlist2):
    # Date of SLC Failure
```

```
    SLC_Failure = date(2003, 0o5, 31)

    # Iterate through the rasters using the year dictionary sublists
    for ym in yearlist2:
        keyL = "L" + str(ym)
        for item in yeardictionary2[keyL]:
            # Determine the year for each image
            date_val = date(int(item[4:8]), int(item[8:10]), int(item[10:12]))
            sensor_val = str(item[1:2])

            # Score the rasters based on their sensor date ranges
            if sensor_val == "C" or sensor_val == "O":
                sensor_score = 1
            elif sensor_val == "T":
                sensor_score = 1
            elif sensor_val == "E":
                if date_val >= SLC_Failure:
                    sensor_score = 0.5
                elif date_val < SLC_Failure:
                    sensor_score = 1
            elif sensor_val == "M":
                sensor_score = 0.5

            # Determine the raster extent to be used in creating a constant
            # raster for the score
            inputExtent = arcpy.sa.Raster(item).extent
            inputCellX = arcpy.GetRasterProperties_management(item,
                                        "CELLSIZEX").getOutput(0)
            SensorRaster_Name = str(item[:12]) + "_Sens.tif"

            # Create constant raster and assign sensor score
            Sensor_Raster = CreateConstantRaster(sensor_score, "FLOAT",
                                        inputCellX, inputExtent)
            Sensor_Raster.save(os.path.join(path, BAPFolder, bap_sensorScore,
                                    sensorScorePost, SensorRaster_Name))
            arcpy.DefineProjection_management(
                os.path.join(path, BAPFolder, bap_sensorScore, sensorScorePost,
                        SensorRaster_Name), projection)

            # Delete variables and objects no longer in use
            del item, Sensor_Raster, SensorRaster_Name, inputCellX, \
                inputExtent, sensor_score

    print("Sensor Score Complete")

# Distance to Cloud or Cloud Shadow Score
```

```
def cloudshadow():
    # Changes environment to BAP Cloud Score – Pre-Analysis
    toProcess = os.path.join(path, preprocessedImages)
    env.workspace = toProcess
    cloudRast = arcpy.ListRasters("*", "TIF")

    currentFolder = os.path.join(path, BAPFolder, bap_cloudScore, cloudScorePRe)
    env.workspace = currentFolder

    # Calculate the cloud score for each image in the cloud score
    # preprocessing folder
    for ras in cloudRast:
        ras = ras[:-4] + "_CLRaw.tif"
        # Extent, cell size, min and max of each raster is obtained
        extentTop = arcpy.GetRasterProperties_management(ras, "TOP")
        extentLeft = arcpy.GetRasterProperties_management(ras, "LEFT")
        extentRight = arcpy.GetRasterProperties_management(ras, "RIGHT")
        extentBottom = arcpy.GetRasterProperties_management(ras, "BOTTOM")
        cellSize = arcpy.GetRasterProperties_management(ras, "CELLSIZEX")

        # Create intermediate save paths
        cloudScore = os.path.join(path, BAPFolder, bap_cloudScore,
                             cloudScorePost, str(ras[:-10]) + "_ScoCl.tif")
        preRe = os.path.join(path, BAPFolder, bap_cloudScore, cloudScorePRe,
                        str(ras[:-4]) + "PrRe.tif")
        sensor = ras[1:2]

        # Determine if there are no clouds present within the raster
        tim = int((arcpy.GetRasterProperties_management(ras,
                                        "ALLNODATA")).getOutput(0))

        # If there is cloud coverage within the raster, conduct the
        # follow BAP score workflow
        if tim == 0:

            # Determine the statistics for the Remap Range and Value Objects
            minValue = int(arcpy.Raster(ras).minimum)
            maxValue = int(arcpy.Raster(ras).maximum)
            newValue = int(0)

            # Remap Value and Range Objects created based on specific
            # raster values
            remapRange = RemapRange([[minValue, maxValue, newValue]])

            # Create extent object from each input raster
            arcpy.env.extent = arcpy.Extent(extentLeft, extentBottom,
```

```
                        extentRight, extentTop)

# Conduct Euclidean Distance for each cloud and cloud shadow layer
outEuc = EucDistance(ras, "", cellSize, "")
outEuc.save(os.path.join(path, BAPFolder, bap_cloudScore,
                    cloudScorePRe,
                    str(ras[:-4]) + "Euc.tif"))

# Reclassify the pixels based on the distance from the clouds
# Pixels identified as cloud or cloud shadow are assigned “NO DATA”
# Pixels greater than 50 pixels away from clouds are assigned a
# score of 1. Pixels that are between 0 and 50 pixels away are
# assigned a score between 0 and 1 using the formula

if sensor == "M":
   maskedEucl = Divide(outEuc, 60)
else:
   maskedEucl = Divide(outEuc, 30)

maskedEucl.save(os.path.join(path, BAPFolder, bap_cloudScore,
                      cloudScorePRe,
                      str(ras[:-4]) + "EucDiv2.tif"))

# Reclassify all clouds to a value of 1
reclassCloud = Reclassify(ras, "VALUE", remapRange)

# Use the mask layer to clip out the area of no data
# Pixels identified as cloud or cloud shadow are
# assigned “NO DATA”
maskedEucl2 = ExtractByMask(maskedEucl, reclassCloud,"OUTSIDE")

# Iterates through pixels and pixels
# that are between 0 and 50 pixels away.
# Pixels are assigned a score between 0 and 1 using the formula.
# Pixels greater than 50 are assigned a 1
rasObj = Raster(maskedEucl2)

rasObj.readOnly = False
for i, y in rasObj:
   if rasObj[i, y] <= 50 and rasObj[i, y] > 0:
      di = rasObj[i, y]
      scoreCalc1 = (50 - 0) / 2
      scoreCalc2 = min(di, 50)
      scoreCalc3 = scoreCalc2 - scoreCalc1
      scoreCalc4 = -0.2 * scoreCalc3
      scoreCalc5 = math.e ** scoreCalc4
```

```
                scoreCalc6 = 1 / (1 + scoreCalc5)
                rasObj[i, y] = scoreCalc6
            elif rasObj[i, y] > 50:
                rasObj[i, y] = 1
        rasObj.save(cloudScore)

        # Delete variables and objects no longer in use
        del extentLeft, extentBottom, extentRight, extentTop, cellSize,\
            minValue, maxValue, outEuc, reclassCloud

    # If there is no cloud coverage, create a constant raster of 1
    else:
        noData = CreateConstantRaster(1, "FLOAT", 30,
                                      arcpy.Extent(extentLeft,
                                      extentBottom, extentRight, extentTop))
        noData.save(cloudScore)
        arcpy.management.DefineProjection(cloudScore, projection)

        # Delete variables and objects no longer in use
        del noData
print("Cloud Shadow Score Complete")

# Opacity Score
def opacityscore():
    # Low, Medium and High Opacity Where Clause Lists
    lowOpacList = [66, 68, 72, 80, 96, 100]
    medOpacList = [130, 132, 136, 144, 160, 164]
    highOpacList = [194, 196, 200, 208, 224, 228]

    # Empty list to iteratively generate accurate Where Clauses
    whereClauseList = []

    currentFolder = os.path.join(path, preprocessedImages)
    env.workspace = currentFolder
    opacRast = arcpy.ListRasters("*", "TIF")

    # List all rasters in the pre-processing folder for the opacity score
    currentFolder = os.path.join(path, BAPFolder, bap_opacityScore,
                                 opacityScorePre)

    env.workspace = currentFolder

    # Calculate the opacity score for each of the rasters in the opacity
    # pre-processing folder
    for pr in opacRast:
        # Determine sensor to apply different Opacity scoring calculations
```

```
# for each sensor group
sensorDate = str(pr[4:-4])
sensor_val = str(pr[1:2])
sensorFull = str(pr[0:4])

if sensorFull == "LM01" or sensorFull == "LM02" or sensorFull == \
      "LM03" or sensorFull == "LM04" or sensorFull == "LM05":
   ras = sensorFull + sensorDate + "_ProxOpac.tif"
else:
   ras = sensorFull + sensorDate + "_OpacRaw.tif"

# Determine if any opacity rasters are ALLNODATA (TM and ETM+ clear)
# If the rasters are clea,r they must be processed separately as to not
# cause an error
tim = int((arcpy.GetRasterProperties_management(ras, "ALLNODATA"))
        .getOutput(0))

# Extent, cell size, min and max of each raster is obtained
extentTop = arcpy.GetRasterProperties_management(ras, "TOP")
extentLeft = arcpy.GetRasterProperties_management(ras, "LEFT")
extentRight = arcpy.GetRasterProperties_management(ras, "RIGHT")
extentBottom = arcpy.GetRasterProperties_management(ras, "BOTTOM")
extentRas = Extent(extentLeft, extentBottom, extentRight, extentTop)
cellSize = arcpy.GetRasterProperties_management(ras, "CELLSIZEX")

# Opacity Scoring for TM and ETM+
# Less than 0.2 assigned 1
# Greater than 0.3 assigned “NO DATA”
# Between 0.2 and 0.3 assigned a score between 1 and 0 based
# on the formula
if sensor_val == "T" or sensor_val == "E":

   # *0.0010
   # If the sky is not clear, calculate the opacity score using the
   # workflow below
   if tim == 0:
      opacRaster = Raster(os.path.join(path, BAPFolder,
                  bap_opacityScore, opacityScorePre, ras))
      opacScore = os.path.join(path, BAPFolder, bap_opacityScore,
                  opacityScorePost,
                  sensorFull + sensorDate + "_OpacScore.tif")

      # Calculate raster statistics
      arcpy.CalculateStatistics_management(opacRaster)

      # Iterates through raster to conditionally
```

```
        # apply reclassification for each pixel for clear
        # and average pixels
        opacRaster.readOnly = False
        for i, y in opacRaster:
            Oi = opacRaster[i, y]
            if 0.3 >= opacRaster[i, y] > 0.2:
                opEquation1 = (0.3 - 0.2) / 2
                opEquation2 = min(Oi, 0.3)
                opEquation4 = -0.2 * (opEquation2 - opEquation1)
                opEquation5 = 1 + math.e ** opEquation4
                opEquation3 = 1 - (1 / opEquation5)
                opacRaster[i, y] = opEquation3
            elif opacRaster[i, y] > 0.3:
                opacRaster[i, y] = math.nan
            elif opacRaster[i, y] < 0.2:
                opacRaster[i, y] = 1
        opacRaster.save(opacScore)

        # Delete variables and objects no longer in use
        del opacRaster

    # If the sky is clear, create a constant raster with a value of 1
    elif tim == 1:
        opacScore = os.path.join(path, BAPFolder, bap_opacityScore,
                                 opacityScorePost, sensorFull +
                                 sensorDate + "_OpacScore.tif")

        noData = CreateConstantRaster(1, "FLOAT", 30,
                                      arcpy.Extent(extentLeft,
                                      extentBottom, extentRight, extentTop))
        noData.save(opacScore)
        arcpy.management.DefineProjection(opacScore, projection)

        # Delete variables and objects no longer in use
        del noData

# TIRS/OLI Opacity
# Based on Opacity Values found in LaSRC (Landsat 8 Surface
# Reflectance Product Guide)
if sensor_val == "C" or sensor_val == "O":
    opacScore = os.path.join(path, BAPFolder, bap_opacityScore,
                             opacityScorePost,
                             sensorFull + sensorDate + "_OpacScore.tif")

    # Convert the input raster into a raster object and calculate
    # statistics
```

```
opacRaster = Raster(ras)
arcpy.CalculateStatistics_management(opacRaster)

# Create empty lists for the true Remap values
lowOpacRemap = []
medOpacRemap = []
highOpacRemap = []
rastObj2 = Raster(opacRaster)

# Create a table to determine which opacity values are within the
# Attribute Table
attTable = arcpy.BuildRasterAttributeTable_management(opacRaster,
                                        "Overwrite")
rasTablePath = os.path.join(path, BAPFolder, bap_opacityScore,
opacityScorePost)
rasTable = arcpy.TableToTable_conversion(attTable, rasTablePath,
                                sensorDate + ".dbf",
                                "", "", "")

# Iterate through the table to create lists for the remap values
searchCurs = arcpy.da.SearchCursor(attTable, ["Value"])
for row in searchCurs:
   whereClauseList.append(row[0])
for i in lowOpacList:
   if i in whereClauseList:
      lowOpacRemap.append([i, 1])
for i in medOpacList:
   if i in whereClauseList:
      medOpacRemap.append([i, 0])
for i in highOpacList:
   if i in whereClauseList:
      highOpacRemap.append([i, "NODATA"])

# Concatenate the formatted remap values
remapComb = (lowOpacRemap + medOpacRemap + highOpacRemap)
# Reclassify the opacity values
opReclass = Reclassify(rastObj2, "Value", RemapValue(remapComb),
                  "NODATA")

# Workaround for issues with Remap integers
outCon = Con(opReclass, 0.5, 1, "VALUE = 0")
outCon.save(opacScore)

# Delete variables and objects no longer in use
del remapComb, rastObj2, attTable, searchCurs, opReclass, outCon
```

```
        # MSS Opacity
        # No opacity information is available
        if sensor_val == "M":
            opacScore = os.path.join(path, BAPFolder, bap_opacityScore,
                                     opacityScorePost,
                                     sensorFull + sensorDate + "_OpacScore.tif")

            # Create constant raster with a value of 1
            outConstant = CreateConstantRaster(1, "FLOAT", cellSize, extentRas)
            arcpy.DefineProjection_management(outConstant, projection)
            outConstant.save(opacScore)
            del outConstant

        # Delete variables and objects no longer in use
        del extentTop, extentLeft, extentRight, extentBottom, extentRas, \
            cellSize

    print("Opacity Score Complete")

# Pixel Brightness Score
def brightnessscore(yearDict):
    # As roads and other urban features are usually bright pixels, favour
    # pixels with higher surface or TOA reflectance
    # Generate empty lists for max and min lists for Surface Reflectance
    # and TOA Reflectance based values
    currentFolder = os.path.join(path, preprocessedImages)
    env.workspace = currentFolder

    # Append maximum surface reflectance values for each year to a list
    for year in yearDict:
        listMax = []
        for i in yearDict[year]:
            listMax.append(arcpy.management.GetRasterProperties(i, "MAXIMUM")
                           .getOutput(0))
        if len(listMax) > 0:
            maxVal = float(max(listMax))
        else:
            maxVal = 0

        # For each raster, assign a brightness DOY score
        for ras in yearDict[year]:
            arcpy.CalculateStatistics_management(ras)
            rasBright = Raster(ras)
            rasBrightF = Float(rasBright)
            normalizedBright = rasBrightF / maxVal
            brightSavePath = os.path.join(path, BAPFolder, bap_brightnessScore,
```

```
                                ras[:-4] + "_Br.tif")
            normalizedBright.save(os.path.normpath(brightSavePath))
            del normalizedBright

# Generate BAP Composites
def finalbap(yeardictionarybap):
    # Calculate pixel scores for each image
    # Use the dictionary to iterate through each year in the study range
    arcpy.env.extent = clipRaster

    for year in yeardictionarybap:
        yearScoreList = yeardictionarybap[year]
        # Iterate through each raster within any given year
        for e in yearScoreList:
            arcpy.env.extent = clipRaster
            # Determine the file paths for each input pixel score component
            brightF = os.path.join(path, BAPFolder, bap_brightnessScore,
                                   e[:-4] + "_Br.tif")
            cloudF = os.path.join(path, BAPFolder, bap_cloudScore,
                                  cloudScorePost, e[:-4] + "_ScoCl.tif")
            opacityF = os.path.join(path, BAPFolder, bap_opacityScore,
                                    opacityScorePost, e[:-4] + "_OpacScore.tif")
            sensorF = os.path.join(path, BAPFolder, bap_sensorScore,
                                   sensorScorePost, e[:-4] + "_Sens.tif")
            yearF = os.path.join(path, BAPFolder, bap_yearScore, yearScorePost,
                                 e[:-4] + "_DOY.tif")

            # Generate raster objects for each input score layer
            brightFRas = Raster(brightF)
            cloudFRas = Raster(cloudF)
            opacityFRas = Raster(opacityF)
            sensorFRas = Raster(sensorF)
            yearFRas = Raster(yearF)

            # Add the scores together to generate a pixel score layer for
            # each image
            scoreFinal = (0.2 * sensorFRas) + (0.2 * cloudFRas) + \
                         (0.2 * opacityFRas) + (0.2 * yearFRas) + \
                         (0.2 * brightFRas)
            nullRas = IsNull(scoreFinal)
            nullRas.save(os.path.join(path, BAPFolder, bap_preprocess, e[:-4]
                                      + "_PixelNULL.tif"))
            outCon = Con(nullRas, 0, scoreFinal, "Value = 1")
            outCon.save(os.path.join(path, BAPFolder, bap_preprocess, e[:-4]
                                     + "_PixelScore.tif"))
```

```
# Use the dictionary to iterate through each year in the study range
for year in yeardictionarybap:
    # Empty lists created for compositing process
    arcpy.env.extent = clipRaster
    yearScoreList = yeardictionarybap[year]

    outHighestList = []
    for y in yearScoreList:
        outHighestList.append(os.path.join(path, BAPFolder, bap_preprocess,
                                    y[:-4] + "_PixelScore.tif"))

    pickList = []
    for y in yearScoreList:
        pickList.append(os.path.join(path, preprocessedImages, y))

    # If there are at least two images per year, execute the compositing
    # process
    if len(yearScoreList) > 1:
        arcpy.env.extent = clipRaster
        # Determine on a cell by cell basis which raster has the highest
        # value for each year
        outHighest = arcpy.sa.HighestPosition(outHighestList)
        outHighest.save(os.path.join(path, BAPFolder, bap_preprocess, year
                                + "_Outhigh.tif"))

# Use output of the highest position raster to generate a BAP composite
        # layer
        pick = Pick(outHighest, pickList)
        pick.save(os.path.join(path, BAPFolder, bap_final, year + name))

        outRas = ExtractByMask(os.path.join(path, BAPFolder, bap_final,
                                year + name), clipRaster, "INSIDE")
        outRas.save(os.path.join(path, BAPFolder, bap_final, year + name))

    # If there is only one image per year, copy the raster to the output BAP
    # folder. More prevalent for pre-1984 years
    elif len(yearScoreList) == 1:
        arcpy.env.extent = clipRaster
        for single in yearScoreList:
            singlePath = os.path.join(path, preprocessedImages, single)
            arcpy.CopyRaster_management(singlePath,
                                (os.path.join(path, BAPFolder,
                                bap_final, year + name)))
            outRas = ExtractByMask(os.path.join(path, BAPFolder, bap_final,
                                year + name),clipRaster, "INSIDE")
            outRas.save(os.path.join(path, BAPFolder, bap_final,
```

```
                                year + name))

        # Clear lists and re-initialize counters
        yearScoreList.clear()
        outHighestList.clear()
        pickList.clear()

    print("BAP Composite Complete", "\n")

################################################################################
# Pre-process the input road files
def referenceprep():
    # Local file paths and variables
    counterRoad = 0
    roadlist = []
    outpath = os.path.join(path, referenceRoad)
    roadMerge = os.path.join(path, referenceRoad, "roadMerge.shp")
    preproRoad = os.path.join(path, referenceRoad, "PreProRoad.shp")
    roadBuff = os.path.join(path, referenceRoad, "roadBuffer30.shp")
    roadBuff60 = os.path.join(path, referenceRoad, "roadBuffer60.shp")
    rastRoad30 = os.path.join(path, referenceRoad, "rr30.tif")
    rastRoad60 = os.path.join(path, referenceRoad, "rr60.tif")
    zoneRaster = os.path.join(path, referenceRoad, "ZonalRaster.tif")
    singleRoad = os.path.join(path, referenceRoad, "SingularSegment.shp")

    # Set workspace environment to the reference road folder
    arcpy.env.workspace = (os.path.join(path, referenceRoad))

    for roads in roadData:
        # Project each of the reference road layers provided by the user
        outputRoad = "roads" + str(counterRoad) + ".shp"
        roadClipp = "roadsC" + str(counterRoad) + ".shp"

        arcpy.management.Project(roads, outputRoad, projection)

        # Clip the merged road layer to the study boundary
        arcpy.analysis.Clip(outputRoad, clipPoly, roadClipp)
        counterRoad += 1

        # Append all projected roads to a list for merging
        roadlist.append(outputRoad)

    # Isolate the arterial and non-arterial road from the ORN dataset
    arcpy.conversion.FeatureClassToFeatureClass("roadsC1.shp", outpath,
                                    "Arterial",
                                    "ROAD_CLASS = 'Arterial'")
```

```
arcpy.conversion.FeatureClassToFeatureClass("roadsC1.shp", outpath,
                        "NonArterial", "ROAD_CLASS <> 'Arterial'")

# Buffer the roads in the MNRF dataset
arcpy.analysis.Buffer("roadsC0.shp", "Buff.shp", "25 Meters", "FULL",
                "FLAT",
                "Dissolve all output features into a single feature")

arcpy.analysis.Erase("roadsC1.shp", "Buff.shp", "ends.shp")

arcpy.management.Merge(["ends.shp", "Arterial.shp", "roadsC0.shp"],
                 "NewRoad.shp")

# Integrate the roads at 30 m to avoid multiple lines representing the
# same road. A 30 m integrate was selected as this is the smallest spatial
# resolution present in the dataset
arcpy.management.Integrate("NewRoad.shp", "30 Meters")

# Extend roads to any line segment within 75 m to address any gaps
arcpy.edit.ExtendLine("NewRoad.shp", "75 Meters", "FEATURE")

# Convert lines into singular entities
arcpy.management.MultipartToSinglepart("NewRoad.shp", preproRoad)

# Delete any duplicate lines that were created in the integrate process
# based on their spatial location
arcpy.management.DeleteIdentical(preproRoad, ["Shape"], "30 Meters", "")

# Add fields to uses for dissolve and conversion tools
arcpy.management.CalculateField(preproRoad, "ValueNull", 0, "", "", "SHORT")
arcpy.management.CalculateField(preproRoad, "Dissolve", 0, "", "", "SHORT")

# Generate an updated road segment raster for MSS zonal statistics
arcpy.conversion.FeatureToRaster(preproRoad, "ValueNull", zoneRaster, "60")

# Create singular part road for joining
arcpy.management.Dissolve(preproRoad, singleRoad, ["Dissolve"], "",
                    "MULTI_PART", "DISSOLVE_LINES")

# Generate Raster road layers in MSS and TM/ETM+/OLI resolutions
# Use 60 m buffers to increase the information used in unsupervised
# classification
arcpy.analysis.Buffer(preproRoad, roadBuff, "60 Meters", "FULL", "FLAT",
                "ALL")
arcpy.analysis.Buffer(preproRoad, roadBuff60, "60 Meters", "FULL", "FLAT",
                "ALL")
```

```
    arcpy.conversion.PolygonToRaster(roadBuff, "FID", rastRoad30,
                                "MAXIMUM_AREA", "", 30, "BUILD")
    arcpy.conversion.PolygonToRaster(roadBuff60, "FID", rastRoad60,
                                "MAXIMUM_AREA", "", 60, "BUILD")

# Extract candidate road pixels using the reference road layers
def extractroad():
    # Required pre-generated layers
    rastRoad30 = os.path.join(path, referenceRoad, "rr30.tif")
    rastRoad60 = os.path.join(path, referenceRoad, "rr60.tif")

    # Changes the environment to the Final BAP Folder
    arcpy.env.workspace = arcpy.env.workspace = (os.path.join(path,
                                          BAPFolder, bap_final))

    # Extract the road area from the BAP images
    arcpy.env.workspace = (os.path.join(path, BAPFolder, bap_final))
    roadDelList = arcpy.ListRasters("L*", "TIF")
    for Ras in roadDelList:
        yearSlice = Ras[1:5]
        # If the data has a spatial resolution of 30 m, use the 30m
        # resolution mask (TM, ETM+, OLI data)
        if int(yearSlice) <= 1984:
            outExtractByMask5 = ExtractByMask(Ras, rastRoad60)
            # Apply convolution (smooth arithmetic mean) to the pre-1984
            # dataset to smooth the datasets
            Convolution_raster = arcpy.sa.Convolution(outExtractByMask5, 10)
            Convolution_raster.save(os.path.join(path, roadExtractSubFolder,
                                    "Ex" + str(yearSlice) + ".tif"))
            del outExtractByMask5, Convolution_raster, Ras
        if int(yearSlice) > 1984:
            outExtractByMask5 = ExtractByMask(Ras, rastRoad30)
            outExtractByMask52 = Float(outExtractByMask5)
            outExtractByMask52.save(os.path.join(path, roadExtractSubFolder,
                                      "EX" + str(yearSlice) + ".tif"))
            del outExtractByMask5, Ras

# Unsupervised road classification
def jenksopt():
    # Local variables and output file paths
    singleRoad = os.path.join(path, referenceRoad, "SingularSegment.shp")
    outputLoc = os.path.join(path, jenksOpti)
    rastRoad30 = os.path.join(path, referenceRoad, "rr30.tif")
    rastRoad60 = os.path.join(path, referenceRoad, "rr60.tif")
    brkList = []
```

```
# Change workspace for the road extraction folder
arcpy.env.workspace = (os.path.join(path, roadExtractSubFolder))

# List masked image composites
roadExtractList = arcpy.ListRasters("*", "TIF")

# Project file used to generate symbology break values
aprx = arcpy.mp.ArcGISProject("D:\RoadsProjects\Roads_23062022\
Roads_23062022.aprx")
m = aprx.listMaps()[0]

# Remove any existing layers open in the map within the given project file
d = m.listLayers()
for rmlayer in d:
   m.removeLayer(rmlayer)

# Add the rasters clipped to the buffered road to the map
for Ras in roadExtractList:
   m.addDataFromPath((os.path.join(path, roadExtractSubFolder, Ras)))
aprx.save()

# List layers within the map
l = m.listLayers()
for lyr in l:
   # File paths with iteratively created file names
   year = lyr.name[2:6]
   Remapp = os.path.join(path, jenksOpti, "NBreaks" + year + ".tif")
   Rass = os.path.join(path, roadExtractSubFolder, "Ex" + str(year) +
                       ".tif")
   polyFinal = os.path.join(path, jenksOpti, "PF" + year + ".shp")
   polyFinalAttDis = os.path.join(path, jenksOpti, "PolyFinalAD" + year +
                                  ".shp")

   # Determine the value of the Jenks Optimization classification
   # generated from the ESRI symbology
   sym = lyr.symbology
   if hasattr(sym, 'colorizer'):
      sym.updateColorizer("RasterClassifyColorizer")
      if sym.colorizer.type == "RasterClassifyColorizer":
         sym.colorizer.classificationField = 'Value'
         sym.colorizer.classificationMethod = "NaturalBreaks"
         sym.colorizer.breakCount = 4
         lyr.symbology = sym
         for brk in sym.colorizer.classBreaks:
            brkList.append(brk.upperBound)
```

```
# Reclassify the road rasters based on the Jenks classes
remappo = RemapRange([[0, brkList[0], 0], [brkList[0],
                brkList[1], 1], [brkList[1], brkList[2], 1],
                 [brkList[2], brkList[3], 1]])
outReclass = Reclassify(Rass, "Value", remappo)
outReclass.save(Remapp)
brkList.clear()

aprx.save()

# Extract the classified layer to 30 m road footprint
# Apply focal statistics on post-1984 data (Rectangular
# Neighbourhood, 4 x 4, Maximum Method)
if int(year) >= 1985:
   outFocal = ExtractByMask(outReclass, rastRoad30)
   outFocal2 = FocalStatistics(outFocal, NbrRectangle(4, 4,
                                    "CELL"), "MAXIMUM")
   outFocal2.save(os.path.join(path, jenksOpti, "onvol"
                        + str(year) + ".tif"))
# Extract the classified layer to 60 m road footprint
else:
   outFocal = ExtractByMask(outReclass, rastRoad60)
   outFocal.save(os.path.join(path, jenksOpti, "onvol" +
                       str(year) + ".tif"))

# To reduce issues with data conversion (raster to polyline,
# duplicate lines, and extension of dangles),
# first convert the raster to a polygon. Add field to convert
# the gridcode (binary) to year.
# Add variable "k" before, as this field will be used to split the
# polygon into polygon segments
arcpy.conversion.RasterToPolygon(outFocal, polyFinal, "NO_SIMPLIFY",
                         "VALUE", "SINGLE_OUTER_PART")

# Add field for the year of each classified road raster
arcpy.management.AddField(polyFinal, "TYear", "TEXT")

fieldO = ["gridcode", "TYear"]

cursorO = arcpy.da.UpdateCursor(polyFinal, fieldO)
for row in cursorO:
   if row[0] == 1:
      row[1] = "K" + year
   elif row[0] == 0:
      row[1] = "K0" + year
   cursorO.updateRow(row)
```

```
        del row, cursorO

        # Dissolve polygon to create the best possible split conditions
        # (multipart segments with dissolved lines)
        arcpy.management.Dissolve(polyFinal, polyFinalAttDis, "TYear", "",
                                  "MULTI_PART", "DISSOLVE_LINES")

        # Split major roads polyline file based on the road polygons with
        # attribute data
        arcpy.analysis.Split(singleRoad, polyFinalAttDis, "TYear",
                             outputLoc)

# Change the workspace to the Jenks Optimization folder
arcpy.env.workspace = (os.path.join(path, jenksOpti))

# Iterate through all the split files (K19**.shp) and assign year and
# binary value (1 = road, 0 = not road)
# attributes
roadSplit = arcpy.ListFeatureClasses("K*")
for ft in roadSplit:
    code = ft[1]
    if int(code) > 0:
        arcpy.management.CalculateField(ft, "TYear", int(ft[1:-4]),
                                        "PYTHON3", "", "LONG")
        arcpy.management.CalculateField(ft, "BinYear", 1, "PYTHON3", "",
                                        "LONG")
    elif int(code) == 0:
        arcpy.management.CalculateField(ft, "TYear", 0, "PYTHON3", "",
                                        "LONG")
        arcpy.management.CalculateField(ft, "BinYear", 0, "PYTHON3", "",
                                        "LONG")

# For all years with convolution rasters, iterate through the following
# logic
roadsList = arcpy.ListRasters("onvol*", "TIF")
for rra in roadsList:
    year = str(rra[5:9])
    # Layers generated through this round of iteration
    merged = os.path.join(path, jenksOpti, "Merged" + year + ".shp")
    smerged = os.path.join(path, jenksOpti, "FF" + year + ".shp")
    smerged2 = os.path.join(path, jenksOpti, "Smerged" + year + ".shp")
    finalPreRoad = os.path.join(path, jenksOpti, "FF2" + year + ".shp")

    # Define the polyline files which define the Road and Non-Road
    # polylines for each year
    valOne = os.path.join(path, jenksOpti, "K" + year + ".shp")
```

```
        valTwo = os.path.join(path, jenksOpti, "K0" + year + ".shp")
        valso = "ValsO" + year + ".shp"
        one = arcpy.Exists(valOne)
        two = arcpy.Exists(valTwo)

        # Merge the road and non-road values
        if one and two:
            arcpy.management.Merge([valOne, valTwo], merged)
        else:
            if one:
                arcpy.conversion.FeatureClassToFeatureClass(valOne,
                                    os.path.join(path, jenksOpti),
                                    "Merged" + year + ".shp")
            else:
                arcpy.conversion.FeatureClassToFeatureClass(valTwo,
                                    os.path.join(path, jenksOpti),
                                    "Merged" + year + ".shp")
        # Ensure the roads are single part
        # Apply the binary logic rules three times
        arcpy.management.MultipartToSinglepart(merged, smerged)
        logicrulesNoD(smerged, "BinYear", "Min" + str(year) + ".dbf", "FBin")
        arcpy.management.Dissolve(smerged, finalPreRoad, "FBin", None,
                            "SINGLE_PART", "UNSPLIT_LINES")
        logicrulesNoD(finalPreRoad, "FBin", "Maj2" + str(year) + ".dbf", "PBin")
        arcpy.management.Dissolve(finalPreRoad, smerged2, "PBin", None,
                            "SINGLE_PART", "UNSPLIT_LINES")
        logicrulesNoD(smerged2, "PBin", "Maj2" + str(year) + ".dbf", "KBin")

        removedisBinary()

        # Convert the polyline roads to rasters for testing purposes
        outRast1 = os.path.join(path, jenksOpti, "q" + year + ".tif")
        outRast = os.path.join(path, jenksOpti, "PP" + year + ".tif")

        if int(year) >= 1985:
            arcpy.conversion.FeatureToRaster(smerged2, "KBin", outRast1, 30)
            outFocal = FocalStatistics(outRast1, NbrRectangle(3, 3, "CELL"),
                                "MAXIMUM")
            outFocal2 = ExtractByMask(outFocal, clipRaster, "INSIDE")
            outFocal2.save(outRast)

        else:
            arcpy.conversion.FeatureToRaster(smerged2, "KBin", outRast, 60)

def difference():
    # Required file paths and variables
```

```
rastRoad30 = os.path.join(path, referenceRoad, "rr30.tif")
rastRoad60 = os.path.join(path, referenceRoad, "rr60.tif")
MSSPoly = os.path.join(path, diffFolder, "MSSFinal.shp")
MSSPolyS = os.path.join(path, diffFolder, "MSSFinalS.shp")
singleRoad = os.path.join(path, referenceRoad, "SingularSegment.shp")
MSSLine = os.path.join(path, diffFolder, "MSSLine.shp")
MSSLineS = os.path.join(path, diffFolder, "MSSLineS.shp")
output = os.path.join(path, diffFolder)
Pre1984 = os.path.join(path, finalFolder,"AllPre1984.tif")
output2 = os.path.join(path, finalFolder)
postMSSFull = os.path.join(path, finalFolder, PostMSS)
PostMSSMergeFull = os.path.join(path, finalFolder, "PMMerge.tif")
postMSSPoly = os.path.join(path, finalFolder, "PostMSSPoly.shp")
postMSSPolyS = os.path.join(path, finalFolder, "PostMSSPolyS.shp")
PostMSSLine = os.path.join(path, finalFolder, "PostMSSLine.shp")
PostMSSLineS = os.path.join(path, finalFolder, "PostMSSLineS.shp")
Pre1984Vals = os.path.join(path, finalFolder, "Pre1984Vals.tif")
reclassRoad = RemapValue([[0, "NoDATA"]])
PostMSSMerge = "PMMerge.tif"
preMSS = "PreMSS.tif"
PostMSS = "PostMSS.tif"

# Local variables and lists
removeYear = [1982, 1985, 1986]
base1974 = os.path.join(path, jenksOpti, "cc1974.tif")
base1984 = ["cc1983.tif", "cc1984.tif"]
remappo2 = RemapValue([[0, "NoData"], [1, 1], [2, 1]])

# Change workspace to Jenks Optimization folder
arcpy.env.workspace = os.path.join(path, jenksOpti)

# Generate mask of roads that were built prior to 1984
pluss = Plus(base1984[0], base1984[1])
outReclass = Reclassify(pluss, "Value", remappo2)
outReclass.save(os.path.join(path, diffFolder, "Year1984.tif"))

# List masked image composites
roadExtractList = arcpy.ListRasters("cc*", "TIF")

# Remove all years without suitable imagery for analysis
for elem in removeYear:
    val = "cc" + str(elem) + ".tif"
    roadExtractList.remove(val)

# List all remaining input extract road layers
# Complete subtraction to generate from-to-change information
```

```
lenList = len(roadExtractList)
count = 0
for road in roadExtractList:
   if count < (lenList - 1):
      year = int(road[2:6])
      minusOne = Raster(roadExtractList[count])
      minusTwo = Raster(roadExtractList[count + 1])
      nameOne = str(roadExtractList[count][2:6])
      nameTwo = str(roadExtractList[count + 1][2:6])
      outMinus = Minus(minusOne, minusTwo)
      outMinus.save(os.path.join(path, diffFolder, "MM" + nameOne +
                          "_" + nameTwo + ".tif"))

      # Apply focal statistics (Rectangular Neighbourhood, 4 x 4,
      # Majority Method)
      outFocal = FocalStatistics(outMinus, NbrRectangle(4, 4,
                                       "CELL"), "MAJORITY")

      # Extract difference regions by road footprints (30 or 60 m)
      if year > 1984:
         extraa = ExtractByMask(outFocal, rastRoad60)
      elif year <= 1984:
         extraa = ExtractByMask(outFocal, rastRoad30)
      extraa.save(os.path.join(path, diffFolder, "d" + nameOne
                          + "_" + nameTwo + ".tif"))
   elif count == (lenList - 1):
      break
   count += 1

# Change workspace to the difference folder
arcpy.env.workspace = os.path.join(path, diffFolder)

# Reclassify all change detection rasters to include only constructed roads
roadDiffList = arcpy.ListRasters("d*", "TIF")
for diff in roadDiffList:
   year = str(diff[6:10])
   remappo = RemapValue([[-1, year], [0, "NoData"], [1, "NoData"]])
   outReclass = Reclassify(diff, "Value", remappo)
   outReclass.save(os.path.join(path, diffFolder, "f" + diff[1:]))

# Reclassify the extracted 1974 roads (first year in analysis) to use in
# the final road network
year1974 = "1974"
remappoo = RemapValue([[0, "NoData"], [1, year1974]])
outReclass1974 = Reclassify(base1974, "Value", remappoo)
outReclass1974.save(os.path.join(path, diffFolder, "Year1974.tif"))
```

```
# Complete the generation of the pre-1984 road network
# List all difference rasters and mosaic them together
# Expand difference rasters by one pixel to ensure there is spatial
# alignment during mosaic
roadPlusList = arcpy.ListRasters("f*", "TIF")
lenList = len(roadPlusList)
countAgain = 0
for img in roadPlusList:
   year = int(img[6:10])
   if countAgain < (lenList - 1):
      if year <= 1984:
         plusOne = Raster(roadPlusList[countAgain])
         if countAgain == 0:
            first = "a" + str(countAgain) + ".tif"
            spando = Expand(plusOne, 1, int(year))
            arcpy.management.MosaicToNewRaster([rastRoad60, spando],
                                    output, first, projection,
                                      "16_BIT_UNSIGNED",
                                      60, 1, "MAXIMUM")
            extractt = ExtractByMask(first, rastRoad60)
            extractt.save(os.path.join(path, diffFolder, first))
         elif countAgain > 0:
            first = "a" + str(countAgain - 1) + ".tif"
            next = "a" + str(countAgain) + ".tif"
            spando2 = Expand(plusOne, 1, int(year))
            arcpy.management.MosaicToNewRaster([first, spando2],
                        output, next, projection, "16_BIT_UNSIGNED",
                        60, 1, "MAXIMUM")
            extractt2 = ExtractByMask(next, rastRoad60)
            extractt2.save(os.path.join(path, diffFolder, next))
   elif countAgain == (lenList - 1):
      break
   countAgain += 1

# List all pre-1984 mosaic images
roadMSSList = arcpy.ListRasters("a*", "TIF")
# Determine the last mosaic raster
maxMSS = max(roadMSSList)

# Mosiac the 1975-1984 change information with the roads already in
# existence
spando3 = Expand(outReclass1974, 1, int(1974))
arcpy.management.MosaicToNewRaster([maxMSS, spando3], output, preMSS,
                           projection, "16_BIT_UNSIGNED",
                            60, 1, "LAST")
```

```
extractt3 = ExtractByMask(preMSS, rastRoad60)
extractt3.save(os.path.join(path, diffFolder, preMSS))


# Convert the raster road network to a polygon to use in creating
# polyline roads
arcpy.conversion.RasterToPolygon(preMSS, MSSPoly, "NO_SIMPLIFY", "VALUE",
                                 "SINGLE_OUTER_PART")

# Dissolve polygon to create the best possible split conditions (multipart
# segments with dissolved lines)
arcpy.management.Dissolve(MSSPoly, MSSPolyS, "gridcode", "", "MULTI_PART",
                          "DISSOLVE_LINES")

# Add field for year of construction attribute
arcpy.management.AddField(MSSPolyS, "TYear", "TEXT")

fieldO = ["gridcode", "TYear"]

cursorO = arcpy.da.UpdateCursor(MSSPolyS, fieldO)
for row in cursorO:
   if str(row[0]) == "0":
      row[1] = "L0000"
   else:
      row[1] = "L" + str(row[0])
   cursorO.updateRow(row)
del row, cursorO

 # Split major roads polyline file based on the road polygons
# with attribute data
arcpy.analysis.Split(singleRoad, MSSPolyS, "TYear", output)

# Merge the road polyline segments
LSegs = arcpy.ListFeatureClasses("L*")

arcpy.management.Merge(LSegs, MSSLine, "", "ADD_SOURCE_INFO")

# Update the year of construction field to remove join information
arcpy.management.AddField(MSSLine, "TYEAR", "TEXT")

fieldO = ["MERGE_SRC", "TYear"]

cursorO = arcpy.da.UpdateCursor(MSSLine, fieldO)
for row in cursorO:
   year = row[0][-8:-4]
   row[1] = str(year)
```

```
    cursorO.updateRow(row)
del row, cursorO

arcpy.management.DeleteField(MSSLine, ["MERGE_SRC"])

# Ensure the roads are single part
# Apply the network logic rules
arcpy.management.MultipartToSinglepart(MSSLine, MSSLineS)
logicrulesNonBin(MSSLineS, "TYear", "FBin", "MSS1.dbf")

# Complete the generation of the post-1984 road network
# List all difference rasters and mosaic them together
# Expand difference rasters by one pixel to ensure there is spatial
# alignment during mosaic

# List all reclassified difference rasters
roadPlusList2 = arcpy.ListRasters("f*", "TIF")

# Reverse the list so the roads are in the latest to earliest order for the
# mosaic process
roadPlusList2.reverse()

# Ensure input mosaic list is only roads constructed post-1984
post1984List = []
for l in roadPlusList2:
    if int(l[1:5]) >= 1984:
        post1984List.append(l)

# Remove any change road segments that have less than 2 500 perimeters to
# reduce the impact of noise in the mosaic process
for i in roadPlusList2:
    # Iteratively created file paths
    year = int(i[-8:-4])
    polyRas = os.path.join(path, diffFolder, "Z" + str(year) + ".shp")
    minProc = os.path.join(path, diffFolder, "X" + str(year) + ".shp")
    minProcRas = os.path.join(path, diffFolder, "W" + str(year) + ".tif")

    # Convert the change regions to polygons
    arcpy.conversion.RasterToPolygon(i, polyRas, "", "Value")
    # Calculate the perimeter of the polygon segments
    arcpy.management.CalculateGeometryAttributes(polyRas, [["PERI",
                                    "PERIMETER_LENGTH"]], "METERS")
    # Remove road segments with a perimeter less than 2500
    arcpy.conversion.ExportFeatures(polyRas, minProc, "PERI >= 2500")

    # If there are features within the polygon change region layer,
```

```
    # then convert the polygons back into rasters
    counnttt = arcpy.management.GetCount(minProc).getOutput(0)
    if int(counnttt) > int(0):
        arcpy.conversion.FeatureToRaster(minProc, "gridcode", minProcRas,
                                         30)

# List filtered change region rasters
roadPlusList = arcpy.ListRasters("W*", "TIF")
roadPlusList.reverse()

# Create a new list of change region rasters post-1984 (some years may
# have been removed because there is not significant change)
post1984List = []
for l in roadPlusList:
    year = int(l[1:5])
    if year >= 1985 and year <= 2019:
        post1984List.append(l)

# Complete the generation of the post-1984 road network.
# List all difference rasters and mosaic them together.
# Expand difference rasters by one pixel to ensure there is spatial
# alignment during mosaic
lenList = len(post1984List)
countAgain2 = 0
for img in post1984List:
    print(img)
    year = int(img[1:5])
    if countAgain2 < (lenList - 1):
        plusOne = Raster(post1984List[countAgain2])
        plusTwo = Raster(post1984List[countAgain2 + 1])
        if countAgain2 == 0:
            first = "b" + str(countAgain2) + ".tif"
            spando2 = Expand(plusOne, 1, int(year))
            spando2.save(os.path.join(path, diffFolder, "s" + str(year) +
                                      ".tif"))

            arcpy.management.MosaicToNewRaster([spando2, plusTwo],
                                               output, first, projection,
                                               "16_BIT_UNSIGNED",
                                               30, 1, "FIRST")
            extractt2 = ExtractByMask(first, rastRoad30)
            extractt2.save(os.path.join(path, diffFolder, first))
        elif countAgain2 > 0:
            plusThree = Raster(post1984List[countAgain2 + 1])
            spando3 = Expand(plusThree, 1, int(year))
            first = "b" + str(countAgain2 - 1) + ".tif"
```

```
        next = "b" + str(countAgain2) + ".tif"
        arcpy.management.MosaicToNewRaster([first, spando3],
                                    output, next, projection,
                                    "16_BIT_UNSIGNED",
                                    30, 1, "FIRST")
        extractt3 = ExtractByMask(next, rastRoad30)
        extractt3.save(os.path.join(path, diffFolder, next))
    countAgain2 += 1

# Reclassify roads that were constructed pre-1984 to change all
# non-road pixels to NoData
re = Reclassify(Pre1984, "Value", reclassRoad)
re.save(Pre1984Vals)

# Determine the final post-1984 mosaic
listnum = []
roadLatList = arcpy.ListRasters("b*", "TIF")
for r in roadLatList:
    listnum.append(int(r[1:-4]))
print(listnum)
maxPostMSS = max(listnum)
maxPostRas = os.path.join(path, diffFolder, "b" + str(maxPostMSS) + ".tif")

# Mosaic the pre- and post-1984 raster roads together and add "0" values
# for areas in the road mask that were not identified as road
arcpy.management.MosaicToNewRaster([maxPostRas, Pre1984Vals], output2,
                            PostMSS, projection, "16_BIT_UNSIGNED",
                             30, 1, "FIRST")
arcpy.management.MosaicToNewRaster([postMSSFull, rastRoad30], output2,
                            PostMSSMerge, projection,
                            "16_BIT_UNSIGNED",
                             30, 1, "FIRST")

# Apply focal statistics on the post-1984 merged dataset (Rectangular
# Neighbourhood, 3 x 3, Majority Method)
extractt3 = FocalStatistics(PostMSSMergeFull, NbrRectangle(3, 3, "CELL"),
                        "MAJORITY", "DATA")
extractt4 = ExtractByMask(extractt3, rastRoad30)
extractt4.save(os.path.join(path, finalFolder, postMSSFull))

# Dissolve polygons to create the best possible split conditions (multipart
# segments with dissolved lines)
arcpy.conversion.RasterToPolygon(postMSSFull, postMSSPoly, "NO_SIMPLIFY",
                           "VALUE", "SINGLE_OUTER_PART")
arcpy.management.Dissolve(postMSSPoly, postMSSPolyS, "gridcode", "", "",
                      "MULTI_PART", "DISSOLVE_LINES")
```

```
# Add year of construction information for the splitting process
arcpy.management.AddField(postMSSPolyS, "TYear", "TEXT")

fieldO = ["gridcode", "TYear"]

cursorO = arcpy.da.UpdateCursor(postMSSPolyS, fieldO)
for row in cursorO:
   if str(row[0]) == "0" or str(row[0]) == "0000":
      row[1] = "H9999"
   else:
      row[1] = "H" + str(row[0])
   cursorO.updateRow(row)
del row, cursorO

 # Split major roads polyline file based on the road
# polygons with attribute data
arcpy.analysis.Split(singleRoad, postMSSPolyS, "TYear", output)

# Change workspace to final road folder
arcpy.env.workspace = os.path.join(path, finalFolder)

# List all split post-1984 roads
HSegs = arcpy.ListFeatureClasses("H*")

# Merge split roads
arcpy.management.Merge(HSegs, PostMSSLine, "", "ADD_SOURCE_INFO")

# Add year of construction to final road network from merge field and
# then delete the merge information field
arcpy.management.AddField(PostMSSLine, "TYEAR", "TEXT")

fieldO = ["MERGE_SRC", "TYear"]

cursorO = arcpy.da.UpdateCursor(PostMSSLine, fieldO)

for row in cursorO:
   year = row[0][-8:-4]
   row[1] = str(year)
   cursorO.updateRow(row)
del row, cursorO

arcpy.management.DeleteField(PostMSSLine, ["MERGE_SRC"])

# Apply network logic rules
arcpy.management.MultipartToSinglepart(PostMSSLine, PostMSSLineS)
```

```
    logicrulesNonBin(PostMSSLineS, "TYear", "FBin", "TM1.dbf")
    removedisNetwork()


################################################################################

# Remove roads disconnected from the main network for the binary road segments
def removedisBinary():

    # File paths for required pre-generated layers
    singleRoad = os.path.join(path, referenceRoad, "SingularSegment.shp")
    edgePoints = os.path.join(path, jenksOpti, "IntersectPoint.shp")
    rastRoad30 = os.path.join(path, referenceRoad, "rr30.tif")
    rastRoad60 = os.path.join(path, referenceRoad, "rr60.tif")
    gdbFile = os.path.join(path, jenksOpti, "RoadRef.gdb")
    onlyVal = os.path.join(path, jenksOpti)
    expression1 = "Join_Count = 1 Or LEN >= 1000"

    # Change the workspace to the Jenks Optimization folder
    arcpy.env.workspace = (os.path.join(path, jenksOpti))

    # Determine the points where the roads within the study boundary connect
    # to the larger road network
    arcpy.analysis.Intersect([singleRoad, clipPoly], edgePoints, "ONLY_FID",
                             "", "POINT")
    arcpy.management.CreateFileGDB(os.path.join(path, jenksOpti),
                                   "RoadRef.gdb")

    # List all polyline binary roads
    roadsList = arcpy.ListFeatureClasses("Smerged*", "Line")
    for rra in roadsList:
        # Iteratively generated file paths
        year = str(rra[-8:-4])
        lineBuff = os.path.join(path, jenksOpti, "LineBuff" + year + ".shp")
        slineBuff = os.path.join(path, jenksOpti, "SLineBuff" + year + ".shp")
        conRoads = os.path.join(path, jenksOpti, "Con" + year + ".shp")
        exteriorIntersect = os.path.join(path, jenksOpti, "EI" + year + ".shp")
        mid = "Mid" + year + ".shp"
        onlyValFull = os.path.join(path, jenksOpti, "Val" + year + ".shp")
        midfull = os.path.join(path, jenksOpti, mid)
        outRast1 = os.path.join(path, jenksOpti, "q" + year + ".tif")
        outRast = "qq" + year + ".tif"
        clipRas = os.path.join(path, jenksOpti, "PP" + year + ".tif")
        outRast2 = os.path.join(path, jenksOpti, "cc" + year + ".tif")

        # Generate a feature class that just has road features (non-road
```

```
# segments are removed)
arcpy.conversion.FeatureClassToFeatureClass(rra, onlyVal, "Val" + year
                                + ".shp", "KBin = 1")

# Buffer the roads and create a singular segment to represent the road
# connectivity
arcpy.analysis.Buffer(onlyValFull, lineBuff, "100 Meters", "FULL",
                  "ROUND", "ALL", "", "PLANAR")
arcpy.management.MultipartToSinglepart(lineBuff, slineBuff)

# Join the buffered roads to the points that intersect with the study
# boundary
arcpy.analysis.SpatialJoin(slineBuff, edgePoints, exteriorIntersect,
                       "JOIN_ONE_TO_ONE", "KEEP_ALL", "",
                       "INTERSECT", "0 Meters")

# Calculate the perimeter of each of the polygon road features
arcpy.management.CalculateGeometryAttributes(exteriorIntersect,
                           [["LEN", "PERIMETER_LENGTH"]], "METERS")

# Retain only road segments that attached to the main network or over
# approximately 500 m
arcpy.conversion.FeatureClassToFeatureClass(exteriorIntersect, onlyVal,
                                    mid, expression1)
arcpy.management.CalculateField(midfull, "Value", "1", "PYTHON3", "",
                           "SHORT")

# Convert roads that are later than 1984 to rasters
if int(year) >= 1985:
   # Convert the polyline roads to raster
   arcpy.env.extent = "MAXOF"
   arcpy.conversion.FeatureToRaster(mid, "Value", outRast1, 30)

   # Merge the raster roads to non-road pixels
   arcpy.management.MosaicToNewRaster([outRast1, rastRoad30],
                                gdbFile, "c" + year, projection,
                                "8_BIT_SIGNED",
                                30, 1, "MAXIMUM", "FIRST")
   arcpy.management.CopyRaster(os.path.join(gdbFile, "c" + year),
                           outRast, "", "", "", "", "",
                           "8_BIT_SIGNED", "", "", "TIFF")

   # Apply focal statistics (Rectangular Neighbourhood, 7 x 7,
   # Maximum Method)
   outFocal = FocalStatistics(outRast, NbrRectangle(7, 7, "CELL"),
                          "MAXIMUM")
```

```
            # Extract to the 30 m road footprint
            outFocal2 = ExtractByMask(outFocal, rastRoad30, "INSIDE")
            outFocal2.save(outRast2)

        # Convert roads that are earlier than 1984 to rasters
        else:

            # Convert the polyline roads to raster
            arcpy.env.extent = "MAXOF"
            arcpy.conversion.FeatureToRaster(mid, "Value", outRast1, 60)

            # Merge the raster roads to non-road pixels
            arcpy.management.MosaicToNewRaster([outRast1, rastRoad60], gdbFile,
                                    "c" + year, projection,
                                    "8_BIT_SIGNED",
                                    60, 1, "MAXIMUM", "FIRST")
            arcpy.management.CopyRaster(os.path.join(gdbFile, "c" + year),
                                outRast, "", "", "", "", "",
                                "8_BIT_SIGNED", "", "", "TIFF")

            # Apply focal statistics (Rectangular Neighbourhood, 4 x 4,
            # Maximum Method)
            outFocal = FocalStatistics(outRast, NbrRectangle(4, 4, "CELL"),
                               "MAJORITY")

            # Extract to the 30 m road footprint
            outFocal2 = ExtractByMask(outFocal, rastRoad60, "INSIDE")
            outFocal2.save(outRast2)

def removedisNetwork():
    # Final paths and required variables
    outputPath = os.path.join(path, finalFolder)
    singleRoad = os.path.join(path, referenceRoad, "SingularSegment.shp")
    edgePoints = os.path.join(path, jenksOpti, "IntersectPoint.shp")
    rastRoad30 = os.path.join(path, referenceRoad, "rr30.tif")
    rastRoad60 = os.path.join(path, referenceRoad, "rr60.tif")
    onlyValFull = os.path.join(path, finalFolder, "Val.shp")
    MSSLineS = os.path.join(path, diffFolder, "MSSLineS.shp")
    lineBuff = os.path.join(path, finalFolder, "LineBuff.shp")
    slineBuff = os.path.join(path, finalFolder, "SLineBuff.shp")
    exteriorIntersect = os.path.join(path, finalFolder, "EI.shp")
    PostMSSLineS = os.path.join(path, finalFolder, "PostMSSLineS.shp")
    midfull = os.path.join(path, finalFolder, mid)
    finalMSS = os.path.join(path, finalFolder, "MSSFinal.shp")
    finalMSSR = os.path.join(path, finalFolder, "MSSFinalR.tif")
```

```
extractt2 = ExtractByMask(Pre1984, rastRoad60)
extractt2.save(os.path.join(path, finalFolder, Pre1984))
onlyValFullPost = os.path.join(path, finalFolder, "ValPost.shp")
lineBuffPost = os.path.join(path, finalFolder, "LineBuffPost.shp")
lineBuffPostS = os.path.join(path, finalFolder, "LineBuffPostS.shp")
exteriorIntersectPost = os.path.join(path, finalFolder, "EIPost.shp")
midFullPost = os.path.join(path, finalFolder, midPost)
PostMSSLineS = os.path.join(path, finalFolder, "PostMSSLineS.shp")
finalTM = os.path.join(path, finalFolder, "FinalTM.shp")
finalTMR = os.path.join(path, finalFolder, "FinalTMR.tif")
finalTMRF = os.path.join(path, finalFolder, "FinalTMRF.tif")
expression1 = "Join_Count = 1 OR LEN > 10"
midPost = "MidPost.shp"
expression1 = "Join_Count = 1"
Pre1984 = "AllPre1984.tif"
mid = "Mid.shp"

# Change working environment folder
arcpy.env.workspace = (os.path.join(path, finalFolder))

# Pre-1984 network level
# Generate a feature class that just has road features (non-road
# segments are removed)
arcpy.conversion.FeatureClassToFeatureClass(MSSLineS, outputPath, "Val.shp",
                                "FBin <> 9999 And FBin <> 0")

# Buffer roads by 100 m to create polygon segments representing road
# connectivity
arcpy.analysis.Buffer(onlyValFull, lineBuff, "100 Meters", "FULL", "ROUND",
                "ALL", "", "PLANAR")

# Ensure that each polygon is a separate polygon that can be evaluated
arcpy.management.MultipartToSinglepart(lineBuff, slineBuff)

# Spatially join the road polygons to the connectivity points with the
# study boundary
arcpy.analysis.SpatialJoin(slineBuff, edgePoints, exteriorIntersect,
                    "JOIN_ONE_TO_ONE", "KEEP_ALL", "",
                    "INTERSECT", "0 Meters")

# Calculate the perimeter of the road segments
arcpy.management.CalculateGeometryAttributes(exteriorIntersect, [["LEN",
                                "PERIMETER_LENGTH"]], "METERS")

# Retain all polygons that are connected to the main network or are
# greater than 10 m
```

```
arcpy.conversion.FeatureClassToFeatureClass(exteriorIntersect, outputPath,
                                mid, expression1)

# Clip the input road network by the mask of roads that are connected to
# the main network
arcpy.analysis.Clip(MSSLineS, midfull, finalMSS)

# Convert the roads connected to the main network to rasters
arcpy.conversion.FeatureToRaster(finalMSS, "FBIN", finalMSSR, 60)

spando3 = Expand(finalMSSR, 2, [1974, 1975, 1976, 1977, 1978, 1979, 1980,
                        1981, 1982, 1983, 1984])

# Mosaic the raster roads with the non-road segments mask
arcpy.management.MosaicToNewRaster([spando3, rastRoad60], outputPath,
                           Pre1984, projection,
                        "16_BIT_SIGNED", 30, 1, "MAXIMUM", "FIRST")

# Pre-1984 network level
# Generate a feature class that just has road features (non-road segments
# are removed)
arcpy.conversion.FeatureClassToFeatureClass(PostMSSLineS, outputPath,
                        "ValPost.shp", "FBin <> 9999 And FBin <> 0")

# Buffer roads by 100 m to create polygon segments representing road
# connectivity
arcpy.analysis.Buffer(onlyValFullPost, lineBuffPost, "100 Meters",
                  "FULL", "ROUND", "ALL", "", "PLANAR")

# Ensure that each polygon is a separate polygon that can be evaluated
arcpy.management.MultipartToSinglepart(lineBuffPost, lineBuffPostS)

# Spatially join the road polygons to the connectivity points with the
# study boundary
arcpy.analysis.SpatialJoin(lineBuffPostS, edgePoints, exteriorIntersectPost,
                        "JOIN_ONE_TO_ONE", "KEEP_ALL", "",
                        "INTERSECT", "0 Meters")

# Calculate the perimeter of the road segments
arcpy.management.CalculateGeometryAttributes(exteriorIntersectPost,
                           [["LEN", "PERIMETER_LENGTH"]], "METERS")

arcpy.conversion.FeatureClassToFeatureClass(exteriorIntersectPost,
                                outputPath, midPost, expression1)

# Clip the input road network by the mask of roads that are connected to
```

```
    # the main network
    arcpy.analysis.Clip(PostMSSLineS, midFullPost, finalTM)

    # Converted roads with connectivity to a raster
    arcpy.conversion.FeatureToRaster(finalTM, "FBIN", finalTMR, 60)

    # Smooth roads with focal statistics
    extractt2 = FocalStatistics(finalTMR, NbrRectangle(3, 3, "CELL"),
                        "MAJORITY", "DATA")

    # Reclassify connected roads to remove pre-1984 values
    remapp = Reclassify(extractt2, "VALUE", RemapRange([[0, 1984, 0]]))
    remapp.save(os.path.join(path, finalFolder, finalTMRF))

    # Finalize pre-1984 layer
    spando3 = Expand(finalMSSR, 2, [1974, 1975, 1976, 1977, 1978, 1979, 1980,
                            1981, 1982, 1983, 1984])
    arcpy.management.MosaicToNewRaster([spando3, rastRoad60], outputPath,
                        Pre1984, projection,
                        "16_BIT_SIGNED", 30, 1, "MAXIMUM", "FIRST")
    extractt2 = ExtractByMask(Pre1984, rastRoad60)
    extractt2.save(os.path.join(path, finalFolder, Pre1984))


# Logic Rules - Network level
def logicrulesNonBin(inroad, infield, AddF, tableName):
    # Local Variables
    AddFie = AddF
    inputRoad = inroad
    inF = "!" + infield + "!"
    nearTable = tableName

    ALlINFID = []

    d = {}
    fd = {}

    arcpy.management.AddField(inroad, "IN_LEN", "FLOAT")
    arcpy.management.AddField(inroad, "NEAR_LEN", "FLOAT")
    arcpy.management.AddField(inroad, "IN_BIN", "FLOAT")
    arcpy.management.AddField(inroad, "NEAR_BIN", "FLOAT")
    arcpy.management.CalculateField(inroad, "IN_BIN", inF)
    arcpy.management.CalculateField(inroad, "NEAR_BIN", inF)
    arcpy.management.CalculateGeometryAttributes(inroad, [["IN_LEN", "LENGTH"]],
                                    "METERS")
    arcpy.management.CalculateGeometryAttributes(inroad, [["NEAR_LEN",
```

```
                                        "LENGTH"]], "METERS")

# Change the workspace to the Network Dataset Folder
arcpy.env.workspace = os.path.join(path, logicRules)

# Generate a near table for the preliminary roads polylines
nearTable = os.path.join(path, logicRules, nearTable)
arcpy.GenerateNearTable_analysis(inputRoad, inputRoad, nearTable,
                                 "0 Meters", "LOCATION", "", "ALL")
arcpy.management.JoinField(nearTable, "IN_FID", inroad, "FID", ["IN_LEN"])
arcpy.management.JoinField(nearTable, "NEAR_FID", inroad, "FID",
                           ["NEAR_LEN"])
arcpy.management.JoinField(nearTable, "IN_FID", inroad, "FID",
                           ["IN_BIN"])
arcpy.management.JoinField(nearTable, "NEAR_FID", inroad, "FID",
                           ["NEAR_BIN"])
arcpy.management.AddField(nearTable, "FINALBIN", "FLOAT")
# Field lists required for Logic Rule execution
fields = ["IN_FID", "NEAR_FID", "FROM_X", "FROM_Y", "NEAR_X", "NEAR_Y",
          "OID", "IN_LEN", "NEAR_LEN", "IN_BIN",
          "NEAR_BIN", "FINALBIN"]

# Determine all input IDs in Near Table
with arcpy.da.UpdateCursor(nearTable, fields) as cursor:
    for row in cursor:
        ALlINFID.append(row[0])
del cursor, row

UniqueFID = list(set(ALlINFID))

for i in UniqueFID:
    d["L" + str(i)] = []

with arcpy.da.UpdateCursor(nearTable, fields) as cursor:
    for item in cursor:
        # print(item)
        fidVal = item[0]
        # print(fidVal)
        for key in d:
            # print(key)
            fidVal2 = "L" + str(fidVal)
            # print("fid val", fidVal2)
            if key == fidVal2:
                d[key] = d.setdefault(key, []) + [item]
            del fidVal2
del cursor, item
```

```
for i in UniqueFID:
   fd["L" + str(i)] = None

for key in d:
   # print(key, '->', d[key])
   numElements = len(d[key])

   # Dead End
   if numElements == 1:
      segmentAYear = d[key][0][10]
      segmentTargetYear = d[key][0][9]
      segmentALength = d[key][0][8] * 0.5
      segmentTargetLength = d[key][0][7]

      # If the dead end base and target segment are not the same
      if segmentAYear != segmentTargetYear:
         # If the target segment has a date earlier than the base and the
         # base is larger than the target year
         if segmentTargetYear < segmentAYear and segmentALength > \
               segmentTargetLength:
            # The target should be the same age as the base
            update = segmentAYear
         else:
            update = segmentTargetYear
      else:
         update = segmentTargetYear
      fd[key] = update
      del update


   # Line or T
   elif numElements == 2:
      # If the adjacent road segments' "NEAR_X"s and
      # "NEAR_Y"s are spatially coincident, the target road segment
      # is part of a "T"
      if (float(d[key][0][4]) == float(d[key][1][4])) and (
            float(d[key][0][5]) == float(d[key][1][5])):

         segmentAYear = d[key][0][10]
         segmentBYear = d[key][1][10]
         segmentTargetYear = d[key][0][9]
         segmentABLength = (d[key][0][8] + d[key][1][8]) * 0.5
         segmentTargetLength = d[key][0][7]

         # If the two road segments making the base of the T are both a
```

```
        # road or both not a road
        if segmentBYear == segmentAYear:
            # If the base of the road is older than the target year and
            # the base is larger than the target year
            if segmentBYear > segmentTargetYear and \
                    segmentTargetLength < segmentABLength:
                # The dead end should be the same age as the base
                update = segmentBYear
            else:
                update = segmentTargetYear
        else:
            update = segmentTargetYear
        fd[key] = update
        del update

    # If the adjacent road segments' "NEAR_X"s and "NEAR_Y"s are
    # not spatially coincident, the
    # target road segment is a line
    if (float(d[key][0][4]) != float(d[key][1][4])) or (
            float(d[key][0][5]) != float(d[key][1][5])):

        segmentAYear = d[key][0][10]
        segmentBYear = d[key][1][10]
        segmentTargetYear = d[key][0][9]
        segmentABLength = (d[key][0][8] + d[key][1][8]) * 0.5
        segmentTargetLength = d[key][0][7]

        # If the two adjacent road segments are the same value
        if segmentBYear == segmentAYear:
            # If the two adjacent segments are the same and the target
            # is different
            # As well as the two adjacent segments being longer than the
            # target segment,
            # the target should be the same as the adjacent roads
            if segmentBYear != segmentTargetYear and \
                    segmentABLength > segmentTargetLength:
                update = segmentBYear
            else:
                update = segmentTargetYear
        else:
            update = segmentTargetYear
        fd[key] = update
        del update
else:
    fd[key] = d[key][0][9]
```

```
        arcpy.management.AddField(inputRoad, AddFie, "LONG")

        with arcpy.da.UpdateCursor(inputRoad, ["FID", AddFie]) as cursor:
            for row in cursor:
                key = "L" + str(row[0])
                if int(key[1:]) in UniqueFID:
                    updateVal = fd[key]
                    if updateVal == None:
                        updateVal = d[key][0][9]
                    row[1] = updateVal
                else:
                    row[1] = 9999
                cursor.updateRow(row)
        del cursor, row

# Logic Rules - Binary Level
def logicrulesNoD(inroad, infield, NTName, AddF):
    AddFie = AddF
    inputRoad = inroad
    nearTableName = NTName
    inF = "!" + infield + "!"

    ALlINFID = []
    d = {}
    fd = {}

    arcpy.management.AddField(inroad, "IN_LEN", "FLOAT")
    arcpy.management.AddField(inroad, "NEAR_LEN", "FLOAT")
    arcpy.management.AddField(inroad, "IN_BIN", "FLOAT")
    arcpy.management.AddField(inroad, "NEAR_BIN", "FLOAT")
    arcpy.management.CalculateField(inroad, "IN_BIN", inF)
    arcpy.management.CalculateField(inroad, "NEAR_BIN", inF)
    arcpy.management.CalculateGeometryAttributes(inroad, [["IN_LEN",
                                            "LENGTH"]], "METERS")
    arcpy.management.CalculateGeometryAttributes(inroad, [["NEAR_LEN",
                                            "LENGTH"]], "METERS")

    # Change the workspace to the Network Dataset Folder
    arcpy.env.workspace = os.path.join(path, logicRules)

    # Generate a near table for the preliminary roads polylines
    nearTable = os.path.join(path, logicRules, nearTableName)
    arcpy.GenerateNearTable_analysis(inputRoad, inputRoad, nearTable,
                             "0 Meters", "LOCATION", "", "ALL")
    arcpy.management.JoinField(nearTable, "IN_FID", inroad, "FID", ["IN_LEN"])
    arcpy.management.JoinField(nearTable, "NEAR_FID", inroad, "FID",
```

```
                    ["NEAR_LEN"])
arcpy.management.JoinField(nearTable, "IN_FID", inroad, "FID", ["IN_BIN"])
arcpy.management.JoinField(nearTable, "NEAR_FID", inroad, "FID",
                    ["NEAR_BIN"])
arcpy.management.AddField(nearTable, "FINALBIN", "FLOAT")

# Field lists required for Logic Rule execution
fields = ["IN_FID", "NEAR_FID", "FROM_X", "FROM_Y", "NEAR_X", "NEAR_Y",
          "OID", "IN_LEN", "NEAR_LEN", "IN_BIN",
          "NEAR_BIN", "FINALBIN"]

# Determine all input IDs in Near Table
with arcpy.da.UpdateCursor(nearTable, fields) as cursor:
    for row in cursor:
        ALlINFID.append(row[0])
del cursor, row

UniqueFID = list(set(ALlINFID))

for i in UniqueFID:
    d["L" + str(i)] = []

with arcpy.da.UpdateCursor(nearTable, fields) as cursor:
    for item in cursor:
        # print(item)
        fidVal = item[0]
        # print(fidVal)
        for key in d:
            # print(key)
            fidVal2 = "L" + str(fidVal)
            # print("fid val", fidVal2)
            if key == fidVal2:
                d[key] = d.setdefault(key, []) + [item]
            del fidVal2
del cursor, item

for i in UniqueFID:
    fd["L" + str(i)] = None

for key in d:
    # print(key, '->', d[key])
    numElements = len(d[key])

    # Dead End
    if numElements == 1:
        segmentAYear = d[key][0][10]
```

```
    segmentTargetYear = d[key][0][9]
    segmentALength = d[key][0][8] * 1.5
    segmentTargetLength = d[key][0][7]


    # If the dead end base and target segment are not the same
    # If the target year is 1 but the base is 0, if the base is
    # larger than the target year
    if segmentAYear == 0 and segmentTargetYear == 1 and \
            segmentAYear != segmentTargetYear and segmentALength > \
            segmentTargetLength:
        # The target year does not exist
        update = 0
        fd[key] = update


# Line or T
elif numElements == 2:
    # If the adjacent road segments' "NEAR_X"s and
    # "NEAR_Y"s are spatially coincident, the target road segment
    # is part of a "T"
    if (float(d[key][0][4]) == float(d[key][1][4])) and (
            float(d[key][0][5]) == float(d[key][1][5])):

        segmentAYear = d[key][0][10]
        segmentBYear = d[key][1][10]
        segmentTargetYear = d[key][0][9]
        segmentABLength = (d[key][0][8] + d[key][1][8]) * 1.5
        segmentTargetLength = d[key][0][7]

        # If the two road segments making the base of the T are both a
        # road or both not a road. If the base of the road is 0 and
        # the target year is 1 and the dead end is smaller
        # than the base road:
        if segmentBYear == 0 and segmentTargetYear == 1 and \
                segmentBYear == segmentAYear and segmentTargetLength\
                < segmentABLength:
            # The target year cannot exist
            update = 0
            fd[key] = update
            del update

    # If the adjacent road segments' "NEAR_X"s and "NEAR_Y"s are not
    # spatially coincident, the
    # target road segment is a line
    if (float(d[key][0][4]) != float(d[key][1][4])) or (
```

```
            float(d[key][0][5]) != float(d[key][1][5])):

        segmentAYear = d[key][0][10]
        segmentBYear = d[key][1][10]
        segmentTargetYear = d[key][0][9]
        segmentABLength = (d[key][0][8] + d[key][1][8]) * 1.5
        segmentTargetLength = d[key][0][7]


        # If the two adjacent road segments are the same value

        # If the two adjacent segments are both 1 and longer than the
        # target segment,
        # the target should exist
        if segmentBYear == segmentAYear and segmentBYear == 1 and \
              segmentABLength >= segmentTargetLength:
           update = 1
           fd[key] = update
           del update

           # If the two adjacent segments are both 0 and longer than
           # the target segment,
           # the target should not exist
        elif segmentBYear == segmentAYear and segmentBYear == 0 and \
              segmentABLength >= segmentTargetLength:
           update = 0
           fd[key] = update
           del update

arcpy.management.AddField(inroad, AddFie, "LONG")

with arcpy.da.UpdateCursor(inputRoad, ["FID", AddFie]) as cursor:
   for row in cursor:
      key = "L" + str(row[0])
      if int(key[1:]) in UniqueFID:
         updateVal = fd[key]
         if updateVal == None:
            updateVal = d[key][0][9]
         row[1] = updateVal
      cursor.updateRow(row)
del cursor, row

arcpy.DeleteField_management(inputRoad, ['Dissolve', infield, 'ORIG_FID',
                             'IN_LEN', 'NEAR_LEN', 'IN_BIN',
                             'NEAR_BIN'])
```

```
################################################################################
# Pre-processing Steps for all Sensors
createfolders(inputPath, outerFile)
OL, T, M = listimages(inputPath)
cloudcover(OL, "OLI")
cloudcover(T, "TM/ETM+")
cloudcover(M, "MSS")
inpputyearnums(inputPath)
preprocessLS8(OL)
preprocessLS457(T)
preprocessLS12345(M, inputPath)
d, y = moasicdictionary()

# BAP Composite
doyscore(d, y)
sensorscore(d, y)
cloudshadow()
opacityscore()
brightnessscore(d)
finalbap(d)

# Road Extraction and Change Detection
referenceprep()
extractroad()
jenksopt()
difference()

# Duration information
print('The script took {0} minutes !'.format((time.time() - startTime) / 60))


################################################################################
```